

# Amplificação do DNA: Clonagem de DNA Utilizando a Célula como Base e PCR

# 6

## CONCEITOS PRINCIPAIS

- As moléculas de DNA são muito grandes e complexas e possuem tendência a mudança de forma quando isoladas de outros constituintes celulares, tornando difícil a sua purificação.
- Clonagem de DNA significa fazer múltiplas cópias idênticas (clones) de uma sequência de DNA de interesse, o DNA-alvo; o aumento no número de cópias é chamado de amplificação. O processo exige uma DNA-polimerase para replicar uma sequência de DNA-alvo repetidamente, tanto em células vivas como *in vitro*.
- Em clonagens que utilizam células como base, o DNA-alvo é fracionado por meio da transferência para células bacterianas ou de levedura (transformação). Cada célula transformada mantém, geralmente, apenas uma molécula de DNA-alvo e a replica usando a DNA-polimerase da célula hospedeira.
- Antes de ser transferido para as células, um DNA-alvo é ligado a moléculas de DNA-vetor, um plasmídeo ou um fago de DNA com capacidade de autoduplicação dentro de células. Os complexos de DNA-alvo-vetor são conhecidos como DNA recombinante.
- As endonucleases de restrição clivam as moléculas de DNA em sequências de identificação curtas e definidas e são necessárias para preparar o DNA-alvo e o vetor para sua união.
- Após serem transferidos para as células, os DNAs recombinantes são, geralmente, duplicados independentemente do(s) cromossomo(s) da célula hospedeira.
- DNAs-alvo longos são mais estáveis nas células se carregados por vetores cujo número de cópias seja restrito a um ou dois por célula.
- Megabases de DNA podem ser clonadas em levedura pela adição de elementos de sequências extremamente curtas, necessárias para o funcionamento do cromossomo, de modo a formar um cromossomo artificial linear.
- Em clonagem de expressão, o DNA-alvo é projetado para ser expresso. Ele é convertido em RNA e, em diversos casos, traduzido em uma proteína recombinante.
- A clonagem de DNA utilizando células como base é lenta e laboriosa; entretanto, pode produzir quantidades extremamente grandes de DNA clonado e é a primeira opção para a clonagem de sequências de DNA longas, genes expressos e para a construção de conjuntos amplos de clones de DNA (bibliotecas de DNA).
- A clonagem de DNA também pode ser realizada rapidamente *in vitro* por meio da reação em cadeia da polimerase (PCR). Geralmente, uma DNA-polimerase purificada é utilizada para replicar de forma seletiva uma sequência específica de DNA dentro de uma amostra complexa de DNA. Oligonucleotídeos são projetados para se ligarem a sítios específicos na amostra de DNA que flanqueiam a sequência-alvo, permitindo que a polimerase inicie a replicação do DNA nestes sítios.
- A PCR também pode ser utilizada para amplificar diversos fragmentos de DNA-alvo diferentes simultaneamente usando, por exemplo, misturas complexas de numerosas sequências de *primers* diferentes. Essa amplificação indiscriminada pode enriquecer muitas das sequências contidas em amostras preciosas onde exista pouco DNA inicial.
- Além de ser um procedimento de clonagem de DNA, a PCR também é amplamente utilizada como um teste para quantificar DNA e RNA após ter sido convertida pela transcriptase reversa em DNA complementar (cDNA). A PCR em tempo real é o método mais eficiente e quantifica produtos de DNA enquanto a reação de PCR está ocorrendo.

Os fundamentos da tecnologia de DNA corrente são largamente baseados em duas abordagens bastante diferentes de estudo de sequências específicas de DNA dentro de uma população de DNA complexa, como listado a seguir (ver também Figura 6.1):

- **Clonagem de DNA.** Sequências de DNA de interesse são *seletivamente replicadas* de alguma maneira para produzir números extremamente grandes de cópias idênticas (**clones**). O vasto aumento no número de cópias resultante (**amplificação**) significa que a clonagem do DNA é, efetivamente, uma maneira de purificar uma sequência de DNA desejada de maneira que ela possa ser amplamente estudada.
- **Hibridização molecular.** O fragmento de interesse não é amplificado ou purificado de qualquer maneira; em vez disso, é *especificamente detectado* dentro de uma mistura complexa de diversas sequências diferentes. Isto será abordado no Capítulo 7.

Antes da clonagem do DNA, o conhecimento sobre o DNA era extremamente limitado. A tecnologia de clonagem do DNA mudou tudo isso e revolucionou o estudo da genética. Em comparação com sequências proteicas, as moléculas de DNA são extremamente grandes e complexas – moléculas individuais de DNA nuclear muitas vezes contêm centenas de milhões de nucleotídeos. Quando o DNA é isolado de células por métodos padrões, aquelas moléculas imensas são fragmentadas por forças de cisalhamento, originando misturas complexas de fragmentos de DNA ainda bastante grandes (cerca de 50 a 100 kb de comprimento com métodos padrões de isolamento de DNA). Devido à complexidade do DNA isolado das células de um eucarioto típico, o desafio era saber como sua análise seria feita.

Para vários eucariotos, uma abordagem inicial havia sido separar diferentes populações de sequências de DNA por centrifugação. A ultracentrifugação de equilíbrio de gradiente de densidade (p. ex., gradientes de densidade CsCl) geralmente fraciona o DNA de uma célula eucariótica em uma banda principal, representando a massa de DNA (*bulk*) mais diversas bandas satélites menores que são compostas de classes de DNA repetitivo. As moléculas de DNA das bandas-satélite (*DNA satélite*) possuem densidades flutuantes diferentes daquelas da massa de DNA (e entre si) porque elas consistem em arranjos muito grandes de repetições curtas em *tandem* cuja composição de bases é significativamente diferente daquela da massa ou *bulk* de DNA. Descobriu-se que as sequências de DNA satélite estão envolvidas em aspectos específicos da estrutura e função cromossômicas. Embora valiosas e interessantes, os DNAs satélite purificados eram um componente menor do genoma e não continham genes.

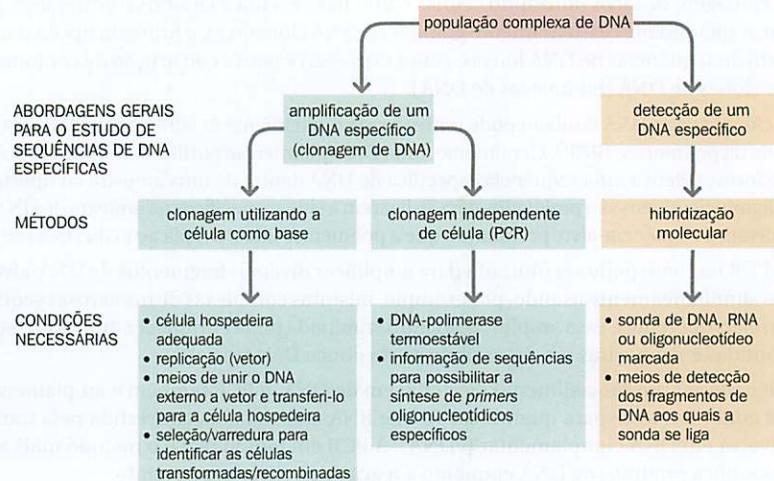
Eram necessários métodos mais gerais que permitissem que qualquer sequência de DNA fosse purificada, não apenas sequências de DNA cuja composição das bases desviasse significativamente daquela do DNA total. Duas metodologias principais de **clonagem de DNA** foram desenvolvidas para tornar isso possível. Não só utilizando DNA-polimerases para fazer cópias múltiplas de sequências de DNA (clones de DNA), como também permitindo, a princípio, a purificação de qualquer sequência de DNA dentro de uma população de DNA inicial complexa.

**Figura 6.1** Abordagens gerais para o estudo de sequências específicas de DNA em populações complexas de DNA.

As moléculas de DNA são muito grandes e complexas, e qualquer gene, éxon ou outra sequência de DNA de interesse, em geral, representa uma pequena fração do DNA inicial que se pode isolar das células. Duas abordagens bastante diferentes podem ser utilizadas para se estudar uma sequência de DNA de interesse. Uma maneira é a purificação de uma sequência específica de DNA por replicação seletiva desta sequência (clonagem de DNA), aumentando, assim, o seu número de cópias (amplificação). O DNA específico pode ser clonado dentro de células de bactérias ou de leveduras pela utilização de DNA-polimerases das células ou de DNA-polimerases purificadas *in vitro*, como na reação em cadeia da polimerase (PCR). Uma abordagem alternativa não envolve a purificação da sequência de DNA, mas, em vez disso, visa detectar essas sequências de forma específica. Esta abordagem envolve a marcação de uma sequência de ácido nucleico previamente isolada a qual é relacionada à sequência de DNA desejada, de maneira que ela possa reconhecer esta sequência específica em uma reação de hibridização molecular.

## Clonagem utilizando células como base

A clonagem utilizando a célula como base foi desenvolvida no início dos anos 1970 e usa as células para fracionar uma amostra complexa de DNA. A técnica consiste na clivagem



da amostra de DNA em pequenos fragmentos e no transporte destes fragmentos em células de bactéria ou de levedura apropriadas. Cada célula hospedeira mantém, geralmente, apenas um fragmento, o qual pode ser replicado diversas vezes dentro da célula pela DNA-polimerase da própria célula hospedeira. Após isso, a célula pode sofrer várias rodadas de divisão celular para produzir um grande número de células com cópias idênticas de apenas um pequeno fragmento da amostra de DNA.

A clonagem utilizando células como base revolucionou a genética e tornou possível a purificação e o estudo de qualquer sequência de DNA. Finalmente, esta técnica permitiu que um vasto número de clones fosse produzido, de forma que o genoma como um todo, bem como seus transcritos de RNA, pudessem ser sequenciados. Com isso, também se tornou possível a produção de proteínas purificadas e RNAs funcionais específicos, de maneira que eles pudessem ser estudados em nível básico, possibilitando várias aplicações médicas e biotecnológicas. Os princípios da clonagem utilizando células como base e os sistemas básicos de clonagem serão considerados na Seção 6.1, e sistemas mais avançados de clonagem nas Seções 6.2 e 6.3.

### A reação em cadeia da polimerase (PCR)

A PCR é um método rápido de clonagem de DNA que não utiliza células e que foi desenvolvido em meados dos anos 1980. A técnica utiliza DNA-polimerases purificadas para replicar sequências de DNA definidas (dentro de uma população complexa de DNA inicial) de forma seletiva. Como este é um método de amplificação de uma sequência específica de DNA extremamente rápido, ele permite uma varredura acelerada de milhares de amostras por cada vez. Devido a isso, a técnica possui diversas aplicações, tanto em pesquisas básicas como em aplicadas; as aplicações médicas são baseadas em seu uso como uma ferramenta de varredura rápida do DNA.

A PCR é um método de clonagem de DNA forte e sensível e tem sido amplamente utilizado em ciência forense bem como na amplificação de DNA de tecidos recuperados de lugares históricos ou arqueológicos. Além de ser um método de clonagem de DNA, a PCR também é amplamente utilizada para quantificar tanto DNA como RNA em vários testes, em pesquisas básicas e aplicadas. A PCR será considerada na Seção 6.4.

## 6.1 PRINCÍPIOS DA CLONAGEM DE DNA UTILIZANDO CÉLULAS COMO BASE

A clonagem de DNA utilizando células como base é conduzida principalmente em células bacterianas e se tornou possível graças à descoberta das **endonucleases de restrição** do tipo II. As endonucleases de restrição servem para proteger a bactéria da invasão de patógenos, principalmente vírus. Após reconhecer elementos de sequência curtos *específicos* no DNA externo, as endonucleases de restrição clivam esse DNA “estrangeiro” nas proximidades de cada um destes elementos, mas o DNA da célula bacteriana não é clivado, pois o mesmo é protegido por metilação; o **Quadro 6.1** fornece mais detalhes e descreve as diferentes classes de enzimas de restrição.

As endonucleases do tipo II oferecem duas grandes vantagens aos geneticistas moleculares. Primeiro, elas permitem que moléculas complexas de DNA sejam clivadas em fragmentos definidos de tamanho manejável. Segundo, elas auxiliam os fragmentos de DNA resultantes a unirem-se às moléculas de **vetores** similarmente clivadas, as quais portam uma origem de replicação que possibilita sua autoduplicação nas células. As moléculas de DNA híbridas resultantes são conhecidas como **DNAs recombinantes**. Isso completa o primeiro de quatro passos essenciais na clonagem de DNA utilizando células como base (**Figura 6.2A**).

O segundo passo essencial na clonagem de DNA utilizando células como base envolve o fracionamento de uma mistura complexa de diferentes moléculas de DNA recombinante, transferindo essas moléculas para células hospedeiras adequadas, geralmente de bactéria ou levedura. Esse processo é conhecido como **transformação**, e cada célula hospedeira normalmente porta apenas *um* tipo de DNA recombinante (**Figura 6.2B**). As células as quais portam um DNA recombinante são conhecidas como **recombinantes**.

São fornecidas condições para que as células transformadas cresçam e se multipliquem por divisões celulares repetidas. Geralmente, durante este tempo o DNA recombinante é replicado de forma independente do(s) cromossomo(s) da célula hospedeira. Devido ao fato de que uma célula transformada normalmente contém apenas um tipo de DNA recombinante, a divisão celular repetida gera uma série de células idênticas (**clones celulares**) que contém o mesmo tipo de DNA recombinante. Populações de clones celulares que surgem de diferentes células transformadas podem ser fisicamente separadas por um procedimento

**QUADRO 6.1 Endonucleases de restrição e modificação de sistemas de restrição**

As endonucleases de restrição (também chamadas de enzimas de restrição) são uma classe de endonucleases bacterianas que clivam DNA de dupla-fita em ambas as fitas após ter reconhecido elementos de sequência curtos específicos no DNA. Sempre que uma enzima de restrição encontra sua sequência de reconhecimento específica, ela pode clivar o DNA dentro da sequência de reconhecimento ou próximo a ela.

Diferentes linhagens bacterianas produzem diferentes enzimas de restrição, e mais de 250 tipos diferentes de especificidade de sequências já foram identificados (ver REBASE, Banco de Dados de Enzimas de Restrição em <http://rebase.neb.com/rebase/rebase.html>). A convenção para nomeação de enzimas de restrição é começar com três letras, uma maiúscula para a primeira letra do gênero bacteriano seguida pelas duas primeiras letras da espécie. Após isso, uma letra é utilizada para denotar a linhagem bacteriana, e, finalmente, um número romano é utilizado para distinguir diferentes enzimas de restrição da mesma linhagem. Por exemplo, *EcoRI*, a qual reconhece a sequência GAATTC, denota a primeira enzima de restrição a ser relatada para *Escherichia coli*, linhagem RY13.

A função natural das enzimas de restrição é proteger as bactérias de patógenos, principalmente bacteriófagos (vírus que matam bactérias – conhecidos como fagos, para abreviação). Elas fazem isso por clivagem seletiva do DNA de patógenos invasores; o DNA bacteriano é protegido da clivagem porque é submetido à metilação sítio-específica por metiltransferases de DNA sequência-específicas da célula hospedeira (*modificação*).

Uma linhagem bacteriana particular produzirá uma metiltransferase de DNA com a *mesma* especificidade de sequência que a enzima de restrição produzida por aquela linhagem. Por exemplo, a *E. coli* linhagem RY13 produz a metiltransferase *EcoRI*, que reconhece especificamente a sequência GAATTC e metila a adenosina central em ambas as fitas. Como resultado, o DNA das células hospedeiras é protegido da clivagem pela enzima de restrição *EcoRI* (os grupos metila se sobressaem dentro da cavidade maior

do DNA nos sítios de ligação, prevenindo, desta maneira, que a enzima de restrição aja no mesmo). A endonuclease *EcoRI*, entretanto, clivará sequências GAATTC não metiladas, tais como as do DNA do fago invasor, as quais não foram previamente modificadas pela metilase *EcoRI*.

Conjuntamente, uma enzima de restrição e sua(s) enzima(s) cognata(s) de modificação formam um sistema de restrição-modificação (R-M). Em alguns sistemas R-M, as duas atividades enzimáticas ocorrem em subunidades separadas ou domínios separados de uma enzima grande que combina restrição e modificação. As enzimas de restrição têm sido classificadas em diversos subgrupos amplos de acordo com a natureza do seu sistema R-M e outras características. Os grupos mais definidos são listados a seguir.

- As enzimas de restrição do tipo I possuem múltiplas subunidades, combinando enzimas de restrição-modificação que clivam em regiões distantes de suas sequências de reconhecimento em localizações variáveis.
- As enzimas de restrição do tipo II são amplamente utilizadas na manipulação e análise do DNA. Elas são, normalmente, endonucleases de restrição de subunidade única, cujas sequências de reconhecimento possuem, frequentemente, de 4 a 8 pb; elas clivam em posições definidas tanto dentro da sequência de reconhecimento como muito próximo a ela. As enzimas de tipo IIR clivam dentro das suas sequências de reconhecimento, as quais são frequentemente *palíndromos* (a sequência na direção 5'→3' é a mesma em ambas as fitas). As enzimas tipo IIS clivam fora de sequências de reconhecimento assimétricas. As enzimas de tipo IIB possuem sequências de reconhecimento bipartidas. Ver Tabela 6.1 para exemplos.
- As enzimas de restrição do tipo III possuem múltiplas subunidades, combinando enzimas de restrição-modificação que reconhecem duas sequências não palindrômicas inversamente orientadas. Elas clivam o DNA cerca de 20-30 pb fora da sequência de reconhecimento.

