Insect Molecular Biology (2004) **13**(5), 495–518

A secondary structural model of the 28S rRNA expansion segments D2 and D3 from rootworms and related leaf beetles (Coleoptera: Chrysomelidae; Galerucinae)

J. Gillespie*, J. Cannone†, R. Gutell† and A. Cognato*

**Department of Entomology, Texas A&M University, College Station, TX, USA; and* †*Institute for Cellular and Molecular Biology and Section of Integrative Biology, University of Texas, Austin, TX, USA*

Abstract

We analysed the secondary structure of two expansion segments (D2, D3) of the 28S rRNA gene from 229 leaf beetles (Coleoptera: Chrysomelidae), the majority of which are in the subfamily Galerucinae. The sequences were compared in a multiple sequence alignment, with secondary structure inferred primarily from the compensatory base changes in the conserved helices of the rRNA molecules. This comparative approach yielded thirty helices comprised of base pairs with positional covariation. Based on these leaf beetle sequences, we report an annotated secondary structural model for the D2 and D3 expansion segments that will prove useful in assigning positional nucleotide homology for phylogeny reconstruction in these and closely related beetle taxa. This predicted structure, consisting of seven major compound helices, is mostly consistent with previously proposed models for the D2 and D3 expansion segments in insects. Despite a lack of conservation in the primary structure of these regions of insect 28S rRNA, the evolution of the secondary structure of these seven major motifs may be informative above the nucleotide level for higher-order phylogeny reconstruction of major insect lineages.

Keywords: rRNA, ribosome, rootworms, secondary structure, expansion segment, homology.

Received 6 April 2004; accepted after revision 11 June 2004. Correspondence: Joseph J. Gillespie, Department of Entomology, Texas A&M University, College Station, TX 77843, USA. Tel.: +1 979 458 0579; fax: +1 979 845 6305; e-mail: pvittata@gmail.com

Note: A website is available at http://hisl.tamu.edu

Introduction

The nuclear-encoded ribosomal large subunit (LSU) rRNA-encoding gene (23S-like rRNA) varies greatly in sequence length and nucleotide composition within the main eukaryote lineages (Ware *et al*., 1983; Clark *et al*., 1984; Hassouna *et al*., 1984). The length heterogeneity in eukaryotic lineages is isolated to specific regions of the LSU rRNA (Clark, 1987; Gorski *et al*., 1987; Michot & Bachellerie, 1987; Hancock & Dover, 1988; Tautz *et al*., 1988; Gutell & Fox, 1988), of which some are referred to as expansion segments (Clark *et al*., 1984). Although these regions of the rRNA are usually not associated with protein translation (Gerbi, 1985), site-directed mutagenesis studies have implicated one of these highly variable regions with function (Sweeney *et al*., 1994). In addition, the structure in these regions with less sequence conservation and more length variation is more variable than the structure in the regions with more sequence conservation and less length variation.

The eukaryotic rDNA occurs as a multigene family of tandemly repeated units of the 23S-like, 16S-like and 5.8S rRNA transcripts that evolve concertedly (Arnheim *et al*., 1980; Dover, 1982; Arnheim, 1983; Flavell, 1986). These tandem arrays, termed nucleolar organization regions (NORs), are located on chromosomes in hundreds to thousands of copies throughout the genome, with copy number dependent on the organism in question. Unequal crossing over and gene conversion keep the many copies of NORs conserved within species (Dover, 1982). The three functional rRNA transcripts are separated by internally transcribed spacers (ITSs) that are spliced out of the transcripts after NOR expression. Although all three transcripts contain regions of variability (in base composition and sequence length), the 23S-like transcript has thirteen expansion segments, as well as nine other identified variable regions (Schnare *et al*., 1996), of rapidly evolving sequence and is the most variable of the nuclear rRNA genes (Mindell & Honeycutt, 1990). This variation is associated with a wide range of phylogenetically informative characters among higher taxonomic levels (De Rijk *et al*., 1995; Schnare *et al*., 1996; Kuzoff *et al*., 1998).

The thirteen expansion segments of the 28S rRNA vary greatly among insect orders (Hwang *et al*., 1998; J. Gillespie,

unpubl. data), as well as within Diptera (Tautz *et al*., 1988; Kjer *et al*., 1994; Schnare *et al*., 1996) and Hymenoptera (Belshaw & Quicke, 2002; J. Gillespie, unpubl. data). As in other eukaryotes, the expansion segments in insects are more variable than the core rRNA, but are constrained structurally, with deleterious mutations often accommodated by compensatory base changes that maintain helical formation (Hancock *et al*., 1988; Tautz *et al*., 1988; Rousset *et al*., 1991; Kjer *et al*., 1994). This duality of variability and conservation makes these regions ideal for phylogenetic reconstruction among insects because the variation yields phylogenetic information and structural conservation helps the assessment of nucleotide homology. For example, the 28S-D1 and D3 regions have been utilized in the reconstruction of Trichoptera phylogeny (Kjer *et al*., 2001), and the 28S-D2 region has been used to resolve tribal relationships within galerucine leaf beetles (Gillespie *et al*., 2001, 2003, 2004). However, their use in phylogeny reconstruction of Insecta is often problematic owing to the difficulty of alignment of multiple sequences from divergent taxa (De Rijk *et al*., 1995). This problem derives from the variability within the expansion segments, particularly in the distal regions of expanding and contracting hairpin-stem loop motifs (Crease & Taylor, 1998; Gillespie, 2004). Thus, unlike the alignment of highly conserved core regions of rRNA molecules, the expansion segments require inspection for compensatory base changes that facilitate the alignment of highly divergent sequences. Co-evolving helices and highly conserved single-stranded regions empirically provide homology assignments that delimit unalignable regions (Kjer, 1995, 1997). After initial exclusion, these subsequent alignment-ambiguous regions can be incorporated into phylogeny reconstruction in a variety of ways. They can be recoded as multistate characters based on nucleotide identity (Lutzoni *et al*., 2000; Kjer *et al*., 2001; Gillespie *et al*., 2003, 2004), and further subjected to a step matrix that implements unequivocal weighting to character transformations (Lutzoni *et al*., 2000; Gillespie *et al*., 2003, 2004; Xia *et al*., 2003; Sorenson *et al*., 2003). Unalignable regions can also be recoded as morphological characters based on the differences these regions impose on the secondary structure of the molecule (Billoud *et al*., 2000; Collins *et al*., 2000; Lydeard *et al*., 2000; Ouvrard *et al*., 2000). Across taxa, transformations from one structure to another can be calculated as a measure of structural variability (Fontana *et al*., 1993; Notredame *et al*., 1997; Moulton *et al*., 2000; Misof & Fleck, 2003). Homologous, yet unalignable structures can even be characterized as phylogenetic trees, with differences in tree topology representing transformations across variable structures (Shapiro & Zhang, 1990; Hofacker *et al*., 1994).

In this study, we present a structural model for the expansion segments D2 and D3 of the 28S rRNA gene from 229 leaf beetles (Coleoptera: Chrysomelidae), the majority of which are found in the subfamily Galerucinae. This model is a refined annotation from previous studies that incorporated secondary structure to improve homology assignment for phylogeny reconstruction of these beetles (Gillespie, 2001; Gillespie *et al*., 2003a, 2004; Kim *et al*., 2003). Using compensatory base change evidence, we define conserved regions of the molecule that provide a custom chrysomelid model for this region of the 28S rRNA gene. Our novel characterization of regions of alignment ambiguity (RAA), slipped-strand compensation (RSC) and expansion and contraction (REC) from structural homology is discussed within taxonomic and phylogenetic contexts. This model will be useful for future studies on related beetle groups that utilize the D2 and D3 expansion segments for phylogeny reconstruction, and for studies that address expansion segment evolution across higher-level insect taxa (Misof & Fleck, 2003).

Results and discussion

Predicted secondary structure

The first nearly complete predicted secondary structural model of the eukaryotic cytoplasmic LSU rRNA from a beetle, the tenebrionid *Tenebrio* sp., is shown here (Fig. 1) in concordance with the conserved 23S and 23S-like structures of the LSU rRNA from the literature (Wool, 1986; Gutell & Fox, 1988; Gutell *et al*., 1990, 1992a,b, 1993; Schnare *et al*., 1996). With existing predicted structures for *Drosophila melanogaster* (Schnare *et al*., 1996, and references therein), *Aedes albopictus* (Kjer *et al*., 1994), and *Acyrthosiphon pisum* (Amako *et al*., 1996), this is the fourth predicted structure of the 28S LSU rRNA from an insect. The expansion segments D2 and D3 are highlighted and correspond, respectively, to the variable regions 545 and 650 of Schnare *et al*. (1996), which refer to the sequence numbering of *E. coli* LSU rRNA (Fig. 1). A multiple sequence alignment spanning the two expansion segments was generated from 229 chrysomelid taxa; however, six sampled taxa are listed for brevity (Fig. 2). The entire alignment is posted in a variety of electronic formats at http://hisl.tamu.edu and http://www.rna.icmb.utexas.edu/.

Of the 864 positions in the *Diabrotica undecimpunctata howardi* reference sequence, we have identified 676 nucleotide positions in the 28S-D2,D3 sequence alignment that can be confidently assigned positional homology across the beetle taxa. Of the remaining length-variable positions, eighteen RAAs, one RSC and two RECs were identified and excluded from primary homology assignment. The thirty conserved helices within the D2 and D3 expansion segments of the 28S rRNA gene are illustrated on a two-dimensional structural model, which also includes the core regions of the 28S between the D2 and D3 and flanking the D3 in the 3′ direction (Fig. 3). Less compensatory base

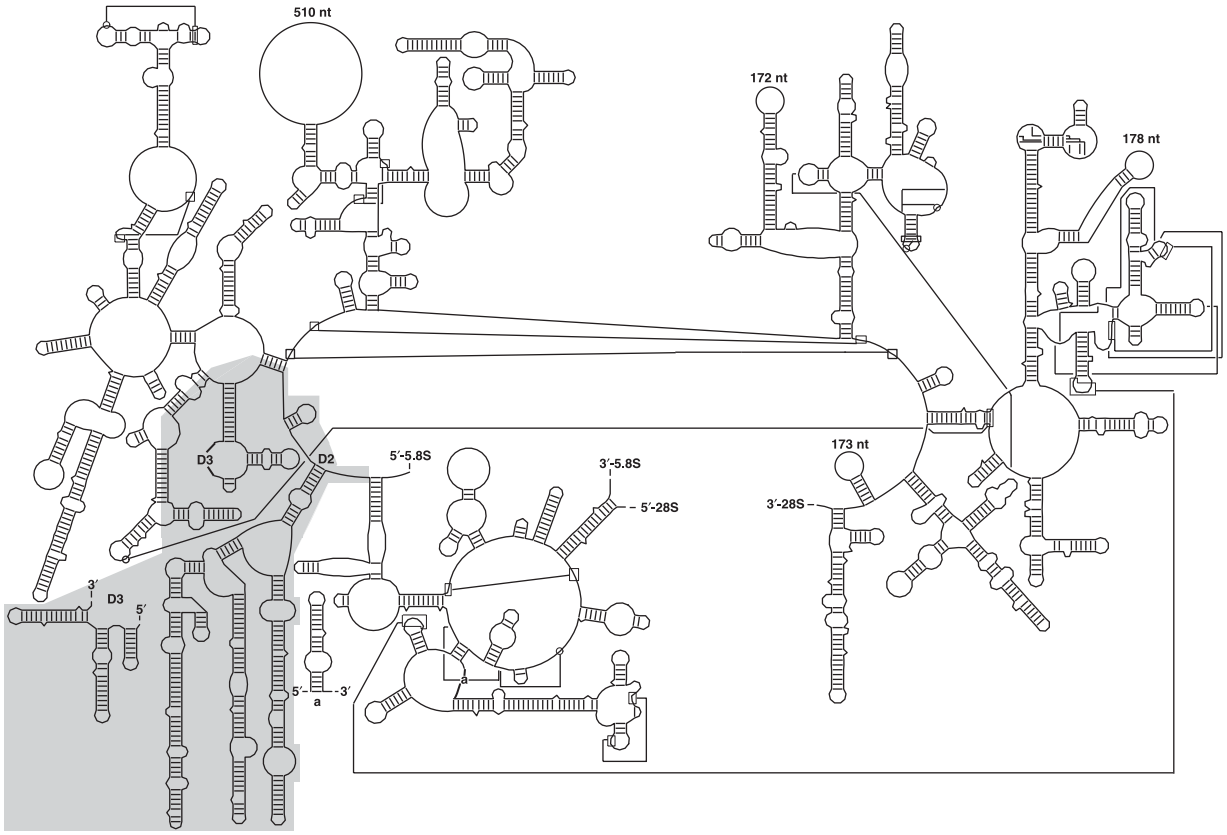


Figure 1. A schematic line drawing of the secondary structure of LSU 28S rRNA from the beetle *Tenebrio* sp. (accession number AY210843). The shaded region shows the expansion segments D2 and D3 (regions 545 and 650, respectively, of Schnare *et al*., 1996) and related core sequence that were analysed in this study. Base-pairing (where there is strong comparative support) and tertiary interactions that link the 5′- and 3′-halves of the molecule are shown connected by continuous lines. Structures for the expansion segments D7a, D7b, D8, D10 and D12 are preliminary at this time (most structures are shown as arcs or loops, with numbers indicating size). These structures will be adjusted when more beetle sequences from these regions are made available.

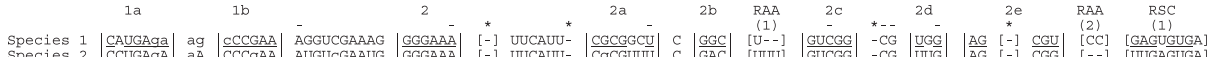
change evidence is found within the D3 expansion segment because many of the analysed sequences are from studies that only included the D2 expansion segment (Gillespie *et al*., 2003, 2004; Kim *et al*., 2003).

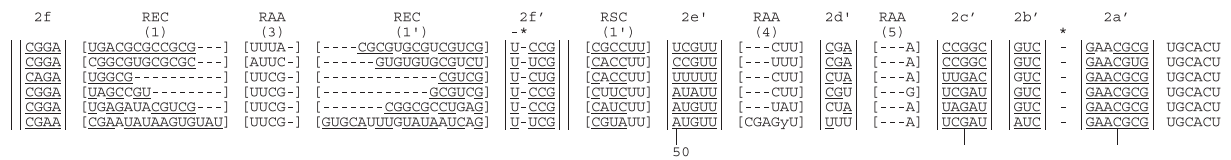
Expansion segment D2

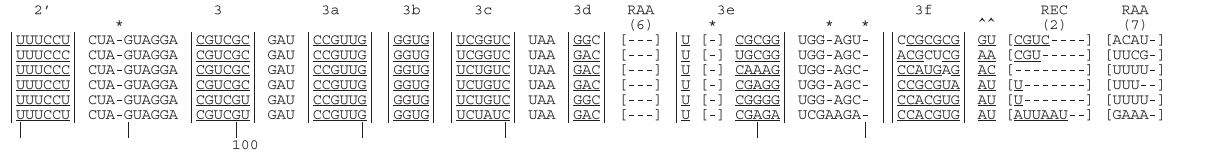
The 28S-D2 segment, corresponding to the 545 variable region of the 23S-like LSU (Schnare *et al*., 1996), comprises four main compound helices that are flanked by highly conserved elements in the 28S core structure. These motifs are labelled 'helix 1', 'helix 2', 'helix 3-1' and 'helix 3-2', and the subcomponents of the compound helices are named a, b, c, etc. (Fig. 3). A total of 26 conserved helical elements comprise the D2 region in chrysomelids (but see below regarding helix 3q in *A. coerulea*). The innermost helix of D2, named here as helices 1a and 1b (helix A in Schnare *et al*., 1996), could not be evaluated for compensatory base changes owing to the prevalence of unknown nucleotide assignments in electropherograms because of the close proximity of the 5′-primer to strand 1.

