

Walter S. Judd *University of Florida*
Christopher S. Campbell *University of Maine*
Elizabeth A. Kellogg *University of Missouri, St. Louis*
Peter F. Stevens *University of Missouri, St. Louis; Missouri Botanical Garden*
Michael J. Donoghue *Yale University*

Sistemática Vegetal

Um Enfoque Filogenético

3ª Edição

Consultoria, supervisão e revisão técnica desta edição:

Rodrigo B. Singer

Doutor em Biologia Vegetal pela Universidade Estadual de Campinas, SP.
Professor adjunto do Departamento de Botânica do Instituto de Biociências da
Universidade Federal do Rio Grande do Sul.

Rosana Farias Singer

Doutora em Biologia Vegetal pela Universidade Estadual de Campinas, SP.



2009

Métodos e Princípios de Sistemática Biológica

2



A sistemática biológica (ou taxonomia) consiste na teoria e na prática de agrupar indivíduos em espécies, organizar tais espécies em conjuntos maiores e dar nomes a esses grupos, conseqüentemente gerando aquilo que é conhecido como uma **classificação**. As classificações são utilizadas para organizar as informações sobre os vegetais e, assim, construir chaves de classificação para a identificação desses organismos.

Existem diversas maneiras de construir uma classificação. Por exemplo, as plantas podem ser classificadas de acordo com suas propriedades medicinais (como ocorre em determinados sistemas de medicina herbalística) ou de acordo com seus habitats preferenciais (como em algumas classificações ecológicas). Uma classificação baseada na filogenia, como a utilizada no presente livro, tenta organizar os organismos em grupos baseados em suas relações evolutivas.

Existem duas etapas básicas que devem ser consideradas para tal classificação. A primeira consiste em determinar a **filogenia**, ou história evolutiva, de um grupo de organismos. A segunda etapa deverá basear a classificação do grupo levando em consideração sua história. As duas etapas podem ser desenvolvidas, e geralmente o são, separadamente, de tal forma que uma nova teoria de relacionamentos não deverá necessariamente levar a uma nova classificação. O presente capítulo salientará os procedimentos para determinação da história de um grupo e, a seguir, discutirá brevemente como podemos construir uma classificação a partir dessa história.

Como as filogenias são construídas?

Como descrito no Capítulo 1, a evolução não é apenas uma sucessão de descendentes modificados, mas também envolve a separação de linhagens. Este processo pode ser visualizado em diagramas como os apresentados nas Figuras 1.3 e 1.4, porém eles são de difícil interpretação. A história evolutiva pode ser mais facilmente sumarizada em diagramas ramificados (Figura 2.1A). (Alguns pesquisadores fazem distinções entre uma árvore evolutiva, uma filogenia e um

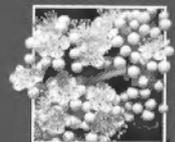


FIGURA 2.1 (A) Um modo simples de redesenhar o padrão de modificações apresentado na Figura 1.4. São fornecidas descrições completas para cada um dos ancestrais e seus descendentes. (B) Uma forma mais simples de redesenhar a Figura 2.1A, ilustrando apenas as modificações que ocorreram em diversas linhagens.

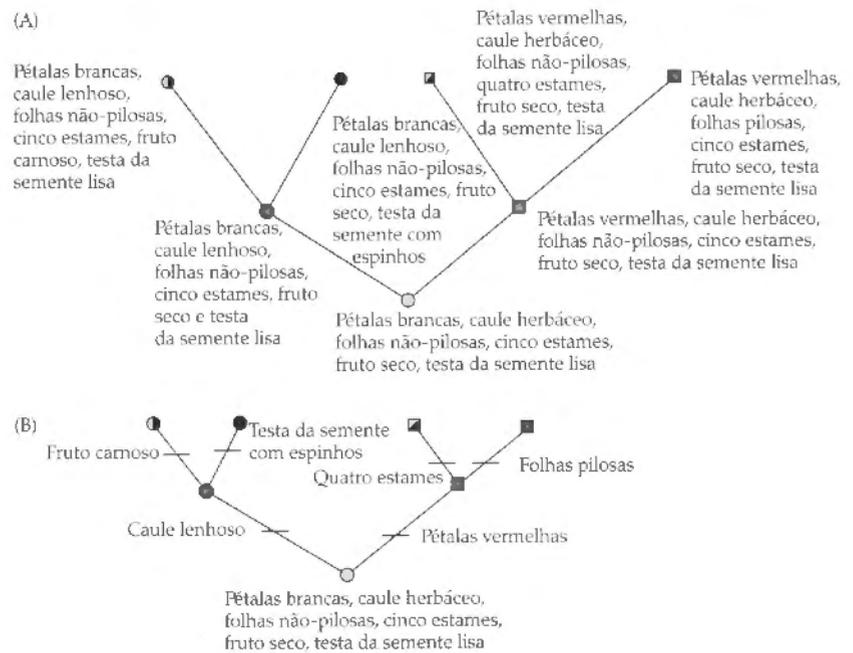


diagrama ramificado ou cladograma, mas no presente texto esses termos são usados como sinônimos). Para evitar a repetição de estados de caracteres ancestrais mantidos em cada grupo, os sistematistas geralmente anotam apenas os caracteres que sofreram modificação e posicionam marcas sobre os ramos apropriados para indicar a ordem relativa na qual os estados de caracteres se originaram (Figura 2.1B).

Os estados de caracteres derivados compartilhados na Figura 2.1B podem ser organizados hierarquicamente de mais inclusivos (p. ex., caule lenhoso ou pétalas vermelhas) para menos inclusivos (p. ex., folhas pilosas, testa da semente com espinhos). Esta organização leva à óbvia conclusão de que as próprias plantas podem ser organizadas sob uma classificação hierárquica que reflete sua história evolutiva. As plantas podem ser divididas em dois grupos: um grupo composto por aquelas que compartilham o estado de caráter derivado pétalas vermelhas e o estado de caráter ancestral de caule herbáceo, e o outro composto por plantas que compartilham o estado de caráter derivado caule lenhoso e o estado de caráter ancestral pétalas brancas. Cada um desses grupos pode, por sua vez, ser dividido em dois outros. Dessa forma, a classificação pode ser derivada diretamente a partir da filogenia.

Observe que a hierarquia não é alterada pela ordem na qual as extremidades dos ramos são desenhadas. A forma, ou **topologia**, da árvore é determinada apenas pelas conexões entre os ramos. Podemos contar a "história" evolutiva a partir de qualquer ponto da árvore, tanto para cima quanto para baixo. Isso nos mostra que os termos, *acima* ou *abaixo* não possuem um significado real, sendo simplesmente um reflexo de como escolhemos desenhar a árvore evolutiva.

A partir desse ponto de vista, um curso de sistemática vegetal pode tanto ser iniciado pelo estudo de Asteraceae, a qual é considerada uma família "derivada" por alguns livros-texto, e a seguir focar os outros integrantes do clado das Astéridae, como pode inicialmente focar as famílias "primitivas", como Magnoliaceae e Nymphaeaceae. Estas últimas famílias compartilham apenas um conjunto de caracteres considera-

dos ancestrais, no entanto estes também estão combinados com um grande número de caracteres derivados.

Determinando a história evolutiva

Nos exemplos apresentados nas Figuras 1.3, 1.4 e 2.1, descrevemos a evolução como se estivéssemos acompanhando o seu desenrolar. Naturalmente, isso raras vezes é possível, de tal forma que um dos desafios da sistemática é que ela deve *inferir* acontecimentos passados. O primeiro passo para poder fazer tais inferências consiste na análise de espécies existentes atualmente que apresentam proximidade em relação a caracteres que acreditamos serem herdáveis. Um **caráter herdável** é qualquer aspecto da morfologia da planta que pode ser transmitido geneticamente ao longo de um período de tempo evolutivo e que permaneça passível de reconhecimento. Por exemplo, sabe-se que a coloração das pétalas em plantas com flores, a estrutura da inflorescência e o hábito (padrão geral de crescimento) encontram-se sob controle genético, e que esses caracteres são geralmente herdados de forma estável de uma geração para a seguinte. Diversos exemplos de caracteres herdáveis são descritos nos Capítulos 4 e 5.

A sistemática envolve necessariamente a observação detalhada e precisa dos organismos. Na ausência de descrições minuciosas e cuidadosas dos caracteres, a reconstrução da filogenia e a descrição da história evolutiva se apresentarão destituídas de significado. Este tipo de classificação é impossível na ausência de uma morfologia comparativa exata. A avaliação da similaridade, em particular, é a base da biologia comparativa e da sistemática. No entanto, determinar quais estruturas de uma planta podem ser adequadamente comparadas a estruturas de uma outra planta através das similaridades pode ser bem mais difícil do que imaginamos inicialmente. Podemos considerar duas estruturas como similares se (1) elas encontram-se em posição similar em ambos os organismos, se (2) apresentam similaridade em nível de estrutura celular e histológica e/ou se (3) estão ligadas por

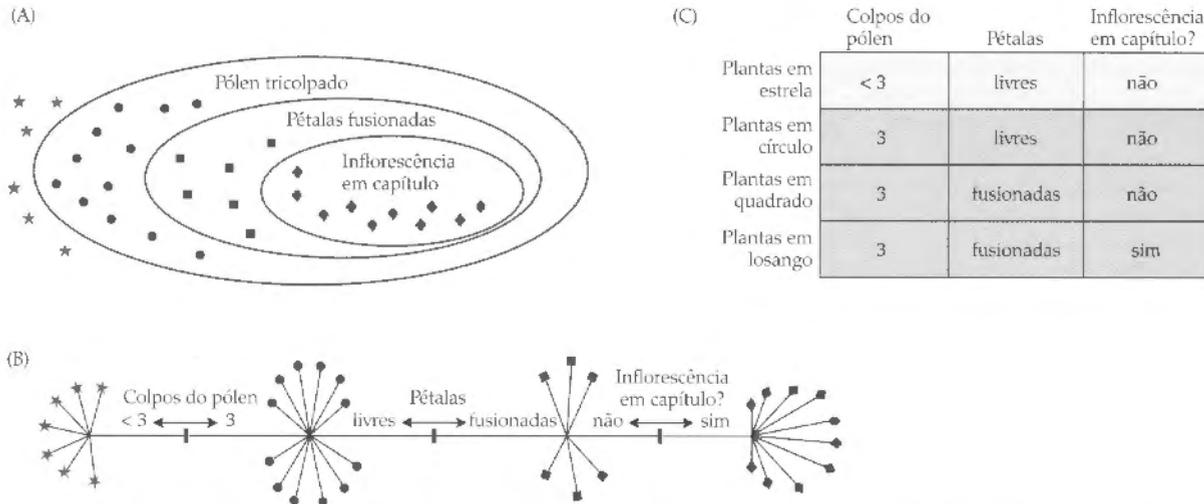


FIGURA 2.2 Cada símbolo representa uma espécie ou grupo de espécies de plantas hipotéticas produtoras de pólen. Um grande subgrupo dessas plantas apresenta pólen tricolpado. Do subgrupo, um grupo menor apresenta pétalas fusionadas e, das plantas com pólen tricolpado e pétalas fusionadas, um subgrupo apresenta as flores organizadas em inflorescências em capítulo. (A) O padrão descrito ilustrado

sob a forma de um diagrama de Venn com os estados de caracteres indicados. (B) O padrão redesenhado sob a forma de uma rede não-enraizada; os caracteres estão indicados com linhas de marcação verde forte, delimitando os diferentes estados de caracteres. (C) O padrão redesenhado sob a forma de uma matriz.

meio de formas intermediárias dessas estruturas (seja pela presença de intermediários em diferentes estágios do desenvolvimento de um mesmo organismo, seja por intermediários em organismos diferentes). Estas três afirmações constituem os **critérios de similaridade de Remane**.

Originalmente, Remane (1952) denominou esta lista de "critérios de homologia". Neste livro, no entanto, o termo *homologia* é usado em um sentido mais restrito, significando **identidade por meio de descendência**. Em outras palavras, se dissermos que um caráter é **homólogo** entre um grupo de espécies, estaremos dizendo que todas estas espécies herdaram tal caráter a partir de um ancestral comum. Sob essa definição, a observação de similaridade é apenas o primeiro passo na determinação da homologia, pois nem todas as similaridades observadas serão resultado de homologia* (p. ex., similaridades estruturais podem evoluir independentemente em plantas não relacionadas que vivem em ambientes semelhantes). Este texto segue o ponto de vista sustentado por muitos sistematistas filogenéticos que argumentam que a homologia só pode ser determinada por meio da construção de uma árvore evolutiva.

Caracteres, estados de caracteres e redes

Grupos de plantas que compartilham determinados estados de caracteres podem ser identificados mediante a observação de caracteres herdáveis. Suponha, por exemplo, que se observe diversidade no número de fendas (um caráter) na superfície do pólen e que o pólen de um grande número de espécies vegetais possua três fendas (um estado de caráter). Essas fendas são na realidade canais de germinação denominados *colpos*, e o pólen que possui três fendas é denominado *tricolpado*.

Dentro do grande grupo de espécies vegetais com pólen tricolpado, existe um grupo menor cujas pétalas (caráter) são fusionadas (estado de caráter) e, no interior desse grupo de pétalas fusionadas, é possível ainda observar um grupo com flores organizadas em uma inflorescência do tipo capítulo. Esses grupos, inseridos uns nos outros, podem ser representados como uma série de estruturas ovais concêntricas em um **diagrama de Venn**, como ilustrado na Figura 2.2A.

A informação no diagrama de Venn pode também ser representada sob a forma de uma **rede** (Figura 2.2B). Aqui os caracteres estão representados sob a forma de linhas verdes verticais, ou "marcas" (uma convenção que é vista ao longo das ilustrações de todo este texto). Ao passo que as formas (espécies) posicionadas à esquerda da linha "pólen" apresentam menos de três colpos, aquelas localizadas à direita da linha possuem pólen tricolpado. De forma semelhante, a linha "pétalas" indica uma mudança entre os estados de caráter livre e fusionado e a linha de inflorescência indica uma mudança entre flores organizadas em inflorescência e flores isoladas. Podemos contar o número de modificações ao longo da rede para determinar seu *comprimento*: da direita para a esquerda, existem modificações para inflorescência, para pétala fusionada e para colpos no pólen, de tal forma que a rede pode ser descrita como apresentando um comprimento de três modificações.

A mesma informação pode ser representada sob a forma de uma **matriz** na qual as linhas correspondem a plantas, e as colunas, a caracteres (Figura 2.2C). Os estados de caracteres são então usados para preencher a matriz. Modificações nos estados de caracteres são, ou hipotetiza-se que sejam, modificações genéticas que potencialmente distinguem grupos de plantas na matriz. Dessa forma, as três modificações na rede da Figura 2.2B representam três modificações na seqüência gênica (e conseqüentemente nas proteínas resultantes), o que altera os estados de caracteres de algumas plantas.

Na Figura 2.2, todas as plantas designadas pela mesma forma estão ilustradas como se tivessem se originado simul-

* Você deve estar ciente de que a palavra *homologia* possui diferentes significados e que, ao ler um texto, é importante conferir sempre o significado específico que o autor em questão confere a esse termo.

taneamente. Este arranjo geralmente indica ambigüidade; por uma questão de simplificação do exemplo, não foi fornecida qualquer informação a respeito da ordem de origem evolutiva das plantas. Além disso, pressupomos que a determinação dos diferentes estados de caracteres era absolutamente clara. No entanto, geralmente essa não é a realidade. Quando descrevemos a variação entre estruturas morfológicas similares mediate a divisão do caráter em estados de caráter, estamos na verdade extrapolando uma hipótese de controle genético subjacente, mesmo considerando que raramente basearemos o pressuposto nestes termos.

Por exemplo, se duas espécies diferem na coloração de suas flores, podemos categorizar o caráter "cor das pétalas" como possuidor de dois estados, vermelho e azul. Através de tal categorização, estamos formando a hipótese de que genes que definem a coloração das pétalas foram modificados, ao longo de um período evolutivo, para a produção de flores vermelhas a partir de um ancestral que possuía flores azuis ou para a produção de flores azuis a partir de um ancestral com flores vermelhas. Neste contexto, sabemos que de fato existem genes (p. ex., envolvidos na via das antocianinas) que controlam a coloração das pétalas, e assim a inferência de dois estados controlados por uma "modificação genética" é uma possibilidade viável. No entanto, em diversos casos, não possuímos qualquer idéia relativa aos mecanismos genéticos que controlam o estado dos caracteres estruturais observados. Ao propormos hipóteses sobre a natureza dos mecanismos relacionados às modificações, freqüentemente o máximo de certeza que podemos ter é que os estados de caráter são realmente distintos. No caso de caracteres quantitativos, como o comprimento das folhas ou o diâmetro do tubo da corola, isto significa determinar os dados quantitativos (p. ex., realizar as medidas) para certificar-se de que as medidas das espécies que estamos estudando não apresentam sobreposição.

Em diversos caracteres, estas medidas não somente apresentam sobreposição como também apresentam alta variabilidade, de tal forma que o pressuposto de existência de modificações genéticas subjacentes – e consequentemente a divisão em estados de caráter – não é apoiado por qualquer evidência. Nesses casos, os caracteres em questão devem ser omitidos de qualquer análise filogenética (a menos que a sobreposição seja causada por um número reduzido de indivíduos, caso em que o caráter poderá ser categorizado como polimórfico para esta espécie e mantido na análise). Mesmo considerando que tais caracteres sobrepostos provavelmente reflitam modificações genéticas ao longo de um tempo evolutivo, dado nosso atual conhecimento, a sobreposição torna difícil a obtenção de informações confiáveis referentes a modificações gênicas subjacentes (mesmo considerando-se que métodos que podem ser utilizados em plantas com caracteres variáveis já foram desenvolvidos).

A variabilidade e a sobreposição de caracteres morfológicos representam boas razões para o fato de diversos sistematistas terem se direcionado para o uso de dados moleculares na construção de filogenias. Com o surgimento de dados de seqüência nucleotídica de diferentes genes, o reconhecimento dos estados de caracteres moleculares (p. ex., se o nucleotídeo presente em uma determinada posição é A, T, G ou C) é geralmente mais exato. No entanto, isto não é sempre verdade quando as seqüências gênicas são de difícil alinhamento ou se os fragmentos de restrição apresentam um tamanho muito semelhante. O uso de estados de caracteres moleculares em sistematia vegetal é abordado detalhadamente no Capítulo 5.

Árvores evolutivas e enraizamento

A Figura 2.2 ilustra três diferentes formas de representação e organização de observações feitas em vegetais. Mesmo que se considere que a rede (Figura 2.2B) se assemelhe de certa forma a uma linha de tempo, ela não corresponde a tal representação. A rede pode ser lida da esquerda para a direita, da direita para a esquerda ou, dependendo do caso, do centro para as extremidades. Para transformar uma rede em uma árvore evolutiva, faz-se necessário determinar quais modificações são relativamente mais recentes e quais ocorreram mais distantes temporalmente. Em outras palavras, a árvore deve ser **enraizada**. O enraizamento *polariza* as modificações nos caracteres, dando a elas uma direção específica.

Se você imagina que uma rede é um pedaço de uma fita, poderá manter as conexões exatamente iguais, mesmo que determine o enraizamento em diferentes locais. A rede apresentada na Figura 2.2B foi redesenhada na Figura 2.3, mas com enraizamento em três diferentes pontos. Observe que o comprimento de cada árvore (ou cladograma) é igual ao comprimento da rede original – 3 – e que todas as conexões são as mesmas, apesar de a ordem dos eventos de modificação de caracteres poder diferir consideravelmente.

