The Origin and Evolution of Variable-Region Helices in V4 and V7 of the Small-Subunit Ribosomal RNA of Branchiopod Crustaceans

Teresa J. Crease and Derek J. Taylor†*

*Department of Zoology, University of Guelph, Ontario, Canada; and †Department of Biological Sciences, State University of New York at Buffalo

We sequenced the V4 and V7 regions of the small-subunit ribosomal RNA (SSU rRNA) from 38 species of branchiopod crustaceans (e.g., *Artemia, Daphnia, Triops*) representing all eight extant orders. Ancestral large-bodied taxa in the orders Anostraca, Notostraca, Laevicaudata, and Spinicaudata (limnadiids and cyzicids) possess the typical secondary structure in these regions, whereas the spinicaudatan *Cyclestheria* and all of the cladocerans (Anomopoda, Ctenopoda, Onychopoda, and Haplopoda) possess three unique helices. Although the lengths and primary sequences of the distal ends of these helices are extremely variable, their locations, secondary structures, and primary sequences at the proximal end are conserved, indicating that they are homologous. This evidence supports the classical view that Cladocera is a monophyletic group and that the cyclestheriids are transitional between spinicaudatans and cladocerans. The single origin and persistence since the Permian of the unique cladoceran helices suggests that births and deaths of variable region helices have been rare. The broad range of sequence divergences observed among the cladoceran helices permitted us to make inferences about their evolution. Although their proximal ends are very GC-biased, there is a significant negative correlation between length and GC content due to an increasing proportion of U at their distal ends. Slippage-like processes occurring at unpaired nucleotides or bulges, which are very U-biased, are associated with both helix origin and runaway length expansion. The overall GC contents and lengths of V4 and V7 are highly correlated. More surprisingly, the lengths of these SSU rRNA variable regions are also highly correlated with the length of the large-subunit rRNA expansion segment, D2, indicating that mechanisms affecting length variation do so both across single genes and across genes in the rRNA gene family.

Introduction

Eukaryotic nuclear genes contain conserved functional units that are usually interspersed with rapidly evolving regions of unknown origin and function. In nuclear ribosomal RNA genes, a conserved core is interrupted by transcribed expansion segments whose evolutionary patterns appear to vary according to gene size. For example, the nuclear large-subunit ribosomal RNA (LSU rRNA) genes from animals (average length, 4,500 nt) tend to be hypervariable with respect to the occurrence of indels, have significant levels of repetition for short sequence motifs, possess novel helical structures, and show correlated changes in length and GC content among expansion segments (Hancock and Dover 1988; Nunn et al. 1996). In contrast, taxonomic groups with short LSU rRNA genes (e.g., plants with an average length of 3,400 nt) tend to have few indels, insignificant levels of repetition, and no unique secondary structures (Michot and Bachellerie 1987; Michot, Qu, and Bachellerie 1990; Kuzoff et al. 1998). This association between length and pattern of sequence divergence may indicate that slippage replication is an important mechanism in the evolution of these regions (Hancock and Dover 1988; Hancock, Tautz, and Dover 1988).

It is unclear if this size-related molecular evolution exists in the nuclear small-subunit ribosomal RNA (SSU rRNA), as size variation is generally modest; most genes are between 1,700 and 1,900 nt. However, unusually long genes (>2,000 nt) are known from a few distantly related taxa, including strepsipteran (Chalwatzis et al. 1995) and sternorrhynchan (Campbell, Steffen-Campbell, and Gill 1994; Kwon, Ogino, and Ishikawa 1991) insects, a cestode flatworm (Piçon et al. 1996), a branchiopod crustacean (Crease and Colbourne 1998), and an amoeba (Hinkle et al. 1994). The SSU rRNA structure of the aphid *Acyrthosiphon pisum* (Sternorrhynca) follows the LSU rRNA pattern, as it shows highly significant repetition (Hancock 1995) and contains unique helical structures (Kwon, Ogino, and Ishikawa 1991). Although the SSU rRNAs of the branchiopod *Daphnia pulex* (Crease and Colbourne 1998) and the amoeba *Phreatamoeba balamuthi* (Hinkle et al. 1994) possess unique helices, they lack significant repetition, suggesting that patterns of tandem repetition generated by past slippage events have been obscured by subsequent point mutations or that length expansion has occurred by some other mechanism.

Understanding the evolution of the unique helices associated with long rRNAs could have implications for deeper-level phylogenetics. If these structures arise and disappear infrequently on an evolutionary timescale, then they might make ideal complex molecular characters for the study of deep phylogenies. Indeed, the first studies of the evolution of rRNA expansion segment secondary structures did identify several kingdom-specific helices (Michot and Bachellerie 1987). Unfortunately, most studies of deeper-level phylogenies have avoided expansion sequences, because the primary sequence and length variation among taxa is usually very large. The result is that little is known about the birth and death of expansion helices or their phylogenetic utility below the kingdom level.

Key words: SSU-rRNA, variable regions, secondary structure, slippage replication, Cladocera, Branchiopoda.

Address for correspondence and reprints: Teresa J. Crease, Department of Zoology, University of Guelph, Guelph, Ontario, Canada N1G 2W1. E-mail: tcrease@uoguelph.ca.

Mol. Biol. Evol. 15(11):1430–1446. 1998
© 1998 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

To date, only a single study has analyzed patterns of expansion in variable regions of the SSU rRNA gene within a monophyletic group of organisms. Vogler, Selsh, and Hancock (1997) examined the sequence of the unusually long V4 region for 27 taxa of the beetle family Cicindelidae. Despite substantial differences in primary sequence and length in the V4 regions of these taxa, secondary-structure predictions indicated that it always folds into a single helical structure. The pattern of sequence variation observed was consistent with the operation of slippage-like mechanisms in the elongation of V4. For example, increasing length of the variable region was correlated with an increasing degree of repetition and was highly correlated with an increasing nucleotide bias. However, because primary sequence divergence was great among the taxa analyzed, Vogler, Selsh, and Hancock (1997) were unable to observe character transformations that would provide direct evidence for the operation of slippage.

The purpose of the present study was to conduct a detailed examination of the origin and evolution of long variable regions and unique secondary structures in the SSU rRNA of branchiopod crustaceans. Crease and Colbourne (1998) showed that the SSU rRNA gene of the branchiopod *D. pulex* is longer than that of any crustacean (2,293 nt) and contains unique helices, whereas this gene is of typical length and lacks these structures in the ancient branchiopod *Artemia.* Length and structural differences occurred mainly in the V4 and V7 variable regions. RT-PCR amplification of these regions from total cellular RNA only produced a fragment of the same size as the one amplified from genomic DNA, indicating that the expanded regions are present in the mature rRNA . In cases in which sequences are spliced out of the mature rRNA, the fragment produced by amplification of RNA is shorter than the one obtained from DNA (Wilcox et al. 1992). In the present study, we sequenced the V4 and V7 regions from 38 branchiopods in all extant orders. Our sampling was designed to allow both shallow (22 anomopod species) and deeper comparisons.

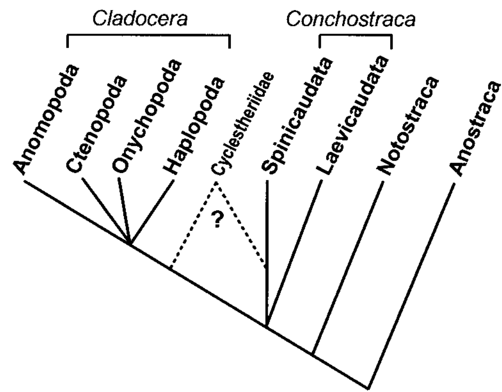

FIG. 1.—A phylogeny of ordinal relationships within the crustacean class Branchiopoda based on traditional ideas (see Walossek 1993; Oleson, Martin, and Roessler 1996). The position of the family Cyclestheriidae is uncertain, and its inclusion in either Conchostraca (traditional) or Cladocera (see Oleson, Martin, and Roessler 1996) have both been proposed. Fryer (1987) divided Cladocera and Conchostraca into four and two orders, respectively, and he, along with Starobogatov (1986), suggested that the cladoceran orders are not monophyletic. However, no alternate phylogeny reflecting the nonmonophyly hypothesis has yet been proposed.

Materials and Methods

Species Analyzed

Regions V4 and V7 of the SSU rRNA gene of representatives of all eight orders in the class Branchiopoda (Fryer 1987) were analyzed. A diagram showing traditional ideas about the phylogenetic relationships among these orders is shown in figure 1. Species names and sampling locations of the specimens used are indicated in table 1. Three sequences, from *Artemia* sp., *Branchinecta packardi,* and *D. pulex,* were obtained from GenBank.

PCR Amplification and Sequencing

Total genomic DNA was extracted from single individuals using the CTAB extraction method (Doyle and Doyle 1987). The primers used to amplify the V4 and V7 regions are from Crease and Colbourne (1998). The V4 primers amplify helices 20 through E23-9 in the model of Van de Peer et al. (1997), while the V7 primers amplify helices 38 through 46 (fig. 2).

Approximately 50 ng of genomic DNA was used as a template in a 50-μl amplification reaction containing 0.2 mM of each dNTP, 1.5 mM $MgCl_2$, 1 × reaction buffer (ID Technologies), 20 pmol of each primer, and 1 U of ID thermostable polymerase (ID Technologies). The reactions were carried out in 500-μl thin-walled tubes in an MJ-PTC 100 thermocycler. Amplification conditions were 37 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 60 s. Amplification products were electrophoresed on 1.8% agarose gels in 1 × TAE buffer and then purified from the agarose using the QIAEX II kit (QIAGEN). Approximately 50–100 ng of purified amplification products was sequenced with 3 pmol of one of the amplification primers and the ABI Prism TAQ FS dye terminator sequencing kit (Perkin-Elmer). Sequence reactions were analyzed on an ABI 377 automated sequencer (Perkin-Elmer). When there were no ambiguities in the data, the amplification product was not sequenced with the second primer. Sequences that were determined from only one strand are indicated in table 2. Direct sequencing of the PCR-amplified fragment was not successful for several taxa due to the occurrence of length variation among copies of the rRNA locus. In three cases (V4 in *Cyclestheria* and V7 in *Daphnia galeata* and *Simocephalus*), it was possible to separate the length variants on a 3% Metaphor agarose (FMC) gel and excise one of them for sequencing. Both strands were sequenced in each of these cases.

Sequence Analysis

A putative secondary structure for the entire SSU rRNA molecule of *Artemia,* based on the model of Van de Peer et al. (1997), was obtained from the SSU rRNA database on the World Wide Web (http://rrna.uia.ac.be/rrna). Based on comparisons with this and other crustacean sequences in the database, Crease and Colbourne

Table 1
Species, Orders, and Sampling Locations of the Taxa Analyzed in this Study

Order	Species	Sampling Location
Anostraca......	*Artemia* sp.	GenBank (accession number X01723)
	Branchinecta packardi	GenBank (accession number L26512)
	Polyartemiella hazeni	Pond near Toolik Lake, Alaska
	Streptocephalus dorothae	Triops Educational Science Inc., Florida
Notostraca	*Lepidurus arcticus*	Igloolik, Nunavit, Canada
	Triops longicaudatus	Triops Educational Science Inc., Florida
Laevicaudata ...	*Lynceus brachyurum*	Cow Creek, North Dakota
Spinicaudata ...	*Caenestheriella setosa*	Zap, North Dakota
	Limnadia sp.	Mount Hampton, Western Australia, Australia
	Cyclestheria hislopi	Karumba, Queensland, Australia
Haplopoda	*Leptodora kindti*	Douglas Lake, Michigan
Onychopoda....	*Bythotrephes longimanus*	Postsee, Germany
	Polyphemus pediculus	Wild Goose Lake, Michigan
Ctenopoda	*Holopedium amazonicum*	Flying Pond, Maine
	Holopedium gibberum	Prospect Lake, Ontario, Canada
	Sida crystallina	Wild Goose Lake, Michigan
Anomopoda	*Acantholeberis curvirostris*	Plastic Lake, Ontario, Canada
	Bosmina sp.	Derwentwater, U.K.
	Ceriodaphnia sp.	First Sister Lake, Michigan
	Chydorus sphaericus	First Sister Lake, Michigan
	Daphnia ambigua	Tucker Pond, Rhode Island
	Daphnia curvirostris	Tuktoyaktuk, Northwest Territories, Canada
	Daphnia ephemeralis	Long Point, Ontario, Canada
	Daphnia exilis	Amarillo, Texas
	Daphnia galeata	Derwentwater, U.K.
	Daphnia longiremis	Melville Peninsula, Northwest Territories, Canada
	Daphnia magna	Crescent Lake, Nebraska
	Daphnia obtusa	Champaign, Illinois
	Daphnia pulex	GenBank (accession number AF014011)
	Daphniopsis quadrangula	Colac, Victoria, Australia
	Daphniopsis truncata	Perth, Western Australia, Australia
	Eurycercus glacialis	Igloolik, Northwest Territories, Canada
	Eurycercus longirostris	First Sister Lake, Michigan
	Ilyocryptus sp.	Ann Arbor, Michigan
	Moina affinis	Valle Santiago, Mexico
	Ofryoxus gracilis	Raven Lake, Ontario, Canada
	Scapholeberis rammneri	Ann Arbor, Michigan
	Simocephalus vetulus	First Sister Lake, Michigan

NOTE.—Ordinal designations are those of Fryer (1987).

