Detection and cDNA cloning of H-strand mitochondrial regulatory region RNAs in cultured human cells and human tissues

Noboru Nakamichi, Mamiko Ito, Takuji Maeda & Toshiharu Matsumura

Meiji Milk Products Co., Cell Technology Center, Meiji Institute of Health Sciences and RITE Meinyu Branch Laboratory, Odawara, Japan 250-0862

(Received 16 March 1999; accepted 23 November 1999)

Key words: cDNA, differentiation, gene expression, growth, human cells, human tissues, mitochondrial DNA, mitochondrial RNA, polyadenylated RNA

Abstract

In the mitochondrion, essential genetic elements for replication and transcription are mostly housed within a short segment of its DNA located between tRNA^{Phe} and tRNA^{Pro} genes, which is called mitochondrial regulatory region (mrr). RNAs are known to be transcribed from mrr, the structures and the functions of which are yet to be fully characterized. We detected ca. 1.3 kb H-strand transcripts of mrr (mrrH-RNAs), and 0.2 kb L-strand transcripts of mrr (mrrL-RNAs) in various human cultured cells and tissues using double stranded mrrDNA probes. The steady state levels of mrrL-RNAs were generally high in cultured cells, while they varied among tissues. On the other hand, the levels of mrrH-RNAs varied among tissues and among cultured cells. A tendency was observed in these cells and tissues that a high level of mrrL-RNA is associated with cell proliferation, and a high level of mrrH-RNA with differentiation. Several cDNA clones to 1.3 kb mrrH-RNA were obtained from human skeletal muscle polyadenylated RNAs. The 5' terminus of the 1.3 kb RNA was determined to be at nucleotide position 15953 which is immediately downstream of tRNA^{Thr} sequence. Polyadenylation site for most of the clones was demonstrated to be at nucleotide position 576 which is immediately upstream of tRNA^{Phe} sequence. The longest cDNA insert obtained was 1177 bps long spanning from nucleotide positions 15969 to 576 which could code for a peptide of 76 amino acids. The cDNAs isolated here are the first cDNA clones reported to human mrrH-RNAs. These results, together with previous results, further substantiate that polyadenylated mrrH- and mrrL-RNAs are commonly present at varying levels among human tissues and cells. The 3' end sequences of the cloned mrrH-cDNA provides with insights into the mechanisms of transcription termination. The cDNA clones will provide tools to further the study of the function of mrr RNAs.

Abbreviations: mrr, mitochondrial regulatory region; mrrL and mrrH, mitochondrial regulatory region L- and H-strand, respectively; tRNA^{Phe}, tRNA^{Thr} and tRNA^{Pro}, transfer RNAs for phenylalanine, threonine and proline, respectively; bp, base pair; kb, kilobase; CSB, conserved sequence block; HSP1, heavy strand promoter for the ribosomal RNAs; HSP2, heavy strand promoter for the whole H-strand transcription unit; CoxIII, cytochrome oxidase III; GAPDH, glyceraldehyde-3-phosphate dehydrogenase

Introduction

The mitochondrion, the energy-production organelle in animal cells, is one of the attractive research objectives as the control of cellular energy level is a key technology in production biotechnology. Yet at present, the current state of art of cellular energy-

production control is far from the desired level, due in part to a paucity of genetic-engineering tools as well as the insufficient knowledge of the mechanistic aspects for mitochondrial biogenesis.

The mammalian mitochondrial DNA (mtDNA) is a ca. 16 kbp closed circular double stranded molecule consisted of two DNA strands called heavy (H)

strand and light (L) strand, respectively (Anderson et al., 1981; Bibb et al., 1981; Gadaleta et al., 1989). The short segment of mtDNA between tRNA^{Phe} and tRNA^{Pro} genes, which will be referred to as mitochondrial regulatory region, or mrr hereafter, plays essential roles in mtDNA replication and gene expression (Walberg and Clayton, 1981; Saccone et al., 1987; Sbisà et al., 1997). This region contains L-strand promoter (LSP) sequence and two H-strand promoter (HSP) sequences, i.e., HSP1 for rRNA transcription and HSP2 (nucleotide position 561) for the transcription of whole H-strand polycistronic transcription unit, respectively.

It also contains origin of H strand replication (OriH), multiple protein binding sequence elements, Displacement loop (D-loop) region, and several conserved sequence blocks (CSBs) (Crews et al., 1979; Walberg and Clayton, 1981, 1983; Montoya et al., 1983; Chang and Clayton, 1984, 1985; Saccone et al., 1987, 1991; Sbisà et al., 1997).

The mrr is of interest not only because it contains these functional sequence elements, but because it transcribes RNAs: the most well known among the transcripts are ca. 0.2 kb heterogeneous L strand transcripts with their 5' end at LSP and their 3' ends in CSBs or at OriH (Ojala et al., 1981; Chang and Clayton, 1985; Chang et al., 1985; Sbisà et al., 1990; Noda, 1992; Nakamichi et al., 1998a). Typically, they are polyadenylated RNA species, and are attributed to prime H-strand DNA replication. Additional regulatory roles have also been suggested to them (Walberg and Clayton, 1981; Chang and Clayton, 1984; Nakamichi et al., 1998a).

Transcripts of mrrH-strand have also been reported in rat, mouse and human (Sbisà et al., 1990; Vijayasarathy et al., 1995), but less intensely. The transcripts are again polyadenylated, and cover the entire or a part of the mrr-sequence spanning from the 3' end of tRNA^{Thr} to the 5' end of tRNA^{Phe} genes. Special interests with mrrH-RNAs are in the variation of their contents in different tissues, suggesting that they are involved in the regulatory mechanisms of cell growth and mitochondrial biogenesis (Vijayasarathy et al., 1995). Studies of mrrRNAs, particularly of mrrH-RNAs, are still insufficient. Intensive sequence analyses have been made only for mouse mrrH-RNA. Knowledge about the transcription levels is limited. The function of the transcripts has little been clarified.

