

## RESEARCH ARTICLE

# A hub-and-spoke nuclear lamina architecture in trypanosomes

Norma E. Padilla-Mejia<sup>1</sup>, Ludek Koreny<sup>1</sup>, Jennifer Holden<sup>1</sup>, Marie Vancová<sup>2</sup>, Julius Lukeš<sup>2</sup>, Martin Zoltner<sup>1,3</sup> and Mark C. Field<sup>1,2,\*</sup>

## ABSTRACT

The nuclear lamina supports many functions, including maintaining nuclear structure and gene expression control, and correct spatio-temporal assembly is vital to meet these activities. Recently, multiple lamina systems have been described that, despite independent evolutionary origins, share analogous functions. In trypanosomatids the two known lamina proteins, NUP-1 and NUP-2, have molecular masses of 450 and 170 kDa, respectively, which demands a distinct architecture from the ~60 kDa lamin-based system of metazoa and other lineages. To uncover organizational principles for the trypanosome lamina we generated NUP-1 deletion mutants to identify domains and their arrangements responsible for oligomerization. We found that both the N- and C-termini act as interaction hubs, and that perturbation of these interactions impacts additional components of the lamina and nuclear envelope. Furthermore, the assembly of NUP-1 terminal domains suggests intrinsic organizational capacity. Remarkably, there is little impact on silencing of telomeric variant surface glycoprotein genes. We suggest that both terminal domains of NUP-1 have roles in assembling the trypanosome lamina and propose a novel architecture based on a hub-and-spoke configuration.

**KEY WORDS:** Lamina, Macromolecular assembly, Trypanosomatid, Nuclear organization, Heterochromatin

## INTRODUCTION

The nucleus is delineated by a double lipid membrane bilayer, the nuclear envelope (NE), and is supported by a proteinaceous lamina that influences nuclear shape, size and resilience to physical forces together with mechano-signalling capability (Gruenbam and Foisner, 2015; Swift and Discher, 2014). The lamina also interacts with the nuclear pore complex (NPC), thereby influencing the position, function, organization and modification of chromatin (Aaronson et al., 1975; Goldberg and Allen, 1996; Liu et al., 2000). Moreover, the lamina governs epigenetic regulation, DNA replication, transcription and the cell cycle, and thus is a major organizing principle within the cell (Verstraeten et al., 2007; Zheng et al., 2018; Kim et al., 2019). Most lamina-dependent processes are important to all eukaryotic lineages, making the ability to build a

lamina from distinct sets of proteins a remarkable example of convergent evolution (Koreny and Field, 2016). Moreover, many organisms lack any known lamina system, implying that yet more diversity remains to be uncovered.

In mammals, the lamina is comprised of ~60 kDa lamin proteins of two major subtypes, lamin A and lamin B. B-type lamins are expressed in all mammalian nucleated cells (including germline and stem cells), whereas A-type lamins (which includes lamins A and C; splice variants that are both encoded by *LMNA*) have a restricted expression profile, and are restricted to differentiated cells (Lehner et al., 1987; Rober et al., 1989; Constantinescu et al., 2006). Lamins form homotypic filaments distributed throughout the nucleus with the separate networks interacting in a complex manner (Goldberg et al., 2008; Shimi et al., 2008, 2015; Turgay et al., 2017; Nmezi et al., 2019). However, lamin B is more intimately associated with the inner NE, whereas lamin A faces the nucleoplasm and avoids regions of the NE proximal to NPCs. B-type lamin filaments are thinner ( $7.3 \pm 0.9$  nm; mean  $\pm$  s.d.) than A-type ( $16 \pm 1.7$  nm) as determined by expression in *Xenopus* oocytes (Goldberg et al., 2008). Lamin B is highly ordered into layers and related to stabilization of nuclear shape, whereas lamin A forms bundles and is more associated with mechanical rigidity (Turgay et al., 2017; Nmezi et al., 2019).

Lamins are composed of an N-terminal domain, or head, a central  $\alpha$ -helical rod and a globular C-terminal domain containing a nuclear localization signal (NLS), an Ig-fold domain and a CAAX-box prenylation motif (Gruenbam and Foisner, 2015; Dechat et al., 2008). These domains are implicated in membrane targeting and diverse contacts with multiple partners, including actin, nesprins, nucleoporins and histones (reviewed in Simon and Wilson, 2013). The central importance of lamins to correct cellular physiology is underscored by the plethora of lamin A mutations associated with heritable syndromes, known as laminopathies, most of which manifest as debilitating diseases (Kang et al., 2018).

Trypanosomes are protists of the Excavata supergroup, which separated from animals and their relatives over a billion years ago. The African trypanosome, *Trypanosoma brucei*, evolved a sophisticated strategy for establishing chronic infection in many mammalian hosts, which principally involves antigenic variation and mono-allelic expression of the superabundant variant surface glycoprotein (VSG). VSG is switched with sufficient frequency to facilitate a population continuing to infect the host (Mugnier et al., 2015; Pinger et al., 2017) despite robust host anti-VSG immune response (Stijlemans et al., 2016; Radwanska et al., 2018). For VSG switching to occur, monoallelic expression utilizes a dedicated transcriptional focus, the expression site body (ESB), together with telomeric silencing and silent VSG loci sequestered within heterochromatin (Figueroide and Cross, 2010). Hi-C analyses (Müller et al., 2018) highlight that subtelomeric regions bearing silent VSGs are folded into highly compact compartments with a high frequency of DNA–DNA contacts, likely important for maintaining a quiescent state. Significantly, monoallelic expression

<sup>1</sup>School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK. <sup>2</sup>Institute of Parasitology, Biology Centre and Faculty of Sciences, University of South Bohemia, 37005 České Budějovice, Czech Republic. <sup>3</sup>Department of Parasitology, Faculty of Science, Charles University in Prague, BIOCEV 252 50, Vestec, Czech Republic.

\*Author for correspondence (mfield@mac.com)

 M.C.F., 0000-0002-4866-2885

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Handling Editor: David Stephens  
Received 16 July 2020; Accepted 10 May 2021

and VSG switching are both impacted by disruption of the trypanosome nuclear lamina (DuBois et al., 2012; Maishman et al., 2016), suggesting a role in regulating subtelomeric surface antigen expression.

There are two known components of the trypanosome lamina, NUP-1 and NUP-2. Both are essential, have predicted coiled-coil structure and have molecular masses of 450 kDa and 170 kDa, respectively. NUP-1 and NUP-2 localize to the NE periphery and have a clear structural role, as depletion leads to abnormalities in nuclear morphology and NPC positioning. In addition to operating in close cooperation with each other, NUP-1 and NUP-2 influence positioning of telomeres and chromosomes, suggesting roles in chromosome and chromatin organization. Significantly, this includes effects on developmentally regulated genes, since knockdown in the mammalian form leads to an increase in levels of both normally silent VSG and procyclin transcripts, with these latter regulated proteins normally only being expressed in the insect stage (DuBois et al., 2012; Maishman et al., 2016). NUP-1 and NUP-2 lack any lamin-related domains (Koreny and Field, 2016) and are substantially larger, suggesting a distinct architecture to the metazoan lamin system, even though it shares many functions.

Here, we exploited a set of NUP-1 deletion mutants to dissect the trypanosome lamina *in vivo*, demonstrating that both terminal domains have crucial roles in lamina assembly. The interactions of different domains with partners such as NUP-2, NPC components and chromosomes suggest that NUP-1 termini constitute hubs in a lamina network with scaffolding properties.

## RESULTS

### NUP-1 domains have distinct spatial distribution

NUP-1 possesses distinct N- and C-terminal domains, separated by an extensive region of near perfect  $\alpha$ -helical repeats (DuBois et al., 2012). If extended as an  $\alpha$ -helix, each NUP-1 polypeptide can span over 400 nm and thus potentially contact much of the trypanosome nuclear volume (Field et al., 2012; DuBois et al., 2012). To monitor *in vivo* the distribution of NUP-1 domains, we chose to independently consider each domain in relation to each other. We tagged the N- and C-termini of NUP-1 with HA and GFP, respectively, and the repeat region was visualized with an in-house affinity-purified polyclonal antibody (DuBois et al., 2012).

In African trypanosomes, the cell cycle stage can be assessed from the number and position of the nuclei and kinetoplasts (the latter a highly organized network of mitochondrial DNA). During interphase a single nucleus and kinetoplast (1K1N cells) are present, the latter becoming elongated (bilobed) during nuclear G<sub>2</sub> phase (1Ke1N cells) to finally divide to produce cells with two kinetoplasts and only one nucleus (2K1N) prior to mitosis. After nuclear division, but prior to cytokinesis, cells with two nuclei are produced (2K2N) (Woodward and Gull, 1990; Benz et al., 2017). We found that both the N- and C-termini of NUP-1 were distributed similarly during interphase, but with distinct distributions in mitosis (Fig. 1). During inter and G<sub>2</sub> phases both termini localized to the nuclear periphery. During mitosis, the N-termini accumulated in the nucleoplasm, while the C-termini localized at the periphery. At later stages, the N-termini were absent from the contractile ring in the NE formed in telophase, while the C-termini remained present across the nuclear periphery. NUP-1 terminal domains were also differentially located during mitosis and cytokinesis, which suggests that they may also have specific functions and independently engage with the machinery separating the two daughter nuclei. This behaviour likely reflects the flexibility/

elasticity properties of NUP-1 as a coiled-coil and filamentous protein, with an ability to reposition during the cell cycle.

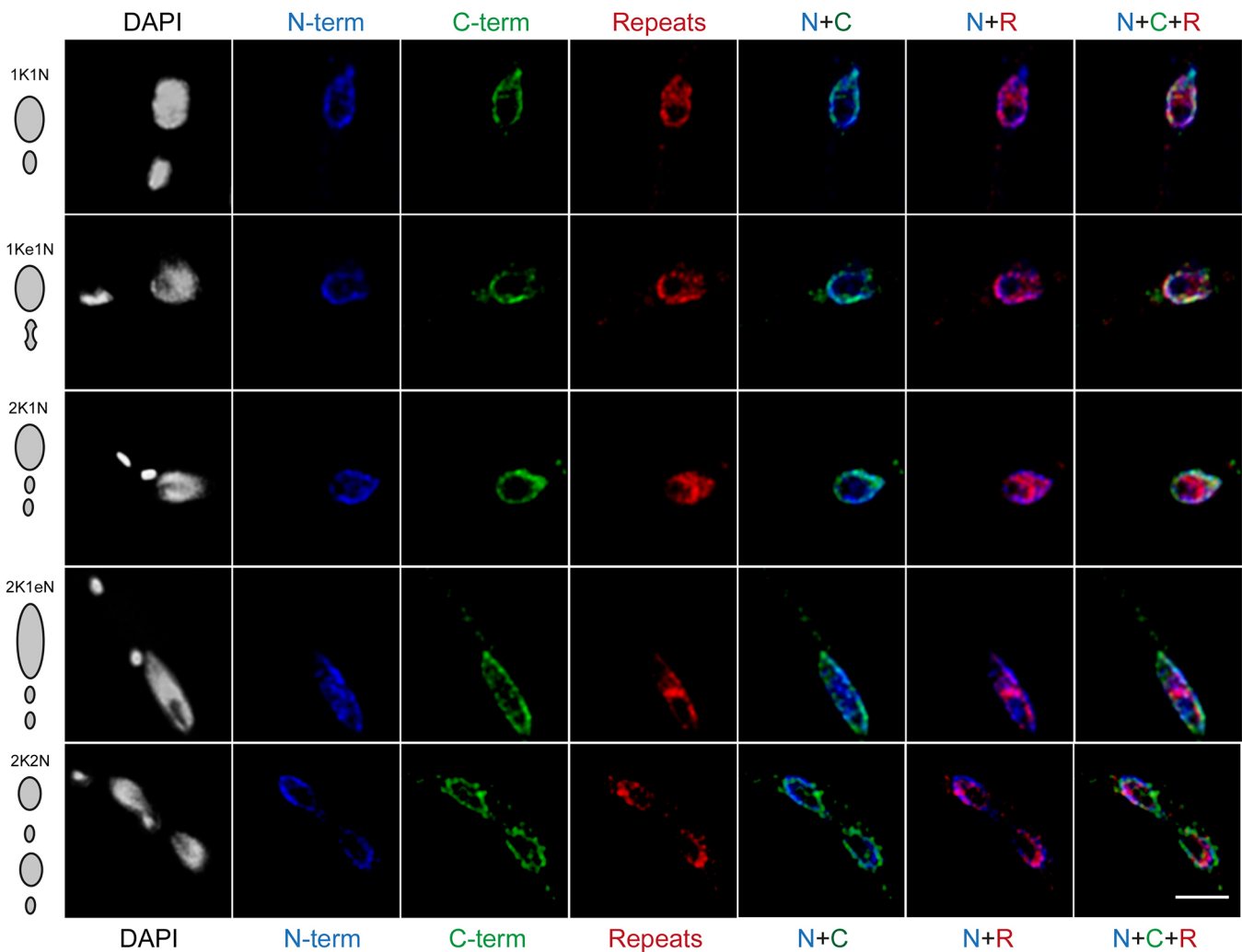
Moreover, we observed that the region constituted by  $\alpha$ -helical coiled coil repeats (NUP-1R, Fig. 2A) also had a unique location throughout the cell cycle; repeats were present at both the nuclear rim but also had a presence within more internal nuclear regions and this latter location became most pronounced at late mitosis/anaphase (Fig. 1; Fig. S3A,B,E). This suggests a dynamic retraction of the repeat domain to the poles as the nucleus completes division, with the possibility that the  $\alpha$ -helical repeats may interact with chromosomes at the 2K1N (early anaphase) stage, potentially being involved with their segregation to the daughter nuclei. This behaviour resembles that of the cohesins, which embrace sister chromatids from S-phase to anaphase. The NUP-1 repeat region shares high structural similarity with structural maintenance of chromosomes (SMC) proteins (Fig. S1), a superfamily of chromosomal DNA compaction proteins with DNA and ATPase activities, engaging in various processes of chromosome organization (Yatskevich et al., 2019). This does not exclude the possibility that the  $\alpha$ -helical repeats of NUP-1 interact with other components of the mitotic machinery.

### The N- and C-terminal domains of NUP-1 assemble as organized structures

The specific folding of individual domains of mammalian lamins facilitates precise assembly and higher order structure. A recent study of the molecular architecture of mammalian lamin A, mapping interactions within lamin dimers and polymers recognizes that head, tail, linkers and rod domains all contribute differentially to the molecular architecture. For example, the linkers and head-tail regions are proposed to act as 'springs' contributing to the dynamic stretch and flexibility of lamin A, with multiple electrostatic interactions between adjacent rods and between head-to-tail and adjacent rods within a lamin dimer (Makarov et al., 2019).

To determine whether there are similar domains with specific functions present in NUP-1, we ectopically expressed the individual terminal domains (DuBois et al., 2012). Three constructs were created, encoding the N-terminal domain, the C-terminal domain and a truncation with the entire repeat region deleted (denoted N+C) (Fig. 2A). All constructs were validated by western blotting using an anti-HA antibody. The protein sizes for the N-terminal, C-terminal and N+C variants were 83 kDa, 61 kDa and 144 kDa, respectively. In the case of the C-terminal variant, a second band of  $\sim$ 80 kDa was also detected, importantly, the presence of the 80 kDa band was clearly not in the parental line and was limited to tetracycline (Tet) induction conditions (Fig. S2A). We amplified and sequenced the tagged ectopic sequence of the C-terminal variant and confirmed that the construct transfected was as expected for the expression of the 61 kDa protein (data not shown). The reason for the presence of the  $\sim$ 80 kDa HA-tagged C-terminal peptide is unknown, although it could be attributable to post-translational modifications, chromatin configuration or transcription utilizing alternate start or stop codons. Nevertheless, the exact cause for the slower migrating form is still unclear.

