Phylogeny of Basal Hexapod Lineages and Estimates of Divergence Times

JEROME C. REGIER,[1] JEFFREY W. SHULTZ,[2] AND ROBERT E. KAMBIC[2]

Ann. Entomol. Soc. Am. 97(3): 411–419 (2004)

ABSTRACT Phylogenetic relationships among basal hexapod lineages were investigated using molecular sequence data derived from three nuclear genes: elongation factor-1α, RNA polymerase II, and elongation factor-2. Nucleotide and amino acids from 12 hexapods and 22 crustacean outgroups were analyzed using maximum parsimony and maximum likelihood methods. The results support most traditional morphology-based relationships, including monophyly of Hexapoda, Diplura, Insecta, and Pterygota. However, placement of Diplura was unstable. Some analyses placed them as the sister group to Ellipura (Collembola + Protura) to form Entognatha. In others, Diplura was recovered as the sister group to Insecta, contrary to the Entognatha hypothesis. The analysis also recovered a monophyletic Thysanura *sensu lato* (Archaeognatha + Zygentoma) as the sister group to Pterygota, a conclusion that is consistent with precladistic notions of hexapod systematics but conflicts with current understanding of morphological evolution. The data were also used to reconstruct divergence times from a Bayesian analysis of sequence changes that also incorporated constraints at several nodes based on our understanding of the fossil record. At one node that was not directly constrained by the fossil record (dictyopteran/orthopteroid divergence), our estimate was inconsistent with fossil evidence, suggesting our results (like those using other dating methods) should be interpreted with caution.

KEY WORDS Hexapoda, Entognatha, Diplura, Thysanura *sensu lato*

RESULTS FROM RECENT MOLECULE- and morphology-based studies have challenged long-standing and widely held hypotheses of arthropod phylogeny, including the once seemingly unassailable view that hexapods and myriapods form a monophyletic group (e.g., Averof and Akam 1995, Boore et al. 1998, Dohle 2001, Regier and Shultz 2001). Similar re-evaluations have targeted traditional groupings of basal hexapod lineages (Fig. 1), with some workers questioning the evidence for monophyly of Dicondylia (Koch 2001) and Entognatha (Kukalová-Peck 1987, Wheeler et al. 2001a, Zhang et al. 2001) and at least one molecular study suggesting that Hexapoda itself is diphyletic (Nardi et al. 2003a, b). In most cases, the alternative phylogenetic systems (if any) offered by these studies are not convincing because they focus on a single character system (e.g., mandibular articulation or ovariole structure), assume Myriapoda to be the closest hexapod outgroup, or use a small or analytically complex sample of gene sequence data. However, the recent critiques have served to highlight weaknesses in the evidence and assumptions on which the received view of basal hexapod phylogeny is based (Klass and Kristensen 2001) and invite new efforts to examine traditional concepts with new data.

Here we examine nucleotides and inferred amino acids from three nuclear genes (elongation factor-1α [*EF-1*α], elongation factor-2 [*EF-2*], and RNA polymerase II [*POL II*]) to assess hypotheses of basal hexapod relationships. Our analysis included multiple representatives from each of five major hexapod lineages (Ellipura, Diplura, Archaeognatha, Zygentoma, and Pterygota) and a broad sampling of crustacean outgroups. The data were analyzed using several phylogenetic methods, as well as a new Bayesian approach to reconstructing phylogenetic divergence times (Kishino et al. 2001, Thorne and Kishino 2002). In contrast to most other molecule-based studies, our results are largely consistent with the traditional morphology-based view of hexapod evolution, including the monophyly of Hexapoda, Entognatha, and Diplura, although bootstrap support for Entognatha was low and we cannot exclude the possibility of a Diplura + Insecta clade. Our results departed from recent orthodoxy in regarding Archaeognatha and Zygentoma (rather than Zygentoma alone) as the monophyletic sister group to Pterygota, a result that harkens back to the precladistic taxonomy that placed all "bristle-tailed" insects in a single order, Thysanura *sensu lato*. The Thysanura *s. lat.* hypothesis is modestly well supported by our analyses and is one of our more controversial results.

Our analysis also included an attempt to reconstruct divergence times of the major hexapod lineages using the Bayesian approach of Thorne and Kishino (2002),

[1] Center for Biosystems Research, University of Maryland Biotechnology Institute, Plant, Sciences Building, College Park, MD 20742 (e-mail: regier@umd.edu).

[2] Department of Entomology, University of Maryland, Plant Sciences Building, College Park, MD 20742.

