# A Comparative Analysis of Mitochondrial Genomes in Coleoptera (Arthropoda: Insecta) and Genome Descriptions of Six New Beetles

*N.C. Sheffield,* H. Song,* S.L. Cameron,† and M.F. Whiting**

*Department of Biology, Brigham Young University; and †Australian National Insect Collection and Commonwealth Scientific and Industrial Research Organization Entomology, Canberra, Australian Capital Territory, Australia

Coleoptera is the most diverse group of insects with over 360,000 described species divided into four suborders: Adephaga, Archostemata, Myxophaga, and Polyphaga. In this study, we present six new complete mitochondrial genome (mtgenome) descriptions, including a representative of each suborder, and analyze the evolution of mtgenomes from a comparative framework using all available coleopteran mtgenomes. We propose a modification of atypical *cox1* start codons based on sequence alignment to better reflect the conservation observed across species as well as findings of TTG start codons in other genes. We also analyze tRNA-Ser(AGN) anticodons, usually GCU in arthropods, and report a conserved UCU anticodon as a possible synapomorphy across Polyphaga. We further analyze the secondary structure of tRNA-Ser(AGN) and present a consensus structure and an updated covariance model that allows tRNAscan-SE (via the COVE software package) to locate and fold these atypical tRNAs with much greater consistency. We also report secondary structure predictions for both rRNA genes based on conserved stems. All six species of beetle have the same gene order as the ancestral insect. We report noncoding DNA regions, including a small gap region of about 20 bp between tRNA-Ser(UCN) and *nad1* that is present in all six genomes, and present results of a base composition analysis.

## Introduction

Animal mitochondrial genomes (mtgenomes) are small, circular DNA with length ranging from 14,000 bp to 17,000 bp (Boore 1999; Cameron, Johnson, and Whiting 2007). They usually encode 37 genes (13 protein-coding, 22 transfer RNA, and 2 ribosomal RNA genes). The number of complete mtgenomes has steadily been on the rise with the technical feasibility of sequencing their entirety (Hwang et al. 2001; Yamauchi et al. 2004). This increasing availability of mtgenome data invites comparative study. In addition to the large amount of nucleotide data that is useful for deep-level phylogenetic studies (Gray et al. 1999; Nardi et al. 2003; Cameron et al. 2004; Cameron, Barker, and Whiting 2006; Cameron, Lambkin, et al. 2007), mtgenomes possess a number of evolutionarily interesting features such as length variation (Boyce et al. 1989), altered tRNA anticodons or secondary structures (Steinberg and Cedergren 1994; Eddy 2002), atypical start codons (e.g., Lavrov et al. 2000), base compositional bias (Gibson et al. 2004; Gowri-Shankar and Rattray 2006), codon usage (Jia and Higgs 2007), and gene rearrangement (Zhang and Hewitt 1997; Shao and Barker 2003; Mueller and Boore 2005). Some of these features appear to be lineage specific (Dowton et al. 2002); however, this insight can only be obtained from comparative analysis at various taxonomic levels.

Insect order Coleoptera contains over 360,000 described species divided into four suborders: Adephaga, Archostemata, Myxophaga, and Polyphaga (Lawrence and Newton 1982). Despite the size and diversity of the group, there are only six published (*Tribolium*, *Crioceris*, *Pyrocoelia*, two species of *Rhagophthalmus*, and *Pyrophorus*) and one unpublished (*Anoplophora*) beetle mtgenomes, all of which belong to suborder Polyphaga (Friedrich and Muqim 2003; Stewart and Beckenbach 2003; Bae et al. 2004; Li et al. 2007; Arnoldi et al. 2007) (table 1). The data from these seven mtgenomes suggest that the gene arrangement of Coleoptera follows that of the ancestral insect, that they all have a derived UCU anticodon and a reduced or missing D-stem in tRNA-Ser(AGN), and that they have atypical *cox1* start codons (Friedrich and Muqim 2003; Stewart and Beckenbach 2003; Bae et al. 2004; Li et al. 2007; Arnoldi et al. 2007). However, there has not been an attempt to describe the possible diversity of mtgenomes across the beetle suborders.

In this paper, we present six new beetle mtgenome descriptions, including representatives of Archostemata, Adephaga, and Myxophaga, and three additional Polyphaga mtgenomes from superfamilies not represented in previous analyses. The comparison of mtgenomes from all four suborders provides unique insights into the evolution of the mtgenome. We use the available 13 coleopteran mtgenomes to highlight unique features and shared characteristics and to point out particular parts of the mtgenome that have caused problems for annotation. We present possible solutions for such difficulties based on the comparative information now available.

## Materials and Methods

### mtgenome Sequencing, Annotation, and Analysis

We extracted total genomic DNA using the DNeasy Tissue kit (Qiagen, Hilden, Germany). Prior to extraction, we removed the entire abdominal segment to avoid possible contamination from gut content and to retain taxonomically important genital structures as vouchers. In all species, we used the entire body without abdomen. We followed primer walking and polymerase chain reaction protocols described in Cameron, Lambkin, et al. (2007). The species-specific primers designed for this study are available upon request from H.S. Morphological voucher specimens and remaining genomic DNA extracts were deposited in the Insect Genomic Collection of the Department of Biology and MLBM Museum, Brigham Young University. Throughout this paper, we refer to all species

Key words: mitochondrial genome, *cox1*, tRNA-Ser(AGN), RNA secondary structure, Coleoptera, Adephaga, Archostemata, Myxophaga, covariance model, base composition.

E-mail: ncs@byu.net.

*Mol. Biol. Evol.* 25(11):2499–2509. 2008
doi:10.1093/molbev/msn198
Advance Access publication September 8, 2008

**Table 1**
**Taxonomic Information and Accession Numbers for the Coleopteran Taxa Used in this Study**

| Species | Classification | Accession | Voucher/Reference | Location |
|---|---|---|---|---|
| This study | | | | |
| *Tetraphalerus bruchi* Heller | Archostemata: Ommatidae | EU877953 | IGC-CO687 | Argentina |
| *Trachypachus holmbergi* Mannerheim | Adephaga: Trachypachidae | EU877954 | IGC-CO843 | Canada |
| *Sphaerius* sp. | Myxophaga: Sphaeriusidae | EU877950 | IGC-CO837 | United States |
| *Chaetosoma scaritides* Westwood | Polyphaga: Cleroidea: Chaetosomatidae | EU877951 | IGC-CO683 | New Zealand |
| *Cyphon* sp. | Polyphaga: Scirtoidea: Scirtidae | EU877949 | IGC-CO838 | United States |
| *Priasilpha obscura* Broun | Polyphaga: Cucujoidea: Priasilphidae | EU877952 | IGC-CO684 | New Zealand |
| Previously reported | | | | |
| *Tribolium castaneum* | Polyphaga Tenebrionoidae: Tenebrionidae | NC_003081 | Friedrich and Muqim (2003) | |
| *Pyrocoelia rufa* | Polyphaga: Elateroidea: Lampyridae | NC_003970 | Bae et al. (2004) | |
| *Crioceris duodecimpunctata* | Polyphaga: Chrysomeloidea: Chrysomelidae | NC_003372 | Stewart and Beckenbach (2003) | |
| *Rhagophthalmus lufengensis* | Polyphaga: Elateroidea: Phengodidae | DQ888607 | Li et al. (2007) | |
| *Rhagophthalmus ohbai* | Polyphaga: Elateroidea: Phengodidae | AB267275 | Li et al. (2007) | |
| *Pyrophorus divergens* | Polyphaga: Elateroidae: Elateridae | NC_009964 | Arnoldi et al. (2007) | |
| *Anoplophora glabripennis* | Polyphaga: Chrysomeloidea: Cerambycidae | NC_008221 | Not applicable | |

by their generic name. GenBank accession numbers, specimen vouchers, classifications, and collecting localities are listed in table 1.

