

CLADÍSTICA NUMÉRICA, ANÁLISIS SIMULTÁNEO Y FILOGENIA DE HEXÁPODOS

James M. Carpenter y Ward C. Wheeler

Division of Invertebrate Zoology,
American Museum of Natural History,
Central Park West at 79th Street,
New York, NY 10024-5192, U.S.A.

Resumen

Se revisan análisis cladísticos de las interrelaciones de los órdenes de hexápodos y análisis simultáneos de datos combinados morfológicos y moleculares.

Palabras clave: Hexapoda, Cladística, Análisis morfológico, Análisis molecular.

Numerical Cladistics, Simultaneous Analysis and Hexapod Phylogeny

Abstract

Numerical cladistic analyses of the interrelationships of hexapod orders, and simultaneous analyses of combined morphological and molecular data, are reviewed.

Key words: Hexapoda, Cladistics, Morphological analyses, Molecular analyses.

INTRODUCCIÓN

La aplicación de métodos cladísticos a las relaciones ordinales de los insectos comenzó con Hennig (1953). En su contribución, Kristensen (1995), enfatizó la importancia del trabajo de Hennig y revisó los avances producidos durante los 40 años anteriores. Entre esos avances se encuentra el artículo del propio Kristensen (1975), seguramente el trabajo más importante sobre la filogenia ordinal de insectos. En él se establece la base de las modernas investigaciones sobre las relaciones filogenéticas ordinales, jugando un destacado papel en el trabajo subsiguiente, incluidos los recientes artículos que discutiremos a continuación centrados en la aplicación de la cladística cuantitativa al problema de definir las relaciones ordinales de los insectos. Junto con nuestros colaboradores Michael Whiting y Quentin Wheeler, hemos emprendido el primer análisis cladístico numérico de este problema, trabajando con datos morfológicos y moleculares, así como el primer análisis simultáneo de datos morfológicos y moleculares (Whiting et al., 1997; Wheeler et al., 1999). Pasaremos revista a estos artículos, y otros en curso, presentándolos a efectos ilustrativos en tres partes: discusión de los datos morfológicos, incluyendo un trabajo anterior y la primera matriz a nivel de orden, discusión de los datos moleculares, contrastando nuestra aproximación con trabajos previos, incluyendo la primera matriz con todos los órdenes representados y finalmente la combinación de estos con los datos morfológicos en un análisis simultáneo.

Históricamente, los sistemas de caracteres únicos han formado la base para la agrupación supraordinal entre hexápodos. Ejemplos clásicos incluyen la división de los Pterygota

(Martynov, 1925) basada en los campos alares y la separación de los Thysanura basada en la articulación mandibular (por Snodgrass, 1938). Ejemplos más recientes incluyen a Baccetti (1979) con espermatozoides, von Paulus (1979) con la ultraestructura ocular, Pass (1991) con las estructuras circulatorias antenales, Štys et al. (1993) con ovarios, etc. El trabajo de Kristensen (1975, 1981, 1991, 1995, 1997) es notable por integrar muchos estudios de caracteres simples de forma tradicional, mediante la argumentación Hennigiana, que es manual, carácter por carácter, nodo por nodo. Todos estos estudios han hecho avanzar en gran medida nuestro conocimiento sobre los hexápodos.

La ventaja de la cladística numérica desde este punto de vista es clara: permite el tratamiento simultáneo de toda la evidencia disponible. A veces se ha afirmado que la investigación funcional de sistemas de caracteres simples es una aproximación mejor (por ejemplo, Manton, 1977), como si las conclusiones hipotéticas sobre las funciones fuesen un indicador filogenético mejor que la similitud observada. La llamada "aproximación histórica" de Kukalová-Peck (1997) es una versión moderna y extrema de consentir conjeturas sobre relaciones para dictar qué constituyen los datos. Si los caracteres analizados en un estudio cladístico son independientes el uno del otro, entonces cualquier carácter es *a priori* tan idóneo para ser un buen indicador de la filogenia como cualquier otro. La teoría básica del muestreo indica que deben combinarse y analizarse tantos caracteres como sea posible, de forma simultánea, para conseguir una hipótesis robusta de relaciones filogenéticas.

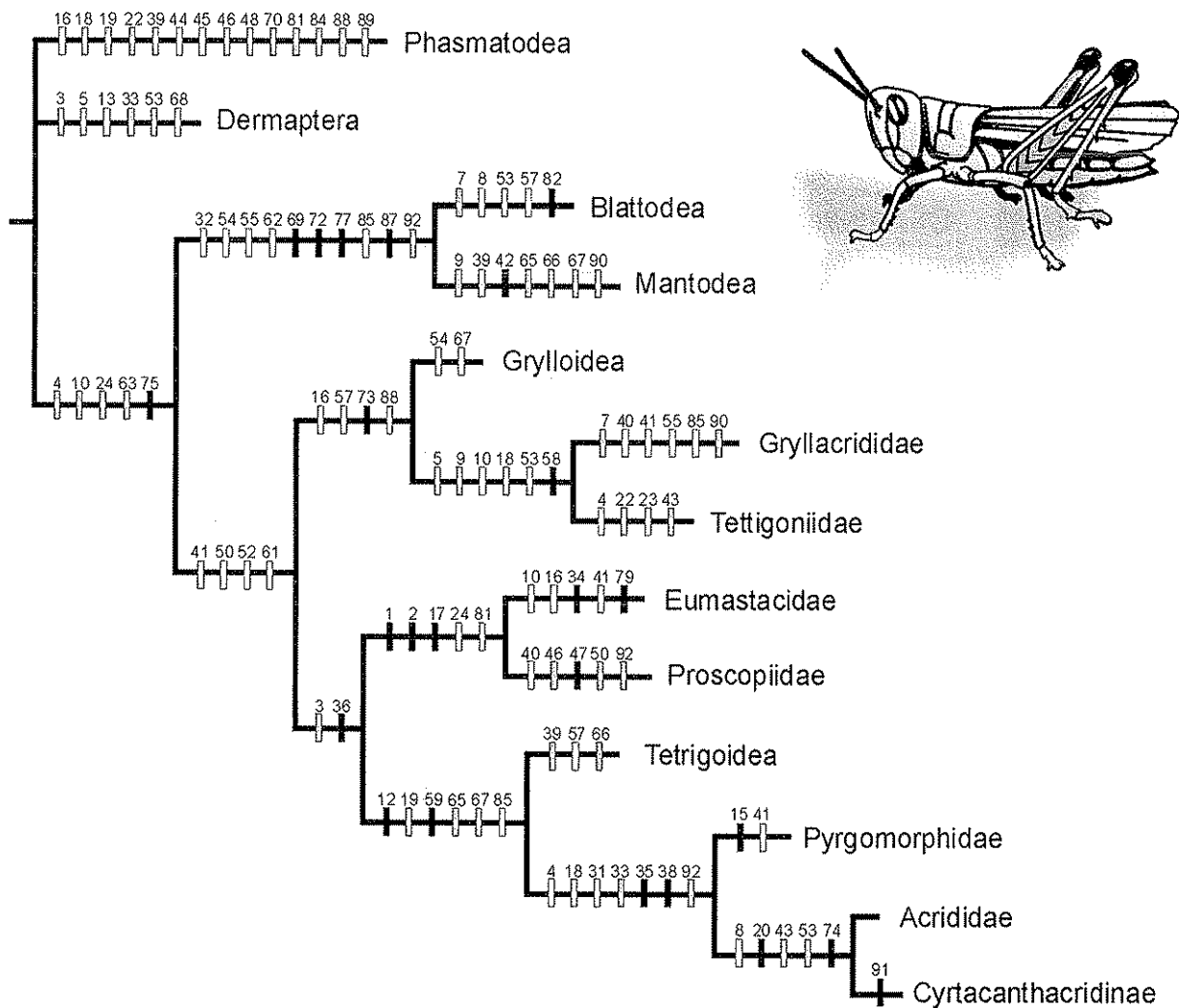


Fig. 1.- Cladograma para los datos de Blackith y Blackith (1968). El cladograma está calculado con la rutina exacta de Hennig86 o NONA y tiene longitud 180, índice de consistencia de 0.51 e índice de retención de 0.57. Las marcas negras representan cambios no homoplásicos y las marcas blancas y grises denotan homoplasias en los caracteres numerados. Las optimizaciones y las figuras fueron dibujados usando las opciones por defecto de CLADOS (Nixon, 1998), que traza únicamente los cambios no ambiguos. El cladograma fue enraizado mediante insectos palo y tijeretas, de esta forma se produce la conclusión de Blackith y Blackith sobre la parafilia de Orthoptera en términos de cucarachas y mántidos.

Fig. 1.- Cladogram for the data of Blackith and Blackith (1968). The cladogram is calculated with the exact routine of either Hennig86 or NONA, and has a length of 180, consistency index of 0.51 and retention index of 0.57. Black hashmarks represent non-homoplastic changes, and the white and gray hashmarks denote homoplasies in the numbered characters. Optimizations and figure were plotted using CLADOS (Nixon, 1998) defaults, which maps only unambiguous changes. The cladogram was rooted to walking sticks and earwigs - thus, allowing Blackith and Blackith's conclusion of parafily of Orthoptera in terms of cockroaches and mantids to be at least a possibility.

La aproximación llamada "total evidence" (Kluge, 1989) o análisis simultáneo (Nixon y Carpenter, 1996) es, en efecto, un argumento en contra de la superioridad de cualquier tipo de dato particular - el viejo argumento a menudo implicado en los estudios de caracteres únicos, ya refutado en los días de Linneo. Independencia de caracteres significa que diferentes tipos de caracteres, como los morfológicos y moleculares, deberían ser también combinados (Kluge, 1989; Nixon y Carpenter, 1996). Así, el análisis simultáneo maximiza el poder explicatorio —parsimonia. Como ha sido apuntado por Nixon y Carpenter (1996), las metodologías llamadas "congruencia taxonómica" y "combinación condicional" (ver, Zrzavý et al., 1997) simplemente abandonan la parsimonia. Por lo tanto abogamos por análisis simultáneos, que son los que aplicamos en nuestro análisis.

