Opinion paper

# Should the enzyme name ‘rhodesain’ be discontinued?

Dietmar Steverding[a,*], Conor R. Caffrey[b]

[a]*Bob Champion Research and Education Centre, Norwich Medical School, University of East Anglia, Norwich, United Kingdom*

[b]*Center for Discovery and Innovation in Parasitic Diseases, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California, USA*

*Corresponding author at: Bob Champion Research and Education Centre, Norwich Medical School, University of East Anglia, Norwich Research Park, James Watson Road, Norwich NR4 7UQ, United Kingdom.

*E-mail address:* d.steverding@uea.ac.uk (D. Steverding)

A B S T R A C T

Rhodesain is the generic name for the cathepsin L-like peptidase of *Trypanosoma brucei rhodesiense*. The term rhodesain was derived from the subspecies epithet *rhodesiense* which itself originated form Rhodesia, a historical region in southern Africa named after the 19th century British imperialist and white supremacist Cecil Rhodes. This tainting could be grounds for discontinuing the name, however, there are also scientific grounds. Specifically, protein sequence comparisons and frequency-based difference profiling reveal that rhodesain is essentially identical (99.87-98.44%) to the cathepsin L-like peptidases of both *T. b. brucei* and *T. b. gambiense*. Accordingly, and based on a previously proposed terminology for kinetoplastid C1 peptidases (Caffrey and Steverding, 2009), we suggest the use of the formal term, *Tbr*CATL, to denote the cathepsin L-like peptidases of the *T. brucei* subspecies. The earlier and informal term, 'brucipain', could also be used.

*Keywords:*

Rhodesain

Brucipain

*Trypanosoma brucei rhodesiense*

*Trypanosoma brucei brucei*

*Trypanozoon*

Trypanosomes express two Clan CA C1 cysteine peptidases that are related to mammalian cathepsins B and L [1]. By tradition, the names of cathepsin L-like peptidases from various parasites, including the trypanosome orthologue, have incorporated the suffix "-pain" or "-ain" to the species or subspecies name based on the archetypal C1 cysteine peptidase, papain, from papaya fruit. For example, brucipain, rhodesain, evansain, congopain and cruzipain/cruzain (EC 3.4.22.51) are the generic names for the cathepsin L-like cysteine peptidases of *Trypanosoma (brucei) brucei*, *T. (brucei) rhodesiense*, *T. evansi*, *T. congolense* and *T. cruzi*, respectively [2-7]. This terminology, however, is sometimes confusing and inconsistent. For instance, no such generic name exists for the cathepsin L-like peptidase of *T. equiperdum* and *T. (brucei) gambiense*, and the term vivapain has been used for the respective enzymes of *T. vivax* as well as of *Plasmodium vivax* [8,9]. Although a unifying nomenclature for kinetoplastid C1 peptidases has been proposed [10], based on a system originally suggested by Clayton et al. [11], informal names for the trypanosome cathepsin L-like peptidase still appear in the academic literature.

The term rhodesain was coined to differentiate the cathepsin L-like peptidase of the *T. b. rhodesiense* subspecies from its orthologue, brucipain, found in the subspecies *T. b. brucei* [3,10]. Rhodesain is a derivation of "*rhodesiense*" itself derived from Rhodesia[1], a former British protectorate in southern Africa comprising the present-day countries of Zambia and Zimbabwe, and where the first case of a human infection with *T. b. rhodesiense* occurred in 1909 [12]. The historical region of Rhodesia[2] was named after the British mining magnate and politician, Cecil John Rhodes (1853-1902). However, Rhodes was a passionate colonialist and a believer in the supremacy of white people, and anything connected to him is contentious. Since the 1950s, protesters have questioned whether it is appropriate to commemorate Cecil Rhodes, and have demanded the removal of his statues and the renaming of buildings bearing his name. In this context, it was recently proposed to either rename the *T. b. rhodesiense* subspecies or to revoke its subspecies status, although this suggestion was also based on

[1] The Latin suffix "-ense" (singular nominative, accusative and vocative form of "-ensis") literally means "of or from [a place]".
[2] Not to be confused with the unrecognised state of Rhodesia/Republic of Rhodesia from 1965-1979 which is equivalent to the territory of present-day Zimbabwe.

scientific reasons [13]. By extension, the link between the name rhodesain and Cecil Rhodes could be used to argue for the discontinuation of the name; again, however, there are also scientific reasons.

