Rearranged Coding Segments, Separated by a Transfer RNA Gene, Specify the Two Parts of a Discontinuous Large Subunit Ribosomal RNA in *Tetrahymena pyriformis* Mitochondria*

(Received for publication, September 4, 1986)

Taisto Y. K. Heinonen‡, Murray N. Schnare‡, Paul G. Young§, and Michael W. Gray‡‡

From the ‡Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia B3H 4H7, Canada and the §Department of Biology, Queen’s University, Kingston, Ontario K7L 3N6, Canada

In the mitochondria of *Tetrahymena pyriformis* ST, the large subunit ribosomal RNA (LSU rRNA) is discontinuous, consisting of two species, α (280 nucleotides in length) and β (~2.45 kilobases long). LSUα is the 5’-terminal portion (i.e. a 5.8 S-like rRNA), and LSUβ constitutes the rest of this LSU rRNA, as judged by both primary and secondary structure homology to other LSU rRNAs. Remarkably, LSUα is encoded downstream of LSUβ, the two genes being separated by a tRNA^16s^ gene. This is the first demonstration of a discontinuous rRNA whose coding sequences are rearranged, deviating from the conventional, highly conserved, 5’→3’ order of sequence domains in the LSU rRNA gene. This novel gene organization (5'-LSUβ-3'-LSUα) is identical in both copies of the subterminal inverted repeat in the linear *T. pyriformis* mitochondrial genome. Sequence heterogeneity in the LSUα transcript data, suggest that both copies of the inverted repeat are expressed.

Primary sequence comparisons of large subunit (LSU) ribosomal RNAs have revealed that these molecules consist of two different kinds of sequence domain, conserved and variable, alternating with one another in a characteristic order (1–9). The conserved domains display a high degree of conservation of both primary and secondary structure between and within species. In contrast, the variable domains are characterized by pronounced differences in length, primary sequence, and potential secondary structure and are largely responsible for the structural diversity observed in these molecules.

LSU rRNAs can be either continuous, consisting of a single RNA molecule, or discontinuous, in which case the covalent ribose phosphate backbone is disrupted at one or more positions. Such discontinuous LSU rRNAs are composed of two or more fragments, which base pair with one another to produce a secondary structure the same as, or very similar to, that of a covalently continuous LSU rRNA. The best known example of LSU rRNA discontinuity is that which divides most eukaryotic cytoplasmic LSU rRNAs into two separate species, designated 5.8 S and 28 S rRNAs. The 5.8 S rRNA is equivalent to the 5’-terminal region of the covalently continuous LSU (23 S) rRNA of *Escherichia coli*, as judged by both primary (10–13) and secondary (3, 8, 9, 13) structure homology. Other discontinuities have been localized in the 5’-region (14, 15), in the middle (16, 17), and in the 3’-region (18–21) of LSU rRNAs. Interestingly, the discontinuities seem to be confined to the variable domains of LSU rRNA.

To the extent that the transcription of discontinuous rRNAs has been studied, the separate portions of the rRNA are co-transcribed as a single long precursor, from which the mature RNA species are generated by spacer excision (14–17, 20, 21). Regardless of the exact number or position of the breaks, the relative order of the various coding portions of the LSU rRNA has been highly conserved at the DNA level. In all previously described cases in which spacer sequences interrupt coding sequences, the latter still follow the conventional 5’ to 3’ order of rRNA domains.

We have been examining the structure and organization of the mitochondrial rRNAs and their genes in *Tetrahymena pyriformis*. The mitochondrial ribosomes in this organism are unusual in that they dissociate into subunits that are indistinguishable on the basis of sedimentation velocity (22) as well as size and morphology (23). Nevertheless, we have identified the constituent SSU and LSU rRNAs of the *T. pyriformis* mitochondrial ribosome, based on clear homologies to the SSU and LSU rRNAs, respectively, of conventional ribosomes. We have recently shown that the SSU rRNA is discontinuous (24). We now show that the LSU rRNA is also discontinuous, consisting of two separate species, LSUα (a 280-nucleotide RNA) and LSUβ (an approximately 2450-nucleotide RNA). Remarkably, although LSUα clearly represents the 5’-terminal portion of the LSU rRNA in this system, we find that it is encoded downstream of LSUβ, the two coding sequences being separated from one another by a gene for tRNA^16s^.