During the course of observation of human mrrL-RNAs in Northern blot analyses using a double stranded cDNA probe to mrrL-RNA, we noted that the

level of the two major mrr-RNAs vary greatly among human cultured cells and tissues (Nakamichi et al., 1998a, b). We cloned and sequenced several cDNAs to mrrH-RNAs, determined the sequence of their 5' ends using a primer extension method, and predicted the largest ORF available in the cDNA sequence. These results confirmed and extended the previous observations on mrrH-RNAs by others (Sbisà et al., 1990; Vijayasarathy et al., 1995), and added new information on the entire nucleotide sequence of human mrrH-RNAs, on the mechanisms of transcription termination, and on the variable balance of mrrL-RNA level and mrrH-RNA level in tissues and cultured cells.

Materials and methods

Cell cultures and RNA preparations

Seven cultured human cell lines were utilized in this study as the sources of total RNA isolation. They were HeLa S3 cervical carcinoma cells, HepG2 hepatocellular carcinoma cells, IMR-90 normal human fibroblasts, KYM rhabdomyosarcoma cells, AT5BIVA Ataxia telangiectasia fibroblasts (Group D) immortalized using SV40, NB-1 neuroblastoma cells, and KATOIII stomach ring cell carcinoma cells. Each of these cell lines was grown in Eagle's MEM medium (GIBCO BRL, Grand Island, NY) supplemented with 10 to 15% fetal bovine serum at 37 °C in a humidified atmosphere containing 5% CO₂.

Total RNA was isolated from a culture using RNazol B reagent (Biotex Laboratories, Houston, Texas), and the poly (A⁺) RNAs were separated using Oligo (dT)-Cellulose type 7 affinity chromatography (Amersham Pharmacia Biotech, Tokyo).

Other human RNA samples used in the present study were purchased from CLONTECH Laboratories, Inc. (Palo Alto, CA). Names and catalog numbers of the total RNAs unless otherwise indicated are as follows; HL-60 human erythroleukemia cell line (Cat. #6530-1, poly (A⁺) RNA, ATCC CCL 240), skeletal muscle (Cat. #64033-1, 25-yr-old, male; Cat. #6541-1, 45-yr-old, male; Cat. #6541-2, poly (A⁺) RNA, 45-yr-old, male; Cat. #6541-48604, poly (A⁺) RNA, 17-61-yr-old, male/female pool), Heart (Cat. #64025-1, 30-yr-old, male), Uterus (Cat. #64029-1, 16-55-yr-old, female pool), Fetal brain (Cat. #64019-1, 20-26-wk-old, male/female pool), Adult brain (Cat. #64020-1, 60-yr-old, male), Salivary (Cat. #64026-1, 6-64-yr-old, male/female pool), Mammary gland

Figure 1. Northern blot analyses of RNAs obtained from human tissues and cultured cells with human *mrr/HeLa-1* cDNA probe. The double stranded human *mrr/HeLa-1* cDNA was labeled with ^{32}P -dCTP and used for hybridization with RNAs indicated below as described under the Materials and method section. Panel A: Northern transfer was carried out as described in the Materials and method section. Each lane was loaded with 20 μg of total RNAs with exception of HL-60 which was loaded with 5 μg of poly (A^+) RNA. Lane a: HeLa, lane b: KYM, lane c: AT5BIVA, lane d: NB-1, lane e: KATO III, lane f: HepG2, lane g: IMR90, lane h: HL-60 (Cat. #6530), lane i: skeletal muscle (Cat. #64033-1), lane j: Heart (Cat. #64025-1), lane k: Uterus (Cat. #64029-1), lane l: Fetal brain (Cat. #64019-1), lane m: Adult brain (Cat. #64020-1), lane n: Salivary gland (Cat. #64026-1), lane o: Mammary gland (Cat. #64037-1), lane p: Lung (Cat. #64023-1), lane q: Fetal liver (Cat. #64018-1), lane r: Adult liver (Cat. #64022-1), lane s: Pancreas (Cat. #64031-1), lane t: Kidney (Cat. #64030-1), and lane u: Placenta (Cat. #64024-1). Size markers indicated on the left hand side are from a 0.2–10 kb RNA ladder (Novagen, Madison WI, U.S.A.).

Panel B: Human multiple tissue northern (MTN) blot II (Cat. #7759-1) containing approximately two micrograms of poly (A^+) RNA per lane was purchased from CLONTECH Laboratories, Inc. (Palo Alto, CA) and processed as indicated by the manufacturer with the ^{32}P -dCTP labeled human *mrr/HeLa-1* cDNA probe. Lane a: Spleen, lane b: Thymus, lane c: Prostate, lane d: Testis, lane e: Ovary, lane f: Small intestine, lane g: Colon (mucosal lining), lane h: Peripheral blood leukocyte. Size markers indicated on the left hand side are from a 1.35–9.5 kb RNA ladder (Gibco BRL).

Figure 2. Northern blot hybridizations of RNAs from human tissues and cultured cells with mitochondrial and nuclear gene probes. The plasmid DNAs, pUCGAP₅₅₀, pUCIB43, pUCMRP4 and pUCCoxIII, were labeled with ³²P-dCTP and used for hybridization with RNAs indicated below. Panel A: Northern transfer was carried out as described in the Materials and methods section. Two labeled probes, pUCGAP₅₅₀, and pUCIB43 were hybridized simultaneously. Each lane was loaded with 20 µg of total RNAs with exception of HL-60 which was loaded with 5 µg of poly (A⁺) RNA. Lane a: IMR-90, lane b: Lung (Cat. #64023-1), lane c: Kidney (Cat. #64030-1), lane d: Skeletal muscle (Cat. #64033-1), lane e: HeLa f: AT5BIVA, lane g: NB-1, lane h: KATO III, lane i: HepG2, lane j: HL-60 (Cat. #6530-1). Size markers indicated on the left hand side are from a 1.35–9.5 kb RNA ladder (Gibco BRL). Panel B: Labeled pUCCoxIII probe was rehybridized to the same filter used in the Panel A which had been stored for two months to reduce the previously hybridized probe radioactivity to the background level. Panel C: Labeled pUCMRP4 probe was rehybridized again to the same filter used in the Panel B which had been stored for 11 months to reduce the previously hybridized probe radioactivity to the background level.

