Supplemental Materials

Molecular Biology of the Cell

Holden et al.

Supplementary material for:

Evolution of the nuclear pore complex: An Mlp analog in trypanosomes functions in chromosomal segregation but lacks transcriptional barrier activity

Jennifer M. Holden*, Ludek Koreny*,†, Samson Obado‡, Alexander V. Ratushny§, Wei-Ming Chen§,||, Steven Kelly¶, Brian T. Chait‡, John D. Aitchison§, Michael P. Rout‡, and Mark C. Field†•

Supplementary figures

Figure S1. Comparison of the coiled-coil patterns among Mlps, Tpr, NUA, TbNup92 and TbNup110 proteins. Coiled-coil-forming probabilities were obtained by Coils/Pcoils in scanning windows of 14, 21 and 28 residues using either a single protein sequence (Mlp1 and Mlp2 of *S. cerevisiae*, human Tpr, NUA of *A. thaliana* and Nup92 and Nup110 of *T. brucei*) or sequence alignments of Nup92 and Nup110 as queries. The major discontinuity in the coiled-coil pattern, which is described for Mlp1 is marked by an arrow.

Figure S2. Evolution and domain structure of nuclear basket components of *T. brucei*. (A) Comparison of the Nup92 protein sequences of *T. brucei* and *Naegleria gruberi*, which is based on the alignment of all TbNup92 proteins identified by homology searches. Regions predicted as coiled-coils are shaded in gray. The circles highlight the first and fourth positions of the predicted heptad pattern. Filled circles stand for uncharged and hydrophobic amino acid residues that are favorable for coiled-coil formation, while open circles indicate less favorable residues. The BRCT domain is highlighted in red and the consensus sequence of the classic monopartite nuclear localization signal is colored in blue. (B) Maximum likelihood phylogenetic tree of TbNup110. Homologs of the *T. brucei* TbNup110 sequence were found in other kinetoplastids but nowhere else. Numbers at nodes represent bootstrap/SH-like aLRT values.

Figure S3. TbNup98 does not localise to the SAS at mitosis. (A) TbNup98::GFP cells were fixed and probed with anti-GFP and imaged by super-resolution microscopy. In contrast to the nuclear basket nucleoporins that localise to the SAS at mitosis (see Figure 2), TbNup98::GFP (green) maintained a peripheral punctate distribution at the NE throughout the cell cycle, corresponding to it's association with the NPC. DAPI was used to visualise DNA (blue).

Figure S4. TbNup92 is not essential for the survival of BSF cells. (A) Quantitation of TbNup92 mRNA in uninduced versus TbNup92 knockdown cells. At 24h induction, ~60% reduction in TbNup92 mRNA is observed. (B) TbNup92 knockdown results in a slight population growth defect over a seven-day period. (C) At 6, 12, 24, 48, and 72 hour time points, cells were fixed and stained with DAPI to visualise DNA. This allowed categorisation of cells into distinct stages of the cell cycle. 100 cells were recorded at each time point and the mean scores of triplicate experiments are shown. TbNup92 knockdown disrupts normal progression through the cell cycle. An initial decrease in 2K1N cells is observed, followed by an increase in unusual cell types.

Figure S5. Single allele deletion and truncation of TbNup92 result in a population growth defect. (A) Population growth curves following the truncation of the C-terminus of TbNup92 (TbNup92tr::GFP), truncation and single allele deletion of TbNup92 (TbNup92tr::KO) and double gene deletion of TbNup92 (TbNup92Δ). A growth defect is observed in TbNup92tr::KO cells, which is even more pronounced in TbNup92Δ cells (data shown in Figure S7A). (B) DAPI counts of parental, TbNup92tr::GFP, TbNup92tr::KO and TbNup92Δ cells. The mean scores of triplicate experiments are shown. An increase in 1K1N (interphase) cells combined with a decrease in 2K1N and 2K2N cells in the TbNup92tr::KO population indicates that the cells may encounter delays in progressing through G1/S phases of the cell cycle. This phenotype is exaggerated in the TbNup92Δ population whereby the delays and errors in progressing through interphase and mitosis result in the accumulation of unusual cell types (data shown in Figure 7A). (C) A single allele of the nuclear basket protein TbNup110 was tagged at the C-terminus with a 3xHA epitope in TbNup92tr::KO cells. Confocal images (shown is the central slice along the z-axis) display the distribution of TbNup110::HA (red) around the nuclear periphery and show that this localisation was not affected in TbNup92tr::KO cells.

Figure S6. Overexpression of TbNup92 lacking the central SMC domain (TbNup92BRCT). The C-terminal BRCT domain of TbNup92 was overexpressed in *T. brucei* under the control of a tetracycline inducible promoter. The construct consisted of the first 90 bases of the TbNup92 ORF (red), immediately followed by the final 579 bases of the ORF (red). A nuclear localisation signal [NLS, yellow, derived from NUP-1 (DuBois *et al.,* 2012)] and 3 x HA epitopes (purple) were added to the C-terminus.

Figure S7. Deletion of both TbNup92 alleles severely impacts on normal cell growth and division. Both alleles of TbNup92 were sequentially replaced with hygromycin and neomycin resistance genes in PCF cells. (A) Whole genome re-sequencing of DNA from TbNup92Δ cells failed to generate any reads that mapped to the TbNup92 (Tb09.160.0340) ORF in the 427 reference genome, thus confirming deletion of TbNup92. (B) Quantitation of TbNup92 mRNA in the parental and TbNup92Δ cells normalised to Rab11 expression values. Expression of TbNup92 is undetectable in TbNup92Δ cells. (C) Population cell counts over a seven-day period show that cell division is severely compromised in TbNup92Δ cells.

