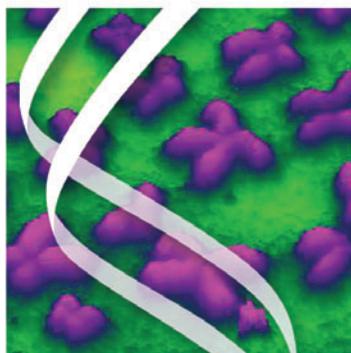


# FUNDAMENTOS DO DNA

Esta seção fornece uma visão geral da genética e das ferramentas utilizadas no DNA recombinante e em análises genômicas. Começamos com Mendel, cujos estudos através de cruzamentos de plantas resultaram nos fundamentos de uma das mais importantes áreas de investigação científica do século XX e, indiscutivelmente, a mais importante do início do século XXI. Os cromossomos foram rapidamente considerados a base física do gene mendeliano, mas a natureza química do gene não tinha sido determinada até 1953. A descoberta do DNA de fita dupla marca o início da idade dourada da genética molecular; foi neste período que o mecanismo de tradução da informação do DNA em proteínas e a regulação da expressão gênica foram decifrados. No final da década de 1960, tinha-se a impressão de que a genética molecular havia se tornado universal e que poucas novas descobertas viriam à tona. Esta visão, entretanto, foi bastante pessimista. No início da década de 1970, uma série de técnicas coletivamente conhecidas como “DNA recombinante” foi desenvolvida, abrindo assim novos campos de investigação. Essas técnicas foram surpreendentes e a partir delas, inesperadamente, detalhes da estrutura molecular, dos arranjos de genes e ainda de fenômenos previamente estudados, como os genes móveis, foram revelados. Atualmente, novas e extraordinárias descobertas têm sido feitas na genética molecular.





## CAPÍTULO 1

---

# O DNA É o Material Genético Primário

**N**ão existe nenhuma substância tão importante quanto o DNA. Por carregar em sua estrutura a informação hereditária que determina as estruturas de proteínas, é a principal molécula da vida. As regras que direcionam as células ao crescimento e à divisão são codificadas por ele; a diferenciação dos ovos fertilizados em uma gama de células especializadas de plantas e animais superiores é determinada por mensagens especializadas emanadas dele. A forma molecular do DNA permite um número infinito de variações estruturais refletido nas variações da informação hereditária que ele transmite. Ele é a base do processo evolutivo que gerou milhões de diferentes formas de vida que têm ocupado a Terra desde que os organismos vivos começaram a surgir 3 a 4 bilhões de anos atrás.

A extraordinária capacidade de moléculas de DNA alteradas darem origem a novas formas de vida mais bem adaptadas em relação aos seus progenitores tornou possível o aparecimento de nossa própria espécie com habilidade de perceber a natureza de nosso ambiente e utilizar esta informação para construir as civilizações humanas. Como resultado de nossa habilidade para um rápido pensamento conceitual, temos perguntado durante vários séculos sobre a natureza dos objetos inanimados como a água, as pedras e o ar, bem como sobre as estrelas do espaço. E a biologia, a ciência dos objetos vivos conhecida há 50 anos como uma ciência inferior, tem se desenvolvido rapidamente. Hoje existe um consenso das mentes mais bem informadas de que a essência da vida pode ser explicada pelas mesmas leis da física e da química, as quais têm nos ajudado a entender, por exemplo, por que a maçã cai para a terra e por que a lua não, ou por que a água é transformada em vapor quando alcança temperaturas superiores ao seu ponto de ebulição.

Do DNA partem todos os comandos que regulam a natureza e o número de praticamente todas as moléculas celulares. Trabalhando na estrutura exata da multiplicidade de manuscritos genéticos codificados nas moléculas de DNA, atualmente estamos dando passos gigantescos para o entendimento da série de complexas reações químicas que se interconectam e que levam os ovos fertilizados a se desenvolverem em organismos multicelulares altamente complexos. Só em meados da década de 1980 foi possível determinar toda a informação genética dos organismos multicelulares, isto é, determinar a seqüência de DNA dos seus genomas. Inicialmente houve muita resistência à execução de um projeto tão grande como o Projeto Genoma Humano, existindo dúvidas sobre sua exequibilidade técnica. O progresso subsequente, porém, foi extraordinariamente rápido. Hoje conhecemos a escrita genética completa de centenas de espécies, incluindo a nossa própria, e veremos ao longo deste livro como estes novos conhecimentos têm transformado nossa visão sobre o mundo vivo e como estamos inseridos nele.

O papel essencial do DNA na genética foi estabelecido somente em meados do século XX, mas nossa fascinação pela hereditariedade remonta há vários séculos. Neste capítulo estudaremos como os padrões de hereditariedade foram determinados antes mesmo de se conhecer suas bases físicas e faremos em seguida uma revisão do DNA como molécula da vida.

## As Experiências de Cruzamento de Mendel com Ervilhas Revelaram os Padrões da Hereditariedade

Para se conhecer muitas das características humanas, como a cor dos olhos e o formato dos rostos, que nos são transferidas pelos nossos pais, é preciso retornar aos primeiros seres humanos. Nos primórdios, os seres humanos domesticaram animais ou cultivaram plantas e sempre procuraram desenvolver algumas características especiais em seus cães e suas cabras por meio de cruzamentos seletivos. Entretanto, apesar do grande sucesso dos cruzamentos através dos séculos – todos realizados de forma empírica – não se conseguia prever, com certeza, os padrões de hereditariedade. A definição desses padrões deve ser atribuída ao esforço de um cientista que, trabalhando sozinho em Brno (agora República Tcheca), elaborou os fundamentos para o entendimento de como os organismos herdavam suas características.

Tal cientista era um monge, Gregor Mendel, que, em 1856, começou suas experiências de cruzamento pelo acasalamento de organismos progenitores com características específicas e pela observação de como essas características eram transmitidas aos seus descendentes. Apesar de outros cien-

tistas terem feito experiências semelhantes, o grande sucesso de Mendel se deveu à escolha do organismo de estudo (um tema recorrente através da história da genética), seu delineamento experimental, seu registro cuidadoso e sua análise numérica dos resultados.

Mendel utilizou a ervilha, *Pisum sativum*, que é fácil de cultivar e tem características bem distintas, ou caracteres (termo de Mendel), que são fáceis de observar. Esses caracteres, presentes em pares contrastantes, incluem sementes que são amarelas ou verdes, lisas ou rugosas; plantas que são baixas ou altas (haste curta ou longa), e as pétalas das flores, que podem ser de cor rosa ou branca. Mendel cruzou linhagens com determinados caracteres por muitas gerações até que elas não mais mostrassem variações naquele caractere – estas “plantas puras” resultantes de vários cruzamentos e seus descendentes foram, por exemplo, ou sempre altos ou sempre baixos. Ele então cruzou linhagens com essas características contrastantes e examinou seus descendentes (*geração F<sub>1</sub>*), contando o número de plantas que mostrou um ou outro caractere. No caso da altura da planta, Mendel notou que toda a geração resultante do cruzamento de uma planta alta com uma baixa era alta; nenhuma foi de estatura mediana. A maior altura foi então considerada *dominante* em relação à baixa estatura, que foi considerada *recessiva*.

Mendel então cruzou a geração *F<sub>1</sub>* para produzir uma *geração F<sub>2</sub>*. De forma marcante, diferente do resultado obtido com o cruzamento das linhagens altas “puras”, a geração *F<sub>2</sub>* originada do cruzamento das plantas altas de *F<sub>1</sub>* era constituída de plantas altas e baixas: na média, havia três plantas altas para cada planta baixa na geração *F<sub>2</sub>*. Esta relação de 3:1 na geração *F<sub>2</sub>* foi mantida para outros pares de caracteres dominantes e recessivos.

Finalmente ele analisou as conseqüências do cruzamento cruzado de linhagens de plantas que tinham *dois* caracteres contrastantes: por exemplo, plantas altas com sementes amarelas foram cruzadas com plantas baixas com sementes verdes. Todos os descendentes da geração *F<sub>1</sub>* foram plantas altas com sementes amarelas (alta dominante em relação a baixa, e amarela dominante em relação a verde). Na segunda geração, todas as quatro possíveis combinações apareceram – plantas altas, sementes amarelas; plantas altas, sementes verdes; plantas baixas, sementes amarelas; plantas baixas, sementes verdes. Parecia que diferentes caracteres eram transmitidos de forma independente (segregação independente). Além disso, esta série de quatro caracteres sempre apareceu, na média, na proporção de 9:3:3:1.

## Os Fatores Discretos de Hereditariedade Obedecem às Leis Mendelianas

A partir destes experimentos simples descritos anteriormente, Mendel deduziu que caracteres como a cor das sementes

e a altura da planta ocorrem como fatores discretos (agora chamados de genes). Esses fatores não surgem espontaneamente nem se misturam entre si, pois no cruzamento apareceram somente sementes que eram verdes ou amarelas e plantas que eram altas ou baixas. A relação observada por Mendel pode ser computada assumindo-se que cada linhagem de ervilha possui dois fatores para cada caractere. Durante a formação do gameta (células sexuais), os dois fatores responsáveis por determinado caractere se separam de forma que cada espermatozóide ou óvulo recebe somente um dos fatores. Cada gameta tem igual chance de receber um ou outro fator para determinado caractere. No processo de fertilização, um descendente recebe um fator do progenitor macho e um do progenitor fêmea.

A característica mostrada pela planta depende dos pares de fatores que ela recebe de seus progenitores. Se ambos os fatores são os mesmos, isto é, dois dominantes (semente amarela) ou dois recessivos (semente verde), então a planta exibirá aquelas características – sementes amarelas para o primeiro e sementes verdes para o segundo. Se um dos fatores é dominante e o outro recessivo, o caractere dominante é expresso; neste caso, a planta com um fator para semente amarela e um para semente verde mostrará o traço dominante, ou seja, semente amarela. Mas o fator para a semente verde persiste mesmo quando a planta é amarela, podendo reaparecer nas gerações subsequentes. Esta foi uma observação primordial, pois no século XIX se acreditava que as características dos progenitores estavam combinadas na progênie e, uma vez misturadas, nunca mais seriam recuperadas. Nesta época, esta teoria causava um sério problema para a teoria da evolução pela seleção natural anunciada por Charles Darwin. Sua crítica era de que qualquer característica favorável que aparecia em um indivíduo não teria um efeito de longa duração, pois poderia ser imediatamente perdida por meio de um cruzamento.

As contribuições de Mendel, publicadas em 1865, foram efetivamente ignoradas até 1900, quando os resultados de sua pesquisa foram independentemente confirmados por renomados cientistas da Europa que realizavam cruzamentos, de plantas como Carl Correns, Hugo de Vries e Erich von Tschermak-Seysenegg. Na Inglaterra, William Bateson tornou-se um entusiasta de Mendel e em 1906 deu o nome de *genética* (do latim *genetikos*, que significa “o produtor”) para esta nova área de conhecimento. Em 1909, um cientista dinamarquês, Wilhelm Johannsen, denominou os fatores mendelianos de *genes* (do latim *genos*, origem). Cada gene pode existir em diferentes formas ou *alelos* (originalmente alelomorfos, outra contribuição de Bateson). Por exemplo, existem alelos verdes ou amarelos para a cor das sementes da ervilha, e quando os dois alelos do gene são iguais, o indivíduo é considerado *homozigoto* para aquele gene; quando os ale-

los são diferentes, o indivíduo é considerado *heterozigoto*. Em 1911, Johannsen também mostrou a diferença entre fenótipo e genótipo. O *genótipo* é a composição genética de um organismo, enquanto o *fenótipo* é a expressão física do genótipo.

## Os Cromossomos São os Portadores Celulares da Hereditariedade

Grandes avanços na microscopia na última metade do século XIX permitiram que as estruturas das células fossem visualizadas com maiores detalhes. Células de plantas e animais mostraram ter um corpo central, o núcleo, circundado por um citoplasma relativamente amorfo. Tanto o núcleo como toda a célula estão envolvidos por membranas e, em plantas, visualizou-se ainda uma parede espessa recobrendo a membrana celular. O citoplasma contém muitas formas de inclusão – como vacúolos e gotas de lipídeos – assim como grânulos na forma de bastão, hoje conhecidos como mitocôndria e cloroplastos. Mas foi o núcleo que atraiu a maior atenção por causa de seu desaparecimento no início da divisão e do concomitante aparecimento de corpos alongados, denominados *cromossomos* (eles podiam ser corados com corantes especiais; *chromo*, do grego, significa “cor”).

Como os embriões são formados pela fusão dos óvulos com o espermatozóide, tornou-se evidente que a hereditariedade é transmitida por estas células especiais. Como o espermatozóide tem pouco material citoplasmático quando comparado com a quantidade no núcleo, a hipótese óbvia foi de que a função do núcleo era de carregar os determinantes hereditários da célula.

Em 1902, observou-se a grande correlação entre o comportamento dos cromossomos e o que era esperado dos fatores de Mendel. Naquele ano, o americano Walter Sutton descreveu a estrutura e o comportamento dos cromossomos do gafanhoto *Brachystola magna* e assinalou que os cromossomos “podem se constituir na base física das leis da hereditariedade de Mendel”. Em um determinado tipo de célula, uma variedade de diferentes cromossomos foi observada, sendo que cada tipo de cromossomo estava normalmente presente em duas cópias (o par de *cromossomos homólogos*). Observou-se que o número de cromossomos apresentava-se duplicado durante a divisão celular (*mitose*), produzindo duas células-filhas, cada uma recebendo um conjunto completo de cromossomos da célula parental. O número de cromossomos das células sexuais do espermatozóide e do óvulo é exatamente a metade daquele observado nas células somáticas. Durante a formação das células germinativas (*meiose*), o número de cromossomos das células diplóides,  $2n$ , é reduzido pela metade nas células *haplóides*,  $n$  (Figura 1-1). O processo de fertilização entre o espermatozóide e o óvulo

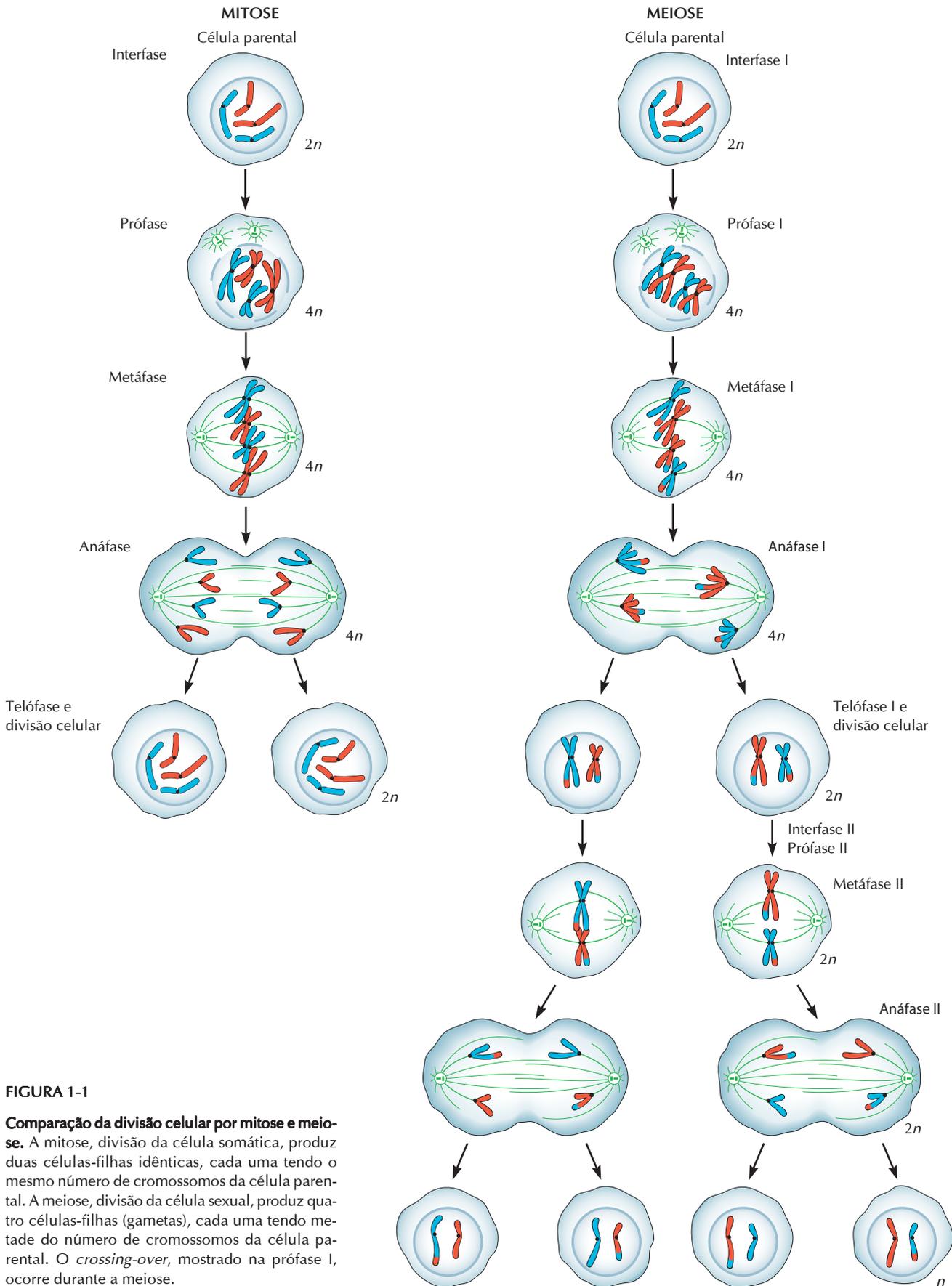


FIGURA 1-1

**Comparação da divisão celular por mitose e meiose.** A mitose, divisão da célula somática, produz duas células-filhas idênticas, cada uma tendo o mesmo número de cromossomos da célula parental. A meiose, divisão da célula sexual, produz quatro células-filhas (gametas), cada uma tendo metade do número de cromossomos da célula parental. O *crossing-over*, mostrado na prófase I, ocorre durante a meiose.

restaura o  $2n$  característico das células diplóides, com um cromossomo de cada par originado do progenitor macho e o outro do progenitor fêmea.

