Appendices Surve et al. *T. brucei* Complex I

Supplementary Methods

Glycerol gradients. Strain 427 BF (2×10^8) parasites isolated from rats were resuspended in 0.5 ml lysis buffer (Tris HCl pH 7.2, 10 mM, $MgCl_2$, 10 mM, KCl, 200 mM) with 1% Triton X-100 and with complete, EDTA-free protease inhibitor cocktail (Roche). The lysate was centrifuged at 16,000 × g at 4°C for 15 min; the supernatant was then centrifuged under the same conditions for an additional 30 min. The resulting clarified supernatant (500 ml) was loaded on a 10-30% linear glycerol gradient in buffer (Tris HCl pH 7.2, 10 mM, $MgCl_2$, 10 mM, KCl, 100 mM) as described (Panigrahi et al., 2001), and centrifuged at 182,400 × g in a Beckman SW40 rotor for 10 hr. Subsequently, 500 ml fractions were collected from the top of the gradients, flash frozen in liquid nitrogen, and stored at -70°C until further use. Gradient fractions (15 μl) were spotted on a nitrocellulose membrane using a dot blot module (BioRad). Membranes were blocked in Odyssey blocking buffer (Li-Cor Biosciences) and incubated with hybridoma supernatants of monoclonal antibodies (mAbs) 52 and 63 (1,3) (1:10 dilution). The primary antibodies were detected with goat anti-mouse IgG IRDye 800CW (Li-Cor Biosciences) upon scanning using a Li-Cor Odyssey imaging system.

Southern analysis. For Southern analysis, genomic DNA cleaved with appropriate enzymes was resolved on a 0.9% agarose gel and transferred to a Nytran N membrane (Schleicher & Schuell). The membrane was then separately probed with randomly primed Klenow radiolabeled probes that were specific for the CDSs and 3'UTRs of the replaced genes, or with a 656 bp probe corresponding to β-tubulin.

Molecular evolutionary analysis. We compared the DNA sequences for all *bona fide* (*i.e.* core or conserved accessory) *T. b. brucei* 927 cI subunits identified in this study to their trypanosomatid orthologues in the TriTrypDB.org database. Codon-based evolutionary distances were calculated using the Nei-Gojobori model (2) in the MEGA5 software (4) and identified the *T. congolense* IL3000 sequences as the ones best suited for a pairwise dN/dS analysis. The Z-test of selection in MEGA5 (again using the Nei-Gojobori model) was used to determine the probability of rejecting the alternative hypothesis ($dN < dS$) in favor of the null hypothesis of strict-neutrality ($dN = dS$). The variance of the difference was computed using the bootstrap method (500 replicates).

Reference List

1. **Allen, T. E., S. Heidmann, R. Reed, P. J. Myler, H. U. Goringer, and K. D. Stuart**. 1998. Association of guide RNA binding protein gBP21 with active RNA editing complexes in *Trypanosoma bruce*i. Mol.Cell.Biol. **18**:6014-6022.

2. **Nei, M. and T. Gojobori**. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol.Biol.Evol. **3**:418-426.

3. **Panigrahi, A. K., A. Zikova, R. A. Dalley, N. Acestor, Y. Ogata, A. Anupama, P. J. Myler, and K. D. Stuart**. 2008. Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex. Mol.Cell.Proteomics **7**:534-545. .

4. **Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar**. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol.Biol.Evol. **28**:2731-2739.