Por exemplo, no enraizamento ilustrado na Figura 2.3A, as plantas ancestrais possuíam pólen com menos de três colpos, pétalas não-fusionadas e flores isoladas, ao passo que podemos concluir, a partir da Figura 2.3B, que as plantas ancestrais apresentavam os estados de caráter exatamente opostos. Na Figura 2.3C, a árvore está enraizada de tal modo que o ancestral possuía pólen tricolpado. Mais tarde, o pólen foi alterado de forma a conter menos de três colpos em uma linhagem, enquanto outra linhagem manteve o estado de caráter do pólen em três colpos e posteriormente adquiriu pétalas fusionadas e flores em inflorescências.

O enraizamento de uma árvore filogenética é crítico para a interpretação de como ocorreu a evolução vegetal, e diferentes enraizamentos sugerem diferentes padrões de modificação (diferentes polarizações de caracteres). Muita discussão já ocorreu entre os sistematistas em relação a como deve ser determinada a posição da raiz. Uma sugestão freqüente consiste no uso de fósseis. No entanto, o simples fato de que uma planta extinta tenha sofrido fossilização não significa que sua linhagem tenha se *originado* antes das linhagens referentes às plantas atuais; a única certeza que temos é que ela morreu antes.

Ao determinarmos a história evolutiva, estamos interessados em determinar quando linhagens divergiram umas das outras (ou seja, quando os *taxa* se originaram). É interessante saber quando um táxon desapareceu, mas este fato *per se* não nos auxilia no estabelecimento de suas origens. (Obviamente, os fósseis são extremamente úteis quando incluídos como *taxa* adicionais em uma filogenia. Freqüentemente eles apresentam combinações de estados de caráter que não mais ocorrem em *taxa* atuais e podem afetar a estrutura geral de uma árvore, algumas vezes de forma surpreendente e informativa.)

Em geral, as árvores evolutivas são enraizadas mediante o uso de um organismo aparentado ao grupo que está sendo estudado: um **grupo-externo**. Quando selecionamos um grupo-externo, devemos assumir apenas que todos os integrantes do grupo-interno (integrantes do grupo que está sob estudo) estejam mais intimamente relacionados entre eles do que com o grupo-externo; em outras palavras, o grupo-externo deve ter se separado da linhagem do grupo-interno antes da diversificação deste. Geralmente, vários grupos-externos são utili-

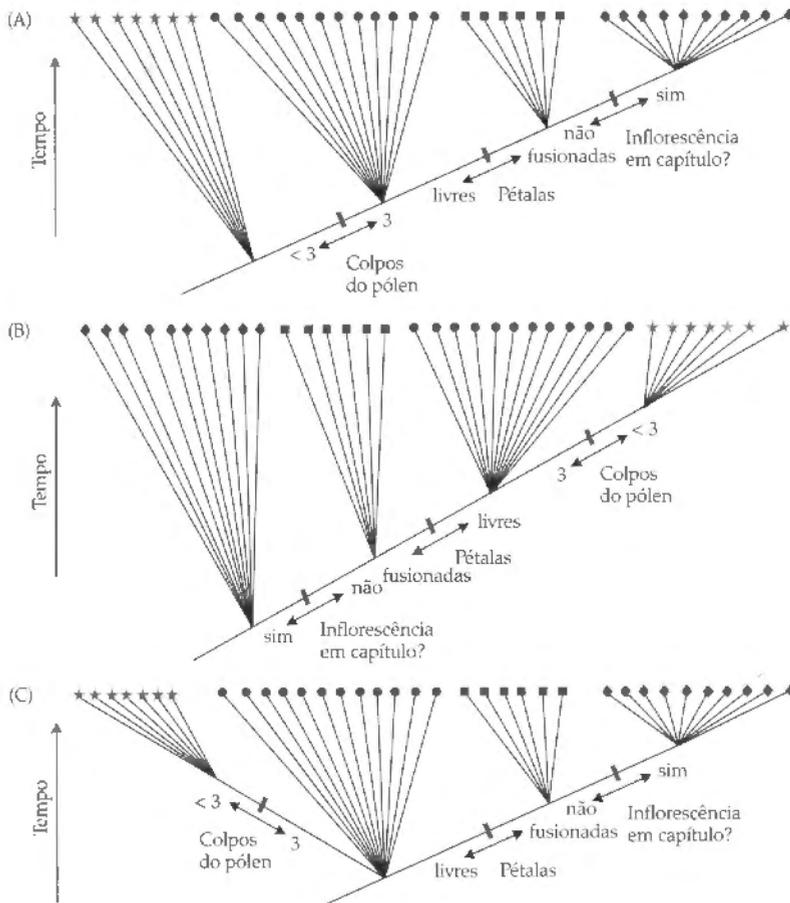


FIGURA 2.3 Três possíveis enraizamentos para a rede da Figura 2.2B. Observe que, em cada caso o número de passos evolutivos (modificações de estados de caracteres) é o mesmo que o apresentado na rede não-enraizada.

zados. Se um grupo-externo é adicionado a uma rede, o ponto no qual ele se posiciona é definido como a raiz da árvore.

No caso das Figuras 2.2 e 2.3, todas as plantas ilustradas são plantas com flores (angiospermas), e seus parentes atuais mais próximos são as coníferas, Cycadales, Gnetales, ginkgos, ou um grupo destes (ver Capítulos 7 e 8). Na Figura 2.4A, uma conífera foi adicionada à matriz da Figura 2.2C. (Poderíamos ter utilizado todas as gimnospermas como grupos-externos, mas, para manter um exemplo simplificado, foi escolhida apenas uma).

Visto que coníferas não possuem pétalas ou flores, dois dos caracteres devem ser categorizados como não passíveis de aplicação, mas sabemos que o pólen de coníferas não possui três colpos. Com esta informação, a conífera pode ser adicionada à rede como grupo-externo, como na Figura 2.4B. Uma vez que a conífera se liga entre as espécies em estrela, a árvore pode ser enraizada e redesenhada como na Figura 2.4C. Essa árvore corresponde à árvore enraizada na Figura 2.3A e fortalece a hipótese de que a Figura 2.3A reflete com precisão a história evolutiva.

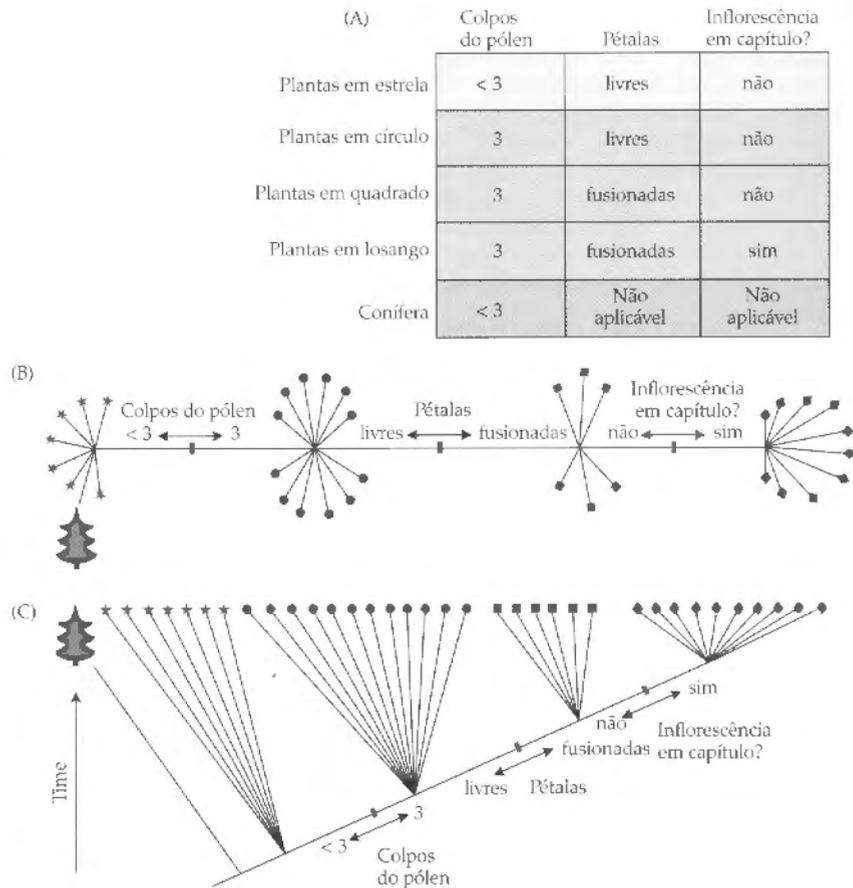
Observe que a árvore pode ser desenhada de diversas maneiras e ainda assim refletir a mesma história evolutiva. Através da comparação entre as Figuras 2.5A e B e a Figura 2.4C mostramos que podemos girar os braços da árvore em torno de qualquer dos pontos de ramificação (nós) sem que a ordem inferida dos eventos seja afetada.

Em uma árvore enraizada (e somente em uma árvore enraizada), podemos determinar quais grupos são monofiléticos

(compostos por um ancestral e todos os seus descendentes). Dessa forma, no exemplo da Figura 2.4C, as plantas representadas por losangos são monofiléticas (i.e., formam um clado). De fato, as plantas com flores de pétalas fusionadas e arranjo floral em inflorescências em capítulo são da família Asteraceae, que é conhecida por formar um grupo monofilético. Assim, o fato de ter flores em inflorescências em capítulo é uma sinapomorfia (i.e., é um caráter derivado compartilhado, ou indica a monofilia) das Asteraceae, o caráter de pétalas fusionadas é um caráter derivado compartilhado (sinapomorfia) que une as espécies representadas em quadrado com as espécies representadas em losango e o fato de apresentar pólen tricolpado indica a monofilia do conjunto formado pelas plantas representadas em círculos mais quadrados mais losangos.

Observe quanto importante é o enraizamento para a determinação de monofilia. Se a Figura 2.3B fosse a opção de enraizamento correto da filogenia de plantas com flores, então pétalas fusionadas e flores em capítulo seriam os estados de caracteres ancestrais (geralmente denominados **simple-siomorfias**) ao invés de serem sinapomorfias. Neste caso, as espécies indicadas por losangos e quadrados não compartilhariam qualquer caráter *derivado* e não incluiriam todos os descendentes de seu ancestral comum; alguns desses descendentes tornaram-se as plantas representadas por círculos e estrelas. Assim, se a Figura 2.3B fosse correta, as espécies em quadrado e losango não representariam um grupo monofilético (como ocorre com o enraizamento na Figura 2.3A). Em

FIGURA 2.4 (A) A matriz da Figura 2.2C, mas com os estados de caracteres adicionados para uma conífera. (B) A rede não enraizada da Figura 2.2B acrescida de uma conífera, de acordo com os estados de caracteres apresentados na Figura 2.4A. (C) A rede da Figura 2.4B enraizada com a conífera. Observe que a história evolutiva é a mesma que a apresentada na Figura 2.3A.



vez disso, elas constituiriam um grupo **parafilético**, o qual inclui um ancestral comum e alguns, mas não todos, descendentes desse ancestral.

Como mencionado anteriormente, um estado de caráter que é derivado (sinapomórfico) em um dado momento pode tornar-se posteriormente ancestral. Na Figura 2.4B, pólen tricolpado é um caráter derivado compartilhado por um grande grupo de plantas com flores. Ele é uma sinapomorfia e indica a monofilia de um grupo às vezes denominado eudicotiledôneas. Em relação ao grupo com pétalas fusionadas, no entanto, pólen tricolpado é um estado de caráter ancestral, ou **plesiomórfico**. Ele corresponde a algo que todas as espécies no grupo herdaram de seu ancestral comum e, assim, não é capaz de dar indicações a respeito das relações entre os integrantes do grupo. Estados de caracteres plesiomórficos não são capazes de indicar relações evolutivas no grupo em estudo, pois evoluíram antes que qualquer dos *taxa* que estão sendo comparados e foram simplesmente mantidos nas diferentes linhagens do grupo.

Algumas vezes, a monofilia de um grupo é indicada pelo fato de que seus estados de caracteres não ocorrem em qualquer outro organismo. Por exemplo, todos os integrantes da família das gramíneas (*Poaceae*) possuem um embrião que difere do embrião de qualquer outra angiosperma. Podemos então gerar a hipótese de que o embrião das gramíneas é exclusivamente derivado constituindo uma sinapomorfia de *Poaceae* e indica que a família é monofilética. Isto equivale a dizer que qualquer enraizamento lógico da árvore filogenética levará à mesma conclusão.

Geralmente é possível encontrar evidências de que um grupo é monofilético mesmo sem o apoio de uma enorme análise filogenética baseada em computação. Na verdade, a maioria das análises filogenéticas (algumas vezes referidas como cladísticas) era realizada manualmente até a metade da década de 1980. Os caracteres são inicialmente divididos em estados de caráter, como em qualquer análise filogenética. A seguir, o estado de caráter do grupo-externo (ou grupos-externos) é assumido como sendo o ancestral (Stevens 1980; Watrous e Wheeler 1981; Maddison et al. 1984). Em outras palavras, cada caráter é polarizado, ou direcionado. O estado de caráter derivado compartilhado, ou sinapomórfico, pode, então, ser utilizado como evidência de monofilia, e pode-se construir um cladograma baseado nos estados de caráter sinapomórficos (Quadro 2A). Esse tipo de raciocínio geralmente é útil em uma primeira formulação de hipóteses, por exemplo, para testar se grupos taxonômicos existentes são monofiléticos e, dessa maneira, nomeá-los adequadamente.

Escolhendo árvores

Como foi demonstrado nas discussões anteriores, a determinação da história evolutiva de um grupo de organismos é conceitualmente bem simples. Inicialmente, caracteres são observados e divididos em estados de caráter. A seguir, a partir desses estados, um diagrama de Venn (ver Figura 2.2A), uma matriz de caracteres \times *taxa* (ver Figura 2.2C) ou uma rede ramificada (ver Figura 2.2B), podem ser construídos. Posteriormente, mediante a inclusão de um grupo-externo, a rede

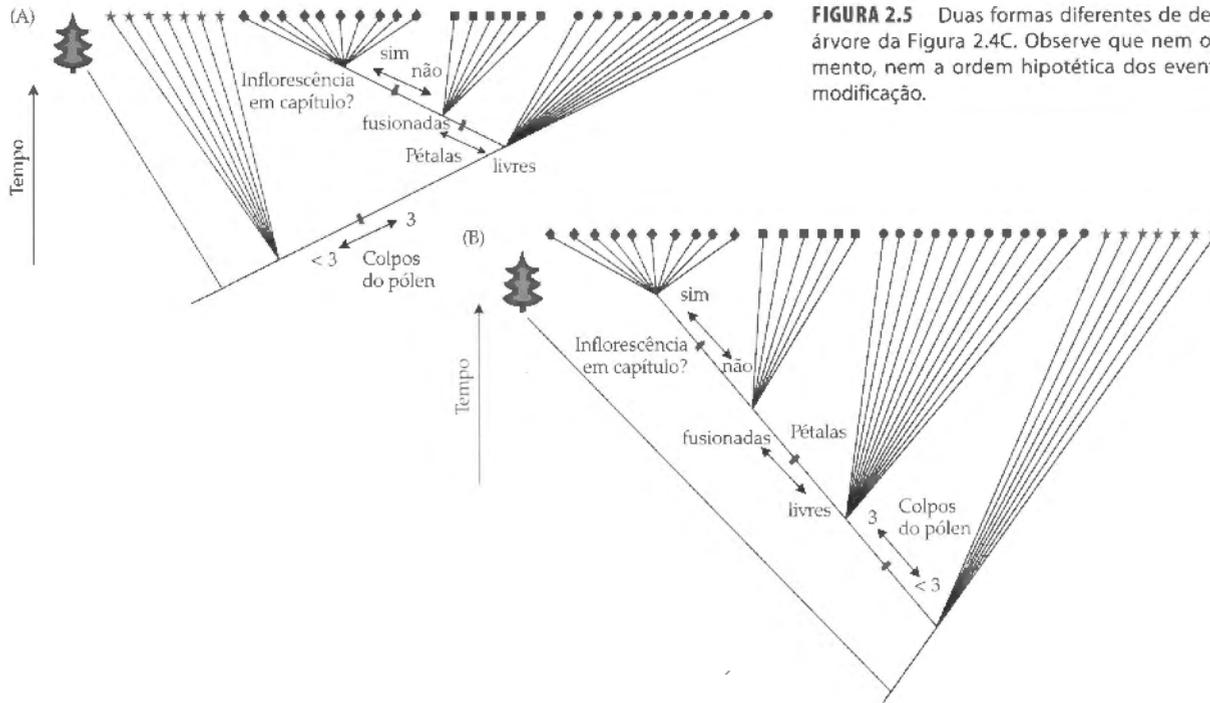


FIGURA 2.5 Duas formas diferentes de desenhar a árvore da Figura 2.4C. Observe que nem o comprimento, nem a ordem hipotética dos eventos sofre modificação.

pode ser enraizada para a produção de uma árvore evolutiva, um cladograma ou uma filogenia.

Dois fenômenos, no entanto, tornam a prática da determinação da história evolutiva uma tarefa muito mais difícil: paralelismo e reversão, que algumas vezes são referidos em conjunto como **homoplasia**. O **paralelismo** é a ocorrência de estados de caráter similares em organismos não relacionados. (Vários autores fazem distinção entre paralelismo e convergência, mas para a presente discussão trataremos ambos como equivalentes.) Uma **reversão** ocorre quando um estado de caráter derivado é revertido para o estado ancestral.

Para fornecer um exemplo claro, vamos dividir o grupo que havíamos denominado "plantas em estrela" em plantas em estrelas vermelhas, plantas em estrelas douradas e plantas em estrelas brancas. Assumiremos que as plantas em estrelas douradas e as brancas possuem um único cotilédone, ao passo que as demais plantas possuem mais de um (incluindo a conífera). Vamos ainda assumir que as plantas em estrelas brancas possuem pétalas fusionadas. Podemos adicionar o caráter número de cotilédones à matriz da Figura 2.4A para criar a matriz da Figura 2.8A, a qual fornecerá a mesma informação que a rede apresentada na Figura 2.8B.

Agora, vemos que, de acordo com essa rede, ocorreram *duas* modificações paralelas na fusão de pétalas. Contando o número de modificações nessa rede (seu comprimento), chegamos a cinco: um para colpos do pólen, um para flores em inflorescências em capítulo, um para número de cotilédones e dois para fusão de pétalas.

