Protist, Vol. 154, 265–279, July 2003 © Urban & Fischer Verlag
http://www.urbanfischer.de/journals/protist
Published online 30 July 2003

Protist

ORIGINAL PAPER

Genomic Organization of *Trypanosoma brucei* Kinetoplast DNA Minicircles

Min Hong[a,2] and Larry Simpson[a,b,1]

[a] Howard Hughes Medical Institute, UCLA, 6780 MacDonald Research Laboratories, 675 Charles E. Young Dr. S., Los Angeles, CA 90095-1662, USA
[b] Department of Microbiology, Immunology and Molecular Genetics, UCLA, Los Angeles, CA 90095, USA

Submitted December 27, 2002; Accepted March 17, 2003
Monitoring Editor: Michael Melkonian

The sequences of seven new *Trypanosoma brucei* kinetoplast DNA minicircles were obtained. A detailed comparative analysis of these sequences and those of the 18 complete kDNA minicircle sequences from *T. brucei* available in the database was performed. These 25 different minicircles contain 86 putative gRNA genes. The number of gRNA genes per minicircle varies from 2 to 5. In most cases, the genes are located between short imperfect inverted repeats, but in several minicircles there are inverted repeat cassettes that did not contain identifiable gRNA genes. Five minicircles contain single gRNA genes not surrounded by identifiable repeats. Two pairs of closely related minicircles may have recently evolved from common ancestors: KTMH1 and KTMH3 contained the same gRNA genes in the same order, whereas KTCSGRA and KTCSGRB contained two gRNA genes in the same order and one gRNA gene specific to each. All minicircles could be classified into two classes on the basis of a short substitution within the highly conserved region, but the minicircles in these two classes did not appear to differ in terms of gRNA content or gene organization. A number of redundant gRNAs containing identical editing information but different sequences were present. The alignments of the predicted gRNAs with the edited mRNA sequences varied from a perfect alignment without gaps to alignments with multiple mismatches. Multiple gRNAs overlapped with upstream gRNAs, but in no case was a complete set of overlapping gRNAs covering an entire editing domain obtained. We estimate that a minimum set of approximately 65 additional gRNAs would be required for complete overlapping sets. This analysis should provide a basis for detailed studies of the evolution and role in RNA editing of kDNA minicircles in this species.

Introduction

The mitochondrial genome of trypanosomatid protists is composed of two types of DNA molecules – minicircles and maxicircles. The approximately 30–50 maxicircles encode two small rRNAs and 18 structural genes, 12 of which are cryptogenes, the transcripts of which must be edited to be translatable (Estévez and Simpson 1999). The maxicircle of *Leishmania tarentolae* also contains approximately 15 guide RNAs (gRNAs) which encode the information for the precise insertion and deletion of Us during editing (Blum et al. 1990). The maxicircle of *Trypanosoma brucei* appears to encode only 3 or 4 putative gRNAs.

[1] Corresponding author;
fax 1310 206 8967
e-mail simpson@kdna.ucla.edu

[2] Current address;
Biochemistry and Molecular Biology Department, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA

1434-4610/03/154/02-265 $ 15.00/0

The approximately 5,000–10,000 minicircles encode the majority of the gRNAs. All the minicircles and maxicircles are catenated together into a single giant DNA network (Shapiro and Englund 1995). The genomic organization and complexity of the minicircle DNA varies from species to species. In *L. tarentolae*, there is a single gRNA gene per ~900 bp minicircle situated within the variable region (Sturm and Simpson 1990). There is also a single conserved region per molecule, which contains three more highly conserved short sequences that are involved in DNA replication initiation (Ray 1989), and a short adjacent region of bent DNA of unknown function. A minicircle sequence class is defined by the encoded gRNA embedded in the identical variable region sequence. The number of minicircle sequence classes, which determines the number of different gRNAs available for editing, varies from 17 in the old UC lab strain (Maslov and Simpson 1992) to over 80 in the recently isolated LEM125 strain (Gao et al. 2001; Thiemann et al. 1994).

The *T. brucei* minicircle is somewhat larger (~1000 bp) but is organized similarly with a single conserved region and a single bend region; however, within the variable region there are 3–4 gRNA genes (Corell et al. 1993; Pollard et al. 1990; Pollard and Hajduk 1991) that are frequently situated between 18-mer inverted repeats of unknown function (Jasmer and Stuart 1986). It has been frequently stated, mainly on the basis of an early Cot analysis (Steinert and Van Assel 1980), that there are over 200 different minicircle sequence classes in this species. However, only 18 complete minicircle sequences are currently available in the database. These minicircles encode a total of 62 putative gRNAs, some of which are redundant, in that they have different sequences but encode the same editing information.

In order to extend our knowledge of the genomic organization and complexity of minicircle DNA in *T. brucei*, we have cloned and sequenced seven new minicircles and have performed a comparative analysis of all the existing sequence data.

Results

Cloning of New *T. brucei* Minicircles

Three minicircle libraries of *T. brucei* kDNA digested with EcoR I, Hind III, and Sst I, respectively, were constructed. Clones with ~1 kb minicircle inserts were selected by colony hybridization using a probe for the CSB-3 conserved region sequence. Due to the presence of several unrelated gRNA genes per molecule, it was not possible to employ the negative selection method that had worked well in the case of *Leishmania* minicircles which have a single gRNA gene per minicircle. Inserts were randomly selected and sequenced in both directions. This method proved to be inefficient since most of the clones represented previously sequenced minicircles. Seven complete sequences and two partial sequences of novel minicircles were obtained.

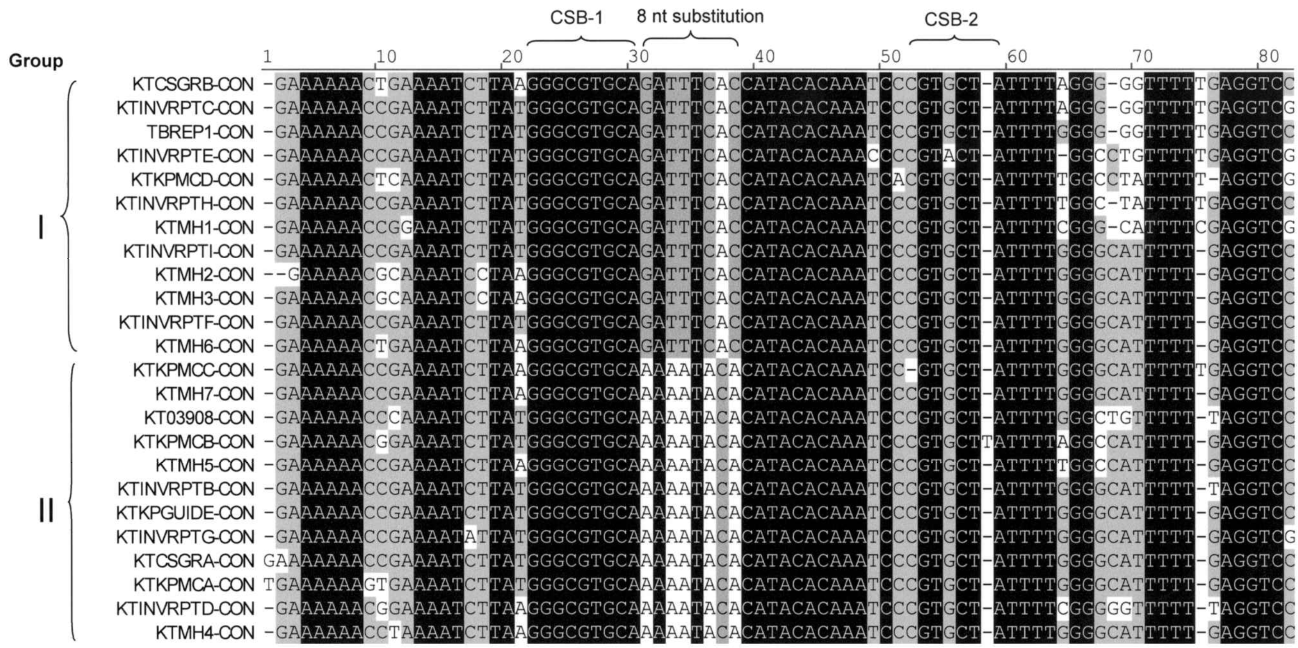
Minicircles can be Separated into Two Groups on Basis of an Eight Nucleotide Substitution in Conserved Region

The alignment of the entire conserved regions (CON) of all minicircles is shown in Figure 1. The three highly conserved sequences, CSB-1–3, are indicated by brackets. Between CSB-1 and CSB-2, there is an 8 nt substitution in some of the minicircle sequences (indicated by bracket in Fig. 3). Based on this substitution, the 25 minicircles fell into two groups. There are however no identifiable differences between the Group I and Group II minicircles in terms of gRNA gene content or organization (Figs 2, 3).

Comparison of Genomic Maps of 25 Minicircles

The sequences of 18 minicircles in GenBank and the seven new minicircles sequenced in our laboratory were analyzed for the presence of the three short conserved sequences, CSB-1, CSB-2 and CSB-3, for putative gRNA genes, and for inverted repeats around 18 nt in length. The results are shown in Figure 2. All sequences were linearized at the CSB-1 sequence for ease of comparison. The overall genomic organization of all the minicircles is similar but not identical. The location and polarity of the CSB-1–3 sequences is identical. The CSB-3 sequence was missing from two minicircles – KTINVRPTF and KTINVRPTI. Since CSB-3 serves as an origin of replication, it is likely that this is due to sequencing errors. The number of identifiable gRNA genes varied from 2 to 5 per circle. Eighty six total putative gRNA genes were identified, including three from a partial minicircle sequence (KTINVRPTA). All genes were in the same polarity.

