Syst. Biol. 58(4):381–394, 2009
Copyright © Society of Systematic Biologists
DOI:10.1093/sysbio/syp037
Advance Access publication on July 15, 2009

Nonstationary Evolution and Compositional Heterogeneity in Beetle Mitochondrial Phylogenomics

NATHAN C. SHEFFIELD[1,2,*], HOJUN SONG[1], STEPHEN L. CAMERON[3], AND MICHAEL F. WHITING[1]

[1]*Department of Biology, Brigham Young University, Provo, UT 84602, USA;*
[2]*Program in Computational Biology & Bioinformatics, Institute for Genome Sciences and Policy, Duke University, Box 90090, Durham, NC 27708, USA;*
[3]*Australian National Insect Collection, Commonwealth Scientific and Industrial Research Organisation, Entomology, PO Box 1700, Canberra, Australian Capital Territory, 2601, Australia;*
[*]*Correspondence to be sent to: Program in Computational Biology & Bioinformatics, Institute for Genome Sciences and Policy, Duke University, Box 90090, Durham, NC 27708, USA; E-mail: nathan.sheffield@duke.edu.*

Nathan C. Sheffield and Hojun Song have contributed equally to this work.

Abstract.—Many published phylogenies are based on methods that assume equal nucleotide composition among taxa. Studies have shown, however, that this assumption is often not accurate, particularly in divergent lineages. Nonstationary sequence evolution, when taxa in different lineages evolve in different ways, can lead to unequal nucleotide composition. This can cause inference methods to fail and phylogenies to be inaccurate. Recent advancements in phylogenetic theory have proposed new models of nonstationary sequence evolution; these models often outperform equivalent stationary models. A variety of new phylogenetic software implementing such models has been developed, but the studies employing the new methodology are still few. We discovered convergence of nucleotide composition within mitochondrial genomes of the insect order Coleoptera (beetles). We found variation in base content both among species and among genes in the genome. To this data set, we have applied a broad range of phylogenetic methods, including some traditional stationary models of evolution and all the more recent nonstationary models. We compare 8 inference methods applied to the same data set. Although the more commonly used methods universally fail to recover established clades, we find that some of the newer software packages are more appropriate for data of this nature. The software packages p4, PHASE, and nhPhyML were able to overcome the systematic bias in our data set, but parsimony, MrBayes, NJ, LogDet, and PhyloBayes were not. [Base compositional heterogeneity; Coleoptera; LogDet; model of evolution; nonstationary evolution; nucleotide composition; phylogeny.]

Most traditional models of DNA evolution invoke the "stationarity assumption," which implies that base composition is constant over all lineages in the data set (Galtier and Gouy 1995). This is a valid assumption for some data sets; however, when the assumption is violated, phylogenetic methods can inaccurately group species whose base composition is similar, regardless of evolutionary history (Collins et al. 1994; Lockhart et al. 1994; Ho and Jermiin 2004; Jermiin et al. 2004; Cox et al. 2008). Although there has been argument as to the magnitude of the problem (van den Bussche et al. 1998; Conant and Lewis 2001; Gruber et al. 2007), its existence is widely accepted. The easiest solution is simply to avoid nonstationary genes (Collins et al. 2005), but this is not always possible. A distance-based transformation method known as LogDet/Paralinear distances (Lake 1994; Lockhart et al. 1994) can correct for compositional bias and is commonly used (e.g., Gibson et al. 2005; Gruber et al. 2007), but its effectiveness is not universal (Foster and Hickey 1999; Tarrío et al. 2001). Data recoding is also widely used in several forms, such as amino acid translations (Loomis and Smith 1990), RY coding (Brown et al. 1982; Phillips and Penny 2003), discarding/downweighting third codon positions (e.g., Chang and Campbell 2000; Delsuc et al. 2003), or Dayhoff coding (Hrdy et al. 2004). Such methods are usually able to overcome bias problems, but they also discard useful signal and can introduce artifactual relationships (Campbell et al. 2000; Tarrío et al. 2000; Cameron et al. 2006; Fenn et al. 2008).

Recently, several complex statistical approaches have been proposed and implemented in software such as p4 (Foster 2004), nhPhyML (Boussau and Gouy 2006), PHASE (Gowri-Shankar and Rattray 2007), and PhyloBayes (Blanquart and Lartillot 2006). nhPhyML is a maximum likelihood method based on the evolutionary model of Galtier and Gouy (1995), which models nonstationary evolution by specifying a different GC content for each branch in the tree. In p4, PHASE, and PhyloBayes, varying numbers of "composition vectors" model nonstationary evolution in a Bayesian framework. These software packages have all been shown to succeed in some cases; however, their efficacy has not been thoroughly tested using real data sets.

Although theoretical and simulation studies have identified and explored the problem relatively thoroughly (Galtier and Gouy 1995; Jermiin et al. 2004), there have only been a handful of studies applying available methods to real data sets. Of the few such studies, most are based on widely divergent taxa because convergent nucleotide composition is more common and noticeable in divergent lineages (e.g., Loomis and Smith 1990; Galtier and Gouy 1995; Herbeck et al. 2005). Less often, studies uncover the issue at lower levels such as within Lepidoptera (Campbell et al. 2000) or Mammalia (van den Bussche et al. 1998; Gibson et al. 2005) or even within the genus *Drosophila* (Rodríguez-Trelles et al. 1999; Tarrío et al. 2000). Most of these studies have focused on the effect of compositional bias in relatively few genes. How compositional bias can affect phylogenetic inference in complex heterogeneous data sets has not been explored thoroughly (but see Collins et al. 2005; Gibson et al. 2005; Rodriguez-Ezpeleta et al. 2007).

We investigated the effect of base compositional heterogeneity in phylogenetic reconstruction using mitochondrial genome (mtgenome) data of the insect order Coleoptera (beetles). We have previously described complete mtgenomes for 6 beetle species (Sheffield et al. 2008) and found that there is considerable variation in base composition among beetles. When we constructed a preliminary phylogeny of Coleoptera using both parsimony and Bayesian methods, we found some unexpected groupings. Because mtgenome data are known to be capable of resolving insect ordinal relationships accurately (see also Rubinoff and Holland 2005; Cameron et al. 2007; Fenn et al. 2008), we investigated the cause of the surprising preliminary results.

In particular, we noticed 3 unusual relationships in our initial analysis. First, the preliminary analyses placed *Tetraphalerus bruchi* within the suborder Polyphaga. *Tetraphalerus* is a member of a family Ommatidae, 1 of 3 extant families of the basal suborder Archostemata (Lawrence and Newton 1982; Beutel et al. 2008). Archostemata is recorded from the Mesozoic, with at least 13 known extinct species from the Jurassic and Cretaceous, and thus represents one of the oldest living lineages of Coleoptera (Crowson 1960; Ponomarenko 1969; Beutel et al. 2008). Archostemata contains mostly wood-boring beetles and its monophyly is well supported by several synapomorphies, including the reduction of the anterior tentorial arms, the reduction of the frontoclypeal suture, the distinctly reduced mentum, a median ridge on the first abdominal ventrite, and other adult and larval features (Beutel et al. 2008). Beutel and Haas (2000) found Archostemata to be sister to the remaining Coleoptera based on morphological data, and Hughes et al. (2006) found the same relationship based on a phylogenomic analysis using 66 ribosomal protein expressed sequence tags (ESTs). The most comprehensive molecular phylogeny of Coleoptera to date (Hunt et al. 2007) found that Archostemata was sister to Myxophaga. Therefore, it is likely that placement of *Tetraphalerus* within Polyphaga is erroneous. Second, our initial analyses found that the monophyletic Tenebrionidae did not group with other cucujiform taxa, thereby rendering the infraorder Cucujiformia paraphyletic. Cucujiformia is a diverse clade within Coleoptera that contains more than half of all beetles and includes weevils, darkling beetles, leaf beetles, and longhorn beetles. Its monophyly is strongly supported morphologically by cryptonephridic Malpighian tubules (Poll 1932; Stammer 1934), nonfunctional and reduced spiracles on eighth abdominal segment (Crowson 1960), and other characters (Wachmann 1977; Caveney 1986). Both the EST study (Hughes et al. 2006) and the molecular phylogeny (Hunt et al. 2007) unambiguously found Cucujiformia monophyletic and Tenebrionidae to belong to this group with high support values. Therefore, a paraphyletic Cucujiformia as found in our preliminary analyses is likely incorrect. Finally, our analyses failed to recover a monophyletic Elateroidea because the elaterid *Pyrophorus divergens* formed a strong group with other polyphagans belonging to different lineages. The monophyly of Elateroidea sensu Lawrence (1988) is supported by a single pair of well-developed stemmata in larvae and the presence of only 4 Malpighian tubules in adults and has also been consistently supported by molecular data (Bocakova et al. 2007; Hunt et al. 2007). Therefore, Elateridae falling outside of the remaining Elateroidea is probably inaccurate. Other relationships derived from the preliminary analyses were difficult to assess given that single representatives from diverse lineages were included. We suspected that the incorrect groupings with high support values might be because a number of divergent taxa share similar base composition.

Herein, we test our data set for compositional heterogeneity (among species and among genes) that may cause incorrect phylogenetic relationships. Using the 3 relationships described above as indicator relationships of correct topology, we examine several commonly used phylogenetic inference methods (distance, parsimony, and Bayesian) and explore a number of methods that are designed to overcome the systematic bias arising from nonstationarity. We focus our comparison on model-based methods because these methods have not been thoroughly compared using empirical data. We also present the result of amino acid recoding here for comparison. We demonstrate that there is compositional heterogeneity among species and that it is distributed in a complex manner among the 13 protein-coding genes of the mtgenome. We show that standard phylogenetic methods consistently fail to recover the correct topology, whereas some of the new methods that explicitly account for the bias can overcome the problem.

