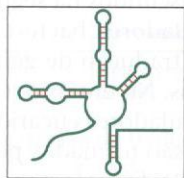


CAPÍTULO 20



RNAs Reguladores

NOS DOIS ÚLTIMOS CAPÍTULOS, DISCUTIU-SE como a transcrição é regulada em procariotos e eucariotos. Aprendeu-se que este controle é alcançado usando proteínas reguladoras – normalmente, proteínas de ligação ao DNA sequência-específicas que ativam ou reprimem a transcrição de genes próximos. Os detalhes mecanísticos da regulação gênica têm sido estudados desde que François Jacob e Jacques Monod propuseram seu modelo de repressão há mais de 50 anos (Cap. 18, Quadro 18-2). Naquela época, eles não sabiam se os fatores em *trans* (repressores) eram proteínas ou RNA. Transparecia que, nos casos que estudaram (e, de fato, na maioria dos casos), os reguladores eram proteínas que atuavam pela ligação a sítios operadores do DNA. Todavia, em seu artigo original, eles sugeriram que os reguladores poderiam facilmente ser moléculas de RNA – de fato, eles favoreciam esta possibilidade.

A ideia de que moléculas de RNA poderiam ser reguladores foi amplamente esquecida à medida que mais e mais proteínas reguladoras eram encontradas em procariotos e eucariotos. Mas, nos últimos anos, houve uma explosão no estudo de reguladores de RNA, particularmente em eucariotos, que operam em nível de transcrição e, sobretudo, tradução. Este novo campo emergiu a partir de duas fontes: a descoberta dos microRNAs, primeiramente descritos no início da década de 1990 e, então, a descoberta do fenômeno conhecido como RNA de interferência no fim da década de 1990. Antes de estas formas de regulação serem descritas – como elas funcionam e as aplicações que permitem aos cientistas –, serão abordados casos de regulação gênica mediada por RNA em bactérias.

REGULAÇÃO POR RNAs EM BACTÉRIAS

Pequenos RNAs têm sido reconhecidos em procariotos há muitos anos. Alguns deles estão envolvidos na regulação da replicação de plasmídeos, e outros estão envolvidos na regulação da expressão gênica (ver discussão sobre *Tn10* no Cap. 12). Alguns dos RNAs deste último grupo controlam a transcrição – o RNA 6S de *Escherichia coli*, por exemplo. Este RNA liga-se à subunidade σ^{70} da RNA-polimerase e controla negativamente a transcrição a partir de vários promotores de σ^{70} . O RNA 6S acumula-se em altos níveis na fase estacionária (fase de crescimento em que as bactérias entram quando os nutrientes se tornam escassos e as células param de se dividir). Na fase estacionária, um fator σ alternativo, σ^S , é produzido. Este σ compete com σ^{70} pelo núcleo da polimerase e dirige a enzima para promotores expressando genes para as múltiplas respostas a estresse necessárias para a sobrevivência na fase estacionária. Ao reduzir a transcrição a partir de promotores de σ^{70} , o RNA 6S ajuda neste redirecionamento de expressão para promotores σ^S .

SUMÁRIO

Regulação por RNAs
em Bactérias, 701

Os RNAs Reguladores estão
Amplamente Distribuídos
em Eucariotos, 711

Síntese e Função das
Moléculas de miRNA, 714

Silenciamento da Expressão
Gênica por Pequenos RNAs, 718

RNAs Longos Não Codificadores
e Inativação do X, 728

Nos últimos anos, a atenção tem sido voltada para pequenas moléculas de RNA de bactérias que regulam a tradução e a degradação de mRNA. O interesse nestes pequenos RNAs foi aumentado por sua semelhança a RNAs que regulam a expressão gênica em eucariotos – os pequenos RNAs de interferência e os microRNAs que serão discutidos na segunda metade deste capítulo.

Uma classe de RNAs reguladores bacterianos (chamados de **sRNAs**) atua em *trans* para controlar a tradução de genes-alvo, parecido com o que microRNAs fazem em eucariotos. No entanto, eles são maiores (80 a 110 nucleotídeos) do que os RNAs reguladores eucarióticos (que variam de 21 a 30 nucleotídeos) e, em geral, não são formados pelo processamento de precursores de RNA de dupla-fita (dsRNA) maiores (como o são os reguladores de RNA eucarióticos); em vez disso, eles são codificados em sua forma final por pequenos genes. Vários destes genes foram identificados por bioinformática, com mais de 100 sRNAs sendo descobertos em *E. coli*. A maioria dos sRNAs atua pelo pareamento de bases com sequências complementares nos mRNAs-alvo e dirigindo a destruição do mRNA, inibindo sua tradução ou até mesmo, em alguns casos, *estimulando* a tradução.

A ligação de um sRNA a seu mRNA-alvo é, na maioria dos casos, auxiliada pela proteína bacteriana Hfq. Esta chaperona de RNA é necessária porque a complementaridade entre os sRNAs e seus mRNAs-alvo é normalmente imperfeita e curta e, assim, sua interação é fraca. A Hfq facilita o pareamento de bases. Além disso, ao ligar-se aos sRNAs mesmo antes que eles estejam pareados com seus alvos, a Hfq aumenta a estabilidade destes reguladores.

Um sRNA bastante estudado de *E. coli* é o RNA RybB, com 81 nucleotídeos. Este sRNA liga-se a vários mRNAs-alvo e desencadeia sua destruição porque o trecho de dupla-fita do heterodúplex formado pelo pareamento é reconhecido como substrato pela nuclease RNase E. A maioria dos mRNAs-alvo de RybB codifica proteínas de armazenamento de ferro. Ferro livre é necessário para a célula sob certas condições, mas altos níveis são tóxicos. RybB regula os níveis de ferro livre pelo controle de proteínas de armazenamento de ferro. RybB é expresso a partir de um promotor reconhecido por um fator σ especial chamado σ^E (assim como σ^S , um fator σ de resposta ao estresse). A própria expressão do gene que codifica σ^E é regulada por RybB e, assim, este sRNA faz parte de um circuito regulador autonegativo por σ^E .

O fator σ de fase estacionária σ^S , mencionado anteriormente, é codificado pelo gene *rpoS* de *E. coli*. A tradução do mRNA de *rpoS* é estimulada por dois sRNAs: DsrA e RprA. A ativação é alcançada por uma troca no pareamento de bases de RNA alternativo: os pequenos RNAs ligam-se a uma região do mRNA que de outro modo parearia com o sítio de ligação ao ribossomo, inibindo a tradução. O gene *rpoS* também é controlado negativamente por outra pequena molécula, OxyS. A Figura 20-1 mostra estes dois mecanismos.

Outros exemplos de RNAs reguladores em bactérias atuam simplesmente como RNAs “antisense”: eles são codificados pela fita oposta à fita codificadora de um gene e atuam pelo pareamento de bases homólogas para inibir a expressão do mRNA produzido a partir do gene. Estes tendem a estar associados com genes que codificam produtos potencialmente tóxicos, e também na regulação de alguns genes de fago (como em λ) (ver Cap. 18). Em geral, diz-se que estes RNAs atuam em *cis* porque agem apenas no gene a partir do qual foram gerados (ao contrário dos sRNAs de atuação em *trans* descritos anteriormente).

Os RNAs reguladores de atuação em *trans* serão retomados na segunda metade deste capítulo, onde será abordado seu papel na regulação da expressão gênica em eucariotos. Mas antes de a atenção ser voltada para este tópico, serão abordados outros exemplos de regulação gênica mediada pelo pareamento de RNA alternativo em bactérias, que operam verdadeiramente em *cis*. Estes são elementos reguladores de RNA que controlam a expressão dos genes em cujos mRNAs eles residem. Os exemplos mais marcantes são os chamados **ribocomutadores**, que controlam operons metabólicos e a **atenuação** em

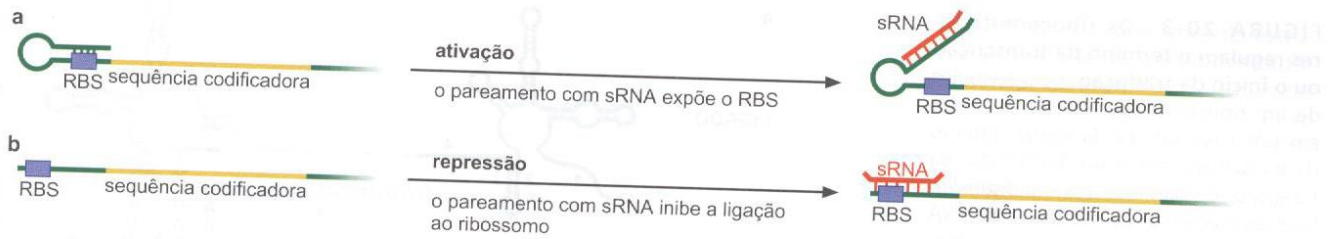


FIGURA 20-1 A ativação e a repressão da tradução por sRNAs. Quando o sítio de ligação ao ribossomo (RBS) está oculto devido ao pareamento de bases com outra molécula de RNA (como ocorre em b) ou por pareamento com outra região da própria molécula de RNA (como ocorre em a), a tradução é inibida. (Adaptada, com permissão, de Gottesman S. et al. 2006. *Cold Spring Harbor Symp. Quant. Biol.* 71: 1-11, Fig. 1. © Cold Spring Harbor Laboratory Press.)

óperons biossintéticos. Os genes *trp* de *E. coli* são o exemplo clássico deste último mecanismo e onde a regulação mediada por RNA foi descoberta (o caso será descrito em detalhes no Quadro 20-1).

Os ribocomutadores residem nos transcritos dos genes cuja expressão eles controlam por meio de alterações da estrutura secundária

Os ribocomutadores controlam a expressão gênica em resposta a alterações nas concentrações de pequenas moléculas. Estes elementos reguladores são normalmente encontrados nas regiões 5' não traduzidas (5'-UTRs) dos genes que eles controlam e podem regular a expressão em nível de transcrição ou tradução. Eles fazem isso por meio de alterações na estrutura secundária do RNA, como será visto.

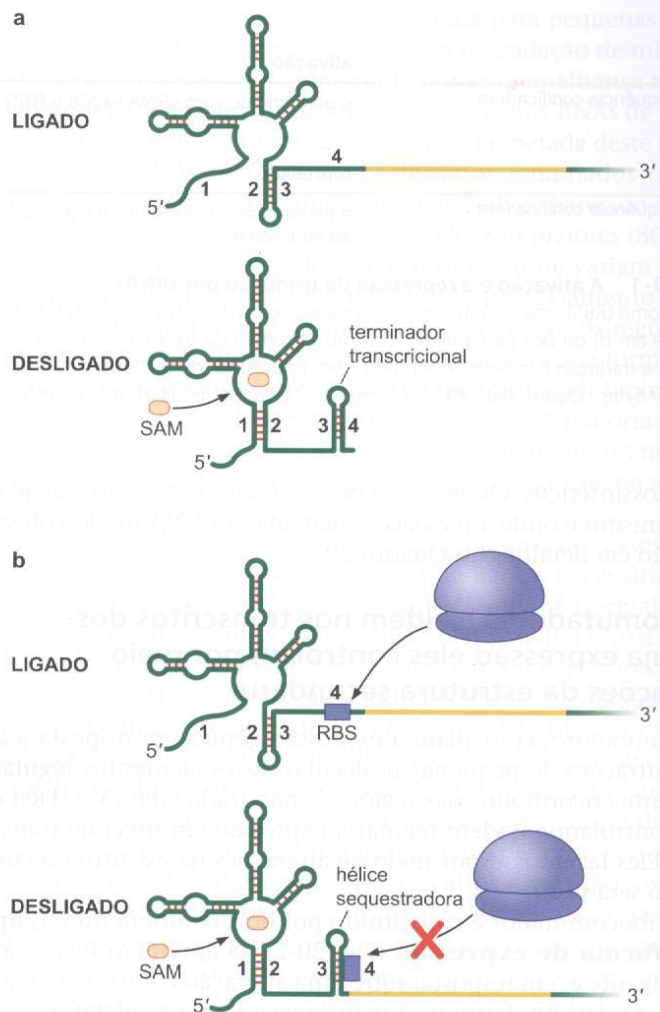
Cada ribocomutador é constituído por dois componentes: o **aptâmero** e a **plataforma de expressão** (Fig. 20-2). O aptâmero liga-se à pequena molécula ligante e, em resposta, sofre uma alteração conformacional que, por sua vez, causa uma mudança na estrutura secundária da plataforma de expressão adjacente. Estas alterações conformacionais alteram a expressão do gene associado pela indução do término da transcrição ou pela inibição do início da transcrição. Ambos os mecanismos estão ilustrados no exemplo apresentado na Figura 20-3, que será descrito agora.

Os ribocomutadores são normalmente encontrados, não surpreendentemente, a montante dos genes envolvidos na síntese do ligante metabólico reconhecido pelo ribocomutador em questão. Em *Bacillus subtilis*, por exemplo, muitos genes envolvidos no uso do aminoácido metionina possuem um RNA-líder não traduzido de 200 nucleotídeos que atua como ribocomutador sensor de SAM (S-adenosilmetionina). A RNA-polimerase inicia a transcrição no promotor e transcreve esta região-líder antes de entrar na sequência codificadora dos genes a jusante. Uma vez transcrita em RNA, a região-líder pode adotar estruturas alternativas por meio de padrões alternativos de pareamento de bases intramoleculares (Fig. 20-3a). Um destes arranjos inclui um



FIGURA 20-2 Organização dos RNAs ribocomutadores. Como descrito no texto, o aptâmero liga-se ao metabólito controlador, causando mudanças na estrutura da plataforma de expressão adjacente. Os aptâmeros identificados até o momento variam de tamanho entre 70 e 200 nucleotídeos; as plataformas de expressão são mais variadas, tanto em tamanho quanto em caráter.

FIGURA 20-3 Os ribocomutadores regulam o término da transcrição ou o início da tradução. Dois exemplos de um ribocomutador sensor de SAM, em um caso (a) regulando o término da transcrição, em outro (b), o início da tradução. Os números 1 a 4 indicam diferentes elementos de sequência do RNA a montante da região codificadora (em amarelo). Na ausência de SAM, as regiões 2 e 3 formam uma haste-alça; na presença de SAM, as regiões 1 e 2 formam uma haste-alça, e as regiões 3 e 4 também. A consequência desta mudança de estrutura secundária controla a transcrição ou a tradução, como mostrado. (a) Uma haste-alça das regiões 3 e 4 produz um terminador transcricional, que faz a RNA-polimerase terminar a transcrição imediatamente após a transcrição destas regiões e antes de entrar na região codificadora a jusante. A haste-alça, neste caso, é seguida por um trecho de Us no mRNA, outra característica do terminador transcricional (Cap. 13, Fig. 13-13). (b) A haste-alça formada pelas regiões 3 e 4 inibe o início da tradução ao sequestrar o sítio de ligação ao ribossomo, como mostrado.



terminador transcricional haste-alça (ver Cap. 13). SAM – o ligante para este ribocomutador – liga-se ao aptâmero e estabiliza a estrutura secundária que inclui este terminador transcricional (como mostrado na base da Fig. 20-3a). Sob estas circunstâncias, a transcrição é terminada antes que a polimerase tenha a chance de transcrever o segmento codificador de proteína do gene a jusante. Esta forma de regulação transcricional é também chamada de **atenuação**. (Observa-se a semelhança mecânica com o sistema *trp* descrito no Quadro 20-1.) Em outro caso – em outro gene – um ribocomutador sensor de SAM pode atuar pela regulação da tradução. Neste caso, como mostrado na Figura 20-3b, a estrutura secundária alternativa estabilizada pela ligação de SAM ao aptâmero inclui uma haste-alça que, embora não seja um terminador transcricional, inclui o sítio de ligação ao ribossomo (RBS; na região 4). Esta alteração conformacional sequestra o RBS e bloqueia o início da tradução pelos ribossomos (ver Cap. 15). Esta forma de inibição da tradução é, portanto, essencialmente idêntica à descrita anteriormente para os sRNAs de atuação em *trans* (Fig. 20-1). Os detalhes das alterações na estrutura secundária do RNA induzidas pela ligação de SAM a um ribocomutador estão mostrados na Figura 20-4.