0013-8746/04/0411–0419$04.00/0

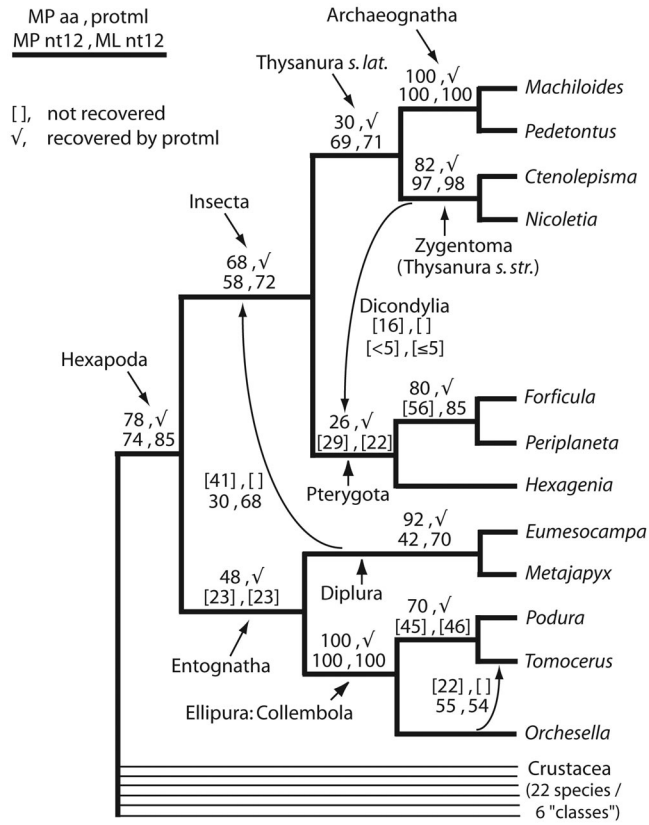


Fig. 1. Phylogenetic analysis of 12 hexapods, using 22 diverse crustaceans as outgroup. Combined *EF-1α*, *Pol II*, and *EF-2* sequences were analyzed by maximum parsimony (MP) and maximum likelihood (ML) for nucleotides (nt12: first two codon positions only) and for amino acids (aa). Amino acids were also analyzed using a modified *protml* procedure (see Materials and Methods). The topology displayed is that recovered by MP aa and protml. Onto this topology and above the interior branches are mapped (in order from left to right and upper followed by lower): bootstrap percentages (BP) for the MP aa analysis; the status of the particular node in the protml analysis (√, recovered; [ ], not recovered); BP values for MP nt12; and BP values for ML nt12. Nodes not recovered are indicated by brackets. Three alternative topologies are displayed with arrows and the corresponding BP values.

which requires an input topology, a maximum age constraint on the basal node, and minimum age constraints on one or more internal nodes. Given the input constraints, the general congruence between our results and the known fossil record is not surprising, although one discrepancy—an unexpectedly recent divergence of dictyopterans from other orthopteroids—is noted. This highlights the caution that must be used in interpreting results generated by this new method.

Materials and Methods

Taxon Sampling and Its Rationale. This study was designed to examine basal relationships in Hexapoda. The outgroup included exemplars from all six classes of Crustacea (Martin and Davis 2001), which were selected over Myriapoda and Chelicerata based on increasingly strong molecular evidence that crustaceans and hexapods form a monophyletic group, called Pancrustacea or Tetraconata (Boore et al. 1998, Friedrich and Tautz 2001, Hwang et al. 2001, Regier and Shultz 2001). However, our phylogenetic results within Hexapoda are not significantly affected by inclusion of more distant outgroups, including Chelicerata and Myriapoda (J. C. R. and J. W. S., unpublished observations). Three genes (*EF-1α*, *Pol II*, and *EF-2*) were sequenced from 12 hexapods and 22 crustaceans. Vouchers are stored in −85°C freezers in the arthropod collection of the Department of Entomology, University of Maryland. The terminal taxa are listed below with our laboratory code names and GenBank accession numbers in brackets.

HEXAPODA. Collembola: *Podura aquatica* L. 1758 (Poduridae) [Paq; AY305474; AY305604, AY305605; AY305518], *Tomocerus* sp. Nicolet 1842 (Tomoceridae) [Tom; U90059; AF139011, AF139012; AF240830], *Orchesella imitari* Snider 1998 (Entomobryidae) [Oim; AY305473; AY305599–AY305601; AY305515, AY305516]. Diplura: *Eumesocampa frigillis* Hilton 1936 (Campodeidae) [Efr; AF137388; AF138978–AF138980; AF240818], *Metajapyx subterraneus* (Packard 1874) (Japygidae) [Jap; AF137389; AF138987, AF138988; AY305503, AY305504]. Zygentoma: *Ctenolepisma lineata* (Lucas 1840) (Lepismati-

dae) [Cli; AF063405; AF138973, AY305553, AY305554; AY305494], *Nicoletia meinerti* Silvestri 1905 (Nicoletiidae) [Nme; AY305472; AY305596–AY305598; AY305514]. Archaeognatha: *Machiloides banksi* Silvestri 1911 (Meinertellidae) [Mba; AF137390; AF138990–AF138992; AF240822], *Pedetontus saltator* Wygodzinsky and Schmidt 1980 (Machilidae) [Psa; U90056; U90041, AY305610; AY305520]. Pterygota: *Hexagenia limbata* (Serville 1829) (Ephemeroptera, Hexageniidae) [May; AY305469; AY305584–AY305588; AY305510]. *Forficula auricularia* L. 1758 (Dermaptera: Forficulidae) [Fau; AY305464; AY305562–AY305564; AY305499, AY305500], *Periplaneta americana* (L. 1758) (Blattodea: Blattidae) [Pam; U90054; AY305602, U90040, AY305603; AY305517].

CRUSTACEA. MALACOSTRACA: *Armadillidium vulgare* Latreille 1804 (Isopoda) [Avu2; U90046; AY305548; AF240816], *Neogonodactylus oerstedii* (Hansen 1895) (Stomatopoda) [Neo; AY305471; AY305591–AY305593; AY305512], *Libinia emarginata* Leach 1815 (Decapoda) [Lem; U90050; AY305572–AY305574; AY305506], *Nebalia hessleri* Martin, Vetter and Cash-Clark 1996 (Leptostraca) [Nhe; AF063413; AF138996, AY305594, AY305595; AY305513]. MAXILLOPODA: Cirripedia: *Semibalanus balanoides* L. 1767 (Thoracica, Sessilia, Balanidae) [Bba; AF063404; AF138971, AF138972, AY305549; AF240817], *Chthamalus fragilis* Darwin 1854 (Thoracica, Sessilia, Chthamalidae) [Cfr; AY305462; AY305550–AY305552; AY305493], *Lepas anserifera* L. 1758 (Thoracica, Pedunculata) [Lean; AY305466; AY305569–AY305571; AY305505], *Loxothylacus texanus* Boschma 1933 (Rhizocephala) [Lox; AY305467; AY305577–AY305580; AY305508]. Copepoda: *Acanthocyclops vernalis* (Fischer 1853) (Cyclopoida) [A369; AY305458; AY305534–AY305536; AY305487], *Mesocyclops edax* (Forbes 1891) (Cyclopoida) [Meso; AY305470; AY305589, AY305590; AY305511], *Eurytemora affinis* (Poppe 1880) (Calanoida) [Eaf; AF063408; AF138977, AY305557, AY305558; AY305497]. Branchiura: *Argulus* sp. Müller 1785 [Arg2; AY305461; AY305544–AY305546; AY305491]. OSTRACODA: *Cypridopsis vidua Müller 1776 (Podocopa) [Ost; AF063414; AF138997–AF138999; AF240825], Harbansus paucichelatus* Kornicker 1958 (Myodocopa, Sarsielloidea) [Hapa; AY305465; AY305566–AY305568; AY305502], *Skogsbergia lerneri* (Kornicker 1958) (Myodocopa, Cypridinoidea) [Skle; AY305477; AY305616–AY305618; AY305522]. CEPHALOCARIDA: *Hutchinsoniella macracantha* Sanders 1955 [Hma; AF063411; AF138984–AF138986; AF240820]. REMIPEDIA: *Speleonectes tulumensis* Yager 1987 [Stu; AF063416; AF139008–AF138010; AF240829]. BRANCHIOPODA: *Triops longicaudatus* (LeConte 1846) (Notostraca) [Tlo; U90058; U90043, AY305622; AY305524], *Lynceus* sp. Müller 1776 (Diplostraca, Laevicaudata) (Lyn; AY305468; AY305581–AY305583; AY305509), *Limnadia lenticularis* (L. 1758) (Diplostraca, Spinicaudata) [Lle; AF063412; AF138989, AY305575, AY305576; AY305507], *Artemia salina* L. 1758 (Anostraca, Artemiidae) [Asa1_2; X03349; AY305547, U10331; AF240815], *Streptocephalus seali* Ryder 1879 (Anostraca, Streptocephalidae) [ufs; AY305480; AY305627–AY305630; AY305526].