Following expression of all three domain constructs (N-terminal, C-terminal and N+C), circular ordered structures were assembled within nuclei (Fig. 2B) as evidenced by immunofluorescence. Interestingly, two distinct distributions were seen for the C-terminal, one forming assemblies and one with a diffuse interior nuclear localization (Fig. 2B; Fig. S2E). The round assemblies from the three constructs presented different sizes, and we named larger structures (Fig. 2B; Fig. S3) as maxi-assemblies (mean $\pm$ s.d., 0.77 $\pm$ 0.2  $\mu$ m in



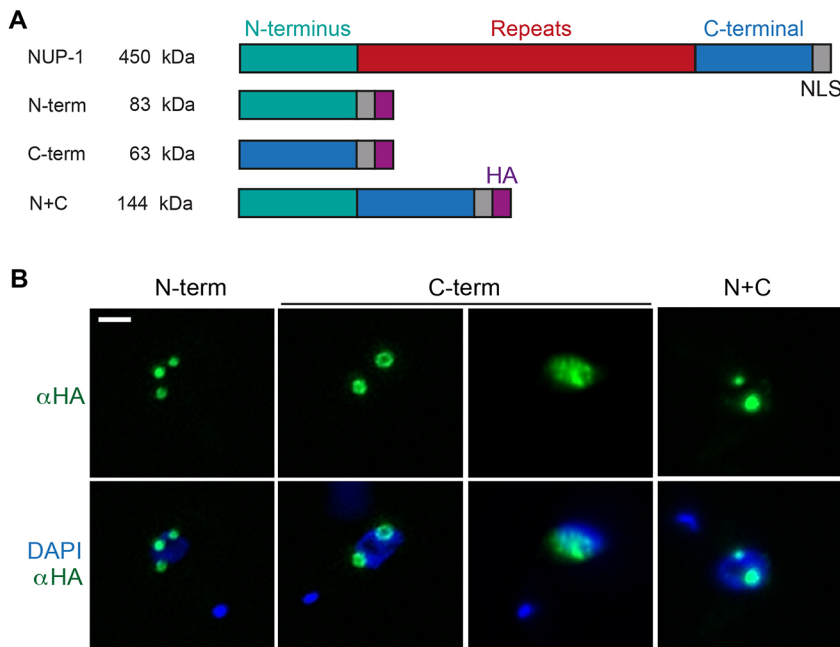
**Fig. 1. NUP-1 domains relocate during the cell cycle.** Cells expressing a doubled-tagged version of NUP-1 were imaged by confocal microscopy. DAPI was used to visualize DNA (white). The N-terminal (N-term) and the C-terminal (C-term) domains were tagged with HA (blue) and GFP (green), respectively. To visualize the repeat region of NUP-1 (repeats), affinity-purified rabbit antibodies raised against the repeat were used (red). The typical distribution of NUP-1 at the nuclear periphery is clear throughout the cell cycle. Number of kinetoplasts and nucleus per cell across the cell cycle are depicted on the right (see text for details). Scale bar: 2  $\mu\text{m}$ .

diameter, range 0.5–1.7  $\mu\text{m}$ ) and the smaller structures (Fig. S3) as mini-assemblies (0.38 $\pm$ 0.05  $\mu\text{m}$  in diameter, range 0.23–0.45  $\mu\text{m}$ ), respectively ( $n=30$  assemblies). Importantly, the diffuse nuclear pattern seen after C-terminal mutant expression is present concurrent only with mini-assemblies (Fig. S2E).

Moreover, the occurrence of assemblies was dependent on the concentration of Tet in the cultures and hence levels of protein produced, as well as the time of induction. We monitored the occurrence of assemblies with two different concentrations of Tet, 0.1 and 1.0  $\mu\text{g/ml}$ , over 24 h (Fig. S2C,D). For the N-terminal construct induced at 0.1  $\mu\text{g/ml}$  Tet, the number of assemblies remained low (mode=1 assembly) compared to inducing with 1  $\mu\text{g/ml}$  Tet (mode=4 assemblies). For mini-assemblies induced with 0.1  $\mu\text{g/ml}$  Tet, one to two assemblies were observed (range 1–10 per nucleus) whereas with 1.0  $\mu\text{g/ml}$  Tet the number ranged from 1–10 assemblies (Fig. S2C,D). For the C-terminal construct at low concentrations of Tet, the predominant phenotype was the diffuse nucleoplasmic pattern (frequency=0.8 in the population), and the frequency of assemblies in the population was low (<0.1 for maxi-assemblies, <0.02 for mini-assemblies). However, this proportion was reversed at high levels of Tet for maxi (mode=2, frequency

0.38), mini-assemblies (mode=1, frequency=0.12) and diffuse nucleoplasmic (frequency=0.26). As mentioned, mini-assemblies and the diffuse pattern can be found in the same nucleus (Fig. S2E), but not with maxi-assemblies. For the N+C construct, the frequency of assemblies with different concentrations of Tet remained similar (maxi and mini assemblies mode=1), with only a small difference in the range of assemblies per nucleus (Fig. S2C,D).

Furthermore, the time of induction also influences the number of assemblies per nucleus. We monitored the number of assemblies after induction with 1  $\mu\text{g/ml}$  Tet at three different times, 12, 24 and 48 h (Fig. S2D). For the N-terminal variant, after 12 h, the number of assemblies ranged from one to seven (maxi) or up to 10 (mini), without a clear mode. After 24 and 48 h, the modes for maxi-assemblies were four and two, respectively (i.e. the number of assemblies was reduced by half). For the C-terminal mutant, one to two assemblies per nucleus was the most predominant phenotype, and this did not change drastically. By contrast the diffuse nucleoplasmic localization did change, with just 8% of the cells in the population having this phenotype at 12 h and gradually increasing to 36% after 48 h (Fig. S2D). In cells showing a nucleoplasmic localized C-terminal mutant, this distribution



**Fig. 2. NUP-1 terminal domains oligomerize.** (A) The fragments of NUP-1 designed for expression. WT NUP-1 is indicated at the top. All mutant sequences were cloned into pDEX-577, a Tet-inducible system expressing HA-tagged proteins. Protein sizes are indicated at left. The endogenous nuclear localization signal (NLS, grey) of NUP-1 was introduced into the N-terminal variant; in other constructs the endogenous NLS was present. The position of the HA-epitope is in purple. (B) Immunofluorescence analysis in bloodstream forms after 24 h of induction with 1.0  $\mu$ g/ml of Tet. Cells were probed with an anti-HA antibody (green) showing round assemblies of expressed NUP-1 domains for the N-terminal (N-term), C-terminal (C-term) and a fusion of N and C terminal domains (N+C). Maxi-assemblies for all mutants are shown and, for the C-terminal mutant, the nucleoplasmic phenotype is also presented. DAPI was used to visualize DNA. Scale bar: 2  $\mu$ m.

co-occurred with mini-assemblies in  $\sim$ 30% of the cells in the population (Fig. S2E,F), as seen after 24 h of induction with 0.1 and 1.0  $\mu$ g/ml Tet. For the N+C truncation, the localization was stable across the time [maxi-assemblies mode=1 (24 h), mode=2 (48 h)]. Moreover, maxi-assembly behaviour was monitored throughout the cell cycle (Fig. S3D). For all the constructs, the mode was one for assemblies in interphase (1K1N) and post-mitosis (2K2N), with a tendency at this stage to remain as low numbers. During  $G_2$  phase (1Ke1N) and mitosis (2K1N) a broader range of number of assemblies per nucleus appeared. We hypothesize that, as the cell prepares for cell division, it probably also divides assemblies for inheritance by daughter nuclei.

Our interpretation of these observations is that for the N-terminal and repeat deletion constructs, mini-assemblies mature into maxi-assemblies as the proteins accumulate. Additionally, this also suggests that assemblies gradually build, and are influenced by time and levels of protein in the nucleoplasm. The C-terminal domain seems more sensitive to these factors, as it can also be found with a nucleoplasm diffuse localization (Fig. 2B; Fig. S2E), with the possibility that this domain extends into the nucleoplasm during the cell cycle. This is consistent with previous observations showing that the C-terminal domain is not required for positioning of NUP-1 at the nuclear membrane (DuBois, et al., 2012). When combined with the N-terminal sequences, the capacity of the C-terminal to assemble is increased since the N+C domain variant is never seen as nucleoplasmic and its behaviour resembles that of the N-terminal variant. This indicates that the ability of the N-terminal domain to self-assemble is likely stronger than that of the C-terminus.

Similar nuclear assemblies obtained with mutant forms of lamins in metazoan cells have been reported (Yang et al., 2013; Sylviois et al., 2008; Hübner et al., 2006; Izumi et al., 2000). The mechanism for formation of these foci is not well understood, but in some cases they are related to disease mutants that disrupt assembly and hence function. Interestingly, not all lamin mutants result in nuclear aggregates (Sylviois et al., 2008 and references therein). Although the mechanism of formation of assemblies from NUP-1 terminal domains is also unclear, the self-assembly properties of the N- and C-terminal domains is very pronounced. Furthermore, no obvious

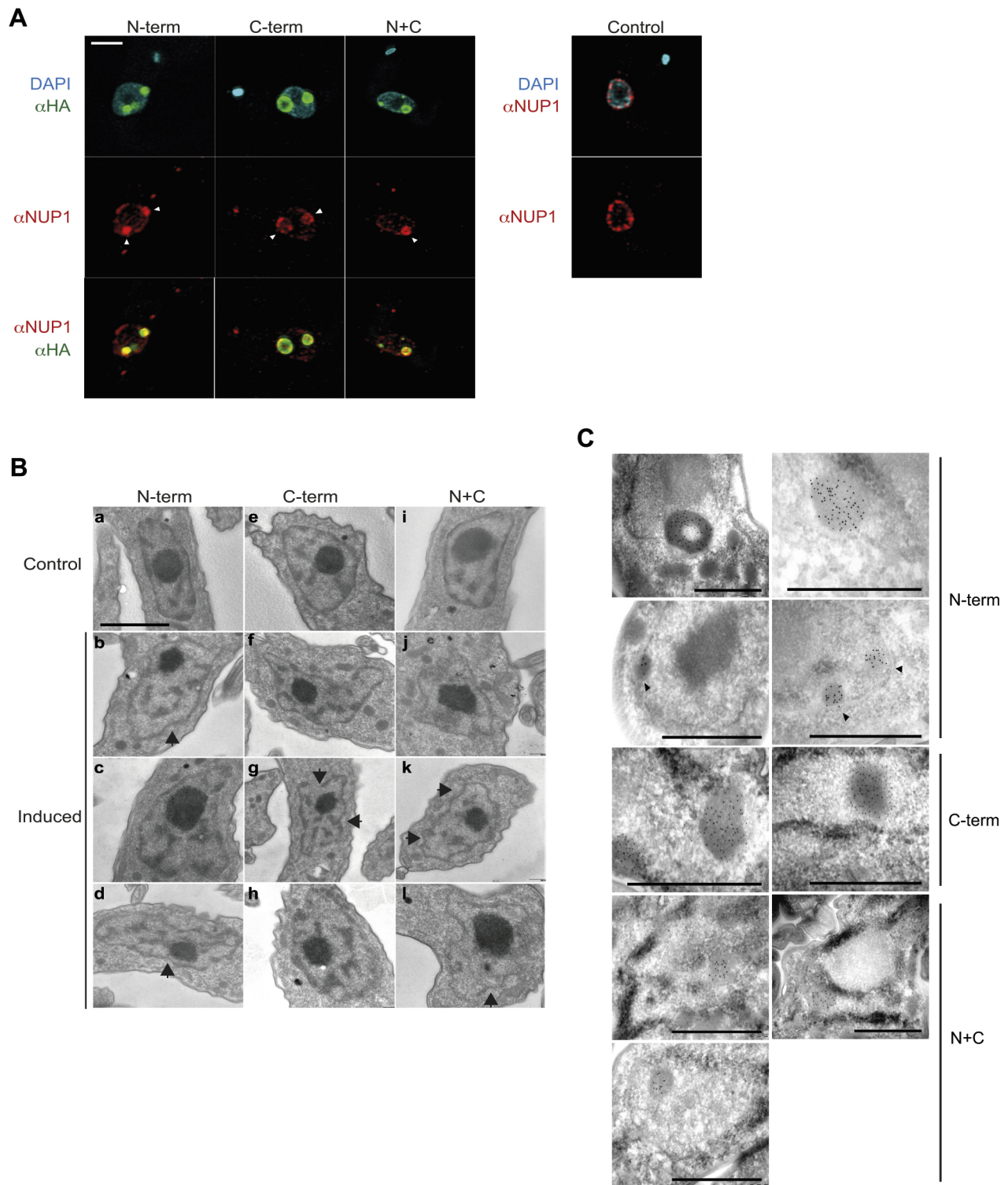
defects to overall cell morphology were observed, although a small increase in cell cycle time in a Tet dose-dependent manner was observed in induced cultures followed for up to 6 days (Fig. S2B). Thus, although there is a detrimental effect in terms of replication rate, indicating a loss of fitness, the presence of these NUP-1 assemblies is non-lethal, at least in the short term, as is the case for many lamin mutants (Sylviois et al., 2008; Muchir et al., 2003, 2004; Mounkes et al., 2005; Stuurman et al., 1999).

#### NUP-1 domains disrupt endogenous NUP-1 localization

To determine in more detail the impact of NUP-1 fragments upon lamina organization, we performed immunofluorescence using the NUP-1  $\alpha$ -helical repeat antibody to visualize the endogenous NUP-1 protein in the presence of the domain constructs. All three NUP-1 domain constructs colocalized with NUP-1 coiled-coil repeats and partially disrupted the nuclear peripheral distribution of NUP-1 (Fig. 3A). Importantly, although assemblies were stable across the cell cycle, associations between NUP-1-domain constructs and NUP-1 repeats (NUP-1R) from the endogenous NUP-1 were seen to be favoured during interphase (Fig. 3A; Fig. S3A–C).

Label-free mass spectrometry of whole-cell lysates indicated that expression levels of endogenous NUP-1 in the N-terminal and N+C domain-expressing cells were essentially unaltered compared to what was seen in wild-type cells (ratios  $0.82 \pm 0.04$  and  $0.97 \pm 0.08$  vs wild type, mean  $\pm$  s.d., respectively) and hence that endogenous protein is recruited to NUP-1 assemblies. By contrast, C-terminal domain-expressing cells accumulated more endogenous NUP-1 than wild-type cells (ratio  $2.21 \pm 0.47$  vs wild type), without significantly affecting proliferation (Fig. S2B), indicating that a modest excess of NUP-1 is tolerated.

With endogenous NUP-1 being recruited to the assemblies, we asked whether sequestering NUP-1 impacted NE integrity. Cells expressing NUP-1 domain constructs possessed an altered nuclear membrane morphology (Fig. 3B) with irregular boundaries (81%, 86% and 83% for N-terminal, C-terminal and N+C terminal variants, respectively,  $n \geq 18$  cells), consistent with altered/modified lamina support (percentage of cells with detectable irregular nuclear boundaries in control cells is 10%,  $n=70$  cells).