### Mapeamento dos Genes nos Cromossomos pela Técnica de Ligação

Se os cromossomos são os portadores dos genes, seria então possível associar a herança de um fenótipo específico com os cromossomos. A primeira característica a ser atribuída aos cromossomos foi a sexual. O trabalho foi feito em 1905 na Universidade de Columbia, na cidade de Nova Iorque, onde Nettie Stevens e Edmund Wilson descobriram a existência dos tão conhecidos *cromossomos sexuais*. Um cromossomo, o X, está presente em duas cópias (XX) nas células somáticas femininas, mas em uma única cópia nas células somáticas masculinas, que também carregam o cromossomo Y, morfologicamente distinto (XY) (Figura 1-2). Durante a divisão do número de cromossomos na meiose, todos os óvulos recebem necessariamente um único cromossomo X, enquanto os espermatozoides podem receber ou o cromossomo X ou o Y. A fertilização por um espermatozoide contendo um cromossomo X vai gerar uma fêmea (XX), enquanto a fertilização por um espermatozoide contendo um cromossomo Y dará origem a um macho (XY). O fato de as características sexuais se localizarem em um simples par de cromossomos explica por que os descendentes machos e fêmeas são pro-

duzidos na relação de 1:1; metade dos fetos recebe um cromossomo X de cada progenitor – originando uma fêmea – e metade recebe o cromossomo X de sua progenitora e um cromossomo Y do progenitor – originando um macho.

Um segundo exemplo de associação das características com os cromossomos foi dado pelo laboratório de Thomas Hunt Morgan, também na Universidade de Columbia. Em 1910, enquanto realizava seus estudos experimentais de evolução usando a mosca-das-frutas *Drosophila* de olhos vermelhos, ele observou um único macho de olhos brancos. Esta variante mostrou ser uma característica tipicamente mendeliana quando os machos de olhos brancos foram cruzados com as fêmeas de olhos vermelhos – toda a progênie tinha olhos vermelhos, concluindo-se que olhos brancos eram recessivos – mas Morgan notou que todas as moscas de olhos brancos da geração  $F_2$  eram machos (Figura 1-3). A cor branca dos olhos não era uma característica limitada aos machos, pois Morgan poderia produzir fêmeas de olhos brancos cruzando machos de olhos brancos com fêmeas de olhos vermelhos da geração  $F_2$ . A simples explicação para estes dados era de que o gene para a cor dos olhos, assim como o que determina o sexo, precisa se localizar no cromossomo X. Nos anos que se seguiram, outras mutações (ou “sports”, como foram então denominadas\*) foram ob-

\*N. de T. “Sports” na língua inglesa pode significar algo incomum ou fora da normalidade.

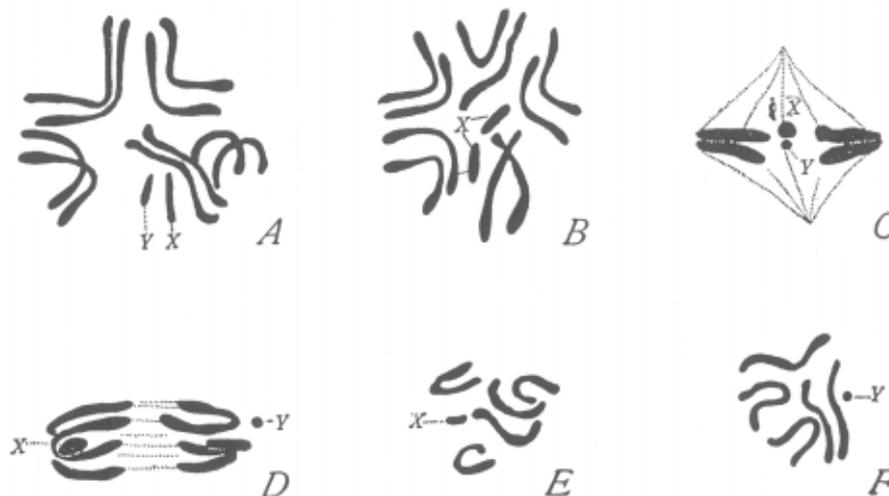


FIGURA 1-2

Fac-símile da figura do trabalho publicado por Nettie Stevens em 1905 mostrando os cromossomos sexuais da mosca-varejeira, *Calliphora vomitoria*. A e B mostram os cromossomos das células somáticas, respectivamente, de machos e fêmeas. Existem cinco pares de cromossomos autossômicos e um par de cromossomos sexuais. O macho tem o cromossomo X mais longo e o cromossomo Y mais curto e a fêmea tem dois cromossomos X. (C) Durante a meiose, os cromossomos X e Y se unem mostrando que são homólogos e (D) se separam de forma que os espermatozoides contêm um número haplóide de autossomos, ou o cromossomo X (E) ou o Y (F).

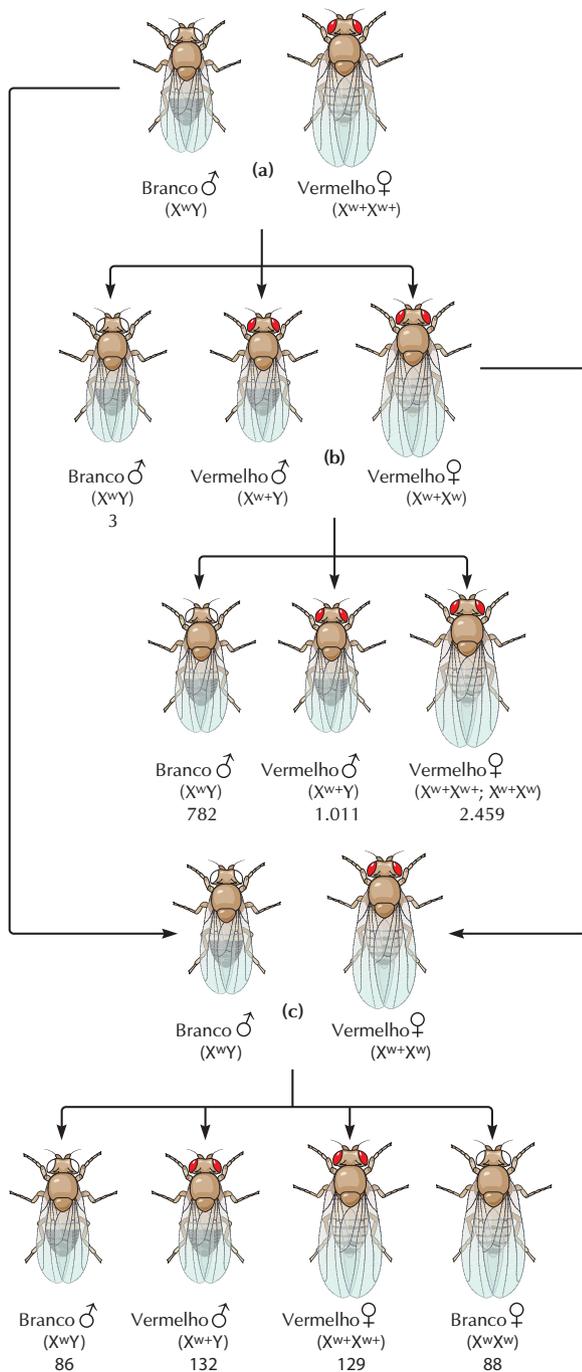


FIGURA 1-3

**Experimento de Morgan com a *Drosophila* macho de olhos brancos.**

A *Drosophila* macho é menor que a fêmea, e a cor dos olhos do tipo selvagem de ambos os sexos é vermelha. O simbolismo é moderno e não aquele usado por Morgan. X e Y representam, respectivamente, os cromossomos X e Y, e os índices superiores  $w+$  e  $w$  representam a cor dos olhos, respectivamente, vermelha e branca. Os números de cada tipo de mosca são aqueles do trabalho original de Morgan da *Science* de 1910. (a) No cruzamento inicial, Morgan descobriu que todas as três gerações originadas do macho de olhos brancos e da fêmea de olhos vermelhos tinham olhos vermelhos, de um total de 1.237 machos e fêmeas. Isso indicava que o branco era recessivo em relação ao vermelho. Existiam três machos de olhos brancos que foram considerados mutantes, assim como o macho de olhos brancos original. (b) Os machos e as fêmeas de olhos vermelhos originados do primeiro cruzamento foram então cruzados um com o outro. Moscas de olhos brancos reapareceram nesta geração, mas eram todos machos. Isso sugeriu que a cor dos olhos estava de alguma forma associada com a masculinidade. (c) Entretanto, no terceiro cruzamento, usando o macho de olhos brancos original e as descendentes fêmeas de olhos vermelhos, apareceram filhas de olhos brancos, mostrando que a associação não estava limitada à masculinidade. Todos estes resultados e aqueles de outros cruzamentos podem ser explicados considerando que o gene para a cor dos olhos é carregado pelo cromossomo X e que as fêmeas são homocigotas (XX) enquanto os machos são heterocigotos para os cromossomos sexuais (XY). Com somente um cromossomo X, os machos mostram um mutante fenotípico mesmo para uma característica como os olhos brancos.

servadas em muitos outros genes de *Drosophila*. Durante outros estudos, muitos exemplos mostraram uma tendência de duas características serem sempre herdadas juntas, ou *ligadas*. Após o estudo de um grande número de variantes, ficou claro que formavam quatro *grupos ligados*, correspondendo a um dos quatro cromossomos da *Drosophila*; isto é, os genes ligados estavam no mesmo cromossomo. Isto foi um marco histórico e um poderoso argumento para a teoria cromossomal da hereditariedade.

Contudo, a ligação entre duas características algumas vezes foi quebrada em consequência da troca física entre pares de cromossomos (denominada *crossing-over*) que ocorre quando cromossomos homólogos ficam pareados durante a meiose (Figura 1-4). Em 1913, Morgan sugeriu que, quanto mais longe dois genes estiverem no mesmo cromossomo, maior a chance de a troca ocorrer e, portanto, haver uma recombinação de genes. Alfred Sturtevant, na época um estudante de graduação no laboratório de Morgan, descobriu que podia estimar quan-

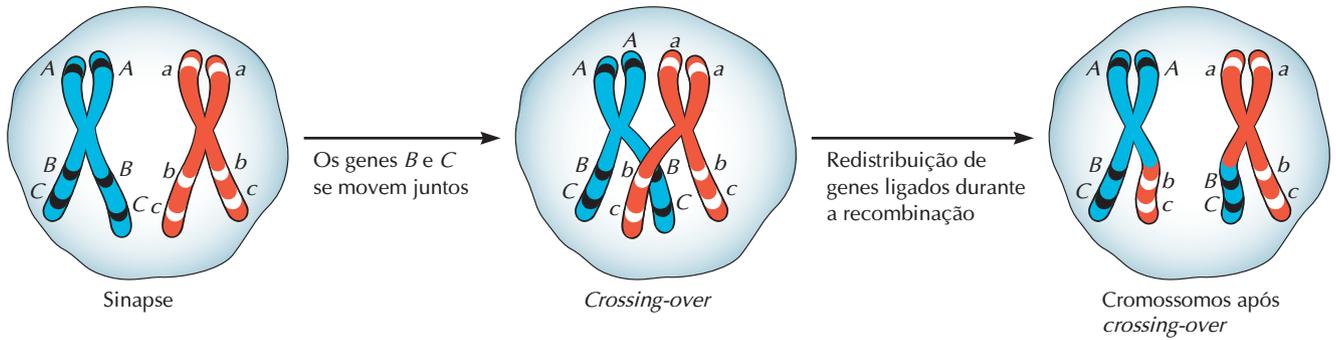


FIGURA 1-4

**Redistribuição dos genes por *crossing-over*.** *B* e *C* (e *b* e *c*) estão ligados porque se movem sempre juntos durante a recombinação. O gene *A* na outra extremidade do cromossomo parece estar fracamente ligado ao *B* pelo fato de o cruzamento ter ocorrido em algum lugar entre *A* e *B*. Duas novas recombinações foram produzidas em decorrência desta recombinação: *Abc* e *aBC*. Em geral, quanto mais afastados estiverem os genes, mais fracamente estarão ligados e maior será a chance de *crossing-over*.

titativamente a distância entre um par de genes calculando a frequência de *crossing-over* entre os dois pares de genes. Sturtevant foi para casa aquela noite e retornou no dia seguinte pela manhã com o primeiro mapa genético – o posicionamento de cinco genes no cromossomo X da *Drosophila* (Figura 1-5).

Um exemplo simples de mapeamento mostra a determinação da distância entre a localização (*loci*) de dois genes, *pr*, para olhos de cor roxa, e *vg*, para asas vestigiais (Figura 1-6). Dos 2.839 descendentes analisados, um total de 305, ou seja, 10,7% mostraram recombinações. Sturtevant definiu a unidade de mapa como a distância

(a)

Cruzamento entre	Recombinantes	Total de meioses	Frequência de recombinação
<i>y w</i>	214	21.736	0,010
<i>w v</i>	491	1.584	0,297
<i>v m</i>	17	573	0,030
<i>v r</i>	109	405	0,269

*y* = amarelo   *w* = branco   *v* = rubro   *m* = miniatura   *r* = rudimentar

(b)

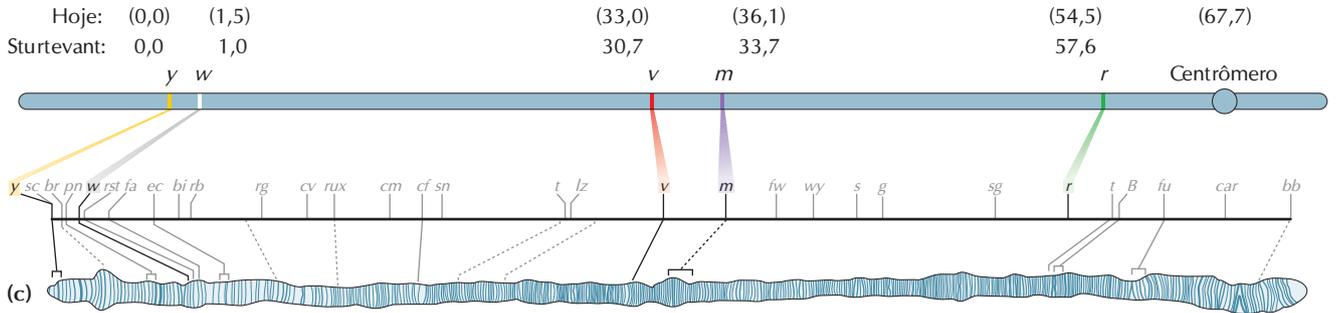


FIGURA 1-5

**O primeiro mapa genético produzido por Sturtevant em 1913.** (a) Seleção dos resultados de Sturtevant do cruzamento de moscas e determinação da frequência de recombinação entre pares de genes. (b) As posições relativas dos genes mapeados por Sturtevant são mostradas na barra representando o cromossomo 1 da *Drosophila*. 1. Os valores de Sturtevant são muito próximos dos valores mais modernos mostrados entre parênteses acima da barra. (c) As posições dos mesmos genes são indicadas no desenho do cromossomo 1 obtido da glândula salivar. Estes cromossomos são muito grandes e mostram notáveis detalhes da organização cromossômica. O mapa foi feito em 1935 por Calvin Bridges, outro protegido de Morgan e colega de Sturtevant. (Os desenhos de Bridges são tão precisos que continuam sendo uma representação definitiva dos cromossomos politênicos da *Drosophila melanogaster*.)