Materials and Methods

Taxon and Character Sampling

Eighteen taxa were analyzed in this study, including 13 coleopteran in-group and 5 out-group (3 lepidopteran and 2 dipteran) taxa (Table 1). We sequenced complete coding regions from 5 new coleopteran species for this study: *Chauliognathus opacus* (Cantharidae), *Adelium* sp. (Tenebrionidae), *Apatides fortis* (Bostrichidae), *Lucanus mazama* (Lucanidae), and *Acmaeodera* sp. (Buprestidae) (GenBank accession numbers FJ613418–FJ613422; Table 1). Descriptions of these new mtgenomes are presented in Appendix Table A1. The in-group sampling represents 2 suborders (Archostemata and Polyphaga), 4 infraorders, 8 superfamilies, and 12 families within Coleoptera.

We generated the complete mtgenomes used in this study using the primer-walking protocols of Cameron et al. (2007) modified for Coleoptera (primers available from H.S.). We annotated the new mtgenomes using MOSAS (http://mosas.byu.edu), following Sheffield et al. (2008). For the present study, we used all 13

TABLE 1. Taxa used in this study

Taxa included	Order	Suborder	Infraorder	Superfamily	Family	GenBank	Reference
Anopheles gambiae	Diptera	—	—	—	Culicidae	NC_002084	Beard et al. (1993)
Drosophila yakuba	Diptera	—	—	—	Drosophilidae	NC_001322	Clary and Wolstenhome (1985)
Ostrinia nubialis	Lepidoptera	—	—	—	Crambidae	NC_003367	Coates et al. (2005)
Bombyx mori	Lepidoptera	—	—	—	Bombycidae	NC_002355	(unpublished)
Antheraea pernyi	Lepidoptera	—	—	—	Saturniidae	NC_004622	(unpublished)
Tetraphalerus bruchi	Coleoptera	Archostemata	—	—	Ommatidae	NC_011328	Sheffield et al. (2008)
Crioceris duodecimpunctata	Coleoptera	Polyphaga	Cucujiformia	Chrysomeloidea	Chrysomelidae	NC_003372	Friedrich and Muqim (2003)
Priasilpha obscura	Coleoptera	Polyphaga	Cucujiformia	Cucujoidea	Phloeostichidae	NC_011326	Sheffield et al. (2008)
Chaetosoma scaritides	Coleoptera	Polyphaga	Cucujiformia	Cleroidea	Chaetosomatidae	NC_011324	Sheffield et al. (2008)
Tribolium castaneum	Coleoptera	Polyphaga	Cucujiformia	Tenebrionoidea	Tenebrionidae	NC_003081	Stewart and Bechenbach (2003)
Adelium sp.	Coleoptera	Polyphaga	Cucujiformia	Tenebrionoidea	Tenebrionidae	FJ613422	This study
Pyrocoelia rufa	Coleoptera	Polyphaga	Elateriformia	Elateroidea	Lampyridae	NC_003970	Bae et al. (2004)
Pyrophorus divergens	Coleoptera	Polyphaga	Elateriformia	Elateroidea	Elateridae	NC_009964	Arnoldi et al. (2007)
Rhagophthalmus lufengensis	Coleoptera	Polyphaga	Elateriformia	Elateroidea	Rhagophthalmidae	NC_010969	Li et al. (2007)
Chauliognathus opacus	Coleoptera	Polyphaga	Elateriformia	Elateroidea	Cantharidae	FJ613418	This study
Acmaeodera sp.	Coleoptera	Polyphaga	Elateriformia	Buprestoidea	Buprestidae	FJ613420	This study
Apatides fortis	Coleoptera	Polyphaga	Bostrichiformia	Bostrichoidea	Bostrichidae	FJ613421	This study
Lucanus mazama	Coleoptera	Polyphaga	Scarabaeiformia	Scarabaeoidea	Lucanidae	FJ613419	This study

protein-coding genes for phylogenetic analysis. We aligned nucleotide sequences for each gene individually with MUSCLE (Edgar 2004) and concatenated the alignments to form a single matrix consisting of 11 633 aligned nucleotide characters, which we used throughout the study. We also aligned amino acid translations of each gene separately and used the concatenated amino acid data matrix (3880 characters) to examine the effect of amino acid recoding. To test how nucleotide alignment might affect our results, we back translated the amino acid alignments into nucleotide sequences using ClustalW as implemented in MEGA 4.0 (Tamura et al. 2007). The resulting matrix after concatenation of individual alignments consisted of 11 655 nucleotide characters

Assessment of the Degree of Heterogeneity

We calculated base composition of each taxon for each of the 13 protein-coding genes as well as that of the final concatenated data set. Because AT% (or reciprocal GC%) represented the compositional bias the best (Fig. 1), we compared the calculated AT% for each gene among the beetle species included in this study. We mapped the total AT% for each species onto the resulting phylogeny to examine the distribution pattern of compositional bias.

To measure the variation in evolutionary pattern, we calculated the disparity index (I_D) (Kumar and Gadagkar 2001) for all 13 genes together and pairwise. We also tested the homogeneity of substitution pattern (I_D test) using a Monte–Carlo method with 1000 replicates as implemented in MEGA 4.0 (Tamura et al. 2007). We calculated the probability of rejecting the null hypothesis that sequences have evolved with the same pattern of substitution at $\alpha < 0.01$. All positions containing gaps and missing data were removed from the data set (complete deletion option).

Phylogenetic Analyses and Comparison of Methods

We analyzed the data using several different inference methods to compare the effect of base compositional heterogeneity. We measured topological accuracy by the presence of the 3 indicator relationships: 1) a sister relationship between monophyletic Archostemata and Polyphaga, 2) a monophyletic Cucujiformia (Tenebrionidae grouping with other cucujiform taxa), and 3) a monophyletic Elateroidea. Because these relationships are well established, deviations serve as an indicator of misleading phylogenetic signal. Due to limited taxon sampling in our analysis, we were less concerned about the internal relationships within these clades. Our aim in this study was not so much to explore the relationships within Coleoptera, which would require more extensive taxon sampling, but to determine the most robust methods against the systematic bias.

For our analysis, we divided the methods into 2 groups. The first group consisted of methods not designed to deal with compositional bias, and we call these

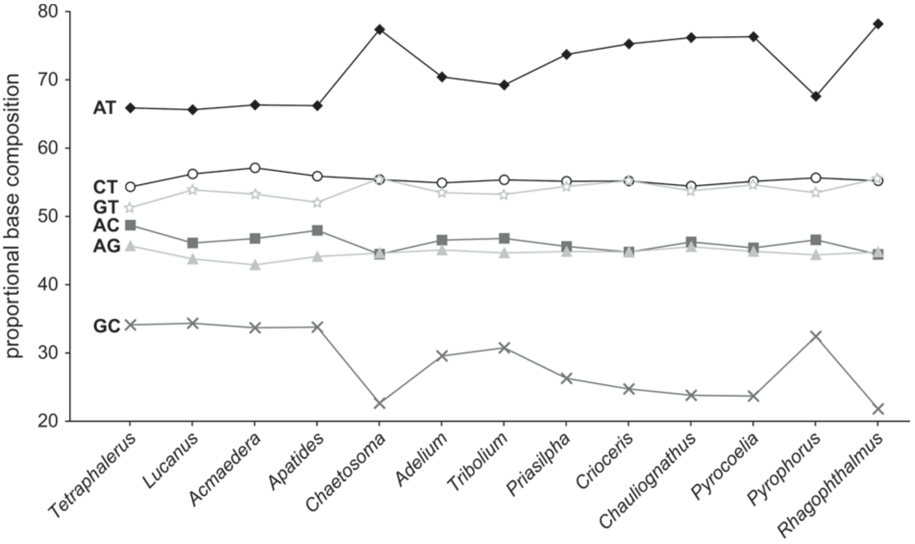

FIGURE 1. Base composition as calculated based on different pairs of bases across Coleoptera for our data set.

approaches "time homogeneous." We analyzed the data set in a conventional manner using distance, parsimony, and Bayesian methods. We performed neighbor joining (NJ) in PAUP* 4b10 (Swofford 2002) under several distance models (JC, HKY85, and GTR), both with and without gamma-distributed among-site rate heterogeneity, each with 1000 bootstrap replicates to assess nodal support. For the parsimony analysis, for each nucleotide alignment, we performed 1000 random addition heuristic search replicates for each of 100 bootstrap iterations in PAUP* 4b10 (gaps were treated as missing). We also ran the parsimony analysis on the amino acid alignment to determine the effect of data recoding. For the Bayesian analysis, we performed a partitioned model analysis with a separate GTR + G + I model for each gene (per recommendation by MrModelTest version 2; Nylander 2004) and ran 4 separate runs each with 4 chains for 10 million generations, sampling every 1000 trees in MrBayes 3.1.1 (Ronquist and Huelsenbeck 2003). We repeated this analysis after partitioning the data by codon position (instead of by gene) as well. We plotted the likelihood trace in R (R Development Core Team 2008) and examined sliding and cumulative split posterior probabilities using AWTY (Nylander et al. 2008) for each run to assess convergence and discarded the first 1 million generations as burn-in.

We also analyzed the data set using robust methods that have been designed to account for compositional bias, which we call "time heterogeneous" approaches. We performed a NJ analysis under a LogDet transformation (Lockhart et al. 1994). Because the LogDet transformation is known to be affected by inclusion of invariable sites (Steel et al. 2000), we estimated the proportion of invariable sites under various models of sequence evolution (F81, F81 + G, HKY85, HKY85 + G, GTR, and GTR + G) and removed the estimated proportions for the LogDet analyses. In a likelihood framework, we used the nonstationary nonhomogeneous model of evolution of Galtier and Gouy (1998) as implemented in nhPhyML (Boussau and Guoy 2006). Although recent improvements have increased the tree-searching capabilities in nhPhyML, search space is still limited. Studies employing this model have been restricted to testing the likelihood of opposing hypotheses (trees) rather than relying on tree search to yield the maximum likelihood estimate. Because this software required an input topology, we supplied 2 rooted trees, one from the LogDet analysis and another from the p4 analysis (see below). We used nhPhyML-Discrete limited to 4 base content frequency categories and with a 4-category discrete gamma model of among-site rate variation. We also tried other settings for nhPhyML such as ignoring among-site rate variation and using a continuous range of equilibrium frequencies. To analyze the data from a nonstationary Bayesian framework, we used 3 different software packages: p4 v.0.85-0.86 (Foster 2004), PhyloBayes 2.3 (Blanquart and Lartillot 2006), and PHASE 2.1alpha (Gowri-Shankar and Rattray 2007). For our p4 analysis, we adjusted tuning numbers by hand during preliminary runs to achieve acceptance proportions near 40% for most proposals (topology changes stayed near 5%). We used a single GTR + I + dG model with 4 gamma categories and 3 composition vectors. We performed 3 runs with 4 chains each for 3 million generations, sampling every 1000 trees. We discarded the first 2 million generations as burn-in. For PhyloBayes, we performed 2 runs for each of 4 analyses using the following parameter combinations: -cat -gtr, -cat -poisson, -ncat 3 -gtr, and -ncat 3 -poisson. Each run continued for 10 000–30 000 points (differences in computational requirements changed the length of the runs). The alpha version of PHASE 2.1 includes a reversible jump algorithm that allows a variable number of composition vectors to be used to describe the data. We performed 4 runs each with 4 chains for 9 million generations, sampling every 500 trees. We used a different starting random seed and initial number of composition vectors (1, 2, 5, and 10) for each iteration. We discarded the first 3 million generations as burn-in. For these Bayesian runs, we assessed convergence by monitoring log likelihood traces using R (R Development Core Team 2008), cumulative and sliding window split frequencies using AWTY (Nylander et al. 2008), across-run topology consensus trees, and the PhyloBayes bpcomp run comparison utility.

