

Dados Comportamentais na Era da Genômica

JOHN W. WENZEL¹ E FERNANDO B. NOLL²

¹*Ohio State University, USA*

²*UNESP - São José do Rio Preto*

Muitos acreditam que com o seqüenciamento do DNA não será mais necessária qualquer outra informação, como a relativa ao comportamento, para inferir passos e origens evolutivas. Em nossa opinião, sempre precisaremos de caracteres mais complexos, como os caracteres comportamentais, para a realização de estudos evolutivos. Estudos evolutivos ou filogenias sem referência aos caracteres fenotípicos são necessariamente incompletos. Não há sentido em estudar-se o DNA sem referência ao fenótipo. Com o advento da genômica, o estudo do comportamento pode parecer estar em segundo plano. Entretanto, argumentos importantes em favor desse estudo podem ser levantados.

Palavras-chave: Comportamento. Evolução. Genômica. Estudos evolutivos. Cladística.

Behavioral data in a genomics age. Many people think that, due to the possibility of DNA sequencing in large quantities, other aspects, such as behavior, are useless for inferring evolutionary paths or origins. In our opinion, we will always need more complex characters, as behavioral characters, for evolutionary studies. Any evolutionary study, any phylogeny, without reference to phenotypic characters are incomplete. There is no point in studying DNA without reference to the phenotype. The advent of genomics may appear to put in a second rank. There are, however, several important arguments on the behalf of behavioral information that need to be raised.

Keywords: Behavior. Evolution. Genomics. Evolutionary studies. Cladistics

O estudo da evolução sempre foi baseado na difícil tarefa de encontrar semelhanças especiais entre as características dos organismos, através de ampla comparação. Esta tarefa tem vários desafios. O primeiro é ter uma boa amostra de espécies disponível, de modo que possamos incluir a diversidade e a complexidade dos organismos atuais, e também representar algum tipo de forma mais ancestral e mais antiga. Outro problema é dividir organismos em partes que possam mostrar histórias evolutivas di-

ferentes em profundidades temporais diferentes, de modo que algumas características possam unir tetrápodes terrestres, características mais recentes unir mamíferos, e características muito mais recentes identificar roedores sul-americanos. Um terceiro problema é definir os caracteres com precisão suficiente para que possa distinguir-se a homologia (origem única compartilhada por um mesmo antepassado) da convergência (aparência semelhante apesar de origens diferentes). Há certamente mais pro-

John W. Wenzel, Museum of Biological Diversity, Department of Entomology, The Ohio State University, 1315 Kinnear Road, Columbus, OH 43212. E-mail: wenzel.12@osu.edu Fernando B. Noll, autor correspondente, Departamento de Zoologia e Botânica; Instituto de Biociências, Letras e Ciências Exatas, UNESP; Rua Cristóvão Colombo, 2265; 15054-000, São José do Rio Preto, SP; Brazil. Voice: 55-17-32212373. Fax: 55-17-32212374. E-mail: noll@ibilce.unesp.br

Trabalho apresentado no XIII Encontro Anual de Etologia - Assis - SP 12 a 15 de Novembro de 2005. Apoio financeiro: The Ohio State University, Fapesp (proc. 01/02491-4). Agradecemos à Fapesp (proc. 01/02491-4) e à Ohio State University pelo financiamento desse e de outros trabalhos. Embora essa pretenda ser uma revisão abrangente, reconhecemos o argumento de um dos revisores de que existem muitas pesquisas que esse trabalho deixou de mencionar.

blemas do que estes, uma vez que biólogos de diferentes áreas procuram questões diferentes, mas estes três desafios devem ser superados em quase todo estudo comparativo de qualidade. É às vezes difícil definir os elementos morfológicos (a base da maior parte da taxonomia) necessários para responder aos problemas mencionados. O crescimento da tecnologia de DNA parece fornecer uma solução. Por muitas razões diferentes, temos todos aprendido a considerar, com respeito, os dados de DNA como facilmente caracterizados, como representativos do delineamento técnico real da herança, etc. Muitas pessoas pensam que porque podemos seqüenciar DNA em grandes quantidades, não temos mais necessidade de recorrer a outros aspectos, como os dados comportamentais, para inferir passos evolutivos ou origens (Scotland, Olmstead & Bennett, 2003). Pensam que se pudemos captar DNA em quantidade suficiente, acharemos a resposta correta a todas as perguntas sem ter que levar em conta outros tipos de variações. Não acreditamos que seja verdadeira essa posição. Sempre necessitaremos medir caracteres mais complexos, como os traços comportamentais, para realizar estudos evolutivos. Frequentemente, quando um pesquisador ignora o comportamento, ele deixa de usar dados valiosos e talvez não consiga resolver o problema que se propôs abordar e fracasse em sua meta original.