Raw sequence files were proofread and aligned into contigs in Sequencher 4.6 (GeneCodes Corporation, Ann Arbor, MI). We used the programs tRNAscan-SE (Lowe and Eddy 1997), INFERNAL (Eddy 2002), DOGMA (Wyman et al. 2004), and our unpublished software MOSAS to annotate the genomes. We located tRNAs with tRNAscan-SE with the default tRNA covariance model (CM), and we developed a new CM for coleopteran mitochondrial tRNA-Ser(AGN). We developed the new CM by creating an alignment of all 13 tRNA-Ser(AGN)s using INFERNAL's cmalign utility and modifying alignments by hand to eliminate the D-stem, in order to create a structural alignment more consistent with what is known about the structure of this tRNA. We used the COVE utility coveb (Eddy and Durbin 1994) to create a new CM and used this model to annotate tRNA-Ser(AGN). We also used the hand-curated INFERNAL alignment to infer a consensus secondary structure for tRNA-Ser(AGN). For questionable tRNAs, we used INFERNAL and Rfam to further investigate and sometimes revise tRNA annotation (Griffiths-Jones et al. 2005). DOGMA and MOSAS facilitate the annotation of organellar genomes by utilizing BLAST against published mtgenomes (e.g., Podsiadlowski 2006). After DOGMA and MOSAS reported general locations for genes based on similarity to other species, we identified start and stop codons to complete the annotation. The end of the small subunit rRNA (12S) was assigned by alignment with the secondary structures of 12S genes of other insects (Gillespie et al. 2006; Cameron and Whiting 2008). Helices were numbered according to the naming system of Gillespie et al. (2006).

For comparison to published genomes, we downloaded the published beetle mtgenome sequences from GenBank. For many of the alignments, such as aligning gap regions and tRNAs across beetle mtgenomes, we used MUSCLE (Edgar 2004). To determine the start codon of *nad1* and *cox1*, we made an alignment based on the translated amino acid sequences using ClustalW (Thompson et al. 1994) as implemented in MEGA version 3 (Kumar et al. 2004). In order to compare base compositional profiles of the six new species, we calculated base composition by codon position for each gene individually.

## Results and Discussion

### mtgenome Organization and Gene Content

The present study reports six new beetle mtgenomes, including sequences belonging to all four suborders of Coleoptera (table 2). Complete mtgenome sequences were obtained for *Tetraphalerus* (15,689 bp) and *Cyphon* (15,919 bp). Entire coding sequences with a partial control region were obtained for *Sphaerius* (15735 bp), *Chaetosoma* (15,511 bp), *Priasilpha* (16,887 bp), and *Trachypachus* (15,991 bp). A comparison of the mtgenome size across all four suborders of Coleoptera based on this study and previous studies suggests that the size of the coding region in Coleoptera is relatively stable around 14,700 bp in length (large intergenic spacers can cause deviations from this pattern). Although the length of coding region is constrained in order for the genes to function properly, the A + T–rich control region, located between the small rRNA subunit (12S) and tRNA-Ile, is free from such functional constraints, and its length variation is considerable. Despite being incomplete, the control region of *Priasilpha* was still longer than that of any other complete beetle mtgenome previously reported (table 3). Based on a restriction site mapping of mtDNA, Boyce et al. (1989) found that the control region of bark weevil *Pissodes* was extremely large (9–13 kb) and reported considerable size variation in the control region of Curculionidae. The size of the control region is therefore not consistent within beetle lineages but varies across them (Zhang and Hewitt 1997).

The six mtgenomes had varying degrees of high A + T content, ranging from 66.2% to 80.4% in the coding region and 78.4% to 91.0% in the control region (table 3). The A + T content of the control region was consistently higher than that of the coding region, which is a well-documented pattern in insect mtgenomes (Clary and Wolstenholme 1985; Zhang and Hewitt 1997). To understand what contributed to this variation in base composition, we examined the base frequency of the protein-coding genes by codon position (fig. 1). Overall, all six beetle mtgenomes followed similar

**Table 2**
**Nucleotide Positions and Anticodons (for tRNAs) for All Genes for Six New Beetle Species**

| Gene | Strand | Anticodon | *Tetraphalerus* | *Trachypachus* | *Sphaerius* | *Cyphon* | *Chaetosoma* | *Priasilpha* |
|---|---|---|---|---|---|---|---|---|
| tRNA-I | + | GAU | 1–63 (0) | 1–64 (0) | 1–64 (0) | 1–64 (0) | 1–66 (0) | 1–66 (0) |
| tRNA-Q | − | UUG | 65–133 (1) | 70–138 (5) | 62–130 (−3) | 62–130 (−3) | 64–132 (−3) | 64–132 (−3) |
| tRNA-M | + | CAU | 138–207 (4) | 141–209 (2) | 132–202 (1) | 129–197 (−2) | 132–200 (−1) | 132–200 (−1) |
| *nad2* | + | | 208–1231 (0) | 210–1238 (0) | 203–1228 (0) | 198–1217 (0) | 201–1206 (0) | 201–1214 (0) |
| tRNA-W | + | UCA | 1232–1296 (0) | 1240–1308 (1) | 1229–1296 (0) | 1395–1462 (177) | 1207–1269 (0) | 1215–1282 (0) |
| tRNA-C | − | GCA | 1289–1351 (−8) | 1354–1418 (45) | 1298–1360 (1) | 1455–1516 (−8) | 1274–1336 (4) | 1282–1343 (−1) |
| tRNA-Y | − | GUA | 1352–1417 (0) | 1431–1499 (12) | 1365–1432 (4) | 1516–1581 (−1) | 1336–1399 (−1) | 1344–1407 (0) |
| *cox1* | + | | 1419–2949 (1) | 1501–3031 (1) | 1434–2964 (1) | 1583–3113 (1) | 1401–2931 (1) | 1611–3144 (203) |
| tRNA-L | + | UAA | 2950–3014 (0) | 3032–3097 (0) | 2965–3029 (0) | 3114–3177 (0) | 2932–2994 (0) | 3145–3209 (0) |
| *cox2* | + | | 3015–3687 (0) | 3101–3788 (3) | 3030–3711 (0) | 3178–3862 (0) | 2995–3682 (0) | 3210–3896 (0) |
| tRNA-K | + | CUU | 3688–3758 (0) | 3789–3859 (0) | 3712–3782 (0) | 3863–3933 (0) | 3683–3752 (0) | 3898–3968 (1) |
| tRNA-D | + | GUC | 3758–3822 (−1) | 3860–3925 (0) | 3799–3865 (16) | 3935–4001 (1) | 3752–3813 (−1) | 3968–4036 (−1) |
| *atp8* | + | | 3823–3981 (0) | 3926–4090 (0) | 3866–4024 (0) | 4002–4163 (0) | 3897–4052 (83) | 4037–4192 (0) |
| *atp6* | + | | 3978–4652 (−4) | 4087–4761 (−4) | 4021–4695 (−4) | 4160–4834 (−4) | 4049–4717 (−4) | 4189–4860 (−4) |
| *cox3* | + | | 4642–5423 (−11) | 4762–5553 (0) | 4695–5482 (−1) | 4840–5619 (5) | 4717–5500 (−1) | 4862–5649 (1) |
| tRNA-G | + | UCC | 5424–5488 (0) | 5560–5625 (6) | 5483–5546 (0) | 5620–5683 (0) | 5501–5563 (0) | 5650–5712 (0) |
| *nad3* | + | | 5489–5840 (0) | 5626–5977 (0) | 5547–5898 (0) | 5684–6035 (0) | 5564–5915 (0) | 5713–6064 (0) |
| tRNA-A | + | UGC | 5841–5904 (0) | 5978–6042 (0) | 5899–5964 (0) | 6036–6103 (0) | 5916–5980 (0) | 6065–6130 (0) |
| tRNA-R | + | UCG | 5904–5970 (−1) | 6042–6106 (−1) | 5973–6037 (8) | 6104–6167 (0) | 5980–6042 (−1) | 6130–6189 (−1) |
| tRNA-N | + | GUU | 5968–6033 (−3) | 6110–6174 (3) | 6040–6105 (2) | 6168–6232 (0) | 6042–6107 (−1) | 6190–6254 (0) |
| tRNA-S | + | GCU/UCU[a] | 6035–6099 (1) | 6174–6242 (−1) | 6107–6171 (1) | 6233–6299 (0) | 6108–6166 (0) | 6255–6321 (0) |
| tRNA-E | + | UUC | 6101–6163 (1) | 6243–6308 (0) | 6174–6240 (2) | 6301–6367 (1) | 6167–6229 (0) | 6322–6386 (0) |
| tRNA-F | − | GAA | 6162–6227 (−2) | 6307–6373 (−2) | 6239–6306 (−2) | 6366–6432 (−2) | 6228–6290 (−2) | 6385–6449 (−2) |
| *nad5* | − | | 6228–7953 (0) | 6374–8102 (0) | 6307–8024 (0) | 6433–8155 (0) | 6291–8001 (0) | 6450–8163 (0) |
| tRNA-H | − | GUG | 7951–8014 (−3) | 8103–8170 (0) | 8025–8089 (0) | 8156–8219 (0) | 8002–8065 (0) | 8164–8230 (0) |
| *nad4* | − | | 8015–9338 (0) | 8171–9509 (0) | 8090–9425 (0) | 8220–9549 (0) | 8066–9386 (0) | 8231–9560 (0) |
| *nad4l* | − | | 9332–9625 (−7) | 9503–9796 (−7) | 9419–9712 (−7) | 9549–9839 (−1) | 9389–9670 (2) | 9557–9838 (−4) |
| tRNA-T | + | UGU | 9628–9689 (2) | 9799–9863 (2) | 9715–9779 (2) | 9842–9906 (2) | 9674–9735 (3) | 9843–9907 (4) |
| tRNA-P | − | UGG | 9690–9756 (0) | 9864–9930 (0) | 9780–9845 (0) | 9907–9972 (0) | 9736–9798 (0) | 9908–9973 (0) |
| *nad6* | + | | 9758–10273 (1) | 9932–10456 (1) | 9847–10356 (1) | 9974–10492 (1) | 9800–10288 (1) | 9975–10478 (1) |
| *cob* | + | | 10273–11408 (−1) | 10456–11590 (−1) | 10356–11490 (−1) | 10492–11626 (−1) | 10288–11422 (−1) | 10478–11615 (−1) |
| tRNA-S | + | UGA | 11409–11474 (0) | 11591–11657 (0) | 11491–11557 (0) | 11627–11693 (0) | 11423–11489 (0) | 11616–11683 (0) |
| *nad1* | − | | 11493–12440 (18) | 11676–12626 (18) | 11580–12530 (22) | 11712–12662 (18) | 11507–12460 (17) | 11701–12651 (17) |
| tRNA-L | − | UAG | 12442–12506 (1) | 12628–12691 (1) | 12532–12594 (1) | 12664–12728 (1) | 12462–12523 (1) | 12653–12717 (1) |
| *rrnL* | − | | 12507–13828 (0) | 12692–14012 (0) | 12595–13909 (0) | 12729–14025 (0) | 12524–13805 (0) | 12718–14000 (0) |
| tRNA-V | − | UAC | 13829–13898 (0) | 14013–14084 (0) | 13910–13980 (0) | 14026–14095 (0) | 13806–13870 (0) | 14001–14071 (0) |
| *RrnS* | − | | 13899–14689 (0) | 14085–14872 (0) | 13981–14764 (0) | 14096–14876 (0) | 13871–14649 (0) | 14072–14859 (0) |
| control | Not applicable | | 14690–15689 (0) | 14873–15991 (0)[b] | 14765–15735 (0)[b] | 14877–15919 (0) | 14650–15511 (0)[b] | 14860–16887 (0)[b] |