In most cases the gRNA genes were situated between short inverted repeat sequences (forward repeats labeled RPa–d, backward repeats labeled RP1–4). However, in several cases, no identifiable repeat sequences were present flanking some gRNA genes, and in several instances no identifiable gRNA genes were found between inverted repeats.



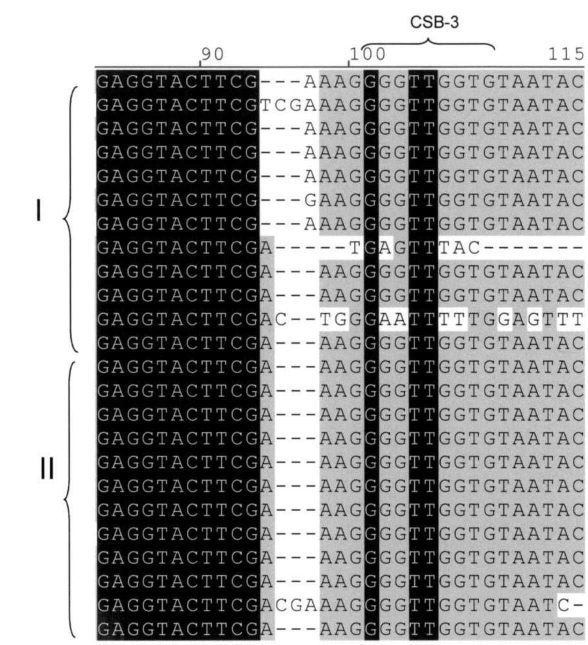


Figure 1. Alignments of conserved regions of 25 minicircles. The locations of the CSB-1, -2 and -3 sequences are indicated by brackets. The sequences have been grouped according to the presence of the 8 nt substitution between CSB-1 and CSB-2. I, group I minicircles. II, group II minicircles.

Seven cassettes contained sequences which were complementary to portions of the never edited region of TbMURF2 (in minicircles KTMH6, TBREP2, KTPMCC, KTUNVRPH, KTMH7, KTINVRPTE. These were not indicated in the maps in Figure 2 since such transcripts, if they exist, would not mediate editing events. We did not investigate if these sequences were actually transcribed.

Two Pairs of Closely Related Minicircle Sequences

The KTMH1 and KTMH3 minicircles appeared to be very similar but not identical in sequence, and had the same five encoded gRNAs in the same organization (Fig. 3A). These minicircles most likely were derived from a recent duplication.

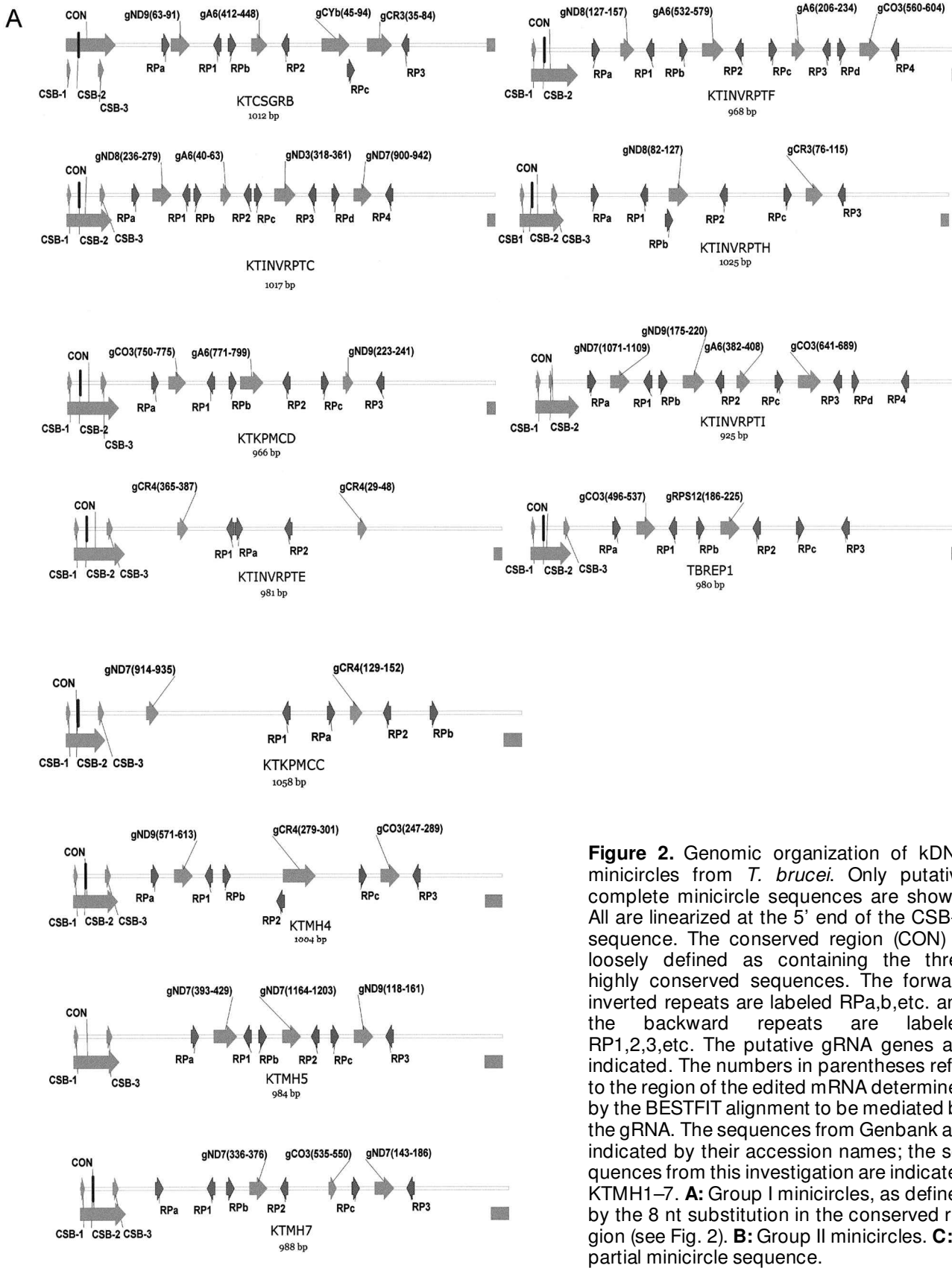
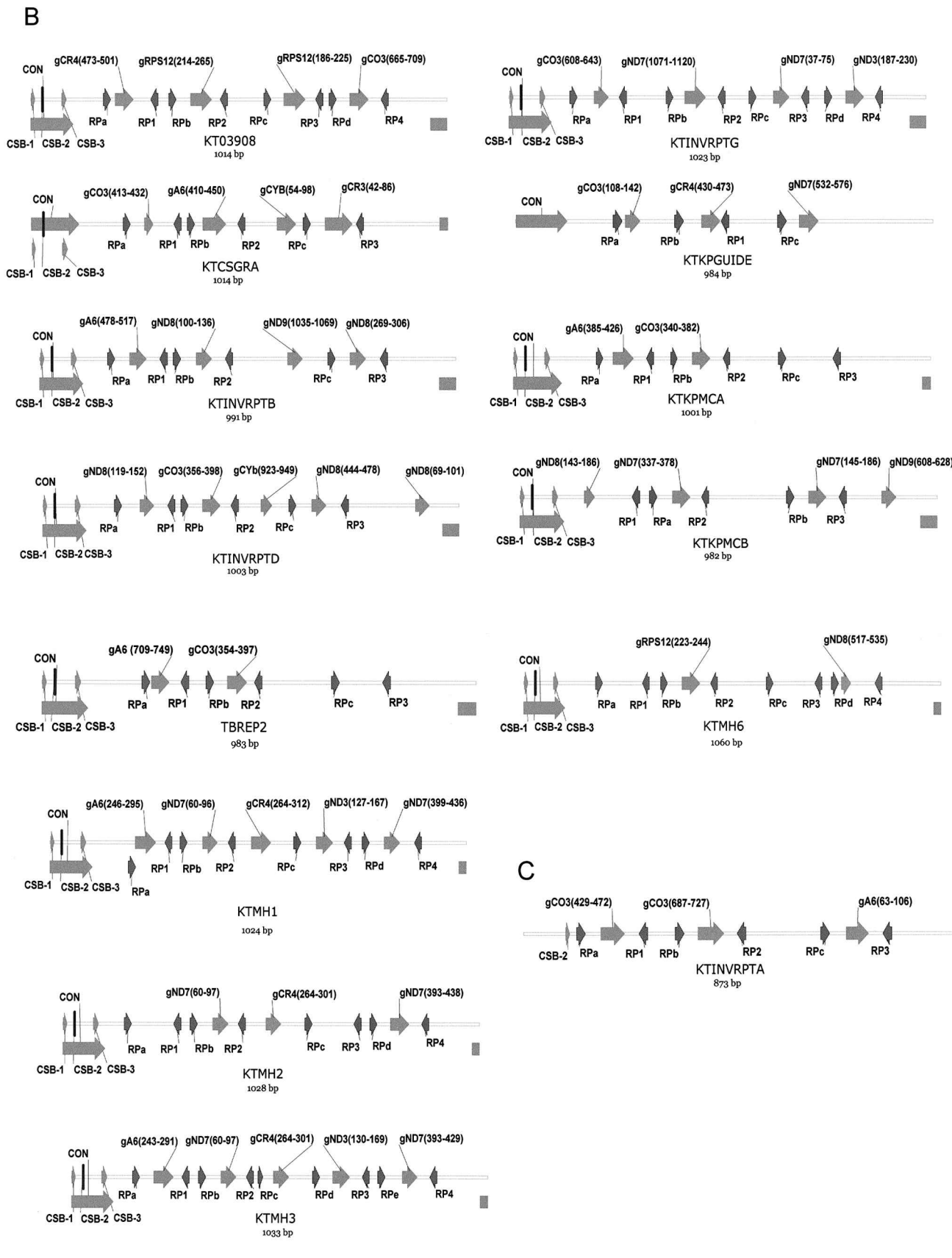