The species/subspecies status of *T. b. rhodesiense* has been a constant matter of debate. Molecular genetic studies (reviewed [14,15]) and phylogenetic relationship analyses [16] have indicated that *T. b. rhodesiense* is only a phenotypic variant of *T. b. brucei*. In fact, the human infectivity of *T. b. rhodesiense*, which distinguishes it from *T. b. brucei*, is based on just one gene, the serum resistance associated (*SRA*) gene [17-19]. In addition, experimental crossing and population genetics studies have provided evidence for gene flow between *T. b. brucei* and *T. b. rhodesiense* [20]. Therefore, should *T. b. rhodesiense* be no longer recognised as a subspecies but rather a human-infectious form of *T. b. brucei*, then there is no longer a reason to use the term rhodesain. Instead, a better option would be to employ the term *Tbr*CATL (for *T. brucei* cathepsin L) according to the previously proposed nomenclature for kinetoplastid C1 peptidases, which states that the first letter of the genus should be followed by the first two letters of the species (all in italics) to indicate the source organism of the peptidase [10].

Further support for dropping the name rhodesain comes from sequence analysis studies. Comparison of the cathepsin L-like peptidase sequences of *T. b. rhodesiense* with those of *T. b. brucei* and *T. b. gambiense* indicates that they are all essentially identical. (Fig. 1). To date, two, six and two different cathepsin L-like peptidase sequences have been published for *T. b. rhodesiense*, *T. b. brucei* and *T. b. gambiense*. The two *T. b. rhodesiense* sequences are 99.33% identical while their identity with the six *T. b. brucei* and the two *T. b. gambiense* sequences ranges between 99.78-98.44% and 99.78-99.33%, respectively. Thus, some *T. b. brucei* and *T. b. gambiense* sequences show higher identity with one of the *T. b. rhodesiense* sequences than the two rhodesain sequences between each other. For example, *T. b. brucei* sequence E and *T. b. gambiense* sequence B are practically identical to *T. b. rhodesiense* sequence B apart of one amino acid replacement at position 115, which is, in any case, a conservative substitution (L for V; Fig. 1). Interestingly, compared to all other *T. brucei* subspecies sequences, the *T. b. rhodesiense* sequence A is the only one that has a third N-glycosylation site motif (Fig. 1, T.b.r./Seq.A: $^{295}$NDS). In addition, the cathepsin L-like peptidase sequences of *T. evansi* and

*T. equiperdum* show high identity with those of the *T. brucei* group members, with percentage values ranging between 99.11-97.56% and 99.78-98.44%, respectively.

The close relationship between the cathepsin L-like peptidase sequences of *T. b. rhodesiense*, *T. b. brucei*, *T. b. gambiense*, *T. evansi* and *T. equiperdum* is also evident from their frequency-based difference profiles which are all very similar (Fig. 2). It should be noted that *T. evansi* and *T. equiperdum* together with the three *T. brucei* subspecies belong to the sub-genus *Trypanozoon*. In contrast, the cathepsin L-like peptidase sequences of *T. (Nannomona) congolense* and *T. (Duttonella) vivax* exhibit less identity with those of the *Trypanozoon* species (70.82-65.92% for *T. congolense* and 60.69-58.71% for *T. vivax*). In addition, their frequency-based difference profiles are quite distinct from those of the *Trypanozoon* species (Fig. 2). Based on the frequency profiles, it can be concluded that the cathepsin L-like peptidase sequences of the *Trypanozoon* species represent the same protein but are clearly different from the cathepsin L-like peptidase sequences of *T. congolense* and *T. vivax*. Thus, these and the protein sequence alignment data, indicate that the cathepsin L-like peptidase sequences of the *T. brucei* subspecies encode an identical protein with just a few, mostly irrelevant, amino acid substitutions.

