## 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

Keywords: cytochrome b; insertional editing; mitochondrial RNA; Nematoda; Teratocephalus lirellus

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-

Reprint requests to: Jacques R. Vanfleteren, Department of Biology, University of Gent, Ledeganckstraat 35, B-9000 Gent, Belgium; e-mail: jacques.vanfleteren@rug.ac.be.

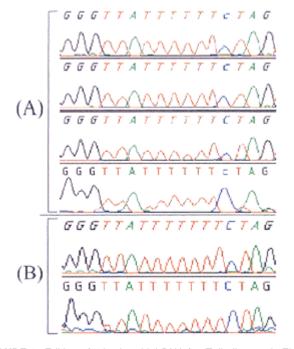
	*	20	4	40	*	60	*		
TERA DNA :	TTAGTTTAGGAAG		TTGGTGTAAA		CATTTGCAG		AGATTTTATTT	ACT :	78
TERA_CDNA :								:	-
	80 *	100	*	120	+	140			
TERA DNA :	TCTTGCCAGCTAA		CATTGTTTTG		AGGCTTTTA		GGTAATCCAAA	TTT :	156
TERA CDNA :							GGTAATCCAAA		51
_						.1			
TERA DNA .	160 * TTAGTGGTTTGTT	• 180 ••••••		200 דמככמכדדכד	• ጥልጥር/ጥጥጥጥ	• 🔶 220		ccc ·	232
	TTAGTGGTTTGTT								129
	240		60		80		00	*	
TERA_DNA :	AGGTGAGGTTTGG' AGGTGAGGTTTGG'	TTGGTTATTAC	GTATTTTGCA	TTTCAATGGA	GCTTCTTTA	TTCTTCCTTT1	TCTCTTTGCTC	ATT :	310 207
IERA_CDNA :	A0010A0011100	11001171170	GIAIIIIGUA		OCTICITIA		10101110010	AII -	201
	₩320	* 🝁	340	÷	360	*	380	*	
	TTTT-AAGGGGTT/								386
TERA_cDNA :	TTTT <b>TAAGGGGT</b> T/	ATTTTTTT <b>T</b> CTA	GTTTTCGTTT	GAAAGAGGTG	TGAGCGAGG	GGTGTTACTAI	TTTGTTGTTTA	TTA :	285
	400	*	420	*	440	*	460		
TERA DNA :	TGGCTGAAGCTTT	CATAGGTTATG	TTTTAGTTTG	AGCTCAAATG	AGATACTGA	GCTAGGGTGGT	AATTACAAGTT	TAC :	464
TERA_cDNA :	TGGCTGAAGCTTT	CATAGGTTATG	TTTTAGTTTG	AGCTCAAATG	AGATACTGA	GCTAGGGTGGI	AATTACAAGTT	TAC :	363
	* 480	+	500		520	•	540		
TERA DNA :	TGAGGGTTATTCC	TTTCTTTGGAG.		TTCTTTTATT		TTTATAGTAAG		TTT :	542
TERA_cDNA	TGAGGGTTATTCC	TTCTTTGGAG	ATGTGATTGT	TTCTTTTATT	TGGGGTGGT	TTTATAGTAAG	AGGGGCTACTC	TTT :	441
_									
-	* 560 CCTTCTTCTTTAC		580 המכתיהכי		60 (11) 000 000 000 000 000 000 000 000 000		620 KATTTTTCTTCC	ለጥጥ •	619
					014010011	TITT OUTT			012
TERA CUNA	CCTTCTTCTTTAC	TTACACTTTC	TTCTTCCTTG	GATTTTGTTG	GTAGTGGTT	TTTTTT <b>T</b> CATTI	AATTTTCTTGC	ATT :	519
TERA_CUNA :	CCTTCTTCTTTAC	TTACACTTTC	TTCTTCCTTG	GATTTTGTTG			AATTTTCTTGC	ATT :	519
-	* (	540	* 6 <sup>,</sup>	60	*	680 🔸	* 70	00	
- TERA DNA :	*	540 GAGTTCTTTAG	* 6 GGGATTTCAG	60 TACGTATTCA	* AAGCTAAGT	680 <b>↓</b> TTTTT-CCTT <b>A</b>	* 70 TTATTGGGGTA	)0 AGG :	696
- TERA DNA :	* (	540 GAGTTCTTTAG	* 6 GGGATTTCAG	60 TACGTATTCA	* AAGCTAAGT	680 <b>↓</b> TTTTT-CCTT <b>A</b>	* 70 TTATTGGGGTA	)0 AGG :	