Figure S8. TbNup92Δ cells no longer show defects in chromosome segregation after several hundred generations. (A) After ~400 generations in culture, the growth rate of TbNup92Δ cells was recorded. TbNup92Δ cells no longer to appear to display a growth defect when compared to the population growth rate of the parental cell line. (B) DAPI counts were used to monitor cell cycle progression in the TbNup92Δrev population. 100 cells were counted and the mean scores of triplicate experiments are shown. There is no cell cycle arrest observed in the TbNup92Δrev population. C) TbNup92Δrev cells were fixed and stained with propidium iodide to measure the karyotype of the population. The FACS profile of the revertant population is compared to the FACs profiles of wild type and TbNup92Δ cells described in Figure 5. The TbNup92Δrev population has an increased number of 1K0N cells (zoids) though aneuploidy is no longer apparent after ~400

generations in culture medium. (D) The fluorescence intensity of telomeres at the posterior and anterior poles of 20 dividing nuclei from TbNup92Δrev cells were recorded. The fluorescence intensity of the telomeres at the posterior and anterior poles was almost equal, indicating that the TbNup92Δrev cells have undergone a compensatory mutation that prevents aneuploidy. The fluorescence analysis was compared to that of wild type and TbNup92Δ cells described in Figure 7F.

Figure S9. Comparison of relative read depth between TbNup92Δ and TbNup92Δrev lines following DNAseq. Shown is the log2 of the ratio of mode-normalised read depth in TbNup92Δrev divided by mode-normalised read depth in the TbNup92Δ line. A value of 0 denotes no difference in read depth (i.e. copy number) between the two lines for that region of the genome; a value of 1 means that read depth in the recovered line is twice that in the mutant line, and a value of -1 means that read depth in the mutant line is twice that in the recovered line. Each point represents the average read depth ratios +/- 10kb of a 20Kb window.

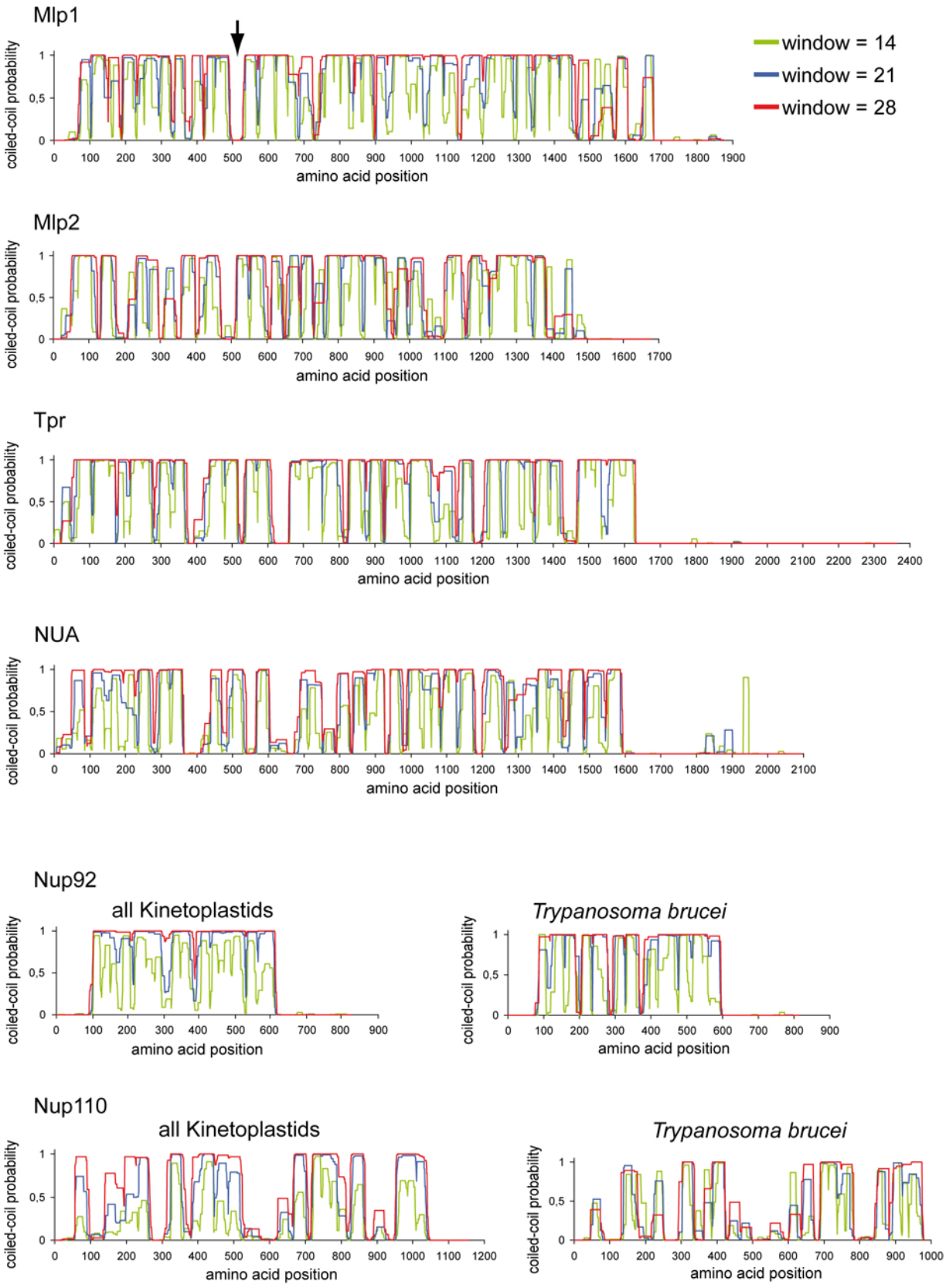
Figure S10. Frequency distributions of the ORF start sites of all genes (blue) and the differentially expressed genes following TbNup92 knockout (red) relative to the transcription start site (TSS) (Siegel *et al.,* 2009). The frequency distribution plots are represented by histograms with 50 equally spaced bins. The differentially expressed genes are not biased in their localisation along individual cistrons (p= 0.31).