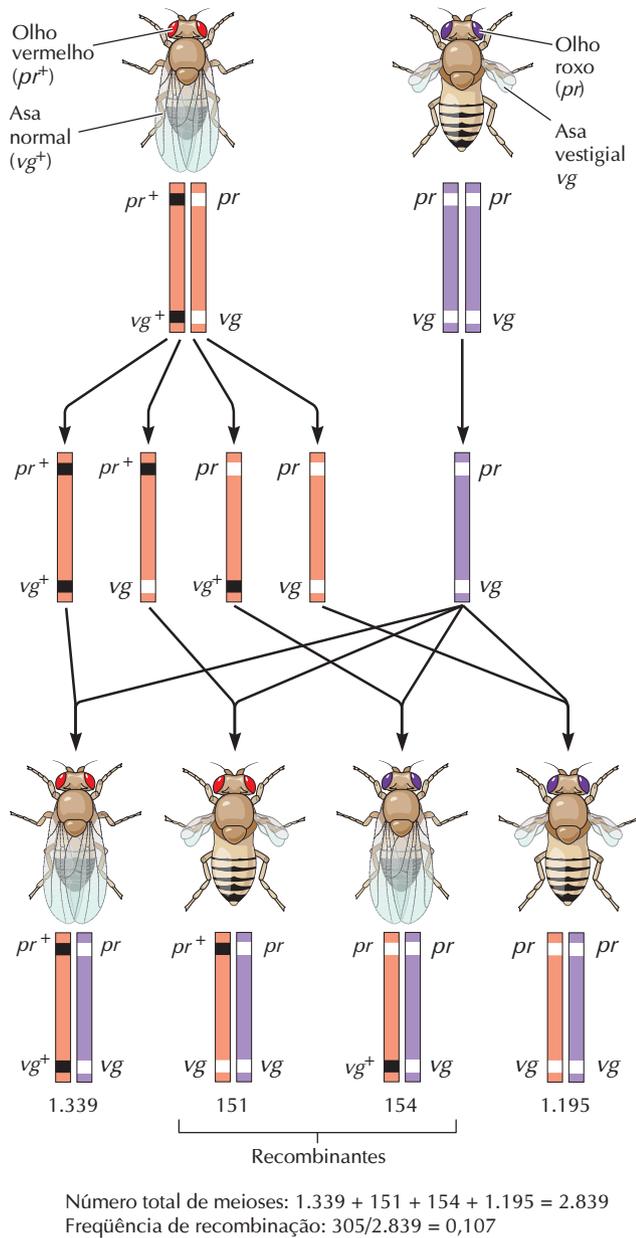


FIGURA 1-6

**Determinando distâncias genéticas utilizando recombinação.** O locus *pr* controla a cor do olho: *pr+* é a cor vermelha dos olhos do tipo selvagem, enquanto *pr* é a cor dos olhos roxa, recessiva. O locus *vg* afeta o desenvolvimento da asa: *vg+* produz asas normais, enquanto *vg* é para asa vestigial. Devido ao fato de *pr+* e *vg+* serem dominantes, moscas com olhos roxos e asas vestigiais são *prvg/prvg*. Sturtevant cruzou estas moscas com moscas de olhos vermelhos e asas normais, mas que eram heterozigotas para estas características, isto é, *pr+vg+/prvg*. São possíveis quatro gametas a partir de moscas heterozigotas: *pr+vg+* e *prvg* e, se sofreram *crossing-over* durante a meiose, *pr+vg* e *prvg+*. Somente um gameta, *prvg*, é possível a partir de moscas duplamente recessivas. Sturtevant contou o número de moscas e descobriu que, das 2.839 moscas, existiam 305 recombinantes: 151 moscas com olhos vermelhos, mas asas vestigiais (*pr+vg*) e 154 com olhos roxos e asas normais (*prvg+*). A frequência de recombinação *pr* e *vg* é  $305/2.389 = 0,107$ . A frequência de recombinação de 0,01 é 1 unidade de mapa e, então, *pr* e *vg* estão distantes 10,7 unidades de mapa.

entre um par de genes, a qual está relacionada à frequência de recombinação entre os genes em cada 100 meioses (esta unidade tem sido desde então denominada centimorgan – cM – em homenagem a Morgan). Neste caso, os loci de *pr* e *vg* estão a 10,7 cM de distância. A análise se torna mais complicada quando são considerados mais loci, mas o princípio continua o mesmo – determina-se o número de *crossing-over* entre os loci e calcula-se a frequência de recombinação. A análise de ligação tem sido mantida como a estratégia fundamental para o mapeamento de genes em todos os organismos que se reproduzem sexualmente, desde as moscas-das-frutas até os seres humanos. Entretanto, é importante lembrar que as frequências de recombinação medem as distâncias genéticas e que podem diferir das distâncias físicas entre os genes. Isto é particularmente verdadeiro para os genes que estão separados por grandes distâncias onde ocorrem múltiplas trocas por *crossing-over*.

Nestes casos, um número de *crossing-over* não será detectado porque os genes serão mantidos ligados, levando assim a uma estimativa mais baixa da verdadeira frequência de recombinação.

### Nem Todos os Genes São Encontrados nos Cromossomos do Núcleo

Na primeira metade do século XX, Morgan, Sturtevant e outros geneticistas concentraram-se na análise das características carregadas pelos cromossomos do núcleo, em parte porque essas características poderiam ser estudadas à luz da herança mendeliana e também porque as mudanças dos genes nucleares produziam diferenças significativas no fenótipo. Contudo, na mesma época, alguns geneticistas estavam estudando o papel da herança citoplasmática, isto é, os padrões da herança que não eram consistentes com a herança nuclear mendeliana. Durante muitos anos, a importância da herança citoplasmática foi matéria de muita controvérsia.

Um exemplo bem precoce, já descrito em 1909, é a herança materna, mostrada por plantas variegadas como a *Mirabilis jalapa* (também conhecida como “four o’clock plant” ou marvel do Peru). Uma única planta apresenta diferentes partes com colorações brancas ou verdes ou com manchas brancas e verdes misturadas, que refletem a presença dos cloroplastos sintetizadores de clorofila (verdes) ou não-sintetizadores (brancas). Flores são observadas em todas as três partes das plantas, que podem então ser cruzadas por transferência do pólen. As análises mostraram que a fonte do pólen não tem nenhum efeito na cor da progênie, que depende somente da flor que fornece o óvulo. Para explicar esta herança não-mendeliana,

foi postulado que os cloroplastos carregavam seu próprio material genético e que a linhagem pura com cloroplastos contendo clorofila dará origem a plantas com cloroplastos verdes, ao passo que os mutantes, contendo cloroplastos sem clorofila, darão origem a plantas com cloroplastos sem clorofila. Os zigotos fertilizados recebem quase todos os seus cloroplastos do citoplasma do óvulo; os espermatozoides, devido à escassez de citoplasma, praticamente não contribuem com nada. São, portanto, os cloroplastos maternos que reproduzem e povoam as plantas em crescimento (Figura 1-7).

De forma semelhante, as mitocôndrias também têm seus próprios genes, e seus padrões de hereditariedade têm sido extensivamente analisados em fungos como a

*Neurospora* e em leveduras como o *Saccharomyces cerevisiae*. Assim como os cloroplastos, as mitocôndrias são predominantemente herdadas do citoplasma carregado com o óvulo. Mitocôndria e cloroplastos têm genes porque são remanescentes de células semelhantes a bactérias que foram assimilados pelas células eucarióticas.

### Mutações São Alterações nos Genes

Os primeiros experimentos em cruzamento foram realizados em variantes que, assim como o macho de olhos brancos de Morgan, apareceram espontaneamente em uma baixa frequência. Sabemos agora que estes *mutantes* espontâ-

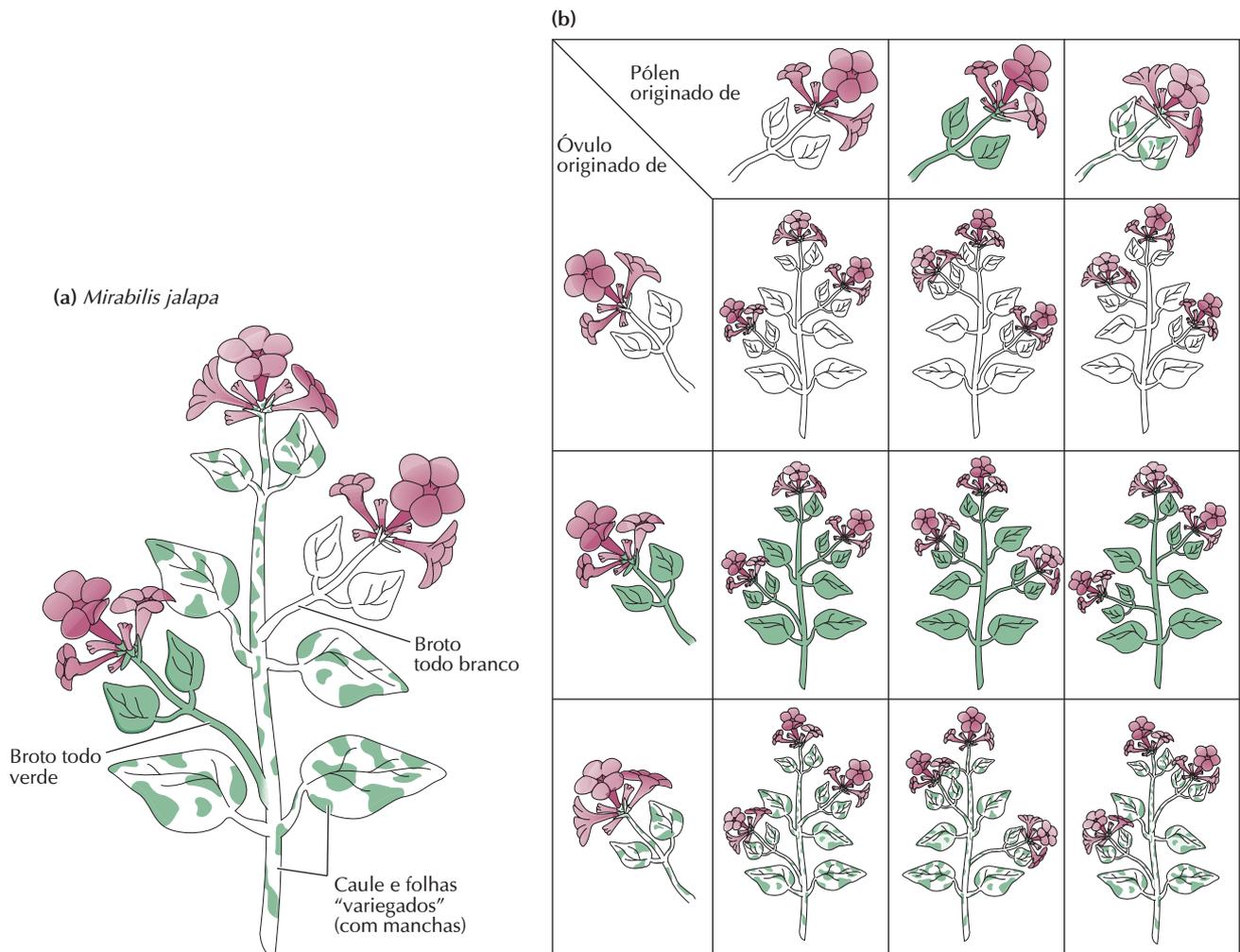


FIGURA 1-7

**Genes extracelulares.** (a) A *Mirabilis jalapa* (planta quatro horas) é uma planta variegada com áreas verdes e outras brancas, caules ramificados completamente verdes, completamente brancos ou variegados. (b) Polens obtidos de flores com estes caules podem ser usados para polinizar flores com caules com pigmentação igual ou diferente. Em todos os casos, a pigmentação das plantas  $F_1$  é dada por aquela que carrega a flor que foi polinizada; a cor do caule que fornece o pólen parece irrelevante. Assim, plantas derivadas de óvulos nascidos de um caule verde serão sempre verdes, independentemente da fonte do pólen. Isto contrasta com a hereditariedade mendeliana, na qual a transmissão de uma característica não é influenciada pela origem parental do gameta (exceto no fenômeno denominado *imprinting*; ver Capítulo 8).

neos resultam da mudança em um simples gene e são geneticamente tão estáveis quanto a versão normal do gene, denominado *tipo selvagem*. Mutações são ferramentas essenciais para a análise genética – a descoberta de uma *Drosophila* de olhos brancos nos revelou que precisa haver um gene que controle a cor dos olhos, e assim fenótipos mutantes atuam como bandeiras sinalizadoras para os genes.

Entretanto, é difícil estudar as mutações, pois elas ocorrem em uma frequência muito baixa – dezenas de milhares de indivíduos têm de ser analisados para que seja descoberto um único mutante. Foi com alívio e certa descrença que geneticistas receberam a observação de Herman Muller, de que os raios X induziram mutações com alta frequência em *Drosophila*; Muller produziu 100 mutantes de *Drosophila* induzidos por raio X em dois meses, em comparação com 200 mutantes naturais de moscas-das-frutas produzidos nos 16 anos anteriores. A descoberta de Muller teve profundas implicações para a saúde do ser humano e levou a um rápido alerta quanto aos perigos que fontes de radiação ionizantes poderiam acarretar se fossem manipuladas sem nenhuma proteção.

Mais tarde, observou-se que as mutações podem ser produzidas por uma grande variedade de agentes – na verdade, qualquer agente que de alguma forma influencie na cópia adequada do gene. Reagentes químicos como o gás de mostarda e o metanossulfonato de etila (EMS – *ethyl methanesulfonate*) são potentes mutagênicos, sendo a presença deles no meio ambiente motivo de muita polêmica. Notavelmente, observou-se que o próprio material genético poderia induzir mutações. Em 1940, no Laboratório Cold Spring Harbor em Long Island, Barbara McClintock observou que o genoma do milho possui pequenos elementos genéticos que podem se mover de um local do cromossomo para outro simplesmente inserindo seu material genético em diferentes sítios do genoma. Hoje sabemos que muitos organismos têm uma ou mais classes de elementos móveis (de transposição) que, quando inseridos em um gene, podem quebrar sua funcionalidade (mutagênese insercional). Os retrovírus – vírus que se propagam inserindo seu material genético nos cromossomos das células hospedeiras – podem da mesma forma causar mutações. Considerando que o risco, com base nos estudos em animais, de ocorrer uma mutagênese insercional é muito baixo, os retrovírus têm sido usados na terapia gênica como vetores para carrear genes em pacientes. Infelizmente, várias crianças que receberam a terapia retroviral para uma imunodeficiência que poderia causar-lhe a morte desenvolveram leucemia (Capítulo 14). (Felizmente, a maior parte das leucemias desenvolvidas pode ser tratada com sucesso.)

Mutagênicos continuam sendo bastante utilizados para a detecção de genes por meio de *mutação* ou *rastreamento genético*; os organismos são expostos a um mutagê-

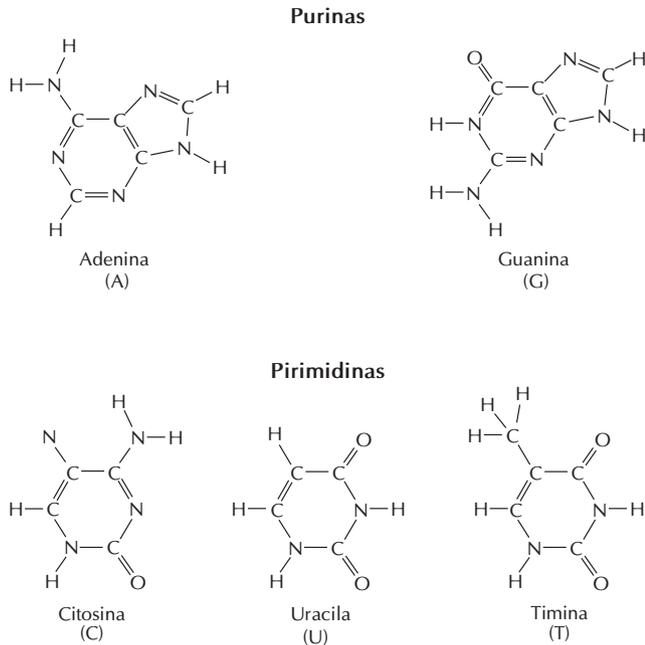
nico e sua progênie é examinada para anormalidades fenotípicas. No caso da *Drosophila*, as mutações têm sido induzidas tanto por agentes químicos como pela utilização de um elemento de transposição da *Drosophila*, denominado elemento P. Quando o peixe-zebra (um pequeno peixe de aquário contendo listras semelhantes às da zebra – *Brachydanio rerio*) foi adaptado para ser utilizado como um organismo modelo, a seleção genética foi realizada usando o agente químico etilnitrosouréia (ENS). Os machos foram expostos à ENS e então cruzados com fêmeas não-tratadas. Para o cruzamento da segunda geração foram necessárias centenas de tanques de peixes, resultando na identificação de mais de 1.000 mutações.

