1696–1700 Nucleic Acids Research, 1994, Vol. 22, No. 9 

# Mitochondrial transcripts are processed but are not edited normally in *Trypanosoma equiperdum* (ATCC 30019) which has kDNA sequence deletion and duplication

Hsiao-Hsueh Shu and Kenneth Stuart[1,*]
Seattle Biomedical Research Institute, 4 Nickerson Street, Seattle, WA 98109 and [1]Pathobiology Department, University of Washington, SC-38, Seattle, WA 98195, USA

Received December 21, 1993; Revised and Accepted March 28, 1994 EMBL accession nos U03738–U03741 (incl.)

## ABSTRACT

**Analyses of the *Trypanosoma equiperdum* (ATCC 30019) maxicircle reveals deletions, duplications and rearrangement compared to *T.brucei*. The genes for 9S rRNA and 12 proteins are absent. The 12S rRNA and cytochrome oxidase subunit I (COI) genes lack their 3′ ends and are adjacent indicating deletion of intervening genes. The remaining two NADH dehydrogenase subunit genes (ND4 and ND5), the ribosomal protein RPS12 gene and the CR5 gene are duplicated and rearranged. ND4, RPS12 and the CR4 transcripts are abundant in steady state RNA while 12S rRNA and COI transcripts are not detected. Full length ND5 transcripts are rare, if present, but chimeric ND5/ND4 transcripts are abundant. The CR4 and RPS12 transcripts are the size of unedited RNAs suggesting that they are processed. However, they are not edited normally, presumably due to the absence of minicircle gRNA genes.**

## INTRODUCTION

The mitochondrial DNA (kinetoplast DNA or kDNA) of African trypanosomes is a concatenated network about 50 maxicircles and thousands of minicircles (1). The 20 kb maxicircles encode mitochondrial rRNAs and genes for mitochondrial respiratory proteins as do other mitochondrial DNAs. The 1 kb minicircles encode guide RNAs (gRNAs) that probably specify edited RNA sequences. *Trypanosoma equiperdum* is morphologically indistinguishable from *T.brucei* but its life cycle has no insect stage (2). While *T.brucei* kDNA contains hundreds of different sequence classes of minicircles, *T.equiperdum* kDNA contains a single major class (3–7). In addition, *T.equiperdum* Pasteur strain maxicircles are 24.5 kb (4), similar to *T.brucei*, but ATCC 30019 strain maxicircles are only 14 kb (5). Thus, *T.equiperdum* appear to be naturally occurring mutants of *T.brucei* that have kDNA alterations. Since cytochrome-mediated respiration is required in insect vector, these alterations are expected to prevent survival of *T.equiperdum* in the insect host.

We examined the maxicircle of *T.equiperdum* (ATCC 30019) and its transcripts because of our interests in RNA editing and control of kDNA gene expression. We show that its maxicircle lacks portions of the 12S rRNA and COI genes and 12 intervening genes. Some remaining gene sequences are duplicated and rearranged. Transcripts of some of the remaining genes are abundant in steady state RNA but others are not detected. Transcripts that are edited in *T.brucei* are not edited normally in this strain presumably due to the absence of minicircle gRNAs gene.

## RESULTS

We confirmed the absence of a 9 kb sequence (5) and localized the maxicircle genes by restriction mapping and DNA sequencing (Fig. 1). A probe for 12S rRNA hybridized to the 2.5 kb R1–R3 maxicircle fragment; ND8 and MURF2 probes did not hybridize to any fragments. Probes for COI, ND4 and the 5′- ND5 all hybridized to the 3.7 kb R3–H3 maxicircle fragment, the same size as in *T.brucei*, suggesting that these genes are on homologous fragments in both species. ND4 and 5′- ND5 probes also hybridized to the 1.2 kb H3–H3x fragment which is not in *T.brucei*, suggesting a sequence duplication. A 3′- ND5 probe hybridized to the 6.2 kb H3x–R1 fragment. This size is similar to the *T.brucei* 164 fragment which contains 3′- ND5 and ~5 kb of the variable region (8,9) implying a similar sequence composition.