Supplementary Table 1. Oligonucleotides

Target	5' primer	3' primer
Coding region amplification		
LIR	5'-ACAAAGCTTATGTCAGCCCTTC TTC-3'	5'-ACAGGATCCACCCTCATTAGAA ACTATGC-3'
NUBM	5'-ATGCAAGCTTATGCTCCGGCGT GTGGGTT-3'	5'-GCATGGATCCGTTCCAATTCCG CACGATC-3'
NUKM	5'-ATGCAAGCTTATGCTTCGTCGC ACGTCG-3'	5'-GCATGGATCCATCTCGAACAGA ATACTTTTG -3'
NB4M	5'-ATGCCCTAGGATGCTTCGTCGA GCATGC-3'	5'-GCATCTCGAGAATCTGCGAGAC AACCAC-3'
NI2M	5'-ATGCCCTAGGATGATGCGCCGT TGCTTT-3'	5'-GCATCTCGAGTCTCCTGGCAAT ATTACT- 3'
NI8M	5'-ATGCCCTAGGATGTCATGGCGA GCGCGT-3'	5'-GCATCTCGAGCGGCATTATTTT AAAACC-3'
NUHM	5'-ATGCAAGCTTATGTGGCGTCGA GCTGGG-3'	5'-GCATGGATCCCTTCTTTGCTGCG CCGCC-3'
Gene disruption		
NUBM 5'	5'-ATAGCGGCCGCGTTTTGTTGTG CGGGACGAA-3'	5'-ATAACGCGTCTCGAGCTACACA CAAGGTTCAATTTGA-3'
NUBM 3'	5'-ATATCTAGAATTTAAATGGTGG TGGTGTAGTATTGAA-3'	5'-ATAAGGCCTGCGGCCGCCGGAA GTCAGCGTTAAGTTA-3'
NUKM 3'	5'-ATATCTAGAATTTAAATGAATC ATGGTACAATAGGAGTTTCT-3'	5'-ATAAGGCCTGCGGCCGCGTTGC GACCGTTACGTTTGA-3'
NUKM 5'	5'-ATAGCGGCCGCCGGCATTTCTT CGAGGCAGT-3'	5'-ATAACGCGTCTCGAGTGCAAAC TATTTATCTATCCTCTTC-3'
Plasmid construction		
3HA	5'-CATGGAAGCTTCTCGAGGGATC CTACCCCTACGACGTGCCCGACTA CGCCTACCCCTACGACGTGCCCGA CTACGCCTACCCCTACGACGTGCC CGACTACGCCTGATAACTGCA-3'	5'-GTTATCAGGCGTAGTCGGGCAC GTCGTAGGGGTAGGCGTAGTCGG GCACGTCGTAGGGGTAGGCGTAGT CGGGCACGTCGTAGGGGTAGGAT CCCTCGAGAAGCTTC-3'

Supplementary Table 2.

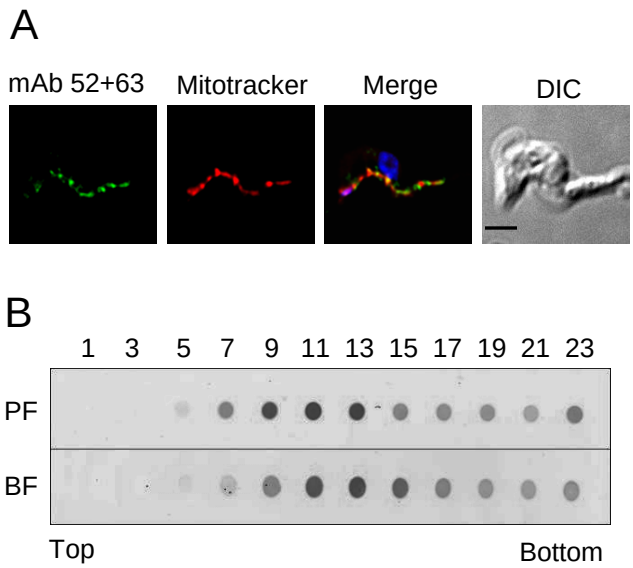
Molecular evolutionary analysis of core and conserved accessory cI subunits in *T. b. brucei* and *T. congolense*