**EXPERIMENTAL PROCEDURES**

Preparation of Mitochondrial Nucleic Acids from *T. pyriformis*—Growth of *T. pyriformis*, amicronucleate strain ST, and preparation of mitochondrial DNA were as described previously (24). Total mitochondrial rRNA was prepared as described (24) except that the CsCl gradient was replaced by a sucrose gradient.

* This investigation was supported by Research Grant MT-4124 from the Medical Research Council of Canada (to M. W. G.) and by Grants A8387 (to M. W. G.), A9035 (to P. G. Y.), and a postdoctoral fellowship (to T. Y. K. H.) from the Natural Sciences and Engineering Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J02688.

‡ To whom correspondence should be addressed.

§ The abbreviations used are: LSU, large subunit; SSU, small subunit; kb, kilobase(s); bp, base pair(s); UCS, upstream conserved sequence; ORF, open reading frame.

gradient step and subsequent dialysis were omitted. RNA, including tRNA (and a small amount of DNA), from phenol/chloroform-extracted mitochondrial lysates was precipitated once with ethanol, redissolved, and loaded directly onto gels.

Cloning of Restriction Fragments of T. pyriformis Mitochondrial DNA—Restriction fragments of isolated mitochondrial DNA were cloned by standard methods (26). HindIII fragments were cloned into the HindIII site in either pAT153 (E. coli host strain HB101) or pUC19 (E. coli strain JM83). The 1.3-kb HindIII fragment (see Fig. 4) was shotgun cloned into pAT153 from total HindIII-digested mitochondrial DNA. The 0.8-kb HindIII fragment was cloned into pUC19 after first isolating EcoRI fragment 1, corresponding to the left end of the genome (26), by electrophoresis in low melting point agarose and subjecting the purified fragment to digestion with HindIII. HaeIII fragments from total HaeIII-digested mitochondrial DNA were shotgun cloned into the Smal site of pUC19. Sau3A1 fragments were shotgun cloned into the BamHI site of pUC19 after enrichment for the larger fragments by precipitation of Sau3A1-digested mitochondrial DNA twice from 4% polyethylene glycol 6000, 0.5 M NaCl.

The cloned HaeIII and HindIII fragments were screened as described (24). The left end-specific 0.9-kb Sau3A1 fragment was identified in the Sau3A1 library by the in situ colony hybridization method (27) using the right end-specific 430-bp EcoRV-HaeIII fragment (Fig. 1C) as a probe.

Northern and Southern Hybridizations—Electrophoresis of DNA in agarose gels and of RNA in agarose-methyl mercury gels and transfer to GeneScreen hybridization membranes (New England Nuclear) were as described previously (24). Labeled hybridization probes were prepared by nick translation of recombinant plasmids (or in some cases, restriction fragments of recombinant plasmids) using DNA polymerase I and DNase I (International Biotechnologies, Inc.) and [α-32P]dATP.

Nucleotide Sequence Analysis of RNA and DNA—Sequence analysis of RNA and DNA was performed as described previously (24).

RESULTS

RNA Sequence Analysis of T. pyriformis LSU rRNA Components—When T. pyriformis mitochondrial RNA is fractionated by electrophoresis under denaturing conditions, it is resolved into four major species that are 203, 280, 1406, and 2450 nucleotides in length (Fig. 1A). The 203- and 1406-nucleotide species are the two separate components of a discontinuous SSU rRNA (24). As we document here, the LSU rRNA is also discontinuous, being composed of the 280-nucleotide species and the ~2450-nucleotide species, which we term LSUA and LSUB, respectively. Each species was subjected to RNA sequence analysis by both chemical and enzymatic hydrolysis, with the entire sequence of LSUA being determined (Figs. 2 and 5). The 5′ terminus of LSUA is completely homogeneous, whereas the 3′ terminus shows minor heterogeneity; the majority of the molecules terminate in U corresponding to position 280, but a small proportion (less than one-fifth) of the population terminates one nucleotide sooner, at U279. A noteworthy feature of the LSUs RNA sequence is the 1-nucleotide internal heterogeneity at position 182, where sequencing ladders clearly show the presence of both G and A (Fig. 3A), as discussed in more detail below.