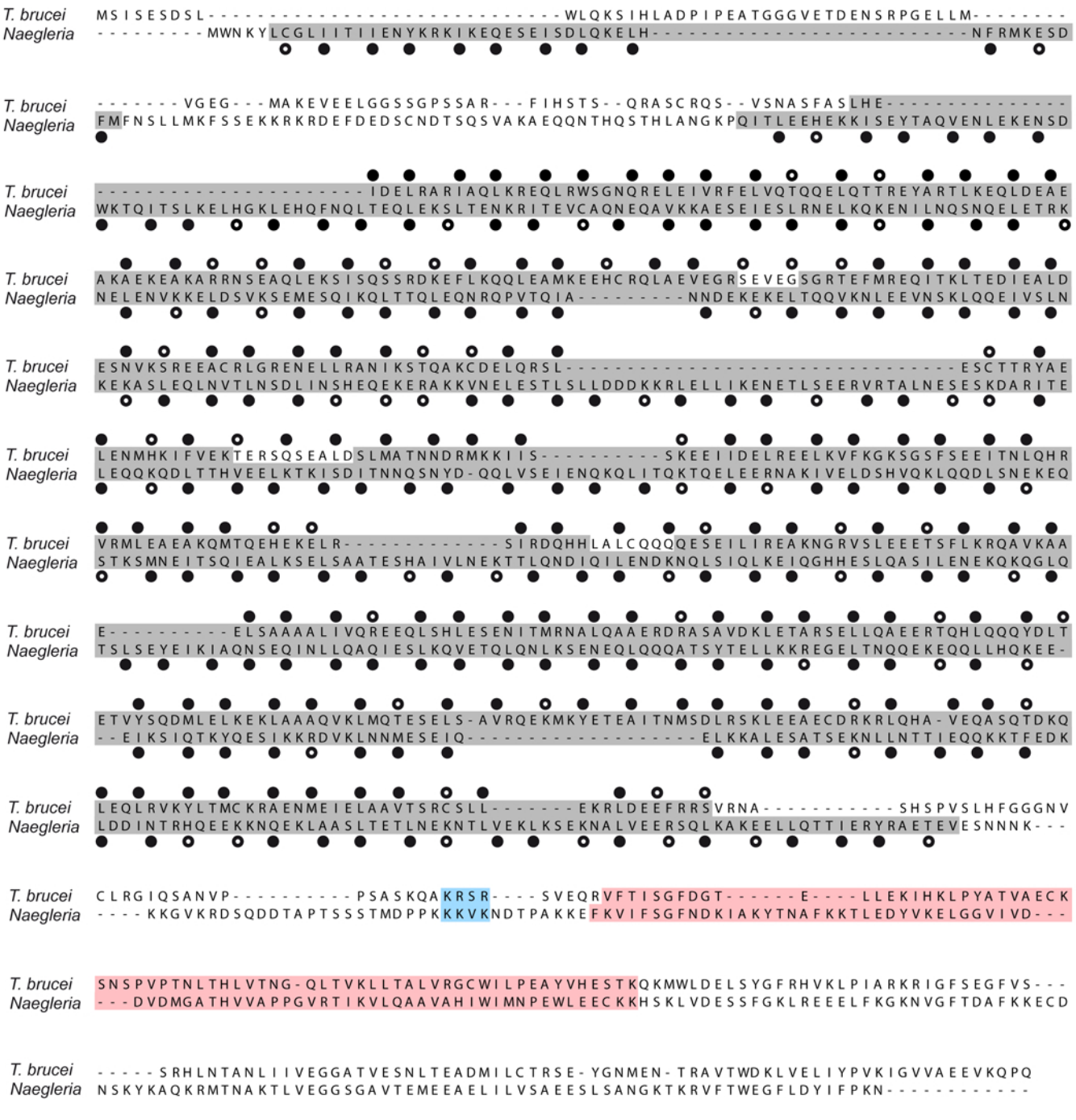
Supplementary table 1. Affinity isolation of TbNup92 complexes. Affinity isolates of TbNup92 were prepared, analysed and identified by mass spectrometry as described in Methods. The above table lists identified proteins including their mass and accession numbers according to the base -10 log of the expectation (Log (e)) that any particular assigned protein was made at random (E-value). The base -10 log of the sum of fragment ion intensities (log(I)) in the tandem mass spectra are provided. The % measured refers to the percentage of amino acid coverage of the protein identified, while % corrected refers to the coverage corrected for peptide sequences that are unlikely to be observed using normal proteomic methods. These include very hydrophobic peptides, very small peptides or very large peptides depending on the enzyme used to digest the protein of interest. In the cases where the % corrected are indicated by 100+, this indicates that all peptide fragments that can be observed by existing techniques are present in the spectra, including other additional peptides likely arising from missed cleavage events. Unique refers to unique/individual peptide sequences associated with each particular protein whilst Total is the total number of peptides observed for each given protein including multiples of each unique peptide.

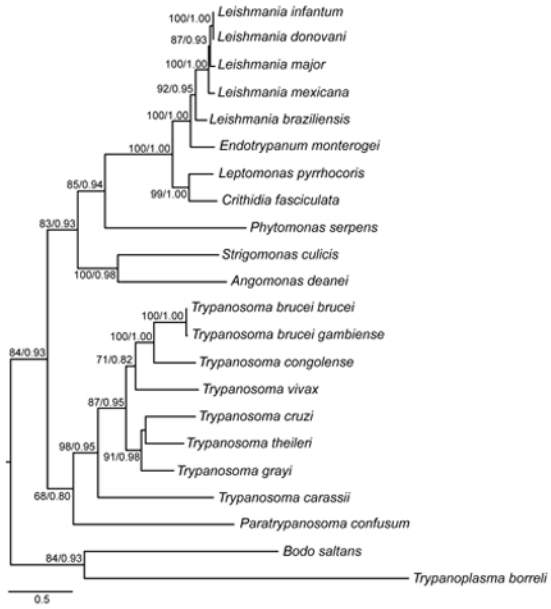
Supplementary table 2. RNAseq raw data. Sheared cDNA synthesised from RNA extracted from TbNup92Δ cells was sequenced by 76bp paired end Illumina sequencing. Reads were mapped to the T. brucei 427 strain reference genome and aligned to annotated transcripts. The FPKM's of annotated transcripts is normalised by quantile normalisation and the statistical significance for transcript level changes is calculated based on (FPKMtreat / FPKMparental) values against null distribution using *t*-test. In the filter column, '1' denotes a significant gene with $p<0.05$ and *vice-versa.* Pie chart displays annotated GO-terms for the significantly up-regulated genes following