É fácil demonstrar que se superestima, atualmente, as possibilidades da biologia genômica. Diz-se que, agora que conhecemos o genoma do camundongo e o do ser humano, sabemos o que é ser um mamífero. Isto foi declarado com orgulho e publicado nas melhores revistas, como confirmação da validade dos enormes projetos de genoma em andamento, atualmente. Mas os dados disponíveis não chegam a tanto porque não sabemos se as semelhanças entre camundongo e Homem são devidas a eles serem mamíferos, ou amniotas, ou tetrápodes, ou vertebrados, ou deuterostômios, etc. O fato de a comunidade científica deixar de perceber quão inútil é ter apenas dois genomas para comparação mostra que os entusiasmos pelos genes é superestimado. Naturalmente, logo teremos mais de dois genomas de mamíferos, mas os genomas de per se não nos fornecerão os dados de que necessitamos por-

que a maioria de biólogos além de tentar entender genomas, procura entender os fenótipos dos organismos existentes.

Qualquer estudo evolutivo, ou qualquer filogenia, não tem valor sem referência a caracteres fenotípicos. O fenótipo é o que queremos explicar, e a única coisa de valor para os morfologistas, ecologistas, behavioristas, e outros biólogos incluindo geneticistas que estudam a regulação dos genes (ou seja, fenótipo), ou RNA (ou seja, fenótipo), ou muitos outros campos considerados mais precisos que a morfologia. A única razão pela qual queremos saber a respeito dos genes de uma mosca é porque estamos interessados pela biologia das moscas. Sabemos o que elas fazem, e isso torna seus genes interessantes. Não há sentido estudar o DNA sem referência a um fenótipo. Quem duvida disto pode executar a seguinte experiência: tome a filogenia de um grupo interessante (digamos: a *Drosophila* havaiana, ou mamíferos sul americanos, ou organismos causadores de doenças), e peça a alguém apagar todos os nomes das espécies das pontas da árvore. Quão interessantes são os genes agora? Ninguém se preocupará com os dados ou com a árvore evolutiva se não houver referência a um fenótipo.

Dados genéticos são úteis por causa dos dados fenotípicos, e a conexão entre eles é muito mais contínua do que costuma acreditar-se. É claro que se quer saber a respeito de características, tais como os aminoácidos, que os genes produzem. Os aminoácidos são caracteres fenotípicos, eles não são exatamente sinônimos da seqüência de DNA (há redundância parcial, uma vez que mais de uma seqüência de DNA pode produzir o mesmo aminoácido). Estruturas morfológicas são presumivelmente codificadas pelo DNA, mas elas também são influenciadas por fatores ambientais ou efeitos epigenéticos que não estão representados na seqüência do gene que mais de perto se relaciona à estrutura morfológica. Pode-se ir mais longe e dizer que é preciso levar em conta que a variação regular não é codificada no DNA. Muitos aspectos comportamentais podem cair nesta categoria, e haverá quem queira excluir de consideração comportamentos que não possam ser relacionados a uma seqüência de DNA.

Pensamos, ao contrário, que as análises evolutivas são aperfeiçoadas quando incluem uma informação histórica que satisfaça o conceito “sine qua non”. Se o caractere de interesse for eliminado, então a espécie de interesse cessaria de existir em qualquer forma significativa. Como mostraremos mais adiante, muitos caracteres ecológicos e comportamentais satisfazem o critério de “sine qua non” sem ser exclusivamente relacionado à seqüência de DNA. O desafio principal é definir um conceito preciso de homologia de caractere, evitando, quando possível, caracteres convergentes, compostos.