TERA_DNA : TERA_cDNA :	* ATAGTGGTAGGAGG ATAGTGGTAGGAGG *	540 GAGTTCTTTAG GAGTTCTTTAG 7₩0	* 6 GGGATTTCAG GGGATTTCAG *	60 TACGTATTCA TACGTATTCA 740♥	* AAGCTAAGT AAGCTAAGT *	680 TTTTT-CCTTA TTTTTTCCTTA 760	* 7( ATTATTGGGGTA ATTATTGGGGTA *	0 AGG : AGG : 780	696 597
- TERA_DNA : TERA_cDNA : TERA_DNA :	* ATAGTGGTAGGAG ATAGTGGTAGGAG * ATTCTTATAATTT	540 GAGTTCTTTAG GAGTTCTTTAG 7¥0 TTT-GTCTTGG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT	* AAGCTAAGT AAGCTAAGT * 'TTAGTCTAT	680	* 70 ATTATTGGGGTA ATTATTGGGGTA * AGGGGATCCGG	AGG : AGG : AGG : 780 AGA :	696 597 772
- TERA_DNA : TERA_cDNA : TERA_DNA :	* ATAGTGGTAGGAGG ATAGTGGTAGGAGG *	540 GAGTTCTTTAG GAGTTCTTTAG 7¥0 TTT-GTCTTGG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT	* AAGCTAAGT AAGCTAAGT * 'TTAGTCTAT	680	* 70 ATTATTGGGGTA ATTATTGGGGTA * AGGGGATCCGG	AGG : AGG : AGG : 780 AGA :	696 597
- TERA_DNA : TERA_cDNA : TERA_DNA :	* ATAGTGGTAGGAG ATAGTGGTAGGAG * ATTCTTATAATTT	540 GAGTTCTTTAG GAGTTCTTTAG 7¥0 TTT-GTCTTGG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT	60 TACGTATTCA TACGTATTCA 740↓ TTTTT-GTCT TTTTTTGTCT	* AAGCTAAGT AAGCTAAGT * 'TTAGTCTAT	680	* 70 ATTATTGGGGTA ATTATTGGGGTA * AGGGGATCCGG	AGG : AGG : AGG : 780 AGA :	696 597 772
TERA_DNA : TERA_cDNA : TERA_DNA : TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTT <u>ATAATTTT</u> TGTTCTTGGAGGC	540 SAGTTCTTTAG SAGTTCTTTAG 7♥0 TTT-GTCTTGG TTTTGTCTTGG 800 FAACTCAATAG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT TGGTTTTTGT * TTAGTCCTGT	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT	* AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT * *	680 ↓ TTTTT-CCTTA TTTTTTCCTTA 760 CCTTTTAGTCT CCTTTTAGTCT 840 TATTTTCTCT	* 7( NTATTGGGGTA NTATTGGGGTA * AGGGGATCCGG AGGGGATCCGG * TGCTTATGCTA	AGG : AGG : 780 AGA : AGA : AGA : 8 .TCT :	696 597 772 675 850
TERA_DNA : TERA_cDNA : TERA_DNA : TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTT <u>ATAATTT</u>	540 SAGTTCTTTAG SAGTTCTTTAG 7♥0 TTT-GTCTTGG TTTTGTCTTGG 800 FAACTCAATAG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT TGGTTTTTGT * TTAGTCCTGT	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT	* AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT * *	680 ↓ TTTTT-CCTTA TTTTTTCCTTA 760 CCTTTTAGTCT CCTTTTAGTCT 840 TATTTTCTCT	* 7( NTATTGGGGTA NTATTGGGGTA * AGGGGATCCGG AGGGGATCCGG * TGCTTATGCTA	AGG : AGG : 780 AGA : AGA : AGA : 8 .TCT :	696 597 772 675