Results

Level of Base Compositional Heterogeneity

The total AT% of all included coleopteran species ranged between 65.63% and 78.19% with a mean of 71.41 (±4.88)%. The 5 out-group taxa had a total AT% between 76.00% and 79.58% with a mean of 78.02 (±1.54)%. Among beetles, *Acmaeodera*, *Apatides*, *Lucanus*, *Pyrophorus*, *Tetraphalerus*, and *Tribolium* had total AT% below 70%. Other beetles with the exception of *Adelium* (70.42%) had higher AT% than the out-groups. When we compared base composition for individual genes, there was a considerable amount of interspecific variation (Fig. 7).

We made 153 pairwise comparisons to calculate the I_D. When we compared all 13 protein-coding genes simultaneously, 142 comparisons had a statistically significant heterogeneous substitution pattern (Table 2). In general, the I_D was low between the taxa that were near the average of 75% AT and very high between the taxa with wider ranges of AT%. Interestingly, the I_D was consistently low between the taxa with low AT%, regardless of their phylogenetic position within Coleoptera, suggesting that the pattern of substitution evolved multiple times. The I_D test on the individual genes suggests a high level of variation in the substitution patterns among different mitochondrial genes, although some of this variation can be explained by gene length (Table 3). The overall pattern was consistent in that the I_D was low between the taxa with low AT% across all 13 protein-coding genes. Additional research into the variation of base composition among genes would be useful for determining exactly how base composition is distributed among genes.

Time-homogeneous Approaches

If base composition were biased enough to group evolutionarily unrelated taxa, we would expect the conventional phylogenetic methods to result in incorrect topologies. The parsimony analysis resulted in a single most parsimonious tree ($L = 33\,854$, consistency index (CI) = 0.41, retention index (RI) = 0.29), and the Bayesian analysis resulted in a tree with almost all nodes with posterior probability of 1.00 (Fig. 2; TreeBASE accession number SN4359). The topology was nearly identical between these 2 inference methods. The parsimony analysis from the back-translated nucleotide alignment also resulted in an identical topology (L = 34574, CI = 0.41, RI = 0.29), which indicates that the effect of alignment was minimal. With AT% of each taxon mapped on to the tree, it is evident that taxa with low AT% (65–70%) grouped together. As a result, unexpected clades were recovered with high support values. First, a sister relationship between monophyletic Archostemata and Polyphaga was not recovered because *Tetraphalerus* grouped with polyphagans with similarly low AT%. Second, a monophyletic Cucujiformia was not recovered because 2 taxa belonging to

TABLE 2. Disparity index results: pairwise base composition bias disparity between sequences (based on all 13 protein-coding genes) and the results of a test of homogeneity of substitution pattern (I_D test) using a Monte–Carlo method with 1000 replications

	Dros	*Anop*	*Osrt.*	*Bomb.*	*Anth.*	**Tetr.**	*Crio.*	*Pria.*	*Chae.*	**Trib.**	**Adel.**	*Pyroc.*	**Pyrop.**	*Rhag.*	*Chau.*	**Acma.**	**Apat.**	**Luca.**
Drosophila	—	0.14	3.13	3.82	1.79	75.11	1.74	7.60	0.74	34.87	24.95	1.99	48.29	1.15	3.99	60.44	65.27	63.72
Anopheles	0.17	—	4.99	5.98	3.29	65.58	0.58	4.85	1.29	28.40	19.61	1.40	40.56	2.33	3.30	51.82	56.28	54.85
Ostrinia	<0.01	<0.01	—	0.02	2.25	97.87	6.99	15.89	1.11	52.34	39.13	4.24	70.03	0.32	5.00	85.43	88.25	90.30
Bombyx	<0.01	<0.01	0.34	—	1.81	103.79	8.42	18.20	1.76	56.34	42.85	5.47	74.49	0.56	6.55	89.59	93.21	95.02
Antheraea	<0.01	<0.01	<0.01	<0.01	—	96.24	6.23	15.27	1.94	49.31	37.93	5.23	64.93	0.89	7.87	76.83	83.57	82.18
Tetraphalerus	<0.01	<0.01	<0.01	<0.01	<0.01	—	53.63	34.90	77.16	**7.64**	**13.12**	60.62	**4.30**	87.55	59.07	**6.36**	**1.26**	**5.44**
Crioceris	<0.01	0.04	<0.01	<0.01	<0.01	<0.01	—	1.76	2.09	21.01	13.28	0.41	32.33	4.01	1.50	43.00	45.67	46.25
Priasilpha	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	8.07	10.20	4.92	3.46	18.85	11.61	3.91	28.03	28.92	30.55
Chaetosoma	0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	37.09	26.31	0.82	52.02	0.11	1.79	65.16	68.08	69.48
Tribolium	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	**0.88**	26.23	**1.19**	44.15	26.36	**5.25**	**4.51**	**6.09**
Adelium	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	17.21	**4.76**	32.48	17.02	**11.20**	**10.02**	**12.20**
Pyrocoelia	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.09	<0.01	0.02	<0.01	<0.01	—	39.39	2.34	0.15	51.34	52.87	55.41
Pyrophorus	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	60.16	40.01	1.49	1.41	1.48
Rhagophthalmus	<0.01	<0.01	0.10	0.04	0.01	<0.01	<0.01	<0.01	0.24	<0.01	<0.01	<0.01	<0.01	—	3.63	73.67	77.44	78.44
Chauliognathus	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.22	<0.01	<0.01	—	52.95	52.68	56.99
Acmaedera	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	**1.65**	**0.51**
Apatides	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	—	**2.04**
Lucanus	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	0.04	<0.01	—

Notes: The estimates of the disparity index are shown for each sequence pair above the diagonal. The probability of rejecting the null hypothesis that sequences have evolved with the same pattern of substitution is shown below the diagonal. Bold are the taxa with low overall AT%. Although the majority of the comparisons resulted in significantly heterogeneous substitution patterns, the disparity index between the taxa with low AT% is relatively low (values shown in bold).

TABLE 3. I_D test summary

Gene	Number of comparisons with significant heterogeneity	Proportion of significant heterogeneity from all comparisons (out of 153) (%)	Aligned length of gene partition
nd2	123	80.39	1099
cox1	121	79.08	1549
cox2	101	66.01	689
atp8	32	20.92	218
atp6	105	68.63	750
cox3	97	63.40	792
nd3	78	50.98	361
nd5	117	76.47	1773
nd4	106	69.28	1391
nd4l	29	18.95	306
nd6	76	49.67	561
cytb	118	77.12	1156
nd1	92	60.13	988

Notes: Summary of the I_D test on individual gene partitions. For each gene, a total of 153 pairwise comparisons were made and shown here are the number of comparisons with the probability of rejecting the null hypothesis that sequences have evolved with the same pattern of substitution at the $\alpha < 0.01$ level. The same number is also shown in the form of proportion. There is a positive correlation between the result of I_D test and the length of gene (Pearson correlation coefficient = 0.8117).

Tenebrionidae (*Tribolium* and *Adelium*) grouped with other low AT% taxa, thereby forming a paraphyletic Cucujiformia with a clade formed by Cleroidea (*Chaetosoma*), Chrysomeloidea (*Crioceris*), and Cucujoidea (*Priasilpha*). Third, a monophyletic Elateroidea was not recovered because 1 of the 4 members of the superfamily included in the analysis (*Pyrophorus*) grouped with other low AT% taxa, resulting in a paraphyletic Elateroidea. Finally, 3 divergent polyphagan taxa, *Apatides*, *Lucanus*, and *Acmaeodera*, were closely grouped, presumably because they share similarly low AT% (although this grouping could be feasible due to small taxon sampling; see Hunt 2007). The codon position partitioned Bayesian analysis yielded a similar topology, although a monophyletic Elateroidea was recovered. The other 2 unexpected relationships (the erroneous *Tetraphalerus* and Cucujiformia groupings) were equivalent.