(Cat. #64037-1, 16–43-yr-old, female), (Cat. #64023-1), Fetal liver (Cat. #64018-1, 26-wk-old, male), Adult liver (Cat. #64022-1, 40-yr-old, male), Pancreas (Cat. #64031-1, 16–40-yr-old, male/female pool), Kidney (Cat. #64030-1), and Placenta (Cat. #64024-1, 26-yr, female).

Northern blot analysis

RNA samples were electrophoresed in a 1.5% agarose gel (Takara Shuzo Co. Shiga, Japan) containing formaldehyde by a reported procedure (Maniatis et al., 1982). After electrophoresis, RNA was transferred from the gel to Hybond N⁺ nylon membrane (Amersham Pharmacia Biotech, Tokyo). The membrane was prehybridized for 2–4 h at 50 °C and then hy-

bridized using ³²P-labeled DNA probe at 50 °C for 16 h. The membrane was then washed and exposed to a film (Biomax MS, Kodak) for 16 h to several days at –80 °C. A nuclear specific gene probe pUCGAP₅₅₀ containing a 550 bp Hind III/Xba I fragment of the human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA (Tso et al., 1985), and a mitochondrial specific gene probe pUCCoxIII containing a portion of human cytochrome oxidase III gene were described previously (Nakamichi et al., 1998a). A nuclear specific gene probe, pUCMRP4, containing mitochondrial RNA processing (MRP) gene, was constructed using pUC19 plasmid and a 271 bp of the MRP RNA gene fragment. This DNA fragment was obtained from HeLa S3 cells by PCR using primers, 5'CAAGCACGACTTCCGGACATAG3'

and 5'AGCCGCGCTGAGAATGAGCCCC3', based on the published human MRP RNA gene sequence (Topper and Clayton, 1990). A double stranded IB-43 cDNA probe containing mtDNA segment between the nucleotide positions 366 and 237, was provided by Dr. A. Noda (pUCIB-43, Noda, 1992). The other cDNA probe, mrr/HeLa-1, was obtained from 0.2 kb mrrL-RNA of HeLa cells, and contains extended 3'- and 5'-sequences in addition to IB43 cDNA sequence obtained from IMR90 cells (Nakamichi et al., 1998a, Figure 4).

Plasmid probes were labeled using [α^{32} P]dCTP (3000 Ci/mmol, Amersham), and a Random Primer DNA labeling kit (Takara Shuzo Co. Shiga, Japan).

cDNA library construction and screening

Human skeletal muscle poly (A⁺) RNA (Cat. #6541-1, CLONTECH Lab. Inc.) was used as the starting material for the construction of a cDNA library. The cDNA was synthesized using a cDNA synthesis system (GIBCO BRL Life Technologies, Tokyo) by following the instruction manual, and inserted into the λ gt10 vector and packaged using Giga pack II Gold (STRATAGENE, La Jolla, CA). Forty-five positive phage clones were selected from an approximately 5×10^5 phages using a 32 P-labelled probe of double stranded mrr/HeLa-1 cDNA encoding the mrr-sequence between the nucleotide positions 229 and 406 (Nakamichi et al., 1998a).

To estimate the size of the cDNA inserts in the plaque-purified positive phage DNAs, each cDNA insert was PCR-amplified using a left arm DNA segment (5'CTGCTTCTCATAGAGTCTTG3') and a right arm DNA segment (5'ATACATATACGGTTCTCTCCAGAG3') of the phage vector as primers. The sizes of the amplified cDNA inserts were estimated by Southern blot hybridization with a 32 P-end labeled synthetic oligonucleotide (5'AAGATAAAATTTGAAATC3') corresponding to the mrr sequence between the nucleotide positions 389 and 406 (Nakamichi et al., 1998b).

DNA sequencing

For the sequence analyses, the above cDNA inserts were subcloned into pUC19 cloning vector, and the recombinant plasmids were purified by a published method (Hattori and Sakaki, 1986) or by the QIAGEN plasmid kit (QIAGEN Inc. Valencia, CA). M13/M4 universal and M13/RV reversal sequencing primers were purchased from Takara Shuzo Co. Other primers shown in the Figure legends were synthesized using

an automated DNA synthesizer (Perkin Elmer, Chiba, Japan). An automated DNA sequencer was used to determine the nucleotide sequences of the cDNAs (Perkin Elmer, Chiba, Japan).

Primer extension analysis

The 5'-end of the putative H-strand coded RNA was determined by primer extension analysis using a 5'-end-labeled synthetic primer corresponding to the mrr sequence between the nucleotide positions 16054 and 16035, i.e., Hu16054R. We used the AMV reverse transcriptase primer extension system essentially following the instruction manual provided by the manufacturer (Promega, Madison, WI). Either skeletal muscle poly (A⁺) RNA (5 μ g, Cat. #6541-48604, 17-61-yr-old, male/female pool), or HepG2 poly (A⁺) RNA (3.4 μ g), was mixed with 1.3 ng of the labelled primer in a reaction mixture and annealed at 58 °C for 20 min. To this annealed mixture, AMV reverse transcriptase, dNTP, sodium pyrophosphate, and extension buffer were added and the primer was extended at 41.5 °C for 30 min. As a template for the molecular weight marker in the primer extension analysis, a pUC19 plasmid containing 481 bp mtDNA fragment amplified by the PCR method using two primers, 5'TGGCTTTGGAGTTGCAGTT3' and 5'CTACCCTTTTACCATCATTG3', i.e., plasmid pUCExt₄₈₁, was used. This 481 bp fragment had completely been sequenced for confirmation before it was used as the reference. The cDNAs extended with Hu16054R primer and the sequencing ladders as markers were fractionated on a 6% denaturing polyacrylamide gel containing 8M urea, and the dried gel was exposed to X-ray film for up to two days in order to visualize the primer extended products for size determination.