NOTE.—Numbers in parenthesis represent the number of intergenic nucleotides before the gene starts.

[a] This tRNA-S has a UCU anticodon for *Cyphon*, *Chaetosoma*, and *Priasilpha* and a GCU anticodon for *Tetraphalerus*, *Trachypachus*, and *Sphaerius*.

[b] Incomplete control region.

**Table 3**
**AT Content Comparison by mtgenome Region in Coleoptera**

| | Coding Region | | Ribosomal RNAs | | Control Region | |
|---|---|---|---|---|---|---|
| Taxon | Size | AT% | Size | AT% | Size | AT% |
| *Trachypachus* | 14,842 | 79.1 | 2,109 | 81.8 | 1,119[a] | 84.9 |
| *Tetraphalerus* | 14,689 | 66.2 | 2,113 | 66.4 | 1,000 | 78.4 |
| *Sphaerius* | 14,764 | 80.4 | 2,099 | 83.8 | 953[a] | 89.6 |
| *Cyphon* | 14,876 | 74.5 | 2,078 | 80.8 | 1,043 | 85.2 |
| *Chaetosoma* | 14,649 | 78.3 | 2,061 | 82.2 | 862[a] | 91.0 |
| *Priasilpha* | 14,859 | 75.2 | 2,071 | 81.1 | 2,028[a] | 87.0 |
| *Tribolium* | 14,642 | 70.8 | 2,054 | 76.1 | 1,239 | 82.5 |
| *Pyrocoelia* | 16,217 | 76.5 | 2,007 | 81.7 | 1,522 | 87.6 |
| *Crioceris* | 14,660 | 76.4 | 2,081 | 81.4 | 1,220 | 83.3 |
| *Rhagophthalmus* | 14,615 | 78.9 | 2,056 | 82.4 | 1,367 | 86.9 |
| *Pyrophorus* | 14,650 | 68.9 | 2,075 | 83.0 | 1,470 | 74.7 |
| *Anoplophora* | 14,659 | 77.6 | 2,148 | 80.0 | 1,115 | 88.0 |

[a] Incomplete control region.

compositional profiles, but *Trachypachus* (Adephaga) and *Sphaerius* (Myxophaga) exhibit extremely low C and G content in the third codon position. The overall A + T content of *Tetraphalerus* (Archostemata) is the lowest of all, and in this species, the C + G content is not as biased toward first and second codon positions. When the compositional profiles of individual protein-coding genes are examined, it becomes evident that a considerable amount of gene-specific variation exists (fig. 2). For instance, in *Cyphon*, the frequencies of A and T in each codon position are relatively stable, whereas those of G and C vary highly across the protein-coding genes. From these observations, we can hypothesize that there is considerable variability in nucleotide content not only among different species but also among genes and codon positions.

The six beetle species we sequenced, like the seven previously reported, retain the inferred ancestral gene complement for insects (Boore 1999). There were no rearrangements, duplications, or deletions of any genes within these mtgenomes. This suggests that there have not been significant gene rearrangements during the diversification of Coleoptera. Given the diversity of beetles, this molecular stability is a remarkable finding because most other major insect orders exhibit diagnostic rearrangements for major taxonomic groups (Dowton and Austin 1999; Thao et al. 2004; Castro et al. 2006; Cameron and Whiting 2008). In fact, only Diptera appears to be as conservative with respect to mtgenome structure as Coleoptera (Cameron, Lambkin, et al. 2007).

## Noncoding DNA

In our annotations, many gene boundaries have been assigned to avoid the implications of noncoding intergenic spacers and gene overlaps. Mitochondrial evolution has traditionally been viewed as favoring genome size reduction (Rand 1993; Macey et al. 1997; McKnight and Shaffer 1997; Boore 1999), possibly by eliminating intergenic spacers (Burger et al. 2003). From an evolutionary perspective, it makes sense that nonfunctional intergenic spacers would be eliminated over time, especially in the highly reduced and efficient mtgenome. Sometimes intergenic spacers are reduced to the point of gene overlap. However, such cases appear to be the exception rather than the rule, due both to posttranscriptional complications (if abutting genes are encoded on the same strand) as well as the low probability that the nucleotides at the end of one gene are also useful as part of an abutting reversed gene. As such, we attempted to avoid both intergenic spacers and overlaps between genes on either strand of the genome in our annotation, but we did identify a number of intergenic spacer regions of variable size.