Figure 2. Genomic organization of kDNA minicircles from *T. brucei*. Only putative complete minicircle sequences are shown. All are linearized at the 5' end of the CSB-1 sequence. The conserved region (CON) is loosely defined as containing the three highly conserved sequences. The forward inverted repeats are labeled RPa,b,etc. and the backward repeats are labeled RP1,2,3,etc. The putative gRNA genes are indicated. The numbers in parentheses refer to the region of the edited mRNA determined by the BESTFIT alignment to be mediated by the gRNA. The sequences from Genbank are indicated by their accession names; the sequences from this investigation are indicated KTMH1–7. **A:** Group I minicircles, as defined by the 8 nt substitution in the conserved region (see Fig. 2). **B:** Group II minicircles. **C:** A partial minicircle sequence.


B
CON
gCR4(473-501)
gRPS12(214-265)
gRPS12(186-225)
gCO3(665-709)
RPa
RP1
RPb
RP2
RPc
RP3
RPd
RP4
CSB-1
CSB-2
CSB-3
KT03908
1014 bp
gCO3(608-643)
gND7(1071-1120)
gND7(37-75)
gND3(187-230)
KTINVRPTG
1023 bp
gCO3(413-432)
gA6(410-450)
gCYB(54-98)
gCR3(42-86)
KTCSGRA
1014 bp
gCO3(108-142)
gCR4(430-473)
gND7(532-576)
KTKPGUIDE
984 bp
gA6(478-517)
gND8(100-136)
gND9(1035-1069)
gND8(269-306)
KTINVRPTB
991 bp
gA6(385-426)
gCO3(340-382)
KTKPMCA
1001 bp
gND8(119-152)
gCO3(356-398)
gCYb(923-949)
gND8(444-478)
gND8(69-101)
KTINVRPTD
1003 bp
gND8(143-186)
gND7(337-378)
gND7(145-186)
gND9(608-628)
KTKPMCB
982 bp
gA6 (709-749)
gCO3(354-397)
TBREP2
983 bp
gRPS12(223-244)
gND8(517-535)
KTMH6
1060 bp
gA6(246-295)
gND7(60-96)
gCR4(264-312)
gND3(127-167)
gND7(399-436)
KTMH1
1024 bp
C
gCO3(429-472)
gCO3(687-727)
gA6(63-106)
KTINVRPTA
873 bp
gND7(60-97)
gCR4(264-301)
gND7(393-438)
KTMH2
1028 bp
gA6(243-291)
gND7(60-97)
gCR4(264-301)
gND3(130-169)
gND7(393-429)
RPe
KTMH3
1033 bp

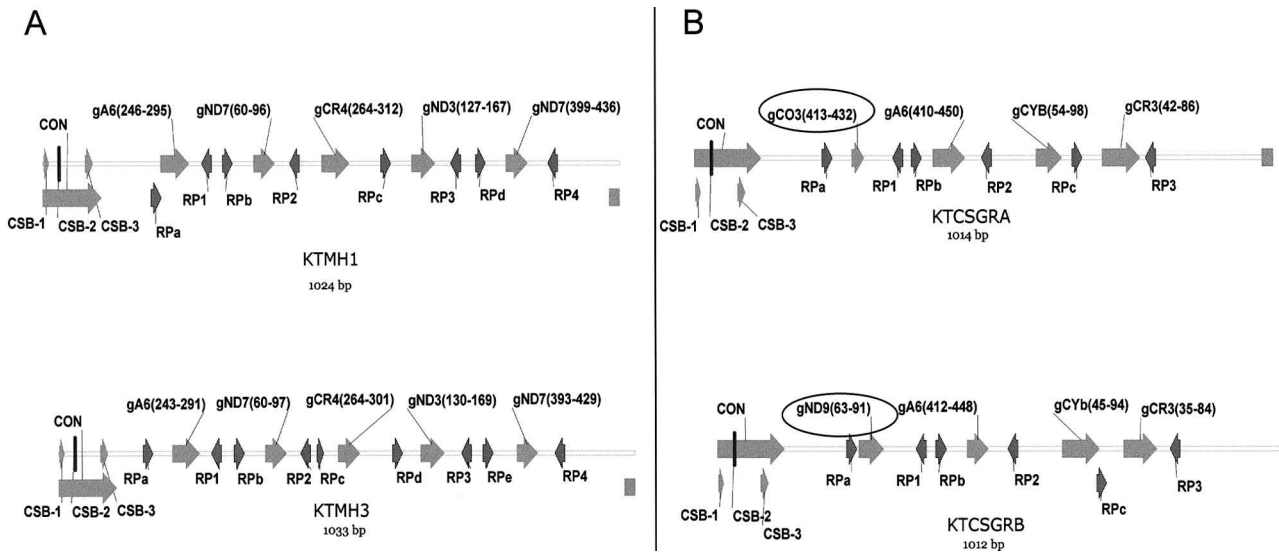


Figure 3. Two pairs of closely related minicircles. Diagrams of the linearized minicircles from Figure 1 are compared side by side. **A:** KTMH1 and KTMH3 minicircles. **B:** KTCSGRA and KTCSGRB minicircles. The two different gRNA genes in cassette 1 are circled.

Likewise, the KTCSGRA minicircle and the KTCSGRB minicircle sequences were highly similar. However, the first cassettes contained different gRNA genes whereas the proximal three cassettes contained the same gRNA genes (Fig. 3B). This may suggest that the gRNA genes are mobile elements that can move between minicircle classes.

Alignment of the Short Inverted Repeats

Identification of the short inverted repeats was often problematical due to the numerous sequence polymorphisms in these sequences. Alignments of all identified forward repeats and all identifiable backward repeats are shown in Figure 4. There are no absolutely conserved motifs other than a single T residue. Again, we cannot distinguish between sequencing errors and actual polymorphisms.

Alignments of Putative gRNAs and Edited mRNA Sequences

Alignments of all of the putative gRNAs with the edited mRNA sequences are shown in the Appendix. In no case is there a complete set of overlapping gRNAs which could account for all of the editing for any gene, as has been obtained with *L. tarentolae*.

The extent of redundant gRNAs, which are defined as gRNAs encoding the same editing information but having different sequences, is striking. There are 13 sets of redundant gRNAs, in most cases consisting of two molecules.

The extent of mismatches between the gRNAs and the edited mRNAs is also striking. Again the contribution of sequencing errors to these mismatches is not known. Nor is it known how many mismatches prevent a gRNA from functioning in mediating editing in vivo.

Discussion

We initially hoped to completely sequence the kinetoplast minicircle genome of *T. brucei* and obtain complete sets of overlapping and redundant gRNAs, but this proved difficult due to the lack of a method for negative selection of the minicircle DNA clones. The seven new sequences we obtained plus the 18 sequences in Genbank contain a total of 86 putative gRNA genes. Many of these are highly overlapping and can be classified as redundant, and several are overlapping sufficiently as to extend an editing cascade in the 5′ direction. Redundant minicircle-encoded gRNAs are also present in a recently isolated strain of *L. tarentolae* but are rare in the old UC lab strain, in which the complement of minicircle-encoded gRNAs has been apparently minimized perhaps due to loss of non-essential minicircle sequence classes on cell division (Savill and Higgs 1999; Simpson et al. 2000).

