Available online at www.sciencedirect.com
ScienceDirect
Zoology 110 (2007) 409–429
ZOOLOGY
www.elsevier.de/zool

Towards an 18S phylogeny of hexapods: Accounting for group-specific character covariance in optimized mixed nucleotide/doublet models

Bernhard Misof[a,*], Oliver Niehuis[a], Inge Bischoff[a], Andreas Rickert[a], Dirk Erpenbeck[b], Arnold Staniczek[c]

[a]*Abteilung für Entomologie, Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, D-53113 Bonn, Germany*
[b]*Department of Coelenterata and Porifera (Zoologisch Museum), Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94766, 1090 GT Amsterdam, The Netherlands*
[c]*Staatliches Museum für Naturkunde Stuttgart, Abt. Entomologie, Rosenstein 1, D-70191 Stuttgart, Germany*

Received 19 May 2007; received in revised form 2 August 2007; accepted 22 August 2007

Abstract

The phylogenetic diversification of Hexapoda is still not fully understood. Morphological and molecular analyses have resulted in partly contradicting hypotheses. In molecular analyses, 18S sequences are the most frequently employed, but it appears that 18S sequences do not contain enough phylogenetic signals to resolve basal relationships of hexapod lineages. Until recently, character interdependence in these data has never been treated seriously, though possibly accounting for the occurrence of biased results. However, software packages are readily available which can incorporate information on character interdependence within a Bayesian approach. Accounting for character covariation derived from a hexapod consensus secondary structure model and applying mixed DNA/RNA substitution models, our Bayesian analysis of 321 hexapod sequences yielded a partly robust tree that depicts many hexapod relationships congruent with morphological considerations. It appears that the application of mixed DNA/RNA models removes many of the anomalies seen in previous studies. We focus on basal hexapod relationships for which unambiguous results are missing. In particular, the strong support for a "Chiastomyaria" clade (Ephemeroptera + Neoptera) obtained in Kjer's [2004. Aligned 18S and insect phylogeny. Syst. Biol. 53, 1–9] study of 18S sequences could not be confirmed by our analysis. The hexapod tree can be rooted with monophyletic Entognatha but not with a clade Ellipura (Collembola + Protura). Compared to previously published contributions, accounting for character interdependence in analyses of rRNA data presents an improvement of phylogenetic resolution. We suggest that an integration of explicit clade-specific rRNA structural refinements is not only possible but an important step in the optimization of substitution models dealing with rRNA data.

Keywords: SSU rRNA; DNA/RNA substitution models; Insect phylogeny; Character interdependence; Bayesian analysis

Introduction

This study focuses on an analysis of insect relationships based on nuclear SSU rRNA sequences. An investigation of insect relationships based exclusively

*Corresponding author.
E-mail address: b.misof.zfmk@uni-bonn.de (B. Misof).

0944-2006/$ - see front matter © 2007 Elsevier GmbH. All rights reserved.
doi:10.1016/j.zool.2007.08.003

on SSU data may appear as a fruitless task since these questions are among the most intensively investigated phylogenetic problems within the field of invertebrate systematics and 18S data have not delivered the desired resolution (Carmean et al., 1992; Liu and Beckenbach, 1992; Pashley et al., 1993; Chalwatzis et al., 1996; Whiting et al., 1997; Huelsenbeck, 1998; Hwang et al., 1998; Carapelli et al., 2000; Caterino et al., 2000; Giribet et al., 2001, 2004; Wheeler et al., 2001; Delsuc et al., 2003; Nardi et al., 2003; Ogden and Whiting, 2003; Luan et al., 2003, 2005; Kjer, 2004; Mallatt and Giribet, 2006).

However, one aspect of rRNA evolution – the interdependence of character variation among paired positions (see for example Fitch and Markowitz, 1970; Galtier, 2004; Felsenstein, 2004) – has been largely ignored (for an exception see Kjer, 2004). But recent work has demonstrated, from a theoretical and an empirical point of view, that character dependence in rRNA or tRNA gene sequences must not be ignored (Stephan, 1996; Schöniger and von Haeseler, 1994; Higgs, 1998; Tillier and Collins, 1998; Savill et al., 2001; Jow et al., 2002; Hudelot et al., 2003; Galtier, 2004). These authors showed that ignoring correlated variance inflates bootstrap support and may lead to biased reconstructions by overemphasizing changes in paired positions (see for example Schöniger and von Haeseler, 1994; Jow et al., 2002; Galtier, 2004). Kjer (2004) published a reanalysis of hexapod SSU rRNA gene sequences based on a structurally aided sequence alignment and an application of mixed DNA/RNA substitution models implemented in MrBayes 3.0 (Huelsenbeck and Ronquist, 2001, 2002). He used a canonical secondary structure consensus model of the 18S rRNA molecule fitted to hexapod sequences (Kjer, 2004).

In this investigation we will build on the published alignment of Kjer (2004), refining the structure model and alignment with a taxon-specific hexapod secondary structure consensus model (Misof et al., 2006) based on a more than twofold extension of the taxon sampling to achieve a better representation of the structural variance, improving the fit of the RNA substitution models by monitoring parameter convergence and scoring model fit with a Bayes factor test.

The idea behind this procedure is the assumption that a biologically more realistic representation of the evolutionary processes will also aid in achieving a more realistic reconstruction of phylogenetic relationships.

Problems in insect systematics and molecular evidence

Hennig (1969) provided the first comprehensive treatment of insect systematics and set the stage for any further work. We will focus on basal hexapod relationships and the early divergence of pterygote insects. Several molecular analyses have been published focusing on arthropod or insect relationships (Liu and Beckenbach, 1992; Chalwatzis et al., 1996; Whiting et al., 1997; Wheeler, 1998; Giribet and Ribera, 2000; Giribet et al., 2001, 2004; Wheeler et al., 2001; Hovmöller et al., 2002; Whiting, 2002a, b; Ogden and Whiting, 2003; Misof and Fleck, 2003; Luan et al., 2003, 2005; Nardi et al., 2003; Delsuc et al., 2003; Kjer, 2004; Mallatt and Giribet, 2006). Most of these studies did not address the relationship of basal hexapod and pterygote lineages, but instead used single representatives of these orders as outgroups (for example, Chalwatzis et al., 1996; Whiting et al., 1997).

Hovmöller et al. (2002) included two representatives of Archaeognatha as outgroups. Representatives of Zygentoma were excluded due to extreme variations in published sequences. The data included 8 ephemeropteran, 18 odonate, and 9 neopteran taxa represented by sequences of the nuclear SSU rRNA gene and a nuclear LSU rRNA fragment. The phylogenetic reconstruction was based on a CLUSTAL X alignment (Thompson et al., 1997) using the parsimony reconstruction method. The analysis yielded a strongly supported monophyly of the Palaeoptera clade based on jackknife support values. This analysis, however, was biased by an uneven taxon sampling, i.e., Neoptera and ''apterygote'' taxa were unevenly represented. Nevertheless, high jackknife support values for a Palaeoptera clade deserve attention. Subsequently, Ogden and Whiting (2003) reanalyzed the data adding taxa to balance the inadequate sampling of Hovmöller et al. (2002). As outgroup taxa they added representatives of Collembola, Diplura and Zygentoma. Furthermore, they included sequence information from the H3 Histone gene. The phylogenetic reconstruction was based on an optimization alignment implemented in POY (Gladstein and Wheeler, 1997–2002, available at ftp.amnh.org/pub/molecular) and a sensitivity analysis to explore potential phylogenetic signal. The molecular data provided unreliable information on the basal diversification of winged insects and only the inclusion of morphological data in a total evidence approach supported the Metapterygote hypothesis. The authors concluded that the presently available molecular data is not suitable to resolve the basal diversification of winged insects. In contrast to this, Kjer (2004) obtained a strongly supported ''Chiastomyaria'' clade. However, jackknife support values (Ogden and Whiting, 2003) and Bayesian support values (Kjer, 2004) cannot be directly compared.

Kjer's results are encouraging, but not convincingly supported on morphological grounds. It will be the task of this investigation to study whether the Chiastomyaria hypothesis is indeed favored by mixed DNA/RNA model-based approaches. If true, it could refresh the discussion on basal winged insect evolution.

Table 1. Taxa list

Diptera
X89496 *Tipula* sp.
AJ244427 *Sergentomyia fallax cypriotica*
AJ244420 *Sergentomyia minuta*
X57172 *Aedes albopictus*
AY325014 *Mydas clavatus*
AY325037 *Apiocera* sp. D054
AY325015 *Promachus bastardii*
AY325039 *Efferia nemoralis*
AY325040 *Proctacanthus nearno*
AY325022 *Philonicus arizonensis*
AY325041 *Machimus* sp. D058
AY325038 *Scenopinus fenestralis*
AY325035 *Ozodiceromyia costalis*
AY325034 *Leptogaster* sp. D051
AY325043 *Psilonyx annulatus*
AY325017 *Maira* sp. D034
AY325024 *Smeryngolaphria* sp. D041
AY325036 *Hemipenthes jaennickeana*
AY325042 *Laphria* sp. D059
AY325028 *Andrenosoma fulvicauda*
AY325029 *Adelodus* sp. D046
AY325021 *Ospriocerus latipennis*
AY325044 *Stenopogon martini*
AY325019 *Diogmites grossus*
AY325023 *Holopogon currani*
AY325032 *Hypenetes critesi*
AY325027 *Laphystia* sp. D044
AY325026 *Saropogon fletcheri*
Siphonaptera
X89486 *Archaeopsylla erinacea*
Mecoptera
X89487 *Boreus* sp.
AF286284 *Apteropanorpa evansi*
AF286287 *Merope tuber*
DQ008167 *Panorpa communis*
DQ008168 *Panorpa acuta*
DQ008169 *Panorpa maculosa*
DQ008170 *Panorpa alpina*
DQ008171 *Panorpa claripennis*
DQ008172 *Panorpa cognata*
DQ008173 *Panorpa fluvicaudaria*
DQ008174 *Panorpa helena*
DQ008175 *Panorpa nebulosa*
DQ008176 *Panorpa similis*
DQ008177 *Panorpa vulgaris*
DQ008178 *Panorpa multifasciata*
DQ008179 *Neopanorpa* sp.
AF423912 *Brachypanorpa oregonensis*
Lepidoptera
X89491 *Galleria mellonella*
AF423786 *Hyles lineata*
AF423783 *Prionoxystus robiniae*
AF286273 *Hemileuca* sp.
AF423785 *Anthocharis sara*
AF423784 *Platyptilia* sp.
X89497 *Sialis* sp.
X89494 *Phaeostigma notata*
Thysanoptera
AF286299 *Papilio troilus*
AF535029 *Attacus ricini*
Trichoptera
X89483 *Hydropsyche* sp.
AF286292 *Pycnopsyche lepida*
AF286300 *Oecetis avara*
AF423801 *Oxyethira dualis*
Hymenoptera
AJ009328 *Ephedrus niger*
AJ009322 *Aphidius funebris*
AJ009353 *Xenostigmus bifasciatus*
AJ009348 *Protaphidius wissmannii*
AJ009344 *Pauesia pini*
AJ009331 *Lysiphlebus confusus*
X77785 *Polistes dominulus*
X89492 *Leptothorax acervorum*
Strepsiptera
X89441 *Mengenilla chobauti*
X89440 *Xylops vesparum*
X77784 *Stylops melittae*
Coleoptera
AF012503 *Loricera foveata*
AF012508 *Pamborus guerinii*
AF002800 *Calosoma scrutator*
AF012509 *Ceroglossus chilensis*
AF012510 *Cychrus italicus*
AF012486 *Laccocenus ambiguus*
AF002803 *Blethisa multipunctata aurata*
AF012500 *Oregus aereus*
AF199527 *Australphilus montanus*
AF012524 *Copelatus chevrolati renovatus*
AF199545 *Hydaticus transversalis*
AF199548 *Hyderodes schuckardi*
AF199581 *Bidessus calabricus*
AF199579 *Bidessodes mjobergi*
AF199582 *Bidessus goudoti*
AF199583 *Hydroglyphus geminus*
AF199525 *Copelatus haemorrhoidalis*
AF199544 *Hydaticus leander*
AF199547 *Rhantaticus congestus*
AF199551 *Megadytes* sp.
AF199550 *Cybister lateralimarginalis*
AF199546 *Notaticus* sp.
AF002802 *Elaphrus clairvillei*
AF002804 *Notiophilus semiopacus*
AF012522 *Systolosoma lateritium*
AF002789 *Diplochaetus planatus*
AF002790 *Pericompsus laetulus*
AF012489 *Batesiana hilaris*
AF002808 *Trachypachus gibbsii*
AF012476 *Catapiesis brasiliensis*
X77786 *Meloe proscarabaeus*
X07801 *Tenebrio molitor*
Neuropterida
X89482 *Anisochrysa carnea*
AF012527 *Oliarces clara*
AY252301 *Cryptostemma* sp. WCW 2003
U06477 *Spissistilus festinus*
U09207 *Prokelisia marginata*