Although most spacers appeared to be unique to individual species (see below), a small intergenic region between the tRNA-Ser(UCN) and *nad1* genes, ranging between 17 and 22 bp in length, was found in all six species. An intergenic spacer of this size at this location has been reported in other insects (e.g., Kim et al. 2006) and arthropods (e.g., Lavrov et al. 2000). Four of the six previously published beetle mtgenomes also have this intergenic spacer, which ranges between 16 and 20 bp in size, and only the two species of genus *Rhagophthalmus* lack it (Li et al. 2007). These latter two species have only about 4 bp in this region. However, we can attribute this disparity to insertions and deletions near the end of *nad1* that may be the result of sequencing errors or correction by posttranslational modification. The *Anoplophora* sequence is also one nucleotide off, shifting the reading frame to avoid the conserved stop codon the other beetles use. In this case, the authors annotated the gene with a partial stop codon (T) to preserve the conserved spacer. According to Taanman (1999), this intergenic spacer region may correspond to

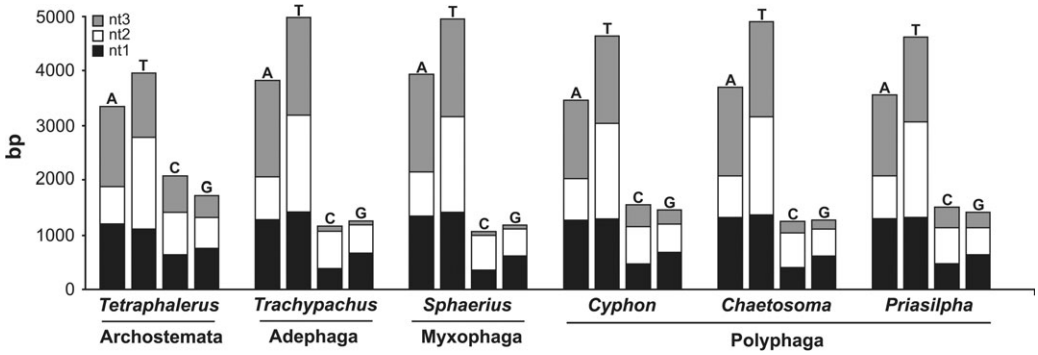


FIG. 1.—Base composition for all protein-coding genes combined. Each column is divided by codon position into three categories.

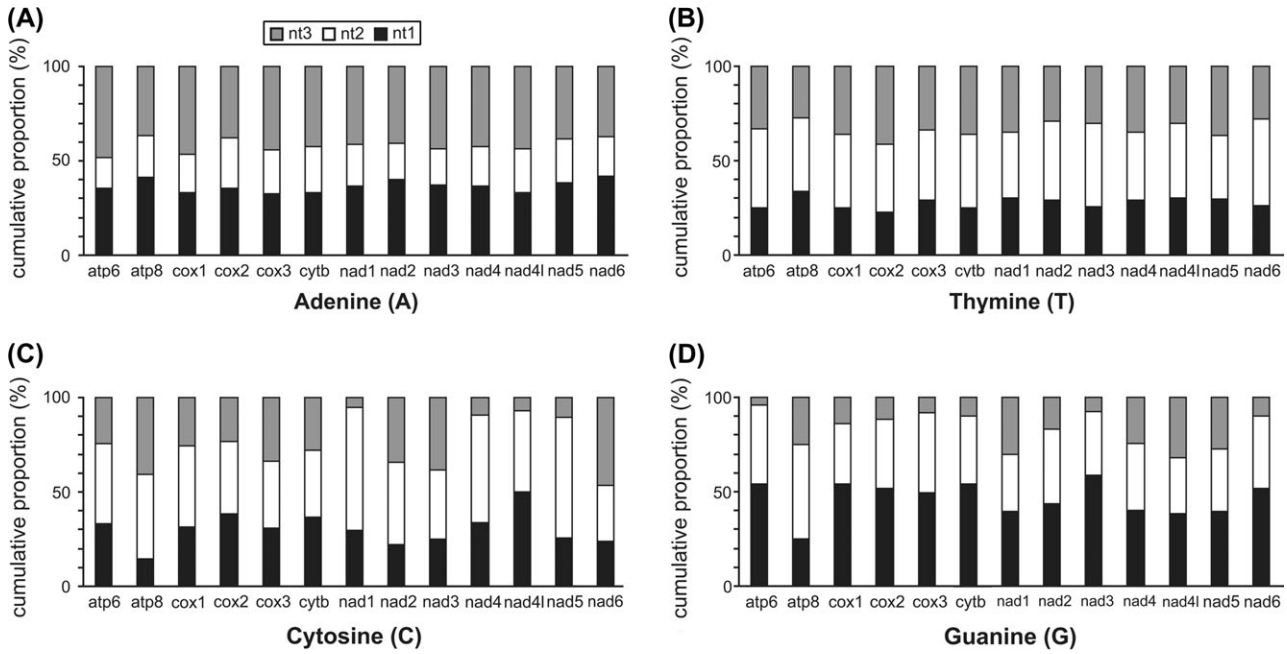


FIG. 2.—Individual base composition for each protein-coding gene in *Cyphon*. Each column is divided by codon position into three categories. To improve visibility, the columns are normalized so that they show proportions rather than counts at each codon position.

the binding site of mtTERM, a transcription attenuation factor, as this position signifies the end of the major-strand coding region. Cameron and Whiting (2008) presented an alignment of several insect orders, highlighting a 7-bp motif (ATACTAA) conserved across Lepidoptera. When we aligned this region across all coleopteran mtgenomes, we found 5 bp (TACTA or its reverse complement TAGTA) to be conserved, and this region matches well the corresponding motif in Lepidoptera (Cameron and Whiting 2008; fig. 3).

In addition to small intergenic regions, there were larger spacer regions of varying A + T content found in different locations in several species (table 2). These regions had no tandem repeats, did not produce any significant BLAST results, did not fold like tRNAs, and did not include open reading frames in either direction, which suggest that they are likely noncoding and nonfunctional. Although noncoding intergenic spacer regions between coding genes have been reported for several insects (e.g., Crozier RH and YC Crozier 1993; Boore 1999; Dotson and Beard 2001; Bae et al. 2004; Cameron, Beckenbach, et al. 2006), their exact origin and function are often unclear. What is evident from this study is that these noncoding regions are lineage specific and common and not conserved at higher taxonomic levels within Coleoptera. Additional sampling will, however, be useful to determine if some of these noncoding regions are conserved across groups of closely related species.

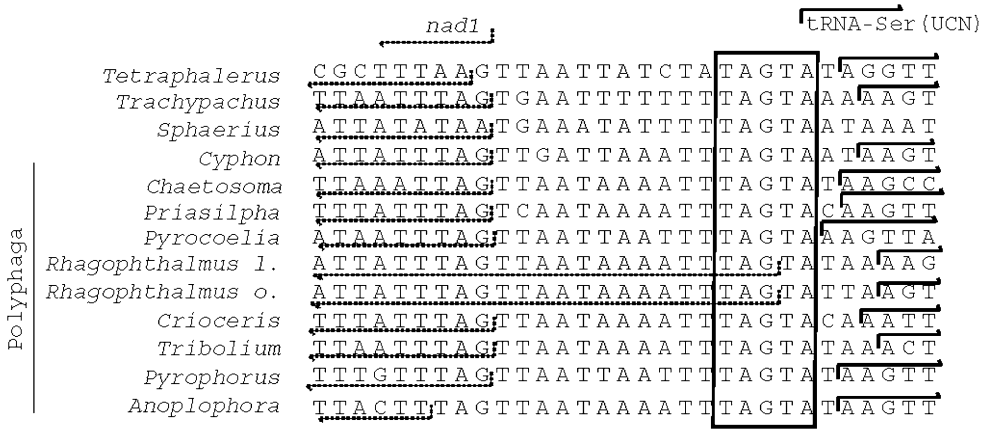


FIG. 3.—An alignment of the gap region between tRNA-Ser(UCN) and *nad1* in all coleopteran genomes. The box indicates a conserved pentanucleotide region (TACTA) across all beetles. The dotted line indicates the location of *nad1* in *Rhagophthalmus* if the current annotation is correct.