Results

Detection of 1.3 kb and 0.2 kb mrrRNAs

In a Northern blot analysis using a 32 P-labeled mrr/HeLa-1 cDNA probe, either a 1.3 kb size band or a 0.2 kb size band, or both, were detected in the RNAs obtained from all of the eight human cultured cell lines and the 21 human tissues (Figures 1A, B and 3). Hereafter, we regard the 0.2 kb band as representing ca. 0.2 kb mrrL-RNA with results of previous cloning and sequencing studies (Noda, 1992; Nakamichi et al., 1998a), and the 1.3 kb band as representing

mrrH-RNA as will be described later in details (see text in the section Results for Figure 5).

In general, the level of the 0.2 kb band was high in the cultured cells, while it varied among tissue cells (Figures 1A and B). On the other hand, the level of 1.3 kb band varied among cultured cells and among tissues (Figure 1A). As a whole, these cells and tissues were roughly classified into the following three groups: the first group with a high level 0.2 kb band and a low level 1.3 kb band included HeLa, KYM, AT5BIVA, NB-1, IMR-90, and mammary tissue, the second group with a high level 0.2 kb band and a high level 1.3 kb band included KATO III, HepG2, HL-60, heart, fetal brain, adult brain, salivary tissue, lung, fetal liver, adult liver, pancreas, kidney, and placenta, and the third group with a low level 0.2 kb band and a high level 1.3 kb band included skeletal muscle and uterus. It was particularly noted that 0.2 kb band was almost undetectable in skeletal muscle tissues (Figure 1A, lane i, Figure 2A, lane d, Figure 3, lane b). We also noted the presence of a 1.3 kb band at significant levels in the tissues of spleen, prostate, testis, ovary, small intestine, colon, and in peripheral blood cells (Figure 1B).

Variation in band intensities for 0.2 kb RNA was further studied by Northern analyses using ten selected RNA samples obtained from cell lines (IMR-90, HeLa, AT5BIVA, NB-1, KATO III, HepG2 and HL-60), and tissues (lung, kidney and skeletal muscle). Here, total RNAs were used, except for HL60 for which poly (A⁺) RNA was used. The RNAs were hybridized respectively with one probe for mrr transcripts (pUCIB43), two probes for nuclear gene transcripts, i.e., GAPDH (pUCGAP₅₅₀) and MRP RNA gene (pUCMRP4), and one probe for mitochondrial gene transcript, CoxIII (pUCCoxIII): Here, pUCIB43 and mrr/HeLa-1 are similar to each other, and we consider that they equally detect mrr transcripts. The probe for GAPDH was used as a standard for RNA content, and the probes for MRP RNA and CoxIII to examine if there is any correlation in transcription among these loci of nuclear and mitochondrial genomes, and mrr region.

Figure 2A shows results with the probes for mrrRNA (pUCIB43) and GAPDH (pUCGAP₅₅₀). The 0.2 kb RNA bands, which we regarded as due to pUCIB43 probe, were shown in all of these sample (Figure 2A, lanes a–c and e–j) except for skeletal muscle (Figure 2A, lane d), confirming the results in Figure 1 using mrr/HeLa-1 probe. The ca. 1.4 kb bands shown in IMR-90, HeLa and NB-1 in Figure 2A

Figure 3. Northern blot hybridization of RNAs from human tissues and cultured cells with human mrr/HeLa-1 cDNA probe. Northern transfer was carried out as described in the Materials and method section. The double stranded human mrr/HeLa-1 cDNA was labeled with ³²P-dCTP and used for hybridization with RNAs. Each lane was loaded with 20 µg of total RNAs with exception of poly (A⁺) RNA selected skeletal muscle sample which was loaded with two micrograms. Lane a: HeLa, lane b: Skeletal muscle (cat. #64033-1), lane c: Skeletal muscle poly (A⁺) (cat. #6541-2), lane d: KYM. Size markers indicated on the left hand side are from a 0.25–9.49 kb RNA ladder (Gibco BRL).

were considered to be due to pUCGAP₅₅₀ probe, since no such bands were shown in Figure 1, in which this probe was not used. Fairly comparable density of the ca. 1.4 kb GAPDH band in all of these lanes indicates that they each had been loaded with an approximately the same amount of RNA (Figure 2A). Apparently however, the 1.3 kb bands, which should appear due to pUCIB43 for mrrH transcripts and the 1.4 kb band for GAPDH were not clearly distinguished from each other in these cells and tissues.

When the same filter was rehybridized with the probe for CoxIII, a ca. 1 kb band was detected in every preparations examined, confirming that all of these ten RNAs contain mitochondrial Cox III transcripts (Figure 2B, lanes a–j). It was noted, however, that the ca. 1 kb band detected here and the 0.2 kb band detected using an mrrRNA probe as above were not equally correlated among these preparations: skeletal muscle RNA contained a high level of ca. 1 kb CoxIII band, in

contrast to its negligible content of 0.2 kb RNA (lane d of Figures 2A and B).

When the same filter was again rehybridized with the probe for MRP RNA, single band was detected in every RNA preparations except for that from HL-60 (Figure 2C). The lack of band in HL-60 was due to the fact that polyadenylated RNA preparation in place of total RNA was used here, while MRP RNA lacks a polyadenyl tail. In the cells and tissues in which the ca. 0.2 kb mrrRNA band was intense, MRP RNA band noted was also intense to some extent.

In addition to the above bands, the probe for mrr RNA (pUCIB-43) detected diffuse bands or smears larger than 2 kb in RNA preparations from most of tissues (Figure 1A, lanes h–k, m–p, and r–u; Figure 1B, lanes f and g; Figure 2A, lanes c–f and h–i; Figure 3, lane b). To test if the bands larger than 2 kb observed in these figures are polyadenylated RNAs or not, we performed Northern blot analysis using the total and poly (A⁺) RNAs of skeletal muscle, respectively. Here, bands larger than 2 kb were detected in the total RNA (Figure 3, lane b), while no corresponding bands in the poly (A⁺) RNAs (Figure 3, lane c), indicating that the bands larger than 2 kb were devoid of polyadenyl tail. These high molecular bands were therefore suggested to be heterogeneous species of D-loop region containing mitochondrial polycistronic transcripts, or simply due to purification difficulties in some of the tissues.

