Systematic Entomology

Systematic Entomology (2010), **35**, 429–448 DOI: 10.1111/j.1365-3113.2009.00517.x

When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics

HOJUN SONG[1], NATHAN C. SHEFFIELD[1,2], STEPHEN L. CAMERON[1,3], KELLY B. MILLER[1,4] and MICHAEL F. WHITING[1]

[1]Department of Biology, Brigham Young University, Provo, UT, U.S.A., [2]Program in Computational Biology & Bioinformatics, Institute for Genome Sciences and Policy, Duke University, Durham, NC, U.S.A., [3]Australian National Insect Collection, CSIRO Entomology, Canberra, Australia and [4]Department of Biology, University of New Mexico, Albuquerque, NM, U.S.A.

Abstract. The ability to generate large molecular datasets for phylogenetic studies benefits biologists, but such data expansion introduces numerous analytical problems. A typical molecular phylogenetic study implicitly assumes that sequences evolve under stationary, reversible and homogeneous conditions, but this assumption is often violated in real datasets. When an analysis of large molecular datasets results in unexpected relationships, it often reflects violation of phylogenetic assumptions, rather than a correct phylogeny. Molecular evolutionary phenomena such as base compositional heterogeneity and among-site rate variation are known to affect phylogenetic inference, resulting in incorrect phylogenetic relationships. The ability of methods to overcome such bias has not been measured on real and complex datasets. We investigated how base compositional heterogeneity and among-site rate variation affect phylogenetic inference in the context of a mitochondrial genome phylogeny of the insect order Coleoptera. We show statistically that our dataset is affected by base compositional heterogeneity regardless of how the data are partitioned or recoded. Among-site rate variation is shown by comparing topologies generated using models of evolution with and without a rate variation parameter in a Bayesian framework. When compared for their effectiveness in dealing with systematic bias, standard phylogenetic methods tend to perform poorly, and parsimony without any data transformation performs worst. Two methods designed specifically to overcome systematic bias, LogDet and a Bayesian method implementing variable composition vectors, can overcome some level of base compositional heterogeneity, but are still affected by among-site rate variation. A large degree of variation in both noise and phylogenetic signal among all three codon positions is observed. We caution and argue that more data exploration is imperative, especially when many genes are included in an analysis.

Introduction

As technology develops, a very large amount of molecular data can be generated for phylogenetic studies (Hwang *et al*., 2001; Ronaghi, 2001; Yamauchi *et al*., 2004; Simison *et al*., 2006). The inclusion of multiple loci is standard practice in phylogenetics, and now 'phylogenomic studies' utilizing over 100 nuclear loci are considered the future of molecular systematics, especially for resolving deep relationships (Philippe *et al*., 2005a; Rokas *et al*., 2005; Delsuc *et al*., 2006; Dunn *et al*., 2008). Although this expansion of data is welcome, it introduces analytical and theoretical problems such as orthology assessment, alignment, conflicting phylogenetic signal, and computational issues (Phillips *et al*., 2004;

Correspondence: Hojun Song, Department of Biology, Brigham Young University, Provo, UT 84602, U.S.A. E-mail: hojun_song@byu.edu

Delsuc *et al*., 2005; Philippe & Telford, 2005; Dunn *et al*., 2008). Undoubtedly debate will continue concerning these problems, but one model system exists with which we can begin to resolve these problems effectively and manageably: mitochondrial genomics.

Mitochondrial genomes (mtgenomes) are the smallest organellar genome (Boore, 1999). A typical metazoan mtgenome is about 14 000–17 000 bp in length in the form of a circular DNA molecule and encodes 13 protein-coding genes, two ribosomal RNAs, 22 transfer RNAs, and a non-coding control region of variable length (Wolstenhome, 1992; Boore, 1999; Taanman, 1999). The small but complex structure of the mtgenome is ideal for investigating problems with multi-gene analyses. Research in analysing mtgenome data is advancing (Gibson *et al*., 2005), and several areas have attracted attention, including: exclusion of certain genes (Nardi *et al*., 2003a), reduced coding of protein-coding genes to cope with saturation of substitution in third codon positions (Nardi *et al*., 2001; Cameron *et al*., 2004; Cameron *et al*., 2007; Castro & Dowton, 2007; Fenn *et al*., 2008), extracting phylogenetic information from gene rearrangements (Boore *et al*., 1995; Boore & Brown, 1998; Boore *et al*., 1998; Dowton & Austin, 1999; Dowton *et al*., 2002) and the alignment of RNAs based on secondary structures (Macey *et al*., 1997; Gillespie *et al*., 2006; Cameron & Whiting, 2008).

Despite these advances, some issues have not been well addressed. Recent mtgenome phylogenetic studies of insects have recovered unexpected relationships, but with high branch support. For example, an unexpected relationship between Crustacea and Collembola leading to a paraphyletic Hexapoda based on mtgenome data (Nardi *et al*., 2003a) spurred controversy (Delsuc *et al*., 2003; Nardi *et al*., 2003b), and a more thorough analysis suggested that mtgenome data alone contain inadequate signal to resolve this relationship unambiguously (Cameron *et al*., 2004). Within Insecta, Stewart & Beckenbach (2003) found Orthoptera (*Locusta*) to group with Hemiptera, Lepidoptera, or to be a sister group to the rest of the holometabolous orders, depending on the analytical treatment used. Castro & Dowton (2007) found Hymenoptera to vary in position within Holometabola, depending on analysis. These studies may reflect inherent problems associated with the data themselves. For example, long-branch attraction (LBA) may cause a parsimony analysis to infer incorrect relationships (Felsenstein, 1978; Bergsten, 2005), and studies show that even model-based approaches are sensitive to LBA (Bergsten, 2005). Other types of problems, such as base compositional heterogeneity (Lake, 1994; Lockhart *et al*., 1994; Galtier & Gouy, 1995; Jermiin *et al*., 2004; Gibson *et al*., 2005; Gruber *et al*., 2007), among-site rate variation (Yang, 1996; Felsenstein, 2001; Mayrose *et al*., 2005) and heterotachy (Kolaczkowski & Thorton, 2004; Philippe *et al*., 2005b), can result in erroneous phylogenetic inference regardless of the inference method used. Evidently, potential biases must be evaluated to avoid incorrect phylogenetic conclusions.

Here we investigate how base compositional heterogeneity and among-site rate variation (ASRV) affect phylogenetic inferences. An implicit assumption often made in modern molecular phylogenetics is base compositional constancy over all lineages (the stationarity assumption: Lake, 1994; Lockhart *et al*., 1994; Galtier & Gouy, 1995). When violated, that is, when extant sequences have evolved a different base composition from the ancestral ones or when a similar base composition has evolved multiple times in distantly related taxa, and this violation is not accounted for, these taxa may group regardless of their true phylogenetic relationships (Hasegawa & Hashimoto, 1993; Lockhart *et al*., 1994; Foster & Hickey, 1999; Gibson *et al*., 2005, but see Ho & Jermiin, 2004; Jermiin *et al*., 2004 for exceptions). Base compositional heterogeneity was recognized long ago (Sueoka, 1962; Hori & Osawa, 1987; Lake, 1994; Lockhart *et al*., 1994; Dowton & Austin, 1997), and methods have been proposed to overcome the problem (Barry & Hartigan, 1987; Reeves, 1992; Lake, 1994; Lockhart *et al*., 1994; Galtier & Gouy, 1998; Foster, 2004; Gibson *et al*., 2005; Jayaswal *et al*., 2005; Gowri-Shankar & Rattray, 2006; Jayaswal *et al*., 2007), but investigation using an empirical dataset is still uncommon (but see Tarrío *et al*., 2001; Gibson *et al*., 2005; Gruber *et al*., 2007; Sheffield *et al*., 2009). The fact that ASRV exists has been known since the 1960s (Fitch & Margoliash, 1967; Yang, 1996), and many phylogenetic models can account for its presence (Felsenstein, 2001; Mayrose *et al*., 2005). However, the degree of rate variation in a given dataset is difficult to assess and it is unclear how these models can account for the bias when applied to a dataset of multiple markers.

We examine these sources of bias in the context of the mitochondrial genome phylogeny of the Coleoptera (Arthropoda: Insecta). Beetles constitute one of the largest radiations in the Tree of Life, containing more than 350 000 species, or *c*. 25% of all described species (Crowson, 1960; Lawrence & Newton, 1982; Hunt *et al*., 2007). Despite this diversity, there are few complete beetle mtgenomes compared with the situation for other holometabolous insects such as Diptera (Cameron *et al*., 2007), rendering Coleoptera one of the least-studied major insect orders in terms of mitochondrial genomics. By combining seven new beetle mtgenomes generated in this study with mtgenomes and data generated as a part of an NSF-funded Beetle Tree of Life Project, we present a preliminary mtgenome phylogeny of Coleoptera based on all protein-coding genes. Our analysis includes all four suborders, thereby providing a unique opportunity to study subordinal-level relationships. We focus particularly on identifying base compositional heterogeneity in different codon positions and explore available methods to overcome systematics bias in molecular data.

Materials and methods

Taxon sampling

A total of 31 taxa were included in this study, including 24 coleopteran ingroup and seven outgroup taxa from the orders Hemiptera, Lepidoptera, Diptera and Hymenoptera (Table S1). Our outgroup taxon sampling was limited by the availability of data and did not include representatives of

the other holometabolous insect orders. However, because the goal of this study was to investigate phylogenetic signal in mtgenome data and not necessarily to establish a phylogenetic position of Coleoptera within Holometabola, we felt that this outgroup sampling was appropriate. Our ingroup taxon sampling included seven new beetle mtgenomes: *Calosoma* sp. (Carabidae, GenBank accession: GU176340), *Cucujus clavipes* (Cucujidae, GU176341), *Euspilotus scissus* (Histeridae, GU176344), *Naupactus xanthographus* (Curculionidae, GU176345), *Necrophila americana* (Silphidae, GU176343), *Sphenophorus* sp. (Curculionidae, GU176342) and *Tropistenus* sp. (Hydrophilidae, GU176339). The remaining coleopteran and outgroup taxa were obtained from previous studies (Table S1). The ingroup sampling included all four coleopteran suborders and 22 families, which included 11 polyphagan superfamilies, thereby representing the most comprehensive coleopteran mitochondrial phylogenomic dataset to date. A hemipteran, *Philaenus spumarinus*, was used as the root. All specimens were collected and stored in 100% ethanol, and the voucher specimens and genomic extracts were stored at −80°C in the Insect Genomic Collection of the Department of Biology and Monte L. Bean Museum, Brigham Young University. Collecting and voucher information and GenBank accession numbers for the taxa used in this study are shown in Table S1.

Mitochondrial genome sequencing

Total genomic DNA was extracted using the DNeasy Tissue kit (Qiagen). We used two dissection techniques, depending on the size of the individual species. For large specimens, we dissected thoracic muscle tissues for extraction. For small specimens, we dissected the entire abdominal segment from the whole body to avoid possible contamination from gut content and to retain taxonomically important genital structures as vouchers. We used the remaining body without the abdomen for extraction. Mtgenome sequencing was accomplished by primer walking. First, we amplified relatively large fragments (3–4 Kb) using conserved primers to generate *cox1-cox3*, *nad4-cytB* and 16*S*-12*S* regions. Both ends of these amplicons were sequenced in order to design species-specific primers to amplify the remainder of the mtgenome to generate *cox3-nad4*, *cytB-16S* and *12S-cox1* regions. A total of 232 primers were designed in this study, and species-specific primers are available as Supplementary Information (File S1). Specific sequencing primers were designed using the already sequenced portions of the mtgenome and used to sequence the remaining regions of the mtgenome. In this manner, the full, double-stranded sequence of the entire mtgenome was determined. Long polymerase chain reactions (PCRs) were performed using the Elongase enzyme (Invitrogen) with the following cycling conditions: 92°C for 2 min; 40 cycles of 92°C for 30 s, 50°C for 30 s, 60°C for 12 min; with a final elongation step of 60°C for 20 min. Short PCRs were performed when necessary using the Elongase enzyme (Invitrogen) with the following cycling conditions: 95°C for 12 min; 40 cycles of 94°C for 1 min, 40°C for 1 min, 72°C for 1 min; with a final elongation step of 72° for 7 min. Sequencing was performed using ABI BigDye 3 dye terminator chemistry and then fractionated on the ABI 3770 or ABI 3740 capillary sequencer. Sequencing PCR conditions were as follows: 96°C for 1 min; 25 cycles of 96°C for 10 s, 50°C for 5 s, 60°C for 75 s. Sequencing of the ambiguous AT-rich control region was accomplished by cloning using the TOPO-TA Cloning kit (Invitrogen) when necessary.