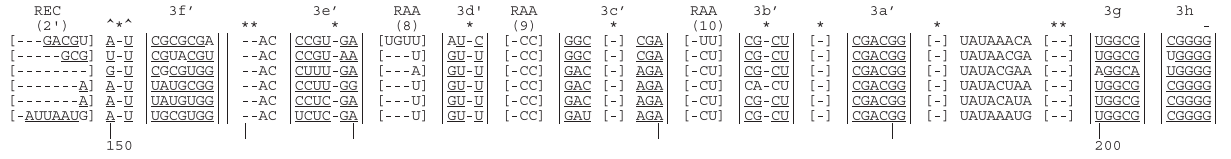
Helix 2 in the D2 region is at the base of the second compound helix and comprises six basepairs across nearly all holometabolous insects (J. Gillespie, unpubl. data). The chrysomelids contain six helices that are apical to helix 2 (2a–2f). Many of the basepairs within these helices are supported with positional covariation. A gallery of structures representing the 'helix 2' motif is presented in Fig. 4. The terminal helix in this motif, helix 2f, has the potential to form additional basepairings beyond the four boxed basepairs; however, a confident homology assignment is not possible here owing to the high sequence and length variation in this region (see REC 1 below). One RSC, one REC and six RAAs occur in 'helix 2' (Fig. 4F).

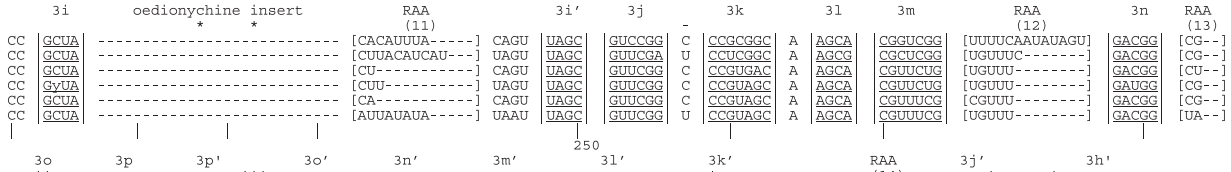
Helix 3 (H2 in Michot & Bachellerie, 1987; E in Schnare *et al*., 1996) is highly conserved in the higher eukaryotes and is the most basal helix to several compound helices (Schnare *et al*., 1996; J. Gillespie, unpubl. data). Helix 3 is six basepairs long in the chrysomelids and most holometabolous insect lineages (J. Gillespie, unpubl. data). The chrysomelids have two compound helices distal to helix 3, 'helix 3-1' (helices 3a–3f) and 'helix 3-2' (helices 3g–3p) (Fig. 3). A gallery of representative 'helix 3-1' structures for different chrysomelids is displayed in Fig. 5. The terminal helix in 'helix 3-1', 3f, has the potential to form additional

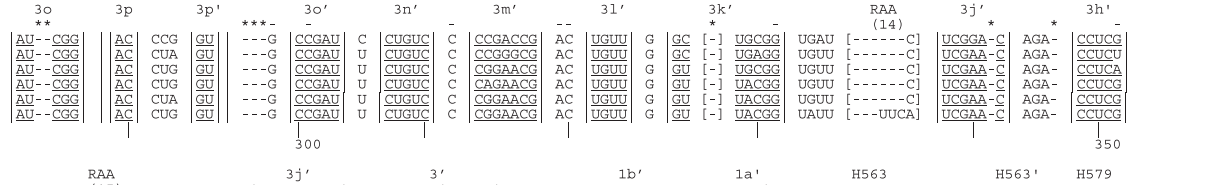


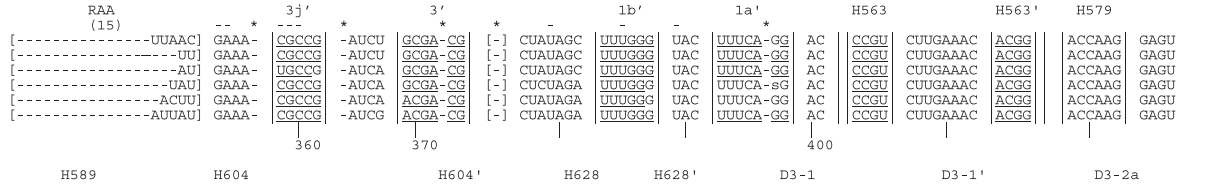


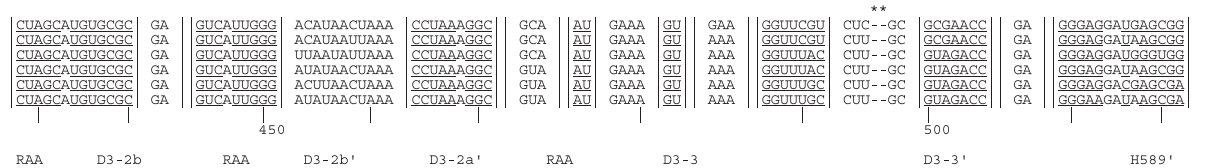


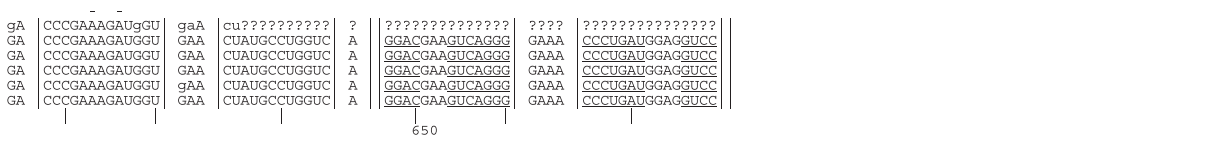












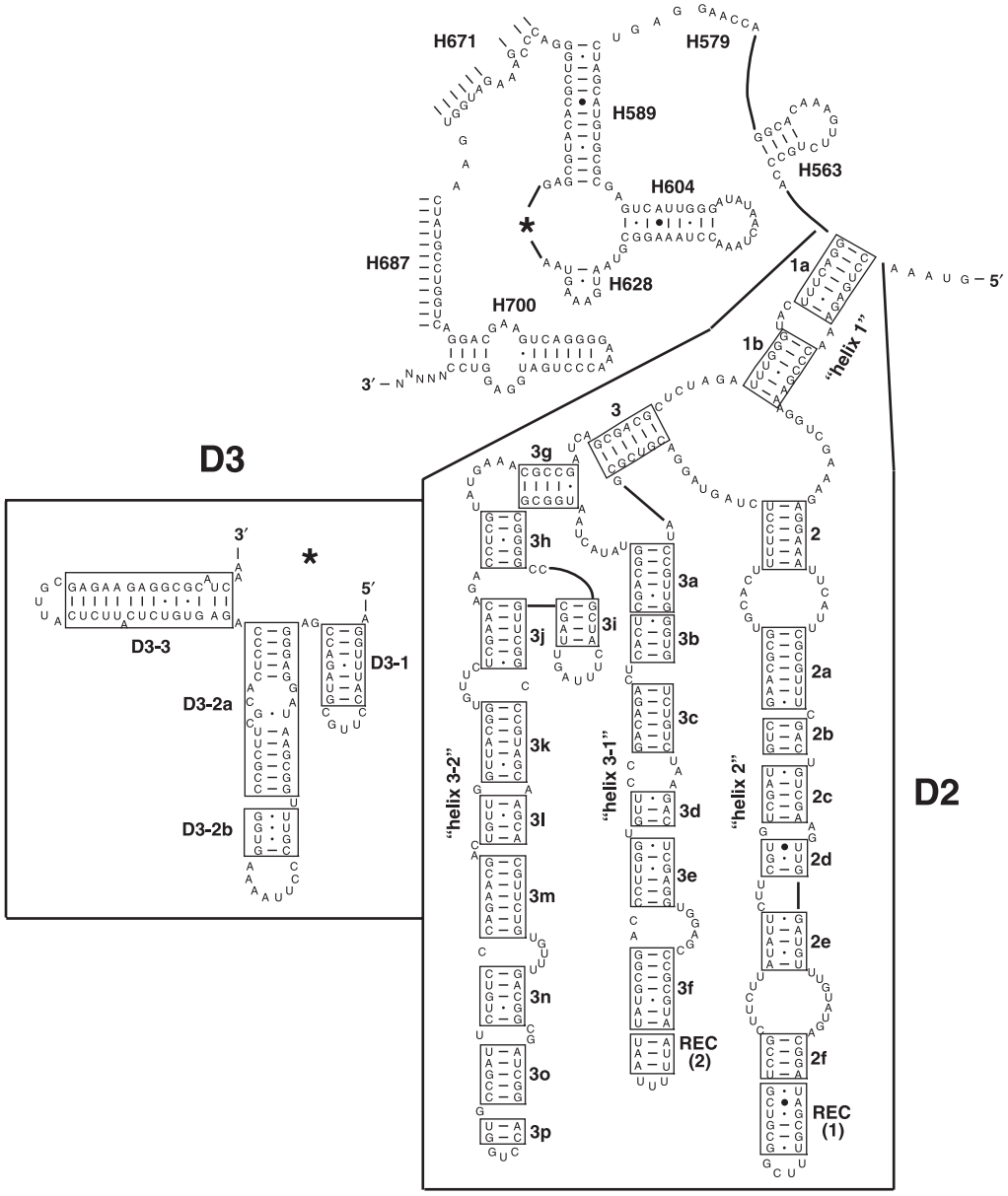


Figure 3. The secondary structure model of the expansion segments D2 and D3 of the LSU 28S nuclear rRNA gene from spotted cucumber beetle (*Diabrotica undecimpunctata howardi*). The thirty conserved, covarying helices present in all of the beetle taxa studied here are boxed. Helix notation is modified from Gillespie *et al.* (2003, 2004) (see Fig. 2). Regions of core rRNA between the two expansion segments and flanking the 3′ end of the D3 are numbered following Cannone *et al.* (2002). Base-pairing is indicated as follows: standard canonical pairs by lines (C-G, G-C, A-U, U-A); wobble G·U pairs by dots (G·U); A·G pairs by open circles (A°G); other non-canonical pairs by filled circles (e.g. C•A). Diagram was generated using the program XRNA (B. Weiser & H. Noller, University of California at Santa Cruz).

Figure 2. Multiple sequence alignment of primary and secondary structure of the expansion segments D2 and D3 of the LSU 28S nuclear rRNA gene from six chrysomelid species (*Lamprosoma* sp., *Metaxyonycha panamensis*, *Epitrix fasciata*, *Diabrotica adelpha*, *Pyrrhalta aenescens*, *Neolochmaea dilatipennis*). Regions of core rRNA between the two expansion segments and flanking the 3′ end of D3 are numbered following Cannone *et al.* (2002). The notation for the twenty-six conserved helices within the expansion segment D2 is modified from Gillespie *et al.* (2003) with slight annotations to the previous predicted structure (Fig. 3). Helices with long range interactions are placed within bars (|) and immediate hairpin-stem loops are placed within double bars (||). All complemenatry strands are depicted with a prime (′; e.g. strand 1 hydrogen bonds with strand 1′ to form helix 1). Regions of alignment ambiguity (RAA), slipped-strand compensation (RSC) and expansion and contraction (REC) are placed within square brackets. Nucleotides within helices involved in hydrogen-bonding are underlined. Single insertions (*) and deletions (–) are noted as in Kjer *et al.* (2001). Positions that can form an expansion of a helix across some but not all taxa are labelled with a caret (^). Every tenth nucleotide assigned positional homology is noted under the alignment with a tick (|), with every 50th position numbered. The sequences are 5′ to 3′ in direction. Missing nucleotides are represented with question marks (?). Lower-case letters depict nucleotides confirmed by one strand only in sequencing. Note: this alignment has not been amended for these six taxa from the original alignment of 229 chrysomelid sequences, and thus gaps and insertions may correspond to taxa not presented in this figure.

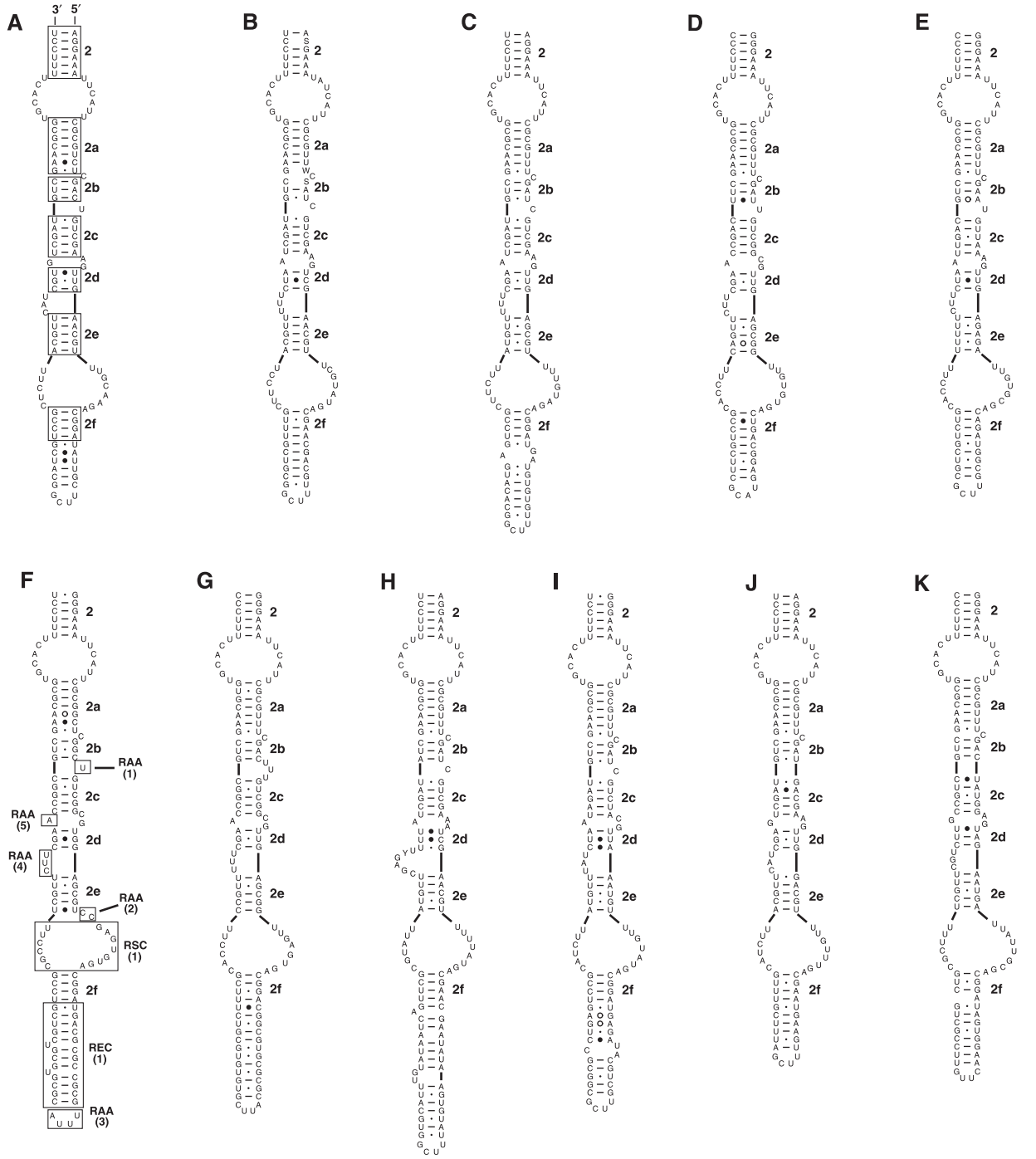


Figure 4. A gallery of diverse secondary structure diagrams from the 'helix 2' compound helix in the D2 region (synonymous with the 545 gallery of Schnare *et al.*, 1996) is shown for the following chrysomelid taxa: (A) *Acalymma vittata*, (B) *Agelastica coerulea*, (C) *Cerochroa brachialis*, (D) *Coptocycla adamantina*, (E) *Epitrix fasciata*, (F) *Lamprosoma* sp., (G) *Metaxyonycha panamensis*, (H) *Neolochmaea dilatipennis*, (I) *Pyrrhalta aenescens*, (J) Thailand specimen 11, (K) *Walterianella bucki*. Notation for the seven helical elements is modified from Gillespie *et al.* (2003, 2004). Helices are boxed in A, and ambiguously aligned regions are boxed in F. The notation for RAAs, RSCs and RECs is described in Fig. 2 and Table 3. The explanations of basepair symbols and reference for software used to construct structure diagrams are given in Fig. 3.