TERA_DNA : TERA_cDNA : TERA_DNA : TERA_DNA : TERA_CDNA :	* ATAGTGGTAGGAG ATAGTGGTAGGAGG * ATTCTTATAATTT ATTCTT <u>ATAATTT</u> TGTTCTTGGAGGC TGTTCTTGGAGGC	540 GAGTTCTTTAG SAGTTCTTTAG 7♥0 TTT-GTCTTGG <u>TTTGTCTTGG</u> 800 FAACTCAATAG TAACTCAATAG	* 6 GGGATTTCAG GGGATTTCAG * TGGTTTTTGT TGGTTTTTGT * TTAGTCCTGT	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 SCATATTATT GCATATTATT	* AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT * *	680 ♥ TTTTT-CCTTA TTTTTTCCTTA 760 CCTTTTAGTCT CCTTTTAGTCT 840 TATTTTCTCTT TATTTTCTCTT	* 7( NTATTGGGGTA NTATTGGGGTA * AGGGGATCCGG AGGGGATCCGG * TGCTTATGCTA	AGG : AGG : 780 AGA : AGA : AGA : 8 .TCT :	696 597 772 675 850
TERA_DNA : TERA_cDNA : TERA_cDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	* CATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTT <u>ATAATTT</u> TGTTCTTGGAGGC TGTTCTTGGAGGC 60 • TGCGAGCGATTCC	540 3AGTTCTTTAG 3AGTTCTTTAG 7♥0 TTT-GTCTTGG 11TTGTCTTGG 800 FAACTCAATAG 880 FAATAAACTTT	* 66 GGGATTTCAG GGGATTTCAG TGGTTTTTGT TGGTTTTTGT TTAGTCCTGT TTAGTCCTGT	60 TACGTATTC2 TAGTATTC2 740♥ TTTTT-GTCT 820 GCATATTATT GCATATTATT 900 TTCTCTTCTT	AAGCTAAGT MAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CTTGAGTATT	680 ↓ TTTTT-CCTTA TTTTTTCCTTA 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTA	* 7(( ATTATTGGGGTA ATTATTGGGGTA AGGGATCCGG AGGGATCCGG AGGGGATCCGG TGCTTATGCTA TGCTTATGCTA ACTTCTTTATCT	AGG : AGG : 780 AGA : AGA : 8 TCT : TCT : TCT :	696 597 772 675 850
TERA_DNA : TERA_cDNA : TERA_cDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTATAATTT TGTTCTTGGAGGC TGTTCTTGGAGGC	540 3AGTTCTTTAG 3AGTTCTTTAG 7♥0 TTT-GTCTTGG 11TTGTCTTGG 800 FAACTCAATAG 880 FAATAAACTTT	* 66 GGGATTTCAG GGGATTTCAG TGGTTTTTGT TGGTTTTTGT TTAGTCCTGT TTAGTCCTGT	60 TACGTATTC2 TAGTATTC2 740♥ TTTTT-GTCT 820 GCATATTATT GCATATTATT 900 TTCTCTTCTT	AAGCTAAGT MAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CTTGAGTATT	680 ↓ TTTTT-CCTTA TTTTTTCCTTA 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTA	* 7(( ATTATTGGGGTA ATTATTGGGGTA AGGGATCCGG AGGGATCCGG AGGGGATCCGG TGCTTATGCTA TGCTTATGCTA ACTTCTTTATCT	AGG : AGG : 780 AGA : AGA : 8 TCT : TCT : TCT :	696 597 772 675 850 753
TERA_DNA : TERA_cDNA : TERA_cDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTATAATTT TGTTCTTGGAGGC 60 TGCGAGCGATTCC TGCGAGCGATTCC	540 SAGTTCTTTAG SAGTTCTTTAG 7♥0 TTT-GTCTTGG <u>FTTTGTCTTGG</u> 800 FAACTCAATAG 880 FAATAAACTTT FAATAAACTTT	* 66 GGGATTTCAG GGGATTTCAG TGGTTTTTGT TGGTTTTTGT TTAGTCCTGT TTAGTCCTGT	60 TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT GCATATTATT 900 TTCTCTTCTT	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CTTGAGTATT	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT CCTTTTAGTCT 840 TATTTTCTCTT 720 TGTTTTATFF TTGTTTTATFF	* 7(( NTTATTGGGTA NTATTGGGTA NGGGGATCCGG NGGGGATCCGG TGCTTATGCTA TGCTTATGCTA ACTTCTTTATCT	AGG : AGG : 780 AGA : AGA : 8 TCT : TCT : TCT :	696 597 772 675 850 753 928
TERA_DNA : TERA_cDNA : TERA_cDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTATAATTT TGTTCTTGGAGGC TGTTCTTGGAGGC 60 TGCGAGCGATTCC TGCGAGCGATTCC 940	540 SAGTTCTTTAG SAGTTCTTTAG 7¥0 TTT-GTCTTGG <u>TTTGTCTTGG</u> 800 TAACTCAATAG 880 TAATAAACTTT TAATAAACTTT * 960	6 GGGATTTCAG GGGATTTCAG TGGTTTTTGT TGGTTTTTGT TAGTCCTGT TAGGAGTTAT TAGGAGTTAT TAGGAGTTAT	60 TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT GCATATTATT 900 TTCTCTTCTT TTCTCTTCTT 980	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG TTGAGTATT	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCCTTF 920 TTGTTTTATTF TTGTTTTATTF * 1000	* 7(( ATTATTGGGTA ATTATTGGGTA AGGGGATCCGG AGGGGATCCGG AGGGGATCCGG AGGGGATCGG AGGGGATCGG AGGGGATCGG AGGGGATCG AGGGGATCGG AGGGGATCG AGGGGATCG AGGGGATCG AGGGGATCG AGGGGGGA AGGGGGATCG AGGGGATCG AGGGGGA AGGGGGA AGGGGA AGGGGGA AGGGGGA AGGGGA AGGGGGA AGGGGG AGGGGA AGGGGGA AGGGGGA AGGGGGA AGGGG AGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGG AGGGGG AGGGG AGGGGG AGGGGG AGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGG AGGGGGG	AGG : AGG : 780 AGA : AGA : AGA : 8 TCT : TCT : TCT :	696 597 772 675 850 753 928 831
TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTATAATTT TGTTCTTGGAGGC 60 TGCGAGCGATTCC TGCGAGCGATTCC 940 GTAATAATTATTA	540 3AGTTCTTTAG GAGTTCTTTAG 7¥0 FTT-GTCTTGG FTAGTCTAGT 800 FAATAAACTTT FAATAAACTTT FAATAAACTTT 960 FTTTCAATTTG	• 6 GGGATTTCAG GGGATTTCAG • TGGTTTTTGT TTAGTCCTGT TTAGTCCTGT TAGGAGTTAT TAGGAGTTAT • ATTTGTTACA	60 TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT GCATATTATT 900 TTCTCTTCTT TTCTCTTCTT 980 CAAGTTTATC	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG TTTGAGTATT	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT CCTTTTAGTCT 840 TATTTTCTCTT 920 TGTTTTATTF TTGTTTTATTF * 1000 TTGTTTCTGGT	* 7(( ATTATTGGGTA ATTATTGGGGAA AGGGGAATCCGG AGGGGATCCGG TGCTTATGCTA TGCTTATGCTA ACTTCTTTATCT ACTTCTTTATCT	AGG : AGG : 780 AGA : AGA : AGA : TCT : TCT : TCT : TCT :	696 597 772 675 850 753 928 831
TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT: ATTCTTATAATTT: TGTTCTTGGAGGC TGTTCTTGGAGGC 60 TGCGAGCGATTCC 940 GTAATAATTATTA	540 SAGTTCTTTAG SAGTTCTTTAG TTT-GTCTTGG <u>FTTTGTCTTGG</u> 800 FAACTCAATAG 880 FAATAAACTTT ANTAAACTTT P60 FTTTCAATTTG.	6 GGGATTTCAG GGGATTTCAG TIGGTTTTTGT TAGTCCTGT TAGGAGTTAT TAGGAGTTAT TAGGAGTTAT ATTTGTTACA ATTTGTTACA	60 TACGTATTCA 740↓ TTTTT-GTCT TTTTTTGTCT 620 SCATATTATT GCATATTATT 900 TTCTCTTCTT TTCTCTTCTT 980 CAAGTTTATC CAAGTTTATC	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTGAGTATT TTGAGTATT GTTTTTACT	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTF TGTTTTATTF * 1000 TGGTTTCGGT	* 7(( ATTATTGGGTA ATTATTGGGTA AGGGGATCCGG AGGGGATCCGG AGGGGATCCGG AGGGGATCCGG AGGCGATCTGT ACTTCTTATGCTA ACTTCTTTATCT ACTTCTTTATCT CCCCCTTTTTAC	AGG : AGG : 780 AGA : AGA : AGA : TCT : TCT : TCT : TCT :	696 597 772 675 850 753 928 831 1006
TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_DNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAGG ATTCTTATAATTT ATTCTTATAATTT TGTTCTTGGAGGC 60 TGCGAGCGATTCC 940 GTAATAATTATTA GTAATAATTATTA	540 3AGTTCTTTAG GAGTTCTTTAG TTT-GTCTTGG' 100 100 100 100 100 100 100 10	<ul> <li>6</li> <li>GGGATTTCAG</li> <li>GGGATTTCAG</li> <li>TGGTTTTTGT</li> <li>TTAGTCCTGT</li> <li>TTAGTCCTGT</li> <li>TAGGAGTTAT</li> <li>TAGGAGTTAT</li> <li>ATTTGTTACA</li> <li>ATTTGTTACA</li> <li>A0</li> </ul>	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT 900 TTCTCTTCTT 900 CAAGTTTATC 980 CAAGTTTATC CAAGTTTATC * 10	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTTGAGTATT TTGAGTATT GTTTTTACT GTTTTTACT	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 640 TATTTTCTCTT 920 TGTTTTATTF 1000 TGTTTTATTF * 1000 TGTTTCTGGT TGTTTCTGGT * 1000	* 7(( ATTATTGGGTA ATTATTGGGTA AGGGGATCCGG AGGGGATCCGG TGCTTATGCTA TGCTTATGCTA ACTTCTTTATCT ACTTCTTTATCT CCTCTTTTTAC	AGG : AGG : 780 AGA : AGA : AGA : 8 TCT : TCT : TCT : TCT : TCT : TCA : *	696 597 772 675 850 753 928 831 1006 909
TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAGG TGTTCTTGGAGGGC TGTTCTTGGAGGGC 60 TGCGAGCGATTCC TGCGAGCGATTCC 940 GTAATAATTATTA GTAATAATTATTAT 1020 CTTGACTCGGTCA	540 SAGTTCTTTAG GAGTTCTTTAG 7₩0 TTT-GTCTTGG TTAGTCTAGT 800 TAATAAACTTT 880 TAATAAACTTT AATAAACTTT 960 TTTTCAATTTG TTTTCAATTTG 10 STGTCTAGTAG	<ul> <li>6</li> <li>GGGATTTCAG</li> <li>GGGATTTCAG</li> <li>TGGTTTTTGT</li> <li>TTAGTCCTGT</li> <li>TTAGTCCTGT</li> <li>TAGGAGTTAT</li> <li>TAGGAGTTAT</li> <li>ATTTGTTACA</li> <li>ATTTGTTACA</li> <li>AGTATCCTTT</li> </ul>	