170–176 *Nucleic Acids Research, 1998, Vol. 26, No. 1*

U-insertion/deletion Edited Sequence Database

Larry Simpson[1,2,3,*], Shirley H. Wang[1], Otavio H. Thiemann[2], Juan D. Alfonzo[1], Dmitri A. Maslov[4] and Herbert A. Avila[5]

[1]Howard Hughes Medical Institute, [2]Department of Molecular, Cell and Developmental Biology, UCLA, Los Angeles, CA 90095-1662, USA, [3]Department of Medical Microbiology and Immunology, UCLA School of Medicine, Los Angeles, CA, USA, [4]Department of Biology, University of California, Riverside, CA 92521, USA and [5]Department of Biochemistry, University of Kentucky Medical Center, Lexington, KY 40536-0084, USA

Received August 20, 1997; Revised and Accepted October 3, 1997

ABSTRACT

Uridine insertion/deletion RNA editing is a post-transcriptional RNA modification occurring in the mitochondria of kinetoplastid protozoa. The U-insertion/deletion Edited Sequence Database is a compilation of mitochondrial genes and edited mRNAs from five kinetoplastid species. It contains separate files with the DNA, mRNA (both unedited and edited) and predicted protein sequences, as well as alignments of the *Leishmania tarentolae* and *Trypanosoma brucei* protein sequences from edited and unedited genes. The sequence files are in GCG format. A 'map' sequence file showing the location of U-deletions, U-insertions and the translated amino acid sequences is also provided for each gene. Genomic maps for each species are also provided with clickable genes, including maxicircle-encoded gRNAs. Sets of aligned nuclear rRNA sequences from kinetoplastid protozoa are also provided, which were used for phylogenetic reconstructions in an analysis of the origin of RNA editing. The database is available through the World Wide Web as an HTML document at the URL http://www.lifesci.ucla.edu/RNA/trypanosome/database.html

INTRODUCTION

Multiple mRNAs from the maxicircle mitochondrial genome of kinetoplastid protozoa undergo uridine (U)-insertion/deletion RNA editing (1–5). In this process, Us are inserted and deleted at precise sites, usually within coding regions. Editing provides a mechanism for the alteration of the reading frame and, in some cases, the creation of translation initiation codons (6–8). Genes whose transcripts are modified by RNA editing are termed 'cryptogenes' (6), and the region in which editing occurs is termed the 'pre-edited region' or the 'editing domain'. In several instances, editing is so extensive as to take an unrecognizable G-rich cryptogene and render it into an open reading frame; this is termed 'pan-editing'. The resulting edited RNAs are homologous to sequences encoding known mitochondrial proteins (9–16).

The information for the precise insertion and deletion events in a cryptogene transcript is provided by guide RNAs (gRNAs). Guide RNAs can base-pair to the pre-edited mRNAs just downstream of the region to be edited, thereby forming the 'anchor duplex' (17). A single gRNA mediates the editing of a 'block' of sequence. Frequently, an editing domain consists of multiple overlapping gRNAs. The 3′→5′ polarity of editing-site selection within a domain is caused by the creation of anchor sequences for upstream gRNAs by downstream editing (18).

The maxicircle molecules correspond to the informational DNA molecules in other organisms, and encode two rRNAs, 18 structural genes and a few gRNAs (4,19,20). The minicircle molecules encode the majority of the gRNAs. The kinetoplastid protozoa consist of at least two major lineages, the trypanosomatids and the bodonids/cryptobiids (21,22). The trypanosomatids, which are the best studied organisms, contain a single network consisting of thousands of catenated minicircles and maxicircles situated within the single mitochondrion adjacent to the basal body of the flagellum. Only one example of the second major lineage has been studied in detail. The cryptobiid, *Trypanoplasma borreli*, contains two types of circular mitochondrial DNA molecules: the 40–80 kb circles contain rRNA genes, structural genes and cryptogenes, and the 180–200 kb circles contain gRNA genes (23–25).