In conclusion, and apart from an association with a contentious past, we propose that the name rhodesain be no longer used as the enzyme is not different from brucipain. Should the subspecies status of *T. b. rhodesiense* be revoked (discussed in [13]), then the term rhodesain would be automatically discontinued. In the meantime, any cathepsin L-like peptidases of the *T. brucei* group should be denoted formally as *Tbr*CATL according to the previously proposed unifying nomenclature system for kinetoplastid C1 peptidases, or informally as brucipain [10].

## CRediT authorship contribution statement

**Dietmar Steverding:** Conceptualization, Formal Analysis, Writing – Original Draft, Visualisation. **Conor R. Caffrey:** Writing – Review & Editing.

## Declaration of Competing Interest

The authors declare that they have no competing interests.

## References


[1] MEROPS, the Peptidase Database. https://www.ebi.ac.uk/merops/, 2020 (accessed 26 March 2021).

[2] C.R. Caffrey, S. Scory, D. Steverding, Cysteine proteinases of trypanosome parasites: novel targets for chemotherapy, Curr Drug Targets 1 (2000) 155-162.

[3] C.R. Caffrey, E. Hansell, K.D. Lucas, L.S. Brinen, A.A. Hernandez, J. Cheng, S.L. Gwaltney 2nd, W.R. Roush, Y.-D. Stierhof, M. Bogyo, D. Steverding, J.H. McKerrow, Active site mapping, biochemical properties and subcellular localization of rhodesain, the major cysteine protease of *Trypanosoma brucei rhodesiense*, Mol. Biochem. Parasitol. 118 (2001) 61-73.

[4] S.C. Yadav, R. Kumar, S. Kumar, U. Tatu, R.K. Singh, A.K. Gupta, Identification and characterization of cysteine proteinases of *Trypanosoma evansi*, Parasitol. Res. 109 (2011) 559-565.

[5] E. Authié, G. Duvallet, C. Robertson, D.J. Williams, Antibody responses to a 33 kDa cysteine protease of *Trypanosoma congolense*: relationship to 'trypanotolerance' in cattle, Parasite Immunol. 15 (1993) 465-474.

[6] J.J. Cazzulo, M.C. Cazzulo Franke, J. Martínez, B.M. Franke de Cazzulo, Some kinetic properties of a cysteine proteinase (cruzipain) from *Trypanosoma cruzi*, Biochem. Biophys. Acta 1037 (1990) 186-191.

[7] A.E. Eakin, A.A. Mills, G. Harth, J.H. McKerrow, C.S. Craik, The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*, J. Biol. Chem. 267 (1992) 7411-7420.

[8] P. Vather, Vivapain: a cysteine peptidase from *Trypanosoma vivax*. https://researchspace.ukzn.ac.za/handle/10413/4775, 2010 (accessed 26 March 2021).

[9] B.-K. Na, B.R. Shenai, P.S. Sijwali, Y. Cloe, K.C. Pandey, A. Singh, C.S. Craik, P.J. Rosenthal, Identification and biochemical characterization of vivapains, cysteine proteases of the malaria parasite *Plasmodium vivax*, Biochem. J. 378 (2004) 529-538.

[10] C.R. Caffrey, D. Steverding, Kinetoplastid papain-like cysteine peptidases, Mol. Biochem. Parasitol. 167 (2009) 12-19.

[11] C. Clayton, M. Adams, R. Almeida, T. Baltz, M. Barrett, P. Bastien, S. Belli, S. Beverley, N. Biteau, J. Blackwell, C. Blaineau, M. Boshart, F. Bringaud, G. Cross, A. Cruz, W. Degrave, J. Donelson, N. El-Sayed, G. Fu, K. Ersfeld, W. Gibson, K. Gull, A. Ivens, J. Kelly, D. Lawson, J. Lebowitz, P. Majiwa, K. Matthews, S. Melville, G. Merlin, P. Michels, P. Myler, A. Norrish, F. Opperdoes, B. Papadopoulou, M. Parsons, T. Seebeck, D. Smith, K. Stuart, M. Turner, E. Ullu, L. Vanhamme, Genetic nomenclature for *Trypanosoma* and *Leishmania*, Mol. Biochem. Parasitol. 97 (1998) 221-224.