Alguns temem que a convergência seja típica dos caracteres fenotípicos, particularmente pela ação da adaptação. Por exemplo, Hedges e Maxon (1996) dizem que

a diferença qualitativa entre dados moleculares e morfológicos envolve uma convergência adaptativa. Os dados morfológicos são mais suscetíveis à convergência adaptativa, como mostram muitos casos notáveis de camuflagem e mimetismo e, por causa disso, são menos apropriados como base para a estimativa da filogenia. (p. 313)

Dentro desta perspectiva, comportamentos ou características ecológicas semelhantes não teriam necessariamente a mesma origem. Mas, a propriedade da “semelhança” é comum a todos os conjuntos de caracteres. O DNA também está sujeito à convergência, talvez mais do que a morfologia ou o comportamento, numa comparação individual, porque as bases do DNA limitam-se a apenas quatro (A, C, G, T), ao passo que o número de morfologias e comportamentos que podem ser medidos talvez não tenha limite. Há, por exemplo, pelo menos 3.000.000 de espécies atualmente distinguidas com base nas abordagens tradicionais da morfologia, do comportamento, ou da ecologia. A alegação mais comum, além disso, é que populações morfológicamente semelhantes podem ser distinguidas a partir de dados de DNA: deve-se inferir que caracteres convencionais são mais conservadores do que convergentes. Decorre que não convém, por medo da convergência, excluir características ecológicas ou comportamentais da análise evolutiva.

Um motivo mais forte para não levar em conta o fenótipo seria que os fenótipos são redundantes em relação aos genótipos (ver a discussão desta questão em Freudenstein, Pickett, Simmons & Wenzel, 2003), e que os genótipos são mais fáceis de se caracterizar (Scotland et al., 2003). Consideremos, contudo, a estrutura secundária das proteínas. Estes produtos adaptáveis e funcionais parecem ser informativos, mas não são simplesmente uma tradução das bases de DNA encontrados nos códons. Os geneticistas parecem concordar que características secundárias das enzimas são significativas, e uma procura simples na literatura sugere que geneticistas freqüentemente preferem seqüências de aminoácidos (ou outras características secundárias) aos dados de DNA. É razoável, então, declarar que seja dado ao comportamento o mesmo status que é dado aos dados de aminoácidos quando o comportamento parece desempenhar um papel semelhante. Em alguns casos, temos boa evidência de que o poder funcional dos comportamentos é que guia os dados genéticos; a genética populacional vem mais tarde e é redundante em relação ao comportamento. Por exemplo, muitas espécies de animais herdam um grande impulso para migrar, mas os locais que constituem o “lar” não são herdados geneticamente. Na América do Norte, ovos de salmões transportados de um local para outro produzem indivíduos fiéis a sua nova situação. Os peixes nadam para locais antigos de alimentação, no oceano, e retornam ao seu rio natal, ano após ano, estabelecendo uma população local. É por isso que os pescadores são capazes de reintroduzir o salmão em rios onde se extinguíram as populações originais estão extintas. A propensão para migrar é herdada, mas não é herdado o caminho. É razoável supor que diferenças na genética populacional decorrerão de diferenças no comportamento de retornar ao rio natal. Se a redundância é um argumento, então às vezes o comportamento suplanta o DNA como indício dos fatores que moldam uma população. Não seria certamente correto dizer que os caracteres comportamentais que influenciam os dados de DNA têm status menor que os do DNA que eles alteraram. No mínimo, se os geneticistas admitem a estrutura secundária (fenotípica) como um meio adequado para a interpretação de seqüências diversas de DNA, então também se deve admitir os pa-

drões comportamentais que norteiam a identidade de uma seqüência de DNA.

Um desafio semelhante surge com a formação de tradições culturais, através de comportamentos aprendidos, persistindo por longos períodos. Em muitos pássaros que aprendem seus cantos, os dialetos são aparentemente mais velhos que a genética populacional dos pássaros que os cantam. De acordo com revisões recentes, em cinco famílias de três Ordens de pássaros, dialetos locais persistentes de canções podem ser mais estáveis que a subjacente variação genética populacional (Freudenstein et al., 2003). Os dialetos podem ter sido constituídos em primeiro lugar, seguindo a estrutura populacional. Neste caso, seria redundante e desnecessário medir diferenças de DNA, porque quando indicam uma estrutura, elas seguem as diferenças mais óbvias do canto, e quando não indicam uma estrutura, então não são significativas. Esta situação não vale igualmente em todos os pássaros, mas vale para um número suficiente deles: a existência e a estabilidade dos cantos aprendidos de pássaros mostra o quanto o estudo do DNA deve ser acompanhado de um estudo do comportamento.