Table 1. (*continued*)

U65123 *Taeniothrips inconsequens*
Psocoptera
AF423793 *Valenzuela* sp.
Phtiraptera
AY077763 *Neophilopterus incomp*
AY077778 *Haematomyzus elephant*
AY077759 *Heterodoxus calabyi*
AY077779 *Liposcelis* sp.
Hemiptera
U06480 *Philaenus spumarius*
U06478 *Okanagana utahensis*
AY252411 *Neacoryphus* sp. WCW 2003
AY252410 *Kleidocerys* sp. WCW 2003
AY324852 *Pachygrontha antennata*
AY324853 *Henestaris oschanini*
AY252262 *Udeocoris nigroaeneus*
AY252323 *Allocoris* sp. WCW 2003
AY252412 *Neoneides muticus*
AY252317 *Campyloneura virgula*
AY252367 *Deraeocoris brevis*
AY252324 *Neurocolpus arizonae*
AY252380 *Pycnocoris ursinus*
AY252291 *Adelphocoris lineolatus*
AY252246 *Leptopterna dolobrata*
AY252244 *Litomiris* sp. WCW 2003
AY252252 *Mecistoscelini* sp. WCW 2003
AY252375 *Slaterocoris* sp. WCW 2003c
AY252248 *Parthenicus* sp. WCW 2003
AY252343 *Brooksetta* sp. WCW 2003a
AY252309 *Parthenicus covilleae*
AY252372 *Melanotrichus* sp. WCW 2003
AY252344 *Lopidea bullata*
AY252286 *Parthenicus juniperi*
AY252336 *Orthotylus* sp. WCW 2003a
AY252347 *Brooksetta* sp. WCW 2003b
AY252383 *Ceratocapsus* sp. WCW 2003c
AY252365 *Lopidea bullata*
AY252345 *Aoplonema* sp. WCW 2003a
AY252359 *Pseudopsallus angularis*
AY252295 *Slaterocoris* sp. WCW 2003b
AY252382 *Paraproba* sp. WCW 2003
AY252397 *Heterotoma meriopterum*
AY252331 *Tupiocoris* sp. WCW 2003a
AY252335 *Phymatopsallus* sp. WCW 2003b
AY252350 *Oligotylus ceanothi*
AY252296 *Psallovius piceicola*
AY252254 *Pilophorus gracilis*
AY252348 *Tuxedo* sp. WCW 2003b
AY252305 *Atractotomus acaciae*
AY252313 *Megalopsallus froeschneri*
AY252349 *Chlamydatus becki*
AY252330 *Rhinacloa forticornis*
AY252265 *Chorotingis* sp. WCW 2003
AY252320 *Orius* sp. WCW 2003b
AY252304 *Hypselosoma hickmani*
AY252303 *Pateena elimata*
AY252300 *Ceratocombus australiensis*
AF220572 *Periplaneta americana*
AF220575 *Polyphaga aegyptiaca*
X89495 *Rhaphigaster nebulosa*
Orthoptera
AY037173 *Oxya chinensis*
Z97573 *Oedipoda coerulescens*
Z97590 *Batrachotetrix* sp.
Z97568 *Glauia terrea*
Z97581 *Rhainopomma montanum*
Z97588 *Xyronotus aztecus*
Z97576 *Physemacris variolosa*
Z97562 *Bullacris membracioides*
Z97585 *Systella rafflesi*
Z97579 *Prosphena scudderi*
Z97571 *Homeomastax dentata*
Z97567 *Euschmidtia cruciformis*
Z97578 *Prosarthria teretrirostris*
Z97583 *Stiphra robusta*
Z97570 *Hemideina crassidens*
Z97566 *Cyphoderris monstrosus*
Z97582 *Ruspolia nitidula*
Z97564 *Comicus campestris*
Z97563 *Ceuthophilus carlsbadensis*
Z97587 *Tettigonia viridissima*
Z97565 *Cylindraustralia kochii*
Z97586 *Tanaocerus koebeli*
Z97560 *Acrida turrita*
X95741 *Acheta domesticus*
Z97589 *Trigonopteryx hopei*
Z97631 *Batrachideidae* sp.
Phasmatodea
Z97561 *Agathemera crassa*
AY121173 *Baculum extradentatum*
AY121175 *Tropidoderus childrenii*
AY121184 *Phobaeticus heusii*
AY121158 *Aretaon asperrimus*
AY121164 *Baculini* sp. WS22
AY121152 *Oreophoetes peruana*
AY121172 *Baculum thaii*
AY121162 *Timema knulli*
Zoraptera
AF372432 *Zorotypus snyderi*
Notoptera
Z97569 *Grylloblatta rothi*
AY121138 *Grylloblatta campodeiformis*
Mantodea
AF220578 *Archimantis latistylus*
AF220577 *Paraoxypilus tasmaniensis*
AF220576 *Creobroter pictipennis*
AF246712 *Kongobatha diademata*
Isoptera
AF220579 *Tenodera angustipennis*
AF220564 *Coptotermes lacteus*
AF220565 *Serritermes serrifer*
AF220566 *Neotermes koshunensis*
AF220567 *Hodotermopsis japonica*
AF220569 *Microhodotermes viator*
AY121141 *Mastotermes darwiniensis*
Blattodea
DQ008186 *Gomphus exilis*
DQ008188 *Arigomphus cornutus*

Table 1. (*continued*)

AF220570 *Cryptocercus relictus*
AF220571 *Cryptocercus punctulatus*
AF220573 *Blattella germanica*
AB036194 *Panesthia cribrata*
Dermaptera
X89490 *Forficula* sp.
AY121131 *Doru spiculiferum*
AY121132 *Echinosoma* sp.
AY121133 *Chelisoches morio*
Embioptera
AY121135 *Teratembian*. sp. EB07
Z97593 *Oligotoma nigra*
AF423802 *Diradius vandykei*
AY121134 *Oligotoma nigra*
AY338693 *Notoligotoma* sp. EB10
Plecoptera
Z97595 *Nemoura meyeri*
AY121149 *Plumiperla diversa*
AF461256 *Isoperla obscura*
Ephemeroptera
X89489 *Ephemera* sp.
AF461255 *Anthopotamus* sp.
AF461251 *Centroptilum luteolum*
AF461252 *Stenonema* sp. RH 2002
AF461254 *Leucrocuta aphrodite*
AY121136 *Hexagenia* sp.
AF370791 *Callibaetis ferrugineus ferrugineus*
AF461253 *Hexagenia rigida*
AY338703 *Behningia* sp.
AY338700 *Ametropus neavei*
AY338701 *Lachlania saskatchewanensis*
AY338699 *Pseudiron centralis*
DQ008181 *Siphlonurus croaticus*
DQ008182 *Ritrogena* sp.
DQ008183 *Ephemerella major*
AY338705 *Polyplocia* sp.
Odonata
X89481 *Aeshna cyanea*
AF461247 *Epiophlebia superstes*
AF461232 *Brachytron pratense*
AF461240 *Leucorrhinia pectoralis*
AF461243 *Sympetrum danae*
AF461233 *Celithemis eponina*
AJ421950 *Lestes macrostigma*
AJ421952 *Lestes numidicus*
AJ421951 *Lestes virens*
AJ421948 *Sympecma fusca*
AJ421949 *Chalcolestes viridis*
AF461238 *Erythromma najas*
AF461239 *Ischnura elegans*
AF461235 *Coenagrion* sp
AF461241 *Pyrrhosoma nymphula*
AJ420944 *Enallagma cyathigerum*
AF461236 *Cordulia aenea*
AF461242 *Somatochlora flavomaculata*
DQ008184 *Gomphus externus*
DQ008185 *Stylurus intricatus*
DQ008186 *Stylurus amnicola*
DQ008189 *Dromogomphus spinosus*
DQ008190 *Onychogomphus forcipatus*
DQ008191 *Onychogomphus forcipatus*
DQ008192 *Ophiogomphus severus*
DQ008193 *Hagenius brevistylus*
DQ008194 *Oxygastra curtisi*
DQ008195 *Macromia splendens*
DQ008196 *Lindenia tetraphylla*
DQ008197 *Caliaeschna microstigma*
DQ008198 *Cordulegaster picta*
DQ008199 *Anaciaeschna isoceles*
DQ008200 *Crocothemis erythraea*
DQ008201 *Sympetrum vulgatum*
DQ008202 *Sympetrum flaveolum*
DQ008203 *Orthetrum albistylum*
DQ008204 *Libellula depressa*
DQ008205 *Libelula fulva*
DQ008206 *Tramea lacerata*
DQ008207 *Platycnemis pennipes*
DQ008208 *Calopteryx splendens*
AF461231 *Aeshna juncea*
Zygentoma
AF005458 *Lepisma* sp.
X89484 *Lepisma saccharina*
AY210811 *Ctenolepisma longicaudata*
AY338726 *Thermobia* sp. ZG01
AF370789 *Tricholepidion gertschi*
AY338728 *Tricholepidion* sp. ZG03
AY338727 *Battigrassiella* sp. ZG02
DQ008209 *Ctenolepisma* sp.
DQ008210 *Thermobia domestica*
Archaeognatha
DQ008211 *Machilidae* sp.
AF370788 *Allomachilis froggarti*
AF005457 *Dilta littoralis*
Diplura
AF173234 *Campodea tillyardi*
AY145138 *Campodea mondainii*
AY145137 *Pseudlibanocampa sinensis*
AY145136 *Lepidocampa takahashii*
AY145135 *Parajapyx isabellae*
AY145134 *Octostigma sinensis*
AY037167 *Lepidocampa weberi*
AY037168 *Parajapyx emeryanus*
Collembola
Z36893 *Crossodonthina koreana*
AF005452 *Podura aquatica*
AY037172 *Neanura latior*
AY145140 *Sphaeridia pumilis*
Z26765 *Hypogastrura dolsana*
AY037171 *Onychiurus yodai*
U61301 *Lepidocyrtus paradoxus*
Protura
AY145139 *Kenyentulus ciliciocalyci*
AY037170 *Neocondeellum dolichotarsum*
AY037169 *Baculentulus tienmushanensis*
AF173233 *Acerentulus traegardhi*