Mutações são quase que invariavelmente deletérias. As vias metabólicas e o desenvolvimento dos organismos são altamente controlados, e a introdução de uma distorção por uma mutação no trabalho celular raramente leva a um melhor desempenho do indivíduo. Mutações nos organismos diplóides em geral são recessivas porque somente uma única cópia das duas cópias do gene é mutada. Ainda assim, mutações são essenciais para a evolução. Sem a variabilidade que as mutações produzem, não teríamos novas formas onde a seleção natural pudesse atuar.

## Os Cromossomos Contêm tanto Ácidos Nucléicos como Proteínas

Embora estivesse definido que os cromossomos carregavam os genes, os antigos geneticistas não focalizaram sua atenção em como os genes eram constituídos; eles não tinham a necessidade de saber qual era a natureza física do gene para realizar sua análise. Na verdade, a análise da natureza química dos cromossomos precedeu qualquer conhecimento de seu papel na hereditariedade. Em 1869, Friedrich Miescher, trabalhando em Tübingen, na Alemanha, isolou do núcleo de células brancas degeneradas de pus (abundante antes da produção de antibióticos) um material contendo fósforo que, por causa de sua origem, ele chamou de nucleína. Por volta de 1874 ele verificou que os espermatozoides de salmão eram mais adequados como matéria-prima para o isolamento da nucleína do que o pus. Miescher mostrou que a nucleína era uma mistura de proteínas carregadas positivamente e um material ácido que, em 1889, Richard Altman chamou de ácido nucléico.

Pela análise de produtos de degradação de ácidos nucléicos utilizando várias fontes, o químico alemão Albrecht Kossel descobriu a existência de bases contendo nitrogênio, guanina (1882), adenina (1886), timina (1893) e citosina (1894). A quinta base, uracila, foi isolada em 1900 (Figura 1-8). Erroneamente, acreditava-se

**FIGURA 1-8**

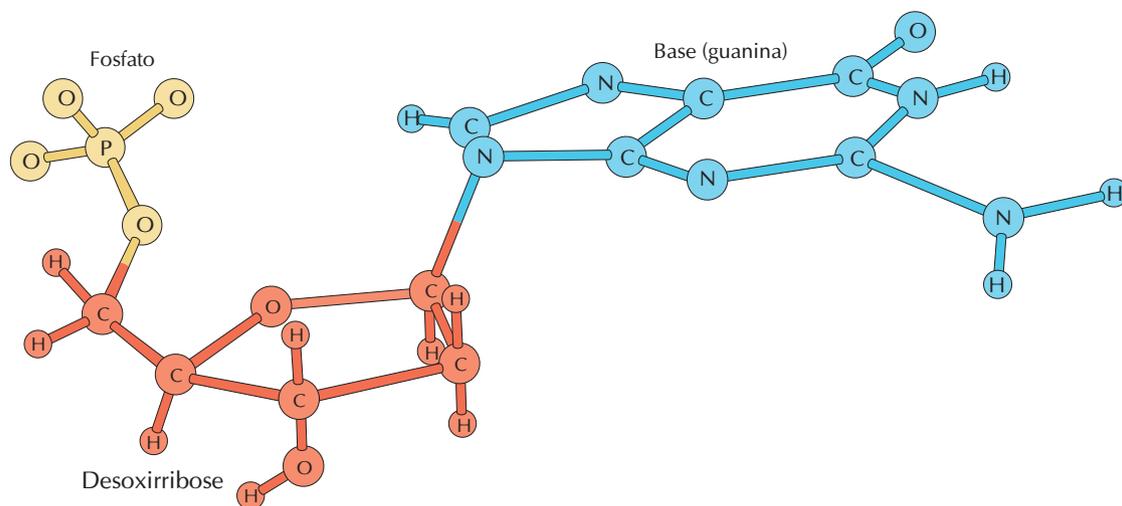
**As bases dos ácidos nucléicos.** Adenina, guanina, citosina e timina são encontradas no DNA. A uracila é um constituinte do RNA, onde toma o lugar da timina.

que algumas células (p. ex., as do timo de bezerro) tinham moléculas de ácidos nucléicos contendo somente timina, enquanto outras fontes de células como as leveduras continham ácidos nucléicos contendo somente uracila. Os componentes glucídicos dos ácidos nucléicos de levedura e do timo de bezerro também eram diferentes. Phoebus Levene, do Instituto Rockefeller, descobriu que o

açúcar no primeiro é a ribose (daí a terminologia ácido ribonucléico – RNA), enquanto no último é a desoxirribose (daí o nome, ácido desoxirribonucléico – DNA). Pouco a pouco, os pesquisadores perceberam que esses ácidos nucléicos não eram exclusivos de leveduras e células animais, mas ocorriam, em diferentes quantidades, em todas as células conhecidas.

Levene mostrou também que tanto o DNA como o RNA eram constituídos de polinucleotídeos – conjunto de blocos construtores, nucleotídeos, covalentemente ligados. Cada nucleotídeo contém também um grupo fosfato, uma porção glucídica (ribose ou desoxirribose), e ou uma purina ou uma pirimidina (moléculas planas contendo carbono e nitrogênio). A Figura 1-9 mostra os três constituintes de um nucleotídeo. Duas purinas e duas pirimidinas são observadas tanto no DNA como no RNA. Duas purinas, adenina e guanina, e uma pirimidina, citosina, são utilizadas tanto para o DNA como para o RNA. A pirimidina timina é encontrada somente no DNA e a uracila, quimicamente semelhante à pirimidina timina, ocorre somente no RNA.

Nesta época, os químicos não se preocupavam em usar condições mais brandas para o isolamento de forma a preservar as estruturas macromoleculares dos ácidos nucléicos; na verdade, estavam mais interessados em determinar os constituintes dos ácidos nucléicos a partir de seus produtos de degradação. Levene erroneamente acreditou que a molécula de ácido nucléico era constituída de uma única cadeia de quatro nucleotídeos e continha somente um de cada tipo de nucleotídeo. Posteriormente, à medida que foram identificados ácidos nucléicos contendo centenas de nucleotídeos, a hipótese do tetranucleotídeo tornou-se cada vez mais insustentável. A possibilidade, no entanto, que não podia ser

**FIGURA 1-9**

**O nucleotídeo de DNA.** A base (azul) está ligada ao anel desoxirribose (laranja) que, por sua vez, está ligado ao grupo fosfato (amarelo). A base mostrada aqui, guanina, pode ser substituída por qualquer uma das três bases do DNA – adenina, citosina, timina. Nas moléculas de DNA, os nucleotídeos formam uma longa cadeia e estão unidos pelo grupo fosfato de um nucleotídeo e a desoxirribose do nucleotídeo adjacente.

descartada era de que o DNA e o RNA eram polímeros uniformes, com cada base sendo repetida a cada quatro nucleotídeos em toda a extensão da cadeia polinucleotídica. Se assim fosse, os ácidos nucleicos não seriam efetivamente específicos e seriam assim diferentes nas proteínas nas quais um número efetivamente ilimitado poderia, em teoria, ser construído pelo alinhamento dos 20 aminoácidos em diferentes ordens.

### A Determinação Quantitativa das Bases do DNA em Diferentes Organismos Sugere uma Variedade Ilimitada de Moléculas de DNA

Somente no final de 1940 foi possível uma quantificação mais precisa de nucleotídeos utilizando uma técnica de cromatografia desenvolvida na Inglaterra por A. J. P. Martin e R. L. M. Synge para a determinação das quantidades de vários aminoácidos em proteínas. As bases do DNA foram, pela primeira vez, quantitativamente analisadas por Erwin Chargaff em seu laboratório no College of Physicians and Surgeons da Universidade de Columbia. Em 1951 estava claro que as quatro bases não estavam presentes em quantidades iguais (Tabela 1-1) e que o DNA não poderia ser feito de unidades de tetranucleotídeos como na hipótese de Levene. Isso também significava que existiria uma grande restrição estrutural para a ordem dos nucleotídeos na molécula de DNA, ao passo que, se a seqüência das quatro bases fosse irregular, o número de diferentes seqüências de DNA seria astronomicamente grande, na ordem de  $4^n$ , onde  $n$  é o número de nucleotídeos na molécula de DNA. Chargaff também notou que as quantidades das quatro bases também não variavam de forma independente (Tabela 1-1). De todas as espécies que ele estudou, as quantidades de adenosina (A) foram muito próximas das quantidades de timina (T). De forma semelhante, a quantidade da segunda base, guanina (G), foi sempre similar à segunda base, citosina (C). Entretanto, existiam alguns casos anômalos que podem ter mascarado de maneira significativa estas relações. Como exemplo, o bacteriófago T2, um vírus que infecta bactéria, não tem citosi-

na. O fato é que Chargaff falhou em considerar que as relações A/T e G/C eram restritas à estrutura do DNA.

Determinações quantitativas de DNA nas células sustentaram o ponto de vista de que o DNA seria o material genético. A mesma quantidade de DNA foi observada em todas as células diplóides de uma dada espécie com a metade dessa quantidade observada nas respectivas células de espermatozóides e óvulo. Igualmente importante foi o fato de que a relação  $A + T = G + C$  não foi similar em todos os organismos. Alguns têm predominância de A e T, enquanto outros, predominância de G e C. Ambos os resultados estão de acordo com a hipótese de que os genes poderiam ser feitos de DNA.

### A Natureza Longa e Linear da Molécula do DNA e seu Diâmetro São Determinados por Microscopia Eletrônica

Apesar de no final de 1930 físicos-químicos suecos mostrarem evidências de que a molécula de DNA era altamente assimétrica devido ao seu comportamento em solução, uma medida direta de seu tamanho só se tornou possível quando o microscópio eletrônico foi popularizado após a Segunda Guerra Mundial. Amostras preparadas cuidadosamente mostraram moléculas extremamente alongadas com milhares de angstroms ( $\text{Å} = 10^{-10} \text{ m}$ ) de comprimento e aproximadamente  $20 \text{ Å}$  de espessura. Todas as moléculas eram lineares e sem ramificação, confirmando a estrutura proposta pelos químicos orgânicos. Considerando o comprimento das moléculas e o fato de um nucleotídeo não medir mais do que  $3 \text{ Å}$ , ficou claro que em sua maioria as moléculas de DNA eram compostas de milhares de nucleotídeos e possivelmente eram mais longas do que qualquer outra molécula polimérica natural.

### Os Nucleotídeos de DNA e RNA Estão Ligados por Ligações Fosfodiéster 3'-5'

Contudo, como os milhares de nucleotídeos, que correspondem aos blocos construtores do DNA, estão ligados para formar moléculas tão grandes? Vários anos após o término da Segunda Guerra Mundial, Alexander Todd, já um químico inglês renomado, decidiu investigar esta questão. Em 1952, seu grande grupo no laboratório de química da Universidade de Cambridge decifrou com precisão as ligações ésteres de fosfato que mantêm os nucleotídeos unidos. Seus resultados foram incontestavelmente simples. Essas ligações foram sempre as mesmas, com um grupo fosfato do carbono 5' de um desoxirribonucleotídeo ligado ao carbono 3' do desoxirribonucleotídeo adjacente (Figura 1-10).

TABELA 1-1. Análise de bases de Chargaff

Base	Fonte de DNA			
	Espermatozóide humano		Timo de boi	
A	0,29		0,28	
T	0,31	A/T 0,94	0,24	A/T 1,16
G	0,18		0,24	
C	0,18	G/C 1,00	0,18	G/C 1,33
				0,15 G/C 1,20

Estas são as três melhores análises de Chargaff, onde a recuperação do material foi de mais de 90% e a relação de bases variou entre 0,94 e 1,33.

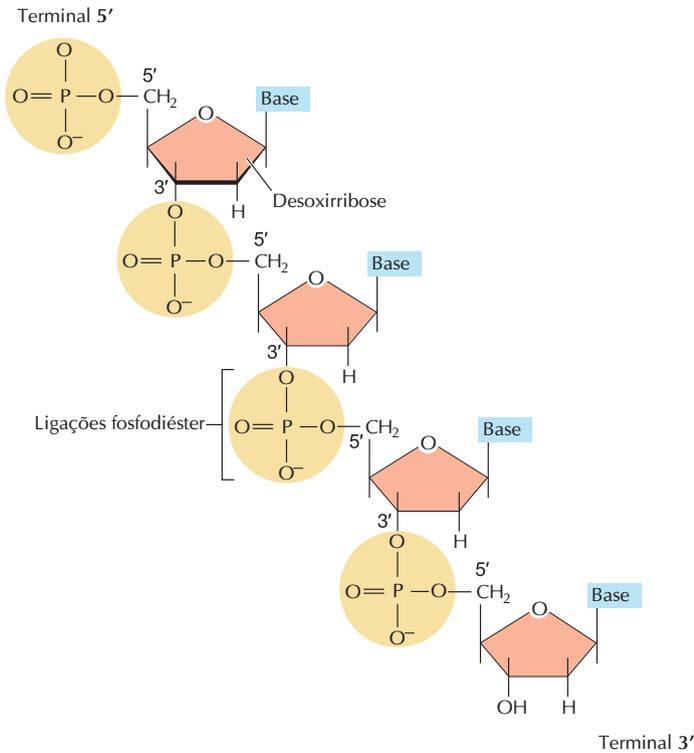


FIGURA 1-10

Os nucleotídeos da molécula de DNA estão ligados por ligações fosfodiéster entre a desoxirribose de um nucleotídeo e o grupo fosfato do nucleotídeo que se segue. Esta cadeia nucleotídeo-fosfato forma o “esqueleto” da molécula de DNA.

Isso significa que a fita de DNA termina com um grupo fosfato 5' em uma extremidade e um grupo hidroxila 3' do outro lado. Nenhum traço de qualquer outra ligação incomum foi observado, e assim o grupo de Todd concluiu que as cadeias de polinucleotídeo do DNA, assim como as cadeias polipeptídicas das proteínas, são moléculas estritamente lineares. Demorou-se muito para estabelecer a natureza da ligação no RNA; mas três anos mais tarde, em 1955, observou-se que o RNA também era uma cadeia linear altamente regular e que também utilizava somente ligações fosfodiéster 3'-5' para manter os seus nucleotídeos juntos.

### Dosagens Biológicas Mostram o DNA como o Material Genético Primário

Em 1928, o microbiologista inglês Frederick Griffith, enquanto investigava a patogenicidade da bactéria causadora da pneumonia, *Diplococcus pneumoniae*, fez uma inesperada observação. Células patogênicas mortas pelo calor, quando misturadas com células vivas não-pato-

gênicas e injetadas em camundongos, foram capazes de transformar uma pequena porcentagem de células não-patogênicas em células patogênicas. Ao se tornarem patogênicas, as células não-virulentas adquiriram uma espessa parede celular rica em polissacarídeos (cápsula), que confere patogenicidade às células. Griffith demonstrou que alguma substância da célula patogênica, não destruída pelo calor, poderia ser absorvida pelas células não-patogênicas, capacitando-as a fazer cápsulas. Posteriormente, estas novas bactérias patogênicas mantinham sua patogenicidade quando isoladas e reinjetadas em camundongos. Esta mudança poderia ser, portanto, herdada.

Griffith não tentou identificar esta substância. Esta tarefa coube a Oswald Avery, cientista do Instituto Rockefeller em Nova Iorque, que trabalhava na química de carboidratos contidos nas paredes das bactérias. Quando começou a trabalhar no “fator de transformação”, Avery pensou que a substância ativa poderia ser um polissacarídeo complexo que de alguma forma induzia a síntese de mais polissacarídeos do mesmo tipo. Na verdade, após uma década de estudo intensivo, Avery e seus jovens colegas, Maclyn McCarty e Colin MacLeod, concluíram, em 1943, que o fator de transformação era o DNA (Figura 1-11). Não somente foi o DNA a molécula predominante em sua preparação purificada do fator de transformação, como a atividade de transformação era destruída por uma preparação altamente purificada de DNase, uma enzima recém-descoberta capaz de degradar o DNA. Ao contrário, a atividade transformante não era afetada por exposição a enzimas que degradam proteínas ou RNA.

Os experimentos de Avery, publicamente anunciados em 1944, tinham sido realizados tão cuidadosamente que muitos cientistas aceitaram sua conclusão de que o DNA era o fator de transformação. Contudo, alguns céticos, incluindo colegas do próprio Instituto Rockefeller, preferiram acreditar que de alguma forma Avery tinha perdido a “proteína genética” e que o DNA era necessário para a atividade na dosagem de transformação somente porque funcionava como um suporte, sem nenhuma especificidade, no qual o verdadeiro material genético – a proteína – estava ligado. Havia outra crítica, relacionada à generalidade da observação de Avery. Talvez suas conclusões obtidas do estudo com bactérias patogênicas não fossem verdadeiras para todos os genes e poderiam não ser aplicáveis a organismos superiores. Oito anos se passaram até que um experimento usando vírus bacterianos (bacteriófagos – fagos vem do grego *phagos*, que significa “comendo”) também mostrasse o DNA como a molécula genética.