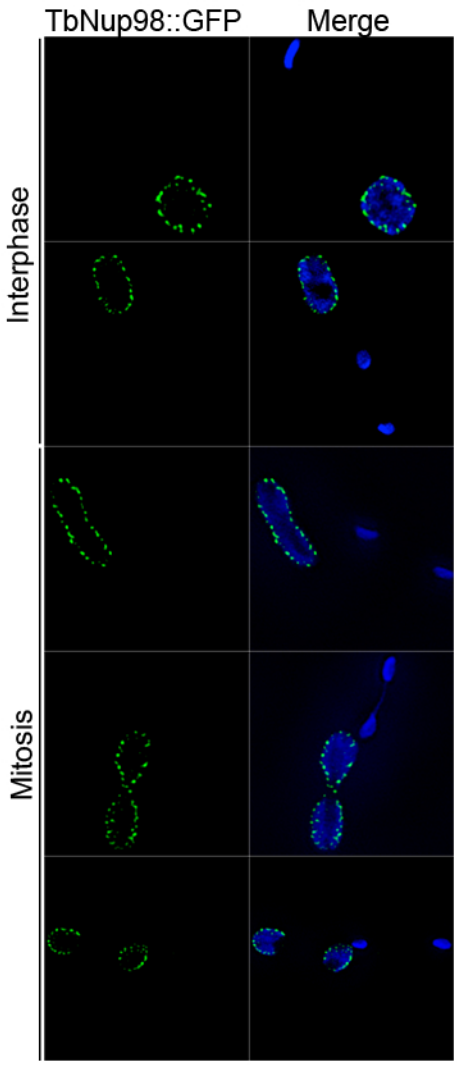
deletion of TbNup92. Significant GO-terms were selected based on the following criteria: 1) #of genes in GO gene pool >=3; 2) #of genes matching to GO gene pool >=2; and 3) p-value <0.05. A number of up-regulated genes were assigned GO-terms that suggested functions in mRNA turnover.

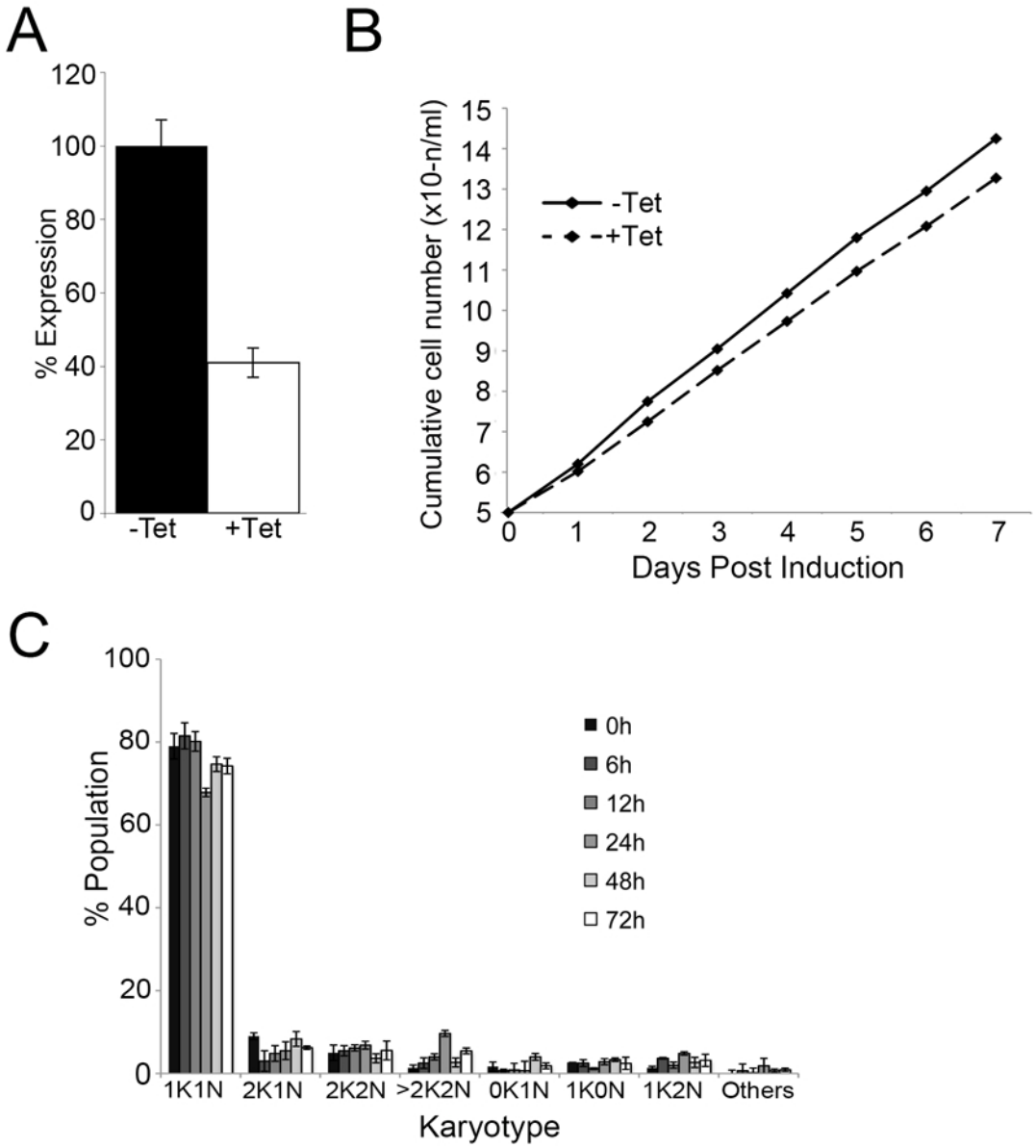
Supplementary movie 1. Cells expressing TbNup92::GFP and TbNup110:cMyc were fixed and probed with anti-GFP and anti-cMyc to visualise TbNup92 (green) and TbNup110 (red) respectively. Cells were imaged by super-resolution microscopy. As well as being localised around the nuclear periphery of mitotic cells, TbNup110 concentrates at the SAS where it co-localises with TbNup92. Shown are serial confocal sections along the z-axis. DAPI was used to visualise DNA.