basepairings beyond the seven boxed positions; however, this homology assignment is ambiguous for the positions identified in REC (two) and RAA (seven) (distal to the 3f boxed basepairs in Fig. 5G) owing to the lack of sequence conservation and the variation in sequence lengths. Although most taxa in the alignment append two more basepairs to helix 3f, the taxon *Eucerotoma* sp. 344 (Fig. 5L) has only seven basepairs in helix 3f. Thus, we limited helix 3f to

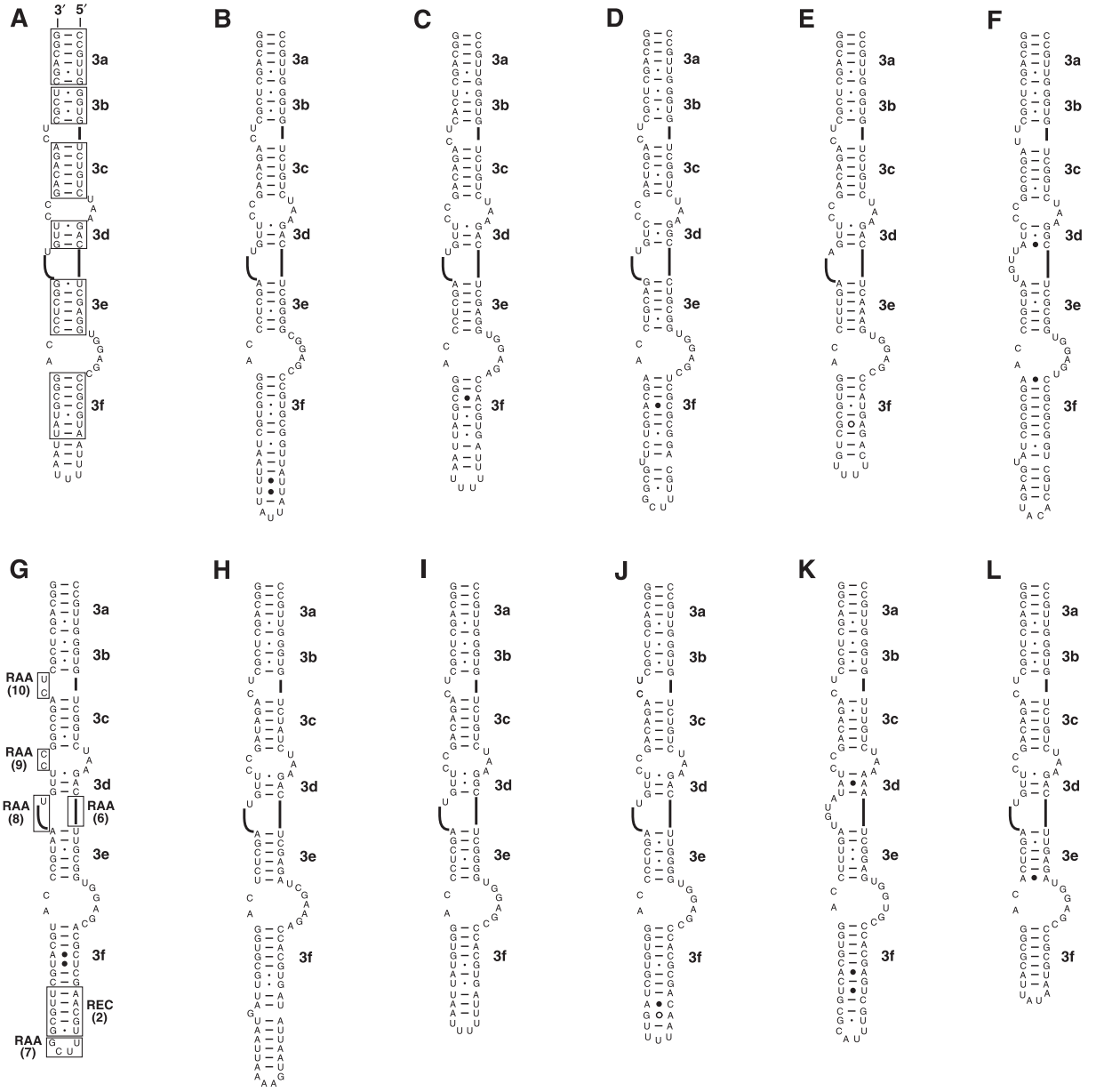


Figure 5. A gallery of diverse secondary structure diagrams from the 'helix 3-1' compound helix in the D2 region (synonymous with the 545 gallery of Schnare *et al.*, 1996) is shown for the following chrysomelid taxa: (A) *Acalymma vittata*, (B) *Agelastica coerulea*, (C) *Cerochroa brachialis*, (D) *Coptocycla adamantina*, (E) *Epitrix fasciata*, (F) *Lamprosoma* sp., (G) *Metaxyonycha panamensis*, (H) *Neolochmaea dilatipennis*, (I) *Pyrrhalta aenescens*, (J) Thailand specimen 11, (K) *Walterianella bucki*, (L) *Eucerotoma* sp. 344. Notation for the six helical elements is modified from Gillespie *et al.* (2003, 2004). Helices are boxed in A, and ambiguously aligned regions are boxed in G. The notation for RAAs and RECs is described in Fig. 2 and Table 3. The explanations of basepair symbols and reference for software used to construct structure diagrams are given in Fig. 3.

seven basepairs because only these positions represent a homologous structure across the alignment. 'Helix 3-1' has one REC and five RAAs (Fig. 5G).

A gallery of different chrysomelid 'helix 3-2' compound helices is shown in Fig. 6. Unlike the first two compound helices in the D2 expansion segment, which contain some length variation, the terminal helices of 'helix 3-2', 3o and 3p, are very conserved in length and base composition. In contrast, helix 3i is variable in length (14–50 nt) and sequence across all taxa (e.g. Fig. 6K). Length variation is also located in the unpaired nucleotides between strands 3h′ and 3g′, ranging from 4 to 24 nt. The chrysomelid sequence with the largest insertion, *Agelastica coerulea*, has the potential to form an eight basepair helix in this region (helix 3q in Fig. 6A). Other large insertions with different sequences in this region in scarab beetles and apocritan Hymenoptera can form a similar helix (J. Gillespie, unpubl. data). 'Helix 3-2' has five RAAs (Fig. 6F).

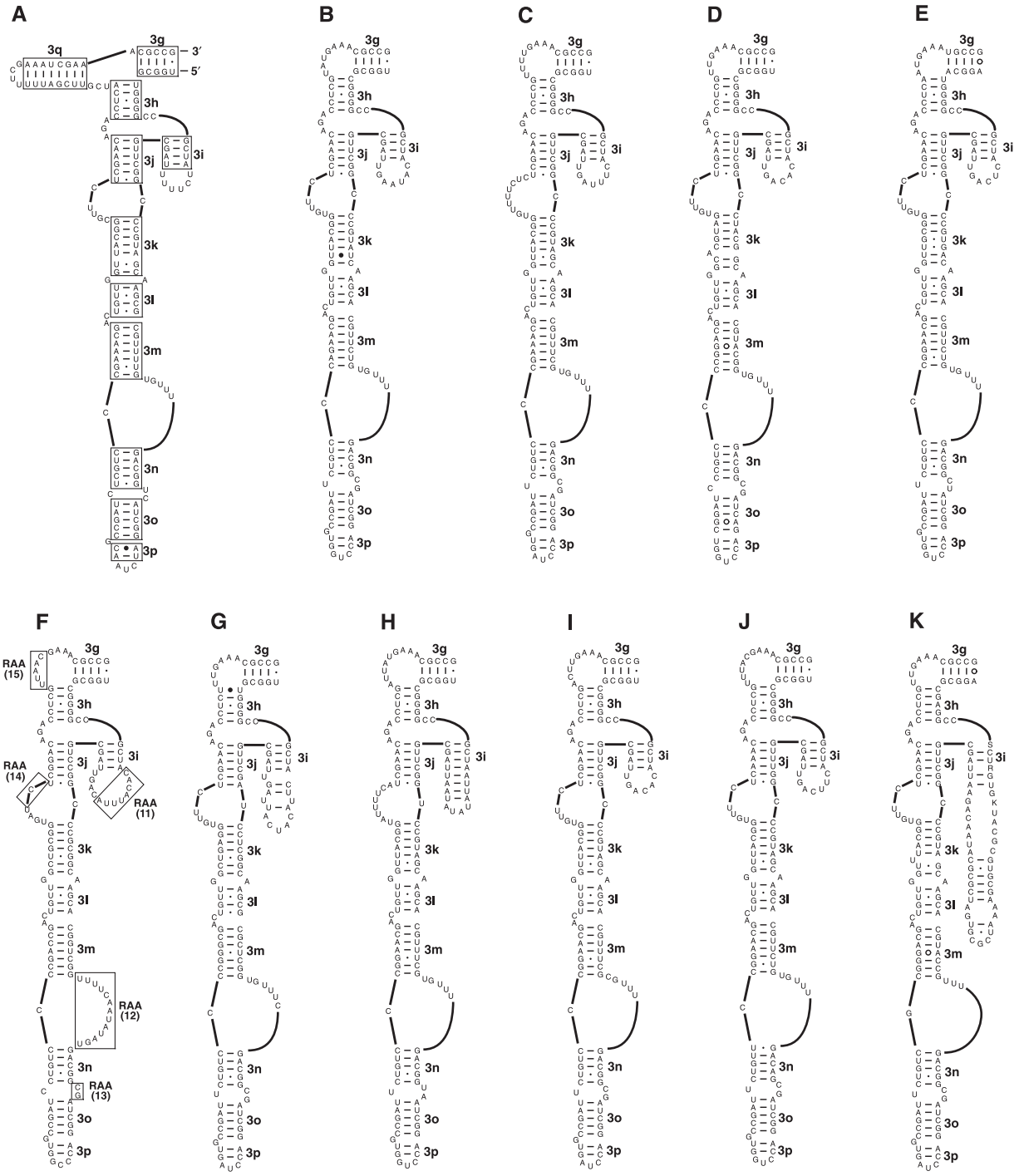


Figure 6. A gallery of diverse secondary structure diagrams from the 'helix 3-2' compound helix in the D2 region (synonymous with the 545 gallery of Schnare *et al.*, 1996) is shown for the following chrysomelid taxa: (A) *Agelastica coerulea*, (B) *Acalymma vittata*, (C) *Cerochroa brachialis*, (D) *Coptocycla adamantina*, (E) *Epitrix fasciata*, (F) *Lamprosoma* sp., (G) *Metaxyonycha panamensis*, (H) *Neolochmaea dilatipennis*, (I) *Pyrrhalta aenescens*, (J) Thailand specimen 11, (K) *Walterianella bucki*. Notation for the ten helical elements is modified from Gillespie *et al.* (2003, 2004), with the potential base pairing region within RAA (fifteen) in *A. coerulea* named helix 3q. Helices are boxed in (A) and ambiguously aligned regions are boxed in (F). The notation for RAAs is described in Fig. 2 and Table 3. The explanations of basepair symbols and reference for software used to construct structure diagrams are given in Fig. 3.

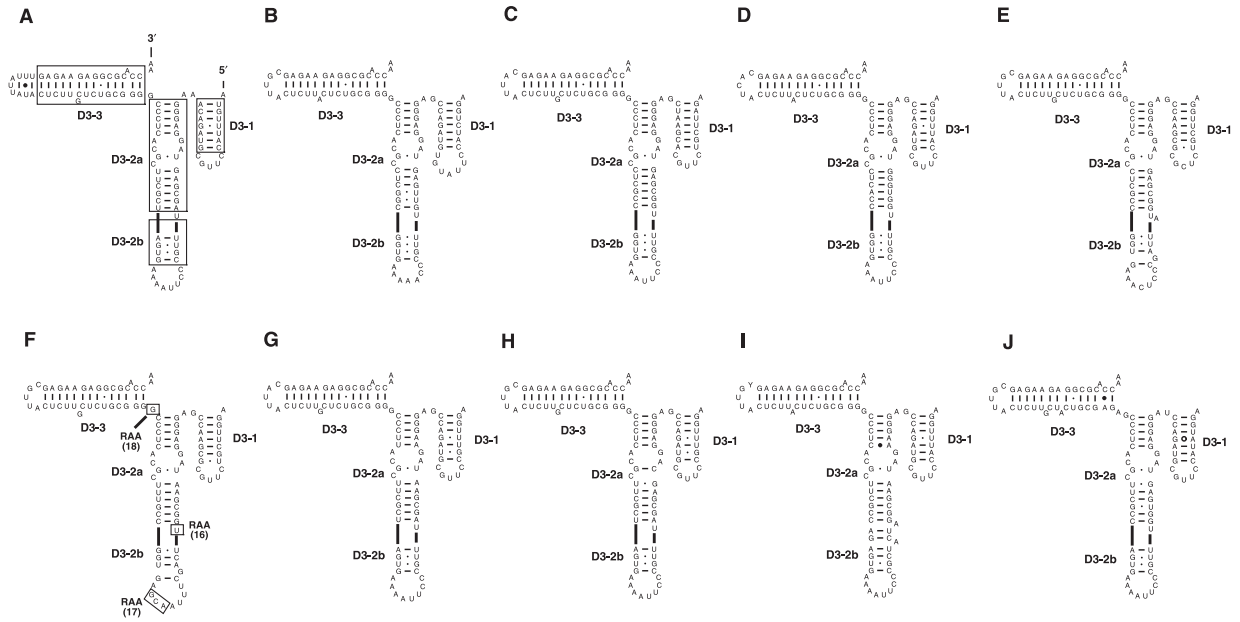


Figure 7. A gallery of diverse secondary structure diagrams for the D3 region (synonymous with the 650 gallery of Schnare *et al*., 1996) is shown for the following chrysomelid taxa: (A) *Cerochroa brachialis*, (B) *Scelidopsis* sp., (C) *Coptocycla adamantina*, (D) *Epitrix fasciata*, (E) *Lamprosoma* sp., (F) *Metaxyonycha panamensis*, (G) *Neolochmaea dilatipennis*, (H) *Pyrrhalta aenescens*, (I) Thailand specimen 11, (J) *Mimastra gracilicornis*. Notation for the three compound helices follows the convention of Kjer *et al*. (2001) with the exception of helix D3-2 being separated into D3-2a and D3-2b. Helices are boxed in (A), and ambiguously aligned regions are boxed in (F). The notation for RAAs is decribed in Fig. 2 and Table 3. The explanations of basepair symbols and reference for software used to construct structure diagrams are given in Fig. 3.

Expansion segment D3

The 28S-D3 region, corresponding to the 650 region of the nuclear LSU (Schnare *et al*., 1996), contains three compound helices in chrysomelids, labelled D3-1, D3-2 and D3-3, following the notation of Kjer *et al*. (2001). In Diptera (Kjer *et al*., 1994; Schnare *et al*., 1996; Hwang *et al*., 1998) and the machilid *Petrobius* sp. (Hwang *et al*., 1998), helix D3-1 is shortened or completely deleted, resulting in only two helices (D3-2 and D3-3) in the D3 expansion segment. The basepairs in helix D3-1 in the chrysomelids are supported by extensive positional covariation for a larger set of sequences that includes the chrysomelids, Trichoptera (Kjer *et al*., 2001), Odonata (K. M. Kjer, Rutgers University, New Brunswick, NJ, pers. comm.) and Hymenoptera (J. Gillespie, unpubl. data). This suggests that a helix that is present in the other holometabolous insect orders is deleted in Diptera. A gallery of structures representing the three motifs of the D3 in chrysomelids is shown in Fig. 7. At least one unpaired nucleotide is flanked by the two helices, D3-2a and D3-2b. Three RAAs occur in the D3 in chrysomelids (Fig. 7F).