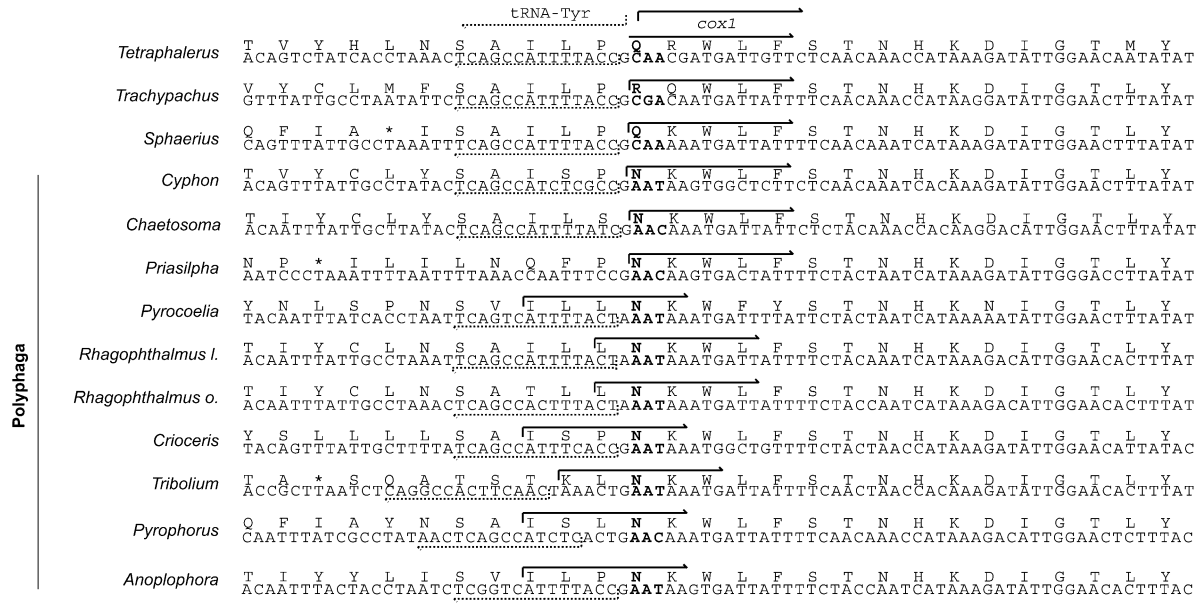


FIG. 4.—An alignment of the 5′ region of *cox1* and the abutting tRNA-Tyr. The dotted line indicates the tRNA; the solid line indicates the beginning of the *cox1* gene as previously proposed. The comparative analysis indicates that the first amino acid after the tRNA (asparagines) is completely conserved across Polyphaga, suggesting a possible molecular synapomorphy for Polyphaga.

## *Cox1* Start Codons

There has been much discussion of potential *cox1* start codons in insects because the beginning of the open reading frame after tRNA-Tyr typically does not have the canonical ATN start codon (e.g., Nardi et al. 2003; Oliveira et al. 2005; Castro et al. 2006; Lee et al. 2006; Fenn et al. 2007; Cameron and Whiting 2008). AAA (lysine), ATT (isoleucine), CTA (leucine), and ATC (isoleucine) have all been proposed as possible start codons in Coleoptera. Without an explicit RNA expression study, it is impossible to determine exactly where *cox1* starts; however, by aligning the region encompassing tRNA-Tyr and *cox1* from all known beetle mtgenomes, we can more accurately determine theoretical start codons for the *cox1* gene in Coleoptera (fig. 4). The possible traditional ATN start codons (isoleucine) near the beginning of cox1 lie either within tRNA-Tyr or 36 bp after the end of the tRNA-Tyr. We argue that it would be most logical to choose a start codon for *cox1* that would minimize intergenic space and gene overlaps. The first nonoverlapping in-frame codon in *cox1* is well conserved throughout all six divergent superfamilies within Polyphaga, and it is possible to choose asparagine (AAT or AAC) as a start codon. At the same site, *Tetraphalerus* and *Sphaerius* have glutamine (CAA) and *Trachypachus* has arginine (CGA). This start location is well conserved, located only a single base pair downstream from the end of the tRNA-Tyr in most species. These codons correspond to the beginning of a highly conserved region, suggesting that this region may be functionally constrained. From our finding, it is possible to hypothesize that asparagine may function as a molecular synapomorphy for Polyphaga.

## Initiation and Termination in Protein-Coding Genes

In insects, most protein-coding genes except *cox1* use typical ATN (methionine or isoleucine) start codons, and we found the same pattern in all six beetle species (table 4). However, there were some genes that varied: *nad1* of *Trachypachus*, *Sphaerius*, *Chaetosoma*, and *Priasilpha* and *nad2* of *Sphaerius*. For these genes, there is no upstream possibility of ATN start codon due to in-frame stop codons, and downstream possibilities all create a considerable intergenic gap. In this study, we propose TTG (leucine) as a start codon for these genes (Okimoto et al. 1990). TTG has been proposed as a start codon for *nad1* in several insects, including *Anopheles quadrimaculatus* (Mitchell et al. 1993), *Tricholepidion gertschi* (Nardi et al. 2003), and *Pyrocoelia rufa* (Bae et al. 2004). In justifying the use of this start codon, Bae et al. (2004) argued from the evolutionary economic perspective that it would minimize intergenic space and avoid overlap with the abutting tRNA. Even more importantly, TTG as a start codon of *nad1* is positionally well conserved as inferred from an alignment of all published beetle mtgenomes (fig. 5). Although some of the previously published mtgenomes (*Crioceris*, *Tribolium*, and *Anoplophora*) annotated *nad1* with a typical ATN start codon which created overlap with tRNA-Leu or a considerable intergenic gap, we suggest that TTG is a more conserved possibility (fig. 5). Additionally, with the revised start codons, the C-terminal end of the peptide is quite conserved with an acidic polar amino acid (D or E) at position 5, and neutral, nonpolar amino acids (I, L, M, V, or F) at positions 1–4 and 6–11 (fig. 5). The evolution of the TTG start codon does not appear to be lineage specific, however. Of the seven polyphagan species, two had the typical ATN (methionine) start codon, whereas the other five had the TTG (leucine) start codon (fig. 5). Different start codons were used in two lineages (*Pyrocoelia* and *Rhagophthalmus*) within the same superfamily (Elateroidea), suggesting that the TTG start codon has evolved multiple times within Coleoptera without much lineage-specific conservation.

**Table 4**
**Start/Stop Codons for Protein-Coding Genes in Six New Beetle Species**

| Gene | *Tetraphalerus* | *Trachypachus* | *Sphaerius* | *Cyphon* | *Chaetosoma* | *Priasilpha* |
|---|---|---|---|---|---|---|
| *nad2* | ATT/T | ATG/TAA | TTG/TAA | ATA/TAA | ATA/T | ATT/TAA |
| *cox1* | CAA/T | CGA/T | CAA/T | AAT/T | AAC/T | AAC/T |
| *cox2* | ATG/T | ATG/T | ATA/T | ATC/T | ATA/T | ATT/TAA |
| *atp8* | ATG/TAA | ATT/TAA | ATT/TAA | ATT/TAA | ATT/TAA | ATT/TAG |
| *atp6* | ATA/TAA | ATA/TAA | ATA/TAA | ATA/TAA | ATA/TAA | ATA/TAA |
| *cox3* | ATG/TA | ATA/TAA | ATG/TA | ATA/TAA | ATG/T | ATG/TA |
| *nad3* | ATT/T | ATT/T | ATT/T | ATT/T | ATA/T | ATT/T |
| *nad5* | ATA/T | ATT/T | ATT/TA | ATT/T | ATT/T | ATA/T |
| *nad4* | ATG/T | ATG/T | ATG/T | ATG/T | ATA/T | ATA/T |
| *nad4l* | ATG/TAA | ATT/TAA | ATT/TAA | ATG/TAA | ATG/TAA | ATG/TAA |
| *nad6* | ATG/TAA | ATT/TAA | ATT/TAA | ATA/TAA | ATA/TAA | ATT/TAA |
| *cob* | ATG/TA | ATG/T | ATG/T | ATG/T | ATG/T | ATG/T |
| *nad1* | ATG/TAA | TTG/TAG | TTG/TAA | ATG/TAG | TTG/TAG | TTG/TAG |

NOTE.—Incomplete stop codons are noted by either T or TA.

The use of incomplete stop codons (T or TA) was frequent in each of the six mtgenomes (table 4), due to ends of protein-coding genes overlapping with the abutting tRNAs. It is hypothesized that a complete stop codon (TAA) is created through posttranscriptional polyadenylation (Ojala et al. 1981). The presence of partial stop codons is well documented in insects (Beard et al. 1993; Coates et al. 2005; Castro et al. 2006). Not surprisingly, complete stop codons were more often TAA than TAG, consistent with patterns found in previously published mtgenomes.