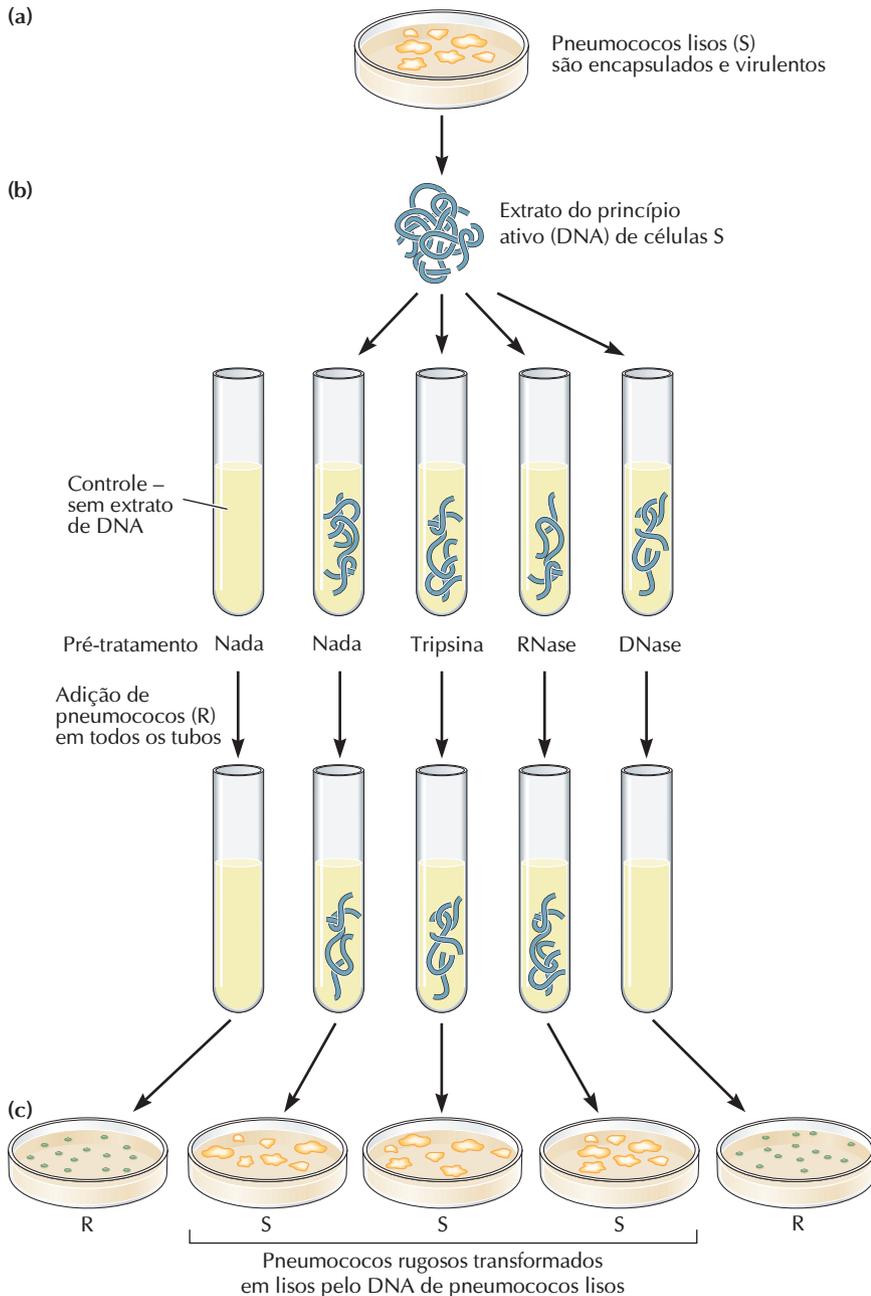


FIGURA 1-11

**Experimento de transformação de Avery.** (a) Formas virulentas da bactéria pneumococo têm uma cápsula de polissacarídeo e formam colônias de superfície lisa (“S”) em cultura. A variante “rugosa” não tem a cápsula de polissacarídeo e forma colônias com características rugosas (“R”). (b) Avery preparou DNA a partir de bactéria do tipo III S. Em alguns casos, ele pré-tratou o DNA com enzimas para determinar a natureza do fator de transformação. Em seguida adicionou bactéria do tipo II R. (c) Sem nenhum extrato da bactéria do tipo S, a bactéria R continuou a produzir colônias rugosas. Com DNA, as bactérias R foram transformadas e produziram a cápsula de polissacarídeo, e então cresceram como colônias lisas. O pré-tratamento com enzimas proteolíticas (tripsina e quimotripsina) ou ribonuclease não produziu nenhum efeito, enquanto a atividade de transformação foi completamente abolida pelo tratamento com um extrato da mucosa intestinal de cachorro (conhecido por destruir DNA).

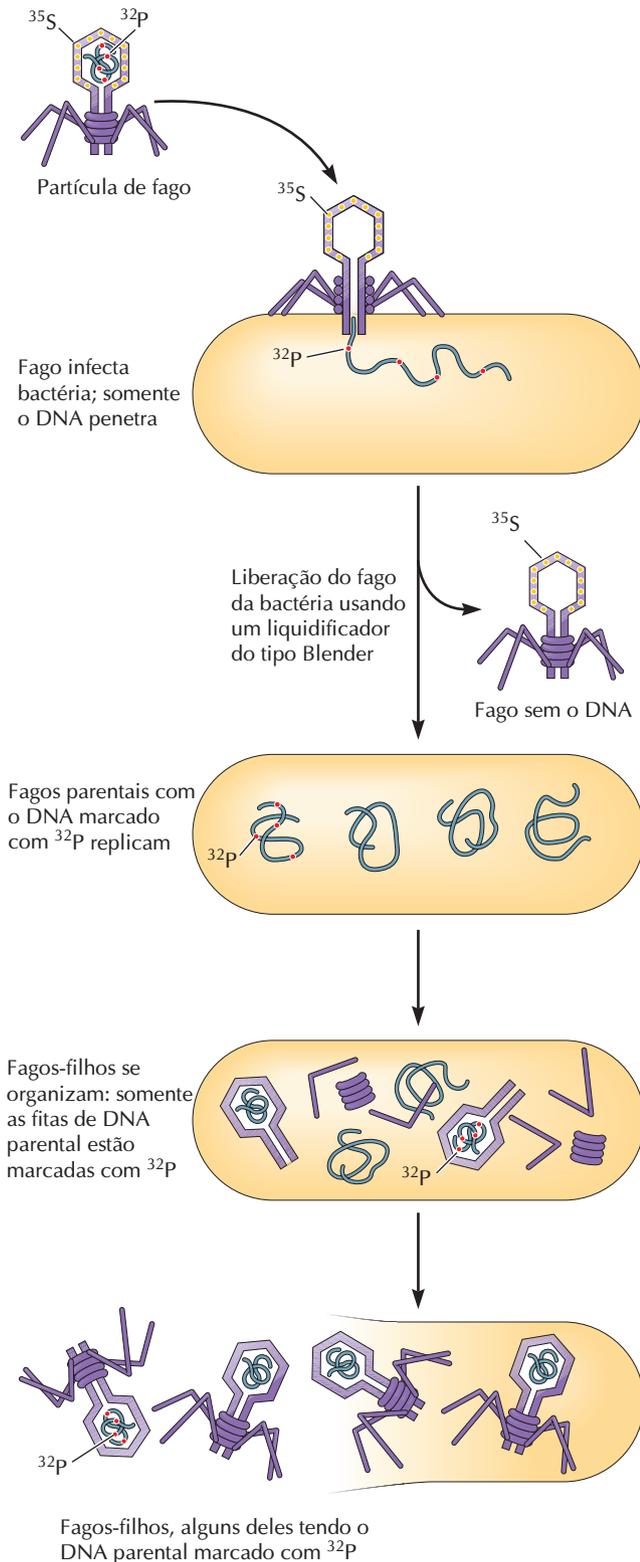
## Vírus São Coleções Empacotadas de Genes que Transmitem suas Instruções de Célula a Célula

O interesse pelo DNA também aumentou, em 1930, com a descoberta de sua ocorrência em vários vírus altamente purificados, dentre eles os bacteriófagos. A natureza destas pequenas partículas, causadoras de doenças e que só se multiplicam em células vivas, foi motivo de muito debate no meio científico. Alguns cientistas consideravam os vírus como um tipo de genes “nus”; outros preferiram considerá-los como a menor forma de vida. Somente quando foi possível purificar os vírus e visualizá-los no microscópio ele-

trônico começou-se a revelar a sua natureza. Obviamente, não eram uma célula minúscula e, como o DNA pode ser liberado por choque osmótico, pareciam ser nada mais do que pacotes de proteínas contendo ácidos nucléicos. A melhor hipótese, portanto, foi de que seriam parasitas obrigatórios em nível genético, incapazes de multiplicação exceto quando dentro de células vivas. Presumivelmente estudando estes vírus simples e descobrindo como se multiplicam, poderíamos entender como os genes são trabalhados e replicados.

Em 1940, o físico Max Delbrück, o médico Salvador Luria e o físico-químico Alfred Hershey estudaram intensivamente um grupo de fagos denominados T1, T2,

T3, T4, etc., os quais se multiplicam dentro da bactéria intestinal *Escherichia coli*. Eles são adequados para a análise genética, pois uma simples partícula infectante necessita somente de 20 minutos para produzir, dentro da célula hospedeira, milhares de partículas virais. Em 1944,



Luria isolou o primeiro fago mutante, somente dois anos antes de Hershey ter demonstrado que os cromossomos dos fagos também participam dos eventos semelhantes aos *crossing-over* que rearranjam genótipos parentais. Esta estratégia genética culminou com os experimentos realizados na Universidade de Purdue por Seymour Benzer em meados de 1950. Seus resultados geraram um grande avanço no entendimento da natureza do gene. Ele isolou milhares de mutantes oriundos de mutações do gene *rII* do fago T4 e mostrou que as recombinações genéticas podem ocorrer tanto dentro dos genes como entre os genes.

### O DNA do Fago por si só Carrega a Informação Genética

Experimentos puramente genéticos não puderam determinar se o responsável pela especificidade genética seria o DNA ou um dos componentes protéicos da partícula viral. Foi em 1951, por fotografias de microscopia eletrônica de um bacteriófago infectando uma célula, que se mostrou que a cápsula do fago ficava retida na parte externa da bactéria. Isto sugeriu que somente o DNA do fago era suficiente para a produção de novos fagos. Presumivelmente, a proteína do fago era apenas uma cápsula protetora e o veículo pelo qual o DNA era introduzido na célula bacteriana. Para testar esta hipótese, em 1952, Alfred Hershey e Martha Chase, do Laboratório Cold Spring Harbor de Nova Iorque, realizaram uma das experiências clássicas da genética molecular (Figura 1-12). Usando  $^{35}\text{S}$  para marcar a proteína e  $^{32}\text{P}$  para marcar o DNA do fago, eles mostraram que a maior parte do  $^{35}\text{S}$  permanecia fora da bactéria infectada e que a maior parte do  $^{32}\text{P}$  era encontrada dentro da bactéria. Posteriormente, cerca de 30% deste  $^{32}\text{P}$  e menos de 1% do  $^{35}\text{S}$  apareciam na progênie do fago, indicando a continuidade do DNA original do fago na progênie. De forma semelhante, o DNA por si só era essencial para uma infecção produtiva.

FIGURA 1-12

**O experimento de Hershey-Chase.** Bacteriófagos T2 foram crescidos em meio contendo  $^{32}\text{P}$  para marcar seu DNA e  $^{35}\text{S}$  para marcar a cápsula de proteína. Os fagos foram adicionados à cultura de *Escherichia coli* e, em vários tempos após a infecção, foram liberados das bactérias por uma vigorosa homogeneização em liquidificador de tipo Blender. A análise mostrou a maior parte da atividade  $^{35}\text{S}$  foi removida da *E. coli*, enquanto a maior parte da atividade de  $^{32}\text{P}$  permaneceu dentro da *E. coli*. Além disso, se a infecção não fosse interrompida e a progênie analisada, o último continha  $^{32}\text{P}$ , indicando assim que existia uma preservação do DNA do fago infectante para a sua progênie. Estes resultados eram esperados se o DNA fosse realmente o material hereditário do fago.

Entretanto, a mesma objeção feita contra os experimentos de Avery e suas conclusões de que o DNA era o material genético foi, igualmente, feita em relação aos experimentos de Hershey e Chase. Eles não descartavam a possibilidade de que uma pequena quantidade da “proteína genética” que estivesse presente na proteína  $^{35}\text{S}$  do fago tivesse sido introduzida nas células bacterianas. Contudo, o experimento de Hershey e Chase, devido à sua ingenuidade e elegância, aumentou a expectativa de que, para se aprender a essência química do gene, o DNA era a molécula a ser estudada.

## O DNA Tem uma Forma Regular e Repetitiva

Para se conhecer profundamente a função de uma molécula biológica e como ela realiza tal função, é necessário conhecer sua estrutura. Atualmente, toda a biologia está voltada para o estudo de como as moléculas interagem entre si e como estas interações ocorrem em nível de átomos. A forma como a hemoglobina capta o oxigênio, por exemplo, só foi entendida quando Max e seus colaboradores na Universidade de Cambridge conheceram a estrutura atômica da molécula de hemoglobina. O mesmo é verdade para o DNA. Assim, as primeiras tentativas para determinar a forma molecular do DNA começaram em 1930.

De uma forma geral, as cadeias de DNA são muito regulares; elas contêm resíduos repetitivos de desoxirribose-fosfato que são unidos pelos mesmos tipos de ligações químicas. Estes grupos repetitivos idênticos formam o esqueleto do DNA. Todavia, se as quatro bases no DNA estão ligadas em qualquer ordem ao longo do esqueleto, esta variabilidade originaria moléculas de DNA com alto grau de individualidade. Assim, dependendo da parte da molécula de DNA focalizada, ela poderia ser vista como regular ou possivelmente como altamente irregular. Mais importante, entretanto, é a questão da configuração global da molécula de DNA. Pode a molécula de DNA dobrar em uma configuração tridimensional (3D) regular determinada pelo esqueleto regular? Neste caso, a configuração seria mais semelhante a uma hélice onde todos os grupos açúcar-fosfato teriam uma vizinhança química idêntica. Por outro lado, se as bases predominam na estrutura da molécula, isso poderia ser terrível, pois nenhuma das duas cadeias teria configuração 3D idêntica, visto que as seqüências das bases são irregulares. Pois este foi o caso: a tarefa de descobrir como cada uma destas moléculas com diferentes formas poderia servir como molde para a formação de outra cadeia de DNA realmente estava fora do nosso alcance.

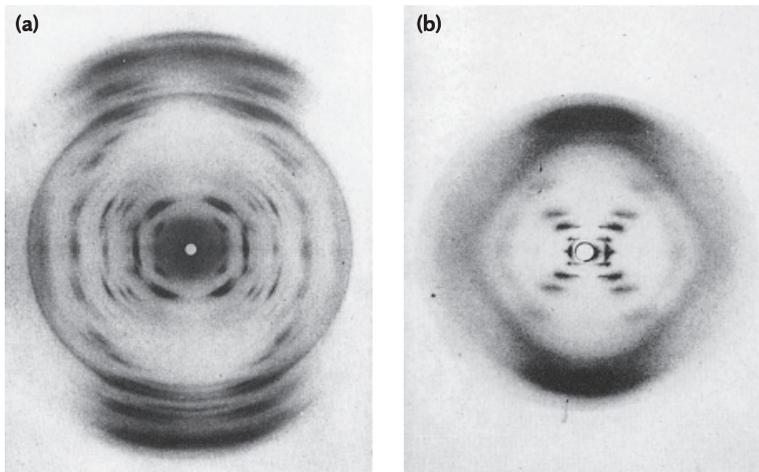
A única forma direta de examinar a estrutura tridimensional do DNA era observar como ele difrata (desvia) os raios X. O DNA seco tem a aparência de filetes

irregulares de algodão, mas se torna altamente viscoso quando colocado em água, podendo assim ser retirado da solução aquosa como finas fibras. Na análise estrutural, as fibras são colocadas no caminho do feixe de raios X e o padrão de difração é registrado em um filme fotográfico. O DNA foi pela primeira vez examinado desta maneira em 1938, por William Astbury na Inglaterra, com material preparado na Suécia por Einar Hammarsten, que usou condições neutras de pH e baixas temperaturas para isolar o DNA a partir de moléculas de alto peso molecular das glândulas do timo. Este DNA na verdade produziu padrões muito distintos de difração, e são tantas as moléculas individuais de DNA que elas precisam compartilhar algumas orientações preferenciais. Entretanto, as fotografias de difração de raios X não foram tão precisas como aquelas produzidas pelas substâncias verdadeiramente cristalinas, o que indica a possibilidade de as cadeias de DNA nunca terem uma configuração definida, comum para todas as cadeias. Um pouco antes da Segunda Guerra Mundial, Astbury e seu aluno Florence Bell propuseram que as bases púricas e pirimídicas estavam posicionadas perpendicularmente ao longo eixo da molécula de DNA como se formassem uma pilha de moedas. Melhores padrões de difração de raios X obtidos após a guerra ainda eram de difícil interpretação; agora sabemos que as preparações de Astbury foram misturas de formas de DNA hoje conhecidas como formas A e B.

