S1A
Topology of phylogenetic relationships within kinesin-2 family (incl. 2A -2D subgroups) from kinesin superfamily motor domain phylogenetic inference analysis by Wickstead, *et al.*, **2010.**

Kinesin-2 Family

Kinesin-2 subgroups*

2C *Homodimeric metazoa‡*

Protist group 4

Protist group 3

2B *KRP95 metazoa*

2A *KRP85 metazoa‡*

Monbr_17662
Protist group 2

2D *Protist group 1*

MEME negative controls†

Seqs excluded MEME analysis†

Sequences included in MEME NST analyses†

S1B

Significant* NST motifs identified for Meme 83

Mtf em#	Wid (aa)	# Site	E-val**	Motif sequence LOGO
1	31	67	-1185	
2	22	56	-320	
3	31	46	-398	
4	12	54	-151	
5	50	15	-192	
6	16	42	-156	
7	22	40	-98	
8	31	39	-89	
9	16	29	-70	
10	12	29	-63	

S1B *(continued)*

Mtf em#	Wid (aa)	# Site	E-val**	Motif sequence LOGO
11	16	13	-61	
12	22	42	-37	
13	14	10	-22	
14	43	3	-21	*(Motif not evaluated)*
15	50	3	-18	
16	35	5	-12	
17	15	8	-12	
18	49	3	-11	
19	35	3	-08	
20	50	3	-12	
21	08	10	-07	
22	31	6	-05	
23	12	21	-05	
24	09	13	-04	
25	42	3	-03	
29	44	5	+01	

S1C Schematic of M83 motifs with motif widths to scale + color -coded motif groups (per S2A), ranked by increasing E-value

Motif number

m1 m3 m2 m5 m6 m4 m7 m8 m9 m10 m11 m12 m13 m15 m16 m20 m18 m19 m21 m22 m23 m24 m25 m29

E-value cutoff=
Scale=50aa

higher ← statistical significance → lower

E-value rank # 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

Fig. S1. Kinesin-2 family phylogeny and NST motifs. **(A)** Kinesin-2 family phylogentic tree topology adapted from the previous analysis by using motor domain sequences; branch lengths shown do not indicate evolutionary distances. Symbols: †Our early MEME analyses showed these NST sequences to be distinct from remaining kinesin-2 sequences; for additional information see Table S3A. ‡Monbr_16629 and Monbr_23354 of the holozoan *M. brevicolis* were previously assigned to subgroups 2A and 2C, respectively (Wickstead et al., 2010). Monbr_23354 had a putative NST domain too short for MEME analysis. **(B)** Kinesin-2 motor domain (MD) sequences are more highly conserved than kinesin-2 NST sequences (see Table S1 for MD sequence comparisons, and Table S2 for NST sequence comparisons). Our NST analyses relied primarily upon the MEME tool suite (see main text), which can identify conserved motifs independently of sequence alignment or motif order. We carried out over 80 MEME runs because of variation between runs due to differences in run parameters and sequences used as described in Materials and Methods. In the figure are the 25 statistically significant NST motifs (based on MEME-determined individual motif overall E-values), plus one additional motif identified, from a single representative run (MEME 83). Shown are: motif number (Mft m#); optimal width (amino acids or aa); the total number of significant, non-overlapping, independent motifs identified in the data set (# sites); the E-value overall statistical significance measure for each motif as defined in (expressed as integer (e.g. -10), equivalent to 1 x 10E-value with lower E-values being more statistically significant), with a motif overall E-value significance cutoff of -02 as recommended in MEME, with the exception of motif 29 which had an E-value = 1; and each motif's MEME-identified amino acid consensus sequence (LOGO) as determined (within MEME) using the sequence display program LOGO (Crooks et al., 2004; Schneider and Stephens, 1990). Motif E-values are calculated in MEME based on combined motif p-values on individual sequences, with only significant motif p-values ($\leq 1 \times 10^{-10}$) on individual sequences used to calculate E-values. **(C)** Schematic of individual motifs from S1B arranged from left to right in order of increasing E-values (from most to least significant). Widths of individual motifs are displayed to scale for motifs having 16 aa or greater; motif widths of 8 - 15 aa are displayed as equal to 16 aa for visibility. Each motif is color-coded according to the motif group to which it is assigned. Orange bar indicates E-value = -02 statistical significance cutoff recommended by MEME authors. Scale bar = 50 aa (with the exceptions stated above).
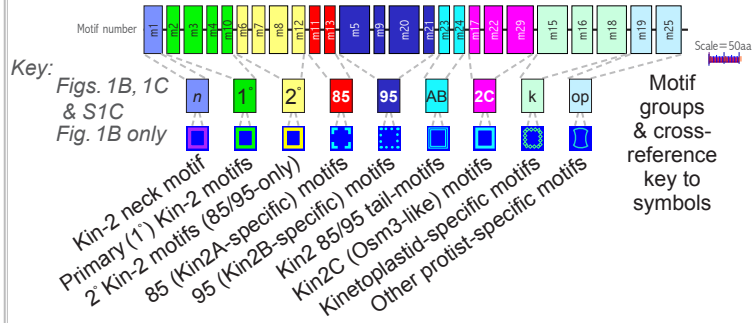
References:

Crooks, G. E., Hon, G., Chandonia, J.-M. and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188–1190.

Schneider, T. D. and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18,** 6097–6100.

Wickstead, B., Gull, K. and Richards, T. A. (2010). Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton. *BMC Evol. Biol.* **10,** 110.

S2A

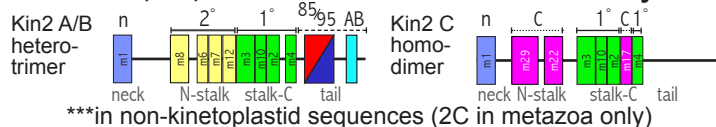**Schematic with cross-reference key: M83 motifs and motif groups, with symbols**

Individual M83 motifs arranged into color-coded motif groups, as used on individual taxa in Figs. 1B, 1C, S1C & S2 (B-F) w/ cross-reference key
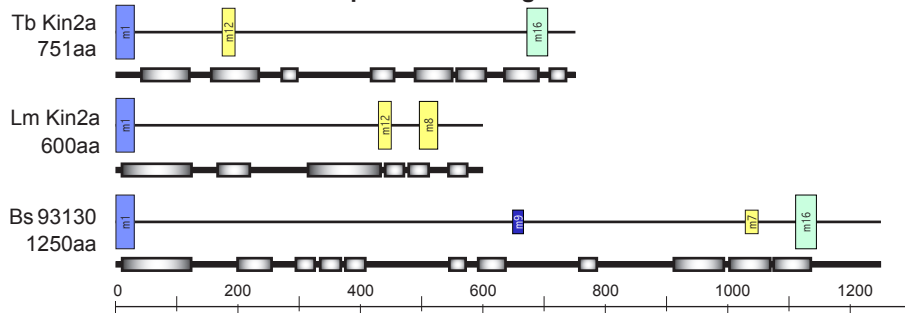
Motif number

Key:
Figs. 1B, 1C & S1C: n 1° 2° 85 95 AB 2C k op
Fig. 1B only

Kin-2 neck motif
Primary (1°) Kin-2 motifs
2° Kin-2 motifs (85/95-only)
85 (Kin2A-specific) motifs
95 (Kin2B-specific) motifs
Kin2 85/95 tail-motifs
Kin2C (Osm3-like) motifs
Kinetoplastid-specific motifs
Other protist-specific motifs

Motif groups & cross-reference key to symbols

Scale = 50aa

S2B Sequence motifs key: motif heights indicate *p*-value ranges

Seq motifs key   Motif height (m_height)   Individual motif *p*-value range:

max ≤ 1 x10 -20.0 -10.0 -8.0 -7.0 -5.0 1
& min > 1 x10 -∞ -20.0 -10.0 -8.0 -7.0 1
Motif height = V | IV | III | II | I