Vários ribocomutadores foram identificados, e resultados atuais de sequenciamento do genoma inteiro sugerem que há provavelmente centenas ou milhares deles em espécies de bactérias. Mesmo os exemplos bem-carac-

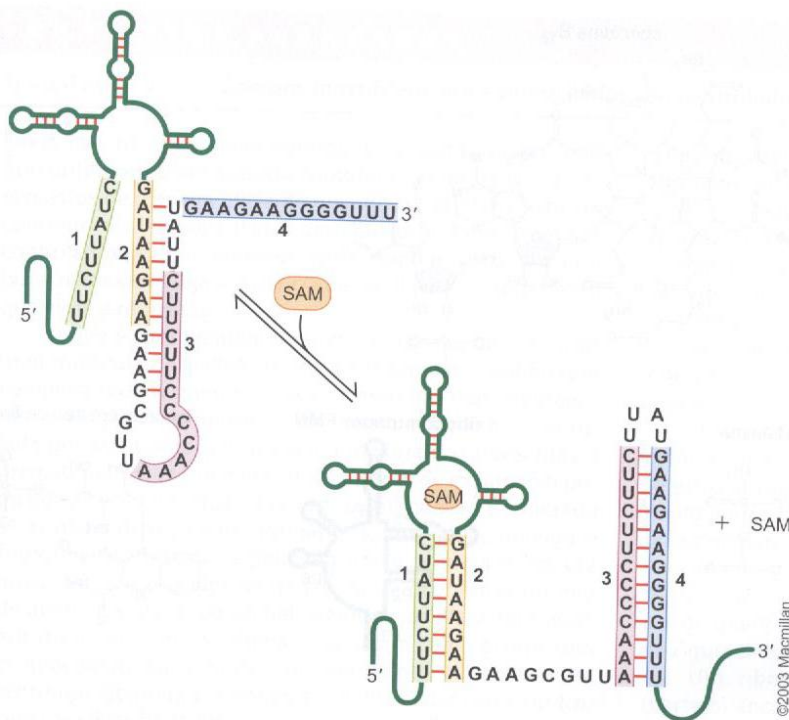


FIGURA 20-4 Alterações da estrutura secundária de um ribocomutador sensor de SAM. As sequências das regiões 1 a 4 (descritas na Fig. 20-3) são aqui mostradas em detalhes e em diferentes cores. Os pareamentos de bases encontrados nas duas estruturas secundárias alternativas – ou seja, com e sem SAM ligado – estão mostrados. (Adaptada, com permissão, de Winkler W.C. et al. 2003. *Nat. Struct. Biol.* **10**: 701-707, Fig. 5b. © Macmillan.)

terizados respondem a uma gama de diferentes metabólitos, incluindo lisina e outros aminoácidos, vitamina B12, coenzima pirofosfato de tiamina (TPP), mononucleotídeo de flavina (FMN) e guanina (Fig. 20-5).

Outro tipo de ribocomutador responde a tRNAs não carregados, em vez de a pequenas moléculas ligantes. Deste modo, determinados genes, sobretudo os genes de aminoacil-tRNA sintetases (ver Cap. 15), são controlados por atenuação mediada por um RNA-líder não traduzido, com 200 a 300 nucleotídeos de comprimento, que interage direta e especificamente com o tRNA correspondente, não carregado, pela sintetase: a forma carregada do tRNA não se encaixa no bolso de ligação fornecido pela estrutura secundária do RNA. A ligação do tRNA não carregado estabiliza o RNA-líder em sua estrutura de antitermino, de modo que a transcrição do gene de sintetase adjacente pode prosseguir. A especificidade é alcançada, em parte, pela interação “códon-anticódon” entre o tRNA e o RNA-líder. Como apenas o tRNA não carregado é capaz de ligar-se ao líder, a leitura de todo o transcrito é estimulada apenas quando o aminoácido correspondente está escasso, e o nível do tRNA não carregado está aumentado na célula.

Os ribocomutadores, apesar de muito prevalente entre bactérias, são também encontrados em arqueobactérias, fungos e plantas. Em alguns casos, nestes organismos mais complexos, os ribocomutadores estão envolvidos até mesmo no controle do processamento alternativo (Cap. 14). Assim, por exemplo, em um caso descrito no fungo *Neurospora crassa*, três aptâmeros de TPP foram identificados, e dois deles inibiram, e o terceiro estimulou, a expressão de genes por meio da regulação do processamento de RNA.

RNAs como agentes de defesa em procariotos e arqueobactérias

Antes de abordar o papel dos RNAs reguladores em eucariotos, há outro sistema a ser considerado em bactérias. Embora ele não seja estritamente um exemplo de regulação gênica (é um sistema de defesa contra vírus e outros

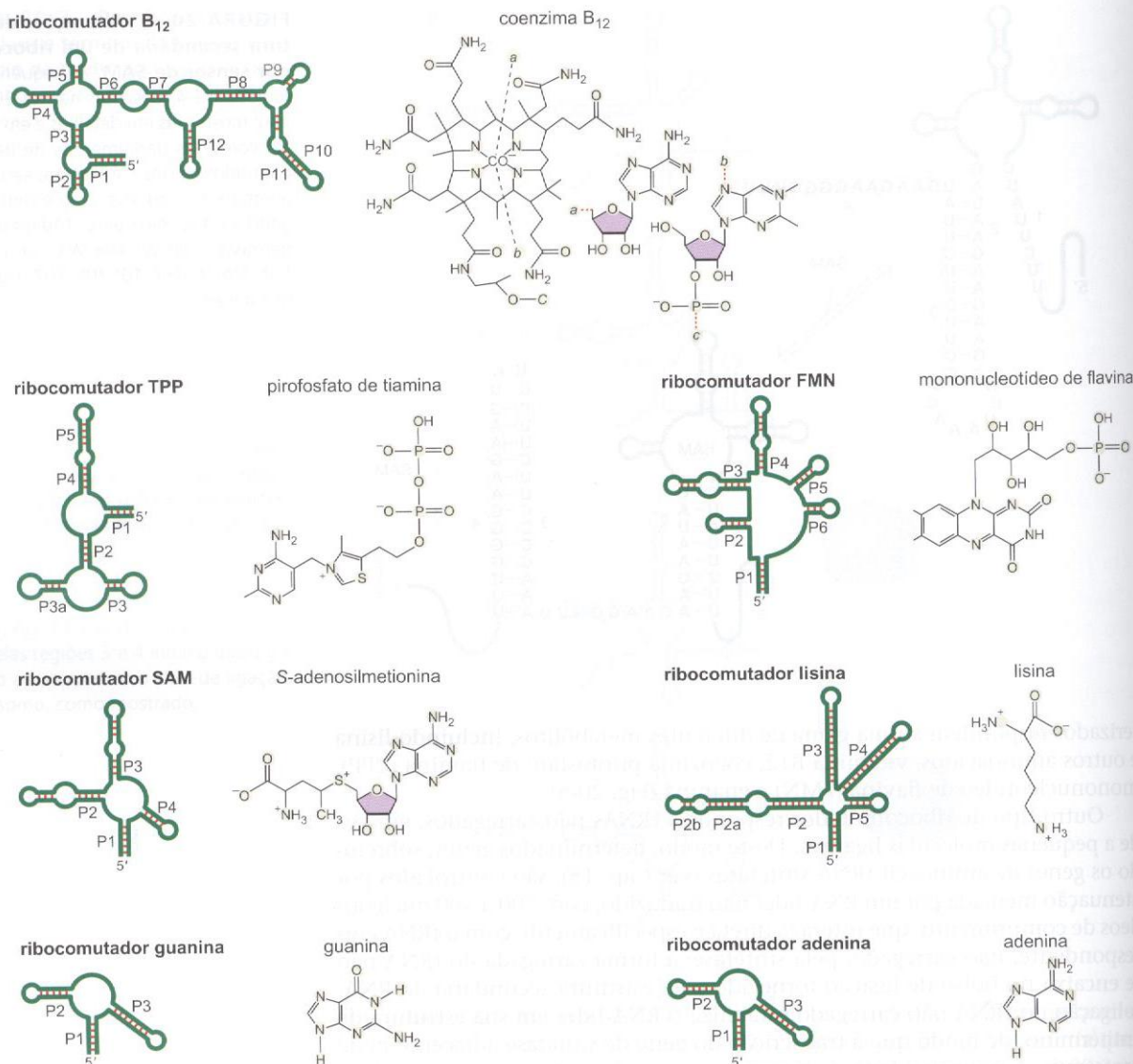


FIGURA 20-5 Ribocomutadores respondem a uma ampla gama de metabólitos. A estrutura secundária de sete ribocomutadores e os metabólitos aos quais eles são sensíveis são mostrados aqui. (Adaptada, com permissão, de Mandal M. et al. 2003. *Cell* **113**: 577-586, Fig. 7A. © Elsevier.)

invasores extracromossômicos), o mecanismo usado é bastante semelhante aos sistemas que serão encontrados em eucariotos, isto é, o RNAi.

CRISPRs são o registro de infecções sobrevividas e resistência adquirida

Anos antes de que qualquer função pudesse ser atribuída a eles, pequenos trechos de sequências incomuns mas caracteristicamente organizadas foram observados nos genomas de várias bactérias. O padrão distinto levou ao nome bastante complicado de Repetições Palindrômicas Curtas Agrupadas e Regularmente Intercaladas (ou **CRISPRs**, *Clustered Regularly Interspaced Short Palindromic Repeats*). As características gerais estão apresentadas na Figura 20-6 e consistem em **sequências repetidas** (cada uma com cerca de 30 pb e altamente conservadas em um dado agrupamento) intercaladas com **sequências**

▶ CONCEITOS AVANÇADOS

Quadro 20-1 Óperons biossintéticos de aminoácidos são controlados por atenuação

Em *E. coli*, os cinco genes contíguos *trp* codificam enzimas que sintetizam o aminoácido triptofano. Esses genes só são expressos de maneira eficiente quando o triptofano está em concentração limitante (Fig. 1 deste quadro). Estes genes são controlados por um repressor, assim como os genes *lac*, embora, neste caso, seja a ausência de seu ligante (o triptofano) que alivie a repressão.

O que é surpreendente, porém, é que depois de iniciar uma molécula de mRNA de *trp*, nem sempre a polimerase completa todo o transcrito. Assim como nos ribocomutadores, a decisão de produzir um transcrito completo é controlada por atenuação; neste caso, a maioria dos transcritos é prematuramente terminada, antes mesmo de incluírem o primeiro gene de *trp* (*trpE*). Contudo, a atenuação é superada se os níveis de triptofano forem baixos na célula; quando o triptofano é limitante, a polimerase não termina e, em vez disso, transcreve todos os genes *trp*. A ocorrência ou não de atenuação depende da habilidade de os RNAs formarem estruturas secundárias alternativas, assim como ocorre com os ribocomutadores. Neste caso, entretanto, a escolha entre estruturas alternativas formadas pelo RNA-líder não é controlada pela ligação do ligante diretamente ao RNA; em vez disso, a escolha de alternativas depende do acoplamento entre transcrição e tradução em bactérias.

A sequência da extremidade 5' do mRNA do óperon *trp* inclui uma sequência-líder de 161 nucleotídeos a montante do primeiro códon de *trpE* (Fig. 2 deste quadro). Próximo ao fim dessa sequência-líder, e antes do *trpE*, existe um terminador de transcrição, constituído por uma alça em grampo característica no RNA (formada pelas sequências das regiões 3 e 4 da Fig. 2 deste quadro), seguida por oito resíduos de uridina (ver Cap. 13, Fig. 13-13). A transcrição geralmente para após este terminador (e poderíamos pensar que a parada deveria ser obrigatória) produzindo um RNA-líder com 139 nucleotídeos. Esse RNA é produzido quando há altos níveis de triptofano.

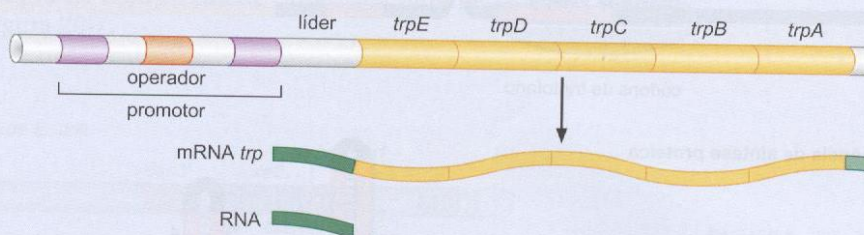
Três características da sequência-líder permitem que o terminador seja transposto pela RNA-polimerase quando a concentração celular de triptofano está baixa. Primeira, um

segundo grampo (além do grampo de término) pode ser formado entre as regiões 1 e 2 do líder (ver Fig. 2 deste quadro). Segunda, a região 2 também é complementar à região 3; podendo, portanto, ser formado mais um grampo com as regiões 2 e 3 e, quando isso ocorre, não há formação do grampo de término (3, 4). Terceira, o RNA do líder contém uma fase de leitura aberta que codifica um pequeno peptídeo-líder, com 14 aminoácidos, a qual é precedida por um forte sítio de ligação ao ribossomo (ver Fig. 2 deste quadro).

A sequência que codifica o peptídeo-líder tem uma característica muito interessante: dois códons de triptofano em sequência. Quando a concentração de triptofano é baixa, há muito pouco tRNA de triptofano carregado disponível, e o ribossomo para ao chegar aos dois códons de triptofano. Nestas circunstâncias, o RNA nas proximidades dos códons de triptofano está no interior do ribossomo e não pode participar do grampo. A consequência desse cenário é apresentada na Figura 3 deste quadro.

Um ribossomo parado nos códons de triptofano (parte b) encobre a região 1, deixando a região 2 livre para parear com a região 3; portanto, o grampo de término (formado pelas regiões 3 e 4) não será formado e a transcrição não é atenuada. Por outro lado, se houver triptofano suficiente (e, portanto, suficientes tRNAs carregados com Trp) para permitir o avanço do ribossomo por meio dos códons de triptofano, o ribossomo bloqueia a sequência 2 e, durante este período, as regiões de RNA 3 e 4 estarão sendo produzidas. Assim, o terminador forma-se, atenuando a transcrição, e os genes *trp* não são transcritos.

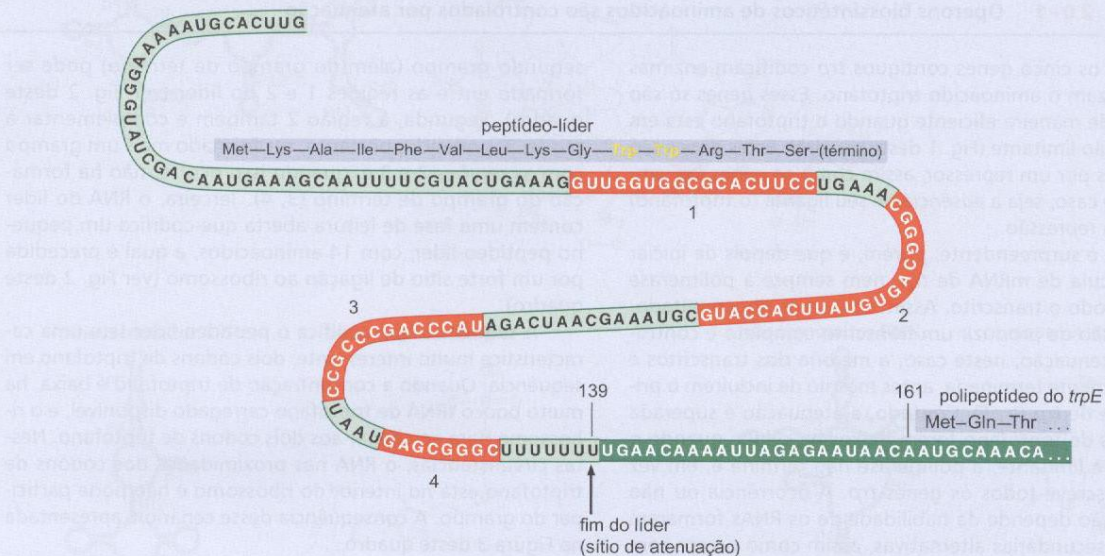
O óperon *trp* é controlado por repressão e atenuação, fornecendo uma resposta de dois estágios à redução progressivamente mais rigorosa dos níveis de triptofano. Contudo, a atenuação, por si, pode fornecer regulação robusta: outros óperons de aminoácidos, como *leu* e *his*, dependem totalmente da atenuação para seu controle. No caso da leucina, o peptídeo-líder do óperon da leucina tem quatro códons de leucina adjacentes, e o peptídeo-líder do óperon da histidina contém sete códons de histidina em sequência.



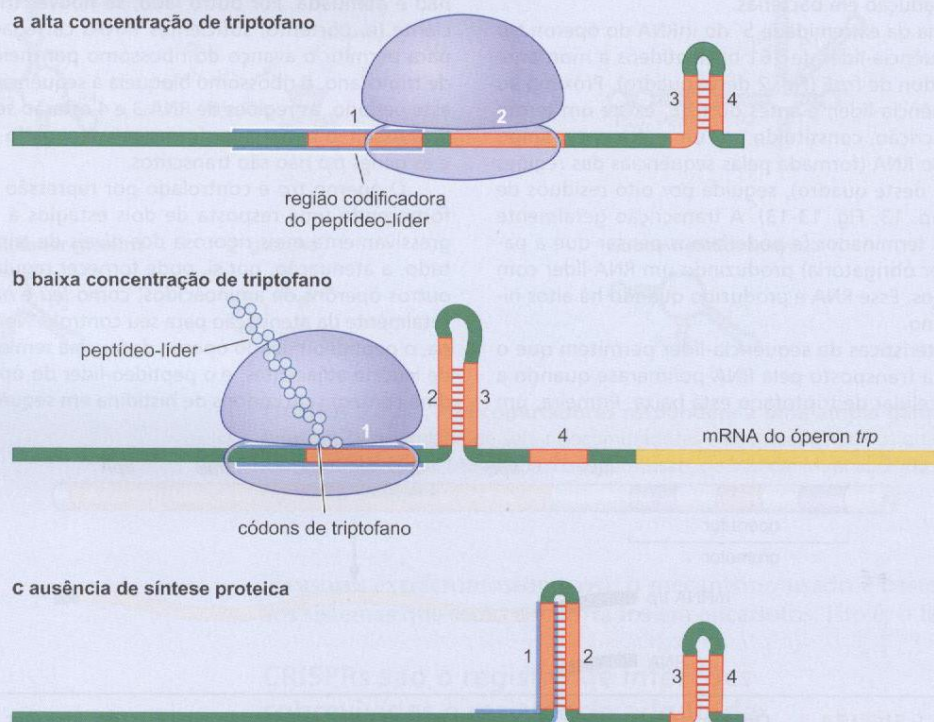
QUADRO 20-1 FIGURA 1 Óperon *trp*. O óperon do triptofano de *E. coli*, mostrando a relação do líder (ver texto) com os genes estruturais que codificam as enzimas do Trp. Os produtos gênicos são a antranilato-sintetase (produto do *trpE*), a fosforribosil-antranilato-transferase (*trpD*), a fosforribosil-antranilato-isomerase-indol-glicerol-fosfato-sintetase (*trpC*), a β -triptofano-sintetase (*trpB*) e a α -triptofano-sintetase (*trpA*).