Isolation and sequencing of cDNA clones to 1.3 kb RNA

To further the structural analyses of 1.3 kb RNAs, we constructed a cDNA library from human skeletal muscle poly (A⁺) RNA, and cloned 45 cDNAs as described in the Materials and methods section. Ten clones among them contained ca. 1.3 kb cDNA insert, as detected by a combination of PCR and Southern blot methods as described, and were sequenced after they were subcloned into pUC19 cloning vector.

The characterization of nine clones out of ten are briefly shown in Table 1. All of these nine clones encoded H-strand sequence of mtDNA spanning between the tRNA^{Thr} and the tRNA^{Phe} genes. Five clones among them (SK 5A, SK 6A, SK 20A, SK 3B and SK 21C) contained the same segment of sequence in which its 5' end mapped to 17 nucleotide downstream to tRNA^{Thr} gene (15969 nucleotide position) and its 3' end mapped to immediately upstream to tRNA^{Phe} gene (576 nucleotide position) (Figures 4 and 5, Table 1). Three other clones slightly

shorter segments with heterogeneous 5' end positions while maintaining its 3' end at 576 nucleotide position (Table 1). We regarded these eight clones are independent because they contained different numbers of polyadenyl residues at their 3' ends (Table 1). Furthermore, these eight cDNAs shared the same 3' end position at 576. From these results, we concluded that these eight cDNAs represent one major species of polyadenylated RNA as detected in Northern blot with an mrr probe in ca. 1.3 kb band (Table 1, Figures 1–3).

We noted one different clone out of the nine clones, which contained cDNA with its 5' end position at nucleotide 15977 and its 3' end position at nucleotide 409. It should be mentioned here that a cDNA clone with its 3' end at nucleotide position 232 was previously described in human liver (Vijayasarathy et al., 1995). The present results together with the previous results indicate that several species of mrrH-strand transcripts are present in human (Table 1, Figure 4). One clone was excluded from the above sequence analyses, because it lacked polyadenyl residues at its 3' end.

The 5' end map position of the 1.3 kb RNA

Because the cloned ca. 1.3 kb cDNAs varied in their 5' end position, we determined the 5' end of the putative mrrH-strand RNA by a primer extension method, using skeletal muscle tissue poly (A⁺) RNA again, and HepG2 poly (A⁺) RNA. As shown in Figure 5A, for skeletal muscle and HepG2, the end-labeled Hu16054R primer yielded single major extension product of 102 nucleotides mapping to nucleotide position 15953 (Figure 5A, lanes 1 and 2). This position corresponds to the nucleotide immediately downstream of the 3' end of the tRNA^{Thr} gene (Figure 4) on the same strand. The primer extension reactions without poly (A⁺) RNA (Figure 5A, lane 3) or with a positive control RNA (Figure 5A, lane 4) produced expected negative and positive bands showing the authenticity of this analysis. Another primer extension study using HL-60 cell poly (A⁺) RNA resulted in the same size extension product as above (data not shown).

We concluded from the results of primer extension analyses that the majority of 1.3 kb RNA spans from the nucleotide positions 15953 to 576, and that they are most probably originated from primary polycistronic transcript by processing at 3' end of the tRNA^{Thr} gene (Figure 5B). However, further studies are necessary to clarify whether or not the tRNA^{Thr}-containing

Table 1. Mapping of human skeletal muscle cDNA clones corresponding to 1.3 kb mitochondrial RNA

cDNA designation	5' End position ^a	3' End position ^a	Insert size (bp) ^b	Poly (A) size (base)	Total cDNA size (bp)
Sk 5A	15969	576	1177	9	1186
Sk 6A	15969	576	1177	6	1183
Sk 20A	15969	576	1177	31	1208
Sk 3B	15969	576	1177	8	1185
Sk 21C	15969	576	1177	45	1222
Sk 9A	15977	576	1168	48	1216
Sk 19A	15981	576	1165	10	1175
Sk 18A	16114	576	1032	8	1040
Sk 8A	15977	409	1001	6	1007

^a Numericals indicate the nucleotide numbers according to Anderson et al. (1981) (EMBL AC#J01415).

^b cDNA sequences homologous to human mitochondrial DNA sequence determined by Anderson et al. (1981) (EMBL AC#J01415).

Figure 4. The genetic and transcriptional map of the human mitochondrial regulatory region. The map positions of the H-strand coded RNAs designated as mrrH-RNA and heterogeneous size class of the L-stranded coded RNAs designated as mrrL-RNA are presented. Human liver cDNA (Vijaysarathy et al., 1995), 7S RNA (Ojala et al., 1981), mrr/HeLa-1 (Nakamichi et al., 1998a) and IB43 (Noda, 1992) are also included as references. The map positions of the tRNAs are according to Anderson et al. (1981).

primary polycistronic transcript is contiguous with tRNA^{Phe} gene sequence.

Prediction of possible polypeptides coded by the 1.3 kb RNA

Eight ORFs coding the peptides with more than 2000 molecular weight were noted in the 1.3 kb RNA sequence. The longest potential ORF encodes a polypeptide of ca. 8600 molecular weight starting at the nucleotide position 16087 and extending to the nucleotide position 16314 (Figure 6; Anderson et al., 1981). Neither AGA nor AGG codons, which have been suggested to act as stop codons in mammals (reviewed in Attardi, 1985), was found within this ORF. As compared with the published sequence (Anderson et al., 1981), we noted nucleotide additions, deletions and point mutations including one at nucleotide 16223 in

ORF. This point mutation is a silent mutation which does not cause any change in amino acid sequence (Figure 6).

