

# Regulação da Expressão Gênica em Eucariotos

## PANORAMA

- ▶ Mecanismos de regulação da expressão gênica eucariótica | Considerações gerais
- ▶ Indução da atividade de transcrição por fatores ambientais e biológicos
- ▶ Controle molecular da transcrição em eucariotos
- ▶ Regulação pós-transcricional da expressão gênica por interferência por RNA
- ▶ Expressão gênica e organização da cromatina
- ▶ Ativação e inativação de cromossomos inteiros

### Tripanossomos africanos | Um baú de disfarces moleculares

No fim do século 19, David Bruce, cirurgião do British Medical Service, resumiu suas observações e experimentos sobre uma doença que acometia animais selvagens e domésticos do sul da África.

A doença, conhecida como nagana, palavra zulu que significa "perda do espírito", é caracterizada por febre, edema, letargia e emagrecimento extremo. Bruce reconheceu que a nagana é transmitida pela mosca tsé-tsé, um tipo de mosca que pica e é comum em planícies abertas de vegetação arbustiva na África. Além disso, o exame dos animais doentes levou-o a concluir que o agente causador é um protozoário unicelular flagelado injetado no sangue do animal durante a picada da mosca tsé-tsé. Esse parasito do sangue, um tipo de tripanossomo, agora é denominado *Trypanosoma brucei* em homenagem a Bruce. Os seres humanos também podem ser infectados pelos tripanossomos transmitidos pela mosca tsé-tsé e desenvolvem a doença debilitante conhecida como doença do sono africana.



Tripanossomos entre as hemácias.

Tanto em animais quanto em seres humanos, as infecções por tripanossomos são muito duradouras. Isso é incrível porque, no sangue, os tripanossomos sofrem ataques repetidos do sistema imune. A cada ataque imune, a maioria dos tripanossomos é destruída; no entanto, alguns sempre sobrevivem para repovoar o sangue e manter a infecção. A base desse ressurgimento é a capacidade do tripanossomo de modificar a proteína que reveste sua superfície. Cada tripanossomo é coberto por cerca de 10 milhões de moléculas de uma mesma glicoproteína. Quando o sistema imune reconhece esse revestimento proteico, o tripanossomo infeccioso é aprisionado e destruído pelas células imunes. No entanto, antes da eliminação de todos os tripanossomos do animal, alguns deles conseguem substituir a glicoproteína de superfície por outra que não é imediatamente reconhecida pelo sistema imune. Esses tripanossomos alterados escapam da destruição e proliferam. Por fim, o sistema imune aprende a reconhecê-los também, mas, enquanto

isso, surge outro grupo de tripanossomos alterados que mantêm a infecção. A oferta aparentemente interminável de disfarces moleculares disponíveis para os tripanossomos deve-se a uma grande série de genes que codificam as glicoproteínas variáveis de superfície (VSG) que revestem esses organismos. Apenas um desses genes é expresso de cada vez; todos os outros são silenciosos. Durante uma infecção, porém, a identidade do gene expresso é modificada. A cada mudança, os tripanossomos adquirem uma nova proteína de superfície e colocam-se um passo à frente das defesas imunes do animal. Assim, a infecção é mantida durante semanas ou até mesmo meses até que o animal morra por exaustão.

# Mecanismos de regulação da expressão gênica eucariótica | Considerações gerais

A regulação da expressão gênica eucariótica pode ocorrer na transcrição, no processamento ou na tradução.

## DIMENSÕES DA REGULAÇÃO GÊNICA EUCARIÓTICA

A história do mecanismo usado por tripanossomos para escapar dos ataques do sistema imune é uma história sobre a regulação gênica. Diferentes genes *vsg* são expressos em diferentes momentos – isto é, há regulação temporal dos genes *vsg*. Entre eucariotos, principalmente organismos multicelulares como nós, os genes também são regulados em uma dimensão espacial. Os organismos multicelulares contêm muitos tipos celulares diferentes organizados em tecidos e órgãos. Determinado gene poderia ser expresso em células do sangue, mas nunca em células nervosas. Outro gene poderia ter perfil de expressão exatamente oposto. A regulação que cria essas diferenças na expressão gênica é a base da complexidade anatômica e fisiológica de eucariotos multicelulares.

Como em procariotos, a expressão de genes em eucariotos requer a transcrição de DNA em RNA e a subsequente tradução desse RNA em polipeptídios. No entanto, antes da tradução, a maior parte do RNA eucariótico é “processada”. Durante o processamento, o RNA é protegido por *cap* na extremidade 5', poliadenilado na extremidade 3' e alterado internamente pela perda das sequências de íntrons não codificadoras (Capítulo 11). Em geral, os RNA procarióticos não sofrem essas modificações terminais e internas.

A expressão gênica é mais complexa em eucariotos que em procariotos, porque as células eucarióticas são divididas em compartimentos por um sistema elaborado de membranas. Essa compartimentalização subdivide as células em organelas separadas, das quais a mais visível é o núcleo; as células eucarióticas também têm mitocôndrias, cloroplastos (no caso das células vegetais) e um retículo endoplasmático. Cada organela tem uma função diferente. O núcleo armazena o material genético, as mitocôndrias e cloroplastos obtêm energia, e o retículo transporta materiais dentro da célula.

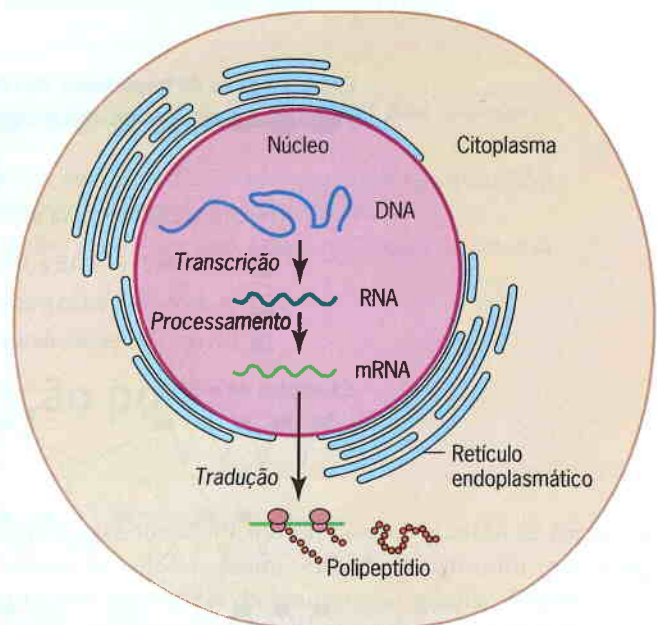
A subdivisão das células eucarióticas em organelas separa fisicamente os processos de expressão gênica. O principal processo, a transcrição de DNA em RNA, ocorre no núcleo. Os transcritos de RNA também são modificados no núcleo por adição do *cap*, poliadenilação e remoção de íntrons. Os RNA mensageiros resultantes são exportados para o citoplasma, onde são associados a ribossomos, muitos deles localizados nas membranas do retículo endoplasmático. Uma vez associados a ribossomos, esses mRNA são traduzidos em polipeptídios. Essa separação física dos processos da expressão gênica torna possível a regulação em diferentes locais

(Figura 19.1). A regulação pode ocorrer no núcleo, em nível de DNA ou RNA, ou no citoplasma, em nível de RNA ou polipeptídio.

## TRANSCRIÇÃO DE DNA CONTROLADA

Em procariotos, a expressão gênica é regulada principalmente por controle da transcrição do DNA em RNA. Um gene que não é transcrito simplesmente não é expresso. A transcrição ocorre em procariotos quando moléculas reguladoras negativas, como a proteína repressora *lac*, são removidas da vizinhança de um gene e moléculas reguladoras positivas, como um complexo proteína ativadora do catabolismo (CAP)/AMP cíclico, ligam-se a ele (Capítulo 18). Essas interações de proteína-DNA controlam o acesso da RNA polimerase a um gene. Além disso, os mecanismos desenvolvidos para controlar a transcrição nesses organismos respondem rapidamente a mudanças ambientais. Como discutimos no Capítulo 18, esse controle imediato é uma estratégia eficiente para a sobrevivência de procariotos.

O controle da transcrição é mais complexo em eucariotos do que em procariotos. Uma razão é que os genes são sequestrados no núcleo. Para que tenham algum efeito sobre o nível da transcrição, os sinais ambientais têm de ser transmitidos da superfície celular, onde geralmente são recebidos, através do citoplasma e da membrana nuclear até os cromossomos. Portanto, as células eucarióticas necessitam de sistemas de sinalização internos muito elaborados para controlar a transcrição de DNA. Ou-



■ FIGURA 19.1 Expressão gênica eucariótica mostrando os estágios em que a expressão pode ser regulada: transcrição, processamento e tradução.



tro fator de complicação é a multicelularidade de muitos eucariotos. Pode ser necessário que os sinais ambientais atravessem camadas de células para influenciar a transcrição de genes em determinado tecido. Portanto, a comunicação intercelular é um aspecto importante da regulação transcricional eucariótica.

Como em procariotos, a regulação transcricional eucariótica é mediada por interações de proteína-DNA. Proteínas reguladoras positivas e negativas ligam-se a regiões específicas do DNA e estimulam ou inibem a transcrição. Esse grupo de proteínas é denominado **fatores de transcrição**. Muitos tipos diferentes foram identificados, e a maioria parece ter domínios característicos que possibilitam sua interação com o DNA. A estrutura dessas proteínas e a natureza de suas interações com o DNA serão comentadas em uma seção a seguir.

## RECOMPOSIÇÃO ALTERNATIVA DE RNA

A maioria dos genes eucarióticos tem íntrons, regiões não codificadoras que interrompem a sequência especificadora dos aminoácidos de um polipeptídeo. É preciso remover todos os íntrons do transcrito de RNA de um gene para que haja expressão adequada da sequência codificadora. Conforme analisado no Capítulo 11, esse processo requer a união precisa das sequências codificadoras, ou éxons, em um RNA mensageiro. A formação do mRNA é mediada por *spliceossomos*, diminutas organelas nucleares.

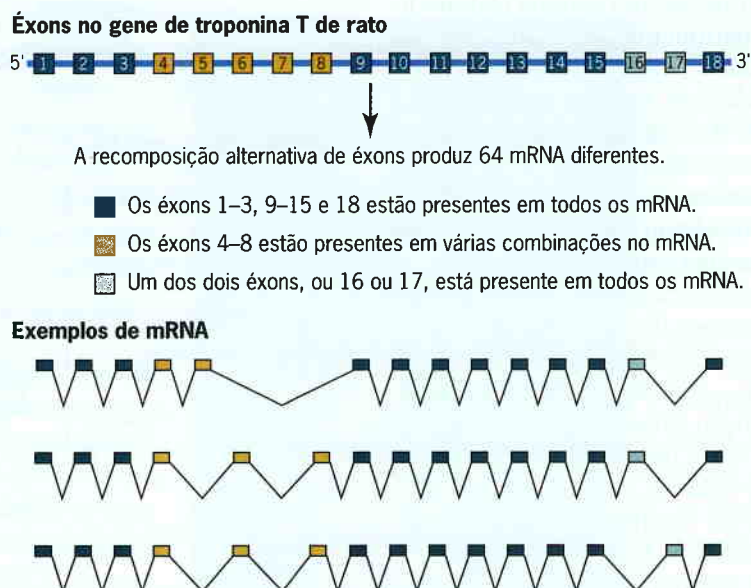
Genes com vários íntrons constituem um problema curioso para o mecanismo de recomposição do RNA. Esses íntrons podem ser removidos separadamente ou em combinação, dependendo da interação entre o mecanismo de recomposição e o RNA. Caso dois íntrons sucessivos sejam removidos juntos, o éxon entre eles também será removido. Assim, o mecanismo de recomposição

tem a oportunidade de modificar a sequência codificadora de um RNA por deleção de alguns de seus éxons. Aparentemente, esse fenômeno de recomposição de um transcrito de RNA de diferentes maneiras é um modo de economizar informações genéticas. Em vez de duplicar genes, ou trechos de genes, a *recomposição alternativa* de transcritos torna possível que um único gene codifique diferentes polipeptídios.

Um exemplo de recomposição alternativa ocorre durante a expressão do gene para troponina T, uma proteína encontrada no músculo esquelético de vertebrados; o tamanho dessa proteína varia de aproximadamente 150 a 250 aminoácidos. No rato, o gene da troponina T tem mais de 16 kb e contém 18 éxons diferentes (**Figura 19.2**). Os transcritos desse gene são recompostos de diferentes maneiras para criar uma grande série de mRNA cuja tradução produz muitos polipeptídios troponina T diferentes. Todos esses polipeptídios têm em comum os aminoácidos dos éxons 1 a 3, 9 a 15 e 18. No entanto, as regiões codificadas pelos éxons 4 a 8 podem estar presentes ou ausentes, dependendo do padrão de recomposição, e aparentemente em qualquer combinação. Outra variação é determinada pela presença ou ausência de regiões codificadas pelos éxons 16 e 17; se 16 estiver presente, 17 não estará, e vice-versa. Essas diferentes formas de troponina T provavelmente atuam de maneiras um pouco diferentes nos músculos, contribuindo para a variabilidade da ação das células musculares. Observe a variação que pode ser gerada por recomposição alternativa de RNA em Resolva!: Contagem de mRNA.

## CONTROLE CITOPLASMÁTICO DA ESTABILIDADE DO RNA MENSAGEIRO

Os RNA mensageiros são exportados do núcleo para o citoplasma, onde servem de molde para a síntese de polipep-



■ FIGURA 19.2 Recomposição alternativa de transcritos do gene de troponina T de rato. Apenas três dos 64 diferentes mRNA possíveis são mostrados.

## Resolva!

### Contagem de mRNA

O transcrito primário de um gene com um íntron é recomposto para produzir um único tipo de mRNA. Com dois íntrons, pode haver recomposição alternativa do transcrito; cada íntron pode ser removido separadamente, ou os dois íntrons podem ser removidos com o éxon entre eles. Assim, dois mRNA diferentes podem ser gerados a partir do transcrito desse gene. Quantos mRNA diferentes podem ser gerados por recomposição alternativa de transcritos de genes com três ou quatro íntrons? Suponha que o primeiro e o último éxons estejam presentes em todos os mRNA, mas que os éxons internos possam estar presentes ou ausentes, dependendo do padrão de recomposição. Qual é a fórmula geral do número de mRNA gerados por recomposição alternativa de um transcrito de um gene com  $n$  íntrons?

► Leia a resposta do problema no site <http://gen-io.grupogen.com.br>.

túdios. Uma vez no citoplasma, determinado mRNA pode ser traduzido por vários ribossomos que se movem sequencialmente ao longo dele. Essa linha de montagem da tradução continua até que o mRNA seja degradado. Portanto, a degradação do RNA mensageiro é outro ponto de controle no processo geral de expressão gênica. Os mRNA de longa duração podem manter vários ciclos de síntese de polipeptídios, ao contrário dos mRNA de curta duração.

Um mRNA que é rapidamente degradado tem de ser repostado por transcrição complementar; caso contrário, cessa a síntese do polipeptídio codificado por ele. É claro que essa interrupção da síntese do polipeptídio pode ser parte de um programa de desenvolvimento. Depois de exercer seu efeito, o polipeptídio pode não ser mais necessário; na verdade, sua síntese contínua pode ser

prejudicial. Nesses casos, a rápida degradação do mRNA seria um modo lógico de evitar a síntese indesejada de polipeptídios.

A longevidade do RNA mensageiro pode ser influenciada por vários fatores. As caudas poli(A) parecem estabilizar o mRNA. A sequência da região 3' não traduzida (3' UTR) que precede uma cauda poli(A) também parece afetar a estabilidade do mRNA. Vários mRNA de curta duração têm a sequência AUUUA repetida várias vezes em suas regiões 3' não traduzidas. Quando essa sequência é artificialmente transferida para a região 3' não traduzida de mRNA mais estáveis, eles também se tornam instáveis. Fatores químicos, como hormônios, também podem afetar a estabilidade do mRNA. No sapo *Xenopus laevis*, o gene *vitellogenin* tem a transcrição ativada pelo hormônio esteroide estrogênio. Entretanto, além de induzir a transcrição desse gene, o estrogênio também aumenta a longevidade de seu mRNA.

Uma pesquisa recente revelou que a estabilidade do mRNA e a tradução do mRNA em polipeptídios também são reguladas por pequenas moléculas de RNA não codificador denominadas pequenos RNA de interferência (siRNA) ou microRNA (miRNA). Essas moléculas de RNA reguladoras, que têm entre 21 e 28 nucleotídeos, são produzidas a partir de RNA bifilamentares maiores em uma grande variedade de organismos eucarióticos, inclusive fungos, vegetais e animais. Os RNA de interferência curtos e os microRNA emparelham suas bases com sequências de mRNA específicos; uma vez emparelhados, eles causam a clivagem e subsequente degradação do mRNA ou impedem a tradução do mRNA em polipeptídio. Em vegetais, essas pequenas moléculas de RNA constituem uma defesa crucial contra infecção por vírus de RNA, e tanto em vegetais quanto em animais elas regulam a expressão de genes participantes da maturação e do desenvolvimento. Esse assunto será apresentado com mais detalhes adiante neste capítulo.