(continua)

Quadro 20-1 (Continuação)



QUADRO 20-1 FIGURA 2 RNA-líder do operador de *trp*. Características da sequência nucleotídica do RNA-líder do operon *trp*.



QUADRO 20-1 FIGURA 3 Término da transcrição no atenuador de *trp*. O término da transcrição no atenuador do operon *trp* é controlado pela disponibilidade de triptofano. A caixa em roxo mostra a região codificadora do peptídeo-líder. (a) Condições de triptofano elevado: a sequência 3 pode parear com a sequência 4, formando um grampo de término de transcrição. (b) Condições de triptofano reduzido: o ribossomo para nos códon de triptofano adjacentes, deixando a sequência 2 livre para pareamento com a sequência 3, impedindo a formação do grampo de término 3-4. (c) Ausência de síntese proteica: se nenhum ribossomo iniciar a tradução do peptídeo-líder AUG, o grampo é formado pelo pareamento das sequências 1 e 2, impedindo a formação do grampo 2-3 e permitindo a formação do grampo de término com as sequências 3-4. As enzimas Trp não são expressas.

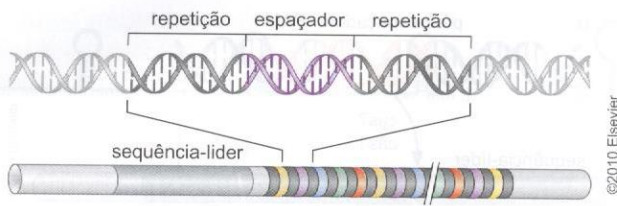


FIGURA 20-6 A organização do *locus* CRISPR. As sequências repetidas conservadas e o espaçador variável estão indicados na parte superior. Na parte inferior da figura, vê-se o arranjo destas sequências (o número delas é extremamente variável); a sequência-líder proximal também é apresentada. (Adaptada, com permissão, de Karginov F.V. e Hannon G.J., 2010. *Mol. Cell* **37**:7–19, Fig. 1A,B, p. 8. © Elsevier.)

espaçadoras de comprimento semelhante mas com sequências altamente divergentes. Em uma das extremidades do arranjo está uma chamada sequência-líder, geralmente rica em A-T e com cerca de 500 pb de comprimento.

Estes agrupamentos não são raros – de fato, as CRISPRs foram encontradas em metade de todos os genomas bacterianos sequenciados, e essencialmente em todos os genomas de arqueobactérias. Em muitos casos, há apenas um agrupamento por genoma, mas não é incomum haver mais, e o número pode chegar a 20 ou mais – e em um caso quase 400 foram detectados em uma espécie de *Chloroflexus*. Mas como eles surgem, e o que eles fazem?

A primeira pista de sua origem veio da curiosa observação (um achado puramente de bioinformática) de que um número significativo de sequências espaçadoras eram idênticas a regiões de fagos ou plasmídeos conhecidos. Isso levou rapidamente à proposta de que estes arranjos estavam envolvidos em algum tipo de mecanismo de defesa contra a entrada de ácidos nucleicos estranhos na célula.

O suporte experimental para este modelo veio quando células de bactérias resistentes que surgiam em populações desafiadas com um dado fago apresentaram sequências espaçadoras incorporadas derivadas do fago. Da mesma maneira, a sensibilidade reduzida a um fago poderia ser conferida ou anulada pela adição ou remoção das sequências espaçadoras relevantes. Além disso, as bactérias tornavam-se mais resistentes à infecção por um dado fago à medida que adquiriam mais sequências espaçadoras deste fago. Mostrou-se também que os vírus recuperam a capacidade de infectar estas células previamente resistentes quando adquirem mutações nestas regiões de seus genomas que doaram – e, portanto, são complementares – às sequências espaçadoras.

Um conjunto de genes codificadores de proteínas conservado está fortemente associado às sequências CRISPR. Os dois membros mais altamente conservados (*cas1* e *cas2*, significando *CRISPR associated* [associados a CRISPR]) são encontrados em todos os casos, mas outros genes *cas*, e genes mais distantemente relacionados, são menos frequentes. Estes genes codificam proteínas envolvidas em diferentes aspectos da função de CRISPR, como será discutido adiante. Um exemplo da complexidade observada em alguns casos reais é apresentado na Figura 20-7.

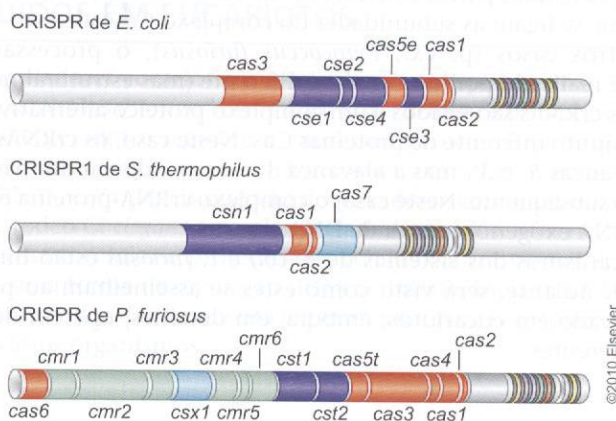
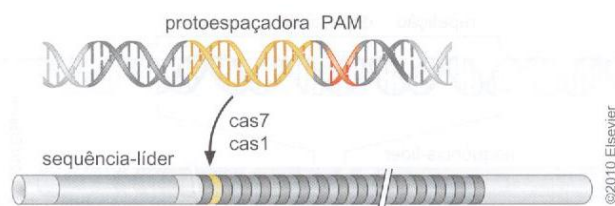


FIGURA 20-7 A organização dos genes *cas* em três *loci* CRISPR. Os diferentes números, orientações e tipos de genes *cas* são mostrados em três *loci* CRISPR bem-estudados. Os genes *cas* centrais são mostrados em vermelho. As sequências repetida e espaçadora mostradas na Figura 20-6 estão aqui na extremidade à direita. (Adaptada, com permissão, de Karginov F.V. e Hannon G.J., 2010. *Mol. Cell* **37**: 7-19, Fig. 1C, p. 8. © Elsevier.)

FIGURA 20-8 Mecanismo de aquisição de sequência espaçadora.

Cada sequência espaçadora nova é inserida próxima à sequência-líder, resultando em um arranjo que é um registro temporal de aquisições passadas. A sequência destinada a tornar-se uma espaçadora é, no genoma do fago, conhecida como "protoespaçadora" e é adjacente a uma sequência PAM, como descrito no texto. (Adaptada, com permissão, de Karginov F.V. e Hannon G.J. 2010. *Mol. Cell* 37: 7-19, Fig. 2B, p. 10. © Elsevier.)

**Sequências espaçadoras são adquiridas de vírus infectantes**

Como já descrito anteriormente, a aquisição de sequências espaçadoras de um dado fago por uma célula confere sensibilidade diminuída a infecções posteriores por este fago. O processo básico está apresentado na Figura 20-8. A sequência do vírus que irá se tornar um novo espaçador é chamada de protoespaçador e é encontrada próxima à sequência PAM (do inglês, *proto-spacer adjacent motif* [motivo adjacente ao protoespaçador]). Quando um novo espaçador é adicionado a um arranjo CRISPR, ele é incorporado à extremidade proximal, próximo à sequência-líder.

Alguns dos genes *cas* codificam proteínas necessárias para o processo de aquisição. Assim, o mecanismo de defesa antiviral não é prejudicado por sua ausência, mas a célula não consegue adquirir resistência a novos vírus. Os produtos dos genes *cas1*, *cas2* e *cas4* estão nesta categoria. Cas1 é uma integrase putativa, enquanto Cas2 é uma ribonuclease. Ao contrário de outras proteínas Cas, a Cas6 está envolvida na expressão e processamento do agrupamento CRISPR e Cas3 está envolvida na interferência à infecção viral.

Uma CRISPR é transcrita como um RNA simples longo, o qual é então processado em espécies de RNA mais curtas que desencadeiam a destruição de DNA ou RNA invasores

A expressão de CRISPR de *E. coli* foi extensamente estudada. O promotor a partir do qual a expressão é iniciada está localizado na região-líder e gera um único transcrito de RNA chamado de pré-crRNA. No caso de *E. coli*, CRISPR está associada a oito genes *cas*, e os produtos de cinco deles formam um complexo chamado Cascade. Este complexo inclui uma subunidade que está implicada no processamento do longo transcrito em **crRNAs** individuais curtos, cada um deles tendo o comprimento de um espaçador e uma sequência repetida. Estes pequenos RNAs permanecem ligados ao complexo Cascade e o direcionam para os genomas de DNA de moléculas invasoras (Fig. 20-9a).

Cada crRNA contém oito nucleotídeos da repetição 5' seguidos pela região espaçadora completa e a maior parte da próxima repetição. As seções de repetição incluídas no crRNA são chamadas de "alavancas" 5' e 3', respectivamente, e representam partes conservadas de cada crRNA, sendo consideradas as regiões que se ligam às subunidades do complexo Cascade.

Em outros casos (p. ex., *Pyrococcus furiosus*), o processamento do pré-crRNA é realizado por uma enzima diferente (mas estruturalmente semelhante), e os crRNAs são ligados a um complexo proteico alternativo formado por um conjunto diferente de proteínas Cas. Neste caso, os crRNAs começam com as alavancas 5' e 3', mas a alavanca 3' é removida em um passo de processamento subsequente. Neste caso, o complexo crRNA-proteína é direcionado contra RNA exógeno, em vez de DNA.

Os mecanismos dos sistemas de *E. coli* e *P. furiosus* estão ilustrados na Figura 20-9. Adiante, será visto como estes se assemelham ao processo de RNAi observado em eucariotos, embora, em detalhes, operem de maneiras bastante diferentes.

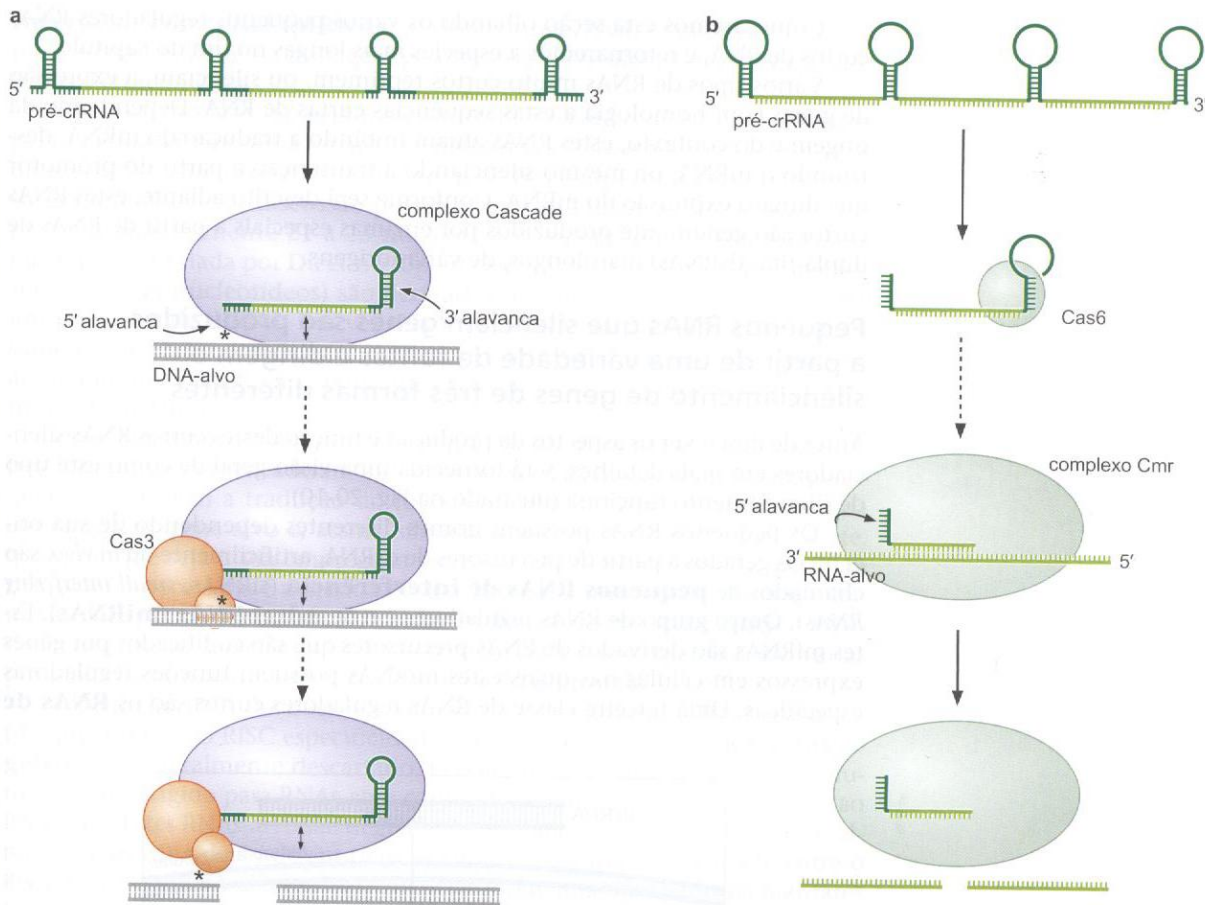


FIGURA 20-9 A operação antiviral dos loci CRISPR de *E. coli* e *P. furiosus*. O sistema de *E. coli* (a) tem o DNA exógeno como alvo, enquanto o de *P. furiosus* (b) tem o RNA como alvo. Embora semelhantes de várias maneiras, o mecanismo de processamento e a operação final dos dois sistemas são diferentes, como destacado no texto. Em a, o pré-crRNA é processado pela subunidade CasE do complexo Cascade (CasE é codificada pelo gene *cse3* na Fig. 20-7). crRNA e Cascade são então dirigidos para, e clivam, o DNA-alvo com auxílio de Cas3 de maneira ainda não totalmente esclarecida. Em b, Cas6 (ver Fig. 20-7) processa o pré-crRNA, e este, unido a um complexo diferente de Cascade, age sobre o RNA viral com um mecanismo análogo ao sistema de RNAi de eucariotos. (Adaptada de Jore M.M. et al. 2012. *Cold Spring Harb. Perspect. Biol.* 4: a003657. © Cold Spring Harbor Laboratory Press.)

OS RNAs REGULADORES ESTÃO AMPLAMENTE DISTRIBUÍDOS EM EUKARIOTOS

Os RNAs reguladores eucarióticos existem em múltiplas formas, caracterizados por seu tamanho ("longo" ou "curto"), sua origem e os mecanismos pelos quais eles são gerados e regulam a expressão gênica. Hoje, se acredita que entre 30 e 70% dos genes em eucariotos complexos sejam regulados até certo ponto por RNAs, com papéis que vão desde o desenvolvimento (bem-estudado no nematódeo *C. elegans* e na planta *Arabidopsis*, descritos no Apêndice 1) até a homeostase celular e proteção das células contra vírus e transposons. Além disso, uma forma de regulação (RNA de interferência) foi adaptada para uso como uma poderosa ferramenta experimental para manipular a expressão gênica em vários organismos.

Começaremos esta seção olhando os vários pequenos reguladores RNAs curtos de RNA, e retornaremos a espécies mais longas no fim do capítulo.

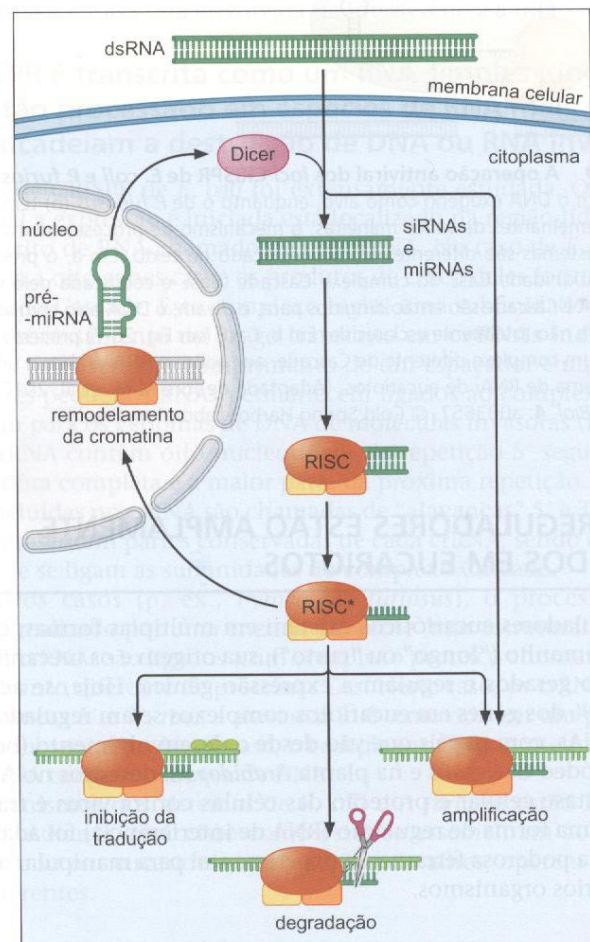
Vários tipos de RNAs muito curtos reprimem, ou silenciam, a expressão de genes com homologia a estas sequências curtas de RNA. Dependendo da origem e do contexto, estes RNAs atuam inibindo a tradução do mRNA, destruindo o mRNA, ou mesmo silenciando a transcrição a partir do promotor que dirige a expressão do mRNA. Conforme será descrito adiante, estes RNAs curtos são geralmente produzidos por enzimas especiais a partir de RNAs de dupla-fita (dsRNAs) mais longos, de várias origens.