LSUB, which corresponds to a previously identified and mapped 21 S RNA (26), was characterized by RNA sequence analysis of its 5′-terminal 110 nucleotides and 3′-terminal 90 nucleotides. Both termini are discrete, with no detectable sequence heterogeneity. End-labeling experiments (data not shown) indicated that both LSUA and LSUB have 5′-P, 3′-OH termini.

LSUA is the 5′-terminal portion of the T. pyriformis mitochondrial LSU rRNA, as judged by primary and secondary structure homology to the 5′-terminal sequences of other LSU rRNAs. In comparisons of several well conserved domains in the 5′-region of the LSU rRNA (Fig. 2), LSUA shows obvious homology to the 5′-regions of eu-bacterial (E. coli) (28), chlo-

![Fig. 1. Electrophoretic and Northern hybridization analysis of T. pyriformis mitochondrial RNA. A, T. pyriformis mitochondrial RNA fractionated by electrophoresis (in a 1.5% agarose-methyl mercury gel, stained with ethidium bromide and visualized under UV light). B and C, Northern blots corresponding to A probed with restriction fragments from the LSUA-β gene complex in the right-hand end of the genome. Autoradiographs are shown. The positions of the probes are indicated at the bottom. Open rectangles denote the genes for the LSU rRNAs, with transcription proceeding from left to right. The solid circles denote tRNA genes (Le, leucine; Me, methionine). The arrow indicates the putative precursor transcript (estimated size, 3.1 kb).]
rRNAs. Thus, the LSUα-LSUβ discontinuity occurs further downstream of the usual 5.8 S–28 S rRNA discontinuity and is localized to the same position in both *T. pyriformis* and *P. primaurelia*. This breakpoint also closely coincides with the break between the 3 S and 23 S rRNA species in *C. reinhardtii* chloroplasts (29). These primary sequence alignments are further supported by secondary structure homologies, as shown below.

Similarly, primary and secondary structure comparisons with other LSU rRNAs demonstrate that LSUβ corresponds to the remainder of the *T. pyriformis* mitochondrial LSU rRNA. In the primary sequence comparisons shown in Fig. 2, sequence block D near the 5' terminus of LSUβ is clearly homologous to a conserved sequence block that in conventional LSU rRNAs is located just downstream from the position of the LSUα-LSUβ breakpoint. In particular, LSUβ is homologous to the 20 S rRNA in *P. primaurelia* mitochondria. Thus, LSUβ corresponds to the 3'-terminal 90% of the LSU rRNA in *T. pyriformis* mitochondria.

**DNA Sequence Analysis of the Coding Regions of the LSU rRNA Components**—The linear 55-kb mitochondrial genome of *T. pyriformis* (32, 33) contains a short (approximately 3 kb) stretch of sequence that is repeated in inverted orientation near each end of the genome (34). The genes for LSUβ (21 S
rRNA) had previously been mapped to this subterminal inverted repeat (26, 33). Our own hybridization studies showed that LSUα also maps to the inverted repeat (see below). In order to delineate the organization of the coding sequences for the LSU rRNAs, we mapped and cloned the appropriate restriction fragments from the inverted repeat at both ends of the genome. Some relevant mapping data are shown in Fig. 4, which shows restriction fragments specific to one end or the other. Although the 1.3-kb HindIII fragment has been mapped to a different location by Suyama et al. (33), our hybridization and sequence data verify that the position shown in Fig. 4 is correct.

Based on the identification of LSUα as the 5'-terminal portion of the LSU rRNA, we expected to find its gene near the 5'-end of the LSUβ coding sequence. However, both Southern (data not shown) and Northern hybridizations (Fig. 1, B and C) indicated that the LSUα coding sequence mapped to the 3'-side of the LSUβ coding sequence. Because of this unexpected finding, we proceeded to sequence portions of the inverted repeat. The sequenced regions, as well as the sequencing strategy, are indicated in Fig. 4, which summarizes our data pertaining to rRNA gene organization in this genome.

DNA sequence analysis confirmed that LSUα is not encoded at the 5'-end of the LSUβ gene in either copy of the inverted repeat. The DNA sequences of the 5'-regions of the LSUβ coding sequences are shown in Fig. 5 (upper half). It is clear that the two sequences diverge very quickly upstream of the LSUβ 5'-terminus. There is a short 20-bp block of conserved sequence (UCS, upstream conserved sequence), which is nearly identical in both copies of the inverted repeat. Upstream of the UCS the sequences are unrelated in the two ends of the genome. We have determined 650 bp of the upstream sequence in the right end and 100 bp in the left end of the genome and have found no other obvious sequence homology.