(1998) proposed a secondary structure for the SSU rRNA gene of *D. pulex* (fig. 2). The sequences for the other species analyzed in this study were first aligned to these two sequences using the DNASTAR software package in order to identify the core structural elements. Analysis of nucleotide composition of the sequences was also performed using DNASTAR. Subsequently, the program RNAdraw (Matzura and Wennborg 1996) was used to suggest a stable secondary structure for the expanded regions in V4 and V7. The default settings for all parameters except temperature were used to fold the sequences. The temperature was set to 20°C.

Sequence simplicity analysis of the length-variable regions in V4 and V7 was carried out using the SIMPLE34 program (Hancock and Armstrong 1994). This program searches for clustering of tri- and tetranucleotide motifs within a sequence. A relative simplicity factor (RSF) is generated by comparing the amount of clustering in the test sequence with the amount observed in random sequences of the same length and nucleotide composition. The RSF is 1.00 for sequences showing the same level of motif clustering as random sequences, and it is greater than 1.00 for sequences showing a high degree of clustering of simple sequence motifs. Statistical significance was calculated by the program as described in Hancock and Armstrong (1994).

To determine whether the predicted helix in the expanded region of V7 is significantly more stable than a random sequence, a Monte Carlo simulation of the distribution of free energies in 100 random sequences with the same nucleotide composition and length as the distal end of helix 43 was performed for four species: *Acantholeberis curvirostris, Holopedium gibberum, Leptodora kindti,* and *Daphniopsis quadrangula.* RNAdraw was used to predict a stable secondary structure and to calculate a corresponding free energy value for each of the 100 sequences.

Results

Size Variation and Secondary Structure of V4 and V7

All representatives of the orders Anostraca, Notostraca, Laevicaudata, and Spinicaudata, except the spinicaudatan *Cyclestheria* possess the typical eukaryotic configuration in both V4 and V7. Length variation is nearly absent in these taxa; the region of V4 between

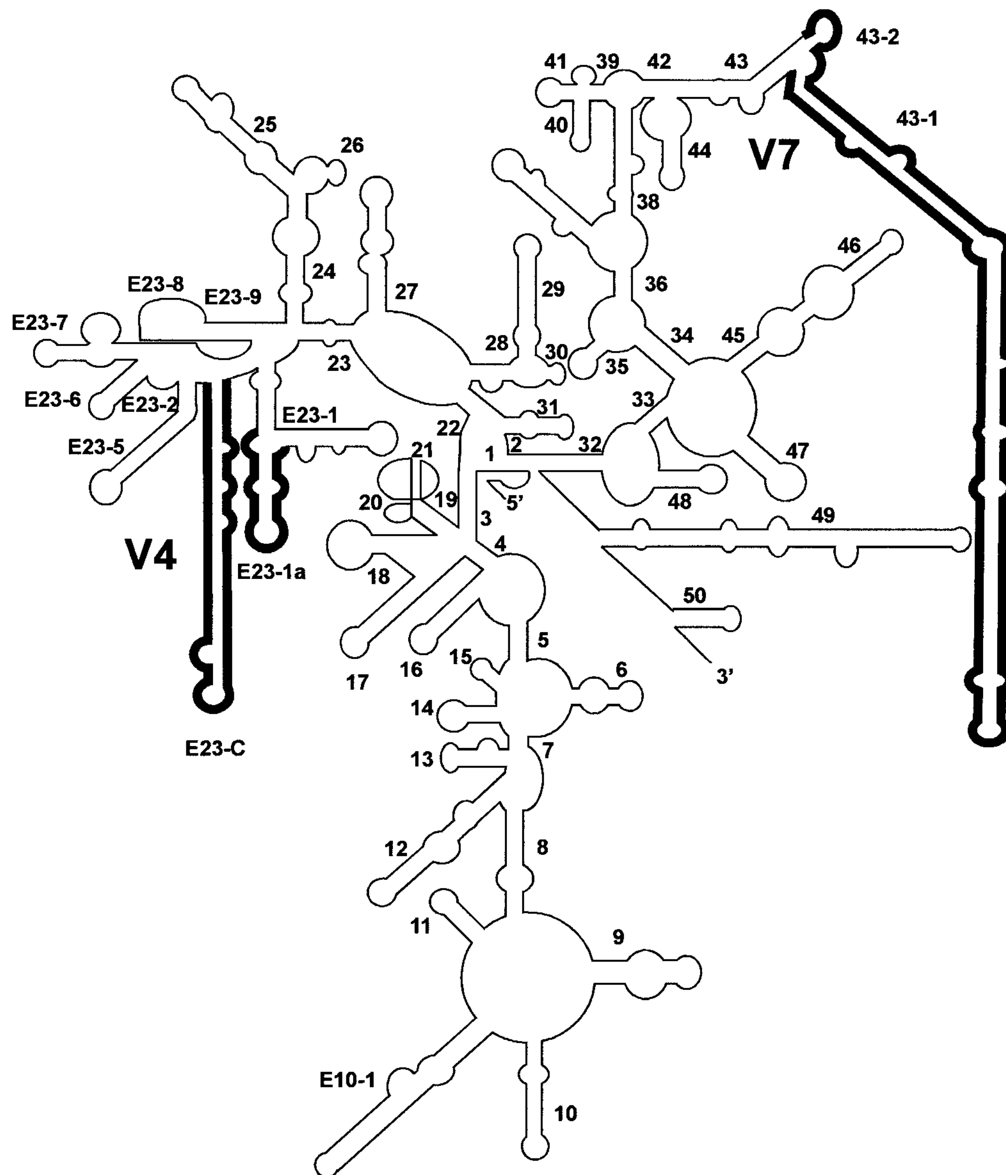

FIG. 2.—Secondary structure of the SSU rRNA of *Daphnia pulex,* redrawn from Crease and Colbourne (1998). Helix numbering and locations of V4 and V7 are from Van de Peer et al. (1997). Paired regions are shown as parallel lines, while curved lines denote unpaired regions. The helices unique to cladoceran branchiopods are indicated with thick lines.

helix 23 and helix E23-2 varies from 59 to 61 nt, while the distal end of helix 43 in V7 varies from 22 to 23 nt (table 2). The corresponding regions in all taxa representing the four orders that previously composed the Cladocera (Anomopoda, Ctenopoda, Haplopoda, and Onychopoda), as well as *Cyclestheria,* show substantial increases in length ranging from 141 to 297 nt in V4 and from 130 to 360 nt in V7. The most extreme increases in length have occurred among the daphniids (*Ceriodaphnia, Daphnia, Daphniopsis, Scapholeberis,* and *Simocephalus*) and the macrothricids (*Acantholeberis, Ilyocryptus,* and *Ofryoxus*) in the order Anomopoda. There is a highly significant positive correlation between length and GC content of the expanded sequences in the two regions (fig. 3 and table 3).

Although lengths and primary sequences of the expanded regions in V4 and V7 vary substantially among the cladocerans and *Cyclestheria,* the secondary structures of both regions are quite conserved and correspond well to the secondary structure initially proposed for *D. pulex* by Crease and Colbourne (1998). One representative from each of the cladoceran orders is shown in figure 4.

The extra length of V4 is due to two helical structures that are not present in the other branchiopods: a "side branch" that occurs near the middle of helix E23-1, hereafter referred to as E23-1a, and an entirely new helix, hereafter referred to as E23-c, that occurs between E23-1 and E23-2 (fig. 4*A*). Most of the length variation in V4 is associated with E23-c, as E23-1a varies only

Table 2
Lengths (nt), Levels of Repetition (RSF), and Nucleotide Compositions (%GC, %U) of Partial SSU rRNA Sequences of Branchiopod Crustaceans

	Core			V4				V7			
Taxon	nt	%GC	%U	nt	RSF	%GC	%U	nt	RSF	%GC	%U
Artemia sp.	633	50.6	25.8	60	—	61.7	25.0	23	—	56.5	26.1
Branchinecta packardi	632	50.3	26.0	60	—	63.4	21.7	22	—	59.1	27.3
Polyartemiella hazeni	610	50.3	25.9	60[a]	—	61.7	23.3	22[a]	—	59.1	27.3
Streptocephalus dorothae	611	50.2	25.9	60[a]	—	61.7	21.7	22[a]	—	59.1	27.3
Lepidurus arcticus	612	51.0	25.2	60[a]	—	65.0	18.3	23[a]	—	60.8	30.4
Triops longicaudatus	611	51.2	25.2	60[a]	—	66.6	18.3	23[a]	—	60.8	30.4
Lynceus brachyurum	589	50.7	26.0	59[a]	—	61.0	25.4	23[a]	—	65.2	30.4
Caenestheriella setosa	611	51.7	25.4	61[a]	—	65.5	19.7	23[a]	—	69.5	21.7
Limnadia sp.	611	51.4	25.4	60[a]	—	68.3	18.3	23[a]	—	69.5	21.7
Mean		50.8	25.6			63.9	21.3			62.2	27.0
Cyclestheria hislopi	603	49.6	26.5	149	0.74	53.0	27.5	233	1.81	61.4	26.6
Leptodora kindti	579	49.8	26.7	174	1.16	58.0	25.9	280	1.67	52.5	35.4
Bythotrephes longimanus	614	50.2	26.7	164[a]	0.70	62.3	20.7	230[a]	1.46	61.8	28.7
Polyphemus pediculus	624	50.0	27.1	171	0.94	57.9	28.1	209[a]	1.25	64.1	30.1
Holopedium amazonicum	526	52.3	25.1	141	0.83	66.7	18.4	161	2.08	65.2	25.5
Holopedium gibberum	596	50.2	26.3	146[a]	1.27	55.4	25.3	173[a]	1.32	71.1	22.0
Sida crystallina	599	50.4	26.4	169[a]	1.35	66.3	21.3	223	1.49	65.5	28.7
Acantholeberis curvirostris	633	52.5	24.2	193[a]	1.06	69.4	19.7	204	1.30	61.3	31.4
Bosmina sp.	NA[b]	—	—	NA	—	—	—	153	1.33	61.4	27.5
Ceriodaphnia sp.	608	52.5	25.0	195	1.29	63.1	22.1	241[a]	1.51	65.6	27.4
Chydorus sphaericus	554	49.6	26.4	161	1.25	60.9	21.7	169	1.86	64.5	26.0
Daphnia ambigua	NA	—	—	NA	—	—	—	288	1.12	53.8	33.3
Daphnia curvirostris	633	52.3	25.6	210	1.24	57.1	27.6	275	1.43	56.8	29.8
Daphnia ephemeralis	617	51.7	25.8	201	1.29	60.2	25.4	261[a]	1.44	59.0	31.0
Daphnia exilis	618	51.8	26.1	223	1.50	56.5	27.8	281[a]	1.15	61.2	29.5
Daphnia galeata	575	52.0	25.6	229[a]	1.60	60.3	27.5	273	1.38	57.9	32.6
Daphnia longiremis	643	52.5	25.0	270	1.17	58.2	29.3	340	1.32	58.3	28.5
Daphnia magna	617	51.7	25.6	236	1.32	63.1	26.3	291[a]	1.36	60.8	28.5
Daphnia obtusa	NA	—	—	204	1.03	59.4	27.5	NA	—	—	—
Daphnia pulex	651	52.7	24.4	215	1.06	56.3	29.8	273	1.25	58.2	30.0
Daphniopsis quadrangula	618	51.8	25.4	279	1.28	59.9	27.6	360	1.36	59.5	30.3
Daphniopsis truncata	592	51.2	26.5	297	1.50	58.0	29.0	323	1.30	55.4	30.7
Eurycercus glacialis	651	52.4	24.7	165	1.08	65.5	23.6	250	1.25	62.0	30.0
Eurycercus longirostris	636	52.7	24.5	166	1.34	67.5	22.3	222	1.72	62.5	29.7
Ilyocryptus sp.	NA	—	—	NA	—	—	—	194	1.24	69.1	19.6
Moina affinis	617	53.5	24.5	156	1.45	71.1	17.3	130	1.85	74.7	16.2
Ofryoxus gracilis	583	51.5	25.8	194	1.26	67.1	21.7	243	1.46	67.4	28.4
Scapholeberis rammneri	620	51.3	26.6	190	1.06	65.3	22.1	251	1.89	63.4	27.1
Simocephalus vetulus	598	52.5	24.9	207	1.16	64.2	23.7	259	1.80	65.2	27.8
Mean		51.5	25.7			61.6	24.6			62.1	28.3