DATOS MORFOLÓGICOS

La primera matriz cuantitativa de datos morfológicos sobre relaciones supraordinales en ser analizada cladísticamente fue

la de Thorne y Carpenter (1992) sobre Dictyoptera (termitas, cucarachas y mántidos). Blackith y Blackith (1968) habían ofrecido previamente lo que llamaron un "cladograma" para algunos órdenes de ortopteroides, mostrando Orthoptera como parafilético respecto a mántidos y cucarachas. Su "cladograma" estaba basado en el método de Camin y Sokal (1967), el cual asume la irreversibilidad de la evolución de los caracteres. Sin embargo, si esto no se asume y se reanalizan sus datos —cladísticamente— se obtiene el cladograma de la Fig. 1, con los Orthoptera monofiléticos.

Thorne y Carpenter estaban influidos por dos arraigadas cuestiones: la ubicación de la cucaracha de la madera, socialmente primitiva, *Cryptocercus* y la de *Mastotermes*, el fósil viviente de la termita semejante a una cucaracha. Ambos taxones juegan un papel clave en la teoría subsocial de la evolución de la sociabilidad en termitas como supuestos estados intermedios (por ejemplo, Wilson, 1971). Thorne y Carpenter llevaron a cabo un escrutinio superficial de caracteres de la bibliografía, y construyeron una matriz de datos con 70 caracteres morfológicos y comportamentales para siete

terminales (tres Isóptera, dos Blattaria, un Mantodea y la raíz). La matriz fue analizada con la rutina exacta del programa Hennig86 (Farris, 1988), lo cual dio como resultado un único cladograma. Sorprendentemente, este cladograma soportaba la hipótesis de Boudreaux (1979) de que cucarachas y mántidos son grupos hermanos, en relación a las termitas —de esta manera se refutaba la antigua teoría subsocial para la evolución de la sociabilidad de las termitas.

Kristensen (1995) cuestionó este resultado y cambió la codificación de los caracteres en la matriz, aunque, como admitió, subjetivamente. El efecto de esas manipulaciones fue que no se podía soportar correctamente ninguna conclusión cuando se aplicó Hennig86 a la matriz cambiada. Consideramos por lo tanto adecuado que Kristensen empleara cladística asistida por ordenador y aunque podríamos ofrecer argumentos sobre las recodificaciones, en lugar de eso demostraremos que los resultados de Thorne y Carpenter están bien apoyados en el contexto de los análisis de todos los órdenes y además en combinación con los datos moleculares. Thorne y Carpenter incluyen un ancestro hipotético en la raíz de su cladograma (pre-enraizamiento, en la terminología de Nixon y Carpenter, 1993), pero mostraremos que las cucarachas son todavía soportadas como un grupo hermano de los mántidos sin este nodo artificial.

Nuestra matriz actual también se basa primordialmente en la bibliografía, especialmente en los artículos de Kristensen, Hennig (1981) y Boudreaux (1979), pero también en algunos artículos recientes de autores tales como Minet y Bourgoin (1986), Štys y Bilinski (1990) y unos pocos de Kukalová-Peck (por ejemplo, 1987) — solamente aquellos que pudimos verificar, los cuales precisaron descartar muchas de sus argumentaciones (por ejemplo, confirmar Kukalová-Peck y Lawrence, 1993, con Whiting y Kathirithamby, 1995). Las codificaciones para los terminales son *groundplans* hipotéticos: hemos intentado codificar con la mayor extensión posible los estados primitivos para los órdenes propuestos por los autores en la bibliografía. El propósito por lo tanto fue el de dar a las afirmaciones sobre relaciones realizadas por estos autores la más favorable de las oportunidades posible en los análisis numéricos. Obviamente, esto también conlleva la excusión de polimorfismos. Parte de la matriz fue publicada en el análisis de Whiting et al. (1997) centrando la atención en Holometabola. La matriz tiene actualmente 256 variables para 32 órdenes de insectos y un miriápodo como grupo externo.

Sobre el grupo externo, debemos reconocer que las relaciones entre las clases de artrópodos son controvertidas, con muchos autores como Manton (1977) defensores de la no monofilia de los hexápodos. No hay razón para tomar alguno de sus argumentos seriamente; no son cladísticos. Superficialmente son más espectaculares los análisis como los de Ballard et al. (1992), que publicó un árbol supuestamente basado en datos de ADNmt 12S, con los artrópodos parafiléticos (en relación a los onicóforos) y los miriápodos en la base del árbol. Pero como han mostrado Carpenter et al. (1998), el árbol publicado por Ballard et al. en realidad estaba basado en manipulaciones de sus datos: el árbol consenso para los datos reales muestra muy poca estructura -grupos tales como los onicóforos y arácnidos son incluso no monofiléticos. El soporte de los Myriapoda como el pariente más cercano de los insectos está fundado en el análisis de la combinación de datos morfológicos y secuencias de datos de Wheeler et al. (1993), aunque "Myriapoda" puede ser parafilético, con sólo una parte de ellos grupo hermano de Hexapoda (Wheeler, 1997; Kraus, 1997; véase también, sin embargo, Giribet et al., 1999 en este volumen).

Cuando se ha analizado con el programa NONA (Goloboff, 1998) la presente matriz morfológica, han resultado dos cladogramas, con una longitud de 450 pasos, índice de consistencia 0,71 e índice de retención 0,81. El árbol de consenso se muestra en la figura 2; una discusión adicional de los resultados puede ser encontrada en Wheeler et al. (1999).

Hay cuatro zonas sobre las cuales debemos llamar la atención:

- 1) La situación de Diplura. El árbol consenso es ambiguo, porque uno de los cladogramas fundamentales sostiene Entognatha (Diplura, Collembola y Protura) como un grupo monofilético (por ejemplo, Hennig, 1981), mientras que el otro sitúa Diplura como grupo hermano de los Insecta *sensu stricto* (por ejemplo, Kukalová-Peck, 1987). Así, hasta donde la matriz morfológica alcanza, las relaciones fundamentales de Hexapoda permanecen ambiguas (por ejemplo, Kristensen, 1991). Podemos añadir que la parafilia de los Diplura fue sugerida por Štys y Bilinski (1990) que propusieron que los Campodeina son más cercanos a los Ellipura que Japygina. No obstante, cuando Diplura es dividido en la matriz morfológica los análisis entonces producen básicamente los dos mismos cladogramas, ambos con Diplura monofiléticos (véase Wheeler et al., 1999). Los caracteres morfológicos sostienen, en efecto, la monofilia de Diplura, pero hasta que los datos moleculares son añadidos (véase más adelante), su situación es inestable.
- 2) Relaciones de Dictyoptera. Los mántidos y cucarachas como grupos hermanos en relación con las termitas están confirmados (Boudreaux, 1979; Thorne y Carpenter, 1992). Las relaciones en el interior de los Dictyoptera han permanecido no obstante controvertidas. Por ejemplo, Willmann (1997:269) describe a Klass (1995) como un "trabajo monumental", que muestra que "los termes son de hecho un subgrupo de cucarachas y que la cucaracha áptera *Cryptocercus* es el grupo hermano de los termes". Esta caracterización contrasta con Grandcolas (1997:234), que afirma: "Klass (1995) presenta un árbol para Isoptera, Blattodea y Mantodea sobre la base del estudio de 14 especies, sin embargo, no fue construido de acuerdo a un genuino análisis filogenético (sin matriz de caracteres, sin grupos externos, caracteres polimórficos no codificados como tales, etc.)..." La declaración de Willmann pasa por alto la demostración de Grandcolas y Deleporte (1992) y Grandcolas (1994, 1996) de que *Cryptocercus* es un pariente derivado, una cucaracha Polyphaginae, y que Polyphagidae es un grupo monofilético - exclusivo de termitas. Nuestros resultados, basados en la morfología, moléculas o ambos (véase más adelante), están en línea con esto.
- 3) Posición de Zoraptera. Este orden es situado como grupo hermano de Eumetabola (Acercaria = hemipteroides + Holometabola). El orden ha sido sugerido, de diferentes modos, como grupo hermano de Acercaria (Hennig, 1981), grupo hermano de Dictyoptera (Boudreaux, 1979), grupo hermano de Embiidina (Minet y Bourgoin, 1986) y recientemente como grupo hermano de Holometabola (Rasnitsyn, 1998). La situación con Eumetabola no está lejos de las relaciones sugeridas por Hennig (1981) o Rasnitsyn (1998), pero, como veremos, los datos moleculares no coinciden con esto.

- 4) Situación de Strepsiptera. El grupo es visto como uno panórpido, grupo hermano de Antliophora (pulgas, moscas y moscas escorpión). Esta solución a la "cuestión de los Strepsiptera" es quizás el problema más sorprendente de los caracteres morfológicos. Tradicionalmente, por supuesto, este orden ha sido considerado cercano a Coleoptera, aunque ciertamente no forma parte de él. Sobre la revisión de los caracteres relevantes, sin embargo, resulta claro que la posición tradicional está basada en una evidencia muy débil: un carácter une Strepsiptera con Coleoptera, el vuelo con las alas posteriores (Kukalová-Peck (1997) se equivocó con la demostración de Whiting y Kathirithamby (1995) de que Kukalová-Peck y Lawrence (1993) no mostraron sinapomorfias en la venación de las alas posteriores de Strepsiptera con Coleoptera, pues no suministró soporte a los datos de su argumentación). Ningún carácter une Strepsiptera con Coleoptera + Neuroptera, pero en contraste, varios caracteres los unen a los órdenes de panorpidos, e incluso algunos los unen de forma próxima a Antliophora (ver detalles en Whiting et al., 1997; Wheeler et al., 1999). Los datos moleculares sostenían previamente una relación de grupo hermano entre Strepsiptera y Diptera (Whiting y Wheeler, 1994; Whiting et al., 1997) y de hecho los datos morfológicos aislados no están lejos en estar de acuerdo con esto —y el análisis combinado (Whiting et al., 1997, y véase más adelante).

