Complete Sequence and Gene Organization of the Mitochondrial Genome of the Land Snail *Albinaria coerulea*

Evi Hatzoglou, George C. Rodakis and Rena Lecanidou

Department of Biochemistry, Cell and Molecular Biology, and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece

Manuscript received January 31, 1995
Accepted for publication May 15, 1995

ABSTRACT

The complete sequence (14,130 bp) of the mitochondrial DNA (mtDNA) of the land snail *Albinaria coerulea* was determined. It contains 13 protein, two rRNA and 22 tRNA genes. Twenty-four of these genes are encoded by one and 13 genes by the other strand. The gene arrangement shares almost no similarities with that of two other molluscs for which the complete gene content and arrangement are known, the bivalve *Mytilus edulis* and the chiton *Katharina tunicata;* the protein and rRNA gene order is similar to that of another terrestrial gastropod, *Cepaea nemoralis.* Unusual features include the following: (1) the absence of lengthy noncoding regions (there are only 141 intergenic nucleotides interspersed at different gene borders, the longest intergenic sequence being 42 nucleotides), (2) the presence of several overlapping genes (mostly tRNAs), (3) the presence of tRNA-like structures and other stem and loop structures within genes. An RNA editing system acting on tRNAs must necessarily be invoked for posttranscriptional extension of the overlapping tRNAs. Due to these features, and also because of the small size of its genes (*e.g.,* it contains the smallest rRNA genes among the known coelomates), it is one of the most compact mitochondrial genomes known to date.

IN the last few years, there has been an accelerated accumulation of sequence data on animal mitochondrial genomes. Most of the complete mitochondrial DNA (mtDNA) sequences that have been published concern deuterostomes [vertebrates and echinoderms: human (ANDERSON *et al.* 1981), mouse (BIBB *et al.* 1981), rat (GADALETA *et al.* 1989), cow (ANDERSON *et al.* 1982), fin whale (ARNASON *et al.* 1991), blue whale (ARNASON and GULLBERG 1993), harbor seal, *Phoca vitulina* (ARNASON and JOHNSSON 1992), American opossum, *Didelphis virginiana* (JANKE *et al.* 1994), chicken (DESJARDINS and MORAIS 1990), Xenopus (ROE *et al.* 1985), carp (CHANG *et al.* 1994), sea urchins *Paracentrotus lividus* (CANTATORE *et al.* 1989) and *Strongylocentrotus purpuratus* (JACOBS *et al.* 1988), sea star, *Asterina pectinifera* (ASAKAWA *et al.* 1991)], while the available sequences for protostome coelomates are limited to four arthropods [*Drosophila yakuba* (CLARY and WOLSTENHOLME 1985a), bee (CROZIER and CROZIER 1993), mosquito, *Anopheles quadrimaculatus* (MITCHELL *et al.* 1993), Artemia (PEREZ *et al.* 1994, EMBL-data bank accession number X69067)] and one mollusc [the chiton *Katharina tunicata* (BOORE and BROWN 1994b)]. Within the phylum of molluscs, considerable information is also available for *Mytilus edulis,* whose partial sequence but complete mtDNA gene content and organization is known (HOFFMAN *et al.* 1992), as well as *Cepaea nemoralis,* whose partial gene order (excluding tRNA genes) has been published (TERRETT *et al.* 1994).

The mtDNA gene content of coelomate metazoans is constant: it consists of 13 protein genes, two genes for the small and large ribosomal RNA subunits, and 22 tRNA genes, some of which are transcribed from one and some from the other mtDNA strand. An exception is observed in molluscs. Mytilus is missing a protein gene (*ATPase8*), while it contains an extra tRNA gene and has all of its genes transcribed from one strand (HOFFMANN *et al.* 1992); Katharina, on the other hand, contains the standard set of 37 mitochondrial genes, as well as two extra tRNA genes, that may or may not be functional (BOORE and BROWN 1994b).

Besides coding regions, there are also noncoding sequences in animal mtDNA and, more specifically, a major noncoding segment, which in deuterostomes (D-loop) contains a combination of sequence elements that are related with control of both replication and transcription (JACOBS *et al.* 1988; CLAYTON 1991, 1992; WOLSTENHOLME 1992; SHADEL and CLAYTON 1993). The length of this region is extremely variable [121 nucleotides (nt), sea urchin (JACOBS *et al.* 1988); over 20 kb, pine weevil (BOYCE *et al.* 1989)]. The length variation is usually due to the presence of short (*e.g.,* 10 bp) (GHIVIZZANI *et al.* 1993) or longer [*e.g.,* 260 bp (BROUGHTON and DOWLING 1994) or 1.2 kb (GJETVAJ *et al.* 1992)] repeated sequences. Size variation in noncoding sequences inevitably results in total length variation, which can be observed even at the level of individ-

Corresponding author: Rena Lecanidou, Department of Biochemistry, Cell and Molecular Biology, and Genetics, University of Athens, Panepistimiopolis, Athens 157 01, Greece.
E-mail: rlecanid@atlas.uoa.ariadne-t.gr

uals (reviewed by MORITZ *et al.* 1987). Extremes in size variation have been documented in molluscs: seven species of scallops have sizes ranging from 16.2 to 42 kb (SNYDER *et al.* 1987; LAROCHE *et al.* 1990; GJETVAJ *et al.* 1992).

The mtDNA gene organization is considered to be relatively constant within each metazoan phylum, where the observed variations mainly concern tRNA gene rearrangements (WOLSTENHOLME 1992) and nucleotide substitutions (CROZIER *et al.* 1989; PALUMBI and BENZIE 1991; AVISE *et al.* 1992; MARTIN *et al.* 1992). However, molluscs are the largest exception to this rule, since large variations in gene organization (*e.g.*, between Mytilus and Katharina) (BOORE and BROWN 1994b) and sequence [*Albinaria turrita vs.* Mytilus (LECANIDOU *et al.* 1994) and Katharina *vs.* Mytilus (BOORE and BROWN 1994b)] have been found.

We have recently pointed out the first indications for this great diversity of molluscan mtDNA, as Albinaria sequences are more similar to the corresponding sequences of Drosophila than of Mytilus, and have suggested that this may be attributed to either a polyphyletic origin or to a high and differentiated evolutionary rate of molluscan mtDNA (DOURIS *et al.* 1995; LECANIDOU *et al.* 1994). BOORE and BROWN (1994b), on the basis of the observation that structural features of Katharina resemble more those of Drosophila than of Mytilus, have proposed a fast mtDNA molecular clock for Mytilus.

In this paper, we present the complete mtDNA sequence of the land snail *A. coerulea.* It is the smallest mtDNA among coelomate metazoans (14,130 bp), does not contain any noncoding sequence longer than 42 nt, and its gene organization seems to have almost nothing in common with that of other known metazoans, with the exception of the major genes of another terrestrial gastropod, *C. nemoralis.*

MATERIALS AND METHODS

The mtDNA of *A. coerulea* was cloned in three consecutive *Hin*dIII segments: 6.55 kb (13A3, bases 1561–8123), 5.4 kb (13A12, bases 8124–13278), and 2.4 kb (13F6, bases 13279–1560) (Figure 2) (DOURIS *et al.* 1995). The following evidence indicates that the entire mitochondrial genome is represented in these three *Hin*dIII clones: (1) sequencing of an *Eco*RI clone (15A6, bases 282–2950) (Figure 2) showed that it overlaps clones 13F6 and 13A3, which therefore must be adjacent; (2) clones 13A12 and 13F6 must also be adjacent, because a sequenced clone from a closely related Albinaria species (A28 from *A. turrita*) (LECANIDOU *et al.* 1994) overlaps them, and (3) the *Hin*dIII site between clones 13A3 and 13A12 (base 8124) (Figure 2) lies in a conserved region of the *ND4* gene (bases 8113 to 8137 correspond to the amino acid sequence FLVKLPIY, which is identical in the chiton *K. tunicata*) (BOORE and BROWN 1994b).

Restriction fragments of the three *Hin*dIII clones, as well as of the overlapping *Eco*RI clone, were subcloned into plasmid vectors pUC8, 9, 18, 19, or pBluescript II KS (Stratagene). Further subcloning was performed after deletions of clones and subclones using the enzyme Exonuclease III (HENIKOFF 1987). Clones and derived subclones were amplified after transformation with $CaCl_2$/RbCl, using as hosts *Escherichia coli* JM83. DNA was extracted by mini plasmid preparations with alkaline lysis (SAMBROOK *et al.* 1989).

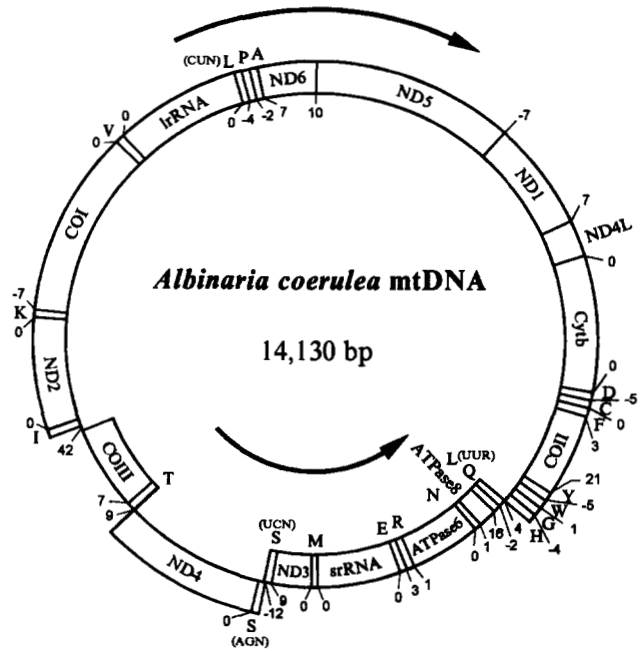

FIGURE 1.—Gene map of the *A. coerulea* mtDNA molecule. The outer circle represents the sense strand for genes transcribed clockwise and the inner circle for genes that are transcribed counterclockwise. tRNA genes are depicted by the one-letter amino acid code. Positive numbers at gene boundaries indicate noncoding nucleotides; negative numbers indicate overlapping nucleotides of adjacent genes. Abbreviations used: *ATPase6-8,* ATP synthase subunits 6 and 8; *COI-III,* cytochrome *c* oxidase subunits I, II, and III; *Cytb,* cytochrome *b* apoenzyme; *ND1-6* and *ND4L,* NADH dehydrogenase subunits 1-6 and 4L; *s-rRNA* and *l-rRNA,* small and large subunits of ribosomal RNA.