Pequenos RNAs que silenciam genes são produzidos a partir de uma variedade de fontes e dirigem o silenciamento de genes de três formas diferentes

Antes de descrever os aspectos da produção e função destes curtos RNAs silenciadores em mais detalhes, será fornecida uma visão geral de como este tipo de silenciamento funciona (ilustrado na Fig. 20-10).

Os pequenos RNAs possuem nomes diferentes dependendo de sua origem. Os gerados a partir de precursores de dsRNA, artificialmente ou *in vivo*, são chamados de **pequenos RNAs de interferência** (siRNAs, *small interfering RNAs*). Outro grupo de RNAs reguladores é o de **microRNAs (miRNAs)**. Estes miRNAs são derivados de RNAs precursores que são codificados por genes expressos em células nas quais estes miRNAs possuem funções reguladoras específicas. Uma terceira classe de RNAs reguladores curtos são os **RNAs de**

FIGURA 20-10 Geração de siRNAs e miRNAs, e seu modo de ação. O processamento de dsRNA para gerar siRNAs e pré-miRNAs para a produção de miRNAs pela enzima Dicer. Outra enzima envolvida apenas na geração de pré-miRNAs – Drosha – não está mostrada aqui, mas é descrita adiante. siRNAs e miRNAs direcionam um complexo chamado RISC (do inglês, *RNA-induced silencing complex* [complexo de silenciamento induzido por RNA]), para reprimir os genes de três modos. Este complexo ataca e digere mRNA com homologia ao siRNA; interfere na tradução desses mRNAs; ou direciona as enzimas modificadoras de cromatina para os promotores que promovem a expressão desses mRNAs (Fig. 20-18). Ao recrutar uma RNA-polimerase RNA-dependente, os siRNAs podem gerar mais dsRNA como substrato para Dicer produzir mais siRNA. Esta é a etapa de “amplificação” mostrada à direita e em detalhes na Figura 20-11. (Adaptada, com permissão, de Hannon G.J. 2002. *Nature* 418: 244-251, Fig. 5. © Macmillan.)



interação com piwi (piRNAs, piwi-interaction RNAs), que são expressos predominantemente na linhagem germinativa e possuem características distintas das de miRNAs.

siRNAs e miRNAs são gerados a partir de moléculas de RNA mais longas pela enzima **Dicer**, uma enzima semelhante à RNase III que reconhece e digere dsRNAs mais longos ou estruturas de haste-alça formadas por precursores de miRNA (ver discussão posterior). Os produtos de siRNA e miRNA possuem normalmente 21 a 23 nucleotídeos; sua produção está mostrada na etapa estimulada por Dicer no topo da Figura 20-10. Os piRNAs (que possuem 24 a 34 nucleotídeos) são derivados de uma forma que não envolve um precursor de dsRNA. Em vez disso, os piRNAs são gerados pelo processamento de transcritos longos de fita simples que abrangem os chamados agrupamentos de piRNA encontrados no genoma. Este processamento não necessita de Dicer.

Estes pequenos RNAs inibem a expressão de genes-alvo homólogos de três formas: eles desencadeiam a destruição do mRNA codificado pelo gene-alvo, eles inibem a tradução do mRNA ou eles induzem modificações da cromatina no gene-alvo e, assim, silenciam sua transcrição. Notavelmente, qualquer que seja a via empregada em determinado caso, grande parte da maquinaria necessária é a mesma. Esta maquinaria inclui um complexo chamado **complexo de silenciamento induzido pelo RNA (RISC, RNA-induced silencing complex)**. Um RISC contém, além do pequeno RNA, várias proteínas incluindo um membro da família **Argonauta**.

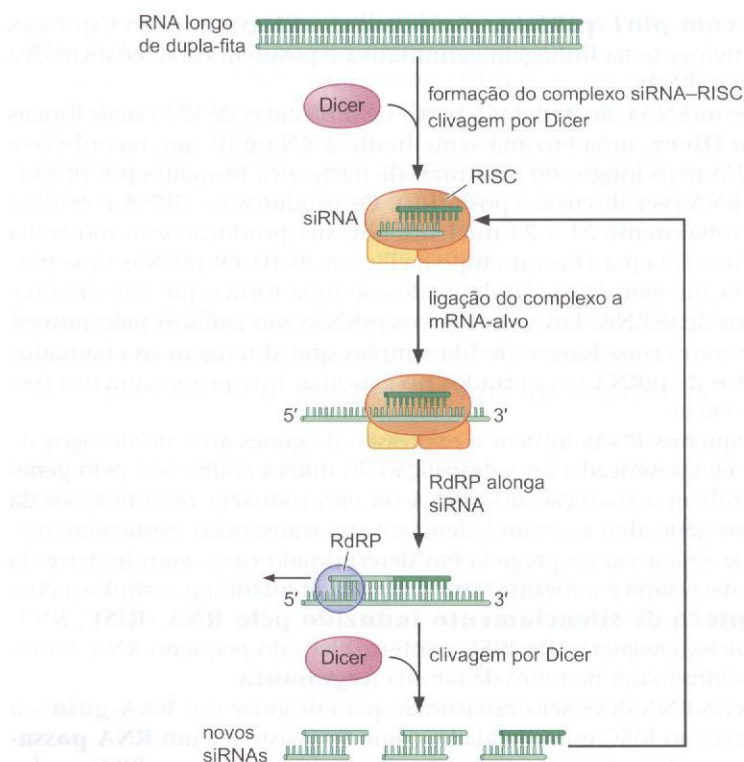
O pequeno RNA deve ser desnaturado para originar um **RNA-guia** – a fita que fornece ao RISC especificidade, como será visto – e um **RNA passageiro**, que é geralmente descartado. O complexo resultante, o RISC maduro, é então dirigido para RNAs-alvo contendo sequências complementares ao RNA-guia. Estes RNAs-alvo são degradados ou sua tradução é inibida. Normalmente, a escolha depende em parte do grau de complementaridade entre o RNA-guia e o mRNA-alvo: se as sequências forem altamente complementares (como geralmente ocorre com os siRNAs), o alvo é degradado; se a correspondência não for tão boa (i.e., se houver vários malpareamentos de bases, como geralmente ocorre com os miRNAs), a resposta é, na maior parte das vezes, a inibição da tradução. Nos casos em que o RNA-alvo é degradado, a Argonauta é a subunidade catalítica que realiza a clivagem inicial do mRNA; por este motivo, a Argonauta é frequentemente chamada de **Slicer** (“fatiador”) e a clivagem do mRNA, de **slicing** (“fatiamento”).

Um RISC pode também ser direcionado para o núcleo, onde recruta outras proteínas que modificam a cromatina em torno do promotor do gene complementar ao RNA-guia (mostrado à esquerda na Fig. 20-10). Essa modificação resulta no silenciamento da transcrição (Cap. 19). O estabelecimento de silenciamento nas regiões centroméricas da levedura *Schizosaccharomyces pombe*, por exemplo, necessita da maquinaria de RNAi, conforme será visto adiante.

Uma distinção que vale fazer entre miRNAs e siRNAs é que os primeiros atuam como reguladores em *trans* tradicionais: eles são codificados por um gene mas atuam sobre outros (como os sRNAs encontrados em sistemas bacterianos). Ao contrário, os siRNAs são normalmente gerados por transcritos das regiões nas quais atuam (formalmente como os RNAs antissenso descritos em bactérias) e são, portanto, descritos como atuando em *cis*.

Outra característica importante no silenciamento por RNAi é a sua extrema eficiência. Quantidades muito pequenas de dsRNA são suficientes para induzir um quase completo desligamento dos genes-alvo. Um fator adicional à eficiência poderia ser a atuação de uma **RNA-polimerase RNA-dependente (RdRP, RNA-dependent RNA polymerase)**, uma enzima adicional necessária em muitos casos de RNAi, incluindo o silenciamento centromérico em leveduras. Esta polimerase pode amplificar o sinal inibidor: a RdRP gera dsRNA após ser recrutada para o mRNA pelo siRNA original (como indicado

FIGURA 20-11 Amplificação do sinal de siRNA por RdRP. Como mostrado na Figura 20-10, o sinal de siRNA pode ser amplificado, gerando mais dsRNA para que Dicer processe em mais siRNAs. Isso é alcançado pelo recrutamento de uma RNA-polimerase dependente de RNA, pelo complexo siRNA-RISC, para o RNA-alvo, e siRNA atua como iniciador para que esta enzima transforme o alvo em dsRNA, que pode então, ele mesmo, ser processado por Dicer. As RdRPs são encontradas em plantas, vermes e na levedura *Schizosaccharomyces pombe* (*Saccharomyces cerevisiae* não possui a maquinaria de RNAi), e a importância desta etapa de amplificação será vista no caso do silenciamento do centrômero de *S. pombe* (Fig. 20-18).



à direita na Fig. 20-10 e mostrado em detalhes na Fig. 20-11). Este processo de retroalimentação gera grandes quantidades de siRNA. A RdRP ainda não foi identificada em células de mamíferos, e sua alta eficiência provavelmente resulta do fato de que o fatiamento (ou *slicing*) é catalítico – isto é, cada RISC pode clivar vários mRNAs.

Assim, embora na primeira seção deste capítulo tenham sido vistos exemplos de pequenos RNAs regulando a expressão gênica em bactérias, os mecanismos de produção e atuação destes RNAs em eucariotos é bastante diferente.

SÍNTESE E FUNÇÃO DAS MOLÉCULAS DE miRNA

Os miRNAs possuem uma estrutura característica que ajuda na sua identificação e na de seus genes-alvo

Conforme mencionado anteriormente, os miRNAs estão codificados no genoma como segmentos de transcritos mais longos. Sua estrutura característica ajuda na sua identificação e na predição dos genes-alvo que eles podem regular.

A forma funcional de um miRNA é normalmente entre 21 ou 22 nucleotídeos (ela pode variar de 19 a 25 nucleotídeos). Estes RNAs curtos são gerados por duas reações de clivagem a partir de um transcrito de RNA mais longo (chamado de pri-miRNA, de “primário”) que possui uma estrutura secundária em formato de grampo. A primeira clivagem libera a haste-alça, chamada de pré-miRNA; a segunda gera o miRNA maduro a partir do pré-miRNA. Um dos primeiros e mais bem caracterizados miRNAs a ser identificado é o *let-7*, que regula o desenvolvimento na transição de larva para adulto no verme *C. elegans* (ver Apêndice 1). As estruturas dos pré-miRNAs para *let-7* e alguns outros miRNAs naturais estão representadas na Figura 20-12.

Inicialmente, pensava-se que um “braço” da estrutura de haste-alça do pré-miRNA seria o miRNA regulador. Porém, numerosos exemplos fo-

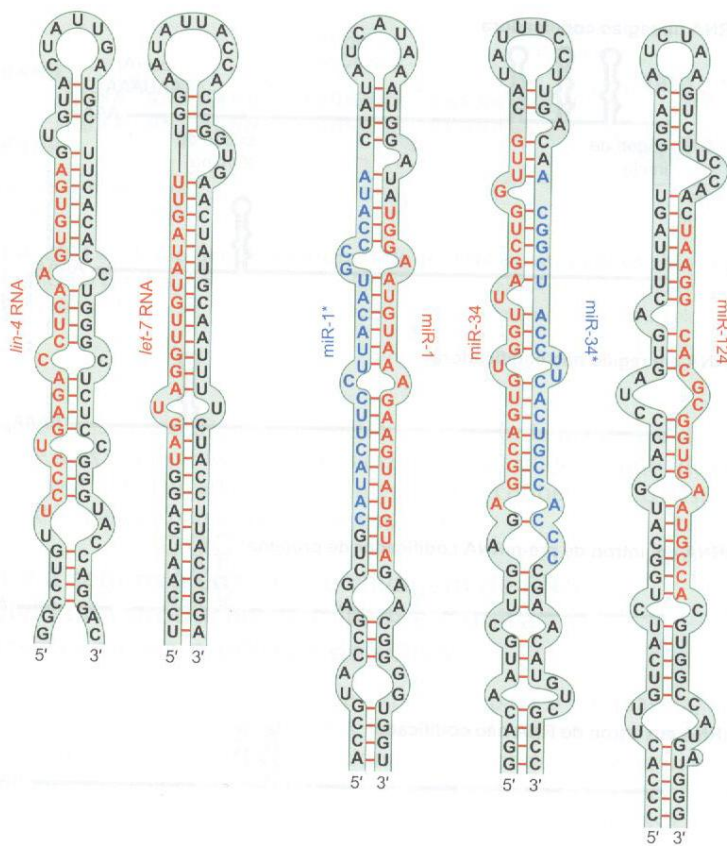


FIGURA 20-12 Estrutura de alguns pré-miRNAs antes do processamento para gerar os miRNAs maduros. As sequências em vermelho são miRNAs. Em alguns casos, ambos os "braços" de uma haste-alça podem gerar um miRNA funcional. Nestes casos, o segundo miRNA é mostrado em azul – por exemplo, miR-1 (vermelho) e miR-1* (azul), bem como em miR-34 (vermelho) e miR-34* (azul). Os miRNAs mostrados são todos de verme. *lin-4* e *let-7* foram geneticamente identificados; os chamados de miR foram encontrados por bioinformática. (Modificada, com permissão, de Lim L.P. et al. 2003. *Genes Dev.* 17: 991, Fig. 6. © Cold Spring Harbor Laboratory Press.)

ram identificados nos quais ambos os braços da estrutura dão origem aos miRNAs funcionais, cada um deles com seu próprio conjunto de genes-alvo (nestes casos, os dois miRNAs são vermelho e azul na Fig. 20-12). Atualmente, considera-se que a produção de miRNAs a partir de ambos os braços é algo comum. Os pré-miRNAs podem ser codificados por qualquer parte de um transcrito: ou seja, eles podem estar em regiões codificadoras, em regiões-líder, ou em íntrons (Fig. 20-13).

A estrutura secundária característica de um transcrito primário portando um miRNA (pri-miRNA) tornou possível prever sua presença com base no dobramento calculado da estrutura secundária da sequência de RNA. Além disso, em muitos casos, candidatos para os genes-alvo regulados podem também ser preditos, porque o silenciamento depende da complementaridade de sequência entre o alvo e o miRNA maduro. O pareamento de bases entre miRNA e o RNA-alvo é iniciado por interações dos chamados resíduos-semente – em geral, a sequência entre as bases 2 e 9 do miRNA de 22 nucleotídeos. Esta é a região de maior complementaridade e, assim, é a região mais útil na identificação de genes-alvo. Obviamente, estabelecer que um miRNA realmente existe requer que sua presença seja detectada em células (p. ex., por *northern blot*) e que a expressão gênica a partir dos mRNAs-alvo seja afetada por sua presença.

As duas reações de clivagem necessárias para gerar o miRNA a partir destes transcritos primários são mediadas por duas RNases distintas. Uma delas é Dicer, que já introduzimos e que também é necessária na geração de siRNAs. A outra, especificamente necessária para o processamento de miRNA, é **Drosha**. Uma característica destas enzimas é que elas reconhecem e clivam RNAs com base na estrutura de seus substratos, em vez de sua sequência específica. Agora a atenção será voltada para como estas enzimas funcionam.

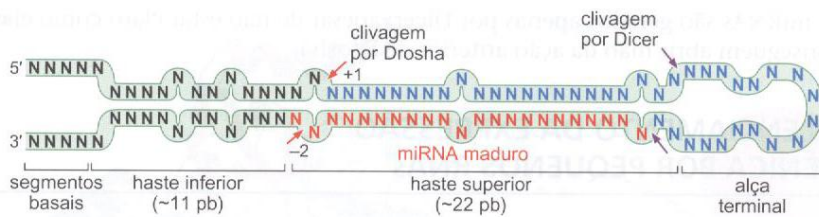


FIGURA 20-14 Visão geral da estrutura do pri-miRNA mostrando os sítios de clivagem de Dicer e Drosha. A região em vermelho torna-se o miRNA maduro. Observa-se que os segmentos basais devem ser de fita simples para o reconhecimento adequado pelo complexo Drosha.

da RNase III são específicas para dsRNA e o clivam de maneira a deixar um trecho de dois nucleotídeos livres nas extremidades 3' do produto de dsRNA. Este trecho 3' não pareado é importante para o reconhecimento desta molécula de RNA pela próxima enzima da via, Dicer.