Note.—With the exception of *Cyclestheria*, taxa in the same order are grouped.
[a] Sequenced on one strand only.
[b] NA = sequence not available.

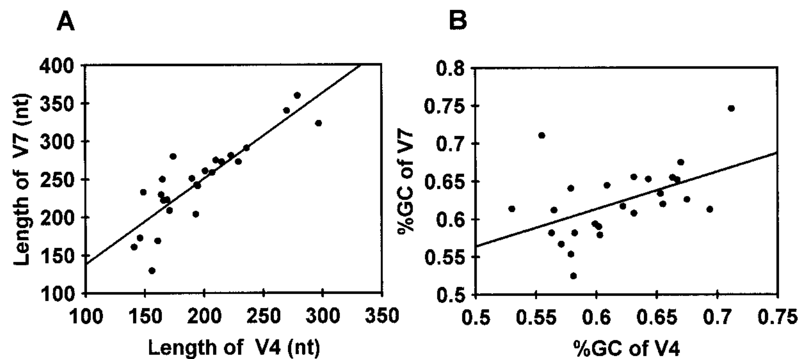

Fig. 3.—Correlation between (*A*) the lengths and (*B*) the nucleotide compositions of regions V4 and V7 in cladoceran crustaceans. See table 3 for full regression analysis results.

Table 3
Regression Analysis of Lengths (nt), Levels of Repetition (RSF), and Nucleotide Compositions (%GC) in Regions V4 and V7 of SSU rRNA in Cladoceran Crustaceans

Comparison	R^2	a	b
V4 nt vs. V7 nt.	0.74***	25.7	1.12
V4 %GC vs. V7 %GC.	0.24**	0.32	0.49
V4			
nt vs. %GC (E23-c only) . .	0.32**	0.79	−0.001
nt vs. %G (E23-c only). . . .	0.28**	0.41	−0.0005
nt vs. %A (E23-c only). . . .	0.02^{NS}	0.04	0.0001
nt vs. %U (E23-c only). . . .	0.43***	0.12	0.001
nt vs. %C (E23-c only). . . .	0.29**	0.39	−0.001
RSF vs. nt	0.20*	0.71	0.003
RSF vs. %GC	0.01^{NS}	0.93	0.42
V7			
nt vs. %GC	0.50***	0.77	−0.0006
nt vs. %G.	0.30**	0.44	−0.0003
nt vs. %A.	0.17*	0.05	0.0002
nt vs. %U.	0.40***	0.17	0.0004
nt vs. %C.	0.41***	0.33	−0.0004
RSF vs. nt	0.16*	1.92	−0.002
RSF vs. %GC	0.10^{NS}	0.46	1.63

NOTE.—The fitted regression line is of the form $y = a + bx$. NS = not significant; * $0.01 < P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

from 26 to 29 nt in the cladocerans (fig. 5*A*). Its length in *Cyclestheria* is 16 nt, and it originates at the site of a single nucleotide bulge in the noncladocerans. The distal end of E23-c is extremely variable, making it difficult to align the sequences unless they belong to species from the same genus. However, there is similarity at the proximal end of this helix, even among distantly related taxa, although *Cyclestheria* is divergent from the cladocerans in that there is a second very short helix next to E23-c (fig. 4*A*). A partial alignment of the length-variable region of V4, based on both primary sequence and secondary structure, is suggested in figure 5*A*. The occurrence of multiple compensatory mutations (indicated by asterisks in fig. 5*A*) that maintain base pairing provides strong support for the existence of helices E23-1a and E23-c (Gerbi 1985 and references within).

The extra length of V7 is due to the substantial elongation of helix 43 from the terminal loop, hereafter referred to as 43-1. In the cladocerans, helix 43-1 is preceded by a short side branch, hereafter referred to as 43-2, whereas a single mismatched nucleotide pair (U-C) occurs at this position in *Cyclestheria* (figs. 4*B* and 5*B*). The length and sequence of 43-2 and of the distal end of 43-1 are highly divergent among taxa, as is the case for E23-c. However, like E23-c, the proximal end of 43-1 is considerably more conserved, and an alignment is suggested in figure 5*B*. Again, the occurrence of compensatory mutations at a substantial number of paired nucleotides at the proximal end of this helix provides strong support for its existence.

The frequency distribution of free energy values for 100 random sequences corresponding in length and nucleotide composition to helices 43, 43-1, and 43-2 from four of the cladocerans shows that the probability of obtaining a structure from a random sequence that is more stable than the structure predicted from the actual sequence is very low ($P \leq 0.01$; fig. 6). These results suggest that the ability to form a stable helix is an important structural constraint on the expanded region.

Nucleotide Composition of V4 and V7

Nucleotide compositions were calculated separately for V4 between helices 23 and E23-2 (E23-1 + E23-c in the cladocerans), for the distal end of helix 43 in V7 (including 43-1 and 43-2 in the cladocerans), and for the remaining core regions (table 2). Although the core sequence obtained by amplification with the V4 and V7 primers is only about 610 nt, its nucleotide composition appears to be representative of the core regions as a whole, because the GC content of this sequence in taxa whose SSU rRNA genes have been completely sequenced is very similar to the value obtained for the complete sequence: *Artemia* (50.6% vs. 50.8%), *Branchinecta* (50.3% vs. 50.4%), and *D. pulex* (52.7% vs. 52.6%).

The mean GC content of the core regions in the noncladocerans (50.8%) is slightly lower ($t = 2.04$, $P = 0.02$) than those of the cladocerans and *Cyclestheria* (51.5%; table 2). Both values are typical for eukaryotic SSU rRNA, which tends to be slightly biased toward GC. However, the variable regions of the noncladocerans are even more biased, with mean GC contents of 63.9% for V4 and 62.2% for V7 (table 2). In the cladocerans and *Cyclestheria,* the GC contents of these two regions are highly variable, ranging from 53.0% to 71.1% for V4 and from 52.5% to 74.7% for V7 (table 2). The tendency for variable regions to have a more biased nucleotide composition than the rest of the SSU rRNA has been observed in other organisms (Nunn et al. 1996; Vogler, Selsh, and Hancock 1997), and, in general, it is expected that the bias will increase with increasing length of the variable region. However, in the case of the cladocerans, the correlation is highly significant and negative for both V4 and V7 (table 3). Thus, the GC bias is most extreme for the shortest variable regions. A more detailed examination of nucleotide composition in these regions shows that the decrease in GC bias is largely due to a highly significant increase in the proportion of U at the expense of both G and C (fig. 7 and table 3). The proportion of A in these regions tends to be very low regardless of their length, although it does increase somewhat as the helices increase in length.

It has been observed that the nucleotide composition of rRNA helices differs between stems and unpaired regions, including bulges, mismatched pairs, and terminal loops (Vawter and Brown 1993; Nunn et al. 1996). In helix 43-1, the mean GC content of stems is 70.7% (range 62.9%–85.5%) which is significantly higher ($t = 19.1$, $P < 0.00001$) than the mean GC content of unpaired regions at 30.2% (range 14.7%–54.6%). The AU bias in the loops and bulges is primarily due to the high proportion of U, at 44.7% on average (table 4). The proportion of A in the stems is extremely low, with a mean of only 4.7%, but the proportions of C (25.4%) and U (24.6%) are very similar. This similarity is reflected in the ratio of G-C to G-U base pairs in the

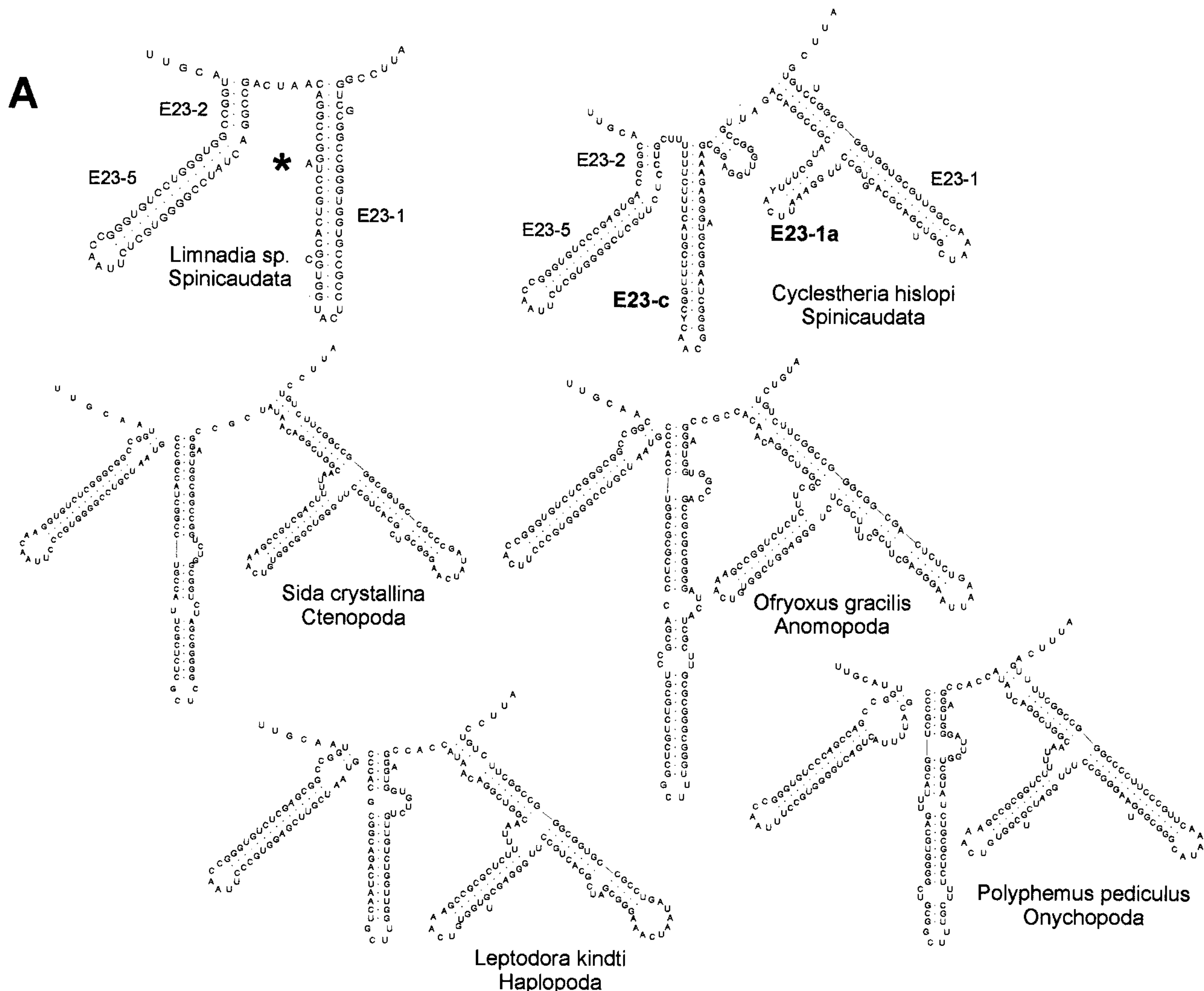

FIG. 4.—Representative secondary structures of (*A*) V4 and (*B*) V7 in branchiopod crustaceans. The diagrams were generated using the computer program CARD (Winnepenninckx et al. 1995). The proposed secondary structures of helices unique to the Cladocera (E23-1a, E23-c, 43-1, and 43-2) were determined using RNAdraw (Matzura and Wennborg 1996). The asterisk next to the V4 sequence of *Limnadia* indicates the unpaired nucleotide from which helix E23-c originates in the cladocerans. The asterisk over the V7 sequence of *Cyclestheria* indicates the U-C mismatch from which 43-2 originates in the cladocerans.