The 12S rRNA and COI genes of *T.equiperdum* are truncated and adjacent. Sequence adjacent to the R3 site shows that the COI gene spans this site as in *T.brucei* (Fig. 2). Thus, the R1–R3 and R3–H3 are adjacent in *T.equiperdum*. The ~400 bp sequence is identical, except for 7 sites (3 nucleotide insertions, 1 deletion and 3 nucleotide substitutions), to the 12S rRNA sequence of *T.brucei* up to an AG, beyond which it matches the COI sequence. The AG occurs 182 bp upstream of the 3′ end of the 12S rRNA gene (10) and 381 bp upstream of the termination codon of the COI gene. These data, along with the hybridization data above indicate that the entire region containing

*To whom correspondence should be addressed

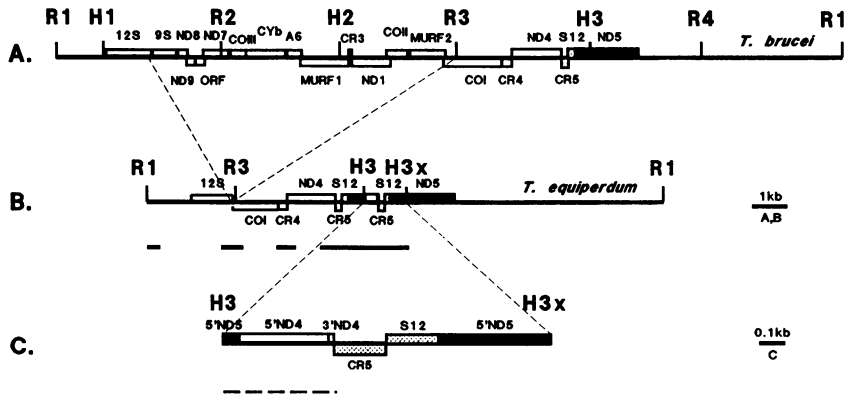


**Figure 1.** Maxicircle maps. The maxicircles of *T.brucei* (**A**) and *T.equiperdum* (**B**) are linearized at an *Eco*RI site and genes are shown as blocks. Blocks above the line represent genes which are transcribed to the right; those below the line are transcribed to the left. *Eco*RI (E) and *Hin*dIII (H) restriction sites are numbered for alignment and H3x is unique to *T.equiperdum*. Dashed lines align the same regions between maps. Sequenced regions in *T.equiperdum* DNA are indicated with solid lines under its map and their accession numbers are U03738, U03739, U03740 and U03741. A portion of the map is enlarged in (**C**) to show the H3−H3x fragment and the ND5 sequences are shaded and CR5 and RPS12 sequences are crosshatched for clarity. The nucleotide sequence indicated by the broken line is shown in Figure 3A. Large and small rRNAs (12S and 9S); NADH dehydrogenase subunits 1, 4, 5, 7, 8 and 9 (ND1, 4, 5, 7, 8 and 9); apocytochrome b (CYb); ATPase subunit 6 (A6); cytochrome oxidase subunits I, II and III (COI, II, and III); ribosomal protein RPS12 (S12); maxicircle unidentified reading frames 1 and 2 (MURF1 and 2); C-rich template sequences 3, 4, and 5 (CR 3,4, and 5).

```
                                12S⇦ ⇨COI                                                    EcoRI (R3)
...ACTATATTCTTTGAAATTATAAAAGAACCAAATCATAATGTGGAAATGAAGAAAAATAATCAAAATGTGTGAAGTTCAATAGGAATTC
```

**Figure 2.** Nucleotide sequence at the junction of 12S rRNA and COI genes. The *T.equiperdum* sequence near the R3 site is shown. The dinucleotide at the junction of 12S rRNA and COI sequences is in boldface.

the 9S rRNA, ND8, ND9, ND7,COIII, CYb, A6, MURF1, CR3, ND1, COII, MURF2 and the 3′ portions of the 12S rRNA and COI genes are absent from the *T.equiperdum* maxicircle (Fig. 1).