The DNA sequences of the 3' terminus of the LSUβ gene and the distal ends of the inverted repeat are also shown in Fig. 5 (lower half). This analysis verified that LSUα is indeed encoded downstream from LSUβ and showed that the two genes are separated by a gene for tRNALeu in both copies of the inverted repeat (the tRNALeu gene was first identified in the inverted repeat of *Tetrahymena thermophila* mitochondrial DNA (35)). Short (2 and 10 bp) spacers separate the three genes, and these spacers are identical in the two copies of the inverted repeat. Thus, the inverted repeat is defined by the unit 5'-(UCS)-(LSUβ)-(2-bp spacer)-(tRNALeu)-(10-bp spacer)-(LSUα)-3'.

Although DNA sequence homology between the two copies of the inverted repeat ends precisely at the 3' terminus of the LSUα coding region, our sequence analyses extend a short distance downstream of this point. These downstream regions contain other tRNA sequences, tRNAThr in the left and tRNASer in the right end of the genome, separated from the LSUα coding sequence by 14- and 15-bp (unrelated) spacers, respectively. These downstream tRNA genes are the most distal known genes in the genome (see Fig. 4). The tRNA genes and their products are discussed in more detail below.

**Potential Secondary Structure Interactions in the LSUα-LSUβ rRNA Complex**—A second criterion identifying LSUα as the 5' terminus of the *T. pyriformis* LSU rRNA is provided by secondary structure modeling. In Fig. 6, the potential interactions between LSUα and LSUβ have been modeled after the secondary structure proposed for the analogous

---

**Fig. 4. Summary of rRNA gene organization in the *T. pyriformis* mitochondrial genome.** The thick line denotes the mitochondrial genome, and the dashed lines indicate the terminal length heterogeneity. Open rectangles denote rRNA genes with the direction of transcription shown by the arrows. The diamonds denote tRNA genes (*Ty* = tyrosine, *Le* = leucine, *Me* = methionine). The lines above and below the rRNA genes indicate cloned restriction fragments (*X* = XbaI, *S* = Sau3A1, *Hi* = HindIII, *Ha* = HaeIII; fragment lengths (kb) are shown), which we have sequenced to delineate the organization of the rRNA genes. For those portions of the LSU rRNA gene complexes discussed in this study, the sequencing strategy is summarized at the bottom of the figure. Some sequence information was obtained from restriction sites (XbaI) within the cloning vector. Overlapping sequence data were obtained for each restriction site. The organization of the SSU rRNA gene complex is from Ref. 24.
Fig. 5. *Primary sequence at the ends of the LSUα,LSUβ rRNA gene complex from the two ends of the* T. pyriformis mitochondrial genome. Dots represent positions in the inserted repeat at which the sequence in the right end (R) of the genome is identical to the sequence in the left end (L). The termini of the LSUα, LSUβ, tRNA^Le^ (Le), tRNA^Ty^ (Ty), and tRNA^Me^ (Me) genes are indicated. Boxed positions represent identical nucleotides in the upstream conserved sequence (UCS). The termination codons of the upstream open reading frames from the left and right ends of the genome are overlined and underlined, respectively. The position of the sequence heterogeneity in the LSUα rRNA genes is denoted by arrowheads.

region (residues 1–525) of *E. coli* LSU (23 S) rRNA. LSUα can associate with LSUβ by means of base-pairing interactions that define four phylogenetically conserved helices (αa, bbc, ccc, dd'; Fig. 6) that are potentially present, with minor variations, in all LSU rRNAs (except, apparently, those of animal mitochondria). Virtually the same secondary structure, involving the same base-pairing interactions, can be constructed between the *P. primaurelia* mitochondrial “5.8 S-like” and 20 S rRNAs (30), as well as between the C. reinhardtii chloroplast 7 S and 23 S rRNAs (29). These base-pairing interactions are supported by the fact that they are flanked by highly conserved primary sequence motifs: e.g. a and b are always separated by GGA; d’ is preceded by A and followed by AA; and c’ is followed by AUAG.