Both mtDNA strands were completely sequenced. Sequence determination was performed from the ends of clones and subclones by the dideoxynucleotide chain termination method (SANGER *et al.* 1977) using Sequenase V 2.0 (US Biochemical) and universal M13/pUC primers. Samples were resolved in 4 or 5% polyacrylamide, 7.5 M urea gels, and autoradiography was performed using Kodak X-Omat film. Sequences were identified by comparison with data from the EMBL data bank or from cited publications. The tRNA genes were identified by their potential to form the characteristic for mt-tRNA stem and loop structures. Sequence analysis was performed using computer programs developed by PUSTELL and KAFATOS (1984, 1986). Other computer programs used are cited in the text.

Northern hybridization was used to identify the location of the $tRNA^{Lys}$ gene. RNA was extracted from the feet (muscle) of 12 snails, treated with Proteinase K, extracted with phenol/chloroform and precipitated with ethanol according to SAMBROOK *et al.* (1989). Samples were resolved by electrophoresis in a 20 × 20 × 0.15 cm 4% polyacrylamide, 6 M urea gel and electroblotted in 0.5× TBE to a Gene Screen Plus (DuPond, NEF-976) hybridization transfer membrane (SRIVASTAVA *et al.* 1993). The membrane was baked at 80° for 2 hr. mtDNA fragments used as probes

were labelled by nick translation (SAMBROOK *et al.* 1989). Hybridization was performed at 65° in 0.4 M NaCl, 1% sodium dodecyl sulfate, 50 mM Tris-HCl pH 7.5. Autoradiography was performed using an intensifying screen (DuPond Cronex Lighting-Plus) and Kodak X-Omat film. Exposure was for 1–4 days.

The sequence of *A. coerulea* mtDNA has been deposited in the EMBL nucleotide sequence data bank under accesion number X83390.

RESULTS AND DISCUSSION

General features: *A. coerulea* mtDNA has the typical features of metazoan mtDNA but several of its characteristics are novel. Its length (14,130 bp) is the smallest among known coelomates with the exception of that belonging to yet another land gastropod, *C. nemoralis*, that is reported to have approximately the same size as Albinaria (TERRETT *et al.* 1994). The pseudocoelomate nematodes *A. suum* and *C. elegans* (OKIMOTO *et al.* 1992) have equally small mtDNA genomes. The small size of Albinaria mtDNA is the consequence of its compact gene organization, the small size of its genes, as well as the absence of a lengthy noncoding region (Figure 1). It contains all 37 genes typical of metazoan mtDNA: 13 protein genes (*ATPase6, ATPase8, COI, COII, COIII, Cytb, ND1, ND2, ND3, ND4, ND4L, ND5, ND6*), two ribosomal RNA genes (*l-rRNA, s-rRNA*) and 22 tRNA genes. The sequence shown in Figure 2 constitutes the sense strand for 24 genes (*major* strand, coding for more genes), while 13 genes are transcribed in the opposite direction (*minor* strand, coding for fewer genes). The base composition of the major strand is T, 37.9%; C, 13.8%; A, 32.8%; G, 15.5%, and the G + T content is 53.4%. However, if we take into account the nucleotide frequencies at fourfold synonymous sites (considered by definition as neutral positions), we find that in these positions the major strand has almost the same G + T percentage as the minor strand (50.3 and 50.2%, respectively).

The A + T content (70.7%) of Albinaria is higher than that of deuterostomes [in echinoderms the A + T content ranges from 58.9%, *S. purpuratus*, (JACOBS *et al.* 1988) to 61.3%, *A. pectinifera* (ASAKAWA *et al.* 1991); in vertebrates from 55.6%, human (ANDERSON *et al.* 1981) to 63.2%, mouse (BIBB *et al.* 1981), as well as of other known molluscs [69.0%, *K. tunicata* (BOORE and BROWN 1994b); 62%, sequenced portions of *M. edulis* (HOFFMANN *et al.* 1992)], but lower than that of insects (78.6%, *D. yakuba* (CLARY and WOLSTENHOLME 1985a); 84.9%, *Apis mellifera* (CROZIER and CROZIER 1993); 77.4%, *A. quadrimaculatus* (MITCHELL *et al.* 1993)] and of pseudocoelomate nematodes (76.2%, *C. elegans;* 72.0%, *A. suum*) (OKIMOTO *et al.* 1992). The A + T content at fourfold synonymous sites (77.5% overall, 77.9% major strand, 76.3% minor strand) is indicative of a clear bias toward A + T.

As far as the dinucleotide composition of the mtDNA is concerned, Albinaria exhibits the same deficiencies as all known metazoan mtDNAs. The double-stranded dinucleotides CC·GG and GC·GC show a high ratio of observed/expected frequency. $\rho^* = 1.44$ and $\rho^* = 1.30$, respectively; ρ^* is the symmetrized dinucleotide odds ratio calculated according to BURGE *et al.* (1992). In contrast, the double-stranded dinucleotide CG·CG is the only under-represented ($\rho^* = 0.70$). At present, there is no widely accepted explanation for the CpG suppression in animal mtDNAs (CARDON *et al.* 1994).

Gene arrangement: Close inspection of the Albinaria gene arrangement map (Figure 1) reveals certain interesting features. The Albinaria gene organization is novel, bearing almost no similarities to any published complete gene arrangements, including those of two other molluscs, Mytilus and Katharina (HOFFMAN *et al.* 1992; BOORE and BROWN 1994a,b). However, the partial gene organization of another terrestrial gastropod, *C. nemoralis* (TERRETT *et al.* 1994), reveals many similarities in the protein and rRNA gene order with Albinaria: only one gene rearrangement is required to interconvert the Albinaria and Cepaea protein and rRNA gene order, namely a transposition of the *ND4* or *COIII* gene. As the Cepaea complete sequence and gene organization has not been published, a comparison of tRNA genes of these two land gastropods cannot be made at present, although it can be inferred from looking at Figure 6 of TERRETT *et al.* (1994) that there are at least two positions in the genome map that differ: a region containing either tRNA genes or noncoding sequences present between genes coding for *ND6* and *ND5* in Cepaea is absent from Albinaria and no tRNA is present between *s-rRNA* and *ATPase6* in Cepaea, whereas these two genes in Albinaria are separated by two tRNA genes.

No gene boundaries are shared between Albinaria and Mytilus and only very few with Katharina and the arthropods. More specifically, $tRNA^{Val}$, *l-rRNA* and $tRNA^{Leu}$*(CUN)* are directly adjacent in Albinaria, Katharina and the arthropods, but in Albinaria these genes are transcribed from the major strand; furthermore, Albinaria and Katharina also share the gene boundaries between $tRNA^{Met}$ and *s-rRNA*, which are transcribed from the minor strand in both molluscs. The only other shared gene boundaries between Albinaria and other organisms are those between genes *COIII* and $tRNA^{Thr}$, which is also observed in the nematodes (OKIMOTO *et al.* 1992), and between the *s-rRNA* and $tRNA^{Glu}$, which is held in common with echinoderms (see SMITH *et al.* 1993).

It has been suggested from gene order comparisons that rearrangements involving tRNAs occur more frequently than rearrangements involving other genes (WOLSTENHOLME 1992). Thus, if the differences in relative locations of tRNAs are ignored, we can discern very limited similarities in protein gene borders between Albinaria and Mytilus involving the *Cytb* and *COII* genes, and the *ND4, COIII* and *ND2* genes, although *COIII* is transcribed from the minor strand in Albinaria (while all Mytilus genes are transcribed from one strand). Perhaps more interesting is the proximity of *ND2* and *COI* genes in many organisms including the molluscs Albinaria, Katharina and Cepaea (but not Mytilus), the arthro-