Core elements

The D2 and D3 expansion regions are flanked by segments of the core rRNA structure. In contrast with the D2 and D3 regions, the core region usually has less insertions and deletions and more sequence conservation. The sequences between D2 and D3, including the 5′ and 3′ halves of helices H589, H604, H628, H700 and H563, and the 5′ half of helices H579, H671 and H687, were determined with the D2 and D3 sequences.

Helical conservation

Characteristic patterns of nucleotide substitutions and positional covariation in the expansion segments D2 and D3 reveal thirty conserved helices in the secondary structure model in the chrysomelids (Table 1). A total of 55.7% of the basepairs within the helical regions of the D2 and D3 chrysomelid expansion segments (not including the core regions sequenced) exhibit some degree of covariation (61.16% in D2, 37.84% in D3; calculated from Table 1). Within the chrysomelid dataset, the more variable positions within helices usually have more positional covariation at a larger percentage of the proposed basepairs, whereas the positions that are more conserved have a minimal amount of covariation among the two positions that are basepaired. Although many of the basepairs in the helices in the D2 and D3 secondary structure model have extensive amounts of positional covariation, some of the sequences underlying the helices, including 2, 2a, 3, 3a, 3h, 3l, 3o, 3p and D3-3, are conserved within the chrysomelids, and thus have minimal or no comparative support. However, sequence variation between the chrysomelids and other insect taxa D2 and D3 sequences contains positional covariations that substantiate the proposed basepairs in the structure model (http://www.rna.icmb.utexas.edu/). The frequency of

Table 1. Composition and degree of compensation for the base pairs of the D2 and D3 expansion segments and related core regions of the 28S rRNA in rootworms and related chrysomelid beetles. For base composition percentages, bold values represent any base pair present at 2% or greater in the alignment. Underlined values show which base pair types strictly covary for that base pair, with the summed underlined numbers providing a percentage of covariation (note: this approach does not account for intermediate GU pairs)

			Base pair composition (%)§																	
			Canonical						Non-canonical											
Helix*	Base pairs†	No.of sequences compared‡	GC	CG	UA	AU	GU	UG	AA	AC	AG	CA	CC	CU	GA	GG	UC	UU	Gap ¶ (–)	Covarying base pair** Y/N
D2																				
2	1	168	**10.1**	0	0	**78.0**	**11.9**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	2	167	**97.6**	0	0	0	1.2	0	0	1.2	0	0	0	0	0	0	0	0	0	Y
	3	173	**99.4**	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	178	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	5	178	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	6	178	0	0	0	**98.9**	0.6	0	0	0	0	0	0	0	0	0	0	0	0.6	N
2a	1	196	0	**99.0**	0	0	0	0	0	0	0	0	0.5	0.5	0	0	0	0	0	N
	2	194	**95.4**	0	0	0	**4.1**	0	0	0	0	0	0.5	0	0	0	0	0	0	N
	3	196	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	197	**99.0**	0	0	0	0	0	0	1.0	0	0	0	0	0	0	0	0	0	N
	5	195	0	0	**97.9**	0	0	0	0	0	0	0	0	0	**2.1**	0	0	0	0	N
	6	196	0	0	**95.4**	0	0	0	0	0	0	**4.6**	0	0	0	0	0	0	0	N
	7	194	0	0	0	0	0	**99.5**	0	0	0	0	0	0	0	0	0	0	0.5	N
2b	1	192	**97.9**	0	0	1.0	0	0	0	0.5	0	0	0.5	0	0	0	0	0	0	Y
	2	199	**2.0**	1.0	0.5	**57.8**	**36.7**	0	0	0.5	0	0	0	0	0	0	0	1.5	0	Y
	3	199	0	**66.8**	**8.0**	0	0	**21.1**	0	0	1.0	0.5	0.5	0	0	0	0	**2.0**	0	Y
2c	1	199	**13.6**	0	0	**4.0**	**79.4**	0.5	0	0	0	0	0	0.5	0	0	0.5	1.5	0	Y
	2	199	0	**3.0**	**88.9**	0.5	1.0	**5.0**	0.5	0	0	0.5	0	0	0	0	0	0.5	0	Y
	3	198	0	**87.9**	1.5	0.5	0	**9.1**	0	0	0.5	0	0	0	0	0	0	0	0.5	Y
	4	194	**94.8**	0	**2.1**	0.5	1.5	0	0	0.5	0	0	0	0	0	0.5	0	0	0	Y
	5	196	**10.7**	0	0	**82.1**	**5.6**	0	0	0	0	0	0.1	0	0.5	0	0	0	0	Y
2d	1	199	1.5	0	**65.8**	0.5	0	0.5	**5.0**	0	0	0	0.5	0.5	1.0	0	0	**24.6**	0	Y
	2	197	0	**4.1**	0.5	1.0	0	**77.7**	0	0	1.0	0	0	**3.0**	0	**6.1**	1.0	**5.6**	0	Y
	3	195	**72.8**	0	0.5	0	**3.6**	0	0	**17.9**	0.5	0	0	0	1.5	1.5	1.0	0	0.5	Y
2e	1	198	**9.6**	0	0	**63.1**	**26.3**	0.5	0	0	0	0	0	0	0	0	0	0.5	0	Y
	2	199	0.5	0	0	**76.4**	**22.1**	0	0	0	0	0	0	0	0	0.5	0	0.5	0	Y
	3	197	0	**58.9**	**19.8**	0.5	0	**20.8**	0	0	0	0	0	0	0	0	0	0	0	Y
	4	198	**43.9**	0	0.5	**3.5**	**50.0**	0	0	0	0	0	0	0	0.5	0	0.5	1.0	0	Y
	5	198	**3.0**	1.5	**81.8**	**5.1**	**2.5**	0.5	0	0	0	0	0	0	0	0	0	**5.6**	0	Y
2f	1	199	0	**99.5**	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	0	N
	2	196	**55.6**	0	0	1.0	**42.9**	0	0	0	0	0	0.5	0	0	0	0.5	0	0	Y
	3	198	**58.1**	0	0	**21.7**	**19.2**	0	0.5	0	0	0	0	0	0	0	0	0	0.5	Y
	4	200	0.5	0	**2.5**	**89.0**	**4.5**	0	0.5	0.5	0	0	0	0	0	0	0	1.0	1.5	Y
3	1	198	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	200	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	201	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	200	0	**98.5**	**1.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	5	201	**99.5**	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	6	197	0	**85.8**	**13.7**	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	Y
3a	1	203	0	**99.5**	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	N
	2	203	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	203	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	202	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	5	203	0	0.5	0	0	0	**99.5**	0	0	0	0	0	0	0	0	0	0	0	N
	6	203	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
3b	1	203	0.5	0	0	0.5	**99.0**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	2	203	**99.5**	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	3	203	0	**3.9**	**9.9**	0	0	**83.7**	0	0	0	0	0	0	0	0	0	**2.5**	0	Y
	4	203	**96.6**	0	0	0	**2.5**	0	0	1.0	0	0	0	0	0	0	0	0	0	Y
3c	1	203	0	0	**99.0**	0	0	1.0	0	0	0	0	0	0	0	0	0	0	0	N
	2	203	0	**94.6**	1.0	0	0	**3.4**	0	0	0	1.0	0	0	0	0	0	0	0	Y
	3	203	**10.3**	0	**89.7**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	4	203	**93.6**	0	0	1.0	**5.4**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	5	203	0	0	**90.6**	0	0	9.4	0	0	0	0	0	0	0	0	0	0	0	N
	6	201	0	**98.0**	0	0	0	2.0	0	0	0	0	0	0	0	0	0	0	0	N

Table 1. (*Continued*)

Helix*	Base pairs†	No.of sequences compared‡	Base pair composition (%)§																Gap ¶ (–)	Covarying base pair** Y/N
			Canonical						Non-canonical											
			GC	CG	UA	AU	GU	UG	AA	AC	AG	CA	CC	CU	GA	GG	UC	UU		
3d	1	203	**31.0**	0	0	1.5	**66.5**	0	0	0	0	0	0	0	0	1.0	0	0	0	Y
	2	203	0	0	0	**64.5**	**34.0**	0	1.5	0	0	0	0	0	0	0	0	0	0	N
	3	203	0	**79.8**	**11.8**	0.5	0.5	1.5	0	0	1.0	**3.0**	0	1.0	0	0	0	1.0	0	Y
3e	1	203	0	**3.9**	**73.9**	0	0	**16.3**	0	0	0	0	0	0	0.5	0	0	**5.4**	0	Y
	2	203	0.5	**75.9**	**3.9**	0	0	**17.7**	0	0	1.5	0	0	0	0	0.5	0	0	0	Y
	3	203	**56.7**	0	0.5	**3.0**	**38.9**	0.5	0	0.5	0	0	0	0	0	0	0	0	0	Y
	4	203	1.5	**7.4**	0	**72.4**	**16.3**	1.0	0	0	0	0	0	0	0	0	0	1.5	0	Y
	5	203	**86.2**	0.5	0	**10.8**	**2.5**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	6	203	**89.2**	0	1.0	0.5	0	0	**8.9**	0	0	0	0	0	0.5	0	0	0	0	Y
3f	1	201	0	**85.6**	**2.0**	**4.5**	0	0	0	0	0	**8.0**	0	0	0	0	0	0	0	Y
	2	202	0	**99.5**	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0	N
	3	203	**39.9**	0	0	**46.3**	**11.8**	0	0	1.5	0.5	0	0	0	0	0	0	0	0	Y
	4	203	0	**81.8**	1.0	0	0	**8.9**	0	0	0	**7.4**	0	0.5	0.5	0	0	0	0	Y
	5	203	**46.8**	0.5	0	**3.0**	**46.8**	0	0	0	0	0	0	0	0	0	0	**3.0**	0	Y
	6	202	0	**29.2**	**51.5**	0	0	**14.9**	1.5	0	**2.0**	0	0	0	0	0	0	1.0	0	Y
	7	201	**30.3**	0	0	**39.8**	**28.4**	0	0	0.5	0	0	0	0	0.5	0	0	0.5	0	Y
3g	1	202	0	1.5	**2.5**	0	0	**89.6**	1.0	0	**5.4**	0	0	0	0	0	0	0	0	Y
	2	203	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	201	**99.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	N
	4	202	0	**97.5**	**2.0**	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	Y
	5	203	**98.0**	0	0	1.0	0.5	0	0	0	0	0	0	0	0	0	0	0	0.5	Y
3h	1	202	0	**86.6**	**7.4**	1.0	0	0	0	0	0	0.5	0.5	0	0	0	0	**3.5**	0.5	Y
	2	203	**96.6**	0	0	1.5	0.5	0	0	1.5	0	0	0	0	0	0	0	0	0	Y
	3	203	1.5	0	0	**29.1**	**69.5**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	4	202	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	5	203	**99.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	N
3i	1	202	**99.5**	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	2	201	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	203	0	0	**99.5**	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	4	202	0	0	0	**99.5**	0.5	0	0	0	0	0	0	0	0	0	0	0	0	N
3j	1	202	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	203	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	203	0	1.5	**98.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	4	203	0	**97.5**	**2.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	5	203	**99.5**	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	N
	6	203	0	0	0	**4.9**	**95.1**	0	0	0	0	0	0	0	0	0	0	0	0	N
3k	1	203	0	**92.6**	**3.4**	0	0	0.5	0	0	1.0	1.0	0	0	0	0	0	0	1.5	Y
	2	203	0	**98.5**	1.0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	Y
	3	202	**95.0**	0	**3.0**	1.5	0.5	0	0	0	0	0	0	0	0	0	0	0	0	Y
	4	203	0	**9.4**	**67.0**	0	0	**23.2**	0	0	0	0	0.5	0	0	0	0	0	0	Y
	5	203	**6.9**	0.5	0	**87.2**	**4.4**	0	0	0.5	0.5	0	0	0	0	0	0	0	0	Y
	6	203	**11.3**	0	0	1.0	**82.3**	0	0	0	0	0	0	0	0	0	0	**5.4**	0	Y
	7	202	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
3l	1	202	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	203	0	0	0	0.5	**99.5**	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	203	0	**98.5**	0	0	0	0.5	0	0	1.0	0	0	0	0	0	0	0	0	N
	4	203	0	0	0	**74.9**	**25.1**	0	0	0	0	0	0	0	0	0	0	0	0	N
3m	1	203	0	**97.5**	**2.0**	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	Y
	2	203	**92.1**	0	0	0.5	**7.4**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	3	203	0.5	**3.0**	**90.6**	0	0	**5.4**	0	0	0	0	0	0	0	0	0	0.5	0	Y
	4	203	0	1.5	**90.1**	0	0	**5.9**	0	0	**2.5**	0	0	0	0	0	0	0	0	Y
	5	202	0	**75.7**	**7.4**	0	0	**15.8**	0	0	0	0	1.0	0	0	0	0	0	0	Y
	6	203	**10.3**	**24.1**	**35.5**	0	0	**30.0**	0	0	0	0	0	0	0	0	0	0	0	Y
	7	203	**99.0**	0	0	0	0	0.5	0	0	0	0	0	0	0.5	0	0	0	0	Y
3n	1	203	**93.1**	0	0	0.5	**5.9**	0	0	0	0	0	0.5	0	0	0	0	0	0	Y
	2	203	0.5	0	0	**99.5**	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	3	203	0	**89.7**	1.5	0	0	**8.9**	0	0	0	0	0	0	0	0	0	0	0	Y
	4	203	**4.4**	0	0	**10.3**	**85.2**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	5	203	**99.0**	0	0	0	1.0	0	0	0	0	0	0	0	0	0	0	0	0	N

Table 1. (*Continued*)

Helix*	Base pairs†	No.of sequences compared‡	Base pair composition (%)§																Gap ¶ (–)	Covarying base pair** Y/N
			Canonical						Non-canonical											
			GC	CG	UA	AU	GU	UG	AA	AC	AG	CA	CC	CU	GA	GG	UC	UU		
3o	1	203	0	0	0	**99.0**	0.5	0	0	0	0	0	0	0.5	0	0	0	0	0	N
	2	203	0	0	**93.6**	0	0	**6.4**	0	0	0	0	0	0	0	0	0	0	0	N
	3	203	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	202	**97.5**	0	0	0	0	0	0	0	1.0	0	0	0	0	0	0	0	1.5	Y
	5	202	**94.1**	0	0	0	**5.9**	0	0	0	0	0	0	0	0	0	0	0	0	N
3p	1	201	0	0	0	**97.0**	0	0	0	**2.5**	0	0	0	0.5	0	0	0	0	0	N
	2	202	0	**97.5**	**2.0**	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	Y
Core																				
H88	1	161	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	161	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	161	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	161	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
27	1	138	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	141	0	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	N
	3	141	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	141	**99.3**	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	0	0	N
	5	141	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	6	142	0	0	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	N
	7	142	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	8	142	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	9	142	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	10	142	1.4	0	0	0	**98.6**	0	0	0	0	0	0	0	0	0	0	0	0	N
	11	144	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	12	144	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	13	144	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
28	1	152	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	152	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	N
	3	152	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	152	0	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	N
	5	152	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	6	153	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	7	153	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	N
	8	153	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	9	153	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
29	1	152	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	152	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	N
D3																				
D3-1	1	151	**99.3**	0	0.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y
	2	151	**94.7**	0	0	**2.0**	**3.3**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	3	151	0	0	**99.3**	0	0	0.7	0	0	0	0	0	0	0	0	0	0	0	N
	4	151	0	**3.3**	**9.9**	0	0	**85.4**	0	0	1.3	0	0	0	0	0	0	0	0	Y
	5	152	0	**9.2**	**77.0**	**11.2**	0	**2.6**	0	0	0	0	0	0	0	0	0	0	0	Y
	6	152	**9.2**	0	0	**86.2**	**4.6**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	7	152	0	**94.7**	1.3	0	0	**3.9**	0	0	0	0	0	0	0	0	0	0	0	Y
D3-2a	1	148	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	149	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	149	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	149	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	5	149	**85.2**	0	0	1.3	0	0	0	**13.4**	0	0	0	0	0	0	0	0	0	Y
	6	149	0	**2.0**	0	0	0	**98.0**	0	0	0	0	0	0	0	0	0	0	0	N
	7	148	**65.5**	0	0	0	0	0	0	**34.5**	0	0	0	0	0	0	0	0	0	N
	8	149	1.3	0	0	**92.6**	**6.0**	0	0	0	0	0	0	0	0	0	0	0	0	Y
	9	148	**97.3**	0	0	0	**2.7**	0	0	0	0	0	0	0	0	0	0	0	0	N
	10	150	0	**92.7**	**3.3**	0	0	**4.0**	0	0	0	0	0	0	0	0	0	0	0	Y
	11	149	**97.3**	0	0	1.3	0.7	0.7	0	0	0	0	0	0	0	0	0	0	0	Y
	12	150	**75.3**	0	0	**22.0**	2.7	0	0	0	0	0	0	0	0	0	0	0	0	Y
D3-2b	1	149	0	0.7	**40.9**	0	0	**55.7**	0	0	0	0	0	0	0	0	0.7	**2.0**	0	N
	2	150	0	**14.0**	**16.0**	0	0	**67.3**	0	0	0	0	0	0	0	0	0	**2.7**	0	Y
	3	150	**2.0**	0	0	**4.7**	**90.7**	0	0	0	0	0	0	0	2.0	0	0	0	7	Y
	4	150	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N