[12] J.W.W. Stephens, H.B. Fantham, On the peculiar morphology of a trypanosome from a case of sleeping sickness and the possibility of its being a new species (*T. rhodesiense*), Proc. Roy. Soc. Lond. B 83 (1910) 28-33.

[13] D. Steverding, K.M. Tyler, Decolonising Parasitology: the case of *Trypanosoma brucei rhodesiense*, Trends Parasitol. 37 (2021) 461-464.

[14] J.R. Baker, The subspecies taxonomy of *Trypanosoma brucei*, Parasite 2 (1995) 3-12.

[15] W. Gibson, Will the real *Trypanosoma brucei rhodesiense* please step forward? Trends Parasitol. 18 (2002) 486-490.

[16] O. Balmer, J.S. Beadell, W. Gibson, A. Caccone, Phylogeography and taxonomy of *Trypanosoma brucei*, PLoS Negl. Trop. Dis. 5 (2011) e961.

[17] C. De Greef, H. Imberechts, G. Matthyssens, N. Van Meirvenne, R. Hamers, A gene expressed only in serum-resistant variants of *Trypanosoma brucei rhodesiense*, Mol. Biochem. Parasitol. 36 (1989) 169-176.

[18] H.V. Xong, L. Vanhamme, M. Chamekh, C.E. Chimfwembe, J. Van Den Abbeele, A. Pays, N. Van Meirvenne, R. Hamers, P. De Baetselier, E. Pays, A VSG expression site-

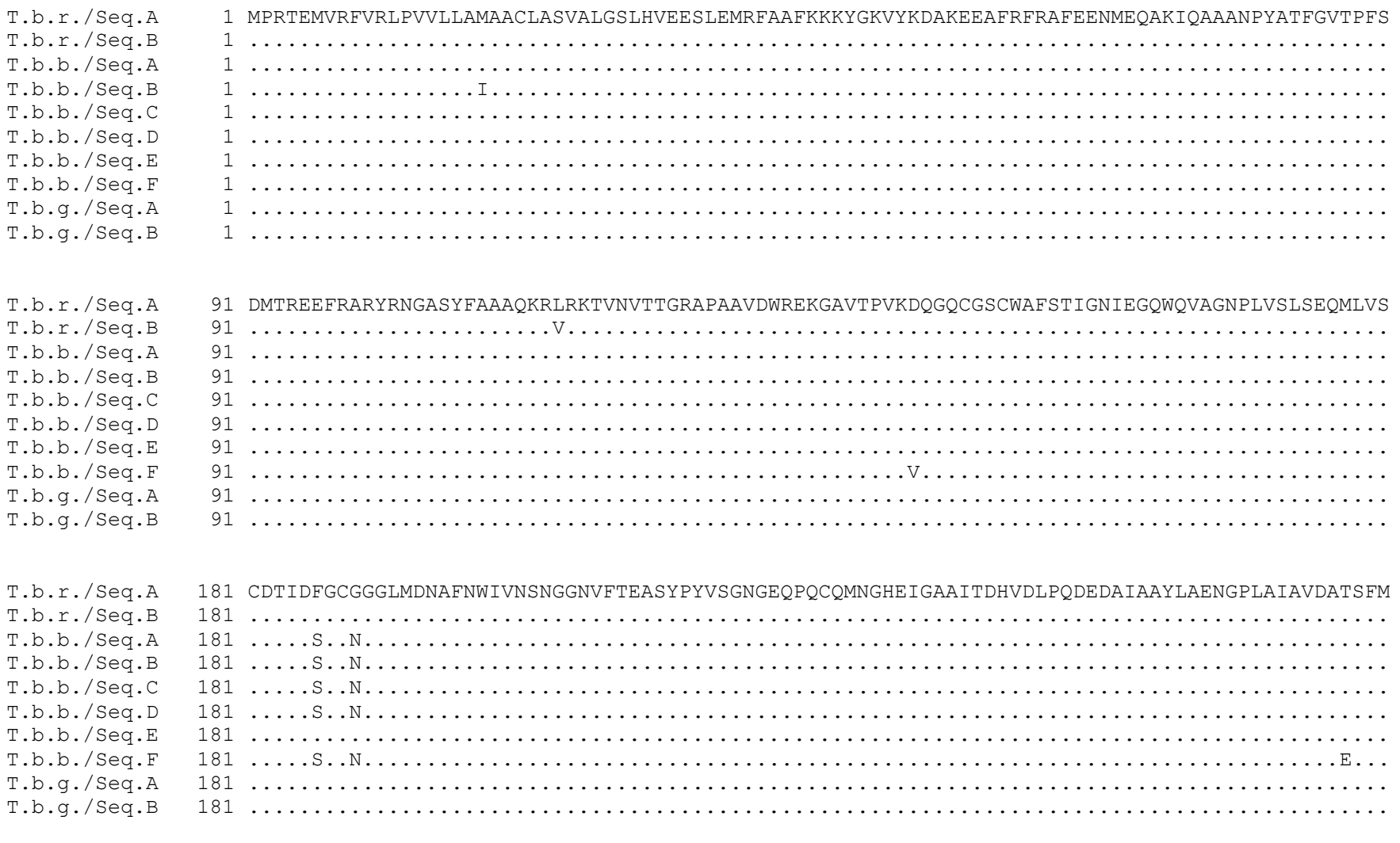
176 associated gene confers resistance to human serum in *Trypanosoma rhodesiense*, Cell
177 95 (1998) 839-846.
178 [19] W.C. Gibson, The SRA gene: the key to understanding the nature of *Trypanosoma*
179 *brucei rhodesiense*, Parasitology 131 (2005) 143-150.
180 [20] W. Gibson, Liaisons dangereuses: sexual recombination among pathogenic
181 trypanosomes, Res. Microbiol. 166 (2015) 459-466.
182