DATOS MOLECULARES

Nos ocuparemos ahora de los datos moleculares. La primera aplicación de datos moleculares a este problema es el artículo de Wheeler (1989), quien analizó una muestra limitada de una secuencia de ADN ribosómico y un mapa de restricción de 11 representantes de diez órdenes de insectos. Encontró soporte para Insecta, agrupación de Odonata con Neuroptera, Eumetabola, Holometabola y Neuroptera como grupo hermano de Coleoptera.

Diversos estudios moleculares subsiguientes han sido un paso atrás en términos de muestreo, aún cuando se han analizado más secuencias. Los artículos de Liu y Beckenback (1992) y Carmean et al. (1992) no respaldan la monofilia de Holometabola; Pashley et al. (1993) no la han comprobado. El primer artículo basado en un muestreo de gran amplitud de Holometabola fue el de Whiting y Wheeler (1994), los cuales describieron que Strepsiptera es el grupo hermano de las moscas. Sus datos de secuencia de ADN son incluidos en el presente estudio. Es suficiente decir aquí que el artículo de Whiting y Wheeler ha sido criticado, por ejemplo por Carmean y Crespi (1995), los cuales sugieren que el resultado fue un artefacto de la llamada atracción de las ramas largas (*long-branch attraction*) — basado en su propio muestreo extremadamente limitado. Whiting et al. (1997) demostraron que esta crítica es engañosa, basada en la amplia muestra que es parte de los resultados que discutimos aquí. Otros ataques relacionados con *long-branch attraction*, basados en análisis de máxima verosimilitud (Huelsenbeck, 1997, 1998) han tenido éxito únicamente en mostrar las deficiencias en los métodos de la máxima verosimilitud (Whiting, 1998; Siddall, 1998), métodos que no han sido en ningún caso justificados lógicamente (Farris, 1999).

Volviendo a los datos moleculares actuales para todos los órdenes de insectos, el estudio de Wheeler et al. (1999) incluye secuencias de datos de dos genes, ADNr 18S y 28S. La muestra consta de 121 secuencias de hexápodos, con cuatro

miriápodos y dos crustáceos secuenciados como grupos externos (Tabla 1). El tamaño del fragmento del 18S fue aproximadamente de 1000 pares de bases, mientras que el tamaño del 28S fue aproximadamente de 350 pares de bases. Todos los órdenes de hexápodos fueron representados con las dos moléculas y para la mayoría de los órdenes se secuenció más de un ejemplar. El total es aproximadamente 200Kb de datos de secuencias. Más detalles, incluido el acceso a los números de las secuencias del *Genbank*, pueden ser encontrados en Wheeler et al. (1999).

El análisis de las secuencias planteó el problema de la alineación, que equivale a especificar la homología en el nivel de la secuencia. Las alineaciones son hechas insertando *gaps* en el interior de secuencias de diferente longitud, comportando inserciones y deleciones de pares de bases. Tanto si se hace subjetivamente ("a simple vista") o a través de una función de coste, el proceso da una interpretación sobre los eventos evolutivos. Dado que las interpretaciones son arbitrarias, hemos usado el análisis de sensibilidad, como lo describe Wheeler (1995). En el análisis de sensibilidad, los parámetros usados para hacer la alineación son variados, construyéndose diferentes alineaciones; el proceso es repetido y el rango completo de alineaciones es analizado cladísticamente. El quid de la cuestión no está en determinar qué combinación de parámetros, tales como la relación entre *gaps* y cambios o el coste de la transición-transversión, es la mejor (más congruente), sino el determinar hasta qué punto los resultados pueden ser sensibles a valores arbitrarios de estos parámetros.

Nuestro procedimiento de alineación también aplicó la aproximación "optimización directa" de Wheeler (1996), o alineación directa. La función de coste fue minimizada mediante la construcción de un hipotético intermedio taxonómico entre las secuencias, así como sobre un cladograma, en lugar de una alineación general. La optimización directa obvia la creación de caracteres *gap* para alinear secuencias en una matriz y forma cladogramas más parsimoniosos que la alineación múltiple (Wheeler, 1996).

Los análisis fueron realizados con el programa POY (Gladstein y Wheeler, 1997) en el ordenador en paralelo del AMNH. Nuestro criterio para elegir de entre los resultados del análisis de sensibilidad el "mejor", es la congruencia. La congruencia, que es concordancia entre los dos conjuntos de datos moleculares, es medida por el índice ILD de Mickevich y Farris (1981; véase Farris et al., 1994), y es simplemente la homoplasia extra o longitud requerida del árbol cuando los datos combinados son analizados cladísticamente.

Los resultados de los análisis de datos moleculares se muestran en la Fig.3. Se encontró un único cladograma de coste 5170, basado en una ratio de transición/transversión de 1:1 y un ratio de *gap*/cambio (*indel*) de 2:1, lo que maximiza la congruencia (minimiza la incongruencia). Para detalles adicionales, véase Wheeler et al. (1999). Respecto a los cuatro puntos mencionados para la morfología:

- 1) Situación de Diplura. Hexapoda no es soportado como un grupo, con Diplura situado como un grupo hermano de los grupos externos de miriápodos.
- 2) Relaciones en Dictyoptera. Se soporta a los Mántidos y cucarachas como grupos hermanos de termitas.
- 3) Situación de Zoraptera. Zoraptera se muestra como parte de Amphiesmenoptera (Lepidoptera + Thrichoptera), un resultado decididamente nuevo que no concuerda con datos morfológicos ni con resultados de análisis combinados.

Tabla 1. Taxones usados en Wheeler et al. (1999); véanse Fig. 3-4**Table 1.** Taxa used in Wheeler et al. (1999); see Figs. 3-4.

Crustacea	<i>Balanus</i> sp. <i>Callinectes</i> sp.	Coleoptera	<i>Cybister fimbriolatus</i> <i>Xyloryctes faunus</i> <i>Octinodes</i> sp. <i>Photuris pennsylvanicus</i> <i>Rhipiphorus fasciatus</i> <i>Meloe proscarabaeus</i> <i>Tenebrio molitor</i> <i>Tetraopes tetropthalmus</i>
Myriapoda		Neuroptera	<i>Lolomyia texana</i> <i>Mantispa pulchella</i> <i>Hemerobius stigmata</i> <i>Chrysoperla plorabunda</i> <i>Myrmeleon immaculatus</i> <i>Myrmeleon</i> sp.
Chilopoda	<i>Scutigera coleoptrata</i> <i>Lithobius forficatus</i>	Megaloptera	<i>Corydalus cognatus</i> <i>Sialis hamata</i>
Diplopoda	<i>Spiroboldus</i> sp. <i>Megaphyllum</i> sp.	Raphidiodea	<i>Agulla</i> sp.
Hexapoda		Hymenoptera	<i>Hartigia cressonii</i> <i>Orussus thoracicus</i> <i>Hemitaxonus</i> sp. <i>Periclista linea</i> <i>Bareogonolus canadensis</i> <i>Evania appendigaster</i> <i>Ichneumon</i> sp. <i>Ophion</i> sp. <i>Mesopolobus</i> sp. <i>Caenochrysis doriae</i> <i>Epyris sepulchralis</i> <i>Priocnemus oregana</i> <i>Dasymutilla gloriosa</i> <i>Apoica</i> sp. <i>Monobia quadridens</i> <i>Polistes fuscatus</i> <i>Polistes dominulus</i> <i>Camponotus ligniperda</i> <i>Chalepoxenus muellerianus</i> <i>Doronomyrmex kutteri</i> <i>Leptothorax acervorum</i> <i>Temnothorax recedens</i> <i>Harpagoxenus sublaevis</i>
Collembola	<i>Pseudachorutes</i> sp. <i>Crossodonthina koreana</i> <i>Hypogastrura dolsana</i> <i>Podura aquatica</i> <i>Lepidocyrtus paradoxus</i>	Lepidoptera	<i>Papilio troilus</i> <i>Galleria mellonella</i> <i>Ascalapha odorata</i>
Protura	<i>Nipponentomon</i> sp.	Trichoptera	<i>Oecetis avara</i> <i>Hydropsyche sparna</i>
Diplura	<i>Metajapyx</i> sp. <i>Campodea tillyardi</i>	Mecoptera	<i>Nannochorista neotropica</i> <i>Nannochorista dipteroides</i> <i>Boreus coloradensis</i> <i>Boreus californicus</i> <i>Merope tuber</i> <i>Bittacus pilicornis</i> <i>Bittacus strigosus</i> <i>Panorpa isolata</i> <i>Panorpa helena</i>
Archaeognatha	<i>Petrobius brevistylis</i> <i>Trigoniophthalmus alternatus</i>	Siphonaptera	<i>Craneopsylla minerva</i> <i>Megarhroglossus divisius</i> <i>Acanthopsylla rothschildi</i> <i>Atyphloceras echis</i> <i>Orchopeas leucopus</i>
Zygentoma	<i>Lepisma</i> sp. <i>Thermobius domestica</i>	Strepsiptera	<i>Triozocera mexicana</i> <i>Caenocholax fenyesi</i> <i>Elenchus japonica</i> <i>Xenos vesparum</i> <i>Xenos pecki</i> <i>Crawfordia n. sp.</i>
Ephemera	<i>Stenonema</i> sp. <i>Ephemerella</i> sp. <i>Heptagenia diabasica</i>	Diptera	<i>Laphria</i> sp. <i>Tipula</i> sp. <i>Drosophila melanogaster</i> <i>Mythicomylia atra</i>
Odonata	<i>Libellula pulchella</i> <i>Agriion maculatum</i>		
Plecoptera	<i>Megarcys stigmata</i> <i>Cultus decius</i> <i>Agnetina</i> sp. <i>Paragnetina media</i> <i>Agnetina capitata</i> <i>Mesoperlina pecircai</i>		
Embiidina	<i>Oligotoma saundersii</i> <i>Clothoda</i> sp.		
Grylloblattaria	<i>Grylloblatta</i> sp.		
Dermaptera	<i>Forficula auricularia</i> <i>Labia</i> sp. <i>Labidura riparia</i>		
Isoptera	<i>Reticulitermes virginiana</i> <i>Anoplotermes</i> sp.		
Blattaria	<i>Blaberus</i> sp. <i>Gromphadorhina portentosa</i>		
Mantodea	<i>Mantis religiosa</i>		
Orthoptera	<i>Ceuthophilus</i> sp. <i>Melanoplus</i> sp. <i>Warramaba picta</i>		
Phasmida	<i>Timema californica</i> <i>Phyllium</i> sp. <i>Anisomorpha buprestoides</i>		
Zoraptera	<i>Zorotypus snyderi</i>		
Phthiraptera	<i>Dennysus hirudensis</i>		
Thysanoptera	<i>Taeniothrips inconsequens</i>		
Psocoptera	<i>Cerastipsocus venosus</i>		
Hemiptera	<i>Oncometopia orbania</i> <i>Tibicen</i> sp. <i>Saldula pallipes</i> <i>Buenoa</i> sp. <i>Belostoma flumineum</i> <i>Lygus lineolaris</i> <i>Oncopeltus fasciatus</i>		