In eukaryotic cytoplasmic LSU rRNAs, sequences a, b, and c are contributed by 5.8 S rRNA (see Fig. 2), whereas helix dd’ is a base-pairing interaction between two parts of the 28 S rRNA sequence. Residues 133–211 in LSUα, therefore, correspond to positions that are normally contained within eukaryotic cytoplasmic 28 S rRNA, not 5.8 S rRNA. Residues 47–132 in LSUα show little homology (either in primary sequence or potential secondary structure) with the corresponding residues (265–352) in *P. primaurelia* mitochondrial “5.8 S-like” rRNA, and *T. pyriformis* LSUα and *P. primaurelia* “5.8 S-like” rRNA are about 40 residues shorter than *E. coli* 23 S rRNA in this region; in neither case can the mitochon-
To understand the complex interplay between the large ribosomal RNA genes in Tetrahymena mitochondria, let's delve into the details. The figure illustrates potential secondary structure interactions between the LSUα and LSUβ rRNAs of Tetrahymena pyriformis mitochondria. The secondary structure is a precise replica of that proposed for the first 525 residues of Escherichia coli LSU (23S) rRNA. Dotted regions denote E. coli motifs that cannot be constructed with the T. pyriformis sequence because of major differences in sequence and/or length in these regions. Solid squares denote deletions in the T. pyriformis sequence. Open circles specify that portion of the E. coli secondary structure that is effectively absent because it represents the discontinuity between the LSUα and LSUβ rRNAs. Numbering is the same as in the legend to Fig. 2.

Residues 212-280 in LSUα are localized to a region that is highly variable in secondary structure models of LSU rRNA (e.g. Ref. 8; region D1); however, part of this region (positions 228-276) does have a secondary structure counterpart in E. coli 23S rRNA (E. coli coordinates 295-342). This structure is supported by numerous compensating base changes between the T. pyriformis LSUα and P. primaurelia "5.8 S-like" rRNA sequences, such that the proposed pattern of base pairing in this region is maintained. It is interesting that this region corresponds precisely to the 3S RNA species identified by Rochaix and Darlix as part of C. reinhardtii chloroplast LSU rRNA.

These comparisons demonstrate that T. pyriformis mitochondrial LSUα is the structural equivalent of the first 343 residues of E. coli 23S rRNA and that it has the potential to interact with LSUβ by base pairing in a manner expected of LSU rRNA 5'-terminal regions, generating a secondary structure equivalent to that of more conventional LSU rRNAs.

Transfer RNA Sequences—Transfer RNA genes identified by DNA sequence analysis (Fig. 5) were assigned their coding specificities on the basis of their anticodon sequence and the universal genetic code. Northern hybridizations (Fig. 1, B and C) showed that the putative tRNALeu and tRNA^Met genes give rise to stable transcripts. The predicted secondary structure of the tRNA^Leu is conventional (Fig. 7) whereas the tRNA^Met sequence is unusually short and its predicted secondary structure (Fig. 7) is atypical in the dihydrouridine stem-loop region.

We had previously isolated tRNA^Ty^r and reported hybridization of the purified species to both ends of T. pyriformis mitochondrial DNA (36). The right end-specific signal was to the same ClaI/HaeIII fragment whose sequence we have now determined and which does not contain a tRNATyr sequence. This signal was probably due to cross-hybridization of tRNA^Ty^r to the tRNA^Leu coding sequence. The tyrosine and leucine tRNA sequences share 51% overall homology, including a block of 18 nucleotides with 15 perfect matches. Our present sequence analysis verifies that only the left end of the genome encodes a tRNA^Ty^r.

Comparison of the Mitochondrial LSU rRNA Gene Complexes in T. pyriformis and P. aurelia—Seilhamer et al. (30) reported that the genes for the LSUα and LSUβ homologs in
The restriction fragment used by Seilhamer et al. (30) in their study can assume virtually the same secondary structure (Fig. 7). Comparison of band intensities suggests that the LSUa (TPY) transcript and its LSUa coding sequences differ at position 182, one copy having a G and the other an A at this position (Fig. 3A). This suggested the existence of two subspecies of the LSUa transcript. Although their relative levels are difficult to estimate, comparison of band intensities suggests that the LSUa(A) species is considerably more abundant than the LSUa(G) species in the steady state population of the one RNA preparation we have analyzed.