Gene	Systematic ID	Nucleotide Identity[a] identical / total / %	Probability[b]	dN/dS[c]
Complex I subunits				
ACP	Tb927.3.860	337 / 447 / 75.4	1.94E-14	0.158
NB4M	Tb927.10.14860	1209 / 1500 / 80.6	8.63E-44	0.045
NI2M	Tb11.01.7460	689 / 882 / 78.1	2.83E-33	0.073
NI8M	Tb11.01.8630	409 / 498 / 82.1	2.32E-25	0.014
NIDM	Tb11.01.1690	600 / 771 / 77.8	7.52E-27	0.119
NUAM	Tb927.10.12540	617 / 795 / 77.6	5.91E-40	0.036
NUBM	Tb927.5.450	1164 / 1491 / 78.1	1.12E-50	0.053
NUEM1	Tb09.244.2620	1110 / 1380 / 80.4	4.96E-40	0.040
NUEM2	Tb927.10.13620	896 / 1122 / 79.9	3.87E-40	0.036
NUFM	Tb927.10.4130	640 / 798 / 80.2	4.45E-32	0.040
NUHM	Tb927.7.6350	523 / 627 / 83.4	1.47E-24	0.026
NUKM	Tb11.47.0017	489 / 609 / 80.3	7.11E-27	0.035
Controls (essential genes)				
REL1	Tb09.160.2970	1052 / 1410 / 74.6	8.74E-40	0.106
ATPase α	Tb927.7.7420	1360 / 1692 / 80.4	6.17E-47	0.040
Pyruvate Kinase 1	Tb927.10.14140	1190 / 1500 / 79.3	4.93E-47	0.112

[a]The entire *T. b. brucei* (strain 927) and *T. congolense* (strain IL3000) genes were used for analysis except for the following (numbers indicate the first and last nt of the CDS region used for analysis, excluding 5' and 3' regions too divergent for reliable analysis): NB4M 301-1800, NI2M 19-903, NUAM 64-858, NUEM1 79-end, NUHM 94-720, ATPase α 64-end.

[b]Codon-based test of purifying selection. The probability P of rejecting the alternative hypothesis (dN < dS, indicating purifying selection) in favor of the null hypothesis of strict neutrality (dN = dS) is shown. Valuces of P less than 0.05 are considered significant. Analyses were conducted using the Nei-Gojobori model in MEGA5 (see Supplementary Methods).

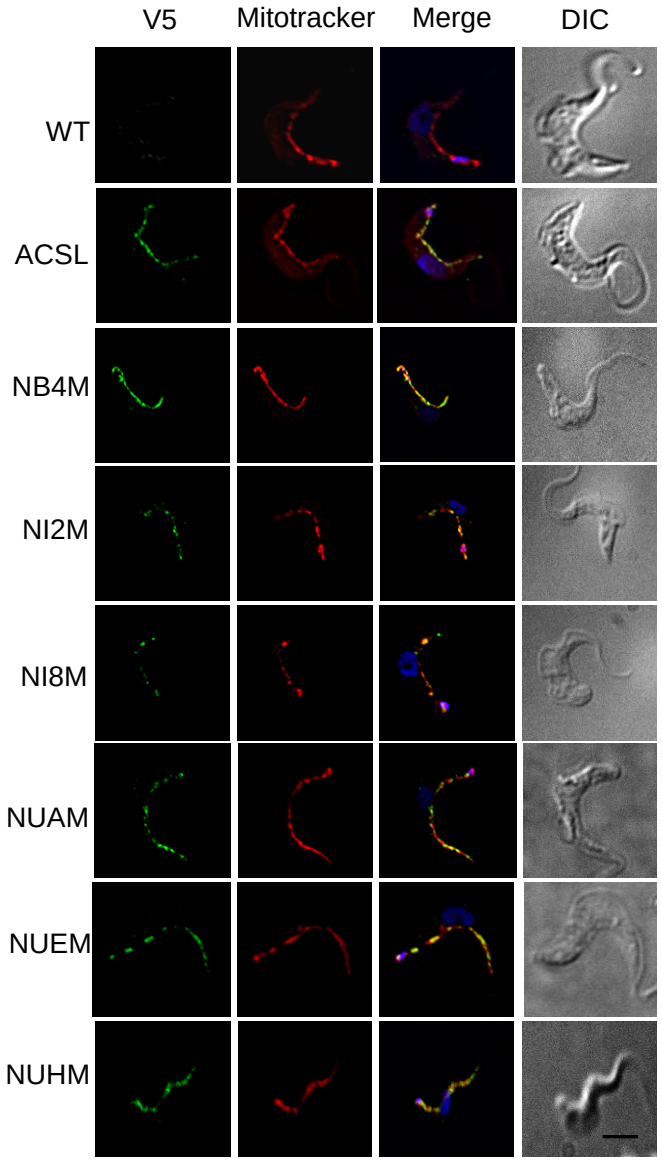
[c]The ratio of non-synonymous differences per non-synonymous site and synonymous differences per synonymous site between the orthologues pairs of *T. b. brucei* and *T. congolense* is shown.