Segments totaling ~2 kb of the 3.7 kb R3−H3 fragment were sequenced and found to have the same gene order of COI, CR4, ND4, CR5, RPS12 and ND5 as in *T.brucei*. (Fig. 1). Sequence adjacent to the H3 site contains the 3′ end of ND4, CR5, RPS12, and 5′ end of ND5, respectively, a second sequence spans the 5′ end of CR4 and ND4, and the sequence adjacent to the R3 site contains the 5′ region of COI (data not shown). Compared to *T.brucei* there were 14 substitutions (5 of which were transversions), 16 deletions and 8 insertions which create frameshifts in the COI, ND4 and ND5 genes. The H3−H3x fragment has ND4, ND5, CR5 and RPS12 genes with ~96% identity with those of *T.brucei* but has gene duplications (see below). The sequences of the CR4 and 12S rRNA genes and a portion of the variable region have ~90% identity with those of *T.brucei*.

The 1.2 kb H3−H3x contains duplicate CR5 and RPS12 genes, duplicated portions of ND4 and ND5 sequences, and $(TA)_{16}$ inserted into the ND4 sequence (Fig. 3A and data not shown) as diagrammed in Fig. 3B. The sequences surrounding the H3 and upstream of the H3x sites match the ND5 gene of *T.brucei*. We did not sequence downstream of H3x but hybridization to a *T.brucei* ND5 probe indicates sequence conservation (data not shown). The duplicate sequence differs from the other copy by 2 deletions, 5 insertions and 4 substitutions over ~1 kb of available sequence. The sequence switches from ND5 to ND4 43 bp downstream of the H3 site (Fig.3A) which corresponds to 17 bp downstream of the ND4 ATG. Thus a sequence from this point in ND4 to a point 43 bp downstream of the H3 site (562 bp downstream of the ND5 ATG) is duplicated into the ND5 gene (Fig. 3B). The duplicated ND4 sequence also lacks ~1050 bp of sequence; from 235 bp downstream of its ATG to 19 bp upstream of its termination codon (Fig. 3A) and contains an insert of $(TA)_{16}$. Thus, gene duplication, amplification of $(TA)_{16}$, and deletion of ~1050 bp led to the duplicate sequence.

Northern analysis with a radio-labelled, *in vitro* RNA probe that hybridizes to both unedited and partially edited CR4 RNAs detects a single band transcripts that are slightly larger than the size of unedited RNA in *T.equiperdum* but a range of sizes in *T.brucei* which reflects RNA editing in *T.brucei* (Fig. 4). A clone obtained from oligo (dT) primed cDNA was unedited (data not shown). Thus, CR4 transcript probably have poly (A) tails and appear processed from polycistronic precursors but they are not edited normally. COI and 12S rRNA transcripts were not detected in *T.equiperdum*. An oligonucleotide probe for 12S rRNA detected this RNA in *T.brucei*, but not in *T.equiperdum* (Fig. 5A). Similarly, a COI oligonucleotide probe hybridized to *T.brucei* mRNA and kDNA and *T.equiperdum* kDNA (data not shown) but did not detect COI mRNA in *T.equiperdum* (Fig. 5B). Re-probing of these blots revealed CR4 (Fig.4), ND4, and ND5 (Figs. 5C and D) transcripts in *T.equiperdum*.