### PONTOS ESSENCIAIS

- Proteínas denominadas fatores de transcrição interagem com o DNA para controlar a transcrição de genes eucarióticos
- Os transcritos de genes eucarióticos também podem ser recompostos para produzir RNA mensageiros que codificam polipeptídios distintos, mas relacionados
- A estabilidade dos RNA mensageiros eucarióticos pode influenciar o nível de síntese de polipeptídios.

## Indução da atividade de transcrição por fatores ambientais e biológicos

A expressão gênica eucariótica pode ser induzida por fatores ambientais como calor e por moléculas sinalizadoras como hormônios e fatores de crescimento.

Ao estudarem o óperon de lactose em *E. coli*, Jacob e Monod descobriram que houve transcrição específica dos

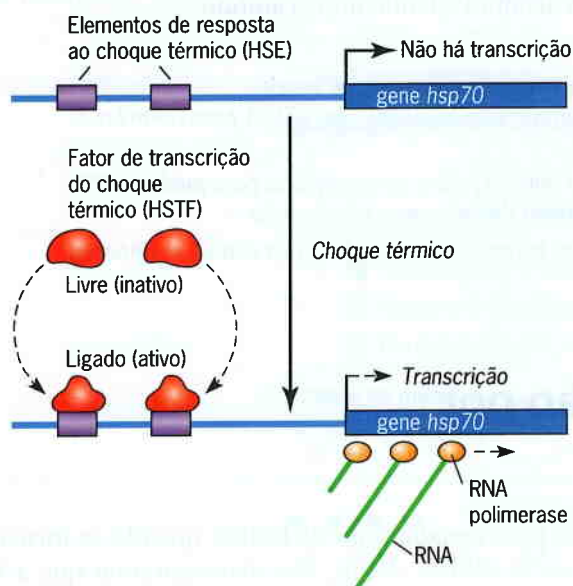
genes para metabolismo da lactose quando se forneceu lactose às células. Assim, eles demonstraram que a lactose era um indutor da transcrição gênica. Seguindo os passos de Jacob e Monod, muitos pesquisadores tentaram identificar indutores específicos da transcrição de genes eucarióticos. Embora essas tentativas tenham tido



considerável sucesso, o grau geral da indução de genes eucarióticos por fatores ambientais e nutricionais parece ser menor que em procariotos. Abordemos aqui dois exemplos de expressão gênica induzível em eucariotos.

## TEMPERATURA | OS GENES DO CHOQUE TÉRMICO

Quando submetidos ao estresse de temperatura elevada, os organismos respondem por meio da síntese de um grupo de proteínas que ajudam a estabilizar o meio celular interno. Essas *proteínas do choque térmico*, encontradas em procariotos e eucariotos, estão entre os polipeptídios mais conservados conhecidos. As comparações das sequências de aminoácidos das proteínas do choque térmico de organismos tão diversos quanto *E. coli* e *Drosophila* mostram que eles são 40 a 50% idênticos – um achado notável considerando-se o tempo evolutivo que separa esses organismos. A expressão das proteínas do choque térmico é regulada na transcrição; isto é, o estresse por calor induz especificamente a transcrição dos genes codificadores dessas proteínas (Figura 19.3). Em *Drosophila*, por exemplo, uma das proteínas do choque térmico, denominada HSP70 (do inglês, *heat shock protein*, com peso molecular de 70 quilodáltons) é codificada por uma família de genes localizados em dois agrupamentos vizinhos em um dos autossomos. Ao todo, existem cinco a seis cópias desses genes *hsp70* nos dois agrupamentos. Quando a temperatura ultrapassa 33°C, como nos dias de verão, cada um desses genes é transcrito em RNA, que é processado e traduzido para produzir polipeptídios HSP70. Essa transcrição dos genes *hsp70* induzida por calor é mediada pelo fator de transcrição do choque térmico (HSTF), um polipeptídio presente nos núcleos



■ FIGURA 19.3 Indução de transcrição do gene *hsp70* de *Drosophila* por choque térmico. Os HSE estão localizados em uma região distante entre 40 e 90 pares de bases do sítio de início da transcrição em direção 5' (seta dobrada).

das células de *Drosophila*. Quando a *Drosophila* é exposta ao calor, o HSTF é quimicamente alterado por fosforilação. Nesse estado alterado, liga-se especificamente às sequências nucleotídicas em posição 5' em relação aos genes *hsp70* e torna os genes mais acessíveis à RNA polimerase II, a enzima que transcreve a maioria dos genes codificadores de proteínas. Então, a transcrição dos genes *hsp70* é vigorosamente estimulada. As sequências a que o HSTF fosforilado se liga são denominadas *elementos de resposta ao choque térmico (HSE)*.

## MOLÉCULAS SINALIZADORAS | GENES QUE RESPONDEM A HORMÔNIOS

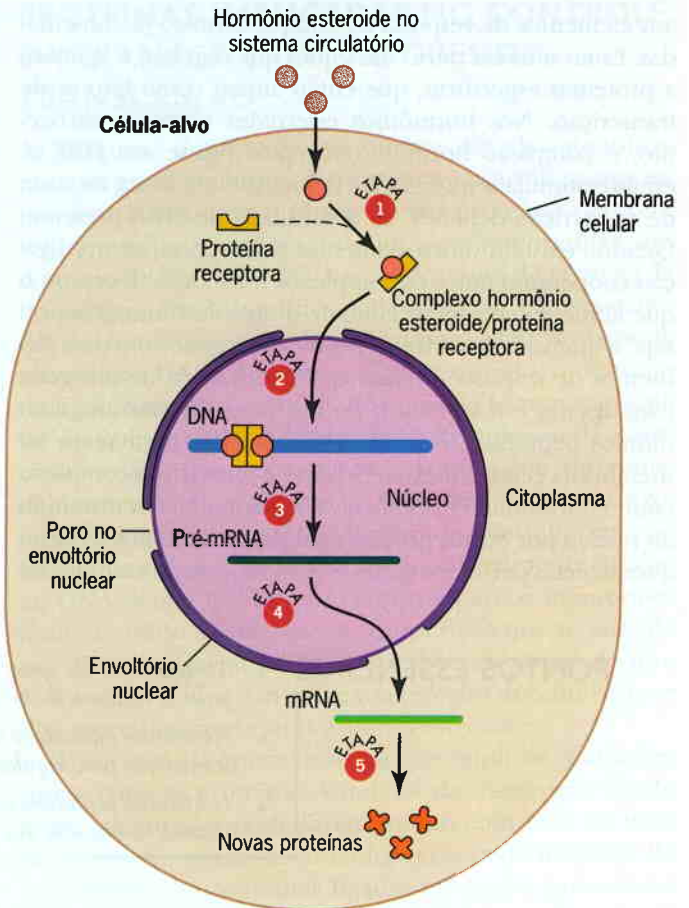
Em eucariotos multicelulares, um tipo de célula pode enviar sinais para outro por secreção de **hormônio**. Os hormônios circulam por todo o corpo, entram em contato com as células-alvo e iniciam uma série de processos que regulam a expressão de determinados genes. Em animais existem duas classes gerais de hormônios. A primeira classe, os *hormônios esteroides*, é constituída de pequenas moléculas lipossolúveis derivadas do colesterol. Em vista de sua natureza lipídica, elas atravessam a membrana celular sem dificuldade ou com pouca dificuldade. Os exemplos são estrogênio e progesterona, que têm papéis importantes no ciclo reprodutivo feminino; testosterona, hormônio de diferenciação e comportamento masculino; glicocorticoides, que participam do controle do nível sanguíneo de glicose; e ecdisona, um hormônio que controla os processos de desenvolvimento em insetos. Uma vez dentro da célula, esses hormônios interagem com proteínas citoplasmáticas ou nucleares denominadas *receptores hormonais*. O complexo receptor/hormônio formado interage com o DNA e atua como fator de transcrição para regular a expressão de determinados genes (Figura 19.4).

A segunda classe de hormônios, os *hormônios peptídicos*, é constituída de cadeias lineares de aminoácidos. Assim como todos os outros polipeptídios, essas moléculas são codificadas por genes. Os exemplos são insulina, que regula os níveis sanguíneos de glicose; somatotropina, um hormônio do crescimento; e prolactina, que atua no tecido mamário das fêmeas. Como os hormônios peptídicos costumam ser grandes demais para atravessarem livremente a membrana celular, é preciso que seus sinais sejam transmitidos para o interior das células por *proteínas receptoras ligadas à membrana* (Figura 19.5). A interação do hormônio peptídico com seu receptor modifica a conformação do receptor, o que acarreta mudanças em outras proteínas dentro da célula. Graças a uma cascata de mudanças, o sinal hormonal é transmitido através do citoplasma da célula até o núcleo, onde finalmente regula a expressão de genes específicos. Esse processo de transmissão do sinal hormonal através da célula até o núcleo é conhecido como **transdução de sinal**.

A expressão gênica induzida por hormônio é mediada por sequências específicas no DNA. Essas sequências, denominadas *elementos de resposta hormonal (HRE)*, são análogas

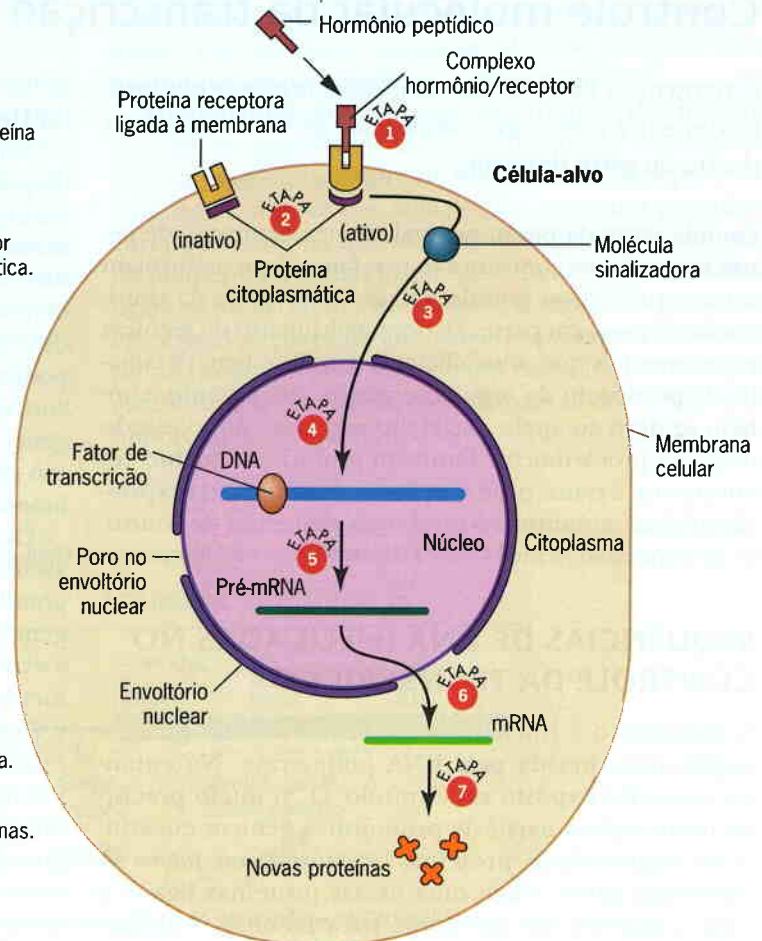
■ FIGURA 19.4 Regulação da expressão gênica por hormônios esteroides. O hormônio interage com um receptor dentro da célula-alvo. Neste exemplo, o receptor está no citoplasma; outros receptores de hormônios esteroides estão localizados no núcleo. O complexo hormônio/receptor desloca-se dentro do núcleo, onde ativa a transcrição de determinados genes.

- ETAPA 1** O hormônio esteroide entra na célula-alvo e combina-se com uma proteína receptora.
- ETAPA 2** O complexo hormônio/receptor liga-se a um elemento de resposta hormonal no DNA.
- ETAPA 3** O complexo ligado estimula a transcrição.
- ETAPA 4** O transcrito é processado e transportado para o citoplasma.
- ETAPA 5** O mRNA é traduzido em proteínas.



■ FIGURA 19.5 Regulação da expressão gênica por hormônios peptídicos. O hormônio (um sinal extracelular) interage com um receptor na membrana da célula-alvo. O complexo hormônio/receptor formado ativa uma proteína citoplasmática que desencadeia uma cascata de modificações intracelulares. Essas modificações transmitem o sinal até o núcleo, onde um fator de transcrição estimula a expressão de determinados genes.

- ETAPA 1** O hormônio liga-se a uma proteína receptora na membrana da célula-alvo.
- ETAPA 2** O complexo hormônio/receptor formado ativa uma proteína citoplasmática.
- ETAPA 3** A proteína citoplasmática ativada transduz um sinal até o núcleo.
- ETAPA 4** O sinal induz a ligação de um fator de transcrição ao DNA.
- ETAPA 5** O fator de transcrição ligado estimula a transcrição.
- ETAPA 6** O transcrito é processado e transportado para o citoplasma.
- ETAPA 7** O mRNA é traduzido em proteínas.





aos elementos de resposta ao choque térmico já comentadas. Estão situadas perto dos genes que regulam e ligam-se a proteínas específicas, que então atuam como fatores de transcrição. Nos hormônios esteroides como o estrogênio, o complexo hormônio/receptor liga-se aos HRE e, então, estimula a transcrição. A intensidade dessa resposta de transcrição depende da quantidade de HRE presente. Quando existem vários elementos de resposta, ocorre ligação cooperativa entre os complexos hormônio/receptor, o que aumenta consideravelmente a taxa de transcrição; ou seja, a intensidade da transcrição de um gene com dois elementos de resposta é maior que o dobro da de um gene com apenas um elemento de resposta. No caso dos hormônios peptídicos, o receptor geralmente permanece na membrana celular, mesmo depois de formar um complexo com o hormônio. Portanto, o sinal hormonal é transmitido ao núcleo por outras proteínas, algumas das quais se ligam a sequências perto dos genes regulados pelo hormônio. Es-

sas proteínas atuam então como fatores de transcrição para controlar a expressão dos genes.

A atividade de transcrição pode ser induzida por muitos outros tipos de proteínas que não são hormônios no sentido clássico – isto é, não são produzidas por uma glândula ou órgão específico. Essas incluem diversas moléculas circulantes secretadas, como o fator de crescimento neural, o fator de crescimento epidérmico e o fator de crescimento derivado das plaquetas, além de outras moléculas não circulantes associadas à superfície celular ou à matriz entre as células. Embora cada uma dessas proteínas tenha suas próprias peculiaridades, o mecanismo geral pelo qual induzem a transcrição assemelha-se ao dos hormônios peptídicos. A interação entre a proteína sinalizadora e um receptor ligado à membrana inicia uma cadeia de processos dentro da célula que acabam por resultar na ligação de fatores de transcrição específicos a determinados genes, que são então transcritos.

### PONTOS ESSENCIAIS

- A transcrição dos genes *hsp70* em resposta ao aumento da temperatura é mediada por um fator de transcrição do choque térmico
- Hormônios esteroides e suas proteínas receptoras formam complexos que atuam como fatores de transcrição para regular a expressão de genes específicos
- Hormônios peptídicos interagem com proteínas receptoras ligadas à membrana e ativam um sistema sinalizador que regula a expressão de genes específicos.

## Controle molecular da transcrição em eucariotos

A transcrição de genes eucarióticos é regulada por interações entre proteínas e sequências de DNA localizadas dentro ou perto dos genes.

Grande parte da pesquisa atual sobre a expressão de genes eucarióticos concentra-se nos fatores que controlam a transcrição. Essa grande ênfase no controle da transcrição deve-se, em parte, ao desenvolvimento de técnicas experimentais que possibilitaram a análise bem detalhada desse aspecto da regulação gênica. No entanto, também se deve ao apelo das ideias surgidas com o estudo de genes procarióticos. Tanto em procariotos quanto em eucariotos, a transcrição é o processo primário da expressão gênica; portanto, é o nível mais elementar de controle da expressão gênica.

### SEQUÊNCIAS DE DNA IMPLICADAS NO CONTROLE DA TRANSCRIÇÃO

A transcrição é iniciada no promotor de um gene, a região reconhecida pela RNA polimerase. No entanto, como foi exposto no Capítulo 11, o início preciso da transcrição a partir de promotores gênicos eucarióticos requer várias proteínas acessórias, ou *fatores de transcrição basais*. Cada uma dessas proteínas liga-se a uma sequência no promotor para facilitar o alinhamento

correto da RNA polimerase no filamento molde do DNA.