Dicer é a segunda enzima de clivagem de RNA envolvida na produção de miRNA e a única necessária para a produção de siRNA

O pré-miRNA liberado por Drosha é exportado para o citoplasma, onde a segunda reação de clivagem do RNA, realizada por Dicer, acontece. Assim como ocorre com Drosha, Dicer seleciona seus sítios de clivagem usando um mecanismo de medida, em vez de um mecanismo sequência-específico. Uma estrutura de alta resolução de Dicer fornece uma ideia de como isso provavelmente acontece.

Dicer é construída por três módulos: dois domínios de RNase III e um domínio de ligação a dsRNA chamado de domínio PAZ (que tem este nome devido a três proteínas que contêm este domínio: Piwi, Argonauta e Zwillie).

A Figura 20-16a mostra um desenho com a organização da proteína Dicer e como se acredita que ela interaja com a molécula de dsRNA. No painel inferior, Figura 20-16b, é apresentada a estrutura de Dicer, modelada com um substrato de RNA. A proteína possui o formato de uma machadinha. O domínio PAZ está na parte inferior do cabo, onde forma um bolso de ligação para a extremidade 3' do substrato de dsRNA. O cabo da machadinha é formado por um domínio de conexão e contém uma superfície de ligação positivamente carregada para a molécula de RNA. A região superior, a "lâmina", compreende os dois domínios de RNase, arranjados em um dímero simétrico. Cada domínio de RNase carrega um sítio ativo e é responsável pela clivagem de uma das duas fitas do substrato de RNA. Portanto, Dicer irá atuar em qualquer dsRNA, independentemente de sua sequência, e irá clivar esta molécula a 22 nucleotídeos de sua extremidade. O domínio PAZ ancora a extremidade 3' terminal do substrato de RNA para posicionar os sítios ativos da enzima a cerca de 22 nucleotídeos de distância como se fosse uma régua (ver Fig. 20-16). De fato, a ocorrência de domínios PAZ de diferentes tamanhos está correlacionada com os diferentes tamanhos dos produtos de Dicer encontrados em diferentes organismos.

Como visto, apenas os miRNAs são produzidos a partir de grandes precursores em formato de grampo. Ao contrário, o RNA precursor para a via de siRNA é um dsRNA mais longo. Como consequência deste substrato inicial diferente, Drosha não é necessária para a geração de siRNAs. A clivagem por Dicer ainda é necessária, no entanto, e gera novamente um RNA com 21 a 22 nucleotídeos adequado para incorporação no RISC. Em plantas, até mesmo

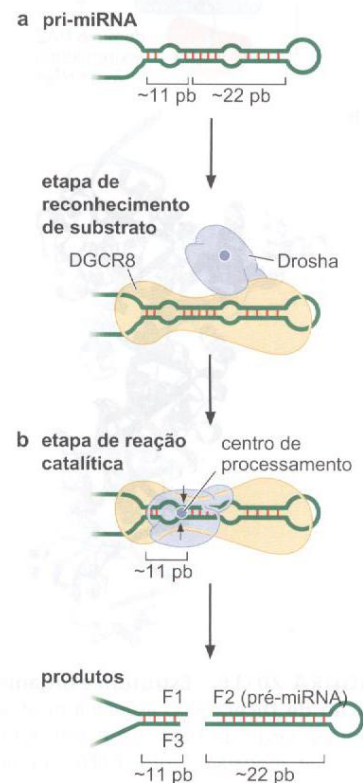


FIGURA 20-15 Reconhecimento e clivagem de pri-miRNA pelo complexo Microprocessador. Três fragmentos são gerados por clivagem, marcados como F1, F2 (pré-miRNA) e F3.

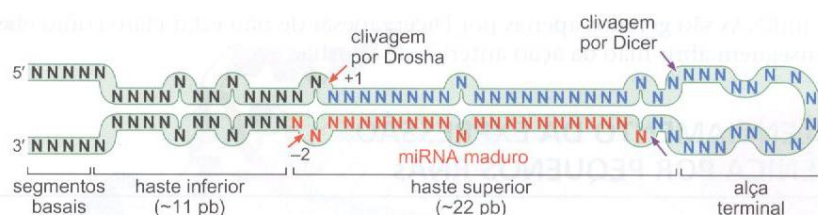


FIGURA 20-14 Visão geral da estrutura do pri-miRNA mostrando os sítios de clivagem de Dicer e Drosha. A região em vermelho torna-se o miRNA maduro. Observa-se que os segmentos basais devem ser de fita simples para o reconhecimento adequado pelo complexo Drosha.

da RNase III são específicas para dsRNA e o clivam de maneira a deixar um trecho de dois nucleotídeos livres nas extremidades 3' do produto de dsRNA. Este trecho 3' não pareado é importante para o reconhecimento desta molécula de RNA pela próxima enzima da via, Dicer.

Dicer é a segunda enzima de clivagem de RNA envolvida na produção de miRNA e a única necessária para a produção de siRNA

O pré-miRNA liberado por Drosha é exportado para o citoplasma, onde a segunda reação de clivagem do RNA, realizada por Dicer, acontece. Assim como ocorre com Drosha, Dicer seleciona seus sítios de clivagem usando um mecanismo de medida, em vez de um mecanismo sequência-específico. Uma estrutura de alta resolução de Dicer fornece uma ideia de como isso provavelmente acontece.

Dicer é construída por três módulos: dois domínios de RNase III e um domínio de ligação a dsRNA chamado de domínio PAZ (que tem este nome devido a três proteínas que contêm este domínio: Piwi, Argonauta e Zwiille).

A Figura 20-16a mostra um desenho com a organização da proteína Dicer e como se acredita que ela interaja com a molécula de dsRNA. No painel inferior, Figura 20-16b, é apresentada a estrutura de Dicer, modelada com um substrato de RNA. A proteína possui o formato de uma machadinha. O domínio PAZ está na parte inferior do cabo, onde forma um bolso de ligação para a extremidade 3' do substrato de dsRNA. O cabo da machadinha é formado por um domínio de conexão e contém uma superfície de ligação positivamente carregada para a molécula de RNA. A região superior, a "lâmina", compreende os dois domínios de RNase, arranjados em um dímero simétrico. Cada domínio de RNase carrega um sítio ativo e é responsável pela clivagem de uma das duas fitas do substrato de RNA. Portanto, Dicer irá atuar em qualquer dsRNA, independentemente de sua sequência, e irá clivar esta molécula a 22 nucleotídeos de sua extremidade. O domínio PAZ ancora a extremidade 3' terminal do substrato de RNA para posicionar os sítios ativos da enzima a cerca de 22 nucleotídeos de distância como se fosse uma régua (ver Fig. 20-16). De fato, a ocorrência de domínios PAZ de diferentes tamanhos está correlacionada com os diferentes tamanhos dos produtos de Dicer encontrados em diferentes organismos.

Como visto, apenas os miRNAs são produzidos a partir de grandes precursores em formato de grampo. Ao contrário, o RNA precursor para a via de siRNA é um dsRNA mais longo. Como consequência deste substrato inicial diferente, Drosha não é necessária para a geração de siRNAs. A clivagem por Dicer ainda é necessária, no entanto, e gera novamente um RNA com 21 a 22 nucleotídeos adequado para incorporação no RISC. Em plantas, até mesmo

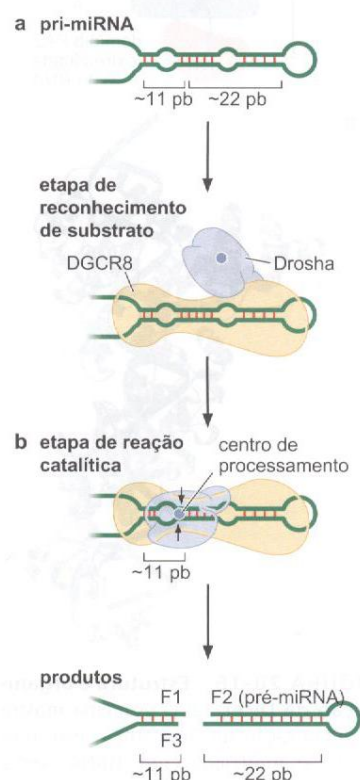


FIGURA 20-15 Reconhecimento e clivagem de pri-miRNA pelo complexo Microprocessador. Três fragmentos são gerados por clivagem, marcados como F1, F2 (pré-miRNA) e F3.

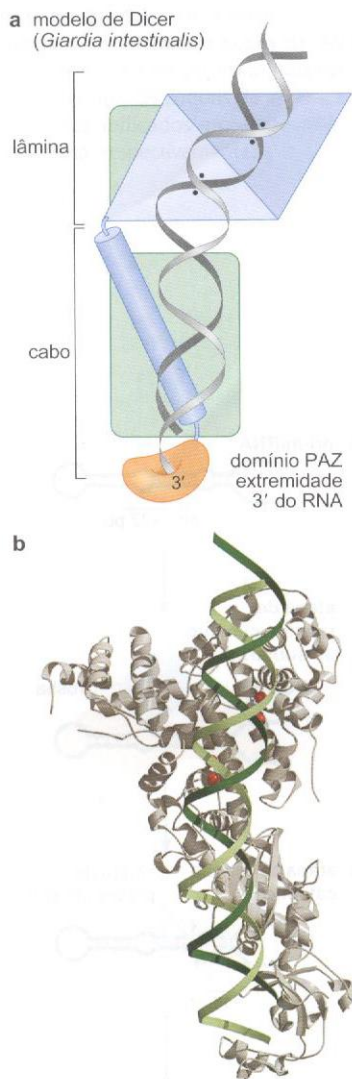


FIGURA 20-16 Estrutura e organização de Dicer. (a) O esquema mostra a organização de Dicer. (b) A estrutura de Dicer modelada com dsRNA revela como o comprimento é medido. A proteína é mostrada em cinza, com os sítios ativos da nuclease indicados por esferas vermelhas (e como pontos pretos na parte a). O RNA está em verde. A estrutura mostrada contém apenas os domínios de RNase III e PAZ. A proteína Dicer também contém ATPase e outros domínios. (b, MacRae I.J. et al. 2006. *Science* **311**: 195-198. Código PDB: 2FFL; observa-se que o RNA foi modelado na estrutura e não era parte da estrutura cristalográfica.) Imagem preparada com MolScript, BobScrip e Raster3D.

os miRNAs são gerados apenas por Dicer; apesar de não estar claro como elas conseguem abrir mão da ação anterior de Drosha.

SILENCIAMENTO DA EXPRESSÃO GÊNICA POR PEQUENOS RNAs

Até o momento, vimos como os pequenos RNAs são gerados a partir de precursores de RNA de dupla-fita ou miRNA. Agora, a atenção é voltada para como estes pequenos RNAs silenciam a expressão de seus genes-alvo.

A incorporação de uma fita de RNA-guia no RISC torna o complexo maduro, pronto para silenciar a expressão gênica

A ação de Dicer gera a pequena molécula de RNA que irá determinar em que RNA-alvo irá atuar. A forma ativa do RNA regulador é de fita simples – neste estágio, chamado de **RNA-guia** – incorporado no complexo proteico RISC. Neste complexo, a fita de RNA-guia recruta RISC para um RNA-alvo. Argumenta-se que o comprimento de 22 nucleotídeos é longo o suficiente para especificar uma única sequência-alvo nos grandes genomas de eucariotos complexos usando o pareamento de bases RNA-RNA.

O componente central de RISC é uma proteína chamada **Argonauta**, que é, em muitos casos, uma enzima de clivagem de RNA. O mecanismo mais bem compreendido de silenciamento gênico é a clivagem (ou fatiamento [*slicing*]) do mRNA-alvo mediado por RISC. Entretanto, muitos organismos possuem múltiplos membros da família de proteínas Argonauta. Por exemplo, há oito Argonautas diferentes em seres humanos, mas nem todas elas, quando incorporadas em um complexo RISC, possuem atividade de fatiador. RISCs que contêm outras Argonautas devem silenciar a expressão gênica usando mecanismos independentes de fatiamento, como a repressão da tradução. Os piRNAs encontrados anteriormente, que não são gerados por Dicer, ainda assim se ligam a uma proteína relacionada a Argonauta, chamada PIWI, em um complexo semelhante ao RISC.

A geração do RISC ativo e o *slicing*, sob a orientação de um siRNA, ocorre da seguinte maneira. O dsRNA curto gerado pela Dicer é incorporado ao RISC, onde é desnaturado para fornecer as fitas-guia e passageira (que é descartada). O RISC resultante – chamado de RISC maduro – com seu RNA-guia de fita simples está agora pronto para reconhecer e fatiar o mRNA-alvo.

Como visto com Dicer, a estrutura de uma proteína Argonauta fornece um quadro para a compreensão do mecanismo de reconhecimento de RNA-alvo e clivagem por RISC (Fig. 20-17). Como Dicer, Argonauta possui um domínio PAZ e um domínio de RNase. O domínio PAZ reconhece especificamente a extremidade 3' do RNA-guia. O RNA-guia ligado é pareado ao RNA-alvo, e a arquitetura do complexo é tal que esta ligação posiciona o sítio ativo do domínio de RNase de maneira adequada para clivar a fita do RNA-alvo. A clivagem ocorre quase no meio do dúplex de RNA-guia–RNA-alvo, entre o 10º e o 11º nucleotídeos a partir da extremidade 5' do RNA-guia.

Como já foi mencionado, em alguns casos, o RISC maduro pode inibir a tradução em vez de fatiar o mRNA e, de fato, é assim que a maioria dos miRNAs parece atuar.

O mecanismo desta repressão traducional ainda está sob investigação, e há grande debate sobre a ordem dos eventos. Assim, embora a tradução seja certamente inibida, o mRNA também sofre declínio. Portanto, é difícil provar qual é a causa e qual é a consequência: se houver decaimento ou degradação do mRNA, não haverá, obviamente, tradução; da mesma maneira, quando a tradução de um mRNA é inibida (por qualquer mecanismo), a célula tende a destruir

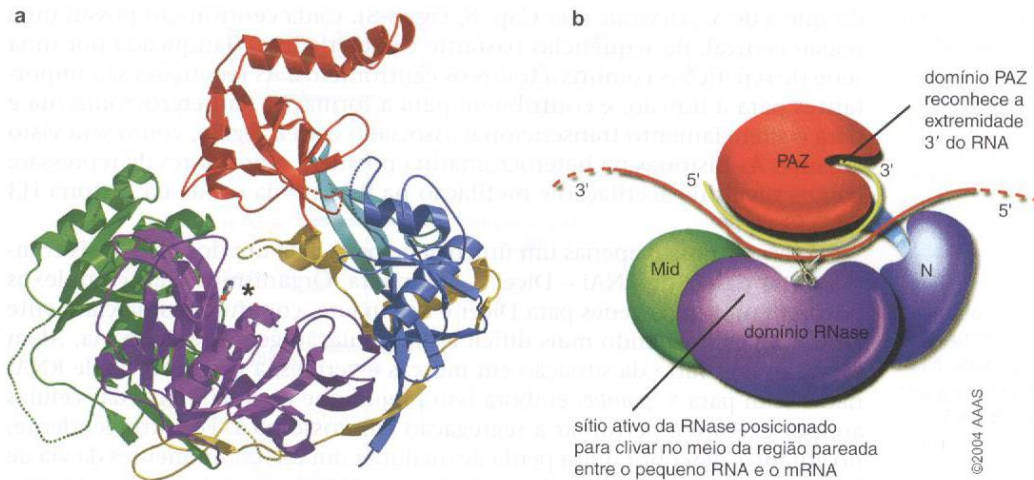


FIGURA 20-17 Estrutura de Argonauta, mostrando as regiões de ligação ao RNA e um domínio semelhante à nuclease RNase H. (a) Estrutura cristalográfica de Argonauta. Os domínios foram coloridos como na parte b, com o domínio em azul sendo a parte amino-terminal da proteína, e o domínio verde, no meio. (b) Ilustração dos domínios de Argonauta. A seta mostra o sítio ativo da RNase posicionado para clivar no meio da região pareada entre o pequeno RNA e mRNA. (a,b, Adaptadas, com permissão, de Song J.J. et al 2004. *Science* 305: 1434-1437, Fig. 4C. Código PDB: 1u04. © AAAS.) Imagens preparadas com MolScript, BobScript e Raster3D.

esta molécula de mRNA. De qualquer forma, isso é diferente do mecanismo de fatiamento ativo desencadeado pelos siRNAs, descrito anteriormente.

O início da tradução é um mecanismo elaborado envolvendo uma série de fatores (ver Cap. 15) e permitindo muitas oportunidades para interferência. Qualquer que seja o mecanismo de inibição da tradução, parece que os miRNAs levam, em alguns casos, ao sequestro do mRNA nos chamados corpúsculos de processamento (corpúsculos P) no citoplasma, onde a tradução é reprimida.

Pequenos RNAs podem silenciar genes em nível transcricional pela coordenação de modificações na cromatina

Vimos agora como miRNAs e siRNAs podem silenciar genes pela inibição da tradução de mRNAs-alvo ou pela sua destruição. RNAs reguladores também podem atuar em nível de transcrição, desligando a expressão de genes-alvo pela modificação de histonas no promotor. Este mecanismo foi mais extensivamente estudado no silenciamento centromérico da levedura *S. pombe*.

No Capítulo 19, observou-se que, em leveduras, os genes posicionados em certas regiões do genoma são normalmente silenciados. No caso descrito em detalhes naquele capítulo, genes posicionados próximos aos telômeros em *Saccharomyces cerevisiae* foram silenciados. Genes no locus de tipo acasalante desta levedura e de *S. pombe* também são silenciados. Em *S. pombe*, os centrômeros são outra região silenciada do genoma. Em ambos os organismos, o silenciamento envolve modificações de histonas. Mas ao contrário dos casos de silenciamento em *S. cerevisiae*, desprovida de uma maquinaria de RNAi, o silenciamento centromérico em *S. pombe* necessita desta via.