Genome annotation and alignment

Raw sequence files were proofread and aligned into contigs in Sequencher 4.6 (GeneCodes Corporation). We annotated the genomes with tRNAscan-SE (Eddy & Durbin, 1994) and mosas (freely available at http://mosas.byu.edu), an online sequence management tool designed specifically to deal with mtgenome data. tRNAs detected by the software were inspected manually to ensure accuracy. After mosas reported general locations for genes based on a blast search, we identified start and stop codons to complete the annotation. The end of the small subunit rRNA (12*S*) was assigned by alignment with other beetle genomes using a conserved tag (5′-ARAATWAAACTHTNH-3′). When the tag did not match completely, the end position of the gene was determined by visual inspection. Annotated mtgenomes available from GenBank were imported into the mosas database in order to create final datasets.

We created a total of 13 datasets, each representing an individual protein-coding gene. Each gene partition was aligned separately based on the conservation of amino acid sequences. First, each gene partition was translated into corresponding amino acids using mega 4 (Tamura *et al*., 2007). The translated amino acid sequences were aligned in muscle (Edgar, 2004) using default parameters. The aligned amino acid sequences were then used as a scaffold for constructing the corresponding nucleotide sequence alignment using RevTrans 1.4 (Wernersson & Pedersen, 2003). The resulting nucleotide alignments were checked thoroughly for possible errors during translation in mega 4 (Tamura *et al*., 2007). The 13 aligned datasets were concatenated in MacClade 4 (Maddison & Maddison, 2005) to compile a final data matrix of a total of 11 904 aligned nucleotides (nt123).

Testing the level of base compositional heterogeneity

We employed two independent methods to determine any effect of base compositional heterogeneity. First, we calculated the disparity index (I_D) (Kumar & Gadagkar, 2001) for all 13 genes together and pairwise. I_D measures the observed differences in substitution pattern for a pair of sequences, thereby indirectly measuring the level of base compositional heterogeneity. We tested the homogeneity of substitution pattern (I_D-test) using a Monte Carlo method with 1000 replications 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 dataset (complete deletion option). Jermiin *et al*. (2004, 2009), however, argued that this method should be used with caution because homology of sites is not considered appropriately while calculating I_D. Thus, we used a second method known as the matched-pairs test of symmetry (Ababneh *et al*., 2006) as implemented in SeqVis (Ho *et al*., 2006), which examines individual sites in a given alignment (Jermiin *et al*., 2004) and tests against the null hypothesis that a pair of sequences has evolved under the same conditions.

We tested the assumptions of stationary, reversible and homogeneous nucleotide evolution for each codon position using the matched-pairs tests of symmetry, marginal symmetry, and internal symmetry (Ababneh *et al*., 2006). The matched-pairs test of marginal symmetry assesses whether a pair of sequences is likely to have evolved under the same stationary conditions. The matched-pairs test of internal symmetry tests whether a pair of sequences is likely to have evolved under the same conditions beyond that due to evolution under stationary or non-stationary conditions (Ababneh *et al*., 2006; Jermiin *et al*., 2009). These three matched-pairs tests allow identification of the evolutionary process responsible for the bias. Furthermore, we examined the effect of base compositional heterogeneity at the level of amino acid sequences using the matched-pairs test of symmetry. These analyses were performed using the alpha version of seqvis (L. Jermiin, unpublished data).

Testing the effectiveness of phylogenetic methods in coping with compositional heterogeneity

First, we tested the effect of standard phylogenetic inference methods, analysing the concatenated dataset in both maximum parsimony (MP) and Bayesian (BA) frameworks (with gaps scored as missing). In the MP analyses, we used tnt (Goloboff *et al*., 2003) to perform a combination of sectorial search, drifting, tree fusing (Goloboff, 1999) and ratchet (Nixon, 1999), to find the most parsimonious trees. We calculated non-parametric bootstrap values using 5000 replicates with 100 tree bisection–reconnection (TBR) random additions per replicate and Bremer support (Bremer, 1994), both in tnt. Relative contributions of individual partitions to the combined dataset were calculated using the partitioned Bremer support (PBS) (Baker & DeSalle, 1997) in tnt (script written by Peña *et al*., 2006). Although the dataset appeared to violate the assumptions of stationarity, reversibility and homogeneity, we tried typical model-based analyses to assess the performance of such methods under assumption violations. In the BA analyses, we used a different, unlinked model for each gene, as recommended by MrModelTest (Nylander, 2004) (see Table S2). Using MrBayes 3.1.1 (Ronquist & Huelsenbeck, 2003), we ran four runs with four chains each for 30 million generations, sampling every 1000 generations. We plotted the likelihood trace for each run to assess convergence, and discarded an average of 25% of each run as burn-in. These analyses were run on the Brigham Young University Life Science's Computational Cluster (http://lsbeast.byu.edu).

Second, we created five additional datasets representing: (i) amino acid (aa) sequences, (ii) first codon position only (nt1), (iii) second codon position only (nt2), (iv) third codon position only (nt3), and (v) first and second codon positions only (nt12). We analysed these datasets in both MP and BA frameworks as described above. Based on the results of the I_D-test and the matched-pairs tests of symmetry, marginal symmetry and internal symmetry, we inferred a phylogeny from the dataset that appeared least likely to violate the assumptions of the models of sequence evolution, which we used as a logical reference topology.

Third, we explored two different methods that have been shown to overcome base compositional heterogeneity: (i) LogDet transformation (Lockhart *et al*., 1994), and (ii) allowing variable compositional vectors during tree construction implemented in the alpha version of phase 2.1 (Gowri-Shankar & Rattray, 2007). A neighbour-joining analysis after LogDet transformation was performed in paup* 4b10 (Swofford, 2002). Because previous studies demonstrated that LogDet transformation could be affected by inclusion of invariable sites (Steel *et al*., 2000), we transformed the dataset after estimating the proportion of invariable sites using maximum likelihood in paup* 4b10 [under General Time Reversible (GTR) model: pinvar = 0.206364]. The phase analysis used a heterogeneous model of sequence evolution with discrete gamma model in the alpha version of phase 2.1, for 6 million generations sampling every 500 generations.

Testing the effect of among-site rate variation

We tested whether ASRV affects the phylogenetic reconstruction. To minimize the confounding effects from other potential biases, we chose the nucleotide dataset least likely to violate the phylogenetic assumptions (nt2 dataset, see below) in studying the effect of ASRV. We examined this in a BA framework by comparing resulting topologies, posterior probabilities and the Bayes factors among the models that assume equal rates across sites (GTR), gamma-shaped rate variation across sites (GTR + G) and gamma-shaped rate variation across sites with a proportion of sites invariable (GTR + G + I). For each analysis, we ran four separate runs with four chains per run for a total of 20 million generations per run with sampling every 1000 generations in MrBayes 3.1.1 (Ronquist & Huelsenbeck, 2003).

Results

Mtgenome description

All coding regions of the mtgenome were obtained from seven newly sequenced beetle genomes (Table S3). We sequenced the complete control region for five species, but were unable to sequence through the control region of the

Table 1. Disparity index values calculated from pairwise comparisons among coleopteran species included in this study.

	nt123		nt12		nt1		nt2		nt3	
Taxa	Sum	Mean	Sum	Mean	Sum	Mean	Sum	Mean	Sum	Mean
Sphaerius	532.47	22.19	120.92	5.04	140.59	5.86	16.39	0.68	388.77	16.20
Macrogyrus	156.28	6.51	37.61	1.57	48.17	2.01	3.46	0.14	311.68	12.99
Calosoma	207.65	8.65	43.89	1.83	55.85	2.33	4.48	0.19	257.48	10.73
Trachypachus	357.70	14.90	67.84	2.83	98.90	4.12	4.90	0.20	381.90	15.91
Tetraphalerus	1055.24	43.97	235.89	9.83	242.21	10.09	45.79	1.91	1225.14	51.05
Cyphon	157.97	6.58	30.36	1.27	35.88	1.49	3.88	0.16	218.15	9.09
Pyrocoelia	199.41	8.31	54.94	2.29	80.02	3.33	6.97	0.29	361.32	15.05
Chauliognathus	198.67	8.28	40.81	1.70	39.72	1.65	10.47	0.44	418.01	17.42
Pyrophorus	643.25	26.80	149.60	6.23	202.83	8.45	9.96	0.41	723.20	30.13
Rhagophthalmus	341.04	14.21	54.70	2.28	69.97	2.92	6.23	0.26	474.61	19.78
Mordella	258.86	10.79	65.78	2.74	89.58	3.73	5.13	0.21	310.57	12.94
Tribolium	436.94	18.21	83.49	3.48	107.97	4.50	6.85	0.29	548.97	22.87
Adelium	308.36	12.85	53.02	2.21	72.09	3.00	4.11	0.17	367.22	15.30
Euspilotus	190.57	7.94	42.56	1.77	43.06	1.79	9.33	0.39	322.56	13.44
Rhopaea	143.86	5.99	31.68	1.32	35.86	1.49	6.11	0.25	199.87	8.33
Naupactus	177.69	7.40	40.61	1.69	53.77	2.24	5.10	0.21	228.40	9.52
Sphenophorus	163.63	6.82	31.67	1.32	42.94	1.79	3.34	0.14	227.26	9.47
Tropisternus	162.70	6.78	46.16	1.92	58.49	2.44	6.80	0.28	225.54	9.40
Necrophila	150.10	6.25	38.48	1.60	40.95	1.71	6.75	0.28	172.61	7.19
Chaetosoma	267.45	11.14	82.45	3.44	92.80	3.87	10.87	0.45	259.24	10.80
Anoplophora	238.00	9.92	37.55	1.56	42.81	1.78	5.31	0.22	340.55	14.19
Cucujus	149.51	6.23	32.57	1.36	37.43	1.56	5.46	0.23	199.74	8.32
Priasilpha	145.68	6.07	38.24	1.59	44.20	1.84	5.67	0.24	222.77	9.28
Crioceris	159.91	6.66	34.16	1.42	41.59	1.73	4.66	0.19	194.39	8.10

'Sum' indicates the sum of all I_D calculated for a particular taxon, and 'Mean' indicates the average of I_D for that taxon. Taxa with high I_D-values have a significantly different base compositional bias from the rest of the species. Numeric values shown next to nt (i.e. nt1) represent codon positions. The values shown here highlight the level of base compositional heterogeneity in our dataset.

genomes from the remaining three species owing to the high AT content and multiple stretches of poly-A and poly-T, which interfered with the accuracy of dye terminator sequencing. The total lengths of the coding region varied from 14 487 bp (*Euspilotus*) to 14 980 bp (*Calosoma*). Most species sequenced in this study had the ancestral insect gene composition and arrangement. However, we found three exceptions to the general pattern. We were unable to locate tRNA-Isoleucine in *Naupactus* and *Sphenophorus*, suggesting that this tRNA might be missing in these species. Moreover, we found a tRNA rearrangement in *Naupactus*. Typically, the ancestral order of tRNAs between the *nad3* and *nad5* genes is ARNSEF, but *Naupactus* exhibited RANSEF, which represents a rare case of gene rearrangement in Coleoptera. The anticodons for tRNA were conserved completely in all tRNAs except for tRNA-Serine in *Calosoma*, which had a GCU anticodon in contrast to all the other beetles, which have a UCU anticodon for this tRNA.

Base compositional heterogeneity

A total of 276 pairwise comparisons were made among beetles to calculate the disparity index (I_D), and closely related taxa generally had a lower I_D than distantly related taxa. For instance, the I_D between two curculionid species, *Naupactus* and *Sphenophorus*, was 0.412, whereas that between *Naupactus* and a distantly related myxophagan *Sphaerius* was 9.292. Within Coleoptera, *Tetraphalerus* had the most divergent base composition (mean $I_D = 43.97$), followed by *Pyrophorus*, *Sphaerius* and *Tribolium* (Table 1). Fifteen of 24 beetle species exhibited a similar base composition, with a mean I_D between 6 and 10, suggesting that the overall dataset exhibited a high level of base compositional heterogeneity. The I_D-values calculated from individual codon partitions suggested that the level of base compositional heterogeneity was the lowest in nt2, followed by nt1 and nt3 (Table 1). As could be expected, the amount of interspecific variation in AT% (a proportion of nucleotides A and T in a given sequence) followed the same pattern (Fig. 1). Interestingly, the taxa that exhibited high levels of compositional heterogeneity compared with other beetle taxa included in the analysis, such as *Tetraphalerus*, had consistently high I_D-values across all codon positions.