stems, which ranges from 0.89 to 2.62, with a mean of 1.36 (table 4). As expected, this ratio is highly correlated with increasing helix length due to the overall increase in the proportion of U (table 3). The mean proportion of nucleotides that are involved in pairing is 85.9%, and there is no correlation ($R^2 = 0.04$, $P = 0.28$) between this value and the length of the helix.

Level of Repetition in V4 and V7

Although the RSF is greater than 1.00 for all the cladoceran variable regions except V4 of *Cyclestheria, Bythotrephes, Holopedium amazonicum,* and *Polyphemus,* none of these values are significant (table 2), indicating the absence of substantial clustering of tri- and tetranucleotide repeats. Regression analysis of RSF against length and GC content shows a significant positive correlation between the length of V4 and RSF but a significant negative correlation between the length of V7 and RSF (fig. 7 and table 3). There is no significant correlation between GC content and RSF in either region. Even though the RSF values are not significant, inspection of the helices shown in figure 3 reveals the occurrence of short arrays of di- and trinucleotide motifs such as GC, GU, GGC, GGU, GCC and GUU, as well as short homopolymer runs, suggesting that slippage replication may have played some role in their elongation. Patterns of repetition are even more evident when the C's and U's are all converted to Y's (pyrimidines), as has been done for helix 43-1 of *Holopedium gibberum* in figure 8*B*. It is clear that there is some repetition of the motifs GYY and GGY in the distal region of this helix.

Patterns of Divergence in E23-c and 43-1 Among Closely Related Species

Comparison of sequences from closely related taxa should provide insight into the types of sequence changes that occur during the early phases of divergence

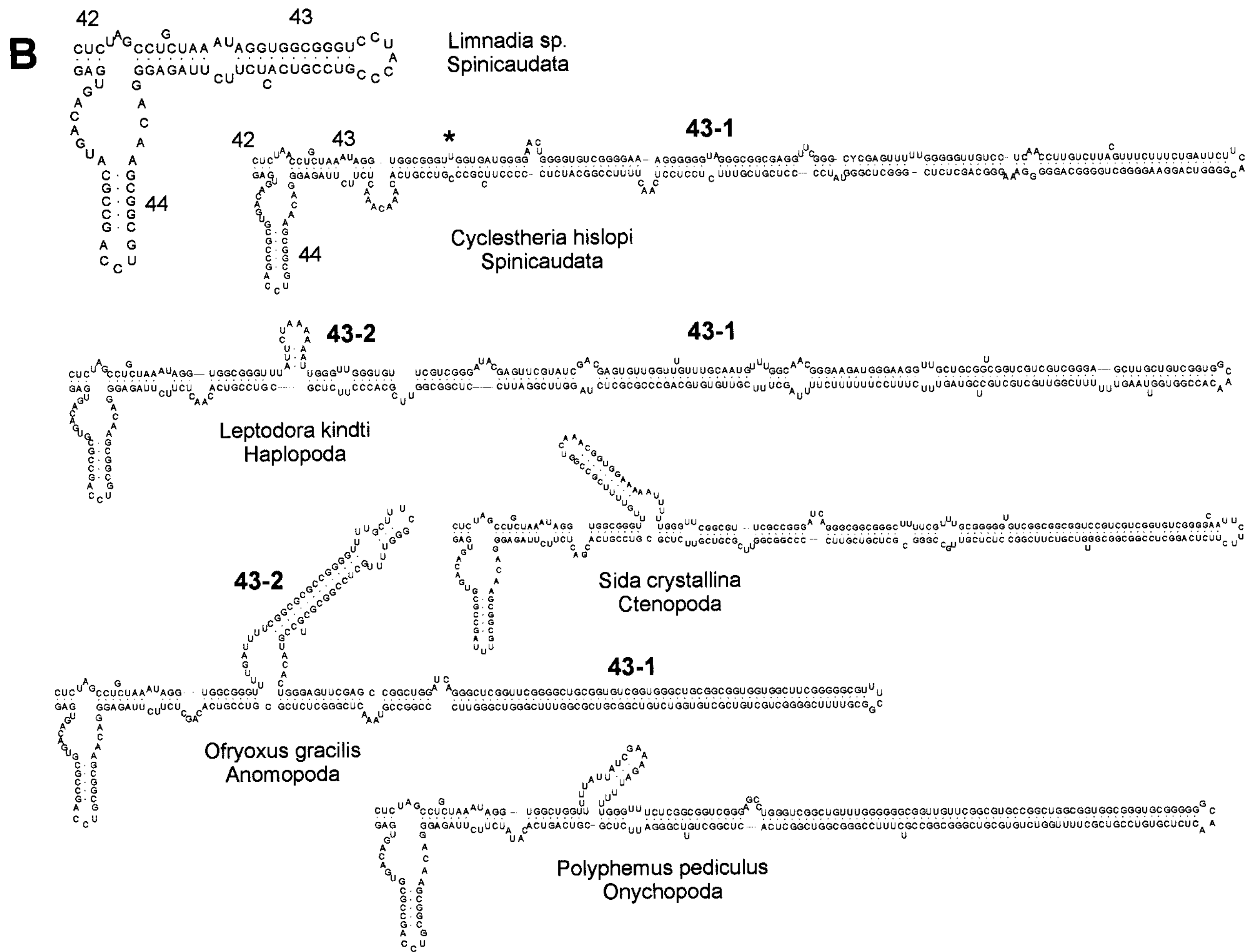

FIG. 4 *(Continued)*

in expansion helices (Vogler, Selsh, and Hancock 1997). Several such opportunities are available with the present data including two or three species of *Holopedium, Eurycercus, Daphnioposis,* and each of the *Daphnia* subgenera, *Daphnia, Hyalodaphnia,* and *Ctenodaphnia.* Such comparisons (fig. 8) show that point mutations occur throughout E23-c and 43-1. The occurrence of compensatory mutations, even at the divergent distal ends of these helices, is also evident among closely related species.

Length increases in E23-c and 43-1 tend to occur in the distal regions of the helices. Some of the length increases seem to have occurred at terminal loops, but a substantial number have not. In such cases, an increase in one strand of a helix has been compensated by a change at the corresponding location on the opposing strand. The result is the insertion of a stem ''module'' that leaves the helix downstream from the insertion site unchanged (E23-c in *Daphniopsis* and *Hyalodaphnia* and 43–1 in *Daphniopsis, Ctenodaphnia,* and *Eurycercus*). A very different module occurs at approximately the same location in several species pairs (E23-c in *Ctenodaphnia* and 43-1 in *Eurycercus, Daphniopsis,* and *Ctenodaphnia*), suggesting that length expansion is more likely to occur at some positions along a helix than at others.

The modules in 43-1 of the two *Holopedium* species appear to consist of the trinucleotide motifs GYY and GGY, with a preference for C over U. In contrast, the modules in *Daphnia* and *Daphniopsis* seem to consist of homopolymer runs of U's on one strand compensated by a run of purines (usually G's) on the other strand. Even so, the compensation is not perfect, leaving an excess of U's in the module as a whole. In addition, the proximal ends of the modules are often associated with unpaired U's or U-U and U-C mismatches (fig. 8). This is especially evident in comparisons involving helices that are very similar to one another, such as E23-c and 43-1 in *Daphniopsis* and the subgenus *Daphnia.* To determine if this pattern could account for the significant trend toward an increase in the proportion of U's with an increase in length (table 3), nucleotide compositions were determined separately for the distal and the proximal ends of E23-c and 43-1 in *Daphnia* and *Daphniopsis.* The ''boundary'' was placed at the point at which it is no longer possible to align the sequences from all 10 species with confidence (fig. 8). The results (table 5) show that the distal ends of both helices have

