LETTER TO THE EDITOR

Insertional RNA editing in metazoan mitochondria: The cytochrome *b* gene in the nematode *Teratocephalus lirellus*

JACQUES R. VANFLETEREN and ANDY R. VIERSTRAETE

Department of Biology, University of Gent, B-9000 Gent, Belgium

Analysis of the sequence of the mitochondrial cyt *b* gene of *Teratocephalus lirellus* revealed multiple frame shift events. As these sites coincided with homopolymeric tracks, sequencing was repeated and chromatograms were screened for poorly resolved peaks. No evidence for sequencing errors was found. We therefore reverse transcribed mRNA, using M-MLV reverse transcriptase and two primers for strongly conserved regions located anteriorly (CbF1: >5′-GTTTTGGAA TGTTGGAAGGCTTTTAGG-3′) and posteriorly (CbR1: <5′-CCCAGATATCACCATAGCCAAATAGACG-3′) in the gene. The cDNA product was PCR amplified and sequenced, and compared with the gDNA sequences. This revealed insertions of a single U in eight sites in the mRNA sequence (Fig. 1).
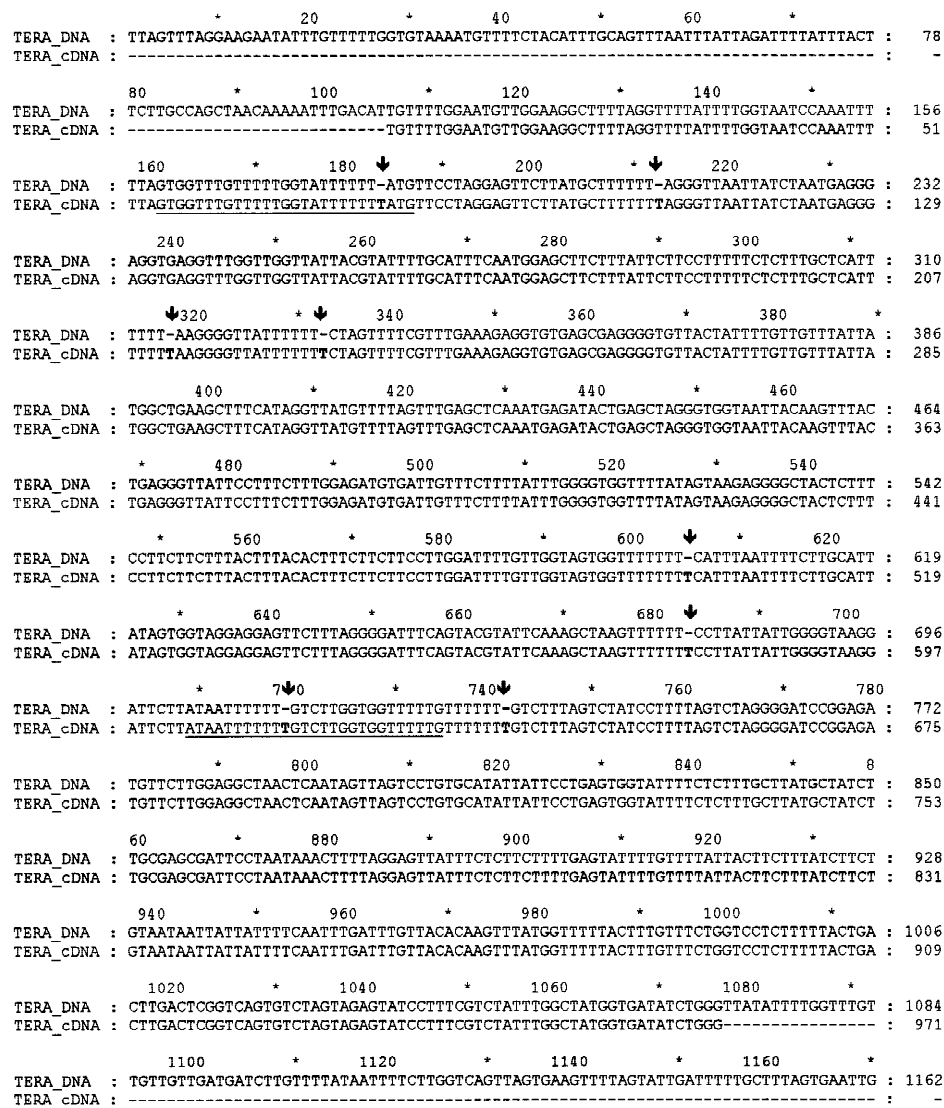
Experimental error is very unlikely for the following reasons. Both sequences were assembled from multiple overlapping fragments covering both strands. Polymerase infidelity is most unlikely, as all PCR products were cycle sequenced. The original gDNA sequence was reconfirmed using genomic DNA isolated simultaneously with RNA processed for the RT-PCR experiment. All differences between the gDNA and cDNA sequences are unambiguous, as demonstrated in Figure 2 for the region 322–337. To control for the possibility that the gDNA sequence might represent a pseudogene, we amplified a portion of the cyt *b* gene using primers that perfectly matched the cDNA sequence, but showed mismatches with the gDNA at their 3′ ends (Fig. 1). An amplicon was still generated by PCR using gDNA as a template, but its nucleotide sequence turned out to be identical with the known cyt *b*