The normal size 1400 nt ND4 transcript is detected in *T.equiperdum* (Fig. 5C) although the ND4 transcript with the longer poly (A) tail (11) is not. A 900 nucleotide transcript is detected and corresponds to the size expected from the recombinant ND5/ND4 gene (Figs. 1 and 3). An 800 nt transcript is also seen and may represent a degradation or aberrant processing product. Re-probing with a ND5 probe revealed 900 and 800 nt transcripts but not normal sized 1800 nt transcripts,

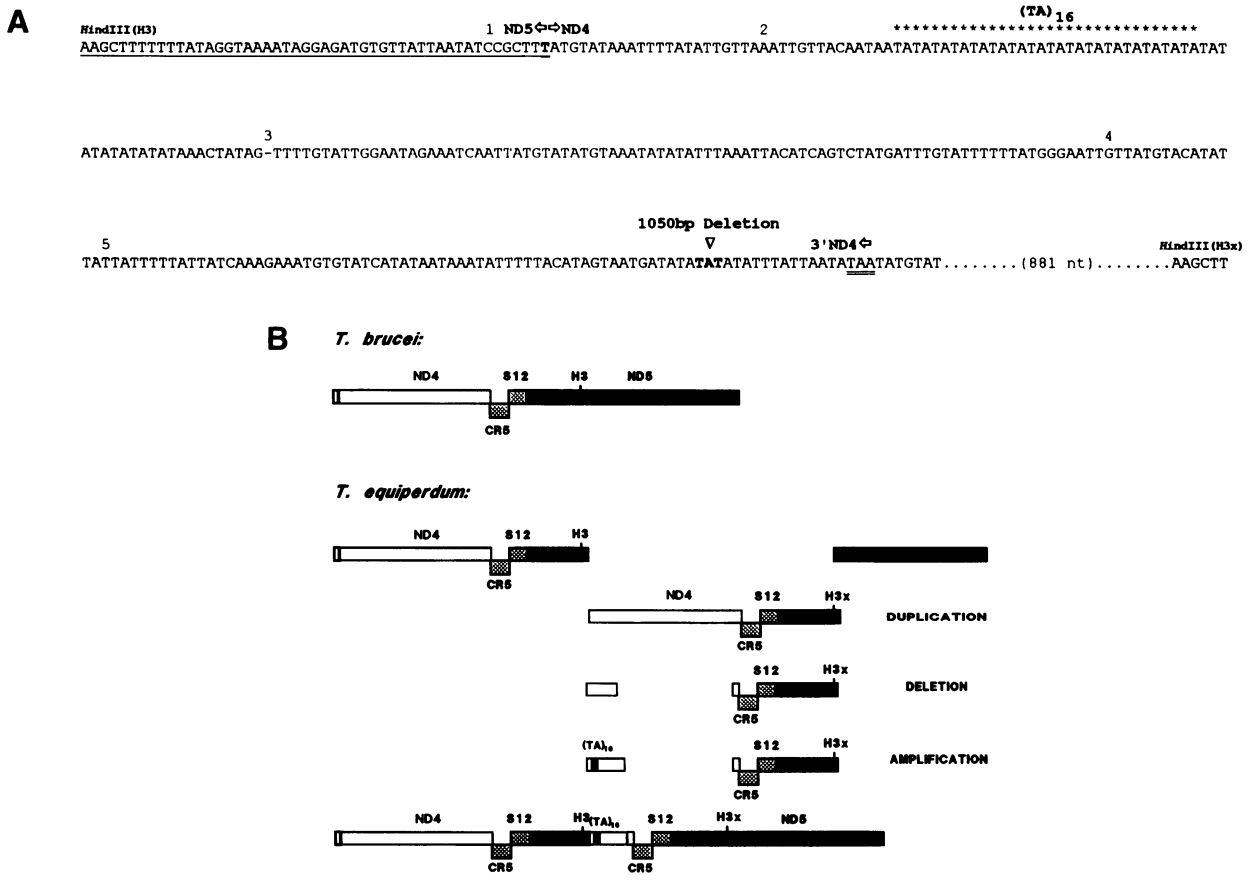


**Figure 3.** The H3−H3x fragment of *T.equiperdum*. (**A**) A 329 nucleotide sequence from the H3 site is shown.The ND5 sequence is underlined; the remaining sequence is ND4. The T at the ND4/ND5 junction region is in boldface. The amplified $(TA)_{16}$ sequence is indicated with *'s, and the deletion site of the 1050 bp is indicated by a triangle and the TAT that occurs at the termini of the deleted sequence is in boldface. The TAA termination codon of ND4 is double underlined. Sequence differences between *T.equiperdum* and *T.brucei* are numbered 1−5: 1, T; 2, a deletion; 3, a T insertion; 4, A; 5, A in *T.brucei*. (**B**) Diagram of possible events. Genes are labelled as in Fig. 1. The duplication, deletion and amplification events are labelled but their order of occurrence is not known.

(Fig. 5D). The low abundance of 1800 nt transcripts in *T.brucei* is not surprising, since procyclic forms have less ND5 RNA than bloodstream forms (12), but is surprising for *T.equiperdum*. Thus, the 900 and 800 nucleotides transcripts contain both ND4 and ND5 sequence. The 780, 440 and 650 nucleotides transcripts in *T.brucei* (Figs. 5C and D) are probably degradation products. Thus, ND4 and recombinant ND4/ND5 transcripts are abundant (Fig. 5C), but there are few ND5 transcripts suggesting that they are either not transcribed or not stable, as discussed below.

Re-probing of the blot from Fig. 5D for RPS12 RNA primarily detects a transcript which is the size of unedited RPS12 RNA. Longer exposure shows faint bands of 1800, 1400 and 900 nt (Fig. 6). The 900 nt transcript may be a RPS12 plus ND5/ND4 processing precursor analogous to that previously reported (13) but it is ~200 nt smaller than expected. The potential 2 kb RPS12/ND5 precursor is not detected. This suggests that, if present, the precursors are processed rapidly. Similarly CR4/COI precursors were not detected (data not shown).