S2C Kin2A, 2B, 2C motifs & motif order are broadly conserved*

Kin2 A/B heterotrimer   n 2° 1° 85 95 AB
neck N-stalk stalk-C tail

Kin2 C homodimer   n C 1° C1°
neck N-stalk stalk-C tail

***in non-kinetoplastid sequences (2C in metazoa only)

S2E TbKin2a and select kinetoplastid homologs

Tb Kin2a 751aa
Lm Kin2a 600aa
Bs 93130 1250aa

0 200 400 600 800 1000 1200

S2F 1 Shared 2C motif on TbKin2b, kinetoplastid homologs & metazoan 2C taxa

Tb Kin2b 754aa
Lm Kin2b 782aa
Bs05805 760aa

0 200 400 600 800

Kinetoplastid taxa with a single 2C motif

Ce Osm3 375aa
Hs Kif17 697aa

0 200 400 600 800

Metazoan taxa classified as 2C

†††*Location of 2C (m17) + 1°Kin2 (m4) motifs (metazoa) or 2C (m17) + low p-value/ partial 1°Kin2 (m4) motif (kinetoplastids)*

S2D Selected sequences with conserved 1° Kin2 & 2° Kin2 + (85 / 95 motifs)†

0 200 400

NgKin2‡ 403aa
Tv 3438‡ 345aa
Cr Fla10 436aa — KRP 85
Cr Fla8 435aa — KRP 95
Tt Kin1 397aa — KRP 85
Tt_731 410aa — KRP 95
Mb16629 374aa — KRP 85
Mb17662 438aa — KRP 95
Hs Kif3a 357aa — KRP 85
Hs Kif3b 410aa — KRP 95

0 200 400

Fig. S2. Kinesin-2 NST motifs and motif groups in individual sequences. **(A)** Schematic of individual motifs arranged into motif groups. Motif groups are cross-referenced to motif group symbols that are used in Fig. 1 (B-D). In Fig. 1 and Fig. S2, references to motif group 2A/ KRP85/ 85 are equivalent, as are references to motif group 2B/ KRP95/ 95, and motif group 2C/ OSM3. See Fig. S1C legend for motif width scaling information. **(B)** Key showing how the height of motif boxes on individual sequence scales with motif p-value ranges. Roman numerals V to I show sequence-specific individual motif p-value ranges as follows: V = p-value $\leq 1 \times 10^{-20}$, IV = p-value $>1 \times 10^{-20}$ and $\leq 1 \times 10^{-10}$, III = p-value $>1 \times 10^{-10}$ and $\leq 1 \times 10^{-08}$, II = p-value $>1 \times 10^{-08}$ and $\leq 1 \times 10^{-07}$, I = p-value $>1 \times 10^{-07}$ and $\leq 1 \times 10^{-05.}$ **(C)** Schematic showing that motif order is generally conserved among kinesin-2 proteins as depicted. **(D)** NST motifs for a selected group of heterotrimeric kinesin-2A and -2B proteins across a broad evolutionary backdrop illustrate sequence motif conservation noted in (C). In particular, we observed a high frequency of 1° motifs 2, 3, 4 on taxa, especially signature residues [F - I - P] that terminate 1° motif 2, and [W/Y - 6X (with 1-3 E/D) - W] in 1° motif 4, which were observed previously in *Xenopus laevis* (De Marco et al., 2001). **(E)** NST motifs for TbKin2a and homologs LmKin2a and Bs93130. One common kinesin-2 motif (2° motif 12) had a significant p-value on 2 of 3 taxa. **(F)** NST motifs for TbKin2b, the related LmKin2b and Bs05805, as well as the kinesin-2C proteins CeOSM3 and HsKif17. Note that 2C motif 17 is followed directly by 1° motif 4 for both metazoan taxa, 1° motif 4 for *T. brucei* and *B. saltans* sequences and a possible (low p-value) amino acid signature for *L. major* (not shown). TbKin2b and kinetoplastid homologs also have a predicted 2° motif 12, here located consistently just after neck domain of the 3 kinetoplastid proteins.

References:

De Marco, V., Burkhard, P., Le Bot, N., Vernos, I. and Hoenger, A. (2001) Analysis of heterodimer formation by Xklp3A/B, a newly cloned kinesin-II from *Xenopus laevis*. *EMBO J* **20**, 3370–3379.
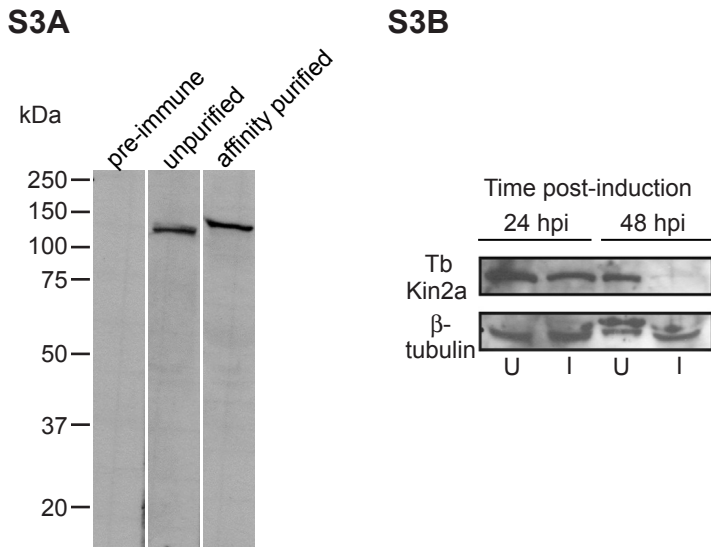
S3A