DNA sequencing of the two LSUa genes revealed that the LSUa coding sequences differ at position 182, one copy having a G and the other an A at this position (Fig. 3, B and C). Taken together, these data indicate that both copies of the LSUa gene are expressed. Since our Northern analyses have identified a putative precursor transcript carrying both LSUa and LSUP sequences, we tentatively identify it as a common precursor to the LSUa and LSUP rRNAs. This putative precursor must also carry in it the tRNALeu sequence, if we are considering the analogous precursor from the left-hand copy of the gene complex) is included in it. We do not know whether this putative precursor is a primary transcript.

The probe specific for LSUa and tRNA^Leu (and a short stretch of LSUa) also appears to hybridize to the 3.1-kb transcript (lane B), although this is difficult to show by Northern analysis because the signal generated by the abundant LSUa transcript masks any fainter signals in the same region of the gel. This probe also hybridizes to the tRNA^Leu, although as noted above, we have observed what appears to be cross-hybridization between tRNA^Leu and tRNA^Tyr sequences.

Sequence Heterogeneity in the LSUa Transcript and Its Genes—RNA sequence analysis of LSUa revealed sequence heterogeneity at one internal position (nucleotide 182), where the sequence consisted of both A and G (Fig. 3A). This suggested the existence of two subspecies of the LSUa transcript. Although their relative levels are difficult to estimate, comparison of band intensities suggests that the LSUa(A) species is considerably more abundant than the LSUa(G) species in the steady state population of the one RNA preparation we have analyzed.

Northern Hybridization Analysis of Transcripts from the LSUa-LSUβ Gene Complex—Fig. 1 shows the results of a Northern hybridization analysis using DNA probes from the right-hand copy of the LSUa-LSUβ gene complex. The probe specific for LSUa and tRNA^Leu (C) hybridized to the 280-nucleotide band representing the mature LSUa transcript as well as to the LSUa band, as predicted (lane C). In addition, a faint band, estimated to be 3.1 kb in length, was observed. Based on the size and relatively low abundance of this transcript, we tentatively identify it as a common precursor to the mature LSUa and LSUP rRNAs. This putative precursor must also carry in it the tRNA^Leu sequence. We have no evidence to show that the tRNA^Leu sequence, if we are considering the analogous precursor from the left-hand copy of the gene complex) is included in it. We do not know whether this putative precursor is a primary transcript.

The probe specific for LSUβ and tRNA^Leu (and a short stretch of LSUa) also appears to hybridize to the 3.1-kb transcript (lane B), although this is difficult to show by Northern analysis because the signal generated by the abundant LSUβ transcript masks any fainter signals in the same region of the gel. This probe also hybridizes to the tRNA^Leu, although as noted above, we have observed what appears to be cross-hybridization between tRNA^Leu and tRNA^Tyr sequences.

Sequence Heterogeneity in the LSUa Transcript and Its Genes—RNA sequence analysis of LSUa revealed sequence heterogeneity at one internal position (nucleotide 182), where the sequence consisted of both A and G (Fig. 3A). This suggested the existence of two subspecies of the LSUa transcript. Although their relative levels are difficult to estimate, comparison of band intensities suggests that the LSUa(A) species is considerably more abundant than the LSUa(G) species in the steady state population of the one RNA preparation we have analyzed.

DNA sequencing of the two LSUa genes revealed that the LSUa coding sequences differ at position 182, one copy having a G and the other an A at this position (Fig. 3, B and C). Taken together, these data indicate that both copies of the LSUa gene are expressed. Since our Northern analyses have identified a putative precursor transcript carrying both LSUa and LSUβ sequences (Fig. 1C), we conclude that both copies of the gene complex (LSUβ-tRNA^Leu-LSUa) are expressed. This is the first direct demonstration of sequence heterogeneity within and transcription from both copies of an inverted repeat. Inverted repeats containing rRNA genes are present in some mitochondrial and most known chloroplast genomes.

Open Reading Frames Upstream of LSU and SSU rRNA Genes—Sequences upstream of both copies of the inverted repeat contain open reading frames (ORFs), but so far we have been unable to find homology between the derived amino acid sequences, if we are considering the analogous precursor from the left-hand copy of the gene complex) is included in it. We do not know whether this putative precursor is a primary transcript.