Quando tentamos reconstituir uma história evolutiva, precisamos reconstruir estados intermediários dentre as alternativas atualmente presentes. Temos um problema sério se as espécies que representam intermediários (parcialmente com características ancestrais, mas com alguns caracteres de todas as espécies modernas) perderam-se devido à extinção, não sendo representadas claramente entre as espécies que temos atualmente. Este é um problema comum quando os intermediários são fósseis que já foram descobertos e colocados em uma seqüência evolutiva. O exemplo mais famoso é o dos dinossauros: Sem fósseis, a fácil conexão entre os répteis modernos e os pássaros seria difícil de ser estabelecida. Entretanto, esse problema é mais grave no caso do DNA. O DNA evolui escrevendo sucessivamente novas seqüências sobre seqüências antigas: os estados intermediários serão perdidos uma vez que novos processos evolutivos os modifiquem. Uma vez que o fenótipo independa de alguma forma do genótipo, poderá preservar alguma informação que se perdeu das seqüências de DNA.

Os fenótipos desempenham parte do papel que os fósseis desempenham fornecendo níveis de conexão que seriam impossíveis de se achar no DNA. É por isso que os caracteres que reconhecemos no nível do fenótipo têm complexidade suficiente para que seja possível encontrar semelhanças especiais entre espécies, mesmo quando elas não são idênticas. Deste ponto de vista, o comportamento pode ser visto como um tipo de fóssil, com respeito à evolução genética que se moveu rápido demais para ser traçada diretamente.

Por exemplo, algumas moscas fitófagas, que são pragas agrícolas, atacam apenas certas plantas. Algumas delas moscas atacam laranjas e outras próximas atacam pastinaca (erva européia da família da cenoura). Em um primeiro momento, essas plantas parecem não ter nada em comum, uma vez que elas são de famílias diferentes. Mas, ambas as plantas produzem, em seus tecidos, aldeído anísico e seus derivados (Ehrlich & Raven, 1967). Sabendo isso, o comportamento faz sentido porque essas moscas procuram derivados do aldeído anísico (outras moscas não), algumas espécies buscam um dos derivados, outras buscam outro. A ecologia química destas espécies produz provas incontestáveis de uma adaptação que evoluiu de acordo com a preferência por derivados do aldeído anísico. Esta característica comportamental é muito informativa, e é difícil acreditar que tal conexão não seja sempre reconhecida no estudo de DNA das moscas. Qual gene seqüenciar? Como saber quando se encontrou algo informativo?

Sejam duas lagartas norte-americanas que parecem semelhantes e ambas comem as folhas de árvores da floresta. Uma come *Quercus*, outra come *Fagus*. São semelhantes, ou diferentes? Se souber que ambas as árvores pertencem à Família Fagaceae, então as espécies são semelhantes porque ambas comem Fagaceae. Mas, podemos acreditar que são diferentes porque não comem Fagaceae idênticas. Então, são semelhantes, mas diferentes, dependendo do que seja uma terceira lagarta. Assim como mostramos com o camundongo e o homem no começo do artigo, uma comparação de dois termos não tem sentido, mas ao menos quando dizemos “come *Fagus*” capturamos os dados eficientemente, o que não é necessariamente o caso

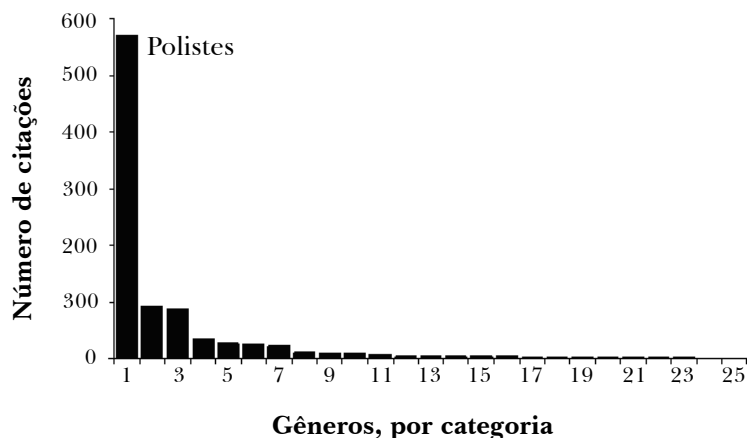


Figura 1. Citações de trabalhos científicos entre 1980-2004 para os 25 gêneros de vespas da subfamília Polistinae, obtidas através da base de dados SCI (Science Citation Index Expanded) utilizando o nome do gênero no título, resumo ou palavras-chave (925 trabalhos no total).

quando tratamos de seqüências genômicas mitocondriais. Este tipo de descoberta da história evolutiva é mais fácil fazer com comportamento e ecologia do que com o DNA. Há mais níveis de ordem disponíveis, há mais o que encontrar rapidamente. O comportamento e a ecologia fornecem uma informação rica que não é facilmente substituída pela tecnologia de DNA.