A transcrição de genes eucarióticos também é controlada por diversos *fatores de transcrição especiais*, como os que participam da regulação dos genes induzíveis por calor e hormônios de que já tratamos. Esses fatores ligam-se a elementos de resposta ou, de modo mais geral, a sequências denominadas *acentuadores* adjacentes a um gene. Os fatores de transcrição especiais que se ligam a esses acentuadores podem interagir com os fatores de transcrição basais e com a RNA polimerase, que se ligam ao promotor de um gene. As interações ocorridas entre os fatores de transcrição especiais, os fatores de transcrição basais e a RNA polimerase regulam a atividade de transcrição de um gene.

Os acentuadores apresentam três propriedades razoavelmente gerais: (1) atuam a distâncias relativamente grandes – até vários milhares de pares de bases do(s) gene(s) regulado(s); (2) a influência que exercem sobre a expressão gênica independe do sentido – eles são igualmente eficazes em sentido normal ou invertido no DNA; e (3) seus efeitos independem da posição – eles podem estar localizados na região 5', na região 3' ou dentro de um íntron de um gene e ainda ter efeitos acentuados sobre a expressão do gene. Essas três características distinguem os acentuadores dos promotores, que geralmente estão localizados em posição imediatamente 5' em relação ao gene e que só atuam em um sentido.

Os acentuadores podem ser relativamente grandes e ter até várias centenas de pares de bases. Às vezes contêm sequências repetidas que têm atividade reguladora parcial própria. A maioria dos acentuadores é específica para o tecido; isto é, só estimula a transcrição em determinados tecidos. Em outros tecidos, são simplesmente ignorados. Um exemplo claro dessa especificidade tecidual provém do estudo do gene *yellow* em *Drosophila* (Figura 19.6). Esse gene é responsável pela pigmentação de muitas partes do corpo – asas, pernas, tórax e abdome. As moscas de tipo selvagem apresentam pigmento escuro, preto-acastanhado, em todas essas estruturas, enquanto as moscas mutantes têm um pigmento castanho-amarelado, mais claro. Alguns mutantes, porém, apresentam um padrão de pigmentação em mosaico, preto-acastanhado em alguns tecidos e castanho-amarelado em outros. Esses mosaicos são causados por mutações que alteram a transcrição do gene *yellow* em alguns tecidos, mas não em outros. Pamela Geyer e Victor Corces mostraram que o gene *yellow* é regulado por vários acentuadores, alguns deles localizados dentro de um íntron, e que cada acentuador ativa a transcrição em um tecido diferente. Se, por exemplo, houver mutação do acentuador para expressão na asa, as cerdas das asas serão castanho-amareladas em vez de preto-acastanhadas. A série de acentuadores associados ao gene *yellow* possibilita o controle de sua expressão de acordo com o tecido. Veja outro método de estudo dos acentuadores em Problema resolvido: Definição das sequências necessárias para expressão de um gene.

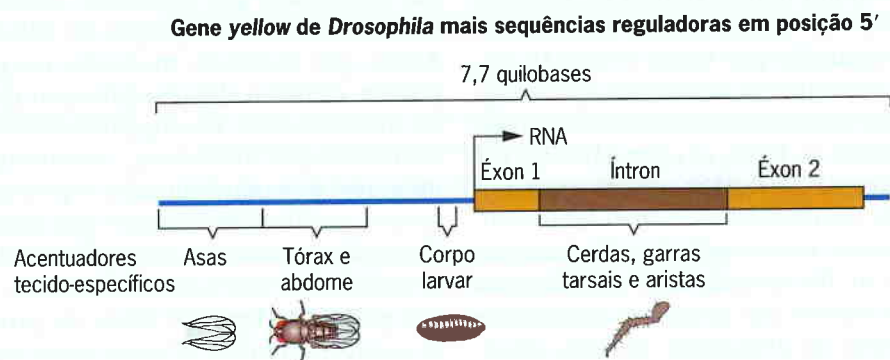
Como os acentuadores influenciam a transcrição de genes? Os resultados de muitos estudos indicam que as proteínas que se ligam aos acentuadores influenciam a atividade das proteínas que se ligam aos promotores, o que inclui os fatores de transcrição basais e a RNA polimerase. Os dois tipos de proteínas são postos em contato físico por um complexo multimérico constituído de pelo menos 20 proteínas diferentes. Esse *complexo mediador* parece curvar o DNA de tal modo que as proteínas ligadas a um acentuador ficam justapostas àquelas ligadas ao promotor. Dessa maneira, as proteínas ligadas ao acentuador controlam a transcrição, que é iniciada no promotor.

## PROTEÍNAS IMPLICADAS NO CONTROLE DA TRANSCRIÇÃO | FATORES DE TRANSCRIÇÃO

As pesquisas durante as três últimas décadas identificaram um grande número de proteínas eucarióticas que estimulam a transcrição. Muitas dessas proteínas parecem ter no mínimo dois domínios químicos importantes: um domínio de ligação ao DNA e um domínio de ativação da transcrição. Esses domínios podem ocupar partes separadas das moléculas ou podem ser superpostos. No fator de transcrição GAL4 de levedura, por exemplo, o domínio de ligação ao DNA está situado perto da terminação amino do polipeptídeo. Dois domínios de ativação da transcrição estão presentes nesse polipeptídeo, um aproximadamente no meio e o outro perto da terminação carboxi. Nas proteínas receptoras do hormônio esteroide, que são fatores de transcrição em animais, o domínio de ligação ao DNA ocupa localização central e parece superpor-se a um domínio de ativação da transcrição que se estende em direção à terminação amino. Os receptores de hormônios esteroides têm ainda um terceiro domínio que se liga especificamente ao hormônio esteroide.

A ativação da transcrição parece implicar interações físicas entre as proteínas. Um fator de transcrição ligado a um acentuador pode fazer contato com uma ou mais proteínas em outros acentuadores ou pode interagir diretamente com proteínas ligadas na região promotora. Por esses contatos e interações, o domínio de ativação da transcrição do fator pode induzir mudanças de conformação nas proteínas montadas, abrindo caminho para que a RNA polimerase inicie a transcrição.

Muitos fatores de transcrição eucarióticos têm motivos estruturais característicos que resultam de associações entre aminoácidos dentro de suas cadeias polipeptídicas. Um desses motivos é o *dedo de zinco*, uma alça peptídica curta que se forma quando duas cisteínas em uma parte do polipeptídeo e duas histidinas em outra parte próxima ligam-se juntas a um íon de zinco; o segmento peptídico entre os dois pares de aminoácidos protraem-se do corpo da proteína como se fosse um dedo (Figura 19.7A). A análise mutacional mostrou que esses dedos têm papéis importantes na ligação do DNA.



■ FIGURA 19.6 Acentuadores tecido-específicos do gene *yellow* de *Drosophila*.



## PROBLEMA RESOLVIDO



## Definição das sequências necessárias para expressão de um gene

## PROBLEMA

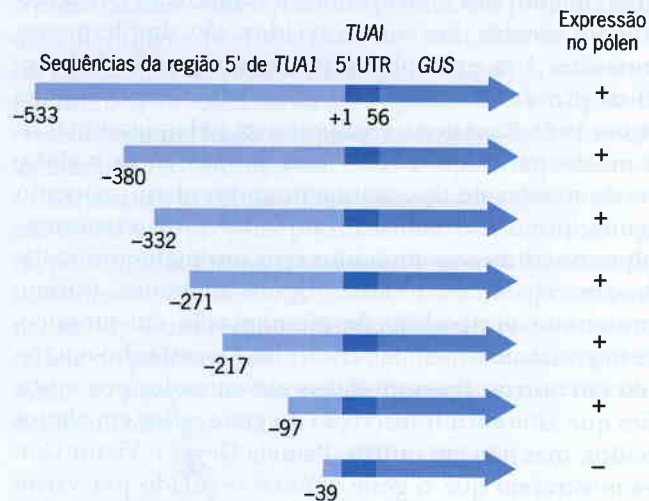
As tubulinas são proteínas importantes do citoesqueleto em eucariotos. Em *Arabidopsis thaliana*, a tubulina codificada pelo gene *TUA1* é expressa principalmente no pólen. Para determinar as sequências responsáveis por essa expressão tecido-específica, 533 pares de bases de DNA em posição 5' em relação ao sítio de início da transcrição de *TUA1* mais os 56 primeiros pares de bases da região não traduzida 5' do gene *TUA1* foram fundidos à sequência codificadora do gene da  $\beta$ -glucuronidase (*GUS*) de *E. coli*. A  $\beta$ -glucuronidase catalisa a conversão de X-gluc, uma substância incolor, em um pigmento azul-escuro. Assim, a observação de pigmento azul em material tratado com X-gluc indica expressão do gene *GUS*. Quando esse ensaio foi aplicado a plantas *Arabidopsis* geneticamente transformadas com o gene *GUS* fundido atrás das sequências na região 5' do gene *TUA1*, o pólen tornou-se azul-escuro; todos os outros tecidos continuaram incolores. Então, todo o experimento foi repetido usando segmentos cada vez mais curtos das sequências da região 5' de *TUA1* para estimular a expressão do gene *GUS*. A partir dos resultados mostrados na Figura 1, que parte da região 5' é necessária para expressão do gene *TUA1*?

## FATOS E CONCEITOS

1. A região 5' do sítio de início de transcrição de um gene contém o promotor desse gene.
2. Essa região também pode conter acentuadores que regulam a expressão do gene por um mecanismo espacial ou temporal específico.
3. A região 5' não traduzida de um gene está entre o sítio de início da transcrição e o sítio de início da tradução.
4. Genes de *E. coli* como *GUS* podem ser expressos em eucariotos como *Arabidopsis* se forem fundidos aos promotores eucarióticos.

## ANÁLISE E SOLUÇÃO

Nessa série de experimentos, *GUS* é um "repórter" que informa se as sequências da região 5' do gene *TUA1* são capazes de estimular



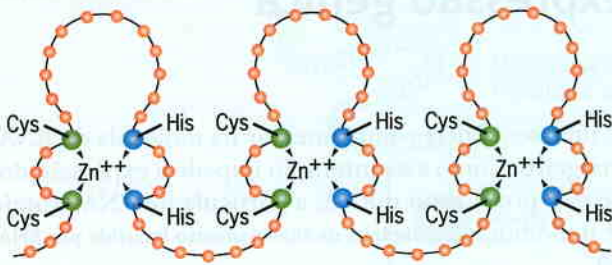
■ FIGURA 1 Expressão de transgenes *TUA1/GUS* no pólen de *Arabidopsis*. Segmentos cada vez mais curtos da região 5' do gene *TUA1* e uma sequência curta da região não traduzida 5' (UTR) desse gene foram fundidos às sequências codificadoras do gene *GUS* de *E. coli*. +1 é o sítio de início da transcrição do gene *TUA1*. Os nucleotídeos à esquerda desse sítio são indicados por números negativos. A atividade de *GUS* no pólen transgênico é indicada por um sinal de mais; a ausência de atividade de *GUS* é indicada por um sinal de menos. Veja mais detalhes em Carpenter, J., S. E. Ploense, D. P. Snustad e C. D. Silflow. 1992. Preferential expression of an  $\alpha$ -tubulin gene of *Arabidopsis* in pollen. *The Plant Cell* 4: 557-571.

a expressão gênica. Todas as sequências da região 5', exceto as menores, podem ser bons estimuladores. Assim, há necessariamente uma sequência entre os pares de bases -97 e -39 na sequência da região 5' de *TUA1* que é crucial para a expressão gênica. Sem essa sequência, o gene *TUA1* não pode ser expresso. Além disso, essa sequência é suficiente para estimular a expressão de *TUA1* no pólen maduro. Desse modo, atua como um acentuador que controla a expressão tecido-específica do gene *TUA1*.

Um segundo motivo em muitos fatores de transcrição é *hélice-volta-hélice*, um trecho de três hélices curtas de aminoácidos separadas por voltas (Figura 19.7B). Análises genéticas e bioquímicas mostraram que o segmento helicoidal mais próximo da terminação carboxi é necessário para ligação ao DNA; as outras hélices parecem participar da formação de dímeros de proteína. Em muitos fatores de transcrição, o motivo hélice-volta-hélice coincide com uma região extremamente conservada de cerca de 60 aminoácidos, denominada *homeodomínio* porque ocorre em proteínas codificadas pelos genes homeóticos de *Drosophila*. Análises clássi-

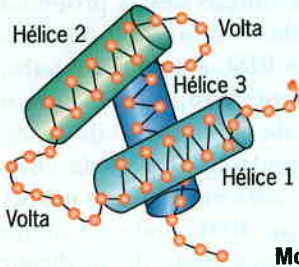
cas mostraram que mutações nesses genes alteram o desenvolvimento de grupos de células (Capítulo 20). Assim, por exemplo, mutações no gene *Antennapedia* podem causar o desenvolvimento de pernas no lugar de antenas. Esse fenótipo bizarro é um exemplo de transformação homeótica – substituição de uma parte do corpo por outra durante o processo de desenvolvimento. Análises moleculares dos genes homeóticos em *Drosophila* mostraram que cada um deles codifica uma proteína com um homeodomínio e que essas proteínas podem se ligar ao DNA. As proteínas de homeodomínio estimulam a transcrição espacial e temporal

**Motivo dedo de zinco**



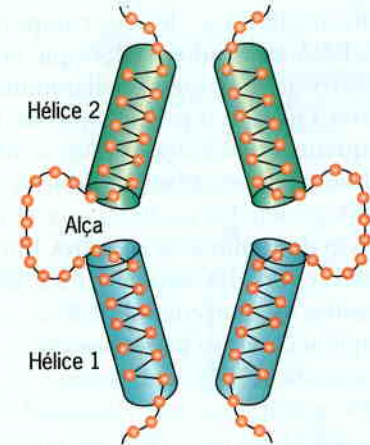
**A**

**Motivo hélice-volta-hélice**



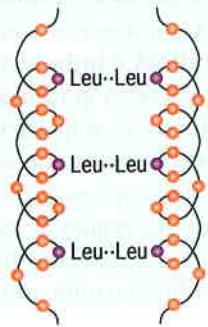
**B**

**Motivo hélice-alça-hélice**



**D**

**Motivo zíper de leucina**



**C**

■ **FIGURA 19.7** Motivos estruturais em diferentes tipos de fatores de transcrição. **A.** Motivos dedos de zinco no fator de transcrição de mamíferos SP1. **B.** Motivo hélice-volta-hélice em um fator de transcrição do tipo homeodomínio. **C.** Um motivo zíper de leucina que possibilita a dimerização de dois polipeptídios, seguida por ligação ao DNA. **D.** Um motivo hélice-alça-hélice que possibilita a dimerização de dois polipeptídios, seguida por ligação ao DNA.

específica de determinados genes durante o desenvolvimento. As proteínas de homeodomínio também foram identificadas em outros organismos, inclusive em seres humanos, nos quais podem ter papel importante como fatores de transcrição.

Um terceiro motivo estrutural encontrado em fatores de transcrição é o *zíper de leucina*, um trecho de aminoácidos com uma leucina a cada sétima posição (**Figura 19.7C**). Os polipeptídios com essa característica podem formar dímeros por interações entre as leucinas em cada uma dessas regiões de zíper. Em geral, a sequência zíper é adjacente a um trecho de aminoácidos com carga positiva. Quando há interação de dois zíperes, essas regiões carregadas se estendem em sentidos opostos, formando uma superfície que pode se ligar ao DNA de carga negativa.

Um quarto motivo estrutural encontrado em alguns fatores de transcrição é a *hélice-alça-hélice*, um trecho de duas regiões helicoidais de aminoácidos separados por uma alça não helicoidal (**Figura 19.7D**). As regiões helicoidais possibilitam a dimerização entre dois polipeptídios. Às vezes, o motivo hélice-alça-hélice está situado adjacente a um trecho de aminoácidos básicos (de carga positiva), de modo que, quando há dimerização, esses aminoácidos podem se ligar ao DNA de carga negativa. Proteínas com essa característica são designadas *proteínas básicas HLH* ou *bHLH*.

Os fatores de transcrição com motivos de dimerização como o zíper de leucina ou a hélice-alça-hélice poderiam, em princípio, combinar-se a polipeptídios semelhantes a eles mesmos para formar homodímeros ou a polipeptídios diferentes para formar heterodímeros. Essa segunda possibilidade sugere um modo de alcançar padrões complexos de expressão gênica. A transcrição de um gene em determinado tecido poderia depender de ativação por um heterodímero, que só seria formado se seus polipeptídios constituintes fossem sintetizados nesse tecido. Além disso, esses dois polipeptídios teriam de estar presentes em quantidades corretas para favorecer a formação do heterodímero em detrimento dos homodímeros correspondentes. Desse modo, seria possível obter modulações sutis da expressão gênica por variação da concentração dos dois componentes de um heterodímero.