All analyses were based on these 34 taxa. However, we do not describe our results for relationships among outgroup taxa. Those will appear in a separate report (unpublished observations) in which outgroups to Pancrustacea (i.e., Myriapoda, Chelicerata, Onychophora, Tardigrada) are also included. Displaying unrooted crustacean relationships in this report would serve little purpose and would greatly expand the scope (and size) of this report.

Specimen Preservation and the Data Set. Live specimens were collected into 100% ethanol at ambient temperature, held for brief periods (days to weeks) at room temperature, and stored long term at −85°C. Total nucleic acids were extracted from wet tissue weighing a few milligrams using the SV Total RNA Isolation kit (Promega, Madison, WI) with the DNase step omitted. The extracts were dissolved in water to a volume of 100 μl. Reverse transcription reactions used 0.1–1.0 μl of this as template. Specific mRNA sequences (5,334 nt total) for *EF-1α* (1,131 nt excluding terminal polymerase chain reaction [PCR] primer sequences), *Pol II* (2,025 nt excluding the terminal PCR primer sequences), and *EF-2* (2,178 nt excluding terminal PCR primer sequences) were reverse transcribed and amplified by the PCR using previously described conditions and oligonucleotide primers (Regier and Shultz 2001). In all cases, nested PCR amplifications were performed.

PCR fragments were sequenced directly from the M13 sequences present at the 5′ ends of all PCR primers using fluorescent-labeled dye terminators and an automated DNA sequencer. The PREGAP and GAP4 programs within the Staden package (Staden et al. 1999) were used to edit and assemble contigs. The Genetic Data Environment software package (version 2.2; Smith et al. 1994) was used to manually align assembled sequences and to construct nucleotide data matrices for phylogenetic analysis. MacClade (Maddison and Maddison 1992) was used to create amino acid matrices. For *EF-1α* and *Pol II*, there were no indels across the 12 hexapods. For *EF-2*, there was a single, autapomorphic, 3-nt deletion in the collembolan *Orchesella imitari.*

Phylogenetic Analysis. Maximum parsimony (MP) analyses of amino acid and nucleotide data sets were conducted with PAUP*4.0 (Swofford 1998) using equally weighted character transformations, both with and without third codon position characters. Analysis consisted of a heuristic search using TBR branch swapping with random sequence addition (100 sequence-addition replicates). Bootstrap analysis (1,000 bootstrap replications) was identical except for 10 sequence-addition replicates per bootstrap replication.

Maximum-likelihood (ML) analysis and bootstrap analysis (250 replications) of a nucleotide data set (third-codon-position nucleotides excluded) was performed with PAUP*4.0 under a general time-reversible model of sequence evolution (Rodriguez et al. 1990) with among-site-rate-heterogeneity modeled by a gamma function approximated by four discrete

rate categories and with invariant sites estimated separately (see Regier and Shultz 2001 for more details of ML tree search strategy). The favored model of sequence evolution was selected from among 56 based on a likelihood ratio test (Huelsenbeck and Rannala 1997) performed using Modeltest (version 3.06; Posada and Crandall 1998).

ML analysis of amino acid data was performed with the *protml* program with the MOLPHY software package (version 2.2; Adachi and Hasegawa 1994) and the empirical transition matrix compiled by Jones et al. (1992). All parsimony trees within eight steps of the MP tree length were read into *protml*, their likelihood scores were calculated, and the topology with the highest likelihood was selected.

Divergence-Time Analysis. Divergence time estimations were performed using Markov chain Monte Carlo procedures for Bayesian analysis as implemented in the programs of Thorne and Kishino (2002) for multiple gene data. Key features of these programs are that evolutionary rates at adjoining nodes are assumed to be autocorrelated and that each of the three genes (*EF-1α*, *Pol II*, *EF-2*) has a separate autocorrelation parameter (Thorne et al. 1998, Kishino et al. 2001). The *estbranches* program estimates branch lengths given a topology, model (we selected the JTT model), and amino acid data set, as well as the variance-covariance structure of the branch length estimates. The *multidivtime* program uses this output to estimate node divergence times for the ingroup, given upper- and lower-bound time constraints, various parameters, and estimated priors. For purposes of this study, we considered Crustacea to be outgroup to Hexapoda. Based on the fossil record, we constrained the following: *Forficula* + *Periplaneta*/Ephemeroptera (*Hexagenia*), >300 mya; Pterygota/Thysanura *s. lat.*, >392 mya; and Collembola, >400 mya (Labandeira 1994). Because both lower and upper bounds are needed to effectively constrain the SD of the estimated time, we added a fourth constraint, namely, the basal divergence within Hexapoda <500 mya. Our rationale for this fourth constraint is as follows: We assume that the basal hexapod clades shared a terrestrial ancestor and that, in the absence of land plants before 500 mya (Shear 1991, Wellman et al. 2003) and other potential food sources, it is unlikely that hexapods would have appeared, let alone diversified. Of course, we make this only as a reasonable claim forced on us by the algorithm itself and not as a certainty.