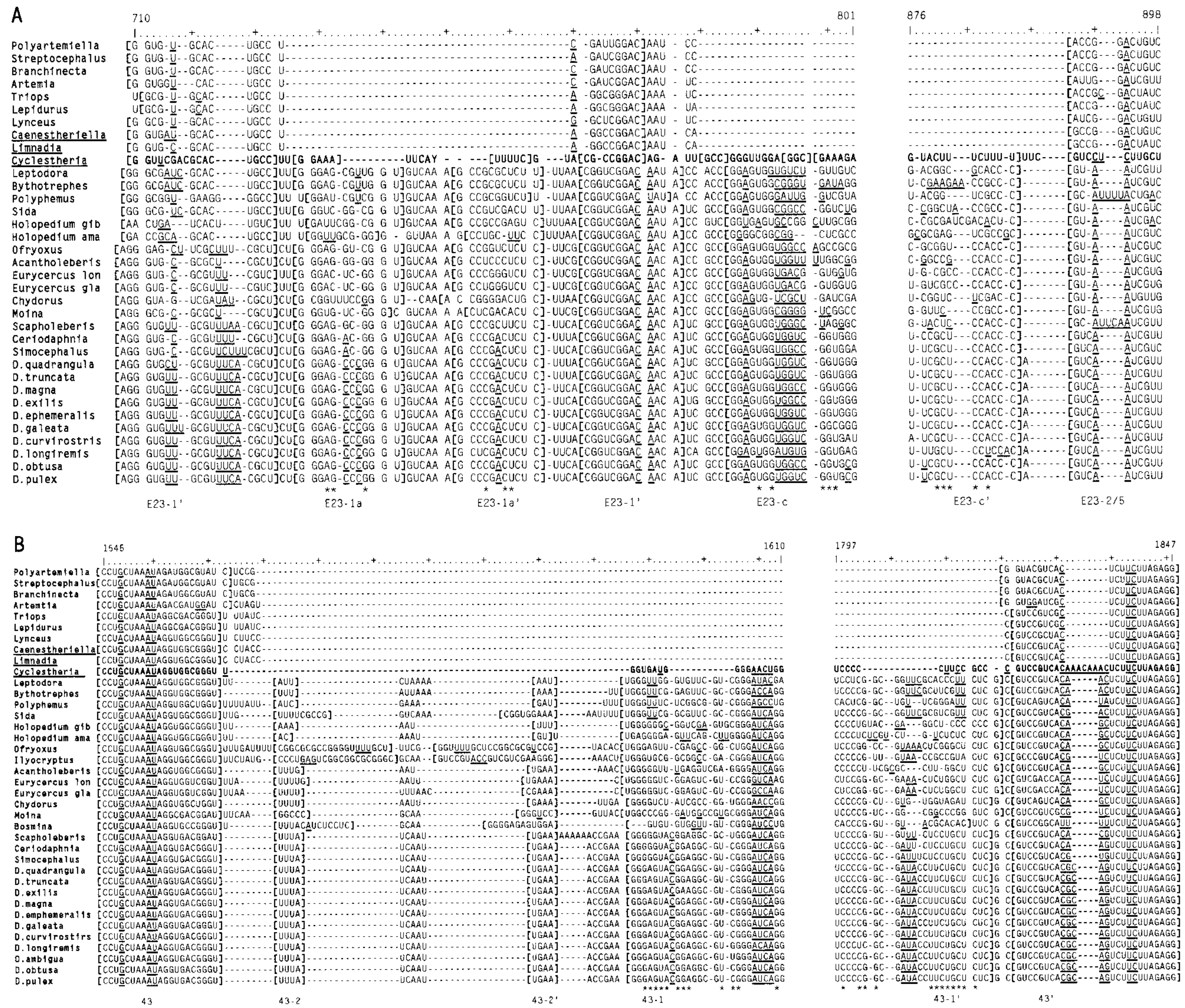

FIG. 5.—Partial alignment of (*A*) V4 and (*B*) V7 in branchiopod crustaceans. Taxa in the same order are grouped. The three Spinicaudata species names are underlined, and the sequence of *Cyclestheria* is shown in bold. The alignment was performed by eye based on primary sequence and secondary structure as suggested by RNAdraw. Nucleotides forming one or the other strand of a helix are enclosed in square brackets. Bulges or mismatched nucleotides within the helix are underlined. Nucleotide positions at which compensatory mutations that maintain base pairing have occurred in at least one taxon are indicated by asterisks along the bases of the alignments. The first and last nucleotide positions in each sequence block are numbered according to the SSU rRNA gene sequence of *D. pulex.* A complete alignment of the sequences is available from tcrease@uoguelph.ca on request. The sequences have been deposited in GenBank under accession numbers AF070093—AF070124 and AF070486—AF070519.

significantly higher mean proportions of U's than do their proximal ends (E23-c: 35.7% vs. 28.6%, $t = -4.30$, $P = 0.001$; 43–1: 33.6% vs. 22.9%, $t = -11.45$, $P < 0.0001$). Unexpectedly, the correlation between the proportions of U's and the lengths of the distal ends of both helices is slightly negative but not significant (data not shown), suggesting that an increase in the lengths of these helices beyond a certain point via the addition of more stem modules does not significantly alter their nucleotide compositions.

Intraindividual length variation occurred in the variable regions of several taxa including E23-c of *Cyclestheria, D. ambigua,* and *Ilyocryptus* and 43-1 of *D. galeata, D. obtusa,* and *Simocephalus.* It was possible to obtain the complete sequence of one variant for some of these taxa (see *Materials and Methods*), but the approach used was not successful for E23-c in *D. ambigua* and *Ilyocryptus* or for 43-1 in *D. obtusa.* Because *D. pulex* and *D. obtusa* are very similar to one another, it is possible to determine exactly where the length variation occurs in 43-1 of the latter species (fig. 8*D*). The sequence becomes unreadable at the same location on *both* strands of the helix, suggesting that an expansion module is present in some copies of the SSU rRNA gene but not in others. The termination of a linear sequence leaves a characteristic peak pattern on the electropherogram generated by the automated sequencer. Because of this, it can be seen that there are at least three length variants present in the PCR-amplified fragment of V7 from *D. obtusa.*

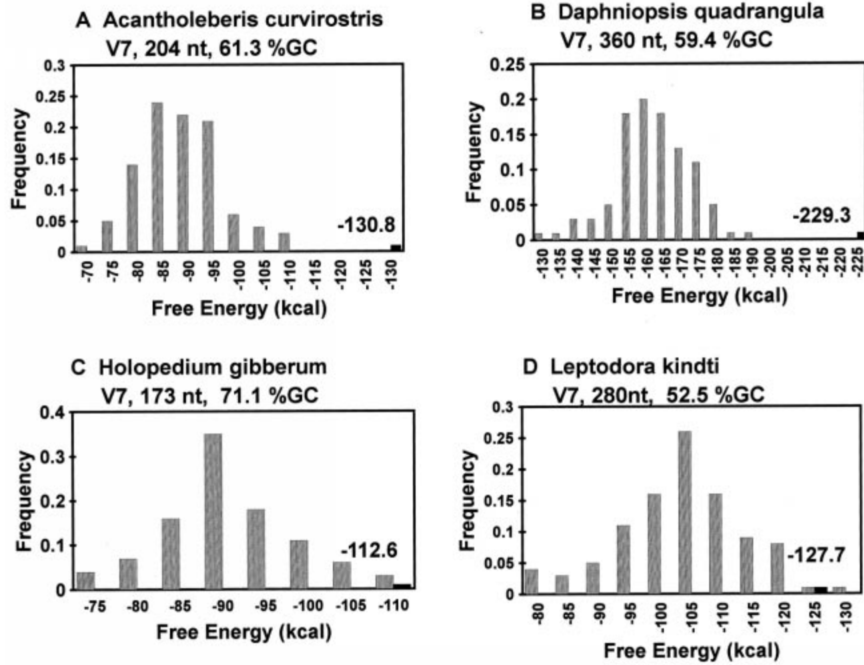

FIG. 6.—Frequency distribution of free energy values for 100 random sequences corresponding in length and nucleotide composition to helices 43, 43-1, and 43-2 in V7 of four cladoceran crustaceans. The length of the sequence (nt), its nucleotide composition (%GC), and the free energy of the actual sequence are indicated on each histogram. *A, Acantholeberis curvirostris. B, Daphniopsis quadrangula. C, Holopedium gibberum. D, Leptodora kindti.*

The mean GC contents of the proximal end of E23-c (61.1%, range 54.8%–67.3%) and 43-1 (64.9%, range 61.1%–74.0%) in *Daphnia* and *Daphniopsis* (table 5) are very similar to the mean GC contents of the corresponding regions in the noncladocerans (63.9% for V4 and 62.2% for V7; table 2). To determine if this is the case in general, the GC content of the proximal end of helix 43-1 (fig. 5*B*) was determined for all of the cladoceran taxa and found to be 67.1%, on average (range 60.4%–80.0%), which is even higher than the value observed for the noncladocerans. The mean value for this region in *Daphnia* and *Daphniopsis,* based on 51 nt, is

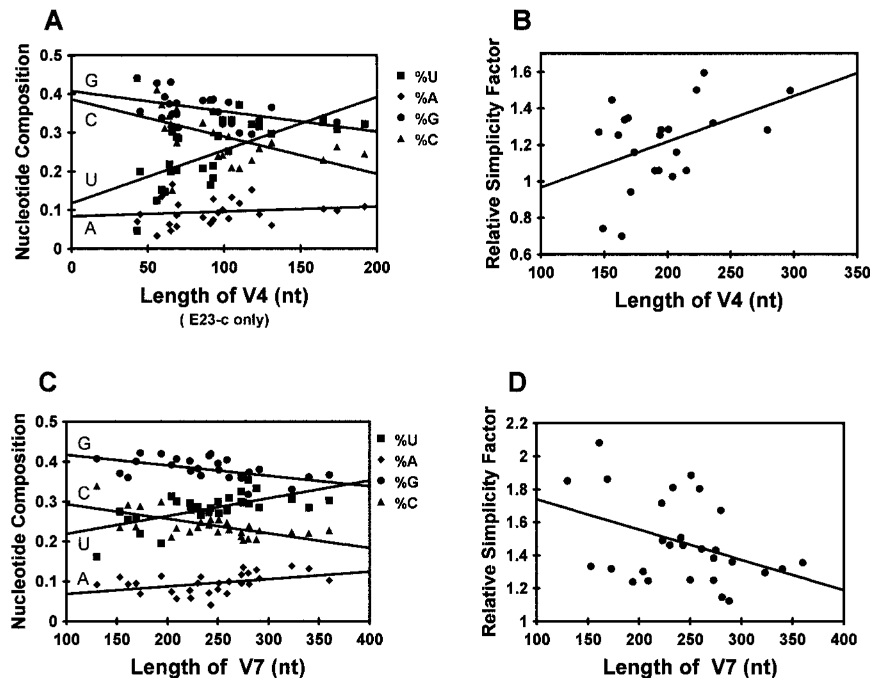

FIG. 7.—Regression analyses of length, nucleotide composition, and level of repetition (RSF) for (*A* and *B*) V4, and (*C* and *D*) V7. See table 3 for full regression analysis results.

Table 4
Nucleotide Compositions of Stems and Loops/Bulges in Helix 43-1 of Cladoceran Crustaceans

	43-1 Stems							43-1 Loops/Bulges				
Taxon	nt	%G	%C	%A	%U	GC/GU[a]	%Stem	nt	%G	%C	%A	%U
Acantholeberis curvirostris	156	46.2	21.8	3.8	28.2	0.89	90.7	16	6.3	18.9	12.5	62.5
Bosmina sp.	90	43.3	22.2	6.7	27.8	1.05	88.2	12	16.7	25.0	33.3	25.0
Bythotrephes longimanus	162	45.1	26.5	4.9	23.5	1.43	81.4	37	21.6	16.2	18.9	43.3
Ceriodaphnia sp.	186	47.8	24.2	2.2	25.8	1.02	91.2	18	11.1	27.8	22.2	38.9
Chydorus sphaericus	110	46.4	23.6	3.6	26.4	1.04	90.9	11	18.2	36.4	36.4	9.0
Cyclestheria hislopi	182	41.8	24.7	8.2	25.3	1.45	84.7	33	3.0	30.3	33.3	33.3
Daphnia ambigua	202	41.6	21.3	8.4	28.7	1.05	79.8	51	7.8	17.6	21.6	53.0
Daphnia curvirostris	202	43.6	20.8	6.4	29.2	0.91	84.2	38	10.5	15.8	39.5	34.2
Daphnia ephemeralis	186	44.1	24.7	5.9	25.3	1.28	83.4	37	8.1	16.2	13.5	62.2
Daphnia exilis	204	45.1	26.0	4.9	24.0	1.36	83.6	40	10.0	15.0	17.5	57.5
Daphnia galeata	202	44.6	22.3	5.4	27.7	1.00	85.6	34	2.9	11.8	17.6	67.7
Daphnia longiremis	300	43.0	20.3	7.0	29.7	0.90	88.0	41	9.8	19.5	31.7	39.0
Daphnia magna	224	44.2	23.2	5.8	26.8	1.11	87.8	31	9.7	22.6	29.0	38.7
Daphnia obtusa	NA[b]	—	—	—	—	—	—	—	—	—	—	—
Daphnia pulex	200	43.5	23.0	6.5	27.0	1.12	85.5	34	2.9	17.6	32.4	47.1
Daphniopsis quadrangula	284	42.6	23.6	7.4	26.4	1.24	88.5	37	5.4	16.2	16.2	62.2
Daphniopsis truncata	242	39.7	24.0	10.3	26.0	1.53	84.6	44	4.5	13.6	25.0	56.9
Eurycercus glacialis	190	46.3	22.6	3.7	27.4	0.96	89.2	23	17.4	21.7	21.7	39.2
Eurycercus longirostris	170	46.5	24.1	3.5	25.9	1.08	90.9	17	11.8	17.6	23.5	47.1
Holopedium amazonicum	104	45.2	32.7	4.8	17.3	2.62	78.2	29	3.4	27.6	20.7	48.3
Holopedium gibberum	120	49.2	34.2	0.8	15.8	2.28	82.8	25	28.0	16.0	24.0	32.0
Ilyocryptus sp.	102	47.1	31.4	2.9	18.6	2.00	85.7	17	5.9	17.6	53.0	23.5
Leptodora kindti	198	40.9	22.2	9.1	27.8	1.19	77.3	58	3.4	13.8	24.1	58.7
Moina affinis	76	50.0	35.5	0.0	14.5	2.45	86.4	12	25.0	16.7	41.6	16.7
Ofryoxus gracilis	154	48.7	24.7	1.3	25.3	1.03	92.2	13	7.7	23.1	38.4	30.8
Polyphemus pediculus	154	48.7	25.3	1.3	24.7	1.08	89.5	18	11.1	27.8	16.7	44.4
Scapholeberis rammneri	178	47.2	27.5	2.8	22.5	1.40	85.2	31	9.7	22.6	12.9	54.8
Sida crystallina	138	48.6	31.9	1.4	18.1	1.91	79.8	35	5.7	22.9	8.6	62.8
Simocephalus vetulus	200	47.5	26.5	2.5	23.5	1.26	90.1	22	4.5	13.6	18.2	63.7
Mean		45.3	25.4	4.7	24.6	1.36	85.9		10.1	20.1	25.1	44.7

[a] Ratio of G-C to G-U base pairs in the stem of 43-1.
[b] Sequence not available.