Table 1. (*Continued*)

Helix*	Base pairs†	No.of sequences compared‡	Base pair composition (%)§ Canonical GC	CG	UA	AU	GU	UG	Non-canonical AA	AC	AG	CA	CC	CU	GA	GG	UC	UU	Gap ¶ (–)	Covarying base pair** Y/N
D3-3	1	144	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	144	**54.9**	0	0	**25.7**	**18.8**	0	0	0.7	0	0	0	0	0	0	0	0	0	Y
	3	144	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	144	0	**75.0**	0	0	0	**25.0**	0	0	0	0	0	0	0	0	0	0	0	N
	5	144	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	6	144	0	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	N
	7	144	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	8	144	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	9	144	0	**97.2**	0	0	0	0	0	0	0	2.8	0	0	0	0	0	0	0	N
	10	146	0	0	**99.3**	0	0	0	0.7	0	0	0	0	0	0	0	0	0	0	N
	11	146	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	12	146	0	**99.3**	0	0	0	0	0	0	0	0.7	0	0	0	0	0	0	0	N
	13	146	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	14	145	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
Core																				
34	1	128	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	2	128	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	3	129	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	4	128	0	**98.4**	0	0	0	1.6	0	0	0	0	0	0	0	0	0	0	0	N
	5	128	0	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	N
	6	129	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	7	128	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	8	129	0	0	0	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	9	126	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	10	129	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N
	11	128	**100**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	N

*Helix numbering refers the nucleotide positions shown in Fig. 2.
†Base pairs are numbered from 5′-end of 5′-strand of each helix.
‡Numbers vary at each position due to missing data (?), deletions (–) and possible presence of IUPAC-IUB ambiguity codes.
§The first nucleotide is that in the 5′-strand.
¶Gaps represent single insertion or deletion events, not indels.
**A covarying position is defined as having substitutions on both sides of the helix across the alignment.

Table 2. Mean percent nucleotides and mean transition/transversion ratios in pairing (stems) and nonpairing (loops) regions of the D2 and D3 expansion segments of the 28S LSU gene of chrysomelids*†‡

	Nucleotide composition (%) A	C	G	U	Substitutions (Ts/Tv)
Stems	0.15	0.24	0.39	0.22	3.66
Loops	0.25	0.25	0.26	0.24	2.30

*Calculated in MacClade 4.0 (Maddison & Maddison, 2000).
†Missing data and gaps not included in calculations.
‡Nucleotides within RAAs, RSCs and RECs were not included in calculations.

the four nucleotides in the unpaired regions of the chrysomelid D2 and D3 sequences is approximately 25% per base, whereas the paired regions have a bias for guanine (40%) and pyrimidines (46%) (Table 2). This unequal nucleotide frequency can be attributed to the ability of guanine to basepair with both cytosine and uracil (reviewed in Gutell *et al*., 1994). An analysis of the ratio of transitions to transversions (ts/tv) in paired and unpaired regions reveals a bias for more transitions in paired regions (Table 2). This is consistent with a mutational mechanism under selection for compensatory base changes repairing deleterious substitutions (Wheeler & Honeycutt, 1988; Rousset *et al*., 1991; Kraus *et al*., 1992; Marshall, 1992; Vawter & Brown, 1993; Gatesy *et al*., 1994; Nedbal *et al*., 1994; Douzery & Catzeflis, 1995; Springer *et al*., 1995; Springer & Douzery, 1996). Although it is expected that transversions should occur in greater frequency than transitions in regions without an expected ts/tv bias (Jukes & Cantor, 1969), we interpret a transition bias in nonpairing regions as a consequence of not including the majority of transversions that probably occur in the hypervariable regions wherein nucleotide homology could not be confidently assigned. In summary, our covariation analyses strongly support our predicted model (Fig. 3) for the expansion segments D2 and D3 from these sampled chrysomelid taxa.

Regions of ambiguous alignment (RAA)

Positional nucleotide homology could not be confidently assigned to twenty-one regions of our multiple sequence

Table 3. A list of the eighteen regions of alignment ambiguity (RAA), one region of slipped-strand compensation (RSC) and two regions of expansion and contraction (REC) created in the multiple sequence alignment of the expansion segments D2 and D3 of the 28S LSU rRNA from 229 sampled chrysomelids

Ambiguous region	Length* (nt)	Nonhomologous position†	General comments
RAA (1)	0–3	24–25	Forms a bulge between strands 2b and 2c
RAA (2)	0–2	40–41	Forms a bulge between strands 2e and RSC (1)
RSC (1)	7–8	40–41	Assignment of homology unclear due to *Acalymma* spp. *sensu stricto* (Gouldi group) forming a different structure, as well as other taxa having unique pairing potentials
RSC (1′)	5–6	49–50	Deletion in RSC (1′) causes a slip in the base-pairing in nine sampled species of *Acalymma s.s.* that results in a different structure
REC (1)	5–15	44–45	REC (1) and its complement REC (1′) form a hairpin-stem loop that is an extension of helix 2f; from 5 to 14 base-pairings occur across alignment with lateral and internal bulges present that make the region up to 15 positions in length
REC (1')	5–18	44–45	REC (1′) and its complement REC (1) form a hairpin-stem loop that is an extension of helix 2f; from 5 to 14 base-pairings occur across alignment with internal bulges present that make the region up to 18 positions in length
RAA (3)	3–5	44–45	Nonpairing terminal bulge formed by hairpin-stem loop REC (1); motif YYYR highly common when 4 nt present
RAA (4)	2–6	54–55	Forms a lateral bulge between strands 2e′ and 2d′
RAA (5)	0–4	57–58	Forms a lateral bulge between strands 2d′ and 2c′
RAA (6)	0–3	126–127	Along with RAA (8), forms an internal bulge between helices 3d and 3e
REC (2)	0–8	149–150	REC (2) and its complement REC (2′) form a hairpin-stem loop that is an extension of helix 3f; from 0 to 6 base-pairings occur across the alignment with lateral and internal bulges present that make the region up to 8 positions in length; some taxa have no extension of helix 3f
REC (2′)	0–8	149–150	REC (2′) and its complement REC (2) form a hairpin-stem loop that is an extension of helix 3f; from 0 to 6 base-pairings occur across the alignment with lateral and internal bulges present that make the region up to 8 positions in length; some taxa have no extension of helix 3f
RAA (7)	3–5	149–150	Nonpairing terminal bulge formed by hairpin-stem loop REC (2) or helix 3f
RAA (8)	0–4	170–171	Along with RAA (6), forms an internal bulge between helices 3e and 3d
RAA (9)	2–3	174–175	Along with positions 121–123, forms an an internal bulge between helices 3c and 3d
RAA (10)	2–3	180–181	Forms a lateral bulge between strands 3c′ and 3b′
RAA (11)	0–13	242–243	Part of the highly variable terminal loop formed by hairpin-stem 3i
RAA (12)	2–13	276–277	Forms a highly variable lateral bulge between strands 3m and 3n
RAA (13)	2–4	281–282	Along with position 305, forms an internal bulge between helices 3n and 3o
RAA (14)	1–7	336–337	(+4 nts 3′ to 3k′) along with position 254, forms an internal bulge between helices 3j and 3k
RAA (15)	0–20	352–353	Highly variable unpaired region joining the 3′ strand of helix 3h with conserved GAAA motif flanking the 3′ strand of helix 3g; forms helix 3q in *Agelastica coerulea*
RAA (16)	1–4	522–523	Forms a lateral bulge separating D3-2a and D3-2b
RAA (17)	0–3	528–529	Part of the variable terminal loop formed by helix D3-2b
RAA (18)	1–2	557–558	Junction between D3-2a and D3-3; AG motif in *Mimastra gracilicornis* causes ambiguous alignment of Gs and As; most likely 1 nt long

*Refers to the range of nucleotides within each ambiguous region.
†Nucleotide positions flanking ambiguous regions are given in Fig. 2.

alignment (Table 3). Eighteen of these unalignable regions are defined as RAA, in which single insertion and deletion events cannot be assessed as homologous characters across all of the sequences in the alignment, and consistent positional covariation (basepairing) is not found. Without secondary structure basepairing to guide the establishment of columnar homology in regions with many insertions and deletions (Kjer, 1995, 1997; Hickson *et al.*, 1996), we did not establish homology statements within RAAs. These nucleotides in the alignment were contained within brackets and were justified to the left (5′-strand) or right (3′-strand). Within the RAA regions, gaps do not represent insertion and deletion events as they do in the unambiguously aligned data. Instead they represent size variation within each RAA.

Regions of slipped-strand compensation (RSC)

The sequence alignment in one region in the D2 expansion segment cannot be aligned with high confidence owing to the inconsistent basepairing in its helix (Table 3). This helix is flanked on both sides by conserved basepairs in which postional homology assessment is unambiguous. Patterns of covariation were used to confirm inconsistent basepairing across the alignment within this RSC, as suggested by Gillespie (2004). As with RAAs, nucleotides in RSCs were bracketed and aligned to approximate homologous basepairs (when basepairs are proposed) or left or right justified, with gaps inserted to adjust for length heterogeneity as in the RAA regions (see above). Underlined positions represent structures that are not consistent across the alignment (Fig. 2).

Regions of expansion and contraction (REC)

The sequence alignment in two other helical regions in the D2 expansion segment also cannot be aligned with high confidence owing to the inconsistent basepairing in their helices (Table 3). Both of these regions have variation in the length of the terminal helix in compound helices

Table 4. Secondary structure characters of the D2, D3 expansion segments from the higher-level chrysomelid taxa sampled in this analysis. General comments describe the conservation of these characters, and whether or not they are found in unrelated taxa

Taxon	Region*	Character†	General comments
Dircema spp.	RAA (2)	GU	Internal bulge absent except for CC in Lamprosoma and single insertions in three flea beetles
Acalymma spp. *s.s.*	RSC (1)	C-UCUU	Deletion causes slippage in the hydrogen-bonding in this region that differs from the rest of the taxa in the alignment
	RSC (1′)	variable	Helix 2f expands and contracts across the alignment with positional homology uncertain; base composition in this helix, as well as sequence length, defines many genera and subtribes of the Luperini
Dircema spp.	RAA (3)	UUU	Triloop formed by extended 2f helix; UCG in *Aplosonyx quadripustulatus* and *Mimastra gracilicornis*; usually a tetraloop with a conserved UUYG motif
Galerucinae *s.s.*	RAA (5)	R	Single base-pair internal bulge is variable outside of the strict subfamily; U in *Medythia suturalis*
	REC (2)	variable	Helix 3f expands and contracts across the alignment with positional homology uncertain; base composition in this helix, as well as sequence length, defines many genera and subtribes of the Luperini
	RAA (3)	UUU	Triloop formed by extended 3f helix; base composition in this loop, as well as sequence length, defines many genera and subtribes of the Luperini, as well as generic groups in other chrysomelid subfamilies; loop is consistently larger in non-galerucine taxa
Oedionychina	pos. 213–239	large insert	These three flea beetles have an insertion within the terminal loop formed by helix 3i
	RAA (11)	variable	Terminal loop formed by helix 3i is informative at the generic level; however, certain motifs, such as CUU, are homoplastic
Agelastica coerulea	RAA (15)	8 bp helix	The ambiguous region between strands 3h′ and 3g′ forms a stable helix (helix 3q); may be a common insertion site as helices form here in other insects

*Regions within the D2 and D3 can be found in Figure 2.
†Illustration of structural characters can be found at http://hisl.tamu.edu/

'helix 2' and 'helix 3-1', and thus the precise placement of nucleotides and indels in the alignment is uncertain. Although consistent homology statements could not be made in these two ambiguous regions across all sequences in the alignment, secondary structure basepairing was used to differentiate between the helical component and the terminal bulge that comprised the enitre hairpin-stem loop structure (see Gillespie, 2004). After bracketing, nucleotides in RECs were treated the same as RSCs (see above).

Taxonomic implications

Structural characters that are unique and characteristic for the tribes, subtribes, sections and genera of the Luperini were identified (Table 4). These signatures in the D2 and D3 regions are consistent with previous taxonomic delineations within the Galerucinae *s.s.* (Leng, 1920; Laboisièrre, 1921; Weise, 1923; Wilcox, 1965; Seeno & Wilcox, 1982). The majority of taxon-specific structural characters in these molecules are located in the hairpin-stem loops of helices 2f and 3f. A more detailed depiction of these taxon-specific structural characters superimposed over our multiple sequence alignment is posted at http://hisl.tamu.edu. Individual secondary structure diagrams are also available (see below) that illustrate taxon-specific structural characters defined by our alignment. Calculated nucleotide frequencies for each higher-level taxon indicate that there are no significant differences between any of the sampled taxa regarding the distribution of the four bases throughout this region of the 28S (data not shown).

Utility for phylogeny reconstruction

The alignment of rDNA sequences becomes progressively more difficult as the sequence and length variation increases. The accuracy of the phylogenetic reconstruction is dependent in part on the accuracy of the alignment of the rDNA sequences. The expansion segments of the eukaryotic LSU rRNA are unique because they accumulate an extreme amount of nucleotide insertions (Veldman *et al.*, 1981; Michot *et al.*, 1984), and yet presumably have little impact on the function of the ribosome in translation (Musters *et al.*, 1989, 1991; Sweeney & Yao, 1989), with the exception of expansion segment D8, which is thought to interact with small nucleolar RNA E2 (Rimoldi *et al.*, 1993; Sweeney *et al.*, 1994). Extraordinary differences in sequence length (Gutell, 1992; De Rijk *et al.*, 1994) and secondary structure in expansion segments, even in recently diverged organisms, are not uncommon (Hillis & Dixon, 1991; Schnare *et al.*, 1996; J. Gillespie, unpubl. data). Thus, severe deviations from a common structure in eukaryotic expansion segments are expected (Schnare *et al.*, 1996), especially among taxa that have diverged over a large evolutionary time-scale.