60 TACGTATTCA 740♥ TTTTT-GTCT TTTTTTGTCT 820 GCATATTATT 900 TTCTCTTCTT 900 CAAGTTTATC 980 CAAGTTTATC CAAGTTTATC 4 10 CGTCTATTTCC	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTTGAGTATT GTTTTTACT GTTTTTACT 60 GCTATGGTG	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTF * 1000 TTGTTTCTGGT TTGTTTCTGGT * 10 ATATCTGGGTT	* 7(( NTTATTGGGTA NTATTGGGTA NGGGGATCCGG NGGGGATCCGG CGGGGATCCGG CGGGGATCCGG NGGGGATCCGG CGGGTCTTTAGGTA CCTCTTTTAGGTA CCTCTTTTAGGTA NGTTCTTTTAGGTA	AGG : AGG : 780 AGA : AGA : AGA : 8 TCT : TCT : TCT : TCT : TCT : TCA : *	696 597 772 675 850 753 928 831 1006 909
TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTAGAGGC TGTTCTTGGAGGC GTGTATGAGGATTCC GTAATAATTATTA GTAATAATTATTA 1020 CTTGACTCGGTCA	540 3AGTTCTTTAG GAGTTCTTTAG TTT-GTCTTGG' 100 100 100 100 100 100 100 10	6 GGGATTTCAG GGGATTTCAG TGGTTTTTGT TTAGTCCTGT TTAGTCCTGT TAGGAGTTAT ATTTGTTACA ATTTGTTACA AGTATCCTTT AGTATCCTTT	60 TACGTATTCA 740♥ TTTTT-GTCT 740♥ TTTTTTGTCT 820 GCATATTATT 900 100 100 100 100 100 100 100	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTGAGTATT TTGAGTATT GTTTTTACT GTTTTTACT 60 GCTATGGTG GCTATGGTG	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTF * 1000 TTGTTTCTGGT TTGTTTCTGGT * 10 ATATCTGGGTT	* 7(( NTTATTGGGTA NTATTGGGTA NGGGGATCCGG NGGGGATCCGG CGGGGATCCGG CGGGGATCCGG NGGGGATCCGG CGGGTCTTTAGGTA CCTCTTTTAGGTA CCTCTTTTAGGTA NGTTCTTTTAGGTA	AGG : AGG : 780 AGA : AGA : AGA : 8 TCT : TCT : TCT : TCT : TCT : TCA : *	696 597 772 675 850 753 928 831 1006 909
TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAGG TGTTCTTGGAGGGC GO TGTCTTGGAGGGATTCC' 940 GTAATAATTATTA 1020 CTTGACTCGGTCAC CTTGACTCGGTCAC 1100	540 SAGTTCTTTAG GAGTTCTTTAG TT-GICTIGG' <u>FITTGTCTTGG</u> 100 100 100 100 100 100 100 10	<ul> <li>6</li> <li>GGGATTTCAG</li> <li>GGGATTTCAG</li> <li>TGGTTTTTGT</li> <li>TTAGTCCTGT</li> <li>TTAGTCCTGT</li> <li>TAGGAGTTAT</li> <li>TAGGAGTTAT</li> <li>ATTTGTTACA</li> <li>AGTATCCTTT-</li> <li>AGTATCCTTT-</li> <li>AGTATCCTTT-</li> <li>1120</li> </ul>	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT 820 GCATATTATT 900 TTCTCTTCTT 980 CAAGTTTATC CAAGTTTATC 4 10 CGTCTATTTCC CGTCTATTTCC CGTCTATTTCC CGTCTATTCC 4	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTTGAGTATT GTTTTTACT 60 GCTATGGTG GCTATGGTG GCTATGGTG 1140	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTF * 1000 TTGTTTCTGGT