Material and methods

In total, we included SSU sequences from 321 specimens representing all major taxa within hexapods (Table 1). Most sequences were drawn from Genbank, except for sequences of several panorpid scorpionflies, dragonflies, mayflies, bristletails and silverfish (see also Misof et al., 2006). In scanning Genbank entries, we rejected sequences of SSU fragments if they spanned less than 2/3 of the entire gene. The data set were more than twofold extension of the SSU data set published by Kjer (2004). Contrary to Kjer (2004), we included the Zoraptera sequence, AF372432_*Zorotypus_snyderi*, in our analysis. This sequence spans only roughly 2/3 of the SSU gene, as does the sequence of *Taeniothrips inconsequens* (U65123).

Secondary structure and sequence alignment

The European Ribosomal RNA database (Van de Peer et al., 2000, http://oberon.fvms.ugent.be:8080/rRNA) was used to retrieve SSU sequences with co-notated structure information for every represented hexapod order. The structural notations from this database rely on a general canonical eukaryotic SSU consensus model presented in Wuyts et al. (2000). This SSU secondary structure consensus model is based on a comparative analysis of more than 3000 SSU sequences and appears to be the most reliable model currently available. We recently used the published annotated SSU sequences from this database to derive a hexapod secondary structure consensus model (Misof et al., 2006). Our analyses were based on this hexapod model.

The hexapod model was used as a structure master file in each insect order to align the remaining sequences drawn from Genbank entries. We aligned sequences using the profile alignment approach of CLUSTAL X. Aligned sequences from Misof et al. (2006) served as alignment profiles. This approach maintains the original alignment of master sequences, thus facilitating a straightforward subsequent concatenation of the sequence groups. In a final run the complete alignment was checked by eye against obvious alignment inconsistencies between groups (the alignment strategy is explained in more detail in Misof et al. (2006).

The final alignment was screened for ambiguously aligned sections which were excluded in subsequent phylogenetic reconstructions (compare also Kjer, 2004). The hexapod secondary structure contains several pseudoknots (Misof et al., 2006). These pseudoknots were removed in structure masks of the PHASE analyses (Jow et al., 2002). Finally, the alignment was compared to the aligned SSU sequences of Kjer (2004) using BioEdit (Hall, 1999).

Phylogenetic reconstructions

The secondary structure mask was used to sort sites into paired and unpaired site classes. Secondly, homogeneity of base compositions within site classes and between site classes was checked prior to phylogenetic reconstructions. In addition to the X^2 test for compositional homogeneity we checked for significant deviations of base composition between groups. The Tukey a posteriori procedure was used to identify taxa groups responsible for significant deviations from homogeneity. In subsequent phylogenetic analyses we cross-checked phylogenetic results against groupings of similar base composition.

We examined whether paired and unpaired site classes showed significantly different base compositions. Base composition was first controlled for normality (Kolmogorov-Smirnov test) and the significance of different base compositions between site classes was checked with a paired *t*-test and a Wilcoxon rank test. For the statistical analyses the SPSS 12.0 software package was used.

Finding suitable substitution models prior to phylogenetic analyses, we opted for a mixed strategy. For unpaired sites, optimal substitution models were fitted employing a NJ tree generated from the complete data set. The software packages PAUP (Swofford, 2001) and MODELTEST3.06 (Posada and Crandall, 1998) were used for these purposes. The likelihood ratio test (LRT) was used to select the optimal model parameters as implemented in MODELTEST3.06. The preference for complex models is a known disadvantage of the LRT (for detailed discussions see for example Nylander et al., 2004) and can lead to the acceptance of models without adequate parameter estimates. Consequently, after final Bayesian analyses we checked for an acceptable convergence of model parameters for the unpaired site substitution model.

The PHASE software package provides implementations of different complex RNA models (for details of RNA substitution models see Savill et al., 2001; Jow et al., 2002; Hudelot et al., 2003). Mismatches and ambiguous pairs (G–, – –, –N) occurred in the data set. Therefore, we applied variations of 16 state models that treat mismatches and ambiguous pairs explicitly. We ran Bayesian analyses for the RNA16A,B,C,D,F,I and K models without perturbing the DNA substitution model for unpaired sites. Initial settings of the Markov Chain Monte Carlo (MCMC) run were: a mixed general time reversible model plus gamma distribution of rate heterogeneity (REV + G)/RNA model, initial branch step proposal parameter = 0.03, branch length upper bound = 1.7, perturbations of the models: REV model priority = 8, RNA model priority = 24, all other proposal priorities set to 1, burn-in iterations = 120,000,

random starting model parameters, random starting trees. We assumed uniform priors on all trees, flat Dirichlet distributions of priors on base frequencies and uniform positive priors on substitution parameters, gamma parameters and branch lengths. The posterior probability density distribution for each model was generated from 200,000 generations sampled every 100th generation. The convergence phase of the ln likelihoods was evaluated by plotting the ln likelihoods. The ln likelihoods of the burn-in phase were discarded before calculating harmonic means of ln likelihoods. If convergence was not detectable in at least one parameter, we rejected the substitution model. Mixed models were compared using the Bayes factor test (Nylander et al., 2004; Niehuis et al., 2006) and the harmonic mean of the ln likelihoods for each run as suggested by Nylander et al. (2004).

The settings of the final MCMC runs in PHASE started with the following parameters: REV + G/RNA16C + G, initial branch step proposal parameter = 0.03, branch length upper bound = 1.7, perturbations of models: REV + G model priority = 8, RNA16C + G model priority = 24, all other proposal priorities set to 1, burn-in iterations = 120,000, random start model parameters, random starting tree. Again, we assumed uniform priors on all trees, flat Dirichlet distributions of priors on base frequencies and uniform positive priors on substitution parameters, gamma parameters and branch length. The final analysis was run for 5,000,000 generations sampled from every 1000th generation. The consensus tree with annotated support values was calculated in PAUP. The maximum posterior probability tree (MAP tree) is presented with estimated branch lengths and Bayesian support values.

Results

Alignment

Table 1 gives names and Genbank accession numbers of sequences. Except for several minor insect orders, all orders were represented by multiple sequences. Within each taxonomic group sequences used as profiles are indicated (Table 1). The alignment contained 321 sequences of 4309 positions compatible with the hexapod SSU structure model (Misof et al., 2006). The alignment contained sections that were characterized by arbitrary placement of indels when judged by eye. These sections exclusively represented unpaired regions. We decided to exclude these arbitrarily aligned sections (371–784, 1467–2363, 3231–3768, 4160–4243) in our phylogenetic reconstructions. The procedure reduced the data from 4309 to 2380 aligned positions. We are aware that the exclusion of data is a debated issue in molecular phylogenetics, however, we are convinced that it is by far better to exclude arbitrarily aligned blocks instead of introducing random noise in phylogenetic analyses (for a similar reasoning see for example Kjer, 2004; compare also Lutzoni et al., 2000).

Positions containing gaps or missing data were included in our analyses. Gaps and missing data were treated similarly. The likelihood approach implemented in PHASE assigns character states in a deterministic fashion to ambiguous sites. The complete alignment is available from the corresponding author upon request.

Our alignment and the alignment of Kjer (2004) were almost identical for most stem and loop regions. However, for several helices we proposed slightly different base pairs, extensions of stems or shorter stem regions. We compared base pairs of the hexapod consensus structure and Kjer's (2004) model and indicated in bold all helices for which we proposed new alignments and different secondary structures (Table 2). Clear differences between the employed consensus structure of Kjer and ours are present in helices 6, 9, 10, helices 23/e1–11 and helices 42–44.

Secondary structure(s)

We used the hexapod consensus model to identify paired positions (Misof et al., 2006). The complete alignment in FASTA format with RNA structure masks can be obtained from the corresponding author upon request. The alignment contains two structure masks, the first one corresponding to the hexapod structure (Misof et al., 2006), the second one containing no pseudoknots. A comparison of our alignment to the SSU alignment of Kjer is also available.