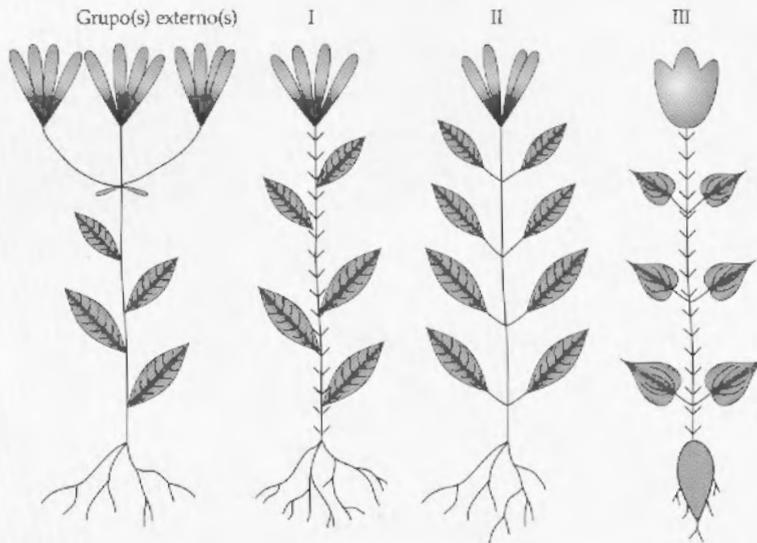
Nesse exemplo, um grupo baseado na fusão de pétalas seria considerado **polifilético**. Grupos polifiléticos possuem duas ou mais linhagens ancestrais nas quais estados de caracteres se desenvolveram paralelamente. (Apesar de distinguirmos aqui grupos parafiléticos de grupos polifiléticos, muitos sistematas têm observado que a diferença é tênue e simplesmente denominam tanto grupos parafiléticos quanto

polifiléticos como grupos não-monofiléticos.) A fusão de pétalas nesse caso não é homóloga, pois não é capaz de passar no teste definitivo de homologia: congruência com outros caracteres em uma análise filogenética.

Por que não desenhamos a rede de tal forma que a fusão de pétalas tenha-se originado apenas uma vez? Uma rede assim montada está ilustrada na Figura 2.8C. Nesta configuração temos apenas uma modificação no padrão de fusão de pétalas, mas são necessárias duas modificações em número de cotilédones e também duas modificações no número de colpos do pólen, o que gera uma rede de seis passos de comprimento.

Cada uma das redes pode ser convertida em uma filogenia por meio do enraizamento mediado pela conífera, mas as filogenias darão diferentes sugestões sobre o modo de evolução das plantas. Na Figura 2.8B, o número de cotilédones e o número de colpos do pólen apresentaram-se estáveis ao longo do tempo evolutivo, ao passo que a fusão de pétalas surgiu duas vezes, independentemente. Na Figura 2.8C, postulamos que tanto o número de cotilédones quanto o número de colpos do pólen sofreram modificação duas vezes ao longo do período evolutivo, ao passo que a fusão de pétalas evoluiu apenas uma vez. Mediante o desenho de qualquer uma dessas redes, seremos capazes de propor uma hipótese acerca de como ocorreu a evolução – ou seja, que modificações genéticas ocorreram, com que frequência e em que ordem.

Como ambas as redes mostram, as duas hipóteses são diferentes. Assim, como determinar qual das duas é a hipótese correta? Não existe um caminho que dê uma resposta inequívoca. A evolução dessas plantas não foi presenciada por qualquer um de nós. Podemos, no entanto, propor, e algumas propostas parecem ser mais prováveis e corretas do que outras. Um dos caminhos possíveis a serem seguidos começa com a questão, "Qual é a explicação mais simples para essas observações?" Por meio dessa questão aplicamos uma regra que é amplamente usada na ciência, conhecida como **nava-**

QUADRO 2A O método Hennigiano**FIGURA 2.6** Três espécies imaginárias (I, II e III) e um grupo-externo.

Nos exemplos apresentados até aqui, uma rede é construída e a seguir é polarizada por meio da determinação do local de conexão do grupo-externo. No entanto, alguns sistematistas preferem polarizar inicialmente os caracteres mediante o uso de um ou mais grupos-externos e, a seguir, construir a filogenia. Isso nos leva ao conceito original de análise filogenética proposto por Willi Hennig (ver Capítulo 3).

Considere, por exemplo, as plantas hipotéticas apresentadas na Figura 2.6. Neste caso, assume-se que os estados de caráter do grupo-externo sejam ancestrais (plesiomórficos) e que são representados por 0; estados de caráter derivados são representados pelo numeral 1 ou por números maiores (Tabela 2.1). Na seqüência, tais estados de caráter são usados para gerar uma matriz de caráter \times táxon (Tabela 2.2).

A seguir, é construída uma árvore filogenética (ou cladograma) na qual os *taxa* serão agrupados (posicionados sobre o

TABELA 2.1 Estados de caráter morfológicos usados na análise cladística das três espécies imaginárias da Figura 2.6

Caráter morfológico	Estado de caráter ^a	
	Plesiomórfico	Apomórfico
1. Raiz	Menos de 1 mm de espessura (0)	Mais de 5 mm de espessura (1)
2. Caule	Glabro (0)	Pubescente (1)
3. Folhas	Alternas (0)	Opostas (1)
4. Venação	Peninérvea (0)	Palmada (1)
5. Pecíolo	Ausente (0)	Presente (1)
6. Base da lâmina	Aguda (0)	Cordada (1)
7. Partes do perianto	4 (0)	3 (1)
8. Partes do perianto	Livres (0)	Fusionadas (1)
9. Flores ^b	Em grupos de 3 (0)	Solitárias (1)

^aA codificação do estado de caráter é dada entre parênteses.

^bObserve que a condição referente à inflorescência (flores solitárias *versus* flores em grupos de 3) só pode ser polarizada se forem utilizados grupos-externos adicionais.

Iha de Occam: não desenvolva uma hipótese mais complexa do que a necessária para explicar os dados. A aplicação deste princípio de simplicidade, ou **parcimônia**, nos conduz a preferir a menor rede. O fato de ela ser mais curta não a torna correta, no entanto ela representa a explicação mais simples para os dados.

No exemplo que apresentamos aqui, no qual existem poucos caracteres e pouca homoplasia, é fácil construir a menor rede capaz de ligar os organismos. Na maioria dos casos reais, no entanto, diversas redes são possíveis e não fica imediatamente óbvia a solução que aponta qual delas será a mais curta. Felizmente, algoritmos computacionais foram desenvolvidos para comparar árvores e calcular seus comprimentos. Entre os

programas mais amplamente utilizados estão PHYLIP (Felsenstein 1989), NONA (Goloboff 1993) e PAUP*4.0 (Swofford 2000). Esses programas avaliam os dados a respeito de árvores possíveis (por meio de uma busca exaustiva) e geram propostas lógicas sobre a topologia das árvores mais curtas (buscas de *branch-and-bound* ou buscas heurísticas).

Se os *taxa* são numerosos, apenas algoritmos heurísticos podem ser usados. No entanto, esses algoritmos podem falhar na identificação da árvore, ou árvores, mais curta(s) tendo em vista o grande número de árvores possíveis. Por exemplo, as relações possíveis entre três *taxa* podem ser expressas por meio de apenas três árvores enraizadas [A(B,C)], [B(A,C)] e [C(A,B)]. No entanto, com um número maior de

TABELA 2.2 Matriz de caráter × táxon para as três espécies hipotéticas da Figura 2.6, baseada nos caracteres da Tabela 2.1

Taxa	Caracteres								
	1	2	3	4	5	6	7	8	9
Espécie I	0	1	0	0	0	0	0	0	1
Espécie II	0	0	1	0	1	0	0	0	1
Espécie III	1	1	1	1	1	1	1	1	1
Grupo(s) externo(s)	0	0	0	0	0	0	0	0	0

mesmo ramo) de acordo com as evidências fornecidas pelo compartilhamento de estados de caráter derivados (sinapomorfias). A presença de um estado de caráter derivado (apomorfia) em dois taxa sugere que eles compartilhem um único ancestral comum no qual houve a evolução inicial da apomorfia; assume-se que os dois taxa tenham herdado a apomorfia (ou novidade evolutiva) desse ancestral. Assim, seguindo o princípio da parcimônia, o cladograma representa a hipótese mais simples que pode explicar o padrão de estados de caráter derivados.

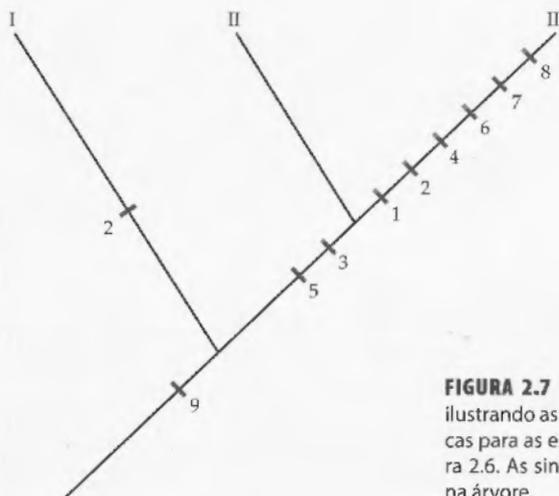
Uma hipótese sobre as relações evolutivas entre as espécies I, II e III da Figura 2.6 está apresentada na Figura 2.7. Parte-se da hipótese de que as espécies II e III compartilham um ancestral comum característico, pois elas compartilham os estados de caráter derivados 3 e 5 (ver Tabela 2.1). Ambas as espécies possuem folhas opostas e pecioladas, que foram consideradas como originadas em seu ancestral comum. De forma semelhante, a presença compartilhada de flores solitárias apóia o reconhecimento de um grupo monofilético mais inclusivo que contém as espécies I, II e III.

A presença de caules pilosos nas espécies I e III é homoplásica, ou seja, pressupõe-se que caules pilosos tenham evoluído em paralelo nestas duas espécies, de

tal forma que esta similaridade não estaria baseada no ancestral comum. Observe, no entanto, que caules pilosos podem ter evoluído em um ancestral comum mais recente dessas três espécies e então ter sido perdido (uma reversão) na espécie II.

O compartilhamento de folhas pinadas com bases agudas (formando um ângulo menor que 90°) e flores com o perianto em quatro partes separadas nas espécies I e II são simpliesimorfias; estas são

características ancestrais compartilhadas. Tais características não são informativas sobre as relações. Em contraste, venação palmada, folhas com base cordada (em forma de coração) e flores com perianto em três partes fusionadas são estados de caráter derivados característicos da espécie III. Esses estados de caráter derivados característicos (autapomorfias) também não são informativos no que diz respeito às relações filogenéticas da espécie III.

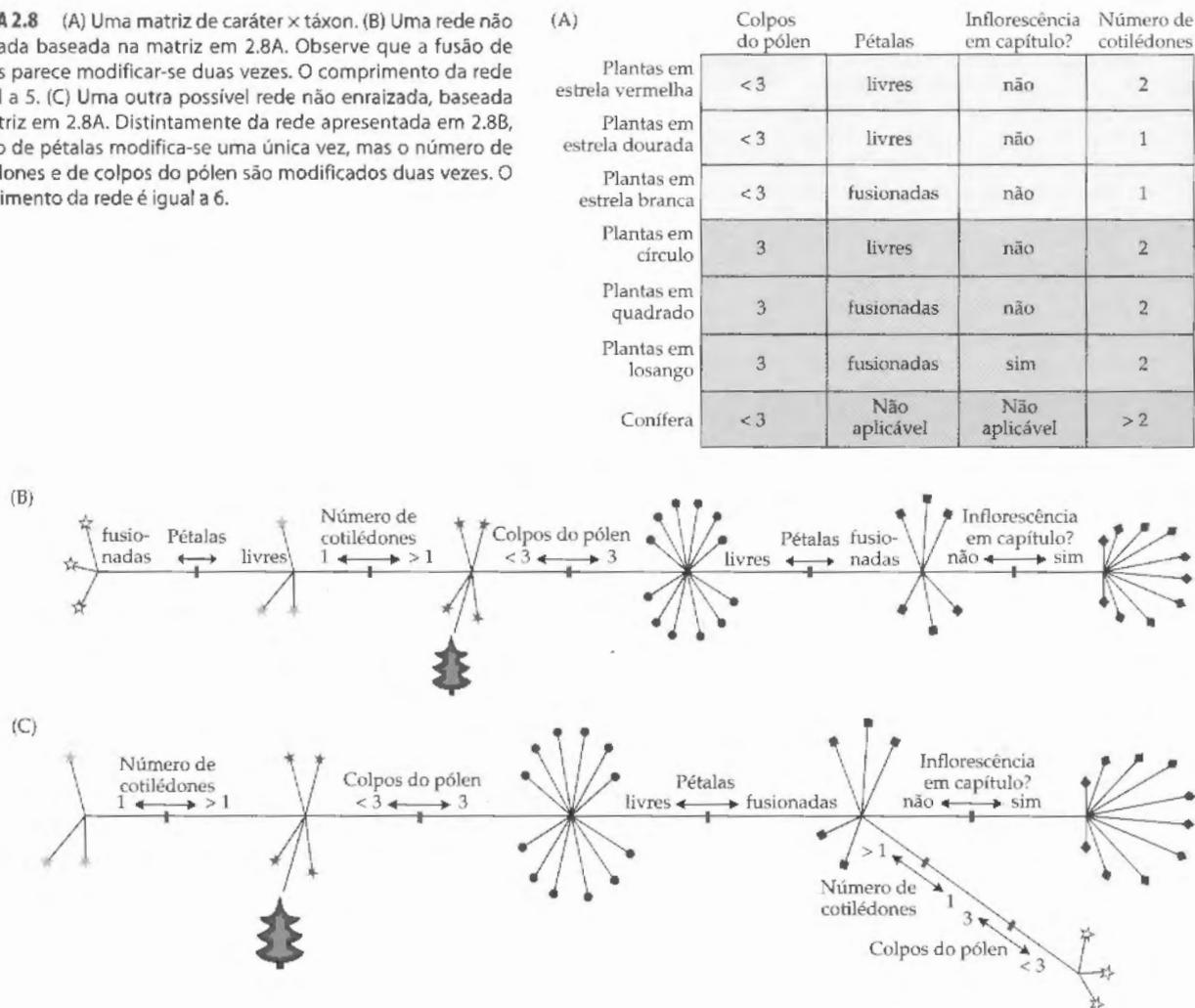
**FIGURA 2.7** Uma árvore filogenética ilustrando as relações evolutivas hipotéticas para as espécies imaginárias da Figura 2.6. As sinapomorfias estão indicadas na árvore.

taxa, o número potencial de árvores se expande rapidamente; por exemplo, quatro taxa geram 15 árvores, cinco dão origem a 105 árvores, seis podem gerar 945 árvores e dez taxa geram 34.459.425 árvores!

O método de parcimônia é amplamente utilizado, fácil de ser aplicado a modificações morfológicas e possivelmente o mais intuitivo dos métodos de reconstrução de árvores. A parcimônia funciona bem quando as taxas evolutivas são lentas o suficiente para que similaridades ao acaso (devido à evolução independente de estados de caráter idênticos em duas ou mais linhagens) não encubram estados de caráter compartilhados a partir de um ancestral comum. Sob taxas de modificação mais elevadas, no entanto, os métodos de par-

cimônia ficam suscetíveis a um fenômeno conhecido como "atração dos ramos longos" (Quadro 2B). Outros métodos de reconstrução de árvores usam diferentes critérios na escolha da melhor (ótima) árvore. Ao invés de escolher a árvore contendo menos modificações evolutivas, podemos converter a matriz de caráter × táxon para uma medida de similaridade ou dissimilaridade entre as plantas e então construir uma rede que minimiza a dissimilaridade; este é conhecido como o **método da mínima distância**. Alternativamente, podemos desenvolver teorias a respeito das probabilidades de modificação de um estado de caráter em outro e então utilizar essas probabilidades para calcular a verossimilhança que um determinado diagrama ramificado vai levar a um determinado

FIGURA 2.8 (A) Uma matriz de caráter × táxon. (B) Uma rede não enraizada baseada na matriz em 2.8A. Observe que a fusão de pétalas parece modificar-se duas vezes. O comprimento da rede é igual a 5. (C) Uma outra possível rede não enraizada, baseada na matriz em 2.8A. Distintamente da rede apresentada em 2.8B, a fusão de pétalas modifica-se uma única vez, mas o número de cotilédones e de colpos do pólen são modificados duas vezes. O comprimento da rede é igual a 6.



grupo de dados observados. A árvore que apresentar a mais alta verossimilhança é selecionada. Assim, esta abordagem é denominada **método de máxima verossimilhança** (Felsenstein 1981; Hillis et al. 1993; Huelsenbeck 1995; Swofford et al. 1996) (Quadro 2C). Para breves descrições de vários métodos atuais de reconstrução filogenética, ver Hall 2005.

Avaliando a homoplasia

As análises de parcimônia minimizam a importância de caracteres que se modificam em paralelismo ou em reversão. Se existem muitos desses caracteres homoplásicos, a árvore filogenética pode resultar de um artefato dos caracteres que escolhemos, e uma pequena modificação nesses caracteres levará a uma árvore diferente. A medida de homoplasia mais simples e comum em uma árvore filogenética é o **índice de consistência (CI)**, o qual considera a quantidade mínima de modificação evolutiva possível (o número de modificações genéticas) dividida pelo comprimento real da árvore (o número de modificações genéticas presentes na árvore).

Na rede ilustrada na Figura 2.2B, cada caráter da árvore representa uma única modificação genética e cada um dos caracteres é modificado uma única vez, de tal forma que o índice de consistência é $3/3 = 1,0$. Na rede da Figura 2.8B,

existem quatro caracteres binários (passíveis de uma modificação), mas um desses caracteres (fusão das pétalas) é modificado duas vezes na árvore, de tal forma que o índice de consistência é $4/5 = 0,80$.

Os índices de consistência podem também ser calculados para caracteres individuais. Nesse caso, o CI é igual ao número mínimo de modificações possíveis (uma, para caracteres binários) dividido pelo número de modificações presentes na árvore. Por exemplo, o CI de fusão das pétalas (ver Figura 2.8B) é $1/2 = 0,50$. Para uma dada matriz de caráter × táxon, a rede ou árvore mais curta também apresentará o mais alto índice de consistência. Baixos índices de consistência indicam a presença de muitos caracteres que contradizem a árvore evolutiva.

A comparação dos índices de consistência em grupos de dados é uma tarefa perigosa, pois o CI apresenta algumas propriedades indesejáveis. Em primeiro lugar, um caráter que sofre modificação uma única vez em apenas um táxon apresentará um índice de consistência igual a 1,0, apesar de tal caráter não ser informativo em termos das relações existentes. Este tipo de caráter derivado e único é algumas vezes denominado **autapomorfia**. Por exemplo, se uma das plantas de estrela vermelha na Figura 2.8B tivesse folhas pilosas

QUADRO 2B A atração dos ramos longos

A atração dos ramos longos foi originalmente identificada por Felsenstein (1978) como um problema potencial para as análises filogenéticas. Se existem grandes diferenças entre as taxas de evolução de caracteres entre linhagens, de tal forma que algumas linhagens apresentam uma evolução muito mais rápida que outras, e se os caracteres possuem apenas um número limitado de estados de caracteres, então ramos muito grandes podem ser conectados a outros ramos longos em uma árvore, mesmo que eles não tenham uma verdadeira relação de proximidade (Figura 2.9). Este problema é particularmente agudo em dados de seqüências de DNA, nos quais cada caráter apresenta apenas quatro possíveis estados, e para os quais as taxas de mutação são amplamente variáveis.

Esse fenômeno ocorre porque diversas modificações aleatórias, algumas das quais ocorrem em paralelo nas linhagens que estão evoluindo rapidamente, superam as modificações que fornecem informações sobre a ancestralidade comum. O problema não pode ser resolvido pela adição de mais caracteres (pares de bases, no caso de seqüências de DNA); isto apenas adicionaria o número de paralelismos que conectam as linhagens com rápida evolução.