gene sequence, with no evidence for the presence of an extra T at any of the five sites where a discrepancy between cDNA and gDNA had been detected. We therefore conclude that the mRNA for cyt *b* in *T. lirellus* is edited by insertion of single uridine residues at eight positions. However, we should add here that we did not perform any Southerns, and that nonamplified additional copies of the gene would have escaped detection.

To our knowledge, this is the first report of insertional editing in a metazoan species. Among eukaryotes, uridine insertion editing is also known to occur in slime mold and kinetoplastid mitochondria. Single uridines are inserted at six sites in the mRNA for cyt *b* in *Physarum polycephalum*, and multiple insertions are common in kRNA (Miller et al., 1993; Smith et al., 1997). However, the eight insertions observed in *T. lirellus* all extend oligo(U) tracts in the primary transcript, whereas the distribution of inserted uridines in kinetoplastids and *Physarum* mitochondria does not require homopolymeric tracts. In this respect, editing in *T. lirellus* somewhat resembles insertional editing in viral systems (Vidal et al., 1990; Volchkov et al., 1995; Sanchez et al., 1996), but is altogether unique in featuring the consistent addition of an extra uridine to all seven motifs corresponding to six thymidines in the mtDNA sequence (and once to seven thymidines; Fig. 1).

Insertional mRNA editing is not known in other nematodes. The entire mitochondrial genomes of *Ascaris suum* (Ascaridida) and *Caenorhabditis elegans* (Rhabditida) have been determined (Okimoto et al., 1992), with no evidence for mRNA editing, although several mitochondrial genes of *C. elegans* were specifically examined to this end (Orr et al., 1997). Furthermore, we have sequenced mitochondrial cyt *b* for 12 other species belonging to the orders Rhabditida and Diplogas-

```
                     *      20        *       40         *       60        *
TERA_DNA  : TTAGTTTAGGAAGAATATTTGTTTTTGGTGTAAAATGTTTTCTACATTTGCAGTTTAATTTATTAGATTTTATTTACT :   78
TERA_cDNA : ----------------------------------------------------------------------------- :    -

             80        *      100        *      120         *      140         *
TERA_DNA  : TCTTGCCAGCTAACAAAAATTTGACATTGTTTTGGAATGTTGGAAGGCTTTTAGGTTTTATTTTGGTAATCCAAATTT :  156
TERA_cDNA : -------------------------TGTTTTGGAATGTTGGAAGGCTTTTAGGTTTTATTTTGGTAATCCAAATTT :   51

            160          *     180  ↓     *      200         *   220
TERA_DNA  : TTAGTGGTTTGTTTTTGGTATTTTTT-ATGTTCCTAGGAGTTCTTATGCTTTTTT-AGGGTTAATTATCTAATGAGGG :  232
TERA_cDNA : TTAGTGGTTTGTTTTTGGTATTTTTTTATGTTCCTAGGAGTTCTTATGCTTTTTTTAGGGTTAATTATCTAATGAGGG :  129

              240         *      260         *      280        *      300         *
TERA_DNA  : AGGTGAGGTTTGGTTGGTTATTACGTATTTTGCATTTCAATGGAGCTTCTTTATTCTTCCTTTTTCTCTTTGCTCATT :  310
TERA_cDNA : AGGTGAGGTTTGGTTGGTTATTACGTATTTTGCATTTCAATGGAGCTTCTTTATTCTTCCTTTTTCTCTTTGCTCATT :  207

                ↓320       *  ↓     340        *      360         *      380            *