Obviamente, o comportamento e a ecologia não são apenas um modo eficiente de resumir dados. Representam os dados que queremos explicar. Como dito anteriormente, não faz sentido estudar o DNA se não sabemos o que ele faz. Devemos então estudar o fenótipo. Mas, o que realmente estudamos? A Figura 1 mostra o número de citações para todas as vespas sociais da subfamília Polistinae, obtido em busca numa base de dados científica. Nota-se que a grande maioria dos trabalhos foi escrita sobre o gênero *Polistes*. *Polistes* é o gênero-modelo, o único que se acredita nos dar uma visão da evolução do comportamento social em vespas em geral. Como alguns behavioristas gostam de trabalhar com modelos de sistemas, eles

tendem a trabalhar sempre e sempre com a mesma espécie. Esse é um método eficiente, mas após 10 ou 20 anos, não explica como estudar uma vespa diferente. Mesmo assim, acredita-se que qualquer coisa que aprendemos sobre *Polistes* nos dirá algo sobre as outras vespas.

Eis uma forma diferente de representar os mesmos dados. A Figura 2 mostra um cladograma dos 25 gêneros de Polistinae, e os ramos estão desenhados com larguras proporcionais ao número de citações de cada gênero. Note que *Polistes* é mais do que a metade, e os 10 gêneros principais perfazem 96% dos trabalhos. A maioria dos gêneros não têm literatura substancial. É difícil imaginar que esse tipo de programa científico seja a melhor forma de explicar a diversidade da vida. Na melhor das hipóteses estamos explicando *Polistes*, e geralmente as mesmas espécies de *Polistes*, não todas as 200 espécies, no mundo todo. É possível que esse problema da restrição a certos modelos explanatórios piore mais e mais, uma vez que a pesquisa tem sido conduzida pela ciência da genômica. Será uma conseqüência negativa

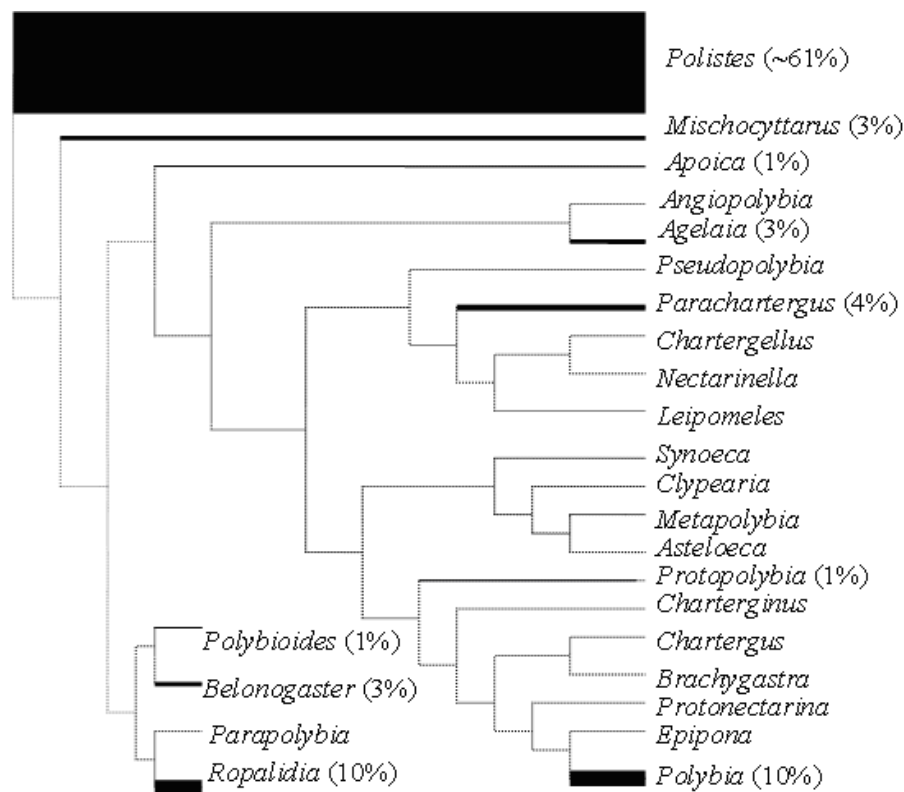


Figura 2. Distribuição dos trabalhos científicos, vistos na Figura 1, de acordo com uma perspectiva evolutiva, exibido na forma de um cladograma. A largura das linhas é igual à proporção das citações. Os principais gêneros correspondem a 96% de todas as citações.