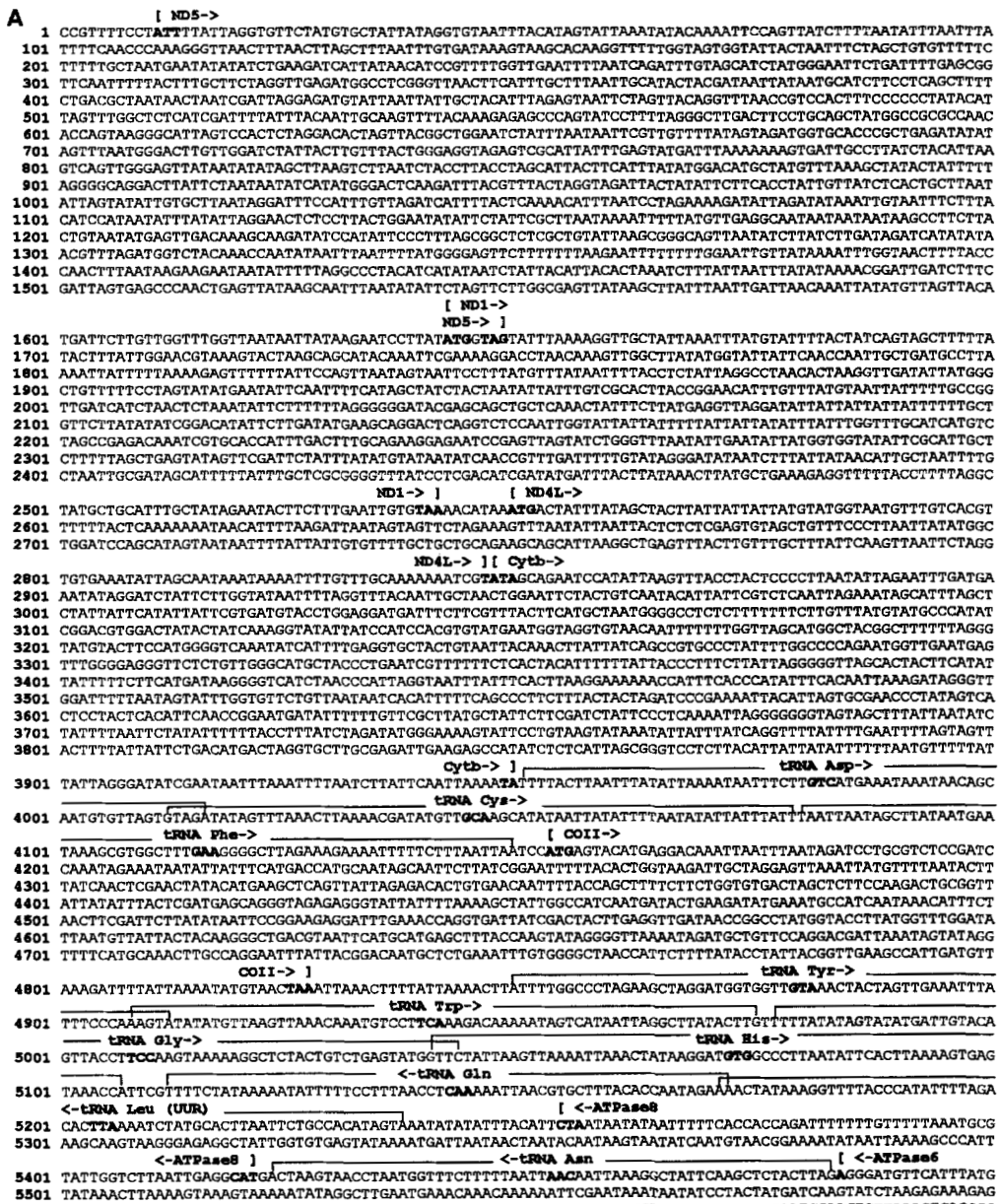

FIGURE 2.—The complete nucleotide sequence of the circular *A. coerulea* mtDNA. Only one strand is shown, which corresponds to the sense strand for the majority of genes. Abbreviations used as in Figure 1; numbering starts at the first nucleotide after the *ND6* termination codon. The putative first and last nucleotide for all genes are indicated; putative initiation and termination codons of translation, as well as the anticodon sequences of the tRNA genes, are in bold face characters. Arrows indicate the direction of transcription. The translated amino acid sequences of protein genes are available from the authors upon request.

pods (CLARY and WOLSTENHOLME 1985a; CROZIER and CROZIER 1993; PEREZ *et al.* 1994) (sequence of Artemia mtDNA submitted to EMBL with accession number X69067) and the vertebrates (see WOLSTENHOLME 1992). In all the above, these two protein genes are either directly adjacent (Katharina) or are separated by different tRNA genes. Interestingly, in sea urchins *ND2* and *COI* are separated by the *1-rRNA* gene (JACOBS *et al.* 1988; CANTATORE *et al.* 1989; DE GIORGI *et al.* 1991; SMITH *et al.* 1993).

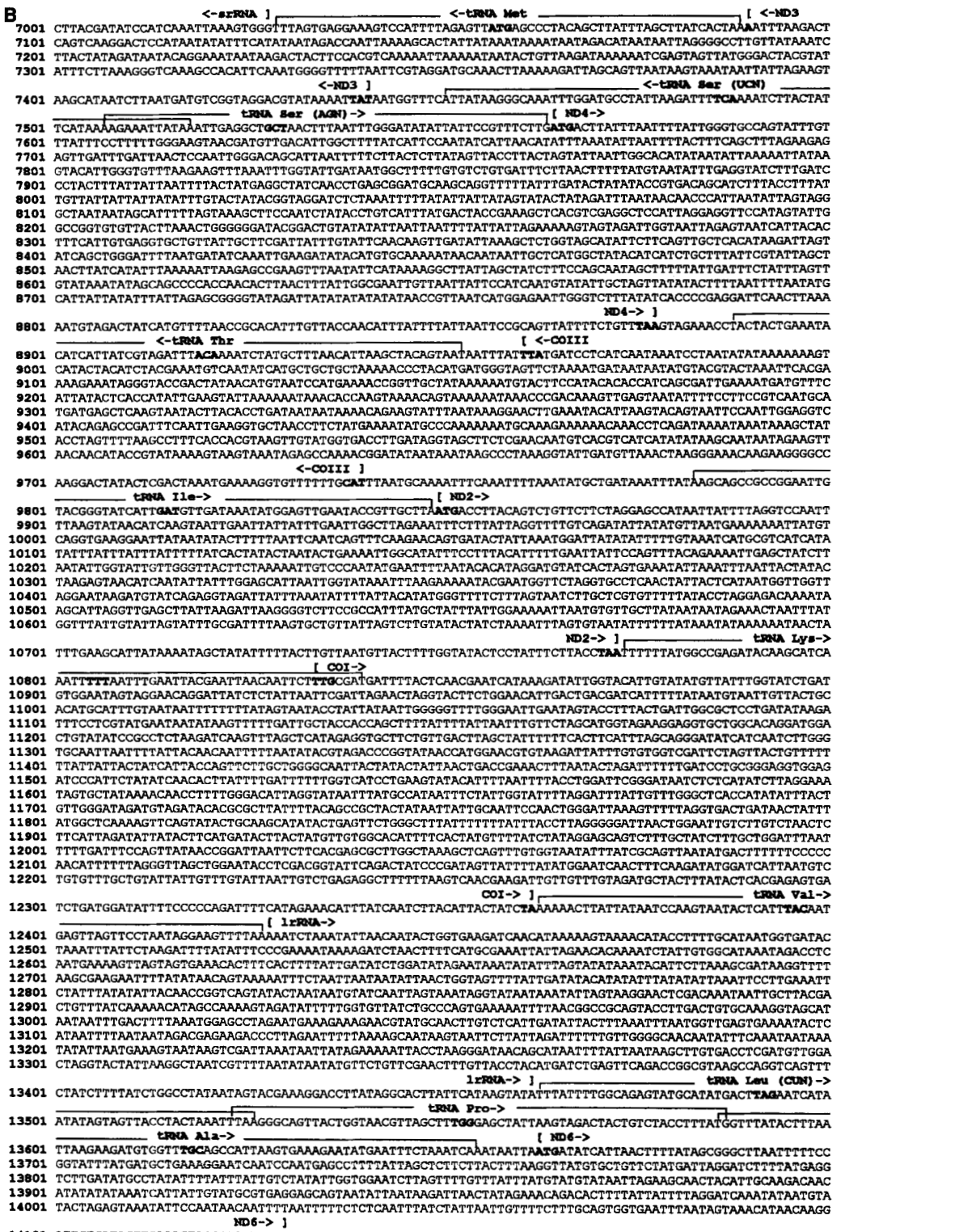

FIGURE 2.—*Continued*

As can be seen in Figures 1 and 2, all Albinaria genes are either directly adjacent or have very few nucleotides separating them. The total number of nucleotides between genes is 141, the longest intergenic sequence of 42 nucleotides separating two genes transcribed in opposite directions (*COIII* and *tRNA*^{Ile}). Five genes end on abbreviated stop codons. Furthermore, there are many cases of overlapping genes coded by the same strand. Three types of such overlap may be identified, involving protein-protein, tRNA-protein, or tRNA-tRNA junctions. Thus, there is one case of overlapping protein genes (*ND5-ND1,* 7-nt overlap), one case of overlap between a tRNA and a protein gene (*tRNA*^{Lys}*-COI,* 7-nt overlap) and six cases of overlapping tRNA genes (*Asp-*

Cys, 5 nt; *Gly-His*, 4 nt; *Leu(CUN)-Pro*, 4 nt; *Leu(UUR)-Gln*, 2 nt; *Pro-Ala*, 2 nt; *Tyr-Trp*, 5 nt). Finally, an extensive overlap of 12 nt is observed between two tRNA genes transcribed in opposite directions [$tRNA^{Ser}(UCN)$ and $tRNA^{Ser}(AGN)$].