Os centrômeros de *S. pombe* possuem uma organização de sequência mais semelhante à de eucariotos superiores (p. ex., moscas e seres humanos)

do que à de *S. cerevisiae* (ver Cap. 8, Fig. 8-8). Cada centrômero possui uma região central, de sequências bastante características, flanqueada por uma série de repetições comuns a todos os centrômeros. As repetições são importantes para a função, e contribuem para a formação da heterocromatina e para o silenciamento transcricional associado com a região, como será visto adiante. As histonas na heterocromatina possuem marcadores de repressão: baixos níveis de acetilação e metilação na lisina-9 da cauda da histona H3 (H3K9).

S. pombe possui apenas um único gene para cada um dos principais componentes da via de RNAi – Dicer e Argonauta. Organismos mais complexos possuem múltiplos genes para Dicer e Argonauta, com funções parcialmente redundantes, tornando mais difícil a manipulação genética desta via. Além disso, ao contrário da situação em moscas e vermes, a perda da via de RNAi não é letal para *S. pombe*, embora isso prejudique o crescimento das células ao, por exemplo, perturbar a segregação cromossômica. Foi surpreendente, no entanto, descobrir que a perda de qualquer um dos componentes da via de RNAi levava à perda da metilação da histona H3K9 e à perda do silenciamento gênico nos centrômeros, sobretudo porque este silenciamento era reconhecidamente transcricional. Até esta descoberta, pensava-se que o RNAi atuava apenas pós-transcricionalmente.

O ponto-chave para entender este silenciamento transcricional parece ser as próprias repetições centroméricas; estes elementos de sequência são transcritos a partir de ambas as fitas pela RNA-polimerase II, produzindo transcritos complementares que podem hibridizar para formar dsRNAs – processo que é amplificado por RdRP (como se viu na Fig. 20-11). Os RNAs, por sua vez, sofrem a ação da maquinaria de RNAi para gerar siRNAs que, de alguma forma – que ainda permanece desconhecida – dirigem um complexo semelhante ao RISC contendo Argonauta (chamado de complexo de silenciamento transcricional induzido por RNA [RITS, *RNA-induced transcriptional silencing*]) para os centrômeros. Os siRNAs poderiam, teoricamente, fazer isso pelo reconhecimento do DNA nos centrômeros, por meio do pareamento de bases sequência-específico diretamente com o molde de DNA. Todavia, mais prováveis são modelos nos quais os siRNAs recrutam RITS para transcritos presos ao centrômero pela RNA-polimerase II. O recrutamento resulta no fatiamento dos transcritos centroméricos, que, por sua vez, é necessário para alastrar o aparato de modificação de histonas ao longo do centrômero (Fig. 20-18). Portanto, a própria transcrição pode alastrar o silenciamento quando os transcritos são alvos do RNAi.

Como foi mencionado anteriormente, os *loci* de tipo acasalante de *S. pombe* são também transcricionalmente silenciados, e aqui o silenciamento não é perdido em linhagens mutantes defectivas para RNAi. Acredita-se que o RNAi atue também neste caso, mas apenas no estabelecimento inicial do estado de silenciamento – ele não é necessário para a manutenção do silenciamento uma vez que este estiver estabelecido. Outros mecanismos baseados em proteínas mantêm o estado reprimido – assim como o fazem em *S. cerevisiae* (Cap. 19). Acredita-se também que o RNAi possua um papel no silenciamento da heterocromatina de outros organismos, abrangendo desde moscas até plantas. O silenciamento da transcrição indesejada de transposons também parece ser mediado por RNA, como descrevemos a seguir.

Uma série fascinante de observações e experimentos levou ao conhecimento atual sobre pequenos RNAs reguladores em eucariotos. Estes iniciaram no fim da década de 1980 com os resultados aparentemente incompreensíveis de tentativas de superexpressar genes de pigmentos em petúnias (para torná-las roxo mais profundo, mas resultando em flores brancas). Depois disso, houve a descoberta dos genes reguladores de vermes cujos produtos mostraram ser miRNAs e, então, experimentos mostrando que a introdução de dsRNAs em vermes silenciava a expressão de genes complementares. Esta história é descrita no Quadro 20-2, Descoberta de miRNAs e RNAi.

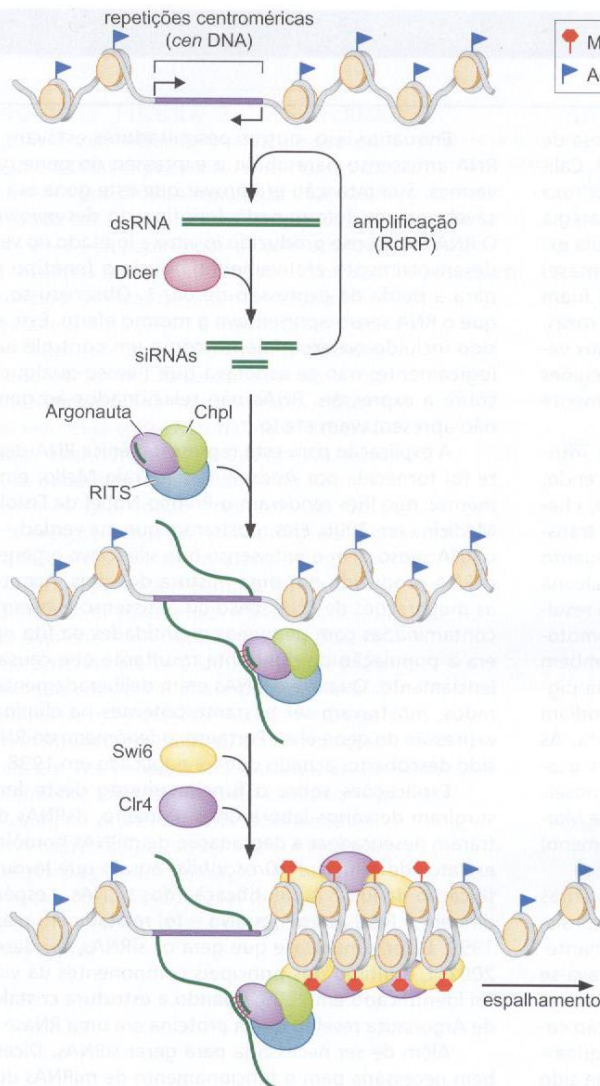


FIGURA 20-18 Modelo para o recrutamento de RITS e o silenciamento dos centrômeros. No topo, estão mostrados os nucleossomos em torno das sequências repetidas em um centrômero de *S. pombe*. As sequências repetidas (cen DNA) são transcritas pela RNA-polimerase II, gerando dsRNA que é substrato para Dicer. Os siRNAs produzidos desta maneira são carregados no complexo RITS contendo ArgonAUTA. Como mostrado no centro, o complexo RITS contendo siRNA é recrutado para os transcritos ligados à Pol II sendo gerados pela transcrição continuada das repetições centroméricas, por meio da complementaridade entre o siRNA e o transcrito. Então, este complexo recruta alguns outros complexos: RDRC, que permite a produção de mais dsRNA por RdRP (ver Fig. 20-11), e Ctr4 e Swi6, que modificam localmente os nucleossomos adicionando marcas de silenciamento H3K9. Outra subunidade de RITS, Chp1, contém um cromodomínio (Cap. 8, Fig. 8-41), que, ao ligar-se aos nucleossomos metilados, provavelmente estabiliza a ligação de RITS. Embora não representado na figura, o “fatiamento” dos transcritos por ArgonAUTA (no interior do RITS) gera substratos de RNA para RdRP, que sintetiza uma fita complementar e gera mais substrato para Dicer. Este processo é necessário para a modificação do nucleossomo – e, assim, a região silenciada seja espalhada. (Redesenhada, com permissão, de Martienssen R. e Moazed D. 2007. *Epigenetics* [ed. D. Allis et al.], p. 157, Fig. 4. © Cold Spring Harbor Laboratory Press.)

O RNAi é um mecanismo de defesa que protege contra vírus e transposons

A maquinaria de RNAi está bastante difundida em eucariotos, embora não seja onipresente. Ela não ocorre em *S. cerevisiae*, por exemplo, como acabamos de ver. Acredita-se, no entanto, que pelo menos o sistema básico existia no ancestral comum mais recente a todos os eucariotos mas foi subsequentemente perdido em algumas linhagens.

Mas o que faz o RNAi, biologicamente? Existem os miRNAs, é claro – e a maquinaria de RNAi é necessária para produzir e usar estes reguladores – mas alguns organismos possuem a maquinaria de RNAi e nenhum miRNA (incluindo *S. pombe*). De fato, acredita-se que os miRNAs evoluíram para tirar proveito da existência da maquinaria de RNAi em vez de serem a razão pela qual a maquinaria existe. Uma função antiga que a maquinaria de RNAi pode ter tido (e ainda tem) é a proteção dos organismos contra transposons e vírus.

Anteriormente, foi descrito o sistema CRISPR-cas em bactérias (ver Fig. 20-9). Neste caso, sequências de DNA derivadas de material exógeno (fago ou plasmídeos) são acumuladas em regiões do genoma que podem ser

▶ EXPERIMENTOS-CHAVE

Quadro 20-2 Descoberta de miRNAs e RNAi

Em 1989, Richard Jorgensen, que trabalhava na empresa de biotecnologia Advanced Genetic Sciences em Oakland, Califórnia, estava tentando gerar petúnias com flores de cor roxa mais profunda do que as linhagens existentes. A estratégia parecia simples: ele iria introduzir nas plantas uma cópia extra do gene de pigmento (que codifica a chalcona sintase) sob o controle de um promotor forte. Estas plantas iriam produzir mais chalcona sintase e as flores seriam mais roxas. O que ele conseguiu, na verdade, foi plantas com graus variados de flores mais pálidas, muitas malhadas – com regiões roxas e brancas – e até mesmo algumas completamente brancas (Fig. 1 deste quadro).

Embora decepcionantes, estes resultados foram intrigantes. Na tentativa de entender o que estava acontecendo, Jorgensen revelou várias características do fenômeno, chamado de **cosupressão** (porque tanto a expressão do transgene quanto a do gene endógeno era reprimida). Quanto maior a expressão do transgene, menor o nível de chalcona sintase; isso era válido quando a expressão aumentada resultava de múltiplas cópias do transgene ou do uso de promotores mais fortes dirigindo o transgene. Observou-se também que algumas plantas possuíam padrões variegados de pigmentação, e que diferentes padrões de variação podiam ser encontrados em diferentes flores na mesma planta. Às vezes, estes padrões eram herdados, mas, em outras ocasiões, eram aparentemente alterados ao acaso. Estas observações sugeriram a Jorgensen e outros (particularmente Marjori Matzke, que também estava investigando este fenômeno) que eles estavam lidando com um fenômeno epigenético.

Outros pesquisadores estavam tentando tornar plantas resistentes à infecção viral. Para isso, uma abordagem consistia em superexpressar em plantas um derivado dominante-negativo de um fator de replicação viral comum: esperava-se que esta proteína fosse bloquear a replicação de qualquer vírus infectante que usasse este mecanismo de replicação comum. Embora o produto viral dominante-negativo bloqueasse a replicação do vírus da batata a partir do qual havia sido derivado, sua especificidade de ação era fortemente restrita ao vírus. Demonstrou-se também que a própria proteína não era necessária – apenas o RNA.



QUADRO 20-2 FIGURA 1 Flor da petúnia. Um exemplo dos efeitos da superexpressão do gene do pigmento chalcona sintase no que seria, no caso de expressão em nível normal, uma flor de petúnia completamente roxa. (Cortesia de Richard A. Jorgensen, University of Arizona.)

Enquanto isso, outros pesquisadores estavam usando RNA antissenso para abolir a expressão do gene *par-1* em vermes. Sua intenção era provar que este gene era responsável por um determinado fenótipo do desenvolvimento. O RNA antissenso produzido *in vitro* e injetado no verme em desenvolvimento efetivamente induzia o fenótipo previsto para a perda da expressão de *par-1*. Observou-se, porém, que o RNA *senso* apresentava o mesmo efeito. Este só havia sido incluído no experimento como um controle negativo, logicamente; não se esperava que tivesse qualquer efeito sobre a expressão. RNAs não relacionados ao gene *par-1* não apresentavam efeito.

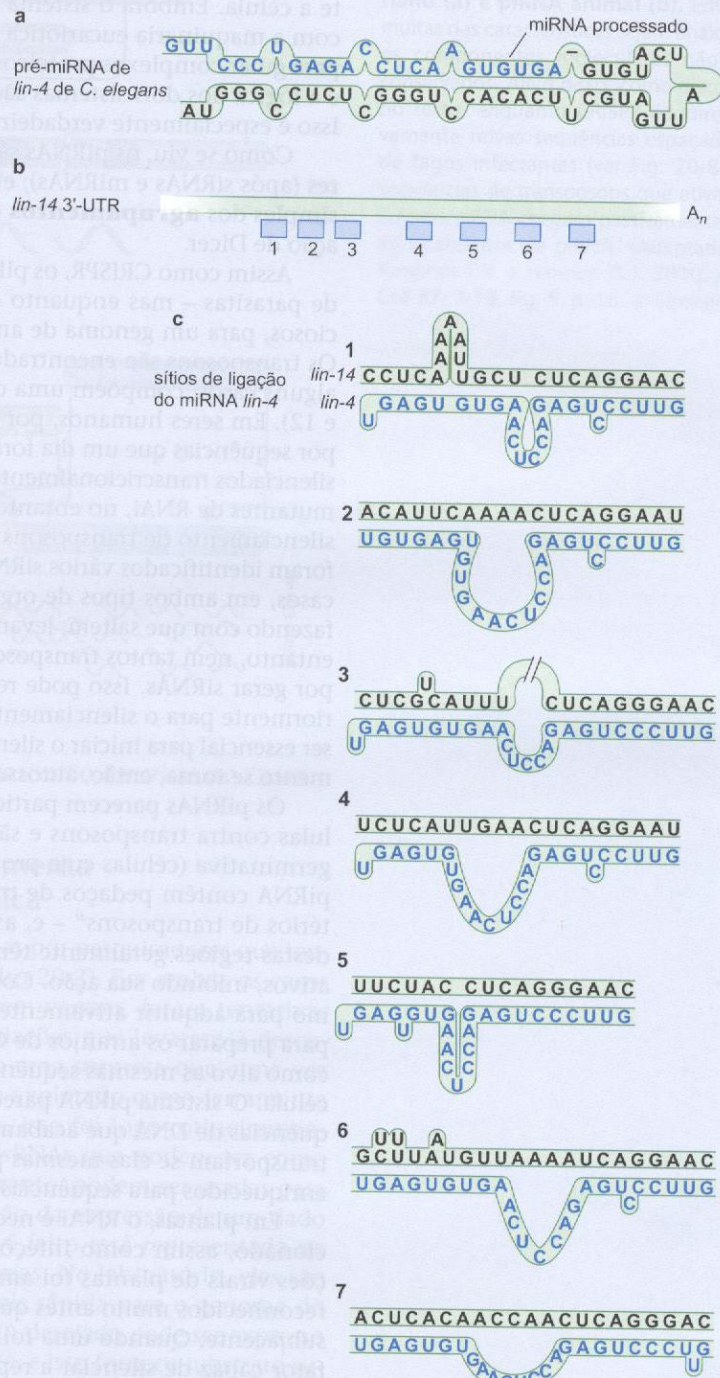
A explicação para esta repressão gênica RNA-dependente foi fornecida por Andrew Fire e Craig Mello, em experimentos que lhes renderam o Prêmio Nobel de Fisiologia ou Medicina em 2006. Eles mostraram que, na verdade, não era o RNA *senso* nem o antissenso que silenciava o gene – era o dsRNA produzido por uma mistura dos dois. Acontece que as preparações de RNA *senso* ou antissenso estavam ambas contaminadas com pequenas quantidades da fita oposta, e era a população de dupla-fita resultante que causava o silenciamento. Quando dsRNAs eram deliberadamente preparados, mostravam ser bastante potentes na eliminação da expressão do gene-alvo. Portanto, o fenômeno do RNAi havia sido descoberto, achado que foi publicado em 1998.

Explicações sobre o funcionamento deste fenômeno surgiram de vários laboratórios. Primeiro, dsRNAs demonstraram desencadear a degradação de mRNAs homólogos em extratos de células de *Drosophila*, ensaio que levou à identificação de RISC. A identificação dos siRNAs – espécies que dirigem o RISC aos genes-alvo – foi relatada em plantas em 1999. Dicer, a nuclease que gera os siRNAs, foi descrita em 2001. E o último dos principais componentes da via, Slicer, foi identificado em 2005, quando a estrutura cristalográfica de Argonauta revelou que a proteína era uma RNase.