**Fig. 3. NUP-1 domains interact with and disrupt normal arrangement of endogenous lamina proteins.** BSF cells containing the NUP-1 domain constructs were fixed, stained and visualized by either confocal immunofluorescence microscopy (A) or electron microscopy (B,C). Overexpression of NUP-1 domains was induced with 1.0  $\mu$ g/ml of Tet for 24 h. (A) The normal NE arrangement of endogenous NUP-1 (control) is disrupted by the expression of NUP-1 N-term, C-term and N+C domain constructs. Cells were co-stained with anti-HA (green) and anti-NUP-1 repeat antibodies (red), and DAPI (cyan), as indicated. Central z-stacks are presented. Arrowheads highlight NUP-1 domain assemblies. Scale bar: 2  $\mu$ m. (B) Disruption of the normal structure of the nuclear membrane visualized by electron microscopy after expression of the N-term (b–d), C-term (f–h) domains and the N+C fusion (j–l). Respective control cells without induction are shown (a, e and i). Disruption is shown as irregular edges in the nuclear membrane. Arrows highlight examples of irregular nuclear membrane boundaries in panels b, d, g, k, l. Scale bar: 1  $\mu$ m. (C) Immunogold localization of NUP-1 mutant variants (N-, C-terminal and N+C fusion). Gold particles are detected in the assemblies of NUP-1 domain constructs, which also appear as electron dense. Black arrowheads show foci of NUP-1 assemblies in proximity to the nuclear membrane. Scale bars: 1  $\mu$ m.

We corroborated the presence of NUP-1 domain constructs by immunogold electron microscopy. We confirmed the presence of well-defined gold-labelled high-density circular structures inside the trypanosome nucleus in Tet-induced cells (Fig. 3C), correlating with the immunofluorescence observations (Figs 2B and 3A). These structures were frequently associated with the NE (Fig. 3C, black arrowheads), supporting our evidence that the constructs interact with endogenous NUP-1 (Fig. 3A) and possibly additional components of the nuclear membrane (see Figs 6 and 7A).

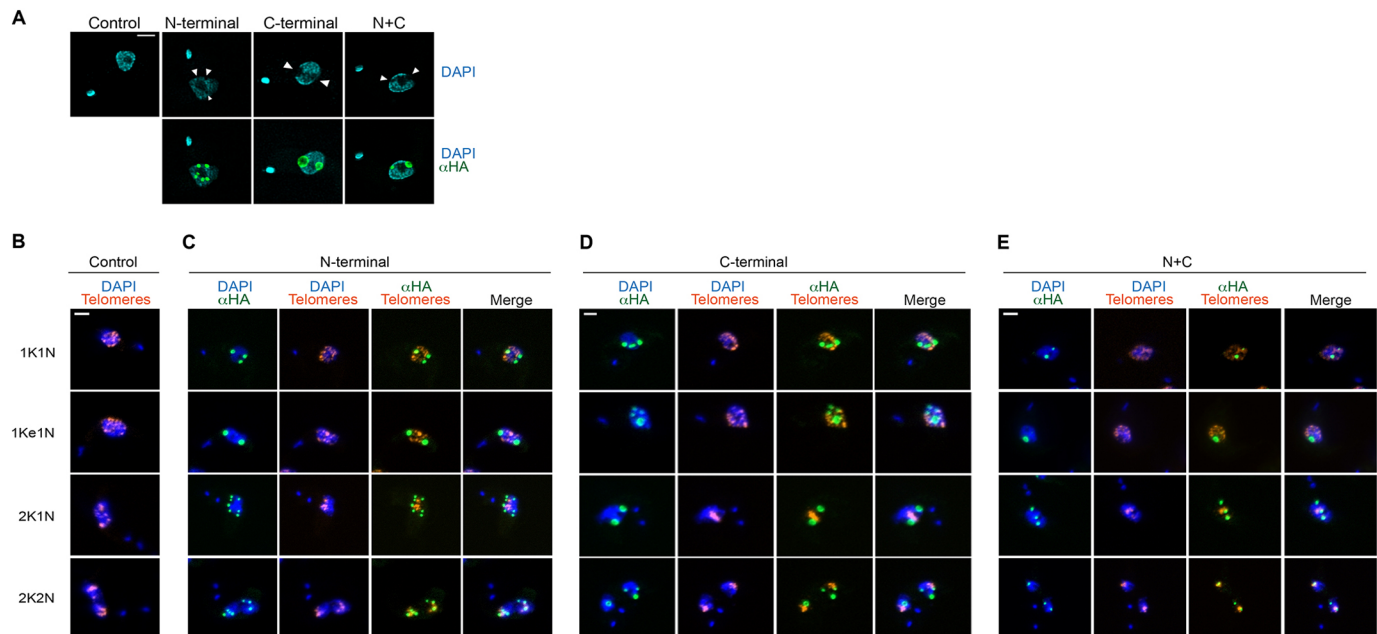
Furthermore, it is known that NUP-1 and NUP-2 are associated with a repressive heterochromatin environment and regulating expression of VSG genes (DuBois et al., 2012; Maishman et al., 2016), which normally organized into heterochromatin when in a quiescent state (Figueroide and Cross, 2010). Importantly, there was a normal retention of heterochromatin observed as electron-dense regions (Fig. 3B,C) indicating no major disruption to heterochromatic regions. This is consistent with transcriptome and proteome analyses (Fig. S5, Table S1), which provided no evidence for disruption to parental VSG 427-3 (alias VSG 224) expression or other VSG genes. Overall, these data suggest that heterochromatin, monoallelic expression and VSG switching are preserved in the subtelomeric VSG loci during expression of NUP-1 domain constructs.

#### NUP-1 domain localizations with chromatin and telomeres

The nuclear lamina controls gene expression by modulating chromatin organization, a mechanism common to all known lamina systems (Gruenbam and Foisner, 2015; Dechat et al., 2009; Koreny and Field, 2016). Expression of NUP-1 domain constructs led to voids in DNA as observed by DAPI staining and revealed by super-resolution immunofluorescence (Fig. 4A). This

alteration of DNA distribution was observed for all three NUP-1 fragments and the phenomenon may also contribute to the disturbed morphology of the NE (Fig. 3B). It is most likely that this is a physical phenomenon, whereby the DNA is simply excluded from dense NUP-1 domain regions, and presumably the free energy of NUP-1 domain self-assembly is sufficient to exclude chromatin. Similar voids in chromatin distributions have also been reported to occur in COS7 cells with a lamin A mutation (Q432X) (Yang et al., 2013), which is a mutation that is also known to lead to cardiac disease (Yang et al., 2013; Sylvius et al., 2008), although the mechanisms causing such voids in the DNA distribution in two different models (COS7 cells and trypanosomes) while expressing mutated versions of a lamina protein lacks clarity.

Given evidence that NUP-1 modulates positioning and movement of the telomeres (DuBois et al., 2012; Field et al., 2012), we performed fluorescence *in situ* hybridization (FISH) to obtain evidence for targeting of NUP-1 terminal domains to telomeres (Fig. 4B–E). We did not detect a strong association between any of the NUP-1 constructs and telomeres across the cell cycle. Nevertheless, during mid- and late mitosis, telomeres (compacted and aligned in the centre of the nucleus) approach NUP-1 assemblies and occasionally occur in the same nuclear foci, yet, no evidence of significant interaction between these nuclear structures was detected and co-occurrence may simply represent segregation of telomeres and NUP-1 termini into the daughter nuclei. In spite of the presence of NUP-1 assemblies, telomeres segregate normally, consistent with cells being tolerant to the presence of the assemblies during several days. Moreover, during mitosis, assemblies also migrate towards opposite poles of the nucleus (Fig. 4B; Fig. S3A–C).



**Fig. 4. NUP-1 domains co-occur with chromatin without disruption.** Cells bearing NUP-1 variants tagged with HA were fixed, stained and visualized by confocal immunofluorescence microscopy with a Leica System (A) or Zeiss system (B–E). In all cases, overexpression of NUP-1 domains was achieved using 1.0 µg/ml of Tet during 24 h. (A) Chromatin is displaced by the overexpression of the three NUP-1 variants. Arrowheads show the void of DNA created by presence of NUP-1 assemblies. SMB cells lacking the overexpression system are used as control. Cells are stained with anti-HA antibody (green) and DAPI is used to visualize DNA (cyan). (B–E) FISH assay. Cells expressing the NUP-1 mutants were used for FISH assay, using a telomere-pairing probe (orange) and an anti-HA antibody to recognize NUP-1 variants (green). DAPI was used to visualize DNA (blue). (B) Control SMB cells displaying the normal telomere arrangement. Interaction of telomeres with the overexpression mutants are shown in (C) N-terminal, (D) C-terminal and (E) N+C mutants. Central z-stacks are shown. All scale bars: 2 µm.

### NUP-1 interacts with specific NPC components

In the mammalian bloodstream trypanosome, multiple mechanisms ensure mono-allelic expression of a single VSG from a telomeric expression site (Mugnier et al., 2015; Faria et al., 2019; Glover et al., 2016; Saura et al., 2019). In some insect stages, the VSG coat is replaced by procyclin (Roditi et al., 1989), and similar to VSG, procyclin genes are transcribed by RNA Pol I, but from chromosomal internal sites rather than a telomere. Importantly, NUP-1 participates in silencing of both VSG and procyclin genes (DuBois et al., 2012). With both N-terminal and C-terminal domains occasionally coincident with telomeres, we asked whether these constructs triggered alterations in the global proteome, and undertook unbiased, label-free mass spectrometry of whole-cell lysates to address this.

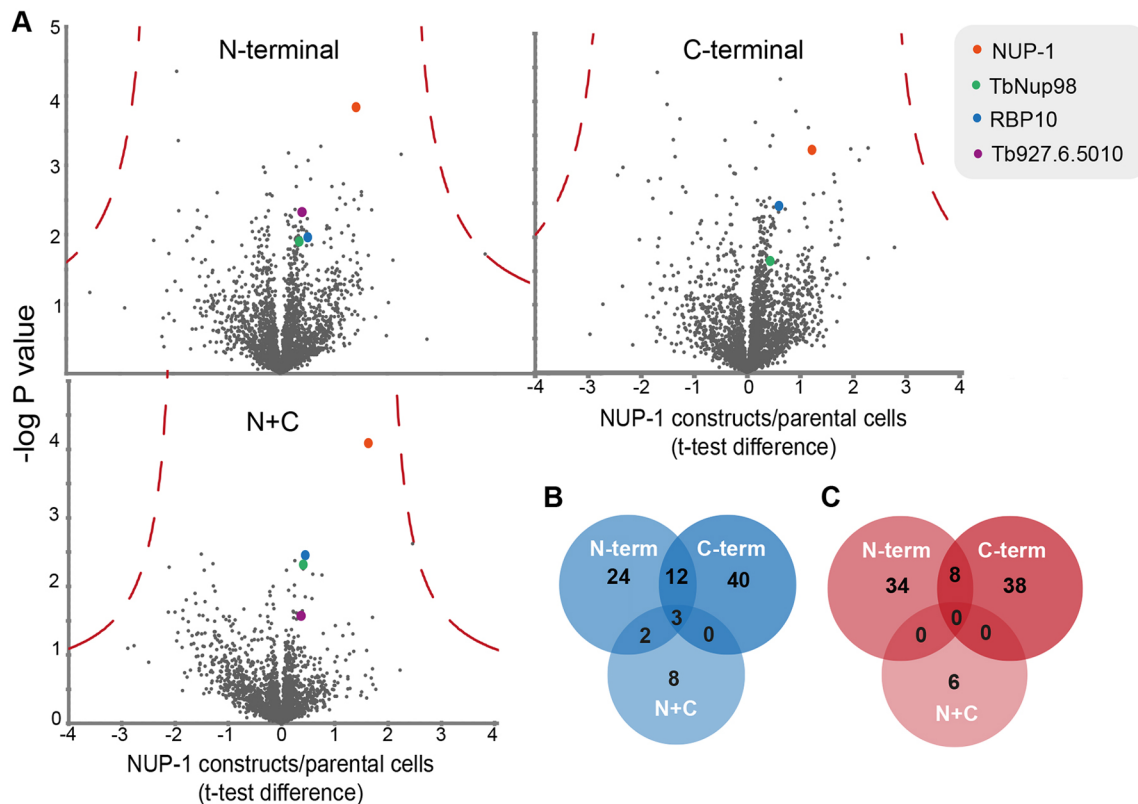
Over 2500 protein groups were identified and quantified (Fig. 5; Table S1). For selection of differentially expressed proteins, we applied the following inclusion filters: (1) at least two unique peptides identified, (2) ratio  $> \pm 0.2$ , (3) statistical significance ( $\log P$ )  $> 1.5$   $-\log P$ -value and (4)  $t$ -test difference of  $\pm 0.3$  with respect to control cells. Following filtering, 83, 101 and 19 differentially expressed protein groups were detected for cells expressing N-terminal, C-terminal and N+C domains, respectively (Fig. 5), and which corresponds to 1%, 1.2% and 0.23% of the predicted proteome (Aslett et al., 2010). Therefore, there is minimal overall impact on the proteome upon expression of NUP-1 constructs. The false discovery rate (FDR) was calculated (FDR=0.01,  $s_0=2$ ), and although none of the proteins met this cut off, some differentially expressed proteins (proteins with changes in

their corresponding expression levels) showed consistency in all three replicates.

As expected, peptides corresponding to NUP-1 were considerably more abundant in all three cell lines. Ratios for NUP-1 termini versus control cells were  $6.33 \pm 2.18$ ,  $3.75 \pm 3.25$  and  $4.67 \pm 0.99$  (mean  $\pm$  s.d.) for N-, C- and N+C terminal constructs, respectively. As described above, compensatory upregulation of endogenous NUP-1 is only observed for the C-terminal variant.

Furthermore, only the nucleoporin TbNup98 (Tb927.3.3180) and RNA-binding protein 10 (RBP10) (Mugo and Clayton, 2017) were upregulated in all three domain cell lines (Table 1; Table S1). TbNup98 is a PheGly (FG) nucleoporin component of the NPC and likely restricted to kinetoplastids (Obado et al., 2016). RBP10 is an RNA-binding protein that functions as a major regulator of development (Mugo and Clayton, 2017).

Among proteins upregulated in cells expressing N-terminal domains were an mRNA-binding protein (Tb927.6.5010) and Tb927.11.2750 (Table 1; Table S2). The Tb927.6.5010 gene product corresponds to a potential mRNA fate regulator, acting as a post-transcriptional repressor (Lueong et al., 2016; Erben et al., 2014; Goos et al., 2017). The gene product of Tb927.11.2750 was also upregulated and is restricted to *T. brucei*, *T. gambiense*, *T. evansi* and *T. cruzi*. Furthermore, downregulated proteins were also detected (Fig. 5C; Table 1, Table S3). Eight such proteins were quantified with high confidence in both the N-terminal and C-terminal domain (Table 1). This list includes an RNA helicase (Florini et al., 2019) and diptine synthase (Aslett et al., 2010), which has been implicated in different aspects of RNA metabolism and localized in the nucleus



**Fig. 5. NUP-1 domain constructs lead to specific changes to the proteome.** (A) Volcano plots of protein abundance changes in cells expressing N-terminal, C-terminal and N+C constructs from three replicates. Cells expressing the N-, C-terminal and N+C mutants were induced with Tet (1.0  $\mu$ g/ml) for 24 h. Statistical differences (as compared to parental SMB cells) are plotted against  $-\log P$  value. An FDR of 1% is shown as a dotted red line. NUP-1 is the orange dot, together with TbNup98 (green), RBP10 (blue) and protein Tb927.6.5010 (purple). (B,C) Venn diagrams show differentially expressed proteins, either (B) upregulated or (C) downregulated in cells expressing NUP-1 variants.