**PONTOS ESSENCIAIS**

- Os acentuadores têm ação independente do sentido e a distâncias consideráveis para regular a transcrição de um promotor do gene
- Os fatores de transcrição reconhecem e se ligam a sequências de DNA específicas nos acentuadores
- Os fatores de transcrição têm motivos estruturais característicos, como o dedo de zinco, a hélice-volta-hélice, o zíper de leucina e a hélice-alça-hélice.



## Regulação pós-transcricional da expressão gênica por interferência por RNA

RNA não codificadores curtos podem regular a expressão de genes eucarióticos por interação com os RNA mensageiros produzidos por esses genes.

Embora grande parte da regulação de genes eucarióticos ocorra na transcrição, uma pesquisa recente mostrou que os mecanismos pós-transcricionais também têm papéis importantes na regulação da expressão de genes eucarióticos. Alguns desses mecanismos contam com pequenos RNA não codificadores. Esses pequenos RNA interferem na expressão gênica mediante emparelhamento de bases com sequências-alvo em moléculas de RNA mensageiro. Portanto, esse tipo de regulação gênica pós-transcricional é denominado **interferência por RNA**, abreviada com frequência como **RNAi**. A maioria dos tipos de organismos eucarióticos é capaz de efetuar a RNAi. Entre os organismos genéticos-modelo, esse fenômeno foi bem estudado no nematódeo *Caenorhabditis elegans*, em *Drosophila* e em *Arabidopsis*. Também existe em mamíferos, inclusive nos seres humanos. Como veremos, a capacidade disseminada de organismos eucarióticos de regular a expressão gênica por RNAi possibilitou que geneticistas analisassem as funções dos genes em organismos que não são receptivos aos métodos genéticos tradicionais.

### VIAS DE RNAi

O fenômeno da interferência por RNA, que é resumido na **Figura 19.8**, conta com a participação de pequenas moléculas de RNA conhecidas como RNA de interferência curtos (siRNA) ou microRNA (miRNA). Essas moléculas, com 21 a 28 pares de bases, são produzidas a partir de moléculas maiores de RNA bifilamentar pela ação enzimática de proteínas que são endonucleases específicas para RNA bifilamentar. Essas endonucleases “dividem” um grande RNA em pedaços pequenos e são denominadas enzimas *Dicer*. O nematódeo *Caenorhabditis elegans* produz um só tipo de enzima *Dicer*; a *Drosophila* produz duas enzimas *Dicer* diferentes; e *Arabidopsis* produz pelo menos três. Em *C. elegans* e *Drosophila*, essas enzimas atuam no citoplasma; em *Arabidopsis*, provavelmente atuam no núcleo. Os siRNA e miRNA produzidos por atividade da enzima *Dicer* apresentam as bases emparelhadas em todo o comprimento, exceto nas extremidades 3', onde há dois nucleotídeos sem par.

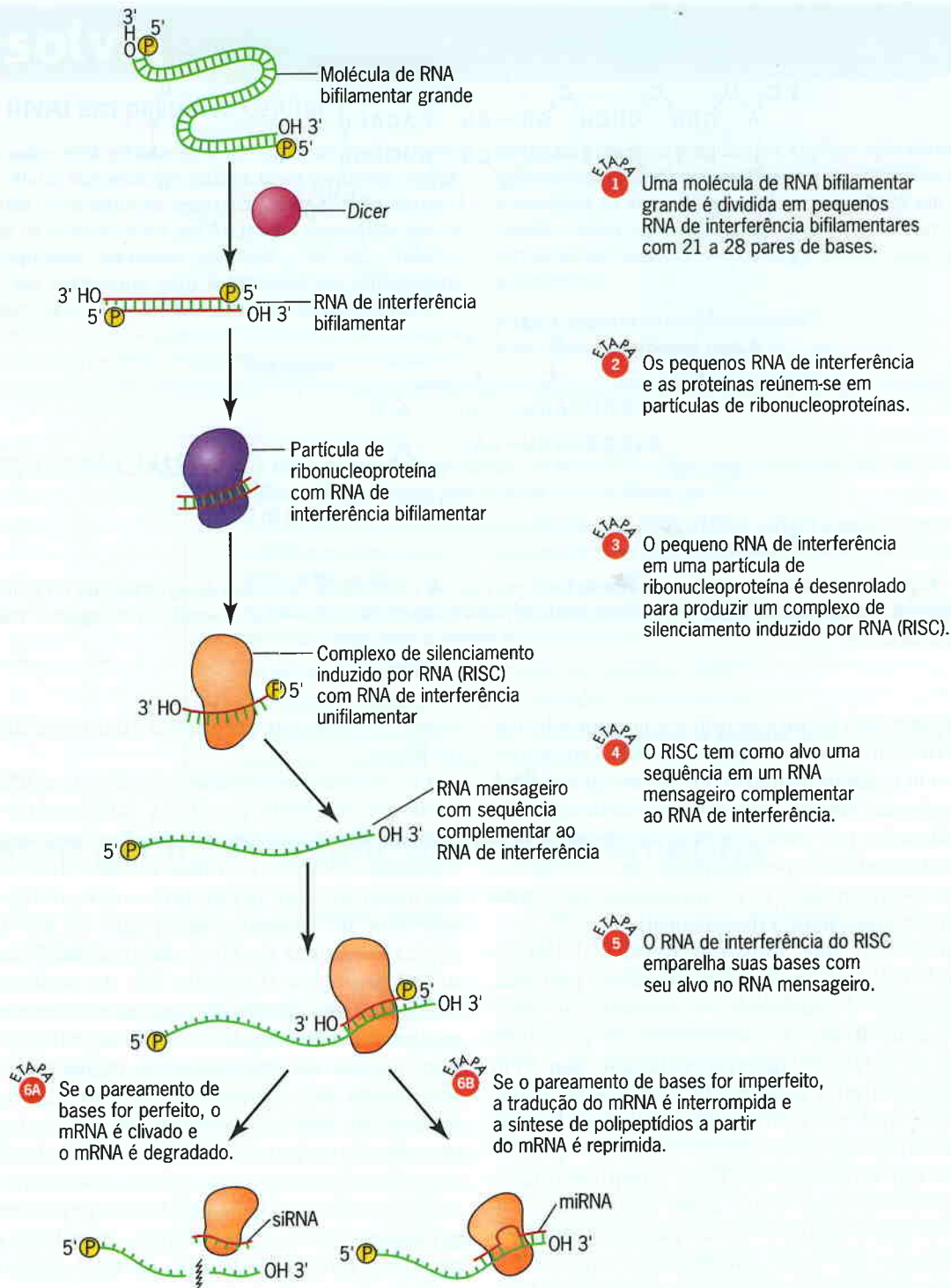
No citoplasma, siRNA e miRNA são incorporados a partículas de ribonucleoproteínas. O siRNA ou miRNA bifilamentar nessas partículas são desenrolados, e um de seus filamentos é preferencialmente eliminado. Então, o filamento simples de RNA remanescente é capaz de interagir com moléculas específicas de RNA mensageiro. Essa interação é mediada por pareamento de bases entre o filamento simples de RNA no complexo RNA-proteí-

na e uma sequência complementar na molécula de RNA mensageiro. Como essa interação impede a expressão do gene que produziu o mRNA, a partícula de RNA-proteína é denominada **complexo de silenciamento induzido por RNA (RISC)**.

Os RISC de diferentes organismos têm tamanhos e composições diferentes. Todos, porém, contêm pelo menos uma molécula da família de proteínas que têm o curioso nome de Argonauta. A função dessas proteínas não é totalmente compreendida. Sempre que o pareamento de bases entre o RNA no RISC e a sequência-alvo no mRNA é perfeito ou quase perfeito, o RISC cliva o mRNA-alvo no meio da região de pareamento de bases. Então, o mRNA clivado é degradado. Depois da clivagem, o RISC pode associar-se a outra molécula de mRNA e induzir sua clivagem. Como um RISC pode ser usado repetidas vezes sem perder a capacidade de se dirigir para o mRNA e clivá-lo, comporta-se como catalisador. Os RNA associados a RISC que ocasionam a clivagem do mRNA geralmente são denominados **RNA de interferência curtos**. Quando o pareamento do RNA no RISC com sua sequência-alvo é imperfeito, o mRNA geralmente não é clivado; em vez disso, a tradução do mRNA é inibida. Os RNA associados ao RISC que têm esse efeito geralmente são denominados **microRNA**. Em animais, as sequências visadas por RISC são encontradas nas regiões 3' não traduzidas de moléculas de mRNA, e, com frequência, essas sequências estão presentes várias vezes na região 3' não traduzida (UTR). Em vegetais, as sequências visadas por RISC geralmente estão localizadas na região codificadora do mRNA ou na UTR 5' do mRNA.

### FONTES DE RNA DE INTERFERÊNCIA CURTOS E microRNA

Algumas das pequenas moléculas de RNA que induzem RNAi são derivadas dos transcritos de genes de microRNA. Esses genes, geralmente indicados pelo símbolo *mir*, são encontrados nos genomas de muitos tipos de eucariotos; os genomas de *C. elegans* e *Drosophila* têm cerca de 100 *mir*s, e os genomas de vertebrados, cerca de 250. A princípio, alguns desses genes foram identificados por análise de mutações que alteravam a regulação de outros genes. Quando os genes *mir* definidos por essas mutações foram analisados em nível molecular, constatou-se que seu potencial codificador de proteínas era nulo ou pequeno. Em vez disso, eles tinham uma estrutura peculiar. Cada um continha um trecho curto de nucleotídeos repetidos em sentidos opostos em torno de um segmento interposto curto de DNA. Quando transcrita, essa estrutura repetida invertida gera um RNA que pode se dobrar sobre si mesmo para formar uma haste bifilamentar curta na base de uma alça unifilamentar (**Figura 19.9A**). Uma



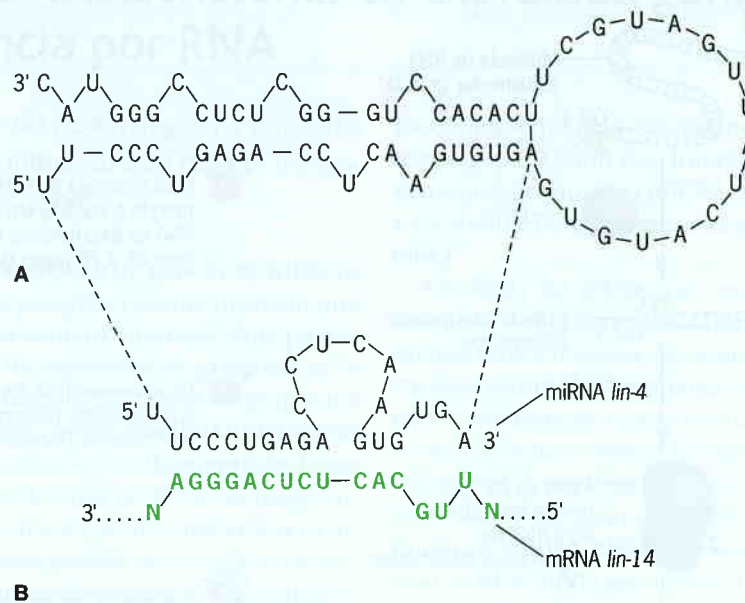
■ FIGURA 19.8 Resumo dos processos das vias de interferência por RNA.

enzima denominada Drosha reconhece essa região de haste-alça e excisa-a do transcrito primário do gene *mir*. A haste-alça liberada é exportada para o citoplasma, onde é clivada por Dicer e forma um miRNA. Em *C. elegans*, no qual esse processo foi descoberto, Dicer remove a alça e aparar a haste até um comprimento de 22 nucleotídeos em cada filamento. Após o amadurecimento em um RISC, o miRNA – agora unifilamentar – pode ter como alvo uma sequência no mRNA produzido por outro gene. A **Figura 19.9B** mostra o pareamento de bases entre o miRNA do

gene *mir* de *C. elegans lin-4* e um desses alvos de miRNA na UTR 3' do mRNA de um gene codificador de proteína, *lin-14*. Por meio desse pareamento de bases, o miRNA *lin-4* reprime a tradução do mRNA *lin-14*.

Desde a descoberta desses genes *mir* definidos por mutação, muitos outros genes *mir* foram encontrados com o auxílio de programas de computador para rastrear as sequências de DNA genômico de *C. elegans*, *Drosophila* e outros organismos-modelo para a estrutura de repetição invertida característica. Muitos dos genes *mir* candidatos





■ FIGURA 19.9 Regulação da expressão gênica por interferência por RNA. A. Estrutura de haste-alça de um transcrito do gene de microRNA *lin-4* de *C. elegans*. B. Pareamento de bases entre o microRNA derivado do transcrito de *lin-4* e uma sequência na região 3' não traduzida do RNA mensageiro de *lin-14*.

identificados por essa técnica genômica computadorizada foram verificados por detecção de miRNA derivados desses genes em extratos celulares. Os genes cujos mRNA contêm sequências visadas por miRNA também estão sendo identificados por uma combinação de uma análise computadorizada e experimentação *in vivo*. Muitos desses genes codificam fatores de transcrição ou outras proteínas importantes para o desenvolvimento.

Alguns dos RNA que induzem RNAi são derivados da transcrição de outros elementos no genoma, como transpósons e transgenes, e também são derivados de vírus de RNA. Os mecanismos de formação desses tipos de RNA de interferência não são bem compreendidos. Algum aspecto do transpósão, transgene ou RNA viral marca-o como incomum. Em vegetais e nematódeos, esses RNA incomuns podem ser copiados em moléculas de RNA complementares por enzimas conhecidas como RNA polimerases dependentes de RNA (RdRP). Se as bases do filamento complementar de RNA continuarem emparelhadas com as bases do molde usado para sintetizá-lo, a molécula de RNA bifilamentar resultante pode ser dividida em siRNA por enzimas tipo *Dicer*; então, os siRNA produzidos por *Dicer* podem entrar na via de RNAi tendo como alvo a população de RNA que deu origem a eles. Desse modo, RNA possivelmente problemáticos derivados de transpósons, transgenes ou vírus podem ser alvos de repressão ou degradação. Essa aplicação de RNAi pode representar sua função mais primitiva – proteger organismos contra infecções virais e transposição descontrolada. Em contrapartida, os intrincados sistemas baseados em miRNA para regulação gênica evidentes em organismos como *C. elegans* pa-

recem representar aplicações altamente desenvolvidas da RNAi.

Os pesquisadores descobriram que a RNAi também pode ser induzida por RNA bifilamentar preparado por transcrição *in vitro* de genes ou segmentos de gene clonados. O DNA é transcrito nos dois sentidos por sua inserção entre promotores em sentidos opostos em um vetor de clonagem adequado ou por inserção de cópias invertidas do DNA em posição 3' em relação a um só promotor (Capítulo 16). As moléculas de RNA bifilamentares derivadas dos transcritos desses clones podem ser introduzidas em células cultivadas; elas também podem ser injetadas em organismos vivos. Uma vez dentro das células, o RNA bifilamentar entra em uma via de RNAi. É dividido em moléculas de siRNA, que são incorporadas a complexos de RNA-proteína e direcionadas para mRNA contendo sequências complementares. Os mRNA-alvo geralmente são degradados. Assim, o tratamento de células ou organismos com determinado tipo de RNA bifilamentar tem o efeito de suprimir ou diminuir a expressão do gene que corresponde a esse RNA. É equivalente, portanto, à indução de uma mutação amórfica ou hipomórfica no gene. Graças a essa técnica, os geneticistas puderam estudar as consequências da ablação ou atenuação da expressão de determinados genes em uma grande variedade de organismos, inclusive alguns nos quais a análise genética é difícil, lenta ou impossível. Portanto, agora a RNAi é usada para analisar a função de genes em peixes, roedores e seres humanos, bem como em organismos-modelo como *C. elegans*, *Drosophila* e *Arabidopsis*. Veja uma aplicação dessa tecnologia em Resolva! Uso de RNAi em pesquisa celular.

## Resolva!

### Uso de RNAi em pesquisa celular

Um pesquisador está estudando a formação de centríolos dentro de células humanas em cultura. Essas pequenas organelas têm papel importante na orquestração da divisão celular. É possível ver os centríolos por coloração apropriada das células. O pesquisador pressupõe que duas proteínas,  $\gamma$ -tubulina e CEP135, são necessárias para a formação do centríolo. Os genes para essas duas proteínas foram clonados a partir do genoma humano, e suas sequências foram analisadas. Descreva como a técnica de interferência por RNA poderia ser usada para

testar a hipótese do pesquisador. Explique que materiais seriam necessários e como você verificaria a ocorrência de bloqueio ou diminuição da síntese de  $\gamma$ -tubulina e CEP135 em células cultivadas. Como você verificaria se houve comprometimento da formação do centríolo? Que controles incluiria nesses experimentos?