66.1%, which is similar to the value obtained above (64.9%) based on 71–75 nt (table 5).

Because extreme increases in the lengths of E23-c and 43-1 have only occurred in *Daphnia* and *Daphniopsis,* and the mechanisms by which this has occurred may not be common to the other cladocerans, the regression analysis of the lengths of V4 and V7 against the proportion of U's was redone excluding the 10 *Daphnia* and *Daphniopsis* species. The resulting correlation was positive ($y = 0.10 + 0.001x$) but not significant for V4 ($R^2 = 0.11$, $P = 0.24$), and positive ($y = 0.12 + 0.0007x$) but highly significant for V7 ($R^2 = 0.47$, $P = 0.002$) suggesting that the elongation of this region via the preferential addition of U's is a general phenomenon in the Cladocera.

Discussion

Phylogenetic Implications

Based on the substantial morphological and developmental diversity among the taxa that comprise the order Cladocera, Fryer (1987) argued that it is not a monophyletic group and divided it into four separate orders (fig. 1). However, all of the taxa analyzed in this study that belong to one of these orders possess the same unique helices in V4 and V7 of the SSU rRNA gene. The fact that they occur at precisely the same location in all cladocerans but do not exist in the other branchiopods (except *Cyclestheria*), along with the fact that the sequence forming the proximal end of each helix is relatively conserved, strongly suggests that they evolved only once in an ancestor that is common to all cladocerans. In addition, phylogenetic analysis of the class Branchiopoda based on the conserved core elements in the V4 and V7 sequences obtained for this study, as well as core elements flanking D2 in the LSU rRNA gene (unpublished data), strongly supports a hypothesis of monophyly. Even so, the analysis also recognized subgroupings within the Cladocera that correspond to the four orders proposed by Fryer (1987).

Cyclestheria is the only noncladoceran to possess the unique helices in V4 and V7. It also reproduces by cyclical parthenogenesis (Roessler 1995; Oleson, Martin, and Roessler 1996), a very rare mode of reproduction among animals that is found in no other group of crustaceans except the Cladocera. It has long been hypothesized that cladocerans evolved from cyclestheriids, which possess several morphological features that show greater similarity to cladocerans than to other spinicaudatans (Oleson, Martin, and Roessler 1996 and references within). The phylogenetic analysis (unpublished data), in which *Cyclestheria* occupied a basal position to the Cladocera, provides additional evidence that the cyclestheriids are the transitional group between spinicaudatans and cladocerans (fig. 2) and is consistent with

Table 5
Nucleotide Compositions of the Proximal and Distal Ends of Expansion Helices in *Daphnia* and *Daphniopsis*

	E23-c Proximal End					E23-c Distal End					43-1 Proximal End					43-1 Distal End				
Taxon	nt	%G	%C	%A	%U	nt	%G	%C	%A	%U	nt	%G	%C	%A	%U	nt	%G	%C	%A	%U
Daphnia ambigua	NA[a]	—	—	—	—	—	—	—	—	—	75	39.4	25.4	11.3	23.9	180	32.8	17.8	11.1	38.3
Daphnia curvirostris	53	35.8	18.9	13.2	32.1	47	34.0	17.1	14.9	34.0	72	37.5	25.0	12.5	25.0	166	38.5	18.1	11.5	31.9
Daphnia ephemeralis	52	40.3	21.2	9.6	28.9	38	35.2	21.6	10.8	32.4	71	39.4	26.8	11.3	22.5	153	37.2	21.6	5.9	35.3
Daphnia exilis	53	37.8	24.5	11.3	26.4	59	25.4	17.0	20.3	37.3	71	38.0	26.8	12.7	22.5	173	39.9	23.1	4.6	32.4
Daphnia galeata	53	39.7	24.5	9.4	26.4	64	32.8	21.9	9.4	35.9	71	39.4	26.8	11.3	22.5	165	38.2	18.1	5.5	38.2
Daphnia longiremis	54	37.0	18.5	13.0	31.5	105	35.2	21.0	8.6	35.2	71	38.0	25.4	14.1	22.5	232	37.5	21.1	11.2	30.2
Daphnia magna	52	42.3	25.0	7.7	25.0	73	37.0	26.0	4.1	32.9	71	39.4	26.8	11.3	22.5	183	40.4	21.3	7.7	30.6
Daphnia obtusa	52	35.9	24.5	9.4	30.2	40	32.5	17.5	10.0	40.0	NA	—	—	—	—	—	—	—	—	—
Daphnia pulex	53	35.9	22.6	9.4	32.1	51	27.5	13.7	13.7	45.1	71	38.0	26.8	12.7	22.5	165	37.6	20.6	8.5	33.3
Daphniopsis quadrangula	52	40.4	23.1	9.6	26.9	116	31.9	25.9	9.5	32.7	71	38.0	26.8	12.7	22.5	252	38.1	21.8	7.5	32.6
Daphniopsis truncata	52	40.4	23.1	9.6	26.9	134	31.3	23.1	11.2	34.4	71	38.0	26.8	12.7	22.5	215	33.0	20.9	12.6	33.5
Mean		38.6	22.6	10.2	28.6		32.4	20.4	11.0	35.7		38.5	26.3	12.3	22.9		37.3	20.5	8.6	33.6

[a] NA = sequence not available.

the hypothesis that the unique secondary structures in V4 and V7 evolved only once in a cylcestheriid lineage that subsequently gave rise to the cladocerans.

Traditionally, rRNA expansion sequences have been used for phylogenetic analysis of closely related taxa and ignored or culled for the construction of phylogenies above the genus level. The present study shows that derived helices in variable regions may provide valuable complex characters for the determination of deep phylogenies (e.g., among orders). The value of complex characters depends on the likelihood that they originate independently. It appears that both the birth and death of helices are rare events in crustacean SSU rRNA, making homoplasy unlikely. The same pattern is true of the kingdom-specific LSU rRNA structures identified by Michot and Bachellerie (1987) and Michot, Qu, and Bachellerie (1990). However, it is known that expansion segment helices can be entirely deleted in rapidly evolving regions of the LSU rRNA (Michot, Qu, and Bachellerie 1990; Nunn et al. 1996). Thus, more comparisons are necessary to determine the general utility of these helices as indicators of deeper-level phylogenetic relationships.

Origin and Elongation of Variable Region Helices in the Cladocera

The nucleotide compositions of the cladoceran variable region helices E23-c and 43-1 differ substantially at the proximal and distal ends (table 5). The proximal ends tend to be biased toward GC, while the distal ends show a marked increase in the proportion of U. Hancock (1995) suggested that sequences with a bias in nucleotide composition are more likely to contain simple sequence motifs that could act as substrates for replication slippage. As slippage occurs and the simple sequences proliferate, the original bias will tend to increase. This is precisely that pattern that has been observed in the variable regions of both SSU and LSU rRNAs across a phylogenetically diverse sample of organisms (Michot and Bachellerie 1987; Hancock and Dover 1988). The similarity between the nucleotide content of the distal end of 43 in the noncladocerans and the proximal end of 43-1 in the cladocerans is consistent with the idea that the original formation of 43-1 occurred via a mechanism that maintained, and even intensified, the original bias toward GC. The proximal ends of both E23-c and 43-1 are composed primarily of short runs of G's paired with short runs of C's (fig. 5), suggesting that slippage of single nucleotides may have been an important factor in the origin of these helices. That slippage of such short motifs has been an important force in genome evolution was suggested by Levinson and Gutman (1987), who showed that single-base repeats form longer runs in mammalian introns than would be expected by chance. The subsequent elongation of E23-c and 43-1 seems to have proceeded through the continued operation of slippage. However, as these helices become longer, the bias in nucleotide composition changes in favor of U (discussed below).

The elongation of E23-c and 43-1 via the insertion of stem modules that leave the downstream portions of

A HELIX E23-C

```
Holopedium gibberum
5'GGUGAGUGCCGGCUUGC  GGC U U
                           C
3'CU CACA  GCU AGCG  CCG  G
Holopedium amazonicum
5'GGGGCG G CGGCUCGC  CGG G C
                           A
3'   CGC C GCUGAGCG CGCC A

Eurycercus longirostris
5'GGAGUGGUGACGGUGGUGC UUCGGUGGGGC U U
                                    C
3'CC CACC    CCGC GUG GGGCUGCCUCUG  G
Eurycercus glacialis
5'GGAGUGGUGACGGUGGUGCCUUCGAUGCGG   U C
                                     C
3'CC CACC    CCGCUGUGGGUUCUGUGCC    G
```

B HELIX 43-1

```
Eurycercus longirostris
UGGGGGUCGGAG  UCGUCCGGGUCAAGGCGGUUGUGGGGGGCGGCCG GGUUCUAGUCUGUUUCGGUGUUCGG UUUUCGGGUUUGCGUU     CGUCUCCU UU
                                                                                                            C
GCUCUCGGUCUCAAAGCGGGCC    UCUGCUGGC UUCCUUGCUGGCUCCA    CGGGUGGGGCUGUAGGCU GGGGGCUCGAGUGCGG     GCGGGGGG G
                                                                       *                                *
Eurycercus glacialis
UGGGGGUCGGAG  UCGUCCGGGCCAAGGCGGUUGCUGGGGGCGGCCG GGUUCUAGUCCUCUUCGGUGUGCGG CAUGUGUCGGUUCUUUUUGGUUUGUGGCGG CAUCUC U UU
                                                                                                                    C
GCUCUCGGUCUCAAAGCGGGCC    UCUGUUAGC UUCUUUGUUGGCUCCA    CGGGUGGGGCUGUGCGCU G GCGUGGUCGGGGGGGGUCGAGCGGUGUC GUGGGG G G

Holopedium amazonicum
5'UGAGGGGAGUUCAG CUUGGGGAUCAGGA GUC GGUUGGGG GCGGGACGUCUGCGUCGUUUGUCCGC C G U C
                                                                          A
3'GCUCCUCUCU-GUCUGCUCUCC    CCU CGG CCGCUCCU CUCUCUGCUGGC CGGCGGCCGGGCG A A A
Holopedium gibberum
5'UGGGGGGCGGUCGA GUGCGGGAUCAGGUCGUCUGGU  GGG CGGUGGUGCCGGCCGUGGCUGUCGUC CGAGGUCC U U
                                                                                 C
3'GCCCCCC UCGGAG CAUGUCC    CCUGCGGGCCG  CCU GCCUCUGCGUCUGGCGGCGGCGGCGGAGG CUGGG  G
    *

5'UGGGGGGCGGUCGA GUGCGGGAUCAGGUCGUCUGGU  GGG yGGyGGyGyyGGyyGyGGyyGyyGyy yGAGGyyy U U
                                                                                 C
3'GCCCCCC UCGGAG CAUGUCC    CCUGCGGGCCG  CCU GyyyyyGyGyyyGGyGGyGGyGGyGGAGG yyGGG  G
```