Essa situação pode afetar todos os diferentes métodos de reconstrução de

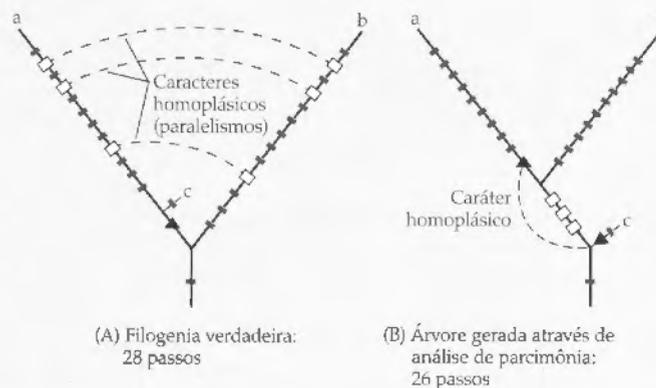


FIGURA 2.9 A atração dos ramos longos, uma situação na qual taxas evolutivas extremamente desiguais provocam uma falha na parcimônia. (A) Uma filogenia verdadeira. As linhas pontilhadas mostram estados de caráter que surgiram em paralelo nas linhagens que originaram a e b. (B) A mesma filogenia reconstruída através de parcimônia. O número de paralelismos compartilhados entre a e b é maior do que o número de caracteres que conectam a e c, de tal forma que a e b aparentam ser *taxa* irmãos, com paralelismos (na filogenia verdadeira) tratados como se fossem caracteres derivados compartilhados por a e b.

árvores. No entanto, com o modelo evolutivo correto, métodos de máxima verossimilhança (ver Quadro 2C) são menos afetados por esse problema (apesar de sabermos que a determinação do modelo correto pode ser bastante difícil). A

atração dos ramos longos é basicamente um problema de amostragem e pode ser atenuado mediante a inclusão de *taxa* que sejam relacionados aos *taxa* mais próximos das extremidades terminais dos ramos longos.

ao passo que todas as outras plantas na rede tivessem folhas não-pilosas, a característica pilosidade das folhas não seria de qualquer ajuda para indicar as relações existentes entre a planta com folhas pilosas e as demais. Em outras palavras, o caráter seria **não-informativo**. No entanto, visto que caracteres não-informativos sofrem modificação uma única vez, eles apresentam um CI de 1,0. Se adicionássemos vários caracteres não-informativos a uma análise, o CI geral seria inchado de acordo com a presença desses caracteres e daria uma falsa impressão de que diversos caracteres estariam dando suporte à árvore. Caracteres não-informativos, portanto, são geralmente omitidos antes do cálculo do índice de consistência.

O índice de consistência é também sensível ao número de *taxa* na análise (Sanderson e Donoghue 1989): análises com muitos *taxa* tendem a apresentar CIs mais baixos do que análises com menos *taxa*. Esta relação é verdadeira tanto para dados morfológicos quanto para dados moleculares e em análises de espécies, gêneros ou famílias.

O uso do **índice de retenção (RI)** evita os problemas sumarizados nos dois parágrafos anteriores, além de outra limitação do CI (Wiley et al. 1991; Forey et al 1992). O CI deveria variar de valores próximos a 0 (um caráter que sofre modificação muitas vezes na árvore) até 1,0 (um caráter que

sofre uma única modificação), mas geralmente a amplitude real é muito menor. Por exemplo, na matriz da Figura 2.8A, apenas dois grupos – as plantas em estrela brancas e as plantas em estrela douradas – possuem um único cotilédono. Se as plantas com um único cotilédono se encontram todas em um único ramo da rede, como na Figura 2.8B, então o CI para número de cotilédones é igual a 1,0. Se elas não são relacionadas, como na Figura 2.8C, então o CI é 0,5 (1/2), que é o menor valor possível na árvore. Assim, em vez de variar entre 0 e 1, o CI varia entre 0,5 e 1,0. O RI corrige esse estreitamento da amplitude do CI comparando o número real de modificações no caráter com o número máximo possível de modificações. O RI é computado por meio do cálculo do comprimento máximo possível da árvore, que é o comprimento que ocorreria se o estado de caráter derivado se originasse independentemente em cada um dos *taxa* no qual ele está presente (i.e., se todos os *taxa* que possuem o estado de caráter derivado não fossem relacionados). O RI então considera o comprimento máximo menos o comprimento efetivo, dividido pelo comprimento máximo menos o comprimento mínimo:

$$(L_{\text{máx}} - L_{\text{efetivo}}) / (L_{\text{máx}} - L_{\text{mín}})$$

Portanto, na Figura 2.8B, o RI é $(9 - 5) / (9 - 4) = 4/5 = 0,80$.

QUADRO 2C Métodos de verossimilhança e Bayesiano

Análises de parcimônia são ainda muito comuns em análises filogenéticas, mas, no caso de análises que usam seqüências de DNA como caracteres, o emprego de métodos de verossimilhança e Bayesiano está se tornando mais rotineiro. Esses métodos se baseiam no pressuposto de que as mutações em uma seqüência de DNA ocorrem de maneira aleatória. Ao longo de um período de tempo evolutivo, se a probabilidade de ocorrer mutação em um determinado nucleotídeo é de 1/100, então esperamos que um nucleotídeo tenha sido mutado em uma seqüência de DNA de 100 nucleotídeos de comprimento. Não temos conhecimento a respeito de qual base em particular sofrerá a mutação, mas sabemos que uma base será alterada. Se o período de tempo considerado for duplicado, teremos uma expectativa de

duas mutações em nossa seqüência hipotética. Em geral, o número esperado de modificações será igual à taxa de mutação multiplicada pelo tempo; essa fórmula é freqüentemente simbolizada por μt . No decorrer de períodos cada vez maiores, mais bases serão modificadas até que, em um determinado ponto, uma segunda mutação ocorrerá sobre um sítio anteriormente já modificado. Novamente, não sabemos que sítio em particular sofrerá essa segunda mutação, mas podemos estimar que isso tenha ocorrido por causa do número total de mutações observadas na seqüência. A teoria básica de probabilidade nos permite estimar o número de mutações "extras" no sítio. Os comprimentos dos ramos usados na criação da árvore filogenética incorporarão estas mutações extras que teremos inferido. Em conjunto,

todos os nossos pressupostos a respeito da probabilidade de mutações específicas constituem um **modelo de evolução**. Os métodos de verossimilhança e Bayesiano são conhecidos como métodos baseados em modelos, pois incorporam idéias a respeito das probabilidades de modificação.

O embasamento estatístico teórico da abordagem dos métodos de verossimilhança e Bayesiano é bastante distinto. No entanto, em termos práticos, uma principal distinção é a velocidade computacional. As análises de máxima verossimilhança precisam de um tempo maior para serem efetuadas, e as análises de *bootstrap* requerem computadores de alto desempenho. O método Bayesiano estima o suporte para a árvore ao mesmo tempo em que computa a árvore e, conseqüentemente, é mais rápido.

Resumindo árvores evolutivas

As análises de parcimônia freqüentemente encontram múltiplas árvores, todas com o mesmo comprimento, mas com diferentes ligações entre os *taxa*. Também, em alguns casos, diferentes métodos de análise resultam em árvores que apresentam topologias diferentes e, conseqüentemente, contam histórias evolutivas diferentes sobre os mesmos *taxa*. Além disso, estudos que usam tipos diferentes de caracteres (p. ex., seqüências gênicas, morfologia) podem chegar a diferentes árvores. Em vez de escolher uma entre as diferentes árvores, neste caso, os sistematas podem simplesmente identificar quais grupos são encontrados em todas as árvores mais curtas, ou que são formados pelo uso de todos os métodos de análise, ou que estão presentes em diferentes tipos de matrizes de caracteres. As informações em comum entre essas árvores podem ser resumidas em uma árvore de consenso.

Árvores de consenso estrito contêm apenas grupos monofiléticos comuns a todas as árvores. Por exemplo, análises de diferentes conjuntos de dados deram origem a diferentes idéias a respeito das relações existentes entre as primeiras angiospermas. Um estudo envolvendo a seqüência de 4 genes originou a árvore evolutiva ilustrada na Figura 2.10A (a qual foi simplificada para facilitar o presente exemplo) (Rydin et al. 2002). A adição de outras seqüências gênicas, e a análise destas por meio de uma metodologia distinta da anteriormente usada, resultou na árvore da Figura 2.10B (Burleigh e Mathews 2004). Ambas as árvores mostram que as angiospermas são irmãs das gimnospermas e que as gimnospermas são monofiléticas. As árvores também mostram que Gnetales e as coníferas (Pinaceae mais coníferas não-Pinaceae) estão intimamente relacionadas. Conseqüentemente, o consenso estrito dos dois cladogramas (Figura 2.10C) mostra as gimnospermas como monofiléticas e as Gnetales mais coníferas como um clado.

No entanto, existem diferenças entre as duas hipóteses evolutivas. A árvore baseada em 4 genes sugere que as

Gnetales são irmãs de todas as coníferas, ao passo que a árvore baseada em 13 genes indica que as Gnetales são irmãs apenas das Pinaceae, as quais constituem um subgrupo das coníferas. Na árvore de consenso estrito (Figura 2.10C), as Gnetales, Pinaceae e coníferas não-Pinaceae aparecem como originadas ao mesmo tempo. Isso significa que os dados disponíveis não são capazes de informar se elas se originaram simultaneamente ou uma após a outra e, assim, não é possível determinarmos em que ordem elas se originaram.

A ocorrência de múltiplas linhagens se originando a partir do mesmo nó aparente no diagrama é, geralmente, uma expressão de ambigüidade. A árvore baseada em 13 genes sugere que as Cycadales são irmãs de todas as demais gimnospermas, mas, na árvore baseada em 4 genes, tanto as Cycadales quanto ginkgo e o clado que contém todas as demais gimnospermas aparecem como se tivessem se originado simultaneamente. A ambigüidade na árvore de 4 genes nos leva a concluir, que na verdade, não sabemos quais linhagens de gimnospermas surgiram primeiro. Esta incerteza se reflete na árvore de consenso estrito, uma vez que todas estas linhagens estão desenhadas como se originando ao mesmo tempo.

Quando muitas árvores são comparadas, pode ser interessante saber se um determinado clado aparece na maioria delas, mesmo que ele não ocorra em todas. Uma árvore **consenso de maioria** mostra todos os grupos que aparecem em 50% ou mais das árvores. Se um determinado clado está presente na maior parte das árvores mais parcimoniosas, ele será representado na árvore consenso de maioria (junto a uma indicação da porcentagem de árvores mais parcimoniosas que apresentam o clado). A árvore consenso de maioria será inconsistente em relação a algumas das árvores originais e, dessa forma, fornecerá apenas um resumo parcial da análise filogenética.

Uma **árvore de consenso semi-estricto** é freqüentemente muito útil, sobretudo em comparações de filogenias com *taxa* terminais que apresentam apenas pequenas diferenças, ou filogenias que tenham sido construídas a partir de diferentes fontes de caracteres. É comum, por exemplo, construir

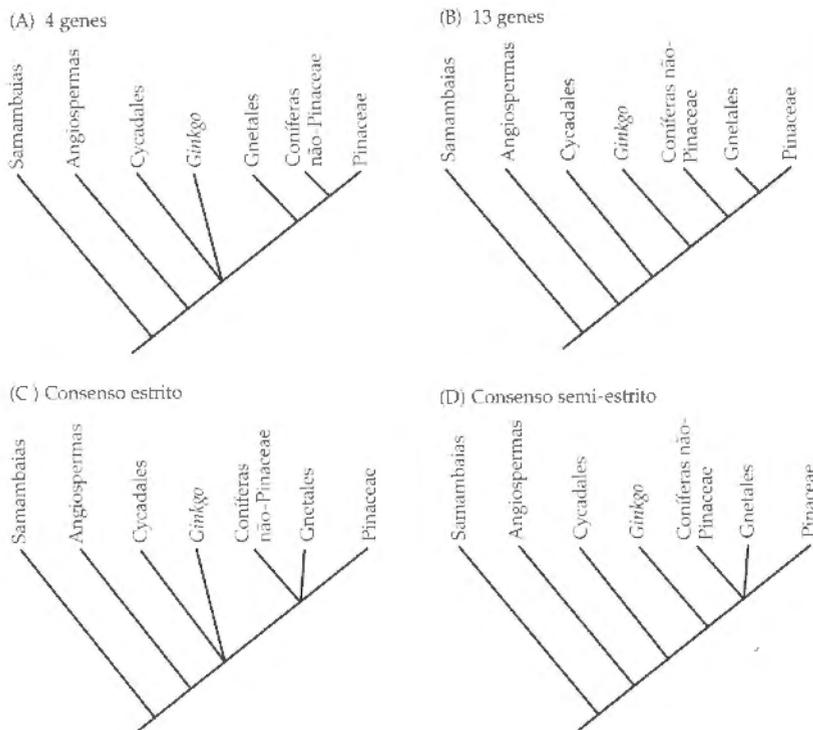


FIGURA 2.10 (A) Filogenia das plantas com sementes baseada em dados de seqüências de DNA provenientes de 4 genes. (B) Filogenia das plantas com sementes baseada em dados de seqüências de DNA provenientes de 13 genes. (C) Consenso estrito das árvores representadas em A e B. (D) Consenso semi-estricto das árvores representadas em A e B. (A baseada em Rydin et al. 2002; D modificada a partir de Burleigh e Mathews 2004.)

árvores a partir de dois conjuntos diferentes de caracteres (p. ex., seqüências gênicas e dados de morfologia) e concluir que ambos os conjuntos de caracteres indicam a monofilia de um grupo determinado de espécies. No entanto, um único conjunto de caracteres pode resolver as relações entre as espécies. A árvore de consenso semi-estricto indica todas as relações apoiadas por uma das árvores, ou por ambas, e não refutadas por qualquer uma delas.

Por exemplo, apesar de a árvore baseada em 4 genes (Figura 2.10A) não fornecer qualquer informação sobre a ordem na qual se originaram as Cycadales, ginkgo e as demais gimnospermas, a árvore baseada em 13 genes (Figura 2.10B) nos dá informações a este respeito. As duas árvores, na verdade, não são conflitantes, apesar de ser verdade o fato de a árvore baseada em 13 genes nos fornecer informações mais exatas. Dessa forma, a árvore de consenso semi-estricto segue o arranjo da árvore baseada em 13 genes no que se refere a esses três grupos (Figura 2.10D).

A probabilidade de modificação evolutiva em caracteres

Ao tentar inferir a história evolutiva de um grupo, dependemos de uma descrição (modelo) implícita ou explícita do processo evolutivo (ver Quadro 2C). Quanto mais acuradamente a descrição refletir o processo evolutivo, maior será nossa capacidade em estimar a história evolutiva. Esse fato é particularmente importante para espécies muito divergentes em filogenias moleculares, para as quais métodos de parcimônia freqüentemente geram resultados conflitantes (ver Quadro 2B). No caso de nucleotídeos em uma seqüência de DNA, assume-se que mutações ocorram aleatoriamente, apesar de esta presunção freqüentemente ser modificada para refletir mecanismos hipotéticos de evolução molecular.

O desenvolvimento de um modelo é muito mais difícil no caso de caracteres morfológicos, pois geralmente não temos informações em relação ao número de genes que estão envolvidos, nem conhecemos que tipos de modificações nesses genes levarão a diferentes estados de caráter. Em todo o caso, certos pressupostos deverão ser considerados se quisermos dar andamento à análise. (E, cabe salientar, *não* existem métodos que sejam totalmente isentos de pressupostos!) Os principais pressupostos deverão estar relacionados com a possibilidade de modificações específicas em estados de caracteres e a possibilidade de reversões e paralelismos.

Ordenando os estados de caracteres Os caracteres na Figura 2.8A apresentam apenas dois estados. Tais caracteres com dois estados (binários) são interpretados como representando uma única mudança genética (p. ex., a partir de um estado com um único sulco, ou monossulcado, para um estado triculpado). Ao longo de um período de tempo evolutivo, naturalmente, tais caracteres podem continuar a sofrer modificações. Por exemplo, o pólen triculpado foi modificado em algumas Caryophyllales de tal maneira que apresenta forma esférica, com muitos poros regularmente espaçados em sua superfície (assemelhando-se bastante a uma bola de golfe); este pólen é dito pantoporado.

Se tivermos que incorporar o caráter colpos do pólen em uma matriz que contém alguns *taxa* com pólen pantoporado, esse caráter apresentará três estados: monossulcado, triculpado e pantoporado. Os colpos do pólen representarão, portanto, um caráter com múltiplos estados (multiestados), contrastando com os caracteres binários discutidos anteriormente. Caracteres multiestados geram uma difícil questão: quantas modificações genéticas estão envolvidas?

É possível que o pólen monossulcado tenha dado origem ao pólen triculpado, o qual, por sua vez, tenha sofrido modi-

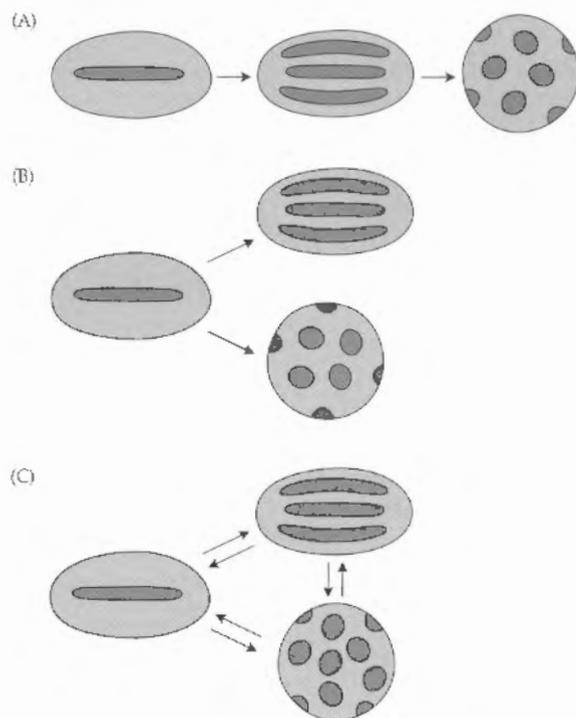


FIGURA 2.11 Três hipóteses alternativas a respeito da evolução da morfologia do pólen. (A) Pólen monossulcado foi modificado em tricolpado, que a seguir foi modificado em pantoporado. Como ilustrado, o caráter é ordenado e irreversível. (B) Pólen monossulcado foi modificado em tricolpado e, independentemente, em pantoporado. Aqui, novamente, o caráter é ordenado e irreversível. Se as flechas forem desenhadas com pontas em ambas as extremidades, poderemos interpretar o caráter como sendo reversível. (C) Qualquer tipo de pólen pode ser modificado em qualquer outro tipo. O caráter é não-ordenado e reversível.

ficações, originando o pólen pantoporado; essa progressão se adequa ao que imaginamos que deve ter acontecido nas angiospermas ao longo do período de tempo evolutivo (Figura 2.11A). (Lembre-se de que o grupo-externo não possui pólen tricolpado.) Esse cenário necessita de duas modificações genéticas. Também necessita que as modificações tenham ocorrido em ordem; ou seja, o pólen pantoporado só poderá originar-se após o surgimento de pólen tricolpado. Se aceitarmos essa série de eventos, teremos **ordenado** os caracteres multiestado.

Mesmo se decidirmos permitir reversões nos estados de caracteres – ou seja, se considerarmos a possibilidade de que um pólen pantoporado possa originar um tricolpado e que o pólen tricolpado possa reverter em monossulcado – o caráter ainda permanecerá ordenado. São necessários dois passos evolutivos (genéticos) para partir do pólen monossulcado e chegar ao pólen pantoporado e dois passos para partir do pólen pantoporado e chegar ao monossulcado. Uma análise filogenética na qual todos os caracteres são tratados como caracteres ordenados é algumas vezes referida na literatura como uma análise de **parcimônia de Wagner**.