## tRNA-Ser(AGN)

In insect mtgenomes, there are typically 22 tRNAs, with tRNA-Ser and tRNA-Leu 8-fold redundant (two sets of 4-fold redundant tRNAs) (Boore 1999). The length of tRNAs ranged between 60 bp and 75 bp. When compared across all beetle species, including the previously published mtgenomes, we found that the tRNAs were highly conserved within Coleoptera and that all the anticodons were identical and completely conserved, with one exception: the tRNA-Ser(AGN). This particular tRNA was also the most difficult to locate and fold using conventional tRNA search methods such as tRNAscan-SE because it often does not fold into a normal cloverleaf structure due to the absence of stem pairings in the DHU arm (fig. 6). This missing D-stem has been reported in insects (Beard et al. 1993; Crozier RH and Crozier YC 1993; Shao and Barker 2003; Bae et al. 2004), mammals (Chimnaronk et al. 2005; Putz et al. 2007), as well as the rest of Metazoa (Steinberg and Cedergren 1994). Garey

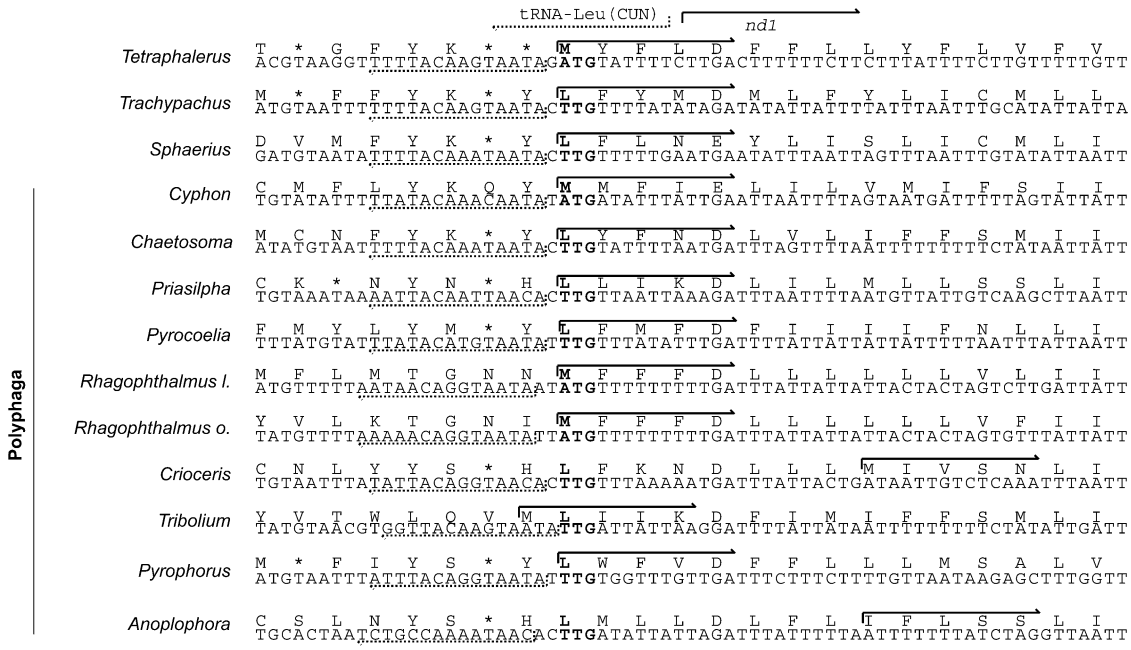


FIG. 5.—An alignment of the tRNA-Leu and *nad1* genes. Dotted line indicates hypothetical amino acid translation of nucleotide sequence that codes for tRNA-Leu. Bold letters indicate the amino acids of the putative start codons that were previously proposed. The box indicates our proposed start codons, which shows that the TTG start codon (leucine) is more common than previously thought.

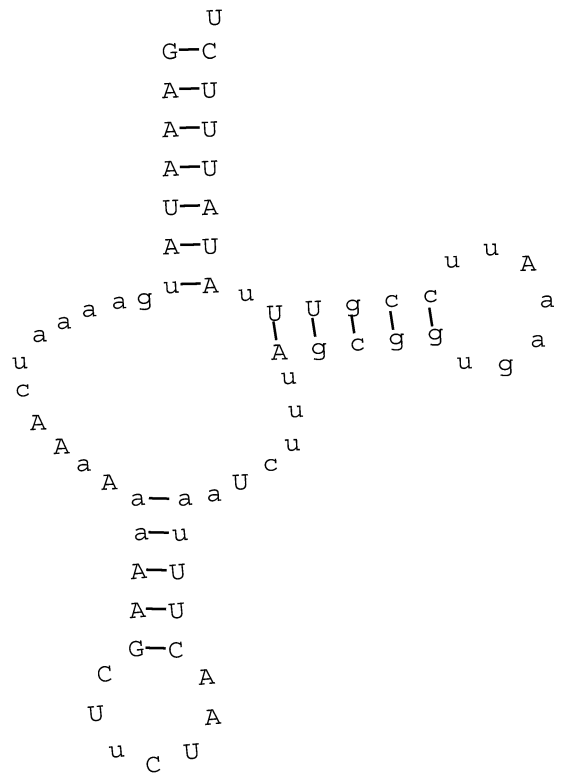

FIG. 6.—Consensus secondary cloverleaf structure for the tRNA-Ser(AGN) gene for all 13 published coleopteran genomes. Capitalized bases are conserved in at least 12 of the 13 sequences; lowercase bases are majority rule. Base pairs may not necessarily match because bases are majority rule.

and Wolstenholme (1989) proposed that the missing D-stem in tRNA-Ser(AGN) evolved very early in the evolution of Metazoa. Despite lacking this stem, however, this tRNA is normally considered to be functional (Steinberg and Cedergren 1994; Stewart and Beckenbach 2003). In an in vitro study, Hanada et al. (2001) found that bovine tRNA-Ser(AGN) (which lacks the D-stem) is functional, although somewhat less effective than other cloverleaf-shaped tRNAs.

tRNAscan-SE is often unable to find tRNA-Ser(AGN) because organellar genome searches in tRNAscan-SE are based on COVE (Eddy and Durbin 1994), which uses a CM to model the structure of typical tRNAs. The general model employed by default is based on a secondary structure alignment of over 1,000 tRNAs from all three domains of life. However, because mitochondrial tRNA-Ser(AGN) is often missing an entire stem, attempting to apply the default CM to this specific class of atypical tRNAs often fails. In order to better understand the consensus structure and ameliorate the problem of finding and folding this tRNA for future mtgenome studies, we constructed a new, specific CM that enables COVE to locate and fold this tRNA in particular. Using COVE with the specific model, we were able to identify and fold tRNA-Ser(AGN) for all 13 species with very good sensitivity (CM available from N.C.S.). Because it is not impressive that a model performs well on the sequences that were used to construct it, we also tested the new CM on additional mtgenomic regions both within Coleoptera (five unpublished mtgenomes encompassing three of the four suborders) and other insect orders including Diptera, Lepidoptera, Hymenoptera, Orthoptera, and Hemiptera (table 5). The new CM was able to identify and fold tRNA-Ser(AGN) in all cases, whereas the default CM often failed to locate it. In cases where the default CM found the tRNA-Ser(AGN), the location was usually slightly different and the resulting secondary structure questionable, whereas the new CM yielded boundaries in greater accordance with published results and secondary structures that match the consensus. We found no false positives in this data set with a COVE score cutoff of 15. Thus, we have shown the new CM to perform well on other insects, despite that fact that it was built using only coleopteran sequences. Perhaps, most importantly, we have demonstrated the utility of specific CMs to facilitate uniform and automated annotation of atypical tRNAs.