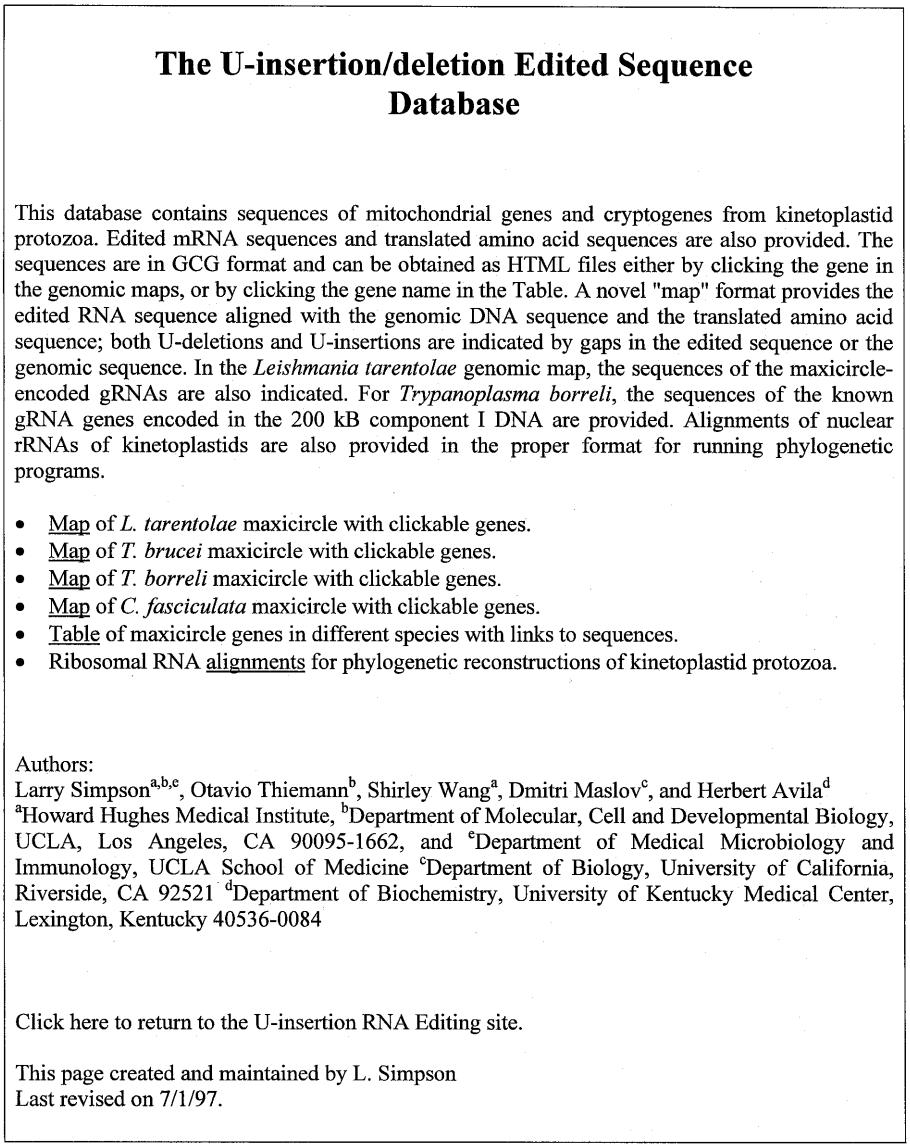
We describe here a database of mitochondrial genes, cryptogenes, and edited mRNA sequences from kinetoplastid protozoa.

DESCRIPTION OF DATABASE

This database (Fig. 1) is a compilation of mitochondrial sequences from four trypanosomatid species: *Leishmania tarentolae* (Fig. 2A), *Trypanosoma brucei* (Fig. 2B), *Trypanosoma cruzi* (not included as a clickable map), *Crithidia fasciculata* (Fig. 2C), and one cryptobiid species, *Trypanoplasma borreli* (Fig. 2D). The database contains files of the 9S and 12S mitochondrial rRNA sequences, the sequences of unedited structural genes, pre-edited cryptogenes, fully edited mRNAs, and the predicted peptide sequences encoded by the unedited genes and the fully edited mRNAs. All files are HTML documents, with the sequences in GCG format, which can be accessed either by clicking the name of the gene in a Table (Fig. 3), or the gene itself in a genomic map (Fig. 2A–D). Preceding each sequence is the gene name, the species, the accession number in GenBank, and notes on sequence corrections or 5′ end localizations.