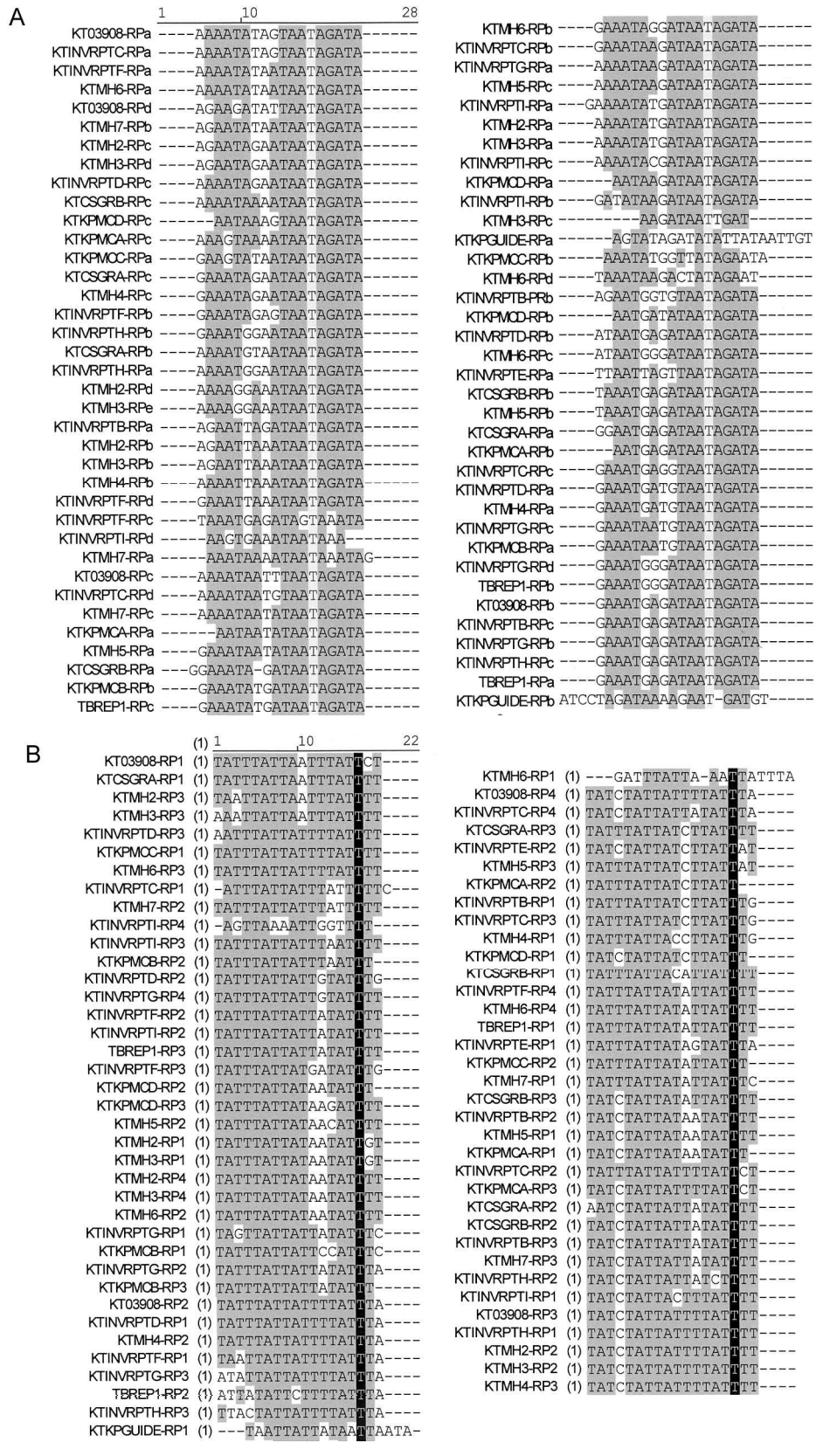


Figure 4. Alignments of inverted repeat sequences. **A:** Forward repeats. **B:** Backward repeats.

From the number and size of the edited sequences not covered by known gRNAs, it can be estimated that approximately an additional 65 gRNAs, or another 16 minicircle classes, would be minimally required to complete the overlaps. Even allowing for the presence of additional multiple redundant gRNAs, the total number of minicircle sequence classes is not likely to be greater than 80–100, which is equivalent to the minicircle complexity found in the LEM125 strain of *L. tarentolae*. It appears from this data that the generally used value of over 200 sequence classes in *T. brucei* is an overestimate.

The overall pattern of gene organization in the 25 minicircles is similar but the extent of genomic plasticity in the coding regions is striking. The number of cassettes defined by short inverted repeats varies from one to five, and in some cases the cassettes appear to be empty of encoded gRNAs. Some gRNA genes are not enclosed within a cassette.

Seven cassettes contained putative gRNA genes for various blocks within the never edited region of TbMURF2. It would be of some interest if these putative nonfunctional genes were actually transcribed, but this was not examined.

The function of the inverted repeats is not clear. They may represent remnants of transposition events or they may represent regulatory sequences. The absence of repeats before some gRNA genes casts doubt on the latter hypothesis, unless these represent non-transcribed genes. Further experimental work is required to determine the biological function of the inverted repeats.

This analysis of all of the known *T. brucei* minicircle sequences in terms of genomic organization should prove valuable for understanding the function and evolution of this unusual molecule.

Methods

Cloning and sequencing: Total kinetoplast DNA was isolated from strain 427 procyclic *T. brucei* growing in SDM medium at 27 °C. The CsCl – ethidium bromide rapid sedimentation method (Sturm et al. 1989) was used to isolate network kDNA. The kDNA was digested with EcoRI, SstI or HindIII, and the linearized minicircles cloned into the Bluescript SK(+) plasmid. Clones containing minicircle inserts were selected by colony hybridization using a probe specific for the CSB-3 conserved sequence. DNA was isolated from each clone and subjected to dye termination sequencing using T3 and T7 primers. Overlaps were obtained by using internal primers determined from the sequences. The ABI trace files were analyzed for overlaps using the Seqman program from the DNASTAR package.

Analysis of minicircle sequences: Putative gRNA genes were identified as described previously. A file containing all *T. brucei* edited mRNA sequences in tandem was first created (available at http://www.hhmi.ucla.edu/simpson/supplement/supple.htm) as well as separate GCG files with the edited sequence of each gene (available at http://www.rna.ucla.edu/trypanosome/database.html). Then the minicircle sequences were analyzed for the presence of CSB-1,-2 and -3 conserved sequences. This defined the conserved region. The entire variable region was then divided into four portions and each was searched for the presence of putative gRNA genes: the minicircle sequence was reversed and ran against the file containing all edited sequences in tandem using the GCG Bestfit alignment program and the “comp” template matrix (available at http://www.hhmi.ucla.edu/simpson/supplement/supple.htm). This led to the identification of the putative gRNA mediating editing of a specific gene. The same alignment procedure was then run using the putative gRNA sequence and the file of the specific edited gene to obtain the coordinates. Each putative gRNA was identified by the 3’ and 5’ nucleotide numbers of the specific edited gene file. Putative gRNA genes were only found in the same polarity as the CSB-1-2-3 sequences.

Multiple sequence alignments were performed using the AlignX program from Vector NTI. All utilized GenBank files of complete *T. brucei* minicircle sequences can be found at http://www.hhmi.ucla.edu/simpson/supplement/supple.htm. The published sequences were from several *T. brucei* strains:

1. TBU03908, strain = IsTat 1.1
2. TRBCSGRA;TRBCSGRB;TRBCSGRC, strain = EATRO 164
3. TRBKPMCA; TRBKPMCB;TRBKPMCC;TRBKPMCD, strain = EATRO 164.
4. TRBKPGUIDE, strain = EATRO 164
5. TBREP1; TBREP2, unknown strain
6. KTINVRPTA – KTINVRPTI, strain = EATRO 164

The short inverted repeats were identified by first searching for short consensus sequences determined from the already identified repeats and were refined by analysis of multiple alignments and locations within the variable region. The sequence TAATA(G/A)ATA was initially used to search for forward repeats, and the sequence (T/A) (A/T) (T/A) (T/C) TATTA was used to search for backward repeats.

References

Blum B, Bakalara N, Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: “Guide” RNA molecules transcribed from maxicircle DNA provide the edited information. Cell **60**: 189–198

Corell RA, Feagin JE, Riley GR, Strickland T, Guderian JA, Myler PJ, Stuart K (1993) *Trypanosoma brucei* minicircles encode multiple guide RNAs which can direct editing of extensively overlapping sequences. Nucleic Acids Res **21**: 4313–4320

Estévez AM, Simpson L (1999) Uridine insertion/deletion RNA editing in trypanosome mitochondria – a review. Gene **240**: 247–260

Gao GH, Kapushoc ST, Simpson AM, Thiemann OH, Simpson L (2001) Guide RNAs of the recently isolated LEM125 strain of *Leishmania tarentolae*: An unexpected complexity. RNA **7**: 1335–1347

Jasmer D, Stuart K (1986) Sequence organization in African trypanosome minicircles is defined by 18 base pair inverted repeats. Mol Biochem Parasitol **18**: 321–332

Maslov DA, Simpson L (1992) The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. Cell **70**: 459–467

Pollard VW, Hajduk SL (1991) *Trypanosoma equiperdum* minicircles encode three distinct primary transcripts which exhibit guide RNA characteristics. Mol Cell Biol **11**: 1668–1675

Pollard VW, Rohrer SP, Michelotti EF, Hancock K, Hajduk SL (1990) Organization of minicircle genes for guide RNAs in *Trypanosoma brucei.* Cell **63**: 783–790

Ray D (1989) Conserved sequence blocks in kinetoplast DNA minicircles from diverse species of trypanosomes. Mol Cell Biol **9**: 1365–1367

Savill NJ, Higgs PG (1999) A theoretical study of random segregation of minicircles in trypanosomatids. Proc R Soc Lond B Biol Sci **266**: 611–620

Shapiro TA, Englund PT (1995) The structure and replication of kinetoplast DNA. Annu Rev Microbiol **49**: 117–143

Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA (2000) Evolution of RNA editing in trypanosome mitochondria. Proc Natl Acad Sci U S A **97**: 6986–6993

Steinert M, Van Assel S (1980) Sequence heterogeneity in kinetoplast DNA: reassociation kinetics. Plasmid **3**: 7–17

Sturm NR, Simpson L (1990) Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. Cell **61**: 879–884

Sturm NR, Degrave W, Morel C, Simpson L (1989) Sensitive detection and schizodeme classification of *Trypanosoma cruzi* cells by amplification of kinetoplast minicircle DNA sequences: use in diagnosis of Chagas Disease. Mol Biochem Parasitol **33**: 205–214