Another unusual feature in Albinaria mtDNA is the presence of four consecutive protein-protein gene junctions (*ND6-ND5*, *ND5-ND1*, *ND1-ND4L* and *ND4L-Cytb*) that are not separated by an intervening tRNA. It has been suggested that the secondary structure of a tRNA gene between pairs of protein genes is needed to act as a signal for the precise cleavage of the polycistronic primary transcript (OJALA *et al.* 1980, 1981). In accordance with this hypothesis, almost all reported cases of protein-protein gene borders with no intervening tRNAs have the potential to form hairpin structures (BIBB *et al.* 1981; CLARY and WOLSTENHOLME 1985a; OKIMOTO *et al.* 1992; BOORE and BROWN 1994b). In Albinaria it is also possible to draw stem and loop structures near or at the protein-protein gene boundaries (Figure 3). In the case of *ND6-ND5* and *ND1-ND4L* borders, the stem and loop structures are positioned more or less similarly with those reported for the nematodes and Katharina with respect to gene termini; the complete termination codons are positioned within the loops. In the case of *ND5-ND1* and *ND4L-Cytb* borders, however, a tRNA-like structure precedes a stable stem and loop structure that is located near the 3′ end of *ND5* and *ND4L* genes, respectively (in *ND4L*, there are two hairpins separated by 29 nt). It should be noted that *ND5* and *ND1* are overlapping genes and that *ND4L* ends with an incomplete termination codon. These tRNA-like structures resemble normal tRNAs with AAG [alternative Leu(CUN)] and TTT (Lys) anticodons respectively. The existence of superfluous tRNA-like genes has been also reported for two other molluscs, Mytilus and Katharina. What is unusual with Albinaria, however, is that, in contrast to other molluscs, the tRNA-like structures are located totally within protein coding genes. Interestingly, an extension to the 5′ end of the *ND5* gene in sea urchins can be folded into a secondary structure resembling a tRNA gene and is thought to be a remnant of a $tRNA^{Leu}(CUN)$ gene (CANTATORE *et al.* 1987, 1989; DE GIORGI *et al.* 1991). The significance of the Albinaria tRNA-like structures is not clear at present. However, we have found that they do not hybridize with low molecular weight RNA in Northern blots, while the two corresponding standard tRNAs do (data not shown).

Control regions: An important question emerging from the gene organization of Albinaria mtDNA concerns the location of regions containing the signals for replication and transcription. In all metazoans, where such sequences have been identified, they are found in noncoding regions (WOLSTENHOLME 1992). What is unique in Albinaria mtDNA, however, is that it contains virtually no unassigned sequences of significant length (the largest unassigned sequence is 42 nt and the second and third largest are 21 and 16 nt, respectively).

Interestingly, the decanucleotide AATATATATT, located between $tRNA^{Leu}(UUR)$ and *ATPase8* within the 16-nt Albinaria unassigned sequence, is reminiscent of sequences present in sea urchins (TTATATATAA) and chicken and duck (ATATATAT). In Albinaria, this palindrome is even longer (tetradecanucleotide, TAAATATATATTTA), if we allow for a 2-nt overlap with the $tRNA^{Leu}(UUR)$ gene. Such palindromes have been implicated to function as bidirectional promoters or as recognition signals for enzymes involved in transcription or processing (JACOBS *et al.* 1988, 1989; CANTATORE *et al.* 1989; L'ABBÉ *et al.* 1991; RAMIREZ *et al.* 1993).

The longest Albinaria noncoding sequence (42 nt) contains yet another decanucleotide perfect palindrome (ATAAATTTAT), which is twice as long (TGCTGATAAATTTATAAGCA) if we overlook a 5-nt overlap with the adjacent $tRNA^{Ile}$ gene.

Synthesis of the second (L) strand in a variety of metazoans (including several mammals, Xenopus, Drosophila and the nematodes) is supposedly initiated within a run of Ts situated in the loops of potential hairpins that can be formed from intergenic sequence (CHANG *et al.* 1985; ROE *et al.* 1985; CLARY and WOLSTENHOLME 1987; OKIMOTO *et al.* 1992). No such structures can be found in the short Albinaria noncoding sequences. However, a sequence with the potential of forming a stable stem and loop structure is found within the *ND5* gene (Figure 3); this secondary structure has a T-rich loop and moreover its 5′ end is identical to that of the O_L loop in Drosophila (boxed sequence ATATAA in Figure 3) (CLARY and WOLSTENHOLME 1987). The significance of this structure is presently unknown.

Protein genes: Identification of protein genes was accomplished by comparison at the nucleotide/protein level (see Table 1 for genetic code) with already known mtDNA genes of a close relative to *A. coerulea* (*A. turrita*) (LECANIDOU *et al.* 1994) as well as with sequences of genes available in the EMBL data bank. Identification of *ND4L* was based on hydropathy profiles, predicted using the computer program SOAP of the PCGENE package (BAIROCH 1988) (data not shown), that were found to be very similar to corresponding profiles of Katharina, Mytilus and Drosophila.

Nine of the 13 protein genes of Albinaria mtDNA start with the orthodox translation initiation codon ATG:two with ATA (*Cytb*, *ND3*), one with ATT (*ND5*) and one with TTG (*COI*). ATT is not used as initiation codon in Katharina and Mytilus but is quite frequent in Drosophila (CLARY and WOLSTENHOLME 1985a) and the nematodes (OKIMOTO *et al.* 1990, 1992). TTG is not used for initiation in any of the known coelomates but is common in pseudocoelomate nematodes (OKIMOTO *et al.* 1990, 1992). Eight Albinaria protein genes terminate with the stop codons TAA or TAG (Figure 2). The other five are inferred to have incomplete translation

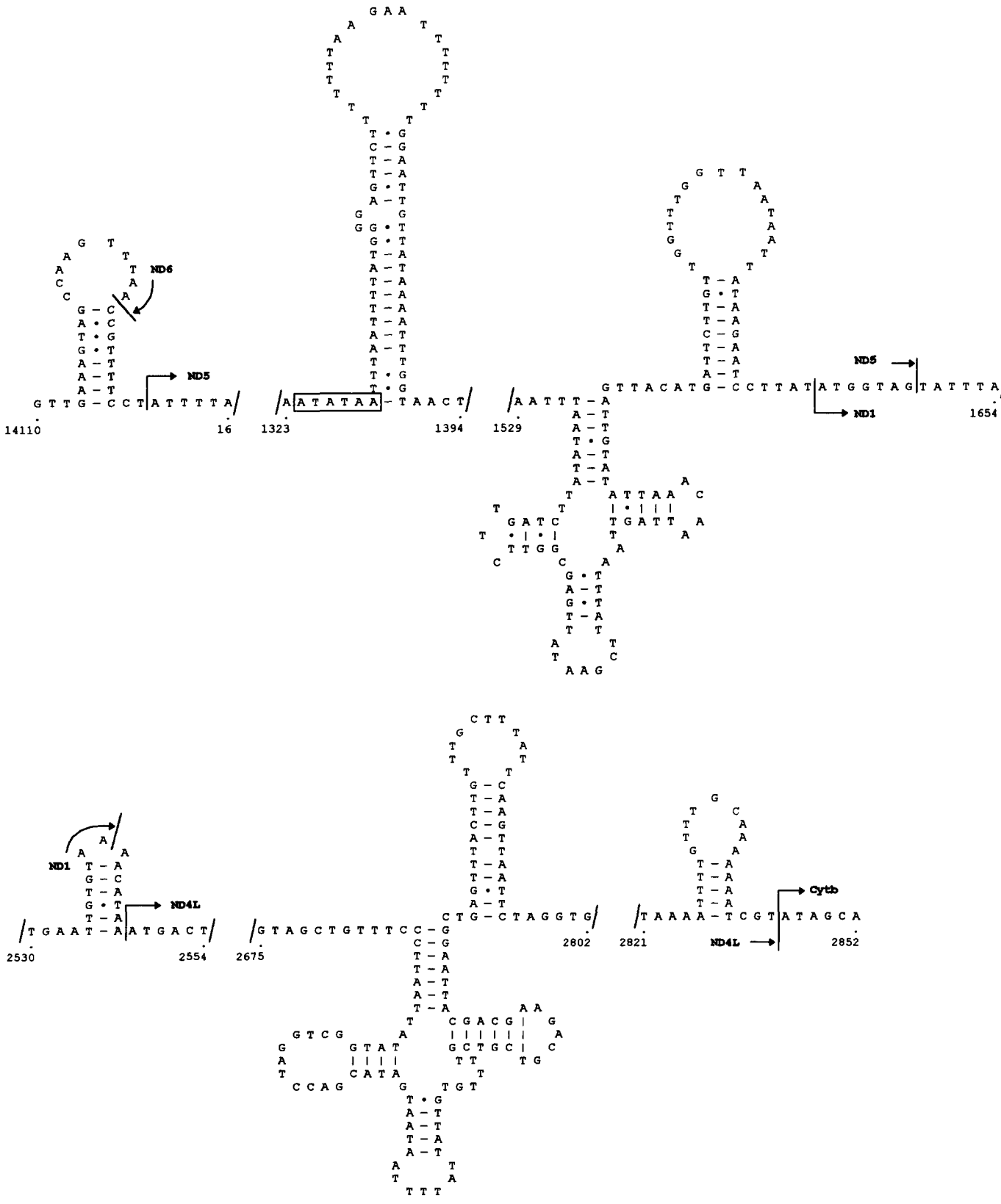

FIGURE 3.—Potential secondary structures near or at the junctions of *ND6-ND5, ND5-ND1, ND1-ND4L,* and *ND4L-Cytb* genes. The start and stop codons are indicated in each case. Within the *ND5* and *ND4L* genes, tRNA-like structures that are followed by a stem and loop structure may be drawn. Moreover, ~300 nt before the end of *ND5* another stem and loop structure may be formed, immediately after the boxed sequence ATATAA; a similar structure has been implicated as the origin of light strand replication in Drosophila (see text).