Além de ser necessária para gerar siRNAs, Dicer é também necessária para o funcionamento de miRNAs durante o desenvolvimento. O primeiro miRNA e seu alvo haviam sido descritos em 1993, por Victor Ambros e Gary Ruvkun, respectivamente. Na época, esta observação foi vista como um resultado correto, mas, ainda assim, estranho; o gene *lin-4* codificava um pequeno RNA que atuava em um gene-alvo, *lin-14*, por complementaridade de sequência entre o miRNA e as regiões da 3'-UTR dos genes-alvo (Fig. 2 deste quadro). Subsequentemente, outros miRNAs foram encontrados em vermes, alguns deles com homologia a genes semelhantes em animais e plantas, sugerindo que este mecanismo de regulação era mais disseminado. Assim, surgiu o quadro de um mundo de pequenos RNAs envolvidos na regulação gênica – alguns fornecidos de maneira exógena, outros gerados como parte dos programas de regulação gênica usados durante o desenvolvimento. O campo desenvolveu-se muito rapidamente, como as datas neste relato revelam, passando de um fenômeno obscuro a um Prêmio Nobel e exigindo seu próprio capítulo em livros-texto, em apenas 15 anos. O progresso acelerado talvez tenha sido, em grande parte, devido à gama de espécies (leveduras, plantas e vermes) estudadas e às abordagens (genética, bioquímica, estudos estruturais e bioinformática) utilizadas.

Quadro 20-2 (Continuação)

QUADRO 20-2 FIGURA 2 O microRNA *lin-4* liga-se à extremidade 3'-UTR de seu gene-alvo, *lin-14*. (a) O pré-miRNA *lin-4* antes do processamento por Dicer. A sequência do miRNA está mostrada em azul. (b) As sete sequências da 3'-UTR de *lin-4* que podem parear com o mRNA de *lin-4* em vários níveis, como mostrado na parte c. A biologia por trás dos eventos reguladores controlados por *lin-4/lin-14* é bastante intrigante. Serão vistos (no Cap. 21) exemplos de como a expressão gênica regula as decisões ao longo do desenvolvimento. Exemplos famosos incluem os genes Hox, cuja expressão define a identidade espacial; assim, membros e outras estruturas formam-se onde deveriam ao longo do eixo corporal, e não no local errado – por exemplo, uma perna não cresce na cabeça. Isso, porém, é o que de fato pode acontecer em mutantes desprovidos de certos genes Hox – como o famoso mutante homeótico de *Drosophila*, antenapédia, em que uma pata cresce na cabeça no lugar de uma antena. A proteína *lin-14* é um exemplo de um regulador que controla a identidade diferencial temporal. Portanto, mutações nos chamados genes heterocrônicos resultam em transformações temporais em vez de espaciais. Ou seja, as células geralmente adotam destinos específicos a elas em estágios iniciais ou tardios do desenvolvimento. A expressão da proteína *lin-14* é, obviamente, regulada pelo microRNA *lin-4*, como descrito no texto. (Modificada de Ha I. 1996. *Genes Dev.* 10: 3041-3050, Fig. 1. © Cold Spring Harbor Laboratory Press.)



subsequentemente expressas na forma de pequenas moléculas de RNA que destroem ácidos nucleicos homólogos caso estes venham a invadir novamente a célula. Embora o sistema CRISPR não possua componentes em comum com a maquinaria eucariótica de RNAi (além do fato de usar pequenos RNAs para guiar complexos proteicos para destruir ácidos nucleicos alvos), a função e a lógica dos dois sistemas são notavelmente parecidas de muitas maneiras. Isso é especialmente verdadeiro para o sistema de piRNA.

Como se viu, os piRNAs são a terceira classe de pequenos RNAs reguladores (após siRNAs e miRNAs); eles surgem a partir de longos transcritos de fita simples dos **agrupamentos de piRNAs** no genoma, sem a necessidade da ação de Dicer.

Assim como CRISPR, os piRNAs parecem ter como alvo os ácidos nucleicos de parasitas – mas enquanto os principais alvos de CRISPR são fagos infecciosos, para um genoma de animal, as principais ameaças são **transposons**. Os transposons são encontrados essencialmente em todos os eucariotos e, em alguns casos, compõem uma quantidade substancial do genoma (ver Caps. 8 e 12). Em seres humanos, por exemplo, cerca de 45% do genoma é composto por sequências que um dia foram transposons. Os transposons são geralmente silenciados transcricionalmente e empacotados na heterocromatina. Em alguns mutantes de RNAi, no entanto, as modificações de histonas associadas com o silenciamento de transposons são perdidas. Além disso, em plantas e vermes, foram identificados vários siRNAs que correspondem a transposons. Em alguns casos, em ambos tipos de organismos, a perda de RNAi reativa transposons, fazendo com que saltem, levando a altos níveis de mutagenese espontânea. No entanto, nem tantos transposons são reativados como quanto são conhecidos por gerar siRNAs. Isso pode refletir uma situação semelhante à descrita anteriormente para o silenciamento do tipo acasalante em *S. pombe*: o RNAi pode ser essencial para iniciar o silenciamento de alguns transposons, mas o silenciamento se torna, então, autossustentável sem necessidade posterior de siRNAs.

Os piRNAs parecem particularmente dedicados à tarefa de proteger as células contra transposons e são predominantemente expressos na linhagem germinativa (células cuja proteção é mais importante). Os agrupamentos de piRNA contêm pedaços de transposons – eles foram descritos como “cemitérios de transposons” – e, assim, os piRNAs gerados a partir de transcritos destas regiões geralmente têm como alvo transcritos feitos por transposons ativos, inibindo sua ação. Como visto, o sistema CRISPR inclui um mecanismo para adquirir ativamente amostras de DNA exógeno invasor e usar isso para preparar os arranjos de CRISPR para produzir pequenos RNAs que terão como alvo as mesmas sequências caso elas venham a aparecer novamente na célula. O sistema piRNA parece ser desprovido desta característica, mas as sequências de DNA que acabam em agrupamentos de piRNA presumivelmente transportam-se elas mesmas para estas regiões, assim os agrupamentos serão enriquecidos para sequências de transposons (ver Fig. 20-19).

Em plantas, o RNAi é necessário para controlar transposons, como mencionado, assim como infecções virais. O efeito protetor do RNAi em infecções virais de plantas foi amplamente observado. De fato, os efeitos foram reconhecidos muito antes que o RNAi fosse reconhecido como o mecanismo subjacente. Quando uma folha de uma planta é infectada por um vírus, um fator capaz de silenciar a replicação do vírus é sistematicamente espalhado ao longo de toda a planta. Este fator não protege a folha originalmente infectada, mas impede que a infecção se espalhe. Em plantas mutadas nos genes de Argonauta ou Dicer, a infecção espalha-se livremente e a replicação viral é muito maior. O sinal protetor inclui siRNAs gerados a partir do próprio genoma viral. Os vírus revidaram: eles geralmente carregam genes cuja função é proteger o vírus infectante contra o RNAi do hospedeiro. Um exemplo disso é o HcPro do vírus Y da batata, que atua reduzindo a produção ou a estabilidade dos siRNAs. Outros produtos virais afetam outras etapas no mecanismo de defesa, incluindo o espalhamento sistêmico de siRNAs.

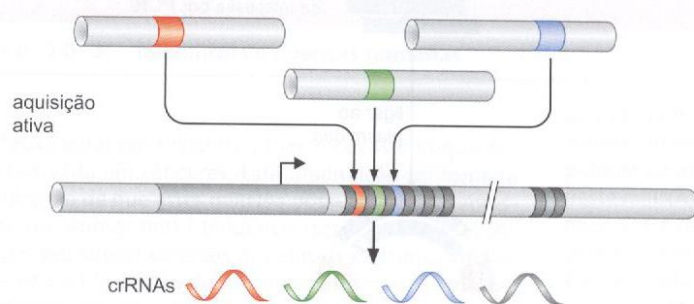
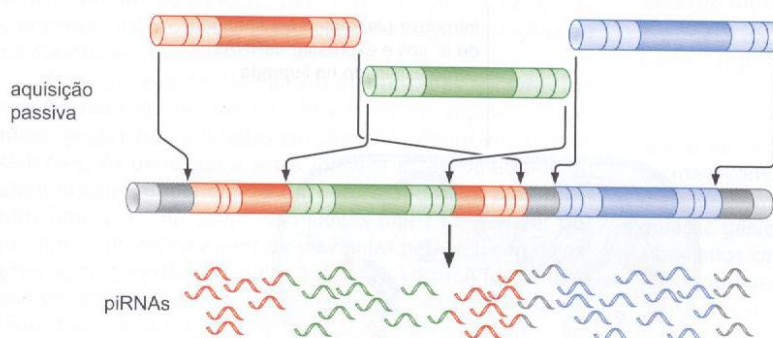
a sistema CRISPR**b sistema piRNA**

FIGURA 20-19 Comparação entre os sistemas de defesa CRISPR bacteriano (a) e piRNA animal (b). Embora muitas das características sejam análogas, os componentes moleculares não são conservados. Além disso, como discutido no texto, enquanto CRISPR adquire ativamente novas sequências espaçadoras de fagos infectantes (ver Fig. 20-8), as sequências de transposons que ativam o sistema piRNA chegam passivamente aos agrupamentos de piRNA. (Adaptada de Karginov F.V. e Hannon G.J. 2010. *Mol. Cell* **37**: 7-19, Fig. 5, p. 16. © Elsevier.)

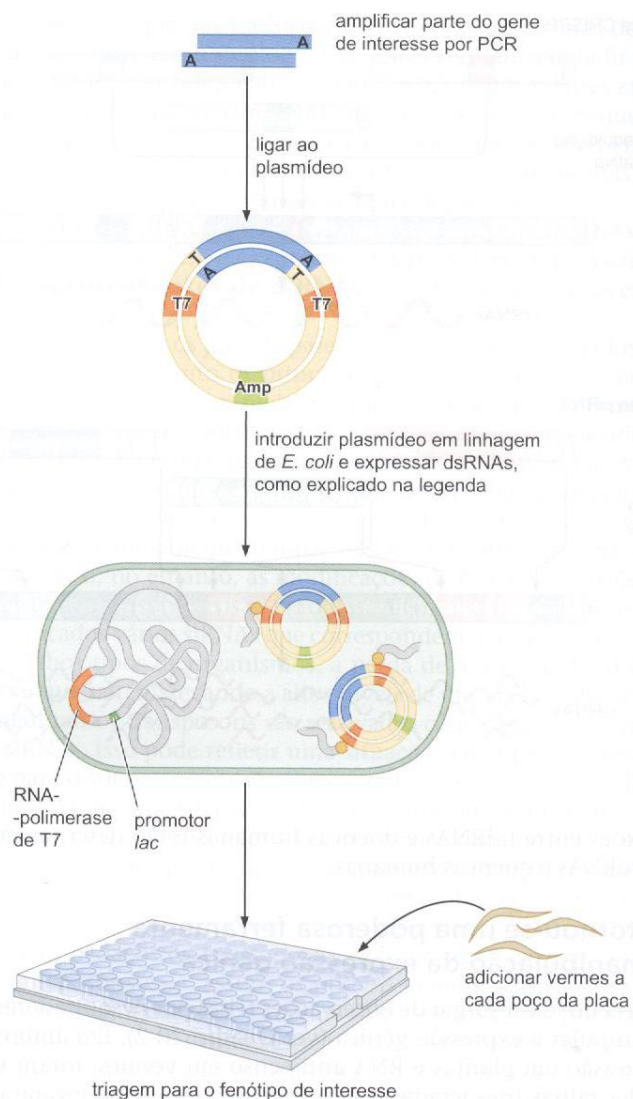
Conexões entre miRNAs e doenças humanas estão descritas no Quadro 20-3, microRNAs e doenças humanas.

O RNAi tornou-se uma poderosa ferramenta para a manipulação da expressão gênica

A descoberta do RNAi surgiu de observações feitas por pesquisadores que tentavam manipular a expressão gênica (ver Quadro 20-2). Em ambos os casos de cossupressão em plantas e RNA antissenso em vermes, foram tentativas de entender falhas inesperadas nestas manipulações que levaram à descoberta do RNAi. Portanto, talvez não tenha sido uma surpresa que, uma vez compreendido, o RNAi tenha sido rapidamente explorado como ferramenta para manipular a expressão gênica. Em vermes, isso foi logo rotineiramente implementado. Bibliotecas que codificam dsRNAs que podem ter como alvo qualquer gene do genoma do verme e, portanto, podem ser usadas para triar vermes quanto às consequências da inibição da expressão de um dado gene qualquer. A maneira geral pela qual isso é feito está representada na Figura 20-20. Os vermes alimentam-se de bactérias. No laboratório, eles são alimentados com *E. coli*, e ocorre que a rota mais rápida para o genoma do verme é através de seu estômago: qualquer dsRNA desejado pode ser expresso em células de *E. coli* das quais o verme se alimenta, e isso fornece substrato suficiente para que a resposta de RNAi seja desencadeada nas células do verme, desligando genes homólogos ao dsRNA original.

Seria, logicamente, de grande interesse fazer uma triagem para genes desta mesma forma em células de mamíferos, em que as triagens genéticas tradicionais não são viáveis. Estabeleceu-se que siRNAs sintetizados artificialmente, produzidos *in vitro* e introduzidos em células de mamíferos em cultivo, desencadeiam uma resposta de RNAi e reduzem a expressão de genes-alvo adequados, mas a eficiência da transfecção (a entrada dos RNAs nas células) é baixa. Moléculas de dsRNA mais longas também são problemáticas porque

FIGURA 20-20 A interferência por RNA pode ser induzida em vermes ao alimentá-los com bactérias que expressam dsRNAs. Ver Capítulo 7 para detalhes da manipulação molecular necessária nas primeiras etapas deste esquema. A expressão do dsRNA a partir do plasmídeo está sob o controle de dois promotores, em orientações opostas, reconhecidos por uma única subunidade da RNA-polimerase de um fago chamado T7. O gene para esta polimerase é artificialmente expresso nas células usadas neste esquema, sob o controle do promotor *lac* (Cap. 18). Assim, a produção do dsRNA pode ser controlada usando um indutor do promotor *lac*.



desencadeiam uma resposta que desliga toda a tradução na célula, resposta evoluída para bloquear a replicação viral, já que muitos vírus possuem genoma de RNA.

Em vez de tentar introduzir o dsRNA nas células, os pesquisadores acharam mais produtivo mimetizar miRNAs. Com este objetivo, foram geradas bibliotecas nas quais genes curtos são sintetizados como oligonucleotídeos e clonados em plasmídeos. Cada gene curto é projetado para gerar um transcrito que irá se enovelar em uma haste-alça. Estas são processadas por Dicer na célula, gerando um siRNA que irá dirigir o silenciamento de seus genes-alvo. Estes genes curtos sintéticos são chamados de genes de **grampos curtos de RNA (shRNAs, short hairpin RNA)**. Usando um shRNA adequadamente projetado, qualquer gene individual do genoma pode ser um alvo. Ou, com uma biblioteca adequada, uma triagem genética pode ser realizada. Em uma biblioteca assim, por exemplo, cada plasmídeo iria codificar um shRNA dirigido contra um gene diferente. A biblioteca inteira é transfectada nas células de maneira que cada célula receba um shRNA diferente. Células com um fenótipo em particular são escolhidas, e o gene cuja repressão levou a este fenótipo pode ser identificado.

CONEXÕES CLÍNICAS

Quadro 20-3 microRNAs e doenças humanas

Câncer

Uma redução geral dos níveis de vários miRNAs é frequentemente observada em cânceres. Esta diminuição foi tomada como indicação de que estes miRNAs possuem efeito supressor de tumor. Apesar desta tendência geral, outros miRNAs específicos são superexpressos em alguns cânceres. Análogos a genes codificadores de proteínas implicados em câncer, os miRNAs em questão são descritos como supressores de tumor (se sua ausência aumentar a incidência de câncer) ou oncogênicos (se sua expressão aumentada levar ao desenvolvimento de câncer). Seus alvos tendem a ser genes envolvidos na progressão do ciclo celular (proliferação) ou na apoptose.

Entre as centenas de miRNAs identificados em seres humanos, mais de metade está localizada em regiões do genoma geralmente alteradas em cânceres. Assim, em muitos cânceres, os genes para estes miRNAs estão deletados ou amplificados, dependendo da natureza do rearranjo cromossômico. Sendo assim, os miRNAs *miR-15* e *miR-16*, por exemplo, induzem a apoptose de células pela redução da expressão do gene *BCL2* (ver Fig. 1 deste quadro). A forma mais comum de leucemia adulta no hemisfério oeste é a leucemia linfocítica crônica (LLC), doença associada com deleções em uma região do cromossomo 13 (13q14). Esta região do genoma contém os genes dos miRNAs *miR-15a* e *miR-16a*; de fato, estes são os dois únicos genes incluídos nas menores deleções associadas a LLC. Assim, quando estes genes estão deletados, a apoptose é reduzida e os tumores podem surgir e desenvolver-se mais facilmente.