**Table 1. Differentially expressed proteins in N-terminal, C-terminal and N+C terminal construct harbouring cells**

| Protein name   | Construct expressed  |           |            |           |       |           |     |
|--|--|-----------|------------|-----------|-------|-----------|-----|
|  | N-terminal   |           | C-terminal |           | N+C   |           |     |
|  | Ratio  | $-\log P$ | Ratio      | $-\log P$ | Ratio | $-\log P$ |     |
| <b>Upregulated proteins</b>                          |  |           |            |           |       |           |     |
| <b>Hits common to N-terminal, C-terminal and N+C</b> |  |           |            |           |       |           |     |
| Tb927.2.4230   | NUP-1  | 2.74      | 3.8        | 2.35      | 3.2   | 3.14      | 4.1 |
| Tb927.8.2780   | RNA-binding protein RBP10  | 1.44      | 1.9        | 1.49      | 2.4   | 1.35      | 2.3 |
| Tb927.3.3180   | Nucleoporin TbNup98  | 1.26      | 2.0        | 1.29      | 1.7   | 1.31      | 2.3 |
| <b>Hits common to N-terminal and N+C</b>             |  |           |            |           |       |           |     |
| Tb927.6.5010   | Hypothetical protein, conserved  | 1.33      | 2.3        |           |       | 1.28      | 1.5 |
| Tb927.11.2750  | POMP12, Present in the outer mitochondrial membrane proteome 12        | 1.29      | 1.7        |           |       | 1.37      | 1.8 |
| <b>Hits common to N-terminal and C-terminal</b>      |  |           |            |           |       |           |     |
| Tb927.1.4050   | Ser/thr protein phosphatase, putative                                  | 1.44      | 3.1        | 1.40      | 3.5   |           |     |
| Tb927.10.12030                                       | Hypothetical protein, conserved  | $\infty$  | 3.1        | $\infty$  | 1.7   |           |     |
| Tb927.10.4900  | TPR-repeat-containing chaperone protein DNAJ, putative                 | 2.33      | 1.9        | 1.35      | 1.6   |           |     |
| Tb927.11.4000  | Hypothetical protein, conserved  | 1.36      | 2.1        | 1.30      | 2.1   |           |     |
| Tb927.3.780  | Proteasome alpha 7 subunit   | 1.31      | 1.9        | 1.38      | 1.7   |           |     |
| Tb927.4.870  | Dynein heavy chain, putative   | 1.27      | 2.0        | 1.30      | 1.9   |           |     |
| Tb927.5.2950   | Component of motile flagella 3   | 1.33      | 2.1        | 1.35      | 2.6   |           |     |
| Tb927.8.1540   | Hypothetical protein, conserved  | 1.48      | 2.8        | 1.27      | 1.5   |           |     |
| Tb927.8.3250   | Dynein heavy chain, putative   | 1.32      | 2.1        | 1.26      | 1.9   |           |     |
| Tb927.9.3280   | Exopolyphosphatase   | 1.39      | 1.9        | 1.34      | 1.5   |           |     |
| Tb927.9.10370  | TAX-1  | $\infty$  | 1.5        | $\infty$  | 1.5   |           |     |
| Tb927.9.6290   | Arginine kinase  | 1.24      | 1.7        | 1.28      | 1.5   |           |     |
| <b>Downregulated proteins</b>                        |  |           |            |           |       |           |     |
| <b>Hits common to N-terminal and C-terminal</b>      |  |           |            |           |       |           |     |
| Tb927.5.4420   | Nucleolar RNA helicase II, putative                                    | 0.81      | 1.5        | 0.75      | 3.4   |           |     |
| Tb927.3.5050   | 60S ribosomal protein L4   | 0.77      | 1.7        | 0.81      | 1.8   |           |     |
| Tb927.8.6330   | WD domain, G-beta repeat/PFU (PLAA family ubiquitin binding), putative | 0.73      | 1.9        | 0.68      | 2.2   |           |     |
| Tb927.4.4910   | 3,2-trans-enoyl-CoA isomerase, mitochondrial precursor, putative       | 0.62      | 1.8        | 0.51      | 1.9   |           |     |
| Tb927.9.1850   | 60S ribosomal protein L35, putative                                    | 0.36      | 2.3        | 0.38      | 1.6   |           |     |
| Tb927.4.4650   | Diphthine synthase, putative   | 0.22      | 1.9        | $\infty$  | 1.9   |           |     |
| Tb927.10.7140  | Membrane-bound acid phosphatase 2                                      | $\infty$  | 2.3        | $\infty$  | 2.1   |           |     |
| Tb927.11.15950                                       | Amino acid transporter, putative                                       | $\infty$  | 1.9        | $\infty$  | 3.0   |           |     |

$\infty$  for upregulated proteins, ratio  $\geq 2.88 \times 10^9$ ;  $\infty$  for downregulated proteins, ratio  $\leq 3.67 \times 10^{-11}$ .

(Dean et al., 2017). Tb927.11.15950, another downregulated protein, is annotated as a transporter (Aslett et al., 2010), and recognized by BLAST to have analogy with nucleobase transporters. Moreover, Tb927.11.15950 shows partial nuclear localization (Dean et al., 2017). These results suggest that the NUP-1 domain constructs may attenuate nuclear RNA processing.

With these changes at proteome level, we performed RNA-seq to detect potential transcriptome alterations as a result of overexpression of NUP-1 domains. No evidence for differential expression was found for any transcript (Fig. S5), and the abundance of VSG and procyclin mRNA were unaltered. Similarly, proteomics revealed no modification in VSG expression compared to the parental cell line. Although originally expressing VSG 427-2 (Tb427.BES40.22) (Wirtz et al., 1999), we uncovered a switch in the parental cell line to VSG 427-3 (Tb427.BES65.13) but beyond this, no alterations were detected.

#### **TbNup98 interacts with both N-terminal and C-terminal domains of NUP-1**

TbNup98 has multiple connections to different regions of the NPC, including other FG Nups, the nuclear basket and the inner and outer rings (Obado, et al., 2016). As proteomics suggested a connection between NUP-1 and TbNup98, we analysed the locations of NUP-1 domain constructs and TbNup98. As previously reported (DuBois et al., 2012; DeGrasse et al., 2009), TbNup98 clearly appeared as punctua at the NE, consistent with an NPC association (Fig. 6A).

After overnight induction, TbNup98 colocalized with the C-terminal assemblies, but not the N-terminal and N+C assemblies. In these latter cases, where these constructs were located there was a zone depleted of nuclear pore complexes (Fig. 6A; Fig. S6). However, after 72 h induction, TbNup98 colocalized with all three NUP-1 domain constructs and the associations were retained across the cell cycle, supporting a potential interaction. This suggests that, although both NUP-1 termini can interact with TbNup98, the C-terminus is more likely to support such interactions.

#### **NUP-1 depends on TbNup98 to maintain NE structure**

With proteomics and immunofluorescence analysis suggesting an interaction between the NUP-1 termini and TbNup98, we decided to explore the relationship further by gene silencing. After 24 h of TbNup98 knockdown the classical distribution of NUP-1 at the nuclear periphery was lost, and the protein instead clustered at specific points of the NE. Moreover, TbNup98-depleted cells possessed nuclei with an altered morphology, including blebs and protuberances (Fig. 6B). After 48 h induction, NUP-1 clustering became more significant in aberrant TbNup98-depleted nuclei (Fig. 6B, arrowheads). Significantly, 'monster' cells appeared, with evident damage to nuclear shape and aberrant foci of chromatin separated from the nucleus (Fig. 6B, white star). The ploidy of cells was altered following TbNup98 depletion as detected by flow cytometry, with a gradual decrease in diploid cells and an increase in tetraploid and higher polyploid (>4n) cells in the population

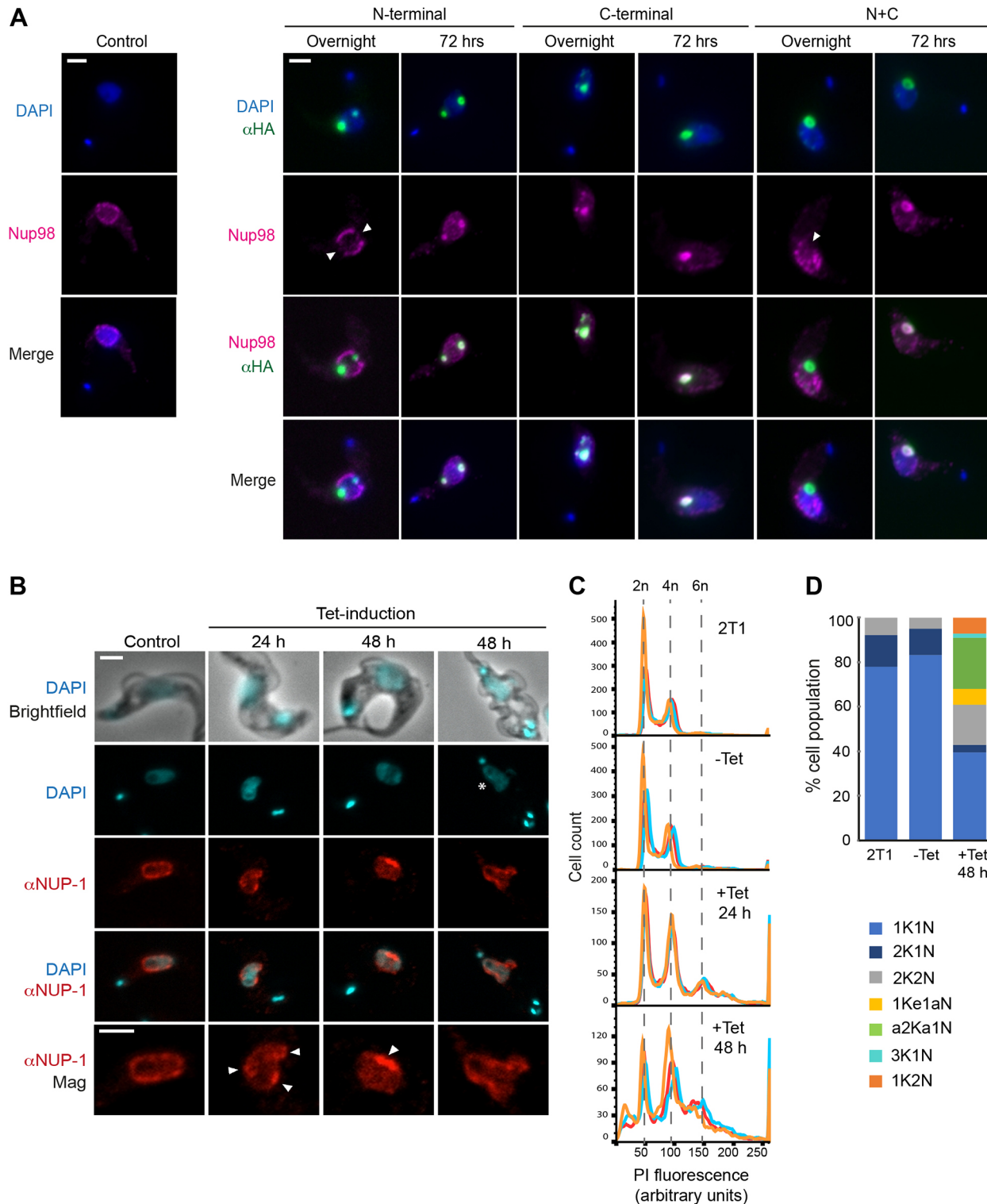


(Fig. 6C). This suggests that DNA duplication takes place, but cells are incapable of completing mitosis and cytokinesis. Moreover, TbNup98 knockdown cells exhibited abnormal DNA-containing bodies, with defects in the segregation, shape and number of kinetoplasts and nuclei (Fig. 6D) as detected by microscopy analysis ( $n=100$  cells). In particular, 2K1N cells bearing extra structures containing chromatin were prevalent among these abnormal cells. These results indicate that TbNup98, apart from

its function as part of the NPC, has an influence on mitosis, cytokinesis and/or normal segregation of chromatin and participates with NUP-1 to maintain NE integrity.

#### The lamina protein NUP-2 mainly interacts with the NUP-1 N-terminus

NUP-2 is the second defined component of the trypanosome lamina. NUP-1 and NUP-2 are intimate interactors and cooperate to



**Fig. 6. NUP-1 depends on TbNup98 to maintain NE structure.** (A) TbNup98 and Tet-induced (1.0 µg/ml) cells bearing NUP-1 constructs were visualized after 16 (overnight) and 72 h by confocal immunofluorescence microscopy. Uninduced cells are used as control to visualize the distribution of TbNup98. After overnight induction, there are regions lacking TbNup98 normal distribution (arrowheads) in the N-terminal and N+C mutants whereas after 72 h, all NUP-1 variants are coincident with TbNup98. Central z-stacks are shown. Scale bars: 2 µm. (B) The impact of TbNup98 knockdown on normal NUP-1 distribution was followed at 24 h and 48 h. Cells were fixed, stained and visualized by confocal immunofluorescence microscopy. Cells are co-stained for NUP-1 repeat (red) and DAPI (blue), as indicated. After induction, NUP-1 starts to localize in foci (arrowheads). After 48 h, 'monster cells' start to appear. Abnormal clustering of NUP-1 distribution is seen (arrowheads) as well as aberrant DNA containing-bodies (white star). Central z-stacks are shown. Scale bars: 2 µm. (C) Flow cytometry analysis of DNA content in TbNup98 depleted cells. Histograms indicating number of cells versus propidium iodide (PI) fluorescence. The analysis was conducted for control cells (2T1 cells), uninduced cells and Tet-induced cells at 24 and 48 h. Histograms represent three independent experiments, represented by orange, blue and red lines. Peaks labelled 2n represent diploid cells; 4n, tetraploid cells and 6n/8n, higher ploidy cells. Flow cytometry profiles for 10,000 propidium iodide-labelled cells are shown. (D) Cell cycle progression after TbNup98 silencing followed by microscopy. 2T1, uninduced and induced RNAi cells ( $n=100$  cells each) were fixed and stained with DAPI. Normal 1K1N, 1Ke1N, 2K1N and 2K2N cells were detected. In the population of induced cells, a series of aneuploid cells were observed: 1Ke1aN (abnormal nuclei), a2Ka1N (cells in 2K1N with abnormal kinetoplasts attached to amorphous nuclei), 3K1N, 1K2N.

maintain NE architecture (Maishman et al., 2016). To better understand interactions between NUP-2 and NUP-1, we used a cell line co-expressing a TY1-tagged version of NUP-2 together with NUP-1 domain constructs. We observed the canonical distribution of NUP-2 across the nuclear periphery in wild-type cells (Fig. 7A). In cells expressing NUP-1 mutants, NUP-2 clearly associated and colocalized with N-terminal and N+C variants, forming specific foci (Fig. 7A). Interestingly, NUP-2 was not found to colocalize strongly with the C-terminal domain or endogenous NUP-1 repeats. Importantly, all these interactions were maintained across the cell cycle (Fig. S7), suggesting that NUP-2 mainly interacts with the N-terminal domain of NUP-1.