S3B

Fig. S3. (A) Western blots of whole-cell extracts from T. brucei probed with pre-immune rabbit serum (left), unpurified post-immune rabbit serum (center), and affinity-purified anti-TbKin2a IgG rabbit polyclonal antibody (right). Predicted molecular weight of full-length TbKin2a (1088 aa) protein = 122.72 kDa. **(B)** Western blots of whole-cell extracts probed with rabbit anti-TbKin2a IgG primary antibody (top), and mouse anti-β-tubulin IgG (KMX) primary antibody (bottom), each followed by applicable secondary anti-IgG antibody conjugated to horseradish peroxidase, from TbKin2a RNAi uninduced (U) and induced (I) cells at 24 and 48 hours post induction (hpi). Note TbKin2a protein levels remain significant at 24 hpi, and are reduced but visible at 48 hpi, indicating residual TbKin2a activity may persist until relatively late during RNAi induction period. See Figure 5A for control (wild-type) cell immunoblots.
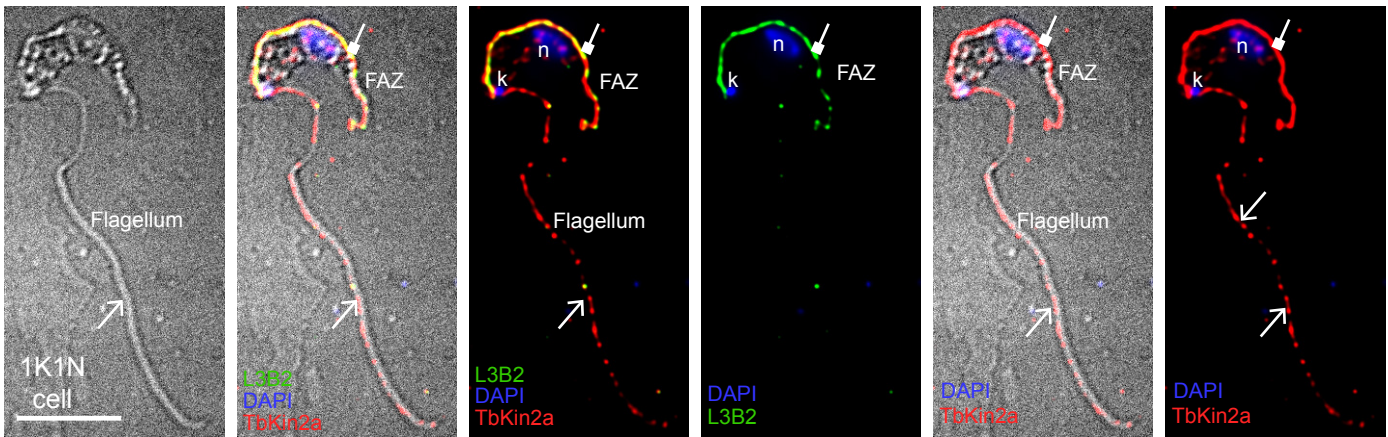
Fig. S4. DIC and immunofluorescence images of cells grown in media with 20 mM citrate and 5 mM MgCl$_2$ for ~30 h to chelate Ca^{2+} and induce flagellar detachment. Cells were stained for TbKin2a (red), FAZ protein TbFAZ1 (green, L3B2) and DNA (blue, DAPI). Immunoflourescence images are 2D maximum intensity projections from 3D deconvolved z-stacks. Symbols are the same as used in Fig. 4. Bar = 7.4 µm
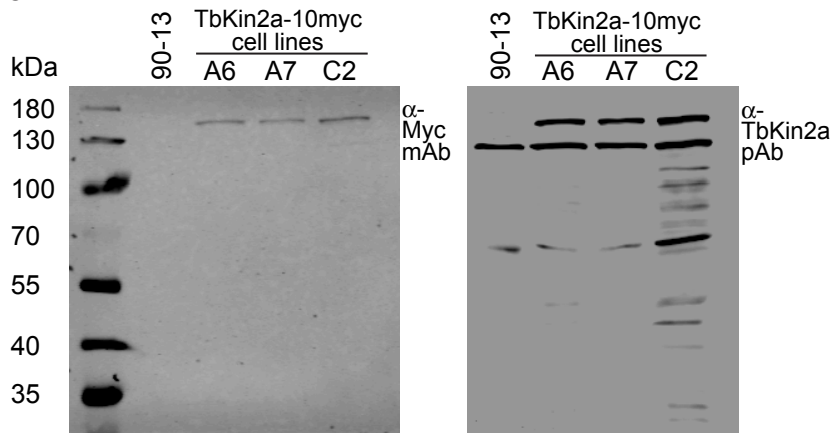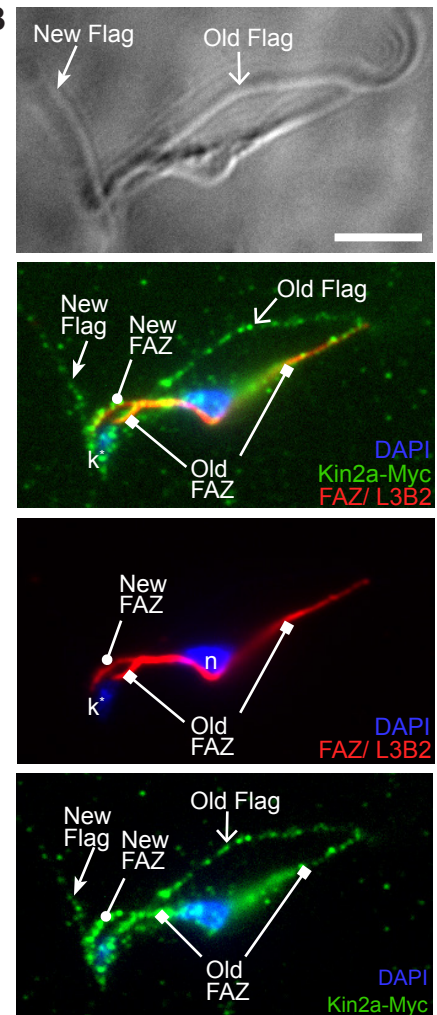
Fig. S5. (A) Western blots showing endogenous expression of full length TbKin2a with a C-terminal Myc fusion tag (TbKin2a-Myc) in three *T. brucei* BSF clonal cell lines (A6, A7, C2) transfected with pPOTv7, and 90-13 negative control, probed with anti-Myc mouse IgG1 monoclonal antibody 9E10 (left) and anti-TbKin2a rabbit polyclonal antibody (right). **(B)** DIC and immunofluorescence images of a detergent extracted *T. brucei* BSF clone-A7 cell, methanol-fixed and stained for TbKin2a-Myc with anti-Myc mouse IgG2a monoclonal antibody 9B11 (catalog #2276, Cell Signaling Technology, Danvers, MA) (red), FAZ protein TbFAZ1 with anti-TbFAZ mouse IgG1 monoclonal antibody L3B2 (green), and DNA with DAPI fluourescent dye (blue). Secondary goat anti-mouse monoclonal antibodies conjugated to AlexaFluor red and green fluorescent dyes were, respectively, mouse IgG2a and IgG1 isotype-specific. Bar = 5 µm.

Table S1. Kinesin-2 Motor Domain Sequences Identity Matrix

Click here to Download Table S1

Table S2. Kinesin-2 NST Domain Sequences Identity Matrix

Click here to Download Table S2

Table S3. Kinesin-2 sequence information and reconciliation, parts 3A and 3B

Click here to Download Table S3A3B

Table S4. Taxa NST FASTA File Names and Sequences

Click here to Download Table S4

Table S5. Primer Sequences Used

Click here to Download Table S5