Based on preliminary runs, we selected the following parameters: sampfreq = 100; numsamps = 40,000; burnin = 2,000,000. The following priors were also chosen after exploring the effect of using a range of priors on the outcome: brownmean = 0.33, brownsd = 0.33, rttm = 5.000 (intended to refer to 500 mya), rttmsd = 1.000, bigtime = 10.000. A fixed feature of *multidivtime* is that the posterior rate (and its prior, called *rtrate*) at the basal ingroup node and one of the two derived, adjoining nodes (decided arbitrarily) must be identical. Preliminary studies demonstrated that the particular value chosen for *rtrate* had a significant effect on the time estimate for nodes near the base. Therefore, we used two distinct values of *rtrate.* In one case, the prior value was assumed to represent the average value across the entire ingroup based on the output from *estbranches* (*rtrate* = 1.00, *rtratesd* = 0.50). For the other, we assumed the rate to be three times faster (*rrate* = 3.00, *rtratesde* = 0.50), based on the idea that rapid changes may be confined to the early history of a lineage (Gould 1989, 1991). Divergence time estimates are based on two topologies—that displayed in Fig. 1 plus another in which Diplura is a sister group to Insecta (Kukalová-Peck 1987, 1991), with all other relationships being the same as in Fig. 1. With two values of *rtrate* and two topologies, four sets of divergence time estimates were obtained. For each node, we display the minimum and maximum values, along with their corresponding SDs.

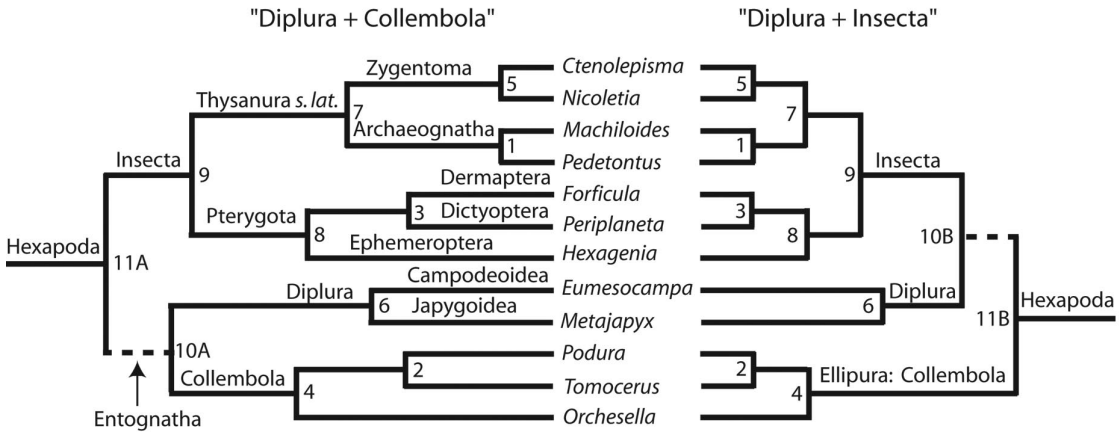
Results

Phylogenetic Inferences. Deep-level relationships across hexapods were inferred based on combined analysis of coding sequence data from three nuclear genes (*EF-1α*, *Pol II*, *EF-2*). The preferred topology from parsimony and *protml* analysis of amino acids is shown, along with node support values for amino acids and separately for nucleotides excluding third codon positions (Fig. 1). The monophyly of Hexapoda is recovered by all approaches with moderate to strong support. Other nodes consistently recovered with moderate to strong support are Insecta, Collembola, Diplura (including a member each of Campodeidae and Japygidae), Thysanura *s. lat.*, and Archaeognatha. Pterygota and Entognatha are also recovered by analysis of amino acids, albeit with low node support, but are not recovered in some nucleotide analyses. Relationships among the three collembolan taxa are ambiguous. Dicondylia is never recovered and there does not seem to be even a weak underlying signal supporting it.

Parsimony analysis of total nucleotides (i.e., with third codon positions included) yields a preferred topology (data not shown) that includes none of the above-mentioned groups except Archaeognatha and Zygentoma.

Divergence-Time Inferences. The Bayesian method for divergence time estimation requires an assumed topology (Fig. 2), a model of amino acid sequence evolution (see Materials and Methods), and age constraints (see list at bottom left of Fig. 2). Because the placement of Diplura was not strongly supported by existing data, two alternative, fully dichotomous topologies were assessed (see nodes 10A and 10B in Fig. 2). The estimation procedure further required that the relative rate at one of the two basal ingroup nodes be preassigned. Our first approach was to set that rate equal to the average rate across the entire tree. To test the sensitivity of that prescription on the time estimate, we also set the rate equal to three times the average rate. The resulting absolute time estimates thus appear as a range with separately estimated SDs for the lower and upper time estimates (see bottom right of Fig. 2).

The basal clades within hexapods diverged ≈480 mya (Ordovician). The insect clade appeared between



node	constraint (mya)
8	>300
9	>392
10A, 11B	>400
11A, 11B	<500

node	min. age (s.d.) - max. age (s.d.)(mya)
1	112 (24) - 152 (34)
2	216 (29) - 324 (30)
3	245 (31) - 279 (30)
4	275 (29) - 395 (22)
5	286 (26) - 319 (26)
6	353 (31) - 394 (20)
7	357 (19) - 383 (20)
8	368 (19) - 385 (19)
9	402 (9) - 415 (17)
10A	422 (16) - 425 (17)
10B	456 (18) - 457 (13)
11A	461 (16) - 483 (12)
11B	485(11) - 488 (10)

Fig. 2. Divergence time estimates for hexapod clades. The topology on the left is the same as that shown in Fig. 1. The topology on the right differs only in the position of Diplura, which now groups with Insecta rather than Collembola The model, parameters, and priors are described in Materials and Methods. Nodes are numbered. Four constraints have been applied (bottom left). Divergence time estimates (bottom right) are displayed as ranges followed by SDs (in parentheses). Time estimates that are underlined correspond to results when priors for *rtrate* on the root node are set equal to the average rate across the entire tree (see Materials and Methods). The other estimates result when priors are set equal to three times the average rate. Ranges for nodes 1–9 are based on separate calculations of both topologies and both *rtrate* values, for a total of four time estimates at each node. Ranges for nodes 10A, 10B, 11A, and 11B are based on separate calculations of different *rtrate* values but for individual topologies only, for a total of two time estimates at each node. The most recent time estimate for node nine (see dashed underline) was recovered from both topologies but using different *rtrate* values.