## DISCUSSION

Several lamina systems are now known, specifically the 'metazoan' lamins, nuclear matrix constituent proteins (NMCPs) in plants and NUP-1 and -2 in kinetoplastids (Koreny and Field, 2016). Lamins assemble to form fibrils with a precise architecture governed by interactions between specific domains, with parallel dimers assembling in antiparallel fashion (Ahn et al., 2019; Makarov et al., 2019) and that are anchored to the NE via C-terminal prenylation. Lamin B is likely capable of supporting lamina functions alone, a conclusion supported by both expression profiles and phylogenetics (Ahn et al., 2019; Koreny and Field, 2016), while both A and B lamins can interact directly with core histones with roles in the formation of lamina-associated domains.

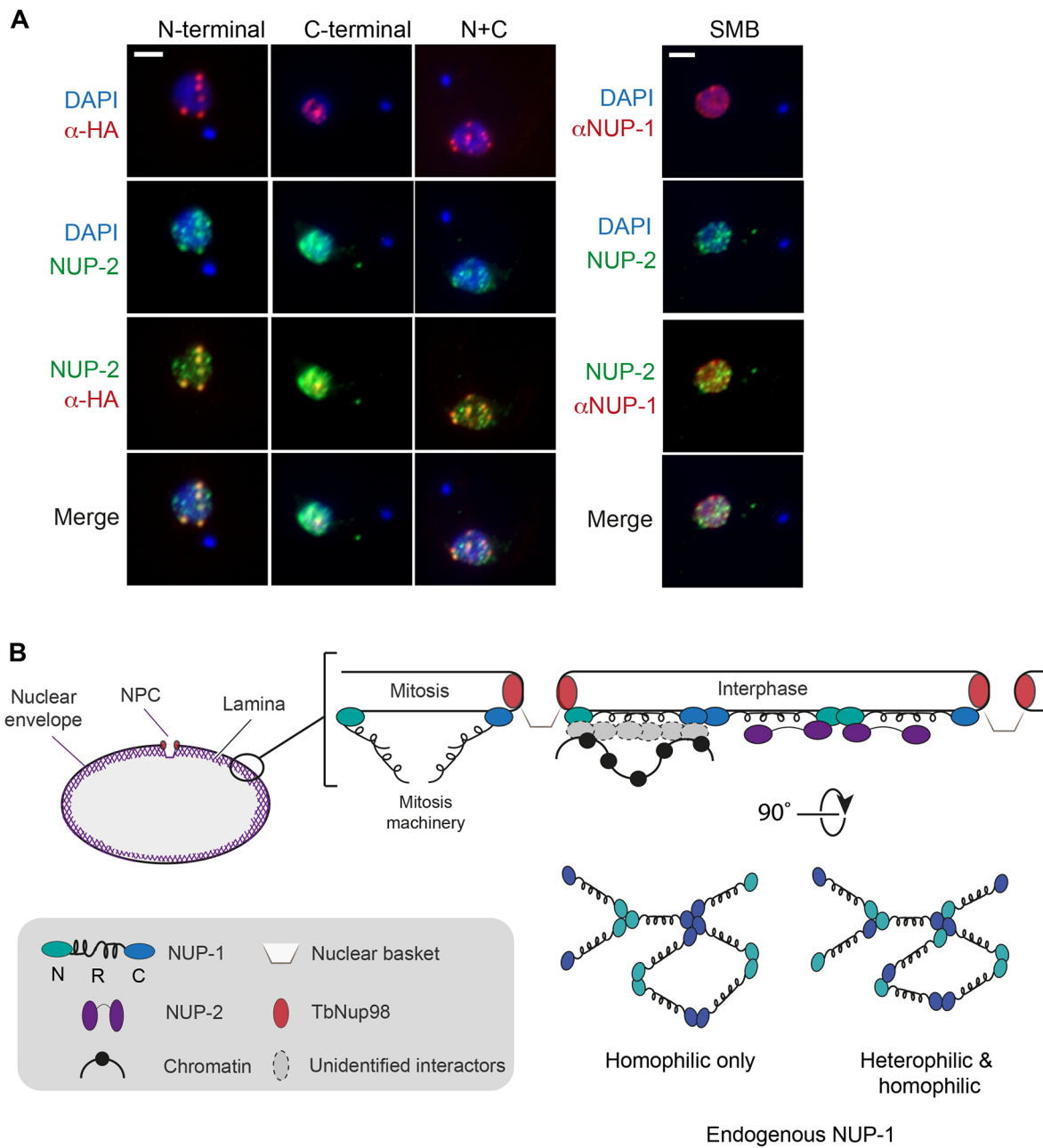
By contrast to A and B lamins, NUP-1 is an elongated protein, spanning a significant fraction of the nuclear volume, is highly flexible with distinct domains targeted differentially and likely presenting intra- and inter-molecular interactions via both termini (Fig. 7B). Whether NUP-1 oligomerizes via the extensive coiled-coil repeat region is unknown, but this is a common property of coiled-coil proteins (Lupas and Gruber, 2005). The observation that telomeric regions have restricted contacts with other genomic regions in trypanosomes is also consistent with the positioning of NUP-1 and the wild-type termini. Partitioning is a common mechanism for nuclear

subdomains, and significantly allows exchange of VSG expression sites from heterochromatin NUP-1 into the expression site body. If these structures resemble membrane-less condensates due to self-assembly remains to be demonstrated.

NUP-1 localizes at the nuclear periphery during interphase with the repeat region migrating into the nuclear interior during mitosis, potentially contributing to chromosome segregation and engaging with the mitotic machinery – the repeat has some sequence similarity to SMC proteins and an interaction between NUP-1 and KKIPI, a kinetochore protein, has been demonstrated (D'Archivio and Wickstead, 2017). SMC proteins, which make multiple contacts with DNA at telomeres, centromeres and chromosome arms are present in trypanosomes (Gluenz et al., 2008; Bessat and Ersfeld, 2009). Furthermore, both the trypanosome-type kinetochore and NUP-1 and NUP-2 are present throughout the Kinetoplastida, but not in relatives of the lineage, e.g. *Euglena gracilis*, which is entirely consistent with a functional connection between NUP-1 and the kinetochore. Significantly, participation of NUP-1 in chromosome segregation may also explain how the very large number of chromosomes are segregated in the trypanosome nucleus in the absence of a highly complex microtubule array within the spindle (Ersfeld and Gull, 1997; Ogbadoyi et al., 2000).

Both N- and C-terminal domains of NUP-1 assemble as circular assemblies that mature into larger structures capable of recruiting endogenous NUP-1. These structures come to lie close to the NE in most cases, and present a somewhat homogenous structure. In mammalian cells, similar structures have been reported to form inside the nucleus as a response to specific mutations in lamin A (Yang et al., 2013; Muchir et al., 2004; Hübner et al., 2006). Moreover, although both N-terminal and C-terminal domains contain coiled-coil regions, they exhibit distinct properties, with the C-terminal having less-efficient self-association and displaying a nucleoplasmic phenotype, suggesting that a fraction of the C-terminal domain of NUP-1 may shift into the nucleoplasm. Two phenotypes can be seen with the C-terminal mutant (circular assemblies and nucleoplasmic form) and further data are required to fully understand the reason for the presence of both phenotypes, as well as for the slower migrating form of the HA-tagged protein. Interestingly, a nucleoplasmic subfraction of lamin A has been implicated in regulating proliferation, differentiation, cell cycle progression and interaction with euchromatin regions (Naetar et al., 2017; Vidak et al., 2018; Gesson et al., 2014) and whether a nucleoplasmic fraction of NUP-1 may exist and have similar functions is still unknown. The NUP-1 N-terminus contains a NUP-2-binding site, but little overlap was detected between NUP-2 and endogenous NUP-1 coiled-coil repeats and therefore these two proteins potentially form separate meshworks connected by hubs within the NUP-1 N-terminal domain.

Proteomics revealed several interactions between NUP-1 and additional nuclear factors. Two mRNA-binding proteins, Tb927.6.5010 and RBP10, were differentially expressed upon induction of assemblies (Table 1). Tb927.6.5010 is a potential post-transcriptional repressor (Lueong et al., 2016; Erben et al., 2014) whereas RBP10 is a regulator of developmental expression and promotes progression from the procyclic form to bloodstream form. RBP10 binds to a 3'-UTR motif in procyclic-specific mRNAs, targeting them for translational repression and degradation (Mugo and Clayton, 2017). Lamins also interact with RNA-binding (Siddam et al., 2018), RNA-processing (e.g. splicing machinery) and RNA transport proteins (Zahr and Jaalouk, 2018; Depreux et al., 2015). An example is the RNA-binding protein Celf1, which participates in a cascade involving kinases required for normal phosphorylation of lamin A/C (Siddam et al., 2018). With RBP10



**Fig. 7. NUP-2 interacts with the N-terminal domain of NUP-1.** (A) Parental SMB cells expressing NUP-2::TY1 were co-stained with anti-TY1 antibodies (green) and anti-NUP-1 repeat antibodies (red). Cells expressing NUP-2::TY1 and domain constructs were co-stained with anti-HA (red) and anti-TY1 (green) antibodies. NUP-2::TY1 colocalizes with N-term::HA and N+C::HA. DAPI was used to visualize DNA (blue). Central z-stacks are shown. Scale bars: 2  $\mu$ m. (B) Model for interactions between NE proteins within the trypanosoma lamina. NUP-1 may form interactions that are head-to-head, tail-to-tail (homophilic interactions) and head-to-tail (heterophilic interactions). NUP-1, via the N-terminal domain, interacts with NUP-2, while both of the NUP-1 terminal domains contact TbNup98 in the NPC. NUP-1 regulates expression of VSG and procyclin genes, although the exact mechanism is still unknown. During mitosis, the NUP-1 repeats (R) localize in the nucleoplasm, suggesting participation/interaction with mitotic machinery. See text for further details.

also capable of recruiting kinases (Mugo and Clayton, 2017) it is likely that an analogous mechanism operates in *T. brucei*.

Levels of TbNup98, an FG nucleoporin and component of the trypanosome NPC (Obado et al., 2016; DeGrasse et al., 2009), were also altered in the NUP-1 domain-expressing cells. Importantly, TbNup98 has an established physical interaction with NUP-1 and NUP-2 by co-immunoprecipitation (Obado et al., 2016), which is fully consistent with the data here. Importantly, there is an increase of TbNup98 in these cells expressing domain constructs, suggesting a compensatory mechanism for sequestration by NUP-1 domains.

Significantly, not all the FG nucleoporins (FG nups) are essential for transport (Strawn et al., 2004), suggesting a role in other NE activities for these NPC components, some of which have already been described to influence mitotic chromosome dynamics and spindle assembly (Wu et al., 2016). For TbNup98, a role in mitosis and/or cytokinesis is possible as those activities are impaired after silencing. Interestingly, KKIPI co-purifies with NUP-1 and some components of the NPC (D'Archivio and Wickstead, 2017), including TbNup92, which interacts with spindle poles during mitosis and with centromeres, contributing to the distribution

of chromosomes during cell division (Holden et al., 2014). Silencing TbNup98 led to NUP-1 clustering and loss of the NE localization, and suggests that TbNup98 is a component of the NPC-mediated anchoring mechanism. Moreover, the abnormal ploidy and nuclear morphology with a failure to complete mitosis in TbNup98-silenced cells is consistent with a role in anchoring NUP-1 and consequent disruption of chromosome segregation. The influence of NPC components on mitosis in other eukaryotes has been already described, proving that the nucleoporins are essential for maintaining the associations of the kinetochores to microtubules, and for promoting spindle assembly and mitotic progression (Ibarra and Hetzer, 2015; Chatel and Fahrenkrog, 2011). Significantly, despite divergent sequence and origins of many components between trypanosomes and metazoan organisms, these comparisons suggest a convergence and retention of overall mechanistic similarity.

In summary, we propose a hub-and-spoke model for NUP-1 assembly (Fig. 7B) within the trypanosoma lamina. As NUP-1 termini can oligomerize, interactions may be occurring in a head-to-head, tail-to-tail or head-to-tail manner through co-occurring homophilic and heterophilic interactions. Furthermore, as terminal domains can recruit the repeats region, a sliding mechanism similar to that reported for lamin A filaments (Makarov et al., 2019) between NUP-1 molecules may be possible. Moreover, in the interaction with NUP-2, the N-terminal domain constitutes the main anchor point, providing additional stability. Additionally, both NUP-1 termini contact nucleoporin TbNup98 in the NPC, with the possibility that other components of the NPC can be contacted by NUP-1. During cell division, the NUP-1  $\alpha$ -helical coiled coil repeats localize to the nucleoplasm, suggesting (1) re-location from the NE and (2) participation/interaction with mitotic machinery. These will require further examination to fully understand the potential role of this trypanosoma lamin in mitosis, a case of closed cell division. Importantly, NUP-1 previously showed participation in the regulation of VSG and procyclin genes, pathogenesis-related genes (DuBois et al., 2012; Maishman et al., 2016), although the mechanism and potential partners await discovery.

## MATERIALS AND METHODS

### Cell culture

Bloodstream form *Trypanosoma brucei brucei* Lister 427 were cultured as previously described (Hirumi and Hirumi, 1989) in HMI-9 medium. Single marker bloodstream form (SMB) and 2T1 bloodstream form (Lister 427, MITat1.2, clone 221a) cells were used for expression of tetracycline-inducible systems, pDEX (Kelly et al., 2007) and RNA interference (RNAi) (Alibu et al., 2005; Alsford and Horn, 2008), respectively. 2T1 cells were maintained in medium containing phleomycin and puromycin (1  $\mu$ g/ml and 0.5  $\mu$ g/ml, respectively). When antibiotic selection was required, drugs were used at the following concentrations: phleomycin 2.5  $\mu$ g/ml, hygromycin 5  $\mu$ g/ml and blasticidin 5  $\mu$ g/ml.

### Recombinant DNA manipulations

Different regions of the NUP-1 coding sequence were HA-tagged in the pDEX-577G vector, a tetracycline-inducible system (Kelly et al., 2007). A modified version of pDEX-577G was used changing the GFP-tag for HA. The inserts were introduced into BamHI and HindIII sites. The regions of NUP-1 used to build the overexpression variants are shown in Fig. S8. The constructs were linearized with NotI and used for transfection of SMB cells.

### Transfection of bloodstream form cells

Approx. 10  $\mu$ g DNA was used for every  $2 \times 10^7$  cells transfected. Usually,  $1.5\text{--}3 \times 10^7$  cells were electroporated in either Cytomix (Ngô et al., 1998; Burkard et al., 2007) or Tb-BSF buffer (Schumann-Burkard et al., 2011) using an Amaxa Nucleofactor II, program X-001 (Lonza Bioscience).

Positive clones were assayed for correct insertion and expression of desired protein by immunofluorescence and immunoblot.

### Western blotting

$5 \times 10^6$  cells were resolved by 4–12% SDS-PAGE (Invitrogen). Proteins were transferred to a PDVF membrane (Millipore). An anti-HA mouse antibody (mouse, Santa Cruz Biotechnology 7392) was used at 1:3000. Detection with secondary anti-mouse IgG peroxidase (Sigma A9044) was performed at a dilution of 1:8000. Visualization was made by chemiluminescence with ECL-detection reagents (GE Amersham RPN 2106). Images were captured using X-ray film (GE Amersham 28906837).