TERA_DNA  : TTTT-AAGGGGTTATTTTTT-CTAGTTTTCGTTTGAAAGAGGTGTGAGCGAGGGGTGTTACTATTTTGTTGTTTATTA :  386
TERA_cDNA : TTTTTAAGGGGTTATTTTTTTCTAGTTTTCGTTTGAAAGAGGTGTGAGCGAGGGGTGTTACTATTTTGTTGTTTATTA :  285

                    400         *      420         *      440         *      460
TERA_DNA  : TGGCTGAAGCTTTCATAGGTTATGTTTTAGTTTGAGCTCAAATGAGATACTGAGCTAGGGTGGTAATTACAAGTTTAC :  464
TERA_cDNA : TGGCTGAAGCTTTCATAGGTTATGTTTTAGTTTGAGCTCAAATGAGATACTGAGCTAGGGTGGTAATTACAAGTTTAC :  363

                  *      480         *      500         *      520         *      540
TERA_DNA  : TGAGGGTTATTCCTTTCTTTGGAGATGTGATTGTTTCTTTTATTTGGGGTGGTTTTATAGTAAGAGGGGCTACTCTTT :  542
TERA_cDNA : TGAGGGTTATTCCTTTCTTTGGAGATGTGATTGTTTCTTTTATTTGGGGTGGTTTTATAGTAAGAGGGGCTACTCTTT :  441

                     *      560         *      580         *      600  ↓      *      620
TERA_DNA  : CCTTCTTCTTTACTTTACACTTTCTTCTTCCTTGGATTTTGTTGGTAGTGGTTTTTTT-CATTTAATTTTCTTGCATT :  619
TERA_cDNA : CCTTCTTCTTTACTTTACACTTTCTTCTTCCTTGGATTTTGTTGGTAGTGGTTTTTTTTCATTTAATTTTCTTGCATT :  519

                     *      640         *      660         *      680  ↓      *      700
TERA_DNA  : ATAGTGGTAGGAGGAGTTCTTTAGGGGATTTCAGTACGTATTCAAAGCTAAGTTTTTT-CCTTATTATTGGGGTAAGG :  696
TERA_cDNA : ATAGTGGTAGGAGGAGTTCTTTAGGGGATTTCAGTACGTATTCAAAGCTAAGTTTTTTTCCTTATTATTGGGGTAAGG :  597

                     *      7↓0        *      740↓        *      760         *      780
TERA_DNA  : ATTCTTATAATTTTTT-GTCTTGGTGGTTTTTGTTTTTT-GTCTTTAGTCTATCCTTTTAGTCTAGGGGATCCGGAGA :  772
TERA_cDNA : ATTCTTATAATTTTTTTGTCTTGGTGGTTTTTGTTTTTTTGTCTTTAGTCTATCCTTTAGTCTAGGGGATCCGGAGA :  675

                        *      800         *      820         *      840         *      8
TERA_DNA  : TGTTCTTGGAGGCTAACTCAATAGTTAGTCCTGTGCATATTATTCCTGAGTGGTATTTTCTCTTTGCTTATGCTATCT :  850
TERA_cDNA : TGTTCTTGGAGGCTAACTCAATAGTTAGTCCTGTGCATATTATTCCTGAGTGGTATTTTCTCTTTGCTTATGCTATCT :  753

              60         *      880         *      900         *      920         *
TERA_DNA  : TGCGAGCGATTCCTAATAAACTTTTAGGAGTTATTTCTCTTCTTTTGAGTATTTTGTTTTATTACTTCTTTATCTTCT :  928
TERA_cDNA : TGCGAGCGATTCCTAATAAACTTTTAGGAGTTATTTCTCTTCTTTTGAGTATTTTGTTTTATTACTTCTTTATCTTCT :  831

              940         *      960         *      980         *     1000         *
TERA_DNA  : GTAATAATTATTATTTTCAATTTGATTTGTTACACAAGTTTATGGTTTTTACTTTGTTTCTGGTCCTCTTTTTACTGA : 1006
TERA_cDNA : GTAATAATTATTATTTTCAATTTGATTTGTTACACAAGTTTATGGTTTTTACTTTGTTTCTGGTCCTCTTTTTACTGA :  909

             1020         *     1040         *     1060         *     1080         *
TERA_DNA  : CTTGACTCGGTCAGTGTCTAGTAGAGTATCCTTTCGTCTATTTGGCTATGGTGATATCTGGGTTATATTTTGGTTTGT : 1084
TERA_cDNA : CTTGACTCGGTCAGTGTCTAGTAGAGTATCCTTTCGTCTATTTGGCTATGGTGATATCTGGG---------------- :  971

             1100         *     1120         *     1140         *     1160         *
TERA_DNA  : TGTTGTTGATGATCTTGTTTTATAATTTTCTTGGTCAGTTAGTGAAGTTTTAGTATTGATTTTTGCTTTAGTGAATTG : 1162
TERA_cDNA : ----------------------------------------------------------------------------- :    -
```