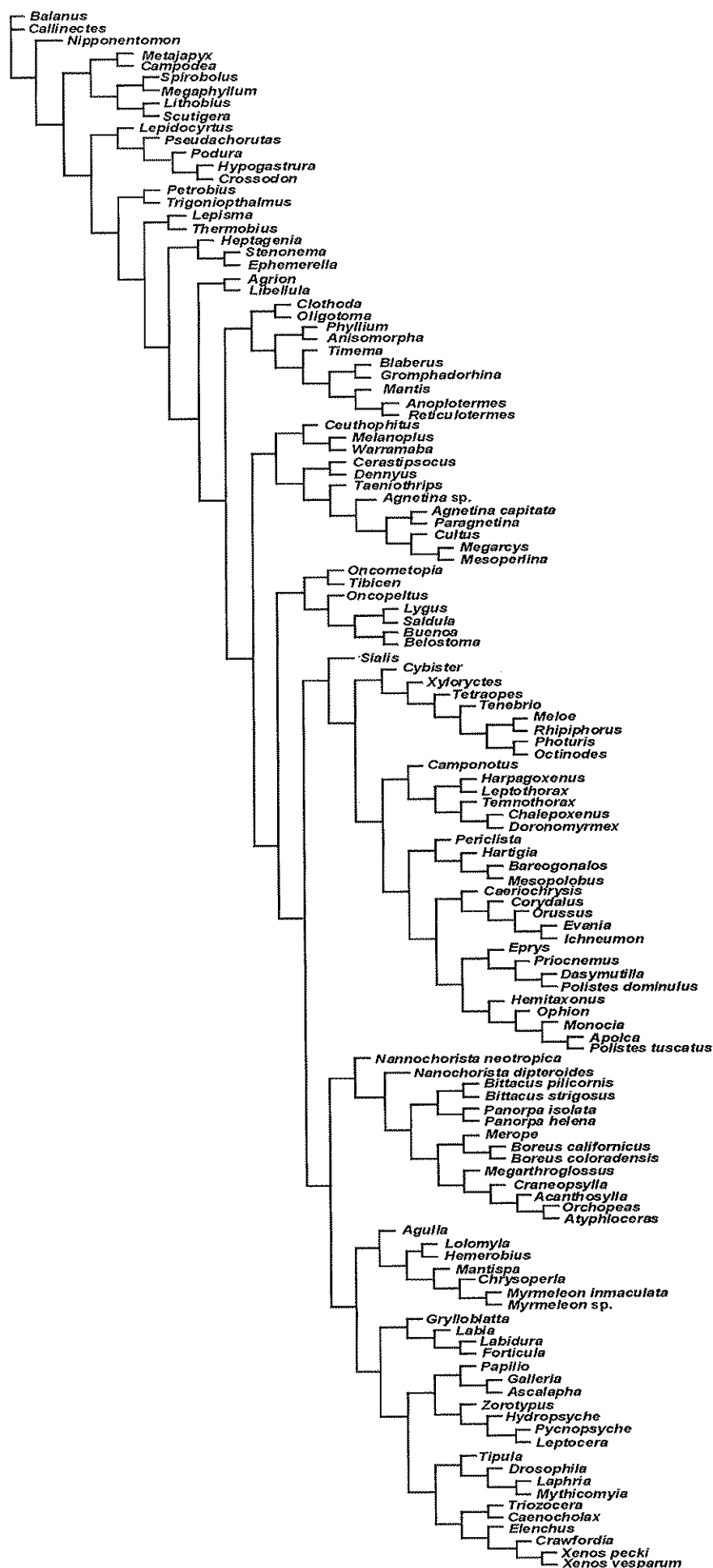


Fig. 3. - Relaciones de los hexápodos basadas en la evidencia molecular, como se describen en Wheeler et al. (1999).

Fig. 3.- Hexapod relationships based upon molecular evidence, as described in Wheeler et al. (1999).

- 4) Situación de Strepsiptera. Se soporta Strepsiptera como grupo hermano de Diptera. De un total de 32 órdenes representados como terminales en la matriz morfológica, 23 "órdenes" son monofiléticos. Mecoptera y Phasmida son parafiléticos y Neuroptera es difilético. Otros resultados nuevos incluyen Dermaptera en el interior de los Holometabola y la situación de Plecoptera con los Psocodea y Thysanoptera. Tal como se muestra más adelante, estas situaciones no están soportadas en el análisis combinado.

La optimización directa también puede permitir diferentes conjuntos de datos que influyan de forma recíproca en las presuntas homologías. Hemos usado esta característica para implementar un análisis simultáneo con POY, para combinar los dos conjuntos de datos moleculares con los caracteres morfológicos. Esto se describe en la siguiente sección.

ANÁLISIS COMBINADO

El primer análisis combinado de datos morfológicos y moleculares en insectos fue el de DeSalle et al. (1992) que secuenció ADNr 18S de un fósil en ámbar de *Mastotermes electrodominicus*, la especie viva de *Mastotermes*, otras dos especies de termitas, un mántido, una cucaracha, un saltamontes, una mosca de las piedras y *Drosophila* como grupo externo. El análisis exclusivo de la secuencia de datos conduce a resultados ambiguos: un cladograma único con Dictyoptera sin resolver como grupo. El análisis simultáneo de la secuencia de datos combinados con la matriz de caracteres de Thorne y Carpenter (1992) mediante la codificación *splice-and-merge* (véase Nixon y Carpenter, 1996) corroboró la agrupación de cucarachas con mántidos, excluyendo las termitas.

Whiting et al. (1997) proveyeron el siguiente ejemplo con insectos y, hasta la fecha, el análisis simultáneo más amplio publicado. El estudio se centró en Holometabola y los datos son un subconjunto de los de Wheeler et al. (1999), por lo que no lo discutiremos más adelante —excepto para una corrección. Hay una diferencia entre los cladogramas de Whiting et al. y Wheeler et al. que debería ser remarcado: a diferencia de Whiting et al., el estudio de Wheeler et al. soporta la monofilia de Coleoptera. El resultado de Whiting et al. (1997) respecto a Coleoptera fue el de parafilia, debido a la gran similitud de las secuencias (más del 90% de identidad) de dos escarabajos primitivos, Megaloptera y Raphidioidea. Polyphaga fue monofilético pero con los archostemmata y adéfagos agrupados con los megalópteros y neurópteros. Ahora sabemos que ha sido producido por un artefacto de contaminación, porque la resecuenciación de estos cuatro taxones de los análisis muestra a Coleoptera y Neuroptera monofiléticos.

Para la inclusión de los caracteres morfológicos en la optimización directa, las codificaciones ordinales fueron combinadas con las respectivos ejemplares secuenciados, esto es, las codificaciones morfológicas para los órdenes fueron tratadas como terminales supraespecíficos (véase Nixon y Carpenter, 1996).

El espacio de parámetros examinado en el análisis simultáneo conllevó unos 100 análisis con POY para examinar diferentes combinaciones de costes, incluyendo costes relativos de combinación de datos de caracteres morfológicos con los datos de las secuencias.

El cladograma de la Fig. 4 es el que maximiza la congruencia entre los tres conjuntos de datos, con un coste de

6211. Está basado en una ratio transición/transversión de 1:1, un coste *indel* (delección/inserción) de dos, y un cambio en caracteres morfológicos igual al coste *indel*. Este cladograma tiene una incongruencia de ILD menor al 3%. Para más detalles, véase Wheeler et al. (1999).