### Proliferation analysis

Cell cultures were adjusted to  $10^5$  cells/ml. If required, cells were induced with tetracycline (Tet) in the culture medium. Cell numbers were determined using a Z1 Coulter counter every 24 h and diluted to  $10^5$  cells/ml. All determinations were performed using triplicate cultures.

### In situ tagging

The pMOTag43M vector system (Oberholzer et al., 2006) was used to introduce a C-terminal *in situ* myc-epitope to TbNup98. The pPOT system (Dean et al., 2015) was used to introduce a N-terminal *in situ* TY1-epitope to Nup-2. The primer sequences are: TbNup98F, 5'-TGGGAATGCTTCAG-CAAGTGGTGAAGAACAATGCTCCACGGAATCCCTTCTCATT-TGGTGCCTTCTTGGGAATGCTGGTACCGGGCCCCCTCGAG-3'; TbNup98R, 5'-ACTAAAGAAGGGTAGAAAACAAAGAAAACACCA-AATAAGGTACCTGACGCAGCGCAACACCACGTCGACTTGCTG-CGGCCGCTCTAGAACTAGTGGAT-3'; Nup2F, 5'-CATTGTTGGGG-TCTCCGTGTCTACACGTCCTTACTCCAGGTGAAGTGAGTGACG-GGAAAGAAGAAAGGGGAAGGAAAACGTATAATGCAGACCTGC-TGC-3'; Nup2R, 5'-CACTGTGAAATGCACGCACTGCTTCCACCACG-CGTTCCCTCCGAGTTTCTGGCATTGCGCTTTCATTGCCCGCAG-CGATCATACTACCCGATCCTGATCC-3'. Linear PCR products were purified and sterilized by ethanol precipitation and use for transfection.

### Immunofluorescence

For microscopy, cells were prepared for microscopy as previously described (Field et al., 2004). Briefly, cells were fixed with 3% paraformaldehyde (v/v) for 15 min at room temperature, washed and allowed to settle onto poly-L-lysine coated slides (VWR International) at room temperature. For permeabilization, cells were incubated with 0.2% Triton X-100 (v/v) in PBS for 10 min and washed three times with excess PBS. Slides were blocked in 20% FBS (Gibco) in PBS for at least 1 h. Cells were incubated with primary and secondary antibodies, successively with washes in excess PBS after antibodies incubations. Slides were mounted with mounting medium plus DAPI (Vectashield Labs). Primary antibodies were used at the following concentrations: anti-HA (1:1000; mouse Santa Cruz Biotechnology sc-7392 or rat Roche 11867423001); anti-Myc (1:400; monoclonal Millipore M4439), anti-TY1 (1:1000; monoclonal mouse Imprint SAB4800032); polyclonal rabbit anti-NUP-1 repeats (1:750; DuBois et al., 2012). Secondary antibodies were goat anti-mouse-IgG Alexa Fluor 488, goat anti-rabbit-IgG Alexa Fluor 568 and goat anti-rat-IgG Alexa Fluor 568 (Invitrogen, A11001, A11011, A11077, respectively) and were used at 1:1000. Confocal microscopy was carried out on a Zeiss microscope and images captured and deconvolved using Zen (Zeiss). For high-resolution microscopy, a Leica System microscope was used and images captured and deconvolved with LAS X software. Image analysis/preparation was made with the OMERO platform (Allan et al., 2012).

### Electron microscopy

Samples for electron microscopy were prepared using a modified protocol previously described (Gadelha et al., 2009). NUP-1 variant cells were induced at 1  $\mu$ g/ml of Tet for 24 h.  $2 \times 10^7$  cells were harvested by centrifugation (800 g, 10 min) and then resuspended in 0.5 ml of HMI-9 medium and fixed by the addition of isothermal glutaraldehyde to a final concentration of 2.5%. Cells were gently rocked for 10 min at room

temperature (RT) culture then harvested at 2000 *g* for 2 min at RT and resuspended in 2.5% glutaraldehyde in PBS for another 30 min at RT. The samples were then post-fixed and processed at the University of Dundee Imaging Facility as previously reported (Maishman et al., 2016). Sections of 70 nm resin were used for imaging; images were taken in a JEOL 1200EX microscope using a SIS Megaview III camera running SIS software. Image analysis and preparation was undertaken with the OMERO platform (Allan et al., 2012).

### Immunogold localization

For immunogold labelling,  $2 \times 10^7$  cells of the following lines were used: SMB cells (control cells), cells expressing the N, C-terminal and N+C mutants (Tet-induced for 24 h). Cells were fixed with 4% formaldehyde with 0.1% glutaraldehyde in 0.1 M HEPES (pH 7.2) for 1 h at RT. After washing in HEPES with 20 mM glycine, pellets of cells embedded in 10% gelatin were immersed in 2.3 M sucrose for 24 h at 4°C and frozen by plunging into liquid nitrogen. Cryosections were cut using an EM UC6 ultramicrotome equipped with an EM FC6 cryochamber (Leica). Cryosections were picked up with a drop of 1.15 M sucrose and 1% methylcellulose. Sections were incubated in blocking solution (1% fish skin gelatine, Sigma-Aldrich) in HEPES with 20 mM glycine for 1 h at RT and incubated with anti-HA antibody, either rabbit (Sigma) or rat (Roche), diluted 1:40 in blocking solution for 15 min at RT. Sections were washed (six times, 2 min each) with blocking solution and incubated with protein A conjugated to 5 nm gold nanoparticles (UMC, Utrecht) diluted in blocking solution 1:50 for 45 min. Samples were washed in HEPES (six times, 2 min each) and dH<sub>2</sub>O, contrasted and embedded in 1.8% methylcellulose and 0.3% uranyl acetate. Samples were observed with a JEOL 1010 transmission electron microscope (TEM) operating at an accelerating voltage 80 kV and equipped with a MegaView III CCD camera (SIS). Refer to Fig. S4 for negative controls.

### Fluorescence *in situ* hybridization

Telomeres were detected using the PNA FISH kit (DAKO K5326) following the manufacturer's instructions. The probe for telomeres is coupled to Cy3. For combined immunofluorescence analysis and the Telomere PNA kit, cells were prepared for immunofluorescence analysis first following the protocol mentioned above. Briefly, after washing the secondary antibody, cells on the slides were fixed with 3.7% formaldehyde during 1 h at room temperature. Slides were then washed twice in TBS, immersed in pre-treatment solution and washed twice again. Slides were immersed in cold (−20°C) ethanol series (70%, 85% and 95%) and then dried. Telomere PNA probe (Cy3) was applied to the slides, moved into a pre-heated incubator at 80°C for 5 min and then placed in the dark at RT for 2–4 h. Slides were rinsed, washed and immersed again in the same cold ethanol series as previously. After drying the slides, DAPI were mounted with Vectashield mounting medium plus DAPI (Vectashield Labs). Confocal microscopy was carried out on a Zeiss microscope (Axiovert 200 M). Images were taken and deconvolved with Zen software (Zeiss). Image processing was performed with the OMERO platform (Allan et al., 2012).

### Proteomics

*T. brucei* bloodstream form cells expressing NUP-1 variants were cultured with 1.0 µg/ml of Tet for 24 h. SMB cells and uninduced cells were used as controls. Cells were washed with PBS containing Complete protease inhibitors (Roche), extracted with 1 × NuPAGE sample buffer and sonicated. Lysates containing  $10^7$  cells were fractionated on a NuPAGE Bis-Tris 4–12% gradient polyacrylamide gel (Thermo Scientific) under reducing conditions at 100 V for 10 min. The migration portion was contained in one slice that was subjected to tryptic digestion and reductive alkylation. Liquid chromatography mass spectrometry (LC-MS2) was performed in-house at the University of Dundee, UK. Samples were analysed on a Dionex UltiMate 3000 RSLCnano System coupled to a Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific). Protein mass spectra were analysed using MaxQuant version 1.6.1.0 (Cox and Mann, 2008) searching the predicted *T. brucei brucei* TREU927 proteome (Aslett et al., 2010). Ratios were calculated from label-free quantification intensities

(NUP-1 variant cell lines versus the uninduced control cells) using only peptides that could be uniquely mapped to a given protein. *P*-values were calculated applying *t*-test-based statistics using Perseus (Tyanova et al., 2016) and the  $-\log P$  and FDR (0.01,  $s_0=2$ ) were calculated. Experiments were conducted in triplicate. Proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino et al., 2016) with the dataset identifier PXD019978. For the selection of differentially expressed proteins (those having abundance shifts after the overexpression of NUP-1 variants), we considered the following criteria: proteins containing at least 2 unique peptides, proteins with a statistical difference with respect to control cells  $\pm 0.3$  and proteins with a statistical significance ( $\log P$ )  $> 1.5$ . For VSG221 quantification the data was re-analysed with the *T. brucei* Lister 427 as search database (Aslett et al., 2010). The repeat region of NUP-1 (absent from the ectopic constructs) was used to distinguish endogenous and ectopic NUP-1.

### TbNup98 silencing

RNAi (Redmond et al., 2003, <https://dag.compbio.dundee.ac.uk/RNAi/>) was used to design primers for TbNup98 RNAi. The sequences for primers were: Nup98F, 5'-AAGCTTGGGCCCCCGGGATTCTTTACGCC-CACCTCG-3' and Nup98R, 5'-AAGCTTTCTAGAGGATCCCTAT-CATCTGGGACCCACGC3'. PCR products were cloned into the pRpA<sup>SL</sup> plasmid (Alsford and Horn, 2008) into sites XmaI/BamHI and XbaI/ApaI. The construct generated was linearized with AscI and used for electroporation. Positive clones were assessed by qPCR.

### Flow cytometry analysis of DNA content

Cells expressing NUP-1 mutants were induced with 1 µg/ml Tet for 24 h. Cells were harvested, resuspended in 1% FBS in PBS and transferred to FACS tubes (Scientific Lab Supplies). Cells were fixed in ice cold 90% methanol for 30 min, washed twice in 1% FBS in PBS and finally resuspended in Staining Buffer (50 mg/ml propidium iodide, 50 mg/ml RNase A in 1% FBS in PBS). Samples were covered from light for 20 min. Samples were analysed for DNA content using a FACS Canto flow cytometer (Becton Dickinson) and DIVA acquisition software. Propidium iodide fluorescence was detected using 488 nm excitation and emission was detected at 585 nm $\pm$ 40 nm. Flow cytometry profiles for 10,000 propidium iodide-labelled cells post induction were obtained. Analysis of data and generation of histograms were performed in FlowJo version 10.6.2.

### RNA-seq analysis

Cells bearing the N+C construct were used for the transcriptomics assay. Cells were induced using 1 µg/ml of Tet (during 24 h) and SMB parental cells were used as control.  $10^8$  cells were used for isolation of total RNA using a Macherey-Nagel NucleoSpin RNA kit (740955) as per manufacturer's instructions and eluted in high purity RNase-free water. The samples were sequenced in triplicates by Global Genomic Services (GBI). Sequencing resulted in paired-ended reads  $2 \times 100$  bp, 12 million reads per sample. RNA-seq reads were mapped to the reference genome *T. brucei* TREU927, release 44, from TriTrypDB database. For VSG and procyclin genes, *T. brucei* 427 genome was used (Aslett et al., 2010). Mapping was done using STAR 2.6.0c aligner (Dobin et al., 2013); 70% of reads were mapped uniquely to the genome. Read counts per gene were found in the same STAR run, using TriTrypDB annotations in a GFF file. Data analysis was done in R environment. RNA-seq data are available in the NCBI BioSample database (<http://www.ncbi.nlm.nih.gov/biosample/>) under accession number PRJNA642306.

### Acknowledgements

We thank Ning (Pax) Zhang and Alex Makarov for fruitful discussions. We acknowledge the core facility LEM supported by MEYS CR (LM2018129 Czech-Biomed).

### Competing interests

The authors declare no competing or financial interests.

**Author contributions**

Conceptualization: M.C.F.; Methodology: N.E.P.-M., L.K., J.H., M.V., M.Z., M.C.F.; Software: M.Z.; Formal analysis: N.P., M.Z., M.C.F.; Investigation: N.E.P.-M., L.K., J.H., M.V., M.Z.; Resources: L.K., J.H., J.L., M.C.F.; Data curation: J.L.; Writing - original draft: N.E.P.-M., M.C.F.; Writing - review & editing: M.Z., M.C.F.; Supervision: J.L., M.C.F.; Project administration: J.L., M.C.F.; Funding acquisition: M.C.F.

**Funding**

This work was funded by the Medical Research Council (MR/N010558/1 to M.C.F.) and benefited from a Wellcome Trust Investigator award (204697/Z/16/Z to M.C.F.), the Czech Ministry of Education, Youth and Science (MEYS) for ERD Funds (project OPVVV/0000759 to J.L.) and a Czech Biolmaging grant (LM2015062 to M.V.). Deposited in PMC for immediate release.

**Data availability**

RNA-seq data are available in the NCBI BioSample database under accession number PRJNA642306. Proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD019978.