A 'map' file is also provided for each gene. The map file consists of the DNA sequence (the pre-edited sequence for cryptogenes), followed by the RNA sequence (fully edited when

*To whom correspondence should be addressed at: 6-780 MacDonald Building, 675 Circle Drive South, Los Angeles, CA 90095-1662, USA. Tel: +1 310 825 4215; Fax: +1 310 206 8967; Email: simpson@hhmi.ucla.edu



The U-insertion/deletion Edited Sequence Database

This database contains sequences of mitochondrial genes and cryptogenes from kinetoplastid protozoa. Edited mRNA sequences and translated amino acid sequences are also provided. The sequences are in GCG format and can be obtained as HTML files either by clicking the gene in the genomic maps, or by clicking the gene name in the Table. A novel "map" format provides the edited RNA sequence aligned with the genomic DNA sequence and the translated amino acid sequence; both U-deletions and U-insertions are indicated by gaps in the edited sequence or the genomic sequence. In the *Leishmania tarentolae* genomic map, the sequences of the maxicircle-encoded gRNAs are also indicated. For *Trypanoplasma borreli*, the sequences of the known gRNA genes encoded in the 200 kB component I DNA are provided. Alignments of nuclear rRNAs of kinetoplastids are also provided in the proper format for running phylogenetic programs.

- Map of *L. tarentolae* maxicircle with clickable genes.
- Map of *T. brucei* maxicircle with clickable genes.
- Map of *T. borreli* maxicircle with clickable genes.
- Map of *C. fasciculata* maxicircle with clickable genes.
- Table of maxicircle genes in different species with links to sequences.
- Ribosomal RNA alignments for phylogenetic reconstructions of kinetoplastid protozoa.

Authors:
Larry Simpson[a,b,e], Otavio Thiemann[b], Shirley Wang[a], Dmitri Maslov[c], and Herbert Avila[d]
[a]Howard Hughes Medical Institute, [b]Department of Molecular, Cell and Developmental Biology, UCLA, Los Angeles, CA 90095-1662, and [e]Department of Medical Microbiology and Immunology, UCLA School of Medicine [c]Department of Biology, University of California, Riverside, CA 92521 [d]Department of Biochemistry, University of Kentucky Medical Center, Lexington, Kentucky 40536-0084

Click here to return to the U-insertion RNA Editing site.