Although seemingly problematic, the above characteristics of the expansion segments of the nuclear LSU rRNA make these markers ideal for phylogeny reconstruction. Conserved regions involved in hydrogen-bonding can be used to delimit regions in which primary assignment of homology is uncertain and indefensible (Kjer, 1997; Lutzoni *et al.*, 2000; Kjer *et al.*, 2001). The assignment of positional homology in length-heterogeneous datasets based on biological criteria has been shown to improve phylogeny estimation (Dixon & Hillis, 1993; Kjer, 1995; Titus & Frost, 1996; Morrison & Ellis, 1997; Uchida *et al.*, 1998; Mugridge *et al.*, 1999; Cunningham *et al.*, 2000; Gonzalez & Labarere, 2000; Hwang & Kim, 2000; Lydeard *et al.*, 2000; Morin, 2000; Xia, 2000; Xia *et al.*, 2003). Recoding RAAs and RECs as complex multistate characters with (Lutzoni *et al.*, 2000;

Xia *et al*., 2003; Gillespie *et al*., 2003a, 2004) or without (Kjer *et al*., 2001; Gillespie *et al*., 2003a, 2004) the implementation of an unequivocal weighting scheme can retain phylogenetic information in these unalignable regions. In addition, the descriptive coding of unalignable positions as morphological characters based on secondary structure can extract information from these regions of rRNA in phylogenetic analysis (Billoud *et al*., 2000; Collins *et al*., 2000; Lydeard *et al*., 2000; Ouvrard *et al*., 2000; J. Gillespie, unpubl. data).

Model applicability

Unpublished data from our laboratories suggest that the structural model presented here for the D2 and D3 expansion segments of the 28S rRNA gene from chrysomelids is applicable for several insect groups, including ichneumonoid, chalcidoid, proctotrupoid and cynipoid Hymenoptera, scaraeboid and curculionoid Coleoptera, and lower level studies on adephagous and other polyphagous beetles, including cassidine Chrysomelidae. All of these insect lineages contain the seven compound helices described in our model, with the majority of the length and structure variation occurring in the most distal regions of these compound helices (J. Gillespie, unpubl. data). Our model is consistent with the predicted structure of the *D. melanogaster* D2 region (Schnare *et al*., 1996). The only significant difference is a reduced 'helix 3-2' in the fruit fly (helix K in Schnare *et al*., 1996). Interestingly, predicted D2 structures for the plant *Arabidopsis thaliana*, the fungus *Cryptococcus neoformans* and the protist *Chlorella ellipsoidea* also share the general four-compound helix model presented here, but contain minor differences in the size of helix 3-1 and helix 3-2 and the length of the unpaired regions linking these motifs to the highly conserved helices 3a and 3 (synonymous with helix H2 of Michot & Bachellerie, 1987). These structural similarities between highly divergent taxa may suggest that similar regions of D2 have the propensity to expand and contract over time, possibly as a consequence of mild structural conservation that limits mutations to these specific locations. These findings are consistent with those of Wuyts *et al*. (2000) for the variable region 4 (V4) of the small subunit (SSU) rRNA across eukaryotes. Lower level studies of mitochondrial rRNA from Odonata (Misof & Fleck, 2003) and Phthiraptera (Page *et al*., 2002) also support this phenomenon of helix birth and death across divergent lineages.

Given the relative conservation within these variable regions of the 28S rRNA, the establishment of primary nucleotide homology across insects may be possible for some groups, particularly those within the Holometabola. However, with increased sequence divergence, it is likely that many regions of the D2 and D3 expansion segments will prove unalignable and noncomparable at the nucleotide level. For instance, published structural models for the expansion segment D3 from Diptera suggest severe deviations from the three compound helices defined by our model (Hancock *et al*., 1988; Tautz *et al*., 1988; Schnare *et al*., 1996; Hwang *et al*., 1998). This could possibly be the result of an accelerated rate of nucleotide substitution that presumably occurred in basal lineages of Diptera (Friedrich & Tautz, 1997). This is supported in part by our D3 model, and the D3 model for Amphiesmenoptera (Kjer *et al*., 2001) and Odonata (K. M. Kjer, pers. comm.), which are more consistent with chordate and nematode D3 structures (compiled in Schnare *et al*., 1996) than those of Diptera (Hancock *et al*., 1988; Tautz *et al*., 1988; Schnare *et al*., 1996; Hwang *et al*., 1998). This accelerated substitution rate, however, does not explain why D2 is so structurally different in lower Diptera (Nematocera) than in derived flies (Brachycera), as our D2 model is not congruent with any structural predictions for this region in *Aedes albopictus* (Kjer *et al*., 1994; Schnare *et al*., 1996). Interestingly, our model and these published dipteran models are quite different than preliminary structures of Strepsipteran D2 (J. Gillespie, unpubl. data) and D3 (Hwang *et al*., 1998) expansion segments.

Experimental procedures

Taxa examined

Table 5 lists the chrysomeloid species analysed in this investigation, with respective GenBank accession numbers for all sequences given. For the 28S-D2 we combined sixty-five new sequences with 137 from a previous study (Gillespie *et al*., 2004). The 153 sequences of the 28S-D3 segment were generated in this investigation. All 229 taxa are represented by the 28S-D2 region, with fifty taxa missing the 28S-D3 expansion segment. Voucher specimens for all sampled taxa can be found in the Texas A&M University, Rutgers University or the University of Delaware insect museums. Information regarding sampled taxa is available at http://hisl.tamu.edu.

Genome isolation, PCR and sequencing

For the sequences generated in this study, total genomic DNA was isolated using DNeasy™ Tissue Kits (Qiagen). PCR conditions followed those of Cognato & Vogler (2001), with primers designed for amplification of both the D2 and the D3 expansion segments found in Gillespie *et al*. (2003, 2004). Double-stranded DNA amplification products were sequenced directly with ABI PRISM™ (Perkin-Elmer) Big Dye Terminator Cycle Sequencing Kits and analysed on an Applied Biosystems (Perkin-Elmer) 377 automated DNA sequencer. Both antisense and sense strands were sequenced for all taxa, and edited manually with the aid of Sequence Navigator™ (Applied Biosystems). During editing of each strand, nucleotides that were readable, but showed either irregular spacing between peaks or had some significant competing background peak, were coded with lower case letters or IUPAC-IUB ambiguity codes. Consensus sequences were exported into Microsoft Word™ for manual alignment.

Multiple sequence alignment

The 28S-D2,D3 sequences were aligned manually according to secondary structure, with the notation following Kjer *et al*. (1994)

Table 5. The chrysomeloid taxa analysed in this investigation

Taxon* (Family/Subfamily/Tribe/Subtribe/Section)	Extract code†	Accession no.
Orsodacnidae		
Orsodacne atra (Ahrens)	JJG114	AY243660
[K]*Orsodacne atra* (Ahrens)	CND114	AY171422
Chrysomelidae		
Lamprosomatinae		
Lamprosoma sp. Kirby	JJG215	AY243651
Clytrinae		
Cyltrasoma palliatum	JJG286	AY646286
Criocerinae		
Lema sp. Fabricius	JJG308	AY243659
Cassidinae		
Coptocycla adamantina (Germar)	JJG214	AY243649
Microrhopala vittata Baly	JJG218	AY243650
Eumolpinae		
Syneta sp.	CND723	AY646287
[K]*Syneta adamsi* Baly	SJK723	AY171441
Megascelis sp. Latreille	JJG244	AY243652
Metaxyonycha panamensis Jacoby	JJG311	AY646288
Metaxyonycha sp. Chevrolat	JJG132	AY243653
Callisina quadripustulata Baly	JJG321	AY243654
Colaspis sp. Fabricius (or nr.)	JJG357	AY646289
Colaspis sp. Fabricius	JJG141	AY243655
Colasposoma sp. Laporte	JJG318	AY243656
Tymnes tricolor (Fabricius)	JJG258	AY243657
Chalcophana sp. Chevrolat	JJG352	AY243658
Chrysomelinae		
Chrysomelini		
Chrysomela knabi Brown	JJG237	AY243661
Chrysomela aeneicollis (Schaeffer)	JJG277	AY243662
Chrysomela populi Linnaeus	JJG236	AY243663
[K]*Chrysomela tremulae* Fabricius	SJK705	AY171423
[K]*Chrysolina coerulans* (Scriba)	SJK703	AY171429
Gastrophysa cyanea Melsheimer	JJG329	AY243664
[K]*Paropsis porosa* Erichson	SJK704	AY171438
[K]*Zygogramma piceicollis* (Stål)	CND334	AY171440
Timarchini		
Timarcha sp. Latreille	CND706	AY646290
[K]*Timarcha tenebricosa* (Fabricius)	SJK707	AY171439
Galerucinae *sensu lato*		
Alticini		
[K]*Altica* sp. Geoffroy	CND221	AY171424
[K]*Allochroma* sp. Clark	CND327	AY171428
[K]*Aphthona nigriscutis* Foudras	SJK700	AY171430
[K]*Chaetocnema* sp. (Stephens) (nr. *costulata*)	SJK720	AY171431
[K]*Disonycha conjuncta* (Germar)	CND061	AY171434
[K]*Blepharida rhois* (Forster)	CND209	AY171435
[K]*Dibolia borealis* Chevrolat	CND419	AY171442
[K]*Sangariola fortunei* (Baly)	SJK721	AY171443
Systena sp. Chevrolat (nr. *lustrans*)	JJG219	AY243665
[K]*Systena bifasciata* Jacoby	SJK219	AY171432
Scelidopsis sp. Jacoby	JJG225	AY243666
Cacoscelis sp. Chevrolat	JJG195	AY243667
Epitrix fasciata Blatchley	JJG328	AY243668
Physodactyla rubiginosa (Gerstaecker)	CND253	AY243671
Alagoasa libentina (Germar)	CND303	AY243670
Walterianella bucki Bechyné	CND039	AY243673
Blepharida ornata Baly	CND209	AY243672
Megistops vandepolli Duvivier	CND002	AY243669
Luperaltica sp. Crotch (or nr.)	JJG253	AY243695
[K]*Orthaltica copalina* (Fabricius)	SJK721	AY171437
Aedmon morrisoni Blake	CND207	AY646291
Galerucinae *sensu stricto*		
Oidini		
Oides decempunctata (Billberg)	JJG334	AY243674
[K]*Oides decempunctata* (Billberg)	SJK718	AY171448
Oides andrewsi Jacoby	JJG409	AY646292
Oides andrewsi Jacoby	JJG439	AY646293
Anoides sp. Weise (or nr.)	JJG380	AY646294
Galerucini		
Galerucini Chapuis 'genus undet.'	JJG387	AY646295
Galerucites		
Galeruca sp. Geoffroy	CND700	AY646297

Table 5. (*Continued*)

Taxon* (Family/Subfamily/Tribe/Subtribe/Section)	Extract code†	Accession no.
[K]*Galeruca rudis* LeConte	CND702	AY171436
Coelomerites		
Caraguata pallida (Jacoby) (or nr.)	JJG139	AY243776
Dircema cyanipenne Bechyné (or nr.)	JJG118	AY243771
Dircema sp. Clark	JJG343	AY243772
Dircema sp. Clark (or nr.)	JJG350	AY646298
Dircema sp. Clark	JJG355	AY646299
Dircema sp. Clark	JJG449	AY646300
Dircemella sp. Weise	JJG202	AY243773
Dircemella sp. Weise	JJG307	AY243774
Trirhabda bacharidis (Weber)	JJG075	AY243769
[K]*Monocesta* sp. Clark	CND710	AY171433
Cerochroa brachialis Stål	JJG405	AY646301
Atysites		
Diorhabda sp. Weise	CND712	AY243784
[K]*Diorhabda elongata* (Brullé)	SJK712	AY171446
Megaleruca sp. Laboisièrre	JJG204	AY243780
Megaleruca sp. Laboisièrre	JJG309	AY243779
Megaleruca sp. Laboisièrre	JJG320	AY646302
Pyrrhalta maculicollis (Motschulsky)	JJG190	AY243781
Pyrrhalta aenescens (Fairmaire)	JJG187	AY646303
Pyrrhalta sp. Joannis	JJG316	AY243782
Schematizites		
Metrogaleruca sp. Bechyné & Bechyné	JJG134	AY243777
Monoxia debilis LeConte	JJG239	AY243778
Neolachmaea dilatipennis (Jacoby)	JJG323	AY243785
Ophraea sp. Jacoby (or. nr.)	JJG131	AY243770
Ophraella notulata (Fabricius)	JJG095	AY243783
Schematiza flavofasciata (Klug)	JJG188	AY243786
[K]*Schematiza flavofasciata* (Klug)	ZSH003	AY171447
Apophyliites (apo)		
Pseudadimonia variolosa (Hope)	JJG312	AY243775
Apophylia pallipes (Baly)	JJG429	AY646304
Metacyclini		
New World genera		
Chthoneis sp. Baly	JJG109	AY243764
Chthoneis sp. Baly (nr. *marginicollis*)	JJG354	AY646305
Chthoneis sp. Baly (nr. *iquitoensis*)	JJG361	AY646306
Masurius violaceipennis (Jacoby) (or nr.)	JJG116	AY243766
Malachorhinus sericeus Jacoby	JJG129	AY243765
Exora obsoleta (Fabricius)	JJG110	AY243762
Exora obsoleta (Fabricius)	JJG353	AY243763
Exora sp. Chevrolat	JJG340	AY646307
Pyesia sp. Clark	JJG246	AY243767
Zepherina sp. Bechyné (or nr.)	JJG342	AY646308
Old World genus		
Palaeophylia sp. Jacoby (or nr.)	JJG222	AY243768
Hylaspini		
Antiphites		
Pseudeusttetha hirsuta	JJG443	AY646309
Emathea subcaerulea	JJG442	AY646310
Sermylites		
Aplosonyx orientalis (Jacoby)	JJG436	AY646311
Aplosonyx quadriplagiatus (Baly)	JJG173	AY243675
Aplosonyx sp. Chevrolat	JJG427	AY646312
Aplosonyx sp. Chevrolat	JJG412	AY646313
Sermylassa halensis (Linnaeus)	JJG179	AY243676
Hylaspites		
Agelasa nigriceps Motschulsky	JJG319	AY243677
Doryidella sp. Laboissière (or nr.)	JJG425	AY646314
Sphenoraia paviei Laboissière	JJG437	AY646315
Agelasticites		
Agelastica coerulea Baly	JJG315	AY243678
[K]*Agelastica coerulea* Baly	SJK701	AY171425
Luperini		
Luperini Chapuis 'genus undet.'	JJG376	AY646338
Aulacophorina		
Aulacophorites		
Paridea sp. Baly (or nr.)	JJG235	AY243696
Chosnia obesa (Jacoby) (or nr.)	JJG201	AY243697
Sonchia sternalis Fairmaire (or nr.)	JJG210	AY243698

Table 5. (*Continued*)