## Figure legends

**Fig. 1.** Sequence alignment of cathepsin L-like peptidases from *T. b. rhodesiense* (T.b.r.), *T. b. brucei* (T.b.b.) and *T. b. gambiense* (T.b.g.). A NCBI Protein database (https://www.ncbi.nlm.nih.gov/protein/) search using the search terms "brucei" and "cathepsin L" gave a total of 2, 35 and 6 entries for cathepsin L-like peptidases sequences for *T. b. rhodesiense*, *T. b. brucei* and *T. b. gambiense*, respectively. However, several entries contained the same sequence so that, in the end, only two, six and two specific cathepsin L-like peptidases sequences for *T. b. rhodesiense*, *T. b. brucei* and *T. b. gambiense*, respectively, were obtained. Species-specific sequences and their corresponding accession numbers are as follows: T.b.r./Seq.A: CAC67416.1; T.b.r./Seq.B: CAA38238.1; T.b.b./Seq.A: XP_845218.1, XP_845219.1, XP_845220.1, XP_845223.1, AAX80351.1, AAX80352.1, AAX80353.1, AAX80356.1, AAZ11659.1, AAZ11660.1, AAZ11661.1 AAZ11664.1; T.b.b./Seq.B: XP_845221.1, AAX80354.1, AAZ11662.1; T.b.b./Seq.C: XP_845222.1, XP_845227.1; XP_845228.1, AAX80355.1, AAX80360.1, AAX80361.1, AAZ11663.1, AAZ11668.1, AAZ11669.1; T.b.b./Seq.D: XP_845225.1, XP_845.226.1, AAX80358.1, AAX80359.1, AAZ11666.1, AAZ11667.1; T.b.b./Seq.E: XP_845224.1, AAX80357.1, AAZ11665.1; T.b.b./Seq.F: CAA34485.1, P14658.1; T.b.g./Seq.A: XP_011773878.1, CBH11593.1; T.b.g./Seq.B: XP_011773880.1, XP_011773883.1, CBH11595.1, CBH11598.1.