► **Leia a resposta do problema no site**  
<http://gen-io.grupogen.com.br>

### PONTOS ESSENCIAIS

- Os RNA de interferência curtos e os microRNA são produzidos a partir de precursores bifilamentares maiores pela ação de endonucleases tipo Dicer
- Nos complexos de silenciamento induzido por RNA (RISC), siRNA e miRNA tornam-se unifilamentares de modo que possam ter como alvos sequências complementares em moléculas de RNA mensageiro
- O RNA mensageiro que é alvo do siRNA é clivado, e o mRNA-alvo do miRNA é impedido de servir de molde para a síntese de polipeptídeos
- Os genomas eucarióticos têm centenas de genes para miRNA
- Transposons e transgenes podem estimular a síntese de siRNA
- A interferência por RNA é usada como instrumento de pesquisa para suprimir ou atenuar a expressão de genes em células e organismos.

## Expressão gênica e organização da cromatina

Vários aspectos da organização da cromatina influenciam a transcrição de genes.

Os cromossomos eucarióticos são constituídos de partes aproximadamente iguais de DNA e proteína. O conjunto desse material é denominado **cromatina**. As características químicas da cromatina variam ao longo do comprimento do cromossomo. Em algumas regiões, por exemplo, as histonas, que constituem a maior parte da proteína na cromatina, são acetiladas, e, em outras regiões, alguns nucleotídeos no DNA são metilados. Essas modificações químicas podem influenciar a atividade de transcrição dos genes. Outros aspectos da organização da cromatina – por exemplo, a presença de proteínas de “empacotamento” – influenciam a regulação gênica. Nesta seção, analisaremos como a composição e a organização da cromatina afetam a expressão gênica.

### EUCROMATINA E HETEROCROMATINA

A variação na densidade da cromatina nos núcleos celulares leva à coloração diferencial de seções dos cromossomos. O material de coloração intensa é a **heterocromatina**, e o material de coloração fraca é a **eucromatina**. Esses diferentes tipos de cromatina têm algum significado funcional? Caso tenham, qual é ele?

Uma combinação de análises genéticas e moleculares mostrou que a grande maioria dos genes eucarióticos está na eucromatina. Além disso, quando são artificialmente transpostos para um meio heterocromático, os genes eucromáticos tendem a apresentar função anormal e, em alguns casos, inatividade total. Esse comprometimento da função pode criar uma mistura de características normais e mutantes no mesmo indivíduo denominada **variação por efeito de posição**. Esse termo é usado porque a variabilidade do fenótipo é causada por mudança da posição do gene eucromático, especificamente por sua realocação na heterocromatina. Muitos exemplos de variação por efeito de posição foram descobertos em *Drosophila*, geralmente em associação com inversões ou translocações que deslocam um gene eucromático para a heterocromatina. O alelo *white motiled* é um bom exemplo. Nesse caso, um alelo selvagem do gene *white* foi realocado por inversão, com uma quebra perto do *locus white* eucromático e outra na heterocromatina basal do cromossomo X. Esse rearranjo interfere na expressão normal do gene *white* e produz um fenótipo de olho manchado (Figura 19.10). Ao que tudo indica, o gene *white* eucromático não tem boa atividade em meio heterocromático. Esse e outros exemplos levaram à noção de que a heterocromatina reprime a função gênica, talvez porque é condensada em uma forma que não é acessível ao mecanismo de transcrição.





■ FIGURA 19.10 Fenótipo de cor dos olhos variegada de *Drosophila* com o alelo *white mottled* em um cromossomo X rearranjado.

O comportamento do gene *white* em moscas com esse cromossomo X rearranjado indica que a expressão gênica pode ser influenciada por condições que não alteram a sequência nucleotídica do gene. Além disso, como o gene *white* é expresso em algumas áreas do olho, mas não em outras, sabemos que uma vez estabelecidas essas condições, sua herança é clonal à medida que as células do olho se dividem. Como essas condições são superpostas à estrutura básica do gene *white*, dizemos que são **epigenéticas**. O prefixo grego “epi” significa “acima” e nesse caso é usado para indicar que a expressão do gene é regulada por um estado hereditário que não a sequência real do gene. Nesse caso, o estado epigenético hereditário implica algum aspecto da organização da cromatina perto do gene *white* reposicionado. Nas seções a seguir, encontraremos outros exemplos de regulação epigenética da expressão gênica.

## ORGANIZAÇÃO MOLECULAR DO DNA TRANSCRICIONALMENTE ATIVO

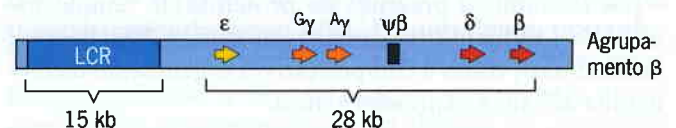
O que é a organização molecular do DNA transcricionalmente ativo? Esse DNA é mais “aberto” que o DNA não transcrito? Essas perguntas foram respondidas medindo-se a sensibilidade do DNA na cromatina à ação da desoxirribonuclease I (DNase) pancreática, enzima que cliva moléculas de DNA e degrada-as em seus nucleotídeos constituintes. Em 1976, Mark Groudine e Harold Weintraub demonstraram que o DNA transcricionalmente ativo é mais sensível à DNase I que o DNA não transcrito. Groudine e Weintraub extraíram cromatina de hemácias de galinha e digeriram-na parcialmente com DNase I. Depois, eles examinaram o material residual da cromatina à procura de sequências de dois genes,  $\beta$ -globina, que é transcrito ativamente em hemácias, e ovalbumina, que não é. Constataram que mais de 50% do DNA do gene de  $\beta$ -globina havia sido digerido pela enzima DNase I em comparação com apenas 10% do DNA do gene de ovalbumina. Esses resultados eram uma forte indicação de que o gene transcrito ativamente estava mais “aberto” ao ataque pela nuclease. Pesquisa subsequente mostrou que

a sensibilidade dos genes transcricionalmente ativos à nuclease depende de pelo menos duas pequenas proteínas não histônicas, HMG14 e HMG17 (HMG é a sigla de *high mobility group* [grupo de alta mobilidade], porque eles têm elevada mobilidade durante a eletroforese em gel). Quando essas proteínas são removidas da cromatina ativa, a sensibilidade à nuclease é perdida; quando elas são acrescentadas de novo, a sensibilidade é restaurada.

O tratamento da cromatina isolada com uma concentração muito baixa de DNase I causa clivagem do DNA em alguns sítios específicos, apropriadamente denominados *sítios hipersensíveis à DNase I*. Demonstrou-se que alguns desses sítios estão em posição 5' em relação aos genes transcricionalmente ativos, na região promotora ou acentuadora. O significado funcional desses sítios hipersensíveis ainda é obscuro, mas há algumas evidências de que podem marcar regiões de desespiralamento local do DNA, talvez por causa do início da transcrição.

No caso dos genes humanos para  $\beta$ -globina, vários sítios hipersensíveis à DNase I estão em uma *região de controle de locus* (LCR) de 15 kb em posição 5' em relação aos próprios genes (Figura 19.11). Os genes da  $\beta$ -globina humana estão em um agrupamento de 28 quilobases no cromossomo 11. Cada gene no agrupamento é cópia de um gene de  $\beta$ -globina ancestral. Durante a evolução, os genes individuais no agrupamento divergiram por mutação aleatória e hoje cada um deles codifica um polipeptídeo ligeiramente diferente. Em um dos genes, uma mutação sem sentido extinguiu a capacidade de produzir um polipeptídeo. Esses genes não codificadores são denominados pseudogenes e geralmente são designados pela letra grega psi ( $\Psi$ ) – portanto, o gene ( $\Psi$ ) $\beta$  nesse agrupamento.

Os genes da  $\beta$ -globina humanos estão submetidos a regulação espacial e temporal. Na verdade, uma característica notável desse agrupamento de genes é que seus membros são expressos em diferentes momentos do desenvolvimento. O gene  $\epsilon$  é expresso no embrião, os dois genes  $\gamma$  são expressos no feto, e os genes  $\delta$  e  $\beta$  são expressos em lactentes e adultos. Aparentemente, essa ativação sequencial de genes de um lado ao outro no agrupamento está relacionada com a necessidade de produzir tipos ligeiramente diferentes de hemoglobina durante o desenvolvimento humano. Embrião, feto e lactente têm ne-



Legenda:	
Tempo de expressão:	
Amarelo	Embrião
Laranja	Feto
Vermelho	Lactente e adulto
Preto	Pseudogene

■ FIGURA 19.11 O agrupamento dos genes da  $\beta$ -globina no cromossomo humano 11.

cessidades diferentes de oxigênio, sistemas circulatórios diferentes e estão em ambientes físicos diferentes. Ao que parece, a variação temporal de expressão do gene da  $\beta$ -globina é uma adaptação a essa variação de condições.

A LCR do agrupamento de genes da  $\beta$ -globina contém sítios de ligação para fatores de transcrição que pré-ativam cada gene para transcrição. A pré-ativação é detectada por aumento da sensibilidade do DNA na LCR à digestão por baixa concentração de DNase I. A transcrição dos genes da  $\beta$ -globina parece exigir essa pré-ativação e é estimulada por fatores de transcrição que se ligam a acentuadores específicos no complexo gênico da  $\beta$ -globina. Entretanto, a especificidade tecidual e temporal da expressão do gene da  $\beta$ -globina depende das sequências inseridas na LCR. Estudos com camundongos transgênicos indicam que a LCR não é apenas uma grande coleção de acentuadores que controlam os vários genes da  $\beta$ -globina. É preciso que a LCR esteja em posição 5' em relação aos genes da  $\beta$ -globina e em seu sentido natural para controlar apropriadamente a expressão gênica. Ou seja, sua ação depende do sentido. Em geral, os acentuadores têm ação independente do sentido e em diferentes posições em relação ao promotor de um gene. A LCR tem outra característica que a distingue de acentuadores simples: é capaz de controlar a expressão do gene da  $\beta$ -globina quando todo o agrupamento gênico (LCR mais genes da  $\beta$ -globina) é inserido em uma posição cromossômica diferente. Os acentuadores, por outro lado, geralmente são inativos quando eles e seus genes associados são transpostos para outra localização cromossômica. Assim, a LCR parece isolar os genes da  $\beta$ -globina da influência da cromatina ao seu redor.

## REMODELAGEM DA CROMATINA

Experimentos que avaliam a sensibilidade do DNA à digestão com DNase I demonstraram que o DNA transcrito é mais acessível ao ataque da nuclease que o DNA não transcrito. O DNA transcrito é empacotado em nucleossomos? Em caso afirmativo, quais são as modificações estruturais do nucleossomo durante a transcrição? Os nucleossomos são “abertos” e “fechados” quando a RNA polimerase passa ao longo do molde de DNA? As tentativas de responder a essas perguntas exigiram uma combinação de métodos genéticos e bioquímicos que demonstraram que o DNA transcrito é realmente empacotado em nucleossomos. No DNA transcrito, porém, os nucleossomos são alterados por complexos multiproteicos que acabam por facilitar a ação da RNA polimerase. Essa alteração de nucleossomos no preparo para transcrição é denominada **remodelagem da cromatina**.

Foram identificados dois tipos gerais de complexos de remodelagem de cromatina. Um tipo é constituído de enzimas que transferem grupos acetila para o aminoácido lisina em posições específicas nas histonas dos nucleossomos. As enzimas dessa classe são denominadas *histona acetiltransferases (HAT)*. Muitos estudos mostraram que a acetilação de histonas está relacionada com o aumento

da expressão gênica, talvez porque o acréscimo dos grupos acetila afrouxa a associação entre o DNA e os octâmeros de histona nos nucleossomos. As *quinases* – enzimas que transferem grupos fosfato para moléculas – também podem ter um papel junto com esses complexos de remodelagem de cromatina. Sabe-se, por exemplo, que a acetilação de lisina-14 em histona H4 frequentemente é precedida por fosforilação de serina-10 nessa molécula. Juntas, essas duas modificações de histona H4 parecem “abrir” a cromatina para aumento da atividade de transcrição.

Outro tipo de complexo de remodelagem de cromatina desorganiza a estrutura do nucleossomo na vizinhança do promotor de um gene. O complexo desse tipo mais estudado é o SWI/SNF encontrado na levedura de pão. Esse complexo recebe o nome dos dois tipos de mutações (*switching-inhibited* [inibido por variabilidade] e *sucrose nonfermenter* [não fermentador de sacarose]) que levaram à descoberta de suas proteínas constituintes. Complexos relacionados foram encontrados nas células de outros organismos, inclusive de seres humanos. O complexo SWI/SNF tem no mínimo oito proteínas. Regula a transcrição por deslizamento de octâmeros de histona ao longo do DNA associado em nucleossomos; também pode transferir esses octâmeros para outros locais em uma molécula de DNA. A substituição do nucleossomo catalisada pelo complexo SWI/SNF aparentemente possibilita o acesso dos fatores de transcrição ao DNA. Então, esses fatores estimulam a expressão de um gene.

Nós analisamos a remodelagem da cromatina do ponto de vista da ativação do gene. No entanto, a cromatina ativa também pode ser remodelada em cromatina inativa. Essa remodelagem invertida parece implicar duas modificações bioquímicas das histonas em nucleossomos: desacetilação, catalisada pelas *histona desacetilases (HDAC)*, e metilação, catalisada pelas *histona metiltransferases (HMT)*. Como será exposto na próxima seção, alguns nucleotídeos no DNA também podem ser metilados por um grupo de enzimas denominadas *DNA metiltransferases (DNMT)*. A cromatina submetida a essas modificações tende a ser transcricionalmente silenciosa.

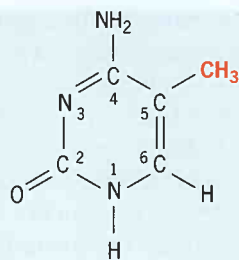
## METILAÇÃO DO DNA

A modificação química dos nucleotídeos também parece ser muito importante para a regulação de genes em alguns eucariotos, sobretudo mamíferos. Dos aproximadamente 3 bilhões de pares de bases no genoma típico de mamífero, cerca de 40% são pares de bases G:C, e cerca de 2 a 7% desses são modificados pelo acréscimo de um grupo metila à citosina (**Figura 19.12**). A maioria das citosinas metiladas é encontrada em díades de pares de bases com a estrutura



em que mC é metilcitosina e o p entre C e G indica a ligação fosfodiéster entre nucleotídeos adjacentes em





■ FIGURA 19.12 Estrutura da 5-metilcitosina.

cada filamento de DNA. Com frequência, essa estrutura é abreviada pela composição de um filamento, portanto, mCpG. Os dinucleotídeos CpG metilados podem ser detectados por digestão do DNA com enzimas de restrição sensíveis a modificações químicas de seus sítios de reconhecimento. Por exemplo, a enzima *HpaII* reconhece e cliva a sequência CCGG; no entanto, quando a segunda citosina nessa sequência é metilada, *HpaII* não é capaz de clivar a sequência. Assim, os DNA metilados e não metilados produzem padrões diferentes de fragmentos de restrição quando são digeridos por essa enzima.

Os dinucleotídeos CpG são menos frequentes que o esperado em genomas de mamíferos, provavelmente porque sofreram mutação em dinucleotídeos TpG ao longo da evolução. Além disso, a distribuição de dinucleotídeos CpG é desigual, com muitos segmentos curtos de DNA que têm uma densidade muito maior de dinucleotídeos CpG que outras regiões do genoma. Esses segmentos ricos em CpG, geralmente com cerca de 1 a 2 kb, são denominados **ilhas de CpG**. No genoma humano, há aproximadamente 30.000 ilhas, a maioria situada perto dos sítios de início da transcrição. A análise molecular demonstrou que as citosinas nessas ilhas raramente, ou nunca, são metiladas e que esse estado não metilado ou submetilado conduz à transcrição. Assim, o DNA na vizinhança de uma ilha de CpG é hipersensível à digestão por DNase I e seus nucleossomos costumam ser um pouco diferentes dos nucleossomos em outras partes do genoma – em geral, há menos histona H1, e algumas histonas centrais são acetiladas.

O DNA metilado, quando presente, está associado à repressão da transcrição. Isso é observado principalmente em fêmeas de mamíferos, cujo cromossomo X inativo tem alto grau de metilação. As regiões do genoma de mamíferos que contêm sequências repetitivas, inclusive as regiões ricas em elementos transponíveis, também são metiladas, talvez como um modo de proteger o organismo contra os efeitos prejudiciais da expressão e do movimento de transpósons. Os mecanismos que tornam o DNA metilado transcricionalmente silencioso não são bem compreendidos; entretanto, pelo menos duas proteínas que reprimem a transcrição ligam-se ao DNA metilado e uma delas, MeCP2, modifica a estrutura da cromatina. Assim, é possível que dinucleotídeos CpG metilados liguem-se a proteínas específicas e que essas proteínas formem um complexo que impeça a transcrição de genes vizinhos.