Time-heterogeneous Approaches

LogDet transformation has been shown to correct for compositional bias (Lockhart et al. 1994), but it did not help recover the correct topology for our data set (Fig. 3). It recovered nearly the same topology as the

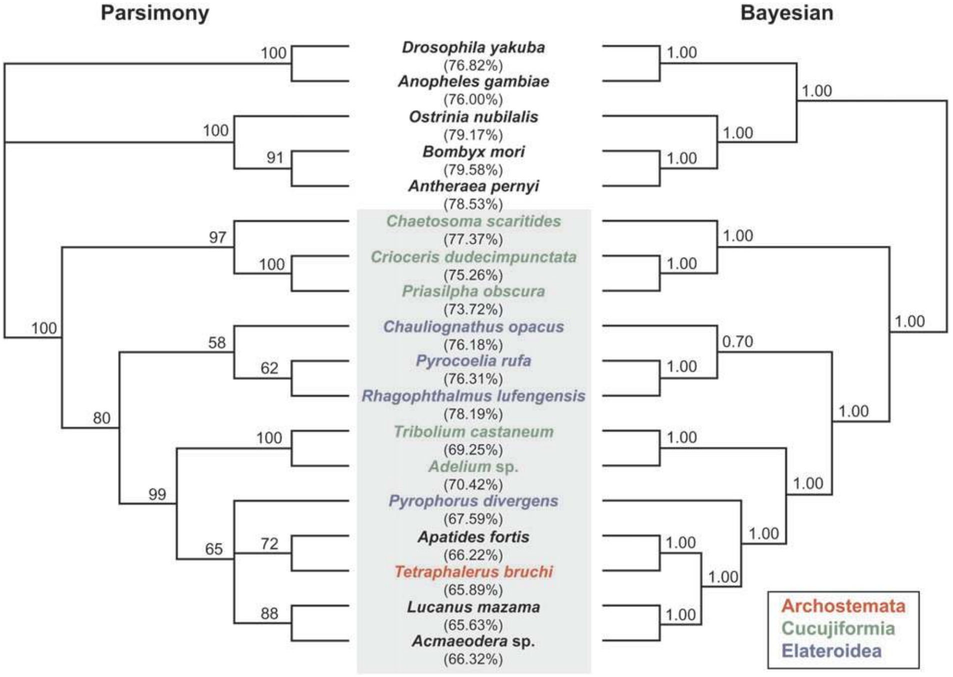

FIGURE 2. The single most parsimonious tree with bootstrap support values and the Bayesian tree with posterior probability. Number in parentheses under taxon name indicates the overall AT% for each taxon. Taxa in light gray box are Coleoptera and each terminal within is color coded to indicate its taxonomical grouping: red = Archostemata, green = Cucujiformia, blue = Elateroidea, and black = other Polyphaga.

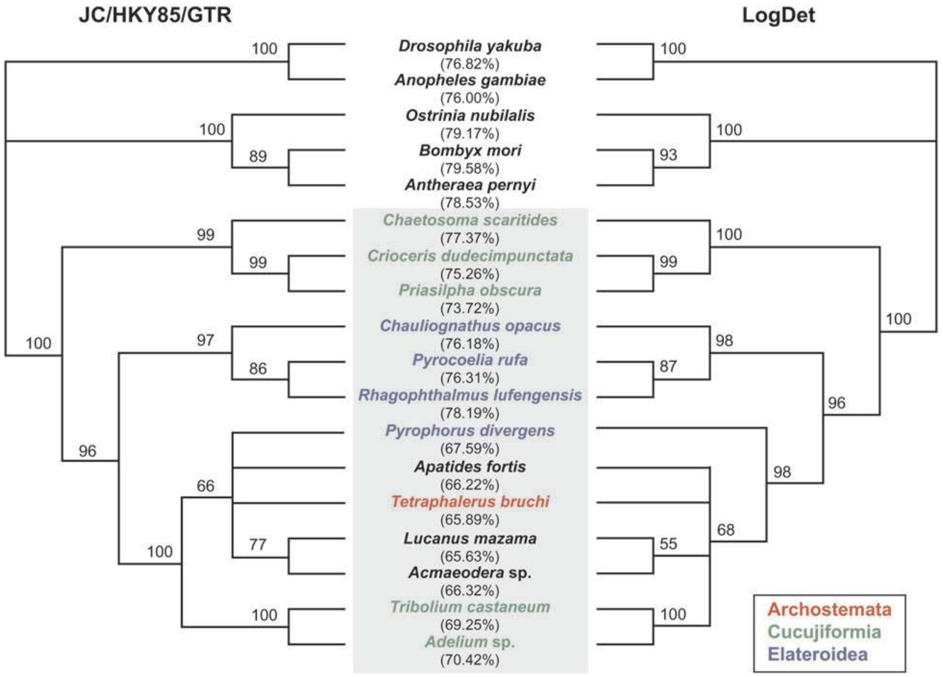

FIGURE 3. Comparison of standard NJ analyses (JC, HKY85, and GTR) and an analysis after LogDet transformation. The identical topology to the LogDet tree shown here was consistently recovered after removing a proportion of invariable sites under different models of sequence evolution. Number above each node indicates bootstrap support value. Number in parentheses under taxon name indicates the overall AT% for each taxon. Taxa in light gray box are Coleoptera and each terminal within is color coded to indicate its taxonomical grouping: red = Archostemata, green = Cucujiformia, blue = Elateroidea, and black = other Polyphaga.

parsimony and Bayesian trees, and there was no substantial difference between LogDet and other distance models, with or without gamma-distributed among-site rate variation. Removal of a proportion of invariable sites estimated under various maximum likelihood models did not have any effect on the resulting topology despite the fact that different models estimated different proportions (F81 = 0.31, F81 + G = 0.02, HKY85 = 0.31, HKY85 + G = 0.16, GTR = 0.31, GTR + G = 0.01). Thus, LogDet did not recover any of 3 indicator relationships correctly. The GG95 model, nhPhyML, had difficulty exploring tree space. Depending on the input topology, the runs sometimes selected a tree with the 3 indicator relationships or a tree similar to the parsimony (or LogDet) tree. Under nhPhyML-Discrete with gamma rates, the “correct” tree was given a higher likelihood than the “incorrect” (LogDet) tree (Fig. 4). However, when we did not account for among-site rate variation, all the runs converged on the incorrect topology regardless of starting topology. The analyses in a nonstationary Bayesian framework performed better in terms of recovering the correct topology. All PHASE runs converged to an identical topology with all 3 indicator relationships (Fig. 5a). Although we had some difficulty with p4 run convergence for particular clades, all p4 runs also converged on the 3 indicator relationships correctly (Fig. 5b). The p4 consensus tree reflects some clade instability with low posterior probabilities for certain groups. Thus, PHASE and p4 recovered similar trees with all 3 indicator relationships. PhyloBayes, however, failed to recover any of the 3 indicator relationships (Fig. 5c).

Amino Acid Recoding

Recoding nucleotide characters into translated amino acid sequences had a positive effect on phylogenetic reconstruction. The resulting parsimony tree (Fig. 6; $L = 13\,878$, CI = 0.57, RI = 0.34) recovered a monophyletic Coleoptera and correctly placed *Tetraphalerus* basal to the remaining Polyphaga. Within Polyphaga, Cucujiformia and Elateroidea were both correctly recovered as monophyletic groups. However, the relationships among *Lucanus* and *Apatides*, *Acmaeodera*, as well as Cucujiformia and Elateroidea, were not resolved.

DISCUSSION

Our study illustrates the possibility of incorrect phylogenetic inference with high support when considering a data set with base compositional heterogeneity. In an analysis of closely related species, one rarely suspects contamination from base compositional heterogeneity a priori. When an initial topology does not seem to make sense, however, the researcher examines the original data more carefully. In this study, our initial topology from conventional phylogenetic methods failed to recover established higher level relationships. These unexpected results can be caused by a number of factors

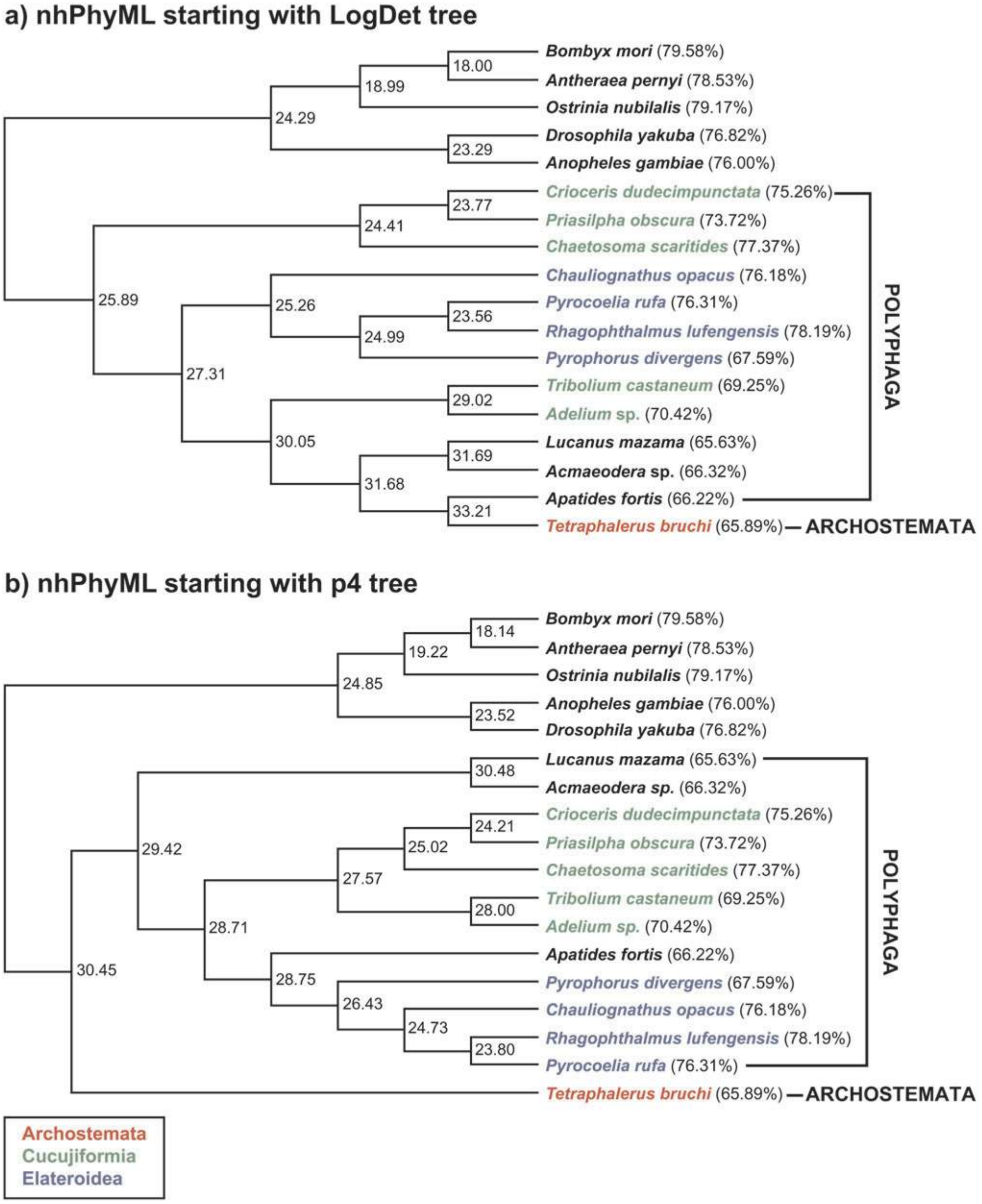
a) nhPhyML starting with LogDet tree
18.00
18.99
24.29
23.29
23.77
24.41
25.89
25.26
23.56
24.99
27.31
29.02
30.05
31.69
31.68
33.21
Bombyx mori (79.58%)
Antheraea pernyi (78.53%)
Ostrinia nubilalis (79.17%)
Drosophila yakuba (76.82%)
Anopheles gambiae (76.00%)
Crioceris dudecimpunctata (75.26%)
Priasilpha obscura (73.72%)
Chaetosoma scaritides (77.37%)
Chauliognathus opacus (76.18%)
Pyrocoelia rufa (76.31%)
Rhagophthalmus lufengensis (78.19%)
Pyrophorus divergens (67.59%)
Tribolium castaneum (69.25%)
Adelium sp. (70.42%)
Lucanus mazama (65.63%)
Acmaeodera sp. (66.32%)
Apatides fortis (66.22%)
Tetraphalerus bruchi (65.89%)
POLYPHAGA
ARCHOSTEMATA
b) nhPhyML starting with p4 tree
18.14
19.22
24.85
23.52
30.48
24.21
29.42
25.02
27.57
28.00
28.71
30.45
28.75
26.43
24.73
23.80
Bombyx mori (79.58%)
Antheraea pernyi (78.53%)
Ostrinia nubilalis (79.17%)
Anopheles gambiae (76.00%)
Drosophila yakuba (76.82%)
Lucanus mazama (65.63%)
Acmaeodera sp. (66.32%)
Crioceris dudecimpunctata (75.26%)
Priasilpha obscura (73.72%)
Chaetosoma scaritides (77.37%)
Tribolium castaneum (69.25%)
Adelium sp. (70.42%)
Apatides fortis (66.22%)
Pyrophorus divergens (67.59%)
Chauliognathus opacus (76.18%)
Rhagophthalmus lufengensis (78.19%)
Pyrocoelia rufa (76.31%)
Tetraphalerus bruchi (65.89%)
POLYPHAGA
ARCHOSTEMATA
Archostemata
Cucujiformia
Elateroidea