**Fig. 2.** Comparison of African trypanosome cathepsin L-like peptidase sequences using the NCBI COBALT Multiple Alignment Tool (https://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi?CMD=Web). The display setting shows frequency-based differences which compares the residue at a position to the position consensus. Darker shades of red indicate a further difference from residues in other sequences in the alignment at that position. Grey indicates identical residues at that position for all sequences while white indicates gaps. The following number of unique sequences were compared: *T. b. rhodesiense*: 2; *T. b. brucei*: 6; T. *b. gambiense*: 2; *T. evansi*: 1; *T. equiperdum*: 2, *T. congolense*: 12; *T. vivax*: 4.

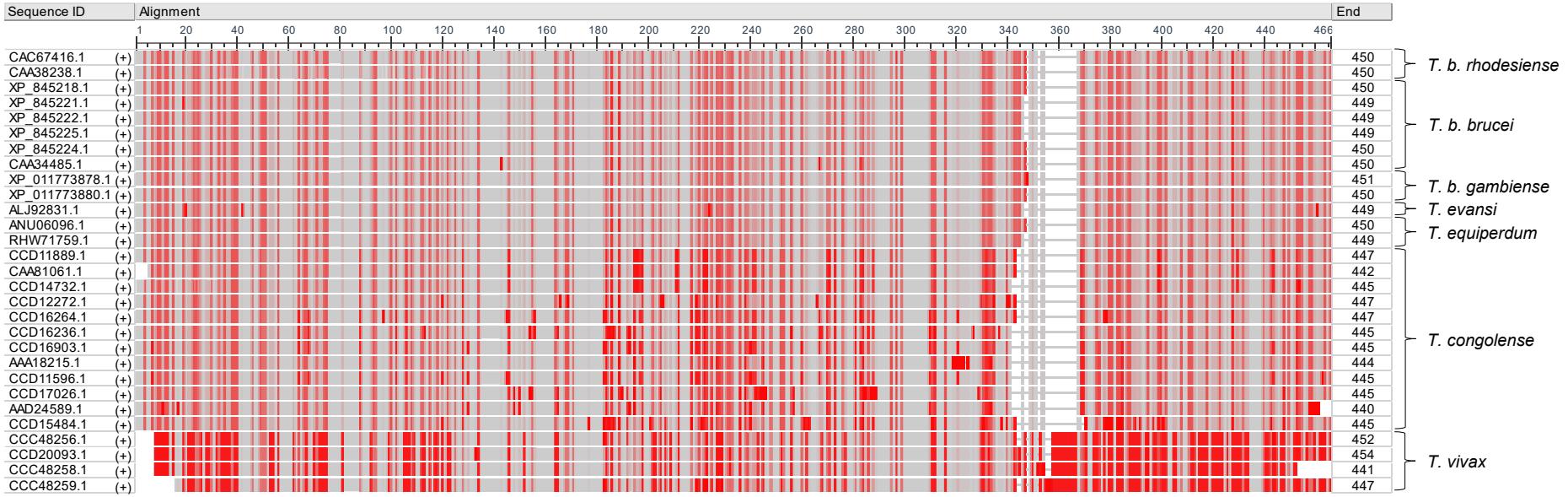
Figure 1


T.b.r./Seq.A 1 MPRTEMVRFVRLPVVLLAMAACLASVALGSLHVEESLEMRFAAFKKKYGKVYKDAKEEAFRFRAFEENMEQAKIQAAANPYATFGVTPFS
T.b.r./Seq.B 1 .....
T.b.b./Seq.A 1 .....
T.b.b./Seq.B 1 .....I.....
T.b.b./Seq.C 1 .....
T.b.b./Seq.D 1 .....
T.b.b./Seq.E 1 .....
T.b.b./Seq.F 1 .....
T.b.g./Seq.A 1 .....
T.b.g./Seq.B 1 .....
T.b.r./Seq.A 91 DMTREEFRARYRNGASYFAAAQKRLRKTVNVTTGRAPAAVDWREKGAVTPVKDQGQCGSCWAFSTIGNIEGQWQVAGNPLVSLSEQMLVS
T.b.r./Seq.B 91 .....V.....
T.b.b./Seq.A 91 .....
T.b.b./Seq.B 91 .....
T.b.b./Seq.C 91 .....
T.b.b./Seq.D 91 .....
T.b.b./Seq.E 91 .....
T.b.b./Seq.F 91 .....V.....
T.b.g./Seq.A 91 .....
T.b.g./Seq.B 91 .....
T.b.r./Seq.A 181 CDTIDFGCGGGLMDNAFNWIVNSNGGNVFTEASYPYVSGNGEQPQCQMNGHEIGAAITDHVDLPQDEDAIAAYLAENGPLAIAVDATSFM
T.b.r./Seq.B 181 .....
T.b.b./Seq.A 181 .....S..N.....
T.b.b./Seq.B 181 .....S..N.....
T.b.b./Seq.C 181 .....S..N.....
T.b.b./Seq.D 181 .....S..N.....
T.b.b./Seq.E 181 .....
T.b.b./Seq.F 181 .....S..N.....E...
T.b.g./Seq.A 181 .....
T.b.g./Seq.B 181 .....

```
T.b.r./Seq.A   271 DYNGGILTSCTSEQLDHGVLLVGYNDSSNPPYWIIKNSWSNMWGEDGYIRIEKGTNQCLMNQAVSSAVVGGPTPPPP-PPPPPSATFTQD
T.b.r./Seq.B   271 ........................N.....................................................-.....
T.b.b./Seq.A   271 ........................N.....................................................-.....
T.b.b./Seq.D   271 ........................N.....................................................--....
T.b.b./Seq.C   271 ........................N.....................................................--....
T.b.b./Seq.D   271 ........................N.....................................................--....