Results from the matched-pairs test of symmetry based on 276 pairwise comparisons were congruent with the I_D-test (Table 2), and more than 93% of the comparisons resulted in a P-value < 0.05, suggesting that most sequence pairs did not evolve under the same conditions (Jermiin *et al*., 2004; Ababneh *et al*., 2006). Additional results from the matched-pairs tests of symmetry based on the aa and individual codon position datasets suggested that stationarity, reversibility and homogeneity assumptions were not met for each dataset (Table 2). Matched-pairs tests of marginal symmetry showed

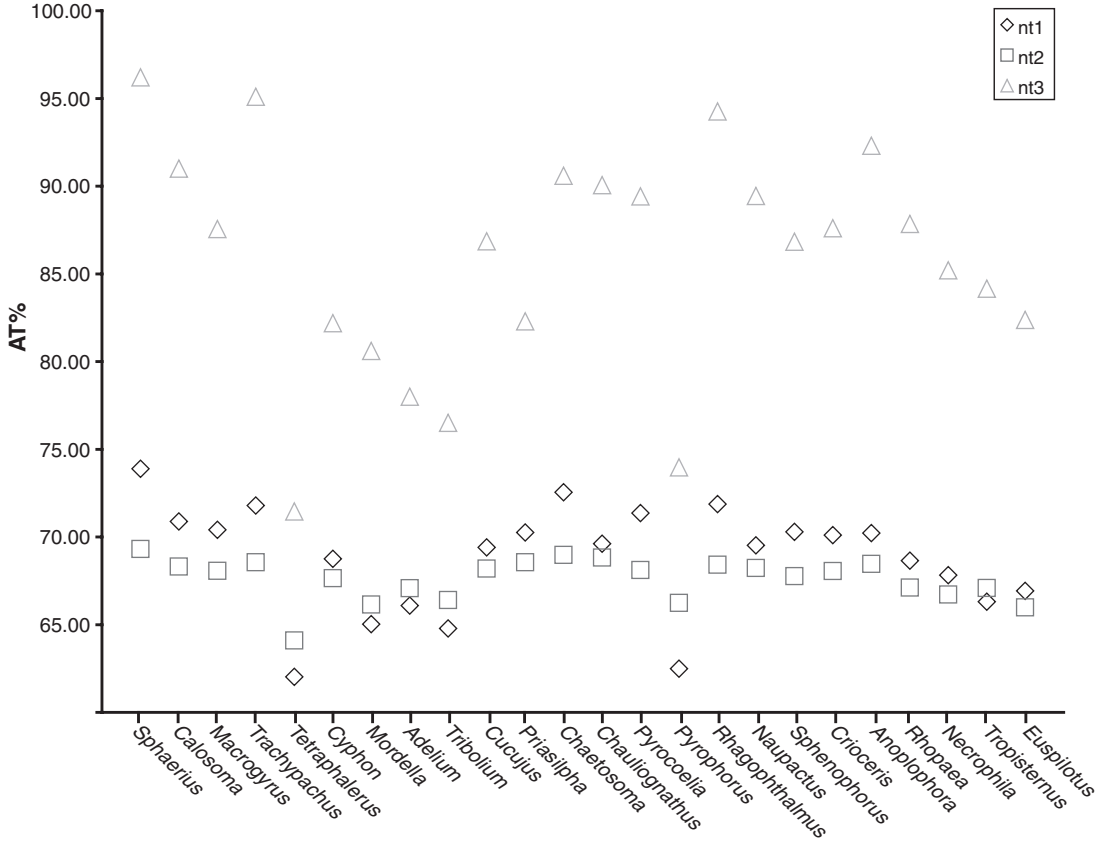

Fig. 1. Variation in base composition in different codon positions across Coleoptera. AT% of individual codon positions is plotted against the ingroup species. The variation in base composition is the highest in the nt3 position, followed by the nt1 and nt2. Different species exhibit different base compositions in different codon positions.

Table 2. Results from matched-pairs tests of symmetry based on coleopteran species included in this study.

P-value	nt123	nt12	nt1	nt2	nt3	aa
<0.05:	259 (0.938)	217 (0.786)	217 (0.786)	135 (0.489)	255 (0.924)	144 (0.522)
<0.01:	249 (0.902)	189 (0.685)	195 (0.707)	93 (0.337)	248 (0.899)	85 (0.308)
<0.005:	248 (0.899)	184 (0.667)	188 (0.681)	86 (0.312)	241 (0.873)	71 (0.257)
<0.001:	238 (0.862)	169 (0.612)	167 (0.605)	61 (0.221)	227 (0.822)	53 (0.192)
<0.0005:	235 (0.851)	164 (0.594)	161 (0.583)	50 (0.181)	225 (0.815)	49 (0.178)
<0.0001:	228 (0.826)	152 (0.551)	146 (0.529)	45 (0.163)	219 (0.793)	38 (0.138)
<0.00005:	223 (0.808)	149 (0.540)	144 (0.522)	41 (0.149)	219 (0.793)	31 (0.112)
<0.00001:	214 (0.775)	142 (0.514)	133 (0.482)	33 (0.120)	212 (0.768)	22 (0.080)

Summary of the distribution of *P*-values out of a total of 276 pairwise comparisons made. The numbers in parentheses describe the proportion of pairwise comparisons that have statistically significant values at each *P*-value level. The majority of tests resulted in highly significant *P*-values, implying that there is a high level of statistical support for rejecting the null hypothesis of evolution under stationary, reversible and homogeneous conditions throughout the datasets. Numeric values shown next to nt (i.e. nt1) represent codon positions, and aa represents the amino acid sequence dataset.

that the evolutionary processes at all three codon positions were unlikely to be stationary, and therefore could not be both reversible and homogeneous (Table 3). The matched-pairs tests of internal symmetry showed that the nt1 and nt3 datasets were likely to have evolved in ways unattributable only to violation of the stationary assumption, whereas the nt2 dataset violated the null hypothesis only marginally (Table 3).

Based on all these tests, we determined that the dataset that was least likely to violate the phylogenetic assumptions was the aa dataset. Among nucleotide sequences, the ranked order of the datasets from best to worst was: nt2 > nt1 > nt12 > nt3 > nt123. Therefore, we use the resulting phylogeny from amino acid sequences as a reference topology in the remainder of this study (Fig. 2).

Table 3. Results from matched-pairs tests of marginal symmetry and internal symmetry based on coleopteran species included in this study.

	Marginal symmetry			Internal symmetry		
P-value	nt1	nt2	nt3	nt1	nt2	nt3
<0.05:	208 (0.754)	126 (0.457)	249 (0.902)	60 (0.217)	24 (0.087)	104 (0.377)
<0.01:	187 (0.678)	92 (0.333)	236 (0.855)	29 (0.105)	6 (0.022)	77 (0.279)
<0.005:	181 (0.656)	80 (0.290)	229 (0.830)	19 (0.069)	5 (0.018)	72 (0.261)
<0.001:	166 (0.601)	55 (0.199)	222 (0.804)	10 (0.036)	0 (0.000)	55 (0.199)
<0.0005:	157 (0.569)	51 (0.185)	217 (0.786)	4 (0.014)	0 (0.000)	48 (0.174)
<0.0001:	143 (0.518)	36 (0.130)	206 (0.746)	3 (0.011)	0 (0.000)	35 (0.127)

Summary of the distribution of *P*-values out of a total of 276 pairwise comparisons made. The numbers in parentheses describe the proportion of pairwise comparisons that have statistically significant values at each *P*-value level. The matched-pairs test of marginal symmetry tests against the null hypothesis that a pair of sequences is likely to have evolved under the same stationary conditions. The results show that the majority of sequences in the nt3 and nt1 dataset do not evolve under stationary conditions, and that about half of sequence pairs in the nt2 dataset evolve under stationary conditions. The matched-pairs test of internal symmetry tests against the null hypothesis that a pair of sequences is likely to have evolved under the same conditions beyond what may be due to evolution under stationary or non-stationary conditions, and we fail to reject the hypothesis in the nt3 and nt1, and only marginally in the nt2 dataset. Numeric values shown next to nt (i.e. nt1) represent codon positions, and aa represents the amino acid sequence dataset.

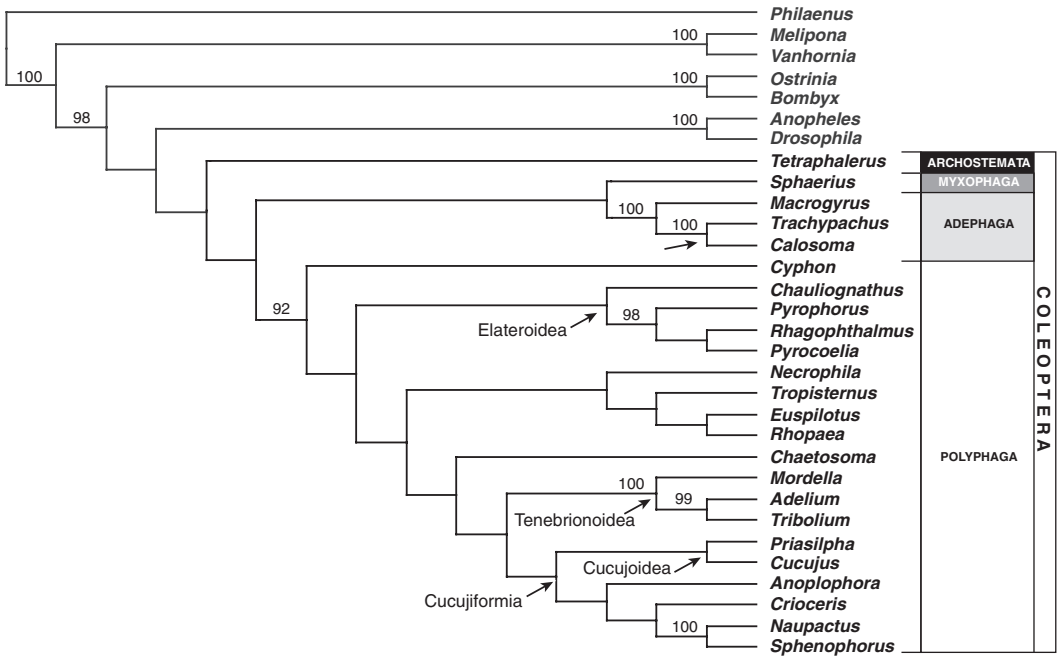

Fig. 2. A reference phylogeny based on the amino acid sequences. The aa dataset was determined to be the least likely to violate the phylogenetic assumptions by the matched-pairs test of symmetry. Shown here is the single most parsimonious tree recovered after translating nucleotide characters to amino acid sequences (L = 3643, CI = 0.43, RI = 0.32). Numbers above the branches indicate bootstrap support values (only values > 80 are shown).

We determined that *Tetraphalerus* and *Pyrophorus* were the two most heterogeneous taxa in our study, and we examined whether removal of these two taxa alleviated the level of base compositional heterogeneity. Thus, we created three additional reduced datasets, one without *Tetraphalerus*, one without *Pyrophorus*, and one without both species, and performed the matched-pairs tests of symmetry on these datasets. In each of these three datasets, we found that the majority of sequence pairs still did not evolve under the same conditions, and that the level of base compositional heterogeneity was comparable to that in the original dataset (Table 4).

Phylogenetic estimate

The combined protein-coding gene (nt123) dataset contained 7424 parsimony-informative characters, and the MP analysis found a single most parsimonious tree (Fig. 3A). The partitioned Bremer support (PBS) values showed that the nt1 and nt2 datasets generally contributed more phylogenetic

Table 4. Effect of removing a taxon affected by base compositional heterogeneity, examined in the context of the matched-pairs tests of symmetry.