Se desconhecemos completamente as plantas envolvidas, podemos considerar a possibilidade de que pólen monossulcado tenha originado pólen tricolpado e que, em um evento independente, pólen monossulcado tenha originado pólen pantoporado (Figura 2.11B). Essa seqüência sugeriria que

existe uma modificação genética que permite a modificação de pólen monossulcado para pólen tricolpado, assim como uma modificação que permite a modificação de pólen monossulcado para pólen pantoporado, mas que uma mudança de pólen tricolpado para pantoporado é impossível. Neste caso, o caráter ainda permanece ordenado, mas segue um caminho diferente do que está ilustrado na Figura 2.11A. Se é possível a reversão, dois passos são necessários para, partindo do pólen tricolpado, chegar-se ao pólen pantoporado, e dois passos são necessários do pólen pantoporado para o tricolpado.

Em relação a caracteres morfológicos e estados de caráter, geralmente não temos certeza das modificações possíveis, de tal forma que é comum tratarmos caracteres multiestado como não-ordenados (Figura 2.11C); esse método é frequentemente denominado **parcimônia de Fitch**. No caso de um caráter não-ordenado, postulamos apenas uma modificação entre quaisquer dois estados. Caracteres de seqüência de DNA são caracteres multiestado com quatro estados (adenina, timina, guanina e citosina). Não é possível tratá-los como ordenados; uma adenina não precisa obrigatoriamente ser substituída por uma citosina antes de ser substituída por uma guanina. Dessa forma, caracteres de DNA são sempre tratados como não-ordenados e plenamente reversíveis.

Reversões, paralelismos e pesagem de caracteres Na rede da Figura 2.8B, foi hipotetizado que a fusão de pétalas teve origem duas vezes, de forma independente. Para produzir a rede da Figura 2.8C, ligeiramente mais longa, foi necessário permitir que o número de cotilédones sofresse modificação de um para mais de um, e novamente fosse alterado para um – ou seja, que sofresse reversão. Mediante a comparação das árvores das Figuras 2.8B e C estaremos, portanto, comparando duas hipóteses: (1) que mutações em genes que levam à fusão de pétalas ocorreram mais de uma vez e (2) que mutações em genes que controlam o número de cotilédones ocorreram e, a seguir, houve reversão de seus efeitos. Ao decidir que a rede da Figura 2.8B é mais curta do que a rede apresentada na Figura 2.8C, contamos todos os passos por igual, independentemente do fato de indicarem paralelismos, reversões ou origens únicas.

Essa abordagem pode ou não ser considerada razoável. A lei de Dollo, por exemplo, sugere que, no caso de caracteres muito complexos, a origem por paralelismo é bastante improvável, ao passo que a reversão pode ser relativamente fácil (Mayr e Ashlock 1991). O pressuposto é que diversos genes devem ser alterados para que uma estrutura morfológica seja criada, mas que uma modificação em apenas um desses genes fará com que a estrutura possa ser perdida.

Podemos introduzir a lei de Dollo no processo de escolha de uma árvore fazendo com que o ganho de estruturas conte mais do que a perda; o processo é denominado **parcimônia de Dollo**. (Obviamente a definição dos termos *ganho* e *perda* necessitarão de uma árvore enraizada; por conseqüência, a parcimônia de Dollo não pode ser aplicada a uma rede não-enraizada.)

Determinados caracteres às vezes são **pesados** nas análises filogenéticas. Este peso reflete o pressuposto de que determinados caracteres deverão ser mais difíceis de alterar do que outros. Pode-se hipotetizar, por exemplo, que a anatomia da folha é uma característica menos provável de sofrer modificação do que a presença de pêlos na folha (pubescência) e que, portanto, uma modificação em um caráter anatômico

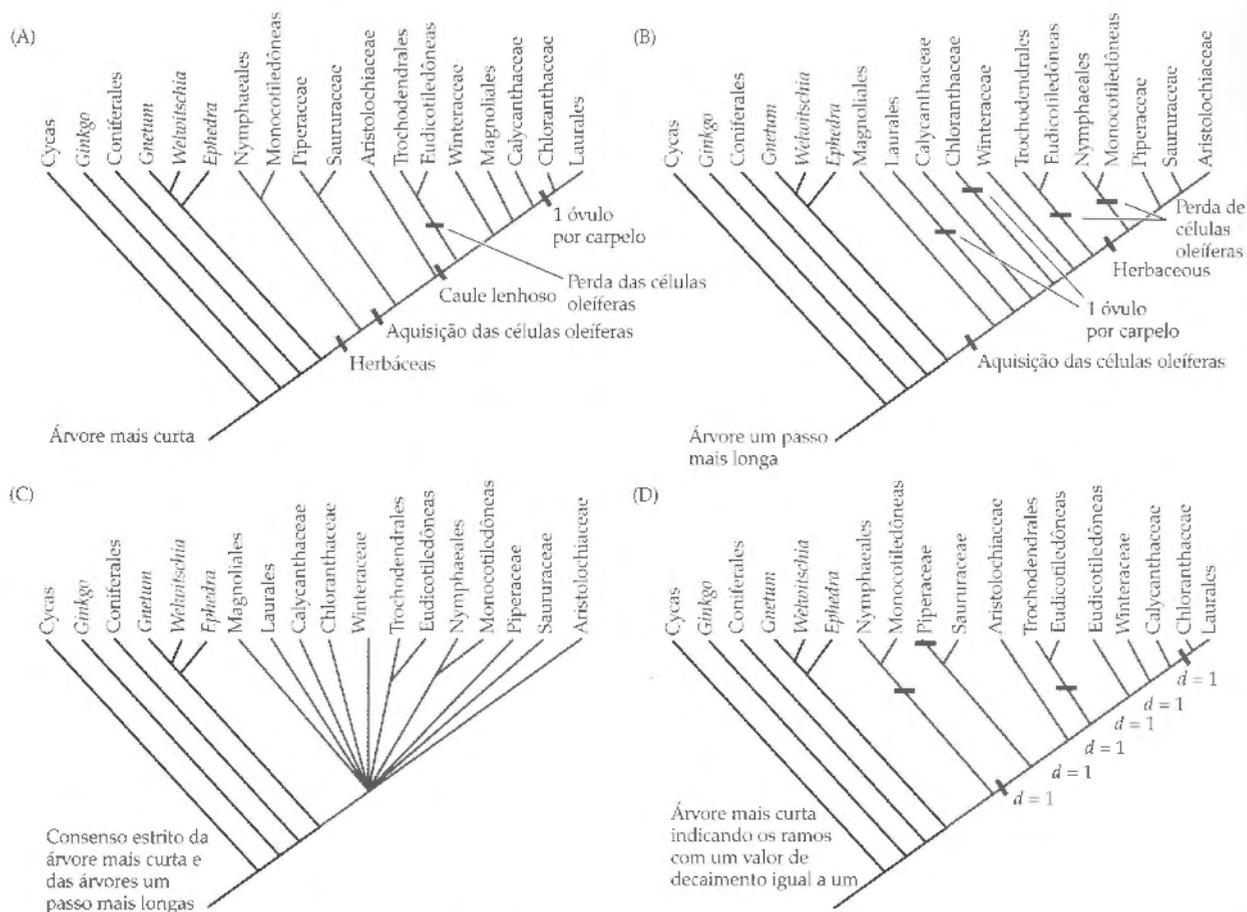


FIGURA 2.13 (A) Filogenia das angiospermas (linhas azuis), indicando os padrões de modificação em relação à presença ou ausência de células oleíferas, número de óvulos por carpelo e hábito da planta (lenhoso ou herbáceo). (B) Uma árvore alternativa, um único passo mais longa do que a árvore apresentada em A, mostrando os padrões de modificação para as mesmas características. Observe que o hábito herbáceo

é agora considerado como havendo evoluído uma única vez, mas que a perda de células oleíferas e a redução do número de óvulos ocorreram duas vezes. (C) Consenso estrito da árvore mais curta e de árvores um passo mais longas (Figuras 2.13A e B). (D) A mesma árvore ilustrada em A, ilustrando ramos com valor de decaimento igual a 1. (Dados de Doyle et al. 1994.)

ou se as modificações evolutivas ocorrerem em frequências demasiadamente baixas na história de um determinado grupo.

Uma forma simples de avaliar o suporte para uma parte específica de uma árvore é pela análise do número de modificações genéticas que ocorrem em um ramo que leva a um determinado grupo, juntamente com os índices de consistência dos caracteres. Por exemplo, uma filogenia de alguns integrantes das Ericaceae baseada em dados de seqüências de DNA (Figura 2.12; Kron e Judd 1997) encontrou 18 modificações no ramo que leva ao clado *Lyonia*. Em uma análise de caracteres morfológicos para os mesmos taxa, foram observados quatro caracteres que sofriam modificação exclusivamente ao longo do ramo de *Lyonia*. Em outras palavras, uma série de modificações que ocorreram durante a origem do clado *Lyonia* levou à produção de novas características que não são encontradas em nenhuma outra parte da família. Grupos como o clado *Lyonia* que compartilham diversos caracteres que não ocorrem em outros pontos do cladograma são mais confiáveis do que grupos que compartilham apenas alguns poucos caracteres altamente homoplásicos.

Outra forma de avaliar o quanto os dados apóiam a árvore é determinando se um grupo de interesse ocorre em outras árvores que sejam quase tão curtas quanto a que está sendo observada. Em outras palavras, suponha que tenhamos questionado se existem outras formas de analisar caracteres homoplásicos que levam a árvores que sejam um, dois ou três passos mais longas.

Por exemplo, em um estudo sobre a diversificação das angiospermas (Doyle et al. 1994), a árvore mais curta indicou que as linhagens que divergiram mais precocemente nas angiospermas eram as monocotiledôneas e os lírios-d'água (Nymphaeaceae; ver Capítulo 9). Isso significa que o caráter caule herbáceo foi adquirido uma vez e foi perdido posteriormente, ao passo que a redução para um óvulo por carpelo ocorreu uma única vez, e o caráter células oleíferas foi adquirido e perdido uma única vez (Figura 2.13A). Por outro lado, árvores um passo mais longas, nas quais as linhagens de angiospermas mais antigas correspondem às magnólias, sugerem que caules herbáceos evoluíram uma vez, mas que a redução no número de óvulos ocorreu duas vezes e que ocorreram três modificações no caráter células oleíferas (aquisição uma única vez e perda duas vezes ou vice-versa) (Figura 2.13B).

Assim, observando árvores um passo mais longas, podemos hipotetizar que alguns caracteres são menos homoplásicos, ao passo que outros são mais homoplásicos. Se agora analisarmos o consenso estrito de todas as árvores, incluindo as mais curtas e aquelas um passo mais longas, todas as linhagens precoces de angiospermas estão ilustradas como se tivessem radiado a partir de um ponto único, indicando incerteza na ordem em que elas realmente evoluíram (Figura 2.13C).

Você pode observar que vários dos ramos presentes nas árvores mais curtas não estão presentes nas árvores que possuem um passo a mais de comprimento. Dessa forma, nenhum destes ramos está ilustrado no consenso estrito; em outros termos, eles “colapsam”, ou “decaem”. Podemos indicar esse fato por meio do posicionamento do numeral 1 próximo a cada um dos ramos colapsados na árvore mais curta (Figura 2.13D). Este número é o **índice de decaimento**, também conhecido como índice de Bremer, e representa o número de passos extras necessários para encontrar árvores que não contenham um grupo determinado. Ele fornece uma medida relativa de quanto a homoplasia nos dados afeta o suporte a um grupo em particular.

O índice de decaimento não é estatístico, o que, dependendo do ponto de vista, pode ser considerado uma vantagem ou uma desvantagem. Visto que a história acontece uma única vez e não pode ser repetida, é impossível replicar um experimento evolutivo. No entanto, decididamente, é possível testar se os dados de caracteres são diferentes do esperado para um processo aleatório, mesmo considerando-se que existem muitas formas diferentes de aleatorizar dados de sistemática. Diferentes testes que utilizam técnicas de aleatorização já foram desenvolvidos. Provavelmente, o método mais amplamente utilizado é a análise de *bootstrap*.

A **análise de bootstrap** aleatoriza os caracteres no que se refere aos *taxa*. Como exemplo, observe a matriz da Figura 2.8A e aleatoriza as colunas ao mesmo tempo em que mantém fixas as linhas. Pegue aleatoriamente uma coluna da matriz original para que esta seja a primeira coluna de uma nova matriz. A seguir, escolha outra coluna para que seja a segunda coluna e assim sucessivamente até que uma nova matriz, contendo o mesmo número de colunas que a matriz original, tenha sido criada. Visto que retornamos à matriz original a cada um dos passos para escolher uma nova coluna, alguns caracteres podem estar representados várias vezes na nova matriz, ao mesmo tempo em que alguns poderão ser omitidos. Esse método é geralmente descrito como amostragem aleatória com reposição.

A Figura 2.14 mostra a matriz da Figura 2.8A amostrada aleatoriamente com reposição; observe que o primeiro caráter da matriz original (colpos do pólen) foi selecionado duas vezes, ao passo que o terceiro caráter (inflorescência em capítulo) foi perdido durante o processo de seleção aleatória. Muitas matrizes aleatorizadas são construídas, e árvores mais parcimoniosas são encontradas para cada nova matriz. Esse processo é usado para criar um conjunto de pelo menos 100 árvores, que podem ser resumidas em uma árvore de consenso (ver páginas 24-25). Na árvore de consenso de *bootstrap*, um clado com valor de *bootstrap* de, digamos, 95% estava presente em 95% das árvores geradas na análise de *bootstrap*.

A filogenia da Figura 2.12 apresenta tanto os índices de decaimento quanto de *bootstrap*, aliados ao comprimento dos ramos. Podemos observar que os valores de decaimento e de *bootstrap* são altos para o gênero *Lyonia*, indicando que os dados dão suporte para a monofilia do gênero, ao passo que a

	Colpos do pólen	Pétalas	Colpos do pólen	Número de cotilédones
Plantas em estrela vermelha	< 3	livres	< 3	2
Plantas em estrela amarela	< 3	livres	< 3	1
Plantas em estrela branca	< 3	fusionadas	< 3	1
Plantas em círculo	3	livres	3	2
Plantas em quadrado	3	fusionadas	3	2
Plantas em losango	3	fusionadas	3	2
Conífera	< 3	não-aplicável	< 3	> 2

FIGURA 2.14 Matriz da Figura 2.8A, amostrada com reposição, como ocorreria para o primeiro passo de uma análise de *bootstrap*. Observe que no processo de amostragem o caráter colpos do pólen foi amostrado duas vezes, ao passo que o caráter inflorescência foi omitido.

ligação entre *Agarista* e *Pieris* é apoiada por apenas 51% das árvores de *bootstrap*, e, em árvores com apenas um passo a mais de comprimento, os dois gêneros não são irmãos, o que é indicado pelo valor de decaimento igual a 1.

Outra excelente forma de conferir confiança aos agrupamentos presentes em uma árvore é comparar filogenias que tenham se baseado em diferentes conjuntos de caracteres. Por exemplo, filogenias baseadas em morfologia, em seqüências de DNA de cloroplasto (cpDNA) e em seqüências de DNA nuclear podem (e freqüentemente são) comparadas. Se essas filogenias apresentam agrupamentos similares, podemos ter uma maior confiança de que elas refletem a ordem real dos eventos. Por exemplo, a monofilia de famílias como Poaceae, Onagraceae, Ericaceae, Asteraceae e Orchidaceae tem sido apoiada por análises filogenéticas realizadas com diferentes tipos de dados incluindo morfológicos, seqüências de genes de DNA de cloroplasto e seqüências de genes nucleares.

A comparação entre árvores freqüentemente mostra-se intrigante quando os dados provêm de diferentes genes, como será discutido em maior detalhe no Capítulo 5. Também é comum que caracteres de DNA e morfológicos sejam combinados em uma única análise filogenética, o que geralmente dá origem a filogenias mais fortemente apoiadas em relação àquelas geradas a partir de um único tipo de dados. Filogenias baseadas em caracteres morfológicos assumem que não ocorreu hibridização, ou pelo menos que este evento foi raro (Quadro 2D); esta questão pode ser testada pelo uso de múltiplas árvores moleculares.

Descrevendo a evolução: o mapeamento de caracteres em árvores

Filogenias podem ser usadas para descrever o processo evolutivo e para desenvolver hipóteses a respeito de adaptação, modificações morfológicas e fisiológicas ou sobre aspectos biogeográficos, entre diversos outros usos. Se uma filogenia será utilizada com o objetivo de descrever a história evolutiva, no entanto, atenção especial deverá ser dada a caracteres e estados de caracteres usados para esta descrição. Na discus-

QUADRO 2D A análise filogenética assume que a evolução pode ser ilustrada na forma de uma árvore ramificada

Os estudos filogenéticos assumem que após duas linhagens divergirem elas nunca mais trocarão informação genética. Este pressuposto pode, no entanto, ser frequentemente transgredido. Se a hibridização é comum, uma planta pode compartilhar caracteres derivados de duas plantas parentais não relacionadas, e a história se assemelhará muito mais a um trabalho de macramê do que a uma árvore. A análise filogenética sempre produzirá um diagrama semelhante a uma árvore, seja ele apropriado ou não. Métodos filogenéticos pressupõem evolução divergente e não podem fornecer

uma filogenia correta para híbridos, os quais apresentam histórias evolutivas reticuladas.

Sabe-se que hibridização interespecífica é comum em plantas, e o tratamento adequado dos híbridos em análises cladísticas já foi objeto de muitas discussões (Bremer e Wanntorp 1979; Wagner 1980, 1983; Bremer 1983; Funk 1985; Kellogg 1989; Kellogg et al. 1996). A maioria dos sistematistas tem sugerido que os híbridos sejam identificados e removidos das análises, pois sua inclusão pode levar a um aumento da homoplasia, a um aumento no número de árvores mais parcimonio-

sas e à distorção dos padrões de relacionamento entre *taxa* não-híbridos.

No entanto, estudos de McDade (1990, 1992, 1997) indicaram ser pouco provável que híbridos causem problemas em análises filogenéticas, a menos que ocorram entre espécies parentais distantes. Quando híbridos são reconhecidos e sua ancestralidade determinada (ver Capítulo 6), eles podem ser inseridos manualmente no cladograma, que então não indicará apenas eventos cladogênicos (contadas por meio da especiação), mas também histórias reticuladas (desenvolvidas por meio da hibridização interespecífica).

são a seguir, nos deteremos em caracteres morfológicos, mas muitos dos pontos abordados podem ser aplicados a qualquer outro tipo de caráter.

Considere um grupo de plantas para o qual a árvore filogenética seja conhecida; um bom exemplo é o grupo das Ericaceae, para o qual um conjunto substancial de informações está disponível (Figura 2.15). Vamos assumir, para fins da presente discussão, que a árvore em questão reflete com acurácia a história e que cada um dos gêneros terminais realmente é monofilético, como demonstrado por meio de estudos com diversas espécies de cada um dos gêneros. A seguir, consideraremos um estudo interessado na análise de ganho ou perda do caráter pétalas fusionadas, o qual está intimamente relacionado com a evolução de sistemas de polinização. Esse é o tipo de estudo geralmente desenvolvido por sistematistas, pois os detalhes da evolução do caráter podem levar a hipóteses sobre a ação da seleção natural. Além disso, quando estamos construindo classificações, freqüentemente temos interesse em determinar que caracteres morfológicos podem ser atribuídos a um grupo monofilético determinado distinguindo-o dos demais grupos.