C HELIX E23-C

```
Daphniopsis quadrangula        ∇
5'GGAGUGGUGGUCGGUGGGUCGCGUUG UAUC GGUUCGCGGACUUUCGGUUCUUUUUUUCCCAUCGGGUGUCUGUCUUU              GUUCCGUCAUCGCU U C
                                                                                 -------------              G
3'CC CACC     UCGCUUAGCGCAGU  U GACUU GCGCC  AAAGCUAG   AGGAGGUGUGGCUC-CGU C GAAU              CG GGC GUUGCGG  C
Daphniopsis truncata
5'GGAGUGGUGGUCGGUGGGUCGCGUUG UAUC GGUUCGCGGACUUUCGGUU UUUUUUCCUCAUCGAGUGU UG-UUCU UUUCUCUUC CUC GUUUUCUCUUCGCU U U
                                                                                                               G
3'CC CACC     UCGCUUAGCGCAGU  U GACUU GCGCC  AAGGCUAG   AGGAGGAGUGGCACACG  C GAGU    GGAAGCGGG CG UGA GUUGCGG  C

Daphnia magna (Ctenodaphnia)   ∇
5'GGAGUGGUGGCCGGUGGGUCGCGUUG UAUC GGUUCG C GUGGC          GUUUC GUUUUCUUCCUCUUCGGGG U U
                                                                                  C
3'CC CACC     UCGCUUGGCGCAGU  U GACUU GC G C UCG          CU AG UAGGGGG GGAGCUGCCUC  G
Daphnia exilis (Ctenodaphnia)
5'GGAGUGGUGGCCGGUGGGUCGCGUUG UAUC GGUUUGUC UUUU  UUUUGUUU  GUUUC CC U A U
                                                                    C
3'CC CACC     UCGCUUAGCGCAGU UUA  CUC AC G AAAGUCAAG  AAG  CGAGG GG C A A

Daphnia curvirostris (Hyalodaphnia)
5'GGAGUGGUGGUCGGUGAUUUGUGGUUGUAUC GGUU UU  UUUUCUCUC  GA    CGGCU U C
                                                                   G
3'CC CACC     UCGCUAAGC UCAGU U G CUAGAGA  AGGAGUGAG  U      UCGG  C
Daphnia galeata (Hyalodaphnia) ∇
5'GGAGUGGUGGUCGGCGGGUCGCGGUUGUAUC GGU UUC            GA   UUCGGUU                     GAGU GCG AGAG  UGAC U C U
                                          ---------             --------------------                            C
3'CC CACC     UCGCUUAGC UCAGU U G CUG UUG            CU   G GCUAG                     CUCUUC U CUUC  UCUG G  G
Daphnia longiremis (Hyalodaphnia)
5'GGAGUGGAUGUGGGUGAGUCGCGGUUGUAUC G UU UU            GAGAGUUCGUUC  AUGUGG GGUGUGUUUCGC GAGU GCGUGAAG  CUGG U U A
                                          ---------                                                             A
3'CACCUCC     UCGUUUAGC UCAGU U G CUCU  A            UUUUCU GCG    UGUGCUGCCGUGCUGCGCG CU  UGCGUUUCC GGCC G  C
                   *

Daphnia obtusa (Daphnia)       ∇
5'GGAGUGGUGGCCGGUGCGUCGUUGUUGUAUCUGCGGUCAGAGUCGGAGG        U U
                                                   -------   U
3'CC CACC     UCGCUUAGCU CAGU U GAUGCUGGUUUUAGUCUUC         C
Daphnia pulex (Daphnia)                        *    *
5'GGAGUGGUGGUCGGUGCGUCGUUGUUGUAUCUGCGGCCAGAGUUGGAAA    GUUC U U
                                                             U
3'CC CACC     UCGCUUAGCU CAGU U GAUGCUGGUUUUAAUUUUU UCUUGGG  C
```

FIG. 8.—Alignment of helices E23-c and 43-1 from pairs or triplets of species in the same genus or subgenus. Each sequence is continuous and is given in the 5′-to-3′ direction. Nucleotides that differ between species pairs (or triplets) are underlined. Sections of helix that are absent in one species relative to the other are indicated by a dashed line. Vertical lines connect nucleotides involved in canonical or G-U base pairing. Open arrow heads indicate the "boundary" between the proximal and distal regions of the helices in *Daphnia* and *Daphniopsis* (see text and table 5). Asterisks indicate positions at which compensatory mutations that maintain base pairing have occurred. *A,* E32-c in *Holopedium* and *Eurycercus. B,* 43-1 in *Holopedium* and *Eurycercus.* The sequence from *H. gibberum* is shown a second time with C and U replaced by Y (pyrimidine). *C,* E23-c in *Daphnia* and *Daphniopsis. D,* 43-1 in *Daphnia* and *Daphniopsis.* The sequence from *D. obtusa* is incomplete due to the occurrence of length polymorphism (see text). Sites at which this sequence differs from the *D. pulex* sequence are underlined.

FIG. 8 *(Continued)*

the helix intact (fig. 8) is consistent with the occurrence of compensatory slippage as proposed by Hancock and Dover (1990). The tendency for these modules to be associated with unpaired or mismatched nucleotides suggests that the original stimulus for their production is the elongation of bulges on one strand of a helix. Indeed, Gonzalez et al. (1985) showed that variants of D6 in the LSU rRNA of a single human differed by the presence/absence of asymmetrical bulges, apparently caused by slippage of GGC and GU motifs in the stems. Initially, the enlargement of a preexisting bulge may have little impact on the stability of a helix. However, increasing asymmetry of the bulge may destabilize the helix to the point where natural selection would act against individuals with a large proportion of rRNA genes that carry it. On the other hand, selection would also favor genes in which elongation of motifs that could pair with at least part of the bulge had occurred in the appropriate location. The creation of such a double mutant could occur if compensatory slippage occurred in a gene copy carrying the original slippage product, which seems unlikely. However, Hancock and Dover (1988) argued that the probability of this would increase as the proportion of the original mutant increased in the rRNA gene family. Alternatively, such compensatory slippage could have occurred in another copy of the rRNA gene, with subsequent generation of the double mutant via unequal crossing over (Hancock and Dover 1990). Either way, the result would be the production of a new section of helix embedded within the original. Nunn et al. (1996) suggested a similar mechanism to explain the fact that nucleotides with high sequence simplicity scores often occurred near bulges in the exceptionally long D3 region of LSU rRNA in isopod crustaceans.

In some cases, differences between E23-c and 43-1 in species from the same genus appear to be due to the addition of stem modules at the very tip of the helix. Such additions were also observed by Hassouna, Michot, and Bachellerie (1984) for the LSU rRNA D domains

of vertebrates, and by Nunn et al. (1996) for D3 of isopods. Extension of the tip of a helix could involve slippage-mediated enlargement of one side of the terminal loop that is compensated by slippage of a complementary motif on the other side.

The association of stem modules with bulges and terminal loops may help explain the increasing bias toward U with an increase in helix length in the cladocerans (fig. 7). The unpaired portions of 43-1 are biased toward A and U, while the stems are biased toward G and C (table 4). If slippage of single-nucleotide motifs in a bulge or a loop is the first step toward helix elongation, then it will most likely involve U or A. Because the entire structure is generally biased toward G and C, compensatory slippage would most likely involve one of these nucleotides, especially if it occurs in part of the stem. A run of G's that is created by slippage could form a stable helix with a run of U's to form a new section of stem. On the other hand, a run of C's would not form a stable helix with a run of A's (Santalucia, Kierzek, and Turner 1991; Wu, McDowell, and Turner 1995) resulting in selection against A-rich bulges that become large enough to destabilize the helix. A run of A's could also compensate a run of U's, and, indeed, there are some short stretches of A-U pairs in the helices of the *Daphnia* species (fig. 8). However, the overall proportion of A remains very low in the variable-region helices (table 2), suggesting that compensatory slippage involving A is not common or that there is strong selection against the accumulation of A-U pairs in the stems even though they should form more stable helices than G-U pairs (Sugimoto et al. 1986). This pattern of helix elongation would tend to perpetuate the U bias, as it is not likely that compensatory slippage would produce a run of G's (or A's) that is exactly the same length as the run of U's with which it pairs. Thus, the ''leftover'' U's in the bulge would maintain or create new sites for future elongation.

The tendency for the expansion of E23-c and 43-1 in the cladocerans to occur via the proliferation of G-U pairs is unusual. For example, the overall frequency of G-U pairs in the AU-biased stems of D1 and D2 in the LSU rRNA of 82 *Drosophila* species is only 6.5% (Rousset, Pelankadis, and Solignac 1991). However, the cladoceran pattern of GC bias in stems versus AU bias in loops and bulges has been observed in the variable regions of both SSU and LSU rRNAs of vertebrates (Hancock and Dover 1988; Vawter and Brown 1993) and in LSU rRNAs of insects and crustaceans (Hancock and Dover 1998; Nunn et al. 1996). For example, the mean GC content of the D3 stems in the 12 isopod species surveyed by Nunn et al. (1996) is 66.5%, while the unpaired regions are only 39.3% GC on average. However, a GC bias in stems generally intensifies as the variable regions expand, resulting in a predominance of G-C pairs. This is precisely the pattern observed for isopods (Nunn et al. 1996) in which the mean G-C/G-U ratio in D3 stems is 4.39 (range 2.00–6.63). This value is substantially higher than the mean G-C/G-U ratio of 1.36 observed for the cladocerans (table 4). A predominance of G-C pairs in the growing stems of variable-region helices is presumably favored (Vawter and Brown 1993) because G-C pairs have a lower free energy than either A-U or G-U pairs (Freier et al. 1986), resulting in increased helix stability. If so, it is interesting that the proportion of nucleotides involved in base pairing averages 85.9% in helix 43-1 of the cladocerans, which contains many G-U pairs (table 4), but only 64.7% (range 58.8–69.5%) in D3 of the isopods, which is rich in G-C pairs.

Like the stem modules of E23-c and 43-1, the two unique cladoceran side branches originate at unpaired nucleotides. E23-1a originates at an unpaired A in E23-1 of the noncladocerans (figs. 4*A* and 5*A*), and 43-2 originates from the U-C mismatch at which 43-1 originates in *Cyclestheria* (figs. 4*B* and 5*B*). This suggests that compensatory slippage may also occur within a growing bulge, leading to the formation of a new helix. However, the formation of new helices seems to be somewhat more constrained in the Cladocera than is the elongation of preexisting ones. For example, the formation of 43-2 seems to have been tolerated at the base of 43-1 (note that 43-2 is quite long in *Ofryoxus* and *Ilyocryptus* compared with the other cladocerans; fig. 5*B*), but energetic criteria suggest that there are no side branches along the length of either E32-c or 43-1. In addition, the length and primary sequence of E23-1a has remained fairly stable across the Cladocera, suggesting that it is under some sort of constraint. Similarly, the formation of side branches may be somewhat constrained in D3 of isopods. Secondary-structure analysis suggests that they occur in several species in subdomain III, but rarely occur in subdomain IV, despite the fact that both domains show extraordinary length increases (Nunn et al. 1996). In contrast, the proposed secondary structure for the extremely long helix 43 of the aphid *A. pisum* (rRNA database) includes three side branches. More data on structural variation in the variable regions of long rRNAs is required to determine if the pattern of length increase without concomitant increase in structural complexity is common. If so, this pattern suggests that the constraints affecting changes in the length of an established expansion helix are relatively relaxed compared with the constraints affecting its initial formation.

Nunn et al. (1996) noted a positive correlation between expansion of D3 in the LSU rRNA of isopods and the increased temperature tolerance that accompanied radiations into terrestrial habitats. No such correlation is apparent among the cladocerans, which generally live in lakes and/or ponds (there are a small number of marine taxa) and whose geographic distribution ranges from the Arctic down into the tropics. Indeed, the two taxa with the longest variable regions (table 2) experience opposite extremes of temperature: *D. quadrangula* inhabits very warm, ephemeral, saline bodies of water in the arid regions of Australia, while *D. longiremis* is restricted to cold-water habitats such as arctic lakes or the hypolimnion of deep temperate lakes.