The probe specific for LSUa and tRNA^Leu (and a short stretch of LSUa) also appears to hybridize to the 3.1-kb transcript (lane B), although this is difficult to show by Northern analysis because the signal generated by the abundant LSUa transcript masks any fainter signals in the same region of the gel. This probe also hybridizes to the tRNA^Leu, although as noted above, we have observed what appears to be cross-hybridization between tRNA^Leu and tRNA^Tyr sequences.

Sequence Heterogeneity in the LSUa Transcript and Its Genes—RNA sequence analysis of LSUa revealed sequence heterogeneity at one internal position (nucleotide 182), where the sequence consisted of both A and G (Fig. 3A). This suggested the existence of two subspecies of the LSUa transcript. Although their relative levels are difficult to estimate, comparison of band intensities suggests that the LSUa(A) species is considerably more abundant than the LSUa(G) species in the steady state population of the one RNA preparation we have analyzed.

Northern Hybridization Analysis of Transcripts from the LSUa-LSUβ Gene Complex—Fig. 1 shows the results of a Northern hybridization analysis using DNA probes from the right-hand copy of the LSUa-LSUβ gene complex. The probe specific for LSUa and tRNA^Leu (C) hybridized to the 280-nucleotide band representing the mature LSUa transcript as well as to the LSUa band, as predicted (lane C). In addition, a faint band, estimated to be 3.1 kb in length, was observed. Based on the size and relatively low abundance of this transcript, we tentatively identify it as a common precursor to the mature LSUa and LSUP rRNAs. This putative precursor must also carry in it the tRNA^Leu sequence. We have no evidence to show that the tRNA^Leu sequence, if we are considering the analogous precursor from the left-hand copy of the gene complex) is included in it. We do not know whether this putative precursor is a primary transcript.

The probe specific for LSUa and tRNA^Leu (and a short stretch of LSUa) also appears to hybridize to the 3.1-kb transcript (lane B), although this is difficult to show by Northern analysis because the signal generated by the abundant LSUa transcript masks any fainter signals in the same region of the gel. This probe also hybridizes to the tRNA^Leu, although as noted above, we have observed what appears to be cross-hybridization between tRNA^Leu and tRNA^Tyr sequences.

Sequence Heterogeneity in the LSUa Transcript and Its Genes—RNA sequence analysis of LSUa revealed sequence heterogeneity at one internal position (nucleotide 182), where the sequence consisted of both A and G (Fig. 3A). This suggested the existence of two subspecies of the LSUa transcript. Although their relative levels are difficult to estimate, comparison of band intensities suggests that the LSUa(A) species is considerably more abundant than the LSUa(G) species in the steady state population of the one RNA preparation we have analyzed.

DNA sequencing of the two LSUa genes revealed that the LSUa coding sequences differ at position 182, one copy having a G and the other an A at this position (Fig. 3, B and C). Taken together, these data indicate that both copies of the LSUa gene are expressed. Since our Northern analyses have identified a putative precursor transcript carrying both LSUa and LSUβ sequences (Fig. 1C), we conclude that both copies of the gene complex (LSUβ-tRNA^Leu-LSUa) are expressed. This is the first direct demonstration of sequence heterogeneity within and transcription from both copies of an inverted repeat. Inverted repeats containing rRNA genes are present in some mitochondrial and most known chloroplast genomes.
acid sequences of these ORFs and known proteins. Interestingly, both ORFs terminate in multiple TAA codons immediately prior to the inverted repeat (Fig. 5). Our preliminary Northern hybridizations have identified a stable transcript from the ORF upstream of the right-hand copy of the inverted repeat. We have recently published the sequence upstream of the SSU rRNA gene complex (24) and have subsequently identified an ORF encoding part of cytochrome oxidase sub-unit II in this sequence. This ORF also ends in TAA just prior to the genes encoding the SSU rRNA. Thus, TAA appears to be used for termination in T. pyriformis mitochondria. Interestingly, some nuclear genes of T. thermophila deviate from the universal code by using TAA to code for glutamine (40).