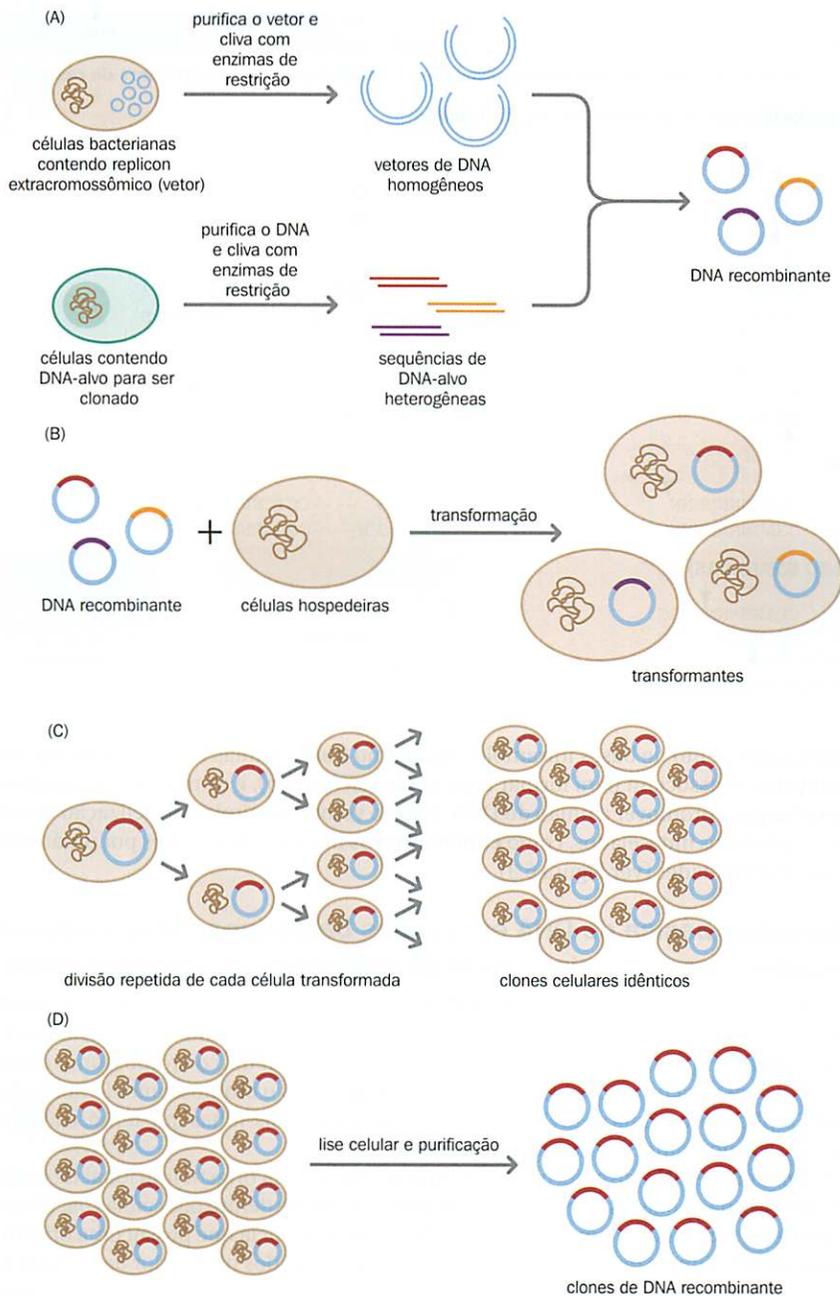
conhecido como *plaqueamento*. Para isso, a mistura de células transformadas é espalhada em uma superfície de ágar, e elas podem crescer para formar colônias de células bem separadas, cada uma se originando de uma única célula transformada. Colônias celulares individuais podem ser escolhidas e passar por uma segunda fase de crescimento em cultura líquida, resultando em um número bastante grande de clones celulares (**Figura 6.2C**).

O passo final é a captura de células das culturas celulares expandidas e a purificação do DNA recombinante (**Figura 6.2D**). As seções a seguir considerarão estes passos mais detalhadamente.

### Fragmentos de DNA-alvo manuseáveis são unidos a moléculas de vetores por meio da utilização de endonucleases de restrição e da DNA-ligase

Das diferentes classes de endonucleases de restrição (ver Quadro 6.1), as enzimas do tipo I não produzem fragmentos de restrição discretos, e as enzimas do tipo III não clivam em todas as sequências de reconhecimento possíveis e, raramente, originam dígestos completos. As enzimas de maior valor para manipulação e análise do DNA são as endonucleases de restrição do tipo II. Elas clivam o DNA em posições definidas dentro de suas sequências de reconhecimento ou muito próximo a elas e produzem, assim, fragmentos de restrição definidos.

Várias endonucleases de restrição do tipo II clivam dentro de sequências que são *palíndromos* (a sequência de bases é a mesma em ambas as fitas quando lidas na direção 5'→3', como resultado de um eixo de simetria dupla). Os fragmentos de restrição resultantes podem ter extremidades sem corte devido a seus pontos de clivagem ocorrerem exatamente no eixo de simetria. Frequentemente, no entanto, a clivagem é assimétrica, gerando tanto extremidades 5' salientes como extremidades 3' salientes (ver Tabela 6.1). Como mostrado para os exemplos de *BamHI* e *PstI* na Tabela 6.1, a clivagem assimétrica de uma sequência palindrômica gera fragmentos de restrição com duas extremidades iminentes que são idênticas em sequência e, *ao mesmo tempo, complementares na sequência de bases*. Eles possuirão uma tendência a se associar uns com os outros (ou com qualquer outra parte saliente similarmente complementar) pela formação de pares de base. Estas extremidades salientes são, então, muitas vezes descritas como extremidades coesivas.



**Figura 6.2 Os quatro passos essenciais da clonagem baseada em células.** (A) Formação do DNA recombinante. Fragmentos de DNA clivados com uma enzima de restrição são misturados com uma população heterogênea de moléculas de vetores que foram clivadas com uma endonuclease de restrição semelhante. O DNA-alvo e o vetor são unidos pela DNA-ligase para formar DNA recombinante. Cada vetor contém uma origem de replicação que permitirá que ele seja copiado em uma célula hospedeira. (B) Transformação. O DNA recombinante é misturado com células hospedeiras as quais, normalmente, absorvem apenas uma molécula de DNA externo. Então, cada célula contém, geralmente, um único DNA recombinante. (C) Amplificação. Células individuais transformadas podem sofrer divisão celular repetida para originar uma colônia de clones celulares idênticos contendo um tipo de DNA recombinante que é mantido fisicamente separado das outras colônias contendo células com diferentes moléculas de DNA recombinante. Por conveniência, é mostrado o exemplo de um vetor cujo número de cópias por célula é altamente restrito, mas diversos vetores plasmidiais podem atingir um número bastante alto de cópias por célula, constituindo um tipo adicional de amplificação. (D) Isolamento de clones de DNA recombinante. Após a separação do DNA recombinante da célula hospedeira.

As endonucleases de restrição do tipo II foram crucialmente importantes para o desenvolvimento da clonagem do DNA utilizando células como base, pois eram capazes de clivar o DNA para produzir fragmentos definidos de um tamanho controlável. Quando o DNA é isolado de tecidos e de células em cultura, o desempacotamento das grandes moléculas de DNA e as rupturas físicas inevitáveis resultam em populações heterogêneas de fragmentos de DNA. Os fragmentos não são facilmente estudados devido ao fato de que seu tamanho médio é ainda muito grande e de eles possuírem comprimentos randomicamente diferentes. Com a utilização de endonucleases de restrição, passou a ser possível a conversão dessas populações de fragmentos de DNA interrompidos em conjuntos de fragmentos de restrição menores de tamanhos definidos.

As endonucleases de restrição do tipo II também preparam o caminho para moléculas de DNA artificialmente recombinantes. Pela clivagem de um vetor e de uma molécula de DNA-alvo com endonuclease(s) de restrição, produzindo o mesmo tipo de extremidade coesiva, a associação vetor-alvo é promovida pelo pareamento de bases entre suas extremidades coesivas (Figura 6.3). A ligação de hidrogênio entre suas terminações facilita a ligação covalente subsequente por enzimas conhecidas como DNA-ligasas. Também pode

TABELA 6.1 Exemplos de endonucleases de restrição comumente utilizadas

Enzima	Fonte	Reconhecimento de sequência e clivagem	Terminações dos fragmentos de restrição
<b>TIPO IIR (CLIVA DENTRO DE UMA SEQUÊNCIA DE RECONHECIMENTO GERALMENTE PALÍNDROMICA)</b>			
Produção de extremidades bruscas		↓ AGCT TCGA ↑	5' CT-----AG 3' 3' GA-----TC 5'
Produção de extremidades 5'		↓ GGATCC CCTAGG ↑	5' GATCC-----G 3' 3' G-----CCTAG 5'
Produção de extremidades 3'		↓ CTGCAG GACGTC ↑	5' G-----CTGCA 3' 3' ACGTC-----G 5'
<b>TIPO IIS (CLIVA FORA DA SEQUÊNCIA ASSIMÉTRICA)</b>			
		↓ CCTCNNNNNN GGAGNNNNNN ↑	5' ---CCTCNNNNNN 3' 3' N---GGAGNNNNNN 5'
<b>TIPO IIB (POSSUI SEQUÊNCIAS DE RECONHECIMENTO BIPARTIDAS)</b>			
		↓ CCANNNNNTGG GGTNNNNNACC ↑	5' NTGG-----CCANNNNN 3' 3' NNNNNACC-----GGTN 5'

Para outras enzimas, ver o banco de dados de nucleases de restrição REBASE em <http://rebase.neb.com/rebase/rebase.html>. N = qualquer nucleotídeo.

ocorrer pareamento de bases intramolecular, como quando duas extremidades salientes de um vetor se associam para formar um círculo (*ciclização*), e as reações de ligação são projetadas para promover a união do DNA-alvo ao vetor e minimizar a ciclização do vetor e outros produtos indesejados, como **concatâmeros** lineares produzidos por união contínua das extremidades (ver Figura 6.3).

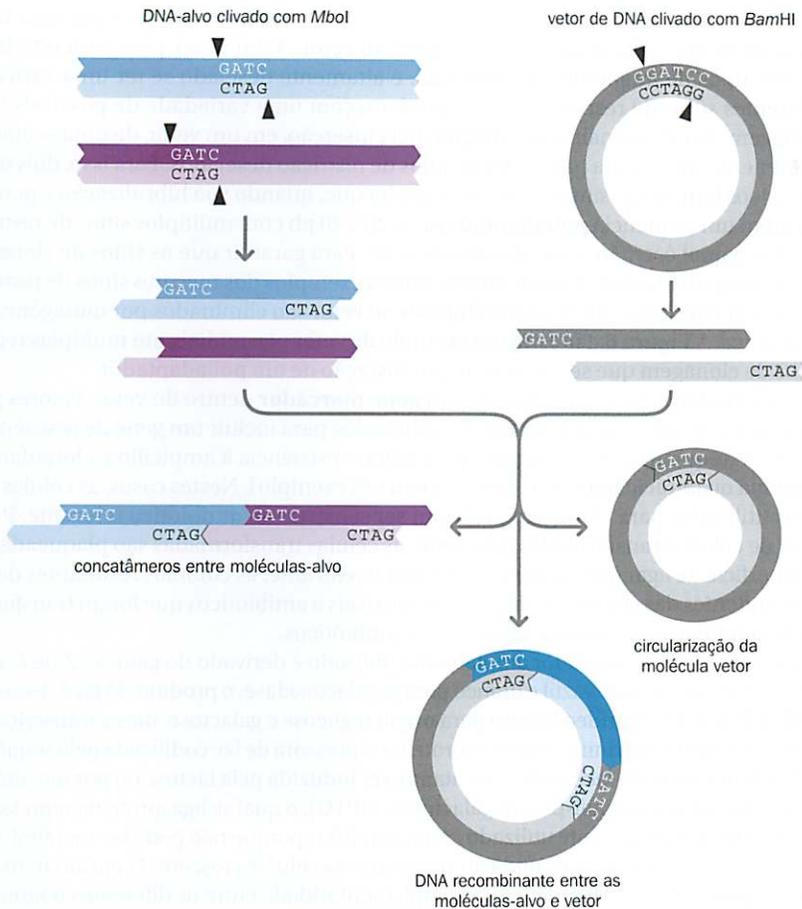
### A clonagem básica de DNA em células bacterianas utiliza vetores baseados na ocorrência natural de replicons extracromossomais

A maioria das clonagens de DNA utilizando células como base utilizam bactérias modificadas como células hospedeiras. O tempo de ciclo celular bacteriano é curto, e, então, elas se multiplicam rapidamente. Elas possuem, geralmente, um único cromossomo circular dupla-fita com uma única origem de replicação. Logo após o cromossomo do hospedeiro ter sido replicado, a célula se divide para dar origem a células-filhas que terão, cada uma, um único cromossomo, como a célula parental.

Fragmentos de DNA-alvo normalmente não possuem uma origem funcional de replicação e, por isso, não podem ser replicados independentemente dentro das células bacterianas. Eles precisam ser ligados a um **replicon** de DNA, uma sequência que possui uma origem de replicação, permitindo, dessa forma, que eles sejam propagados no interior das células. Um replicon que porta um DNA-alvo como “passageiro” dentro das células e permite que este DNA-alvo seja replicado é conhecido como um **vetor** molecular. O modo mais eficiente de propagar o DNA-alvo dentro das células é utilizar um vetor com uma origem de replicação derivada de um replicon extracromossômico natural, que replique independentemente do cromossomo da célula hospedeira. Replicons extracromossômicos frequentemente passam por vários ciclos de replicação durante um único ciclo celular, podendo então atingir um alto número de cópias (ao contrário do número de cópias do DNA cromossômico, o qual é restrito a, em geral, um por célula). Consequentemente, grandes quantidades de DNA-alvo podem ser produzidas ao serem coreplicadas com um replicon extracromossômico.

Os replicons extracromossômicos encontrados em células bacterianas geralmente pertencem a uma de duas classes: plasmídeos e bacteriófagos (fagos). **Plasmídeos** são moléculas de DNA não essenciais que replicam independentemente do cromossomo da célula hospedeira. Eles são distribuídos verticalmente para as células-filhas após a divisão de célula hospedeira, mas podem ser transferidos horizontalmente para células vizinhas durante eventos de conjugação bacteriana. Frequentemente consistem em DNA circular pequeno de dupla-fita, mas alguns plasmídeos possuem DNA linear. Eles variam em tamanho, de 2 kb a mais de 200 kb, e contêm, individualmente, poucos genes.

O número de cópias de plasmídeos varia significativamente: plasmídeos com alto número de cópias podem atingir mais de cem cópias por célula, mas outros podem estar



**Figura 6.3** Formação de DNA recombinante. Sequências de DNA-alvo heterogêneas clivadas com uma enzima de restrição do tipo II (*MboI*, neste exemplo) são misturadas com vetores de DNA clivados com outra enzima de restrição, tal como a *BamHI*, que gera o mesmo tipo de extremidade coesiva (aqui, uma sequência GATC 5' saliente). As condições de ligação escolhidas vão permitir que as moléculas de DNA-alvo e vetor se combinem para originar DNA recombinante. No entanto, outros produtos de ligação são possíveis. Concatâmeros intermoleculares de DNA-alvo são mostrados à esquerda, e ciclicização intramolecular de vetor à direita. Outros produtos de ligação possíveis não mostrados aqui incluem concatâmeros intermoleculares de vetores, ciclicização de múltiplos e eventos de coligação que envolvem duas sequências-alvo diferentes sendo incluídas com um vetor na mesma molécula de DNA recombinante.

restritos a apenas uma ou duas cópias por célula. Diferentes plasmídeos podem coexistir em uma célula. Um isolado típico de *Escherichia coli*, por exemplo, pode possuir três plasmídeos pequenos diferentes presentes em cópias múltiplas e um plasmídeo grande de cópia única. Exemplos naturais de plasmídeos bacterianos incluem plasmídeos que portam fatores sexuais (F) e aqueles que portam genes de resistência a drogas. Alguns plasmídeos às vezes inserem o seu DNA em cromossomos bacterianos (*integração*). Estes plasmídeos, os quais podem existir em duas formas – replicons extracromossomais ou plasmídeos integrados –, são conhecidos como **epissomos**.

Um **bacteriófago** (ou **fago**, por abreviação) é um vírus que infecta células bacterianas. Ao contrário dos plasmídeos, os fagos podem existir extracelularmente, mas, como outros vírus, eles devem invadir uma célula hospedeira para se reproduzirem e podem infectar e se reproduzir em apenas alguns tipos de células hospedeiras. Os fagos podem ter genoma de DNA ou de RNA, e, em fagos de DNA, o genoma pode ser constituído de DNA dupla-fita ou de fita simples, além de poder ser linear ou circular. Existe uma variação considerável no tamanho do genoma, desde poucas quilobases a algumas centenas delas.

No fago maduro, o genoma é revestido por uma capa proteica que pode auxiliar na ligação a receptores de proteínas específicas na superfície de uma nova célula hospedeira e a entrar na nova célula. Para escapar de uma célula e infectar outras, os bacteriófagos produzem enzimas que causam a explosão celular (*lise celular*). A habilidade de infectar novas células depende do número de partículas de fago produzidas e do *tamanho da lise*; então, um alto número de cópias é geralmente alcançado. Como plasmídeos epissomais, alguns bacteriófagos, tais como o fago  $\lambda$ , também podem se integrar ao cromossomo da célula hospedeira, onde a base da sequência do fago integrado é conhecida como *prófago*.

### A clonagem em células bacterianas utiliza plasmídeos geneticamente modificados ou vetores de bacteriófagos e células hospedeiras modificadas

Para plasmídeos naturais e fagos serem utilizados como moléculas vetores, eles precisam ser geneticamente modificados de maneiras diferentes. Primeiro, é importante projetar

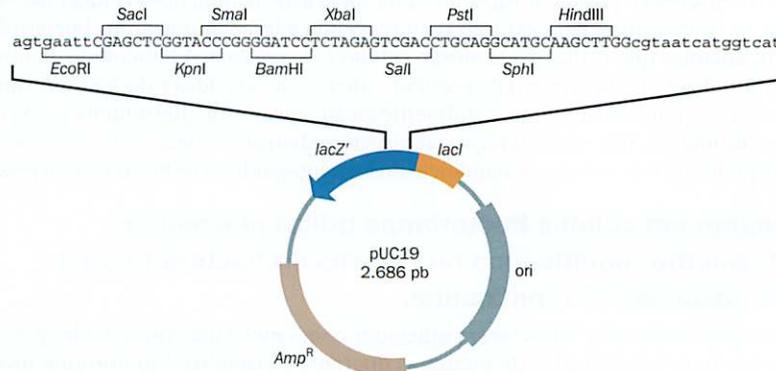
**Figura 6.4** Um exemplo de um vetor plasmidial com alto número de cópias, pUC19. A origem de replicação (*ori*) possibilita mais de 100 cópias de pUC19 na célula hospedeira adequada de *E. coli*. O gene de resistência à ampicilina (*Amp<sup>R</sup>*) permite a seleção das células que contêm a molécula do vetor. *lacZ'* é um fragmento 5' do gene *lacZ* que pode ser expresso para originar um fragmento N-terminal curto de  $\beta$ -galactosidase, o fragmento  $\alpha$ . O *lacI* codifica uma proteína repressora de *lac* que reprime a transcrição do *lacZ'* até que o indutor IPTG seja adicionado ao meio de cultura para se ligar ao repressor da proteína e inativá-lo. A célula hospedeira possui um gene *lacZ* mutante que produz um fragmento C-terminal de  $\beta$ -galactosidase, o qual pode, no entanto, ser complementado pelo fragmento  $\alpha$  produzido pelo vetor para produzir uma  $\beta$ -galactosidase funcional. O método utiliza um substrato incolor [5-bromo-4-cloro-3-indolil- $\beta$ -D-galactopiranosil (X-Gal)] que é convertido em um produto de reação não tóxico com uma cor azul intensa. Um poliadaptador inserido perto do início da sequência codificadora do *lacZ'* do vetor fornece uma região de múltiplos sítios de clonagem (MCS), uma série de sítios de restrição diferentes e únicos que permite possibilidades alternativas para clonagem de fragmentos de restrição. O poliadaptador é pequeno e inserido de maneira que seu sistema de leitura do *lacZ'* seja preservado e funcional. Entretanto, a clonagem de um inserto comparativamente grande dentro do MCS normalmente causa inativação por inserção e ausência da atividade da  $\beta$ -galactosidase e, então, as células com DNA recombinante serão incolores.