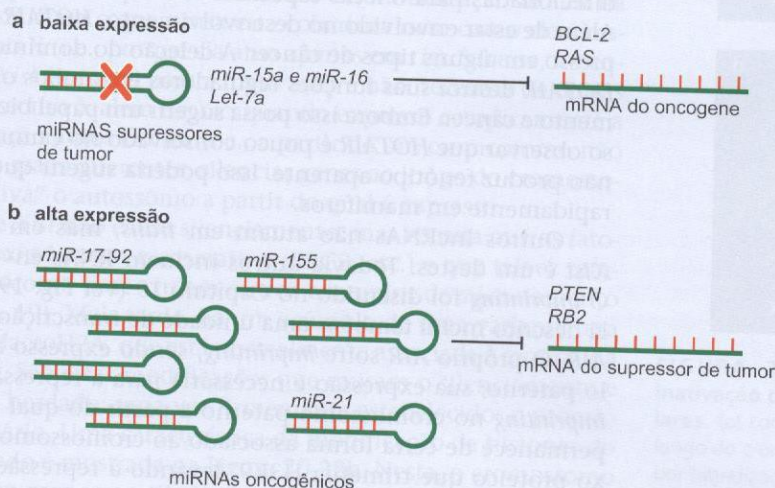
Em outra região do cromossomo 13 (13q31) está o *miR-17-92*, um miRNA oncogênico. Comparada ao tecido normal, a expressão deste gene está significativamente au-

mentada em vários cânceres, incluindo câncer de pulmão, sobretudo em suas formas mais agressivas (p. ex., câncer de pulmão de células pequenas). Além disso, a superexpressão deste miRNA em camundongos transgênicos induz tumorigênese. Entre os vários alvos preditos para *miR-17-92* estão dois genes supressores de tumor, *PTEN* e *RB2*. Um alvo definido é o regulador da progressão do ciclo celular, *E2F1*. Estes e outros exemplos de miRNAs em câncer estão mostrados na Figura 1 deste quadro.

Retardo mental associado ao X frágil

Por meio da análise bioquímica do complexo RISC, várias proteínas a ele associadas foram identificadas. Uma delas é a proteína do retardo mental associado ao X frágil (FMRP, *Fragile X mental retardation protein*). O gene que codifica esta proteína (*FMR1*) é ligado ao X, e sua mutação é a causa da mais comum das formas hereditárias de retardo mental. A FMRP é uma proteína de ligação ao RNA envolvida na regulação gênica; sabe-se que ela interage com alguns miRNAs associados com neuroplasticidade. Pacientes desprovidos de FMRP possuem uma gama de defeitos de desenvolvimento, bem como o retardo mental, devido à perturbação da expressão gênica.

Drosophila possui uma homóloga de FMRP. Em moscas deficientes para este gene, foram observadas conexões sinápticas incomuns entre neurônios e músculos. Uma das proteínas Argonauta de *Drosophila* foi associada à FMRP, enquanto estudos separados de Argonauta também descobriram que FMRP estava ligada a este componente da maquinaria de RNAi. Achados semelhantes foram relatados também em células humanas, indicando uma conexão intrigante entre a condição de X frágil e o amadurecimento e função de miRNA.



QUADRO 20-3 FIGURA 1 miRNAs como supressores de tumor ou oncogênicos. (a) Neste modelo, um miRNA que normalmente reduz a expressão de um oncogene pode atuar como gene supressor de tumor. A perda de função do miRNA por mutação ou deleção, por exemplo, pode resultar em expressão anormal do oncogene-alvo, o que contribuiria para a formação de tumor. (b) Aqui, a amplificação ou a superexpressão de um miRNA que reduz a expressão de um supressor de tumor ou outros genes importantes envolvidos na diferenciação podem contribuir para a formação de tumor por estimulação da proliferação, da angiogênese e da invasão. (Redesenhada, com permissão, de Garzon R. et al. 2006. *Trends Mol. Med.* 12: 580-587, Fig. 2. © Elsevier.)

RNAs LONGOS NÃO CODIFICADORES E INATIVAÇÃO DO X

Os RNAs longos não codificadores exercem vários papéis na regulação gênica, incluindo os efeitos em *cis* e *trans* da transcrição

Nos últimos anos, técnicas de sequenciamento de alto rendimento (descritas no Cap. 7) revelaram a presença de vários RNAs não codificadores expressos em células animais e vegetais. Até o momento, foram discutidos RNAs curtos, mas há também uma classe de RNAs não codificadores com mais de 200 nucleotídeos, conhecida como **RNAs longos não codificadores (lncRNAs, long non-coding RNAs)**. Estes possuem vários papéis em processos que vão desde a tradução e o processamento até a regulação transcricional. O caso mais bem estudado deste último é o RNA de *Xist* envolvido no processo de inativação do X em mamíferos, que discutiremos a seguir. Mas antes disso, serão abordados outros dois lncRNAs reguladores com papéis no desenvolvimento.

Existe grande debate sobre exatamente o quão prevalente são os lncRNAs na célula. Como recém-mencionado, a maioria deles é detectada usando poderosos métodos de alto rendimento que podem detectar até mesmo espécies muito raras. Deve-se ter cuidado ao atribuir significado biológico a estes na ausência de testes diretos. Porém, alguns lncRNAs específicos (além de *Xist*) claramente têm funções reguladoras específicas, e alguns deles serão abordados aqui.

HOTAIR é um lncRNA cujo gene é encontrado no agrupamento *HoxC* de seres humanos, mas ele atua na regulação da expressão dos genes *HoxD* em outro cromossomo (em *trans*) ao recrutar o complexo repressivo Polycomb 2 (PRC2, *Polycomb Repressive Complex 2*; ver Cap. 19) para este locus. O PRC2 adiciona grupos trimetil à lisina-27 da histona H3, uma marca (H3K27me3) associada com expressão gênica reprimida (ver Fig. 19-29). *HOTAIR* também recruta um segundo complexo que remove modificações de histonas normalmente associadas a genes ativos. Estes dois complexos proteicos ligam-se a regiões separadas da molécula de RNA de *HOTAIR* e são presumivelmente direcionadas para o locus específico *HoxD* por meio de uma terceira região. Além de estar envolvido no desenvolvimento, *HOTAIR* também está superexpresso em alguns tipos de câncer. A deleção do domínio de ligação a PRC2 de *HOTAIR* destrói suas funções reguladoras em ambos os eventos, desenvolvimento e câncer. Embora isso possa sugerir um papel biológico crucial, é curioso observar que *HOTAIR* é pouco conservado no camundongo, e sua deleção não produz fenótipo aparente. Isso poderia sugerir que ele deve ter evoluído rapidamente em mamíferos.

Outros lncRNAs não atuam em *trans*, mas em *cis*. Como será visto, *Xist* é um destes. Todavia outros incluem RNAs envolvidos no *imprinting*. O *imprinting* foi discutido no Capítulo 19 (ver Fig. 19-31). O locus *Igf2/H19* lá descrito inclui também uma unidade de transcrição que produz o lncRNA *AIR*. O próprio *AIR* sofre *imprinting*, sendo expresso apenas a partir do alelo paterno; sua expressão é necessária para a repressão de vários genes sob *imprinting* no cromossomo paterno a partir do qual ele é expresso. O RNA permanece de certa forma associado ao cromossomo e recruta um complexo proteico que trimetila H3K9, levando à repressão da transcrição nestes promotores-alvo.

A inativação do X gera indivíduos mosaicos

Agora será dada mais atenção à função do lncRNA chamado *Xist* e ao processo de inativação do X. Fêmeas de mamíferos possuem dois cromossomos X, enquanto machos possuem apenas um X e um cromossomo Y. Embora esta seja a base da determinação do sexo – o que permite que machos e fêmeas

sejam diferentes – ela também gera um problema: qualquer gene codificado pelo cromossomo X seria, se não fosse controlado, expresso em dobro nas fêmeas, quando comparadas aos machos. Este desbalanço iria potencialmente causar distúrbios em processos metabólicos e celulares. Evitar estes problemas requer o que se chama de **compensação de dose**. Em mamíferos, isso é obtido pela **inativação** de uma das cópias do cromossomo X das fêmeas. Esta ação resulta na ausência da expressão de todos os genes da cópia do cromossomo. Em mamíferos placentados, a inativação ocorre no estágio de 32 para 64 células, e a escolha de qual cromossomo X será inativado – a cópia materna ou a paterna – é aparentemente feita ao acaso em cada célula. Uma vez selecionada em cada célula, a mesma cópia permanece inativada em todas as células descendentes.

Uma consequência de a inativação ser aleatória em cada célula é que as fêmeas são mosaicos – algumas de suas células expressam o cromossomo X paterno e outras, o materno. Isso geralmente não tem consequências significativas, embora possa influenciar na gravidade dos sintomas de doenças ligadas ao X, dependendo da proporção de células nas quais o gene mutado é expresso ou silenciado. Um exemplo mais familiar é o gato cálico (ou casco de tartaruga) (Fig. 20-21). Em gatos, um gene do cromossomo X determina se sua pelagem será cor de laranja ou preta – um alelo deste gene dá origem à pelagem cor de laranja e o outro, à pelagem preta. Em gatos heterozigotos para este gene, as diferentes manchas de pelagem preta ou cor de laranja revelam regiões compostas por células nas quais um ou outro cromossomo X foi inativado. Esta observação também explica porque os gatos cálicos são fêmeas. A pelagem branca vem do efeito de um gene autossômico.

Xist é um longo RNA não codificador que inativa um único cromossomo X em fêmeas de mamíferos

Como um cromossomo X é inativado, e como a inativação é herdada ao longo do desenvolvimento? O regulador inicial é uma molécula de RNA chamada *Xist*. Este RNA é codificado no *locus* reconhecidamente vital para a inativação do X, chamado *Xic* (do inglês, *X-inactivation center* [centro de inativação do X]), no cromossomo X. O RNA de *Xist* cobre o cromossomo X a partir do qual é expresso. Isso pode ser visto no resultado da hibridização *in situ* na Figura 20-22a. Não está claro o que causa esta cobertura nem como ela é restrita a um cromossomo X (i.e., porque ela atua apenas em *cis*). Sabe-se, no entanto, que a ação de *Xist* é central para a inativação e não necessita de outras sequências do cromossomo X além de *Xic*: quando expresso ectopicamente a partir de uma localização autossômica (i.e., a partir de um cromossomo não sexual), *Xist* pode, em graus diferentes, silenciar genes ao longo do cromossomo. Ou seja, ele “inativa” o autossomo a partir do qual é expresso.

O próprio RNA de *Xist* não causa silenciamento, mas recruta outros fatores que modificam e condensam a cromatina (PRC2, etc.) e que talvez também metilem o DNA (como já se viu em outros exemplos de silenciamento de mamíferos no Cap. 19). Mais tarde, há um acúmulo de uma variante rara de histona chamada MacroH2A, que está normalmente associada à cromatina compacta e silenciada. São estas modificações que causam o silenciamento e garantem que ele seja herdado: uma vez firmemente estabelecido, o próprio *Xist* não é mais necessário. Uma característica da modificação de histonas do cromossomo X inativado é mostrada na Figura 20-22b. Nesta, o cromossomo X inativado está muito menos acetilado do que o resto do genoma. Como foi visto em capítulos anteriores (Caps. 8 e 19), histonas desacetiladas estão associadas a regiões do genoma que não são transcritas.

Como uma célula escolhe qual cromossomo X será inativado? A resposta ainda é vaga, mas outro RNA regulador pode ser a chave. Este outro RNA também é codificado pelo *locus Xic* mas na fita oposta e sobreposto ao gene de *Xist*. Ele é chamado de *Tsix* (*Xist* escrito ao contrário) e atua como regulador negativo de *Xist* (Fig. 20-23). De fato, se *Tsix* for mutado em um dado



FIGURA 20-21 Visualização da inativação do X: o gato cálico. As manchas de pelagem cor de laranja e preta fornecem uma visualização indireta da inativação do cromossomo X, como descrito no texto. (Cortesia de VG.)

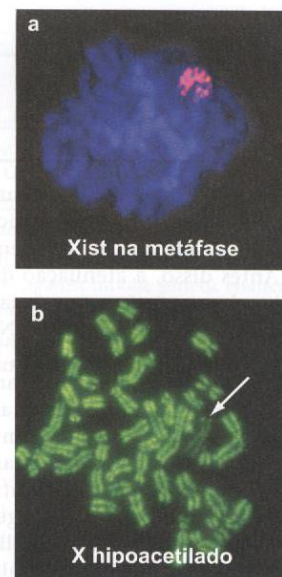


FIGURA 20-22 Visualização da inativação do X: marcadores moleculares. (a) Localização do RNA de *Xist* ao longo do cromossomo X inativo mostrada por hibridização *in situ* de células metafásicas. (b) Os cromossomos estão marcados para acetilação na histona H4. A seta indica cromossomo X inativado, que possui níveis muito mais baixos de acetilação do que outros cromossomos. (Reproduzida, com permissão, de Brockdorff N. e Turner B.M. 2007. *Epigenetics* [ed. D. Allis et al.], p. 327. © Cold Spring Harbor Laboratory Press.)

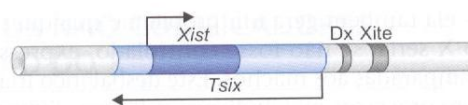


FIGURA 20-23 *Tsix* antagoniza a expressão de *Xist*. *Tsix* (sombreado em azul-claro) é expresso como um RNA antissenso de *Xist* (sombreado em azul-escuro) e é mais longo que *Xist*. O grau de sobreposição está indicado na região em azul-escuro. *Xite* e *DxPas34* (*Dx*) são elementos reguladores que controlam a expressão dos genes. No início da inativação, *Xist* e *Tsix* são expressos a partir de ambos os cromossomos X, mas, após um tempo, o cromossomo que será inativado aumenta a expressão de *Xist*, enquanto a expressão a partir do cromossomo destinado a permanecer ativo diminui. Ainda não está claro como esta alteração nos níveis de *Xist* é regulada por *Tsix*, mas se *Tsix* for deletado de um dos cromossomos, será sempre esta cópia que se tornará inativa.

cromossomo X, este será o cromossomo escolhido para inativação. Assim, um balanço entre a produção e a estabilidade dos RNAs de *Xist* e *Tsix* pode inclinar o resultado para um lado ou para outro em cada célula.

A compensação de dose é necessária em todos os animais (p. ex., vermes e moscas), da mesma maneira como nos mamíferos. Em cada caso, porém, os mecanismos para atingir a compensação são diferentes. Por exemplo, em *Drosophila*, ela é alcançada pelo aumento da expressão de genes ligados ao X no macho (ao invés de sua diminuição na fêmea). Mas aqui o mecanismo também envolve RNAs reguladores não codificadores. Neste caso, os RNAs (chamados de *roX1* e *roX2*) estão envolvidos no recrutamento de complexos modificadores de cromatina para os genes do cromossomo X de machos, onde auxiliam a ativar a transcrição.

RESUMO

Apesar da proposta de que moléculas de RNA seriam provavelmente agentes de regulação gênica datar de 1961, foi apenas na última década que sua ocorrência e significância vieram à tona. Antes disso, a atenuação do operon *trp* em *E. coli* era considerado um caso raro no qual se sabia que as sequências de RNA na região 5' de um mRNA controlavam a expressão dos genes a jusante. Neste caso, padrões alternativos de pareamento de bases intramoleculares nesta região do RNA dão origem a estruturas secundárias alternativas que comunicam diferentes destinos aos genes. Em uma conformação, a transcrição é terminada antes de entrar na região codificadora dos genes a jusante, enquanto em outra conformação, ela permite que a transcrição continue, e os genes são expressos.

Os ribocomutadores controlam genes de maneira semelhante: estruturas secundárias alternativas nas 5'-UTRs dos genes determinam se a transcrição destes genes continua (ou, em outros casos, se a tradução é iniciada). Com os ribocomutadores, a escolha da estrutura secundária alternativa depende da ligação direta ao RNA dos ligantes que controlam o gene em questão.

E. coli também codifica pequenos RNAs (sRNAs) que atuam em *trans* para regular genes. Assim, pequenos genes codificam RNAs curtos que pareiam com mRNAs portadores de sequências complementares. Esta situação inibe a tradução destes mRNAs-alvo, desencadeia sua destruição, ou até mesmo, em alguns casos, estimula sua tradução. As ações de pequenos RNAs bacterianos são semelhantes em várias maneiras às de sRNAs que regulam genes em células eucarióticas, embora a maquinaria usada para produzir estes reguladores de RNA e

a maquinaria usada para alcançar seus efeitos em genes-alvos sejam bastante diferentes.

Um sistema bacteriano reflete mais de perto o que se vê em eucariotos, e este é o sistema CRISPR. Agrupamentos de sequências de DNA repetitivas características (diferentes em cada agrupamento) dão origem a uma classe especial de sRNAs que dirigem uma maquinaria proteica para destruir fagos e plasmídeos infecciosos – de fato, qualquer DNA exógeno que entra na célula. A habilidade para distinguir entre “próprio” e “não próprio” vem do fato de que as regiões CRISPR acumulam fragmentos de genomas de fago de infecções anteriores e estes determinam as sequências-alvo dos RNAs.

Os equivalentes mais próximos a CRISPR nos eucariotos são os chamados agrupamentos de piRNA. Estas regiões contêm fragmentos de transposons (e são, às vezes, chamadas de “cemitérios de transposons”). Os curtos piRNAs gerados a partir destas regiões silenciam transposons homólogos no genoma, particularmente na linhagem germinativa, garantindo que eles permaneçam inativos. Pequenos RNAs de interferência (siRNAs) e microRNAs (miRNAs) são as duas classes de RNAs reguladores curtos encontradas em eucariotos. Ao contrário dos piRNAs, os siRNAs e miRNAs são gerados a partir de regiões de RNA de dupla-fita por meio do processamento por uma enzima chamada Dicer. Os siRNAs são derivados diretamente de dsRNA endógeno ou exógeno em uma única etapa. Os miRNAs são codificados no genoma e seu caráter de dupla-fita deriva de regiões de estrutura secundária que são reconhecidas primeiramente por uma enzima chamada Drosha, que os processa até um estágio em que Dicer possa atuar. Em ambos