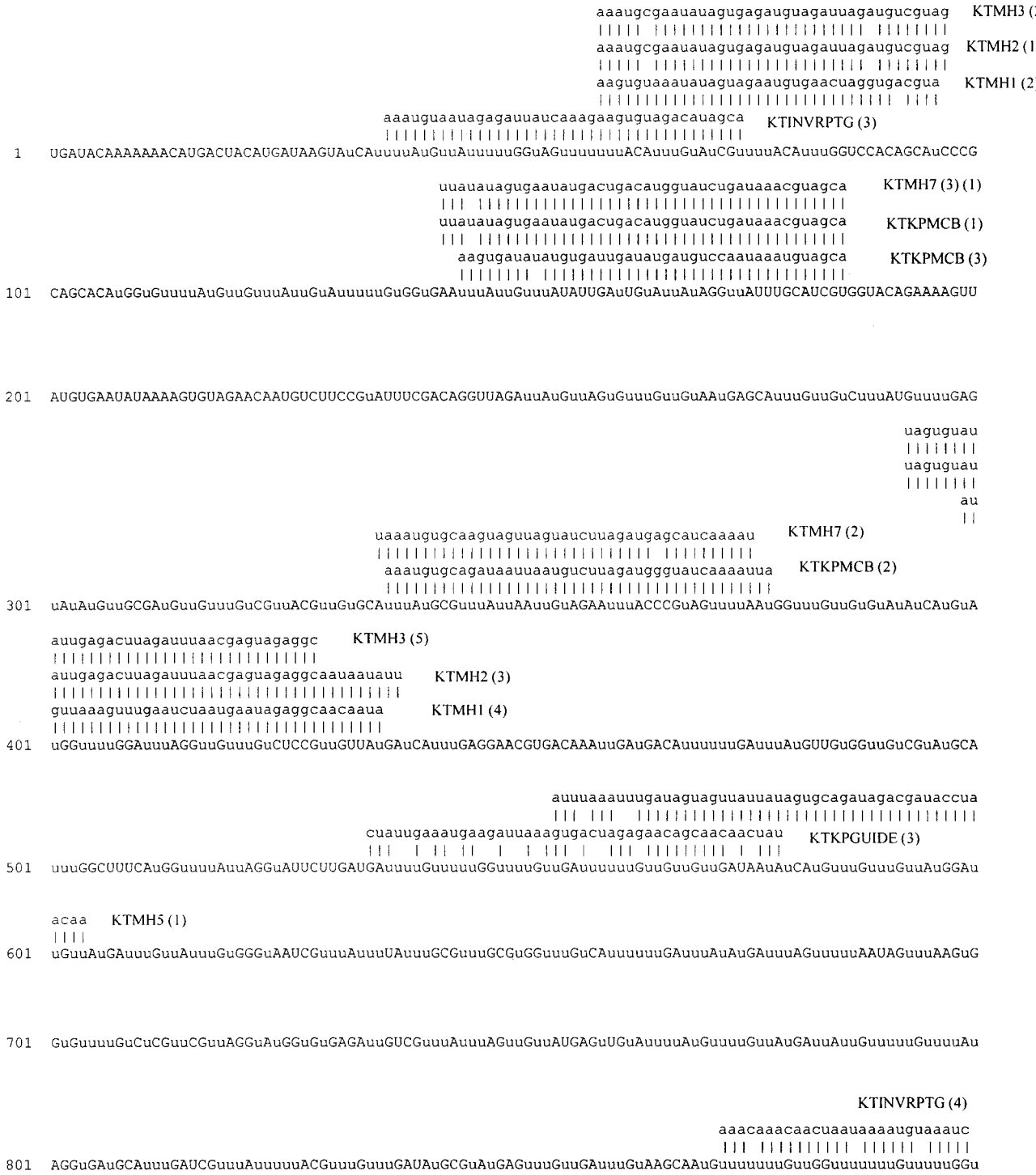
Thiemann OH, Maslov DA, Simpson L (1994) Disruption of RNA editing in *Leishmania tarentolae* by the loss of minicircle-encoded guide RNA genes. EMBO J **13**: 5689–5700

Appendix: Alignments of putative gRNAs and edited mRNA sequences. A–J. Specific edited mRNA sequences aligned with the putative gRNAs mediating editing. Genomically encoded nucleotides are in upper case and Us added by editing are in lower case. U-deletions are not indicated. The specific minicircle encoding the putative gRNA is shown adjacent to the gRNA sequence and the relative number of the gRNA cassette is indicated in parenthesis. A–U, G–C and G–U base pairs are shown by |.
See http://164.67.60.200/tbrucei/default.htm for a wordfile of the Appendix.

TbND3

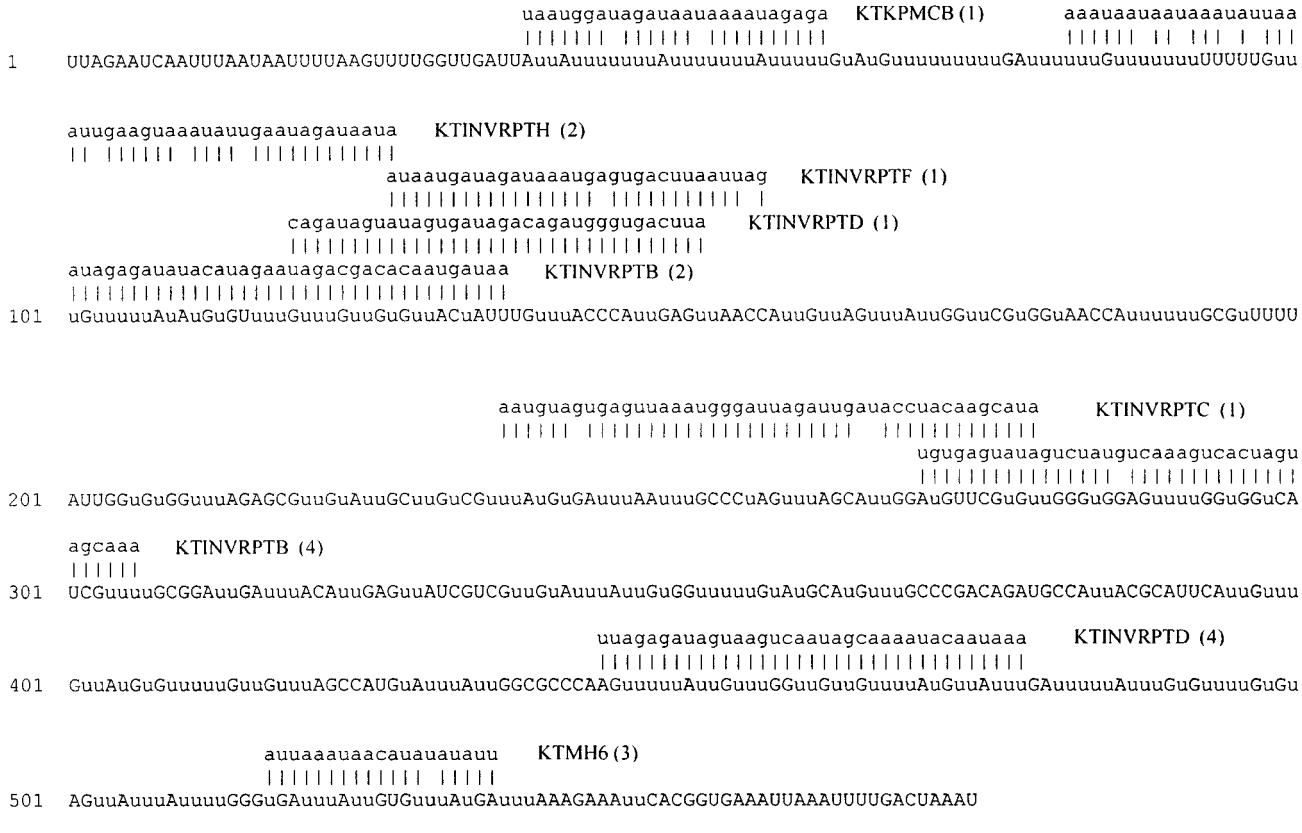
```
1   UCAAAAAAUCCUCGCCUUUUUACUUUAGUUUGUUAUCAuuAuuuuuAuAuuuGuuuuUGAUAuuGuGGuuuAUUAuuuuAuuuAuAGGuuuuuuuuuAuG

                              aauugauagaaguaacgguagauacuaaaggcacacuaaac    KTMH3 (4)
                              ||||||||||||||||||||||||||||||||||||||||||
                              aauugauagaaguaacgguagauacuaaaggcacacuaaac    KTMH1 (4)
                              ||||||||||||||||||||||||||||||||||||||||||
101 uuuuuuAuGuuuuuuAuuGCAuuuuuuuuGAuuGuuuuCGuuGuuGuuuGuGGuuuuCGuGuGGuUUGuAuGAuAuGAAuUCACGuuuGGUGuuuuAuACA

201 uuGGAuuuAUGuuuuGuuAGuUGuUUGuuuuuuGuAuuGuuAAAuuCCAuUAuuuGuGuUUUGuuGuuuGuuuuuGUGAuAGuGuuGuuuuAuuuuuGuu

                         auuaagaguagaagguagcauguauauucuaaguacaccauaaa      KTINVRPTC (3)
                         ||||||||||||||||||||||||||||||||||||||||||||
301 AUGGuuuuuuGuUUUUGuGGuuuuuGuuuuuuGuuGuAuGuAuAGGAuuUGuGuGGuAuuuuuGGGAUCACGuAuAUUUGUGUGGUGUAAUUUUAuuuuG

401 uuuAuGAUGuuuUUUGUUGUAUUAUACAUAUUAUAUUAAUAAAUAUAUAAAAUAAUAAAAUACUA
```