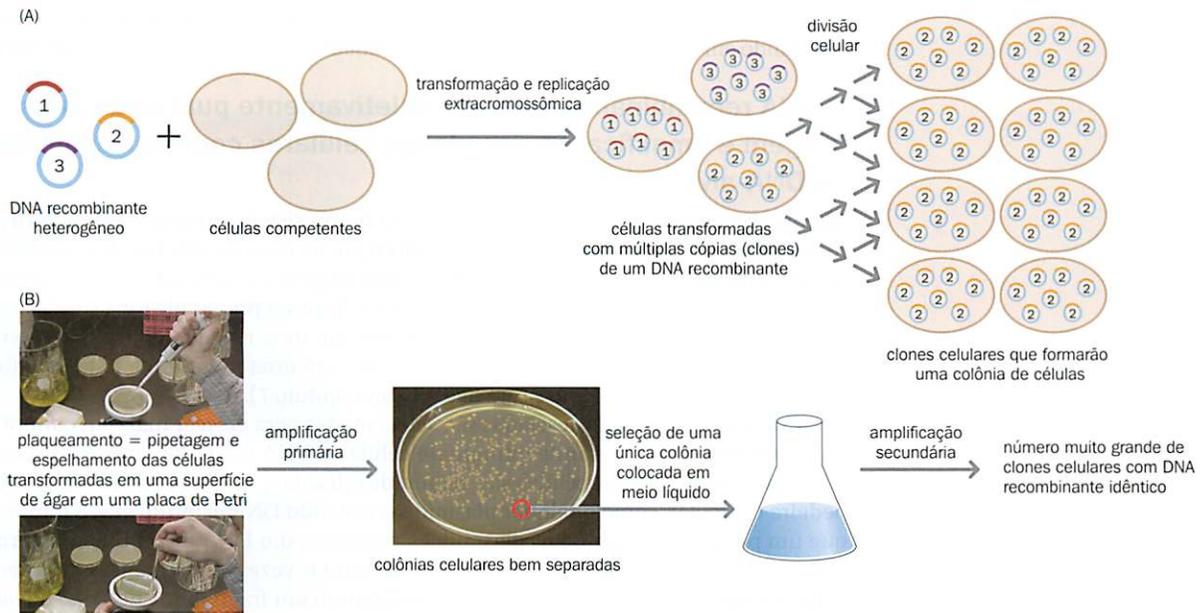
o vetor de maneira que os fragmentos de restrição alvo sejam inseridos em uma única localização (o *sítio de clonagem*) na molécula do vetor. Além disso, para se acumularem diferentes tipos de fragmentos de restrição, é altamente desejado se ter uma variedade de diferentes sítios de restrição únicos que forneçam uma variedade de possíveis sítios de clonagem. Isso é normalmente atingido pela inserção, em um vetor, de uma sequência artificialmente sintetizada para se ter os sítios de restrição desejados. Para isso, dois oligonucleotídeos longos são sintetizados de maneira que, quando sua hibridização é permitida, formem um sequência **poliadaptadora** de 20 a 40 pb com múltiplos sítios de restrição desejados, a qual é, então, inserida em um vetor. Para garantir que os sítios de clonagem desejados no poliadaptador sejam únicos, outros exemplos dos mesmos sítios de restrição que ocorrem naturalmente em outros lugares no vetor são eliminados por mutagênese sítio-específica. A **Figura 6.4** mostra um exemplo de vetor plasmidial com múltiplas regiões de sítios de clonagem que se originaram por inserção de um poliadaptador.

Outra modificação é a inserção de um **gene marcador** dentro do vetor. Vetores plasmidiais, por exemplo, são amplamente modificados para incluir um gene de resistência a antibiótico, frequentemente um gene que confere resistência à ampicilina, cloranfenicol, canamicina ou tetraciclina (ver Figura 6.4 para um exemplo). Nestes casos, as células bacterianas utilizadas para clonagem precisam ser sensíveis ao antibiótico relevante. Para a triagem de células transformadas pelo vetor, as células transformadas são plaqueadas em uma superfície de ágar que contém o antibiótico relevante; as colônias resultantes devem ser descendentes das células originalmente sensíveis a antibióticos que foram transformadas pelo vetor para se tornarem resistentes a antibióticos.

Um outro sistema marcador amplamente utilizado é derivado do gene *lacZ* de *E. coli* e envolve um ensaio de cores azul e branco para  $\beta$ -galactosidase, o produto do *lacZ*. Esta enzima pode clivar o dissacarídeo lactose para originar glicose e galactose, mas a transcrição do *lacZ* é normalmente reprimida por uma proteína repressora de *lac* codificada pela sequência do *lacI*. A transcrição do *lacZ* pode, no entanto, ser induzida pela lactose ou por um análogo à lactose à tal como isopropil- $\beta$ -D-tiogalactoside (IPTG), o qual se liga, proteína repressora e a inativa. O IPTG é geralmente utilizado como o indutor porque não pode ser metabolizado, e então sua concentração não muda ao passo que as células crescem. O ensaio marcador também é projetado para depender da complementaridade entre os diferentes fragmentos de  $\beta$ -galactosidase inativos produzidos pela célula hospedeira e pelo vetor. Então, o hospedeiro *E. coli* é um mutante no qual a extremidade 5' do gene *lacZ* foi deletada, resultando na produção de um grande, mas inativo, fragmento C-terminal de  $\beta$ -galactosidase. Os vetores tais como pUC19 (ver Figura 6.4) foram modificados para conter *lacZ'*, um componente 5' de gene *lacZ* que, quando transcrito, produz um pequeno fragmento N-terminal inativo de  $\beta$ -galactosidase. Este fragmento, o fragmento  $\alpha$ , complementa o fragmento C-terminal da célula hospedeira para produzir uma enzima funcional (*complementaridade  $\alpha$* ). A enzima ativa é convenientemente analisada por uma reação química, como descrito na Figura 6.4.

Um requinte adicional do ensaio da complementaridade do *lacZ* permite a triagem por células transformadas que tenham absorvido DNA recombinante (vetor com inserto de DNA). Neste caso, o poliadaptador é inserido dentro da sequência codificadora que especifica o fragmento N-terminal da  $\beta$ -galactosidase, como no exemplo do plasmídeo pUC19 (ver Figura 6.4). Neste caso, o número de nucleotídeos no poliadaptador é um múltiplo exato de três, de maneira que o fragmento N-terminal da  $\beta$ -galactosidase do vetor pode continuar complementando o fragmento da  $\beta$ -galactosidase do hospedeiro para produzir colônias azuis. Entretanto, os recombinantes com um inserto no poliadaptador podem causar inativação por inserção, de maneira que  $\beta$ -galactosidase funcional não é produzida.





### A transformação é o passo-chave de fracionamento na clonagem de DNA que utiliza células como base

A membrana plasmática das células é seletivamente permeável, de maneira que apenas algumas pequenas moléculas podem passar para dentro e fora das células; moléculas tais como grandes fragmentos de DNA são normalmente incapazes de atravessar a membrana. Entretanto, as células podem ser tratadas de diferentes maneiras, de modo que a permeabilidade seletiva da membrana plasmática seja temporariamente distorcida. Uma maneira, por exemplo, é utilizar **eletroporação**, na qual as células são expostas a um pulso de eletricidade de alta voltagem curto. Como resultado da alteração temporária da membrana da célula, algumas das células tratadas se tornam *competentes*, ou seja, elas são capazes de absorver DNA proveniente do ambiente extracelular.

Somente uma pequena porcentagem de células competentes irá absorver o DNA recombinante para se tornar transformada. No entanto, aquelas que passam por isso frequentemente absorvem apenas uma única molécula de DNA recombinante. Esta é a base do passo crítico de fracionamento na clonagem que utiliza células como base. A população de células transformadas pode ser vista como um posto de classificação, no qual uma mistura complexa de fragmentos de DNA é ordenada a partir do depósito de *moléculas individuais de DNA em células recipientes individuais*. A replicação extracromossômica do DNA recombinante permite a presença de várias cópias idênticas de um único tipo de DNA recombinante em uma célula (**Figura 6.5**).

Devido ao fato de que o DNA circular é transformado de forma muito mais eficiente que o DNA linear, a maioria das células transformadas conterá produtos cíclicos em vez de concatâmeros lineares de DNA recombinante. Geralmente, moléculas vetoriais são tratadas para reduzir a chance de ciclização, e, então, a maioria dos transformantes conterá DNA recombinante. Note, no entanto, que a **cotransformação**, a ocorrência de mais de um tipo de molécula de DNA introduzida dentro de um clone celular, pode acontecer algumas vezes.

As células transformadas podem, então, se multiplicar. Em clonagens utilizando vetores plasmidiais e uma célula hospedeira bacteriana, uma solução contendo as células transformadas é espalhada de forma simples sobre uma superfície de ágar em uma placa de Petri, e o crescimento celular é permitido (um processo chamado de **plaqueamento** – Figura 6.5). Isso geralmente resulta na formação de colônias bacterianas que consistem em **clones celulares** (todas as células em uma mesma colônia são idênticas porque todas descendem de uma única célula fundadora).

O plaqueamento é geralmente projetado para produzir colônias bem separadas na superfície de ágar, permitindo a separação física das colônias contendo diferentes tipos de DNA recombinante. A seleção de uma colônia individual, para ser posta em um frasco de cultura para subsequente crescimento em cultura líquida, permite uma expansão secundária do número de células que pode resultar em grandes produções de clones, todos idênticos a uma única célula ancestral (ver Figura 6.5). Se a célula original continha um único tipo de frag-

**Figura 6.5** Transformação, o passo crítico de fracionamento do DNA, e amplificação em larga escala de DNA recombinante. (A) A clonagem utilizando células como base começa, geralmente, com uma população complexa de fragmentos de DNA-alvo que são unidos a um único tipo de vetor, produzindo uma coleção heterogênea de moléculas de DNA recombinante. Durante a transformação, as moléculas de DNA recombinante são transferidas para células competentes de maneira que uma célula absorva, normalmente, apenas um tipo de molécula de DNA recombinante. Como resultado, a população de células age como um sistema fracionado, alocando DNAs recombinantes individuais a células individuais. Dentro de cada célula, um DNA recombinante pode ser replicado como DNA extracromossômico, produzindo, frequentemente, diversos clones idênticos de DNA. Após isso, cada célula transformada sofre repetidas divisões para originar populações de clones celulares idênticos que formarão colônias celulares. (B) As colônias celulares podem ser fisicamente separadas por **plaqueamento**, um procedimento que envolve o crescimento das células transformadas em uma superfície de ágar, de forma a permitir a separação das colônias celulares. Colônias individuais são escolhidas e colocadas em meio líquido em frascos para permitir a amplificação secundária. (Cortesia de James Stock, King's College. Utilizada sob licença de Creative Commons Attribution ShareAlike 3.0.)

mento de DNA externo ligado a um replicon, então as células descendentes terão o mesmo, resultando em uma vasta amplificação na quantidade do fragmento externo específico.

### O DNA recombinante pode ser seletivamente purificado após triagem e amplificação dos clones celulares com os fragmentos de DNA-alvo desejados

Alguns sistemas de genes marcadores fornecem um sistema geral para identificação de recombinantes, tais como inativação por inserção, no caso do sistema de complementaridade  $\alpha$  do *lacZ* descrito acima. Para fazer uma triagem em um indivíduo recombinante que possui um DNA-alvo de interesse específico, deve ser preparada uma sonda marcada que pode detectar especificamente o DNA-alvo em uma reação de hibridização contra colônias bacterianas que tenham sido separadas em uma membrana (isto é conhecido como *hibridização de colônia* e será descrito no Capítulo 7).

Colônias bacterianas desejadas são selecionadas para crescer mais em cultura líquida, resultando em um grande número de células e DNA recombinante. Para recuperar o DNA recombinante, as células são estouradas (lisadas). O cromossomo da célula hospedeira é circular, como qualquer plasmídeo contendo DNA externo, mas é muito maior que um plasmídeo. Embora os plasmídeos menores com DNA recombinante permaneçam circulares, o grande DNA cromossômico muitas vezes sofre quebras durante a lise celular e subsequente extração de DNA, resultando em fragmentos de DNA linear com extremidades livres. Os métodos de purificação, geralmente, exploram a diferença física entre DNA cromossômico (contendo extremidades livres de DNA) e DNA recombinante (sem extremidades livres) nas células lisadas. Para isso, o DNA é, primeiro, sujeito a um passo de desnaturação, por exemplo, por meio de tratamento com um álcali, seguido por um passo de renaturação. O DNA cromossômico linearizado se desnatura facilmente e precipita fora da solução. Entretanto, o plasmídeo com DNA recombinante está na forma de DNA circular covalentemente fechado e, após o passo de renaturação, passa espontaneamente para a forma de DNA superespiralado e permanece em solução.

Conforme necessário, uma purificação adicional é geralmente realizada com uma cromatografia por troca de ânions. A resina de troca iônica Qiagen, amplamente utilizada, consiste em esferas de sílica de tamanho definido que possuem uma superfície hidrofílica de revestimento, permitindo o acoplamento entre grupos dietilaminoetil (DEAE) positivamente carregados. Devido ao grande número de grupos fosfato negativamente carregados do eixo principal do DNA que possuem afinidade com os grupos DEAE na resina de sílica, o DNA se liga fortemente à resina, enquanto impurezas se ligam fracamente ou nem se ligam. O DNA fortemente ligado pode, subsequentemente, ser eluído com tampões altamente salinos para romper as ligações fosfato-DEAE. O DNA-alvo pode ser removido de um DNA recombinante plasmidial e subclonado em outros vetores como requisições que podem oferecer a possibilidade de sequenciar o DNA ou expressar qualquer gene dentro do alvo, por exemplo. Algumas destas aplicações serão consideradas nas próximas seções.

## 6.2 CLONAGEM DE INSERTOS GRANDES E SISTEMAS DE CLONAGEM PARA PRODUÇÃO DE DNA FITA SIMPLES

As primeiras tentativas bem-sucedidas de clonagem de fragmentos de DNA de humanos em células bacterianas tirou vantagem de transcritos gênicos que eram naturalmente enriquecidos em alguns tecidos. Por exemplo, grande parte do mRNA produzido nos eritrócitos consiste em mRNA de  $\alpha$  e  $\beta$ -globina. Quando a transcriptase reversa é utilizada para copiar mRNA de eritrócitos, o DNA complementar (cDNA) de eritrócitos resultante é extremamente enriquecido em cDNA de globina, facilitando seu isolamento.

Para possibilitar um método mais geral de clonagem de genes e de todas as seqüências de DNA, tecnologias foram, subsequentemente, desenvolvidas com o objetivo de clonar todas as seqüências de DNA constituintes de uma população inicial. As amplas coleções de clones de DNA resultantes são conhecidas como **bibliotecas de DNA**. Bibliotecas de DNA genômico são preparadas a partir do DNA genômico total de qualquer célula nucleada de um organismo; bibliotecas de cDNA são preparadas a partir de cópias de cDNA do RNA ou mRNA total de tecidos ou células específicos de um organismo. As bibliotecas de DNA têm sido ferramentas vitalmente importantes no projeto Genoma e na resolução da estrutura gênica e da relação entre os genes e seus transcritos. Este assunto será visto em detalhes no Capítulo 8.

Genes de vertebrados são, frequentemente, muito grandes, com um elevado conteúdo de DNA repetitivo, fazendo deles menos sujeitos a clonagem. DNA repetitivo é ex-

**TABELA 6.2** Propriedades de sistemas de clonagem que utilizam células como base, organizados de acordo com a capacidade de inserção

Sistema de clonagem baseado em		Tamanho do inserto	Nº de cópias (por célula)	Aplicações
<b>VETORES PLASMIDIAIS</b>				
Vetores plasmidiais permitindo replicação de alto número de cópias	Plasmídeo padrão e vetores fagemídeos	Geralmente até 5-10 kb	frequentemente alto	Muito versáteis. Rotineiramente utilizados na subclonagem de fragmentos amplificados por PCR. Também amplamente utilizados na produção de bibliotecas de DNA e para expressar genes a fim de dar origem a produtos de RNA ou proteínas. Fagemídeos podem produzir DNA recombinante de fita simples.
	Vetores	30-44 kb	alto	Utilizados para a produção de bibliotecas de DNA com insertos de tamanho modesto. Entretanto, os altos números de cópias tornam os insertos comparativamente instáveis.
Vetores plasmidiais cujo número é restrito	Vetores	30-44 kb	1-2	Como uma alternativa para clonagem de cosmídeos, fornecendo uma estabilidade muito maior ao inserto à custa de produções mais baixas de DNA recombinante.
	Vetores PAC (cromossomo artificial P1)	130-150 kb	1-2	Utilizado para produção de bibliotecas de DNA genômico, frequentemente como um auxiliar suplementar à clonagem com BAC.
	Vetores BAC (cromossomo bacteriano artificial)	até 300 kb	1-2	Extensivamente utilizado para a produção de bibliotecas genômicas para uso nos projetos do genoma.
	Vetores YAC (cromossomo artificial de levedura)	0,2-2 Mb	baixo	Utilizado como um sistema de clonagem intermediário para alguns projetos do genoma (p. ex., Projeto Genoma Humano), mas os insertos são, frequentemente, instáveis. Também são muito úteis no estudo das funções de grandes genes e de <i>clusters</i> gênicos.
<b>VETORES DE FAGOS</b>				
M13		pequeno	baixo a médio	Muito útil no passado à produção de DNA recombinante de fita simples para moldes de sequenciamento e mutagênese sítio-direcionada. Foi substituída por fagemídeos.
Um vetor $\lambda$ de inserção (p. ex., $\lambda$ gt11)		até 10 kb	baixo	Muito útil para a produção de bibliotecas de cDNA devido ao fato de que podem receber insertos bastante grandes e porque grandes números de recombinantes podem ser produzidos.
Vetores $\lambda$ de inserção		9-23 kb <sup>a</sup>	baixo	Raramente utilizado nos dias de hoje.
P1		70-100 kb	baixo	Utilizado para a produção de bibliotecas de DNA genômico no passado. Raramente utilizado nos dias de hoje.

<sup>a</sup> O tamanho do inserto aumenta comparado com vetores de inserção devido à remoção de uma parte central não essencial do genoma do  $\lambda$  e à substituição desta pelo DNA a ser clonado. Note que há uma preferência pelos vetores plasmidiais; vetores de fagos são, comparativamente, mais difíceis de se trabalharem.