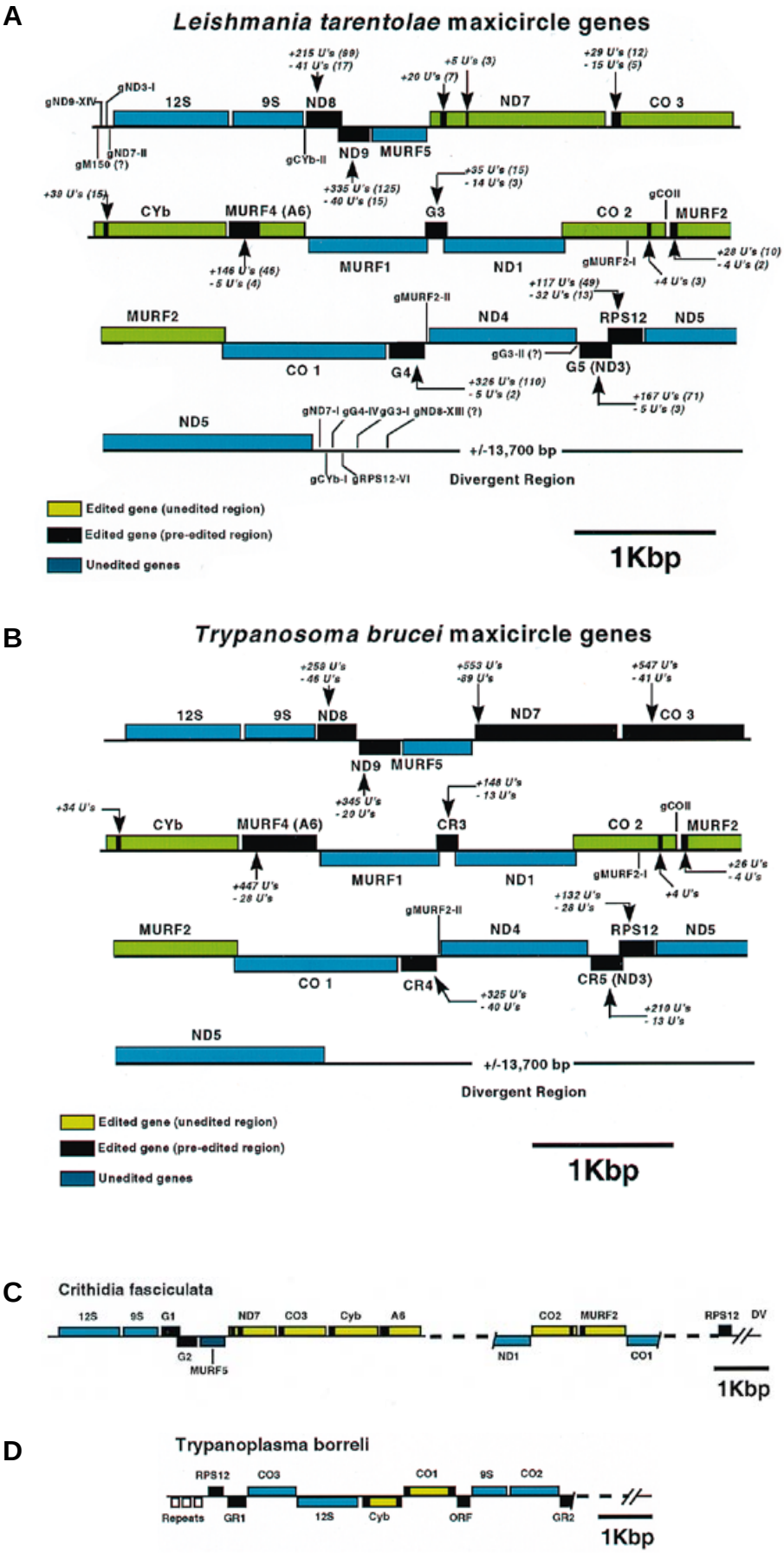
This page created and maintained by L. Simpson
Last revised on 7/1/97.

Figure 1. The U-insertion/deletion Edited Sequence Database web site, located at the following URL: http://www.lifesci.ucla.edu/RNA/trypanosome/database.html . This site provides access to clickable genomic maps of the maxicircle DNAs from *L.tarentolae, T.brucei, C.fasciculata* and *T.borreli* (Fig. 2A–D), a Table of maxicircle genes from these species (Fig. 3), and several sets of aligned nuclear rRNA sequences from several kinetoplastid species.

appropriate), and the translated amino acid sequence underneath the edited RNA sequence. The DNA sequence of edited genes contains gaps at the positions corresponding to U insertions, and the edited RNA sequence contains gaps at positions of U deletions. This is useful since the specific U-deletion information is otherwise difficult to determine from the separate sequence files. Nucleotide and amino acid numbering is provided, the top line of numbers corresponding to nucleotide positions in the DNA sequence, the second row of nucleotide numbers (located below the RNA sequence) to the mature RNA, and the bottom line to the predicted amino acid sequence. This format is useful for the design of primers for various analyses.