A transmissão do estado metilado é clonal por divisão celular. Quando uma sequência de DNA é metilada, os

dois filamentos da sequência adquirem grupos metila. Depois da replicação do DNA, cada dúplex-filho tem uma sequência de DNA parental metilada e uma sequência não metilada. As DNA metiltransferases, enzimas que ligam os grupos metila ao DNA, podem reconhecer essa assimetria e acrescentar um grupo metila à sequência não metilada. Desse modo, o estado totalmente metilado é restabelecido nos dúplex-filhos de DNA. Dessa maneira, o padrão de metilação é transmitido de modo mais ou menos fiel a cada ciclo de replicação do DNA – ou seja, a cada divisão celular. Nesse sentido, a metilação do DNA é uma modificação epigenética da cromatina. A acetilação da histona também é considerada uma modificação epigenética, embora não esteja claro como o padrão de acetilação é transmitido por divisão celular. O texto O futuro: A epigenética de gêmeos discorre sobre o possível significado dessas modificações em seres humanos.

## IMPRINTING

A metilação do DNA em mamíferos também é responsável por casos incomuns nos quais a expressão de um gene é controlada por sua origem parental. Por exemplo, em camundongos, o gene *Igf2*, que codifica um fator de crescimento similar à insulina (*insulin-like growth factor*), é expresso quando é herdado do pai, mas não da mãe. Já o gene *H19* é expresso quando é herdado da mãe, mas não do pai. Sempre que a expressão de um gene é condicionada por sua origem parental, os geneticistas afirmam que o gene foi *imprinted* – termo usado para transmitir a ideia de que o gene foi marcado de algum modo para “lembrar” que é oriundo do pai ou da mãe.

Análise molecular recente demonstrou que a marca que condiciona a expressão de um gene é a metilação de um ou mais dinucleotídeos de CpG na vizinhança do gene. A princípio, esses dinucleotídeos metilados são formados na linhagem germinativa parental (Figura 19.13). Assim, por exemplo, o gene *Igf2* é metilado na linhagem germinativa feminina, mas não na linhagem germinativa masculina. Por ocasião da fertilização, um gene *Igf2* metilado de origem materna é combinado a um gene *Igf2* não metilado de origem paterna. Durante a embriogênese, os estados metilado e não metilado são preservados a cada replicação do gene. Como um gene metilado é silencioso, apenas o gene *Igf2* de origem paterna é expresso no animal em desenvolvimento. Acontece exatamente o oposto com o gene *H19*, que é metilado na linhagem germinativa masculina, mas não na linhagem germinativa feminina. Já foram identificados mais de 20 diferentes genes *imprinted* em camundongos e seres humanos. O *imprint* por metilação de cada gene é estabelecido na linhagem germinativa parental. Entretanto, um gene metilado herdado de um sexo pode ser desmetilado quando passa por uma prole do sexo oposto. Assim, os *imprints* por metilação são redefinidos a cada geração, dependendo do sexo do animal. O fato de que alguns genes são metilados em um sexo, mas não no outro, implica que fatores sexo-específicos controlam o mecanismo de metilação.

## FUTURO



### EPIGENÉTICA DE GÊMEOS

Muitos gêmeos humanos têm aparência e atitudes semelhantes, de tal modo que é difícil diferenciá-los. Mas os pais de gêmeos "idênticos" sabem que cada um deles é diferente, e suas diferenças tornam-se mais aparentes com a idade. Um gêmeo pode se tornar autoconfiante e o outro, tímido. Um pode se tornar atleta e o outro, artista. Mais tarde, embora tenham aparência semelhante, um gêmeo pode sucumbir a uma doença crônica, como diabetes, mas o outro, não; na idade avançada, um pode ter doença de Alzheimer e o outro, não. Essas diferenças despertaram nossa curiosidade porque sabemos que esses tipos de gêmeos começaram a vida com genótipos idênticos. O ovócito fertilizado dividiu-se e formou dois embriões, que deram origem a pessoas diferentes. Para ressaltar a origem do mesmo ovócito fertilizado, dizemos que esses gêmeos são **monozigóticos**.

Em 2005, uma equipe de pesquisa internacional explorou a possibilidade de que gêmeos idênticos poderiam ter diferenças epigenéticas.<sup>1</sup> Eles estudaram 40 pares de gêmeos monozigóticos espanhóis. Esses gêmeos tinham de 3 a 74 anos de idade e graus variados de compartilhamento das experiências de vida. Os pesquisadores examinaram dois tipos de modificações epigenéticas na cromatina de leucócitos retirados dos gêmeos: metilação de DNA e acetilação de histona.

A maioria dos pares de gêmeos apresentou perfis epigenéticos espantosamente semelhantes. No entanto, em 35% dos pares, houve diferenças notáveis dos níveis gerais de metilação de DNA e acetilação de histona. Essas diferenças foram mais prevalentes nos pares mais velhos e nos que passaram menos tempo de vida juntos ou que tinham diferentes histórias de saúde. Uma análise mais detalhada das diferenças de metilação do DNA mostrou que cerca de

metade delas estava associada a retrotranspósons no genoma dos gêmeos; a outra metade estava associada a genes conhecidos ou suspeitos. O mapeamento citogenético mostrou que as diferenças eram distribuídas em todo o genoma. Elas foram localizadas nos telômeros dos cromossomos e em determinadas regiões ricas em genes, como os braços longo e curto do cromossomo 1, o braço curto do cromossomo 3 e o braço longo do cromossomo 8. Quando se analisaram os níveis de RNA, as sequências de DNA hipermetiladas eram silenciosas ou subexpressas. Assim, as diferenças epigenéticas entre os gêmeos pareciam ter significado funcional.

Esse estudo – o primeiro desse tipo – demonstrou que gêmeos com o mesmo genótipo podem ter "epigenótipos" diferentes, e isso sugeriu que algumas diferenças fenotípicas entre os gêmeos poderiam ser causadas por diferenças epigenéticas, o que poderia, por sua vez, ser causado por diferentes histórias de vida dos gêmeos. Portanto, esse estudo significa que, com o tempo, as experiências de uma pessoa – alimentação, atividades sociais e físicas, tratamentos clínicos, exposição a diferentes ambientes, e assim por diante – poderiam participar da moldagem do "epigenoma", o que pode influenciar o modo de expressão do genoma.

No outono de 2010, outra equipe internacional se formou para estudar as diferenças epigenéticas em gêmeos. Esse projeto "EpiTwin" é liderado por cientistas do Reino Unido e da China e pesquisa modificações epigenéticas que influenciam a suscetibilidade a vários distúrbios e doenças, como obesidade, diabetes, osteoporose e longevidade. Cinco mil gêmeos serão analisados. Portanto, esse estudo em larga escala pode revelar como a expressão gênica regulada de modo epigenético afeta a etiologia de traços complexos.

<sup>1</sup>Fraga, M. F. et al., 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. USA* 102: 10604-10609.

### PONTOS ESSENCIAIS

- A heterocromatina está associada à repressão da transcrição
- A variação por efeito de posição é um exemplo da regulação epigenética da expressão gênica
- A transcrição ocorre preferencialmente na cromatina com organização frouxa
- O DNA transcricionalmente ativo tende a ser mais sensível à digestão por DNase I
- Durante a ativação da transcrição, a cromatina é remodelada por complexos multiproteicos
- A metilação de DNA está associada ao silenciamento gênico em mamíferos
- A expressão de um gene imprinted é condicionada pela origem parental do gene.

## Ativação e inativação de cromossomos inteiros

Mamíferos, moscas e vermes têm mecanismos diferentes de compensação das diferentes doses de cromossomos X em machos e fêmeas.

Organismos com um sistema de determinação de sexo XX/XY ou XX/XO enfrentam o problema de igualar a atividade de genes ligados ao X nos dois sexos. Nos ma-

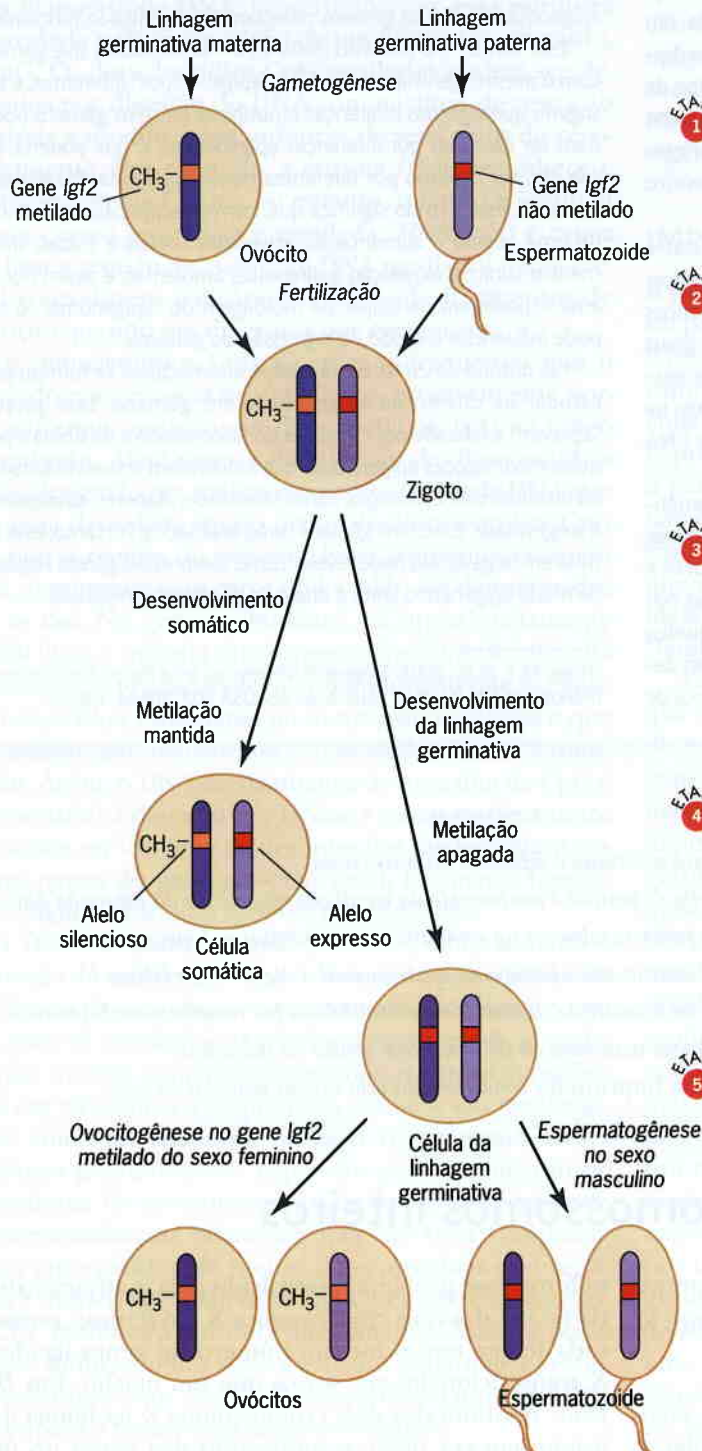
míferos, esse problema é resolvido pela inativação aleatória de um dos dois cromossomos X em fêmeas; portanto, cada fêmea tem o mesmo número de genes ligados ao X transcricionalmente ativos que um macho. Em *Drosophila*, nenhum dos dois cromossomos X na fêmea é inativado; em vez disso, a transcrição dos genes no único cromossomo X do macho é intensificada para que seu



produto final esteja de acordo com os genes nos dois cromossomos X de uma fêmea. Ainda outra solução para o problema de números diferentes de genes ligados ao X foi encontrada no nematódeo *Caenorhabditis elegans*. Nesse organismo, indivíduos XX são hermafroditas (atuam como macho e fêmea) e indivíduos XO são machos. A atividade de transcrição ligada ao X é igualada nesses dois genótipos por repressão parcial dos genes nos dois cromossomos X em hermafroditas. Portanto, mamíferos,

moscas e vermes resolveram o problema da dose de gene ligado ao X de diferentes maneiras (Figura 19.14). Em mamíferos, um dos cromossomos X nas fêmeas é inativado; em *Drosophila*, o único cromossomo X em machos é hiperativado; e em *C. elegans*, os dois cromossomos X em hermafroditas são hipoativados.

Esses três diferentes mecanismos de compensação da dose – inativação, hiperativação e hipoativação – têm uma característica importante em comum: muitos genes diferentes



**ETAPA 1** Alelos do gene *Igf2* são *imprinted* nas linhagens germinativas parentais – metilados na linhagem germinativa feminina e não metilados na linhagem germinativa masculina.

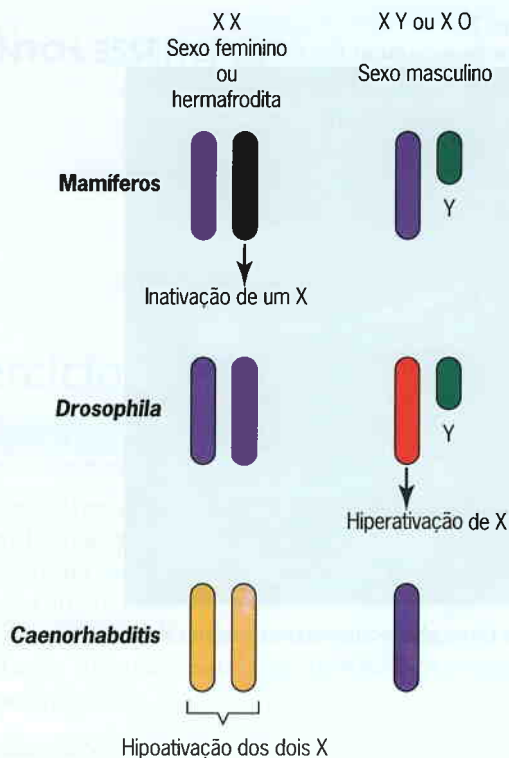
**ETAPA 2** Alelos *imprinted* do gene *Igf2* do pai e da mãe são combinados no zigoto na fertilização.

**ETAPA 3** Durante o desenvolvimento dos tecidos somáticos, o alelo de origem materna permanece metilado enquanto o alelo de origem paterna permanece não metilado. Em células somáticas, só é expresso o alelo não metilado de origem paterna. O alelo metilado de origem materna é silencioso.

**ETAPA 4** Durante o desenvolvimento da linhagem germinativa, o *imprint* por metilação é apagado.

**ETAPA 5** A metilação é restabelecida durante a ovocitogênese, mas não durante a espermatogênese. Assim, se o camundongo for fêmea, todos os genes *Igf2* serão metilados, mesmo que sejam cópias do alelo *Igf2* não metilado do pai. Se o camundongo for macho, nenhum dos genes *Igf2* será metilado, mesmo que seja cópia do alelo *Igf2* metilado herdado da mãe.

■ FIGURA 19.13 Metilação e imprinting do gene *Igf2* em camundongos. O gene é metilado em fêmeas, mas não em machos.

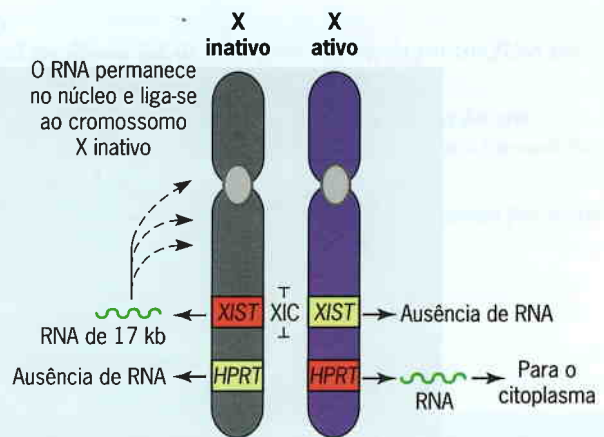


■ FIGURA 19.14 Três mecanismos de compensação de dose para genes ligados ao X: inativação, hiperativação e hipotivação.

são regulados coordenadamente porque estão no mesmo cromossomo. Essa regulação global do cromossomo é superposta a todos os outros mecanismos reguladores usados na expressão espacial e temporal desses genes. Qual seria o responsável por esse sistema regulador global? Durante décadas, os geneticistas vêm tentando elucidar a base molecular da compensação de dose. A hipótese de trabalho é a de que existe um ou mais fatores que se ligam especificamente ao cromossomo X e alteram suas atividades de transcrição. Descobertas recentes indicam que essa ideia está correta.