**References**

- Aaronson, R. P. and Blobel, G. (1975). Isolation of nuclear pore complexes in association with a lamina. *Proc. Natl. Acad. Sci. USA* **72**, 1007-1011. doi:10.1073/pnas.72.3.1007
- Ahn, J., Jo, I., Kang, S.-M., Hong, S., Kim, S., Jeong, S., Kim, Y.-H., Park, B.-J. and Ha, N.-C. (2019). Structural basis for lamin assembly at the molecular level. *Nat. Commun.* **10**, 3757. doi:10.1038/s41467-019-11684-x
- Alibu, V. P., Storm, L., Haile, S., Clayton, C. and Horn, D. (2005). A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **139**, 75-82. doi:10.1016/j.molbiopara.2004.10.002
- Allan, C., Burel, J.-M., Moore, J., Blackburn, C., Linkert, M., Loynton, S., Macdonald, D., Moore, W. J., Neves, C., Patterson, A. et al. (2012). OMERO: flexible, model-driven data management for experimental biology. *Nat. Methods* **9**, 245-253. doi:10.1038/nmeth.1896
- Alsford, S. and Horn, D. (2008). Single-locus targeting constructs for reliable regulated RNAi and transgene expression in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **161**, 76-79. doi:10.1016/j.molbiopara.2008.05.006
- Aslett, M., Aurrecochea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X. et al. (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res.* **38**, D457-D462. doi:10.1093/nar/gkp851
- Benz, C., Dondelinger, F., McKean, P. G. and Urbaniak, M. D. (2017). Cell cycle synchronisation of *Trypanosoma brucei* by centrifugal counter-flow elutriation reveals the timing of nuclear and kinetoplast DNA replication. *Sci. Rep.* **7**, 17599. doi:10.1038/s41598-017-17779-z
- Bessat, M. and Ersfeld, K. (2009). Functional characterization of cohesin SMC3 and separate and their roles in the segregation of large and minichromosomes in *Trypanosoma brucei*. *Mol. Microbiol.* **71**, 1371-1385. doi:10.1111/j.1365-2958.2009.06611.x
- Burkard, G., Fragoso, C. M. and Roditi, I. (2007). Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **153**, 220-223. doi:10.1016/j.molbiopara.2007.02.008
- Chatel, G. and Fahrenkrog, B. (2011). Nucleoporins: leaving the nuclear pore complex for a successful mitosis. *Cell. Signal.* **23**, 1555-1562. doi:10.1016/j.cellsig.2011.05.023
- Constantinescu, D., Gray, H. L., Sammak, P. J., Schatten, G. P. and Csoka, A. B. (2006). Lamin A/C expression is a marker of mouse and human embryonic stem cell differentiation. *Stem Cells* **24**, 177-185. doi:10.1634/stemcells.2004-0159
- Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367-1372. doi:10.1038/nbt.1511
- D'Archivio, S. and Wickstead, B. (2017). Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J. Cell Biol.* **216**, 379-391. doi:10.1083/jcb.201608043
- Dean, S., Sunter, J., Wheeler, R. J., Hodgkinson, I., Gluenz, E. and Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol.* **5**, 140197. doi:10.1098/rsob.140197
- Dean, S., Sunter, J. D. and Wheeler, R. J. (2017). TrypTag.org: a Trypanosome genome-wide protein localisation resource. *Trends Parasitol.* **33**, 80-82. doi:10.1016/j.pt.2016.10.009
- Dechat, T., Pfeleghaar, K., Sengupta, K., Shimi, T., Shumaker, D. K., Solimando, L. and Goldman, R. D. (2008). Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev.* **22**, 832-853. doi:10.1101/gad.1652708
- Dechat, T., Adam, S. A. and Goldman, R. D. (2009). Nuclear lamins and chromatin: when structure meets function. *Adv. Enzyme Regul.* **49**, 157-166. doi:10.1016/j.advenzreg.2008.12.003
- DeGrasse, J. A., DuBois, K. N., Devos, D., Siegel, T. N., Sali, A., Field, M. C., Rout, M. P. and Chait, B. T. (2009). Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Mol. Cell. Proteomics* **8**, 2119-2130. doi:10.1074/mcp.M900038-MCP200
- Depreux, F. F., Puckelwartz, M. J., Augustynowicz, A., Wolfgeher, D., Labno, C. M., Pierre-Louis, D., Cicka, D., Kron, S. J., Holaska, J. and McNally, E. M. (2015). Disruption of the lamin A and matrin-3 interaction by myopathic LMNA mutations. *Hum. Mol. Genet.* **24**, 4284-4295. doi:10.1093/hmg/ddv160
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21. doi:10.1093/bioinformatics/bts635
- DuBois, K. N., Alsford, S., Holden, J. M., Buisson, J., Swiderski, M., Bart, J.-M., Ratushny, A. V., Wan, Y., Bastin, P., Barry, J. D. et al. (2012). NUP-1 is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol.* **10**, e1001287. doi:10.1371/journal.pbio.1001287
- Erben, E. D., Fadda, A., Lueong, S., Hoheisel, J. D. and Clayton, C. (2014). A genome-wide tethering screen reveals novel potential post-transcriptional regulators in *Trypanosoma brucei*. *PLoS Pathog.* **10**, e1004178. doi:10.1371/journal.ppat.1004178
- Ersfeld, K. and Gull, K. (1997). Partitioning of large and minichromosomes in *Trypanosoma brucei*. *Science* **276**, 611-614. doi:10.1126/science.276.5312.611
- Faria, J., Glover, L., Hutchinson, S., Boehm, C., Field, M. C. and Horn, D. (2019). Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex. *Nat. Commun.* **10**, 3023. doi:10.1038/s41467-019-10823-8
- Field, M. C., Allen, C. L., Dhir, V., Goulding, D., Hall, B. S., Morgan, G. W., Veazey, P. and Engstler, M. (2004). New approaches to the microscopic imaging of *Trypanosoma brucei*. *Microsc. Microanal.* **10**, 621-636. doi:10.1017/S1431927604040942
- Field, M. C., Horn, D., Alsford, S., Koreny, L. and Rout, M. P. (2012). Telomeres, tethers and trypanosomes. *Nucleus* **3**, 478-486. doi:10.4161/nucl.22167
- Figueiredo, L. M. and Cross, G. A. M. (2010). Nucleosomes are depleted at the VSG expression site transcribed by RNA Polymerase I in African Trypanosomes. *Eukaryot. Cell* **9**, 148-154. doi:10.1128/EC.00282-09
- Florini, F., Naguleswaran, A., Gharib, W. H., Bringaud, F. and Roditi, I. (2019). Unexpected diversity in eukaryotic transcription revealed by the retrotransposon hotspot family of *Trypanosoma brucei*. *Nucleic Acids Res.* **47**, 1725-1739. doi:10.1093/nar/gky1255
- Gadelha, C., Rothery, S., Morphew, M., McIntosh, J. R., Severs, N. J. and Gull, K. (2009). Membrane domains and flagellar pocket boundaries are influenced by the cytoskeleton in African trypanosomes. *Proc. Natl. Acad. Sci. USA* **106**, 17425-17430. doi:10.1073/pnas.0909289106
- Gesson, K., Vidak, S. and Foisner, R. (2014). Lamina-associated polypeptide (LAP)2 $\alpha$  and nucleoplasmic lamins in adult stem cell regulation and disease. *Semin. Cell Dev. Biol.* **29**, 116-124. doi:10.1016/j.semdb.2013.12.009
- Glover, L., Hutchinson, S., Alsford, S. and Horn, D. (2016). VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proc. Natl. Acad. Sci. USA* **113**, 7225-7230. doi:10.1073/pnas.1600344113
- Gluenz, E., Sharma, R., Carrington, M. and Gull, K. (2008). Functional characterization of cohesin subunit SCC1 in *Trypanosoma brucei* and dissection of mutant phenotypes in two life cycle stages. *Mol. Microbiol.* **69**, 666-680. doi:10.1111/j.1365-2958.2008.06320.x
- Goldberg, M. W. and Allen, T. D. (1996). The nuclear pore complex and lamina: three-dimensional structures and interactions determined by field emission in-lens scanning electron microscopy. *J. Mol. Biol.* **257**, 848-865. doi:10.1006/jmbi.1996.0206
- Goldberg, M. W., Huttenlauch, I., Hutchison, C. J. and Stick, R. (2008). Filaments made from A- and B-type lamins differ in structure and organization. *J. Cell Sci.* **121**, 215-225. doi:10.1242/jcs.022020
- Goos, C., Dejung, M., Janzen, C. J., Butter, F. and Kramer, S. (2017). The nuclear proteome of *Trypanosoma brucei*. *PLoS ONE* **12**, e0181884. doi:10.1371/journal.pone.0181884
- Gruenbaum, Y. and Foisner, R. (2015). Lamins: nuclear intermediate filament proteins with fundamental functions in nuclear mechanics and genome regulation. *Annu. Rev. Biochem.* **84**, 131-164. doi:10.1146/annurev-biochem-060614-034115
- Hirumi, H. and Hirumi, K. (1989). Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J. Parasitol.* **75**, 985-989. doi:10.2307/3282883
- Holden, J. M., Koreny, L., Obado, S., Ratushny, A. V., Chen, W.-M., Chiang, J.-H., Kelly, S., Chait, B. T., Aitchison, J. D., Rout, M. P. et al. (2014). Nuclear pore complex evolution: a trypanosome Mlp analogue functions in chromosomal segregation but lacks transcriptional barrier activity. *Mol. Biol. Cell* **25**, 1421-1436. doi:10.1091/mbc.e13-12-0750
- Hübner, S., Eam, J. E., Hübner, A. and Jans, D. A. (2006). Laminopathy-inducing lamin A mutants can induce redistribution of lamin binding proteins into nuclear aggregates. *Exp. Cell Res.* **312**, 171-183. doi:10.1016/j.yexcr.2005.10.011

- Ibarra, A. and Hetzer, M. W. (2015). Nuclear pore proteins and the control of genome functions. *Genes Dev.* **29**, 337-349. doi:10.1101/gad.256495.114
- Izumi, M., Vaughan, O. A., Hutchison, C. J. and Gilbert, D. M. (2000). Head and/or CaaX domain deletions of lamin proteins disrupt preformed Lamin A and C but not lamin B structure in mammalian cells. *Mol. Biol. Cell* **11**, 4323-4337. doi:10.1091/mbc.11.12.4323
- Kang, S.-M., Yoon, M.-H. and Park, B.-J. (2018). Laminopathies: mutations on single gene and various human genetic diseases. *BMB Rep.* **51**, 327-337. doi:10.5483/BMBRep.2018.51.7.113
- Kelly, S., Reed, J., Kramer, S., Ellis, L., Webb, H., Sunter, J., Salje, J., Marinsek, N., Gull, K., Wickstead, B. et al. (2007). Functional genomics in *Trypanosoma brucei*: a collection of vectors for the expression of tagged proteins from endogenous and ectopic gene loci. *Mol. Biochem. Parasitol.* **154**, 103-109. doi:10.1016/j.molbiopara.2007.03.012
- Kim, Y., Zheng, X. and Zheng, Y. (2019). Role of lamins in 3D genome organization and global gene expression. *Nucleus* **10**, 33-41. doi:10.1080/19491034.2019.1578601
- Koreny, L. and Field, M. C. (2016). Ancient eukaryotic origin and evolutionary plasticity of nuclear lamina. *Genome Biol. Evol.* **8**, 2663-2671. doi:10.1093/gbe/evw087
- Lehner, C. F., Stick, R., Eppenberger, H. M. and Nigg, E. A. (1987). Differential expression of nuclear lamin proteins during chicken development. *J. Cell Biol.* **105**, 577-587. doi:10.1083/jcb.105.1.577
- Liu, J., Rolef Ben-Shahar, T., Riemer, D., Treinin, M., Spann, P., Weber, K., Fire, A. and Gruenbaum, Y. (2000). Essential roles for *Caenorhabditis elegans* lamin gene in nuclear organization, cell cycle progression, and spatial organization of nuclear pore complexes. *Mol. Biol. Cell* **11**, 3937-3947. doi:10.1091/mbc.11.11.3937
- Lueong, S., Merce, C., Fischer, B., Hoheisel, J. D. and Erben, E. D. (2016). Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol. Microbiol.* **100**, 457-471. doi:10.1111/mmi.13328
- Lupas, A. N. and Gruber, M. (2005). The structure of  $\alpha$ -helical coiled coils. *Adv. Protein Chem.* **70**, 37-38. doi:10.1016/S0065-3233(05)70003-6
- Maishman, L., Obado, S. O., Alford, S., Bart, J.-M., Chen, W.-M., Ratushny, A. V., Navarro, M., Horn, D., Aitchison, J. D., Chait, B. T. et al. (2016). Co-dependence between trypanosome nuclear lamina components in nuclear stability and control of gene expression. *Nucleic Acids Res.* **44**, 10554-10570. doi:10.1093/nar/gkw751
- Makarov, A. A., Zou, J., Houston, D. R., Spanos, C., Solovyova, A. S., Cardenal-Peralta, C., Rappsilber, J. and Schirmer, E. C. (2019). Lamin A molecular compression and sliding as mechanisms behind nucleoskeleton elasticity. *Nat. Commun.* **10**, 3056. doi:10.1038/s41467-019-11063-6
- Mounkes, L. C., Kozlov, S. V., Rottman, J. N. and Stewart, C. L. (2005). Expression of an LMNA-N195K variant of A-type lamins results in cardiac conduction defects and death in mice. *Hum. Mol. Genet.* **14**, 2167-2180. doi:10.1093/hmg/ddi221
- Muchir, A., Medioni, J., Laluc, M., Massart, C., Arimura, T., van der Kooi, A. J., Desguerre, I., Mayer, M., Ferrer, X., Briault, S. et al. (2004). Nuclear envelope alterations in fibroblasts from patients with muscular dystrophy, cardiomyopathy, and partial lipodystrophy carrying lamin A/C gene mutations. *Muscle Nerve* **30**, 444-450. doi:10.1002/mus.20122
- Muchir, A., van Engelen, B. G., Lammens, M., Mislow, J. M., McNally, E., Schwartz, K. and Bonne, G. (2003). Nuclear envelope alterations in fibroblasts from LGMD1B patients carrying nonsense Y259X heterozygous or homozygous mutation in lamin A/C gene. *Exp. Cell Res.* **291**, 352-362. doi:10.1016/j.yexcr.2003.07.002
- Mugnier, M. R., Cross, G. A. M. and Papavasiliou, F. N. (2015). The *in vivo* dynamics of antigenic variation in *Trypanosoma brucei*. *Science* **347**, 1470-1473. doi:10.1126/science.aaa4502
- Mugo, E. and Clayton, C. (2017). Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in *Trypanosoma brucei*. *PLoS Pathog.* **13**, e1006560. doi:10.1371/journal.ppat.1006560
- Müller, L. S. M., Cosentino, R. O., Förstner, K. U., Guizzetti, J., Wedel, C., Kaplan, N., Janzen, C. J., Arampatzis, P., Vogel, J., Steinbiss, S. et al. (2018). Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121-125. doi:10.1038/s41586-018-0619-8
- Naetar, N., Ferraioli, S. and Foisner, R. (2017). Lamins in the nuclear interior – life outside the lamina. *J. Cell Sci.* **130**, 2087-2096. doi:10.1242/jcs.203430
- Ngô, H., Tschudi, C., Gull, K. and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA* **95**, 14687-14692. doi:10.1073/pnas.95.25.14687
- Nmezi, B., Xu, J., Fu, R., Armiger, T. J., Rodriguez-Bey, G., Powell, J. S., Ma, H., Sullivan, M., Tu, Y., Chen, N. Y. et al. (2019). Concentric organization of A- and B-type lamins predicts their distinct roles in the spatial organization and stability of the nuclear lamina. *Proc. Natl. Acad. Sci. USA* **116**, 4307-4315. doi:10.1073/pnas.1810070116
- Obado, S. O., Brillantes, M., Uryu, K., Zhang, W., Ketaren, N. E., Chait, B. T., Field, M. C. and Rout, M. P. (2016). Interactome mapping reveals the evolutionary history of the nuclear pore complex. *PLoS Biol.* **14**, e1002365. doi:10.1371/journal.pbio.1002365
- Oberholzer, M., Morand, S., Kunz, S. and Seebeck, T. (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol. Biochem. Parasitol.* **145**, 117-120. doi:10.1016/j.molbiopara.2005.09.002
- Ogbadanyi, E., Ersfeld, K., Robinson, D., Sherwin, T. and Gull, K. (2000). Architecture of the *Trypanosoma brucei* nucleus during interphase and mitosis. *Chromosoma* **108**, 501-513. doi:10.1007/s004120050402
- Pinger, J., Chowdhury, S. and Papavasiliou, F. N. (2017). Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. *Nat. Commun.* **8**, 828. doi:10.1038/s41467-017-00959-w
- Radwanska, M., Vereecke, N., Deleuw, V., Pinto, J. and Magez, S. (2018). Salivarian Trypanosomosis: a review of parasites involved, their global distribution and their interaction with the innate and adaptive mammalian host immune system. *Front. Immunol.* **9**, 2253. doi:10.3389/fimmu.2018.02253
- Redmond, S., Vadivelu, J. and Field, M. C. (2003). RNAi: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **128**, 115-118. doi:10.1016/S0166-6851(03)00045-8
- Rober, R. A., Weber, K. and Osborn, M. (1989). Differential timing of nuclear lamin A/C expression in the various organs of the mouse embryo and the young animal: a developmental study. *Development* **105**, 365-378. doi:10.1242/dev.105.2.365
- Roditi, I., Schwartz, H., Pearson, T. W., Beecroft, R. P., Liu, M. K., Richardson, J. P., Bühring, H. J., Pleiss, J., Bülow, R. and Williams, R. O. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. *J. Cell Biol.* **108**, 737-746. doi:10.1083/jcb.108.2.737
- Saura, A., Iribarren, P. A., Rojas-Barros, D., Bart, J. M., López-Farfán, D., Andrés-León, E., Vidal-Cobo, I., Boehm, C., Alvarez, V. E. et al. (2019). SUMOylated SNF2PH promotes variant surface glycoprotein expression in bloodstream trypanosomes. *EMBO Rep.* **20**, e48029. doi:10.15252/embr.201948029
- Schumann Burkard, G., Jutzi, P. and Roditi, I. (2011). Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Mol. Biochem. Parasitol.* **175**, 91-94. doi:10.1016/j.molbiopara.2010.09.002
- Shimi, T., Pflieger, K., Kojima, S.-i., Pack, C.-G., Solovei, I., Goldman, A. E., Adam, S. A., Shumaker, D. K., Kinjo, M., Cremer, T. et al. (2008). The A- and B-type nuclear lamin networks: microdomains involved in chromatin organization and transcription. *Genes Dev.* **22**, 3409-3421. doi:10.1101/gad.1735208
- Shimi, T., Kittisopikul, M., Tran, J., Goldman, A. E., Adam, S. A., Zheng, Y., Jaqaman, K. and Goldman, R. D. (2015). Structural organization of nuclear lamins A, C, B1 and B2 revealed by superresolution microscopy. *Mol. Biol. Cell* **26**, 4075-4086. doi:10.1091/mbc.E15-07-0461
- Siddam, A. D., Gautier-Courteille, C., Perez-Campos, L., Anand, D., Kakrana, A., Dang, C. A., Legagneux, V., Méreau, A., Viet, J., Gross, J. M. et al. (2018). The RNA binding protein Celf1 post-transcriptionally regulates p27Kip1 and Dnase2b to control fiber cell nuclear degradation in lens development. *PLoS Genet.* **14**, e1007278. doi:10.1371/journal.pgen.1007278
- Simon, D. N. and Wilson, K. L. (2013). Partners and post-translational modifications of nuclear lamins. *Chromosoma* **122**, 13-31. doi:10.1007/s00412-013-0399-8
- Stijlemans, B., Caljon, G., Van Den Abbeele, J., Van Ginderachter, J. A., Magez, S. and De Trez, C. (2016). Immune evasion strategies of *Trypanosoma brucei* within the mammalian host: progression to pathogenicity. *Front. Immunol.* **7**, 233. doi:10.3389/fimmu.2016.00233
- Strawn, L. A., Shen, T., Shulga, N., Goldfarb, D. S. and Wente, S. R. (2004). Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* **6**, 197-206. doi:10.1038/ncb1097
- Stuurman, N., Delbecq, J. P., Callaerts, P. and Aebi, U. (1999). Ectopic overexpression of Drosophila lamin C is stage-specific lethal. *Exp. Cell Res.* **248**, 350-357. doi:10.1006/excr.1999.4396
- Swift, J. and Discher, D. E. (2014). The nuclear lamina is mechano-responsive to ECM elasticity in mature tissue. *J. Cell Sci.* **127**, 3005-3015. doi:10.1242/jcs.149203
- Sylvius, N., Hathaway, A., Boudreau, E., Gupta, P., Labib, S., Bolongo, P. M., Rippstein, P., McBride, H., Bilinska, Z. T. and Tesson, F. (2008). Specific contribution of lamin A and lamin C in the development of laminopathies. *Exp. Cell Res.* **314**, 2362-2375. doi:10.1016/j.yexcr.2008.04.017
- Turgay, Y., Eibauer, M., Goldman, A. E., Shimi, T., Khayat, M., Ben-Harush, K., Dubrovsky-Gaupp, A., Sapro, K. T., Goldman, R. D. and Medalia, O. (2017). The molecular architecture of lamins in somatic cells. *Nature* **543**, 261-264. doi:10.1038/nature21382
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M. and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731-740. doi:10.1038/nmeth.3901
- Verstraeten, V. L. R. M., Broers, J. L. V., Ramaekers, F. C. S. and van Steensel, M. A. M. (2007). The nuclear envelope, a key structure in cellular integrity and