Então em 1950 o físico Maurice Wilkins e seu aluno de pesquisa Raymond Gosling do King's College, Londres, trabalhando com o DNA que tinha sido cuidadosamente preparado em Bern pelo químico suíço R. Singer, obtiveram o verdadeiro padrão de difração da estrutura cristalina, o padrão “A” (Figura 1-13a). As moléculas individuais de DNA que se juntam para formar as fibras cristalinas precisam ter formas muito similares ou não seriam capazes de se empacotar tão regularmente. Assim, tivemos a certeza de que o DNA tem uma estrutura precisa, uma definição essencial que começaria a revelar a maneira pela qual as moléculas de DNA são copiadas com tanta precisão durante o processo de replicação.

## Duas Cadeias Polinucleotídicas Entrelaçadas Constituem a Unidade Fundamental do DNA (a Dupla Hélice)

Os dados obtidos dos padrões de difração de raios X de uma molécula tão complicada como o DNA em geral não fornecem informação suficiente para revelar a estrutura molecular em estudo. Uma análise de tais padrões, entretanto, freqüentemente fornece parâmetros-chave que marcam bastante as características estruturais da molé-



**FIGURA 1-13**

**Padrão de difração de raios X.** (a) A forma cristalina A; (b) a forma B produzida na presença de alta umidade. Os pontos que aparecem no padrão de difração de raios X são característicos de uma estrutura de hélice.

cula em investigação. Esse foi o caso do DNA. Os padrões de raios X, contudo, não foram os obtidos das fibras cristalinas do DNA, mas sim de formas menos agregadas obtidas quando as fibras são expostas a uma alta umidade com mais água de hidratação. Estes parâmetros paracristalinos, a forma “B” (Figura 1-13b), foram vistos pela primeira vez no verão de 1951 pela físico-química inglesa Rosalind Franklin, uma colega de Wilkins do King’s College. Suas fotografias mostraram um padrão dominante de ligações cruzadas, uma figura típica de hélice. Assim, apesar da presença de uma seqüência irregular de bases, o esqueleto açúcar-fosfato do DNA de alguma forma assumia a configuração de hélice. Um nucleotídeo foi observado a cada 3,4 Å ao longo do eixo da fibra, com 10 nucleotídeos, ou 34 Å, sendo necessários para cada volta da hélice.

Uma conclusão importante originada das determinações de raios X foi obtida quando se mediu o diâmetro da hélice. A largura de 20 Å estava longe de corresponder a uma molécula de DNA contendo somente uma cadeia de DNA. Mas realmente quantas cadeias estavam na molécula de DNA? A dificuldade de se medir a quantidade exata de moléculas de água na fibra de DNA levou à confusão inicial sobre se existiam duas ou três cadeias entrelaçadas. Felizmente a simetria observada na difração de raios X é favorável a moléculas contendo duas cadeias que se dispõem em direções opostas. Antes da utilização dos métodos de difração de raios X, nenhum químico suspeitou que o DNA fosse constituído de uma molécula com cadeias múltiplas.

### As Duas Cadeias Polinucleotídicas São Mantidas Juntas por Pontes de Hidrogênio Formadas entre as Bases Complementares

A forma como as duas cadeias são mantidas juntas na molécula de DNA não poderia ser deduzida somente pelos

dados de raios X. Com esta finalidade, o biólogo James Watson e o físico Francis Crick do laboratório de Cavendish na Universidade de Cambridge idealizaram, em fevereiro de 1953, um modelo tridimensional do DNA com uma configuração mais favorável e mais compatível com os parâmetros da hélice fornecidos pelos dados de raios X obtidos pelos pesquisadores do King’s College. Esta estratégia os conduziu a um modelo molecular no qual os açúcares-fosfatos da cadeia estivessem voltados para o lado externo da molécula de DNA e as bases púricas e pirimídicas para o seu interior, orientadas de tal forma que as bases de uma fita pudessem formar pontes de hidrogênio com as bases da outra fita. Embora as interações por pontes de hidrogênio sejam bem mais fracas do que as interações covalentes, existe uma quantidade tão grande de pontes entre as duas cadeias que faz com que elas sejam mantidas juntas de forma bem estável.

O padrão exato de pontes de hidrogênio utilizado no DNA surgiu somente após a construção do modelo no qual ficou estabelecido que, se uma purina de uma cadeia forma pontes de hidrogênio com uma pirimidina da outra cadeia, as dimensões das bases pareadas são as mesmas ao longo de todo o comprimento do DNA. Também importante, as duas purinas adenina e guanina ligam-se seletivamente às duas pirimidinas timina e citosina, de tal forma que a adenina (A) pode parear somente com a timina (T) e a guanina (G) somente com a citosina (C), o que fornece uma explicação química para as regras de Chargaff:  $A = T$  e  $G = C$  (Figura 1-14). A observação anterior de que o bacteriófago T2 não continha citosina serviu de suporte para uma nova estrutura do DNA. Em 1952, uma citosina modificada – 5-hidroximetil citosina – foi descoberta no DNA de T2. Assim como a citosina em outros DNAs, a quantidade de metil citosina em T2 foi exatamente igual à de guanina, mostrando que ela forma um par de bases tal como a citosina.

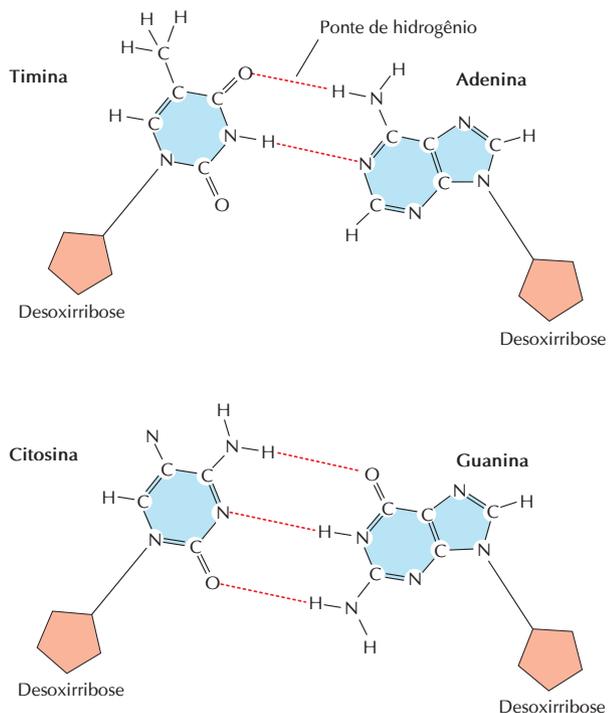


FIGURA 1-14

**Pontes de hidrogênio formadas entre os pares de bases adenina-timina e guanina-citosina.** Existem duas pontes de hidrogênio que mantêm unidas A:T e três entre G:C.

Cada um destes pares de bases possui uma simetria que permite a inserção na dupla hélice sob duas orientações (A:T e T:A; G:C e C:G); assim, em qualquer cadeia de DNA, as quatro bases podem existir em todas as variações possíveis de seqüência. Devido ao pareamento específico entre as bases, se a seqüência de uma cadeia (p. ex., TCGCAT) é conhecida, a seqüência da outra cadeia (AGCGTA) também o será. A seqüência da cadeia oposta é denominada *complementar* e as cadeias pareadas são conhecidas como *cadeias complementares*. Somente com o pareamento de bases seria possível ter uma cadeia de grupos açúcar-fosfato com orientações idênticas e assim permitir que o DNA tenha a mesma estrutura com qualquer seqüência de bases (Figura 1-15). Ficou também evidente que para se ter uma molécula simétrica é necessário que as duas fitas da dupla hélice estejam em direções opostas; o fosfato 5' no final de uma fita sendo oposto ao hidroxil 3' da outra fita. (Por convenção, as fitas únicas de DNA em geral são desenhadas com a porção 5' no lado esquerdo.)

### Uma Cadeia Complementar do DNA Funciona como Molde para a Síntese da Nova Cadeia Durante a Duplicação

Antes de sabermos como os genes se apresentavam era praticamente impossível especular de forma intelligen-

te como eles poderiam ser exatamente duplicados durante a divisão celular. Qualquer proposição tinha que ser geral. Linus Pauling e Max Delbrück sugeriram em 1940 que a superfície de um gene de alguma maneira atua como uma matriz positiva, ou molde, para a formação de uma molécula complementar (negativa), da mesma maneira que um material pode ser moldado a partir de um pedaço de escultura para se fazer uma fôrma. A molécula complementar matriz poderia então servir como um molde para a formação de seu próprio complemento, produzindo assim uma cópia idên-

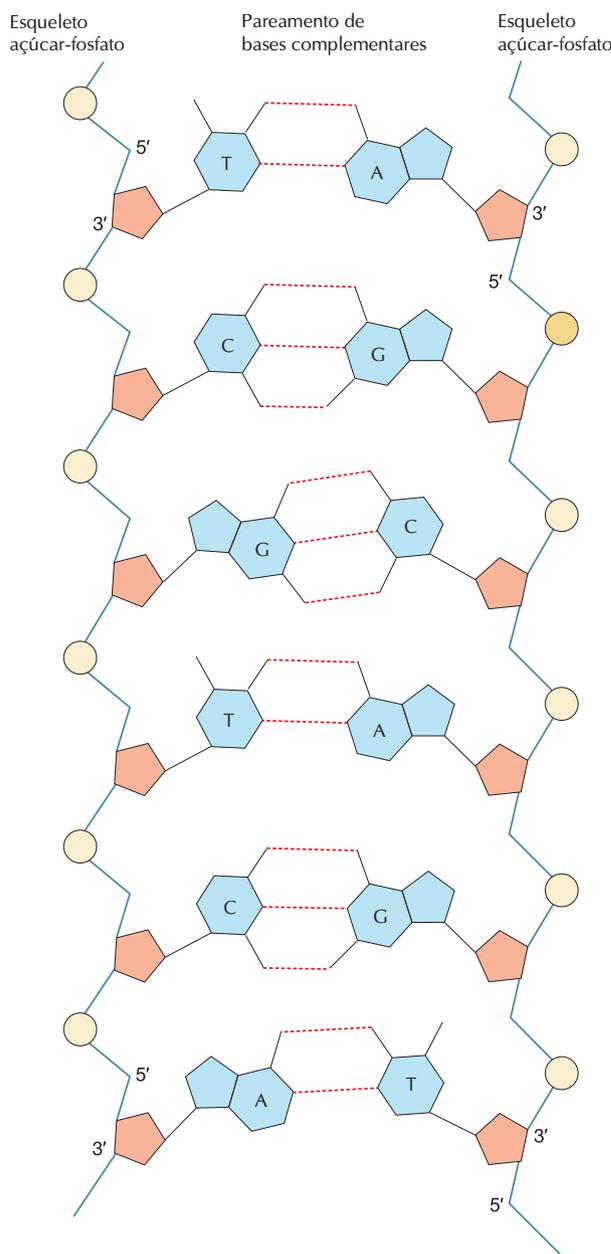


FIGURA 1-15

**A molécula de DNA tem duas cadeias açúcar-fosfato mantidas juntas por pontes de hidrogênio formadas entre as bases complementares, A:T e G:C.**

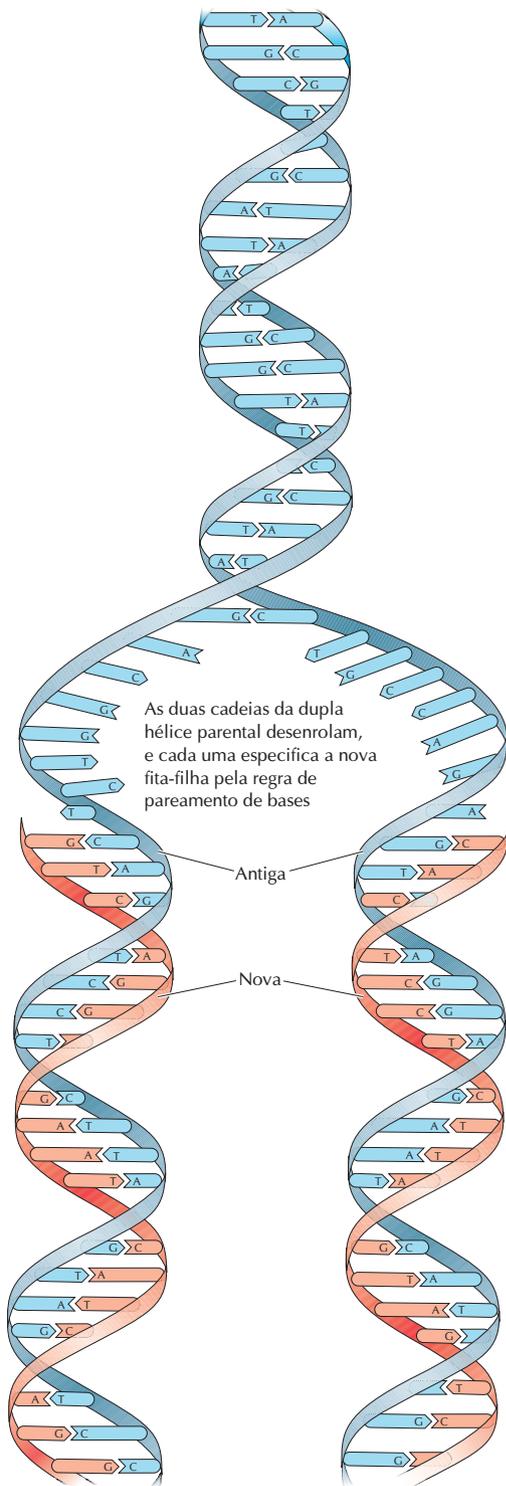


FIGURA 1-16

**Duplas hélices filhas idênticas são geradas através da replicação semiconservativa do DNA.** As fitas da hélice original se separam e atuam como moldes para a síntese das fitas-filhas complementares.

tica ao molde original. Assim, a descoberta de que as duas cadeias tinham forma complementar causou grande entusiasmo, sendo imediatamente proposto que as duas fitas da dupla hélice estariam relacionadas como moldes positivos e negativos, cada uma especificando o seu complemento e, portanto, capaz de gerar duas moléculas de DNA filhas com seqüências idênticas às da dupla hélice parental (Figura 1-16). O que não se sabia era se a duplicação ocorria de maneira *conservativa*, na qual as duas fitas originais são mantidas juntas e novas fitas complementares formam uma nova dupla hélice, ou de maneira *semiconservativa*, em que as duas fitas originais se separam completamente durante a síntese das novas fitas complementares, de tal forma que a nova dupla hélice contém uma fita parental e uma nova fita.

### A Replicação do DNA É Semiconservativa, Produzindo uma Nova Fita Sintetizada Pareada com uma Fita Parental

Somente cinco anos depois apareceu uma evidência de que as duas fitas se separavam. A prova definitiva veio com os experimentos de Matthew Meselson e Franklin Stahl do California Institute of Technology, em 1958. Eles tiveram a brilhante idéia de usar diferenças de densidade para separar as moléculas de DNA parental das moléculas-filhas (Figura 1-17). Inicialmente cresceram culturas da bactéria *E. coli* em meio enriquecido com o isótopo pesado  $N^{15}$ . Em função de seu conteúdo isotópico, o DNA nestas bactérias era muito mais pesado do que o DNA produzido quando estas células eram crescidas na presença de um meio enriquecido com o isótopo natural  $N^{14}$ . Por causa da maior densidade, o DNA mais pesado poderia ser separado do DNA mais leve por uma ultracentrifugação utilizando gradiente de densidade formado por cloreto de césio. Quando as células contendo o DNA mais pesado foram transferidas para um meio “leve” normal ( $N^{14}$ ) e se multiplicaram por uma geração, todo o DNA pesado foi substituído por um DNA cuja densidade estava entre leve e pesada. O desaparecimento do DNA pesado mostrou que os produtos da replicação não incluem a dupla hélice original, e então o processo era semiconservativo, com cada molécula-filha tendo uma fita pesada (parental) e uma fita leve (recém-sintetizada).