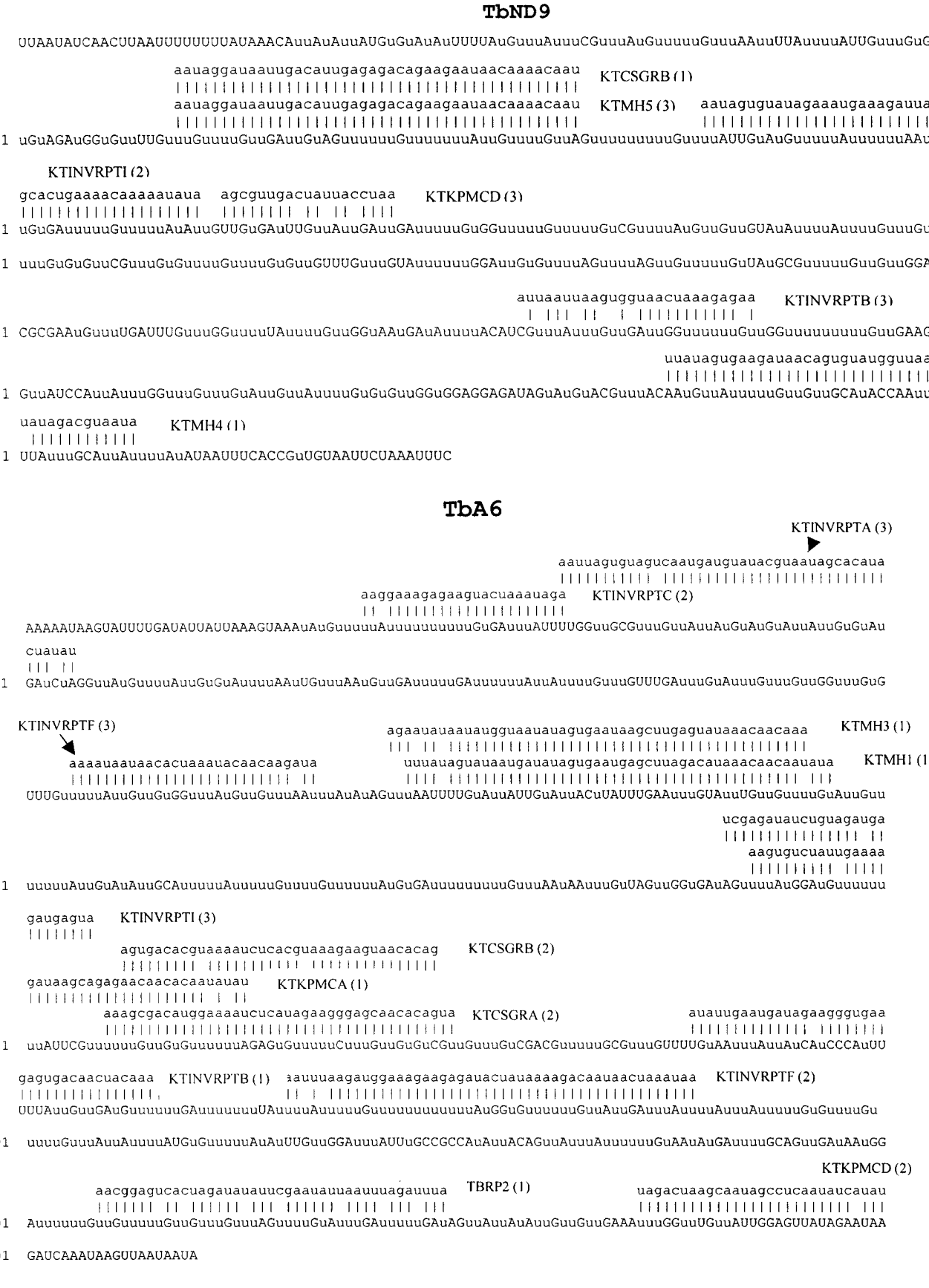
TbND7

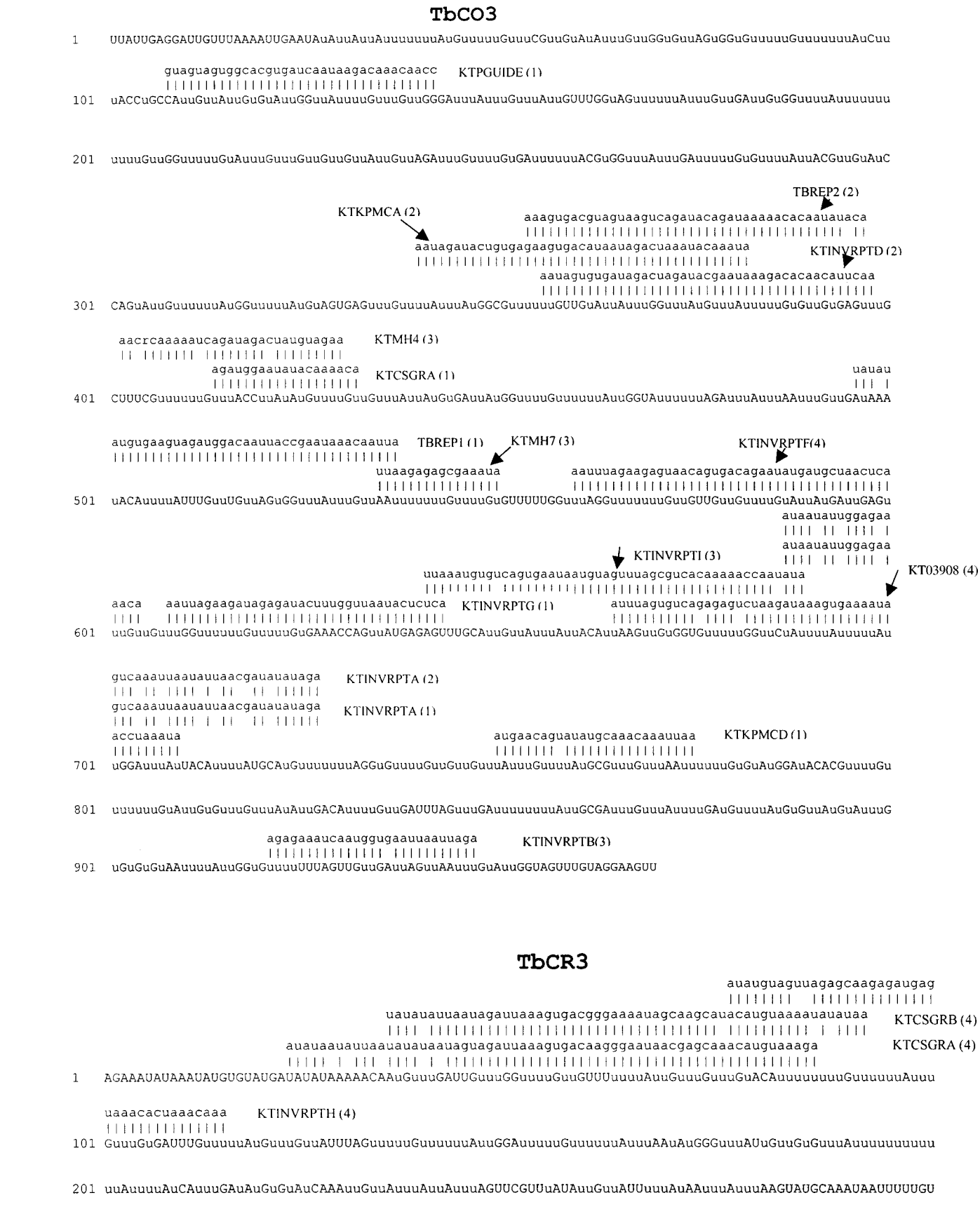

aaaugcgaauauagugagauguagauuagaugucguag KTMH3 (2)
aaaugcgaauauagugagauguagauuagaugucguag KTMH2 (1)
aaguguaaauauaguagaaugugaacuaggugacgua KTMH1 (2)
aaauguaauagagauuaucaaagaaguguagacauagca KTINVRPTG (3)
1 UGAUACAAAAAAACAUGACUACAUGAUAAGUAuCAuuuuAuGuuAuuuuuGGuAGuuuuuuuuACAuuuGuAuCGuuuuACAuuuGGUCCACAGCAuCCCG
uuauauagugaauaugacugacaugguaucugauaaacguagca KTMH7 (3) (1)
uuauauagugaauaugacugacaugguaucugauaaacguagca KTKPMCB (1)
aagugauauaugugauugauaugauguccaauaaaauguagca KTKPMCB (3)
101 CAGCACAuGGuGuuuuAuGuuGuuuAuuGuAuuuuuGuGGuGAAuuuAuuGuuuAUAUUGAuUGuAuuAuAGGuuAUUUGCAUCGUGGUACAGAAAAGUU
201 AUGUGAAUAUAAAAGUGUAGAACAAUGUCUUCCGuAUUUCGACAGGUUAGAuuAuGuuAGuGuuuGuuGuAAuGAGCAuuuGuuGuCuuuAUGuuuuGAG
uaguguau
uaguguau
au
uaaaugugcaaguaguuaguaucuuagaugagcaucaaaau KTMH7 (2)
aaaugugcagauaauuaaugucuuagaugggguaucaaaauua KTKPMCB (2)
301 uAuAuGuuGCGAuGuuGuuuGuCGuuACGuuGuGCAuuuuAuGCGuuuAuuAAuuGuAGAAuuuACCCGuAGuuuuAAuGGuuuGuuGuGuAuAuCAuGuA
auugagacuuagauuuaacgaguagaggc KTMH3 (5)
auugagacuuagauuuaacgaguagaggcaauaauauu KTMH2 (3)
guuaaaguuugaaucuaaugaauagaggcaacaaua KTMH1 (4)
401 uGGuuuuGGAuuuAGGuuGuuuGuCUCCGuuGUUAuGAuCAuuuGAGGAACGUGACAAAuuGAuGACAuuuuuuGAuuuAuGUUGuGGuuGuCGuAuGCA
auuuaaauuugauaguaguuauuauagugcagauagacgauaccua
cuauugaaaugaagauuaaagugacuagagaacagcaacaacuau KTKPGUIDE (3)
501 uuuGGCUUUCAuGGuuuuAuuAGGuAUUCUUGAUGAuuuuGuuuuuGGuuuuGuuGAuuuuuuGuuGuuGuuGAUAAuAuCAuGuuuGuuuGuuAuGGAu
acaa KTMH5 (1)
601 uGuuAuGAuuuGuuAuuuGuGGGuAAUCGuuuAuuuUAuuuGCGuuuGCGuGGuuuGuCAuuuuuuGAuuuAuAuGAuuuAGuuuuuAAUAGuuuAAGuG
701 GuGuuuuGuCuCGuuCGuuAGGuAuGGuGuGAGAuuGUCGuuuAuuuAGuuGuuAUGAGuUGuAuuuuAuGuuuuGuuAuGAuuAuuGuuuuuGuuuuAu
KTINVRPTG (4)
aaacaaacaacuaauaaaauguaaauc
801 AGGuGAuGCAuuuGAUCGuuuAuuuuuACGuuuGuuuGAUAuGCGuAuGAGuuuGuuGAuuuGuAAGCAAuGuuuuuuuGuuGGuuuuuuuGuuuuuGGu