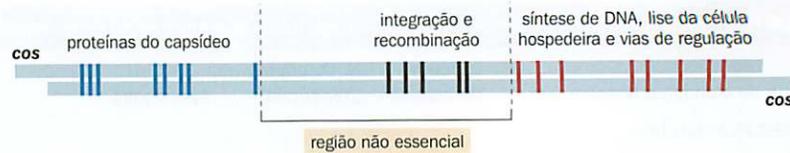
tremamente raro em genomas bacterianos, e vastas quantidades de DNA recombinante introduzidas que sejam ricas em sequências de DNA repetitivo não são bem toleradas. Como resultado, plasmídeos-padrão raramente podem ser utilizados para clonar grandes (> 10 kb) fragmentos de DNA animal; em vez disso, os insertos são, geralmente, menores que 5 kb. Ainda assim, os plasmídeos continuam sendo os instrumentos de clonagem para diversos propósitos de pesquisa diferentes. Vetores tais como pUC19 (ver Figura 6.4) permitem replicação de DNA recombinante de alto número de cópias e em grandes quantidades, e, como será visto na Seção 6.3, vetores plasmidiais são amplamente utilizados para expressar genes de interesse. Para clonar fragmentos de DNA maiores, foram necessários sistemas de clonagem de DNA diferentes (ver Tabela 6.2 para uma visão global). Além disso, algumas aplicações para o estudo de DNA recombinante, particularmente sequenciamento de DNA e mutagênese sítio-específica *in vitro*, requerem moldes de DNA de fita simples. Consequentemente foram desenvolvidos vetores de clonagem que pudessem também ser utilizados para gerar DNA recombinante de fita simples.

### Os primeiros vetores de clonagem de insertos exploravam as propriedades do bacteriófago $\lambda$

O bacteriófago  $\lambda$  (lambda) possui um genoma de DNA dupla-fita de 48,5 kb. Dentro de uma célula hospedeira de *E. coli*, o DNA do  $\lambda$  pode ser replicado independentemente do cromossomo ou pode se integrar ao cromossomo da célula hospedeira e ser replicado como um componente deste. A porção central do genoma  $\lambda$  é uma região não essencial: os genes lá contidos são necessários apenas para integrar o  $\lambda$  ao cromossomo hospedeiro. A clonagem de DNA é então possível por meio da remoção de uma grande seção central do genoma  $\lambda$  e pela substituição desta por um DNA-alvo de até 23 kb (Figura 6.6).

**Figura 6.6** Mapa do genoma do fago  $\lambda$ .

Os genes estão ilustrados sob a forma de barras verticais. A região central de 48,5 kb do genoma do fago  $\lambda$  não é essencial para a replicação extracromossômica, para a produção de proteínas do capsídeo ou para a lise da célula hospedeira – ela contém genes necessários apenas no caso de integração de  $\lambda$  a um cromossomo de seu hospedeiro. Quando se utilizam vetores  $\lambda$  de substituição, a clonagem ocorre pela substituição desta região central não essencial por segmentos demarcados por sítios de restrição que podem ter tamanho de 9 a 23 kb. As sequências *cos* terminais possuem extremidades 5' fita simples com um tamanho igual a 12 nucleotídeos e sequências complementares que permitem a circularização. O genoma  $\lambda$  deve ser linear para caber em seu capsídeo, mas após a adesão do fago à membrana de uma nova célula hospedeira e inserção do DNA linear, o DNA  $\lambda$  inserido imediatamente é circularizado por pareamento de bases entre as sequências *cos* das extremidades 5' terminais, permitindo a replicação pelo sistema de círculo rolante, o qual gera múltímeros lineares da unidade de 48,5 kb de comprimento. As proteínas do capsídeo são sintetizadas e os múltímeros  $\lambda$  são clivados assimetricamente na sequência *cos*, dando origem a múltiplas cópias da unidade genômica, que são individualmente inseridas no revestimento gerado pelas proteínas do capsídeo.



Um refinamento adicional que aumentou o limite superior do tamanho dos fragmentos que poderiam ser clonados explorou as propriedades das sequências terminais *cos* do  $\lambda$ . As sequências *cos* terminam em extremidades 5' salientes com um comprimento de 12 nucleotídeos complementares em sequência, a fim de formar extremidades coesivas; por isso o nome *cos* (ver Figura 6.6). O genoma  $\lambda$  precisa ser linear para se encaixar dentro de seu revestimento proteico, mas após o fago se ligar à membrana de uma nova célula hospedeira e inserir seu DNA linear, o DNA  $\lambda$  inserido é imediatamente ciclizado por pareamento entre as sequências *cos*. A replicação extracromossômica do DNA  $\lambda$  ocorre por um modelo de círculo rolante que gera múltímeros lineares de comprimento da unidade. Os múltímeros  $\lambda$  são clivados na sequência *cos* a fim de gerar genomas lineares de comprimento da unidade para empacotamento nos capsídeos virais proteicos.

Para o adequado empacotamento do DNA  $\lambda$  dentro do capsídeo proteico do fago, o espaçamento entre as sequências *cos* é restrito a 35 a 53 kb. Por meio da inserção da sequência *cos* em um plasmídeo, um novo tipo de vetor de clonagem – o cosmídeo – foi produzido; este, quando utilizado em um sistema de empacotamento  $\lambda$  *in vitro*, poderia receber fragmentos externos de até 40 kb de comprimento. Entretanto, recombinantes de cosmídeos que contêm DNA de vertebrados de 40 kb são instáveis. Os insertos possuem, geralmente, sequências repetitivas, e, quando o DNA recombinante está presente em uma célula bacteriana em grande número de cópias, a recombinação torna os insertos propensos a rearranjos e deleção.

Para que a clonagem de grandes insertos fosse possível, os sistemas de clonagem precisavam utilizar replicons de baixo número de cópias. Uma abordagem bem-sucedida utiliza hospedeiros bacterianos, mas com replicons extracromossômicos que possuem um número restrito de cópias. Outra solução envolve a utilização de células de levedura e visa à construção de cromossomos artificiais para replicar DNA-alvo. Nesse caso, os vetores incluem sequências essenciais para a função cromossômica, incluindo uma origem de replicação que se origina de um cromossomo, resultando em baixo número de cópias. Para originar grandes fragmentos de DNA-alvo para clonagem, o DNA precisa ser isolado das células sob condições suaves e, então, sujeito a *digestão parcial* com uma endonuclease de restrição, de maneira que apenas uma pequena minoria de sítios de restrição disponíveis seja realmente clivada.

### Grandes fragmentos de DNA podem ser clonados em células bacterianas por meio da utilização de replicons extracromossômicos de baixo número de cópias

Diversos vetores utilizados para clonagem de DNA em células bacterianas são baseados em replicons extracromossômicos que permitem a replicação de alto ou médio número de cópias em células hospedeiras. O DNA de eucariotos complexos, especialmente mamíferos, contém sequências repetitivas que podem causar instabilidade quando propagadas em alto número de cópias em células bacterianas. Clones contendo grandes insertos de DNA eucariótico com múltiplas sequências repetitivas são particularmente instáveis, e propagá-los em grandes números de cópias resulta em deleção do inserto ou rearranjos do DNA clonado. Para superar tamanha limitação de tamanho, vetores plasmidiais e de fagos foram desenvolvidos, os quais utilizavam replicons de baixo número de cópias.

### Cromossomo bacteriano artificial (BAC) e vetores fosmídeos

O fator F, um plasmídeo de fertilidade de *E. coli*, contém dois genes, *parA* e *parB*, que mantêm o seu número de cópias em um ou dois por célula de *E. coli*. Os vetores plasmidiais baseados no sistema do fator-F são capazes de receber grandes fragmentos de DNA externo (até 300 kb), e os recombinantes resultantes podem ser transferidos com eficiência considerável para células bacterianas por meio da **eletroporação**. Os plasmídeos recombinantes resultantes são bastante grandes e, embora não tão grandes quanto um cromossomo bacteriano (o qual é também circular, dupla-fita e, na sua maior parte, livre de proteínas), eles se tornaram conhecidos como **cromossomos bacterianos artificiais (BACs)**. Os BACs contêm replicons de baixo número de cópias, e, então, apenas pequenas produções de DNA recombinante podem ser recuperadas das células hospedeiras. No entanto, devi-

do à grande estabilidade do inserto, eles foram os clones mais amplamente utilizados para o sequenciamento no Projeto Genoma Humano.

Vetores como os cosmídeos que foram adaptados para replicação com baixo número de cópias pela utilização de genes do fator-F são conhecidos como vetores fosmídeos. Os insertos possuem a mesma amplitude de tamanho que em uma clonagem com cosmídeo, mas são muito mais estáveis.

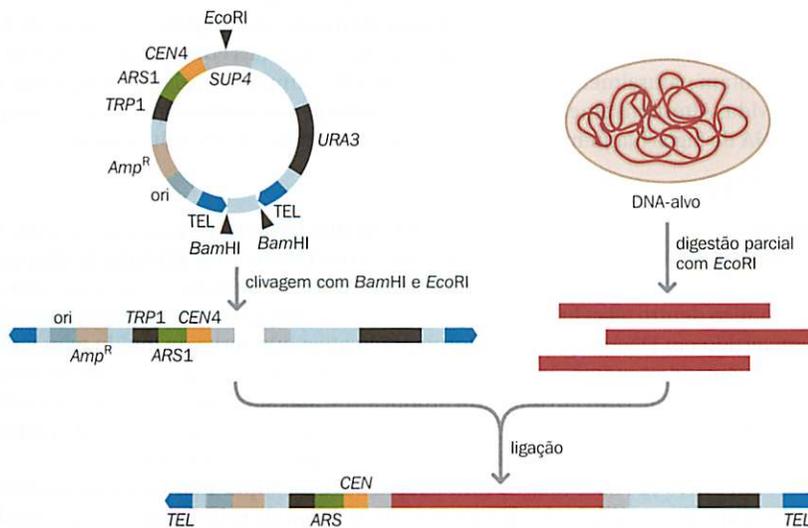
### Vetores de bacteriófago P1 e cromossomos artificiais P1

Uma abordagem alternativa para a clonagem de grandes insertos em *E. coli* foi desenvolver vetores baseados em bacteriófagos que possuem genomas naturalmente grandes. Por exemplo, o bacteriófago P1 possui um DNA genômico linear dupla-fita de 110 a 115 kb que, como no fago  $\lambda$ , é empacotado em um capsídeo proteico. Vetores de clonagem do bacteriófago P1 têm sido, portanto, projetados em componentes nos quais o P1 é incluído em um plasmídeo circular. Subsequentemente, características do P1 e do sistema do fator-F foram combinados em outros sistemas de clonagem baseados em plasmídeos, nos quais os recombinantes são conhecidos como **cromossomos artificiais P1 (PACs)**. Como nos BACs, os insertos de PACs são muito estáveis e também têm sido utilizados como moldes diretos para sequenciamento em alguns projetos do genoma.

### Cromossomos artificiais de levedura (YACs) possibilitam a clonagem de megabases de fragmentos de DNA

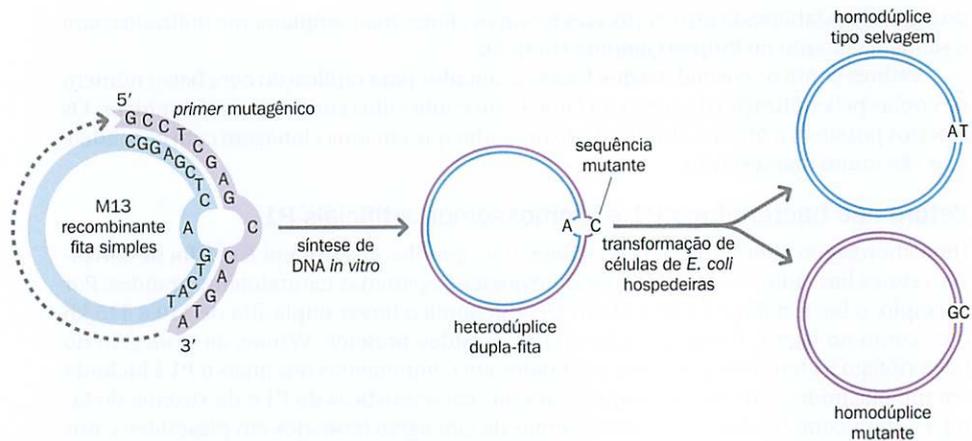
O método mais popular de clonagem de fragmentos de DNA extremamente grandes é baseado na produção de cromossomos artificiais que podem ser propagados em células de brotamento de levedura. Os cromossomos de *Saccharomyces cerevisiae* são lineares e variam em tamanho, de 200 kb a 1,5 Mb. Além de portarem um grande número de genes, eles também possuem elementos de sequência curtos que são essenciais para suas funções cromossômicas básicas: manter a molécula de DNA, replicá-la e garantir uma segregação fiel nas células descendentes após a divisão celular. Apenas três tipos de sequências curtas, cada uma com no máximo algumas poucas centenas de pares de bases, são necessários para uma molécula de DNA se comportar como um cromossomo em *S. cerevisiae* (ver Figura 2.10). São necessários **centrômeros** para a disjunção das cromátides-irmãs na mitose e também para a disjunção dos cromossomos homólogos na primeira divisão meiótica. São necessários **telômeros** para a replicação completa de moléculas lineares e para a proteção das extremidades dos cromossomos contra ataques de nucleases. Elementos de **sequência de replicação autônoma (ARS)** são necessários para a replicação autônoma do DNA cromossômico e acredita-se que ajam como origens de replicação específicas.

Para fazer um **cromossomo artificial de levedura (YAC)**, tudo o que é necessário é combinar um fragmento de DNA externo de tamanho adequado com quatro sequências curtas que podem funcionar em células de levedura (dois telômeros, um centrômero e um elemento ARS). A molécula de DNA linear resultante deve ter as sequências teloméricas corretamente posicionadas nas regiões terminais (**Figura 6.7**). Os YACs não podem ser



**Figura 6.7** A produção de cromossomos artificiais de levedura (YACs).

Um gene que confere resistência para ampicilina (*Amp<sup>R</sup>*) e uma origem de replicação derivada de um plasmídeo (*ori*) permitem que o vetor YAC seja replicado originando um grande número de cópias em uma célula hospedeira de *E. coli*. Além disso, o vetor contém os três tipos de elementos necessários para que o DNA se comporte como um cromossomo de levedura: um centrômero (*CEN4*), uma sequência de replicação autônoma (*ARS1*), e dois telômeros (*TEL*). Os alelos dominantes *TRP1* e *URA3* são incluídos como genes marcadores: eles complementam os alelos recessivos (*trp1* e *ura3*) na célula de levedura hospedeira, fornecendo um sistema de seleção para a identificação de células transformadas que contenham o vetor YAC. O sítio de clonagem para o DNA-alvo está localizado no interior do gene *SUP4*. Este gene compensa uma mutação na célula hospedeira que causa o acúmulo de pigmento vermelho. As células transformadas são anormalmente vermelhas, e células transformadas apenas com o vetor formarão colônias incolores. A clonagem de um fragmento de DNA-alvo no interior do gene *SUP4* causa uma inativação insercional, restaurando o fenótipo mutante (vermelho). O DNA-alvo é parcialmente digerido com a endonuclease de restrição *EcoRI* de forma a originar fragmentos muito grandes (centenas de quilobases) e é misturado com moléculas linearizadas do vetor clivadas com as endonucleases de restrição *BamHI* e *EcoRI*. Ocorre ligação entre as extremidades *EcoRI*, com sequências vetor flanqueando os fragmentos de DNA-alvo. O DNA recombinante pode ser usado para a transformação de células hospedeiras especializadas e os sistemas de seleção acima descritos podem ser usados para a identificação de células transformadas e de células que contêm o DNA recombinante.



**Figura 6.8 Mutagenese por oligonucleotídeos imperfeitos.** Vários métodos diferentes de mutagenese por oligonucleotídeos imperfeitos podem ser utilizados para criar uma mutação de ponto desejada em um único sítio predeterminado dentro da molécula de DNA clonada. O exemplo ilustra o uso de um oligonucleotídeo mutagênico para direcionar a substituição de um único nucleotídeo em um gene. O gene é clonado em um vetor apropriado para gerar DNA recombinante fita simples, tal como um vetor M13 (como mostrado aqui) ou um vetor fagemídeo. Um *primer* de oligonucleotídeo é desenhado para ser complementar em sequência à porção da sequência do gene contendo o nucleotídeo a ser mutado (neste caso, uma adenina A) e para conter a base não complementar desejada nesta posição (C, e não T). Apesar do mau pareamento interno, o anelamento do *primer* mutagênico é possível, e a síntese de uma segunda fita pode ser estendida pela DNA-polimerase, e os intervalos selados, pela DNA-ligase. O heterodúplce resultante pode ser transformado em *E. coli*, no qual duas populações de recombinantes podem ser recuperadas: homodúplces do tipo selvagem e mutantes. Os mutantes podem ser identificados por métodos de amplificação alelo-específica baseados em PCR (ver Figura 6.17).

diretamente transferidos para células de levedura; em vez disso, as paredes externas da célula devem ser removidas. Os *esferoplastos* resultantes podem receber fragmentos endógenos, mas são osmoticamente instáveis e precisam ser incorporados em ágar. A eficiência geral da transformação é bastante baixa, assim como a produção de DNA clonal (cerca de uma cópia por célula). Todavia, a capacidade para clonar grandes fragmentos (com até 2 Mb) fez dos YACs uma ferramenta vital no mapeamento físico de alguns genomas complexos, especialmente o genoma humano.

### Produção de DNA fita simples para uso em sequenciamento e mutagenese sítio-específica *in vitro*

Os clones de DNA fita simples têm sido os moldes preferidos para sequenciamento de DNA: as sequências obtidas são mais claras e mais fáceis de serem lidas do que as de DNA dupla-fita. Os aspectos de sequenciamento de DNA serão considerados no Capítulo 8.