## DISCUSSION

We show that the *T.equiperdum* (ATCC 30019) maxicircle deleted 9kb including 12 genes and the 3′ portions of the 12S rRNA and COI genes. The remaining genes retain the same order as in *T.brucei* except a 1.2 kb sequence with ND4, ND5, CR5 and RPS12 genes is duplicated and deletion and amplification modified a portion of the duplicated segment creating a recombinant ND4/ND5 gene. The genes that remain are transcribed and the transcripts are processed but are not edited normally, presumably due to the absence of most gRNA genes.

Multiple events and processes probably produced the *T.equiperdum* (ATCC 30019) maxicircle and it is likely that recombination played a role unlike processes related to those proposed for yeast mitochondrial DNA rearrangement (14). Amplified $(TA)_{16}$ implies unequal cross over or polymerase slippage during replication (15). The 9 kb and 1050 bp deletions were probably separate events, and the sequence duplication probably preceded the latter deletion.

The alterations in *T.equiperdum* (ATCC 30019) disrupt the function of all remaining maxicircle genes. COI and ND5 genes have frameshifts and their full length transcripts are essentially absent. The ND4 gene has no initiation codon in its open reading frame. Production of functional CR4, CR5 and RPS12 transcripts is unlikely since they are extensively edited in *T.brucei* but not in *T.equiperdum*. Thus, *T.equiperdum* (ATCC 30019) does not appear to produce functional maxicircle encoded mRNAs. Furthermore, 12S and 9S rRNAs are absent indicating that mitochondrial translation is not functional. The kDNA alterations

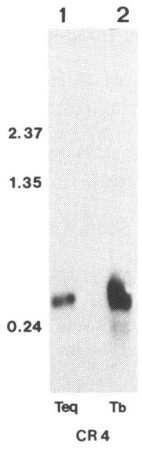


**Figure 4.** CR4 is transcribed in *T.equiperdum*. Equivalent amounts of RNA from a mitochondria-enriched fraction of *T.equiperdum* (lane 1) or from *T.brucei* procyclic form mitochondria (lane 2) was separated in a 1.5% agarose-2.2 M formaldehyde gel, transferred to nylon membrane and hybridized with the CR4 RNA probe (see Methods). The positions of RNA size markers are shown in kb.