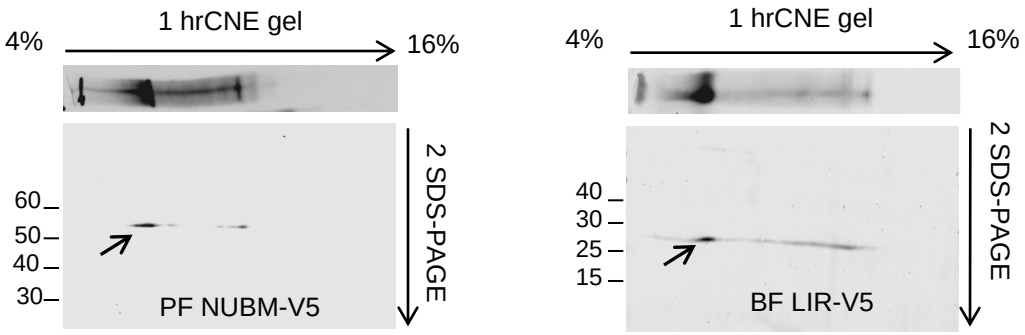
<u>Supplementary Figure 1.</u> Complex I proteins are present and localize to the mitochondrion in *T. brucei* BF

A) Immunofluorescence staining. BF were fixed, permeabilized and stained with mAb52+63. mAbs were visualized by FITC-labeled secondary antibody (green) and mitochondria were identified using Mitotracker (red). The nucleus and kinetoplast were stained with DAPI. Scale bar = 2 µm.

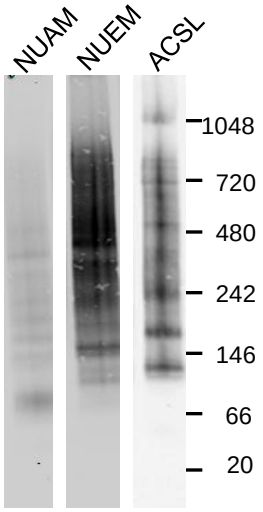
B) Dot blot analysis of glycerol gradient fractions. BF and PF lysates were centrifuged on 10-30% glycerol gradients. Fractions were spotted onto a nitrocellulose membrane. The dot blot was probed with mAbs 52+63. Both stages show a peak between fractions 9-13. Similar results were obtained using the individual antibodies (data not shown).



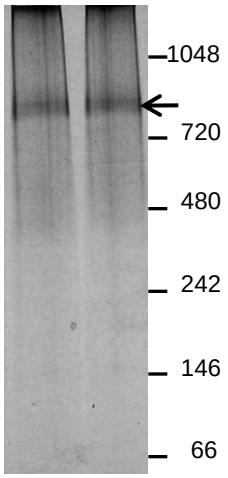
<u>Supplementary Fig. 2.</u> Tagged putative cI subunits ACSL, NB4M, NI2M, NI8M, NUAM, NUEM, and NUHM localize to the mitochondrion in BF T. *brucei.*
Immunofluorescence analysis. The BF cell transfectants above were induced with Tet, permeabilized, and stained with mouse anti-V5 mAb (green), Mitotracker (red) and DAPI (blue). In this and other experiments strain 427 or its "single marker" derivative were used as wild type (WT) BF parasites. Scale bar = 2 μm.