FIGURE 1. Comparison of gDNA and cDNA sequences for *Teratocephalus lirellus* cyt *b*. The arrows indicate the location of eight thymidine insertions in the cDNA, relative to the gDNA. Insertional editing was confirmed by sequence analysis of an amplified segment of gDNA using cDNA-specific primers spanning the underlined sequence segments.

terida, none of which exhibits any trace of changing reading frames. However, all these species belong to the subclass Secernentea, and could have secondarily lost editing: although the classification of *T. lirellus* is controversial, SSU rRNA analysis places it as a sister taxon to Secernentea (Blaxter et al., 1998), allowing for the possibility that insertional editing is actually common in non-secernentean nematodes. The only other cyt *b* sequence known so far from a non-secernentean nematode was determined using a cDNA library (Powers et al., 1993).

Widely different mechanisms for insertional editing are known to occur outside Metazoa. For example, kinetoplastid editing uses guide RNA sequences, while editing in viral systems results from stuttering of RNA polymerase at slippery homopolymer tracts within the viral DNA template (Vidal et al., 1990; Volchkov et al., 1995; Sanchez et al., 1996; Smith et al., 1997). The evolutionary implications of insertional mRNA editing in *T. lirellus* therefore depend on an understanding of the mechanism involved. It is nevertheless intriguing that kinetoplastids and nematodes do share other remarkable biochemical features, such as extensive trans-splicing of short leader sequences to nascent mRNA (Sutton & Boothroyd, 1986; Krause & Hirsh, 1987; Huang & Hirsh, 1989) and nematodes possess a functional glyoxylate cycle and are capable of synthesizing *de novo* polyunsaturated fatty acids (Rothstein, 1970), two features shared with plants and many protists. Persistence of these various, presumably ancestral features in nematodes may therefore tie in with other indications (Sidow & Thomas, 1994; Vanfleteren et al., 1994), suggesting that this phylum is one of the most ancient among Metazoa.
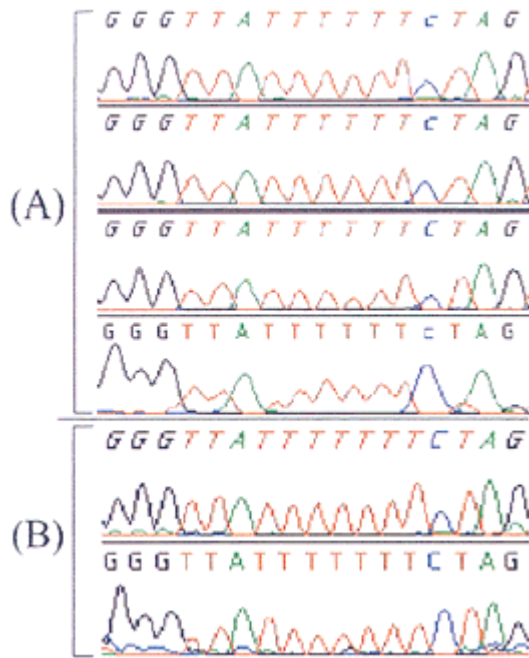
FIGURE 2. Editing of mitochondrial RNA for *T. lirellus* cyt *b*. Electropherogram data showing a portion of 16 (**A**) and 17 (**B**) bases of sequence of gDNA (**A**) and cDNA (**B**) obtained by cycle-sequencing. Italicized letters indicate that the sequence was obtained using the reverse primer.