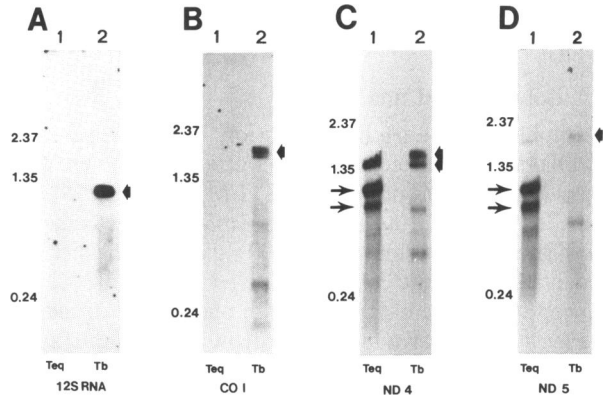


**Figure 5.** Transcripts of 12S, COI, ND4 and ND5 genes are at different steady state levels. Northern blots prepared as described in Fig. 4 were hybridized with: **A.** Oligonucleotide ATCC-12S. **B.** Oligonucleotide TbCOI-2. **C.** Oligonucleotide ND4-41. **D.** Oligonucleotide ND5-1. Arrowheads indicate the full length transcripts. Arrows indicate the 900 and 800 nucleotide transcripts with both ND4 and ND5 sequences. Size markers are as described in Figure 4.

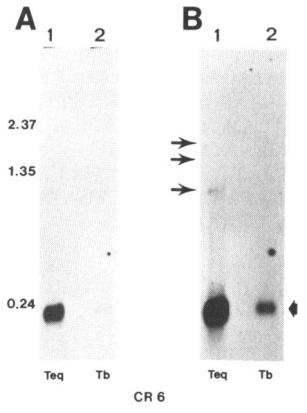


**Figure 6.** No potential precursor of RPS12 is detected in *T.equiperdum*. Northern blot prepared as described in Figure 4 was hybridized with radiolabelled CR6P oligonucleotide. **A.** 7 hours exposure. **B.** 20 hours exposure. The arrowhead indicates the unedited RPS12 transcript. Transcripts of 1800, 1400 and 900 nucleotides are marked with arrows. Size markers are as described in Figure 4.

and functional losses are tolerated since kDNA is not essential in the mammalian host (16) and *T.equiperdum* is restricted to this life cycle stage (17). Hence the mutations appear free from selective pressure. The absence of maxicircles in *T.evansi* (18) and all kDNA in spontaneous and induced mutants (19,20) may reflect a parallel situation.

The presence in *T.equiperdum* of maxicircle transcripts and gRNAs with 3′ U-tails (21) despite substantial sequence deletion indicates that mitochondrial RNA polymerase and terminal uridylyl transferase are not encoded in the deleted sequences. The normal sizes of the maxicircle transcripts indicates that genes responsible for their processing also have a nuclear location.

The low abundance or absence of 12S rRNA, COI and ND5 transcripts despite abundant ND4/ND5 transcripts probably reflects deletion of 3′ sequence, perhaps resulting in transcript instability. Alternatively, it may reflect promoter alteration since sequence differences occur upstream of the ND5/ND4 and ND5 genes. Maxicircle promoters are not identified; there may be several per maxicircle (22) or one per DNA strand.

The abundance of full length CR4 and ND4 unedited transcripts indicates normal transcription and cleavage and thus the intact gene and flanking sequences contain the processing signals. The abundance of chimeric ND5/ND4 transcripts suggests that sequences for processing this transcript are retained. This gene has 562 bp of ND5 5′ sequence and 19 bp of ND4 3′ sequence and normal flanking sequences. Thus, if the chimeric gene is co-transcribed with RPS12 analogous to the RPS12 and ND5 genes in *T.brucei* (13), the retained sequences are sufficient for processing. Alternatively, if the recombinant gene is a single transcription unit, the retained sequence is sufficient for initiation and termination of the transcription.