aauuaauauauauaauacugca KTKPMCC (1)
aaguagauaaaauaggcugauagaugugacacuauaauggaaa KINVRPTC (4)
901 uuuGuuuGuuuGuuuGAuuAuuuAuAuuGuGAuAuuACCAuuGAGACCAuuAuuAuGuuAuuuuAuAGuuuGuGGuGuuGuuGuuuGCCGGGuAuAUCAu
aauuguaaguguaauaugcucgagagacgu
aaauguagaugugacacggcugagagacgu
aauuguaaguguaauaugcucgagagacgu
1001 uuGCUUGUGuuGAACACCCCAAAGGuGAGuAuuGuuuGuuAuuAUGuuuuuGuGuuGGuuuAuGuuCUCGuuuACGuuuGCGuuGuGCGGAuuuuuuGCA
auaugcaaauaaccuauaaaa KINVRPTG (2)
auaaacaaa KTINVRPTI (1)
auaugcaaauaaccuauaaaa KTMH5 (2)
auuauaguuacaaugagagucuaggcuuauagcaauc
1101 UAUUUGuuuAuuGGAuGuuuGuuuGCGuGGuuuuuuAuuGCAuGAuuuAGuuGCCGuuuuAGGuAAuAuuGAuGuuGuuuuuGGAuCCGUAGAUCGuuAG
aaa KTINVRPTG (2)
1201 uuuuAuAuGuGAGGUUAUUGuAGGAUUGUUUAAAAUUGAAUAAAAA

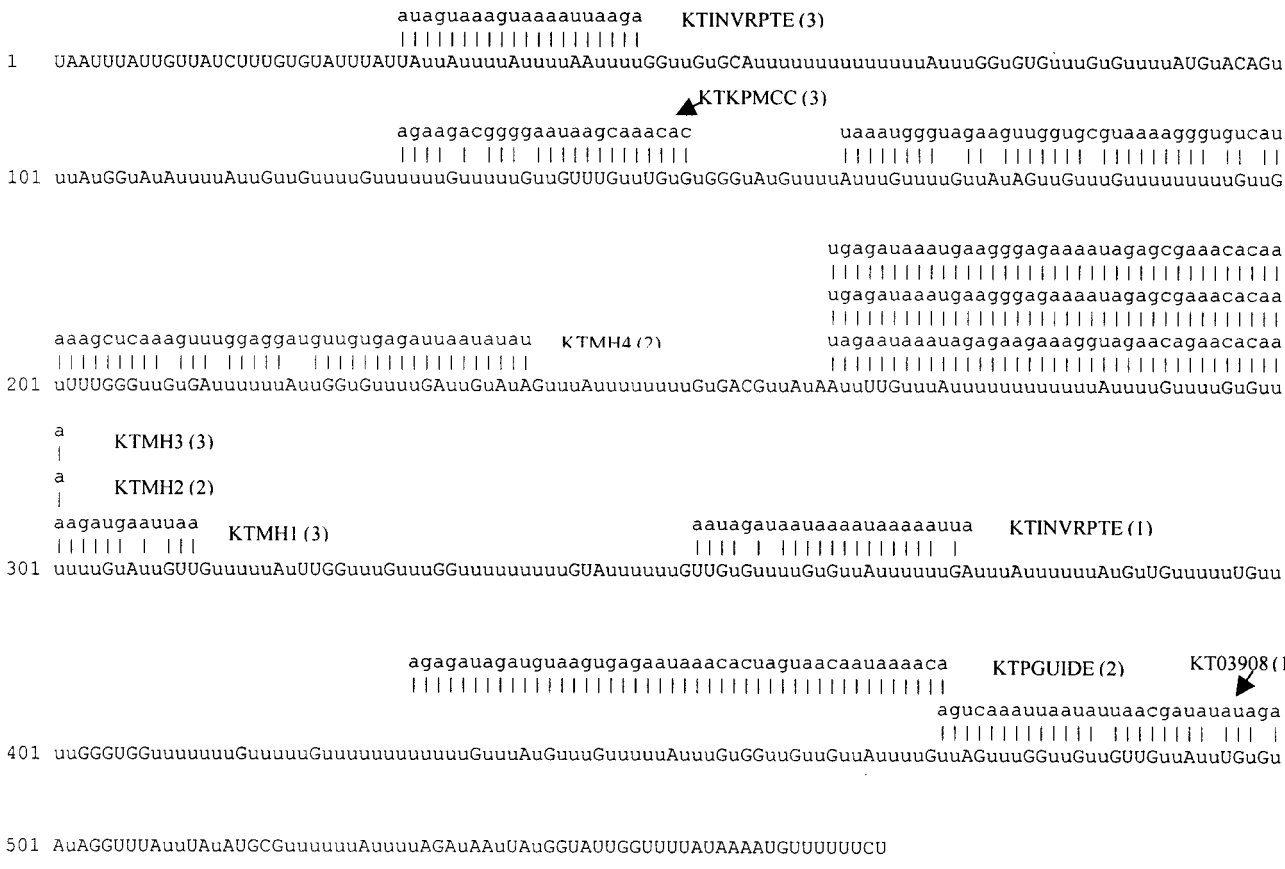
TbND8


uaauggauagauaauaaaauagaga KTKPMCB (1)
aaauaauaauaaauauuaa
1 UUAGAAUCAAUUUAAUAAUUUUAAGUUUUGGUUGAUUAuuAuuuuuuuAuuuuuuuAuuuuuGuAuGuuuuuuuuuGAuuuuuuGuuuuuuuUUUUUGuu
auugaaguaaaauauugaauagauaaua KTINVRPTH (2)
auaaugauagauaaaugagugacuuaauuag KTINVRPTF (1)
cagauaguauaugugauagacagaugggugacuua KTINVRPTD (1)
auagagauauacauagaauagacgacacaaugauaa KTINVRPTB (2)
101 uGuuuuuAuAuGuGUuuuGuuuGuuGuGuuACuAUUUGuuuACCCAuuGAGuuAACCAuuGuuAGuuuAuuGGuuCGuGGuAACCAuuuuuuGCGuUUUU
aauguagugaguuaaaugggauuagauugauaccuacaagcaua KTINVRPTC (1)
ugugaguauagucuaugucaaagucacuagu
201 AUUGGuGuGGuuuAGAGCGuuGuAuuGCuuGuCGuuuAuGuGAuuuAAuuuGCCCuAGuuuAGCAuuGGAuGUUCGuGuuGGGuGGAGuuuuGGuGGuCA
agcaaa KTINVRPTB (4)
301 UCGuuuuGCGGAuuGAuuuACAuuGAGuuAUCGUCGuuGuAuuuAuuGuGGuuuuuGuAuGCAuGuuuGCCCGACAGAUGCCAuuACGCAUUCAuuGuuu
uuagagauaguaagucaauagcaaaauacaauaaa KTINVRPTD (4)
401 GuuAuGuGuuuuuGuuGuuuAGCCAUGuAuuuAuuGGCGCCCAAGuuuuuAuuGuuuGGuuGuuGuuuuAuGuuAuuuGAuuuuuAuuuGuGuuuuGuGu
auuaaauaacauauauauu KTMH6 (3)
501 AGuuAuuuAuuuuGGGuGAuuuAuuGUGuuuAuGAuuuAAAGAAAuuCACGGUGAAAUUAAAUUUUGACUAAAU