## INATIVAÇÃO DE CROMOSSOMOS X EM MAMÍFEROS

Em mamíferos, a inativação do cromossomo X começa em um sítio específico denominado *centro de inativação X* (*XIC*) e propaga-se em sentidos opostos até as extremidades do cromossomo. Curiosamente, nem todos os genes do cromossomo X inativado são transcricionalmente silenciosos. Um que permanece ativo é denominado *XIST* (*X inactive specific transcript* [transcrito específico do X inativo]); esse gene está dentro do *XIC* (Figura 19.15). Em seres humanos, o gene *XIST* codifica um transcrito de 17 kb que não tem nenhuma matriz aberta de leitura relevante. Portanto, parece improvável que o gene *XIST* codifique uma proteína. Em vez disso, provavelmente o próprio RNA é o produto funcional do gene *XIST*. Embora poliadenilado, esse RNA é restrito ao núcleo e está localizado especificamente nos cromossomos X inativa-



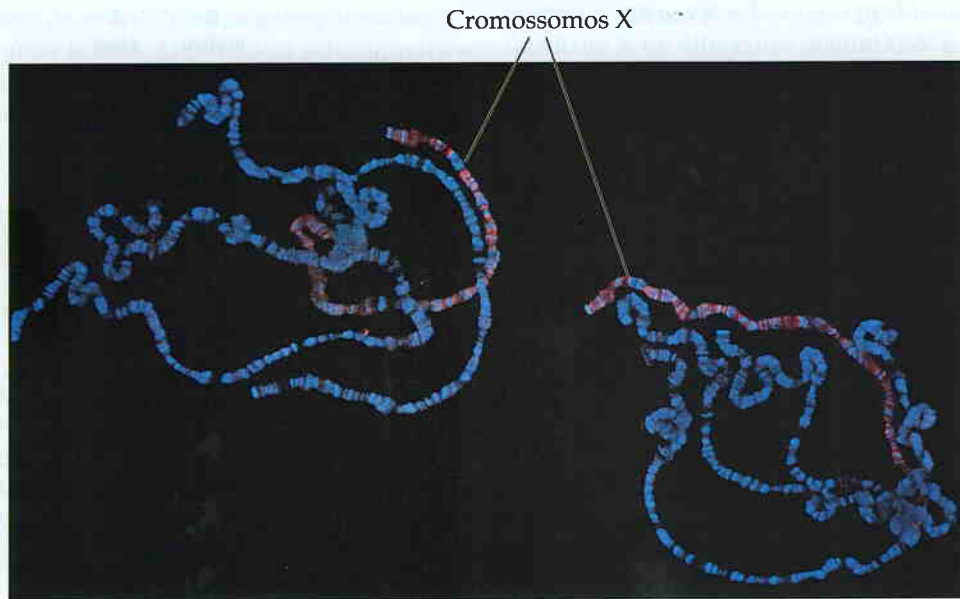
■ FIGURA 19.15 Expressão do gene *XIST* no cromossomo X inativo de mulheres. Para comparação, é mostrada a expressão do gene *HPRT* no cromossomo X ativo. Esse gene codifica a hipoxantina fosforribosil-transferase, enzima que participa do metabolismo de purinas.

dos; não parece estar associado a cromossomos X ativos em machos nem em fêmeas.

Em camundongos, nos quais foi possível fazer uma análise experimental detalhada, os pesquisadores constataram que o homólogo do gene *XIST* humano é transcrito durante os estágios iniciais do desenvolvimento embrionário em baixo nível de ambos os cromossomos X presentes nas fêmeas. Os transcritos de cada gene *Xist* da fêmea de camundongo são instáveis e permanecem intimamente associados a seus respectivos genes. Ao longo do desenvolvimento, os transcritos de um dos genes estabilizam-se e acabam por envolver todo o cromossomo X no qual esse gene está localizado; os transcritos do outro gene *Xist* desintegram-se, e a transcrição adicional desse gene é reprimida por metilação de nucleotídeos no promotor do gene. Assim, na fêmea do camundongo, um cromossomo X – aquele cujo gene *Xist* continua a ser transcrito – é envolvido pelo RNA *Xist* e o outro, não. Ao que tudo indica, a escolha do cromossomo a ser revestido é aleatória. Embora o mecanismo de revestimento ainda não seja compreendido, a consequência é clara: a maioria dos genes no cromossomo revestido é reprimida e esse cromossomo torna-se o cromossomo X inativo. No sistema de compensação de dose de mamíferos, portanto, o cromossomo X que permanece ativo é, paradoxalmente, o que reprime seu gene *Xist*.

Os cromossomos X inativos são facilmente identificados em células de mamíferos. Durante a intérfase, eles se condensam em uma massa escura associada à membrana nuclear. Essa massa, o corpúsculo de Barr, descondensa-se durante a fase S para possibilitar a replicação do cromossomo X inativo. Entretanto, como a descondensação leva algum tempo, a replicação do cromossomo X inativo ocorre mais tarde que a dos demais cromossomos. Desse modo, os cromossomos X inativos têm obrigatoriamente uma estrutura de cromatina muito diferente da estrutura dos outros cromossomos. Essa diferença é determinada em parte pelos tipos de histonas associadas ao DNA. Uma das quatro histonas centrais, H4, pode ser quimicamente





■ FIGURA 19.16 Ligação do produto proteico de um dos genes *mSl* de *Drosophila* ao único cromossomo X de machos.

modificada pelo acréscimo de grupos acetila a qualquer uma das várias lisinas na cadeia polipeptídica. A H4 acetilada está associada a todos os cromossomos do genoma humano. No entanto, no cromossomo X inativo parece estar restrita a três bandas bem estreitas, cada uma delas correspondente a uma região que contém alguns genes ativos. Também há depleção da H4 acetilada em áreas de heterocromatina nos outros cromossomos. Esses achados sugerem que a depleção de H4 acetilada é uma característica básica do cromossomo X inativo.

## HIPERATIVAÇÃO DE CROMOSSOMOS X EM DROSOPHILA

Em *Drosophila*, a compensação de dose requer os produtos proteicos de pelo menos cinco genes diferentes. As mutações nulas nesses genes acarretam letalidade específica do macho porque o único cromossomo X dos machos não é hiperativado. Os machos mutantes costumam morrer durante o fim do estágio larvar ou o início do estágio pupal. Portanto, esses genes de compensação de dose são denominados *loci* letais macho-específicos (*mSl*) e seus produtos são denominados proteínas MSL. Os anticorpos preparados contra essas proteínas foram usados como sondas para localizar as proteínas dentro das células. O achado significativo é que cada proteína MSL liga-se especificamente ao cromossomo X em machos (Figura 19.16). Essas proteínas não se ligam aos outros cromossomos no genoma masculino e não se ligam a nenhum cromossomo, inclusive os X, no genoma feminino. A ligação das proteínas MSL ao cromossomo X masculino é facilitada por dois tipos de moléculas de RNA denominadas *roX1* e *roX2* (do inglês, RNA on the X

chromosome [RNA no cromossomo X]) que são transcritos de genes no cromossomo X.

O modelo atual propõe que as proteínas MSL formam um complexo ao qual se unem os RNA *roX*. Então, esse complexo liga-se a 30 a 40 sítios ao longo do cromossomo X masculino, inclusive aos *loci* que contêm os dois genes *roX*. A partir de cada um desses sítios de entrada, o complexo MSL/*roX* propaga-se nos dois sentidos até alcançar todos os genes no cromossomo X masculino que precisam ser hiperativados. O processo de hiperativação pode implicar remodelagem da cromatina pelo complexo MSL/*roX*. Uma das proteínas MSL é uma histona acetiltransferase, e uma versão acetilada específica da histona H4 está exclusivamente associada aos cromossomos X hiperativados.

## HIPOTATIVAÇÃO DE CROMOSSOMOS X EM CAENORHABDITIS

Em *C. elegans*, a compensação de dose implica a repressão parcial de genes ligados ao X nas células somáticas de hermafroditas. O mecanismo não é totalmente compreendido, mas há participação de produtos de vários genes. Assim como as proteínas MSL em *Drosophila*, as proteínas codificadas por esses genes ligam-se especificamente ao cromossomo X. No entanto, ao contrário da situação em *Drosophila*, elas só se ligam quando há dois cromossomos X. Aparentemente, as proteínas não se ligam ao cromossomo X único em machos nem se ligam a nenhum autossomo em machos ou hermafroditas. Portanto, a compensação de dose em *C. elegans* parece implicar um mecanismo exatamente oposto ao que ocorre em *Drosophila*. Um complexo proteico liga-se aos cromossomos X e reprime a transcrição em vez de estimulá-la.

**PONTOS ESSENCIAIS**

- A inativação de um cromossomo X em fêmeas XX de mamíferos é mediada por um RNA não codificador transcrito do gene XIST nesse cromossomo
- A hiperativação do único cromossomo X em machos de *Drosophila* é mediada por um complexo RNA-proteína que se liga a muitos sítios nesse cromossomo e estimula a transcrição de seus genes
- A hipoativação dos dois cromossomos X em hermafroditas de *C. elegans* é mediada por proteínas que se ligam a esses cromossomos e reduzem a transcrição de seus genes.

**Exercícios****Aplique a análise genética básica**

1. Organize os processos a seguir em ordem cronológica, partindo do que ocorre primeiro: (a) recomposição de uma molécula de RNA, (b) migração de uma molécula de mRNA para o citoplasma, (c) transcrição de um gene, (d) degradação de uma molécula de RNA, (e) síntese de polipeptídeo.

**Resposta:** c-a-b-e-d.

2. Que fator induz a expressão do gene *hsp70* em *Drosophila*?

**Resposta:** O gene *hsp70* é induzido pelo estresse por calor.

3. Indique se cada um destes fenômenos relacionados com a regulação da expressão gênica ocorre no núcleo ou no citoplasma de uma célula eucariótica.
  - (a) Estimulação da expressão gênica por um fator de transcrição.
  - (b) Recomposição alternativa do transcrito primário de um gene.
  - (c) Poliadenilação do transcrito primário de um gene.
  - (d) Tradução de um RNA mensageiro.
  - (e) Inibição da tradução por ligação de um microRNA a um RNA mensageiro.
  - (f) Degradação de um RNA mensageiro induzido por um RNA de interferência curto.
  - (g) Ligação de um hormônio peptídico a seu receptor.
  - (h) Ligação de um hormônio esteroide a seu receptor.
  - (i) Silenciamento da expressão gênica pela heterocromatina.
  - (j) Inativação de todo o cromossomo.

**Resposta:** O item (h) pode ocorrer no citoplasma ou no núcleo, dependendo do hormônio esteroide específico. Os itens (a), (b), (c), (i) e (j) ocor-

rem no núcleo. Todos os outros itens ocorrem no citoplasma.

4. Cite algumas diferenças entre eucromatina e heterocromatina.

**Resposta:** A heterocromatina adquire coloração escura durante todo o ciclo celular; a eucromatina não adquire coloração escura durante a intérfase. A heterocromatina é rica em sequências repetidas de DNA e em elementos transponíveis; a eucromatina pode conter sequências repetidas e transpósons, mas geralmente não no mesmo grau que a heterocromatina. A heterocromatina tem poucos genes codificadores de proteínas; a eucromatina tem muitos genes codificadores de proteínas.

5. Indique se os itens a seguir estão associados à atividade ou inatividade do gene: (a) metilação do DNA, (b) acetilação de histona, (c) metilação de histona, (d) heterocromatina, (e) região de controle de *locus*, (f) proteína GAL4, (g) sensibilidade à DNase I.

**Resposta:** (a) inatividade, (b) atividade, (c) inatividade, (d) inatividade, (e) atividade, (f) atividade, (g) atividade.

6. Qual é o mecanismo de igualação do nível de expressão de genes ligados ao X nos dois sexos em (a) seres humanos, (b) moscas e (c) vermes?

**Resposta:** (a) Em seres humanos, um dos cromossomos X nas fêmeas é inativado aleatoriamente. (b) Em moscas, o único cromossomo X em machos é hiperativado. (c) Em vermes, os dois cromossomos X em hermafroditas são hipoativados.

**Autoavaliação****Integre diferentes conceitos e técnicas**

1. O gene *lacZ* bacteriano para  $\beta$ -galactosidase foi inserido em um elemento *P* transponível de *Drosophila* (Capítulo 17) de modo que pudesse ser transcri-

to do promotor do elemento *P*. Esse gene de fusão foi então injetado na linhagem germinativa de um embrião de *Drosophila* junto com uma enzima que

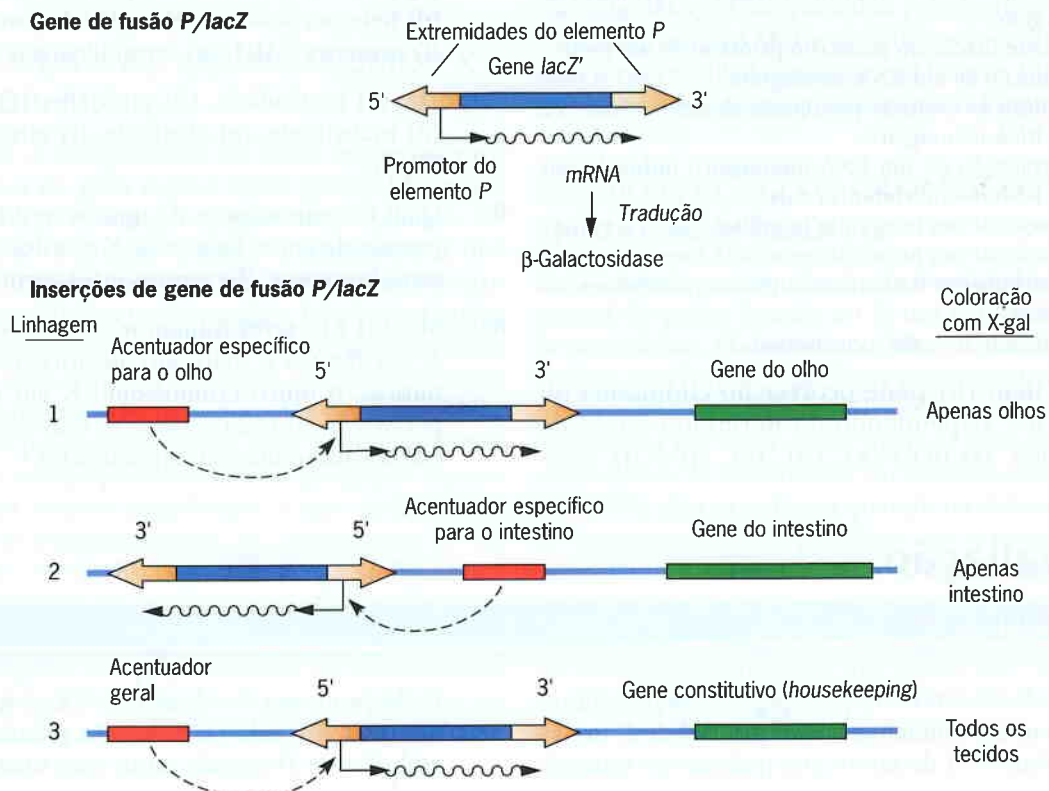


catalisa a transposição de elementos *P*. Durante o desenvolvimento, o elemento *P* modificado foi inserido nos cromossomos de algumas células da linhagem germinativa. Então, a prole desse animal foi cruzada individualmente com moscas de um estoque padrão do laboratório para criar linhagens que tivessem o gene de fusão *P/lacZ* em seus genomas. Em três dessas linhagens, analisou-se a expressão de *lacZ* por coloração de tecidos dissecados de moscas adultas com X-gal, substrato cromogênico que se torna azul na presença de  $\beta$ -galactosidase. Na primeira linhagem, apenas os olhos coraram-se de azul; na segunda, apenas o intestino corou-se de azul; e na terceira, todos os tecidos coraram-se de azul. Como você explica esses resultados?

**Resposta:** Sem dúvida, as três linhagens tinham diferentes inserções do gene de fusão *P/lacZ* (ver diagrama). Em cada linhagem, a expressão do gene de fusão *P/lacZ* foi necessariamente influenciada por uma diferente sequência reguladora, ou acentuador, capaz de interagir com o promotor *P* e iniciar a transcrição no gene *lacZ*. Na primeira linhagem, o elemento *P* modificado certamente se inseriu perto de um acentuador específico para o olho, o que promoveu a transcrição apenas no tecido ocular. Na segunda linhagem, certamente se inseriu perto de um acentuador que promove a transcrição nas células intestinais e, na terceira linhagem, inseriu-se perto de um acentuador que promove a transcrição em todas ou quase todas as células, qualquer que seja a associação tecidual. Provavelmente todos esses diferentes acentuadores estão perto de um gene que

normalmente seria expresso sob seu controle. Por exemplo, o acentuador específico para o olho estaria perto de um gene necessário para algum aspecto da função ou do desenvolvimento ocular. Esses resultados mostram que inserções aleatórias do gene de fusão *P/lacZ* podem ser usadas para identificar diferentes tipos de acentuadores e, por meio deles, os genes que eles controlam. Portanto, essas inserções de gene de fusão frequentemente são denominadas capturas de acentuador (*enhancer traps*).