The sizes of the CR4 and RPS12 transcripts indicates that they are not edited normally in *T.equiperdum*. This is evidence that the gRNAs specify edited sequences. The lack of extensive editing is correlated with only three gRNA genes in *T.equiperdum* minicircles for ND7, COIII and CR4 mRNAs. (21). Only the CR4 gene is retained but the CR4 gRNA cannot form an anchor duplex with unedited mRNA. (23). Thus, most mitochondrial mRNA is not edited in *T.equiperdum*.

## MATERIALS AND METHODS

Bloodstream form *T.equiperdum* (ATCC 30019) and normal and Dk *T.brucei* (EATRO 164) were grown in rats or as procyclic forms (12,24).

### DNA isolation, cloning and analysis

kDNA was isolated (25,26), digested with restriction enzymes, separated in 1% agarose gels, and transferred to Nytran membrane (Schleicher and Schuell). End labelled oligonucleotides complementary to *T.brucei* 12S rRNA, ND8 [CR1.8 (27)], MURF2, COI, ND4 and ND5 genes and a random primed probe of a *Hin*dIII−*Nde*I fragment of clone pTKHR34 (28) which contains the 3′ ND5 gene sequence were used as probes. Hybridizations were done as described previously (29). *Eco*RI, *Hin*dIII and *Eco*RI−*Hin*dIII double-digested restriction fragments of *T.equiperdum* (ATCC 30019) kDNA were cloned into

pBluesript II SK(−) (Stratagene). Plasmids from each clone were sequenced using an ABI DNA sequencer.

### RNA isolation and analysis

A mitochondria-enriched fraction was prepared from *T.equiperdum* by differential centrifugation. Cells were disrupted by grinding with carborundum in STEM buffer (0.25M sucrose, 20mM Tris−HCl, 10mM EDTA, 1mM mercaptoethanol, pH 7.9) and debris was removed by 5000×g centrifugation (30) RNA was isolated from the 30,000×g pellet by lysis in 0.2% SDS, followed with three extractions with phenol and digestion with DNase I (Pharmacia), following manufacturer's recommendations. Isolated RNA was separated in 1.5% agarose-2.2 M formaldehyde gels, and transferred to Nytran membrane.

End labelled oligonucleotides complementary to *T.equiperdum* 12S rRNA, *T.brucei* ND4, ND5 and RPS12 transcripts were used as probes (29). A partially edited CR4 *T.brucei* cDNA clone (CR4B-19) (Corell and Stuart unpublished), was used to synthesize RNA probe as described for random primed probes at 65 °C.

*Oligonucleotides.* The following oligonucleotides were used as probes for hybridizations:

12S rRNA, GGTACATATAGAACAACTGT;
ATCC-12S, AATTTCAAAGAATATAGTTCTTATGCTATATAAATAA-ATATAGAT; MURF2-D, CACGACTCAATCAAAGCCAATC;
TbCOI-2, GTCACATGCTAAGCTAGAATG;
ND4-41, CATATACATAATTGATTTCTATTCCAATACAAAAACTAG;
ND5-1, CCTACCAAACATAAATGAACCTGATATAAACCC;
CR6P, GTTCCCCCCACCCAAATC

## ACKNOWLEDGEMENTS

We thank T.Morales and J.Bertles for excellent technical assistance, Dr E.J.Bienen for communicating his cell fractionation protocol, and Drs J.E.Feagin and M.Zamora for critical comments on this manuscript. This work was supported by NIH grant AI14102 to KS who is a Burroughs Wellcome Scholar in Molecular Parasitology.