Isolation and nucleotide sequence determination of cDNA clones to 0.2 kb RNA

As shown in the above, the level of ca. 0.2 kb RNA was extremely low in comparison with ca. 1.3 kb RNA in the skeletal muscle (Figures 1–3). This, however, does not mean it lacks 0.2 kb RNA, since cDNAs to ca. 0.2 kb RNA were successfully isolated from the skeletal muscle tissue RNA. The sequence analyses of four such cDNA clones showed that they were all transcripts of the L-strand of the mrrDNA, polyadenylated and were mapped to the CSB containing region with heterogeneous 3' and 5' map positions, and that they

Figure 5. Determination of the 5' end of the mitochondrial regulatory region H-strand coded RNA. Panel A: the 5' terminus of the H-strand coded RNA was mapped by primer extension method which utilized γ -³²P-ATP end-labeled Hu16054R. The labeled primer was annealed with the sample RNAs to obtain extended products as described under the Materials and method section. Lane 1: 5 μ g of skeletal muscle Poly (A⁺) RNA (Cat. #48604), lane 2: 3.4 μ g of HepG2 Poly (A⁺) RNA, lane 3: without RNA sample, and lane 4: 2 ng of internal positive control RNA (87 bases of cDNA is the major product). The sequence ladder marked GACT was synthesized using Hu16054R primer and pUCExt481 DNA. The G* indicates that the original template sequence was confirmed to be G residue which differ from the reported sequence (Anderson et al., 1981). Panel B: Mapping of ca. 1.3 kb mrrH-RNA to the human mitochondrial regulatory region. The 5' terminus (nucleotide position 15969) of the H-strand coded cDNA was extended by a solid line to the terminus determined by primer extension method (nucleotide position 15953). The connected sequence named as ca. 1.3 kb mrrH-RNA was aligned to the standard sequence of mitochondrial RNA segment between the tRNA^{Thr} and tRNA^{Phe} genes (Anderson et al., 1981). The cloverleaf configuration was based on Crews et al. (1980) for the tRNA^{Phe} and tRNA^{Thr} genes, and on a computer assisted tRNA prediction software for tRNA^{Phe} gene (DNAsis, Hitachi Co., Tokyo).

Figure 6. Nucleotide sequences and the longest putative reading frame of 1.3 kb cDNA. Entire nucleotide sequence of the chimeric cDNA to 1.3 kb mrrH-RNA and the largest putative open reading frame were shown where the amino acids corresponding to individual codons are indicated as three-letter symbols. Nucleotide sequences were compared with the published mtDNA sequence (Anderson et al., 1981), where point mutations were indicated by the lower case letters, nucleotide additions by the underlined lower case letters, and nucleotide deletions by space bars. The rightward arrow with solid line above the sequence show the 3' end of the tRNA^{Thr} gene on the same strand. The leftward arrow with dashed line below the sequence show the entire tRNA^{Pro} gene on the antisense strand. Nucleotide positions, 15953 and 576, are the same as in Figures 4 and 5.

belong to the same ca. 0.2 kb *mrrL*-RNA as previously reported (Figure 4, ref. to Nalamichi et al., 1989a).

Discussion

In the results, we demonstrated that the polyadenylated RNAs of ca. 0.2 kb size and ca. 1.3 kb size were detected in Northern analyses using mitochondrial regulatory region (*mrr*) DNA probes. Sequence analyses of cDNA clones to these RNAs indicated that they are the transcripts, respectively, of the L-strand and H-strand of *mrr*DNAs. They were expressed at various levels in all human tissue cells and cultured cells examined. These results confirmed and extended the previous observations that L-strand transcripts of mitochondrial regulatory region (*mrrL*-RNA) and H-strand transcripts of mitochondrial regulatory region (*mrrH*-RNA) are present in mammalian cells and tissues (Sbisa et al., 1990; Vijayasarathy et al., 1995).

The results of primer extension analysis and nucleotide sequencing of the cDNA clones have identified a major species of human *mrrH*-RNAs as mapped to the non-coding region spanning from the 3' end of tRNA^{Thr} to the 5' end of tRNA^{Phe} genes, being a homologue of previously detected RNAs in the rat and mouse (Sbisa et al., 1990; Vijayasarathy et al., 1995). A particular interest here will be that the major *mrrH*-RNA terminates its 3' end, except for its polyadenyl tail, at immediately adjacent to the 5' end of tRNA^{Phe}. This may suggest that its 3' end is formed according to the tRNA punctuation model for the processing of polycistronic mitochondrial RNA transcript. Here, the model hypothesizes that the secondary structure formed by the tRNAs scattered throughout the mitochondrial polycistronic transcript provides the necessary signal for the transcript to be processed to liberate mitochondrial mRNA species (Manam and Van Tule, 1979; Montoya et al., 1981; Ojala et al., 1981). It has been known that the 3' end of a mitochondrial mRNA, which was processed adjacent to the 5' end of a tRNA, is either U or UA, except for its polyadenyl tail. In the present results for *mrrH*-RNA, however, the terminal except for its polyadenyl tail is C or CA, differing from the general rule. Moreover, if *mrrH*-RNA is terminated following the tRNA punctuation model, then it predicts that H-strand transcription should continue over one complete round of mtDNA. To the author's knowledge, the termination of polycistronic mitochondrial transcript has not been studied in details. Since the whole transcription unit of H-strand RNA is tran-

scribed from HSP2 (Montoya et al., 1983), *mrrH*-RNA may easily be recognized as the terminal portion of the whole transcription unit. However, the possibility has not completely rejected that the H-strand polycistronic RNA exceeds one complete round of mitochondrial genome. It is of particular interests, therefore, if there are any polycistronic RNAs extending beyond the base 576.

There is a minor cDNA, SK8A with its 3' terminus at nucleotide position 409, and another cDNA clone reported previously in liver (Vijayasarathy et al., 1995) at nucleotide position 232. Because there are no tRNA sequence around these nucleotide positions in *mrr*, these 3' end positions suggest that they terminate their 3' ends by alternative mechanisms rather than the tRNA punctuation mechanism.