P-value	All beetles	No *Tetraphalerus*	No *Pyrophorus*	No *Tetraphalerus* and *Pyrophorus*
<0.05:	259 (0.938)	236 (0.933)	236 (0.933)	214 (0.926)
<0.01:	249 (0.902)	226 (0.893)	226 (0.893)	204 (0.883)
<0.005:	248 (0.899)	225 (0.889)	225 (0.889)	203 (0.879)
<0.001:	238 (0.862)	215 (0.850)	215 (0.850)	193 (0.835)
<0.0005:	235 (0.851)	212 (0.838)	212 (0.838)	190 (0.823)
<0.0001:	228 (0.826)	205 (0.810)	205 (0.810)	183 (0.792)
<0.00005:	223 (0.808)	200 (0.791)	200 (0.791)	178 (0.771)

The second column (All Beetles) is a summary of the distribution of *P*-values out of a total of 276 pairwise comparisons made. The numbers in parentheses describe the proportion of pairwise comparisons that have statistically significant values at each *P*-value level. The third column is a summary of the distribution of *P*-values after removing *Tetraphalerus*, which is the most heterogeneous taxon in the present study, and the fourth column is a summary after removing *Pyrophorus*, which is the second most heterogeneous taxon. The last column is a summary of the distribution of *P*-values after removing both *Tetraphalerus* and *Pyrophorous*. The analyses show that over 92% of the pairwise comparisons even after removing two most heterogeneous taxa have statistically significant values at $\alpha < 0.05$, suggesting that the majority of the data still do not evolve under stationary, reversible and homogeneous conditions of sequence evolution.

signal than the nt3 dataset (Fig. 3A). Some relationships were supported positively by all three codon positions, but other relationships were unsupported by at least one codon position (Fig. 3A). Using a hemipteran as the root, the analysis recovered (Hymenoptera (Lepidoptera (Diptera + Coleoptera))) at the ordinal level, but this grouping is likely to be artifactual because other holometabolous lineages were not included. Within the ingroup, the analysis recovered Coleoptera as monophyletic as well as monophyly of the suborder Adephaga. Myxophaga (*Sphaerius*) was sister to Adephaga. Within Adephaga, *Calosoma* (Carabidae) and *Trachypachus* (Trachypachidae) formed a sister clade with a basal *Macrogyrus* (Gyrinidae). Polyphaga was not monophyletic because the sole representative of Archostemata included in the analysis (*Tetraphalerus*) was sister to an elaterid *Pyrophorus*. A scirtid *Cyphon* was placed at the base of Polyphaga, although it formed a clade with (*Tetraphalerus* + *Pyrophorus*). Most of the remaining polyphagan superfamilies represented by multiple taxa were monophyletic, including Tenebrionoidea, Cucujoidea, Curculionoidea and Chrysomeloidea, and the infraorder Cucujiformia was recovered. Support values generally were high for more recent bifurcations, but low for earlier bifurcations within Coleoptera (Fig. 3A). The mixed model BA analysis recovered identical outgroup relationships to the MP tree, but resulted in different ingroup relationships from the MP tree (Fig. 3B). Coleoptera was monophyletic with the (Myxophaga + Adephaga) clade at the base. Archostemata was sister to a monophyletic Polyphaga. *Cyphon* was sister to the rest of Polyphaga, and all four elateroid representatives formed a strong monophyletic group within Polyphaga. Other ingroup relationships were similar to those of the MP analysis.

The aa dataset under parsimony resulted in a single most parsimonious tree, our reference phylogeny [Fig. 2, length (L) = 3643, consistency index (CI) = 0.43, retention index (RI) = 0.32]. The topology was quite different from that of the combined protein-coding gene dataset in which Polyphaga was monophyletic, as were the other suborders. The resulting subordinal relationship within Coleoptera was (Archostemata (Polyphaga (Myxophaga + Adephaga))). Within the monophyletic Polyphaga, *Cyphon* was sister to the rest of the species, and each superfamily represented by multiple taxa was monophyletic.

Analyses based on individual codon partitions were revealing about the phylogenetic signal in each partition. The MP analysis of the nt1 dataset did not recover a monophyletic Coleoptera because of the (Diptera + Lepidoptera) clade nested within beetles. *Tetraphalerus* and *Pyrophorus* had the lower AT% in the nt1 dataset, but these two did not group (Figure S1A); instead, *Tetraphalerus* grouped with a histerid *Euspilotus*, and *Pyrophorus* correctly grouped with other elateroid species. The BA analysis of the nt1 dataset recovered a different topology from the MP tree, and likewise failed to recover the monophyly of Coleoptera (Figure S1B). The nt2 analyses resulted in topologies most congruent to the reference phylogeny. The MP analysis recovered the monophyletic Coleoptera and each of the suborders (Figure S2A). The subordinal relationships were ((Myxophaga + Adephaga) + (Archostemata + Polyphaga)). The BA analysis recovered the monophyletic Coleoptera, but with different subordinal relationships: (Archostemata (Polyphaga (Myxophaga + Adephaga))) (Figure S2B). The nt3 analyses resulted in the relationships that were least similar to the reference phylogeny. In the MP analysis, the monophyly of Coleoptera was not supported because the (Diptera + Lepidoptera) clade was nested deep within beetles (Figure S3A). A sister relationship between Myxophaga and Adephaga was strongly supported, but Polyphaga was largely paraphyletic. A strong sister relationship between *Tetraphalerus* and *Pyrophorus* was again found. Similarly, the BA analysis resulted in paraphyletic Coleoptera and Polyphaga, strong sister relationships between Myxophaga and Adephaga, and *Tetraphalerus* and *Pyrophorus* (Figure S3B). The MP analysis of the nt12 dataset recovered a monophyly of Coleoptera, but not of Polyphaga, because *Cyphon* and *Tetraphalerus* formed a clade at the base of Polyphaga (Figure S4A). On the other hand, the BA analysis of the same dataset did recover monophyletic Coleoptera and all four suborders, although the placement of Archostemata was different from in the reference phylogeny (Figure S4B).

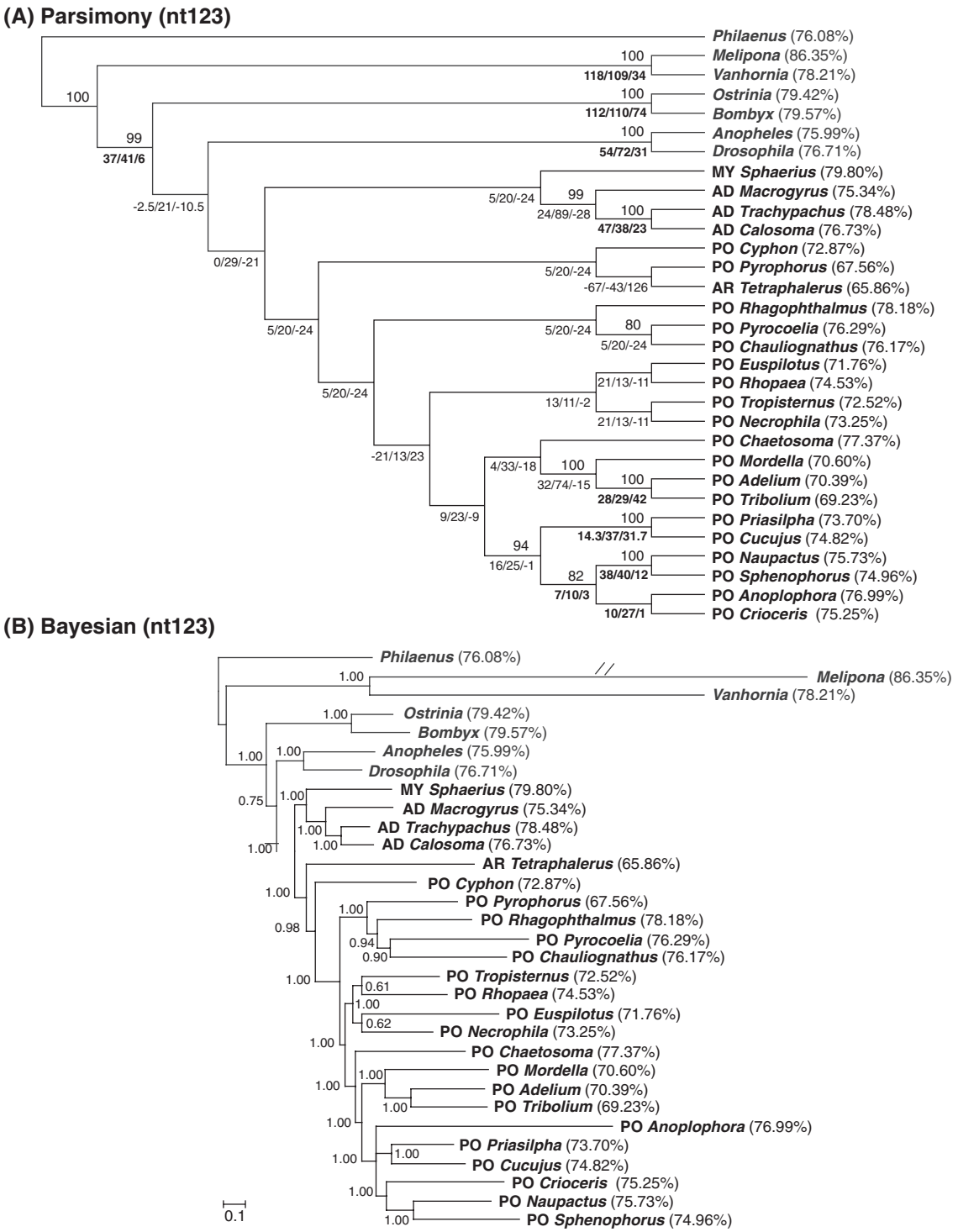

Fig. 3. Phylogeny of Coleoptera based on 13 protein-coding genes (nt123). (A) A single most parsimonious tree (L = 58736, CI = 0.29, RI = 0.26). Numbers above the branches indicate bootstrap support values (only values > 80 are shown) and numbers below the branches indicate partitioned Bremer support of individual codon positions (first/second/third). PBS values in bold indicate the clades that are positively supported by all three codon positions. (B) A Bayesian tree based on mixed models of individual gene partitions. Numbers above the branches indicate posterior probability values. Numbers in parentheses next to terminal indicate the total AT% from the nt123 dataset. The code before the taxon name indicates the taxonomic grouping: MY, Myxophaga; AD, Adephaga; AR, Archostemata; PO, Polyphaga.

Effect of phylogenetic methods in coping with compositional heterogeneity

When we mapped base composition onto the MP topology of the nt123 phylogeny (Fig. 3A), it became apparent that two distantly related taxa, an archostematan *Tetraphalerus* and an elaterid *Pyrophorus*, were grouped owing to shared low AT bias, which was observed in our previous study based on smaller taxon sampling (Sheffield *et al*., 2009). The neighbor-joining analysis after LogDet transformation (Lockhart *et al*., 1994) resulted in paraphyletic Coleoptera and paraphyletic Polyphaga (Fig. 4A). This treatment recovered a strong clade of (Myxophaga + Adephaga), but this clade was sister to a (Lepidoptera + Diptera) clade. A scirtid *Cyphon* formed a polytomy with the aforementioned clade and the remaining Coleoptera, rendering Polyphaga paraphyletic. *Tetraphalerus* also formed a polytomy with Elateroidea and the rest of Polyphaga. LogDet transformation after removing invariable sites recovered different relationships within Polyphaga, and *Cyphon* was grouped correctly with other coleopterans, although the position of *Tetraphalerus* was unresolved (Fig. 4B). The PHASE analysis allowing variable compositional vectors resulted in a topology similar to that from the BA analysis of the nt123 dataset (Fig. 5), in that the monophyletic Coleoptera was recovered and the subordinal relationships were ((Myxophaga + Adephaga) + (Archostemata + Polyphaga)). Within Polyphaga, *Cyphon* was sister to the rest, and the major superfamilies were all monophyletic. Neither LogDet nor PHASE analyses recovered a topology identical to the reference phylogeny.

Among-site rate variation

The least likely nucleotide dataset to violate the phylogenetic assumptions was the nt2 dataset. Because this dataset was relatively free from the bias originating from the base compositional heterogeneity, we examined the effect of another source of bias (ASRV) with minimal confounding factors. By comparing the topologies resulting from the BA analyses with or without the rate variation parameter, we could identify the taxa affected by ASRV indirectly. When assuming equal rate among sites (GTR), an archostematan *Tetraphalerus* was sister to Polyphaga, and, within the monophyletic Elateroidea, the ((*Pyrocoeilia* + *Chauliognathus*) + (*Pyrophorus* + *Rhagophthalmus*)) clade was recovered (Fig. 6A). When assuming gamma-shaped rate variation across sites (GTR + G or GTR + G + I), *Tetraphalerus* was placed at the base of Coleoptera, and, within Elateroidea, the (*Chauliognathus* (*Pyrophorus* (*Pyrocoelia* + *Rhagophthalmus*))) clade was recovered (Fig. 6B, C). The posterior probabilities of the clades affected by changes in model parameters were mostly robust. We calculated the harmonic mean estimator using MRBAYES in order to compare between models in the context of Bayes factors (Table 5). Based on the interpretation recommended by Kass & Raftery (1995), we determined that the Bayes factor analysis found decisive evidence against the model assuming equal rate among sites (GTR) when compared to either GTR + G or GTR + G + I, suggesting that implementing gamma-shaped rate variation among sites was a better model for the data. The implementation of the invariable sites (GTR + G + I) was a stronger model than GTR + G.