Ainda não se sabia se as fitas complementares se separavam completamente antes da replicação. Mais tarde, muitas evidências reveladas por microscopia eletrônica mostraram que a forma de Y que aparecia na forquilha de replicação indicava que a separação das duas fitas e a replicação ocorrem simultaneamente. Tão logo a dupla

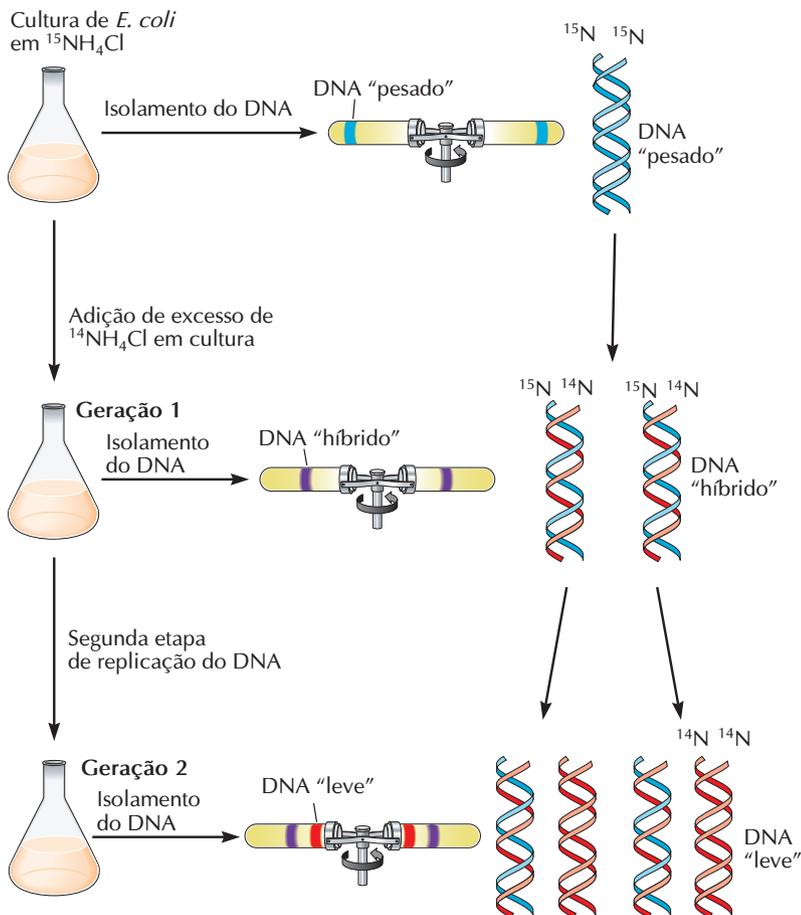


FIGURA 1-17

O experimento de Meselson-Stahl demonstrando a replicação semiconservativa do DNA. *Escherichia coli* foram crescidas em meio contendo  $^{15}\text{NH}_4\text{Cl}$  para produzir moléculas de DNA "pesado". Amostras foram coletadas desta cultura em vários tempos após a adição em excesso de  $^{14}\text{NH}_4\text{Cl}$ . O  $^{14}\text{N}$  foi então incorporado no DNA que estava sendo replicado, produzindo moléculas com menor densidade do que aquelas contendo  $^{15}\text{N}$ . Moléculas com diferentes densidades podem ser separadas por ultracentrifugação em gradiente de cloreto de césio – moléculas com maior densidade formam uma banda mais baixa no gradiente do que moléculas de menor densidade. Após uma geração crescendo em  $^{14}\text{NH}_4\text{Cl}$ , uma única banda foi observada com menor densidade do que a primeira amostra coletada antes da adição do  $^{14}\text{NH}_4\text{Cl}$ . O desaparecimento da banda correspondente ao DNA contendo  $^{15}\text{N}$  mostrou que não existiam moléculas de DNA contendo duas fitas marcadas com  $^{15}\text{N}$  – uma única banda precisa conter moléculas híbridas de DNA, uma fita marcada com  $^{15}\text{N}$  e a outra fita marcada com  $^{14}\text{N}$ . Após duas gerações observou-se que, além da banda do DNA híbrido, aparecia uma banda ainda mais leve. Esta nova banda correspondia a uma molécula de DNA onde as duas fitas estavam marcadas com  $^{14}\text{N}$ .

hélice começa a separar para a replicação, as regiões da fita simples são imediatamente utilizadas como molde, tornando-se novas regiões de fita dupla.

## A Replicação de DNA em Tubos-Teste Exige DNAs-Molde Preexistentes

Uma prova independente de que as cadeias de DNA servem como moldes para a formação do DNA da progênie foi fornecida pelos experimentos no laboratório de Arthur Kornberg na Universidade de Washington em St. Louis entre 1956 e 1958. Trabalhando com extratos livres de células de *E. coli*, Kornberg inicialmente observou a incorporação de desoxirribonucleotídeos radiativos em moléculas características de DNA; isto é, eles eram insolúveis em ácidos, destruídos pela DNase e resistentes à RNase. Através de um extensivo fracionamento dos extratos livres de células, não só a enzima de polimerização do DNA, DNA polimerase, foi descoberta, como também foi observada a necessidade da presença de cadeias preexistentes de DNA para servir como moldes na formação da nova cadeia de seqüência complementar. De forma interessante, a

composição de bases do DNA produzido no tubo-teste foi idêntica à composição do DNA-molde, mostrando que a última determina a composição da primeira (Tabela 1-2). Isto foi um marco para os bioquímicos, pois aqui uma enzima estava tomando as *instruções* a partir de um molde, em vez de simplesmente estar promovendo uma reação catalítica.

TABELA 1-2. Composição de nucleotídeo de um DNA enzimaticamente sintetizado comparada com aquela do DNA-molde

DNA	A	T	G	C
Bacteriófago T2				
Molde	1,31	1,32	0,67	0,70
Produto	1,32	1,29	0,69	0,70
<i>Mycobacterium phlei</i>				
Molde	0,65	0,66	1,35	1,34
Produto	0,66	0,65	1,34	1,37
Timo de bezerro				
Molde	1,14	1,05	0,90	0,85
Produto	1,12	1,08	0,85	0,85
Copolímero A:T				
Produto	1,99	1,93	<0,05	<0,05

A composição de bases do DNA produto assemelha-se à do DNA-molde, enquanto o DNA sintetizado quimicamente composto somente de A e T direciona a síntese de uma nova molécula de DNA contendo somente Ts e As.

Curiosamente, a DNA polimerase de Kornberg não é a polimerase que a *E. coli* utiliza para sintetizar DNA *in vivo*. John Cairns e Paula De Lucia do Laboratório Cold Spring Harbor fizeram um mutante de *E. coli* e dosaram a atividade de DNA polimerase destes mutantes. Após testar mais de 3.000 mutantes, observaram que um dos mutantes, apesar de não ter a atividade de polimerase mensurável, mantinha a capacidade de crescimento. Este mutante era mais suscetível a danos do DNA por luz ultravioleta, sugerindo que provavelmente a polimerase de Kornberg, hoje chamada de polimerase I, estava envolvida no mecanismo de reparo do DNA. Mais tarde, duas outras DNA polimerases de *E. coli* foram descobertas por Tom Kornberg, um dos filhos de Arthur Kornberg. Uma delas, a DNA polimerase III, é a principal enzima utilizada para replicar o DNA.

Antes de iniciar este trabalho, Kornberg, assim como seus colegas, focalizaram sua atenção na bioquímica enzimática, considerando a dupla hélice como uma hipótese ainda não estabelecida. Existiam muitas dúvidas sobre se a realidade da dupla hélice se tornaria intelectualmente insustentável após a divulgação dos experimentos de Kornberg e Meselson-Stahl.

## A Replicação do DNA É Realizada por um Conjunto Altamente Complexo de Proteínas

A DNA polimerase III é uma maquinaria molecular bem grande composta de nove proteínas, condizente com a variedade de atividades que ela tem de desempenhar durante a replicação. Essas atividades incluem ligar-se à fita de DNA, mover-se ao longo da fita, sintetizar uma nova fita, verificar se a síntese está sendo feita de maneira correta e ainda interagir com outras proteínas essenciais para o processo de replicação. Existem helicases, responsáveis pela desespiralização da hélice, topoisomerases, que diminuem a tensão criada pela desespiralização, e proteínas, que se ligam ao DNA de forma a estabilizar o DNA desespiralizado. O ponto onde as duas fitas são separadas é chamado de forquilha de replicação, o qual avança ao longo do DNA quando ele é desespiralizado e uma nova fita é sintetizada.

Contudo, a replicação do DNA é complicada. Todas as DNA polimerases adicionam o novo nucleotídeo no grupo 3'-OH do último nucleotídeo da cadeia em crescimento. Ele é sintetizado na direção de 5' para 3', de forma que a polimerase se movimenta ao longo da fita-molde de DNA na direção de 3' para 5'. Para uma das duas fitas (*líder*), ele trabalha muito bem, mas para a outra fita (*atrasada*) ele deve correr em direção reversa. Reiji Okazaki mostrou que a nova fita é sintetizada na fita atrasada em seções curtas, denominadas *fragmentos de Oka-*

*zaki*. A síntese das duas fitas é realizada pelo mesmo complexo da polimerase, sendo que a fita atrasada deve formar uma alça através do complexo. Os espaços entre os fragmentos de Okazaki são preenchidos pela DNA polimerase I e os fragmentos são unidos pela DNA ligase. Levando-se em conta a complexa coreografia de tantas enzimas na forquilha de replicação, a síntese de DNA é relativamente rápida, cerca de 1.000 nucleotídeos por segundo.

## As Moléculas de DNA Podem Ser Desnaturadas e Renaturadas

As moléculas de DNA contêm tantos pares de bases que as cadeias complementares do DNA nunca se separam espontaneamente sob condições fisiológicas. Se, entretanto, as duplas hélices são expostas a temperaturas próximas à ebulição ou a extremos de pH, elas se separam em suas fitas simples (isto é, elas são desnaturadas ou se fundem – “melt”). A temperatura na qual metade das moléculas de DNA de uma amostra é desnaturada é chamada de temperatura de fusão (*melting temperature* –  $T_m$ ); duplas hélices com um excesso de pares de bases G:C são mais estáveis do que duplas hélices com predominância de pares de bases A:T e têm conseqüentemente maior  $T_m$ . Isto ocorre devido ao fato de que G e C no DNA têm um aumentado poder de *empilhamento* dos pares de bases adjacentes, levando a uma maior estabilização da dupla hélice, assim como a presença de três pontes de hidrogênio entre G:C em comparação com a formação de somente duas pontes entre os pares A:T. De fato, a proporção de A:T e G:C pode ser diretamente avaliada pela medida da  $T_m$  da amostra.

No início, a desnaturação foi considerada essencialmente irreversível, mas em 1960 Julius Marmur, Paul Doty e seus colaboradores, trabalhando na Universidade de Harvard, mostraram que duas fitas complementares podem se recombinar e formar a dupla hélice nativa quando são mantidas por várias horas em condições subdesnaturantes, como a 65°C (Figura 1-18). Esta renaturação de uma molécula tão grande é notável. Mais surpreendente é que a renaturação é muito específica e produzirá uma dupla hélice perfeita quando as seqüências de bases das duas fitas forem exatamente complementares. Estas moléculas reaneladas são biologicamente funcionais quando introduzidas em células.

O anelamento é agora chamado de *hibridização* e constitui uma ferramenta poderosa e essencial utilizada no rol das técnicas da genética molecular. Por exemplo, tendo em mãos um pedaço de DNA, ele pode ser usado para se descobrir seu complemento, até mesmo no DNA do genoma total. Marcando-se o fragmento de DNA com

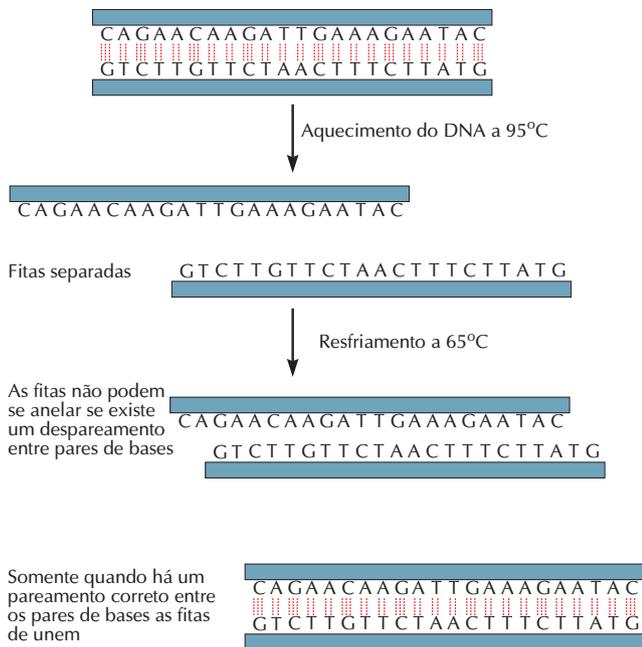


FIGURA 1-18

**O pareamento de bases de um pequeno fragmento de uma molécula de DNA.** As três pontes de hidrogênio entre guanina e citosina e as duas pontes de hidrogênio entre adenina e timina são representadas, respectivamente, por três linhas ou duas linhas. Quando aquecidas a aproximadamente 95°C, as duas fitas se separam. Com um resfriamento lento até aproximadamente 65°C, as duas fitas voltam a se parear, mas o reanelamento ocorre somente quando há uma complementaridade perfeita ou quase perfeita entre as duas fitas.

um composto fluorescente, o local do seu complemento pode ser descoberto em uma seção de tecido; e pela observação da extensão entre as duas fitas de DNA em solução, os pesquisadores podem determinar o grau de similaridade existente entre moléculas de DNA de diferentes espécies e então mostrar o quão próximo duas espécies estão geneticamente relacionadas.

## Cromossomos São Moléculas de DNA muito Longas

Quando a estrutura da dupla hélice foi determinada, não se sabia se os cromossomos do tamanho da *E. coli* continham uma ou mais moléculas de DNA. Da mesma forma, existiria uma única molécula de DNA em cromossomos maiores como os de plantas e animais superiores? O último contém em média 20 vezes mais DNA do que aquele que ocorre nos cromossomos da *E. coli*. Inicialmente acreditava-se que mesmo o mais simples dos cromossomos era constituído por muitas moléculas separadas de DNA, ligadas através de proteínas. Este quadro mudou com a descoberta de que as longas moléculas de DNA eram muito frágeis e fáceis de serem quebradas em fragmentos menores. Quando tais quebras foram minimizadas, foi possível observar o DNA dos cromossomos do fago T2 contendo cerca de 200.000 pares de bases.

Agora que os genomas de muitos organismos já foram totalmente seqüenciados, sabemos que o cromossomo da *E. coli* é constituído de uma molécula simples de DNA, contendo 4.639.221 pares de bases, enquanto o DNA humano do cromossomo 6 tem 166.880.988 pares de bases, além de aproximadamente 480.000 pares de bases adicionais que ainda não foram seqüenciados (Tabela 1-3). Um cromossomo é então apropriadamente definido como uma molécula simples de DNA, geneticamente específica de cada espécie ligada a um grande número de proteínas envolvidas na manutenção da estrutura dos cromossomos e na regulação da expressão gênica.

## Alguns Vírus Têm um Genoma de RNA

Nem todos os organismos usam o DNA como seu material genético. Na verdade, o primeiro vírus que foi

TABELA 1-3. Moléculas de DNA em cromossomos

Organismos e cromossomos	Tipo	Tamanho (bases)	Comprimento (mm) <sup>a</sup>
Bacteriófago λ	Vírus	48.502 <sup>b</sup>	0,016
<i>Mycoplasma genitalium</i>	Micoplasma	580.074 <sup>b</sup>	0,197
<i>Escherichia coli</i> K12-MG1655	Bactéria	4.639.221 <sup>b</sup>	1,577
<i>Saccharomyces cerevisiae</i> Cromossomo IV <sup>c</sup>	Levedura	1.531.974 <sup>d</sup>	0,521
<i>Caenorhabditis elegans</i> Cromossomo V <sup>c</sup>	Nematoda	20.922.238 <sup>d</sup>	7,11
<i>Arabidopsis thaliana</i> Cromossomo 1 <sup>c</sup>	Planta	30.494.425 <sup>d</sup>	9,90
<i>Drosophila melanogaster</i> Cromossomo III <sup>c</sup>	Inseto	51.243.003 <sup>d</sup>	17,42
<i>Homo sapiens</i> Cromossomo 1 <sup>c</sup>	Mamífero	246.047.941 <sup>d</sup>	83,65

<sup>a</sup> O comprimento dos cromossomos foi calculado usando o valor de 0,34 nm por nucleotídeo. (Existem dez nucleotídeos por volta na dupla hélice, e cada volta corresponde a 3,4 nm.)

<sup>b</sup> Estes cromossomos já foram completamente seqüenciados.

<sup>c</sup> Estes são os cromossomos mais longos destes organismos.

<sup>d</sup> Estes cromossomos ainda não foram completamente seqüenciados; algumas regiões que são muito difíceis de ser seqüenciadas (p. ex., os centrômeros) ainda não foram seqüenciadas.

isolado e cristalizado – vírus do mosaico do tabaco – é constituído de RNA ao invés de DNA. Outros vírus que contêm genomas de RNA incluem bacteriófago R17, vírus da influenza e poliovírus; vale ressaltar a existência de vírus com uma fita simples de RNA e outros contendo um RNA de fita dupla. Estes vírus replicam seus genomas de RNA da mesma forma que os outros organismos replicam seu DNA – o RNA de fita simples serve como molde para fazer cadeias com seqüências complementares.