TABLE 1

Codon usage of *A. coerulea* mtDNA-encoded proteins

Amino acid	Codon	N	%	Amino acid	Codon	N	%	Amino acid	Codon	N	%	Amino acid	Codon	N	%
Phe	TTT	240	6.7	Ser	TCT	77	2.1	Tyr	TAT	148	4.1	Cys	TGT	34	0.9
	TTC	25	0.7		TCC	15	0.4		TAC	38	1.1		TGC	10	0.3
Leu	TTA	340	9.4		TCA	68	1.9	Term	TAA	6	0.2	Trp	TGA	68	1.9
	TTG	63	1.8		TCG	14	0.4		TAG	2	0.1		TGG	20	0.6
Leu	CTT	83	2.3	Pro	CCT	49	1.4	His	CAT	61	1.7	Arg	CGT	17	0.5
	CTC	14	0.4		CCC	22	0.6		CAC	13	0.4		CGC	4	0.1
	CTA	79	2.2		CCA	50	1.4	Gln	CAA	41	1.1		CGA	22	0.6
	CTG	24	0.7		CCG	11	0.3		CAG	10	0.3		CGG	6	0.2
Ile	ATT	266	7.4	Thr	ACT	77	2.1	Asn	AAT	106	2.9	Ser	AGT	49	1.4
	ATC	36	1.0		ACC	23	0.6		AAC	32	0.9		AGC	28	0.8
Met	ATA	213	5.9		ACA	73	2.0	Lys	AAA	73	2.0		AGA	64	1.8
	ATG	50	1.4		ACG	12	0.3		AAG	13	0.4		AGG	30	0.8
Val	GTT	90	2.5	Ala	GCT	91	2.5	Asp	GAT	47	1.3	Gly	GGT	73	2.0
	GTC	20	0.6		GCC	24	0.7		GAC	15	0.4		GGC	24	0.7
	GTA	97	2.7		GCA	66	1.8	Glu	GAA	48	1.3		GGA	79	2.2
	GTG	22	0.6		GCG	12	0.3		GAG	25	0.7		GGG	43	1.2

N, total number of particular codon in all proteins. The total number of codons is 3595; the incomplete termination codons were excluded.

termination codons (either T for *ATPase6, ND3* and *ND4L* or TA for *COI* and *Cytb*).

In general, proteins coded by Albinaria mtDNA are quite short (Table 2). Five of them (COI, COII, COIII, Cyt*b*, ND5) are the shortest among known metazoans (compare with Table II of WOLSTENHOLME 1992). Also, five others (ATPase6, ND1, ND2, ND4, ND6) are the shortest among coelomate metazoans, since these are shorter only in the nematodes (the corresponding lengths in *C. elegans* are 199, 291, 282, 409, and 145 amino acids). The only Albinaria proteins that are not the shortest among known molluscs are ATPase8 and ND4L; both are among the least conserved mitochondrial proteins and exhibit a high degree of length variation (WOLSTENHOLME 1992). It is worth noting that ATPase8 differs in length by >5% between two Albinaria species (*A. coerulea*, 55 amino acids; *A. turrita*, 52 amino acids).

Among compared molluscs of Table 2, there is extensive variation not only in the size of their mitochondrial

TABLE 2

Comparison of the *A. coerulea* mitochondrial protein coding genes with those of *A. turrita, C. nemoralis, K. tunicata, M. edulis* and *D. yakuba*

	Protein length (amino acids)						Percentage of identity				
Protein	*A.c.*	*A.t.*	*C.n.*	*K.t.*	*M.e.*	*D.y.*	*A. coerula vs. A.t.*	*A. coerulea vs. C.n.*	*A. coerulea vs. K.t.*	*A. coerula vs. M.e.*	*A. coerula vs. D.y.*
ATPASE6	214	—	—	230	238	224	—	—	34.0	30.2	34.0
ATPase8	55	52	54	53	A	53	50.9	19.6	28.1	A	26.8
COI	509	P	—	513	P	512	98.9*	—	72.6 (49.0)	— (36.0)	71.5 (46.5)
COII	224	224	—	229	P	228	92.4 (94.7)	—	55.0 (54.1)	— (37.4)	53.9 (54.9)
COIII	259	—	—	259	264	262	—	—	62.7	42.9	59.5
Cytb	367	—	—	379	P	378	—	—	53.6 (50.4)	— (46.0)	53.6 (52.3)
ND1	299	—	—	316	P	324	—	—	46.1 (47.3)	— (40.2)	47.4 (49.8)
ND2	307	—	—	338	P	341	—	—	24.9 (25.2)	— (31.6)	30.9 (32.4)
ND3	117	—	—	120	116	117	—	—	39.2	35.5	39.7
ND4	437	—	—	442	P	446	—	—	35.2 (29.9)	— (33.1)	43.0 (39.5)
ND4L	99	—	79	100	93	96	—	24.2	22.7	26.2	22.5
ND5	545	—	—	571	P	573	—	—	35.0 (37.0)	— (33.9)	34.2 (35.7)
ND6	155	P	180	166	158	174	75.3*	19.9	23.9	17.4	27.8

Percentage of amino acid identity was calculated by dividing the number of identical amino acid positions by the common length of the compared sequences. Numbers in parentheses show the percentage of identity among the partially sequenced portions of *M. edulis* protein coding genes and the corresponding regions of *A. coerulea, A. turrita, K. tunicata* and *D. yakuba* genes. Asterisk denotes the percentage identity values for the partially sequenced *A. turrita COI* and *ND6* genes. A, absent; P, partial. Alignment of compared amino acid sequences is not shown.

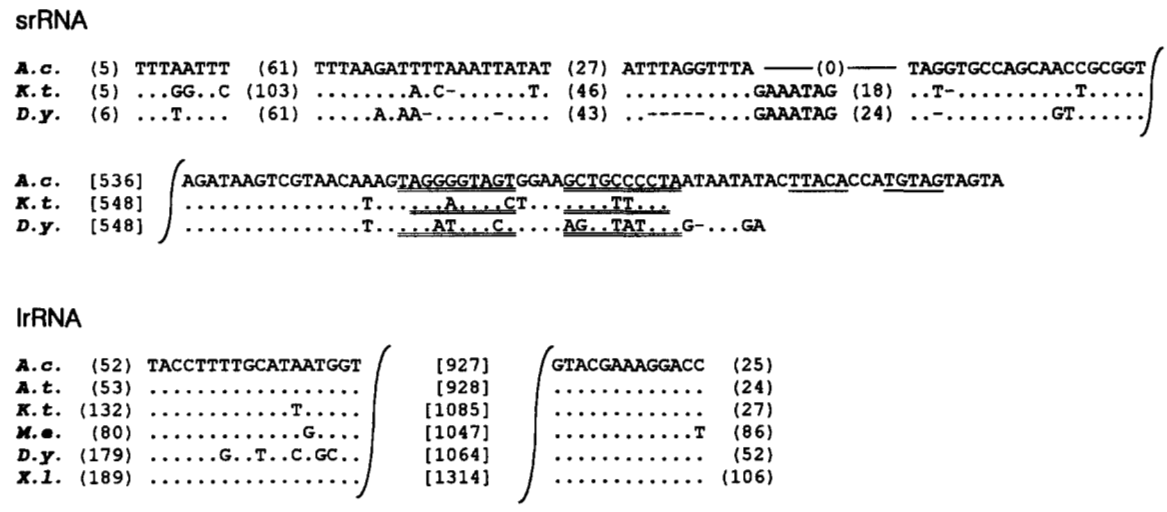

FIGURE 4.—Alignment of the ends of the two ribosomal RNA genes. *A.c., A. turrita; K.t., K. tunicata; M.e., M. edulis; D.y., D. yakuba; X.l., X. laevis.* Numbers in parentheses denote regions where alignment is umbiguous; numbers in brackets denote the total intervening length of unaligned sequences between the ends of the molecules. Double underlining emphasizes inverse repeats forming a stem and loop structure; in the proposed secondary structure of Drosophila *s-rRNA* this is the last hairpin (CLARY and WOLSTENHOLME 1985b). Single underlining indicates an additional hypothetical stem and loop structure in Albinaria.

proteins but also in the percentage of identity they exhibit. As expected, the greatest identity is observed between the two Albinaria species. When more distantly related species are compared, there are some discordances with traditional taxonomic relationships, a conclusion we have previously reached from analysis of *A. turrita* mtDNA sequences (LECANIDOU *et al.* 1994) and which is also reported by BOORE and BROWN (1994b).

Codon usage and codon bias: The genetic code of Albinaria (Table 1) has been presented previously (LECANIDOU *et al.* 1994) and is the same in all known molluscs (HOFFMANN *et al.* 1992; BOORE and BROWN 1994 a,b; TERRETT *et al.* 1994) as well as in Drosophila (CLARY and WOLSTENHOLME 1985a). It differs from the universal code in that ATA codes for methionine, TGA for tryptophane, and AGA and AGG for serine.

As can be seen in Table 1, codons ending at A or T are much more frequent (~81%) than those ending in C or G. In fourfold synonymous codon families 77.5% of the codons end at A or T. Since any nucleotide substitution in these sites does not lead to an amino acid replacement, the AT bias in the third codon position should not be attributed to pressure of natural selection at the protein level.

Seven out of 62 amino acid codons, which consist exclusively of A and/or T, represent 38.4% of the total number of codons. Six of these are the most frequently used codons and correspond to the amino acids Leu, Ile, Phe, Met, Tyr, and Asn, which constitute 16.8, 8.4, 7.4, 7.3, 5.2 and 3.8% of the total number of amino acids, respectively. With the exception of Leu, this order of amino acid frequencies is different from that actually observed in Albinaria mitochondrial proteins, where Ser, and not Ile, is the second most frequent amino acid (9.7%), while Val (6.4%), Gly (6.1%), Ala (5.4%) and Thr (5.2%) are more (or equally, Thr) frequent than Tyr and Asn. Thus, there is no direct correlation between the most commonly used codons and the most frequent amino acids. If the predominance of codons composed solely of A and/or T is attributed to AT bias (CROZIER and CROZIER 1993; BOORE and BROWN 1994b), then the relatively high percentage of certain amino acids that are not coded by strictly A and/or T codons (such as Ser, Thr and the aliphatic Val, Gly, Ala with similar physicochemical properties) must be attributed to selective pressure.