**DISCUSSION**

In this report, we show that the *T. pyriformis* mitochondrial LSU rRNA is a discontinuous molecule, consisting of two separate RNA components, which we term “α” (280 nucleotides in length) and “β” (approximately 2450 nucleotides in length). Primary sequence comparisons with other LSU rRNAs as well as secondary structure modeling prove that LSUα is the 5’-terminal portion and that LSUβ represents the remaining 3’-portion of this LSU rRNA. Thus, LSUα is a “5.8 S-like” rRNA, although it is considerably longer (280 nucleotides) than conventional 5.8 S rRNAs (approximately 165 nucleotides). The LSUα-LSUβ discontinuity occurs further downstream from the typical eukaryotic 5.8 S-28 S rRNA discontinuity, and it corresponds to a region that shows pronounced primary and secondary structural variation in broad phylogenetic comparisons. Although most known LSU rRNAs are covalently continuous in this region, the mitochondrial LSU rRNA of *P. primaurelia*, a ciliate closely related to *T. pyriformis*, contains a break in the identical position (Figs. 2 and 8). Of additional note is the occurrence of a break in virtually the same position in a distantly related system, the chloroplast LSU rRNA of *C. reinhardtii* (Figs. 2, 6, and 8).

Our analysis of the coding sequences for LSUα and LSUβ led to the surprising result that LSUα, the 5’-terminal portion of the LSU rRNA, is encoded downstream of LSUβ, the two genes being separated by a gene for tRNA<sub>Leu</sub> (Figs. 4, 5, and 8). This is the first reported example in which a discontinuity in a rRNA is associated with a rearrangement of the coding sequences, resulting in an organization that deviates from the conventional 5′ to 3′ order of rRNA domains at the DNA level. As such rearrangements have not been observed previously, it appears that the conventional organization is maintained by strong selection, perhaps due to constraints involving coupled rDNA transcription, post-transcriptional processing, and assembly of the ribosomal particle. The location of the discontinuity in the *T. pyriformis* mitochondrial LSU rRNA in a region known to be involved in ribosome assembly in some other systems (41-44) may indicate that these processes are largely uncoupled in the *T. pyriformis* mitochondrion, permitting gene rearrangements of the type described here.

Some insight into the molecular events that gave rise to the rearrangement of the two components of the LSU rRNA gene in *T. pyriformis* mitochondria may be obtained by comparison with the homologous genes in *P. primaurelia* mitochondrial DNA (Fig. 8). It appears that the discontinuous LSU rRNA was already present in the common ancestor of these two organisms, whereas the rearrangement occurred in the *T. pyriformis* lineage after the *T. pyriformis-P. primaurelia* divergence. Whatever the exact mechanism of rearrangement,
it may have been associated with the duplication event that gave rise to the inverted repeat (containing these genes) in the *T. pyriformis* mitochondrial genome. The rearrangement event also involved tRNA genes, the likely unit of transposition being 5'-((LSUα)-(tRNAMet)-3'), so that the overall pattern of tRNA sequences in the gene complexes was maintained. This is particularly striking since the rearrangement in *T. pyriformis* resulted in the loss of a tRNA<sub>Met</sub> gene from the LSUα-LSUβ junction, with the new LSUβ-LSUα junction containing a tRNA<sub>Met</sub> gene instead (the origin of the tRNA<sub>Met</sub> gene is unclear, as it does not appear to be present in or near the LSU RNA gene complex in *P. primaurelia* (30)). The similar organization of tRNA sequences in the mitochondrial LSU rRNA gene complexes of these two organisms suggests a functional role for these sequences, perhaps in the processing of precursor transcripts, as is the case in animal (45-47) and some fungal (48, 49) mitochondria.

Our sequence analyses have defined the inverted repeat of the *T. pyriformis* mitochondrial genome as the unit 5'-((UCS)-(LSUβ)-(2-bp spacer)-(tRNA<sub>Met</sub>)-(10-bp spacer)-(LSUα))-3'. The two versions of the upstream conserved sequence, UCS, differ only by a single base substitution and a 2-bp deletion in the left-hand sequence relative to the right-hand sequence. We consider the UCS a potential signal in the regulation of LSU rRNA synthesis in this system, for two reasons. First, its location precisely at the 5'-end of the gene complex suggests a potential regulatory role in either transcription initiation or post-transcriptional processing, depending on whether the putative precursor represents a primary transcript or is in turn derived from longer (undetected) transcripts. Second, the slight sequence differences in the two versions of the UCS may be expected to provide a basis for differential expression of the two copies of the LSU rRNA gene complex, which could account for our observation that the two LSUα subspecies are present in unequal amounts.

Acknowledgments—We wish to thank Gregg Morin for communicating data prior to publication and Lisa Laskey for typing this manuscript.

REFERENCES