Taxon* (Family/Subfamily/Tribe/Subtribe/Section)	Extract code†	Accession no.
Aulacophora indica (Gmelin)	JJG220	AY243701
[K]*Aulacophora indica* (Gmelin)	SJK711	AY171444
Aulacophora lewisii Baly	JJG158	AY243700
Aulacophora lewisii Baly	JJG228	AY243699
Aulacophora lewisii Baly	JJG127	AY646316
Leptaulaca fissicollis Thomson (or nr.)	JJG234	AY243703
Diacantha fenestrata Chapuis (or nr.)	JJG232	AY243704
Idacanthites		
Prosmidia conifera Fairmaire (or nr.)	JJG212	AY243702
Diabroticina		
Diabroticites		
Diabroticites Chapuis 'genus undet.'	JJG345	AY646339
Isotes multipunctata (Jacoby)	JJG300	AY243723
Isotes sp. Weise	JJG145	AY243724
Isotes sp. Weise	JJG349	AY243722
Isotes sp. Weise	JJG351	AY243720
Isotes sp. Weise	JJG363	AY243721
Isotes sp. Weise	JJG372	AY243725
Isotes sp. Weise	JJG373	AY243726
Paranapiacaba tricincta (Say)	JJG322	AY243753
Paranapiacaba sp. Bechyné	JJG094	AY243752
Acalymma vittatum (Fabricius)	JJG413	AY646317
Acalymma fairmairei (Baly)	JJG016	AY243708
Acalymma bivittatum (Fabricius)	JJG297	AY243709
Acalymma blomorum Munroe & R. Smith (or nr.)	JJG229	AY243710
Acalymma trivittatum (Mannerheim)	JJG059	AY243711
Acalymma hirtum (Jacoby)	JJG053	AY243712
Acalymma albidovittatum (Baly)	JJG305	AY243713
Acalymma sp. Barber	JJG359	AY243714
Acalymma sp. Barber	JJG360	AY243715
Acalymma sp. Barber	JJG399	AY646318
Paratriarius subimpressa (Jacoby)	JJG128	AY243727
Paratriarius sp. Schaeffer	JJG147	AY243728
Paratriarius sp. Schaeffer	JJG348	AY243729
Paratriarius sp. Schaeffer	JJG374	AY243730
Amphelasma nigrolineatum (Jacoby)	JJG227	AY243754
Amphelasma sexlineatum (Jacoby)	JJG295	AY243755
Diabrotica balteata LeConte	JJG288	AY243731
Diabrotica biannularis Harold	JJG010	AY243732
Diabrotica decempunctata (Latreille)	JJG299	AY243733
Diabrotica speciosa (Germar)	JJG306	AY646319
Diabrotica speciosa speciosa (Germar)	JJG125	AY271865
Diabrotica virgifera virgifera LeConte	JJG060	AY243734
Diabrotica adelpha Harold	JJG046	AY243735
Diabrotica porracea Harold	JJG292	AY243737
Diabrotica undecimpunctata howardi Barber	JJG370	AY243739
Diabrotica undecimpunctata howardi Barber	JJG223	AY243738
[K]*Diabrotica undecimpunctata howardi* Barber	SJK223	AY171445
Diabrotica tibialis Jacoby	JJG170	AY243746
Diabrotica limitata (Sahlberg)	JJG313	AY243747
Diabrotica l. quindecimpunctata (Germar)	JJG180	AY243736
Diabrotica viridula (Fabricius)	JJG314	AY243748
Diabrotica sp. Chevrolat	JJG335	AY243740
Diabrotica sp. Chevrolat	JJG336	AY243741
Diabrotica sp. Chevrolat	JJG341	AY243742
Diabrotica sp. Chevrolat	JJG356	AY243743
Diabrotica sp. Chevrolat	JJG362	AY243744
Diabrotica sp. Chevrolat	JJG365	AY243745
Gynandrobrotica nigrofasciata (Jacoby)	JJG152	AY243717
Gynandrobrotica lepida (Say)	JJG298	AY243718
Gynandrobrotica sp. Bechyné	JJG358	AY243716
Gynandrobrotica sp. Bechyné	JJG371	AY243719
Gynandrobrotica ventricosa (Jacoby)	JJG135	AY646321
Cerotomites		
Neobrotica caeruleofasciata Jacoby	JJG117	AY243749
Neobrotica sp. Jacoby	JJG337	AY243750
Neobrotica sp. Jacoby	JJG375	AY243751
Eucerotoma sp. Laboissière	JJG344	AY243756
Eucerotoma sp. Laboissière	JJG346	AY243759
Eucerotoma sp. Laboissière	JJG347	AY243757

Table 5. (*Continued*)

Taxon* (Family/Subfamily/Tribe/Subtribe/Section)	Extract code†	Accession no.
Eucerotoma sp. Laboissière	JJG364	AY243758
Cerotoma arcuata (Olivier)	JJG048	AY243760
Cerotoma sp. Chevrolat	JJG339	AY243761
Cerotoma ruficornis (Olivier)	JJG172	AY646322
Cerotoma facialis Erichson	JJG161	AY646323
Phyllecthrites		
Trichobrotica nymphaea Jacoby	JJG226	AY243706
Phyllecthris gentilis LeConte	JJG366	AY243707
Phyllecthrites Dejean 'genus undet.'	JJG377	AY646324
Trachyscelidites		
Trachyscelida sp. Horn	JJG224	AY243705
Luperina		
Adoxiites		
Medythia suturalis (Motschulsky)	JJG434	AY646325
Medythia suturalis (Motschulsky)	JJG448	AY646326
Scelidites		
Scelolyperus lecontii (Crotch)	JJG099	AY243684
Scelolyperus meracus (Say)	JJG257	AY243686
Scelolyperus sp. Crotch	JJG054	AY243685
Lygistus streptophallus Wilcox	JJG367	AY243687
Keitheatus blakeae (White)	JJG414	AY646327
Stenoluperus nipponensis Laboissière	CND717	AY243694
Phyllobroticites		
Phyllobrotica sp. Chevrolat	JJG076	AY243690
[K]*Phyllobrotica* sp. Chevrolat	SJK076	AY171427
Mimastra gracilicornis Jacoby	JJG287	AY243691
Mimastra sp. Baly	JJG430	AY646328
Hoplasoma unicolor Illiger	JJG419	AY646329
Ornithognathites		
Hallirhotius sp. Jacoby	JJG206	AY243689
Exosomites		
Pteleon brevicornis (Jacoby)	JJG415	AY646330
Liroetiella bicolor Kimoto	JJG368	AY646331
Cassena indica (Jacoby)	JJG416	AY646332
Monoleptites		
Monoleptites Chapuis 'genus undet.'	JJG422	AY646333
Monoleptites Chapuis 'genus undet.'	JJG431	AY646334
Monoleptites Chapuis 'genus undet.'	JJG440	AY646335
Monoleptites Chapuis 'genus undet.'	JJG338	AY646296
Monolepta nigrotibialis Jacoby	JJG044	AY243681
[K]*Monolepta nigrotibialis* Jacoby	SJK044	AY171426
Monolepta sp. Chevrolat	JJG183	AY243682
Monolepta sp. Chevrolat	JJG310	AY243679
Monolepta sp. Chevrolat	JJG317	AY243680
Monolepta sp. Chevrolat	JJG369	AY243683
Metrioidea sp. Fairmaire (or nr.)	JJG301	AY243688
Luperites		
Spilocephalus bipunctatus Allard	JJG205	AY243692
Palpoxena sp. Baly	JJG230	AY243693
Luperus longicornis Fabricius	JJG407	AY646336
Megalognathites		
Megalognatha sp. Baly	JJG303	AY646337
Unidentified specimens		
Thailand specimen 4	JJG411	AY646340
Thailand specimen 7	JJG417	AY646341
Thailand specimen 8	JJG418	AY646342
Thailand specimen 10	JJG420	AY646343
Thailand specimen 11	JJG421	AY646344
Thailand specimen 13	JJG423	AY646345
Thailand specimen 14	JJG424	AY646346
Thailand specimen 22	JJG432	AY646347
Thailand specimen 25	JJG435	AY646348
Thailand specimen 31	JJG441	AY646349
Thailand specimen 36	JJG446	AY646350
Thailand specimen 37	JJG447	AY646351

*Taxonomic groupings follow Seeno & Wilcox (1982).
†DNA extraction codes for all taxa are listed as recorded on all vouchered specimens.
[K]Sequence from Kim *et al.* (2003).

and Kjer (1995), with slight modifications (Fig. 2). Alignment initially followed the secondary structural models of Gutell *et al.* (1994), which were obtained from http://www.rna.icmb.utexas.edu (Cannone *et al.*, 2002), and was further modified according to an existing chrysomelid D2 model (Gillespie *et al.*, 2003, 2004) and a trichopteran D3 model (Kjer *et al.*, 2001). Individual sequences, especially hairpin-stem loops, were evaluated in the program *mfold* (version 3.1; http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi), which folds rRNA based on free energy minimizations (Matthews *et al.*, 1999; Zuker *et al.*, 1999). These free-energy-based predictions were used to facilitate the search for potential basepairing stems, which were confirmed only by the presence of compensatory base changes across all taxa.

Regions in which positional homology assessments were ambiguous across all taxa were defined according to structural criteria, as in Kjer (1997), and described as regions of alignment ambiguity (RAA) or regions of slipped-strand compensation (RSC; Levinson & Gutman, 1987; for reviews regarding rRNA sequence alignment see Schultes *et al.*, 1999; Hancock & Vogler, 2000). Briefly, ambiguous regions in which basepairing was not identifiable were characterized as RAAs. For ambiguous regions in which basepairing was observed (RSCs), compensatory base change evidence was used to confirm structures that were not consistent across the alignment owing to the high occurrence of unknown insertion and deletion events (indels). For two ambiguous regions in the alignment caused by the expanding and contracting of hairpin-stem loops, RSCs were further characterized as RECs (regions of expansion and contraction) based on structural evidence used to identify separate nonpairing ambiguous regions of the alignment (terminal bulges). A recent paper addresses the characterization of RAAs, RSCs and RECs with a discussion on phylogenetic methods accommodating these regions (Gillespie, 2004).

Our alignment was entered into the alignment editor AE2 (developed by T. Macke; see Larsen *et al.*, 1993) for comparison with established eukaryotic secondary structural models (Gutell & Fox, 1988; Gutell *et al.*, 1990, 1992a,b, 1993; Schnare *et al.*, 1996; Cannone *et al.*, 2002). This process searched for compensating base changes using computer programs developed within the Gutell laboratory (University of Texas at Austin, http://www.rna.icmb.utexas.edu/ discussed in Gutell *et al.*, 1985; 1992a,b) and used subsequent information to infer additional secondary structural features. This refined alignment was reanalysed for positional covariations and the entire process was repeated until the proposed structures were entirely compatible with the alignment. Secondary structure diagrams were generated interactively with the computer program XRNA (developed by B. Weiser and H. Noller, University of Santa Cruz). Individual secondary structure diagrams are available at http://www.rna.icmb.utexas.edu/ and http://hisl.tamu.edu. Our complete multiple sequence alignment is posted at http://hisl.tamu.edu, with specific explanations regarding the rRNA structural alignment. The reader is encouraged to check J.J.G.'s homepage (http://hisl.tamu.edu) for continuing updates to the alignment and availability of secondary structure diagrams.

Comparative sequence analysis

The nucleotide frequency data and covarying positions were obtained with the Sun Microsystems Solaris-based program query (Gutell lab, unpublished software). Positional covariation was identified by several methods, including mutual information (Gutell *et al.*, 1992a,b), a pseudo-phylogenetic event scoring algorithm (Gautheret *et al.*, 1995) and an empirical method (Cannone *et al.*, 2002). This output was filtered to include only mutual best scores, i.e. pairs of positions that share a high covariation score, and examined for nested patterns that could represent helical regions (Goertzen *et al.*, 2003). These patterns included Watson–Crick (G:C and A:U), wobble (G:U) and other (e.g. C:A) base pairings that are adjacent and antiparallel to one another in helical regions. Nucleotide frequency tables for all positions (excluding RAAs, RSCs and RECs) within the putative 'stem-loop' regions were prepared to assess the quality and consistency of the predicted base pairing. In general, we accepted only those base pairs that exhibit near-perfect positional covariation in the dataset or invariant nucleotides with the potential to form Watson–Crick pairings within the same helix (Goertzen *et al.*, 2003).

Our alignment was also modified as a NEXUS file to estimate transition/transversion (ts/tv) ratios. In PAUP* (Swofford, 1999), a heuristic parsimony search implementing 100 random sequence additions, saving 100 trees per replicate (all other settings were left as default), generated 500 equally parsimonious trees. These trees were then used to calculate the mean ts/tv ratios in pairing and nonpairing regions across the entire alignment using the 'state changes and statistics' option in the chart menu of MacClade 4.0 (Maddison & Maddison, 2000).

Acknowledgements

We thank Matt Yoder for helpful comments on this manuscript. We are grateful to Markus Friedrich and Karl Kjer for providing unpublished structural alignments for Holometabola and Odonata, respectively. We are indebted to Shawn Clark, Catherine Duckett, Elizabeth Grobbelaar, Luciano Moura, Chris Reid and Ed Riley for collecting and identifying chrysomelids used in this study, and to Charles Bartlett, Dan Duran, Astrid Eben, Brian Farrell, David Furth, Annika Gillespie, Al Gillogly, April Harlin, Ting Hsiao, Sung Jin Kim, Karl Kjer, Doug LeDoux, Jim Plyler, Andrew Short, Rob Snyder, Doug Tallamy, Brian Urbain and Don Windsor for contributing specimens. J.J.G. specifically thanks Karl Kjer for training in structural alignment. Travel grants from the ALAS Project (Costa Rica) and Discover Life in America supported the collection of a minority of the taxa sampled in this study. This project was supported by grants to R.R.G. (the National Institutes of Health (GM48207 and GM067317, startup funds from The Institute for Cellular and Molecular Biology at the University of Texas at Austin, and the Dean's Boyer Fellow grant), and start-up funds from the Department of Entomology at Texas A&M University to A.I.C.

References

Amako, D., Kwon, O.-Y. and Ishikawa, H. (1996) Nucleotide sequence and presumed secondary structure of the 28S rRNA of pea aphid: implication for diversification of insect rRNA. *J Mol Evol* **43**: 469–475.

Arnheim, N. (1983) Concerted evolution of multigene families. In *Evolution of Genes and Proteins* (Nei, M. and Koehn, R.K., eds), pp. 38–61. Sinauer, Sunderland, MA.

Arnheim, N., Krystal, M., Shmickel, R., Wilson, G., Ryder, O. and Zimmer, E. (1980) Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proc Natl Acad Sci USA* **77**: 7323–7327.

Belshaw, R. and Quicke, D.L.J. (2002) Robustness of ancestral state estimates: evolution of life history strategy in ichneumonoid parasitoids. *Syst Biol* **51**: 450–477.

Billoud, B., Geurrucci, M.-A., Masselot, M. and Misof, B. (2000) Cirripede phylogeny using a novel approach: molecular morphometrics. *Mol Biol Evol* **17**: 1435–1445.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Müller, K.M., Pande, N., Shang, Z., Yu, N. and Gutell, R.R. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BioMed Central Bioinformatics* **3**: 2. [Correction: *BioMed Central Bioinformatics* **3**: 15.]

Clark, C.G. (1987) On the evolution of ribosomal RNA. *J Mol Evol* **25**: 343–350.

Clark, C.G., Tague, B.W., Ware, V.C. and Gerbi, S.A. (1984) *Xenopus laevis* 28S ribosomal RNA: a secondary structural model and its evolutionary and functional implications. *Nucleic Acids Res* **12**: 6197–6220.

Cognato, A.I. and Vogler, A.P. (2001) Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Scolytinae). *Syst Biol* **50**: 758–780.

Collins, L.J., Moulton, V. and Penny, D. (2000) Use of RNA secondary structure for studying the evolution of RNase P and RNase MRP. *J Mol Evol* **51**: 194–204.

Crease, T.J. and Taylor, D.J. (1998) The origin and evolution of variable-region helices in V4 and V7 of the small-subunit ribosomal RNA of branchiopod crustaceans. *Mol Biol Evol* **15**: 1430–1446.

Cunningham, C.O., Aliesky, H. and Collins, C.M. (2000) Sequence and secondary structure variation in the *Gyrodactylus* (Platyhelminthes: Monogenea) ribosomal RNA gene array. *J Parasitol* **86**: 567–576.

De Rijk, P., Van de Peer, Y., Chapelle, S. and De Wachter, R. (1994) Database on the structure of large ribosomal subunit RNA. *Nucleic Acids Res* **22**: 3495–3501.

De Rijk, P., Van de Peer, Y., Van den Broeck, I. and De Wachter, R. (1995) Evolution according to large ribosomal subunit RNA. *J Mol Evol* **41**: 366–375.

Dixon, M.T. and Hillis, D.M. (1993) Ribosomal secondary structure: Compensatory mutations and implications for phylogenetic analysis. *Mol Biol Evol* **10**: 256–267.

Douzery, E. and Catzeflis, F.M. (1995) Molecular evolution of the mitochondrial 12S rRNA in Ungulata (Mammalia). *J Mol Evol* **41**: 622–636.

Dover, G. (1982) Molecular drive: a cohesive mode of species evolution. *Nature* **299**: 111–117.

Flavell, R.B. (1986) Structure and control of expression of ribosomal RNA genes. *Oxf Surv Plant Mol Cell Biol* **3**: 252–274.

Fontana, W., Konings, D.A.M., Stadler, P.F. and Schuster, P. (1993) Statistics of RNA secondary structures. *Biopolymers* **33**: 1389–1404.

Friedrich, M. and Tautz, D. (1997) An episodic change of rDNA nucleotide substitution rate has occurred during the emergence of the insect order Diptera. *Mol Biol Evol* **14**: 644–653.