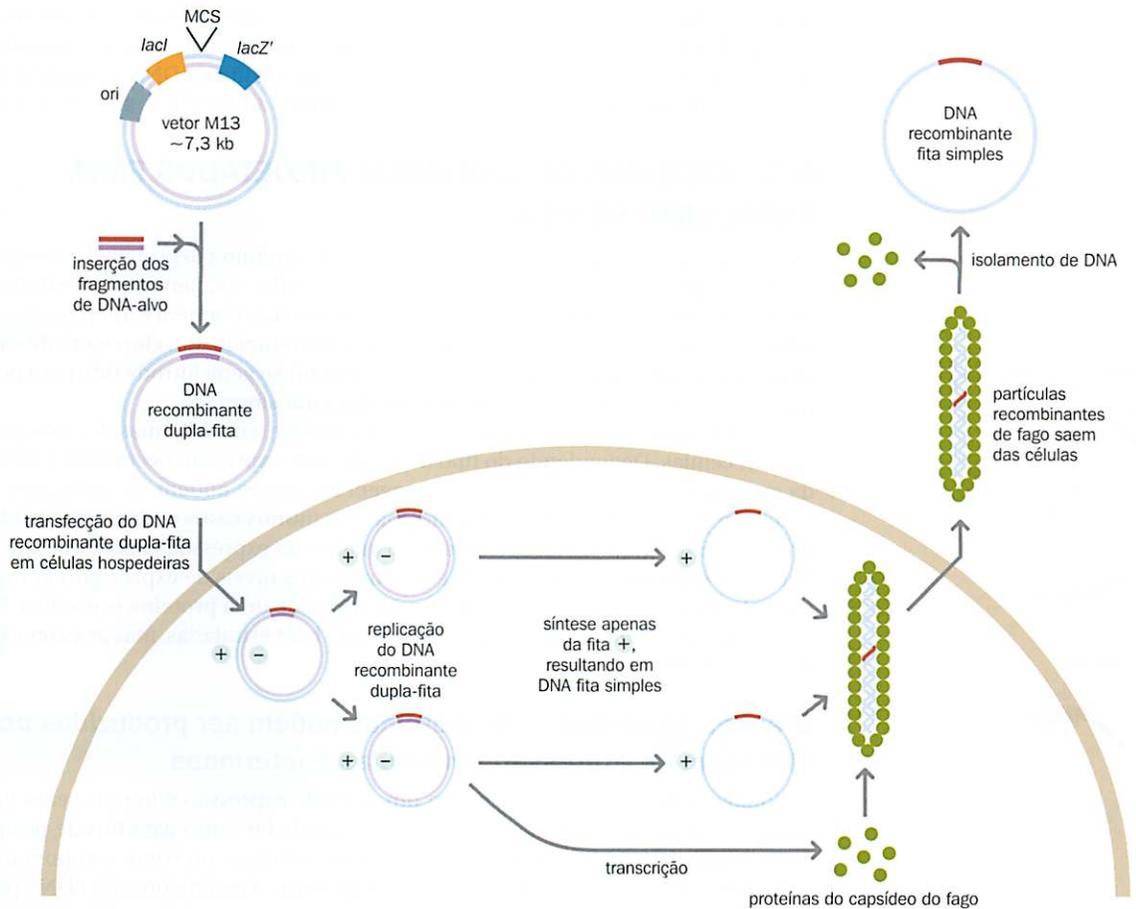
Os moldes de DNA fita simples também são substratos ideais para **mutagenese sítio-específica *in vitro***, uma ferramenta experimental inestimável para esmiuçar as contribuições funcionais feitas por sequências curtas de nucleotídeos ou aminoácidos. Por exemplo, a contribuição esperada de um aminoácido específico para a função de sua proteína pode ser testada por meio da sua deleção ou da sua mutação, dando origem a um aminoácido diferente com uma cadeia lateral diferente. Para isso, se permite que um DNA clonado fita simples com a sequência do tipo selvagem forme pares de base com um oligonucleotídeo sintético que é designado para ser perfeitamente complementar à sequência de bases, exceto por algumas mudanças que corresponderão à mutação desejada. O oligonucleotídeo pode prover síntese de DNA novo para criar uma sequência completamente complementar à mutação desejada. O recém-formado *heterodúplce* tipo selvagem-mutante é utilizado para transformar células, e a sequência mutante desejada pode ser identificada por uma varredura pela mutação (Figura 6.8).

O DNA recombinante fita simples é frequentemente obtido pela utilização de sistemas de clonagem baseados em bacteriófagos filamentosos tais como os fagos M13, fd e f1, os quais adotam naturalmente uma forma de DNA fita simples em alguns estágios de seus ciclos de vida. Alternativamente, são utilizados vetores plasmidiais que contêm sequências de DNA de fagos filamentosos que regulam a produção de DNA fita simples.

### Vetores M13

O bacteriófago M13 possui um genoma de DNA circular fita simples de 6,4 kb revestido por um capsídeo proteico, formando uma longa estrutura filamentososa. Ele infecta algumas linhagens de *E. coli*. Após o fago M13 se ligar à parede de uma célula bacteriana, ele injeta seu DNA genômico dentro da célula. Logo após o DNA fita simples do M13 penetrar na célula, ele é convertido em DNA dupla-fita, a forma replicativa que serve de molde para fabricar numerosas cópias do genoma. Após um curto espaço de tempo, um produto codificado pelo fago desvia a síntese de DNA em direção à produção de fitas simples que migram até a membrana da célula. Aqui, elas são revestidas em um capsídeo proteico, e centenas de partículas maduras de fago são expulsas da célula infectada sem que ocorra lise celular.

Os vetores M13 são baseados na forma replicativa de dupla-fita, a qual foi modificada para conter um sítio múltiplo de clonagem e um sistema de seleção *lacZ* para identifica-



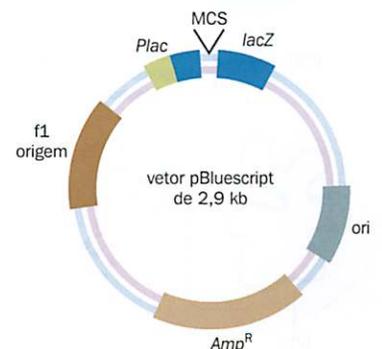
**Figura 6.9** Produção de DNA recombinante fita simples. Vetores M13 utilizam o mesmo teste de *lacZ* para triagem de recombinantes utilizado para pUC19 (ver Figura 6.4). O DNA-alvo é inserido em um sítio de clonagem múltipla (MCS). O recombinante de DNA dupla-fita M13 entra no ciclo normal de replicação do fago para gerar numerosas cópias do genoma, antes de alternar para a produção de DNA fita simples (apenas a fita +). Após o empacotamento com o capsídeo do fago M13, as partículas recombinantes maduras de fago saem da célula sem que haja lise. O DNA recombinante fita simples é recuperado por precipitação das partículas de fago expulso, e as proteínas são quimicamente removidas.

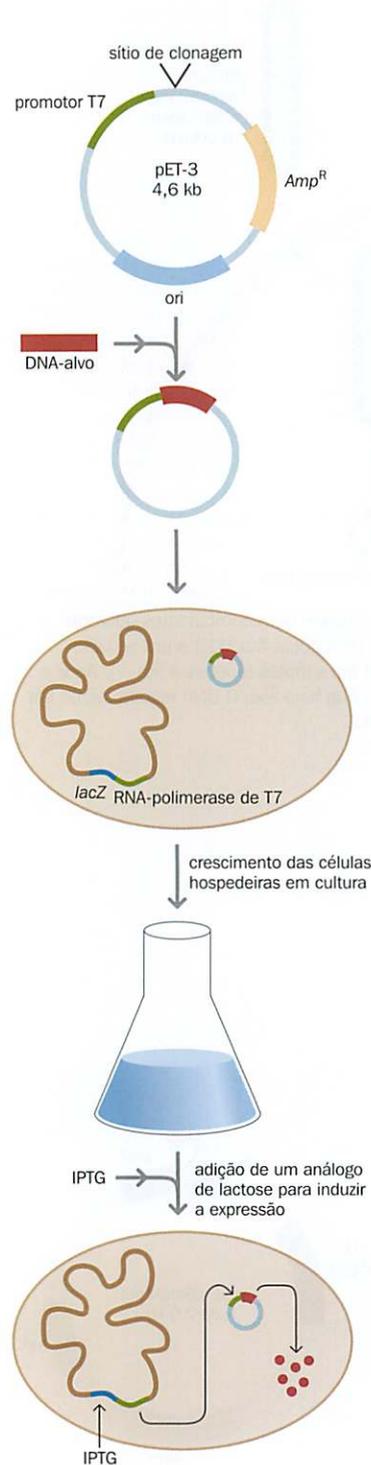
ção dos recombinantes. Recombinantes M13 dupla-fita podem ser transferidos para uma linhagem adequada de *E. coli*; após certo período, partículas do fago são selecionadas e utilizadas para recuperar DNA recombinante de fita simples (Figura 6.9).

### Vetores fagemídeos

Um pequeno segmento de um genoma de fita simples e filamentosos de um bacteriófago pode ser inserido em um plasmídeo para formar um vetor híbrido conhecido como **fagemídeo**. A Figura 6.10 mostra o vetor pBluescript como um exemplo. O segmento escolhido do genoma do fago contém todos os elementos *cis*-atuantes necessários para a replicação do DNA e montagem dentro das partículas de fago. Eles permitem a clonagem bem-sucedida dos insertos com várias quilobases de comprimento, ao contrário dos vetores M13, nos quais os insertos deste tamanho tendem a ser instáveis. Após a transformação de uma linhagem adequada de *E. coli* com o fagemídeo recombinante, as células bacterianas são *superinfectadas* com um *fago auxiliar* filamentosos, tal como f1, o qual é requerido para fornecer o capsídeo proteico. As partículas de fago secretadas a partir das células su-

**Figura 6.10** pBluescript, um vetor fagemídeo. A série pBluescript dos vetores fagemídeos contém, cada uma, duas origens de replicação, uma origem padrão de plasmídeos (*ori*) e uma segunda originária do fago filamentosos f1. A origem de replicação de f1 permite a produção de DNA fita simples, mas o vetor não possui os genes para a produção das proteínas do capsídeo. O DNA-alvo é inserido dentro do vetor fagemídeo, o qual é, então, utilizado para transformar células hospedeiras de *E. coli*. A superinfecção de células transformadas com um fago auxiliar M13 permite que o DNA do fagemídeo recombinante seja empacotado dentro de capsídeos proteicos do fago, e, então, o DNA recombinante fita simples pode ser recuperado como na Figura 6.9.





perinfectedas serão uma mistura de fagos auxiliares e fagemídeos recombinantes. A população de DNA fita simples misturada pode ser diretamente utilizada para sequenciamento devido ao fato de o *primer* para iniciação da síntese da fita de DNA ser projetado para se ligar especificamente à sequência do vetor fagemídeo adjacente ao sítio de clonagem.

### 6.3 SISTEMAS DE CLONAGEM PROJETADOS PARA EXPRESSÃO GÊNICA

Os sistemas de clonagem descritos até agora neste capítulo são projetados, simplesmente, para amplificar o DNA-alvo a fim de obter quantidades suficientes para estudos estruturais e funcionais. Entretanto, em diversas circunstâncias, também é interessante a possibilidade de expressar, de alguma maneira, o gene introduzido. Em **clonagem de expressão**, sinais apropriados precisam ser fornecidos próximo ao gene introduzido para possibilitar que o mesmo seja expresso na célula hospedeira transformada.

A clonagem de expressão é realizada, geralmente, com sistemas de clonagem baseados em células. Dependendo do tipo do produto de expressão necessário e do propósito da expressão, diversos sistemas de clonagem diferentes podem ser utilizados. Algumas vezes um produto de RNA é suficiente, mas em muitos casos o objetivo é produzir uma proteína. Em alguns casos, é necessária a análise da expressão gênica quando os níveis baixos de expressão são aceitáveis; em outros, altos níveis de expressão são necessários para alcançar, por exemplo, grandes quantidades de uma proteína específica. É comum expressar o produto em células que tenham sido bem estudadas, mas às vezes expressar o produto *in vitro* pode ser suficiente.

#### Grandes quantidades de proteínas podem ser produzidas por clonagem de expressão em células bacterianas

A clonagem de cDNA eucariótico em um vetor de expressão é muitas vezes necessária para a produção de proteínas em grandes quantidades tanto para fins de pesquisa, tais como estudos estruturais, como para fins biotecnológicos ou como componentes medicinalmente relevantes, tais como proteínas terapêuticas. Geralmente, um cDNA portando a informação genética que especifica a sequência da proteína é inserido em um vetor juntamente com sinais separados de expressão, tais como promotores de expressão fortes e outros elementos regulatórios. Devido ao sistema de expressão ser baseado em DNA recombinante, as proteínas resultantes são descritas, às vezes, de forma um tanto quanto inadequada, como **proteínas recombinantes**.

As células bacterianas possuem a vantagem de crescer rapidamente e podem ser facilmente expandidas a grandes volumes em cultura. A *Escherichia coli* tem sido a célula hospedeira favorita para a expressão de proteínas induzidas.

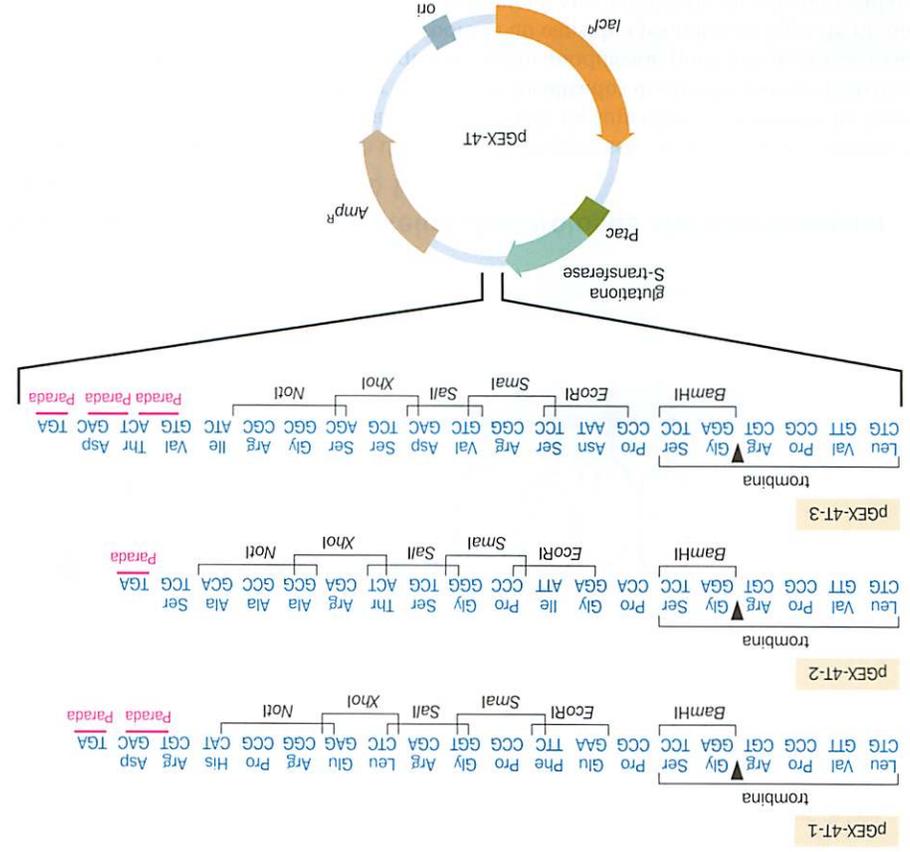
A produção de grandes quantidades de uma proteína não endógena pode, no entanto, ser prejudicial ao crescimento da célula hospedeira, além de ser, algumas vezes, tóxica. Então, é vantajoso utilizar um *promotor induzível*, de maneira que a expressão possa ser atrasada até que as células transformadas tenham sido identificadas e crescido em tamanho. Por exemplo, a série pET de vetores contém um promotor do bacteriófago T7 o qual não é reconhecido pelo promotor endógeno de RNA-polimerase de *E. coli* (Figura 6.11). Tais vetores são utilizados para transformar uma linhagem de *E. coli* geneticamente modificada que contém uma RNA-polimerase de T7 regulada por um promotor *lac*. As células transformadas podem ser selecionadas e crescidas em grandes quantidades sem que haja expressão do gene externo. A adição do indutor de  $\beta$ -galactosidase IPTG ativa o promotor *lac* e a expressão do gene exterior adjacente, e as células podem ser selecionadas posteriormente.

As bactérias possuem diversas vantagens em expressar proteínas heterólogas (estranheiras), mas também possuem limitações. Muitas proteínas eucarióticas são modificadas

**Figura 6.11 Vetores de indução bacteriana induzíveis.** A série de vetores plasmidiais pET-3 contém um promotor de bacteriófago T7. Eles são utilizados com uma linhagem de *E. coli* geneticamente modificada para conter um gene para a RNA-polimerase de T7 sob controle de um promotor *lac* induzível. Após a transformação com DNA recombinante pET-3, as células de *E. coli* podem crescer e originar grandes números de células em cultura. Em um estágio desejado, o análogo à lactose IPTG (isopropil  $\beta$ -D-tiogalactosídeo) é adicionado para induzir a expressão do gene de RNA-polimerase de T7 do hospedeiro. A RNA-polimerase de T7 induzida se liga especificamente ao promotor T7 no recombinante de DNA para originar altos níveis de expressão do inserto. *Amp<sup>R</sup>*, gene de resistência à ampicilina; ori, origem de replicação.