2. Em seu artigo seminal sobre interferência por RNA, Andrew Fire e colaboradores (1988 *Nature* 391: 806-811) descreveram os resultados de experimentos nos quais o RNA derivado do gene *mex-3* foi injetado em hermafroditas de *C. elegans*. Embriões obtidos desses hermafroditas foram analisados por hibridização *in situ* usando sondas para RNA *mex-3*. As sondas foram projetadas para se ligarem ao RNA mensageiro de *mex-3*, que normalmente se acumula nas gônadas de hermafroditas e em seus embriões. A ligação das moléculas de sonda ao mRNA nos embriões é facilmente detectada se as moléculas de sonda forem marcadas. Ao fazer esses experimentos de hibridização *in situ*, Fire e seus colegas constataram que embriões de vermes nos quais foi injetado RNA *mex-3* bifilamentar não foram marcados pelas moléculas de sonda, enquanto embriões de vermes nos quais foi injetado RNA unifilamentar complementar ao mRNA *mex-3* – isto é, com RNA *mex-3 antisense* – foram marcados, embora não tão intensamente quanto os embriões de vermes que não receberam injeção. O



que esses resultados indicam sobre a eficácia do RNA *antisense* bifilamentar *versus* unifilamentar para silenciar a expressão gênica?

**Resposta:** Os resultados desses experimentos de hibridização *in situ* indicam que o RNA bifilamentar é um forte silenciador da expressão do gene *mex-3* em embriões de *C. elegans*. Em contraposição, o RNA *antisense* unifilamentar quase não tem efeito sobre a expressão do gene *mex-3*. Os embriões de vermes nos quais foi injetado o RNA *mex-3* bifilamentar não apresentam RNA mensageiro *mex-3* detectável. A ausência de RNA mensageiro *mex-3* nesses embriões é consequência de interferência por RNA induzida pelo RNA bifilamentar injetado. Os embriões de vermes nos quais foi injetado o RNA *mex-3 antisense* unifilamentar apresentavam algum RNA mensageiro *mex-3*. Assim, o RNA *mex-3 antisense* unifilamentar não é tão eficaz quanto o RNA *mex-3* bifilamentar na indução de RNAi.

3. O fenótipo desigual de gatos com pelagem tartaruga (Capítulo 5) é consequência da inativação aleatória de cromossomos X em fêmeas heterozigotas para diferentes alelos de um gene ligado ao X para cor da pelagem; um alelo produz pelagem clara e o outro, pelagem escura. O fenótipo desigual de ginandromorfos em *Drosophila* (Capítulo 6) é consequência da não disjunção dos cromossomos X durante uma das divisões iniciais de clivagem. Se um zigoto XX for heterozigoto para os alelos selvagem e mutante do

gene *white* ligado ao X, a não disjunção pode produzir uma linhagem de células XO que têm apenas o alelo mutante, e se essas células formam um olho, ou parte de um olho, esse tecido ocular será branco. Em contraposição, o tecido derivado de células XX será vermelho porque essas células têm o alelo selvagem do gene *white*. Algum desses fenótipos desiguais é um exemplo da regulação epigenética da expressão gênica? Explique sua resposta.

**Resposta:** O fenótipo desigual de gatos com pelagem tartaruga é consequência de um fenômeno epigenético – inativação aleatória de um cromossomo X em cada célula destinada a formar células produtoras de pigmento no adulto. Todas as células produtoras de pigmento são geneticamente equivalentes – isto é, têm o mesmo conteúdo de DNA. O fenótipo tartaruga não é causado por alteração do genótipo durante o desenvolvimento embriológico do animal. Na verdade, a causa é a mudança no estado de um dos cromossomos X, o X que é inativado, e a herança desse estado é clonal por divisão celular. Assim, as manchas claras e escuras da pelagem do gato têm diferenças epigenéticas, não genéticas. Já o fenótipo desigual de ginandromorfos de *Drosophila* deve-se a uma mudança genética ocorrida durante o desenvolvimento. Um dos cromossomos X é perdido. As manchas vermelhas e brancas de tecido no olho de um ginandromorfo não são geneticamente equivalentes. Portanto, a diferença entre elas é genética, não epigenética.

## Avaliação adicional

### Entenda melhor e desenvolva a capacidade analítica

- 19.1 Os óperons são comuns em bactérias, mas não em eucariotos. Sugira uma razão para isso.
- 19.2 Em bactérias, a tradução de um mRNA começa antes da conclusão da síntese desse mRNA. Por que esse “acoplamento” da transcrição e tradução não é possível em eucariotos?
- 19.3 A distrofia muscular em seres humanos é causada por mutações em um gene ligado ao X que codifica uma proteína denominada distrofina. Que técnicas poderiam ser usadas para verificar se esse gene é ativo em diferentes tipos de células, por exemplo, células cutâneas, células nervosas e células musculares?
- 19.4 Por que os hormônios esteroides interagem com receptores dentro da célula, enquanto hormônios peptídicos interagem com receptores na superfície celular?
- 19.5 Nos cromossomos politênicos de larvas de *Drosophila* (Capítulo 6), há formação de grandes *pufes* em algumas bandas quando as larvas são expostas a temperatura elevada. Como você poderia demonstrar que esses *pufes* contêm genes que são intensamente transcritos em resposta ao tratamento de choque térmico?
- 19.6 Como é possível distinguir um acentuador de um promotor?
- 19.7 As tropomiosinas são proteínas mediadoras da interação de actina e troponina, duas proteínas que participam das contrações musculares. Em animais superiores, as tropomiosinas existem como uma família de proteínas intimamente relacionadas que têm algumas sequências de aminoácidos em comum e outras diferentes. Explique como essas proteínas poderiam ser criadas a partir do transcrito de um único gene.
- 19.8 Um polipeptídeo é constituído de três segmentos de aminoácidos, A–B–C. Outro polipeptídeo contém os segmentos A e C, mas não o segmento B. Como seria possível determinar se esses dois polipeptídeos são produzidos por tradução de versões alternativamente recompostas de RNA de um único gene ou por tradução de mRNA de dois genes diferentes?



- 19.9** Que técnicas poderiam ser usadas para mostrar que um gene vegetal é transcrito quando a planta é iluminada?
- 19.10** Quando os íntrons foram descobertos pela primeira vez, acreditava-se que fossem o “lixo” genético – isto é, sequências sem função útil. Na verdade, pareciam ser piores que lixo porque interrompiam as sequências codificadoras dos genes. Entre eucariotos, porém, os íntrons são disseminados e tudo que é disseminado em biologia geralmente tem uma função. Que função poderiam ter os íntrons? Que benefício eles poderiam conferir a um organismo?
- 19.11** O fator de transcrição GAL4 em leveduras regula dois genes adjacentes, *GAL1* e *GAL10*, por ligação a sequências de DNA entre eles. Esses dois genes são transcritos em sentidos opostos no cromossomo, um para a esquerda do sítio de ligação da proteína GAL4 e o outro para a direita desse sítio. Que propriedade dos acentuadores essa situação ilustra?
- 19.12** Usando as técnicas de engenharia genética, um pesquisador construiu um gene de fusão contendo os elementos de resposta ao choque térmico de um gene *hsp70* de *Drosophila* e a região codificadora de um gene de água-viva (*gfp*) para proteína fluorescente verde. Esse gene de fusão foi inserido nos cromossomos de *Drosophila* vivas pela técnica de transformação mediada por transpóson (Capítulo 17). Em que condições a proteína fluorescente verde será sintetizada nessas moscas transformadas geneticamente? Explique.
- 19.13** Suponha que o segmento do gene *hsp70* usado para produzir o gene de fusão *hsp70/gfp* do problema anterior tivesse mutações em todos os seus elementos de resposta ao choque térmico. A proteína fluorescente verde codificada por esse gene de fusão seria sintetizada nessas moscas transformadas geneticamente?
- 19.14** Os produtos polipeptídicos de dois genes diferentes, *A* e *B*, atuam como fatores de transcrição. Esses polipeptídios interagem e formam dímeros: homodímeros AA, homodímeros BB e heterodímeros AB. Se os polipeptídios *A* e *B* são igualmente abundantes nas células e se a formação de dímeros é aleatória, qual é a razão esperada entre homodímeros e heterodímeros nessas células?
- 19.15** Um determinado fator de transcrição liga-se a acentuadores em 40 genes diferentes. Preveja o fenótipo de indivíduos homocigotos para uma mutação por mudança de matriz de leitura na sequência codificadora do gene que especifica esse fator de transcrição.
- 19.16** As formas alternativamente recompostas do RNA do gene *doublesex* de *Drosophila* codificam proteínas necessárias para bloquear o desenvolvimento de um ou outro conjunto de características sexuais.
- A proteína produzida nas fêmeas bloqueia o desenvolvimento de características masculinas, e a proteína produzida em machos bloqueia o desenvolvimento de características femininas. Preveja o fenótipo de animais XX e XY homocigotos para uma mutação nula no gene *doublesex*.
- 19.17** O RNA do gene *Sex-lethal* (*Sxl*) de *Drosophila* é recomposto alternativamente. Em machos, a sequência do mRNA derivado do transcrito primário contém todos os oito éxons do gene *Sxl*. Em fêmeas, o mRNA só contém sete éxons porque, durante a recomposição, o éxon 3 é removido do transcrito primário junto com seus íntrons flanqueadores. Portanto, a região codificadora no mRNA feminino é mais curta que no mRNA masculino. No entanto, a proteína codificada pelo mRNA feminino é mais longa que a proteína codificada pelo mRNA masculino. Como você explicaria esse paradoxo?
- 19.18** Em *Drosophila*, a expressão do gene *yellow* é necessária para a formação de pigmento escuro em muitos tecidos diferentes; sem essa expressão, a cor do tecido é amarela. Nas asas, a expressão do gene *yellow* é controlada por um acentuador em posição 5' em relação ao sítio de início da transcrição do gene. Nas garras tarsais, a expressão é controlada por um acentuador localizado no único íntron do gene. Suponha que, por técnicas de engenharia genética, o acentuador da asa seja posto no íntron e o acentuador da garra, em posição 5' em relação ao sítio de início da transcrição. Uma mosca que tivesse esse gene *yellow* modificado em vez do gene *yellow* natural teria asas e garras escuras? Explique.
- 19.19** Um pesquisador suspeita de que um íntron com 550 pb contenha um acentuador que promova a expressão específica de um gene de *Arabidopsis* no tecido da extremidade da raiz. Descreva um experimento para testar essa hipótese.
- 19.20** Qual é a natureza de cada uma das seguintes classes de enzimas? Qual é a ação de cada tipo de enzima sobre a cromatina? (a) HAT, (b) HDAC e (c) HMT.
- 19.21** Em larvas de *Drosophila*, o único cromossomo X de machos parece difuso e inflado nas células politênicas das glândulas salivares. Essa observação é compatível com a ideia de que os genes ligados ao X são hiperativados em machos de *Drosophila*?
- 19.22** Suponha que a LCR do agrupamento de genes da  $\beta$ -globina fosse deletada de um dos dois cromossomos 11 de um homem. Que doença essa deleção poderia causar?
- 19.23** O RNA bifilamentar derivado de um íntron seria capaz de induzir interferência por RNA?
- 19.24** A interferência por RNA foi implicada na regulação de elementos transponíveis. Em *Drosophila*, duas das principais proteínas implicadas na interferên-

cia por RNA são codificadas pelos genes *aubergine* e *piwi*. Moscas homozigotas para alelos mutantes desses genes são letais ou estéreis, mas moscas heterozigotas para esses alelos são viáveis e férteis. Suponha que você tenha linhagens de *Drosophila* heterozigotas para alelos mutantes *aubergine* ou *piwi*. Por que a taxa de mutação genômica nessas linhagens mutantes poderia ser maior que a taxa de mutação genômica em uma linhagem selvagem?

- 19.25** Suponha que fêmeas de camundongos homozigotas para o alelo *a* do gene *Igf2* sejam cruzadas com machos homozigotos para o alelo *b* desse gene. Qual desses dois alelos será expresso na prole  $F_1$ ?
- 19.26** A transmissão dos estados epigenéticos é clonal por divisão celular. Que tipos de observações indicam que esses estados podem ser revertidos ou redefinidos?
- 19.27** Um pesquisador formula a hipótese de que em camundongos o gene *A* é transcrito ativamente em células hepáticas, enquanto o gene *B* é transcrito ativamente em células encefálicas. Descreva procedimentos que possibilitariam que o pesquisador testasse essa hipótese.
- 19.28** Suponha que a hipótese mencionada na questão anterior esteja correta e que o gene *A* seja transcrito ativamente em células hepáticas e o gene *B* seja transcrito ativamente em células encefálicas. Agora, o pesquisador extrai quantidades equivalentes dos tecidos hepático e encefálico e trata esse extrato separadamente com DNase I durante um período limitado. Se o DNA remanescente após o tratamento for fracionado por eletroforese em gel, transferido para uma membrana por *Southern blotting* e hibridizado com uma sonda marcada radioativamente específica para o gene *A*, que mostra (figado ou encéfalo) deve exibir maior sinal na autorradiografia? Justifique sua resposta.
- 19.29** Por que as mutações nulas no gene *msl* em *Drosophila* não afetam as fêmeas?
- 19.30** Suponha que uma mulher tenha um cromossomo X com deleção do locus *XIST*. O outro cromossomo X tem um locus *XIST* intacto. Que padrão de inativação de X seria observado em todo o corpo da mulher?
- 19.31** Em *Drosophila*, o fenótipo desigual do alelo *white mottled* é suprimido por uma mutação autossômica dominante que elimina a função do gene da proteína 1 da heterocromatina (HP1), um fator importante na formação da heterocromatina. Moscas com o alelo *white mottled* e a mutação supressora têm cor vermelha quase uniforme nos olhos; sem a mutação supressora, os olhos são mosaicos de tecido vermelho e branco. Sugira uma explicação para o efeito da mutação supressora.
- 19.32** A ovelha Dolly (Capítulo 2) foi o primeiro mamífero clonado. Dolly foi criada por implantação do núcleo de uma célula retirada do úbere de uma ovelha em um ovócito enucleado. Esse núcleo tinha dois cromossomos X, e, por ser originado de uma célula diferenciada, um deles foi necessariamente inativado. Na célula do úbere heterozigota para no mínimo um gene ligado ao X cuja expressão você poderia analisar, como se poderia determinar se todas as células de Dolly tinham o mesmo cromossomo X inativado? Se, na análise, as células de Dolly mostrassem ser mosaicos da atividade do cromossomo X – isto é, atividade de diferentes cromossomos X em diferentes clones de células – o que teria acontecido durante o desenvolvimento embriológico?

## Genômica na Web em <http://www.ncbi.nlm.nih.gov>

Os genes da  $\beta$ -globina humana estão localizados em um agrupamento no braço curto do cromossomo 11.

1. Pesquise o gene homônimo do agrupamento, o gene da  $\beta$ -globina de adulto, no banco de dados do genoma humano. Qual é o símbolo oficial desse gene? Quantos éxons ele contém?
2. Use a função *Map Viewer* para localizar o agrupamento do gene da  $\beta$ -globina no ideograma do cromossomo 11. Em que banda citológica ele está localizado? Está mais próximo do telômero do braço curto ou do centrômero?
3. Use o *Sequence Viewer* para examinar em detalhes o gene da  $\beta$ -globina do adulto. O gene transcrito está próximo do centrômero ou do telômero? Qual é o comprimento do transcrito do gene? Qual é o comprimento do mRNA maduro? Quantos aminoácidos o mRNA específica? Quais são os três primeiros aminoácidos e por que códons são especificados?
4. Obtenha a sequência de texto do gene da  $\beta$ -globina do adulto clicando no botão ATGC na página *Sequence Viewer*. Localize o códon de iniciação para metionina no primeiro éxon. Como a sequência mostrada na janela é a do filamento-molde do DNA, a leitura desse códon é 5'-CAT-3' da esquerda para a direita na tela.
5. GATA1 e MyoD são dois fatores de transcrição que reconhecem sequências curtas nos genomas de mamíferos. A sequência reconhecida por GATA1 é 5'-TGATAG-3', e a sequência reconhecida por MyoD é 5'-CAAATG-3'. Copie a sequência da porção transcrita do gene da  $\beta$ -globina do adulto em um arquivo de texto e procure cada uma dessas sequências de reconhecimento. Onde estão localizadas? Qual desses dois fatores de transcrição poderia participar da regulação da expressão do gene da  $\beta$ -globina do adulto?