FIGURE 4. Results of nhPhyML-Discrete with gamma-distributed rate variation. The analysis results in different topologies, depending on which input tree is used. Each coleopteran terminal within is color coded to indicate its taxonomical grouping: red = Archostemata, green = Cucujiformia, blue = Elateroidea, and black = other Polyphaga.

other than nonstationary evolution, such as long-branch attraction (Felsenstein 1978; Bergsten 2005), heterotachy (Kolaczkowski and Thornton 2004; Philippe et al. 2005; Ruano-Rubio and Fares 2007), and among-site rate variation (Yang 1996; Steel et al. 2000). A combination of these and other factors is at work in any evolutionary analysis. As such, it would be overly simplistic to assume that nonstationary evolution is the only factor affecting this or any other data set, even when mapping composition onto the topology seems to make it so clear. Our nhPhyML results illustrate this in our data set: using this method, it is not enough to solely account for compositional bias; one must also account for among-site rate heterogeneity. Along the same lines, our Bayesian analyses that account for among-site rate heterogeneity but not compositional bias end up in almost the same place as the parsimony analysis. In this data set, nonstationary evolution does appear to be one of the major confounding factors, as evidenced by the results of the I_D test as well as the fact that some of the new Bayesian methods that employ nonstationary evolutionary models succeed where alternatives fail.

Unfortunately, it is not always straightforward to predict if a data set will have a problem with bias. Studies that have given exact numbers of where to expect problems, such as 10% base compositional difference (Eyre-Walker 1998), have been rebutted by other studies with data where lower differences have caused

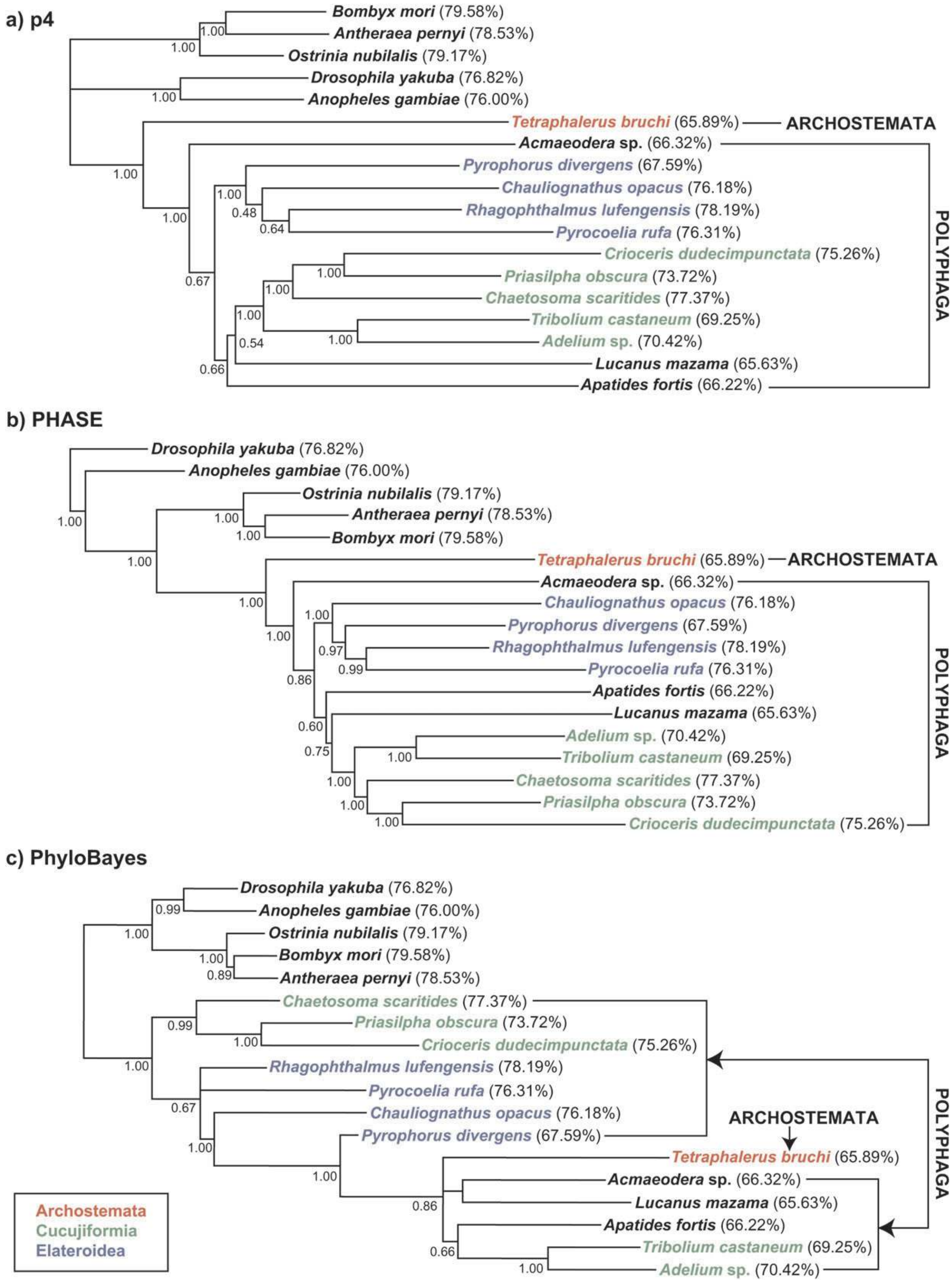
a) p4
Bombyx mori (79.58%)
Antheraea pernyi (78.53%)
Ostrinia nubilalis (79.17%)
Drosophila yakuba (76.82%)
Anopheles gambiae (76.00%)
Tetraphalerus bruchi (65.89%)
ARCHOSTEMATA
Acmaeodera sp. (66.32%)
Pyrophorus divergens (67.59%)
Chauliognathus opacus (76.18%)
Rhagophthalmus lufengensis (78.19%)
Pyrocoelia rufa (76.31%)
Crioceris dudecimpunctata (75.26%)
Priasilpha obscura (73.72%)
Chaetosoma scaritides (77.37%)
Tribolium castaneum (69.25%)
Adelium sp. (70.42%)
Lucanus mazama (65.63%)
Apatides fortis (66.22%)
POLYPHAGA
1.00
1.00
1.00
1.00
1.00
1.00
0.48
0.64
0.67
1.00
1.00
1.00
1.00
0.54
0.66
b) PHASE
Drosophila yakuba (76.82%)
Anopheles gambiae (76.00%)
Ostrinia nubilalis (79.17%)
Antheraea pernyi (78.53%)
Bombyx mori (79.58%)
Tetraphalerus bruchi (65.89%)
ARCHOSTEMATA
Acmaeodera sp. (66.32%)
Chauliognathus opacus (76.18%)
Pyrophorus divergens (67.59%)
Rhagophthalmus lufengensis (78.19%)
Pyrocoelia rufa (76.31%)
Apatides fortis (66.22%)
Lucanus mazama (65.63%)
Adelium sp. (70.42%)
Tribolium castaneum (69.25%)
Chaetosoma scaritides (77.37%)
Priasilpha obscura (73.72%)
Crioceris dudecimpunctata (75.26%)
POLYPHAGA
1.00
1.00
1.00
1.00
1.00
1.00
1.00
0.97
0.99
0.86
0.60
0.75
1.00
1.00
1.00
1.00
c) PhyloBayes
Drosophila yakuba (76.82%)
Anopheles gambiae (76.00%)
Ostrinia nubilalis (79.17%)
Bombyx mori (79.58%)
Antheraea pernyi (78.53%)
Chaetosoma scaritides (77.37%)
Priasilpha obscura (73.72%)
Crioceris dudecimpunctata (75.26%)
Rhagophthalmus lufengensis (78.19%)
Pyrocoelia rufa (76.31%)
Chauliognathus opacus (76.18%)
Pyrophorus divergens (67.59%)
ARCHOSTEMATA
Tetraphalerus bruchi (65.89%)
Acmaeodera sp. (66.32%)
Lucanus mazama (65.63%)
Apatides fortis (66.22%)
Tribolium castaneum (69.25%)
Adelium sp. (70.42%)
POLYPHAGA
0.99
1.00
1.00
0.89
0.99
1.00
1.00
0.67
1.00
1.00
0.86
0.66
1.00
Archostemata
Cucujiformia
Elateroidea

FIGURE 5. Comparison of 3 Bayesian methods implementing algorithms to account for compositional heterogeneity. a) p4, b) PHASE, and c) PhyloBayes. Number below each node indicates posterior probability. The trees shown are 50% majority rule consensus tree for (a) and (b) and –cat –poisson analysis tree for (c). Number in parentheses next to taxon name indicates the overall AT% for each taxon. Each coleopteran terminal within is color coded to indicate its taxonomical grouping: red = Archostemata, green = Cucujiformia, blue = Elateroidea, and black = other Polyphaga.