**Figura 6.12** Vetores de proteínas de fusão. A série de vetores pGEX-4T possui

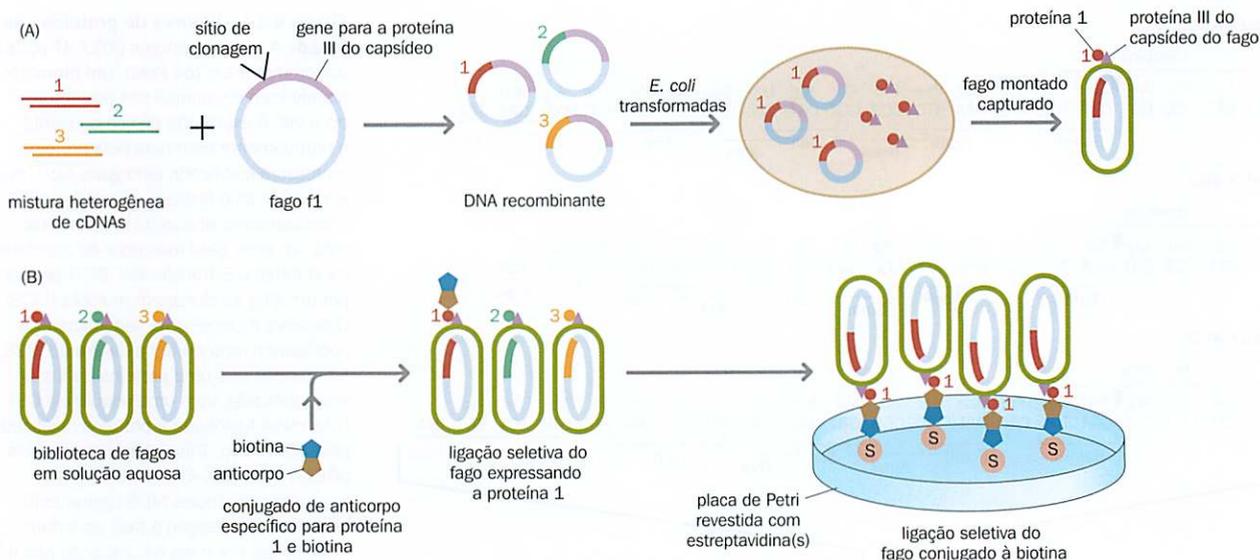
um promotor *lac* (ou *P<sub>lac</sub>*), um promotor híbrido com elementos dos promotores *trp* e *lac*. A expressão gênica a jusante é normalmente reprimida pela proteína repressora codificada pelo gene *lacI*, mas é induzível pelo análogo da lactose, IPTG. Imediatamente abaixo do promotor *lac* está um gene para marcador de afinidade da glicona S-transferase (GST) seguido por um sítio de clonagem múltipla (MCS). O objetivo é clonar uma sequência-alvo codificando uma proteína dentro do MCS, de maneira que uma proteína de fusão seja produzida, com uma sequência GST N-terminal fusionada à proteína codificada pelo cDNA-alvo. Três vetores alternativos, pGEX-4T-1, pGEX-4T-2 e pGEX-4T-3, possuem sequências MCS ligeiramente diferentes em relação à fase de leitura tradicional. Por meio da utilização dos três vetores alternativos, os insertos clonados podem ser expressos em cada uma das três fases de leitura, de forma que, em pelo menos um dos três casos, o DNA-alvo esteja na mesma fase da sequência GST. A proteína de fusão expressa pode ser facilmente purificada em uma coluna com afinidade por glicona, tal como a glicona Sepharose 4B. Devido ao fato de o MCS ser projetado para conter um sítio de clivagem por trombina, a proteína desejada pode ser purificada após a clivagem neste sítio. *Amp<sup>R</sup>*, gene de resistência a ampicilina; ori, origem de replicação; STOP, códon de terminação.



por adição de grupos fosfato, lipídicos ou de açúcar após a tradução. Essas modificações são essenciais para a função biológica da proteína. Células bacterianas não possuem as enzimas necessárias para seu processamento pós-traducional, e então algumas proteínas eucarióticas produzidas em células bacterianas se tornam instáveis ou não possuem ou mostram atividade biológica limitada. Diversas proteínas eucarióticas, especialmente as gomas proteínas de mamíferos, são também muito maiores que proteínas bacterianas e não podem ser facilmente sintetizadas em *E. coli*. A superexpressão leva à produção de agregados insolúveis de proteínas configuradas de forma anormal (*corpos de inclusão*). Os corpos de inclusão podem ser facilmente purificados, mas é difícil solubilizar as proteínas e obter uma reconfiguração eficiente *in vitro*.

Esforços para aumentar o rendimento e a solubilidade envolvem, frequentemente, a produção de **proteínas de fusão**. Por exemplo, o vetor pode ser modificado de maneira que ele contenha, imediatamente adjacente ao sítio de clonagem, sequências de DNA para toda ou parte de uma proteína endógena. Os recombinantes expressarão, dessa forma, a proteína desejada fusionada a uma sequência de proteína endógena. Diversos vetores modernos de expressão de proteínas são modificados para conter uma sequência de DNA codificante para um peptídeo ou proteína especificos que são facilmente purificados por  *Cromatografia de afinidade*. Nestes casos, os recombinantes são expressos para dar origem à proteína desejada, ligada a um peptídeo ou proteína curtos, conhecidos como **marcadores de afinidade** porque se ligam à proteína de interesse para auxiliar na purificação das proteínas recombinantes por cromatografia por afinidade.

Dois dos sistemas favoritos que permitem a purificação de proteínas expressas são baseados na afinidade por glicona-GST e por íon de poli-histidina-níquel. A glicona S-transferase (GST) é uma pequena proteína com uma grande afinidade por substrato de glicona. O clone do vetor de expressão posiciona o DNA-alvo logo após o gene codificante GST, de maneira que a proteína de fusão seja produzida na célula transformada (Figura 6.12). Esta proteína de fusão pode ser purificada por ligação seletiva a uma coluna contendo glicona. Alternativamente, um marcador de afinidade de seis resíduos de histidina consecutivos podem se ligar a uma proteína. As cadeias laterais do marcador His<sub>6</sub> ligam de forma seletiva a íons de níquel, auxiliando na purificação por cromatografia por afinidade com uma matriz de ácido níquel-tiacético-níquel.



**Figura 6.13** Expressão em fagos. (A) Construção da biblioteca. Uma mistura heterogênea de cDNAs é clonada dentro de um vetor de fago, tal como a que é baseada nos fagos M13 ou f1, para expressar proteínas externas na superfície do fago. Neste exemplo, o DNA é inserido em um sítio de clonagem localizado dentro da parte inicial da sequência codificante para o gene do fago f1 que produz uma das proteínas do capsídeo do fago (proteína III). A ideia é que recombinantes devem expressar uma proteína de fusão com a proteína-alvo fusionada a esta proteína do capsídeo do fago, possibilitando, assim, que as partículas do fago liberem a proteína externa na sua superfície. Após a transformação das células de *E. coli*, o DNA do fago é replicado, e as suas partículas são montadas, liberadas e espalhadas. A mistura de fagos recombinantes é conhecida como biblioteca de expressão de fagos. (B) Triagem da biblioteca com um anticorpo específico. Um anticorpo específico para apenas um tipo de proteína se ligará apenas a partículas do fago que expressem esta proteína. Se o anticorpo é complexado com biotina, as partículas do fago portando a proteína desejada podem ser purificadas por meio da utilização de estreptavidinas revestidas em placas de Petri: os recombinantes desejados possuirão um anticorpo complexado a um grupo biotina, e o grupo biotina se ligará fortemente à estreptavidina.

### Na expressão em fagos, proteínas heterólogas são expressas na superfície das partículas do fago

A **expressão em fagos** é uma forma de clonagem de expressão na qual o gene clonado é expresso para originar uma proteína a qual é exibida na superfície da partícula do fago. Para isso, sítios de clonagem no vetor de fago são localizados dentro do gene do fago que codifica uma proteína de capsídeo adequada. O gene modificado pode então ser expresso como uma proteína de fusão que é incorporada ao capsídeo proteico do fago, de modo que ela seja liberada na superfície do fago sem afetar sua habilidade de infectar células. Se há um anticorpo disponível para a proteína expressa, os fagos que exibem esta proteína podem ser selecionados por ligação preferencial ao anticorpo: a purificação por afinidade das partículas do vírus que portam tal proteína pode ser realizada a partir de  $10^8$  vezes o excesso de fagos que não exibem a proteína, utilizando até mesmo quantidades minúsculas do anticorpo relevante (Figura 6.13).

Em engenharia de proteínas, a expressão em fagos é uma ferramenta auxiliar poderosa para programas de mutagênese randômica como uma maneira de seleção das variantes desejadas a partir de uma biblioteca de mutantes. A expressão em fagos se mostrou, também, uma poderosa fonte alternativa de construção de anticorpos, passando por cima de técnicas de imunização normais e até de tecnologias de Híbridomas. Bibliotecas de fagos também podem ser utilizadas para identificar proteínas que interagem com uma proteína específica. Da mesma maneira que anticorpos podem ser utilizados em investigações de afinidade, uma proteína conhecida (ou qualquer outra molécula à qual uma proteína possa se ligar) pode ser utilizada como uma isca para selecionar os fagos que exibem quaisquer outras proteínas que se liguem à proteína isca.

### A expressão de genes eucarióticos é realizada com maior fidelidade em linhagens celulares eucarióticas

Sistemas de expressão bacteriana oferecem a grande vantagem de a expressão proteica ocorrer em níveis extremamente elevados. Entretanto, as propriedades biológicas de várias proteínas eucarióticas sintetizadas em bactérias podem não representar as moléculas nativas. Quando produzidas em células bacterianas, elas não sofrem seu processamento pós-traducional normal, e a conformação da proteína eucariótica pode ser incorreta ou ineficiente. Como alternativa, células eucarióticas, incluindo células de insetos e mamíferos, são amplamente utilizadas para expressar proteínas recombinantes. Os vetores de expressão são geralmente projetados com a capacidade para serem replicados tanto nas células eucarióticas desejadas como em células de *E. coli*, em que a intenção é, simplesmente, produzir grandes quantidades de DNA recombinante que possam ser, subsequentemente, utilizadas para estudos de expressão em células eucarióticas.

Uma vez transferido para dentro de células animais, o DNA recombinante pode ser expresso por curtos ou longos períodos de tempo, de acordo com o vetor de expressão utilizado. Para alguns propósitos, só é necessária a *expressão temporária*. Aqui, o DNA re-

combinante permanece como um elemento genético extracromossômico independente (um *episômo*) dentro das células transferidas. A expressão do gene introduzido atinge seu nível máximo cerca de dois ou três dias após a transfecção do vetor de expressão dentro de uma linhagem celular de mamífero; após isso, a expressão pode diminuir rapidamente como um resultado da morte celular ou perda da arquitetura de expressão.

A alternativa é utilizar sistemas de *expressão estável* nos quais o DNA recombinante possa se integrar a um cromossomo da célula hospedeira. A vantagem aqui é que, se o DNA recombinante se integra de maneira que o gene introduzido seja expresso, todas as células descendentes conterão este gene, e a expressão poderá ser mantida ao longo de várias gerações celulares. Alguns vírus são altamente eficientes em se integrar no DNA cromossômico, e serão considerados no contexto de terapia gênica em um capítulo posterior. Mas vetores de expressão plasmidial também podem, randomicamente, se integrar em frequências mais baixas, e linhagens germinativas estáveis podem ser desenvolvidas com um gene de interesse integrado.

### Expressão temporária em células de insetos utilizando baculovírus

A expressão gênica em baculovírus é um método popular para produzir grandes quantidades de proteínas recombinantes em células hospedeiras de insetos; as produções de proteínas são mais altas que em sistemas de expressão em células de mamíferos, e os custos são mais baixos. Na maioria dos casos, o processamento pós-traducional de proteínas eucarióticas expressas em células de insetos é similar ao processamento que ocorre em células de mamíferos, e é possível a expressão de proteínas muito grandes. As proteínas produzidas em células de insetos possuem atividade biológica e reatividade imunológica comparáveis às das proteínas expressas em células de mamíferos.

O vírus da poliedrose nuclear, *Autographa californica* (AcMNPV), é um baculovírus que pode ser propagado em certas linhagens celulares e é frequentemente utilizado como vetor de clonagem para sistemas de expressão de proteínas. A proteína viral poliedrina é transcrita em altos níveis e, embora essencial para a propagação viral em seu habitat natural, ela não é necessária em cultura. Consequentemente, sua sequência codificante pode ser substituída por um DNA codificante para uma proteína eucariótica. O vetor de clonagem é projetado para expressar a proteína inserida de um promotor forte de poliedrina, resultando em níveis de expressão de mais de 30% do total da proteína na célula.

### Expressão temporária em células de mamíferos

Expressar proteínas de mamíferos em células de mamíferos possui a vantagem óbvia de que o dobramento proteico correto e as modificações pós-traducionais em células humanas ou de outros mamíferos não são, geralmente, um problema, e é possível analisar sinais e efeitos celulares. Sistemas de expressão estáveis em células de mamíferos, normalmente baseados em sequências de plasmídeos que se integram dentro de um cromossomo, fornecem quilogramas de proteínas complexas em biorreatores de escala industrial, mas requerem grandes investimentos em tempo, recursos e equipamento. Como alternativa, sistemas de expressão temporária em larga escala têm sido desenvolvidos para fornecer proteínas recombinantes em células de mamíferos.

Além de proliferar a expressão proteica, algumas linhagens celulares de mamíferos possuem aplicações ainda maiores em investigações dos efeitos de manipulações *in vitro* em sequências de controles transcricional e pós-transcricional. Um bom exemplo é fornecido pelas células COS, linhagens celulares estáveis derivadas de uma linhagem celular de fibroblastos de rim de macaco verde africano, CV-1. Quando células CV-1 são infectadas com o vírus SV40 de macaco, o ciclo lítico normal do SV40 se segue. Entretanto, quando células CV-1 foram transformadas com uma linhagem de SV40 com uma origem de replicação defeituosa, um segmento do genoma do SV40 se tornou integrado ao cromossomo de CV-1. As células COS resultantes (CV-1 com origens defeituosas de SV40) expressam de maneira estável o grande antígeno T codificado por SV40, a única proteína viral que é necessária para a ativação da origem de replicação do SV40.

Por meio da expressão do grande antígeno T de SV40, de maneira estável, as células COS permitem que qualquer DNA circular introduzido que possua uma origem de replicação SV40 funcional se replique independentemente do cromossomo da célula hospedeira, sem limitação clara de tamanho. Quando vetores de expressão temporária são transferidos para células COS, linhagens celulares permanentes não são originadas, pois a replicação massiva do vetor torna a célula inviável. Ainda que uma pequena proporção de células seja transformada de forma bem-sucedida, a amplificação do DNA introduzido para altos números de cópias naquelas células compensa a baixa taxa de absorção.



Devido à sua simplicidade, a PCR é uma técnica popular com uma ampla gama de aplicações que confiam principalmente em três vantagens: rapidez, sensibilidade extrema e robustez. Sua sensibilidade permite a amplificação de quantidades minúsculas de DNA-alvo – até mesmo DNA de uma única célula –, e várias aplicações já foram descobertas em diagnósticos, pesquisas e ciência forense. Por exemplo, um indivíduo pode ser identificado a partir de análises de PCR de tecidos tais como cabelo ou células descartadas da pele. Sua robustez significa que é frequentemente possível amplificar DNA de tecidos ou células que estejam degradados ou imersos em um meio que torne difícil o isolamento do DNA por métodos-padrão. A amplificação bem-sucedida por PCR é possível a partir de pequenas quantidades de DNA degradado extraído de tecidos decompostos em locais históricos ou arqueológicos e de amostras de tecidos fixadas em formalina.

### A PCR pode ser utilizado para amplificar de forma seletiva um DNA-alvo raro dentro de uma população complexa

Geralmente, a PCR é projetado para permitir replicação seletiva de uma ou mais sequências de DNA-alvo específicas dentro de uma coleção heterogênea de sequências de DNA, tornando possível um vasto aumento em seu número de cópias (amplificação). Muitas vezes, o DNA inicial é o DNA genômico total de uma amostra tecidual ou de células de cultura, e o DNA-alvo é, com frequência, uma pequena fração de DNA inicial. Por exemplo, se quiser amplificar um gene de  $\beta$ -globina de 1,6 kb de DNA genômico humano (com um tamanho de genoma haploide de 3.200 Mb), a relação de DNA-alvo para DNA inicial é de 1,6 kb:3.200 Mb ou 1:2.000.000. Muitas reações de PCR envolvem a amplificação de alvos ainda menores. Por exemplo, éxons humanos simples possuem, em média, 290 pb e, então, representam menos de 0,000001% de uma população inicial de DNA genômico.

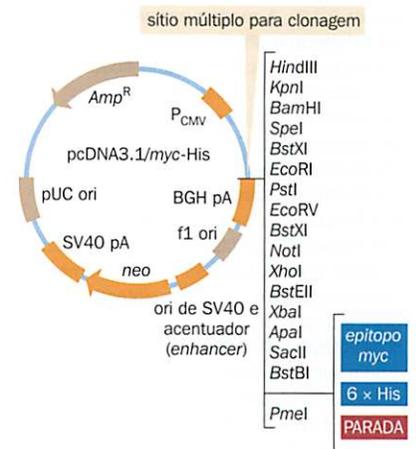
Como uma alternativa ao DNA genômico, o DNA inicial pode ser totalmente preparado como cDNA, por meio do isolamento de RNA de um tecido ou linhagem celular adequada e posterior conversão em DNA com a enzima transcriptase reversa. Isso é conhecido como PCR por transcriptase reversa ou RT-PCR, por abreviação. Se a população original de RNA possui várias cópias do transcrito do DNA-alvo, pode haver um enriquecimento significativo da sequência-alvo em comparação com sua representação no DNA genômico. Sob condições ótimas, a RT-PCR também pode ser usada, algumas vezes, para amplificar sequências-alvo a partir de tecidos nos quais são expressas somente em nível basal.

### A natureza cíclica da PCR leva à amplificação exponencial do DNA-alvo

A PCR depende da ação de síntese de sequências de oligonucleotídeos como *primers* para sintetizar um novo DNA em sequências-alvo específicas contidas no DNA inicial; a ideia é replicar seletivamente apenas a(s) sequência(s)-alvo. Ela consiste em uma série de ciclos de três reações sucessivas conduzidas em temperaturas diferentes, por isso o processo é descrito, algumas vezes, como *termociclagem*. Primeiro, o DNA inicial é aquecido a uma temperatura que é alta o suficiente para quebrar as pontes de hidrogênio que ligam as duas fitas complementares de DNA. Como resultado, a estrutura de dupla-fita se separa para dar origem ao DNA fita simples (**desnaturação**). Para o DNA genômico humano, a mistura de reação é geralmente aquecida em torno de 93 a 95°C para o passo de desnaturação. Após o resfriamento, os *primers* oligonucleotídicos sintéticos podem se ligar, por pareamento de bases, a sequências complementares no DNA fita simples (**anelamento**). Em seguida, na presença dos quatro desoxinucleosídeos trifosfato dATP, dCTP, dGTP e dTTP, uma DNA-polimerase purificada inicia a síntese de novas fitas de DNA que são complementares às fitas individuais de DNA do segmento de DNA-alvo.

A orientação dos *primers* é deliberadamente escolhida de maneira que a direção da síntese das novas fitas de DNA seja em direção ao sítio de anelamento do outro *primer*. Sendo assim, as fitas recentemente sintetizadas podem, uma por vez, servir de molde para a síntese de DNA novo, causando uma reação em cadeia com aumento exponencial de produto (**Figura 6.16**). As três reações sucessivas de desnaturação, anelamento dos *primers* e síntese de DNA são repetidas até 30 a 40 vezes em uma reação de PCR padrão.