**Table 5**
**Results Comparing COVE's Default CM versus Beetle-Specific CM for tRNA-Ser(AGN) Using tRNAscan-SE**

| | Default CM | | | Specific CM | | | Published | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Organism | Start | End | Score | Start | End | Score | Start | End | Classification | Reference |
| *Tribolium castaneum* | — | — | — | 6077 | 6135 | 33.69 | 6077 | 6135 | Coleoptera | Friedrich and Muqim (2003) |
| *Pyrophorus divergens* | — | — | — | 6048 | 6114 | 61.33 | 6048 | 6114 | Coleoptera | Arnoldi et al. (2007) |
| *Drosophila yakuba* | 6199 | 6268 | 9.38 | 6200 | 6267 | 48.97 | 6200 | 6267 | Diptera | Clary and Wolstenholme (1985) |
| *Culicoides arakawae* | — | — | — | 7975 | 8040 | 32.75 | 7985 | 8040 | Diptera | NA |
| *Adoxophyes honmai* | — | — | — | 6180 | 6246 | 50.12 | 6180 | 6246 | Lepidoptera | Lee et al. (2006) |
| *Bombyx mori* | 927 | 995 | 21.56 | 928 | 994 | 49.94 | 928 | 994 | Lepidoptera | NA |
| *Locusta migratoria* | 6115 | 6183 | 5.02 | 6116 | 6182 | 36.25 | 6116 | 6182 | Orthoptera | Flook et al. (1995) |
| *Gryllotalpa orientalis* | 6062 | 6130 | 8.96 | 6063 | 6129 | 55.45 | 6063 | 6129 | Orthoptera | Kim et al. (2005) |
| *Apis mellifera* | 103 | 166 | 3.5 | 117 | 177 | 25.1 | 116 | 178 | Hymenoptera | Crozier RH and Crozier YC (1993) |
| *Philaenus spumarius* | — | — | — | 5992 | 6054 | 34.77 | 5991 | 6055 | Hemiptera | Stewart and Beckenbach (2005) |
| Additional coleopterans | | | | | | | | | | |
| *Hydroscapha* | 6086 | 6154 | 9.27 | 6087 | 6153 | 65.84 | NA | | Coleoptera | NA |
| *Necrophila* | — | — | — | 6082 | 6148 | 61.47 | NA | | Coleoptera | NA |
| *Naupactus* | — | — | — | 6053 | 6121 | 35.52 | NA | | Coleoptera | NA |
| *Calosoma* | 6229 | 6297 | 7.6 | 6230 | 6296 | 69.05 | NA | | Coleoptera | NA |
| *Rhopaea* | — | — | — | 6081 | 6147 | 65.97 | NA | | Coleoptera | NA |

NOTE.—Dashes (—) indicate that tRNAscan-SE with the default CM did not find tRNA-Ser(AGN). NA, not applicable.

Although most arthropods use a GCU anticodon in tRNA-Ser(AGN), all beetle mtgenomes published so far have the UCU anticodon for this tRNA, suggesting that this anticodon may be a molecular synapomorphy for Coleoptera. Outside of Coleoptera, there are a few arthropods that reportedly use a UCU anticodon in tRNA-Ser(AGN), including the sea firefly *Vargula hilgendorfii* (Ogoh and Ohmiya 2004), the hermit crab *Pagurus longicarpus* (Hickerson and Cunningham 2000), and all species of lice studied to date (Cameron, Johnson, and Whiting 2007). With an expanded taxon sampling including all four coleopteran suborders, we found that while all the species belonging to the Polyphaga had the UCU anticodon, *Trachypachus*, *Sphaerius*, and *Tetraphalerus*, representing the smaller suborders Adephaga, Myxophaga, and Archostemata, respectively, had the common GCU anticodon instead (fig. 7). Except for the single base difference, the sequences for anticodon and anticodon loop, as well as the distal three paired bases, were identical across all beetles. Given that most arthropods have the GCU anticodon in the tRNA-Ser(AGN), it is possible to speculate that the ancestral anticodon for Coleoptera was GCU, which mutated to UCU in the common ancestor of Polyphaga, thus serving as a molecular synapomorphy for this suborder.

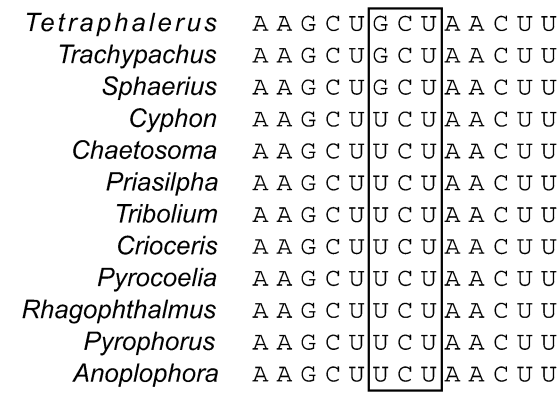


FIG. 7.—An alignment of tRNA-Ser(UCN) anticodon loops (and 3 paired stem nucleotides). Among beetles, Adephaga, Archostemata, and Myxophaga have the common GCU anticodon; all polyphagan species reported to date have the uncommon UCU anticodon, which suggests that this particular anticodon might be a possible molecular synapomorphy for Polyphaga.

### Ribosomal RNAs

The mitochondrial ribosomal RNA genes of beetles are largely uniform across the suborders and similar in secondary structure to those proposed for other insect orders (Gillespie et al. 2006; Cameron and Whiting 2008) (supplementary fig. S1 and S2, Supplementary Material online). The published annotations of 12S for beetles all included additional bases at the 5′ end that would play no functional role in the mature rRNA and so are likely not part of the gene. The 5′ end of the 12S molecule was made up of a short, unpaired leader sequence (4–5 bp) followed by a pseudoknot formed by stem H9 and the 5′ portion of stem H17. This pseudoknot can thus be used in the annotation of the 12S gene with the consensus sequence AAGTT-TDA-TYWT-DRYTT; the first and last 5 bp form the 5′ and 3′ portions, respectively, of stem H9. There was some length variation across the rest of 12S within Coleoptera, with most of variability located in the H47 stem and in the loop regions between H577 and H673. H47 is highly variable between different insect groups—it consists of a short stem and large loop in Hymenoptera (Gillespie et al. 2006) or a long stem and short loop in Lepidoptera (Cameron and Whiting 2008). Most beetles had the long stem form similar to Lepidoptera; however, the elateroid genera (*Pyrocoelia*, *Rhagophthalmus*, and *Pyrophorus*) had the short stem form found in Hymenoptera.

The 16S is more variable than the 12S both across insects and across beetles. The 16S is traditionally annotated as the entire region between adjacent tRNA genes (tRNA-Val and tRNA-Leu(CUN)). This results in considerable length variability in the 5′ end of the gene, approximately 150 bp upstream of the H533 stem. We were able to identify the three stems in this region (H183, H235, and H461); however, there was considerable sequence variability in these stems and length variability in the regions between them. At the 3′ end, there was some length variability between different beetle species; however, all beetles had truncated 16S genes relative to Lepidoptera and Hymenoptera, lacking the 3′ half and most of the loop region of the H2735 stem–loop. The major regions of length variation in beetles were the H837 and H2077 stems–loops as well as the bulge region in the middle of the H991 stem–loop. The large insertion regions and microsatellite regions that distinguish the 16S genes of Lepidoptera were absent, resulting in a much shorter 16S gene in beetles.

## Conclusion

Our study represents the first comprehensive comparative analysis of beetle mtgenomes. We find that Coleoptera follows the ancestral insect arrangement with no deletions or duplications. There are several common features that many beetle lineages share, such as a noncoding region of about 18 bp between *nad1* and tRNA-Leu(CUN) and the usage of a noncanonical TTG start codon. To cope with the atypical structure of tRNA-Ser(AGN), we present a new specific CM for use with COVE and tRNAscan-SE that allows for more consistent identification and secondary structure prediction of this tRNA. We also find that smaller beetle suborders have the common GCU anticodon for tRNA-Ser(AGN), whereas all polyphagans share a rare UCU anticodon variant. We hypothesize that this UCU anticodon of tRNA-Ser(AGN) and asparagine as a start codon for *cox1* are possible molecular synapomorphies for the suborder Polyphaga. Our study demonstrates the importance of comparative analysis in understanding the evolution of mtgenome.

## Acknowledgments

D.R. Maddison, A.L. Wild, and R.A. Leschen kindly provided specimens. This research was funded by the NSF Assembling the Tree of Life (EF-0531665): Beetle Tree of Life Project to M.F.W.

## Literature Cited


Arnoldi FGC, Ogoh K, Ohmiya Y, Viviani VR. 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 JS, Kim I, Sohn HD, Jin BR. 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.

Beard CB, Hamm DM, Collins FH. 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.

Boore J, Brown W. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr Opin Genet Dev. 8:668–674.

Boyce TM, Zwick ME, Aquadro CF. 1989. Mitochondrial DNA in the bark weevils: size, structure, and heteroplasmy. Genetics. 123:825–836.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. Trends Genet. 19:709–716.

Cameron SL, Barker SC, Whiting MF. 2006. Mitochondrial genomics and the relationships and validity of the new insect order Mantophasmatodea. Mol Phylogenet Evol. 38: 274–279.

Cameron SL, Beckenbach AT, Dowton M, Whiting MF. 2006. Evidence from mitochondrial genomics on interordinal relationships in insects. Arthropod Syst Phylogeny. 65:27–34.

Cameron SL, D′Hearse CA, Miller KB, Whiting MF, Barker SC. 2004. Mitochondrial genome data alone are not enough to unambiguously resolve the relationships of Entognatha, Insecta and Crustacea *sensu lato* (Arthropoda). Cladistics. 20:543–557.