422–457 mya (Silurian–Ordovician), whereas pterygotes diverged from thysanurans *s. lat.* ≈410 mya (Devonian–Silurian). The basal splits within pterygotes, thysanurans *s. lat.*, and diplurans were approximately contemporaneous at 375 mya (Devonian).

Discussion

Monophyly of Hexapoda. Hexapoda has long been regarded as monophyletic, and this view has been corroborated repeatedly by a variety of evidence. Kristensen (1981, 1991, 1995, 1997) has reviewed the putative morphological synapomorphies of hexapods and found few "strong" hexapod synapomorphies (e.g., unique tagmosis) but no compelling phylogenetic alternatives to hexapod monophyly. Still, the observation that most hexapod synapomorphies are "weak" (i.e., homoplasious) has led some traditional Hennigian systematists to suggest that hexapod monophyly is not well supported (e.g., Klass and Kristensen 2001). For example, it can be argued that segmental composition of the head is a "weak" hexapod synapomorphy because it also occurs in Myriapoda; segmentation of the postcephalic trunk is "weak" because it occurs in certain crustaceans, and so forth. Significantly, the alternative phylogenetic hypotheses implied by these homoplasies not only conflict with hexapod monophyly but also with each other and thereby fail to support a compelling alternative. By weighting the mere existence of homoplasy over the phylogenetic implications of sets of characters (whether homoplasious or not), character-by-character critiques ignore the pivotal phylogenetic information represented by the persistent association of independent apomorphies within a group of organisms. In contrast, quantitative parsimony-based methods can recover clades despite homoplasy being present in independent morphological characters and such analyses tend to recover Hexapoda unambiguously

(Wheeler et al. 1993, Edgecombe et al. 2000). Furthermore, molecule-based analyses may recover Hexapoda or not, but few have offered a convincing alternative to hexapod monophyly.

However, Nardi et al. (2003a) recently proposed that hexapods are diphyletic, with Insecta being more closely related to certain crustaceans than to Collembola. They based this conclusion on an initial maximum likelihood analysis of 1305 amino acids from the mitochondrial cytochrome oxidase complex (cox1, cox2, cox3, coxb) from 30 arthropods and five distant outgroup taxa (i.e., three annelids, a mollusk, and a tunicate). The sequences were selected from complete mitochondrial genomes for having alignments with a high number (>20%) of invariable sites and a low number (<10%) of sites with indels. Likelihood analysis recovered most insects as a monophyletic sister group to crustaceans (i.e., malacostracans and branchiopods) and with collembolans as the sister group to insects and crustaceans. The presence of a mixed chelicerate-myriapod-insect clade with placement of the honey bee and a louse as the sister group to ticks was dismissed as artifacts of biased amino acid composition and heterogeneous evolutionary rates. In an attempt to control for these biases, Nardi et al. assembled a second data set that eliminated those taxa with sequences that appeared to be biased in comparison to two phylogenetically pivotal hexapods, the plesiomorphic zygentoman *Tricholepidion* Wygodzinsky 1961 and the collembolan *Gomphiocephalus* Carpenter (1908). This procedure yielded a data set of 14 arthropods and one outgroup taxon. Their likelihood and Bayesian analyses recovered trees congruent with the those obtained from analysis of the full data set; that is, insects and collembolans were diphyletic with insects being the sister group to crustaceans. The second analysis seemed to confirm the authors' contention that peculiarities in the first analysis were caused by anomalies in the full data set, and they concluded that hexapody and terrestrialization must have occurred twice.

Delsuc et al. (2003) criticized the data and analysis used by Nardi et al. (2003a). Specifically, they questioned the use of an amino acid substitution matrix based largely on mammalian sequences, the failure to use a likelihood model that accommodates rate variation across sites, and the presence of bias in nucleotide composition that can affect amino acid composition. They argued further that existing likelihood models designed for nucleic acids are substantially more flexible and powerful than those currently available for amino acids. They suggested that the detrimental analytical effects of rapid evolution at third codon positions could be ameliorated by using ambiguity codes for purines and pyrimidines at these positions and thereby preserve potential phylogenetic information offered by transversions. To test their proposal, Delsuc et al. analyzed nucleic acid data for the same genes and taxa used by Nardi et al. but used GTR + Γ + I models within both maximum likelihood and Bayesian analyses. They recovered a monophyletic Hexapoda with an internal structure that approximated those relationships generally accepted by entomologists, even eliminating the artifactual placements of bees and lice. This result seemed to support the view that there can be more phylogenetic information within the nucleotide data than could be recovered with amino acids. Nardi et al. (2003b) responded to this critique by reaffirming the value of amino acids and by demonstrating the persistent recovery of a diphyletic Hexapoda even after accounting for rate heterogeneity among sites and biases in taxon sampling. They summarized by stating that hexapod diphyly is one of several competing hypotheses that must be tested by further evidence, but that even now it is consistent with one fossil and the results of some molecule-based studies; that is, their own work (Nardi et al. 2001, 2003a) and a tendentious interpretation of the ambiguous results of one study based on 18S rDNA (Spears and Abele 1997)

In defending their hypothesis, Nardi et al. 2003b noted fossils of a presumed marine stem hexapod, *Devonohexapodus* Haas et al. 2003 from the Lower Devonian (≈315 Ma) Hunsrück Slate in Germany. The animal has long filiform antennae, large compound eyes, an ill-defined "thorax" bearing three pairs of leg-like appendages, and a long "abdomen" containing over 20 leg-bearing somites and terminating in a pair of short cercus-like appendages. However, evidence for this specimen being a hexapod is not convincing given the preservational distortion of the anterior portion of the body, the reliance by Haas et al. (2003) on the highly questionable "Atelocerata" concept, and the sometimes strained comparisons of features in the fossil to those of extant hexapods (e.g., gonopods). Indeed, in an earlier paper, Briggs and Bartels (2001) described similar fossil arthropods from the same geological formation but placed them among crustaceans. In *Cambronatus* (Briggs and Bartels 2001) and *Eschenbachiellus* (Briggs and Bartels 2001), the last three cephalic appendages are long and leg-like. In a third, *Wingertshellicus* (Briggs and Bartels 2001), the anterior "legs" are very long but were also interpreted as cephalic rather than thoracic appendages. We strongly doubt that *Devonohexapodus* is a stem-group hexapod and question the implication by Nardi et al. (2003a, b) that extant hexapods, and even malacostracan and branchiopod crustaceans, are derived from a marine hexapod ancestor. In summary, we do not regard the analyses and evidence marshaled by Nardi et al. (2003a, b) as a convincing refutation of hexapod monophyly.