TTGTTTCTGGT * 10 ATATCTGGGT ATATCTGGGT * 10	* 7(( NTTATTGGGTA NTATTGGGTA NGGGGATCCGG NGGGGATCCGG TGCTTATGCTA TGCTTATGCTA NCTTCTTTATGTA CCTCCTTTTATCT CCTCCTTTTTAC NCTCTTTTTACT NCTTCTTTATCT NCTTCTTTATCT NCTTCTTTACT NCTTCTTTACT NCTTCTTTACT NCTTCTTTTACT NCTTCTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT	AGG : AGG : 780 AGA : AGA : AGA : TTCT : TTCT : TTCT : TTCT : TTGA : *	696 597 772 675 850 753 928 831 1006 909 1084 971
TERA_DNA : TERA_CDNA :	ATAGTGGTAGGAG ATAGTGGTAGGAG ATAGTGGTAGGAG ATTCTTATAATTT ATTCTTAGAGGC TGTTCTTGGAGGC GTGTATGAGGATTCC GTAATAATTATTA GTAATAATTATTA 1020 CTTGACTCGGTCA	540 SAGTTCTTTAG GAGTTCTTTAG TT-GICTIGG' <u>FITTGTCTTGG</u> 100 100 100 100 100 100 100 10	<ul> <li>6</li> <li>GGGATTTCAG</li> <li>GGGATTTCAG</li> <li>TGGTTTTTGT</li> <li>TTAGTCCTGT</li> <li>TTAGTCCTGT</li> <li>TAGGAGTTAT</li> <li>TAGGAGTTAT</li> <li>ATTTGTTACA</li> <li>AGTATCCTTT-</li> <li>AGTATCCTTT-</li> <li>AGTATCCTTT-</li> <li>1120</li> </ul>	60 TACGTATTCA TACGTATTCA 740♥ TTTTT-GTCT 820 GCATATTATT 900 TTCTCTTCTT 980 CAAGTTTATC CAAGTTTATC 4 10 CGTCTATTTCC CGTCTATTTCC CGTCTATTTCC CGTCTATTCC 4	AAGCTAAGT AAGCTAAGT TTAGTCTAT TTAGTCTAT CCTGAGTGG CCTGAGTGG CCTGAGTGG TTTGAGTATT GTTTTTACT 60 GCTATGGTG GCTATGGTG GCTATGGTG 1140	680 ↓ TTTTT-CCTTF TTTTTTCCTTF 760 CCTTTTAGTCT 840 TATTTTCTCTT 920 TTGTTTTATTF * 1000 TTGTTTCTGGT TTGTTTCTGGT * 10 ATATCTGGGT ATATCTGGGT * 10	* 7(( NTTATTGGGTA NTATTGGGTA NGGGGATCCGG NGGGGATCCGG TGCTTATGCTA TGCTTATGCTA NCTTCTTTATGTA CCTCCTTTTATCT CCTCCTTTTTAC NCTCTTTTTACT NCTTCTTTATCT NCTTCTTTATCT NCTTCTTTACT NCTTCTTTACT NCTTCTTTACT NCTTCTTTTACT NCTTCTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT NCTTCTTTTACT	AGG : AGG : 780 AGA : AGA : AGA : TTCT : TTCT : TTCT : TTCT : TTGA : *	696 597 772 675 850 753 928 831 1006 909 1084 971

**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 transsplicing 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.

### ACKNOWLEDGMENTS

This work was supported by grant G.2023.94N from the Fund for Scientific Research–Flanders, and grant BOF 01105097 from the University of Gent. The senior author is a Research Director with the Fund for Scientific Research–Flanders.

Received January 13, 1999; returned for revision February 16, 1999; revised manuscript received March 1, 1999

## 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 *Physa-rum. 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.