A Figura 2.15 ilustra os estados de caráter observados nos gêneros. Observando a distribuição de caracteres e estados de caráter, parece extremamente óbvio que pétalas livres devem ter evoluído na linhagem que deu origem a *Ledum* e uma segunda vez na linhagem que originou *Vaccinium* seção *Oxycoccum*. Em outras palavras, o ancestral de *Vaccinium* seção *Oxycoccum* e todas as outras espécies de *Vaccinium* possuem pétalas fusionadas, da mesma forma que o ancestral de *Ledum* e de *Rhododendron* sect. 3.

Examinaremos esta conclusão "óbvia" um pouco mais atentamente. Se estivéssemos estudando apenas espécies de *Vaccinium*, não teríamos condições de saber se pétalas fusionadas são ancestrais ou derivadas (Figura 2.16A). Deve ter ocorrido uma modificação genética, mas ela poderia ter acontecido tanto na linhagem que originou as espécies de *Vaccinium* seção *Oxycoccum*, quanto na linhagem que originou as demais espécies *Vaccinium*.

Apenas pela referência do grupo-externo *Epacris* é que podemos determinar quando pétalas fusionadas foram per-

didadas. Visto que *Epacris* possui pétalas fusionadas, pétalas livres devem ter-se originado dentro de *Vaccinium*; é mais simples (mais parcimonioso) assumir apenas uma modificação genética, de fusionada para livre (Figura 2.16B). Isso equivale a dizer que o ancestral de *Vaccinium* possuía pétalas fusionadas. Se postulássemos que o ancestral possuía pétalas livres, seriam necessárias duas modificações para pétalas fusionadas: uma em *Epacris* e uma nas espécies de *Vaccinium* que não pertençam a seção *Oxycoccum*. O mesmo argumento pode ser aplicado no caso de *Rhododendron* e *Ledum*.

Agora, suponha que estamos estudando apenas espécies de *Vaccinium*, mas que, ao invés de utilizarmos *Epacris* ou ou-

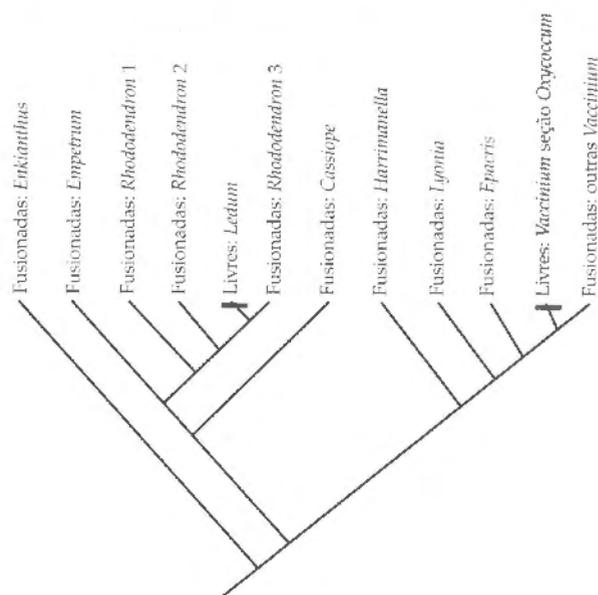


FIGURA 2.15 Filogenia de uma parte das Ericaceae. O gênero *Rhododendron* é parafilético e está representado por três linhagens separadas, numeradas de 1 a 3. Duas modificações para pétalas livres são hipotetizadas. (Dados de Stevens 1998.)

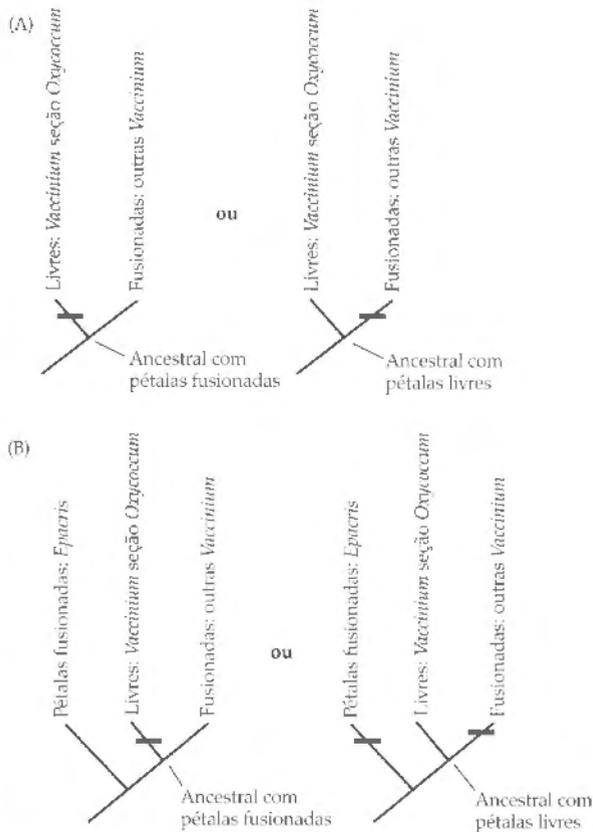


FIGURA 2.16 (A) Dois taxa de *Vaccinium* diferem em relação a estados de caracteres. É impossível determinar, a partir desta única informação, qual é o estado de caráter ancestral, pois qualquer das possibilidades envolve uma modificação em uma linhagem descendente. (B) A adição de um grupo-externo determina o estado de caráter do ancestral. Neste caso, é mais simples (requer menos passos) assumir que o ancestral possuía pétalas fusionadas.

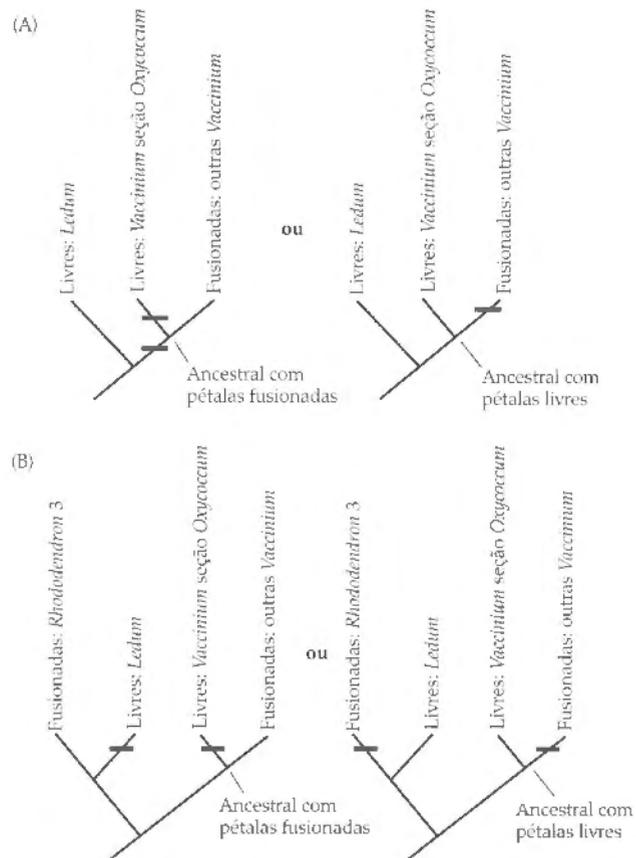


FIGURA 2.17 (A) Análise de modificação de estado de caráter em *Vaccinium* usando um grupo-externo diferente. Observe que a inferência do estado de caráter ancestral é absolutamente oposta àquela quando *Epacris* foi usada como grupo-externo. (B) Análise de modificação de estado de caráter em *Vaccinium* usando dois grupos-externos que diferem em relação ao estado. Neste caso é impossível determinar o estado de caráter do ancestral.

tra Ericaceae como grupo-externo, tenhamos usado apenas *Ledum*. Essa situação é fácil de ser imaginada no caso de ser difícil a obtenção de material de outros gêneros, ou no caso de tais gêneros estarem extintos ou mesmo se desconhecermos sua existência. Assim, concluiríamos que o ancestral de todas as espécies de *Vaccinium* possuía pétalas livres e que, em resposta a uma pressão seletiva desconhecida, houve modificação para pétalas fusionadas (Figura 2.17A). Essa conclusão é absolutamente oposta à conclusão a qual havíamos chegado no exemplo anterior, no entanto a única diferença entre ambos foram os gêneros incluídos na análise.

Podemos tentar resolver a situação mediante o uso de grupos-externos adicionais. Por exemplo, consideraremos o mesmo estudo com *Vaccinium*, mas desta vez utilizaremos tanto *Ledum* quanto *Rhododendron* como grupos-externos. Neste caso, a direção da modificação é totalmente ambígua (Figura 2.17B). É tão simples postular que o ancestral do grupo possuía pétalas fusionadas e que ocorreram duas modificações para pétalas livres quanto postular que o ancestral possuía pétalas livres e que ocorreram duas modificações para pétalas fusionadas.

Estas duas possibilidades são denominadas **reconstruções igualmente parcimoniosas**. Para muitos caracteres em muitas árvores, podem ocorrer múltiplas reconstruções

igualmente parcimoniosas. Em outras palavras, existem várias hipóteses igualmente boas para descrever a direção e o tempo de modificação do estado de caráter. Se retornarmos ao exemplo da Figura 2.13, seremos capazes de encontrar reconstruções igualmente parcimoniosas diferentes das que estão ilustradas.

Uma ambigüidade também pode ser resultante da inclusão de taxa para os quais o estado de caráter é desconhecido. Suponha, por exemplo, que dois novos taxa sejam descobertos de tal forma que, tendo como base outros caracteres, um deles seja claramente irmão de *Vaccinium* seção *Oxyococcum* e o outro seja irmão das demais *Vaccinium* (Figura 2.18). Além disso, suponha que não existam dados exatos em relação a eles possuírem pétalas livres ou fusionadas. (Esse tipo de ambigüidade é mais comum do que você pode imaginar; ele pode ocorrer quando a descrição original é vaga e/ou as ilustrações não são claras, ou quando a planta original só é conhecida através de seus frutos.) Neste exemplo, desconhecemos como era o estado ancestral para *Vaccinium*, de tal forma que podemos formular qualquer hipótese a respeito da direção da modificação evolutiva. Também não poderemos ter certeza de que o caráter "pétalas fusionadas" seja sinapomórfico para o gênero.

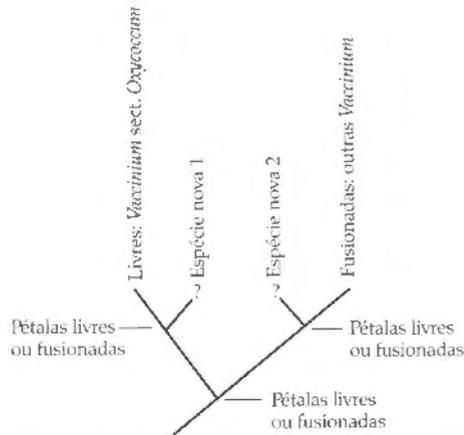


FIGURA 2.18 A adição de espécies para as quais desconhecemos o estado de caráter pode evitar inferências sobre o estado ancestral.

Diversos algoritmos foram desenvolvidos para designar modificações de estados de caráter a porções específicas de árvores (ver Capítulos 3 e 4 de Maddison e Maddison 2000 para uma discussão clara e compreensível deste assunto). Dependendo do algoritmo utilizado, as modificações do estado de caráter podem ser direcionadas em prol de paralelismos (o algoritmo de “transformação retardada” ou DELIRAN, do inglês “*delayed transformation*”) ou em prol de reversões (o algoritmo de “transformação acelerada” ou ACCTRAN, do inglês “*accelerated transformation*”). Os resultados podem interferir – algumas vezes de forma bastante importante – em hipóteses que dizem respeito ao processo evolutivo e podem também afetar a forma como os organismos são descritos em uma classificação.

Construindo uma classificação

A teoria da classificação é um tópico com o qual os sistematas vêm brigando há centenas de anos; tais combates resultaram em uma literatura ampla e, freqüentemente, conflitante (ver Capítulo 3). Os princípios da classificação filogenética aqui delineados são amplamente, apesar de não universalmente, aceitos. É preciso salientar que uma classificação costuma ter diversos objetivos. Uma classificação é um vocabulário comum desenvolvido para auxiliar na comunicação. Conseqüentemente, uma classificação deve ser estável; nomes que são freqüentemente alterados não são úteis em termos de comunicação. Além disso, uma classificação deve ser preditiva; ou seja, o nome de uma planta deve auxiliar a compreender melhor a planta e deve guiar na busca de literatura apropriada.

Os sistematas geralmente concordam em relação aos objetivos de um processo de classificação, apesar de discordarem profundamente em relação a como esses objetivos serão atingidos. No presente texto, adotamos um ponto de vista particular, usando classificações filogenéticas em todos os seus aspectos. Assim, sempre que possível, reconhecemos grupos monofiléticos e evitamos grupos parafiléticos ou polifiléticos. Nos poucos casos em que famílias ou ordens não-monofiléticos ainda não foram divididos em unidades monofiléticas, o nome do táxon foi colocado entre aspas. A monofilia de diversos gêneros de angiospermas ainda é bastante questionável, mas relativamente poucas análises filoge-

néticas estão disponíveis a este nível, de tal forma que, em geral, não tentamos indicar uma possível ou provável parafilia ou polifilia dos gêneros.

A diversidade biológica na Terra é resultante de descendência genealógica com modificações, e grupos monofiléticos devem sua existência a esse processo. Conseqüentemente, é apropriado usar grupos monofiléticos em classificações biológicas de forma a refletir com maior precisão sua história genealógica. Classificações baseadas em grupos monofiléticos são mais preditivas e de maior valor heurístico do que aquelas baseadas em similaridades gerais ou em valorações subjetivas de caracteres específicos (Farris 1979; Donoghue e Cantino 1988).

Classificações filogenéticas, tendo em vista refletirem a genealogia, apresentarão maior utilidade em áreas biológicas, tais como o estudo da distribuição de plantas (fitogeografia), interações parasita-hospedeiro e herbívoros-plantas, biologia da polinização e dispersão de frutos, bem como na abordagem de questões relativas à origem de caracteres adaptativos (Nelson e Platnick 1981; Humphries e Parenti 1986; Brooks e McLennan 1991; Foey et al. 1992). Considerando-se seu potencial preditivo, uma classificação filogenética pode direcionar a busca por genes, produtos biológicos, agentes de controle biológico e espécies com potencial para cultivo. A informação filogenética é também muito útil na tomada de decisões de conservação. Em conclusão, classificações filogenéticas fornecem um arcabouço para o conhecimento biológico e a base para estudos comparativos, conectando todos os diferentes campos da biologia (Cracraft e Donoghue 2004).

A construção de classificações envolve duas etapas. A primeira consiste na delimitação e nomeação dos grupos. Em uma classificação filogenética, esta etapa não apresenta controvérsias: grupos que recebem nomes devem ser monofiléticos. A segunda etapa envolve o ordenamento dos grupos e o posicionamento destes de acordo com uma hierarquia. Esta fase permanece problemática.

Agrupando: grupos que recebem nomes são monofiléticos

Uma classificação filogenética reflete a história evolutiva e visa nomear *apenas* os grupos que são monofiléticos – ou seja, compostos por um ancestral e todos os seus descendentes. No exemplo da Figura 2.4C, inferimos que Asteraceae (as plantas representadas por um losango) são monofiléticas, pois possuem flores em capítulos. As plantas representadas por um quadrado mais as Asteraceae são também monofiléticas, pois compartilham o estado de caráter derivado pétalas fusionadas; esse grupo também possui um nome, Asteridae (ou clado das asterídeas). De forma semelhante, o grupo completo de plantas com pólen tricolpado (plantas representadas por círculos mais as Asteridae) é monofilético e denominado eudicotiledôneas (ou o clado das tricolpadas). Esse grupo poderia ter recebido um nome formal em latim, no entanto, até o momento, não recebeu esse tipo de denominação e provavelmente não há necessidade de que venha a receber.

Na classificação filogenética, grupos parafiléticos não são nomeados. Na Figura 2.4C, um grupo composto pelas plantas representadas por círculos mais as plantas representadas por quadrados seria parafilético. O ancestral comum mais recente compartilhado por qualquer planta representada por um quadrado e uma planta em círculo também é o ancestral comum mais recente de qualquer planta representada por círculo e

uma planta em losango. Em outras palavras, as plantas em círculo são tão longinquamente relacionadas às plantas representadas em quadrado quanto qualquer uma delas o é em relação às plantas em losango. Nomear um grupo que inclua as plantas em quadrado mais as plantas em círculo significaria dizer que esses dois tipos de planta são intimamente relacionados mesmo que isso não esteja ocorrendo.

Existem diversos exemplos neste livro de grupos de plantas que receberam denominação, mas que atualmente se acredita que sejam parafiléticos. Um exemplo bem conhecido envolve as "briófitas", um grupo que inclui tradicionalmente plantas terrestres avasculares (hepáticas, antóceros e musgos; ver Figura 1.1). No entanto, a distância entre as hepáticas, antóceros e musgos é maior do que a distância entre os musgos e as plantas vascularizadas (traqueófitas). Sem a marcação com aspas, o nome *briófitas* implicaria em uma relação de proximidade entre estas plantas maior do que a relação realmente existente.

Diversas famílias de plantas tradicionalmente reconhecidas, como Apocynaceae e Capparaceae em um senso amplo, são parafiléticas. No presente texto, estas famílias foram recircunscritas de modo a torná-las grupos monofiléticos: Apocynaceae foi combinada com Asclepiadaceae, e Capparaceae foi inserida em Brassicaceae (apesar de alguns sistematas aceitarem Capparaceae s.s. e Cleomaceae).

Nomeando: nem todos os grupos são nomeados

Uma classificação filogenética visa dar nomes unicamente a grupos monofiléticos, mas o fato de um grupo ser monofilético não significa que ele deva obrigatoriamente receber um nome. As razões para essa situação são de ordem prática. Poderíamos dispor cada par de espécies em seu próprio gênero, cada par de gêneros em sua própria família, cada par de famílias em sua própria superfamília e assim por diante. No entanto, este tipo de classificação seria inconvenientemente grande; além disso, não seria estável, pois nossa visão em relação a espécies irmãs se alteraria cada vez que uma nova espécie fosse descrita, o que levaria a uma nova visão de toda a classificação, a qual, por sua vez, deveria ser reorganizada de acordo com a nova visão.

Na prática, vários grupos monofiléticos não recebem nomes. Por exemplo, o gênero *Stenanthium* (Melanthiaceae) é monofilético e contém quatro espécies (Zomlefer et al. 2001; Zomlefer e Judd 2002; Wofford 2006). Apesar de estar claramente definido que estas quatro espécies se organizam em dois pares monofiléticos, nenhum dos pares recebeu nome especial, e poucos sistematas consideram a necessidade de fazê-lo. Em outro exemplo, praticamente a metade dos gêneros da família das gramíneas se concentra em um único grande clado que contém quatro grandes subfamílias tradicionalmente reconhecidas e duas outras menores. Apesar dos agrostólogos se referirem ao clado por meio da denominação de clado PACCAD (um acrônimo para Panicoideae-Arundinoideae-Centothecoideae-Chloridoideae-Aristidoideae-Danthonioideae), ele não possui um nome formal em latim.