Discussion

Mtgenome phylogeny of Coleoptera

Our study investigates the issues of base compositional heterogeneity and ASRV evident in our dataset, and how best to overcome such bias. Different treatments resulted in a wide range of different relationships within Coleoptera, suggesting that our current taxon sampling may be insufficient for a thorough understanding of beetle evolution. Although all our datasets appeared to be biased, the aa dataset was the least likely to violate the phylogenetic assumptions, so we used it as a reference phylogeny (Fig. 2). Our taxon sampling includes all four beetle suborders as well as some of the major superfamilies for which the mtgenome data find some interesting relationships, which we elaborate here briefly.

The subordinal-level relationships of Coleoptera remain contentious (Caterino *et al*., 2002) and several hypotheses have been proposed. One often-cited hypothesis based on morphological characters is (Archostemata (Adephaga (Myxophaga + Polyphaga))) (Crowson, 1960; Beutel & Haas, 2000). Molecular phylogeny based on the 18S ribosomal gene (Caterino *et al*., 2002) proposed (Archostemata (Myxophaga (Adephaga + Polyphaga))), but 66 expressed sequence tag (EST) sequences (Hughes *et al*., 2006) recovered (Archostemata (paraphyletic Adephaga (Myxophaga + Polyphaga))). The most comprehensive molecular phylogeny of Coleoptera (Hunt *et al*., 2007) found the relationship ((Myxophaga + Archostemata) + (Adephaga + Polyphaga)). Our findings (Fig. 2) suggest yet another novel relationship, namely (Archostemata ((Myxophaga + Adephaga) + Polyphaga)). The strong sister relationship between Myxophaga and Adephaga has been suggested by Kukalová-Peck & Lawrence (1993) and Maddison *et al*. (1999). A denser taxon sampling is necessary to resolve the subordinal-level relationships, as the present study included only single members of Myxophaga and Archostemata.

Within the Adephaga, Carabidae (*Calosoma*) and Trachypachidae (*Trachypachus*) form a clade, which is in turn sister to Gyrinidae (*Macrogyrus*). This relationship corresponds well with the traditional division within the suborder, the terrestrial Geadephaga and aquatic Hydradephaga (Kavanaugh, 1986; Beutel, 1993), as well as with recent molecular phylogenetic studies (Shull *et al*., 2001; Hunt *et al*., 2007). The Polyphaga is the largest suborder within the Coleoptera, including more than 90% of beetle species. Traditionally, the suborder has been considered to comprise five infraorders: Staphyliniformia, Scarabaeiformia, Elateriformia, Bostrichiformia and Cucujiformia (Crowson, 1960; Kukalová-Peck & Lawrence, 1993). Hunt *et al*. (2007) established the five early-branching lineages in Polyphaga in a molecular phylogeny: Decliniidae,

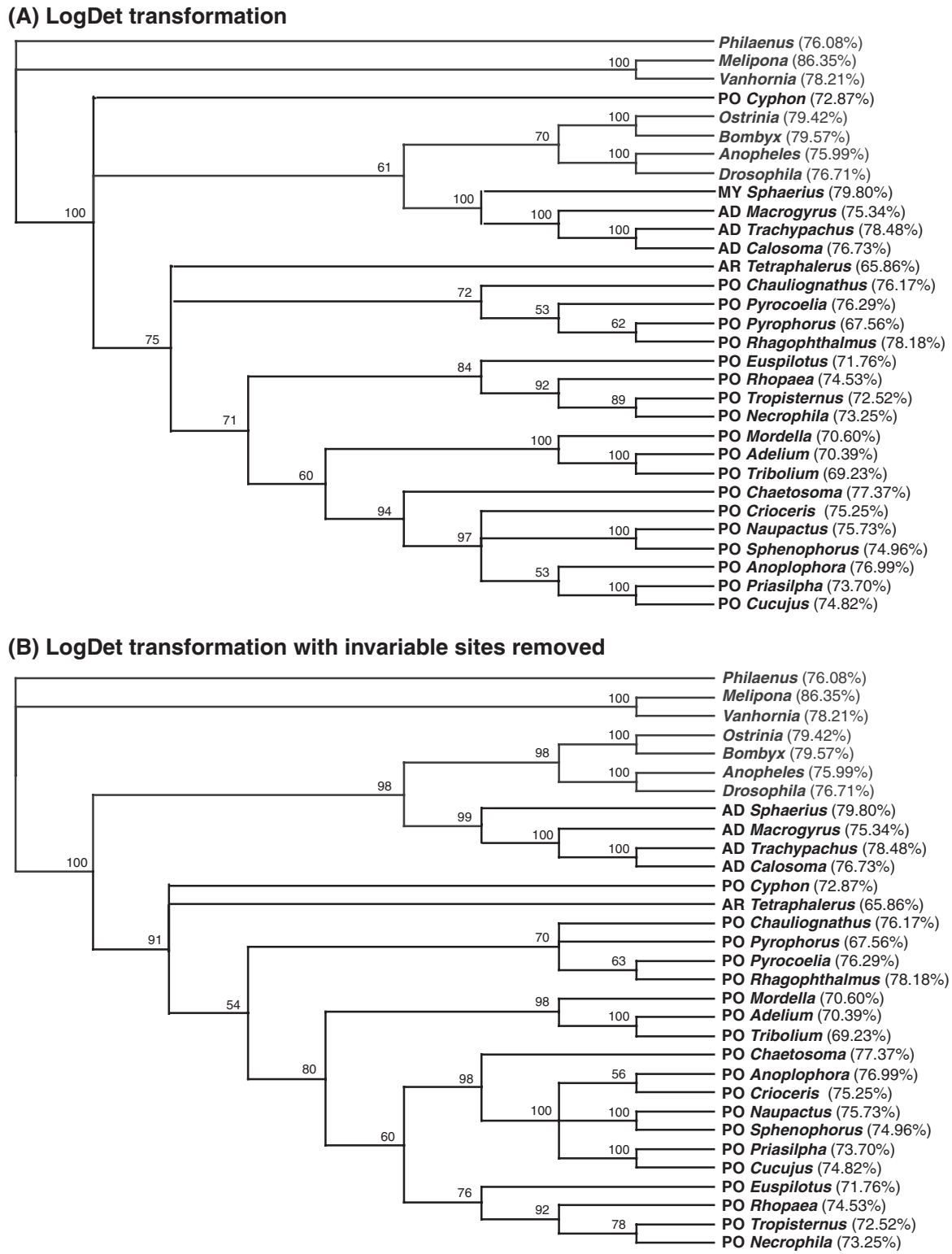

Fig. 4. The effect of the LogDet transformation. (A) A neighbor-joining tree after LogDet transformation of the nt123 dataset. (B) A neighbor-joining tree after LogDet transformation of the reduced dataset without invariable sites. Numbers above the branches indicate bootstrap support values. Numbers in parentheses next to terminals indicate (A) AT% from the all 13 protein-coding genes and (B) AT% after removal of invariable sites. The code before the taxon name indicates the taxonomic grouping: MY, Myxophaga; AD, Adephaga; AR, Archostemata; PO, Polyphaga.

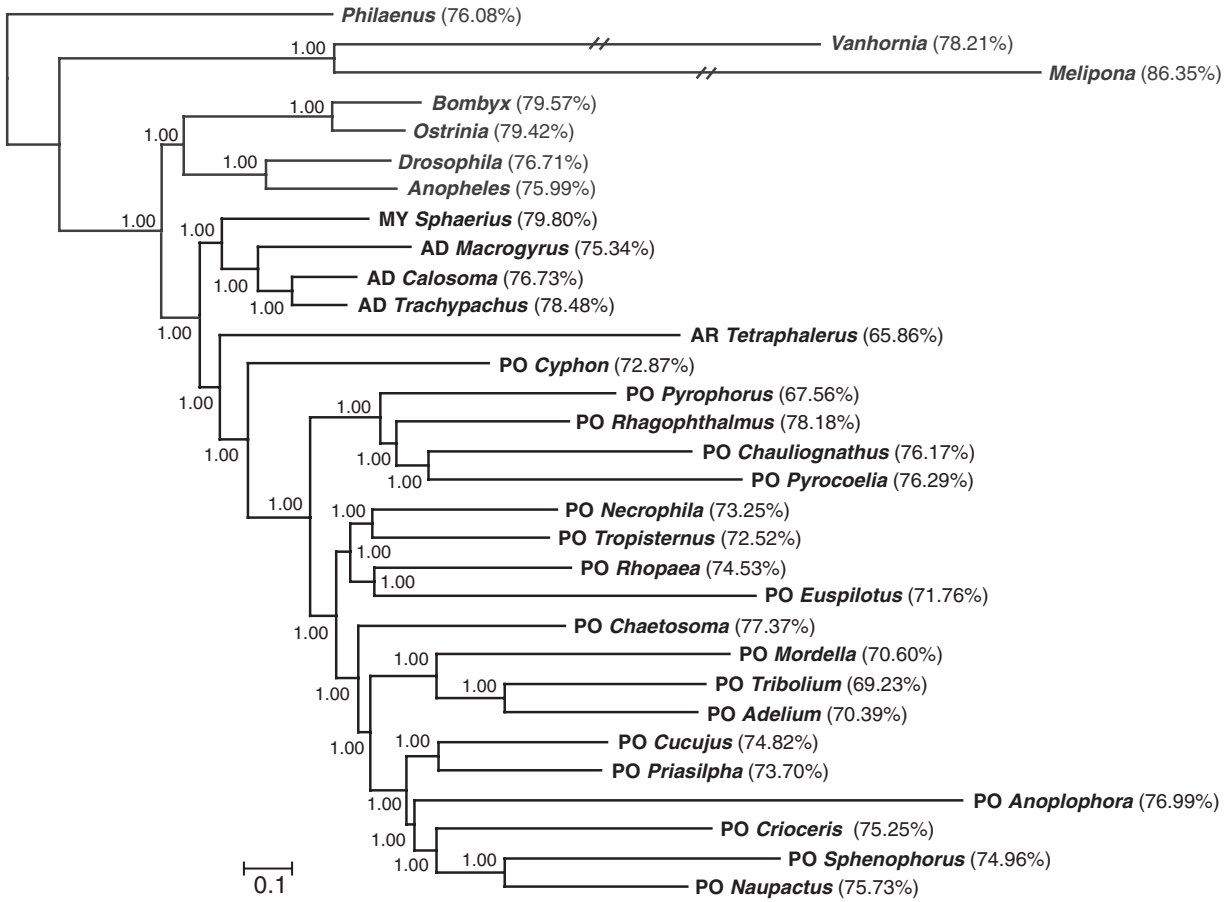

Fig. 5. The effect of modelling variable compositional vectors. A Bayesian tree implementing variable compositional vectors during tree search recovered using PHASE. Numbers above the branches indicate posterior probability values. Numbers in parentheses next to terminals indicate the total AT% from all 13 protein-coding genes. The code before the taxon name indicates the taxonomic grouping: MY, Myxophaga; AD, Adephaga; AR, Archostemata; PO, Polyphaga.

Scirtidae, Derodontidae, Eucinetidae and Clambidae. Our analysis includes one of these five polyphagan lineages, Scirtidae (*Cyphon*), and all infraorders but one (Bostrichiformia). Traditionally, Scirtidae has been placed as sister to the rest of Elateriformia because of several plesiomorphic characters (Lawrence & Newton, 1982), but our analysis suggests that it is sister to the rest of Polyphaga, corroborating Hunt *et al*. (2007). Elateridae (*Pyrophorus*), Cantharidae (*Chauliognathus*), Lampyridae (*Pyrocoelia*) and Rhagophthalmidae (*Rhagophthalmus*) form a monophyletic group in our study, which corresponds to the traditional Elateroidea (Branham & Wenzel, 2001; Bocakova *et al*., 2007). The relationship between Staphyliniformia (Hydrophiloidea, Histeroidea, and Staphylinoidea) and Scarabaeiformia (Scarabaeioidea) is not robust in our analysis, but a monophyletic relationship consisting of Silphidae (*Necrophila*), Hydrophilidae (*Triopisternus*), Histeridae (*Euspilotus*) and Scarabaeidae (*Rhopaea*) is consistently recovered, suggesting possible support for Haplogastra (Caterino *et al*., 2005). Cucujiformia is the largest infraorder in Polyphaga, comprising about 90 families, and its monophyly is strongly supported by the presence of cryptonephridic Malpighian tubules (Poll, 1932; Stammer, 1934) and non-functional and reduced spiracles on the eighth abdominal segment (Crowson, 1960), among other morphological characters (Wachmann, 1977; Caveney, 1986). Our study included ten members of Cucujiformia, which consistently formed a monophyly. Within Cucujiformia, the monophyly of Tenebrionoidea (*Mordella*, *Adelium* and *Tribolium*), Curculionoidea (*Naupactus* and *Sphenophorus*) and of Cucujoidea (*Cucujus* and *Priasilpha*) was strongly supported.