Mlp1
window = 14
window = 21
window = 28
coiled-coil probability
amino acid position
Mlp2
Tpr
NUA
Nup92
all Kinetoplastids
Trypanosoma brucei
Nup110
all Kinetoplastids
Trypanosoma brucei

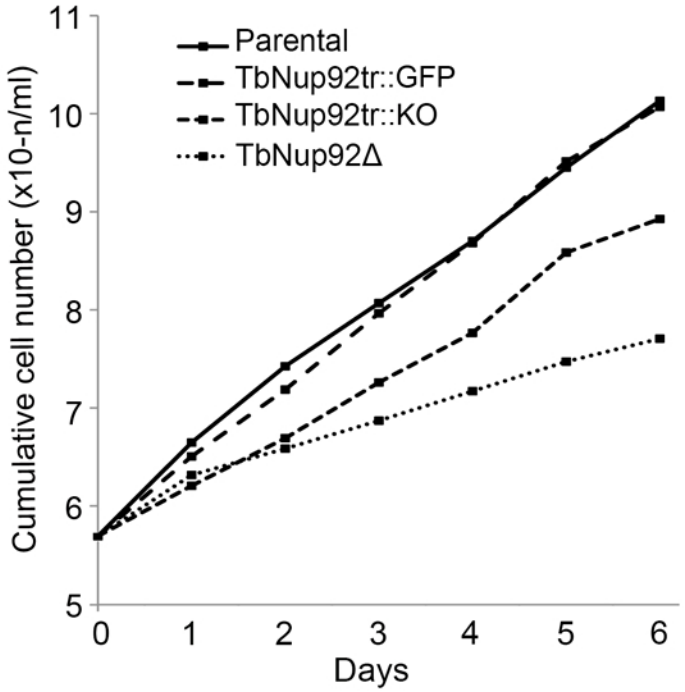

T. brucei
Naegleria
MSISESDSL-----------------------------WLQKSIHLADPIPEATGGGVETDENSRPGELLM-------
---------MWNKYLCGLIITIIENYKRKIKEQESEISDLQKELH-------------------------NFRMKESD
-------VGEG---MAKEVEELGGSSGPSSAR---FIHSTS--QRASCRQS--VSNASFASLHE--------------
FMFNSLLMKFSSEKKRKRDEFDEDSCNDTSQSVAKAEQQNTHQSTHLANGKPQITLEEHEKKISEYTAQVENLEKENSD
----------------------IDELRARIAQLKREQLRWSGNQRELEIVRFELVQTQQELQTTREYARTLKEQLDEAE
WKTQITSLKELHGKLEHQFNQLTEQLEKSLTENKRITEVCAQNEQAVKKAESEIESLRNELKQKENILNQSNQELETRK
AKAEKEAKARRNSEAQLEKSISQSSRDKEFLKQQLEAMKEEHCRQLAEVEGRSEVEGSGRTEFMREQITKLTEDIEALD
NELENVKKELDSVKSEMESQIKQLTTQLEQNRQPVTQIA---------NNDEKEKELTQQVKNLEEVNSKLQQEIVSLN
ESNVKSREEACRLGRENELLRANIKSTQAKCDELQRSL-----------------------------ESCTTRYAE
KEKASLEQLNVTLNSDLINSHEQEKERAKKVNELESTLSLLDDDKKRLELLIKENETLSEERVRTALNESESKDARITE
LENMHKIFVEKTERSQSEALDSLMATNNDRMKKIIS----------SKEEIIDELREELKVFKGKSGSFSEEITNLQHR
LEQQKQDLTTHVEELKTKISDITNNQSNYD-QQLVSEIENQKQLITQKTQELEERNAKIVELDSHVQKLQQDLSNEKEQ
VRMLEAEAKQMTQEHEKELR-------------SIRDQHHLALCQQQQESEILIREAKNGRVSLEEETSFLKRQAVKAA
STKSMNEITSQIEALKSELSAATESHAIVLNEKTTLQNDIQILENDKNQLSIQLKEIQGHHESLQASILENEKQKQGLQ
E----------ELSAAAALIVQREEQLSHLESENITMRNALQAAERDRASAVDKLETARSELLQAEERTQHLQQQYDLT
TSLSEYEIKIAQNSEQINLLQAQIESLKQVETQLQNLKSENEQLQQQATSYTELLKKREGELTNQQEKEQQLLHQKEE-
ETVYSQDMLELKEKLAAAQVKLMQTESELS-AVRQEKMKYETEAITNMSDLRSKLEEAECDRKRLQHA-VEQASQTDKQ
--EIKSIQTKYQESIKKRDVKLNNMESEIQ-------------------ELKKALESATSEKNLLNTTIEQQKKTFEDK
LEQLRVKYLTMCKRAENMEIELAAVTSRCSLL-------EKRLDEEFRRSVRNA----------SHSPVSLHFGGGNV
LDDINTRHQEEKKNQEKLAASLTETLNEKNTLVEKLKSEKNALVEERSQLKAKEELLQTTIERYRAETEVESNNNK---
CLRGIQSANVP----------PSASKQAKRSR----SVEQRVFTISGFDGT------E----LLEKIHKLPYATVAECK
----KKGVKRDSQDDTAPTSSSTMDPPKKKVKNDTPAKKEFKVIFSGFNDKIAKYTNAFKKTLEDYVKELGGVIVD---
SNSPVPTNLTHLVTNG-QLTVKLLTALVRGCWILPEAYVHESTKQKMWLDELSYGFRHVKLPIARKRIGFSEGFVS---
---DVDMGATHVVAPPGVRTIKVLQAAVAHIWIMNPEWLEECKKHSKLVDESSFGKLREEELFKGKNVGFTDAFKKECD
-----SRHLNTANLIIVEGGATVESNLTEADMILCTRSE-YGNMEN-TRAVTWDKLVELIYPVKIGVVAEEVKQPQ
NSKYKAQKRMTNAKTLVEGGSGAVTEMEEAELILVSAEESLSANGKTKRVFTWEGFLDYIFPKN----------


100/1.00
Leishmania infantum
87/0.93
Leishmania donovani
100/1.00
Leishmania major
92/0.95
Leishmania mexicana
100/1.00
Leishmania braziliensis
Endotrypanum monterogei
100/1.00
Leptomonas pyrrhocoris
85/0.94
99/1.00
Crithidia fasciculata
83/0.93
Phytomonas serpens
Strigomonas culicis
100/0.98
Angomonas deanei
Trypanosoma brucei brucei
100/1.00
Trypanosoma brucei gambiense
100/1.00
84/0.93
Trypanosoma congolense
71/0.82
Trypanosoma vivax
87/0.95
Trypanosoma cruzi
Trypanosoma theileri
98/0.95
91/0.98
Trypanosoma grayi
68/0.80
Trypanosoma carassii
Paratrypanosoma confusum
Bodo saltans
84/0.93
Trypanoplasma borreli
0.5