Como os sistematas decidem quais grupos monofiléticos devem receber nomes? Não existe um conjunto de regras codificado, mas diversos critérios têm sido sugeridos, alguns dos quais são usados comumente apesar de não estarem totalmente ajustados. Um dos principais critérios – e talvez o critério principal – é a força das evidências que sustentam um

dado grupo. Idealmente, apenas clados ligados por muitos caracteres derivados compartilhados deveriam ser formalmente reconhecidos e receber nomes em processos de classificação. Isso faz sentido quando queremos que a classificação funcione como um vocabulário comum.

Os nomes tornam-se mais úteis quando podem ser definidos e, quanto maior a exatidão da definição, melhor será o nome. Em outras palavras, para receber um nome, um clado deve apresentar um conjunto de caracteres que possa distingui-lo dos demais clados, podendo ser **diagnosticado**. Esse critério também é importante para a estabilidade da nomenclatura: se o significado do nome sofre modificações cada vez que uma nova filogenia é construída ou cada vez que um novo caráter é examinado, o nome torna-se efetivamente sem significado.

Um segundo critério consiste na presença de um caráter morfológico óbvio. Apesar de os sistematas talvez não concordarem com a importância desse critério, ele representa uma extensão importante da ideia de um grupo bem sustentado e é também relevante quando do uso das classificações por não-sistematas, com o objetivo de identificação. Se, por exemplo, a única forma de uma bióloga de campo identificar um dado organismo é por meio de uma alanina ou de uma serina na posição 281 de sua molécula ribulose 1,5-bisfosfato carboxilase/oxigenase, ela provavelmente não considerará a classificação muito útil para fazer inferências a respeito do organismo. Se, por outro lado, ela souber que o organismo em questão é uma gramínea contendo uma estrutura de espiguetas específica, poderá facilmente e fidedignamente inferir diversos aspectos de sua biologia. (A ausência de sinapomorfia morfológica óbvia é uma das muitas razões de o clado PACCAD não ter recebido um nome.) Os caracteres utilizados para a classificação não necessitam ser os mesmos usados para a identificação, porém muitos sistematas preferem dar nomes a clados que sejam facilmente reconhecidos por meio da morfologia.

Outro critério consiste no tamanho do grupo. A memória humana tem grande facilidade em recordar pequenas quantidades de itens (em um intervalo entre 3 e 7) (Stevens 1998); no entanto, mecanismos mnemotécnicos adicionais são necessários para organizar e recordar grandes quantidades de itens. (Por exemplo, considere quantos CPFs contendo 11 dígitos você consegue se lembrar em comparação com números de telefone com até 8 dígitos). Uma das formas de organizar o pensamento, em termos de *taxa* muito numerosos, é pela divisão de grupos grandes em grupos menores. Nas palavras de Davis e Heywood (1963: 83), "Devemos ser capazes de posicionar *taxa* em *taxa* superiores de tal forma que possamos reencontrá-los." O gênero *Stenanthium* pode ser redefinido para agrupar apenas *Stenanthium gramineum* e *S. diffusum*, e um novo gênero pode ser descrito para incluir *S. densum* e *S. leimanthoides*. No entanto, são poucas as razões para fazer essa divisão, visto ser fácil acompanhar um conjunto de quatro espécies. Além disso, não existem muitas razões para dividir um grupo grande se clados bem sustentados não puderem ser identificados em seu interior.

Um quarto critério consiste na estabilidade da nomenclatura. Uma classificação é, no final das contas, um vocabulário, uma forma de comunicação. Essa comunicação não ocorrerá se os significados dos nomes constantemente sofrerem modificações. Assim, considerando-se conjuntos de grupos monofiléticos diagnosticáveis e bem sustentados que recebe-

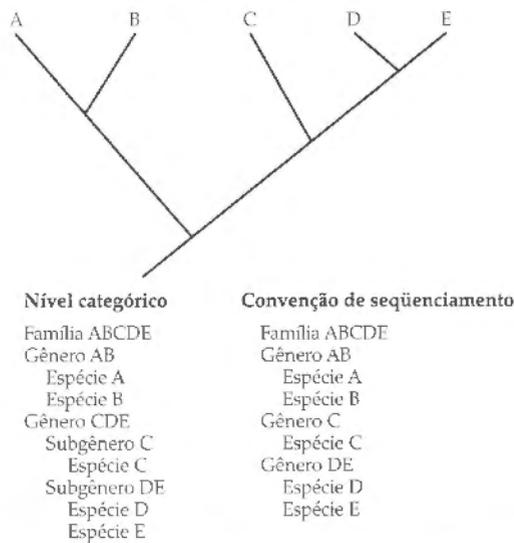


FIGURA 2.19 Classificações alternativas baseadas na filogenia de um grupo hipotético dos *taxa* A, B, C, D e E. A classificação à esquerda utiliza quatro categorias, ou níveis (família, gênero, subgênero e espécie); a classificação à direita usa apenas três níveis (família, gênero e espécie) acrescida de uma convenção de seqüenciamento.

ram um nome no passado, estes podem – e argumentamos que devem – permanecer com a nomeação. Esse é também um argumento contrário à aferição de nome formal ao clado PACCAD das gramíneas, o que levaria a um conjunto de modificações desnecessárias que atingiria usos taxonômicos padronizados e estabelecidos há bastante tempo (Backlund e Bremer 1998; Stevens 1998).

Categorias: as categorias são arbitrárias

Após decidir quais grupos monofiléticos receberão nomes, ainda é necessário definir exatamente como será realizado o processo de nomeação. Os grupos podem, por exemplo, ser numerados, e um índice geral pode listar quais elementos serão englobados por cada grupo numerado. Esta abordagem é similar ao sistema utilizado por companhias telefônicas para organizar os números de telefone. Naturalmente, a dificuldade encontrada é que, sem a existência de uma lista telefônica (um índice geral) e/ou uma excelente memória, o sistema torna-se inacessível.

A classificação biológica visa fornecer um vocabulário de trabalho que transmita informação filogenética, mas que mesmo assim possa ser compreendido por biólogos que não sejam prioritariamente sistematas. Visto que uma filogenia apresenta similaridades estruturais a uma hierarquia, na qual grupos menores estão inseridos em grupos maiores, os quais por sua vez estão incluídos em grupos ainda maiores, é sensato que a classificação reflita este fato sob a forma de uma hierarquia.

A classificação botânica usa um sistema desenvolvido no século XVIII, no qual *taxa* são definidos em posições (ou categorias) determinadas, como reino, filo, classe, ordem, família, gênero e espécie (i.e., as categorias de Linneu) (ver Capítulo 3 e Apêndice 1). Uma classificação dos grupos monofiléticos que foram nomeados deve ser logicamente consistente com as relações filogenéticas hipotetizadas para os organismos que estão sendo classificados (como expresso através da se-

qüência de pontos de ramificação em um cladograma). Ou seja, as categorias da classificação de Linneu são usadas para expressar relações entre grupos-irmãos.

Mesmo considerando-se que *taxa* monofiléticos representam grupos reais que existem na natureza como resultado do processo histórico da evolução, as categorias *per se* são apenas construções mentais.

Estes patamares de classificação apresentam significado apenas relativo (e não absoluto) (Stevens 1998). Em outras palavras, o nível de família é menos inclusivo do que o nível de ordem e mais inclusivo do que o nível de gênero, porém não existem critérios disponíveis para indicar que um determinado táxon, por exemplo, angiosperma, seja reconhecido no nível de filo, classe ou ordem.

Na Figura 2.19, um cladograma de *taxa* imaginários de A até E é primeiramente convertido para uma classificação hierárquica estabelecida de acordo com as categorias definidas por Linneu. Observe que o subgênero DE está alocado no interior do gênero CDE, que por sua vez está alocado no interior da família ABCDE. (Mas poderíamos ter tratado o clado ABCDE como ordem, o clado CDE como uma família e o clado DE como um gênero.) Muitas vezes, contudo, para expressar completamente as relações de grupos-irmãos (em um cladograma), são necessárias mais categorias do que as que se encontram disponíveis (em uma taxonomia hierárquica), mesmo após a criação de categorias adicionais mediante o uso dos prefixos *super-* e *sub-*.

Uma modificação do método de classificação aqui salientado é a **convenção de seqüenciamento**, a qual estabelece que *taxa* que formam uma porção assimétrica de um cladograma devem ser posicionados em uma mesma categoria e organizados de acordo com sua ordem de ramificação (Wiley 1979, 1981). Assim, na Figura 2.19, os conjuntos AB, C e DE podem ser designados como gêneros. A seqüência dos nomes na classificação denota a seqüência de ramificação no cladograma. Observe que essa convenção equivale a dizer que nem todos os grupos monofiléticos receberão nomes.

Mesmo considerando-se que a categorização é arbitrária, os critérios aqui descritos para a decisão dos grupos que devem ser nomeados podem também ser aplicados para decidir o nível no qual um determinado grupo será incluído (ver Stevens 1998 para uma discussão detalhada). Neste ponto, novamente a estabilidade nomenclatural assume um papel im-

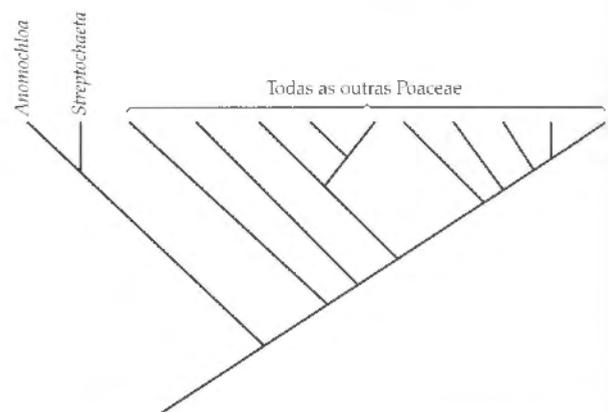


FIGURA 2.20 Filogenia das Poaceae indicando a posição dos gêneros *Anomochloa* e *Streptochaeta*.

portante. Por exemplo, recentemente foi demonstrado que a linhagem da família Poaceae que divergiu mais precocemente inclui apenas dois gêneros atuais, *Anomochloa* e *Streptochaeta* (Figura 2.20). Dessa forma, poderíamos, em princípio, criar uma nova família para *Anomochloa* e *Streptochaeta*, pois ela seria monofilética e também deixaria Poaceae como monofilética. Para fins de estabilidade, no entanto, é sensato deixar os dois gêneros em Poaceae, onde lhes foi conferido um nome em nível de subfamília: Anomochlooideae.

Alguns sistematas propuseram o abandono do sistema de Linneu e a substituição deste por uma "taxonomia filogenética". Uma completa exploração desta possibilidade não faz parte do escopo de nosso texto, mas abordaremos brevemente alguns dos argumentos contrários ao uso das categorias de Linneu.

Visto que as categorias são arbitrárias, um gênero (grupo de espécies) em uma família pode não possuir a mesma idade, compreender a mesma quantidade de variabilidade ou mesmo compartilhar qualquer outra coisa – além do fato de ambos serem grupos monofiléticos – com um gênero de outra família. Sistematas experientes em geral estão cientes disso (Darwin, por exemplo, estava ciente dessa situação) e sabem que gêneros, famílias, e assim por diante, não são unidades comparáveis (Stevens 1997). Alguns cientistas, no entanto, freqüentemente utilizam tais categorias como se fossem reais. Por exemplo, é comum medir a diversidade vegetal por meio da listagem do número de famílias representadas pela flora local, mesmo sabendo-se que a unidade *família* não possui um significado real em particular.

Se a categoria é arbitrária, uma conseqüência lógica seria a eliminação completa das categorias. Os *taxa* seriam arranjados em grupos, mas esses grupos não receberiam qualquer denominação como gênero, família, ordem ou outra categoria qualquer. Esse tipo de categorização já existe informalmente, sobretudo entre grupos acima do nível de ordem. As eudicotiledôneas, por exemplo, são amplamente reconhecidas como monofiléticas, mas não receberam uma categoria específica na hierarquia de Linneu. De forma similar, poucos sistematas se preocupam em definir se as angiospermas devem ser reconhecidas como divisão, classe, subclasse, superordem ou alguma outra categoria; elas são nitidamente monofiléticas e designadas pelo nome não-linneano de *angiospermas*.

A eliminação dos níveis ou categorias torna-se, no entanto, mais problemática entre ordens, famílias e gêneros. Grupos designados a tais categorias são corriqueiros e seus nomes usados comumente, o que torna pouco provável que uma forma inteiramente nova de nomenclatura seja aceita de imediato e sem protestos. Não obstante, um sistema alternativo de nomenclatura filogenética, conhecido como *PhyloCode*, está sendo desenvolvido. O *PhyloCode* é estruturado inteiramente à margem das regras do Código Internacional de Nomenclatura Botânica (ICBN), que governa o uso das categorias de Linneu e é, há muito tempo, usado por todos os taxonomistas vegetais (ver Apêndice 1). O *PhyloCode* é um sistema de nomenclatura alternativo, e não uma revisão do sistema atualmente existente (ver *Website* do *PhyloCode* em www.ohiou.edu/phylocode).

Um outro resultado dos estudos filogenéticos é a observação de que diversas filogenias encontram-se apenas parcialmente resolvidas, de tal forma que, considerando-se os dados disponíveis, um posicionamento exato dos *taxa* é impossível. Isto quer dizer que algumas espécies não podem ser posicio-

nadas com certeza em um gênero, e que alguns gêneros não podem ser fidedignamente designados como pertencentes a uma dada família. O sistema atual permite posicionamentos duvidosos acima da categoria de espécie, refletidos na categoria *incertae sedis* – literalmente, "de posição incerta". Uma alternativa seria um sistema sem categorias, no qual não seria necessário nem o posicionamento em um grupo maior, nem a outorga de nomes a todos os ramos de uma dicotomia ou politomia.

Os autores do presente texto estão envolvidos na reclassificação de gêneros, famílias e ordens com base em dados filogenéticos e consideram que – se uma filogenia está clara – o uso da hierarquia-padrão de Linneu é bastante simples (principalmente quando suplementado por nomes informais não-categorizados). Quando a filogenia não está nítida, é geralmente razoável esperar a disponibilização de novos dados antes de modificar a classificação.

Para uma discussão a respeito dos problemas encontrados no uso do sistema de Linneu para a classificação filogenética, consulte Wiley 1981; de Queiroz e Gauthier 1990, 1992; Wiley e colaboradores 1991; Forey e colaboradores 1992; e Hibbett e Donoghue 1998.

Comparando as classificações filogenéticas com as classificações derivadas do uso de outros métodos taxonômicos

Apesar de os métodos filogenéticos constituírem uma abordagem bastante disseminada, nem todos os taxonomistas os utilizam. Algumas sistematas postulam que, ao longo da evolução, eventos de paralelismo e reversão foram tão comuns que os detalhes da história evolutiva nunca serão decifrados. Este ponto de vista deu origem a uma escola de sistematas conhecida como *fenética*. Os feneticistas argumentam que, considerando-se que a história evolutiva nunca pode ser detectada de forma inequívoca, a melhor classificação para os organismos será a realizada por meio de sua similaridade geral. Dessa forma, organismos similares são posicionados em conjunto dentro de um mesmo grupo, ao passo que organismos muito diferentes são posicionados em grupos distintos (Sneath e Sokal 1973).

Uma grande dificuldade da abordagem fenética reside no fato de muitos sistematas terem produzido diagramas semelhantes a árvores, agrupando organismos por sua similaridade geral, e esses diagramas terem sido interpretados como se refletissem a história evolutiva. Algumas vezes esta abordagem leva a resultados similares aos resultados encontrados por uma análise filogenética, mas, em alguns casos, leva à produção de "grupos" compostos por organismos que compartilham apenas o fato de serem diferentes dos demais, inclusive dos próprios integrantes do grupo. Tais grupos mostram-se parafiléticos ou polifiléticos.

O desenvolvimento de métodos fenéticos foi um importante prelúdio para a aceitação e uso das abordagens filogenéticas. Um taxonomista que construía uma classificação fenética inicialmente observava com cuidado todos os caracteres possíveis. Esses caracteres eram divididos em estados, ou o valor quantitativo do caráter era anotado (p. ex., uma série de medidas de comprimento foliar era realizada e registrada a média para cada táxon). Esta informação era organizada

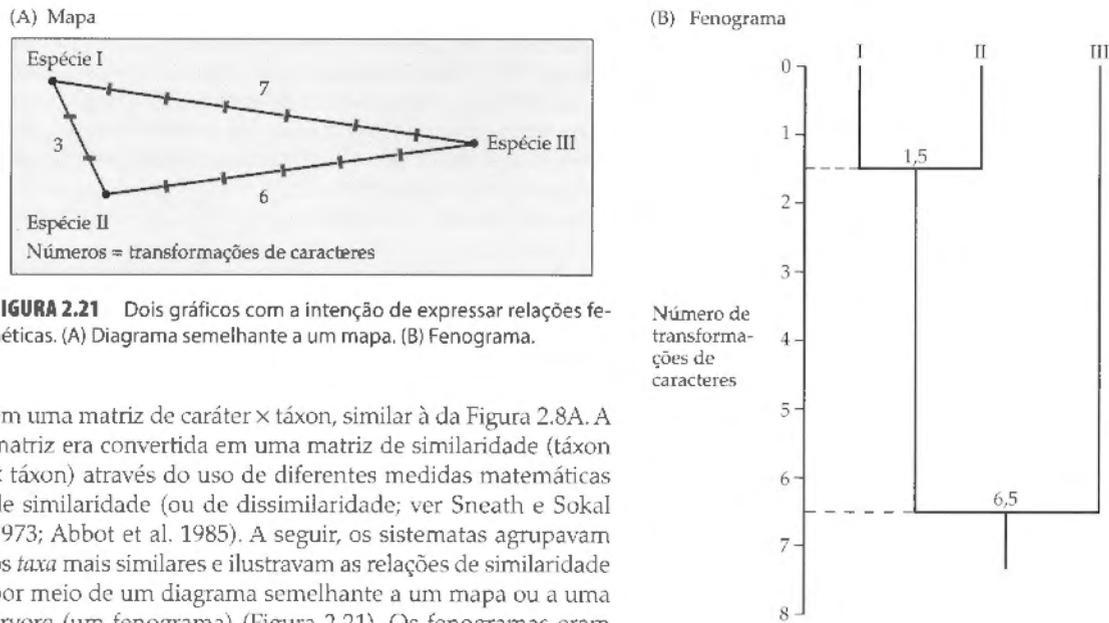


FIGURA 2.21 Dois gráficos com a intenção de expressar relações fenéticas. (A) Diagrama semelhante a um mapa. (B) Fenograma.

em uma matriz de caráter \times táxon, similar à da Figura 2.8A. A matriz era convertida em uma matriz de similaridade (táxon \times táxon) através do uso de diferentes medidas matemáticas de similaridade (ou de dissimilaridade; ver Sneath e Sokal 1973; Abbot et al. 1985). A seguir, os sistematistas agrupavam os *taxa* mais similares e ilustravam as relações de similaridade por meio de um diagrama semelhante a um mapa ou a uma árvore (um fenograma) (Figura 2.21). Os fenogramas eram construídos usando-se algoritmos de agrupamento, ao passo que os diagramas semelhantes a mapas eram provenientes de estudos de ordenação que empregavam análises estatísticas multivariadas (ver Abbot et al. 1985).