Os retrovírus – por exemplo, o vírus da imunodeficiência humana (HIV) – têm genomas de RNA, mas se replicam de forma completamente diferente dos outros vírus RNA. Os retrovírus foram inicialmente conhecidos por causar câncer em animais; quando infectam as células, os retrovírus as *transformam* em células tumorais. Assim como a transformação dos pneumococos de um tipo celular em outro foi acompanhada de mudanças nas características da bactéria, as células de mamíferos transformadas pelos vírus tumorais de RNA adquirem novas características – um rápido crescimento, diminuição da necessidade de nutrientes, além da capacidade de induzir tumores quando injetadas em animais. Estas mudanças são hereditárias e os cientistas observaram que as células transformadas continuavam a expressar novas proteínas, presumivelmente de origem viral. Esta observação levou à proposição de que os vírus tumorais de RNA de alguma forma se integram no DNA cromossomal das células transformadas. Como isto é possível – ou seja, como o RNA pode ser convertido em DNA antes da integração no genoma do hospedeiro – foi decifrado em 1970, quando, por meio de duas pesquisas independentes, Howard Temin e Satoshi Mizutani de um lado e David Baltimore do outro descobriram uma enzima capaz de promover a *transcrição reversa* do RNA cromossomal dos vírus infectantes em uma fita complementar de DNA.

A enzima envolvida, *transcriptase reversa*, é codificada por todos os genomas virais e está presente na partícula viral de forma que é capaz de agir no momento que o RNA infectante entra nas células hospedeiras apropriadas. A primeira cadeia de DNA produzida pela transcriptase reversa – complementar ao RNA viral – atua como molde para fazer a segunda fita de DNA, produzindo assim um DNA de fita dupla, que pode agora ser integrado como um *provírus* no DNA cromossomal da célula hospedeira. O provírus de DNA é agora replicado com o DNA da célula hospedeira. O

ciclo de vida do vírus é completado pela transcrição do DNA do provírus em uma fita de RNA idêntica àquela do vírus infectante e pela síntese de proteínas de superfície do vírus. Os retrovírus podem se integrar nas células germinativas e perpetuarem-se de geração a geração. Eles em geral podem sofrer mutações sem capacidade de replicar, tornando-se retrovirais “fósseis” no genoma do hospedeiro. Existem centenas de milhares de seqüências relacionadas aos retrovírus dispersas nos cromossomos de muitos organismos.

Os retrovírus têm se mostrado ferramentas bastante úteis para os geneticistas moleculares. O comportamento dos retrovírus – infectando células, integrando seu DNA nos cromossomos do hospedeiro – é explorado na geneterapia, onde o vetor retroviral pode carrear um gene corrigido para as células defeituosas e o gene se torna permanentemente integrado no genoma (Capítulo 6). A transcriptase reversa permite aos pesquisadores sintetizarem uma cópia de DNA a partir de qualquer RNA e constitui hoje, como veremos no Capítulo 4, um elemento-chave para o desenvolvimento de técnicas para o isolamento e a clonagem de genes.

No final da década de 1960, um considerável progresso foi alcançado usando vírus e bactérias como sistemas-modelo para o estudo da regulação gênica e do câncer, mas parecia ser impossível entendermos, em nível molecular, como as células trabalham nos organismos superiores. Mesmo sabendo que os genes eram feitos de DNA, parecia não existir um caminho para explorar tal conhecimento – nenhuma forma de determinar a estrutura e a seqüência de bases de um determinado gene. Além disso, havia a idéia de que seria impossível isolar e manipular segmentos específicos do DNA celular, e o entendimento de como os genes são controlados, como o embrião se desenvolve e como os cânceres surgem parecia ser necessariamente um objetivo para um futuro bastante distante. Tudo isso foi mudado com o desenvolvimento das técnicas do DNA recombinante (Capítulo 4). Como ilustram muitos capítulos deste livro, essas técnicas conduziram ao nosso entendimento da vida – uma abundância de conhecimentos que continua hoje e mostra sinais de que nunca acabará.

Em 1953, o primeiro objetivo de trabalho era descobrir como a informação genética era armazenada no DNA, e como essa informação era usada pelas células para sintetizar proteínas. Aqueles que esperavam que isso fosse um trabalho para toda uma vida ficaram surpresos quando todos estes aspectos foram decifrados em uma única década.

## Referências para Leitura

### Geral

Os três livros-texto indicados abaixo fornecem uma cobertura bem geral da genética moderna e da biologia molecular.

- Alberts B., Johnson A., Lewis J., Raff M., Roberts K., and Walter P. 2002. *Molecular Biology of the Cell*. Garland, New York.
- Lodish H., Berk A., Zipursky S.L., Matsudaira P., Baltimore D., and Darnell J. 2003. *Molecular Cell Biology*. W.H. Freeman, New York.
- Watson J.D., Baker T.A., Bell S.P., Gann A., Levine M., and Losick R. 2004. *Molecular Biology of the Gene*, 5th ed. Benjamin Cummings, San Francisco, em conjunto com Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

### História

A origem e o desenvolvimento da genética produzem uma história fascinante.

- Dunn L.C., ed. 1951. *Genetics in the 20th Century: Essays on the Progress of Genetics during Its First 50 Years*. Macmillan, New York. (Uma coleção de dosagens escrita por alguns fundadores da genética.)
- Sturtevant A.H. 1965. *A History of Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. (Reimpresso em 2001.)
- Judson H.F. 1980. *The Eighth Day of Creation: Makers of the Revolution in Biology*. Simon and Schuster, New York. (Edição expandida, reimpressa com um novo pós-escrito, Cold Spring Harbor Laboratory Press, 1996; uma narrativa da descoberta da dupla hélice do DNA e do desenvolvimento da genética molecular.)
- Carlson E. 2004. *Mendel's Legacy: The Origins of Classical Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

### Herança Mendeliana

- Mendel G. 1866. Versuche über Pflanzen-Hybriden. *Verh. Naturforsch. Ver. Brunn* 4: 3-47. (Esta é uma citação do trabalho original de Mendel. Versão alemã e inglesa, comentários com referências a grandes textos poderão ser vistos na página de Mendel: <http://www.mendelweb.org/>)
- Olby R.C. 1966. *Origins of Mendelism*. Constable, London. (Uma discussão sobre estudos de hibridização antes de Mendel.)
- Stern C. and Sherwood E.R., eds. 1966. *The Origin of Genetics*. W.H. Freeman, San Francisco. (Tradução inglesa do trabalho de Mendel, junto com interessantes documentos.)

### A Teoria Cromossomal da Hereditariedade

- Sutton W.S. 1902. On the morphology of the chromosome group in *Brachystola magna*. *Biol. Bull.* 4: 124-139.
- Stevens N.M. 1905. Studies in spermatogenesis with especial reference to the "accessory" (sex) chromosome. *Carnegie Inst. Wash. Publ.* 36: 1-33.
- Wilson E.B. 1928. *The Cell in Development and Heredity*. Macmillan, New York. (Reimpresso, Garland Publishing, New York, 1987; uma revisão clássica de citogenética.)
- Voeller B.R., ed. 1968. *The Chromosome Theory of Inheritance: Classic Papers in Development and Heredity*. Appleton-Century-Crofts, New York. (Citações de trabalhos clássicos.)

### Ligação

- Morgan T.H. 1910. Sex-limited inheritance in *Drosophila*. *Science* 32: 120-122.
- Sturtevant A.H. 1913. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.* 14: 43-59.
- Crow J.F. 1988. A diamond anniversary: The first chromosome map. *Genetics* 118: 1-3.
- Allen G.E. 1978. *Thomas Hunt Morgan: The Man and His Science*. Princeton University Press, Princeton, New Jersey. (Uma descrição detalhada de Morgan e sua pesquisa.)
- Kohler R.E. 1994. *Lords of the Fly: Drosophila Genetics and the Experimental Life*. University of Chicago Press, Chicago. (Uma descrição interessante de como a *Drosophila* se tornou um organismo de escolha para a análise genética.)

### Mutações e Rastreamento Genético

- Muller H.J. 1927. Artificial transmutation of the gene. *Science* 46: 84-87.
- Stadler L.J. 1928. Mutations in barley induced by X-rays and radium. *Science* 68: 186-187.
- Nüsslein-Volhard C. and Wieschaus E. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287: 795-801.
- Driever W., Stemple D., Schier A., and Solnica-Krezel L. 1994. Zebra fish: Genetic tools for studying vertebrate development (review). *Trends Genet.* 10: 152-158.
- Carlson E.A. 1981. *Genes, Radiation, and Society: The Life and Work of H.J. Muller*. Cornell University Press, Ithaca, New York. (A bibliografia de um cientista complexo.)

### Genes Extranucleares

- Correns C. 1909. Vererbungsversuche mit blaß(gelb) grünen und buntblättrigen sippem bei *Mirabilis jalapa*, *Urtica*

- pilulifera* und *Lunaria annua*. *Z. Indukt. Abstammungs-Vererbungslehre*. 1: 291-329.
- Gray M.W., Burger G., and Lang B.F. 1999. Mitochondrial evolution. *Science* 283: 1476-1481.
- McFadden G.I. 2001. Chloroplast origin and integration. *Plant Physiol.* 125: 50-53.
- Sapp J. 1987. *Beyond the Gene: Cytoplasmic Inheritance and the Struggle for Authority in Genetics*. Oxford University Press, New York. (Uma descrição da pesquisa da herança citoplasmática com uma interessante análise sociológica.)

### O DNA como o Material Genético Primário

- Griffith F. 1928. The significance of *Pneumococcal* types. *J. Hyg.* 27: 13-59.
- Avery O.T., MacLeod C.M., and MacCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of *Pneumococcal* types. *J. Exp. Med.* 79: 137-158. (O DNA como o material genético.)
- Hershey A.D. and Chase M. 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* 36: 39-56. (Um experimento clássico.)
- Cairns J., Stent G.S., and Watson J.D., eds. 1966. *Phage and the Origins of Molecular Biology*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. (Uma coleção de ensaios em homenagem a Max Delbrück; reimpresso em 1992 com uma nova introdução e materiais adicionais.)
- McCarty M. 1985. *The Transforming Principle: Discovering That Genes Are Made of DNA*. Norton, New York. (Uma narrativa da descoberta por um dos participantes.)
- Portugal F.H. and Cohen J.S. 1977. *A Century of DNA: A History of the Discovery of the Structure and Function of the Genetic Substance*. MIT Press, Cambridge, Massachusetts. (A história começando com os estudos prévios de Miescher, Kossel e Levene.)

### A Química e o Tamanho do DNA

- Astbury W.T. and Bell F.O. 1938. X-ray study of thymonucleic acid. *Nature* 141: 747-748. (O primeiro modelo para a estrutura do DNA.)
- Chargaff E. 1950. Chemical specificity of nucleic acids and the mechanism of their enzymatic degradation. *Experientia* 6: 201-209. (Chargaff chama a atenção para a relação de bases.)
- Brown D.M. and Todd A.R. 1952. Nucleotides. Part X. Some observations on structure and chemical behaviour of the nucleic acids. *J. Chem. Soc.* 52-58.
- Williams R.C. 1952. Electron microscopy of sodium desoxyribonucleate by use of a new freeze-drying method. *Biochim. Biophys. Acta* 9: 237-239.
- Dekker C.A., Michaelson A.M., and Todd A.R. 1953. Nucleotides. Part XIX. Pyrimidine deoxyribonucleoside diphosphates. *J. Chem. Soc.* 947-951.
- Wyatt G.R. and Cohen S.S. 1953. The bases of the nucleic acids of some bacterial and animal viruses: The occurrence of 5-hydroxymethylcytosine. *Biochem. J.* 55: 774-782.

### A Dupla Hélice

- Franklin R.E. and Gosling R.G. 1953. Molecular configuration in sodium thymonucleate. *Nature* 171: 740-741.
- Watson J.D. and Crick F.H.C. 1953. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* 171: 737-738.
- \_\_\_\_\_. 1953. Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171: 964-967.
- Wilkins M.H.F., Stokes A.R., and Wilson H.R. 1953. Molecular structure of deoxypentose nucleic acids. *Nature* 171: 738-740.
- Chambers D.A., ed. 1995. DNA: The double helix—Perspective and prospective at forty years. *Ann. N.Y. Acad. Sci.* 758: 1-472. (Narrativas de pessoas importantes no desenvolvimento da biologia molecular, incluindo publicações dos três trabalhos originais do *Nature*.)
- Watson J.D. 1968. *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. Atheneum, New York. (Reimpresso como *The Double Helix* [Norton critical edition, ed. G.S. Stent, Norton, New York, 1980], com várias revisões que fornecem diferentes perspectivas.)
- Crick F.C. 1988. *What Mad Pursuit: A Personal View of Scientific Discovery*. Basic Books, New York. (Uma autobiografia descrevendo toda a vida de Crick até o presente.)
- Maddox B. 2002. *Rosalind Franklin: The Dark Lady of DNA*. HarperCollins, New York.
- Wilkins M. 2003. *The Third Man of the Double Helix: The Autobiography of Maurice Wilkins*. Oxford University Press, Oxford.
- Ridley M. 2006. *Francis Crick: Discover of the Genetic Code*. HarperCollins, London.
- Olby R. 1994. *The Path to Double Helix: The Discovery of DNA*. Dover Publications, New York. (Uma narrativa detalhada da descoberta.)

### Replicação do DNA

- Pauling L. and Delbrück M. 1940. The nature of the intermolecular forces operative in biological processes. *Science* 92: 77-79.
- Meselson M. and Stahl F.W. 1958. The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci.* 44: 671-682.
- Kornberg A., Lehman I.R., Bessman M.J., and Simms E.S. 1956. Enzymatic synthesis of deoxyribonucleic acid. *Biochim. Biophys. Acta* 21: 197-198.

- Lehman I.R., Zimmerman S.B., Adler J., Bessman M.J., Simms E.S., and Kornberg A. 1958. Enzymatic synthesis of deoxyribonucleic acid. V. Chemical composition of enzymatically synthesized deoxyribonucleic acid. *Proc. Natl. Acad. Sci.* **44**: 1191-1196.
- De Lucia P. and Cairns J. 1969. Isolation of an *E. coli* strain with a mutation affecting DNA polymerase. *Nature* **224**: 1164-1166.
- Kornberg T. and Gefter M.L. 1970. DNA synthesis in cell-free extracts of a DNA polymerase-defective mutant. *Biochem. Biophys. Res. Commun.* **40**: 1348-1355. (Esta é a DNA polimerase II, outra DNA polimerase de reparo.)
- . 1971. Purification and DNA synthesis in cell-free extracts: Properties of DNA polymerase II. *Proc. Natl. Acad. Sci.* **68**: 761-764. (Esta foi chamada de polimerase III de *E. coli*, a enzima de replicação do DNA.)
- Friedberg E.L. 2006. The eureka enzyme: The discovery of DNA polymerase. *Nat. Rev. Mol. Cell Biol.* **7**: 143-147.
- Okazaki R., Okazaki T., Sakabe K., Sugimoto K., and Sugino A. 1968. Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc. Natl. Acad. Sci.* **59**: 598-605.
- Kornberg A. and Baker T.A. 1992. *DNA Replication*. W.H. Freeman, San Francisco.
- Brush G.S. and T.J. Kelly. 1996. Mechanisms for replicating DNA. In *DNA Replication in Eukaryotic Cells* (ed. M.L. DePamphilis), pp. 1-43. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Desnaturação e Renaturação do DNA*
- Doty P., Marmur J., Eigner J., and Schildkraut C. 1960. Strand separation and specific recombination in deoxyribonucleic acids: Physical chemical studies. *Proc. Natl. Acad. Sci.* **46**: 461-476.
- Marmur J. and Lane L. 1960. Strand separation and specific recombination in deoxyribonucleic acids: Biological studies. *Proc. Natl. Acad. Sci.* **46**: 453-461.
- Hall B.D. and Spiegelman S. 1961. Sequence complementarity of T2-DNA and T2-specific RNA. *Proc. Natl. Acad. Sci.* **47**: 137-146.
- Marmur J. 1994. DNA strand separation, renaturation and hybridization. *Trends Biochem. Sci.* **19**: 343-346. (Uma narrativa histórica).
- Tamanho dos Cromossomos*
- Eukaryote genomes: The European Bioinformatics Institute, Project Ensembl: <http://www.ensembl.org/>
- Prokaryote and viral genomes: The Institute for Genome Research: [http://www.tigr.org/tigr-scripts/CMR2/CMR\\_Genomes.spl](http://www.tigr.org/tigr-scripts/CMR2/CMR_Genomes.spl)
- Viral genomes: The National Center for Biotechnology Information: <http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/viruses.html>
- Genomas Virais*
- Baltimore D. 1970. Viral RNA-dependent DNA polymerase. *Nature* **226**: 1209-1211.
- Temin H.M. and Mizutani S. 1970. Viral RNA-dependent DNA polymerase. *Nature* **226**: 1211-1213.
- Levine A.J. 1992. *Viruses*. Scientific American Library, New York.