Base composition, model fitting, MCMC runs

Prior to fitting substitution models we analyzed the base composition of the data. After removal of invariant sites the sequences showed a significantly inhomogeneous base composition (X^2 test). Separate analyses of paired and unpaired sites, invariants excluded, showed non-homogeneous base composition for unpaired sites (paired sites: ntax = 261, mean GC% = 57, $X^2 = 558.8$ (df = 780), $p = 1.000$, unpaired sites: ntax = 261, mean GC% = 42, $X^2 = 1187.6$ (df = 780), $p = 0.0000$). Base composition among insect orders is clearly non-homogeneous (Fig. 1). Diptera had the lowest GC% content in paired and unpaired sites (paired: mean GC% = 46, unpaired: mean GC% = 32). Interestingly, "apterygote" insects, mayflies and dragonflies had the highest GC% content of all hexapods (total mean of "apterygotes" paired sites: GC% = 58, mayflies and dragonflies GC% = 59). Sequence variation was not normally distributed (Kolmogorov–Smirnov test) for

Table 2. Base pairings

Helix	5′Stem	3′Stem	Helix	Alignment
H4	2–5	1070–1073	*yes	yes
H5	18–20,22–24	1022–1027	*mod	yes
H6			***no***	***mod***
H7	114–116	976–978	yes	yes
H8	124–126,129,130,133,135,141,142,145	882–884,888–891,894–896	mod	yes
H9	***230,232,234–236***	***242–244,246,247***	***mod***	***yes***
H10	***271,272,275,276,280,281***	***325,326,334–337***	***mod***	***yes***
H10\e1	***347–351***	***820–824***	***mod***	***mod***
H11	843–846,849–852	868–872,874–876	yes	mod
H12	898–902,906–908,913–919,920–922	929–940,945–949	mod	mod
H13	953–960	965–968,972–975	yes	yes
H14	981,982,984–987	999–1004	yes	yes
H15	1006–1009	1015–1018	yes	yes
H16	1037,1039–1041,1046–1049	1056,1057,1059,1060,1063–1066	yes	yes
H17	***1083–1085,1087,1088,1090–1093,1095–1097***	***1109–1115,1117–1121***	***mod***	***yes***
H18	1130,1131	1154,1155	mod	yes
H19	1166–1171	1208–1211,1213,1214	yes	yes
H20			omi	yes
H21	1179–1182	1204–1207	yes	yes
H22	1235–1240	2949–2951,2954–2956	mod	yes
H23	***1248–1250,1252,1255,1256,1258***	***2784–2786,2793,2794,2796,2797***	***mod***	***yes***
H23\e2	***1288–1290,1292,1294–1296,1298,1299,1302–1304***	***1324–1328,1351–1355,1360,1361***	***mod***	***mod***
H23\e1	***1275–1279,1281–1283,1285,1286***	***1384–1386,1388,1389,1393–1397***	***mod***	***mod***
H23\e4	***1460–1466***	***2470–2472,2475–2478***	***mod***	***mod***
H23\e7	***2365–2368,2395,2403–2405,2407–2409,2411,2412***	***2434,2436,2438–2440,2443,2444,2462–2467***	***mod***	***mod***
H23\e8	***2483–2485***	***2571–2573***	***mod***	***yes***
H23\e9	***2498,2499***	***2554,2555***	***mod***	***yes***
H23\e10	***2503,2504***	***2512,2513***	***mod***	***yes***

H23\e11			***omi***	***yes***
H23\e12	2514–2517	2544,2547,2548,2550	mod	yes
H23\e13			omi	yes
H23\e14	2583–2588,2590,2593,2594,2596,2597	2623–2627,2631,2634–2638	mod	yes
H24	2648–2655,2664–2667	2774,2776–2780,2782,2783	mod	yes
H25	2670–2672,2675–2678,2680–2682,2689,2690,2709	2716–2718,2725–2727,2729–2732,2735,2737,2738	mod	yes
H26	2751,2752	2758,2759	yes	yes
H27	2799–2805,2808,2812,2814–2816	2826–2829,2833–2840	mod	yes
H28	2851–2853,2856,2857	2937,2938,2945–2947	yes	yes
H29	***2865–270,2873,2874***	***2903–2910***	***mod***	***mod***
H30	2925,2926	2931,2932	yes	yes
H31	***2959,2961,2962,2964–2968***	***2973,2974,2976–2978,2981–2983***	***mod***	***yes***
H32	2994–2997,2999–3006	4103–4112,4114,4115	yes	yes
H33	3013–3017	4057–4061	yes	yes
H34	3020–3029	3874,3875,3877–3884	mod	yes
H35	3034–3037	3047–3050	yes	yes
H36	3061–3067	3863–3869	yes	yes
H37	3075,3077–3079,3086,3088–3096	3120–3125,3128–3131,3133–3136	mod	mod
H38	3140,3141,3143,3144,3147–3154,3156–3160	3835–3840,3842–3845,3850–3856	mod	yes
H39	3162–3164	3206,3207,3209	mod	yes
H40	3168–3171	3176–3179	yes	yes
H41	3183,3185,3187	3195,3196,3198	mod	yes
H42	***3215,3216***	***3830,3831***	***mod***	***mod***
H44	3799–3801,3803,3805	3815–3819	mod	mod
H45	3890–3895,3904–3906	3978–3980,3988–3990,4009–4011	yes	yes
H47	4025–4029	4042–4046	yes	yes
H48	4067–4070,4072–4075	4084–4091	yes	yes
H49	***4129–4135,4142–4147,4149–4151,4154–4157***	***4279–4282,4285–4287,4289,4290,4292–4294,4296,4301–4306,4308***	***mod***	***mod***

*yes – similar to Kjer, 2004; mod – modified compared to Kjer, 2004; **helices in bold** – different compared to Kjer, no – not folded

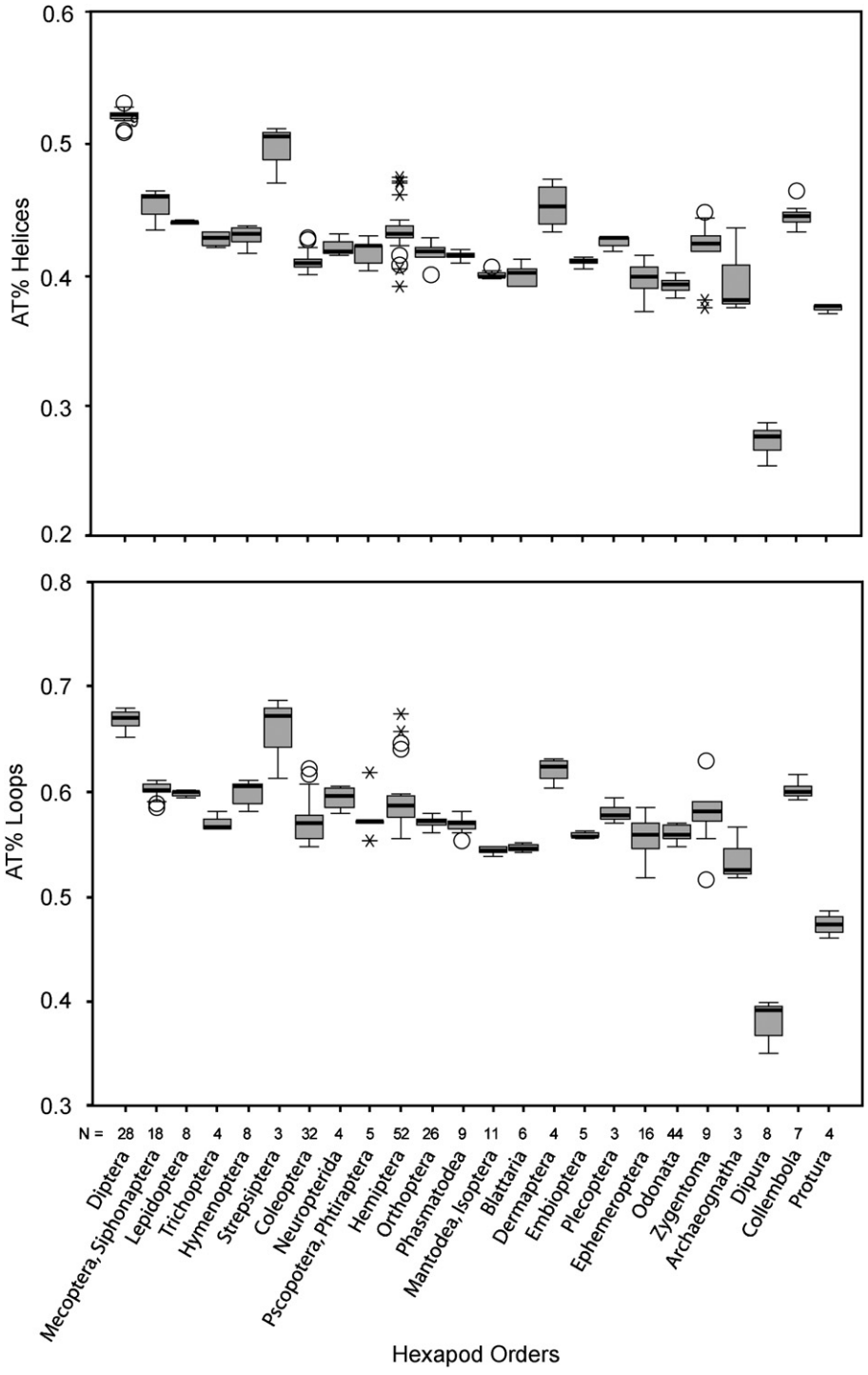

Fig. 1. Base composition within hexapod orders.

most insect orders. The GC% content was significantly different between insect orders (Kruskal–Wallis test, $p<0.0001$, df = 12). A post hoc Tukey test identified Diptera as the most deviating insect order for both paired and unpaired sites. However, after removal of dipteran sequences, differences in base composition between insect orders were still significant for unpaired sites.

Paired sites show a much higher GC% content compared to unpaired sites. We checked whether this difference was significant. First, the GC% content of paired and unpaired sites within sequences deviated from normality (Kolmogorov–Smirnov test, paired: $p = 0.015$, $n = 261$, unpaired: $p = 0.002$, $n = 261$). A visual inspection of the frequency histogram shows a roughly symmetrical distribution. We used non-parametric and

parametric tests to show that paired and unpaired sites have significantly different GC% content (Wilcoxon sign rank test, $p < 0.0001$, paired t-test, $p < 0.0001$).

The comparison of harmonic means of overall likelihoods between runs employing the Bayes factor test favored a GTR + G/RNA16C + G model HM: $-\ln L = 52902.40$, (RNA16A + G: $-\ln L = 53058.29$, RNA16B + G: $-\ln L = 53791.43$, RNA16D + G: $-\ln L = 53234.09$, RNA16F + G: $-\ln L = 53542.33$, RNA16I + G: $-\ln L = 53633.47$, RNA16K + G: $-\ln L = 53452.37$; $2B_{01} = 2(\text{RNA16A + G–RNA16C + G}) = 311.78$). The presented nucleotide pair substitution parameters show that the exchange of Watson–Crick pairs is by far the most frequent substitution in paired sites. In comparison, the number of intermediate non-Watson–Crick pairs is much lower. The observable frequency of mismatches is quite high in our data. This is probably best explained by the local inadequacy of the general hexapod SSU model (also see examples of two helices with frequency of mismatches in Table 3). For example, in proturans and diplurans several helices are different.

Phylogenetic reconstructions

A majority rule consensus tree inferred from the set of sampled trees was rooted between Entognatha and Ectognatha (Fig. 2). The tree shows the expected paraphyly of basal insects *sensu stricto* with Archaeognatha as sister group to Dicondylia. Silverfish are monophyletic to winged insects. Dragonflies are a sister group to a monophyletic clade of Ephemeroptera + Neoptera, corresponding to the *Chiastomyaria* hypothesis. Neopterous insects split into two monophyletic clades corresponding to the traditional concepts of Hemimetabola and Holometabola. The tree showed that several taxa displayed long branches. Several long-branch taxa showed strong deviations from average substitution patterns or were hard to align in several sections; this is true in particular for strepsipteran and dipteran taxa. Overall, the tree resembles the reconstruction of Kjer (2004) in many respects.