Monophyly of Entognatha. Our results are consistent with the monophyly of Entognatha (Diplura, Collembola, Protura) and its placement as sister group to Insecta, although support for this placement is not strong (Fig. 1). Entognatha is characterized by entognathy (i.e., enclosure of mouthparts by lateral extensions of the head). There are other apparent morphological synapomorphies, but these tend to be simplifications or absences (e.g., no median or compound eyes, malpighian tubules) that may be prone to convergence. Questions about entognathan monophyly have been raised repeatedly over the decades and generally focus on the phylogenetic placement of

Diplura. Manton (1977) argued that entognathy evolved independently in Diplura and Ellipura (Collembola + Protura) (see also Koch 2001), but this is a statement about evolutionary morphology not phylogeny per se. Later, Kukalová-Peck (1987) hypothesized Diplura to be more closely related to Insecta than to Ellipura based principally on the morphology of a dipluran-like Carboniferous hexapod, *Testajapyx* (Kukalová-Peck 1987), and speculation about the homology of basal podomeres in the thoracic legs (Kristensen 1997). Still, Kukalová-Peck's (1987) hypothesis has been corroborated in recent analyses that combined morphology and molecular data (Edgecombe et al. 2000, Wheeler et al. 2001a, b). Results from most of our analyses favor a monophyletic Entognatha, but greater support for a Diplura + Insecta clade is evident under certain analytical conditions (Fig. 1). In our view, the problem of entognathan phylogeny remains effectively unresolved.

Thysanura *sensu lato* Versus Dicondylia. Our analyses indicate that the two extant orders of bristletailed insects, Archaeognatha and Zygentoma, form a monophyletic Thysanura *sensu lato* that forms the sister group to Pterygota. Thysanuran monophyly is one of the more strongly supported conclusions of our study and probably one of the more controversial. In the precladistic past, hexapod systematists tended to place all insectan "bristletails" in the order Thysanura. Snodgrass (1938, 1951) considered Zygentoma to be more closely related to Pterygota than to Archaeognatha, even though he continued to use "Thysanura" to encompass all bristletails. In applying phylogenetic methods to hexapods, Hennig (1981) emphasized the plesiomorphic nature of the bristletail body form and formalized the relationship between Zygentoma and Pterygota with the clade Dicondylia, a name inspired by the dicondylic articulation of each mandible with the head. Dicondylia is widely treated as an effectively irrefutable clade in most entomological texts, and indeed, Kristensen (1997) has defended its monophyly with 13 synapomorphies. It must be admitted at the outset that our results do not readily accommodate the distribution of morphological characters as they are currently understood. However, there are several aspects of the morphological data that need to be highlighted when evaluating our conclusions.

A principal character uniting Dicondylia is the dicondylic mandible. Specifically, the primitive hexapod mandible is generally thought to have had a single lateral articulation with the head, with the ancestors of Dicondylia evolving a second, more anterior articulation near or on the base of the clypeus. However, Koch (2001) has conducted a detailed comparative study of mandibular mechanisms and concluded that an effective anterior articulation is widespread in Hexapoda and is probably the ancestral condition for the group. Koch found that anterior manibular articulations are modified in various ways among the basal hexapod lineages, with their degree of development being correlated with the strength of transverse biting. In fact, there seems to be a suite of characters associated with this trait, especially increased skeletal surface area and skeletal strengthening to accommodate powerful mandibular muscles. Indeed, those hexapods with well-defined dicondylic mandibles also have large head capsules strengthened by postoccipital sutures and posteriorly fused anterior tentoria, features also regarded as synapomorphic for Dicondylia by Kristensen (1997). Likewise, those features characteristic of Archaeognatha (namely, "monocondylic" mandibles, small "microcoryphian" head, incomplete postocciptal suture, anterior tentoria without transverse connection) may reflect specialization for feeding on minute particles and concomitant reduction in transverse biting force, not the retention of a primitive "archaeognathous" condition (Koch 2001).

Other purported synapomorphies of Dicondylia are also problematic. For example, presence of a gonangulum, a sclerite linking the base of the ovipositor with the ninth tergite, is often cited as synapomorphic for Dicondylia, but a small sclerite is present in the corresponding position in extant Archaeognatha (Bitsch 1994) and a fully developed gonangulum has been reported in fossil Monura (Kukalová-Peck 1987), a group that is sometimes regarded as the sister to Archaeognatha. Anastomosing abdominal tracheae are considered unique to Dicondylia by some authors (e.g., Kristensen 1997) but are said to be present in Archaeognatha by others (Hinton 1958, Hennig 1981). Short maxillary palps have also been proposed as synapomorphic for Dicondylia, but this assumes that the pediform palps of Archaeognatha and Monura are primitive and ignores the large palps of the fossil "zygentoman" *Ramsdelepidion* (Kukalová-Peck 1987). Other characters are generally known to be homoplasious (e.g., loss of postcephalic spina, tarsi with five segments). Thus, there are some reasons for not accepting the Dicondylia hypothesis as conclusive.