O passo de síntese de DNA da PCR ocorre, geralmente, a cerca de 70 a 75°C, mas a DNA-polimerase utilizada também precisa suportar uma temperatura muito mais alta no passo de desnaturação. A necessidade de uma enzima termoestável levou os pesquisadores a isolarem a DNA-polimerase de microrganismos cujo hábitat natural são águas termais. Um exemplo antigo e amplamente utilizado é a Taq-polimerase de *Thermus*

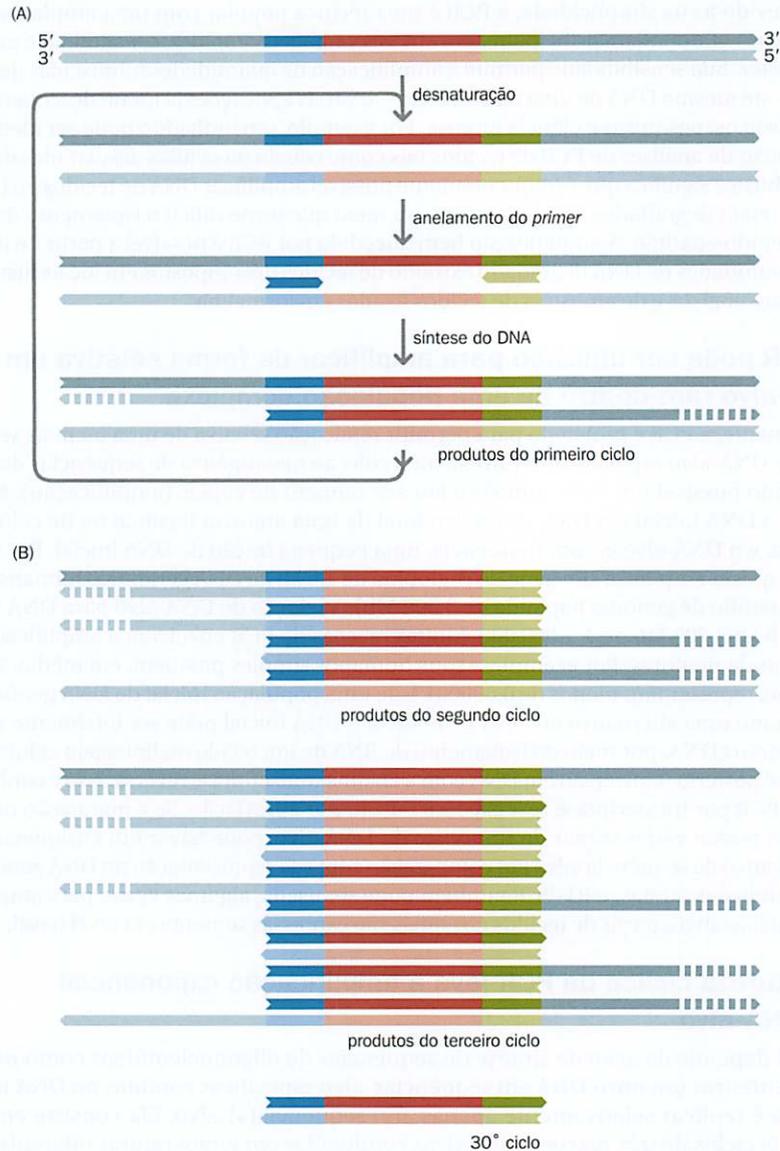


**Figura 6.15 Um vetor de expressão em mamíferos, pcDNA3.1/myc-His.** A série de vetores de expressão plasmidiais pcDNA da Invitrogen proporciona a expressão em altos níveis de proteínas constitutivas em células de mamíferos. Aqui é mostrado o vetor de expressão pcDNA3.1/myc-His. Insetos de DNA clonados podem ser transcritos a partir do promotor forte de citomegalovírus ( $P_{CMV}$ ), o qual garante altos níveis de expressão em células de mamíferos, e os transcritos são poliadenilados com o auxílio de um elemento de sequência de poliadenilação de um gene de hormônio de crescimento bovino (BGH pA). Os sítios múltiplos de clonagem são seguidos por duas sequências codificantes curtas que podem ser expressas para originar marcadores dos peptídeos. O epítipo marcador *myc* permite uma varredura por recombinantes com um anticorpo específico para o peptídeo Myc, e um marcador de afinidade de seis resíduos de histidina consecutivos ( $6 \times His$ ) facilita a purificação da proteína recombinante por cromatografia por afinidade com uma matriz de ácido níquel-nitrilotriacético. A tradução termina em um códon de terminação determinado (PARADA). Um gene *neo* marcador (regulado por um promotor SV40 e uma sequência poli(A) SV40) permite a seleção por crescimento de células hospedeiras transformadas com o antibiótico G418. Componentes para propagação em *E. coli* incluem uma origem de replicação permissiva de plasmídeo ColE1 (ori pUC) e um gene de resistência a ampicilina ( $Amp^R$ ) mais uma origem de replicação do fago f1 (ori f1), a qual fornece uma opção para a produção de DNA recombinante fita simples.

**Figura 6.16** A reação em cadeia da polimerase (PCR).

(A) DNA contendo sequência a ser amplificada (a sequência-alvo, mostrada aqui em vermelho) é, primeiramente, desnaturado por aquecimento. Os *primers* de oligonucleotídeos, projetados para serem complementares às sequências de DNA que flanqueiam a sequência-alvo, na extremidade 3' em fitas opostas (setas azul e verde), podem então se ligar (anelamento dos *primers*). Após, ocorre a síntese de DNA, e os *primers* são incorporados nas novas fitas de DNA sintetizadas. O primeiro ciclo resulta em duas novas fitas de DNA cujas extremidades 5' são fixadas pela posição do *primer* de oligonucleotídeo, mas cujas extremidades 3' se estendem além do outro *primer* (indicado por linhas pontilhadas).

(B) Após o segundo ciclo, as quatro novas fitas consistem em mais dois produtos com extremidades 3' variáveis, mas agora as duas fitas possuem seu tamanho em ambas extremidades 5' e 3' definido pelas sequências dos *primers*. Após o terceiro ciclo, seis das oito novas fitas consistem apenas em DNA-alvo e região flanqueadora; após 30 ciclos, ou mais, com um aumento exponencial, este será o produto predominante.



*aquaticus*. Entretanto, a Taq-polimerase não possui atividade de exonuclease 3'→5' para promover uma revisão. Todas as polimerases cometem erros ocasionais por meio da inserção de uma base errada, mas estes erros de cópia podem ser corrigidos por uma atividade exonuclease de revisão 3'→5'. Consequentemente, outras bactérias termofílicas foram testadas como fontes de polimerases termoestáveis como com atividade de revisão 3'→5', tal como a Pfu-polimerase de *Pyrococcus furiosus*.

### A amplificação seletiva de sequências-alvo depende da ligação altamente específica de sequências de *primers*

Para permitir a amplificação seletiva, algumas informações sobre a sequência do DNA-alvo são necessárias. As informações são utilizadas para projetar dois *primers* oligonucleotídeos com um comprimento de 18 a 25 nucleotídeos que são específicos para duas sequências imediatamente flanqueadoras ao DNA-alvo.

Para a maioria das reações de PCR, o objetivo é amplificar uma única sequência de DNA, então é importante reduzir as chances dos *primers* se ligarem a outras localidades no DNA que não as desejadas. É importante evitar sequências repetitivas de DNA. A fim de não permitir a ligação dos *primers* uns aos outros, a sequência 3' de um *primer* não deve ser complementar em sequência a nenhuma região do outro *primer* na mesma reação. Para garantir que um *primer* não sofra pareamento de bases internas, sua sequência não

deve incluir repetições invertidas ou qualquer sequência complementar a si mesma que seja maior em comprimento que 3 pb.

Para garantir que o *primer* se ligue com um alto grau de especificidade à sua sequência complementar, a temperatura para o passo de anelamento do *primer* deve ser estabelecida de maneira que seja tão alta a ponto de permitir o pareamento entre ele e seu sítio de ligação. Uma medida útil de estabilidade de qualquer dúplice de ácido nucleico é a **temperatura de fusão** (do inglês  $T_m$  - *melting temperature*), a temperatura que corresponde ao ponto médio nas transições observadas da forma de dupla-fita para a forma de fita simples. Para garantir uma alta especificidade na ligação do *primer*, a temperatura de anelamento é em geral estabelecida para ser em torno de 5°C abaixo da temperatura de fusão calculada; se a temperatura de anelamento fosse muito mais baixa, os *primers* tenderiam a se ligar a outras regiões no DNA que possuíssem sequências parcialmente complementares.

As temperaturas de fusão e de anelamento do *primer* dependem da composição de bases. Isto porque os pares de base GC possuem três pontes de hidrogênio e os pares de base AT possuem apenas duas. Fitas com uma alta porcentagem de pares de base GC são, então, muito mais difíceis de se separar do que as que possuem uma baixa composição de pares de base GC. Os melhores *primers* possuem um conteúdo de GC entre 40 e 60% com uma distribuição de todos os quatro nucleotídeos. Além disso, os valores de  $T_m$  calculados para os dois *primers* utilizados juntamente não devem diferir por mais de 5°C, e o valor de  $T_m$  no DNA-alvo amplificado não deve diferir dos valores dos *primers* por mais de 10°C.

O desenvolvimento de *primers* é auxiliado por diversos *softwares* comerciais e alguns programas gratuitos, os quais podem ser acessados pela internet em sites de compilação, tais como [http://www.humgen.nl/primer\\_design.html](http://www.humgen.nl/primer_design.html). Como visto a seguir, é particularmente importante que as extremidades 3' dos *primers* sejam perfeitamente combinadas às sequências desejadas. Diversas modificações são frequentemente utilizadas para reduzir as chances de ligação não específica de *primers*, incluindo *hot start* PCR, PCR *touch-down* e a utilização de *primers* agrupados (Quadro 6.2).

### A PCR é desvantajosa como um método de clonagem de DNA por pequenos comprimentos e comparativamente baixos rendimentos de produtos

A PCR é um método de escolha para amplificar seletivamente quantidades muito pequenas de DNA-alvo, de maneira rápida e acurada. Ela é ajustada de forma ideal para conduzir análises rápidas de sequências genômicas ou de cDNA. Entretanto, possui desvantagens como método de clonagem de DNA. Em particular, a gama de variação de tamanho dos produtos de amplificação em uma reação de PCR padrão é raramente maior que 5 kb, enquanto a clonagem de DNA que utiliza células como base torna possível a clonagem de fragmentos de até 2 Mb. Ao passo que o tamanho do produto desejado aumenta, torna-se mais difícil obter uma amplificação eficiente com PCR. Para contornar esse problema, protocolos de PCR de longo alcance que utilizam uma mistura de dois tipos de polimerases termoestáveis foram desenvolvidos em um esforço para fornecer níveis ótimos de DNA-polimerase e atividade de exonuclease de revisão 3'→5'. Utilizando protocolos modificados, podem ser obtidos produtos de PCR com quilobases de comprimento; DNAs maiores clonados podem ser obtidos somente por meio de clonagem utilizando células como base.

A PCR geralmente envolve apenas 30 a 40 ciclos, tempo no qual a reação atinge uma fase estável, ao passo que os reagentes vão se esgotando e os inibidores vão se acumulando. Quantidades em nível de microgramas do produto de DNA desejado podem ser obtidas, mas de forma demorada e bastante cara para fazer a reação atingir quantidades maiores de DNA. Além disso, o produto de PCR pode não estar em uma forma adequada, que permita determinados estudos subsequentes. Então, quando quantidades particularmente grandes de DNA são necessárias, é frequentemente conveniente clonar o produto do PCR em uma clonagem que utiliza células como base. Sistemas de clonagem plasmidiais são utilizados para propagar em células bacterianas DNA clonado por PCR. Uma vez clonado, o inserto pode ser clivado com endonucleases de restrição apropriadas e transferido para outros plasmídeos que, por exemplo, permitam sua expressão para originar um produto proteico ou de RNA.

Diversas polimerases termoestáveis, rotineiramente utilizadas, possuem uma atividade terminal de desoxinucleotidil-transferase que modifica de forma seletiva fragmentos gerados por PCR pela adição de um único nucleotídeo, geralmente desoxiadenosina, à extremidade 3' dos fragmentos de DNA amplificados. As extremidades 3' livres podem, às vezes, tornar difícil a clonagem de produtos de PCR em células, e vários métodos são utilizados

## QUADRO 6.2 Alguns métodos de PCR comuns

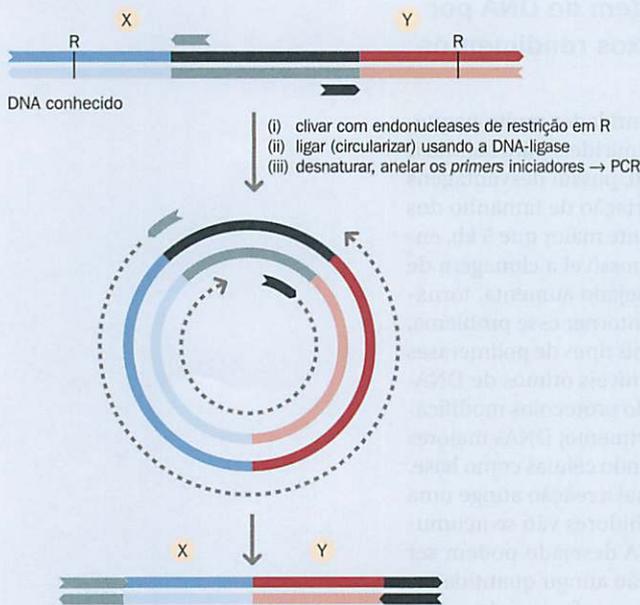
**PCR alelo-específico.** Projetado para amplificar uma sequência de DNA excluindo a possibilidade de amplificar outros alelos. Baseado na necessidade de pareamento preciso entre a extremidade 3' de um *primer* de PCR e o DNA-alvo. Ver Figura 6.17.

**PCR ancorado.** Utiliza um *primer* de sequência específica e um *primer* universal para amplificar sequências adjacentes a uma sequência conhecida. O *primer* universal reconhece e se liga a uma sequência comum que é artificialmente ligada a todas as diferentes moléculas de DNA.

**PCR-DOP (PCR com primers de oligonucleotídeos degenerados).** Utiliza *primers* de oligonucleotídeos parcialmente degenerados (conjuntos de sequências oligonucleotídicas que foram sintetizadas em paralelo para obter a mesma base em certas posições nucleotídicas, diferindo em outras posições) para amplificar uma variedade de DNAs-alvo relacionados.

**PCR hot-start.** Uma maneira de aumentar a especificidade da reação de PCR. A mistura de todos os reagentes do PCR antes de um passo inicial de desnaturação por aquecimento possibilita mais oportunidades de ligação não específica de *primers*. Para reduzir essa possibilidade, um ou mais componentes da PCR são fisicamente separados até o primeiro passo de desnaturação.

**PCR inverso.** Uma maneira de acessar o DNA que fica imediatamente adjacente a uma sequência conhecida. Neste caso, a população inicial de DNA é digerida com uma endonuclease de restrição, diluída em baixas concentrações de DNA e, então, tratada com DNA-ligase para auxiliar a formação de moléculas de DNA circular por ligação intramolecular. Os *primers* de PCR são posicionados de modo a se ligarem a uma sequência de DNA conhecida e, então, iniciarem a síntese de DNA novo na direção contrária, levando à amplificação da sequência desconhecida. Ver Figura 1 (X e Y são sequências não caracterizadas flanqueando uma sequência conhecida).



**PCR com primers específicos de conectores.** Uma forma de amplificação indiscriminada que envolve a ligação de nucleotídeos conectores a ambas as extremidades de todos os fragmentos de DNA em um DNA inicial e a amplificação de todos os fragmentos por meio da utilização de um *primer* com ligação específica (ver Figura 6.18).

**PCR com primers específicos.** Uma maneira de aumentar a especificidade de uma reação de PCR. Os produtos de uma reação de amplificação inicial são diluídos e utilizados como fonte de DNA inicial para uma segunda reação na qual um conjunto diferente de *primers* é utilizado, correspondendo às sequências localizadas próximas, mas internas, das utilizadas na primeira reação.

**PCR competitivo.** Uma forma de PCR com *primers* ancorados (veja anteriormente) para a amplificação rápida das extremidades de cDNA. Isso será descrito em um capítulo posterior.

**PCR em tempo real (também chamado de PCR quantitativo e qPCR).** Utiliza um termociclador de detecção por fluorescência para amplificar sequências específicas de ácidos nucleicos e para realizar uma quantificação simultânea de suas concentrações. Existem duas aplicações de pesquisa principais: (1) quantificar expressão gênica (e confirmar a expressão diferencial de genes por meio de análises de hibridização por microarranjos) e (2) procurar por mutações e polimorfismos de nucleotídeos simples. Em laboratórios de análises também é utilizado como medida de abundância de sequências de DNA e RNA em amostras clínicas e industriais.

**RT-PCR (PCR com transcriptase reversa).** Uma PCR na qual a população inicial é RNA total ou mRNA poli(A)<sup>+</sup> e um passo inicial com transcriptase reversa é necessário para produzir cDNA.

**PCR touch-down.** Uma maneira de aumentar a especificidade de uma reação de PCR. A maioria dos termocicladores pode ser programada para realizar corridas nas quais a temperatura de anelamento seja diminuída de forma incrementada durante a ciclagem da PCR de um valor inicial acima do  $T_m$  esperado para um valor abaixo do  $T_m$ . Por meio da manutenção da estrigência da hibridização bastante alta no início, a formação de produtos espúrios é desfavorecida, permitindo que a sequência esperada seja predominante.

**PCR de todo o genoma.** PCR indiscriminada. Antigamente, era realizada utilizando-se *primers* amplamente degenerados (ver *PCR degenerado*) ou por ligação de *primers* de oligonucleotídeos conectores a uma população complexa de DNA, utilizando, então, *primers* de oligonucleotídeos de ligação específica para amplificar todas as sequências (*PCR com primers de conectores*). Entretanto, esses métodos não amplificam todas as sequências, pois estruturas secundárias de DNA representam um problema para a polimerase padrão utilizada em PCR, causando deslizamento ou dissociação da enzima e do molde, resultando em artefatos de amplificação não específica e compreensão incompleta de *loci*. Para compensar essas dificuldades, um procedimento de amplificação isotérmica sem ser por PCR é hoje comumente utilizado, conhecido como *amplificação de deslocamento múltiplo* (MDA). No método de MDA, uma DNA-polimerase do fago phi29 que desloca a fita é utilizada em uma amplificação de DNA em forma de círculo rolante a uma temperatura constante de cerca de 30°C (ver Leituras adicionais).

Figura 1 PCR inversa.

para aumentar a eficiência da clonagem. Vetores especializados, como pGEM-T-*easy*, podem ser tratados para possuírem extremidades 3' livres complementares em seus sítios de clonagem que estimularão o pareamento de bases com as extremidades 3' livres dos produtos de PCR a serem clonados (clonagem TA). Alternativamente, os nucleotídeos livres nos produtos de PCR são removidos com enzimas de polimerase tais como polimerase T4 ou polimerase Pfu. *Primers* de PCR também podem ser modificados por meio da criação de uma extensão de 10 nucleotídeos contendo um sítio de restrição adequado na extremidade 5'. A extensão nucleotídica não forma pares de base com o DNA-alvo durante a amplificação, mas, mais tarde, o produto amplificado pode ser digerido com a enzima de restrição apropriada a fim de gerar extremidades livres para clonagem em um vetor apropriado.

## Uma grande variedade de abordagens de PCR foi desenvolvida para aplicações específicas

As diversas aplicações da PCR e a necessidade de aperfeiçoar a eficiência e a especificidade incitaram a criação de uma grande variedade de abordagens de PCR. A seguir e no Quadro 6.2 serão consideradas algumas aplicações selecionadas.