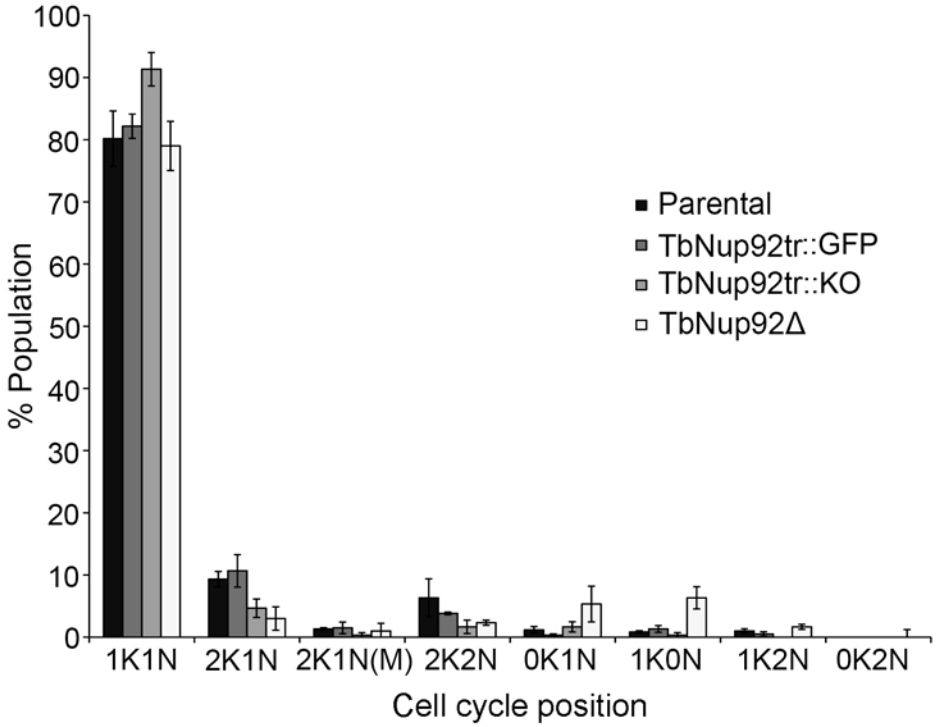

TbNup98::GFP
Merge
Interphase
Mitosis


A
120
100
80
60
40
20
0
% Expression
-Tet
+Tet
B
Cumulative cell number (x10-n/ml)
15
14
13
12
11
10
9
8
7
6
5
-Tet
-+Tet
0
1
2
3
4
5
6
7
Days Post Induction
C
100
80
60
40
20
0
% Population
0h
6h
12h
24h
48h
72h
1K1N
2K1N
2K2N
>2K2N
0K1N
1K0N
1K2N
Others
Karyotype

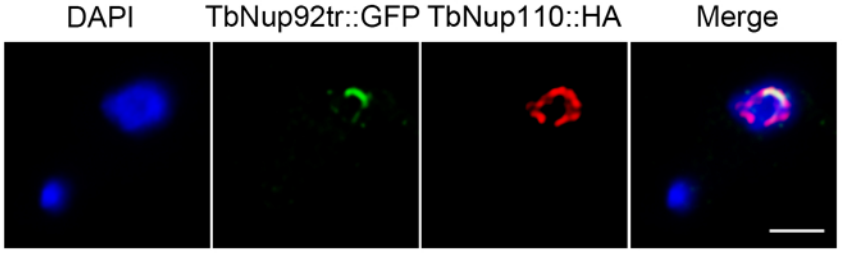
A


Parental
TbNup92tr::GFP
TbNup92tr::KO
TbNup92Δ
Cumulative cell number (x10-n/ml)
11
10
9
8
7
6
5
0
1
2
3
4
5
6
Days