Still, as we have noted above, acceptance of Thysanura *sensu lato* requires corroborating evidence, not simply a critique of characters that support alternative hypotheses. Other than similar results from a phylogenetic analysis of 18S rDNA (Wheeler et al. 2001a, b), which did not survive within a total evidence analysis by the same authors, there seems to be little corroborating information available. It is unclear whether this represents an actual absence of such information or a bias toward accumulating data that supports Dicondylia; only additional morphological work would provide an answer. However, the absence of epipodial "protowings" and reduction of the pleuron could be an interesting (albeit highly speculative) synapomorphy of Thysanura *s.lat.* given the emerging view that insect wings evolved from epipodial gills of crustacean-like hexapod ancestors (Kukalová-Peck 1983, 1987, 1991, Averof and Cohen 1997). If the epipodial wing hypothesis is valid, the absence of functional epipodia or their derivatives in Archaeognatha and Zygentoma must be accounted for in any phylogenetic scheme. In the context of our phylogenetic results, it would be accounted for as a synapomorphic loss.

Estimated Phylogenetic Divergence Times. The known hexapod fossil record is fairly extensive (Labandeira and Sepkoski 1993, Labandeira 1994) but

does not provide enough information to determine the timing of major events in hexapod evolution, such as the origin of terrestriality or flight. It is clear from fossils that terrestrial Hexapoda existed in the Lower Devonian (≈390 mya) and had already undergone the basal diversification of its extant members (Whalley and Jarzembowski 1981, Labandeira et al. 1988). Furthermore, winged insects had originated and given rise to Neoptera by the Middle Carboniferous (≈320 mya) (Labandeira 1994). These fossil-based dates provide a lower limit on the age of these evolutionary events but provide no upper limit; for example, the existence of winged insects in the Silurian cannot be excluded. The possibility that nucleotide and amino acid substitutions occur in a clocklike manner has inspired many workers to suggest that questions left open by paleontology can be filled by results of molecule-based analyses of extant organisms.

Before our study, Gaunt and Miles (2002) attempted to reconstruct the temporal structure of early hexapod diversification using molecular data. Detailed comparison of their results with ours is not possible given differences in taxon sampling, but several points are noteworthy. First, Gaunt and Miles predicted that hexapods diverged from an assumed sister group, Branchiopoda, during the Silurian (428–419 mya), whereas our analysis suggested that Hexapoda was already in existence and had undergone a basal divergence by this time (Ordovician, 488–461 mya). This discrepancy may be because of the conflation by Gaunt and Miles of the origin of the branchiopodan sister group (which may include other unrepresented crustaceans as well as Hexapoda) and the origin and terrestrialization of Hexapoda alone. Second, Gaunt and Miles reconstructed the divergence of dictyopterans from other orthopteroids as having occurred sometime between the Middle Devonian and early Lower Carboniferous (380–354 mya), and our analysis indicated a more recent Permian divergence (279–245 mya). The extensive fossil record of Dictyoptera during the Carboniferous (Labandeira 1994) is congruent with the predictions of Gaunt and Miles and effectively falsifies the relevant hypothesis derived from our analysis. Both analyses predicted a Devonian origin for insect flight. Perhaps a central conclusion should be that skepticism toward all methods of dating is warranted, although disregarding all results because a subset is clearly incorrect would be counterproductive. Molecular data remain a treasure trove of historical information, but extracting that treasure will require persistent exploration.

Acknowledgments

This work was financially supported by National Science Foundation Grant DEB 9981970 and the Maryland Agricultural Experiment Station.

References Cited

Adachi, J., and M. Hasegawa. 1994. Programs for molecular phylogenetics, version 2.2. Institute of Statistical Mathematics, Tokyo, Japan.

Averof, M., and M. Akam. 1995. Insect-crustacean relationships: insights from comparative developmental and molecular studies. Phil. Trans. R. Soc. Lond. B. Biol. Sci. 347: 293–303.

Averof, M., and M. Cohen. 1997. Evolutionary origin of insect wings from ancestral gills. Nature 385: 627–630.

Bitsch, J. 1994. The morphological groundplan of Hexapoda: critical review of recent concepts. Ann. Soc. Entomol. Fr. 30: 103–129.

Boore, J. L., D. V. Lavrov, and W. M. Brown. 1998. Gene translocation links insects and crustaceans. Nature 392: 667–668.

Briggs, D.E.G., and C. Bartels. 2001. New arthropods from the Lower Devonian Hunsrück Slate (Lower Emsian, Rhenish Massif, western Germany). Palaeontology 27: 843–855.

Delsuc, F., M. J. Phillips, and D. Penny. 2003. Comment on "Hexapod origins: monophyletic or paraphyletic?" Science 301: 1482D.

Dohle, W. 2001. Are the insects terrestrial crustaceans? Ann. Soc. Entomol. Fr. 37: 85–103.

Edgecombe, G. D., G.D.F. Wilson, D. J. Colgan, M. R. Gray, and G. Cassis. 2000. Arthropod cladistics: combined analysis of histone H3 and U2 snRNA sequences and morphology.Cladistics 16: 155–203.

Friedrich, M., and D. Tautz. 2001. Arthropod rDNA phylogeny revisited: A consistency analysis using Monte Carlo simulation. Ann. Soc. Entomol. Fr. 37: 21–40.

Gaunt, M. W., and M. A. Miles. 2002. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol. Biol. Evol. 19: 748–761.

Gould, S. J. 1989. Wonderful life. W. W. Norton & Co., New York.

Gould, S. J. 1991. The disparity of the Burgess Shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. Paleobiology 17: 411–423.

Haas, F., D. Waloszek, and R. Hartenberger. 2003. *Devonohexapodus bocksbergensis*, a new marine hexapod from the Lower Devonian Hunsruck Slates, and the origin of Atelocerata and Hexapoda. Org. Divers. Evol. 3: 39–54.

Hennig, W. 1981. Insect phylogeny. Wiley, Chichester, United Kingdom.

Hinton, H. E. 1958. The phylogeny of the panorpoid orders. Annu. Rev. Entomol. 3: 181–206.

Huelsenbeck, J. P., and B. Rannala. 1997. Phylogenetic methods come of age: Testing hypotheses in an evolutionary context. Science 276: 227–232.

Hwang, U. W., M. Friedrich, D. Tautz, C. J. Park, and W. Kim. 2001. Mitochondrial protein phylogeny joins myriapods with chelicerates. Nature 413: 154–157.

Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8: 275–287.

Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18: 352–361.