It may be relevant to note that the minor *mrrH*-RNA with its 3' end at 409 (SK8A) contains a U-rich sequence, 5'UUUU3', in the region 12 bases downstream of the polyadenylation site of the *mrrH*-transcript. In addition, the putative polyadenylation signal 5'AATTAATTAATT3', suggested previously in human (Vijayasarathy et al., 1995), is located about 190 bases upstream of the polyadenylation site of the SK8A containing primary transcript. These two cis-acting elements could come closer to 40 bases when a computer predicted cloverleaf with a 160 bases segment flanked by the two cis-elements is introduced (Figures not shown). Since in the case of HTLV-1 mRNAs, where these two cis-acting elements are separated more than 250 bases apart from each other, cleavage and polyadenylation at a destined site occur by the formation of such a secondary structure that brings these two elements closer to about 30 bases (Ahmed et al., 1991).

Another alternative model of transcription termination and polyadenylation has been proposed in plant mitochondrial mRNAs in which the 3'-inverted repeats act as processing signals (Dombrowski et al., 1997). We did not detect inverted repeat sequences in the 3' portion of the major *mrrH*-RNA, however. We believe therefore that the 3' end formation by the secondary structure model is the likely mechanism in SK8A *mrrH*-RNA.

In the Results section, the longest possible ORF among multiple polypeptide coding frames ranging from 18 to 76 amino acids long was described. An ORF with 36 amino acids long has already been noted in the D-loop region H-strand sequence of human mtDNA (Saccone et al., 1987). This short ORF starts 50 bases downstream from the terminator codon of 76

amino acids long peptide detected here. These results suggest possible functions of mrrH-RNA as a messenger RNA. However, no critical studies to test the presence of predicted products have been attempted. This is partly due to the difficulty in introducing recombinant technologies in mitochondria. Apparently further studies are necessary.

In the present study, we specifically focused our attention to the distribution of polyadenylated mrr-RNAs in human tissues and cultured cells. Strikingly, the results of Northern analyses using mrr probes indicated that both ca. 0.2 kb band, i.e., mrrL-RNA, and ca. 1.3 kb band, i.e., mrrH-RNA, distributed very differently among tissue cells and cultured cells. With respects to the levels of intensity, they were roughly classified, as described in the Results section, into the following three groups: the first group with high level 0.2 kb band and low level 1.3 kb band, the second group with high level 0.2 kb band and high level 1.3 kb band, and the third group with low level 0.2 kb band and high level 1.3 kb band.

Since all cultured cell lines, which are in growing states, showed intense 0.2 kb band, while tissues showed varying levels on the bands, we speculate that the level of 0.2 kb band may be related to the proliferation of cells and mitochondria. The lowest level of 0.2 kb band shown of all the tissues examined was the typically non-growing tissue, skeletal muscle, which may further support this speculation. In fact, one of the functions of mrrL-RNA is attributed to serve as a primer for H-strand DNA synthesis (Chang et al., 1985; Chang and Clayton, 1985; Clayton, 1991). Here, it has been postulated that newly transcribed mrrL-RNA from LSP forms persistent RNA-DNA hybrid in the region containing CSBs, where a site-specific mitochondrial RNA processing endonuclease, RNaseMRP, creates a donor 3'-hydroxyl group for DNA synthesis to commence by DNA polymerase (Xu and Clayton, 1996; Lee and Clayton, 1996, 1997). Supporting this postulation, free RNA relevant to mrrL-RNA and plasmid DNA template containing mouse mtDNA D-loop region were shown to associate stably together forming an R-loop (a structure of two DNA strands and one RNA strand) in the absence of protein factors in an *in vitro* study (Lee and Clayton, 1996), and such RNA-DNA hybrids were implicated to serve as primers for mtDNA replication in human (Xu and Clayton, 1996). It should be recalled here that rat mtDNA replication and expression may be regulated by an antisense mechanism with mrrL-RNA (Sbisa et al., 1990), and that the content of triple

stranded D-loop structure in mitochondria and cellular growth rate and/or the oxidative capacity of the tissues are closely related (Annex and Williams, 1990; King and Low, 1987; Williams et al., 1986). Essentially, however, little is known about the mechanisms regulating the transcription of mrrL-RNA. The present results, together with the above lines of evidence suggest that the regulation of mrrL-RNA transcription could be a key step for the proliferation control of mitochondria. Apparently further studies are awaited, and the present results showing a variety of expression levels in tissues and cell lines will provide suitable subjects here.

As to the function of mrrH-RNA, the present state of knowledge is very limited. The previous observation that H-strand RNA species accumulates by a 10–20 fold during *in vitro* induced myogenesis (Vijayarathy et al., 1995), suggests that they are involved either in the proliferation and/or in the activated function of mitochondria, and provide suitable tools to study the function of mrrH-RNA. The present results show that the level of ca. 1.3 kb mrrH-RNA was high in most tissues, and in addition, in some of the cultured cell lines including HepH2, KATO III and HL-60 as well. It is of particular interests that HepG2 cells, KATO III and HL-60 cells maintain functions such as secreting proteins or differentiating themselves. These cell lines, together with the cDNA clones isolated here to mrrH-RNAs, will present additional suitable models and tools for the study of regulation of mrrH-RNA transcription. Among the problem to be solved are if the ORFs shown in mrrH-RNAs are in fact functioning or not, if mrrH-RNAs are in fact the terminal portion of polycistronic H-strand transcripts, and if they interact with D-loop region of mtDNA or mrrL-RNAs as modulators of mitochondrial replication and gene expression.

Acknowledgements

This work was supported in part by the New Energy and Industrial Technology Development Organization (NEDO).

References

- Ahmed Y, Gilmartin G, Hanly S, Nevins J and Greene W (1991) The HTLV-1 rex responsive element mediates a novel form of mRNA polyadenylation. *Cell* 64: 727–737.