Implications for General Genome Evolution

Hancock (1995) argued that elongation of variable regions in the rRNA genes is the byproduct of a ge-

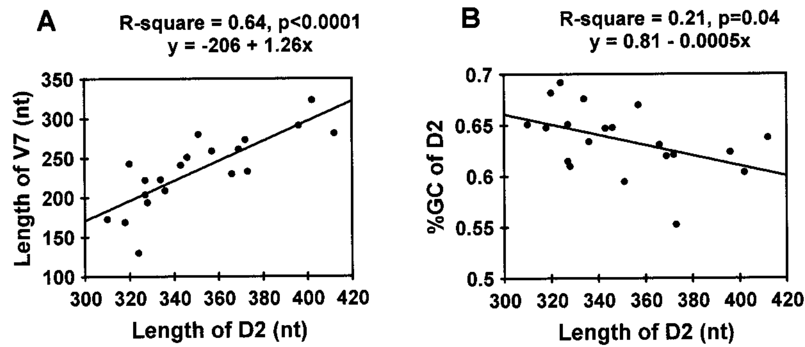

FIG. 9.—Regression analysis of (*A*) the length of D2 versus the length of V7 and (*B*) the length of D2 versus its nucleotide composition in cladoceran crustaceans.

nome-wide tendency for sequences to expand via mechanisms such as slippage replication and that this accounts for the observation that different expansion segments in the LSU rRNA gene tend to coevolve with respect to sequence and nucleotide composition (Michot and Bachellerie 1987; Hancock and Dover 1988). The highly significant correlation between length and nucleotide composition for E23-c and 43-1 in the cladocerans (fig. 3) is consistent with this idea. However, it could also be argued that there is some sort of functional relationship between V4 and V7 that requires a correlated increase in the size of the variable-region helices. If the coexpansion of these helices is a byproduct of elongation mechanisms, then one might expect a correlation between the lengths of expanded regions in the SSU rRNA and the LSU rRNA genes, which both occur within the same transcription unit. Taylor (unpublished data) sequenced the D2 region of the LSU rRNA gene for many of the taxa used in this study. The correlation between the length of D2 and V7 in the cladocerans is highly significant (fig. 9*A*). Moreover, the same tendency toward a decrease in GC content with increasing helix length, due to an increase in the proportion of U, also occurs (fig. 9*B*), suggesting that the mechanism responsible for elongation is similar in both regions. Even so, it is still not clear that the tendency toward elongation of variable regions in the rRNA genes reflects a genome-wide phenomenon. For example, *Daphnia* possess unusually long rRNA genes but have the smallest genomes yet reported for crustaceans (reviewed in Lécher, Defaye, and Noel 1995). Beaton (1995) found that the haploid genomes of 35 species of *Daphnia* range in size from 0.21 to 0.58 pg. Similarly, the genome size of the laevicaudatan *Lynceus brachyurum* is 0.30 pg (Beaton 1988). In contrast, the genome sizes of the anostracans *Artemiopsis stephanssoni* and *Branchinecta paludosa* are 0.87 and 2.77 pg, respectively (Beaton 1988). Genome sizes for nonbranchiopod crustaceans range from 0.7 pg in cirripedes to over 20 pg in some decapods (Lécher, Defaye, and Noel 1995), yet SSU rRNA genes longer than 1,900 nt have not been reported in any other crustacean group (Crease and Colbourne 1998). These preliminary results suggest that genome size may not be strongly correlated with the size of variable regions in rRNA, and that the unusual expansion of the SSU rRNA gene in the cladocerans may oppose the more general tendency for their genomes to be compact.

It is interesting to note that even though the mean length of D2 in the noncladocerans (330 nt, range 292–364 nt) is significantly shorter ($t = -1.84$, $P = 0.04$) than the mean length of D2 in the cladocerans (350 nt, range 310–412 nt), the ranges overlap substantially. This is definitely not the case for E23-c and 43-1, which do not even exist in the noncladocerans. Thus, an important question that still remains to be answered is why new helices evolve in the SSU rRNAs of some lineages but not in those of others, even though the mechanisms that could create them are operating in the D domains of the LSU rRNA genes within the same transcription unit.

Acknowledgments

Financial support for this project was provided by an NSERC Research Grant to T.J.C. We thank Paul Hebert and Chris Wilson for *Cyclestheria, Limnadia,* and *Daphniopsis; Chad Rowe for Holopedium*; Tim Collins for *Leptodora*; and John Colbourne for many stimulating discussions about rRNA secondary structure. Comments by two anonymous reviewers improved the manuscript.

LITERATURE CITED

BEATON, M. J. 1988. Genome size variation in the Cladocera. M.Sc. thesis, University of Windsor, Ontario, Canada.

———. 1995. Patterns of endopolyploidy and genome size variation in *Daphnia.* Ph.D. dissertation, University of Guelph, Ontario, Canada.

CAMPBELL, B. C., J. D. STEFFEN-CAMPBELL, and R. J. GILL. 1994. Evolutionary origin of whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) inferred from 18S rDNA sequences. Insect Mol. Biol. **3**:73–88.

CHALWATZIS, N., A. BAUR, E. STETZER, R. KINZELBACH, and F. K. ZIMMERMAN. 1995. Strongly expanded 18S rRNA genes correlated with a peculiar morphology in the insect order Strepsiptera. Zoology **98**:115–126.

CREASE, T. J., and J. K. COLBOURNE. 1998. The unusually long small-subunit ribosomal RNA of the Crustacean, *Daphnia pulex*: sequence and predicted secondary structure. J. Mol. Evol **46**:307–313.

DOYLE, J. J., and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**:11–15.

FREIER, S. M., R. KIERZEK, J. A. JAEGER, N. SUGIMOTO, M. H. CARUTHERS, T. NEILSON, and D. H. TURNER. 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.

FRYER, G. 1987. Morphology and the classification of the so-called Cladocera. Hydrobiologia **145**:19–28.

GERBI, S. A. 1985. Evolution of ribosomal DNA. Pp. 419–517 *in* R. J. MACINTYRE, ed. Molecular evolutionary genetics. Plenum Press, New York.

GONZALEZ, I. L., J. L. GORSKI, T. J. CAMPEN, D. J. DORNEY, J. M. ERICKSON, J. E. SYLVESTER, and R. D. SCHMICKEL. 1985. Variation among human 28S ribosomal RNA genes. Proc. Natl. Acad. Sci. USA **82**:7666–7670.

HANCOCK, J. M. 1995. The contribution of DNA slippage to eukaryotic nuclear 18S rRNA evolution. J. Mol. Evol. **40**: 629–639.

HANCOCK, J. M., and J. S. ARMSTRONG. 1994. SIMPLE34: an improved and enhanced implementation for VAX and Sun computers of the SIMPLE algorithm for analysis of clustered repetitive motifs in nucleotide sequences. Comput. Appl. Biosci. **10**:67–70.

HANCOCK, J. M., and G. A. DOVER. 1988. Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. Mol. Biol. Evol. **5**:377–391.

———. 1990. Compensatory slippage in the evolution of ribosomal RNA genes. Nucleic Acids Res. **18**:5949–5954.

HANCOCK, J. M., D. TAUTZ, and G. A. DOVER. 1988. Evolution of the secondary structures and compensatory mutations of ribosomal RNAs of *Drosophila melanogaster*. Mol. Biol. Evol. **5**:393–414.

HASSOUNA, N., B. MICHOT, and J.-P. BACHELLERIE. 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.

HINKLE, G., D. D. LEIPE, T. A. NERAD, and M. L. SOGIN. 1994. The unusually long small subunit ribosomal DNA of *Phreatamoeba balamuthi.* Nucleic Acids Res **22**:465–469.

KUZOFF, R. K., J. A. SWEERE, D. E. SOLTIS, P. S. SOLTIS, and E. A. ZIMMER. 1998. The phylogenetic potential of entire 26S rDNA sequences in plants. Mol. Biol. Evol. **15**:251–263.

KWON, O.-Y., K. OGINO, and H. ISHIKAWA. 1991. The longest 18S ribosomal RNA ever known. Eur. J. Biochem. **202**: 827–833.

LÉCHER, P., D. DEFAYE, and P. NOEL. 1995. Chromosomal and nuclear DNA of Crustacea. Invertebr. Reprod. Dev. **27**:85–114.

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

MATZURA, O., and A. WENNBORG. 1996. RNAdraw: an integrated program for RNA secondary structure calculation and analysis under 32-bit Microsoft Windows. CABIOS **12**: 247–249.

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

MICHOT, B., L.-H. QU, and J.-P. BACHELLERIE. 1990. Evolution of large-subunit rRNA structure: the diversification of divergent D3 domain among major phylogenetic groups. Eur. J. Biochem. **188**:219–229.

NUNN, G. B., B. F. THEISEN, B. CHRISTENSEN, and P. ARCTANDER. 1996. Simplicity-correlated size growth of the nuclear 28S ribosomal RNA D3 expansion segment in the Crustacean order Isopoda. J. Mol. Evol. **42**:211–223.

OLESON, J., J. W. MARTIN, and E. W. ROESSLER. 1996. External morphology of the male of *Cyclestheria hislopi* (Baird, 1859)(Crustacea, Branchiopoda, Spinicaudata), with a comparison of male claspers among the Conchostraca and Cladocera and its bearing on phylogeny of the 'bivalved' Branchiopoda. Zool. Scr. **25:**291–316.

PIÇON, M., R. R. GUTELL, R. EHRLICH, and A. ZAHA. 1996. Characterization of a flatworm ribosomal RNA-encoding gene: promoter sequence and small subunit rRNA secondary structure. Gene **171**:215–220.

ROESSLER, E. W. 1995. Review of Colombian Conchostraca (Crustacea)—ecological aspects and life cycles—family Cyclestheriidae. Hydrobiologia **298**:113–124.

ROUSSET, F., M. PELANKADIS, and M. SOLIGNAC. 1991. Evolution of compensatory substitutions through G-U intermediate state in *Drosophila* rRNA. Proc. Natl. Acad. Sci. USA **88**:10032–10036.

SANTALUCIA, J., R. KIERZEK, and D. H. TURNER. 1991. Stabilities of consecutive A-C, C-C, C-G, U-C, and U-U mismatches in RNA internal loops: evidence for stable hydrogen-bonded U-U and C-C+ pairs. Biochemistry **30**:8242–8251.

STAROBOGATOV, Y. I. 1986. Systema rakoobrazayh. Zool. J. **65**:1769–1781.

SUGIMOTO, N., R. KIERZEK, S. M. FREIER, and D. H. TURNER. 1986. Energetics of internal G-U mismatches in ribooligonucleotide helices. Biochemistry **25**:5755–5759.

VAN DE PEER, Y., J. JANSEN, P. DERIJK, and R. DE WACHTER. 1997. Database on the structure of small ribosomal subunit RNA. Nucleic Acids Res. **25**:111–116.

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

VOGLER, A. P., A. SELSH, and J. M. HANCOCK. 1997. Phylogenetic analysis of slippage-like sequence variation in the V4 rRNA expansion segment in tiger beetles (Cicindelidae). Mol. Biol. Evol. **14**:6–19.

WALOSSEK, D. 1993. The Upper Cambrian *Rehbachiella* and the phylogeny of Branchiopoda and Crustacea. Fossils Strata **32**:1–202.

WILCOX, L. W., L. A. LOUISE, P. A. FUERST, and G. L. FLOYD. 1992. Group I introns within the nuclear-encoded small subunit rRNA gene of three green algae. Mol. Biol. Evol. **9**: 1103–1118.

WINNEPENNINCKX, B., Y. VAN DE PEER, T. BACKELJAU, and R. DE WACHTER. 1995. CARD: a drawing tool for RNA secondary structure models. Biotechniques **18**:1060–1063.

WU, M., J. A. MCDOWELL, and D. H. TURNER. 1995. A periodic table of symmetric tandem mismatches in RNA. Biochemistry **34**:3204–3211.

DAVID M. RAND, reviewing editor

Accepted July 21, 1998