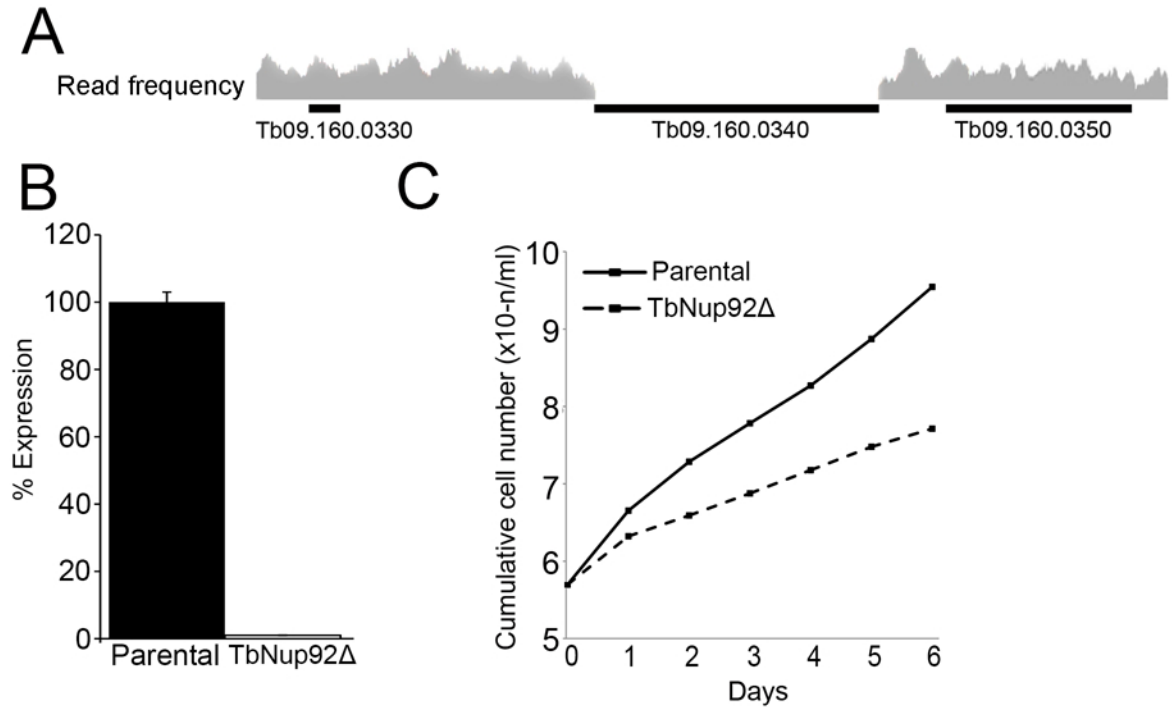
B

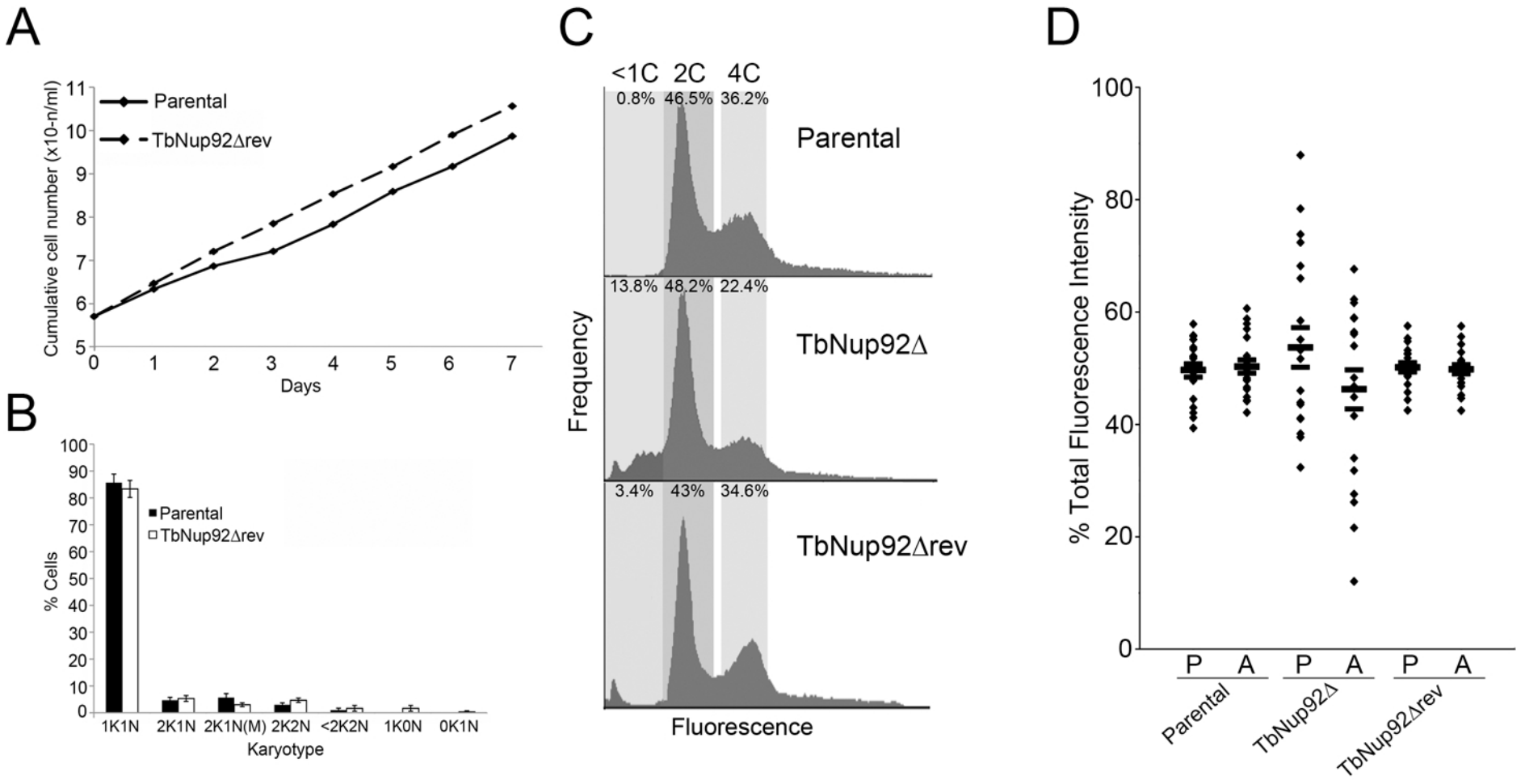

% Population
100
90
80
70
60
50
40
30
20
10
0
Parental
TbNup92tr::GFP
TbNup92tr::KO
TbNup92Δ
1K1N
2K1N
2K1N(M)
2K2N
0K1N
1K0N
1K2N
0K2N
Cell cycle position