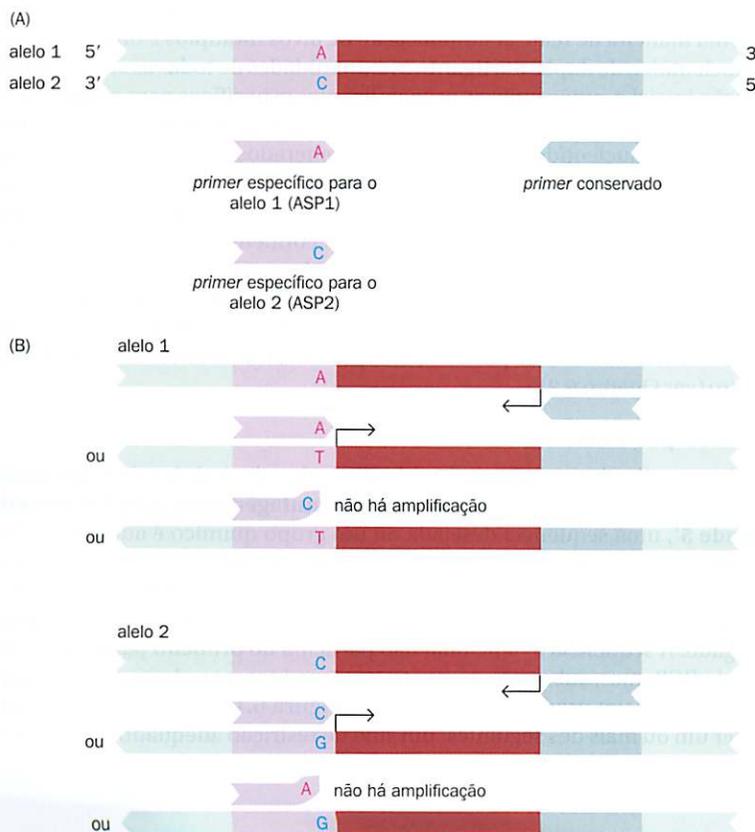
### PCR alelo-específico

Várias aplicações requerem a habilidade de distinguir entre dois alelos que se podem diferir por apenas um único nucleotídeo. Aplicações de diagnósticos incluem a habilidade de distinguir entre uma mutação de ponto causadora de doenças e o alelo normal. O PCR alelo-específico tira vantagem da dependência crucial de pareamento correto entre bases na extremidade 3' dos *primers* de ligação. No método popular ARMS (sistema de amplificação refratária de mutações), *primers* alelo-específicos são desenhados de forma que os nucleotídeos na sua extremidade 3' pareiem com os nucleotídeos variáveis que distinguem os alelos. Sob condições experimentais adequadas, a amplificação não acontecerá quando os nucleotídeos da extremidade 3' não estiverem perfeitamente pareados, distinguindo, assim, os dois alelos (Figura 6.17).

### Amplificação de alvos múltiplos e métodos de PCR de todo o genoma

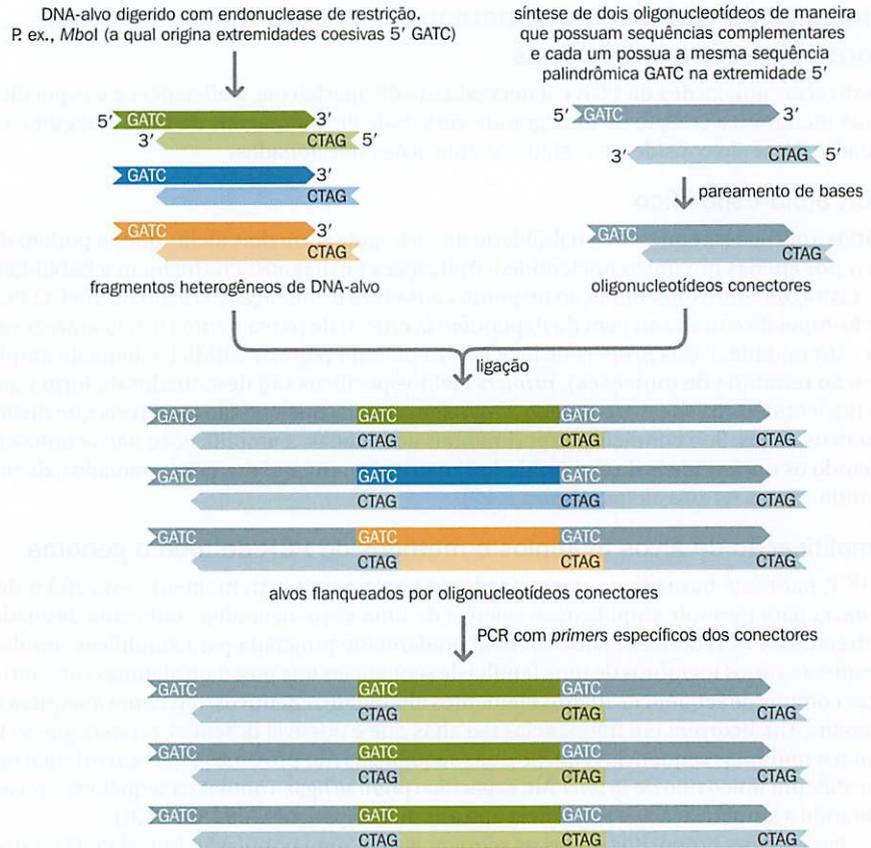
A PCR padrão é baseada na necessidade de anelamento extremamente específico dos *primers* para permitir amplificação seletiva de uma sequência-alvo conhecida desejada. Entretanto, a PCR também pode ser deliberadamente projetada para amplificar simultaneamente vários membros de uma família de sequências que possuam algumas características comuns de sequência. Alguns elementos altamente repetitivos, tais como a sequência humana Alu, ocorrem em frequências tão altas que é possível desenhar *primers* que se ligem a múltiplas sequências Alu. Se duas sequências Alu próximas estão em orientações opostas, um único tipo de *primer* Alu específico pode se ligar a ambas as sequências, possibilitando a amplificação da sequência entre as duas repetições Alu (Alu-PCR).

Para alguns propósitos, todas as sequências em uma população inicial de DNA precisam ser amplificadas. Essa abordagem pode ser utilizada para reabastecer uma fonte refinada de DNA que esteja presente em quantidades bastante limitadas (*amplificação*



**Figura 6.17** O PCR alelo-específico é dependente de pareamento perfeito entre os nucleotídeos das extremidades 3' dos *primers*. (A) Os alelos 1 e 2 diferem-se por uma única base (A→C). Dois *primers* de oligonucleotídeos alelo-específicos (ASP1 e ASP2) são desenhados para serem idênticos em sequência aos dois alelos em uma região precedendo a posição do nucleotídeo variável, mas que se diferem e terminam neste nucleotídeo variável. ASP1 e ASP2 são utilizados como *primers* alternativos em reações de PCR com outro *primer* que é desenhado para se ligar a uma região conservada na fita de DNA oposta para ambos os alelos. (B) ASP1 que se ligará perfeitamente à fita complementar da sequência do alelo 1, permitindo a amplificação com *primers* conservados. Entretanto, a extremidade 3' C-terminal do *primer* ASP2 possui falhas no pareamento com o T da sequência do alelo 1, tornando a amplificação impossível. Similarmente, ASP2, mas não ASP1, pode se ligar perfeitamente ao alelo 2 e iniciar a amplificação.

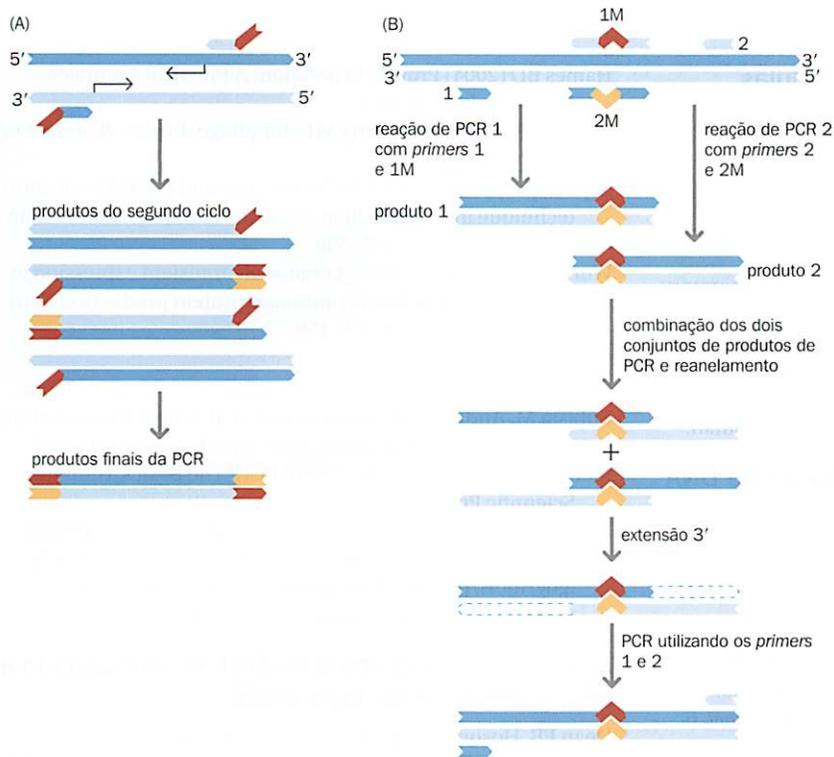
**Figura 6.18 Utilização de oligonucleotídeos conectores para amplificar múltiplos DNAs-alvo simultaneamente.** O DNA a ser amplificado é digerido com uma endonuclease de restrição que produz extremidades coesivas. O conector é preparado por síntese individual de dois oligodesoxirribonucleotídeos que são complementares em sequência, de modo que, quando passíveis de combinação, formem uma sequência de DNA dupla-fita com o mesmo tipo de extremidade coesiva que os fragmentos de restrição. A ligação dos oligonucleotídeos conectores aos fragmentos de restrição do DNA-alvo pode resultar em fragmentos sendo flanqueados por um conector de cada lado. Os *primers* específicos para os conectores podem, então, permitir a amplificação das moléculas de DNA-alvo com conectores em ambos os lados (mostrado no caso de apenas um dos diferentes fragmentos neste exemplo). Como resultado, numerosas sequências de DNA em uma população inicial podem ser amplificadas simultaneamente.



*de todo o genoma*). Como será visto em capítulo posterior, ela também tem sido utilizada para amplificação indiscriminada com o objetivo de produzir moldes ao sequenciamento de DNA. Uma maneira de realizar amplificação de alvos múltiplos é unir covalentemente um oligonucleotídeo de dupla-fita **ligante** às extremidades de todas as sequências de DNA na população inicial e, então, utilizar *primers* ligantes específicos para amplificar todas as sequências de DNA (Figura 6.18). Uma alternativa à utilização de oligonucleotídeos ligantes é o uso de oligonucleotídeos amplamente degenerados como *primers*. Isto é, quando oligonucleotídeos são sintetizados com um nucleotídeo de cada vez, todos os quatro nucleotídeos podem ser inseridos em posições específicas, em vez de um único, para originar um conjunto paralelo de diversos oligonucleotídeos diferentes. Como resultado, números muito grandes de diferentes oligonucleotídeos são sintetizados e podem se ligar a numerosas sequências diferentes no genoma, possibilitando a amplificação de grande parte do DNA genômico por PCR. Note, entretanto, que a amplificação de todo o genoma não é frequentemente realizada por PCR, mas por uma forma alternativa de clonagem de DNA *in vitro* (ver Quadro 6.2).

### Mutagênese por PCR

A PCR pode ser utilizada para construir vários tipos de substituições de bases predeterminadas, deleções e inserções em um DNA-alvo. Em **mutagêneses causadas por adição na extremidade 5'**, uma sequência desejada ou um grupo químico é adicionado de forma muito semelhante ao que seria a ligação de um oligonucleotídeo. Um *primer* mutagênico é desenhado de forma a complementar a sequência-alvo na sua extremidade 3', e a extremidade 5' a conter uma sequência nova desejada ou uma sequência com um grupo químico agregado. A sequência 5' adicional não participa no primeiro passo de anelamento da reação de PCR, mas, subsequentemente, se torna incorporada ao produto amplificado, gerando, assim, um produto recombinante (Figura 6.19A). A sequência 5' adicional pode conter um ou mais dos seguintes: um sítio de restrição adequado que possa facilitar



**Figura 6.19 Mutagenese por PCR.** (a) Mutação por adição na extremidade 5'. Os primers podem ser modificados na extremidade 5' para introduzirem uma sequência nova desejada ou um grupo químico (barras vermelhas) que não faça parte do anelamento inicial ao DNA-alvo, mas seja copiado nos ciclos subsequentes. Os produtos do segundo ciclo e dos ciclos finais são mostrados neste exemplo. (B) Mutagenese por falha no pareamento do primer. Uma mutação local específica predeterminada localizada em um segmento central pode ser introduzida por duas reações de PCR separadas 1 e 2, cada uma amplificando segmentos de DNA que se sobrepõem. Primers complementares mutantes, 1M e 2M, introduzem mau pareamento de bases deliberadamente no sítio da mutação. Após os dois produtos de PCR serem combinados, desnaturados e reanelados, cada uma das duas fitas do produto 1 pode parear suas bases com fitas complementares do produto 2 para formar heterodúplexes com extremidades 3' recuadas. A DNA-polimerase pode estender as extremidades 3' para formar produtos de comprimento completo com a mutação introduzida em um segmento central, e eles podem ser amplificados por meio da utilização dos primers externos 1 e 2, apenas.

a clonagem de DNA baseada em células; um nucleotídeo modificado contendo um grupo repórter ou marcado, tal como um nucleotídeo biotilado ou um fluoróforo; ou um fago promotor para conduzir a expressão gênica.

Em *mutação por primers imperfeitos*, o primer é desenhado para ser apenas parcialmente complementar ao sítio-alvo, mas de maneira que ele ainda se ligue especificamente ao alvo. Inevitavelmente, isso significa que a mutação é introduzida próximo ao final da extremidade do produto de PCR. Essa abordagem pode ser explorada para introduzir um sítio artificial de enzima de restrição diagnóstica que permita a procura por uma mutação conhecida. Mutações também podem ser introduzidas em qualquer ponto dentro de uma sequência escolhida utilizando primers imperfeitos. Duas reações mutagênicas são concebidas, nas quais os dois produtos de PCR separados possuem sequências parcialmente sobrepostas contendo a mutação. Os produtos desnaturados são combinados para gerar um produto maior com a mutação em uma localização mais central (Figura 6.19B).

### PCR em tempo real (qPCR)

PCR em tempo real (ou qPCR, uma abreviação de PCR quantitativo) é um método amplamente utilizado para quantificar DNA ou RNA. No último caso, o RNA é copiado com o auxílio de uma transcriptase reversa para originar fitas complementares de DNA. Reações de PCR padrão, nas quais os produtos de amplificação são analisados no final do procedimento, não são muito adequadas para quantificação. Amostras são removidas dos tubos de reação, fracionadas por tamanho em géis de agarose e detectadas por ligação com brometo de etídeo, o qual floresce sob radiação violeta. Este procedimento, além de demorado, não é muito sensível. Em PCR em tempo real (qPCR), entretanto, a quantificação é feita automaticamente, enquanto a reação está realmente ocorrendo, em um equipamento de PCR especialmente projetado. Os produtos de reação são analisados em um estágio mais inicial do processo de amplificação, quando a reação ainda está exponencial, permitindo uma quantificação mais precisa do que a do final da reação. Existem diversas aplicações, tanto para as quantificações absolutas de um ácido nucleico como para as quantificações comparativas. Uma das mais importantes é a do rastreamento da expressão gênica, e será descrita em detalhes no Capítulo 8 como qPCR funciona e é utilizado na construção de perfis de expressão gênica.

## LEITURAS ADICIONAIS

### Clonagem de DNA geral baseada em células

Primrose SB, Twyman RM, Old RW & Bertola G (2006) Principles Of Gene Manipulation And Genomics, 7th ed. Blackwell Publishing Professional.

Roberts RJ (2009) Official REBASE Homepage.  
<http://rebase.neb.com/rebase/rebase.html> [banco de dados de endonucleases de restrição.]

Sambrook J & Russell DW (2001) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Banco de dados sobre vetores: para uma compilação de sites da Web, ver [http://mybio.wikia.com/wiki/Vector\\_databases](http://mybio.wikia.com/wiki/Vector_databases)

Watson JD, Caudy AA, Myers RM & Witkowski JA (2007) Recombinant DNA: Genes And Genomes, 3rd ed. Freeman.

### Clonagem de insertos grandes e produção de DNA recombinante fita simples

Iouannou PA, Amemiya CT, Garnes J et al. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.* 6, 84–89.

Mead DA & Kemper B (1988) Chimeric single-stranded DNA phage-plasmid cloning vectors. *Biotechnology* 10, 85–102.

Schlessinger D (1990) Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet.* 6, 248–258.

Shizuya H, Birren B, Kim U-J et al. (1992) Cloning and stable maintenance of 300 kilobase-pair fragments of human DNA in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 89, 8794–8797.

### Clonagem de expressão gênica e marcadores selecionáveis em células de mamíferos

Colosimo A, Goncz KK, Holmes AR et al. (2000) Transfer and expression of foreign genes in mammalian cells. *Biotechniques* 29, 314–331.

Hames BD (2004) Protein Expression: A Practical Approach. Oxford University Press.

Sidhu SS (2001) Engineering M13 for phage display. *Biomol. Eng.* 18, 57–63.

Szybalski W (1992) Use of the HPRT gene and the HAT selection technique in DNA-mediated transformation of mammalian cells. *BioEssays* 14, 495–500.

Wurm F & Bernard A (1999) Large-scale transient expression in mammalian cells for recombinant protein production. *Curr. Opin. Biotechnol.* 10, 156–159.

### PCR

Kubista M, Andrade JM, Bengtsson M et al. (2006) The real-time polymerase chain reaction. *Mol. Aspects Med.* 27, 95–125.

McPherson MJ & Moller SG (2006) PCR: The Basics. Garland Scientific Press.

National Center for Biotechnology Information (2008) Primer-BLAST: primer designing tool. <http://www.ncbi.nlm.nih.gov/tools/primer-blast/> [para identificar iniciadores específicos para uma sequência definida.]

### Amplificação isotérmica de DNA *in vitro* usando a DNA-polimerase do fago phi29

Dean FB, Hosono S, Fang L et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl Acad. Sci. USA* 99, 5261–5266.

### Mutagênese de DNA

Ling MM & Robinson BH (1997) Approaches to DNA mutagenesis: an overview. *Anal. Biochem.* 254, 157–178.