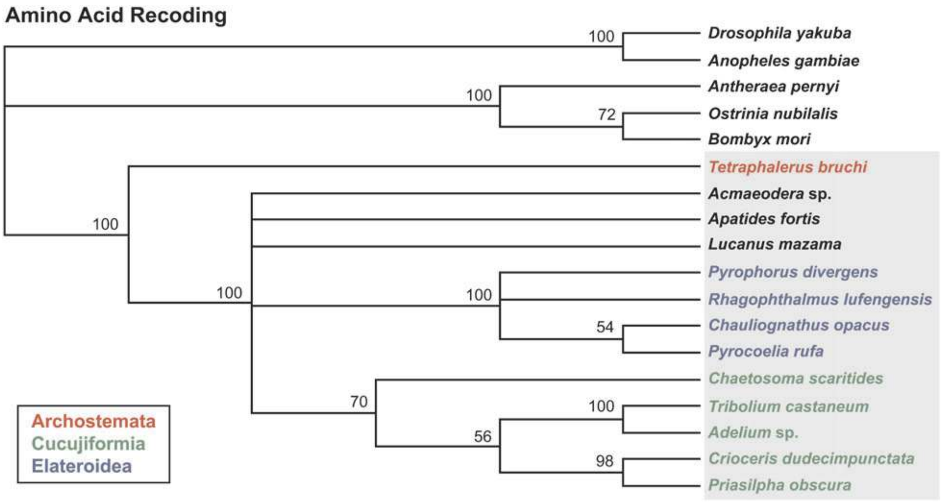

FIGURE 6. A parsimony tree with bootstrap support values after recoding nucleotide data into amino acid sequences. Taxa in light gray box are Coleoptera and each terminal within is color coded to indicate its taxonomical grouping: red = Archostemata, green = Cucujiformia, blue = Elateroidea, and black = other Polyphaga.

problems (Gruber et al. 2007). Jermiin et al. (2004) provide a thorough review of methods for assessing degree of compositional heterogeneity.

Our study demonstrates that the parsimony analysis severely suffers from the high level of homoplasy resulting from convergent evolution (Fig. 2). In many cases, model-based methods can account for confounding factors and infer topologies that are more accurate; however, in this study, the Bayesian analysis with even the most parameter-rich models for individual gene partitions recovered a topology almost identical to the parsimony analysis (Fig. 2). This suggests that the homoplasy is not due to factors accounted for in that model (such as among-site rate heterogeneity). Other methods that use similar models but also account for nonstationarity are able to overcome the problem. Therefore, we affirm that base compositional heterogeneity affects both parsimony and likelihood approaches.

Data recoding can be an effective means to overcome the homoplasy from nonstationarity (Loomis and Smith 1990; Nardi et al. 2001; Cameron et al. 2004), although some have cautioned that it can introduce artifactual relationships (Yang and Roberts 1995; Cameron et al. 2007; Fenn et al. 2008). In this data set, amino acid recoding is an effective method of dealing with the problem. When we apply amino acid recoding to our data set in a parsimony framework, all 3 indicator relationships are recovered correctly, suggesting that the compositional

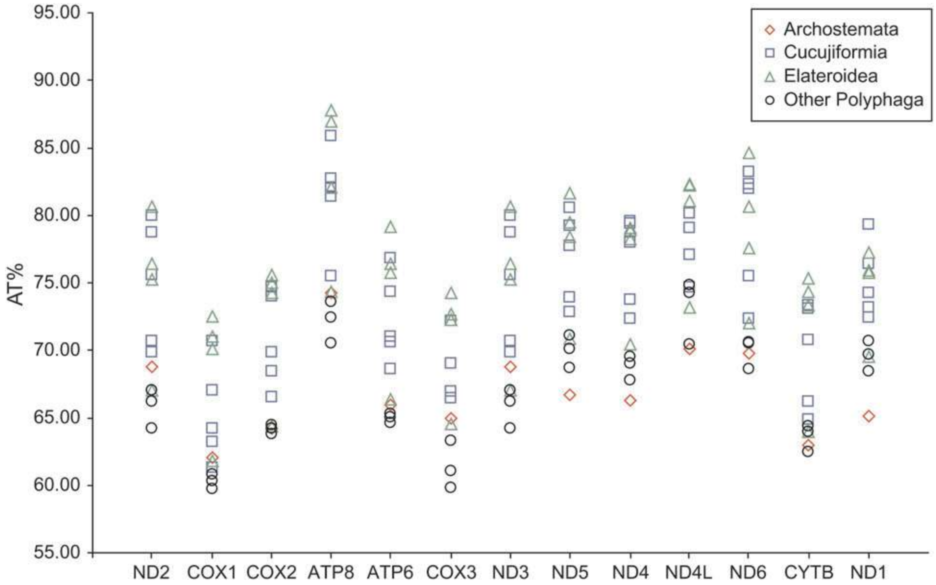

FIGURE 7. Gene by gene comparison of compositional bias. Gene order shown here follows that of a hypothetical insect ancestor. Each shape is color coded to indicate its taxonomic grouping: red diamond = Archostemata (*Tetraphalerus*), blue square = Cucujiformia, green triangle = Elateroidea, and black circle = other Polyphaga with low AT%.

bias does not extend to the amino acid level (Fig. 6). However, some relationships are unresolved, which indicates that a by-product of the reduced coding may be the loss of phylogenetic signal.

Many studies that recognize compositional bias in the data set employ LogDet as a means to correct the problem (van den Bussche et al. 1998; Nishiyama and Kato 1999; Waddell et al. 1999; Tarrío et al. 2000. Phillips et al. 2004; Barrowclough et al. 2006). However, LogDet does not correct the bias in our data set, and, in fact, the transformation does not appear to be much better than the other NJ trees (Fig. 3). Removal of the proportion of invariable sites with LogDet (Steel et al. 2000) did not have any effect of the final topology regardless of the model of sequence evolution used to estimate the proportion. This finding suggests that our data set contains a high level of bias that is difficult to overcome by a simple transformation of the data.

The 13 protein-coding genes may be under different selective pressures and exhibit variations in composition despite the fact that mtgenomes evolve as a single unit. The comparisons of the AT% in individual protein-coding genes results in a pattern that is difficult to generalize (Fig. 7). Sometimes, a species has a consistently low AT% across all genes (e.g., *Tetraphalerus* in Fig. 7). Other times, only some genes have low AT%. The I_D test on individual genes also confirms that the pairwise substitution patterns are highly variable across genes (Table 3). These observations suggest that although 2 mtgenomes may convergently evolve lower compositional bias as a whole, individual genes within each mtgenome can evolve differently, resulting in different substitution patterns. Whether or not these differences are biologically meaningful or simply the realization of stochastic variation is difficult to know, but this complex pattern of compositional bias introduces a number of issues in multigene phylogenetic analyses. Calculations that are based on the mean nucleotide composition of the combined data set do not correct the bias in individual genes that have a different bias from the whole data set (Bevan et al. 2007). This is one possible reason for the failure of LogDet in this data set.

Although LogDet is a simple transformation of data, several recently developed models more realistically model nonstationary evolution. These new models allow the composition to differ over the tree during phylogenetic reconstruction. Thus, they can accommodate localized compositional bias and apply different model parameters accordingly. When such models are used, the compositional bias can be overcome and correct relationships can be recovered. In our analysis, both p4 and PHASE correctly recovered the 3 indicator relationships with high support. However, a similar method, PhyloBayes, failed to recover the indicator relationships. It is unclear why PhyloBayes failed with this data set. We used several of the available options implemented in the software and each gave a different incorrect result. In theory, the implementation of the nonstationary model in PhyloBayes is more realistic than nonstationary models implemented in other software because changes in composition vector are not restricted to speciation events. In many cases, this type of model is known to fit real data better (Blanquart and Lartillot 2006). However, for our data set, this model did not perform well. It is possible that the additional complexity of the model makes it more difficult to fit the parameters, resulting in less efficient exploration of tree space (Steel 2005).

It is difficult to speculate why and how the bias might have originated. However, we can hypothesize that similar compositional bias can arise independently many times throughout a lineage. For instance, ~65% AT content has evolved at least twice in 2 widely divergent lineages, *Tetraphalerus* (Archostemata) and *Lucanus* (Polyphaga). To what degree compositional bias is inherited is difficult to determine. On one hand, 2 closely related species belonging to the same family (*Tribolium* and *Adelium*) have similar compositional bias, which may suggest that this bias is a shared character originating in the tenebrionid common ancestor. On the other hand, only 1 of 4 taxa (*Pyrophorus*) in the superfamily Elateroidea has a distinctly lower AT bias, which suggests that this bias must have evolved after this superfamily diversified.