Node support values were deduced from a total of 5000 sampled trees (Table 4). The consensus tree shows that several of the presented relationships appear robust. The monophyly of each entognathous order is maximally supported. The relationship between the three orders appears resolved. Without a non-hexapod outgroup the placement of the root is arbitrary; however, the analysis can support monophyletic Entognatha. The tree cannot be rooted with proturans and springtails as monophyletic (Ellipura). A clade Insecta *sensu stricto* received high support; the relationship of bristletails and silverfishes is resolved and strongly supported (Table 4).

Table 3. Two selected helices and their respective substitution patterns

X	Y	M(x, y)	A|A	A|C	A|G	A|U	A|–	C|A	C|C	C|G	C|U	C|–	G|A	G|C	G|G	G|U	G|–	U|A	U|C	U|G	U|U	U|–	–|A	–|C	–|G	–|U	–|–
Helix 9 M(x, y)																												
X positions: 230 232 234–236																												
Y positions: 247 246 244–242																												
230	247	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
232	246	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
234	244	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
235	243	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
236	242	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Helix 23/e4 M(x, y)																												
Y positions: 2478–2475 2472–2470																												
X positions: 1460–1466																												
1460	2478	0.08	0.01	0.00	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.89	0.02	0.00	0.01	0.01	0.00	0.01	0.00	0.00
1461	2477	0.47	0.00	0.00	0.00	0.18	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.76	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
1462	2476	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.23	0.00	0.06	0.00	0.00	0.01	0.00	0.01	0.00	0.00
1463	2475	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00
1464	2472	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.92	0.02	0.02	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1465	2471	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.02	0.92	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
1466	2470	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.02	0.90	0.00	0.01	0.00	0.01	0.00	0.01

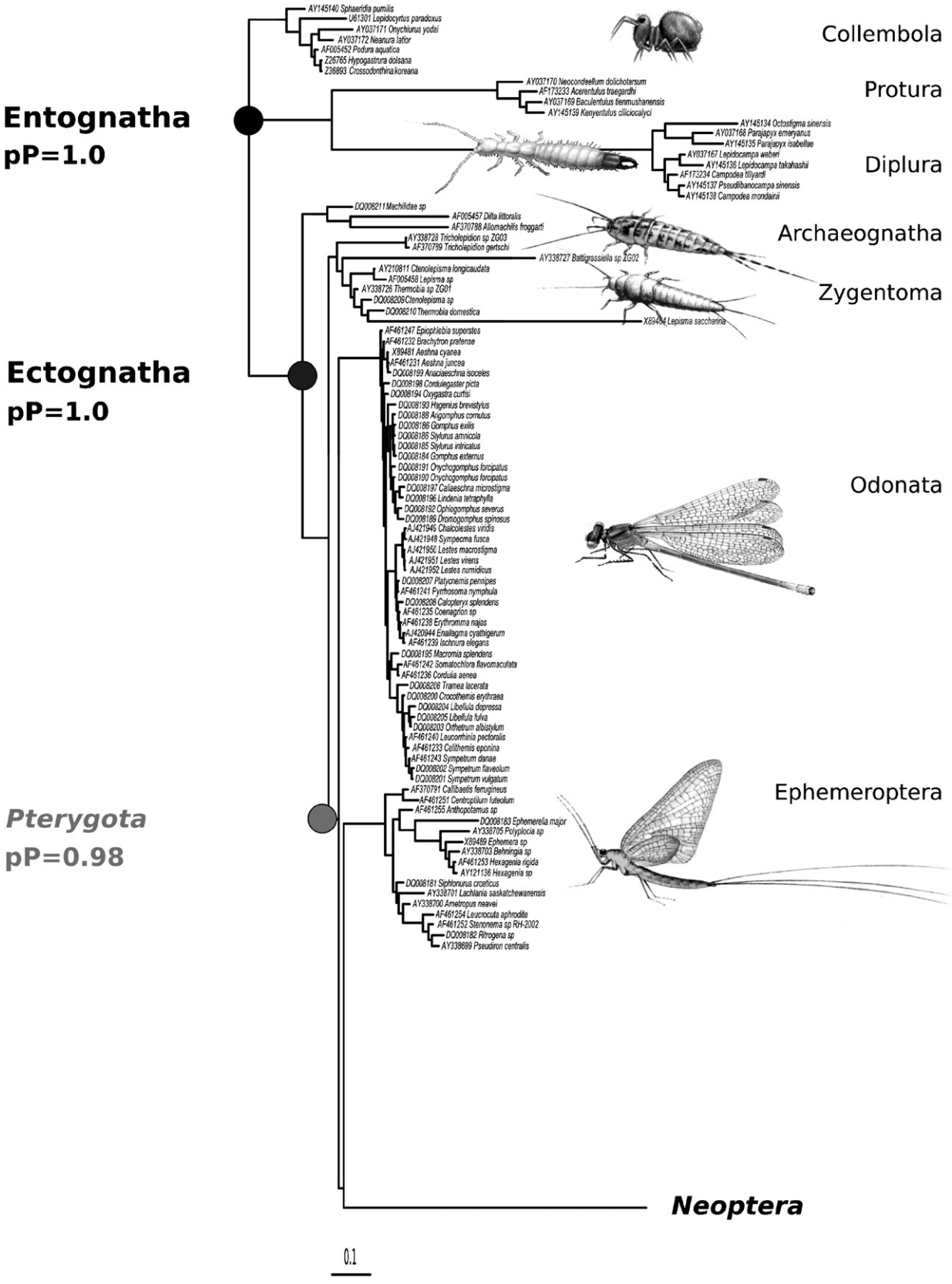

Fig. 2. Bayesian majority rule consensus tree. (a) Basal part of the tree, (b) Hemimetabola, (c) Holometabola. Pictures are modified after Chinery (1993).

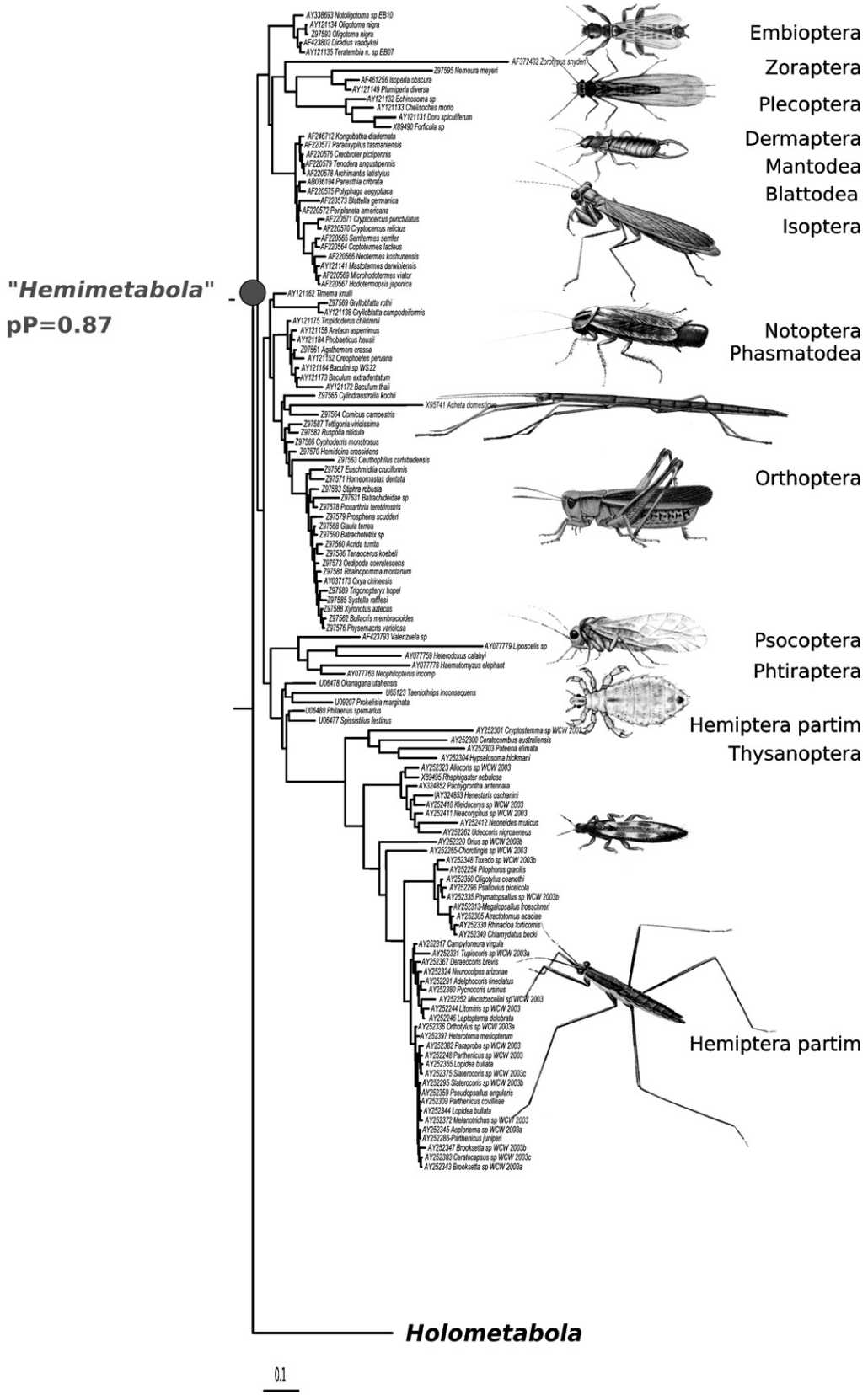

Fig. 2. (*Continued*)