Ribosomal RNA genes: Identification of the Albinaria *s-rRNA* and *l-rRNA* genes was accomplished by comparison with other known mitochondrial ribosomal RNA genes. The *l-rRNA* gene had been previously identified in two cloned segments of *A. turrita* mtDNA (LECANIDOU *et al.* 1994). Using the complete *A. coerulea* mtDNA sequence as a guide, it is inferred that these two segments are actually consecutive and that the 5′ and of the *A. turrita l-rRNA* gene is at residue 631 of segment I (the sequence preceding the 5′ end contains the $tRNA^{Val}$ gene), while the 3′ end is at nucleotide 923 of segment II [due to revision of the *A. turrita* $tRNA^{Leu}$-*(CUN)*] (see discussion on tRNAs).

The ends of the two Albinaria rRNAs cannot be precisely determined because of uncertainties in the alignments with known rRNAs. Figure 4 shows a comparison of the 5′ and 3′ regions, which is based on conserved sequence elements showing a significant degree of similarity. Since the size of the compared regions is smaller in Albinaria, we assume at present that each Albinaria rRNA gene occupies all of the available space between the $tRNA^{Met}$ and $tRNA^{Glu}$ (*s-rRNA*) and the $tRNA^{Val}$ and $tRNA^{Leu}$ genes (*l-rRNA*).

At the 5′ end of the Albinaria large rRNA gene, the first conserved region (18 nt) is located 52 nt down-

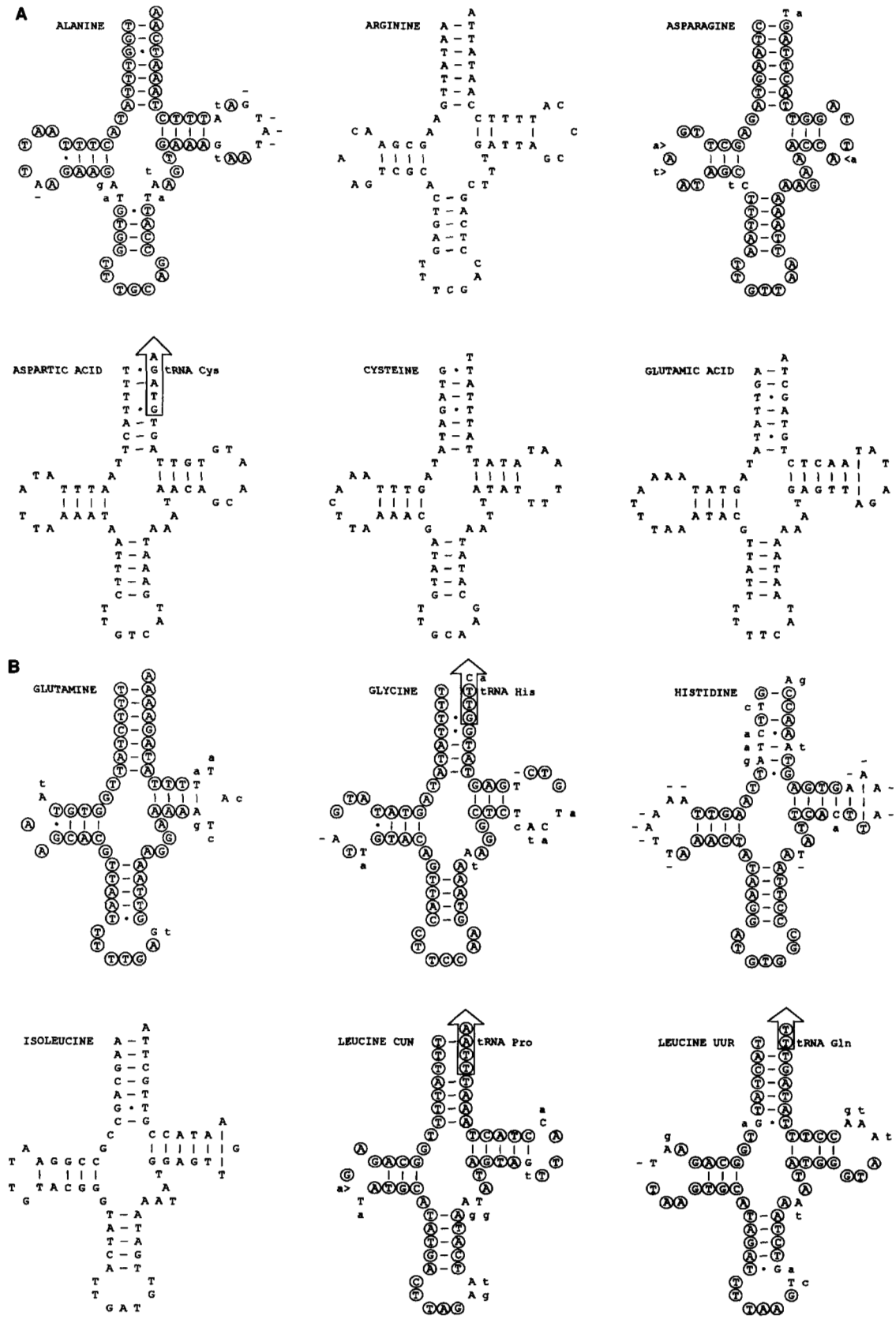

FIGURE 5.—Cloverleaf representation of 22 putative *A. coerulea* tRNA genes. In 11 of them a direct comparison is made with the corresponding *A. turrita* putative tRNAs. O, identical nucleotides; small leters, differences; ---, gaps in *A. turrita*. Arrow-shaped boxes encompassing a variable number of nucleotides denote overlapping regions that are supposedly modified posttranscriptionally by RNA editing using the opposite T nucleotides of the 5′ end as an internal guide.

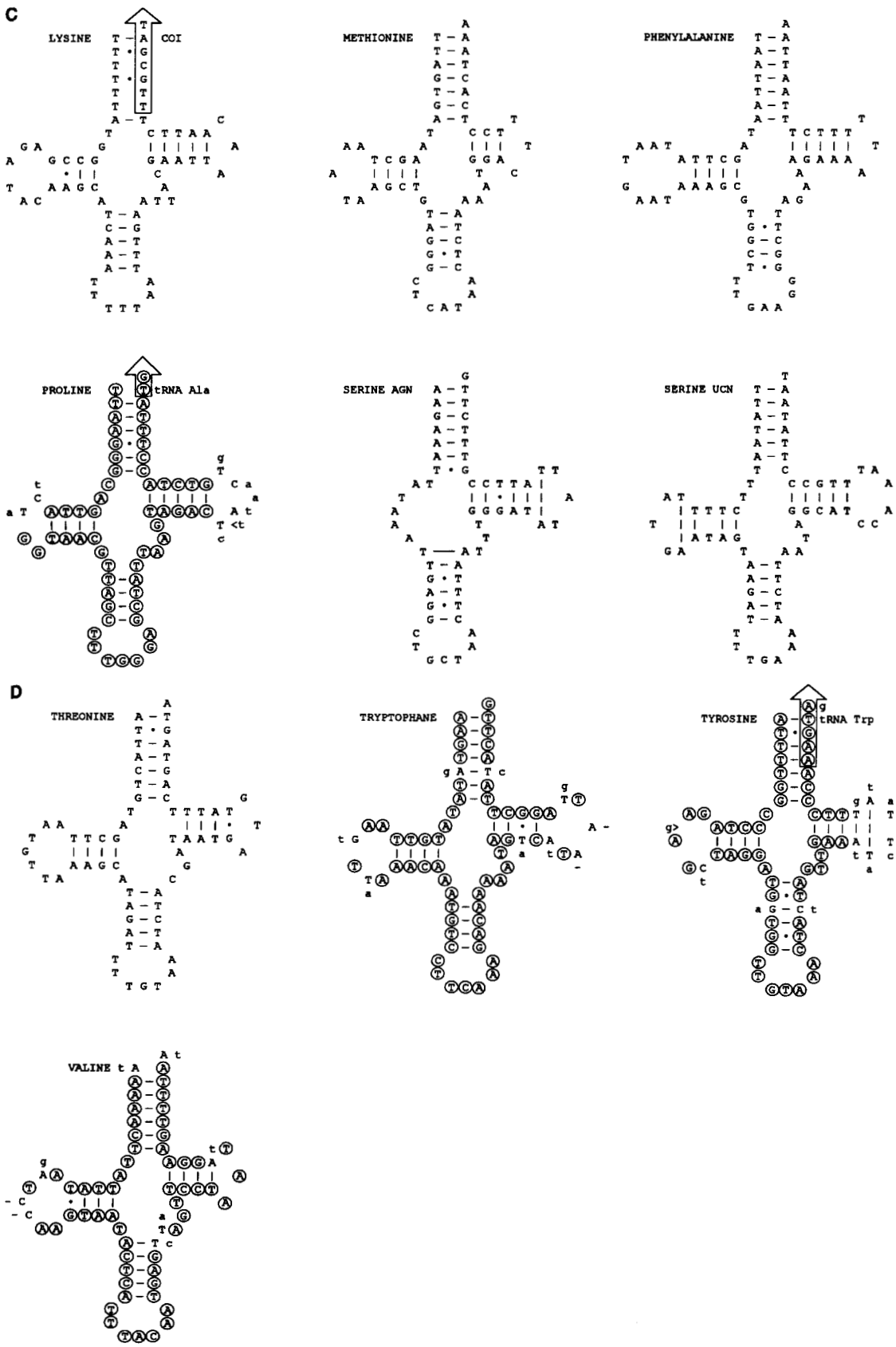

FIGURE 5.—*Continued*

stream from the 3′ end of the *tRNA^{Val}* gene, this being the smallest length among all compared molecules (Figure 4). A similar observation can be made for the 3′ end, where a conserved sequence of 13 nt is located only 24–25 nt from the junction with *tRNA^{Leu}*. The presence of the conserved heptamer sequence TGGCAGA, defined as the rRNA transcription termination signal in animal mitochondria (VALVERDE *et al.* 1994) in positions 8–14 of the adjacent *tRNA^{Leu}* gene, should be noted.