Our study serves as a reminder that base compositional heterogeneity can lead to incorrect phylogenetic inference. Traditionally used methods, such as LogDet, can fail to correct the bias due to lack of data partitioning, failure to account for among-site rate heterogeneity, or other problems. Reduced coding techniques can be useful in correcting the bias at the level of nucleotides but can also result in the loss of phylogenetic signal. Some newer models that account for bias in a complex data set, however, are effective at correcting the problem. Because this bias causes a systematic error, high nodal supports are in fact no indication of topological accuracy: incorrect clades are often highly supported in our study. In particular, Bayesian posterior probabilities can be misleading in the event of model misspecification (Erixon et al. 2003). We recommend that future phylogenetic studies, especially ones based on complex data sets consisting of a large number of heterogeneous loci, examine base composition before analysis thoroughly to minimize incorrect inferences due to compositional bias.

Supplementary Material

Supplementary material can be found at: http://www.sysbio.oxfordjournals.org.

Funding

This project has been supported by National Science Foundation grants AToL EF-0531665 to M.F.W. and DEB-0444972 to S.L.C. and M.F.W. and a Brigham Young University Office of Research and Creative Activities grant to N.C.S.

Acknowledgments

Thanks are due to Vivek Gowri-Shankar for helpful advice running PHASE. Thanks also to James Robertson for comments on an earlier version of this manuscript. Helpful comments from J. Sullivan, T. Buckley, M. Phillips, G. Naylor, and C. Lakner improved our manuscript.

References

Arnoldi F.G.C., Ogoh K., Ohmiya Y., Viviani V.R. 2007. Mitochondrial genome sequence of the Brazilian luminescent click beetle Pyrophorus divergens (Coleoptera: Elateridae): mitochondrial genes utility to investigate the evolutionary history of Coleoptera and its bioluminescence. Gene 405:1–9.

Bae J.S., Kim I., Sohn H.D., Jin B.R. 2004. The mitochondrial genome of the firefly, Pyrocoelia rufa: complete DNA sequence, genome organization, and phylogenetic analysis with other insects. Mol. Phylogenet. Evol. 32:978–985.

Barrowclough G.F., Groth J.G., Mertz L.A. 2006. The rag-1 exon in the avian order caprimulgiformes: phylogeny, heterozygosity, and base composition. Mol. Phylogenet. Evol. 41:238–248.

Beard C.B., Hamm D.M., Collins F.H. 1993. The mitochondrial genome of the mosquito Anopheles gambiae. DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Mol. Biol. 2:103–124.

Bergsten J. 2005. A review of long-branch attraction. Cladistics 21: 163–193.

Beutel R.G., Ge S., Hornschemeyer T. 2008. On the head morphology of *Tetraphalerus*, the phylogeny of Archostemata and the basal branching events in Coleoptera. Cladistics 24:270–298.

Beutel R.G., Haas F. 2000. Phylogenetic relationships of the suborders of Coleoptera (Insecta). Cladistics 16:103–141.

Bevan R.B., Bryant D., Lang B.F. 2007. Accounting for gene rate heterogeneity in phylogenetic inference. Syst. Biol. 56:194–205.

Blanquart S., Lartillot N. 2006. A Bayesian compound stochastic process for modeling nonstationary and nonhomogeneous sequence evolution. Mol. Biol. Evol. 23:2058.

Bocakova M., Bocak L., Hunt T., Teravainen M., Vogler A.P. 2007. Molecular phylogenetics of Elateriformia (Coleoptera): evolution of bioluminescence and neoteny. Cladistics 23:477–496.

Boussau B., Gouy M. 2006. Efficient likelihood computations with nonreversible models of evolution. Syst. Biol. 55:756–768.

Brown W.M., Prager E.M., Wang A., Wilson A.C. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. 18:225–239.

Cameron S.L., Barker S.C., Whiting M.F. 2006. Mitochondrial genomics and the relationships and validity of the new insect order Mantophasmatodea. Mol. Phylogenet. Evol. 38:274–279.

Cameron S.L., Lambkin C.L., Barker S.C., Whiting M.F. 2007. A mitochondrial genome phylogeny of Diptera: whole genome sequence data accurately resolve relationships over broad timescales with high precision. Syst. Entomol. 32:40–59.

Cameron S.L., Miller K.B., D'Haese C.A., Whiting M.F., Barker S.C. 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea sensu lato (Athropoda). Cladistics 20:534–557.

Campbell D.L., Brower A.V.Z., Pierce N.E. 2000. Molecular evolution of the wingless gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera: Papilionoidea). Mol. Biol. Evol. 17:684–696.

Caveney S. 1986. The phylogenetic significance of ommatidium structure in the compound eyes of polyphagan beetles. Can. J. Zool. 64:1787–1819.

Chang B.S.W., Campbell D.L. 2000. Bias in phylogenetic reconstruction of vertebrate rhodopsin sequences. Mol. Biol. Evol. 17:1220–1231.

Clary D.O., Wolstenholme D.R. 1985. The mitochondrial DNA molecule of Drosophila yakuba: Nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22:252–271.

Coates B.S., Sumerford D.V., Hellmich R.L., Lewis L.C. 2005. Partial mitochondrial genome sequences of Ostrinia nubilalis and Ostrinia furnicalis. Int. J. Biol. Sci. 1:13.

Collins T.M., Fedrigo O., Naylor G.J.P. 2005. Choosing the best genes for the job: the case for stationary genes in genome-scale phylogenies. Syst. Biol. 54:493–500.

Collins T.M., Wimberger P.H., Naylor G.J.P. 1994. Compositional bias, character-state bias, and character-state reconstruction using parsimony. Syst. Biol. 43:482–496.

Conant G.C., Lewis P.O. 2001. Effects of nucleotide composition bias on the success of the parsimony criterion in phylogenetic inference. Mol. Biol. Evol. 18:1024–1033.

Cox C.J., Foster P.G., Hirt R.P., Harris S.R., Embley T.M. 2008. The archaebacterial origin of eukaryotes. Proc. Natl. Acad. Sci. U.S.A. 105: 20356–20361

Crowson R.A. 1960. The phylogeny of Coleoptera. Annu. Rev. Entomol. 5:111–134.

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

Edgar R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792.

Erixon P., Svennblad B., Britton T., Oxelman B. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52:665–673.

Eyre-Walker A. 1998. Problems with parsimony in sequences of biased base composition. J. Mol. Evol. 47:686–690.

Felsenstein J. (1978). Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401–410.

Fenn J.D., Song H., Cameron S.L., Whiting M.F. 2008. A mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. Mol. Phylogenet. Evol. 49:59–68.

Foster P.G. 2004. Modeling compositional heterogeneity. Syst. Biol. 53:485–495.

Foster P.G., Hickey D.A. 1999. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. J. Mol. Evol. 48:284–90.

Friedrich M., Muqim N. 2003. Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle Tribolium castanaeum. Mol. Phylogenet. Evol. 26:502–512.

Galtier N., Gouy M. 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc. Natl. Acad. Sci. U.S.A. 92:11317–11321.

Galtier N., Gouy M. 1998. Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. Mol. Biol. Evol. 15:871–879.

Gibson A., Gowri-Shankar V., Higgs P.G., Rattray M. 2005. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol. Biol. Evol. 22:251–264.

Gowri-Shankar V., Rattray M. 2007. A reversible jump method for Bayesian phylogenetic inference with a nonhomogeneous substitution model. Mol. Biol. Evol. 24:1286–1299.

Gruber K.F., Voss R.S., Jansa S.A. 2007. Base-compositional heterogeneity in the RAG1 locus among didelphid marsupials: implications for phylogenetic inference and the evolution of GC content. Syst. Biol. 56:83–96.

Herbeck J.T., Degnan P.H., Wernegreen J.J. 2005. Nonhomogeneous model of sequence evolution indicates independent origins of primary endosymbionts within the enterobacteriales (gamma-Proteobacteria). Mol. Biol. Evol. 22:520–532.

Ho S.Y., Jermiin L. 2004. Tracing the decay of the historical signal in biological sequence data. Syst. Biol. 53:623–637.

Hrdy I., Hirt R.P., Dolezal P., Bardonova L., Foster P.G., Tachezy J., Embley T.M. 2004. Trichomonas hydrogenosomes contain the nadh dehydrogenase module of mitochondrial complex I. Nature 432:618–622.

Hughes J., Longhorn S.J., Papadopoulou A., Theodorides K., de Riva A., Mejia-Chang M., Foster P.G., Vogler A.P. 2006. Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (beetles). Mol. Biol. Evol. 23:268–278.

Hunt T., Bergsten J., Levkanicova Z., Papadopoulou A., John O.S., Wild R., Hammond P.M., Ahrens D., Balke M., Caterino M.S. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. Science 318:1913.

Jermiin L., Ho S.Y., Ababneh F., Robinson J., Larkum A.W. 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53:638–643.

Kolaczkowski B., Thornton J.W. 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature 431:980–984.

Kumar S., Gadagkar S.R. 2001. Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics 158:1321–1327.

Lake J.A. 1994. Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances. Proc. Natl. Acad. Sci. U.S.A. 91:1455–1459.

Lawrence J.F. 1988. Rhinorrhipidae, a new beetle family from Australia, with comments on the phylogeny of the Elateriformia. Invertebr. Taxon. 2:1–53.

Lawrence J.F., Newton A.F. Jr. 1982. Evolution and classification of beetles. Annu. Rev Ecol. Syst. 13:261–290.

Li X., Ogoh K., Ohba N., Liang X., Ohmiya Y. 2007. Mitochondrial genomes of two luminous beetles, Rhagophthalmus lufengensis and R. ohbai (Arthropoda, Insecta, Coleoptera). Gene 392: 196–205.

Lockhart P.J., Steel M.A., Hendy M.D., Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605.

Loomis W.F., Smith D.W. 1990. Molecular phylogeny of *Dictyostelium discodeum* by protein sequence comparison. Proc. Natl. Acad. Sci. U.S.A. 87:9093–9097.

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

Nishiyama T., Kato M. 1999. Molecular phylogenetic analysis among bryophytes and tracheophytes based on combined data of plastid coded genes and the 18s rrna gene. Mol. Biol. Evol. 16: 1027–1036.

Nylander J.A.A. 2004. MrModeltest. Version 2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Nylander J.A.A., Wilgenbusch J.C., Warren D.L., Swofford D.L. 2008. AWTY (are we there yet?): a system for graphical exploration of mcmc convergence in Bayesian phylogenetics. Bioinformatics 24:581–583.