Volviendo a los cuatro puntos ya remarcados para cada análisis por separado:

- 1) Situación de Diplura. Se soporta un Entognatha monofilético. Esta es una de las posibilidades de la matriz morfológica.
- 2) Relaciones en Dictyoptera. Se soporta a Mántidos y cucarachas como grupo hermano de termitas. Este esquema de relaciones fue soportado en cada uno de los diferentes análisis.
- 3) Situación de Zoraptera. Zoraptera está soportado como grupo hermano de Dermaptera. Este resultado no fue obtenido bajo ninguno de los análisis por separado. Esta relación particular no ha sido sugerida anteriormente, pero una situación similar de Zoraptera como una parte de Polyneoptera fue propuesta por Boudraux (1979) y Kukalová-Peck y Peck (1993).
- 4) Situación de Strepsiptera. Se soporta Strepsiptera como grupo hermano de Diptera. Este resultado está fuertemente avalado por datos moleculares, y como se dijo antes, los resultados del análisis morfológico no son muy diferentes.

Hay algunas otras diferencias con el cladograma morfológico —al margen de las relaciones internas de Paraneoptera que son inciertas— con Thysanoptera y Hemiptera que no aparecen como grupos hermanos; los tisanópteros aparecen sólo con los piojos. El único orden no soportado como monofilético es Mecoptera, que es parafilético con referencia a Halteria (Diptera + Strepsiptera). No obstante, como discuten Wheeler et al. (1999), el análisis de sensibilidad muestra que este resultado particular es altamente dependiente de las asunciones, resultando usualmente un grupo Mecoptera monofilético (véase también Whiting et al., 1997).

CONCLUSIÓN

Las relaciones filogenéticas esbozadas en la Fig. 4 representan el esquema actualmente disponible más apoyado. Esto no significa que sea concluyente. Las relaciones apoyadas más fuertemente en el análisis simultáneo son los Hexapoda, Neoptera, Insecta s. str., Dictyoptera, Amphiesmenoptera y Halteria (ver detalles en Wheeler et al., 1999). Deberíamos esperar que estos resultados sean estables a la nueva evidencia, porque las pruebas que lo soportan son abundantes. Grupos menos soportados incluyen Entognatha y Polyneoptera (incluyendo Zoraptera). Nuevas evidencias podrían cambiar estas agrupaciones, ya que la evidencia que ahora las soportan es escasa. La inclusión de nuevos datos de otras moléculas, la adición de taxones fósiles o la rotura de los terminales representados como órdenes en la matriz morfológica pueden producir una revisión. En efecto, estamos siguiendo algunas investigaciones en esas líneas, especialmente con hexápodos basales.

AGRADECIMIENTOS

Agradecemos a Antonio Melic por solicitar esta contribución y a John Wenzel por la información facilitada.

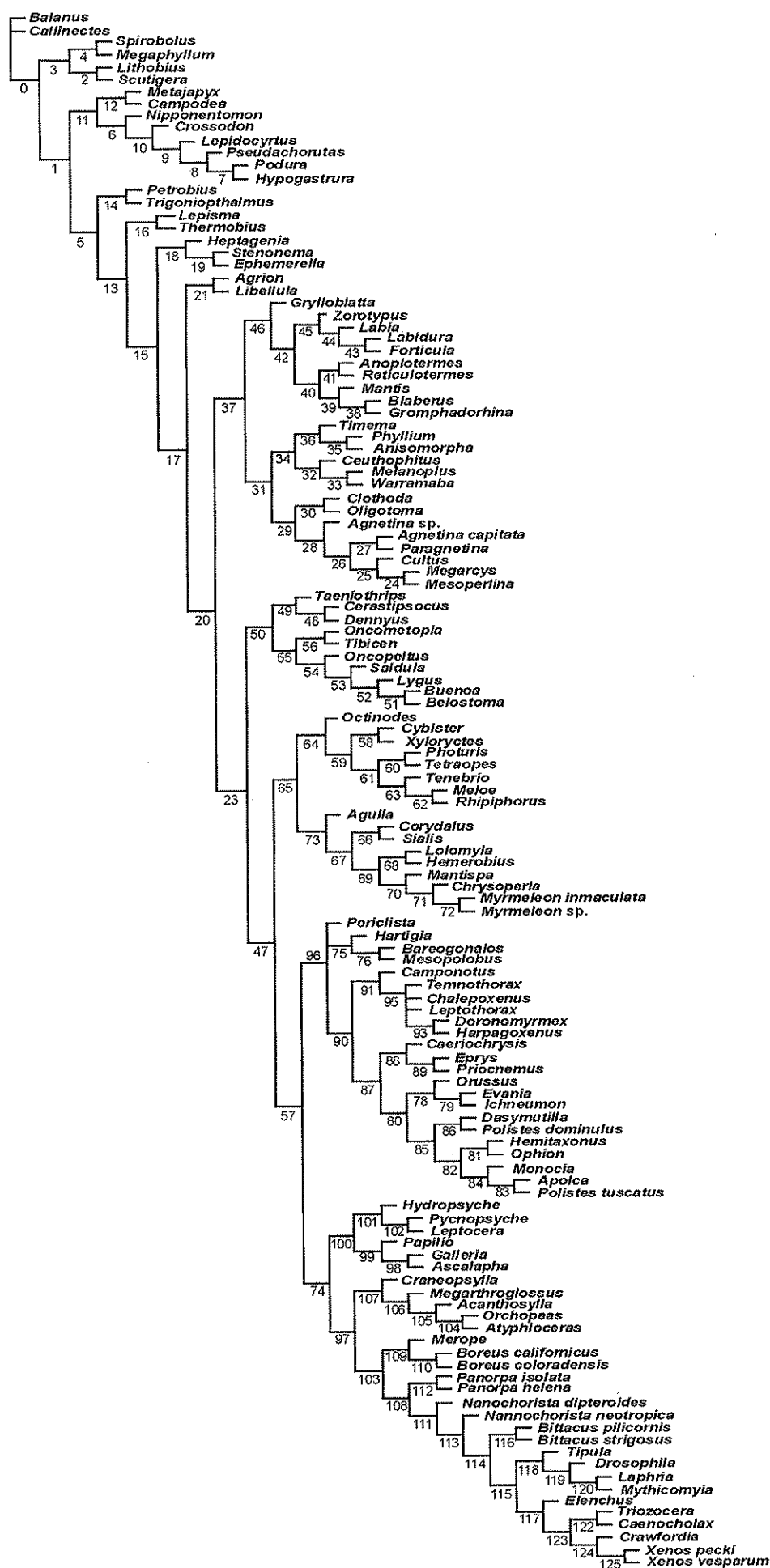


Fig. 4. - Relaciones de los hexápodos basadas en el análisis simultáneo de caracteres morfológicos y moleculares, como se describe en el texto y en Wheeler et al. (1999). Los números de los nodos bajo las ramas hacen referencia a los valores de soporte de Bremer en Wheeler et al. (1999: tabla 4).

Fig. 4.- Hexapod relationships based upon simultaneous analysis of morphological and molecular characters, as described in the text and Wheeler et al. (1999). Node numbers below branches reference Bremer support values in Wheeler et al. (1999: table 4).