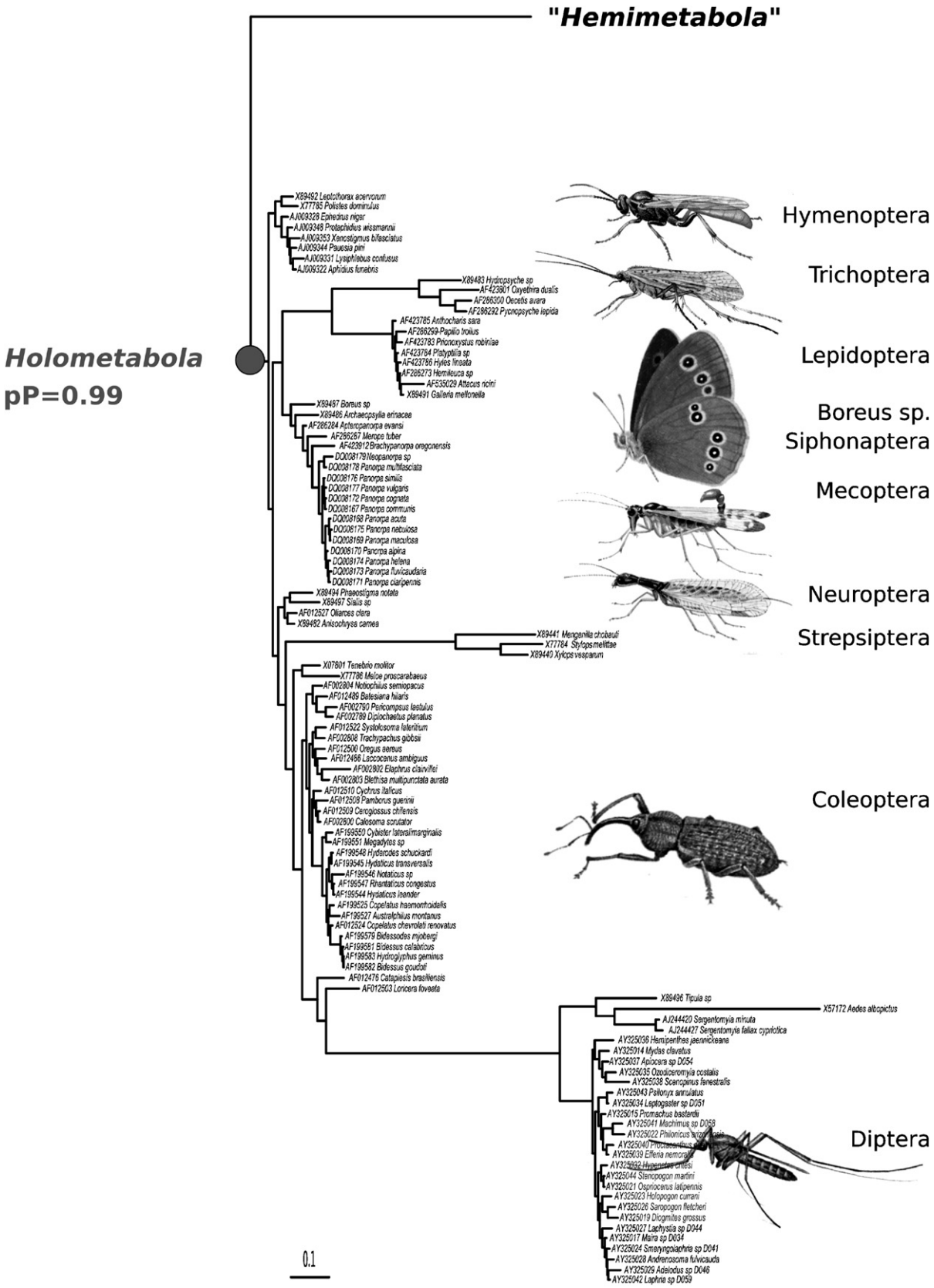

Fig. 2. (*Continued*)

Table 4. Bayesian support values for selected clades

Clade probability (*pP*)	Posterior
Ectognatha	1.00
Zygentoma	0.95
Dicondylia	0.97
Chiastomyaria	0.43
Hemimetabola	0.87
(Embioptera((Zoraptera(Plecoptera,Dermaptera)),(Mantodea(Blattodea,Isoptera)))	0.72
(Zoraptera(Plecoptera,Dermaptera))	0.45
(Plecoptera,Dermaptera)	0.55
(Mantodea(Blattodea,Isoptera))	1.00
(Blattodea,Isoptera)	0.61
(Phasmatodea,Notoptera,Orthoptera)	0.74
((Psocoptera,Phtiraptera),(Hemiptera,Thysanoptera))	0.95
(Psocoptera,Phtiraptera)	0.96
(Hemiptera,Thysanoptera)	0.57
Holometabola	0.99
(((Trichoptera,Lepidoptera),(Boreus(Siphonaptera,Mecoptera))),(Neuropterida(Strepsiptera(Coleoptera,Diptera))))	0.57
((Trichoptera,Lepidoptera),(Boreus(Siphonaptera,Mecoptera)))	0.82
(Trichoptera,Lepidoptera)	0.97
(Boreus(Siphonaptera,Mecoptera))	1.00
(Neuropterida(Strepsiptera(Coleoptera,Diptera)))	0.72
(Strepsiptera(Coleoptera,Diptera))	0.63
(Coleoptera,Diptera)	0.64

Within winged insects the *Chiastomyaria* clade is only poorly supported, but nevertheless appears as the favored arrangement among the sampled topologies. Within the neopterous insects, Holometabola were strongly supported, much less a clade Hemimetabola.

Within the Holometabola, Hymenoptera split off first, followed by a clade ((Trichoptera, Lepidoptera), (Siphonaptera, Mecoptera)) which is sister group to (Neuropterida (Strepsiptera (Coleoptera, Diptera))).

Hemimetabola split into two sister groups, a clade (Embioptera ((Zoraptera (Plecoptera, Dermaptera)), (Mantodea (Blattodea, Isoptera))) and a clade ((*Timema*, Notoptera), (Phasmatodea, Orthoptera))+ Paraneoptera (Fig. 2, Table 4).

Discussion

Applying consensus structure models

Our structure model is strictly applicable only to hexapod sequences. Sequences of crustaceans or myriapods will display several deviations from this model that makes them problematic as outgroups in such structured RNA-based analyses. Some differences between the structure model of Kjer and ours are probably best explained by the fact that Kjer included chelicerates, myriapods and crustaceans. His model had to cover base pairing in these taxa as well.

Consensus models are limited when dealing with deep phylogenetic problems. Even within hexapods we found structural deviations not covered by the hexapod consensus structure model. Among hexapods, structure variation is mostly limited to expansion segments, comprising helices 23/e#, 43 and 49. All three helices are notoriously hard to align. These expansion segments cannot be included in any consensus model and as has been shown by Hancock and Vogler (2000) are prone to frequent homoplasious base pairing. We will always have to trade off between conserved substructures and structure variation reducing the applicability of RNA substitution models. There is currently no reconstruction algorithm available that could cope with structure variation. If we want to fully exploit the phylogenetic signal of paired changes we would need such an approach. Typically, we see deviations from the consensus hexapod structure model in deep splits. In particular, proturans and diplurans deviate in several otherwise more or less conserved helices. Assuming that both orders belong to the hexapod clade, we see an increased structure variation of the SSU rRNA molecule at divergences of at least 300 MYAs. Comparisons with other arthropod sequences will show whether this is a general pattern of the SSU rRNA molecule.

Substitution patterns for unpaired and paired positions (Table 3) demonstrate that changes between canonical base pairings are by far the most frequent ones. Ignoring these biased substitution patterns will

lead to biased results in phylogenetic reconstructions or at least will inflate measures of support based on paired positions (Galtier, 2004). RNA substitution models can cope with this problem. These substitution models have been in the literature for some time, but they have found little application in structured RNA-based phylogenies (for example Jow et al., 2002; Hudelot et al., 2003). Two important aspects might account for this phenomenon. First, taxon-specific consensus structures have not been available. The reliable reconstruction of consensus structures relies on a dense taxon sampling of more or less complete sequences that is still not available for many groups. Secondly, the large numbers of free parameters in RNA substitution models are difficult to fit to natural data.

Structured RNA molecules and particularly rRNA sequences tend to be difficult to align. These difficulties have caused discussions about the correct alignment strategy in rRNA sequences with partially dramatic effects on phylogenetic reconstructions (Kjer, 2004; Simmons, 2004). Several authors favor a structurally aided alignment procedure (for example Lydeard et al., 2000; Hickson et al., 2000; Misof and Fleck, 2003; Misof et al., 2001; Kjer, 1995, 1997, 2004) emphasizing the possibility of maximizing homology over identity. This approach was criticized by demonstrating that slippage-derived sequence evolution can cause frequent homoplasy of stem-loop regions thus invalidating the structurally aided procedure (for example, see Hancock and Vogler, 2000). The analysis of slippage-derived sequence evolution further implies that evolutionary robust stem-loop structures could be a by-product of the process itself instead of being maintained by stabilizing selection. If true, this puts into question the approach of structurally aided alignment procedures that assume structure maintenance through effects of stabilizing selection. However, there is ample empirical evidence that RNA structures are maintained by the effects of stabilizing selection and correct folding of molecules is important for proper functioning (for a review see Higgs, 2000).

New phylogenies

In this investigation we focused on basal and inter-ordinal relationships of hexapods. In the following, we will briefly discuss our results in comparison to previously published molecular analyses. We are aware that our analysis has been restricted to 18S sequences and a Bayesian approach for which inflated support values can be expected (Holder and Lewis, 2003; Felsenstein, 2004). But we think it is worthwhile to discuss general patterns that emerged from the analysis.

The Entognatha

Resolving the relationships of the basal apterygote hexapods is a key problem in understanding the evolution of hexapods. Representatives of all three orders are predominantly small soil arthropods that are superficially characterized by reduced eyes, limbs, tracheal systems or Malpighian tubules (Bitsch, 1994; Kristensen, 1998; Bitsch and Bitsch, 2000).

Traditionally, two alternative concepts are favored, firstly the monophyly of Entognatha and secondly a paraphyly of Ellipura and Diplura in relation to true insects, Ectognatha (Kristensen, 1998; Bitsch and Bitsch, 2000). Important arguments in favor of paraphyletic Entognatha are characters diplurans share with true insects. Among others, these characters comprise paired claws, sperm ultrastructure, the loss of the Tömösváry organ in diplurans and true insects or the loss of the pseudotentorium (Kristensen, 1998). Additionally, Koch (1997) asserts that the entognathous condition in diplurans is different from the situation in proturans and collembolans, making a common origin unlikely.

Molecular evidence supporting paraphyletic Entognatha is equivocal (Carapelli et al., 2000; Giribet and Ribera, 2000; Giribet et al., 2001; D'Haese, 2002; Luan et al., 2003; Kjer, 2004; Giribet et al., 2004). Regier et al. (2004) found Entognatha monophyletic, based on gene fragments of Pol II, Ef 1-a and EF 2, Giribet et al. (2001) and Wheeler et al. (2001) obtained polyphyletic Entognatha based on nuclear and mitochondrial genes, Nardi et al. (2003) reconstructed polyphyletic hexapods with springtails as members of crustaceans. The analysis of Luan et al. (2003) suggested paraphyletic Entognatha. Kjer (2004) used representatives of chelicerates, myriapods and crustaceans as outgroups in his SSU DNA/RNA-based analysis of hexapod relationships. His analysis supported monophyly of Hexapoda and monophyly of Entognatha. Proturans and diplurans were sister groups. The sister group relationship of proturans and diplurans has been confirmed by Giribet et al. (2004), Mallatt and Giribet (2006) and Luan et al. (2005) based on multiple gene markers. However, Giribet et al. (2004) were unable to support monophyly of the entire hexapod clade but instead suggested polyphyly of the group. Our analysis was not designed to address monophyly of hexapods, since we used Entognatha as outgroup taxa. However, the reconstructed tree cannot be rooted with monophyletic Ellipura. This result is congruent with other published analyses based on different markers (Luan et al., 2003, 2005; Kjer, 2004; Giribet et al., 2004; Mallatt and Giribet, 2006).