The alignment of the *T.brucei* and *L.tarentolae* predicted amino acid sequences of each mitochondrial gene from both edited and unedited RNA sequences is provided. These pairwise alignments

Figure 2. (Overleaf) Clickable genomic maps of maxicircle DNAs from four kinetoplastid species. (**A**) *L.tarentolae*, (**B**) *T.brucei*, (**C**) *C.fasciculata* and (**D**) *T.borreli*. The molecules are linearized. Genes above the line are in the 5′→3′ direction, left to right, and genes below are in the opposite orientation. The number of Us added and deleted are indicated for each editing domain in the *L.tarentolae* and *T.brucei* maps (A and B), and the number of sites shown in parenthesis. The locations of maxicircle-encoded gRNAs are indicated and the gRNAs identified by gene and editing block (3′→5′ within the editing domain). 12S and 9S, mitochondrial rRNAs; CO, cytochrome oxidase; MURF, maxicircle unidentified reading frame; ND, NADH dehydrogenase; CYb, cytochrome b; A6, ATPase subunit 6; RPS12, ribosomal protein S12; G1–G5 and CR1–CR5, G-rich regions 1–5. Clicking on the gene gives an HTML file of the unedited or edited sequence. Clicking on the gene name gives the map file. Clicking on the species name gives the entire maxicircle sequence from GenBank.


A
Leishmania tarentolae maxicircle genes
+215 U's (99)
- 41 U's (17)
+5 U's (3)
+20 U's (7)
+29 U's (12)
- 15 U's (5)
gND3-I
gND9-XIV
12S
9S
ND8
ND7
CO 3
gND7-II
gM150 (?)
gCYb-II
ND9
MURF5
+335 U's (125)
- 40 U's (15)
+35 U's (15)
- 14 U's (3)
+39 U's (15)
CYb
MURF4 (A6)
G3
CO 2
gCOII
MURF2
+146 U's (46)
- 5 U's (4)
MURF1
ND1
gMURF2-I
+4 U's (3)
+28 U's (10)
- 4 U's (2)
+117 U's (49)
- 32 U's (13)
gMURF2-II
MURF2
ND4
RPS12
ND5
CO 1
G4
gG3-II (?)
G5 (ND3)
+326 U's (110)
- 5 U's (2)
+167 U's (71)
- 5 U's (3)
ND5
gND7-I gG4-IV gG3-I gND8-XIII (?)
+/-13,700 bp
gCYb-I gRPS12-VI
Divergent Region
Edited gene (unedited region)
Edited gene (pre-edited region)
Unedited genes
1Kbp
B
Trypanosoma brucei maxicircle genes
+258 U's
- 46 U's
+553 U's
-89 U's
+547 U's
- 41 U's
12S
9S
ND8
ND7
CO 3
ND9
MURF5
+345 U's
- 20 U's
+148 U's
- 13 U's
+34 U's
CYb
MURF4 (A6)
CR3
CO 2
gCOII
MURF2
+447 U's
- 28 U's
MURF1
ND1
gMURF2-I
+4 U's
+26 U's
- 4 U's
+132 U's
- 28 U's
gMURF2-II
MURF2
ND4
RPS12
ND5
CO 1
CR4
CR5 (ND3)
+325 U's
- 40 U's
+210 U's
- 13 U's
ND5
+/-13,700 bp
Divergent Region
Edited gene (unedited region)
Edited gene (pre-edited region)
Unedited genes
1Kbp
C
Crithidia fasciculata
12S
9S
G1
ND7
CO3
Cyb
A6
CO2
MURF2
RPS12
DV
G2
MURF5
ND1
CO1
1Kbp
D
Trypanoplasma borreli
RPS12
CO3
CO1
9S
CO2
Repeats
GR1
12S
Cyb
ORF
GR2
1Kbp

U insertion/deletion editing in kinetoplastid mitochondria Sequence Database

Total Maxicircle Sequences:

L. tarentolae (partial - lacks complete divergent region)
T. brucei
C. fasciculata (partial - two files)
File 1 (9S and 12S rRNAs, COII, 5' end of Cyb)
File 2 (ND1, COII, Murf2 and the 3' portion of COI)

Gene:

Mitochondrial ribosomal RNA genes
9S rRNA
- ***L. tarentolae***
- ***T. brucei***
- ***C. fasciculata***