BIBLIOGRAFÍA

- BACCETTI, B., 1979. Ultrastructure of sperm and its bearing on arthropod phylogeny. In A. P., GUPTA, (ed.), *Arthropod Phylogeny*: 609-644. Van Nostrand-Reinhold, New York.
- BALLARD, J. W. O., OLSEN, G. J., FAITH, D. P., ODGERS, W. A., ROWELL, D. M. & ATKINSON, P. W., 1992. Evidence from 12S Ribosomal RNA sequences that onychophorans are modified arthropods. *Science*, **258**: 1345-1348.
- BLACKITH, R. E. & BLACKITH, R. M., 1968. A numerical taxonomy of orthopteroid insects. *Austr. J. Zool.*, **16**: 111-131.
- BOUDREAUX, H. B., 1979. *Arthropod phylogeny with special reference to insects*. John Wiley & Sons, New York.
- CARMEAN, D., KIMSEY, L. S. & BERBEE, M. L., 1992. 18S rDNA sequences and holometabolous insects. *Mol. Phyl. Evol.*, **1**: 270-278.
- CARMEAN, D. & CRESPI, B. J., 1995. Do long branches attract flies? *Nature*, **373**: 666.
- CARPENTER, J. M., GOLOBOFF, P. A. & FARRIS, J. S., 1998. PTP is meaningless, T-PTP is contradictory: A reply to Trueman. *Cladistics*, **14**: 105-116.
- DESALLE, R., GATESY, J., WHEELER, W. & GRIMALDI, D., 1992. DNA sequences from a fossil termite in Oligo-Miocene amber and phylogenetic implications. *Science*, **257**: 1933-1936.
- FARRIS, J. S. 1988. *Hennig86, version 1.5*. Program and documentation. Port Jefferson Station, New York.
- FARRIS, J. S. 1999. Likelihood and inconsistency. *Cladistics*, **15**: 199-204.
- FARRIS, J. S., KÄLLERSJÖ, M., KLUGE, A. G. & BULT, C., 1994. Testing significance of incongruence. *Cladistics*, **10**: 315-319.
- GIRIBET, G., WHEELER, W. C. & EDGEcombe, G. D., 1999. Sistemática y filogenia de artrópodos, estado de la cuestión con énfasis en análisis de datos moleculares. *Bol. SEA*, **26**: 197-212 (este volumen).
- GLADSTEIN, D. S. & WHEELER, W. C., 1997. *POY: The Optimization of Alignment Characters*. Program and documentation. American Museum of Natural History, New York. Available at "ftp.amnh.org" /pub/molecular.
- GOLOBOFF, P. A., 1998. *NONA, version 1.9*. Program and documentation. Instituto Miguel Lillo, Tucumán, Argentina.
- GRANDCOLAS, P., 1994. Phylogenetic systematics of the subfamily Polyphaginae, with the assignment of *Cryptocercus* Scudder, 1862 to this taxon (Blattaria, Blaberoidea, Polyphagidae). *Syst. Entomol.*, **19**: 145-158.
- GRANDCOLAS, P., 1996. The phylogeny of cockroach families: a cladistic appraisal of morpho-anatomical data. *Can. J. Zool.*, **74**: 508-527.
- GRANDCOLAS, P., 1997. What did the ancestors of the woodroach *Cryptocercus* look like? A phylogenetic study of the origin of subsociality in the subfamily Polyphaginae (Dictyoptera, Blattaria). *Mém. Mus. Natl. Hist. Nat.*, **173**: 231-252.
- GRANDCOLAS, P. & DELEPORTE, P., 1992. La position systématique de *Cryptocercus* Scudder au sein des Blattes et ses implications évolutives. *C. R. Acad. Sci. Paris (III)*, **315**: 317-322.
- HENNIG, W., 1953. Kritische Bemerkungen zum phylogenetischen System der Insekten. *Beitr. Ent.*, **3** Sonderheft: 1-85.
- HENNIG, W., 1981. *Insect Phylogeny*. Academic Press, New York.
- HUELSENBECK, J. P., 1997. Is the Felsenstein zone a fly trap? *Syst. Biol.*, **46**: 69-74.
- HUELSENBECK, J. P., 1998. Systematic bias in phylogenetic analysis: Is the Strepsiptera problem solved? *Syst. Biol.*, **47**: 519-537.
- KLASS, K.-D., 1995. *Die Phylogenie der Dictyoptera*. Cuvillier Verlag, Göttingen.
- KLUGE, A. G., 1989. A concern for evidence and a phylogenetic hypothesis for relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.*, **38**: 1-25.
- KRAUS, O., 1997. Phylogenetic relationships between higher taxa of tracheate arthropods. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 295-303. Chapman and Hall, London.
- KRISTENSEN, N. P., 1975. The phylogeny of hexapod "orders." A critical review of recent accounts. *Z. Zool. Syst. Evolutionsforsch.*, **13**: 1-44.
- KRISTENSEN, N. P., 1981. Phylogeny of insect orders. *Annu. Rev. Entomol.*, **26**: 135-157.
- KRISTENSEN, N. P., 1991. Phylogeny of extant hexapods. In I. D. NAUMANN, P. B. CARNE, J. F. LAWRENCE, E. S. NIELSEN, J. P. SPRADBERRY, R. W. TAYLOR, M. J. WHITTEN, & M. J. LITTLEJOHN (eds.), *The Insects of Australia*, 2nd ed.: 125-140. CSIRO, Melbourne Univ. Press, Melbourne.
- KRISTENSEN, N. P. 1995. Forty years' insect phylogenetic systematics. *Zool. Beitr. (N. F.)*, **36**: 83-124.
- KRISTENSEN, N. P. 1997. The groundplan and basal diversification of the hexapods. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 281-293. Chapman and Hall, London.
- KUKALOVÁ-PECK, J., 1987. New Carboniferous Diplura, Monura, and Thysanura, the hexapod groundplan and the role of thoracic side lobes in the origin of wings (Insecta). *Can. J. Zool.*, **65**: 2327-2345.
- KUKALOVÁ-PECK, J., 1997. Arthropod phylogeny and "basal" morphological structures. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 249-268. Chapman and Hall, London.
- KUKALOVÁ-PECK, J. & LAWRENCE, J. F., 1993. Evolution of the hind wing in Coleoptera. *Can. Entomol.*, **125**: 181-258.
- KUKALOVÁ-PECK, J. & PECK, S. B., 1993. Zoraptera wing structures: evidence for new genera and relationship with the blattoid orders (Insecta: Blattellodea). *Syst. Entomol.*, **18**: 333-350.
- LIU, H. & BECKENBACH, A. T., 1992. Evolution of the mitochondrial oxidase II gene among 10 orders of insects. *Mol. Phyl. Evol.*, **1**: 41-52.
- MANTON, S. M., 1977. *The Arthropoda: Habits, Functional Morphology, and Evolution*. Clarendon Press, Oxford.
- MARTYNOV, A. V., 1925. Über zwei Grundtypen der Flügel bei den Insekten und ihre Evolution. *Z. Morph. Ökol. Tiere*, **4**: 465-501.
- MICKEVICH, M. F., & FARRIS, J. S., 1981. The implications of congruence in *Menidia*. *Syst. Zool.*, **30**: 351-370.
- MINET, J., & BOURGOIN, T., 1986. Phylogénie et classification de Hexapodes (Arthropoda). *Cah. Liason*, **63**: 23-28.
- NIXON, K. C., 1998. *CLADOS, version 1.9*. Program and documentation. Bailey Hortorium, Ithaca, New York.
- NIXON, K. C. & CARPENTER, J. M., 1993. On outgroups. *Cladistics*, **9**: 413-426.
- NIXON, K. C. & CARPENTER, J. M., 1996. On simultaneous analysis. *Cladistics*, **12**: 221-241.
- SNODGRASS, R. E., 1938. Evolution of the Annelida, Onychophora and Arthropoda. *Smiths. Misc. Coll.*, **97**: 1-159.
- PASHLEY, D. P., MCPHERON, B. A. & ZIMMER, E. A., 1993. Systematics of holometabolous insect orders based on 18S ribosomal RNA. *Mol. Phyl. Evol.*, **2**: 132-142.
- PASS, G., 1991. Antennal circulatory organs in Onychophora, Myriapoda, and Hexapoda: Functional morphology and evolutionary implications. *Zoomorphology*, **110**: 145-164.
- PAULUS, H. F. VON, 1979. Eye structure and the monophyly of the Arthropoda. In A. P., GUPTA, (ed.), *Arthropod Phylogeny*: 299-383. Van Nostrand-Reinhold, New York.
- RASNITSYN, A. P. 1998. On the taxonomic position of the insect order Zorotypida = Zoraptera. *Zool. Anz.*, **237**: 185-194.
- SIDDALL, M. E., 1998. Success of parsimony in the four-taxon case: long-branch repulsion by likelihood in the Farris zone. *Cladistics*, **14**: 209-220.
- ŠTYS, P. & BILINSKI, S., 1990. Ovariole types and the phylogeny of hexapods. *Biol. Rev.*, **65**: 401-429.
- ŠTYS, P., ZRZAVÝ, J. & WEYDA, F., 1993. Phylogeny of the Hexapoda and ovarian metamerism. *Biol. Rev.*, **68**: 365-379.
- THORNE, B. L. & CARPENTER, J. M., 1992. Phylogeny of the Dictyoptera. *Syst. Entomol.*, **17**: 253-268.
- WHEELER, W. C., 1989. The systematics of insect ribosomal DNA. In B. FERNHOLM, K. BREMER, AND H. JÖRNVALL (eds.), *The Hierarchy of life. Molecules and Morphology in Phylogenetic Analysis*: 307-321. Elsevier, Amsterdam.
- WHEELER, W. C., 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst. Biol.*, **44**: 321-331.
- WHEELER, W. C., 1996 Optimization Alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics*, **12**: 1-10.
- WHEELER, W. C., 1997. Sampling, groundplans, total evidence and the systematics of arthropods. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 87-96. Chapman and Hall, London.

- WHEELER, W. C., CARTWRIGHT, P. & HAYASHI, C. Y., 1993. Arthropod phylogeny: A combined approach. *Cladistics*, **9**: 1-39.
- WHEELER, W. C., WHITING, M. F., WHEELER, Q. D. & CARPENTER, J.M., 1999. The phylogeny of the extant hexapod orders. *Cladistics* (en prensa).
- WHITING, M. F., 1998. Long-branch distraction and the Strepsiptera. *Syst. Biol.*, **47**: 134-138.
- WHITING, M. F. & WHEELER, W.C., 1994. Insect homeotic transformation. *Nature*, **368**: 696.
- WHITING, M. F., CARPENTER, J. M. , WHEELER, Q. D. & WHEELER, W.C., 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst. Biol.*, **46**: 1-68.
- WHITING, M. F. & KATHIRITHAMBY, J., 1995. Strepsiptera do not share hind-wing venational synapomorphies with the Coleoptera. *J. N. Y. Entomol. Soc.*, **103**: 1-14.
- WILLMANN, R. 1997. Advances and problems in insect phylogeny. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 269-279. Chapman and Hall, London.
- WILSON, E. O., 1971. *The Insect Societies*. Harvard University Press, Cambridge.
- ZRZAVÝ, J., HYPŠA, V. & VLÁŠKOVÁ, M., 1997. Arthropod phylogeny: taxonomic congruence, total evidence and conditional combination approaches to morphological and molecular data sets. In R. A. FORTEY AND R. H. THOMAS (eds.), *Arthropod Relationships*. Syst. Assoc. Spec. Vol. 55: 97-107. Chapman and Hall, London.

Evolution and Phylogeny of Arthropoda

NUMERICAL CLADISTICS, SIMULTANEOUS ANALYSIS AND HEXAPOD PHYLOGENY

James M. Carpenter & Ward C. Wheeler

Introduction

The application of cladistic methods to insect ordinal relationships began with Hennig (1953). Kristensen's (1995) contribution to the field emphasized the importance of Hennig's paper, and reviewed advances during the 40 years since that work. Among these advances is of course Kristensen's noted (1975) paper, perhaps still the landmark work on insect ordinal phylogenetics. That paper truly set the stage for modern work on insect ordinal relationships, playing a major role in subsequent work, including the recent work we will discuss here, which is the application of quantitative cladistics to the entire problem of insect ordinal relationships. We, together with our collaborators Michael Whiting and Quentin Wheeler, have undertaken the first numerical cladistic analyses of this problem, working with both morphological and molecular data, as well as the first simultaneous analyses of morphological and molecular data (Whiting et al., 1997; Wheeler et al., 1999). We will review that, and other ongoing work, organizing it for the illustrative purposes into three parts: discussion of morphological data, including previous work and the first comprehensive ordinal matrix, discussion of molecular data, contrasting our approach with previous work, including the first matrix with all orders represented, and finally combination of that with the morphological data in a simultaneous analysis.