## REFERENCES

1. Stuart,K. and Feagin,J.E. (1992) *Int. Rev. Cytol.*, **141**, 65−88.
2. Vickerman,K. (1976) In Lumsden,W.H.R. and Evans,D.A. (eds.), Biology of kinetoplastida. Academic Press, London, pp. 1−34.
3. Riou,G. (1979) *Biochem. Biophy. Res. Com.*, **90**, 405−409.
4. Riou,G. and Saucier,J. (1979) *J. Cell Biol.*, **82**, 248−263.
5. Frasch,A.C.C., Borst,P., Hajduk,S.L., Hoeijmakers,J.H.J., Brunel,F. and Davison,J. (1980) *Biochim. Biophys. Acta*, **607**, 397−401.
6. Barrois,M., Riou,G. and Galibert,F. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3323−3327.
7. Silver,L.E., Torri,A.F. and Hajduk,S.L. (1986) *Cell*, **47**, 537−543.
8. Myler,P.J., Glick,D.L., Feagin,J.E., Morales,T.H. and Stuart,K. (1993) *Nucl. Acids Res.*, **21**, 687−694.
9. Sloof,P., de Haan,A., Eier,W., van Iersel,M., Boel,E., van Steeg,H. and Benne,R. (1992) *Mol. Biochem. Parasitol.*, **56**, 289−300.
10. Adler,B.K., Harris,M.E., Bertrand,K.I. and Hajduk,S.L. (1991) *Mol. Cell. Biol.*, **11**, 5878−5884.
11. Bhat,G.J., Souza,A.E., Feagin,J.E. and Stuart,K. (1992) *Mol. Biochem. Parasitol.*, **52**, 231−240.
12. Jasmer,D.P., Feagin,J.E. and Stuart,K. (1985) *Mol. Cell. Biol.*, **5**, 3041−3047.
13. Read,L.K., Myler,P.J. and Stuart,K. (1992) *J. Biol. Chem.*, **267**, 1123−1128.
14. Clark-Walker,G.D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 8847−8851.
15. Tautz,D. (1989) *Nucl. Acids Res.*, **17**, 6463−6471.
16. Vickerman,K. (1965) *Nature*, **208**, 762−766.
17. Woo,P.T.K. (1977) In Kreier,J.P. (ed.), Parasitic Protozoa. Academic Press, New York, pp. 269−296.
18. Borst,P. and Hoeijmakers,J.H.J. (1979) *Plasmid*, **2**, 20−40.
19. Wenyon,C.M. (1928) *Trans.Roy.Soc.Trop.Med.Hyg.*, **22**, 85−87.
20. Tobie,E.J. (1951) *Trans. Am. Microscopic. Soc*, **70**, 251−254.
21. Pollard,V.W. and Hajduk,S.L. (1991) *Mol. Cell. Biol.*, **11**, 1668−1675.
22. Simpson,L., Neckelmann,N., de la Cruz,V.F. and Muhich,L. (1985) *Nucl. Acids Res.*, **13**, 5977−5993.
23. Corell,R.A., Myler,P.J., and Stuart,K.D., (1993) *Mol. Biochem. Parasitol.*, (in press)
24. Brun,R. and Schonenberger,M. (1979) *Acta Trop.*, **36**, 289−292.
25. Stuart,K. (1979) *Plasmid*, **2**, 520−528.
26. Feagin,J.E., Abraham,J.M. and Stuart,K. (1988) *Cell*, **53**, 413−422.
27. Souza,A.E., Myler,P.J. and Stuart,K. (1992) *Mol. Cell. Biol.*, **12**, 2100−2107.
28. Stuart,K.D. and Gelvin,S.B. (1982) *Mol. Cell. Biol.*, **2**, 845−852.
29. Shu,H., Wise,C.A., Clark-Walker,G.D. and Martin,N.C. (1991) *Mol. Cell. Biol.*, **11**, 1662−1667.
30. Grady,R.W., Bienen,E.J. and Clarkson,A.B.,Jr. (1986) *Mol. Biochem. Parasitol.*, **19**, 231−240.