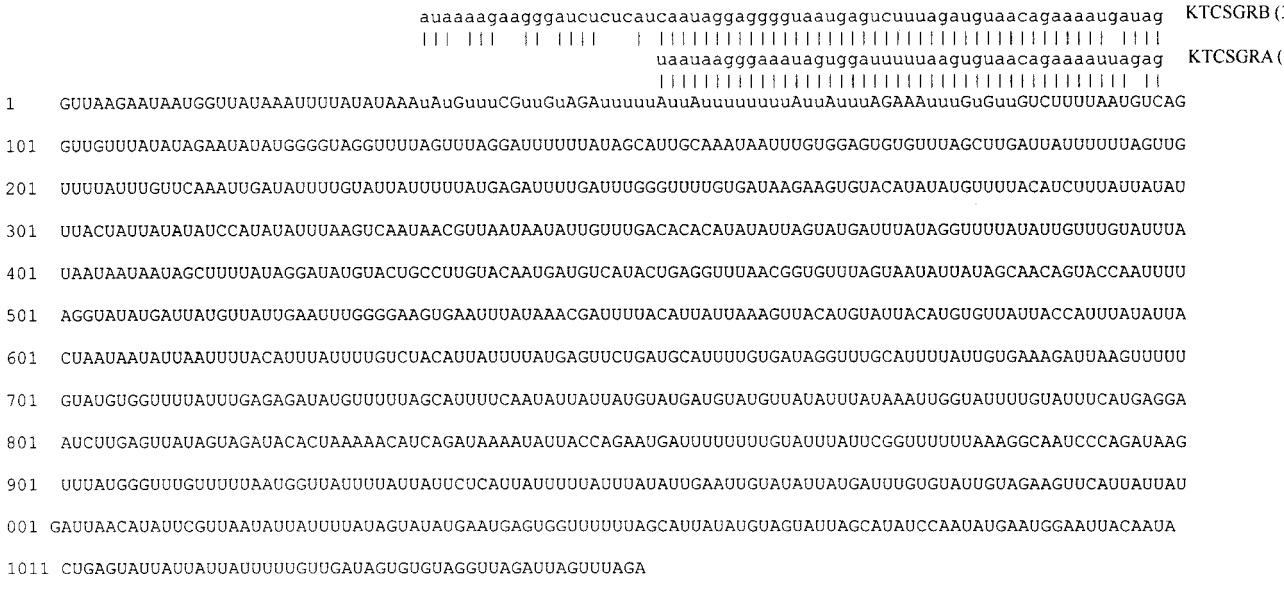

TbND9
KTCSGRB (1)
KTMH5 (3)
KTINVRPTI (2)
KTKPMCD (3)
KTINVRPTB (3)
KTMH4 (1)
TbA6
KTINVRPTA (3)
KTINVRPTC (2)
KTINVRPTF (3)
KTMH3 (1)
KTMH1 (1)
KTINVRPTI (3)
KTCSGRB (2)
KTKPMCA (1)
KTCSGRA (2)
KTINVRPTB (1)
KTINVRPTF (2)
KTKPMCD (2)
TBRP2 (1)


TbCO3
KTPGUIDE (1)
KTKPMCA (2)
TBREP2 (2)
KTINVRPTD (2)
KTMH4 (3)
KTCSGRA (1)
TBREP1 (1)
KTMH7 (3)
KTINVRPTF(4)
KTINVRPTI (3)
KT03908 (4)
KTINVRPTG (1)
KTINVRPTA (2)
KTINVRPTA (1)
KTKPMCD (1)
KTINVRPTB(3)
TbCR3
KTCSGRB (4)
KTCSGRA (4)
KTINVRPTH (4)

TbCR4


auaguaaaguaaaauuaaga
KTINVRPTE (3)
1 UAAUUUAUUGUUAUCUUUGUGUAUUUAUUAuuAuuuuAuuuuAAuuuuGGuuGuGCAuuuuuuuuuuuuuuuAuuuGGuGUGuuuGuGuuuuAUGuACAGu
KTKPMCC (3)
agaagacggggaauaagcaaacac
uaaauggguagaaguuggugcguaaaagggugucau
101 uuAuGGuAuAuuuuAuuGuuGuuuuGuuuuuuGuuuuuGuuGUUUGuuUGuGuGGGuAuGuuuuAuuuGuuuuGuuAuAGuuGuuuGuuuuuuuuuGuuG
ugagauaaaugaagggagaaaauagagcgaaacacaa
ugagauaaaugaagggagaaaauagagcgaaacacaa
aaagcucaaaguuuggaggauguugugagauuaauauau
KTMH4 (2)
uagaauaaauagagaagaaagguagaacagaacacaa
201 uUUUGGGuuGuGAuuuuuuAuuGGuGuuuuGAuuGuAuAGuuuAuuuuuuuuuGuGACGuuAuAAuuUGuuuAuuuuuuuuuuuuuAuuuuGuuuuGuGuu
a
KTMH3 (3)
a
KTMH2 (2)
aagaugaauuaa
KTMH1 (3)
aauagauaauaaaauaaaaauua
KTINVRPTE (1)
301 uuuuGuAuuGUUGuuuuuAuUUGGuuuGuuuGGuuuuuuuuuuGUAuuuuuuGUUGuGuuuuGuGuuAuuuuuuGAuuuAuuuuuuAuGuUGuuuuuUGuu
agagauagauguaagugagaauaaacacuaguaacaauaaaaca
KTPGUIDE (2)
KT03908 (1)
agucaaauuaaauauuaacgauauauaga
401 uuGGGUGGuuuuuuuGuuuuuGuuuuuuuuuuuuuGuuuAuGuuuGuuuuuAuuuGuGGuuGuuGuuAuuuuGuuAGuuuGGuuGuuGUUGuuAuuUGuGu
501 AuAGGUUUAuuUAuAUGCGuuuuuuAuuuuAGAuAAuUAuGGUAUUGGUUUUAUAAAAUGUUUUUUCU

TbCyb


auaaaagaagggaucucucaucaauaggagggguaaugagucuuuagauguaacagaaaaugauag
KTCSGRB (3)
uaauaagggaaauaguggauuuuuaaguguaacagaaaauuagag
KTCSGRA (3)
1 GUUAAGAAUAAUGGUUAUAAAUUUUAUAUAAAuAuGuuuCGuuGuAGAuuuuuAuuAuuuuuuuuAuuAuuuAGAAAuuuGuGuuGUCUUUUAAUGUCAG
101 GUUGUUUAUAUAGAAUAUAUGGGGUAGGUUUUAGUUUAGGAUUUUUUAUAGCAUUGCAAAUAAUUUGUGGAGUGUGUUUAGCUUGAUUAUUUUUUAGUUG
201 UUUUAUUUGUUCAAAUUGAUAUUUUGUAUUAUUUUUAUGAGAUUUUGAUUUGGGUUUUGUGAUAAGAAGUGUACAUAUAUGUUUUACAUCUUUAUUAUAU
301 UUACUAUUAUAUAUCCAUAUAUUUAAGUCAAUAACGUUAAUAAUAUUGUUUGACACACAUAUAUUAGUAUGAUUUAUAGGUUUUAUAUUGUUUGUAUUUA
401 UAAUAAUAAUAGCUUUUAUAGGAUAUGUACUGCCUUGUACAAUGAUGUCAUACUGAGGUUUAACGGUGUUUAGUAAUAUUAUAGCAACAGUACCAAUUUU
501 AGGUAUAUGAUUAUGUUAUUGAAUUUGGGGAAGUGAAUUUAUAAACGAUUUUACAUUAUUAAAGUUACAUGUAUUACAUGUGUUAUUACCAUUUAUAUUA
601 CUAAUAAUAUUAAUUUUACAUUUAUUUUGUCUACAUUAUUUUAUGAGUUCUGAUGCAUUUUGUGAUAGGUUUGCAUUUUAUUGUGAAAGAUUAAGUUUUU
701 GUAUGUGGUUUUAUUUGAGAGAUAUGUUUUUAGCAUUUUCAAUAUUAUUAUGUAUGAUGUAUGUUAUAUUUAUAAAUUGGUAUUUUGUAUUUCAUGAGGA
801 AUCUUGAGUUAUAGUAGAUACACUAAAAACAUCAGAUAAAAUAUUACCAGAAUGAUUUUUUUUGUAUUUAUUCGGUUUUUUAAAGGCAAUCCCAGAUAAG
901 UUUAUGGGUUUGUUUUUAAUGGUUAUUUUAUUAUUCUCAUUAUUUUUAUUUAUAUUGAAUUGUAUAUUAUGAUUUGUGUAUUGUAGAAGUUCAUUAUUAU
001 GAUUAACAUAUUCGUUAAUAUUAUUUUAUAGUAUAUGAAUGAGUGGUUUUUUAGCAUUAUAUGUAGUAUUAGCAUAUCCAAUAUGAAUGGAAUUACAAUA
1011 CUGAGUAUUAUUAUUAUUUUUGUUGAUAGUGUGUAGGUUAGAUUAGUUUAGA


TbRPS12
1 CUAAUACACUUUUGAUAACAAACUAAAGUAAAuAuAuuuuGuuuuuuuuuGCGuAuGuGAUUUUUGuAuGGuuGuuGuuuACGuuuuGuuuuAuuuGuuuu
uauaauguagcagag
101 AuGuuAuuAuAuGAGuCCGCGAuuGCCCAGuuCCGGuAACCGACGuGuAuuGuAuGCCGuAuuuuAuuUAuAuAAuuuuGuuuGGAuGuuGCGuuGuuuu
aauauaguaaugagugauaucuuagacc
KTMH6 (1)
aagaugacgaaauagucaaaucaau
KT03908 (3)
aauuaauauuauauaguagugagugauaucuuuuaccgucaaaacaacua
KT03908 (2)
201 uuuuGuuGuuuuAuuGGuuuAGuuAuGUCAuuAuuuAuuAuAGAGGGUGGuGGuuuuGuuGAuuuACCCGGuGUAAAGuAuuAuACACGUAuuGuAAGuu
301 AGAUUUAGAuAUAAGAUAUGUUUUU