Comparison of the 5′ ends of the small rRNA genes shows that a sequence of 152 nt in Albinaria corresponds to a sequence of 236 nt in Katharina or 192 nt in Drosophila. In contrast, the Albinaria 3′ end appears to be longer than the equivalent regions of Katharina and Drosophila. A sequence of 18 nt, which is almost identical among compared species, is immediately followed by an inverted repeat (double underlining in Figure 4) corresponding to the final stem and loop structure of Drosophila *s-rRNA* (CLARY and WOLSTENHOLME 1985b). In Albinaria this region is followed by a sequence of 27 nt that could potentially form an extra stem and loop structure, as it contains a 5-nt inverted repeat (single underlining in Figure 4). In any case, the exact points of the rRNA gene ends must be determined by more direct methods (VAN ETTEN *et al.* 1980; CLARY and WOLSTENHOLME 1985b).

Even if we assume that the Albinaria ribosomal RNA genes occupy all of the available space between adjacent tRNA genes, they still represent the shortest rRNA genes among coelomate metazoans (*s-rRNA/l-rRNA:* Albinaria, 759/1035 bp; Katharina 826/1275 bp; Mytilus, 945/1244 bp; Drosophila, 789/1326) but are larger than those of pseudocoelomate nematodes (compare with Table VI of WOLSTENHOLME 1992; OKIMOTO *et al.* 1992).

Transfer RNA genes: Identification of the standard set of 22 Albinaria tRNA genes was based on their predicted cloverleaf structures, which define unambiguous anticodons (Figure 5). Some of the putative secondary structures of ten previously reported *A. turrita* tRNA-gene sequences (LECANIDOU *et al.* 1994) have been redrawn to conform with the *A. coerulea* tRNA secondary structures (see discussion below and Figure 5). In addition to these tRNA genes, two more sequences that can be folded into tRNA-like structures were detected (see gene arrangement) within protein coding genes (Figure 3).

All standard Albinaria tRNAs have the same anticodons as those reported for Mytilus and Katharina; these are preceded by T and followed by either A or G (purine). Five cases of mismatched base pairs at exactly equivalent positions are evident in the anticodon stem (at the top of five-membered stems; in *Ala, Gly, Phe, Pro, Trp*). Albinaria anticodon stems consist of 5 and sometimes 6 bp [*Ser(AGN),* Tyr, and probably *Ser(UCN)*]; in the case of *Ser(AGN),* as many as 9 bp are possible. Although not shown, the Katharina and Mytilus *tRNA^{Ser}(AGN)* also have the potential of forming 9-bp anticodon stems, all three molluscan sequences being very conserved in this region. Anticodon stems with an increased potential of base pairing have been also drawn for *Ser(AGN)* and *Ser(UCN)* of *C. elegans* (WOLSTENHOLME 1992). The actual presence of a six-membered anticodon stem has been demonstrated by direct RNA sequencing of a mammalian *tRNA^{Ser}(UCN)* (YOKOGAWA *et al.* 1991; JANKE *et al.* 1994). Interestingly, the Albinaria *tRNA^{Tyr}* gene conforms very well with the secondary structure of mammalian *tRNA^{Ser}(UCN),* in that, in addition to the six-membered anticodon stem, it contains only one nucleotide between the 7-bp acceptor stem and the 4-bp D stem. Finally, the Albinaria *tRNA^{Ser}(UCN)* gene may equally well be drawn to conform with the secondary structure of its mammalian counterpart.

Amino-acyl stems consist of 7 bp. As can be noted in Figure 5, mismatching is observed in several amino-acyl stems [*Asp, Gly, His, Leu(CUN), Leu(UUR), Lys, Pro, Ser(UCN), Val*]. What is interesting is that in most cases of mismatches, we are dealing with overlapping genes: *Asp* with *Cys, Gly* with *His, Leu(CUN)* with *Pro, Leu(UUR)* with *Gln, Lys* with *COI, Pro* with *Ala.* It is also worth noticing that in all such cases the 3′ ends of the amino-acyl stems are almost invariably composed of T residues. Recently, it was demonstrated that in *A. castellanii* tRNA mitochondrial genes, certain bases of the 5′ end (confined to the first 3 bp of the acceptor arm, where correct base pairing is presumably essential for biological activity) are modified posttranscriptionally by a process of RNA editing (LONERGAN and GRAY 1993); the specificity of editing, rather than being provided by guide RNAs, could be provided by the 3′ end of the acceptor stem itself. In Albinaria we believe that a similar mechanism of RNA editing is operating, but that in this case what we probably have is the 5′ end of the acceptor stem acting as an internal guide for editing of the 3′ end. This editing might thus resemble a primitive polyadenylation mechanism. Final verification must await direct tRNA sequence determination.

We thank professors R. A. D. CAMERON, M. J. SMITH and an anonymous reviewer for helpful comments on the manuscript. We thank VASSILIS DOURIS and LARA KRAVARITI for their excellent technical assistance. This work was supported by the University of Athens and by the General Secretariat of Research and Technology of the Greek Ministry of Industry, Energy and Technology.

LITERATURE CITED

ANDERSON, S., A. T. BANKIER, B. G. BARRELL, M. H. L. DE BRUIJN, A. R. COULSON *et al.*, 1981 Sequence and organization of the human mitochondrial genome. Nature **290:** 457–465.

ANDERSON, S., M. H. L. DE BRUIJN, A. R. COULSON, I. C. EPERON, F. SANGER *et al.*, 1982 Complete sequence of bovine mitochondrial DNA: conserved features of the mammalian mitochondrial genome. J. Mol. Biol. **156:** 683–717.

ARNASON, U., and E. JOHNSSON, 1992 The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina.* J. Mol. Evol. **34:** 493–505.

ARNASON, U., and A. GULLBERG, 1993 Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. J. Mol. Evol. **37:** 312–322.

ARNASON, U., A. GULLBERG and B. WIDEGREN, 1991 The complete

nucleotide sequence of the mitochondrial DNA of the fin whale, *Balaenoptera physalus*. J. Mol. Evol. **33:** 556–568.

ASAKAWA, S., Y. KUMAZAWA, T. ARAKI, H. HIMENO, K. MIURA *et al.*, 1991 Strand-specific nucleotide composition bias in echinoderm and vertebrate mitochondrial genomes. J. Mol. Evol. **32:** 511–520.

AVISE, J. C., B. W. BOWEN, T. R. LAMB, A. B. MEYLAN and E. BERMINGHAM, 1992 Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the testudines. Mol. Biol. Evol. **9:** 457–473.

BAIROCH, A., 1988 PC/Gene version 5.15. University of Geneva, Switzerland.

BIBB, M. J., R. A. VAN ETTEN, C. T. WRIGHT, M. W. WALBERG and D. A. CLAYTON, 1981 Sequence and gene organization of mouse mitochondrial DNA. Cell **26:** 167–180.

BOORE, J. L., and W. M. BROWN, 1994a Mitochondrial genomes and the phylogeny of mollusks. Nautilus (Suppl.) **2:** 61–78.

BOORE, J. L., and W. M. BROWN, 1994b Complete DNA sequence of the mitochondrial genome of the black chiton, *Katharina tunicata*. Genetics **138:** 423–443.

BOYCE, T. H., M. E. ZWICK and C. F. AQUADRO, 1989 Mitochondrial DNA in the pine weevils: size, structure and heteroplasmy. Genetics **123:** 825–836.

BROUGHTON, R. E., and T. E. DOWLING, 1994 Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics **138:** 179–190.

BURGE, C., A. M. CAMBELL and S. KARLIN, 1992 Over- and under-representation of short oligonucleotides in DNA sequences. Proc. Natl. Acad. Sci. USA **89:** 1358–1362.

CANTATORE, P., M. N. GADALETA, M. ROBERTI, C. SACCONE and A. C. WILSON, 1987 Duplication and remoulding of tRNA genes during the evolutionary rearrangement of mitochondrial genomes. Nature **329:** 853–855.

CANTATORE, P., M. ROBERTI, G. RAINALDI, M. N. GADALETA and C. SACCONE, 1989 The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of *Paracentrotus lividus*. J. Biol. Chem. **264:** 10965–10975.

CARDON, L. R., C. BURGE, D. A. CLAYTON and S. KARLIN, 1994 Pervasive CpG suppression in animal mitochondrial genomes. Proc. Natl. Acad. Sci. USA **91:** 3799–3803.

CHANG, D. D., T. W. WONG, J. E. HIXON and D. A. CLAYTON, 1985 Regulatory sequences for mammalian mitochondrial transcription and replication, pp. 135–144 in *Achievements and Perspectives of Mitochondrial Research, Vol. II. Biogenesis,* edited by E. QUAGLIARIELLO, E. C. SLATOR, F. PALMIERI, C. SACCONE and A. M. KROON. Elsevier, Amsterdam.