Muitas das classificações produzidas a partir de métodos fenéticos são úteis para a identificação e localização de informações. No entanto, essas classificações não foram desenvolvidas com o objetivo de recuperar a história evolutiva e, portanto, não são adequadas para o posicionamento de questões evolutivas. Sistemas fenéticos não fazem distinção entre sinapomorfias e evolução paralela ou convergente.

A **taxonomia evolutiva** também diferia da taxonomia filogenética em termos da abordagem de classificação. A similaridade morfológica de um grupo era de extrema importância, ao passo que monofilia e parafilia (no estrito senso cladístico destas palavras) eram secundárias. Assim, um grupo podia ser reconhecido com base em uma combinação única de caracteres derivados e ancestrais compartilhados (Figura 2.22). Um importante fato era o reconhecimento de "lacunas" no padrão de variação entre grupos filogeneticamente adjacentes (Simpson 1961; Ashlock 1979; Cronquist 1987; Mayr e

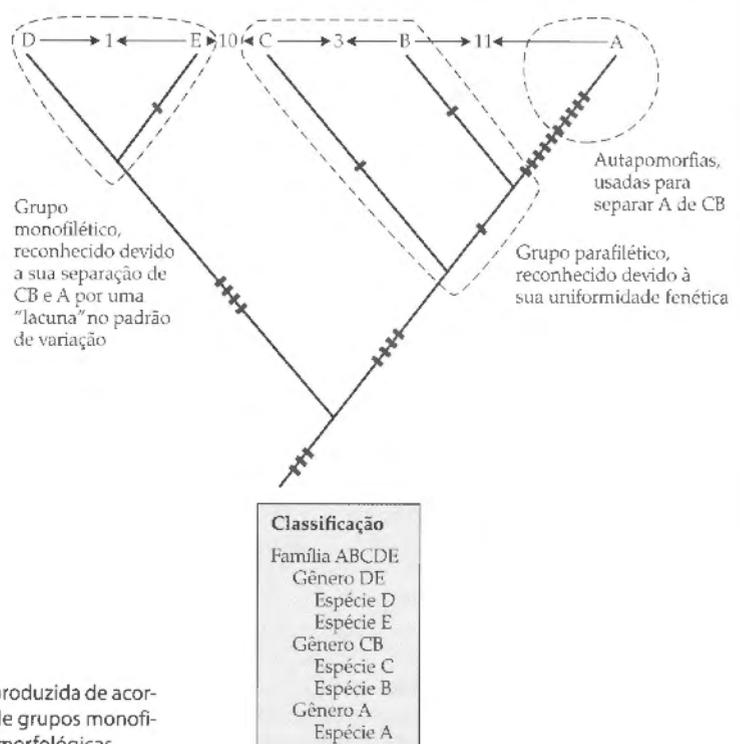


FIGURA 2.22 Filogenia e uma classificação não-filogenética produzida de acordo com a escola evolutiva. A classificação inclui uma mistura de grupos monofiléticos e parafiléticos, separados uns dos outros por "lacunas" morfológicas.

Ashlock 1991). Caracteres considerados evolutivamente (ou ecologicamente) significativos foram extremamente discutidos, e a competência, a autoridade e a intuição dos sistematistas foram fundamentais neste debate. Em conclusão, apesar das classificações evolutivas geralmente se referirem à evolução e apesar de os grupos reconhecidos nestas classificações serem geralmente monofiléticos, era esperado que os taxa fossem morfologicamente homogêneos e que estivessem separados uns dos outros por lacunas discretas (Ashlock 1979; Stuessy 1983, 1990; Stevens 1986; Mayr e Ashlock 1991).

Já foi dito que a sistemática é tanto uma arte quanto uma ciência (apesar da afirmação trazer à tona o questionamento

sobre as definições de arte e ciência), de um lado porque existem muitos aspectos desta disciplina que ainda parecem não possuir uma base objetiva. Um fato importante na sistemática filogenética é que pelo menos um dos principais aspectos da sistemática – a delimitação de grupos – sofreu uma formalização que é geralmente aceita e que pode ser seguida. Ao contrário das classificações evolutivas e fenéticas, que apresentam ambigüidades em termos de critérios de agrupamento, as classificações filogenéticas são precisas. Um grupo denominado pode ser considerado monofilético, incluindo todos os descendentes a partir de um único ancestral comum.

BIBLIOGRAFIA CITADA E LEITURAS RECOMENDADAS

Itens marcados com asterisco são especialmente recomendados para os leitores interessados em informações adicionais sobre os assuntos discutidos neste capítulo.

- Abbot, L. A., F. A. Bisby and D. A. Rogers. 1985. *Taxonomic analysis in biology*. Columbia University Press, New York. [Uma introdução a diversos métodos fenéticos.]
- Albert, V. A., A. Backlund, K. Bremer, M. W. Chase, J. R. Manhart, B. D. Mishler and K. C. Nixon. 1994. Functional constraints and *rbcL* evidence for land plant phylogeny. *Ann. Missouri Bot. Gard.* 81: 534-567.
- Ashlock, P. D. 1979. An evolutionary systematist's view of classification. *Syst. Zool.* 28: 441-450. [Apresentação de um método explícito e de simples compreensão para construir uma classificação taxonômica evolutiva.]
- *Backlund, A. and K. Bremer. 1998. To be or not to be: Principles of classification and monotypic plant families. *Taxon* 47: 391-400.
- Bremer, K. 1983. Angiosperms and phylogenetic systematics. *Verh. Naturwiss. Verein Hamburg* 26: 343-354. [Uma discussão sobre evolução reticulada e cladística.]
- Bremer, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795-803. [O índice de decaimento.]
- Bremer, K. and H.-E. Wänntorp. 1979. Hierarchy and reticulations in systematics. *Syst. Zool.* 28: 624-627.
- *Brooks, D. R. and D. A. McLennan. 1991. *Phylogeny, ecology, and behavior*. University of Chicago Press, Chicago. [Uma apresentação excelente dos usos biológicos para hipóteses filogenéticas.]
- Burleigh, J. G. and S. Mathews. 2004. Phylogenetic signal in nucleotide data from seed plants: Implications for resolving the seed plant tree of life. *Am. J. Bot.* 91: 1599-1613.
- Cantino, P. D., H. N. Bryant, K. de Queiroz, M. J. Donoghue, T. Eriksson, D. M. Hillis and M. S.Y. Lee. 1999. Species names in phylogenetic nomenclature. *Syst. Biol.* 48: 790-807.
- *Cracraft, J. and M. J. Donoghue. 2004. *Assembling the tree of life*. Oxford University Press, Oxford and New York.
- Cronquist, A. 1987. A botanical critique of cladism. *Bot. Rev.* 53:1-52.
- *Dahlgren, R. and F. N. Rasmussen. 1983. Monocotyledon evolution: Characters and phylogenetic estimate. In *Evolutionary biology*, vol. 16, M. K. Hecht, B. Wallace and G. T. Prance (eds.), 255-395. Plenum Press, New York. [Uma introdução simples para a metodologia cladística.]
- Davis, P. D. and V. H. Heywood. 1963. *Principles of angiosperm taxonomy*. Krieger, New York.
- *de Queiroz, K. and J. Gauthier. 1990. Phylogeny as a central principle in taxonomy: Phylogenetic definitions of taxon names. *Syst. Zool.* 39: 307-322. [Uma proposta para abandonar o sistema linneano.]
- de Queiroz, K. and J. Gauthier. 1992. Phylogenetic taxonomy. *Annu. Rev. Ecol. Syst.* 23: 449-480.
- de Queiroz, A., M. J. Donoghue and J. Kim. 1995. Separate versus combined analyses of phylogenetic evidence. *Annu. Rev. Ecol. Syst.* 26: 567-581.
- *Donoghue, M. J. and P. D. Cantino. 1988. Paraphyly, ancestors, and the goals of taxonomy: A botanical defense of cladism. *Bot. Rev.* 54: 107-128.
- Doyle, J. A., M. J. Donoghue and E. A. Zimmer. 1994. Integration of morphological and ribosomal RNA data on the origin of angiosperms. *Ann. Missouri Bot. Gard.* 81: 419-450.
- *Eldredge, N. and J. Cracraft. 1980. *Phylogenetic patterns and the evolutionary process: Methods and theory in comparative biology*. Columbia University Press, New York.
- *Farris, J. S. 1974. Formal definitions of paraphyly and polyphyly. *Syst. Zool.* 23: 548-554. [Aqui um grupo polifilético é definido como "um grupo onde o ancestral em comum mais recente é designado não ao próprio grupo, mas em outro".]
- Farris, J. S. 1979. The information content of the phylogenetic system. *Syst. Zool.* 28: 458-519.
- Farris, J. S. 1989. *Hennig86*, Version 1.5. Port Jefferson Station, NY.
- Felsenstein, J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.* 27: 401-410.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* 17: 368-376.
- Felsenstein, J. 1989. *PHYLIP 3.2 manual*. University of California Herbarium, Berkeley.
- *Forey, P. L., C. J. Humphries, I. L. Kitching, R. W. Scotland, D. J. Siebert and D. M. Williams. 1992. *Cladistics: A practical course in systematics*. Oxford University Press, Oxford. [Um resumo dos métodos cladísticos em uso naquela época.]
- *Frohlich, M. W. 1987. Commonis-primitive: A partial validation by tree counting. *Syst. Bot.* 12: 217-237. [O princípio diz que havendo um grupo interno e um grupo externo, ambos com estados A e B de um caráter homólogo, se o estado A é muito mais comum que o B no grupo externo, então o estado A é provavelmente ancestral no grupo interno.]
- Funk, V. A. 1985. Phylogenetic patterns and hybridization. *Ann. Missouri Bot. Gard.* 72: 681-715.
- Gift, N. and P. F. Stevens. 1997. Vagaries in the delimitation of character states in quantitative variation: An experimental study. *Syst. Biol.* 46: 112-125.
- *Givnish, T. J. and K. J. Sytsma. 1997. Homoplasy in molecular vs. morphological data: The likelihood of correct phylogenetic inference. In *Molecular evolution and adaptive radiation*, T. J. Givnish and K. J. Sytsma (eds.), 55-101. Cambridge University Press, Cambridge.
- Goloboff, P. A. 1993. *NONA*, Version 1.5.1. P. A. Goloboff, Tucumán, Argentina.
- Hall, B. G. 2005. *Phylogenetic trees made easy*, 2nd ed. Sinauer Associates, Sunderland, MA.
- *Hennig, W. 1966. *Phylogenetic systematics*. University of Illinois Press, Urbana.
- Hibbett, D. and M. J. Donoghue. 1998. Integrating phylogenetic analysis and classification in fungi. *Mycologia* 90: 347-356.
- Hillis, D. M., M. W. Allard and M. M. Miyamoto. 1993. Analysis of DNA sequence data: Phylogenetic inference. *Methods Enzymol.* 224: 456-87.
- *Huelsenbeck, J. P. 1995. Performance of phylogenetic methods in simulation. *Syst. Biol.* 44: 17-48. [Comparação de máxima verossimilhança, parcimônia e semelhança geral em simulações de computador para gerar árvores filogenéticas.]
- *Huelsenbeck, J. P. and D. M. Hillis. 1993. Success of phylogenetic methods in the fourtaxon

- case. *Syst. Biol.* 42: 247-264. [Muitos tipos de métodos baseados na parcimônia são relativamente bem sucedidos na reconstrução de filogenias]
- Humphries, C. J. and V. A. Funk. 1982. Cladistic methodology. In *Current concepts in plant taxonomy* (Systematics Association Special Volume, No. 25), V. H. Heywood and D. M. Moore (eds.), 323-362. Academic Press, London. [Uma introdução às metodologias básicas da cladística.]
- Humphries, C. J. and L. R. Parenti. 1986. *Cladistic biogeography*. Clarendon, Oxford.
- Kellogg, E. A. 1989. Comments on genomic genera in the Triticeae. *Am. J. Bot.* 76: 796-805.
- Kellogg, E. A., R. Appels and R. J. Mason-Gamer. 1996. When genes tell different stories: The diploid genera of Triticeae (Gramineae). *Syst. Bot.* 21: 321-347.
- Kitching, I. J., P. L. Forey, C. J. Humphries and D. M. Williams. 1998. *Cladistics: The theory and practice of parsimony analysis*, 2nd ed. Oxford University Press, Oxford. [Um resumo dos métodos para construir hipóteses filogenéticas baseadas na parcimônia, incluindo tópicos relacionados com os caracteres.]
- Kron, K. A. and W. S. Judd. 1997. Systematics of the *Lyonia* group (Andromedeae, Ericaceae) and the use of species as terminals in higher level cladistic analyses. *Syst. Bot.* 22: 479-492.
- *Maddison, D. R. and W. P. Maddison. 2000. *MacClade: Analysis of phylogeny and character evolution*, Version 4.0. Sinauer Associates, Sunderland, MA. [Um programa muito útil para explorar a evolução de caracteres em um cladograma.]
- Maddison, W. P. M. J. Donoghue and D. R. Maddison. 1984. Outgroup analysis and parsimony. *Syst. Zool.* 33: 83-103.
- Mayr, E. and P. D. Ashlock. 1991. *Principles of systematic zoology*, 2nd ed. McGraw-Hill, New York.
- *McDade, L. A. 1990. Hybrids and phylogenetic Systematics. I. Patterns of character expression in hybrids and their implications for cladistic analyses. *Evolution* 44: 1685-1700.
- *McDade, L. A. 1992. Hybrids and phylogenetic Systematics. II. The impact of hybrids on cladistic analyses. *Evolution* 46: 1329-1346.
- McDade, L. A. 1997. Hybrids and phylogenetic Systematics. III. Comparison with distance methods. *Syst. Bot.* 22: 669-683.
- Nelson, G. and N. Platnick. 1981. *Systematics and biogeography*. Columbia University Press, New York.
- Pankhurst, R. J. 1991. *Practical taxonomic computing*. Cambridge University Press, Cambridge. [A utilização de computadores e bases de dados em análises cladísticas e fenéticas.]
- *Quicke, D. L. J. 1993. *Principles and techniques of contemporary taxonomy*. Blackwell, London.
- Remane, A. 1952. *Die Grundlagen des natürlichen Systems, der vergleichenden Anatomie und der Phylogenetik* (The principles of the natural system, comparative anatomy, and phylogenetics). Geest & Portig, Leipzig.
- Rydin, C., M. Källersjö and E. M. Friis. 2002. Seed plant relationships and the systematic position of Gnetales based on nuclear and chloroplast DNA. *Int. J. Plant Sci.* 163: 197-214.
- *Sanderson, M. J. and M. J. Donoghue. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781-1795.
- Schuh, R. T. 2000. *Biological Systematics: Principles and applications*. Cornell University Press, Ithaca, NY. [Uma cobertura detalhada da metodologia cladística e de sua aplicação em biogeografia, ecologia e conservação.]
- Simpson, G. G. 1961. *Principles of animal taxonomy*. Columbia University Press, New York.
- Sneath, P. H. A. and R. R. Sokal. 1973. *Numerical taxonomy*. Freeman, San Francisco.
- Sokal, R. R. 1986. Phenetic taxonomy: Theory and methods. *Annu. Rev. Ecol. Syst.* 17: 423-442.
- *Sokal, R. R. and F. J. Rohlf. 1980. An experiment in taxonomic judgment. *Syst. Bot.* 5: 341-365.
- Stevens, P. F. 1980. Evolutionary polarity of character states. *Annu. Rev. Ecol. Syst.* 11: 333-358.
- *Stevens, P. F. 1984. Homology and phylogeny: Morphology and Systematics. *Syst. Bot.* 9: 395-409.
- Stevens, P. F. 1986. Evolutionary classification in botany, 1960-1985. *J. Arnold Arb.* 67: 313-339.
- *Stevens, P. F. 1991. Character states, morphological variation, and phylogenetic analysis: A review. *Syst. Bot.* 16: 553-583.
- Stevens, P. F. 1997. How to interpret botanical classifications: Suggestions from history. *BioScience* 47: 243-250.
- Stevens, P. F. 1998. What kind of classification should the practising taxonomist use to be saved? In *Plant diversity in Malesia III: Proceedings of the 3rd International Flora Malesiana Symposium 1995*, J. Dransfield, M. J. E. Coode and D. A. Simpson (eds.), 295-319. Royal Botanical Gardens, Kew, London.
- Stuessy, T. F. 1983. Phylogenetic trees in plant Systematics. *Sida* 10: 1-13.
- Stuessy, T. F. 1990. *Plant taxonomy*. Columbia University Press, New York.
- *Swofford, D. L. 1993. *PAUP: Phylogenetic analysis using parsimony*, Version 3.1.1. Distributed by the Illinois Natural History Survey, Champaign.
- *Swofford, D. L. 2000. *PAUP*: Phylogenetic analysis using parsimony and other methods*, Version 4.0. [Edição teste Beta, distribuída por Sinauer Associates, Sunderland, MA.]
- *Swofford, D. L., G. J. Olsen, P. J. Waddell and D. M. Hillis. 1996. Phylogenetic inference. In *Molecular Systematics*, 2nd ed., D. M. Hillis, C. Moritz and B. K. Mable (eds.), 407-514. Sinauer Associates, Sunderland, MA. [Um excelente resumo dos métodos de construção de árvores filogenéticas.]
- Sytsma, K. J. and J. C. Pires. 2001. Plant systematics in the next 50 years-Remapping the new frontier. *Taxon* 50: 713-732.
- *Wagner, W. H., Jr. 1980. Origin and philosophy of the groundplan-divergence method of cladistics. *Syst. Bot.* 5: 173-193.
- Wagner, W. H., Jr. 1983. Reticulistics: The recognition of hybrids and their role in cladistics and classification. In *Advances in cladistics: Proceedings of the second meeting of the Willi Hennig Society*, N. I. Platnick and V. A. Funk (eds.), 63-79. Columbia University Press, New York.
- Watrous, L. E. and Q. D. Wheeler. 1981. The outgroup comparison method of character analysis. *Syst. Zool.* 30: 1-11.
- Weins, J. J., ed. 2000. *Phylogenetic analysis of morphological data*. Smithsonian Institution Press, Washington, DC.
- Wiley, E. O. 1979. An annotated Linnaean hierarchy, with comments on natural taxa and competing systems. *Syst. Zool.* 28: 308-337.
- *Wiley, E. O. 1981. *Phylogenetics*. Wiley, New York. [Uma detalhada discussão dos princípios da cladística.]
- *Wiley, E. O., D. Siegel-Causey, D. R. Brooks and V. A. Funk. 1991. *The complete cladist: A primer of phylogenetic procedures* (University of Kansas, Museum of Natural History, Special Publication, No. 19). University of Kansas, Museum of Natural History, Lawrence. [Um resumo dos métodos cladísticos em uso nessa época.]
- Wofford, B. E. 2006. A new species of *Stenanthium* (Melanthiaceae) from Tennessee, U.S.A. *Sida* 22: 447-459.
- Zomlefer, W. B. and W. S. Judd. Resurrection of segregates of the polyphyletic genus *Zigadenus* s.l. (Liliales: Melanthiaceae) and resulting new combinations. *Novon* 12: 299-308.
- Zomlefer, W. B., N. H. Williams, W. M. Whitten and W. S. Judd. 2001. Generic circumscription and relationships in the tribe Melanthieae (Liliales, Melanthiaceae), with emphasis on *Zigadenus*: Evidence from ITS and *trnL-F* sequence data. *Am. J. Bot.* 88: 1657-1669.