T.b.b./Seq.E   271 ........................N.....................................................-.....
T.b.b./Seq.F   271 ...........K............N.....................................................-.....
T.b.g./Seq.A   271 ........................N.....................................................P.....
T.b.g./Seq.B   271 ........................N.....................................................-.....

T.b.r./Seq.A   360 FCEGKGCTKGCSHATFPTGECVQTTGVGSVIATCGASNLTQIIYPLSRSCSGLSVPITVPLDKCIPILIGSVEYHCSTNPPTKAARLVPHQ 450
T.b.r./Seq.B   360 .....................................................P..................................... 450
T.b.b./Seq.A   360 .....................................................P..................................... 450
T.b.b./Seq.B   359 .....................................................P..................................... 449
T.b.b./Seq.C   359 .....................................................P..................................... 449
T.b.b./Seq.D   359 ..................................................................................... 449
T.b.b./Seq.E   360 .....................................................P..................................... 450
T.b.b./Seq.F   360 .....................................................P..................................... 450
T.b.g./Seq.A   361 ..................................................................................... 451
T.b.g./Seq.B   360 ..................................................................................... 450
```

Figure 2


Sequence ID
Alignment
End
1 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360 380 400 420 440 466
CAC67416.1 (+) 450
CAA38238.1 (+) 450
XP_845218.1 (+) 450
XP_845221.1 (+) 449
XP_845222.1 (+) 449
XP_845225.1 (+) 449
XP_845224.1 (+) 450
CAA34485.1 (+) 450
XP_011773878.1 (+) 451
XP_011773880.1 (+) 450
ALJ92831.1 (+) 449
ANU06096.1 (+) 450
RHW71759.1 (+) 449
CCD11889.1 (+) 447
CAA81061.1 (+) 442
CCD14732.1 (+) 445
CCD12272.1 (+) 447
CCD16264.1 (+) 447
CCD16236.1 (+) 445
CCD16903.1 (+) 445
AAA18215.1 (+) 444
CCD11596.1 (+) 445
CCD17026.1 (+) 445
AAD24589.1 (+) 440
CCD15484.1 (+) 445
CCC48256.1 (+) 452
CCD20093.1 (+) 454
CCC48258.1 (+) 441
CCC48259.1 (+) 447
T. b. rhodesiense
T. b. brucei
T. b. gambiense
T. evansi
T. equiperdum
T. congolense
T. vivax