Overall, the mtgenome data recover relationships that are mostly congruent with traditional classifications of Coleoptera and previous phylogenetic studies, despite the small taxon sampling. Although mtgenome data are unable to resolve relationships among major arthropod lineages robustly (Cameron *et al*., 2004), previous studies have suggested that these data can resolve intraordinal relationships within Diptera (Cameron *et al*., 2007), Hymenoptera (Castro & Dowton, 2007; Cameron *et al*., 2008; Dowton *et al*., 2009) and Orthoptera (Fenn *et al*., 2008), and interordinal relationships within Polyneoptera

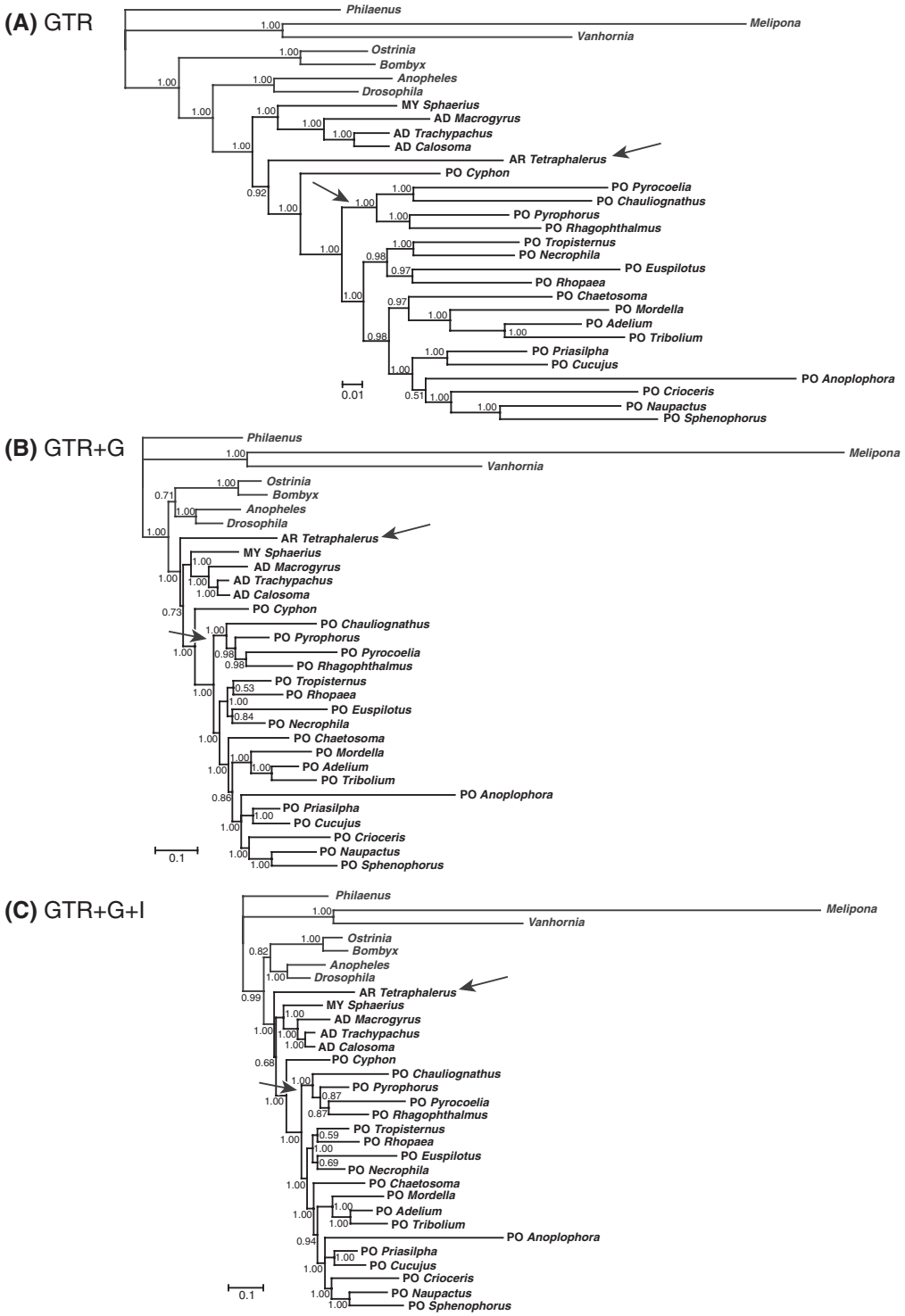

Fig. 6. The effect of accounting for among-site rate variation. The trees shown here are analysed based on the nt2 dataset, which is determined to be the least likely to violate the phylogenetic assumptions among the nucleotide datasets. (A) A Bayesian tree implementing a GTR model, assuming an equal rate among sites. (B) A Bayesian tree implementing a GTR + G model, assuming a gamma-shaped rate among sites. (C) A Bayesian tree implementing a GTR + G + I model, assuming a gamma-shaped rate among sites with invariable sites. Arrows indicate the terminals whose phylogenetic positions shift after implementing different models of sequence evolution. Numbers above the branches indicate posterior probability values. Numbers in parentheses next to terminals indicate the total AT% from all 13 protein-coding genes. The code shown before the taxon name indicates the taxonomic grouping: MY, Myxophaga; AD, Adephaga; AR, Archostemata; PO, Polyphaga.

Table 5. Bayes factor analysis showing the effect of applying different models of rate variation on likelihood.

Model comparison (M_1/M_0)	Model likelihood (harmonic mean)		Bayes factor		Evidence against M_0
	$\log_e f(X \mid M_1)$	$\log_e f(X \mid M_0)$	$\log_e B_{10}$	$2\log_e B_{10}$	
GTR + G/GTR	−46372.03	−51325.74	4953.71	9907.42	Decisive
GTR + G + I/GTR	−46341.21	−51325.74	4984.53	9969.06	Decisive
GTR + G + I/GTR+G	−46341.21	−46372.03	30.82	61.64	Strong

Model likelihood was estimated using harmonic means calculated from MrBayes. This analysis is based on the nt2 dataset, which is determined to be the least likely to violate the phylogenetic assumptions owing to base compositional heterogeneity.

(Cameron *et al*., 2006). Our study again confirms the phylogenetic utility of the mtgenome data in resolving higher-level relationships in insect systematics.

Base compositional heterogeneity in mtgenomes

Based on the disparity index values calculated from individual codon positions (Table 1) and the matched-pairs tests of symmetry (Table 2), several interesting patterns of base compositional heterogeneity are revealed. First, if a taxon experiences non-stationary evolution, it can be expressed in all three codon positions. For example, one of the most heterogeneous taxa within our dataset is *Tetraphalerus*, which has high I_D-values across different codon positions compared with other taxa and the highest number of significant P-values in the matched-pairs tests, which suggests that the mtgenome as a whole is affected by the compositional bias. Second, there can be a variation from the first pattern in that a taxon with compositional bias may express more heterogeneity in certain codon positions than others. For instance, *Pyrophorus*, which groups with *Tetraphalerus* owing to a similar base composition, has high I_D-values in the nt1 and nt3 position whereas the nt2 position has an I_D-value similar to that of other taxa. The number of significant P-values of a given species varies among different codon positions in the matched-pairs tests. Third, base compositional heterogeneity is most obvious in the nt3 position, followed by the nt1 and nt2. In fact, the nt2 position is the least affected by compositional bias, probably owing to functional constraint; a similar pattern was observed in mammalian mtgenomes (Gibson *et al*., 2005).

An examination of AT% of different codon positions across the species also reveals an interesting pattern (Fig. 1). In general, the nt2 position has the lowest AT bias, followed by the nt1 and nt3. In fact, the differences in AT% between the nt1 and nt2 position are small compared with those between these two positions and the nt3 position. When the dataset is dissected further into the different codon positions of individual protein-coding genes, another interesting pattern emerges (Fig. 7). In all genes, the variation of AT% among different species is the highest in the nt3 position, followed by the nt1 and nt2. Whereas AT% varies little within the nt2 position, there is a fair amount of variation across different genes. For example, the nt2 position of *cox1* has a low AT% whereas that of *nd4l* has a much higher AT%. These findings collectively demonstrate the complexity of the dataset.

Performance of standard phylogenetic inference methods when the data are affected by base compositional heterogeneity

Our study demonstrates that mtgenome data can be affected severely by base compositional heterogeneity. Both the I_D-test (Table 1) and the matched-pairs tests of symmetry (Table 2) reveal that our dataset, regardless of how it is partitioned or recoded, violates the phylogenetic assumptions.

When the data are biased, standard phylogenetic inference methods consistently perform poorly. Among these methods, we find that a parsimony analysis without any data transformation is the most severely affected (Collins *et al*., 1994). The single most parsimonious tree recovered from the nt123 dataset grouped two distantly related taxa, *Tetraphalerus* and *Pyrophorus*, that have similarly low AT base composition and many homoplasious characters from the nt3 position (Fig. 3A). Among the smaller datasets consisting of individual codon position partitions, we find that the parsimony analysis performs poorly in the nt1 (Figure S1A), nt3 (Figure S3A) and nt12 (Figure S4A) datasets. Coleoptera is not recovered as a monophyletic group in the nt1 and nt3 datasets, and Polyphaga is not recovered as a monophyletic group in the nt12 dataset. The nt2 analysis finds a monophyletic Coleoptera and the monophyly of each of the suborders (Figure S2A). Although the matched-pairs test of symmetry reveals that the nt2 position does exhibit some biases (Table 2), the parsimony appears to perform well for this dataset, suggesting that this inference method is robust against a slight detectable level of base compositional heterogeneity.

A standard Bayesian analysis proceeds first by selecting appropriate models for gene partitions using statistically based tests such as the hierarchical likelihood ratio tests (hLRTs) or the Akaike information criterion (AIC) (Posada & Buckley, 2004), as implemented in model-selecting software such as MrModelTest (Nylander, 2004). MrModelTest chooses one of 24 models in the family of the time-reversible Markov models (Nylander, 2004), which assume that the sequences evolve under stationary, reversible and homogenous conditions (Jermiin *et al*., 2008). However, these assumptions are tested infrequently. Because these steps are standardized and typically a single majority-rule consensus tree is obtained with