Supplementary Fig. 3. Verification of tagged proteins in complexes by two-dimensional gel analysis.
Lanes from native gels run (hrCNE, first dimension) were cut out and loaded onto 10% SDS-PAGE gels (second dimension). Blots were probed with mouse anti-V5. Left panel, PF NUBM-V5, right panel, BF LIR-V5. Arrows indicate spot derived from large complex of interest.



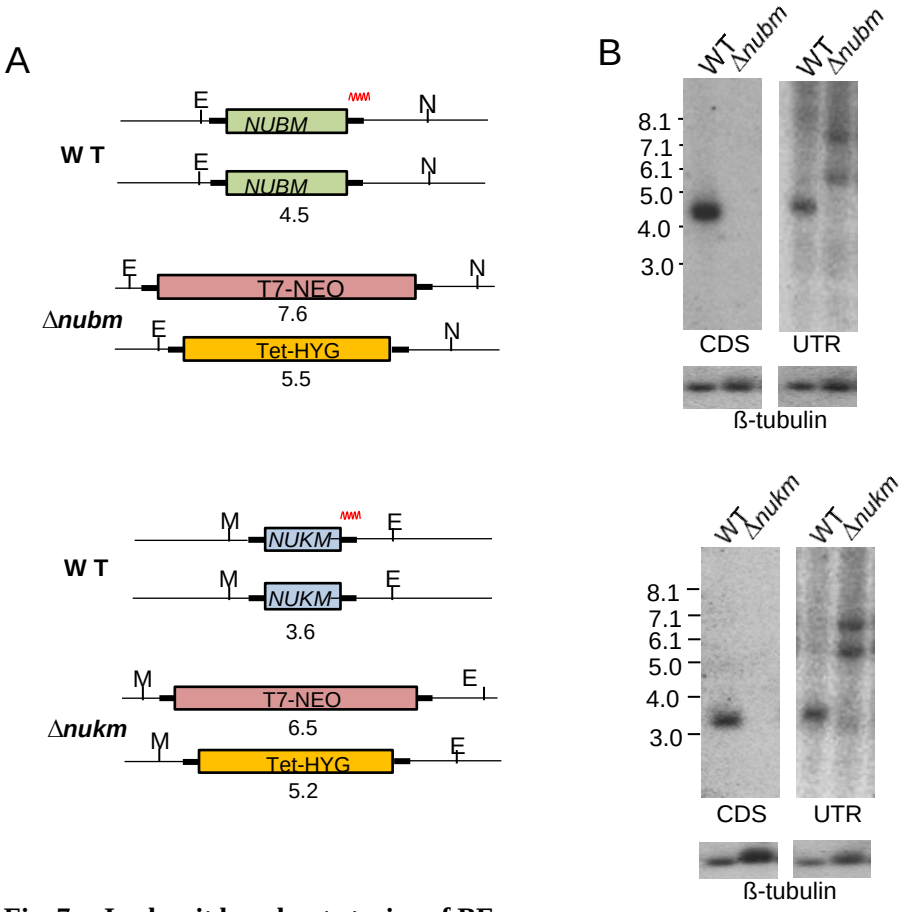
Supplementary Fig. 4. Tagged putative cI subunits NUAM, NUEM and ACSL do not form stable complexes in BF *T. brucei.*
Crude organellar preparations of induced (+ Tet) BF lines expressing NUAM-V5, NUEM-V5 and ACSL-V5 were solublized and separated by hrCNE. Blot were incubated with anti-V5 mAb. At higher expression levels, a small amount of the ACSL-V5 was observed in a complex which is larger than the NUBM or LIR complexes (~1100 kDa).