12S rRNA
- ***L. tarentolae***
- ***T. brucei***
- ***C. fasciculata***

COI
Edited mRNA
- ***T. borreli***

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***
- ***C. fasciculata***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***
- ***C. fasciculata***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***
- ***C. fasciculata***

COII
Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

COIII
Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Cyb
Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***C. fasciculata***
- ***L. tarentolae***

Figure 3. Table of mitochondrial genes. The sequence of each gene is accessible by clicking the species name. The sequences are in GCG format, with the exception of the map files. *T.cruzi* sequences are also included.

- ***T. brucei***
- ***T. borreli***

G3
Edited mRNA
- ***L. tarentolae***

Unedited mRNA
- ***L. tarentolae***

Amino acid sequence of mRNA
- ***L. tarentolae***

Map
- ***L. tarentolae***

G4 (CR4)
Edited mRNA
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***L. tarentolae***
- ***T. brucei***

MURF1
Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***L. tarentolae***
- ***T. brucei***

MURF2
Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

A6 (MURF4)
Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. cruzi***

MURF5
Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***L. tarentolae***
- ***T. brucei***

ND1
Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***
- ***C. fasciculata***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***
- ***C. fasciculata***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***
Map
- ***L. tarentolae***
- ***T. brucei***
- ***C. fasciculata***

ND3 (G5,CR5)
Edited mRNA
- ***L. tarentolae***

Figure 3. *continued*

- ***T. brucei***

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***L. tarentolae***
- ***T. brucei***

ND4

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***L. tarentolae***
- ***T. brucei***

ND5

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***L. tarentolae***
- ***T. brucei***

ND7

Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***

ND8 (G1,CR1)

Edited mRNA
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***L. tarentolae***
- ***T. brucei***

ND9 (G2,CR2)

Edited mRNA
- ***L. tarentolae***
- ***T. brucei***

Unedited mRNA
- ***L. tarentolae***
- ***T. brucei***

Amino acid sequence of mRNA
- ***L. tarentolae***
- ***T. brucei***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***L. tarentolae***
- ***T. brucei***

RPS12 (G6,CR6)

Edited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Unedited mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Amino acid sequence of mRNA
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Bestfit alignment of proteins - *L. tarentolae* vs ***T. brucei***

Map
- ***C. fasciculata***
- ***L. tarentolae***
- ***T. brucei***
- ***T. borreli***

Figure 3. *continued*

were performed using the BESTFIT local homology program in the GCG package (26). Identities are designated by '|' and conserved substitutions by '**:**' and '**.**'.

Nuclear ribosomal RNA alignments are provided for multiple kinetoplastid species. These alignments are included in this database since they were used for phylogenetic reconstructions in an analysis of the origin of this type of RNA editing (21,22,27).

A variety of selected information is also available in the linked U-insertion/deletion RNA editing web site. For example, a table with the taxonomy of kinetoplastid protozoa, diagrams of the various models for RNA editing, electron micrographs of kinetoplast DNA, a comprehensive list of literature references, and selected recent research results from papers in this field are provided. This list will be updated frequently with current information from the literature. In addition, there is a comprehensive list of investigators in this field with their URLs and electronic mail addresses, and announcements of upcoming scientific meetings.

AVAILABILITY

The U-insertion/deletion RNA editing database is accessible by any Web reader at:

http://www.lifesci.ucla.edu/RNA/trypanosome/database.html

The linked U-insertion RNA editing web site and the main RNA editing web site are at the URLs:

http://www.lifesci.ucla.edu/RNA/trypanosome/index.htm and

http://www.lifesci.ucla.edu/RNA/index.html

The administrator of the database and these web sites (Dr Larry Simpson) can be contacted by electronic mail (simpson@hhmi.ucla.edu) or by mail at the address given above. Users of the database should cite this publication. Corrections, new entries, errors and/or omissions and other material for inclusion in the database are welcome.