REFERENCES

Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, Thomas WK. 1998. A molecular evolutionary framework for the phylum Nematoda. *Nature 392*:71–75.

Huang X-Y, Hirsh D. 1989. A second *trans*-spliced RNA leader sequence in the nematode *Caenorhabditis elegans*. *Proc Natl Acad Sci USA 86*:8640–8644.

Krause M, Hirsh D. 1987. A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell 49*:753–761.

Miller D, Mahendran R, Spottswood M, Costandy H, Wang S, Ling M-L, Yang N. 1993. Insertional editing in mitochondria of *Physarum*. *Cell Biol 4*:261–266.

Okimoto R, Macfarlane JL, Clary DO, Wolstenholme DR. 1992. The mitochondrial genomes of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics 130*:471–498.

Orr AT, Rabets JC, Horton TL, Landweber LF. 1997. RNA editing missing in mitochondria. *RNA 3*:335–336.

Powers TO, Harris TS, Hyman BC. 1993. Mitochondrial DNA sequence divergence among *Meloidogyne incognita, Romanomermis culicucivorax* and *Caenorhabditis elegans*. *J Nematol 25*:565–571.

Rothstein M. 1970. Nematode biochemistry. XI. Biosynthesis of fatty acids by *Caenorhabditis briggsae* and *Panagrellus redivivus*. *Int J Biochem 1*:422–428.

Sanchez A, Trappier SG, Mahy BWJ, Peters CJ, Nichol ST. 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc Natl Acad Sci USA 93*:3602–3607.

Sidow A, Thomas WK. 1994. A molecular evolutionary framework for eukaryotic model organisms. *Curr Biol 4*:596–603.

Smith HC, Gott JM, Hanson MR. 1997. A guide to RNA editing. *RNA 3*:1105–1123.

Sutton RE, Boothroyd JC. 1986. Evidence for *trans* splicing in trypanosomes. *Cell 47*:527–535.

Vanfleteren JR, Van de Peer Y, Blaxter ML, Tweedie SAR, Trotman C, Lu L, Van Hauwaert M-L, Moens L. 1994. Molecular genealogy of some nematode taxa as based on cytochrome *c* and globin amino acid sequences. *Mol Phylogenet Evol 3*:92–101.

Vidal S, Curran J, Kolakofsky D. 1990. Editing of the Sendai Virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. *J Virol 64*:239–246.

Volchkov VE, Becker S, Volchkova VA, Ternovoj VA, Kotov AN, Netesov SV, Klenk H-D. 1995. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and Vaccinia Virus polymerases. *Virology 214*:421–430.