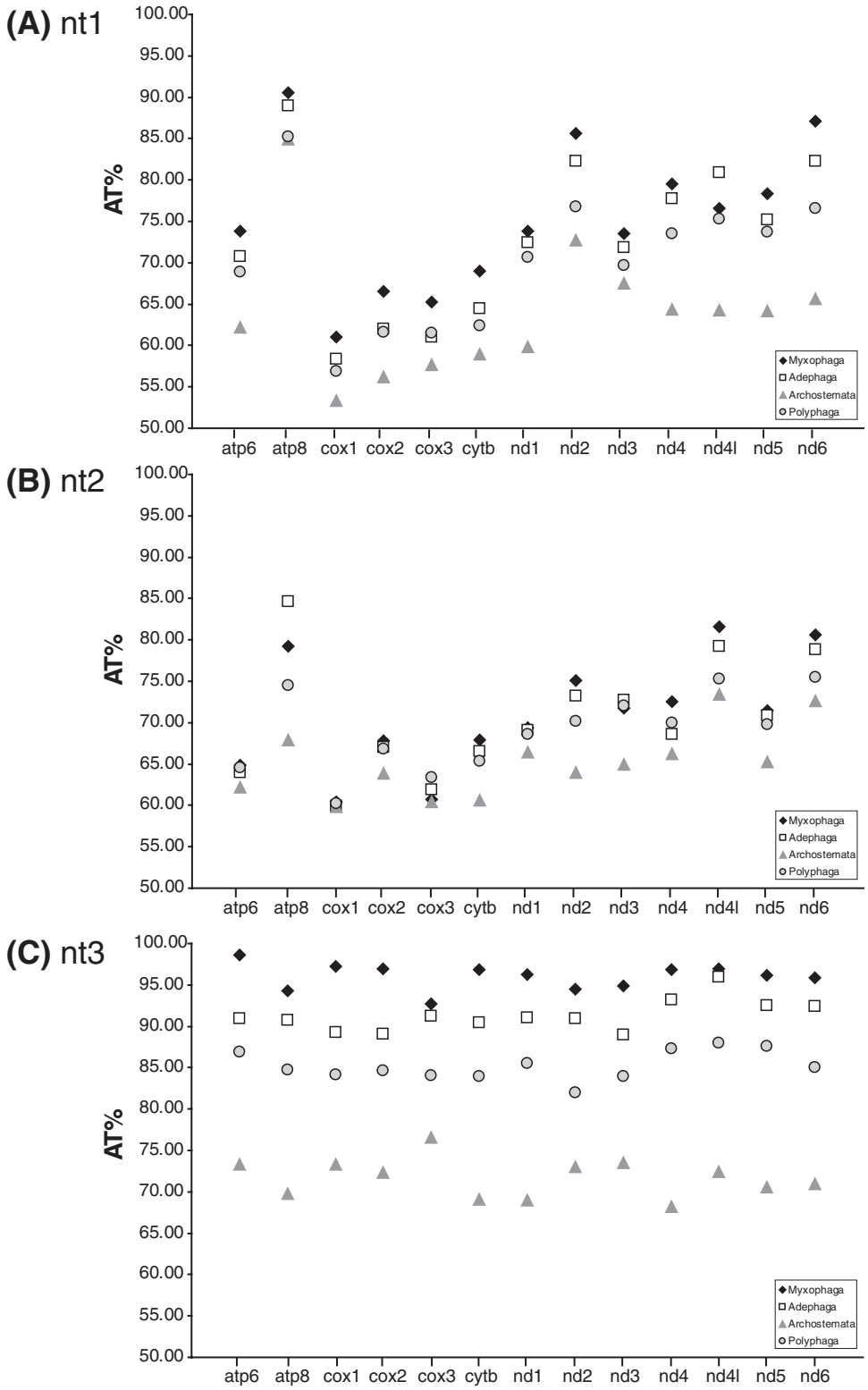

Fig. 7. Variation in base composition in different codon positions across 13 protein-coding genes. Shown here are the mean AT% for each taxonomic group. (A) AT% of the nt1 position plotted against individual protein-coding genes. (B) AT% of the nt2 position plotted against individual protein-coding genes. (C) AT% of the nt3 position plotted against individual protein-coding genes. The variation in base composition is the highest in the nt3 position, followed by the nt1 and nt2 across all genes.

high posterior probability (Simmons *et al*., 2004), it is difficult to identify how violations to the implicit assumptions present in the data might have affected the results. We show that our data evolve under non-stationary conditions, which by definition means that they evolve under non-reversible and heterogeneous conditions (Jermiin *et al*., 2008). Therefore, by applying models specified by MrModelTest to the Bayesian analyses of our data, we violate the implicit assumptions of time-reversible models. The question then becomes how robust the models are to violations of the assumptions. Our mixed-model BA analysis (Fig. 3B) recovered the monophyly of each of the four beetle suborders, but was different from the reference phylogeny in the placement of *Tetraphalerus*, which is the genus most severely affected by base compositional heterogeneity. However, *Pyrophorus*, which also has a low AT%, groups correctly with other elateroid beetles in this analysis. The negative effect of base compositional heterogeneity may also depend on taxon sampling. Sheffield *et al*. (2009) failed to recover taxonomically accepted relationships in a Bayesian framework using a dataset affected by base compositional heterogeneity. Their dataset included only 18 terminals, including *Tetraphalerus* and *Pyrophorus*, and the taxa with low AT composition were grouped strongly under typical phylogenetic methods. Therefore, an increase in taxon sampling may alleviate the bias, as seen in the present study. The analyses based on codon position exhibit a pattern similar to that for the parsimony analyses. The nt1 (Figure S1B) and nt3 (Figure S3B) datasets do not recover Coleoptera as a monophyletic clade, whereas the nt2 dataset (Figure S2B) recovers a monophyletic Coleoptera, and the monophyly of each of the suborders is identical to the reference topology.

The phylogenetic analyses based on individual codon positions under standard inference methods collectively demonstrate that there is much variation in terms of the influence from base compositional heterogeneity and phylogenetic signal among all three codon positions, and that the standard methods generally perform poorly with biased data.

Performance of phylogenetic methods designed to cope with base compositional heterogeneity

LogDet transformation is the most popular method of addressing base compositional heterogeneity (Lockhart *et al*., 1994). We find that a LogDet transformation seems to account partially for the bias because the two taxa with the lowest AT%, *Tetraphalerus* and *Pyrophorus*, are not grouped. However, it does not recover a monophyletic Coleoptera, which suggests that it is affected by some other source of bias (Fig. 4). Thus LogDet transformation does what it is supposed to do – corrects the bias arising from base compositional heterogeneity (Lockhart *et al*., 1994) – but it does not appear to be immune to ASRV. Another recent method to deal with compositional heterogeneity is the use of compositional vectors during tree search in a Bayesian framework (Gowri-Shankar & Rattray, 2007), which can be performed using the programPHASE. Sheffield *et al*. (2009) compared several model-based approaches known to deal with base compositional heterogeneity and showed that the algorithm implemented in PHASE performs fairly well against the bias. Our PHASE analysis can correct the bias from base compositional heterogeneity because taxa with low AT% do not group, but the recovered relationship among the suborders differs from the reference topology (Fig. 5). We conclude that the PHASE analysis can cope well with the base compositional heterogeneity, but appears to be influenced by other possible sources of bias.

Performance of standard phylogenetic inference methods when the data are affected by among-site rate variation

ASRV is a more difficult bias to identify than base compositional heterogeneity because there is no simple indicator of such bias, and typical models with gamma-shaped rate variation (i.e. +G) already filter out this bias. Nevertheless, it is important to understand this source of bias because it can cause incorrect phylogenetic inference if unaccounted for.

Inclusion of the gamma-shaped rate variation in the GTR model recovers a different topology than the GTR model alone (Fig. 6). The Bayes factor analysis shows that the addition of the rate variation parameter results in a better fit to the data (Table 5), suggesting that our dataset (nt2) is affected by the presence of ASRV and that the taxa whose phylogenetic positions shift for different models are probably those affected most by this bias. The phylogenetic position of the archostematan *Tetraphalerus* appears to be the most affected by ASRV, followed by the members of Elateroidea (Fig. 6).

The problem of ASRV can be easily remedied in a model-based approach by applying the rate variation parameter, but not so in a parsimony framework. The MP tree topology based on the nt2 dataset (Figure S2A) is identical to the BA tree under the GTR model of the same dataset (Fig. 6A), suggesting that parsimony at the nucleotide level is vulnerable to the bias originating from ASRV. However, when the nucleotide data are translated to the amino acid sequences the effect of ASRV appears to be diminished, as shown by the MP topology based on the aa dataset (Fig. 2).

Conclusion

In a typical molecular phylogenetic study, assumptions are rarely tested rigorously. Researchers are more likely to check for possible violations if unexpected relationships are recovered. Here, two such violations in the mtgenome phylogeny of Coleoptera are identified, and various ways of overcoming these biases explored. Standard phylogenetic inference methods are severely affected and the methods known to cope with one type of bias are not immune to other violations.

One of the ideas behind modern phylogenomics is that the inclusion of more data will eventually overcome the conflicting signal present in a few genes. In the context of mitochondrial phylogenomics, even with a large amount of data our difficulty

in overcoming systematic bias implies that the addition of data creates additional complexity in a final dataset already laden with biases. Simply adding more genes to the data will not necessarily solve the phylogenetic problem at hand because the addition of each marker can create another bias needing correction. The important and necessary current trend of including extensive molecular data is welcome, but we urge caution and argue that more data exploration is imperative, especially when many genes are included in the analysis.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: DOI: 10.1111/j.1365-3113.2009.00517.x

Figure S1. The first codon position only (nt1) phylogeny.

Figure S2. The second codon position only (nt2) phylogeny.

Figure S3. The third codon position only (nt3) phylogeny.

Figure S4. The first and second codon positions only (nt12) phylogeny.

Table S1. Taxon sampling and GenBank accession numbers

Table S2. Models used for mixed-model Bayesian analyses

Table S3. Mitochondrial genome description of the seven new beetle species included in this study.

File S1. A list of species-specific primers used in primer walking.

Please note: Neither the Editors nor Wiley-Blackwell are responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Acknowledgements

We would like to thank our collaborators who kindly provided specimens: G. Svenson, M. Caterino, A. Marvaldi, A. Ames, K. Jarvis, T. Waite, M. Terry, A. Wild and J. Sullivan. We also thank C. Shepherd for technical assistance. Comments by P. S. Cranston and L. Jermiin improved the clarity of the manuscript. We wish to express special thanks to L. Jermiin for examining our data in detail and providing numerous constructive ideas. This work is supported by the National Science Foundation grants EF-0531665 (AToL: Beetle Tree of Life), DEB 0120718 and DEB 0444972 to MFW.

References

Ababneh, F., Jermiin, L.S., Ma, C. & Robinson, J. (2006) Matched-pairs tests of homogeneity with applications to homologous nucleotide sequences. *Bioinformatics*, **22**, 1225–1231.

Baker, R.H. & DeSalle, R. (1997) Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. *Systematic Biology*, **46**, 654–673.

Barry, D. & Hartigan, J.A. (1987) Statistical analysis of hominoid molecular evolution. *Statistical Science*, **2**, 191–210.

Bergsten, J. (2005) A review of long-branch attraction. *Cladistics*, **21**, 163–193.

Beutel, R.G. (1993) Phylogenetic analysis of Adephaga (Coleoptera) based on characters of the larval head. *Systematic Entomology*, **18**, 127–147.

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

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

Boore, J.L. (1999) Animal mitochondrial genomes. *Nucleic Acids Research*, **27**, 1767–1780.

Boore, J.L. & Brown, W.M. (1998) Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Current Opinion in Genetics and Development*, **8**, 668–674.

Boore, J.L., Collins, T.M., Stanton, D., Daehler, L.L. & Brown, W.M. (1995) Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements. *Nature*, **376**, 163–165.

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

Branham, M.A. & Wenzel, J.W. (2001) The evolution of bioluminescence in cantharoids (Coleoptera: Elateroidea). *Florida Entomologist*, **84**, 565–586.

Bremer, K. (1994) Branch support and tree stability. *Cladistics*, **10**, 295–304.

Cameron, S.L. & Whiting, M.F. (2008) The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta*, (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene*, **408**, 112–123.

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.

Cameron, S.L., Barker, S.C. & Whiting, M.F. (2006) Mitochondrial genomics and the new insect order Mantophasmatodea. *Molecular Phylogenetics and Evolution*, **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. *Systematic Entomology*, **32**, 40–59.

Cameron, S.L., Dowton, M., Castro, L.R. *et al*., (2008) The sequence of the mitochondrial genomes of two vespid wasps reveals a number of derived tRNA gene rearrangements. *Genome*, **51**, 800–808.

Castro, L.R. & Dowton, M. (2007) Mitochondrial genomes in the Hymenoptera and their utility as phylogenetic markers. *Systematic Entomology*, **32**, 60–69.

Caterino, M.S., Shull, V.L., Hammond, P.M. & Vogler, A.P. (2002) Basal relationships of Coleoptera inferred from 18S rDNA sequences. *Zoologica Scripta*, **31**, 41–49.

Caterino, M.S., Hunt, T. & Vogler, A.P. (2005) On the constitution and phylogeny of Staphyliniformia (Insecta: Coleoptera). *Molecular Phylogenetics and Evolution*, **34**, 655–672.

Caveney, S. (1986) The phylogenetic significance of ommatidium structure in the compound eyes of polyphagan beetles. *Canadian Journal of Zoology/Revue Canadienne de Zoologie*, **64**, 1787–1819.

Collins, T.M., Wimberger, P.H. & Naylor, G.J.P. (1994) Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Systematic Biology*, **43**, 482–496.

Crowson, R.A. (1960) The phylogeny of Coleoptera. *Annual Review of Entomology*, **5**, 111–134.

Delsuc, F., Philips, M.J. & Penny, D. (2003) Comment of "Hexapod origins: Monophyletic or paraphyletic?". *Science*, **301**, 1482e.