There are no synapomorphies which could support a clade Protura + Diplura (see Koch, 1997, 2000, 2001; Klass and Kristensen, 2001; Willmann, 2002, 2003). Instead, the monophyly of Ellipura based on details of

the entognathous situation and the possible homologies of abdominal appendages appears more plausible. The contradicting well-supported results of molecular and morphological studies challenge both levels of analysis and demand a closer look at both data.

The Ectognatha

Archaeognatha, Zygentoma and Pterygota most likely form a monophyletic group, the Insecta *sensu stricto* (or Ectognatha) with Zygentoma forming a sister group to winged insects (Hennig, 1969; Kristensen, 1981).

Several authors consider Zygentoma paraphyletic in relation to winged insects, with the Lepidothrichidae as sister taxon to other silverfish and winged insects (see Kristensen, 1998). Molecular studies based on different genetic markers never unequivocally supported the monophyly of Zygentoma and true insects. In particular, the relationship of bristletails, silverfish and winged insects, the clade Ectognatha, was not reconstructed with notable support (e.g., Wheeler, 1998; Kjer, 2004).

Giribet et al. (2004) presented molecular data comprising sequences of five loci which cannot deliver robust reconstructions of Ectognatha relationships. In their analysis, bristletails and silverfish are a monophyletic group related to a clade Odonata + Neoptera. Only a combined analysis of morphological and molecular data delivers reconstructions of a more conventional view. Kjer (2004) recovered Insecta *sensu stricto*, but the position of *Tricholepidion gertschi* (Lepidothrichidae) was not resolved. The analysis placed *Tricholepidion* as sister group to dragonflies, with low support. Again, the monophyly of winged insects was not recovered. The relationship of bristletails and silverfish was congruent with conventional views based on morphological analyses.

Our analysis with a refined and extended alignment of 18S sequences recovered monophyletic Insecta *sensu stricto* and monophyletic Zygentoma (Fig. 2).

Ephemeroptera and Odonata

Relationships of mayflies (Ephemeroptera), dragonflies (Odonata) and Neoptera are intensively discussed (e.g., Ogden and Whiting, 2003; Kjer, 2004). Molecular analyses could not yet answer all questions left open by morphological studies (Staniczek, 2000; Fürst von Lieven, 2000). Today, the "Metapterygota" concept is supported by the most compelling morphological evidence (but compare arguments for a "Paleoptera" hypothesis in Kukalova-Peck and Lawrence, 2004). This concept assumes that Odonata and Neoptera together form a monophyletic clade. The morphology of the orthopteroid mandible, its articulation and musculature in Odonata and Neoptera supports this view. However, direct sperm transfer in Ephemeroptera and Neoptera could support the alternative concept coined "Chiastomyaria concept" (Boudreaux, 1979) in which Ephemeroptera and Neoptera form a monophyletic clade. Molecular analyses based on identical genetic markers can support either of the two alternative concepts, depending on reconstruction methods (Hovmöller et al., 2002; Ogden and Whiting, 2003; Kjer, 2004). Particularly interesting in this context is the strong support for the Chiastomyaria clade in Kjer's analysis (2004). This result is perplexing since 18S sequences do not seem to harbor much signal at this level (Table 4). The absence of a signal at this level of divergence is discouraging.

Neopterous insects

Only few molecular investigations have studied the inter-ordinal relationships of winged insects (e.g., Chalwatzis et al., 1996; Whiting et al., 1997; Hwang et al., 1998; Wheeler et al., 2001). Molecular investigations of Chalwatzis et al. (1996), Whiting et al. (1997) and Wheeler et al. (2001) are hallmarks in this field. But Chalwatzis et al. (1996) reported some unexpected relationships of insects solely based on 18S data. The analysis of Whiting et al. (1997) is plagued by obvious sequence errors (see discussion in Kjer, 2004) and besides this problem does not establish relationships among holometabolous insects which could withstand the evidence of morphological analyses (Wheeler et al., 2001). Similarly, the analysis of Wheeler (1998) yields no robust relationship based on molecular characters. It appeared that insect inter-ordinal relationships cannot be addressed with molecular characters (compare the discussion in Caterino et al., 2000). The analysis of Kjer (2004) presents a reconstruction of neopterous relationships which is congruent with many views based on morphological analyses (compare Kristensen, 1981; Willmann, 2003). Congruent with our extended and refined analysis of the 18S data, several clades appear robust in the reconstructions (Fig. 2, Table 4).

Plecoptera are consistently recovered as sister taxon to Dermaptera, and this clade is embedded in a clade of dictyopteran insects + Embioptera + Zoraptera. However, the clade Embioptera + Zoraptera + Plecopera + Dermaptera + Dictyoptera is not robust ($pP = 0.72$). Plecoptera are probably not a sister taxon to all other neopterous insects.

The clade Holometabola receives significant support ($pP = 0.99$) as it does in Kjer's study. Morphological studies traditionally assume the monophyly of holometabolous insects based on the holometaboly itself (e.g., Kristensen, 1998; Willmann, 2003). However, studies using different molecular markers had been unable to confirm or clearly contradict this hypothesis (Caterino et al., 2000). Our results based on structurally aided analyses appear reassuring.

In our study, most inter-ordinal relationships within the Holometabola are well resolved but they are in unexpected contrast to other results (see Fig. 2). Here, Hymenoptera are reported as sister taxon to all other Holometabola, but this relationship is not strongly supported. Morphological characters are completely arbitrary about the position of the Hymenoptera, however, a sister group relationship of Hymenoptera + Mecopterida is favored by many authors (Beutel and Gorb, 2001, 2006; Kristensen, 1981; Willmann, 2003).

Two additional results are certainly worth being mentioned: first, the monophyly of Diptera + Siphonaptera + Mecoptera (= Antliophora) was not recovered, instead, Siphonaptera + Mecoptera group with Amphiesmenoptera (= Lepidoptera + Trichoptera; Fig. 2). The monophyly of Antliophora is generally considered to be supported by the presence of a sperm pump in all three orders (see for a review Kristensen, 1981; Willmann, 2003). However, Wood and Borkent (1989) point at the problem of an unclear homology of sperm pumps in these orders and additionally at the fact that in many taxa this sperm pump is totally absent. Hünefeld and Beutel (2005) showed that the sperm pump is most likely not a synapomorphy of Antliophora. Additionally, it must be mentioned that taxa like the Nannochoristidae have not been included in the molecular analyses yet. With the inclusion of this and other potential early branching taxa the picture might well change.

Likewise, the present analysis supports a clade of Neuropterida + Coleoptera + Strepsiptera + Diptera, which is in contrast to morphological character analyses. Currently, we are not aware of a morphological character set supporting such a clade. Recent molecular evidence also does not support this clade (Bonneton et al., 2006). Extensive new morphological and molecular data will be necessary to clarify these contrasting hypotheses. For molecular studies, it will be particularly important to include sequences of key taxa, for example from Mecoptera, Neuropterida.

Conclusions

With this study we intended to pursue model-based phylogenetic reconstructions to generate the most realistic pictures of the past. Our approach can only work if the models are realistic. Simulations showed that ML reconstructions are robust against violations of model assumptions (Felsenstein, 2004), but simulations and theoretical considerations also showed that character dependence realized in rRNA-based data can produce misleading results (Jow et al., 2002; Hudelot et al., 2003; Galtier, 2004). In this case, modeling the character dependence is the solution and can be realized via RNA doublet models. We showed that the implementation of more realistic RNA doublet models is readily accommodated, in line with a recently published analysis by Kjer (2004).

We have been engaged in deriving a hexapod SSU rRNA consensus model based on explicit criteria (Misof et al., 2006) and have used this structure model to fit a doublet substitution model within a Bayesian approach. This protocol of modeling covariation patterns observed in RNA molecules increases the biological relevance of applied evolutionary models and will recover more reliably phylogenetic signal, if present.

Concerning the evolution of hexapods, we inferred relationships that were in many aspects congruent with morphological analyses. This makes us confident that model-based analyses capturing biological realism can indeed provide a grip on solving the puzzle of hexapod diversification.

Acknowledgments

This publication is dedicated to Prof. K.P. Sauer on the occasion of his 66th birthday. As the corresponding author (B.M.), I wish to express my admiration and respect to K.P. Sauer, and my gratefulness for four years of working as a Post Doc in his Institute of Evolutionary Biology from which I profited greatly.

We would like to acknowledge valuable input from discussions within the ZFMK lab, from F. Frati, University of Siena, from the G. Pass group, University of Vienna and two anonymous reviewers. This work was financed by the DFG grant MI649/4-1 to B.M.

References

Beutel, R., Gorb, S., 2001. Ultrastructure of attachment specializations of hexapods (Arthropoda): evolutionary patterns inferred from a revised ordinal phylogeny. J. Zool. Syst. Evol. Res. 39, 1–31.

Beutel, R.G., Gorb, S.N., 2006. A revised interpretation of the evolution of attachment structures in Hexapoda with special emphasis on Mantophasmatodea. Arthropod Syst. Phyl. 64, 3–25.

Bitsch, J., 1994. The morphological groundplan of hexapoda: critical review of recent concepts. Ann. Soc. Entomol. Fr. (N.S.) 30, 103–129.

Bitsch, C., Bitsch, J., 2000. The phylogenetic interrelationships of the higher taxa of apterygote hexapods. Zool. Scripta 29, 131–156.

Bonneton, F., Brunet, F.G., Kathirithamby, J., Laudet, V., 2006. The rapid divergence of the ecdysone receptor is a synapomorphy for Mecopterida that clarifies the Strepsiptera problem. Insect Mol. Biol. 15, 351–362.

Boudreaux, H., 1979. Arthropod Phylogeny. Wiley, New York, Chichester, Brisbane, Toronto.

Carapelli, A., Frati, F., Nardi, F., Dallai, R., Simon, C., 2000. Molecular phylogeny of the apterygotan insects based on nuclear and mitochondrial genes. Pedobiologia 44, 361–373.

Carmean, D., Kimsey, L.S., Berbee, M.L., 1992. 18S rDNA sequences and holometabolous insects. Mol. Phylogenet. Evol. 1, 270–278.

Caterino, M.S., Cho, S., Sperling, F.A.H., 2000. The current state of insect molecular systematics: a thriving tower of babel. Ann. Rev. Entomol. 45, 1–54.

Chalwatzis, N., Hauf, J., van de Peer, Y., Kinzelbach, R., Zimmermann, F.K., 1996. 18S ribosomal RNA genes of insects: primary structure of the genes and molecular phylogeny of Holometabola. Ann. Entomol. Soc. Amer. 89, 775–787.

Chinery, M., 1993. Pareys Buch der Insekten. Ein Feldführer der europäischen Insekten. Übers. u. bearb. von I. Jung und D. Jung. Verlag Paul Parey - Hamburg und Berlin.