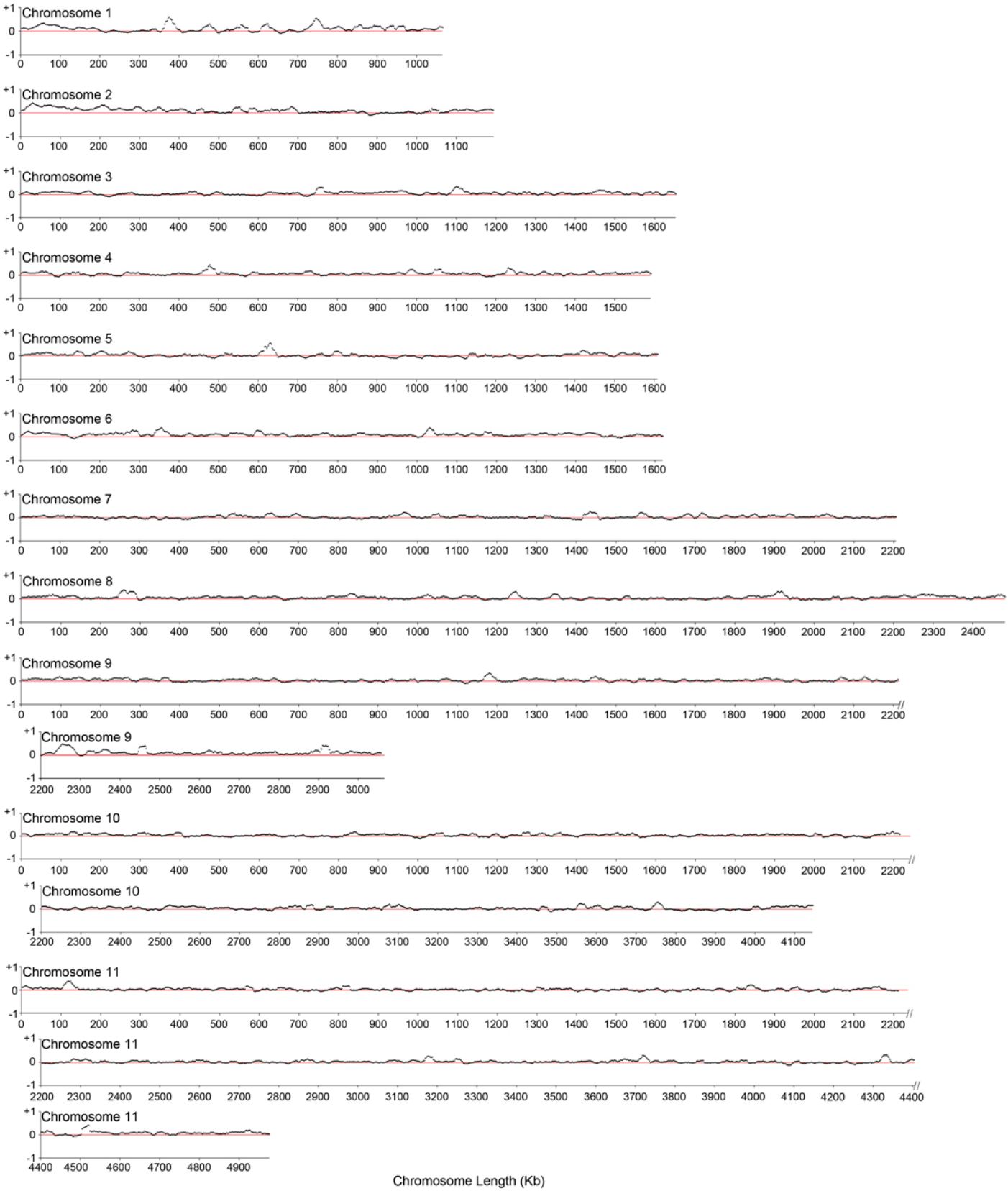
C

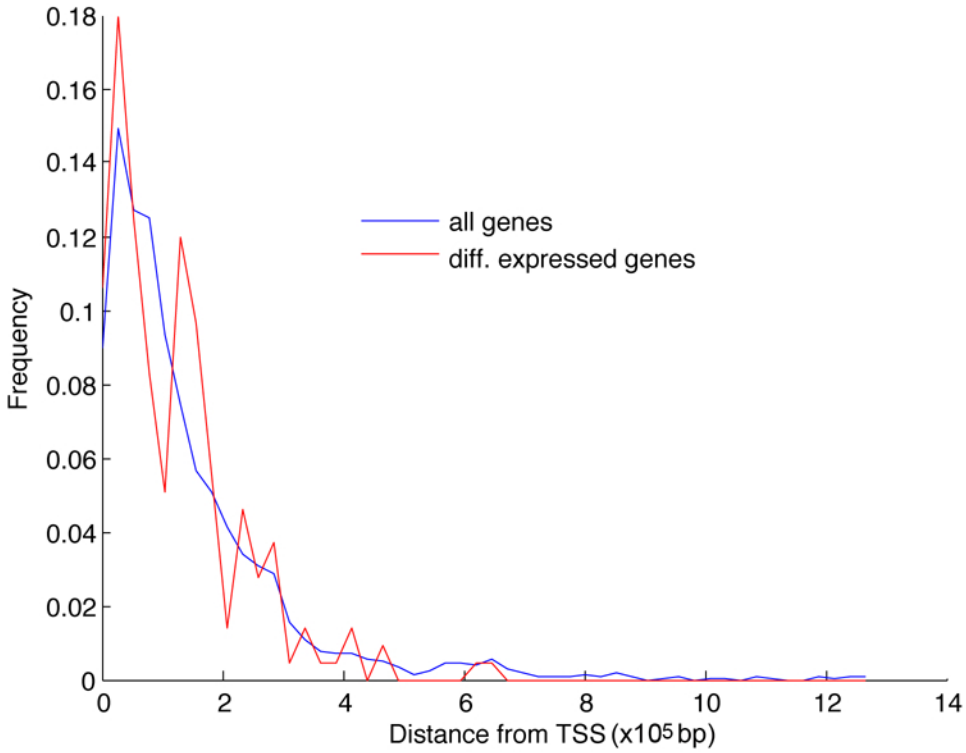

DAPI
TbNup92tr::GFP
TbNup110::HA
Merge

MSISESDSLWLQKSIHLADPIPEATGGGVEHMSANVPPSASKQAKRSRSVEQRVFTIS
GFDGTELLEKIHKLPYATVAECKSNSPVPTNLTHLVTNGQLTVKLLTALVRGCWILPEA
YVHESTKQKMWLDELSYGFRHVKLPIARKRIGFSEGFVSSRHLNTANLIIVEGGATVES
NLTEADMILCTRSEYGNMENTRAVTWDKLVELIYPVKIGVVAEEVKQPQQLMHIRSTR
KRSRSAMHLIKYPYDVPDYAYPYDVPDYAYPYDVPDYAYPYDVPDYA*


A
Read frequency
Tb09.160.0330
Tb09.160.0340
Tb09.160.0350
B
% Expression
120
100
80
60
40
20
0
Parental
TbNup92Δ
C
Cumulative cell number (x10-n/ml)
10
9
8
7
6
5
Parental
TbNup92Δ
0
1
2
3
4
5
6
Days


A
Cumulative cell number (x10-n/ml)
Parental
TbNup92Δrev
Days
B
% Cells
Parental
TbNup92Δrev
1K1N
2K1N
2K1N(M)
2K2N
<2K2N
1K0N
0K1N
Karyotype
C
<1C
2C
4C
0.8%
46.5%
36.2%
Parental
13.8%
48.2%
22.4%
TbNup92Δ
3.4%
43%
34.6%
TbNup92Δrev
Frequency
Fluorescence
D
% Total Fluorescence Intensity
P
A
Parental
TbNup92Δ
TbNup92Δrev


Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 9
Chromosome 10
Chromosome 10
Chromosome 11
Chromosome 11
Chromosome 11
+1
0
-1
Chromosome Length (Kb)


0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Frequency
all genes
diff. expressed genes
0
2
4
6
8
10
12
14
Distance from TSS (x10^5 bp)