para países como Brasil que têm muita história natural para ser feita. Quando tivermos o genoma da abelha produtora de mel (*Apis mellifera*), quantos outros genomas das 20.000 outras abelhas serão financiados? Já vemos a comunidade científica sendo sugada pelo estudo de moléculas de espécies-exemplo, freqüentemente de laboratórios de cultura, mais do que estudar os reais fenótipos de interesse em espécies que queremos conhecer.

Parece que se cria facilmente confusão sobre o que é um estudo de comportamento. Por exemplo, Cameron (1993) apresentou o que

se supõe ser um estudo da evolução do complexo comportamento social de abelhas. As abelhas têm um comportamento muito sofisticado, e há muitas variedades de abelhas com comportamentos muito variados. Atualmente, elas são representadas por pelo menos 20.000 espécies, algumas das quais tem evoluído há mais de 90 milhões de anos. Temos inclusive fósseis excelentes em âmbar, de modo que é uma boa condição para se levar adiante estudos evolutivos. O estudo de Cameron está, contudo, limitado a dados derivados de três genes mitocondriais e a um gene nuclear, e de não mais que 15 espécies, sete das quais não mos-

tram comportamento complexo. Surgem pelo menos dois problemas: o primeiro é que o DNA mitocondrial talvez tenha sido sobrescrito muitas vezes nos últimos 90 milhões de anos (o DNA não seria então o melhor dado, para esta análise), e o segundo é que inexistem dados reais relacionados ao comportamento. Noll (2002), por outro lado, apresentou um cladograma de abelhas, incluindo 23 espécies, baseado em 65 comportamentos evolutivamente novos. Este estudo é mais substancial, mesmo sem dados de DNA. Não pode haver dúvida que a árvore representada na figura resulta de um estudo real do comportamento. Dado que o comportamento seja o foco, em que seria melhor embasar o estudo, nas seqüências de DNA, ou na variação comportamental?

Em conclusão: são várias as razões para a importância de se levar em conta o comportamento, em estudos básicos. A mais óbvia é que comportamentos representam os fenômenos naturais que queremos explicar. Temores sobre convergência não se aplicam apenas ao comportamento (dados de DNA podem sofrer mais convergência). O comportamento é importante porque ele tanto pode se relacionar estreitamente ao genótipo, como não. O comportamento pode ilustrar etapas intermediárias que foram perdidas ao nível de seqüência de bases de DNA, ou retratar princípios gerais que a análise genômica não encontra. Enquanto a biologia for, em muitos casos, guiada pelo estudo de organismos-modelo, não será possível explicar a diversidade da vida ao nosso redor, e quanto mais atenção for dada aos modelos, maior a

disparidade. Fazemos um apelo geral ao retorno dos estudos de comportamento para o cerne dos programas de biologia evolutiva básica. Como a maioria dos biólogos está caminhando em direção aos estudos genômicos, ou de DNA, precisamos preservar nossa perícia em interpretar fenótipos.

Referências

- Cameron, S. A. (1993). Multiple origins of advanced eusociality in bees inferred from mitochondrial DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 90(18), 8687-8691.
- Ehrlich, P. R., & Raven, P. H. (1967). Butterflies and plants. *Scientific American*, 216, 104-113.
- Freudenstein, J. V., Pickett, K. M., Simmons, M. P., & Wenzel, J. W. (2003). From basepairs to birdsongs: phylogenetic data in the age of genomics. *Cladistics*, 19, 333-348.
- Hedges, B., & Maxson, L. (1996). Re: Molecules and morphology in amniote phylogeny. *Molecular Phylogenetics and Evolution*, 2, 312-314.
- Noll, F. B. (2002). Behavioral phylogeny of corbiculate Apidae (Hymenoptera; Apinae), with special reference to social behavior. *Cladistics*, 18(2), 137-153.
- Scotland, R. W., Olmstead, R. G., & Bennett, J. R. (2003). Phylogeny reconstruction: The role of morphology. *Systematic Biology*, 52, 539-548.

Recebido em 14 de dezembro de 2005

Revisão recebida em 20 de março de 2006

Aceito em 24 de agosto de 2006