ACKNOWLEDGEMENTS

This work was supported in part by a research grant from the National Institutes of Health (AI09102). We thank Drs Rob Benne and Paul Sloof for supplying the *C.fasciculata* sequences.

REFERENCES

1 Simpson,L. and Thiemann,O.H. (1995) *Cell*, **81**, 837–840.
2 Simpson,L., Maslov,D.A. and Blum,B. (1993) In Benne,R. (ed.) *RNA Editing - The Alteration of Protein Coding Sequences of RNA*. Ellis Horwood, New York, pp. 53–85.
3 Kable,M.L., Heidmann,S. and Stuart,K.D. (1997) *Trends Biochem. Sci.*, **22**, 162–166.
4 Sloof,P. and Benne,R. (1997) *Trends Microbiol.*, **5**, 189–195.
5 Alfonzo,J.D., Thiemann,O. and Simpson,L. (1997) *Nucleic Acids Res.*, **25**, 3751–3759.
6 Simpson,L. and Shaw,J. (1989) *Cell*, **57**, 355–366.
7 Benne,R., Van den Burg,J., Brakenhoff,J., Sloof,P., Van Boom,J. and Tromp,M. (1986) *Cell*, **46**, 819–826.
8 Shaw,J., Feagin,J.E., Stuart,K. and Simpson,L. (1988) *Cell*, **53**, 401–411.
9 Maslov,D.A., Sturm,N.R., Niner,B.M., Gruszynski,E.S., Peris,M. and Simpson,L. (1992) *Mol. Cell. Biol.* **12**, 56–67.
10 Sturm,N.R. and Simpson,L. (1990) *Cell*, **61**, 871–878.
11 Souza,A.E., Shu,H.-H., Read,L.K., Myler,P.J. and Stuart,K.D. (1993) *Mol. Cell. Biol.*, **13**, 6832–6840.
12 Corell,R.A., Myler,P. and Stuart,K. (1994) *Mol. Biochem. Parasitol.*, **64**, 65–74.
13 Read,L.K., Wilson,K.D., Myler,P.J. and Stuart,K. (1994) *Nucleic Acids Res.*, **22**, 1489–1495.
14 Read,L.K., Myler,P.J. and Stuart,K. (1992) *J. Biol. Chem.*, **267**, 1123–1128.
15 Souza,A.E., Myler,P.J. and Stuart,K. (1992) *Mol. Cell. Biol.*, **12**, 2100–2107.
16 Koslowsky,D.J., Bhat,G.J., Perrollaz,A.L., Feagin,J.E. and Stuart,K. (1990) *Cell*, **62**, 901–911.
17 Blum,B., Bakalara,N. and Simpson,L. (1990) *Cell*, **60**, 189–198.
18 Maslov,D.A. and Simpson,L. (1992) *Cell*, **70**, 459–467.
19 Simpson,L. (1997) *Mol. Biochem. Parasitol.*, **86**, 133–141.
20 Stuart,K. (1993) In Benne,R. (ed.) *RNA editing - The Alteration of Protein Coding Sequences of RNA*. Ellis Horwood, New York, Vol.1. pp. 25–52.
21 Maslov,D.A., Lukes,J., Jirku,M. and Simpson,L. (1996) *Mol. Biochem. Parasitol.*, **75**, 197–205.
22 Maslov,D.A., Avila,H.A., Lake,J.A. and Simpson,L. (1994) *Nature*, **365**, 345–348.
23 Yasuhira,S. and Simpson,L. (1996) *RNA*, **2**, 1153–1160.
24 Maslov,D.A. and Simpson,L. (1994) *Mol. Cell. Biol.*, **14**, 8174–8182.
25 Lukes,J., Arts,G.J., Van den Burg,J., De Haan,A., Opperdoes,F., Sloof,P. and Benne,R. (1994) *EMBO J.*, **13**, 5086–5098.
26 Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucleic Acids Res.*, **12**, 387–395.
27 Lukes,J., Jirku,M., Dolezel,D., Kral'ová,I., Hollar,L. and Maslov,D.A. (1997) *J. Mol. Evol.*, **44**, 521–527.