- Anderson S, Bankier A, Barrel B, de Bruijn M, Coulson A, Drouin J, Eperon I, Nierlich D, Roe B, Sanger F, Schreier P, Smith A, Staden R and Young I (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290: 457–465.
- Annex B and Williams R (1990) Mitochondrial DNA structure and expression in specialized subtypes of mammalian striated muscle. *Mol Cell Biol* 10: 5671–5678.
- Attardi G (1985) Animal mitochondrial DNA: An extreme example of genetic economy. *Int Rev Cytol* 93: 93–145.
- Attardi G (1993) The human mitochondrial genetic system. In: DiMauro S and Wallace DC (eds) *Mitochondrial DNA in Human Pathology* (pp. 9–25). Raven Press, Ltd., New York.
- Bibb M, van Etten R, Wright C, Walberg M and Clayton D (1981) Sequence and gene organization of mouse mitochondrial DNA. *Cell* 26: 167–180.
- Brown G, Gadaleta G, Pepe G, Saccone C and Sbisà E (1986) Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *J Mol Biol* 192: 504–511.
- Chang D and Clayton D (1984) Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell* 36: 635–643.
- Chang D and Clayton D (1985) Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc Natl Acad Sci USA* 82: 361–355.
- Chang D, Hauswirth W and Clayton D (1985) Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. *EMBO* 4: 1559–1567.
- Clayton D (1991) Replication and transcription of vertebrate mitochondrial DNA. In: Palade G, Alberts B and Spudich J (eds) *Ann Rev Cell Biol* 7: 453–478.
- Crews S and Attardi G (1980) The sequences of the small ribosomal RNA gene and the phenylalanine tRNA gene are joined end to end in human mitochondrial DNA. *Cell* 19: 775–784.
- Crews S, Ojala D, Posakony J, Nishiguchi J and Attardi G (1979) Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. *Nature* 277: 192–198.
- Dombrowski S, Brennicke A and Binder S (1997) 3'-inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO* 16: 5069–5076.
- Fitzgerald M and Shenk T (1981) The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. *Cell* 24: 251–260.
- Gadaleta G, Pepe G, De Candia G, Quagliariello C, Sbisà E and Saccone C (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.
- Hattori M and Sakaki Y (1986) Dideoxy sequencing method using denaturated plasmid template. *Anal Biochem* 152: 232–238.
- King T and Low R (1987) Mitochondrial DNA displacement loop structure depends on growth state in bovine cells. *J Biol Chem* 262: 6214–6220.
- Lee D and Clayton D (1996) Properties of a primer RNA-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication. *J Biol Chem* 271: 24262–24269.
- Lee D and Clayton D (1997) RNase mitochondrial RNA processing correctly cleaves a novel R loop at the mitochondrial DNA leading-strand origin of replication. *Genes Dev* 11: 582–592.
- MacDonald C, Wilusz J and Shenk T (1994) The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol* 14: 6647–6654.
- Manam S and van Tuyle G (1987) Separation and characterization of 5'- and 3'-tRNA processing nucleases from rat liver mitochondria. *J Biol Chem* 262: 10272–10279.
- Montoya J, Ojala D and Attardi G (1981) Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. *Nature* 290: 465–470.
- Montoya J, Gaines G and Attardi G (1983) The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell* 34: 151–159.
- Nakamichi N, Rhoads D, Hayashi J-I, Kagawa Y and Matsumura T (1998a) Detection, localization, and sequence analyses of mitochondrial regulatory region RNAs in several mammalian species. *J Biochem* 123: 392–398.
- Nakamichi N, Ito M and Matsumura T (1998b) Detection of mitochondrial regulatory region RNA in cultured cells and differentiated tissue cells: The implications for cellular growth control. In: Ikura K et al. (eds) *Animal Cell Technology: Challenges for the 21st Century* (pp. 319–324). Kluwer Academic Press, Dordrecht, The Netherlands.
- Noda A (1992) Gene expression in Ataxia Telangiectasia cells as perturbed by bleomycin treatment. *Somatic Cell Mol Genet* 18: 113–122.
- Ojala D, Crews S, Montoya J, Gelfand R and Attardi G (1981) A small polyadenylated RNA (7S RNA), containing a putative ribosome attachment site, maps near the origin of human mitochondrial DNA replication. *J Mol Biol* 150: 303–314.
- Proudfoot N and Brownlee G (1976) 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* 263: 211–214.
- Saccone C, Attimonelli M and Sbisà E (1987) Structural elements highly preserved during the evolution of the D-loop containing region in vertebrate mitochondrial DNA. *J Mol Evol* 26: 205–211.
- Saccone C, Pesole G and Sbisà E (1991) The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. *J Mol Evol* 33: 83–91.
- Sbisà E, Nardelli M, Tanzariello F, Tullo A and Saccone C (1990) The complete and symmetric transcription of the main non coding region of the rat mitochondrial genome: in vivo mapping of heavy and light transcripts. *Curr Genet* 17: 247–253.
- Sbisà E, Tanzariello F, Reyes A, Pesole G and Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: Identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205: 125–140.
- Takagaki Y and Manley J (1997) RNA recognition by the human polyadenylation factor CstF. *Mol Cell Biol* 17: 3907–3914.
- Topper J and Clayton D (1990) Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res* 18: 793–799.
- Tso Y, Sun X, Kao T, Reece K and Wu R (1985) Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNAs: genomic complexity and molecular evolution of the gene. *Nucleic Acids Res* 13: 2485–2502.
- Vijayasarathy C, Zheng Y, Mullick J, Basu A and Avadhani N (1995) Identification of a stable RNA encoded by the H-strand of the mouse mitochondrial D-loop region and a conserved sequence motif immediately upstream of its polyadenylation site. *Gene Expr* 4: 125–141.
- Walberg M and Clayton D (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Res* 9: 5411–5421.
- Walberg M and Clayton D (1983) In vitro transcription of human mitochondrial DNA. *J Biol Chem* 258: 1268–1275.

- Williams R, Salmons S, Newsholme E, Kaufman R and Mellor J (1986) Regulation of nuclear and mitochondrial gene expression by contractile activity in skeletal muscle. *J Biol Chem* 261: 376–380.
- Xu B and Clayton D (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication of RNA-DNA hybrids serving as primers. *EMBO J* 15: 3135–3143.