Historically, single character systems formed the basis for supraordinal groupings among hexapods. Classic examples include the divisions of Pterygota (Martynov, 1925) based on wing folding, and Snodgrass' (1938) breakup of Thysanura based on jaw articulation. More recent examples include Baccetti (1979) on spermatozoa, von Paulus (1979) on eye ultrastructure, Pass (1991) on antennal circulatory structures, Štys et al. (1993) on ovarioles, etc. Kristensen's (1975, 1981, 1991, 1995, 1997) work is notable for integrating many single character studies in traditional, Hennigian argumentation, that is manually, character by character, node by node. All of these studies have greatly advanced our knowledge of hexapods.

The compelling advantage of numerical cladistics relative to this approach is simple: simultaneous treatment of all available evidence. It has sometimes been claimed that detailed functional investigation of single character systems is a better approach (e. g., Manton, 1977), as if conjectural conclusions about function is a better phylogenetic indicator than observed similarity. The so-called "historical approach" of Kukalová-Peck (1997) is a modern, extreme version of allowing conjectures on relationships to dictate what constitutes data. If the characters analyzed in a cladistic study are independent of each other, than any character is a priori equally likely to be as good an indicator of phylogeny as any other. Basic sampling theory then dictates that as many characters as possible should be combined and analyzed simultaneously, to achieve a robust hypothesis of phylogenetic relationships. We have attempted just this.

The "total evidence" (Kluge, 1989) or simultaneous analysis (Nixon and Carpenter, 1996) approach is in effect an argument against superiority of any particular data type - the old argument often implicit in single character studies, already refuted in the days of Linnaeus. Independence of characters means different character types, like morphological and molecular, should also be combined (Kluge, 1989; Nixon and Carpenter, 1996). Simultaneous analysis thus maximizes explanatory power-parsimony. As pointed out by Nixon and Carpenter (1996), the so-called "taxonomic congruence" and "conditional combination" approaches (see, e. g., Zrzavý et al., 1997) merely abandon parsimony. We therefore advocate simultaneous analysis, and apply that in our analyses.

Morphological Data

The first quantitative morphological matrix on supraordinal relationships to be analyzed cladistically was that of Thorne and Carpenter (1992), on Dictyoptera (termites, cockroaches and mantids). To be sure, Blackith and Blackith (1968), offered what they called a "cladogram" for some orthopterid orders, showing Orthoptera itself as paraphyletic in terms of mantids and cockroaches. Their "cladogram" was however based on the method of Camin and Sokal (1967), which assumes irreversibility in character evolution. If one does not assume that, and reanalyzes their data - cladistically - one obtains the cladogram of Fig. 1, with a monophyletic Orthoptera.

Thorne and Carpenter were motivated by two inveterate questions: the placement of the primitively social wood roach *Cryptocercus*, and the placement of *Mastotermes*, the "roachlike," living fossil termite. Both taxa play a key role in the subsocial theory for the evolution of sociality in termites (e. g., Wilson, 1971), as suggestive intermediates. Thorne and Carpenter carried out a literature survey for characters, with limited checking, and assembled a data matrix of 70 morphological and behavioral characters for seven terminals (three Isoptera, two Blattaria, one Mantodea and root). The matrix was analyzed with the exact routine of the Hennig86 program (Farris, 1988), which resulted in a single cladogram. Surprisingly, this cladogram supported Boudreaux's (1979) hypothesis of roaches and mantids as sister-groups, relative to termites - thus refuting the hoary subsocial theory for the evolution of termite sociality.

Kristensen (1995) questioned this result, and changed character codings in the matrix, as he admitted, subjectively. The effect of these manipulations was that no conclusions were well supported when Hennig86 was applied to the changed matrix. We regard it as commendable that Kristensen thereby employed computer-assisted cladistics, and whereas we could offer arguments about the recodings, instead we will show that Thorne and Carpenter's results are well-supported in the context of analysis of all the orders, and in combination with molecular data. Thorne and Carpenter included a hypothetical ancestor to root their cladogram (pre-rooting, in the terminology of Nixon and Carpenter, 1993), but we will show that roaches are still supported as the sister-group of mantids without such a synthetic node.

Our present matrix also relies primarily on the literature, especially the papers of Kristensen, Hennig (1981) and Boudreaux (1979), but also some recent papers by authors such as Minet and Bourgoin (1986) Štys and Bilinski (1990), and a few from Kukalová-Peck (e. g., 1987) - only those we could verify, which necessitated discarding many of her claims (e. g., cf. Kukalová-Peck and Lawrence, 1993, with Whiting and Kathirithamby, 1995). The scorings for terminals are putative ground-plans: we attempted to score to the extent possible what the authors in the literature claimed as the primitive state for an order. The purpose thereby was to give the claims made about relationships by these authors the most favorable chance possible in the numerical analysis. Obviously, this also entailed eschewing polymorphisms. Part of the matrix was published in the analysis focusing on Holometabola by Whiting et al. (1997). The matrix is presently 256 variables for 32 insect orders and a myriapod outgroup.

About the outgroup, we must acknowledge that relationships among arthropod classes have been controversial, with authors such as Manton (1977) claiming that hexapods are not monophyletic. There is no reason to take any of her arguments seriously; they are not cladistic. Superficially more impressive are analyses such as that by Ballard et al. (1992), who published a tree purportedly based on 12S mtDNA data, with Arthropoda

paraphyletic in terms of velvet worms, and Myriapoda at the base of the tree. But as Carpenter et al. (1998) showed, Ballard et al.'s published tree was actually based on manipulations of their data: the consensus tree for the real data shows very little structure - groups such as velvet worms and arachnids are not even monophyletic. Support for Myriapoda as the closest relatives of insects is found in analysis of combined morphological and sequence data by Wheeler et al. (1993), although "Myriapoda" may be paraphyletic, with only part of it in fact the sister-group of Hexapoda (Wheeler, 1997; Kraus, 1997; cf. however Giribet et al., 1999, this volume).

When the present morphological matrix is analyzed with the Nona program (Goloboff, 1998), two cladograms result, with a length of 450, consistency index 0.71 and retention index 0.81. The consensus tree is given in Fig. 2; further discussion of the results may be found in Wheeler et al. (1999).

There are four parts to which we will call attention here:

- 1) Placement of Diplura. The consensus tree is ambiguous, because one of the underlying cladograms supports Entognatha (Diplura, Collembola and Protura) as a monophyletic group (e. g., Hennig, 1981), while the other places Diplura as sister-group to Insecta s. str. (e. g., Kukalová-Peck, 1987). Thus, as far as the morphological matrix is concerned, the basal relationships in Hexapoda remain uncertain (e. g., Kristensen, 1991). We might add that paraphyly of Diplura was suggested, by Štys and Bilinski (1990), who proposed that Campodeina are more closely related to Ellipura than Japygina. However, when Diplura is broken up in the morphological matrix, analysis then results in basically the same two cladograms, both with a monophyletic Diplura (see Wheeler et al., 1999). The morphological characters thus in fact support the monophyly of Diplura, but until the molecular data are included (see below), its placement is unstable.
- 2) Relationships within Dictyoptera. Mantids and roaches as sister-groups relative to termites is confirmed (Boudreaux, 1979; Thorne and Carpenter, 1992). Relationships within Dictyoptera have nevertheless remained controversial. For example, Willmann (1997: 269) described Klass (1995) as a "monumental work," which showed that "termites are in fact a subgroup of the cockroaches and that the wingless roach *Cryptocercus* is the sister-group of the termites." Contrast this characterization with Grandcolas (1997: 234), who stated: "Klass (1995) presented a tree for Isoptera, Blattodea and Mantodea on the basis of the study of 14 species, however, it was not constructed according to a genuine phylogenetic analysis (no matrix of characters, no outgroups, polymorphic characters not coded as such, etc.)..." Willmann's statement overlooks the demonstration by Grandcolas and Deleporte (1992) and Grandcolas (1994, 1996) that *Cryptocercus* is a relatively derived, polyphagine cockroach, and that Polyphagidae is a monophyletic group - exclusive of termites. Our results, based on either morphology or molecules or both (see below), are in line with this.
- 3) Placement of Zoraptera. This order is placed as sister-group to Eumetabola (Acercaria = hemipterids + Holometabola). The order has variously been suggested as sister-group of Acercaria (Hennig, 1981), sister-group of Dictyoptera (Boudreaux, 1979), sister-group of Embiidina (Minet and Bourgoin, 1986) and recently as sister-group of Holometabola (Rasnitsyn, 1998). The placement with Eumetabola is not far off from the relationships suggested by Hennig (1981) or Rasnitsyn (1998), but the molecular data do not accord with this, as will be seen.
- 4) Placement of Strepsiptera. The group is shown as a panorpida, sister-group to Antliophora (fleas, flies and scorpionflies). This solution to the "Strepsiptera problem" is perhaps the most surprising result for the morphological characters. Traditionally, of course, this order has been considered close to Coleoptera, if indeed not part of it. Upon review of the relevant characters, however, it becomes clear that the traditional placement is based upon very little evidence: one character links Strepsiptera with Coleoptera, flight by the hind wings. (Kukalová-Peck (1997) equivocated with Whiting and Kathirithamby's (1995) demonstration that Kukalová-Peck and Lawrence (1993) had not shown synapomorphies in the hind wing venation of Strepsiptera with Coleoptera, but failed to provide data supporting her conten-

tions.) No characters link Strepsiptera with Coleoptera + Neuroptera, but in contrast, several characters do link them to the panorpida orders, and there are even some that link them closer to Antliophora (see details in Whiting et al., 1997; Wheeler et al., 1999). Molecular data previously supported a sister-group relationship between Strepsiptera and Diptera (Whiting and Wheeler, 1994; Whiting et al., 1997), and in fact the morphological data in isolation are not far from agreeing with this - and the combined analysis (Whiting et al., 1997, and see below).