Delsuc, F., Brinkmann, H. & Philippe, H. (2005) Phylogenomics and the reconstruction of the tree of life. *Nature Reviews: Genetics*, **6**, 361–375.

Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. (2006) Tunicates and not cephalochordates are the closest living relatives of vertabrates. *Nature*, **439**, 965–968.

Dowton, M. & Austin, A.D. (1997) The evolution of strand-specific compositional bias. A case study in the hymenopteran mitochondrial 16S rRNA gene. *Molecular Biology and Evolution*, **14**, 109–112.

Dowton, M. & Austin, A.D. (1999) Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. *Molecular Biology and Evolution*, **16**, 298–309.

Dowton, M., Castro, L.R. & Austin, A.D. (2002) Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome 'morphology'. *Invertebrate Systematics*, **16**, 345–356.

Dowton, M., Cameron, S.L., Austin, A.D. & Whiting, M.F. (2009) Phylogenetic approaches for the analysis of mitochondrial genome sequence data in the Hymenoptera – a lineage with both rapidly and slowly evolving mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **52**, 512–519.

Dunn, C.W., Hejnol, A., Matus, D.Q., *et al*. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, **452**, 745–749.

Eddy, S.R. & Durbin, R. (1994) RNA sequence analysis using covariance models. *Nucleic Acids Research*, **22**, 2079–2088.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Felsenstein, J. (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology*, **27**, 401–410.

Felsenstein, J. (2001) Taking variation of evolutionary rates between sites in account in inferring phylogenies. *Journal of Molecular Evolution*, **53**, 447–455.

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. *Molecular Phylogenetics and Evolution*, **49**, 59–68.

Fitch, W.M. & Margoliash, E. (1967) A method for estimating the number of invariant amino acid coding positions in a gene using cytochrome *c* as a model case. *Biochemical Genetics*, **1**, 65–71.

Foster, P.G. (2004) Modeling compositional heterogeneity. *Systematic Biology*, **53**, 485–495.

Foster, P.G. & Hickey, D.A. (1999) Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. *Journal of Molecular Evolution*, **48**, 284–290.

Galtier, N. & Gouy, M. (1995) Inferring phylogenies from DNA sequences of unequal base compositions. *Proceedings of the National Academy of Sciences of the United States of America*, **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. *Molecular Biology and Evolution*, **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. *Molecular Biology and Evolution*, **22**, 251–264.

Gillespie, J.J., Johnston, J.S., Cannone, J.J. & Gutell, R.R. (2006) Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization and retrotransposable elements. *Insect Molecular Biology*, **15**, 657–686.

Goloboff, P.A. (1999) Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics*, **15**, 415–428.

Goloboff, P.A., Farris, J.S. & Nixon, K.C. (2003) *T.N.T.: Tree Analysis Using New Technology*. Program and documentation, available from the authors, and at www.zmuc.dk/public/phylogeny.

Gowri-Shankar, V. & Rattray, M. (2006) On the correlation between composition and site-specific evolutionary rate: Implications for phylogenetic inference. *Molecular Biology and Evolution*, **23**, 352–364.

Gowri-Shankar, V. & Rattray, M. (2007) A reversible jump method for Bayesian phylogenetic inference with a non-homogeneous substitution model. *Molecular Biology and Evolution*, **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. *Systematic Biology*, **56**, 83–96.

Hasegawa, M. & Hashimoto, T. (1993) Ribosomal RNA trees misleading? *Nature*, **361**, 23.

Ho, J.W.K. & Jermiin, L.S. (2004) Tracing the decay of the historical signal in biological sequence data. *Systematic Biology*, **53**, 623–637.

Ho, J.W.K., Adams, C.E., Lew, J.B., *et al*. (2006) SeqVis: Visualization of compositional heterogeneity in large alignments of nucleotides. *Bioinformatics*, **22**, 2162–2163.

Hori, H. & Osawa, S. (1987) Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. *Molecular Biology and Evolution*, **4**, 445–472.

Hughes, J., Longhorn, S.J., Papadopoulou, A., *et al*. (2006) Dense taxonomic EST sampling and its applications for molecular systematics of the Coleoptera (Beetles). *Molecular Biology and Evolution*, **23**, 268–278.

Hunt, T., Bergsten, J., Levkanicova, Z., *et al*. (2007) A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science*, **318**, 1913–1916.

Hwang, U.W., Park, C.J., Yong, T.S. & Kim, W. (2001) One-step PCR amplification of complete arthropod mitochondrial genomes. *Molecular Phylogenetics and Evolution*, **19**, 345–352.

Jayaswal, V., Jermiin, L.S. & Robinson, J. (2005) Estimation of phylogeny using a general Markov model. *Evolutionary Bioinformatics,* **1**, 62–80.

Jayaswal, V., Robinson, J. & Jermiin, L.S. (2007) Estimation of phylogeny and invariant sites under the general Markov model of nucleotide sequence evolution. *Systematic Biology,* **56**, 155–162.

Jermiin, L.S., Ho, S.Y., Ababneh, F., Robinson, J. & Larkum, A.W. (2004) The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. *Systematic Biology,* **53**, 638–643.

Jermiin, L.S., Jayaswal, V., Ababneh, F. & Robinson, J. (2008) Phylogenetic model evaluation. *Bioinformatics, Volume 1: Data, Sequence Analysis, and Evolution* (ed. by J. M. Keith), pp. 331–364. Humana Press, Totowa, New Jersey.

Jermiin, L.S., Ho, J.W.K., Lau, K.W. & Jayaswal, V. (2009) SeqVis: A tool for detecting compositional heterogeneity among aligned nucleotide sequences. *Bioinformatics for DNA Sequence Analysis* (ed. by D. Posada), pp. 65–91. Humana Press, Totowa, New Jersey.

Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kavanaugh, D.H. (1986) A systematic review of amphizoid beetles (Amphizoidae: Coleoptera) and their phylogenetic relationships to other Adephaga. *Proceedings of the California Academy of Sciences*, **44**, 67–109.

Kolaczkowski, B. & Thorton, J.W. (2004) Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature*, **431**, 980–984.

Kukalová-Peck, J. & Lawrence, J.F. (1993) Evolution of the hind wing in Coleoptera. *Canadian Entomologist*, **125**, 181–258.

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. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 1455–1459.

Lawrence, J.F. & Newton, A.F. Jr (1982) Evolution and classification of beetles. *Annual Review of Ecology and Systematics*, **13**, 261–290.

Lockhart, P.J., Steel, M.A., Hendy, M.D. & Penny, D. (1994) Recovering evolutionary trees under a more realistic model of sequence. *Molecular Biology and Evolution*, **11**, 605–612.

Macey, J.R., Larson, A., Ananjeva, N.B. & Papenfuss, T.J. (1997) Evolutionary shifts in three major structural features of the mitochondrial genome among iguanian lizards. *Journal of Molecular Evolution*, **44**, 660–674.

Maddison, D.R. & Maddison, W.P. (2005) *MacClade 4*. Sinauer Associates, Inc., Sunderland, Massachusetts.

Maddison, D.R., Baker, M.D. & Ober, K.A. (1999) Phylogeny of carabid beetles as inferred from 18S ribosomal DNA (Coleoptera: Carabidae). *Systematic Entomology*, **24**, 103–138.

Mayrose, I., Friedman, N. & Pupko, T. (2005) A gamma mixture model better accounts for among site rate heterogeneity. *Bioinformatics*, **21** (Suppl. 2), ii151–ii158.

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. *Molecular Biology and Evolution*, **18**, 1293–1304.

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

Nardi, F., Spinsanti, G., Boore, J.L., Carapelli, A., Dallai, R. & Frati, F. (2003b) Response to comment on "Hexapod origins: monophyletic or paraphyletic?" *Science*, **301**, 1482e.

Nixon, K.C. (1999) The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics*, **15**, 407–414.

Nylander, J.A.A. (2004) *MrModeltest v.2. Program distributed by the author*. Evolutionary Biology Centre, Uppsala University, Uppsala.

Peña, C., Wahlberg, N., Weingartner, E., Kodandaramaiah, U., Nylin, S., Freitas, A.V.L. & Brower, A.V.Z. (2006) Higher level phylogeny of Satyrinae butterflies (Lepidoptera: Nymphalidae) based on DNA sequence data. *Molecular Phylogenetics and Evolution*, **40**, 29–49.

Philippe, H. & Telford, M.J. (2005) Large-scale sequencing and the new animal phylogeny. *Trends in Ecology and Evolution*, **21**, 614–620.

Philippe, H., Delsuc, F., Brinkmann, H. & Lartillot, N. (2005a) Phylogenomics. Annual Review of Ecology. *Evolution and Systematics*, **36**, 541–562.

Philippe, H., Zhou, Y., Brinkmann, H., Rodrigue, N. & Delsuc, F. (2005b) Heterotachy and long-branch attraction in phylogenetics. *BMC Evolutionary Biology*, **5**, 50.

Phillips, M.J., Delsuc, F. & Penny, D. (2004) Genome-scale phylogeny and the detection of systematic biases. *Molecular Biology and Evolution*, **21**, 1455–1458.

Poll, M. (1932) Note sur la fonction des tubes de Malpighi des Coléoptères. *Bulletin & Annales de la Société Entomologique de Belgique*, **72**, 103–109.

Posada, D. & Buckley, T.R. (2004) Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53**, 793–808.

Reeves, J.H. (1992) Heterogeneity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA. *Journal of Molecular Evolution*, **35**, 17–31.

Rokas, A., Krüger, D. & Carroll, S.B. (2005) Animal evolution and the molecular signature of radiations compressed in time. *Science*, **310**, 1933–1938.

Ronaghi, M. (2001) Pyrosequencing sheds light on DNA sequencing. *Genome Research*, **11**, 3–11.

Ronquist, F. & Huelsenbeck, J.P. (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Sheffield, N.C., Song, H., Cameron, S.L. & Whiting, M.F. (2009) Nonstationary evolution and compositional heterogeneity in beetle mitochondrial phylogenomics. *Systematic Biology*, **58**, 381–394.

Shull, V.L., Vogler, A.P., Baker, M.D., Maddison, D.R. & Hammond, P.M. (2001) Sequence alignment of 18S ribosomal RNA and the basal relationships of adephagan beetles: Evidence for monophyly of aquatic families and the placement of Trachypachidae. *Systematic Biology*, **50**, 945–969.

Simison, W.B., Lindberg, D.R. & Boore, J.L. (2006) Rolling circle amplification of metazoan mitochondrial genomes. *Molecular Phylogenetics and Evolution* **39**, 562–567.

Simmons, M., Pickett, K.M. & Miya, M. (2004) How meaningful are Bayesian support values? *Molecular Biology and Evolution*, **21**, 188–199.

Stammer, H.J. (1934) Bau und Bedeutung der malpighischen Gefässe der Coleopteren. *Zoomorphology*, **29**, 196–217.

Steel, M., Huson, D. & Lockhart, P.J. (2000) Invariable sites models and their use in phylogeny reconstruction. *Systematic Biology*, **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*. *Molecular Phylogenetics and Evolution*, **26**, 513–526.

Sueoka, N. (1962) On the genetic basis of variation and heterogeneity of DNA base composition. *Proceedings of the National Academy of Sciences of the United States of America*, **48**, 166–169.

Swofford, D.L. (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland, Massachusetts.

Taanman, J.W. (1999) The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta*, **1410**, 103–123.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596.

Tarrío, R., Rodríguez-Trelles, F. & Ayala, F.J. (2001) Shared nucleotide composition biases among species and their impact on phylogenetic reconstructions of the Drosophilidae. *Molecular Biology and Evolution*, **18**, 1464–1473.

Wachmann, E. (1977) Vergleichende Analyse der feinstrukturellen Organisation offener Rhabdome in den Augen der Cucujiformia (Insecta, Coleoptera), unter besonderer Berücksichtigung der Chrysomelidae. *Zoomorphology*, **88**, 95–131.

Wernersson, R. & Pedersen, A.G. (2003) RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Research*, **31**, 3537–3539.

Wolstenhome, D.R. (1992) Animal mitochondrial DNA: Structure and evolution. *International Review of Cytology*, **141**, 173–216.

Yamauchi, M.M., Miya, M.U. & Nishida, M. (2004) Use of a PCR-based approach for sequencing whole mitochondrial genomes of insects: two examples (cockroach and dragonfly) based on the method developed for decapod crustaceans. *Insect Molecular Biology*, **13**, 435–442.

Yang, Z. (1996) Among-site rate variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, **11**, 367–372.

Accepted 16 November 2009
First published online 1 March 2010