Klass, K.-D., and N. P. Kristensen. 2001. The ground plan and affinities of hexapods: recent progress and open problems. Ann. Soc. Entomol. Fr. 37: 265–298.

Koch, M. 2001. Mandibular mechanisms and the evolution of hexapods. Ann. Soc. Entomol. Fr. 37: 129–174.

Kristensen, N. P. 1981. Phylogeny of insect orders. Annu. Rev. Entomol. 26: 135–157.

Kristensen, N. P. 1991. Phylogeny of extant hexapods, pp. 125–140. *In* I. D. Naumann (ed.), The insects of Australia, vol. 1, 2nd ed. Cornell University Press, Ithaca, NY.

Kristensen, N. P. 1995. Forty years insect phylogenetic systematics: Hennig's "Kritische Bemerkungen" and subsequent developments. Zool. Beitr. 36: 83–124.

Kristensen, N. P. 1997. The groundplan and basal diversification of the hexapods, pp. 281–293. *In* R. A. Fortey and R. H. Thomas (eds), Arthropod relationships. Chapman & Hall, London, England.

Kukalová-Peck, J. 1983. Origin of the insect wing and wing articulation from the arthropodan leg. Can. J. Zool. 61: 1618–1669.

Kukalová-Peck, J. 1987. New Carboniferous Diplura, Monura, and Thysanura, the hexapod ground plan, and the role of thoracic side lobes in the origin of wings (Insecta). Can. J. Zool. 65: 2327–2345.

Kukalová-Peck, J. 1991. Fossil history and the evolution of hexapod structures, pp. 141–179. *In* I. D. Naumann (ed.), The insects of Australia, vol. 1, 2nd ed. Cornell University Press, Ithaca, NY.

Labandeira, C. C. 1994. A compendium of fossil insect families. Contrib. Biol. Geol. Milwaukee Public Mus. 88: 1–71.

Labandeira, C. C., and J. J. Sepkoski Jr. 1993. Insect diversity in the fossil record. Science 261: 310–315.

Labandeira, C. C., B. S. Beall, and F. M. Hueber. 1988. Early insect diversification: evidence from a Lower Devonian bristletail from Quebec. Science 242: 913–916.

Maddison, W. P., and D. R. Maddison. 1992. MacClade: analysis of phylogeny and character evolution, version 4.04 for Mac OS X. Sinauer, Sunderland, MA.

Manton, S. M. 1977. The arthropoda: habits, functional morphology, and evolution. Clarendon Press, Oxford, United Kingdom.

Martin, J. W., and G. E. Davis. 2001. An updated classification of the recent crustacea. Nat. Hist. Mus. Los Angeles County Sci. Series 39: 1–124.

Nardi, F., A. Carapelli, P. P. Fanciulli, R. Dallai, and F. Frati. 2001. The complete mitochondrial DNA sequence of the basal hexapod *Tetrodontophora bielanensis:*evidence for heteroplasmy and tRNA translocations. Mol. Biol. Evol. 18: 1293–1304.

Nardi, F., G. Spinsanti, J. L. Boore, A. Carapelli, R. Dallai, and F. Frati. 2003a. Hexapod origins: monophyletic or paraphyletic? Science 299: 1887–1889.

Nardi, F., G. Spinsanti, J. L. Boore, A. Carapelli, R. Dallai, and F. Frati. 2003b. Response to coment on "Hexapod origins: monophyletic or paraphyletic?" Science 301: 1482E.

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14(9): 817–818.

Regier, J. C., and J. W. Shultz. 2001. Elongation factor-2: a useful gene for arthropod phylogenetics. Mol. Phylogen. Evol. 20: 136–148.

Rodriguez, F., J. L. Oliver, A. Marin, and J. R. Medina. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142: 485–501.

Shear, W. A. 1991. The early development of terrestrial ecosystems. Nature (Lond.). 351: 283–289.

Smith, S. W., R. Overbeek, C. R. Woese, W. Gilbert, and P. M. Gillevet. 1994. The genetic data environment and expandable GUI for multiple sequence analysis. Comput. Appl. Biosci. 10: 671–675.

Snodgrass, R. E. 1938. Evolution of the Annelida, Onychophora, and Arthropoda. Smithson. Misc. Coll. 97: 1–159.

Snodgrass, R. E. 1951. Comparative studies of the head of mandibulate arthropods. Comstock Publishing Co, Ithaca, NY.

Spears, T., and L. G. Abele. 1997. Crustacean phylogeny inferred from 18S rDNA, pp. 169–187. *In* R. A. Fortey and R. H. Thomas (eds.), Arthropod relationships. Chapman & Hall, London, England.

Staden, R., K. R. Beal, and J. K. Bonfield 1999. The Staden package, 1998, pp. 115–130. *In* S. Misener and D. Krawetz (eds.), Bioinformatics methods and protocols. Humana Press, Totowa, NJ.

Swofford, D. L. 1998. PAUP*, version 4.0 vB10. Sinauer, Sunderland, MA.

Thorne, J. L., and H. Kishino. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51: 689–702.

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15: 1647–1657.

Wellman, C. H., P. L. Osterloff, and U. Mohiuddin. 2003. Fragments of the earliest land plants. Nature (Lond.) 425: 282–285.

Whalley, P., and E. A. Jarzembowski. 1981. A new assessment of *Rhyniella,* the earliest known insect, from the Devonian of Rhynie, Scotland. Nature 291: 317.

Wheeler, W. C., P. Cartwright, and C. Y. Hayashi. 1993. Arthropod phylogeny: a combined approach. Cladistics 9: 1–39.

Wheeler, W. C., M. Whiting, Q. D. Wheeler, and J. M. Carpenter. 2001a. The phylogeny of the extant hexapod orders. Cladistics 17: 113–169.

Wheeler, W. C., M. Whiting, Q. D. Wheeler, and J. M. Carpenter. 2001b. Erratum: The phylogeny of the extant hexapod orders. Cladistics 17: 403–404.

Zhang, Y. Z., Y. P. Zhang, Y. X. Luan, Y. J. Chen, and W. Y. Lin. 2001. Phylogeny of higer taxa of Hexapoda according to 12s RNA sequences. Chin. Sci. Bull. 46: 840–842.

Received 14 October 2003; accepted 23 December 2003.