CHANG, Y.-S., F.-L. HUANG and T.-B. LO, 1994 The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. J. Mol. Evol. **38:** 138–155.

CLARY, D. O., and D. R. WOLSTENHOLME, 1985a The mitochondrial DNA molecule of *Drosophila yakuba:* nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. **22:** 252–271.

CLARY, D. O., and D. R. WOLSTENHOLME, 1985b The ribosomal RNA genes of *Drosophila* mitochondrial DNA. Nucleic Acids Res. **31:** 4029–4044.

CLARY, D. O., and D. R. WOLSTENHOLME, 1987 *Drosophila* mitochondrial DNA: conserved sequences in the A+T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. J. Mol. Evol. **25:** 116–125.

CLAYTON, D. A., 1991 Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. **7:** 453–478.

CLAYTON, D. A., 1992 Transcription and replication of animal mitochondrial DNAs. Int. Rev. Cyt. **141:** 217–232.

CROZIER, R. H., and Y. C. CROZIER, 1993 The mitochondrial genome of the honeybee *Apis mellifera:* complete sequence and genome organization. Genetics **133:** 97–117.

CROZIER, R. H., Y. C. CROZIER and A. G. MACKINLAY, 1989 The COI and COII region of honey-bee mitochondrial DNA: evidence for variation in insect mitochondrial rates. Mol. Biol. Evol. **6:** 399–411.

DE GIORGI, C., C. LANAVE, M. D. MUSCI and C. SACCONE, 1991 Mitochondrial DNA in the sea urchin *Arbacia lixula:* evolutionary inferences from nucleotide sequence analysis. Mol. Biol. Evol. **8:** 515–529.

DESJARDINS, P., and R. MORAIS, 1990 Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher in higher vertebrates. J. Mol. Biol. **212:** 599–634.

DOURIS, V., G. C. RODAKIS, S. GIOKAS, M. MYLONAS and R. LECANIDOU, 1995 Mitochondrial DNA and morphological differentiation of *Albinaria* populations (Gastropoda: Clausiliidae). J. Moll. Studies **61:** 65–78.

GADALETA, G., G. PEPE, G. DE CANDIA, C. QUAGLIARIELLO, E. SBISA *et al.*, 1989 The complete nucleotide sequence of the *Rattus norvegicus* mitochondrial genome: cryptic signals revealed by comparative analysis between vertebrates. J. Mol. Evol. **28:** 497–516.

GHIVIZZANI, S. C., S. L. D. MACKAY, C. S. MADSEN, P. J. LAIPIS and W. W. HAUSWIRTH, 1993 Transcribed heteroplasmic repeated sequences in the porcine mitochondrial DNA D-loop region. J. Mol. Evol. **37:** 36–47.

GJETVAJ, B., D. I. COOK and E. ZOUROS, 1992 Repeated sequences and large scale size variation of mitochondrial DNA: a common feature among scallops (Bivalvia: Pectinidae). Mol. Biol. Evol. **9:** 106–124.

HENIKOFF, S., 1987 Exonuclease III generated deletions for DNA sequence analysis. Promega Notes No. 8.

HOFFMANN, R. J., J. L. BOORE and W. M. BROWN, 1992 A novel mitochondrial genome organization for the blue mussel *Mytilus edulis*. Genetics **131:** 397–412.

JACOBS, H., D. ELLIOT, V. MATH and A. FARGUHARSON, 1988 Nucleotide sequence and gene organization of sea urchin mtDNA. J. Mol. Biol. **201:** 185–217.

JACOBS, H. T., E. R. HERBERT and J. RANKINE, 1989 Sea urchin egg mitochondrial DNA contains a short displacement loop (D-loop) in the replication origin region. Nucleic Acids Res. **17:** 8949–8965.

JANKE, A., G. FELDMAIER-FUCHS, W. K. THOMAS, A. VON HAESELER and S. PÄÄBO, 1994 The marsupial mitochondrial genome and the evolution of placental mammals. Genetics **137:** 243–256.

L'ABBÉ, D., J. F. DUHAIME, B. F. LANG and R. MORAIS, 1991 The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. J. Biol. Chem. **266:** 10844–10850.

LAROCHE, J., M. SNYDER, D. I. COOK, K. FULLER and E. ZOUROS, 1990 Molecular characterization of a repeat element causing large size variation in the mitochondrial DNA of the sea scallop *Placeopecten magellanicus*. Mol. Biol. Evol. **7:** 45–64.

LECANIDOU, R., V. DOURIS and G. C. RODAKIS, 1994 Novel features of metazoan mtDNA revealed from sequence analysis of three mitochondrial DNA segments of the land snail *Albinaria turrita* (Gastropoda: Clausiliidae). J. Mol. Evol. **38:** 369–382.

LONERGAN, K. M., and M. W. GRAY, 1993 Editing of transfer RNAs in *Acanthamoeba castelanii* mitochondria. Science **259:** 812–816.

MARTIN, A. P., G. J. P. NAYLOR and S. R. PALUMBI, 1992 Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature **357:** 153–155.

MITCHELL, S. E., A. F. COCKBURN and J. A. SEAWRIGHT, 1993 The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization. Genome **36:** 1058–1073.

MORITZ, C., T. E. DOWLING and W. M. BROWN, 1987 Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. **18:** 296–292.

OJALA, D., C. MERKEL, R. GELFAND and G. ATTARDI, 1980 The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell **22:** 393–403.

OJALA, D., J. MONTOYA and G. ATTARDI, 1981 tRNA punctuation model of RNA processing in human mitochondria. Nature **290:** 470–474.

OKIMOTO, R., J. L. MACFARLANE and D. R. WOLSTENHOLME, 1990 Evidence for the frequent use of TTG as the translation initiation codon of mitochondrial protein genes in the nematodes, *Ascaris suum* and *Caenorhabditis elegans*. Nucleic Acids Res. **18:** 6113–6118.

OKIMOTO, R., J. L. MACFARLANE, D. O. CLARY and D. R. WOLSTENHOLME, 1992 The mitochondrial genomes of two nematodes *Caenorhabditis elegans* and *Ascaris suum*. Genetics **130:** 471–498.

PALUMBI, S. R., and J. BENZIE, 1991 Large mitochondrial DNA differences between morphologically similar Paneid shrimb. Mol. Mar. Biol. Biotech. **1:** 27–34.

PEARSON, W. R., and D. J. LIPMAN, 1988 Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:** 2444–2448.

PEREZ, M. L., J. R. VALVERDE, B. BATUECAS, F. AMAT, R. MARCO *et al.*, 1994 Speciation in the *Artemia* genus: mitochondrial DNA analysis of bisexual and parthenogenetic shrimps. J. Mol. Evol. **38:** 156–168.

PUSTELL, J., and F. C. KAFATOS, 1984 A convenient and adaptable

package of computer programs for DNA and protein sequence management, analysis and homology determination. Nucleic Acids Res. **12:** 643–655.

PUSTELL, J., and F. C. KAFATOS, 1986 A convenient and adaptable microcomputer environment for DNA and protein manipulation and analysis. Nucleic Acids Res. **14:** 479–488.

RAMIREZ, V., P. SAVOIE and R. MORAIS, 1993 Molecular characterization and evolution of a duck mitochondrial genome. J. Mol. Evol. **37:** 296–310.

ROE, B. A., D.-P. MA, R. K. WILSON and J. F.-H. WONG, 1985 The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. J. Biol. Chem. **260:** 9759–9774.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 Molecular cloning: a laboratory manual. Cold Spring Laboratory Press, Cold Spring Harbor, NY.

SANGER, F., S. KICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SHADEL, G. S., and D. A. CLAYTON, 1993 Mitochondrial transcription initiation. J. Biol. Chem. **268:** 16083–16086.

SMITH, M. J., A. ARNDT, S. GORSKI and E. FAJBER, 1993 The phylogeny of echinoderm classes based on mitochondrial gene arrangements. J. Mol. Evol. **36:** 545–554.

SNYDER, M., A. R. FRASER, J. LAROCHE, K. E. GARTNER-KEPKAY and E. ZOUROS, 1987 Atypical mitochondrial DNA from the deep-sea scallop *Placopecten magellanicus*. Proc. Natl. Acad. Sci. USA **84:** 7595–7599.

SRIVASTAVA, K. K., E. E. CABLE and H. L. BONKOVSKY, 1993 Improved resolution and sensitivity of Northern blots using polyacrylamide-urea gels. Biotechniques **15:** 227–231.

TERRETT, J., S. MILES and R. H. THOMAS, 1994 The mitochondrial genome of *Cepaea nemoralis:* gene order, base composition, and heteroplasmy. Nautilus (Suppl.) **2:** 79–84.

VALVERDE, J. R., R. MARCO and R. GARESSE, 1994 A conserved heptamer motif for ribosomal RNA transcription termination in animal mitochondria. Proc. Natl. Acad. USA **91:** 5368–5371.

VAN ETTEN, R. A., M. W. WALBERG and D. A. CLAYTON, 1980 Precise localization and nucleotide sequence of the two mouse mitochondrial rRNA genes and three immediately adjacent novel tRNA genes. Cell **22:** 157–170.

WOLSTENHOLME, D. R., 1992 Animal mitochondrial DNA: Structure and evolution. Int. Rev. Cyt. **141:** 173–216.

YOKOGAWA, T., Y. WATANABE, Y. KUMAZAWA, T. UEDA, I. HIRAO *et al.*, 1991 A novel cloverleaf structure found in mammalian mitochondrial $tRNA^{Ser}$ (UCN). Nucleic Acids Res. **19:** 6101–6105.

Communicating editor: W.-H. LI