Gatesy, J., Hayashi, C., DeSalle, R. and Vrba, E. (1994) Rate limits for pairing and compensatory change: the mitochondrial ribosomal DNA of antelopes. *Evolution* **48**: 188–196.

Gautheret, D., Damberger, S.H. and Gutell, R.R. (1995) Identification of base-triples in RNA using comparative sequence analysis. *J Mol Biol* **248**: 27–43.

Gerbi, S.A. (1985) Evolution of ribosomal DNA. In *Molecular Evolutionary Genetics* (MacIntyre, R.J., ed.), pp. 419–517. Plenum, New York.

Gillespie, J.J. (2001) Inferring phylogenetic relationships among basal taxa of the leaf beetle tribe Luperini (Chrysomelidae: Galerucinae) through the analysis of mitochondrial and nuclear DNA sequences. Unpublished Masters Thesis, University of Delaware.

Gillespie, J.J. (2005) Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal RNA molecules. *Mol Phylogenet Evol* in press.

Gillespie, J.J., Duckett, C.N. and Kjer, K.M. (2001) Identification of a gene region that gives good phylogenetic signal for determining high level divergences within alticine and galerucine chrysomelids. *Chrysomela* **40/41**: 10–11. Available at. http://www.coleopsoc.org/chrys/chrysomela_4041r.pdf.

Gillespie, J.J., Kjer, K.M., Duckett, C.N. and Tallamy, D.W. (2003) Convergent evolution of cucurbitacin feeding in spatially isolated rootworm taxa (Coleoptera: Chrysomelidae; Galerucinae, Luperini). *Mol Phylogenet Evol* **29**: 161–175.

Gillespie, J.J., Kjer, K.M., Riley, E.R. and Tallamy, D.W. (2004) The evolution of cucurbitacin pharmacophagy in rootworms: insight from Luperini paraphyly. In *New Contribuitons to the Biology of Chrysomelidae* (Jolivet, P.H., Santiago-Blay, J.A. and Schmitt, M., eds), pp. 37–58. Kluwer Academic Publishers, Boston.

Goertzen, L.R., Cannone, J.J., Gutell, R.R. and Jansen, R.K. (2003) ITS secondary structure derived from comparative analysis: implications for sequence alignment and phylogeny of the Asteraceae. *Mol Phylogenet Evol* **29**: 216–234.

Gonzalez, P. and Labarere, J. (2000) Phylogenetic relationships of *Pleurotus* species according to the sequence and secondary structure of the mitochondrial small-subunit rRNA V4, V6 and V9 domains. *Microbiology* **146**: 209–221.

Gorski, J.L., Gonzalez, L.L. and Schmickel, R.D. (1987) The secondary structure of human 28S rRNA: the structure and evolution of a mosaic rRNA gene. *J Mol Evol* **24**: 236–251.

Gutell, R.R. (1992) Evolutionary characteristics of 16S and 23S rRNA structures. In *The Origin and Evolution of Prokaryotic and Eukaryotic Cells* (Hartman, H. and Matsuno, K., eds), pp. 243–309. World Scientific Publishing Co, New Jersey.

Gutell, R.R. and Fox, G.E. (1988) A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res* **16S**: r175–r269.

Gutell, R.R., Gray, M.W. and Schnare, M.N. (1993) A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures. *Nucleic Acids Res* **20S**: 2095–2109.

Gutell, R.R., Larsen, N. and Woese, C.R. (1994) Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol Rev* **58**: 10–26.

Gutell, R.R., Power, A., Hertz, G.Z., Putz, E.J. and Stormo, G.D. (1992a) Identifying constraints on the higher-order structure of RNA: continued development and application of comparative sequence analysis methods. *Nucleic Acids Res* **20**: 5785–5795.

Gutell, R.R., Schnare, M.N. and Gray, M.W. (1990) A compilation of large subunit (23S-like) ribosomal RNA sequences presented in a secondary structure format. *Nucleic Acids Res* **18S**: 2319–2330.

Gutell, R.R., Schnare, M.N. and Gray, M.W. (1992b) A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures. *Nucleic Acids Res* **21S**: 3055–3074.

Gutell, R.R., Weiser, B., Woese, C.R. and Noller, H.F. (1985) Comparative anatomy of 16S-like ribosomal RNA. *Prog Nucleic Acid Res Mol Biol* **32**: 155–216.

Hancock, J.M. and Dover, G.A. (1988) Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. *Mol Biol Evol* **5**: 377–392.

Hancock, J.M., Tautz, D. and Dover, G.A. (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Mol Biol Evol* **5**: 393–414.

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

Hassouna, N., Michot, B. and Bachellerie, J.-P. (1984) The complete nucleotide sequence of mouse 28S rRNA gene: implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Res* **12**: 3563–3583.

Hickson, R.E., Simon, C., Cooper, A., Spicer, G.S., Sullivan, J. and Penny, D. (1996) Conserved sequence motifs, alignment, and secondary structure for the third domain of animal 12S rRNA. *Mol Biol Evol* **13**: 150–169.

Hillis, D.M. and Dixon, M.T. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Q Rev Biol* **66**: 411–453.

Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M. and Schuster, P. (1994) Fast folding and comparison of RNA secondary structures. *Monatsh Chem* **125**: 167–188.

Hwang, S.K. and Kim, J.G. (2000) Secondary structure and phylogenetic implications of nuclear large subunit ribosomal RNA in the ectomycorrhizal fungus *Tricholoma matsutake*. *Curr Microbiol* **40**: 250–256.

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

Jukes, T.H. and Cantor, C.R. (1969) Evolution of protein molecules. In *Mammalian Protein Metabolism* (Munro, N.H., ed.), pp. 21–132. Academic Press, New York.

Kim, S.J., Kjer, K.M. and Duckett, C.N. (2003) Comparison between molecular and morphological-based phylogenies of galerucine/alticine leaf beetles (Coleoptera, Chrysomelidae: Galerucinae). *Insect Syst Evol* **34**: 53–64.

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

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

Kjer, K.M., Baldridge, G.D. and Fallon, A.M. (1994) Mosquito large subunit ribosomal RNA: simultaneous alignment of primary and secondary structure. *Biochim Biophys Acta* **1217**: 147–155.

Kjer, K.M., Blahnik, R.J. and Holzenthal, R.W. (2001) Phylogeny of Trichoptera (Caddisflies): characterization of signal and noise within multiple datasets. *Syst Biol* **50**: 781–816.

Kraus, F., Jarecki, L., Miyamoto, M., Tanhauser, S. and Laipis, P. (1992) Mispairing and compensational changes during the evolution of mitochondrial ribosomal RNA. *Mol Biol Evol* **9**: 770–774.

Kuzoff, R.K., Swere, J.A., Soltis, D.E., Soltis, P.S. and Zimmer, E.A. (1998) The phylogenetic potential of entire 26S rDNA sequences in plants. *Mol Biol Evol* **15**: 251–263.

Laboisièrre, V. (1921) Etude des Galerucini de la collection du Musee Congo belge. *Rev Zool Africaine* **9**: 33–86.

Larsen, N., Olsen, G.J., Maidak, B.L., McCaughey, M.J., Overbeek, R., Macke, T.J., Marsh, T.L. and Woese, C.R. (1993) The ribosomal database project. *Nucleic Acids Res* **21**: 3021–3023.

Leng, C.W. (1920) *Catalogue of the Coleoptera of America, north of Mexico*. Mount Vernon, New York.

Levinson, G. and Gutman, G.A. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol* **4**: 203–221.

Lutzoni, F., Wagner, P. and Reeb, V. (2000) Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses using unequivocal coding and optimal character-state weighting. *Syst Biol* **49**: 628–651.

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

Maddison, D.R. and Maddison, W.P. (2000) *Macclade 4: Analysis of Phylogeny and Character Evolution*, Version 4.0. Sinauer Associates, Sunderland, MA.

Marshall, C.R. (1992) Substitution biases, weighted parsimony, and amniote phylogeny as inferred from 18S-ribosomal-RNA sequences. *Mol Biol Evol* **9**: 370–377.

Matthews, D.H., Sabina, J., Zuker, M. and Turner, D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* **288**: 911–940.

Michot, B. and Bachellerie, J.-P. (1987) Comparisons of large subunit rRNAs reveal some eukaryote-specific elements of secondary structure. *Biochimie* **69**: 11–23.

Michot, B., Hassouna, N. and Bachellerie, J.-P. (1984) Secondary structure of mouse 28S rRNA and general model for the folding of the large rRNA in eukaryotes. *Nucleic Acids Res* **12**: 4259–4279.

Mindell, D.P. and Honeycutt, R.L. (1990) Ribosomal RNA in vertebrates: evolution and phylogenetic implications. *Annu Rev Ecol Syst* **21**: 541–566.

Misof, B. and Fleck, G. (2003) Comparative analysis of mt LSU rRNA secondary structures of odonates: structural variability and phylogenetic signal. *Insect Mol Biol* **12**: 535–547.

Morin, L. (2000) Long branch attraction effects and the status of 'basal eukaryotes': phylogeny and structural analysis of the ribosomal RNA gene cluster of the free-living diplomonad *Trepomonas agilis*. *J Eukaryot Microbiol* **47**: 167–177.

Morrison, D.A. and Ellis, J.T. (1997) Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of Apicomplexa. *Mol Biol Evol* **14**: 428–441.

Moulton, V., Zuker, M., Steel, M., Pointon, R. and Penny, D. (2000) Metrics on RNA secondary structures. *J Comput Biol* **7**: 277–292.

Mugridge, N.B., Morrison, D.A., Johnson, A.M., Luton, K., Dubey, J., Votypka, J. and Tenter, A.M. (1999) Phylogenetic relationships of the genus *Frenkelia*: a review of its history and new knowledge gained from comparison of large subunit ribosomal ribonucleic acid gene sequences. *Int J Parasitol* **29**: 957–972.

Musters, W., Goncalves, P.M., Boon, K., Gaué, H.A., van Heerikhuizen, H. and Planta, R.J. (1991) The conserved GTPase center and variable region V9 from *Saccharomyces cerevisiae* 26S rRNA can be replaced by their equivalents from other prokaryotes or eukaryotes without detectable loss of ribosomal function. *Proc Natl Acad Sci USA* **88**: 1469–1473.

Musters, W., Venema, J., van der Linden, G., van Heerikhuizen, H., Klootwijk, J. and Planta, R.J. (1989) A system for the analysis of yeast ribosomal DNA mutations. *Mol Cell Biol* **9**: 551–559.

Nedbal, M.A., Allard, M.W. and Honeycutt, R.L. (1994) Molecular systematics of hystricognath rodents: Evidence from the mitochondrial 12S rRNA gene. *Mol Phylogenet Evol* **3**: 206–220.

Notredame, C., O'Brien, E.A. and Higgins, D.G. (1997) RAGA: RNA sequence alignment by genetic algorithm. *Nucleic Acids Res* **25**: 4570–4580.

Ouvrard, D., Campbell, B.C., Bourgoin, T. and Chan, K.L. (2000) 18S rRNA secondary structure and phylogenetic position of Peloridiidae (Insecta, Hemiptera). *Mol Phylogenet Evol* **16**: 403–417.

Page, R.D.M., Cruickshank, R. and Johnson, K.P. (2002) Louse (Insecta: Phthiraptera) mitochondrial 12S rRNA secondary structure is highly variable. *Insect Mol Biol* **11**: 361–369.

Rimoldi, O.J., Raghu, B., Mag, M.K. and Eliceiri, G.L. (1993) Three new small nucleolar RNAs that are psoralen cross-linked *in vivo* to unique regions of pre-rRNA. *Mol Cell Biol* **13**: 4382–4390.

Rousset, F., Pelandakis, M. and Solignac, M. (1991) Evolution of compensatory substitutions through GU intermediate state in *Drosophila* rRNA. *Proc Natl Acad Sci USA* **88**: 10032–10036.

Schnare, M.N., Damberger, S.H., Gray, M.W. and Gutell, R.R. (1996) Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. *J Mol Biol* **256**: 701–719.

Schultes, E.A., Hraber, P.T. and LaBean, T.H. (1999) Estimating the contributions of selection and self-organization in RNA secondary structure. *J Mol Evol* **49**: 76–83.

Seeno, T.N. and Wilcox, J.A. (1982) Leaf beetle genera (Coleoptera: Chrysomelidae). *Entomography* **1**: 1–222.

Shapiro, B.A. and Zhang, K. (1990) Comparing multiple RNA secondary structures using tree comparisons. *CABIOS* **6**: 309–318.

Sorenson, M.D., Oneal, E., Garcia-Moreno, J. and Mindell, D.P. (2003) More taxa, more characters: the Hoatzin problem is still unresolved. *Mol Biol Evol* **20**: 1484–1499.

Springer, M.S. and Douzery, E. (1996) Secondary structure and patterns of evolution among mammalian mitochondrial 12S rRNA molecules. *J Mol Evol* **43**: 357–373.

Springer, M.S., Hollar, L.J. and Burk, A. (1995) Compensatory substitutions and the evolution of the mitochondrial 12S rRNA gene in mammals. *Mol Biol Evol* **12**: 1138–1150.

Sweeney, R., Chen, L. and Yao, M.-C. (1994) An rRNA variable region has an evolutionary conserved essential role despite sequence divergence. *Mol Cell Biol* **14**: 4203–4215.

Sweeney, R. and Yao, M.-C. (1989) Identifying functional regions of rRNA by insertion mutagenesis and complete gene replacement in *Tetrahymena thermophila*. *EMBO J* **8**: 933–938.

Swofford, D.L. (1999) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, MA.

Tautz, D.J., Hancock, J.M., Webb, D.A., Tautz, C. and Dover, G.A. (1988) Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Mol Biol Evol* **5**: 366–376.

Titus, T.A. and Frost, D.R. (1996) Molecular homology assessment and phylogeny in the lizard family Opluridae (Squamata: Iguania). *Mol Phylogenet Evol* **6**: 49–62.

Uchida, H., Kitae, K., Tomizawa, K.I. and Yokota, A. (1998) Comparison of the nucleotide sequence and secondary structure of the 5.8S ribosomal RNA gene of *Chlamydomonas tetragama* with those of green algae. *DNA Seq* **8**: 403–408.

Vawter, L. and Brown, W.M. (1993) Rates and patterns of base change in the small subunit ribosomal RNA gene. *Genetics* **134**: 597–608.

Veldman, G.M., Klootwijk, J., De Regt, V.C.F.H., Planta, R.J., Branlant, C., Krol, A. and Ebel, J.-P. (1981) The primary and secondary structure of yeast 26S rRNA. *Nucleic Acids Res* **9**: 6935–6952.

Ware, V.C., Tague, B.W., Clark, C.G., Gourse, R.L., Brand, R.C. and Gerbi, S.A. (1983) Sequence analysis of 28S ribosomal DNA from the amphibian *Xenopus laevis*. *Nucleic Acids Res* **11**: 7795–7817.

Weise, J. (1923) Chrysomeliden und Coccinelliden aus Queensland. *Ark Zool* **15**: 1–150.

Wheeler, W.C. and Honeycutt, R.L. (1988) Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Mol Biol Evol* **5**: 90–96.

Wilcox, J.A. (1965) A synopsis of North American Galerucinae (Coleoptera: Chrysomelidae). *Bull NY St Mus Surv* **400**: 1–226.

Wool, I.G. (1986) Studies of the structure of eukaryotic (mammalian) ribosomes. In *Structure, Function and Genetics of Ribosomes* (Hardesty, J. and Kramer, G., eds), pp. 391–411. Springer-Verlag, New York.

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

Xia, X. (2000) Phylogenetic relationship among horseshoe crab species: the effect of substitution models on phylogenetic analyses. *Syst Biol* **49**: 87–100.

Xia, X., Xie, Zheng and Kjer, K.M. (2003) 18S ribosomal RNA and tetrapod phylogeny. *Syst Biol* **52**: 283–295.

Zuker, M., Mathews, D.H. and Turner, D.H. (1999) Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In *RNA Biochemistry and Biotechnology* (Barciszewski, J. and Clark, B.F.C., eds), pp. 11–43. Academic Publishers, Boston, MA.