Molecular Data

We now take up molecular data. The earliest application of molecular data to this problem is Wheeler's (1989) paper, which applied a limited sample of ribosomal DNA sequence and restriction map data from 11 representatives of ten insect orders. He found support for Insecta, grouping of Odonata with Neoptera, Eumetabola, Holometabola, and Neuroptera as sister-group to Coleoptera.

Several subsequent molecular studies have been a step back in terms of sampling, even if more sequences have been analyzed. The papers by Liu and Beckenback (1992) and Carmean et al. (1992) did not support monophyly of Holometabola; Pashley et al. (1993) did not test for it. The first paper based on a comprehensive sampling of Holometabola was that by Whiting and Wheeler (1994), which reported that Strepsiptera are the sister-group to flies. Their DNA sequence data are included in the present study. Suffice to say here that Whiting and Wheeler's paper has been criticized, for example by Carmean and Crespi (1995), who suggested that the result was an artifact of so-called long branch attraction - based on their own extremely limited sample. Whiting et al. (1997) showed that this criticism is specious, based on the very large sample that is part of the results we will discuss here. Related charges of long-branch attraction based upon maximum likelihood analysis (Huelsenbeck, 1997, 1998) have succeeded only exposing deficiencies of maximum-likelihood methods (Whiting, 1998; Siddall, 1998), methods that were in any event never logically justified (Farris, 1999).

Turning to the present molecular data for all insect orders, Wheeler et al.'s (1999) study includes sequence data from two genes, ribosomal 18S and 28S rDNA. The sample consists of 121 hexapod sequenced, with four myriapods and two Crustacea sequenced as outgroups (Table 1). The 18S fragment size was approximately 1000 base pairs, while the 28S size was approximately 350 base pairs. All hexapod orders were included for both molecules, and for the majority of the orders more than one exemplar was sequenced. The total is about 200 kb of sequence data. Further details, including Genbank accession numbers to the sequences, may be found in Wheeler et al. (1999).

Analysis of sequence data raises the issue of alignment, which is how homology is specified at the sequence level. Alignments are made by inserting gaps into sequences of differing lengths, entailing insertions and deletions of base pairs. Whether done subjectively ("by eye") or through a cost function, the process amounts to claims about evolutionary events. Because such claims are arbitrary, we have used sensitivity analysis, as described by Wheeler (1995). In sensitivity analysis, the parameters used in making the alignment are varied, and different alignments constructed, the process is repeated and the entire range of alignments analyzed cladistically. The point is not to conclude that a given combination of parameters, such as gap to change ratio or transition-transversion cost, is necessarily best, as much as it is to determine how sensitive the results are to arbitrary values of these parameters. This allows one to avoid assumption-specific, unstable conclusions.

Our alignment procedure also applied Wheeler's (1996) "optimization alignment" approach, or direct alignment. The cost function was minimized by constructing hypothetical taxonomic intermediates among the sequences, as on a cladogram, instead of an overall alignment as such. Direct alignment obviates the creation of gap characters to align sequences in a matrix, and appears to generate more parsimonious cladograms than multiple alignment (Wheeler 1996).

Analyses were conducted with the program POY (Gladstein and Wheeler, 1997) on the parallel computer cluster at the AMNH. Our criterion for choosing among the results of the sensitivity analysis for the "best" result is congruence. Congruen-

ce, that is, agreement among the two molecular data sets, is measured by the incongruence length difference of Mckevich and Farris (1981; see Farris et al., 1994), and is simply the extra homoplasy or tree length required when the combined data are analyzed cladistically.

The results of the analyses of molecular data are shown in Fig. 3. A single cladogram of cost 5170 was found, based on a transition/transversion ratio of 1:1, and a gap (indel) to change cost ratio of 2:1 which maximizes congruence (minimizes incongruence). For further details, see Wheeler et al. (1999). Regarding the four points mentioned for morphology:

- 1) Placement of Diplura. Hexapoda is not supported as group, with Diplura placed as sister-group to the myriapod outgroups.
- 2) Relationships within Dictyoptera. Mantids and roaches as sister-groups relative to termites is supported.
- 3) Placement of Zoraptera. Zoraptera is shown as part of the Amphiesmenoptera (Lepidoptera + Trichoptera), a decidedly novel result, and one which accords with neither morphological data nor the results of combined analysis.
- 4) Placement of Strepsiptera. Strepsiptera is supported as the sister-group to Diptera.

Of a total of 32 orders represented as terminals in the morphological matrix, twenty-three were sampled for molecular data by multiple exemplars. In Fig. 3, twenty of these twenty-three "orders" are monophyletic. Mecoptera and Phasmida are paraphyletic and Neuroptera are diphyletic. Other novel results include Dermaptera within Holometabola, and placement of Plecoptera with the Psocodea and Thysanoptera. As shown below, these placements are not supported in the combined analysis.

Optimization alignment can also allow different data sets to influence the putative homologies of each other simultaneously. We have used this feature to implement a simultaneous analysis with POY, to combine the two molecular data sets with the morphological characters. This is described in the next section.

Combined Analysis

The first combined analysis of morphological and molecular data in insects was that of DeSalle et al. (1992), who sequenced 18S rDNA from an amber fossil of *Mastotermes electrodominicus*, the living species of *Mastotermes*, two other species of termites, a mantid, a cockroach, a grasshopper, a stonefly and *Drosophila* as outgroup. Analysis of the sequence data alone led to ambiguous results: a single cladogram with Dictyoptera not resolved as a group. Simultaneous analysis of the sequence data combined with Thorne and Carpenter's (1992) character matrix by splice-and-merge coding (see Nixon and Carpenter, 1996) corroborated the grouping of roaches with mantids, exclusive of termites.

Whiting et al. (1997) provided the next insect example, and thus far most comprehensive simultaneous analysis published. That study concentrated on Holometabola, and the data are a subset of those in Wheeler et al. (1999), so we will not discuss that study further - except for one correction. There is one difference between the cladograms of Whiting et al. and Wheeler et al. that should be stressed: unlike Whiting et al., Wheeler et al.'s study supports monophyly of Coleoptera. The Whiting et al. (1997) result for Coleoptera was paraphyly, due to the great sequence similarity (more than 90% identity) of two primitive beetles, Megaloptera and Raphidioidea. Polyphaga was monophyletic, but the single archostematan and adephagan clustered with the dobsonfly and snakefly. We now know that to be an artifact of contamination, because resequencing of these four taxa from the analysis gives a monophyletic Coleoptera and Neuroptera.

For inclusion of the morphological characters in the direct alignment, the ordinal scores were combined with the respective

sequenced exemplars, that is, the morphological scores for the orders were treated as summary terminals (see Nixon and Carpenter, 1996).

The parameter space examined for the simultaneous analysis entailed about 100 computer runs with POY to examine different combinations of costs, including relative costs of combining the morphological characters with the sequence data.

The cladogram of Fig. 4 is that which maximizes congruence among the three data sets, at a cost of 6211. It is based on a transition/transversion ratio of 1:1, a indel cost of two, and change in morphological characters weighted equal to indel cost. This cladogram has an incongruence length difference of less than 3%. For further details, see Wheeler et al. (1999).

So to return to the four points already highlighted for each of the separate analyses:

- 1) Placement of Diplura. A monophyletic Entognatha is supported. This is one of the possibilities for the morphological matrix.
- 2) Relationships within Dictyoptera. Mantids and roaches as sister-groups relative to termites is supported. This scheme of relationships was supported in each of the separate analyses.
- 3) Placement of Zoraptera. Zoraptera is supported as sister-group to Dermaptera. This result was obtained under neither of the separate analyses. This particular relationship has not been suggested previously, but similar placements of Zoraptera as a part of Polyneoptera were made by Boudreaux (1979) and Kukalová-Peck and Peck (1993).
- 4) Placement of Strepsiptera. Strepsiptera is supported as the sister-group to Diptera. This result is strongly supported by molecular data, but as noted before, the results for analysis of morphology alone are not very dissimilar.

There are few other differences from the morphology-only cladogram, except that relationships within Paraneoptera are uncertain, with Thysanoptera and Hemiptera not appearing as sister-groups, rather the thrips cluster with lice. The only order not supported as monophyletic is Mecoptera, which is shown as paraphyletic relative to Halteria (Diptera + Strepsiptera). However, as discussed in Wheeler et al. (1999), sensitivity analysis shows that this particular result is highly assumption-dependent, with a monophyletic Mecoptera usually resulting (see also Whiting et al., 1997).

Conclusion

The phylogenetic relationships depicted in Fig. 4 represent the best supported scheme presently available. This is not to say it is conclusive. The most strongly supported relationships in the simultaneous analysis are Hexapoda, Neoptera, Insecta s. str., Dictyoptera, Amphiesmenoptera, and Halteria (see details in Wheeler et al., 1999). We would expect these results to remain robust to new evidence, because the evidence supporting them is abundant. Weakly supported groupings include Entognatha and Polyneoptera (inclusive of Zoraptera). New evidence may well change such groupings, because the evidence now supporting them is scarce. Inclusion of new data from other molecules, or addition of fossil taxa, or breaking up of the terminals represented as orders in the morphological matrix may well be expected to lead to revision. In fact, we are pursuing some investigations along those lines now, particularly with basal hexapods.

Acknowledgments

We thank Antonio Melic for soliciting this contribution, and John Wenzel for information.