Cameron SL, Johnson KP, Whiting MF. 2007. The mitochondrial genome of the screamer louse *Bothriometopus* (phthiraptera: Ischnocera): effects of extensive gene rearrangements on the evolution of the genome. J Mol Evol. 65:589–604.

Cameron SL, Lambkin CL, Barker SC, Whiting MF. 2007. Utility of mitochondrial genomes as phylogenetic markers for insect intraordinal relationships—a case study from flies (Diptera). Syst Entomol. 32:40–59.

Cameron SL, Whiting MF. 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.

Castro LR, Ruberu K, Dowton M. 2006. Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vanhorniidae) and *Primeuchroeus spp*. (Aculeata: Chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). Genome. 49:752–766.

Chimnaronk S, Jeppesen MG, Suzuki T, Nyborg J, Watanabe K. 2005. Dual-mode recognition of noncanonical tRNAs-Ser by seryl-tRNA synthetase in mammlian mitochondria. EMBO J. 24:3369–3379.

Clary DO, Wolstenholme DR. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J Mol Evol. 22:252–271.

Coates BS, Sumerford DV, Hellmich RL, Lewis LC. 2005. Partial mitochondrial genome sequences of *Ostrinia nubilalis* and *Ostrinia furnicalis*. Int J Biol Sci. 1:13–18.

Crozier RH, Crozier YC. 1993. The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. Genetics. 133:97–117.

Dotson EM, Beard CB. 2001. Sequence of organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Mol Biol. 10:205–215.

Dowton M, Austin AD. 1999. Evolutionary dynamics of a mitochondrial rearrangement “hot spot” in the Hymenoptera. Mol Biol Evol. 16:298–309.

Dowton M, Castro LR, Austin AD. 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome ‘morphology’. Invertebr Syst. 16:345–356.

Eddy SR. 2002. A memory-efficient dynamic programming algorithm for optimal alignment of a sequence to an RNA secondary structure. BMC Bioinformatics. 3:18.

Eddy SR, Durbin R. 1994. RNA sequence analysis using covariance models. Nucleic Acids Res. 22:2079–2088.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–97.

Fenn JD, Cameron SL, Whiting MF. 2007. The complete mitochondrial genome sequence of the Mormon cricket (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability. Insect Mol Biol. 16:239–252.

Flook PK, Rowell CHF, Gellissen G. 1995. The sequence, organization, and evolution of the *Locusta migratoria* mitochondrial genome. J Mol Evol. 41:928–941.

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.

Garey JR, Wolstenholme DR. 1989. Platyhelminth mitochondrial DNA: evidence for early evolutionary origin of a tRNA(ser-AGN) that contains a dihydrouridine arm replacement loop, and of serine-specifying AGA and AGG codons. J Mol Evol. 28:374–387.

Gibson A, Gowri-Shankar VG, Higgs PG, Rattray M. 2004. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. Mol Biol Evol. 22:251–264.

Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR. 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 Mol Biol. 15:657–686.

Gowri-Shankar V, Rattray M. 2006. On the correlation between composition and site-specific evolutionary rate: implications for phylogenetic inference. Mol Biol Evol. 23:352–364.

Gray MW, Burger G, Lang BF. 1999. Mitochondrial evolution. Science. 283:1476–1481.

Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A. 2005. Rfam: annotating non-coding RNAs in complete genomes. Nucleic Acids Res. 33:D121–D124.

Hanada T, Suzuki T, Yokogawa T, Takemoto-Hori C, Sprinzl M, Watanabe K. 2001. Translation ability of mitochondrial tRNAs-Ser with unusual secondary structures in an *in vitro* translation system of bovine mitochondria. Genes Cells. 6:1019–1030.

Hickerson MJ, Cunningham CW. 2000. Dramatic Mitochandrial Gene Rearrangements in the Hermit Crab *Pagurus longicarpus* (Crustacea, Anomura). Mol Biol. Evol. 17:639–644.

Hwang UW, Park CJ, Yong TS, Kim W. 2001. One-step PCR amplification of complete arthropod mitochondrial genomes. Mol Phylogenet Evol. 19:345–352.

Jia W, Higgs PG. 2007. Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. Mol Biol Evol. 25:339–351.

Kim I, Cha SY, Yoon MH, Hwang JS, Lee SM, Sohn HD, Jin BR. 2005. The complete nucleotide sequence and gene organization of the mitochondrial genome of the oriental mole

cricket, *Gryllotalpa orientalis* (Orthoptera: Gryllotalpidae). Gene. 2:155–168.
Kim I, Lee EM, Seol KY, Yun EY, Lee YB, Hwang JS, Jin BR. 2006. The mitochondrial genome of the Korean hairstreak, *Coreana raphaelis* (Lepidoptera: Lycaenidae). Insect Mol Biol. 15:217–225.
Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform. 5:150–163.
Lavrov DV, Boore JL, Brown WM. 2000. The complete mitochondrial DNA sequence of the horseshoe crab *Limulus polyphemus*. Mol Biol Evol. 17:813–824.
Lawrence JF, Newton AF. 1982. Evolution and classification of beetles. Ann Rev Eco Syst. 13:261–290.
Lee ES, Shin KS, Kim MS, Park H, Cho S, Kim CB. 2006. The mitochondrial genome of the smaller tea tortrix *Adoxophyes honmai* (Lepidoptera: Tortricidae). Gene. 373:52–57.
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.
Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genome sequences. Nucleic Acids Res. 25:955–964.
Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ. 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol. 14:91–104.
McKnight ML, Shaffer HB. 1997. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). Mol Biol Evol. 14:1167–1176.
Mitchell SE, Cockburn AF, Seawright JA. 1993. The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization. Genome. 36(6):1058–1073.
Mueller RL, Boore JL. 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Mol Biol Evol. 22:2104–2112.
Nardi F, Carapelli A, Dallai R, Frati F. 2003. The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations. Insect Mol Biol. 12:605–611.
Ogoh K, Ohmiya Y. 2004. Complete mitochondrial DNA sequence of the sea-firefly *Vargulahilgendorfii* (Crusacea Ostracoda) with duplicate control regions. Gene. 327:131–139.
Okimoto R, Macfarlane JL, Wolstenholme DR. 1990. Evidence for the frequent use of TTG as the translation initiation codon of mitochondrial protein genes in the nematodes, *Ascaris suum* and *Caenorhabditis elegans*. Nucleic Acids Res. 18:6113–6118.
Oliveira MT, Azeredo-Espin AML, Lessinger AC. 2005. Evolutionary and structural analysis of the cytochrome c oxidase subunit I (COI) gene from *Haematobia irritans*, *Stomoxys calcitrans*, and *Musca domestica* (Diptera: Muscidae) mitochondrial DNA. DNA Seq. 16(2):156–160.
Podsiadlowski L. 2006. The mitochondrial genome of the bristletail *Petrobius brevistylis* (Archaeognatha: Machilidae). Insect Mol Biol. 15:253–258.
Putz J, Dupuis B, Sissler M, Florentz C. 2007. Mamit-tRNA, a database of mammalian mitochondiral tRNA primary and secondary structures. RNA. 13:1184–1190.
Rand DM. 1993. Endotherms, ectotherms, and mitochondrial genome-size variation. J Mol Evol. 37:281–295.
Shao R, Barker SC. 2003. The highly rearranged mitochondrial genome of the plague thrips, *Thrips imagines* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol Biol Evol. 20:362–370.
Steinberg S, Cedergren R. 1994. Structural compensation in atypical mitochondrial tRNAs. Nat Struct Biol. 1:507–510.
Stewart JB, Beckenbach AT. 2003. Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*. Mol Phylogenet Evol. 26:513–526.
Stewart JB, Beckenbach AT. 2005. Insect mitochondrial genomics: the complete mitochondrial genome sequence of the meadow spittlebug *Philaenus spumarius* (Hemiptera: Auchenorrhyncha: Cercopoidae). Genome. 48:46–54.
Taanman JW. 1999. The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta. 1410:103–123.
Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 20:3252–3255.
Yamauchi MM, Miya MU, 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 Mol Biol. 13:435–442.
Zhang DX, Hewitt GM. 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. Biochem Syst Ecol. 25:99–120.

Dr. Jennifer Wernegreen, Associate Editor

Accepted August 30, 2008