Philippe H., Zhou Y., Brinkmann H., Rodrigue N., Delsuc F. 2005. Heterotachy and long-branch attraction in phylogenetics. BMC Evol. Biol. 5:50.

Phillips M.J., Delsuc F., Penny D. 2004. Genome-scale phylogeny and the detection of systematic biases. Mol. Biol. Evol. 21:1455–1458.

Phillips M.J., Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Mol. Phylogenet. Evol. 28: 171–185.

Poll M. 1932. Note sur la fonction des tubes de Malpighi des Coléoptères. Bull. Ann. Soc. Ent. Belg. 72:103–109.

Ponomarenko A.G. 1969. Historical development of the Coleoptera-Archostemata. Trudy Paleontol. Inst. Akad. Nauk SSSR 125:1–240.

R Development Core Team. 2008. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: URL http://www.R-project.org.

Rodriguez-Ezpeleta N., Brinkmann H., Roure B., Lartillot N., Lang B.F., Philippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst. Biol. 56:389–399.

Rodríguez -Trelles F., Tarrío R., Ayala F.J. 1999. Molecular evolution and phylogeny of the *Drosophila saltans* species group inferred from the Xdh gene. Mol. Phylogenet. Evol. 13:110–121.

Ronquist F., Huelsenbeck J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Ruano-Rubio V., Fares M.A. 2007. Artifactual phylogenies caused by correlated distribution of substitution rates among sites and lineages: the good, the bad, and the ugly. Syst. Biol. 56:68–82.

Rubinoff D., Holland B.S. 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54:952–961.

Sheffield N.C., Song H., Cameron S.L., Whiting M.F. 2008. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. Mol. Biol. Evol. 25: 2499–2509.

Stammer H.J. 1934. Bau und bedeutung der malpighischen gefässe der coleopteren. Zoomorphology 29:196–217.

Steel M. 2005. Should phylogenetic models be trying to 'fit an elephant'? Trends Genet. 21:307–309.

Steel M., Huson D., Lockhart P.J. 2000. Invariable sites models and their use in phylogeny reconstruction. Syst. Biol. 49:225–232.

Stewart J.B., Beckenbach A.T. 2003. Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle Crioceris duodecimpunctata. Mol. Phylogenet. Evol. 26:513–526.

Swofford D.L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland (MA): Sinauer Assoicates.

Tamura K., Dudley J., Nei M., Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596.

Tarrío R., Rodríguez-Trelles F., Ayala F.J. 2000. Tree rooting with outgroups when they differ in their nucleotide composition from the ingroup: the Drosophila saltans and willistoni groups, a case study. Mol. Phylogenet. Evol. 16:344–349.

Tarrío R., Rodríguez-Trelles F., Ayala F.J. 2001. Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. Mol. Biol. Evol. 18: 1464–1473.

van den Bussche R.A., Baker R.J., Huelsenbeck J.P., Hillis D.M. 1998. Base compositional bias and phylogenetic analyses: a test of the "flying DNA" hypothesis. Mol. Phylogenet. Evol. 10:408–416.

Wachmann E. 1977. Vergleichende analyse der feinstrukturellen organisation offener rhabdome in den augen der Cucujiformia (lnsecta, Coleoptera), unter besonderer berücksichtigung der Chrysomelidae. Zoomorphology 88:95–131.

Waddell P.J., Okada N., Hasegawa M. 1999. Towards resolving the interordinal relationships of placental mammals. Syst. Biol. 48:1–5.

Yang Z. 1996. Among-site rate variation and its impact on phylogenetic analyses. Trends. Ecol. Evol. 11:367–372.

Yang Z., Roberts D. 1995. On the use of nucleic acid sequences to infer early branchings in the tree of life. Mol. Biol. Evol. 12:451–458.

Received 17 October 2008; reviews returned 7 January 2009; accepted 21 May 2009

Associate Editor: Thomas Buckley

APPENDIX TABLE A1. Nucleotide positions and anticodons (for transfer RNAs) for all genes for 5 beetle species

Gene	Strand	Anticodon	*Adelium*	*Acmaeodera*	*Apatides*	*Chauliognathus*	*Lucanus*
tRNA-Ile	+	gau	1–64 (0)[a]	1–66 (0)	1–67 (0)	1–66 (0)	1–62 (0)
tRNA-Gln	–	uug	62–130 (−3)	71–139 (4)	65–133 (−3)	64–132 (−3)	60–128 (−3)
tRNA-Met	+	cau	130–195 (−1)	139–207 (−1)	133–201 (−1)	132–197 (−1)	129–196 (0)
ND2	+		196–1204 (0)	208–1228 (0)	202–1204 (0)	198–1208 (0)	197–1210 (0)
tRNA-Trp	+	uca	1205–1268 (0)	1229–1294 (0)	1205–1270 (0)	1209–1273 (0)	1548–1614 (337)
tRNA-Cys	–	gca	1268–1329 (−1)	1287–1351 (−8)	1263–1323 (−8)	1266–1326 (−8)	1607–1667 (−8)
tRNA-Tyr	–	gua	1330–1393 (0)	1353–1416 (1)	1323–1386 (−1)	1328–1393 (1)	1667–1729 (−1)
COX1	+		1395–2928 (1)	1418–2951 (1)	1388–2918 (1)	1395–2925 (1)	1731–3261 (1)
tRNA-Leu	+	uaa	2929–2993 (0)	2952–3016 (0)	2919–2981 (0)	2926–2989 (0)	3262–3326 (0)
COX2	+		2994–3681 (0)	3017–3698 (0)	2982–3657 (0)	2991–3669 (1)	3327–4010 (0)
tRNA-Lys	+	cuu	3682–3751 (0)	3699–3768 (0)	3658–3729 (0)	3670–3741 (0)	4012–4083 (1)
tRNA-Asp	+	guc	3751–3814 (−1)	3769–3833 (0)	3745–3809 (15)	3741–3804 (−1)	4083–4144 (−1)
ATP8	+		4547–4705 (732)	3834–3992 (0)	3810–3965 (0)	3805–3960 (0)	4145–4300 (0)
ATP6	+		4705–5370 (−1)	3989–4660 (−4)	3962–4622 (−4)	3957–4628 (−4)	4297–4971 (−4)
COX3	+		5370–6153 (−1)	4660–5448 (−1)	4623–5408 (0)	4628–5412 (−1)	4964–5747 (−8)
tRNA-Gly	+	ucc	6154–6215 (0)	5454–5519 (5)	5409–5470 (0)	5413–5474 (0)	5748–5809 (0)
ND3	+		6216–6567 (0)	5520–5871 (0)	5471–5822 (0)	5475–5826 (0)	5810–6161 (0)
tRNA-Ala	+	ugc	6568–6632 (0)	5872–5936 (0)	5823–5887 (0)	5827–5891 (0)	6162–6225 (0)
tRNA-Arg	+	ucg	6632–6694 (−1)	5938–6001 (1)	5889–5954 (1)	5891–5956 (−1)	6226–6289 (0)
tRNA-Asn	+	guu	6694–6759 (−1)	6014–6079 (12)	5955–6019 (0)	5954–6018 (−3)	6289–6352 (−1)
tRNA-Ser	+	ucu	6760–6817 (0)	6080–6146 (0)	6020–6086 (0)	6019–6076 (0)	6353–6419 (0)
tRNA-Glu	+	uuc	6818–6879 (0)	6149–6211 (2)	6088–6152 (1)	6087–6149 (10)	6421–6482 (1)
tRNA-Phe	–	gaa	6878–6941 (−2)	6211–6274 (−1)	6151–6217 (−2)	6148–6213 (−2)	6481–6542 (−2)
ND5	–		6942–8652 (0)	6275–7988 (0)	6218–7928 (0)	6214–7913 (0)	6543–8256 (0)
tRNA-His	–	gug	8653–8715 (0)	7989–8051 (0)	7929–7993 (0)	7914–7977 (0)	8257–8319 (0)
ND4	–		8716–100 46 (0)	8052–9387 (0)	7994–9327 (0)	7978–9298 (0)	8320–9652 (0)
ND4L	–		10 043–10 330 (−4)	9381–9671 (−7)	9321–9602 (−7)	9295–9570 (−4)	9646–9882 (−7)
tRNA-Thr	+	ugu	10 333–10 395 (2)	9674–9738 (2)	9605–9666 (2)	9573–9634 (2)	9924–9987 (41)
tRNA-Pro	–	ugg	10 396–10 459 (0)	9738–9802 (−1)	9667–9730 (0)	9635–9698 (0)	9991–10 053 (3)
ND6	+		10 462–10 959 (2)	9804–10 310 (1)	9735–10 223 (4)	9700–10 194 (1)	10 140–10 544 (86)
CYTB	+		10 959–12 087 (−1)	10 310–11 450 (−1)	10 223–11 366 (−1)	10 194–11 292 (−1)	10 544–11 684 (−1)
tRNA-Ser	+	uga	12 088–12 154 (0)	11 451–11 517 (0)	11 367–11 433 (0)	11 293–11 357 (0)	11 685–11 749 (0)
ND1	–		12 172–13 122 (17)	11 537–12 481 (19)	11 452–12 402 (18)	11 374–12 324 (16)	11 769–12 719 (19)
tRNA-Leu	–	uag	13 123–13 184 (0)	12 482–12 546 (0)	12 404–12 467 (1)	12 326–12 389 (1)	12 720–12 782 (0)
l-rRNA	–		13 185–14 464 (0)	12 547–13 840 (0)	12 468–13 743 (0)	12 390–13 654 (0)	12 783–14 040 (0)
tRNA-Val	–	uac	14 465–14 533 (0)	13 841–13 910 (0)	13 744–13 809 (0)	13 655–13 722 (0)	14 041–14 106 (0)
s-rRNA	–		14 534–15 292 (0)	13 911–14 684 (0)	13 810–14 545 (0)	13 723–14 580 (0)	14 107–14 851 (0)
Control	n/a		15 293–16 449 (0)	14 685–16 217 (0)	14 546–16 171 (0)	14 581–14 893 (0)[b]	14 852–15 261 (0)

[a] Number in parentheses represents the number of intergenic nucleotides before the gene starts.

[b] Incomplete control region.