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

D'Haese, C.A., 2002. Phylogeny of the apterygote hexapods. Abstracts, 20th Annual Meeting of the Willi Hennig Society. Cladistics 18, 220.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Inc., Sunderland, Massachusetts.

Fitch, W.M., Markowitz, E., 1970. An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. Biochem. Genet. 4, 579–593.

Fürst von Lieven, A., 2000. The transformation from monocondylous to dicondylous mandibles in the insecta. Zool. Anz. 239, 139–146.

Galtier, N., 2004. Sampling properties of the bootstrap support in molecular phylogeny: influence of nonindependence among sites. Syst. Biol. 53, 38–46.

Giribet, G., Ribera, C., 2000. A review of arthropod phylogeny: new data based on ribosomal DNA sequences and direct character optimization. Cladistics 16, 204–231.

Giribet, G., Edgecombe, G.D., Wheeler, W.C., 2001. Arthropod phylogeny based on eight molecular loci and morphology. Nature 413, 157–161.

Giribet, G., Edgecombe, G.D., Carpenter, J.M., D'Haese, C.A., Wheeler, W.C., 2004. Is Ellipura monophyletic? A combined analysis of basal hexapod relationships with emphasis on the origin of insects. Org. Div. Evol. 4, 319–340.

Gladstein, D.S., Wheeler W.C., 1997–2002. POY: the optimization of alignment characters. Program and documentation. New York. Available at 〈ftp.amnh.org/pub/molecular〉.

Hall, T.A. (Ed.), 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41, 95–98.

Hancock, J.M., Vogler, A.P., 2000. How slippage-derived sequences are incorporated into rRNA variable-region secondary structure: implications for phylogeny reconstruction. Mol. Phylogen. Evol. 14, 366–374.

Hennig, W., 1969. Die Stammesgeschichte der Insekten. Verlag Waldemar, Kramer, Frankfurt am Main.

Hickson, R.E., Simon, C., Perrey, S.W., 2000. The performance of several multiple-sequence alignment programs in relation to secondary-structure features for a rRNA sequence. Mol. Biol. Evol. 17, 530–539.

Higgs, P.G., 1998. Compensatory neutral mutation and the evolution of RNA. Genetica 102, 91–101.

Higgs, P.G., 2000. RNA secondary structure: physical and computational aspects. Quart. Rev. Biophys. 33, 199–253.

Holder, M., Lewis, P.O., 2003. Phylogeny estimation: traditional and Bayesian approaches. Nat. Rev. Genet. 4, 275–284.

Hovmöller, R., Pape, T., Källersjö, M., 2002. The Palaeoptera problem: basal pterygote phylogeny inferred from 18S and 28S rDNA sequences. Cladistics 18, 313–323.

Hudelot, C., Gowri-Shankar, V., Jow, H., Rattray, M., Higgs, P.G., 2003. RNA-based phylogenetic methods: application to mammalian mitochondrial RNA sequences. Mol. Phylogen. Evol. 28, 241–252.

Huelsenbeck, J.P., 1998. Systematic bias in phylogenetic analysis: is the Strepsiptera problem solved? Syst. Biol. 47, 519–537.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Huelsenbeck, J.P., Ronquist, F., 2002. MrBayes 3: Bayesian analysis of phylogeny. Computer program distributed by the authors. Department of Ecology, Behavior and Evolution, University of California.

Hünefeld, F., Beutel, R.G., 2005. The sperm pumps of Strepsiptera and Antliophora (Hexapoda). J. Zool. Syst. Evol. Res. 43, 297–306.

Hwang, U.W., Kim, W., Tautz, D., Friedrich, M., 1998. Molecular phylogenetics at the Felsenstein zone: approaching the Strepsiptera problem using 5.8S and 28SrDNAsequences. Mol. Phylogenet. Evol. 9, 470–480.

Jow, H., Hudelot, C., Rattray, M., Higgs, P.G., 2002. Bayesian phylogenetics using an RNA substitution model applied to early mammalian evolution. Mol. Biol. Evol. 19, 1591–1601.

Kjer, K.M., 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from frogs. Mol. Phyogenet. Evol. 4, 314–330.

Kjer, K.M., 1997. An alignment template for amphibian 12S rRNA, domain III: conserved primary and secondary structural motifs. J. Herpetol. 31, 599–604.

Kjer, K.M., 2004. Aligned 18S and insect phylogeny. Syst. Biol. 53, 1–9.

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

Koch, M., 1997. Monophyly and the phylogenetic position of the Diplura (Hexapoda). Pedobiologia 41, 9–12.

Koch, M., 2000. The cuticular cephalic endoskeleton of primarily wingless hexapods: ancestral state and evolutionary changes. Pedobiologia 44, 374–385.

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

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

Kristensen, N.P., 1998. The groundplan and basal diversification of hexapods. In: Fortey, R.A., Thomas, R.H. (Eds.), Arthropod Relationships. Systematics Association Special Volume Series 55. Chapman & Hall, London, pp. 281–293.

Kukalova-Peck, J., Lawrence, J.F., 2004. Relationships among coleopteran suborders and major endoneopteran lineages: evidence from hind wing characters. Eur. J. Entomol. 101, 95–144.

Liu, H., Beckenbach, A.T., 1992. Evolution of the mitochondrial oxydase II gene among 10 orders of insects. Mol. Phylogenet. Evol. 1, 41–52.

Luan, Y.X., Zhang, Y.P., Yue, Q.Y., Pang, J.F., Xie, R.D., Yin, W.Y., 2003. Ribosomal DNA gene and phylogenetic relationships of Diplura and lower hexapods. Sci. Chin. Ser. C 46, 67–76.

Luan, Y., Mallatt, J.M., Xie, R., Yang, Y., Yin, W., 2005. The phylogenetic positions of three basal-hexapod groups (Protura, Diplura and Collembola) based on ribosomal RNA sequences. Mol. Biol. Evol. 22, 1579–1592.

Lutzoni, F., Wagner, P., Reeb, V., Zoller, S., 2000. Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. 49, 628–651.

Lydeard, C., Holznagel, W.E., Schnare, M.N., Gutell, R.R., 2000. Phylogenetic analysis of molluscan mitochondrial LSU rDNA sequences and secondary structures. Mol. Phylogen. Evol. 15, 83–102.

Mallatt, J., Giribet, G., 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. Mol. Phylogen. Evol. 40, 772–794.

Misof, B., Fleck, G., 2003. Comparative analysis of mt 16S rRNA secondary structures of Odonates and its relevance to phylogenetic problems in insect systematics. Insect Mol. Biol. 12, 535–547.

Misof, B., Rickert, A.M., Buckley, T.R., Fleck, G., Sauer, K.P., 2001. Phylogenetic signal and its decay in mitochondrial SSU and LSU rRNA gene fragments of Anisoptera. Mol. Biol. Evol. 18, 27–37.

Misof, B., Niehuis, O., Bischoff, I., Rickert, A., Erpenbeck, D., Staniczek, A., 2006. A Hexapod nuclear SSU rRNA secondary structure model and catalogue of taxon specific structural variation. J. Exp. Zool. B (Mol. Dev. Evol.) 306B, 70–88.

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

Niehuis, O., Yen, S.-H., Naumann, C.M., Misof, B., 2006. Higher phylogeny of zygaenid moths (Insecta: Lepidoptera) inferred from nuclear and mitochondrial sequence data and the evolution of cuticular cavities for chemical defence. Mol. Phylogen. Evol. 39, 812–829.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

Ogden, T.H., Whiting, M.F., 2003. The problem with "the Paleoptera problem": sense and sensitivity. Cladistics 19, 432–442.

Pashley, D.P., McPheron, B.A., Zimmer, E.A., 1993. Systematics of holometabolous insect orders based on 18S ribosomal RNA. Mol. Phylogenet. Evol. 2, 132–142.

Posada, D., Crandall, K.A., 1998. ModelTest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Regier, J.C., Shultz, J.W., Kambic, R.E., 2004. Phylogeny of basal hexapod lineages and estimates of divergence times. Ann. Entomol. Soc. Am. 97, 411–419.

Savill, N., Hoyle, D., Higgs, P.G., 2001. RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum likelihood methods. Genetics 157, 399–411.

Schöniger, M., von Haeseler, A., 1994. A stochastic model and the evolution of autocorrelated DNA sequences. Mol. Phylogenet. Evol. 3, 240–247.

Simmons, M.P., 2004. Independence of alignment and tree search. Mol. Phylogenet. Evol. 31, 874–879.

Staniczek, A.H., 2000. The mandible of silverfish (Insecta: Zygentoma) and mayflies (Ephemeroptera): its morphology and phylogenetic significance. Zool. Anz. 239, 147–178.

Stephan, W., 1996. The rate of compensatory evolution. Genetics 144, 419–426.

Swofford, D.L., 2001. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Tillier, E., Collins, D., 1998. High apparent rate of simultaneous compensatory base pair substitutions in ribosomal RNA. Genetics 148, 1993–2002.

Van de Peer, Y., De Rijk, P., Wuyts, J., Winkelmans, T., De Wachter, R., 2000. The European small subunit ribosomal RNA database. Nucleic Acids Res. 28, 175–176.

Wheeler, W.C., 1998. Sampling, groundplans, total evidence and the systematics of arthropods. In: Fortey, R.A., Thomas, R.H. (Eds.), Arthropod Relationships. Chapman and Hall, London, pp. 87–96.

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

Whiting, M.F., 2002a. Mecoptera is paraphyletic: multiple genes and phylogeny of Mecoptera and Siphonaptera. Zool. Scripta 31, 93–104.

Whiting, M.F., 2002b. Phylogeny of the holometabolous insect orders based on 18S ribosomal DNA: when bad things happen to good data. In: DeSalle, R., Giribet, G., Wheeler, W.C. (Eds.), Molecular Systematics and Evolution: Theory and Practice. Birkhäuser Verlag, Basel, pp. 69–83.

Whiting, M.F., Carpenter, J.C., 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.

Willmann, R., 2002. Phylogenese und System der Insecta. In: Kaestner, A., Dathe, H. (Eds.), (Hrsg.): Lehrbuch der

Speziellen Zoologie. Band 1/5, Insecta. G. Fischer, Stuttgart, Jena, pp. 1–65.

Willmann, R., 2003. Die phylogenetischen Beziehungen der Insecta: offene fragen und probleme. Verh. Westd. Entom. Tag. 2001, 1–64.

Wood, D., Borkent, A., 1989. Phylogeny and classification of the Nematocera. In: McAlpine, J.F., (Ed.), Manual of Nearctic Diptera, Vol. 3. Research Branch, Agriculture Canada. Monograph 32, pp. 1333–1370.

Wuyts, J., De Rijk, P., Van de Peer, Y., Pison, G., Rousseeuw, P., De Wachter, R., 2000. Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA. Nucleic Acids Res. 28, 4698–4708.