Supplementary Fig. 5. Silver-stained native gel of immunopurified LIR-V5 complex from BF. A crude organellar preparation of induced BF expressing LIR-V5 was immunoprecipitated with anti-V5 antibody coupled to beads, and the bound complex was eluted with V5 peptide. The eluate was loaded into two wells of a 4-16% gel and subjected to hrCNE . After silver staining, the 855 kDa complex (arrow) was excised, digested with trypsin and subjected to mass spectrometry.

```
N.crassa_Nuo_12.3/NIDM  VPPTFDGV-DYNDTKRLKQAQDAIIREQWVRVMMGRLVREELSKCYYREG-VNHLEKCGHLRER  77
Tb11.01.1690            VPPPVDG--KYNDYYHYLAYKNWMERERDVYVAHANLVNEMAMRCLVKEGQYNAAKNCRHLYHK 198
Ricinus_12kDa_subunit   GPDDFDPSDPYKDPVAMLEMREHIVREKWIQIEKSKILREKLRWCYRIEG-VNHLQKCRHLVQQ  76
                         *  .*    *:*       :: : **: : :  ..::.*    *   **  *  ::* ** .:
```

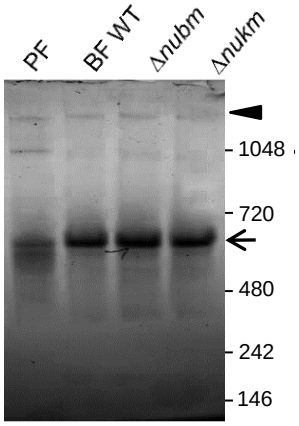
Supplementary Fig. 6. Identification of the putative NIDM subunit in *T. brucei*. Partial ClustalW alignment of the NIDM subunits of *Neurospora crassa* (GenBank CAA48768.1) and *Ricinus communis* (GenBank EEF28972.1) with Tb11.01.1690. Note that the *T. brucei* sequence has an N-terminal extension compared to the *Neurospora* and *Ricinus* sequences. Tb11.01.1690 is the highest scoring protein in a BLAST analysis of *Neurospora* NIDM against the *T. brucei* database (Expect value 0.00048) and, likewise, *Neurospora* NIDM is the highest scoring match for the C-terminal part of Tb11.01.1690 using NCBI BLAST (Expect value 6e-04). Asterisks, colons and periods indicate full conservation, conservation between groups of strongly similar properties, and conservation between groups of weakly similar properties, respectively.



Supplementary Fig. 7. cI subunit knockout strains of BF.

A. Cartoon depicting the *NUBM* and *NUKM* loci in BF 427 (WT) and double knockouts. The boxes indicate the CDS or the inserted cassettes, while the thicker lines correspond to the flanking regions used for homologous recombination. Restriction enzyme sites relevant to the Southern analyses are marked: E, EcoRV, M, MluI, and N, NotI. The size of the predicted fragments is indicated in kb. The red wavy line shows the location of the 3' UTR probe for each gene.

B. Southern blots of *NUBM* and *NUKM* loci in BF 427 (WT), double allele knockout lines. The blots were probed with radiolabeled probes specific for the *NUBM* and *NUKM* CDSs, 3' untranslated regions corresponding to those genes, and β-tubulin as a control. For analysis of Δ*nukm*, genomic DNA was digested with EcoRV and MluI, while for Δ*nubm*, EcoRV and NotI were used. The size (kb) and migration of marker DNAs are indicated.



<u>Supplementary Fig. 8.</u> In-gel NADH dehydrogenase activity is unaltered in NUBM or NUKM deletion mutants and dyskinetoplastic *T. brucei*

In-gel activity assay following hrCNE on 3-12% gels. An arrow marks the major band of activity, and a triangle marks the fainter high molecular weight bands. Half as much PF sample ($2.5x10^6$ cell equivalents) was loaded on the gel on the right as compared to the left. PF samples showed a weak band of activity migrating close to the 1048 kDa marker; a faint, similarly sized band was occasionally seen in BF. BF samples had four-fold more parasites per lane.