- gene expression. *Curr. Med. Chem.* **14**, 1231-1248. doi:10.2174/092986707780598032
- Vidak, S., Georgiou, K., Fichtinger, P., Naetar, N., Dechat, T. and Foisner, R.** (2018). Nucleoplasmic lamins define growth-regulating functions of lamina-associated polypeptide 2 $\alpha$  in progeria cells. *J. Cell Sci.* **131**, jcs208462. doi:10.1242/jcs.208462
- Vizcaíno, J. A., Csordas, A., del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T. et al.** (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447-D456. doi:10.1093/nar/gkv1145
- Wirtz, E., Leal, S., Ochatt, C. and Cross, G. A. M.** (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89-101. doi:10.1016/S0166-6851(99)00002-X
- Woodward, R. and Gull, K.** (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *J. Cell Sci.* **95**, 49-57. doi:10.1242/jcs.95.1.49
- Wu, Z., Jin, Z., Zhang, X., Shen, N., Wang, J., Zhao, Y. and Mei, L.** (2016). Nup62, associated with spindle microtubule rather than spindle matrix, is involved in chromosome alignment and spindle assembly during mitosis. *Cell Biol. Int.* **40**, 968-975. doi:10.1002/cbin.10633
- Yang, L., Munck, M., Swaminathan, K., Kapinos, L. E., Noegel, A. A. and Neumann, S.** (2013). Mutations in LMNA Modulate the Lamin A - Nesprin-2 interaction and cause LINC complex alterations. *PLoS ONE* **8**, e71850. doi:10.1371/journal.pone.0071850
- Yatskevich, S., Rhodes, J. and Nasmyth, K.** (2019). Organization of chromosomal DNA by SMC complexes. *Annu. Rev. Genet.* **53**, 445-482. doi:10.1146/annurev-genet-112618-043633
- Zahr, H. C. and Jaalouk, D. E.** (2018). Exploring the crosstalk between LMNA and splicing machinery gene mutations in dilated cardiomyopathy. *Front. Genet.* **9**, 231. doi:10.3389/fgene.2018.00231
- Zheng, X., Hu, J., Yue, S., Kristiani, L., Kim, M., Sauria, M., Taylor, J., Kim, Y. and Zheng, Y.** (2018). Lamins organize the global three-dimensional genome from the nuclear periphery. *Mol. Cell* **71**, 802-815.e7. doi:10.1016/j.molcel.2018.05.017



Phyre2

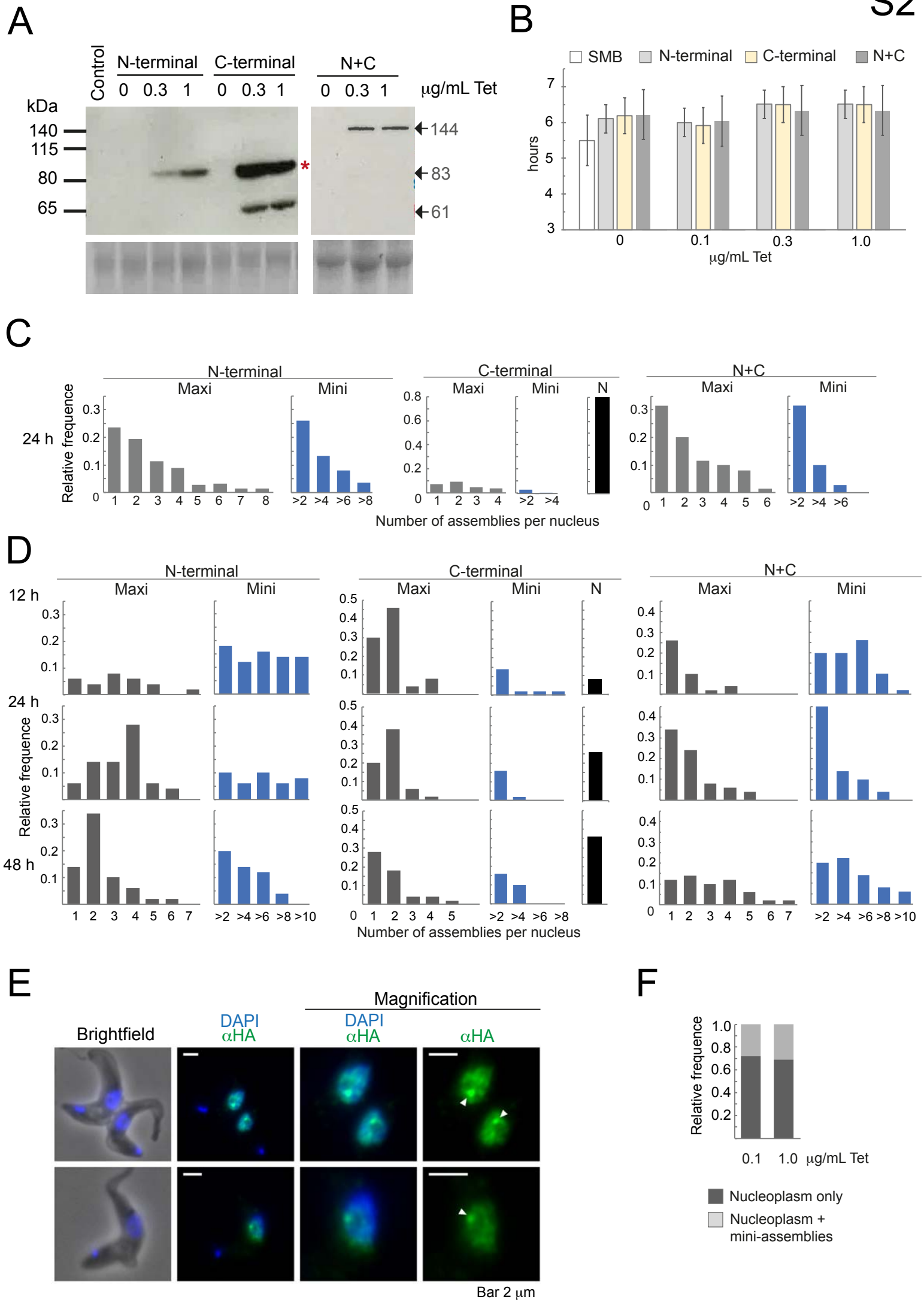
|               |                              |
|---------------|------------------------------|
| Email         | hepadillamejia@dundee.ac.uk  |
| Description   | NUP-1_repeats                |
| Date          | Fri Mar 20 14:07:09 GMT 2020 |
| Unique Job ID | 428f1c12be9033eb             |

Detailed template information

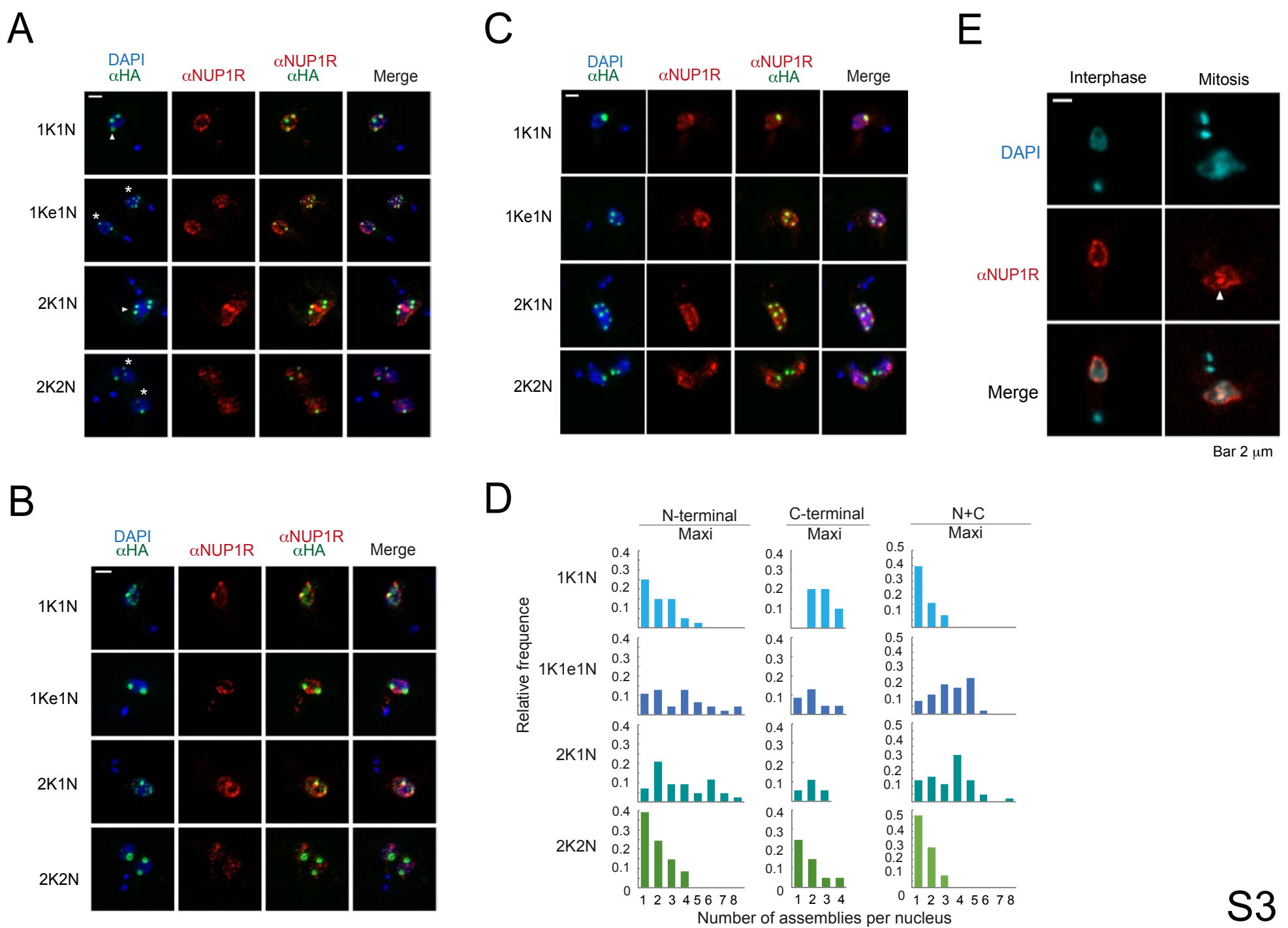
| #  | Template                | Alignment Coverage | 3D Model | Confidence | % i.d. | Template Information  |
|----|-------------------------|--------------------|----------|------------|--------|---|
| 1  | <a href="#">c6h2xA_</a> | Alignment          |          | 98.6       | 11     | <b>PDB header:</b> dna binding protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> chromosome partition protein mukb,chromosome partition<br><b>PDBTitle:</b> mukb coiled-coil elbow from e. coli  |
| 2  | <a href="#">c1c1gA_</a> | Alignment          |          | 98.5       | 14     | <b>PDB header:</b> contractile protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> tropomyosin;<br><b>PDBTitle:</b> crystal structure of tropomyosin at 7 angstroms resolution in the2 spermine-induced crystal form                             |
| 3  | <a href="#">c5m2B_</a>  | Alignment          |          | 98.4       | 10     | <b>PDB header:</b> hydrolase<br><b>Chain:</b> B; <b>PDB Molecule:</b> tyrosine-protein phosphatase non-receptor type 23;<br><b>PDBTitle:</b> crystal structure of hd-ptp phosphatase  |
| 4  | <a href="#">c5xg2A_</a> | Alignment          |          | 98.1       | 17     | <b>PDB header:</b> dna binding protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> chromosome partition protein smc;<br><b>PDBTitle:</b> crystal structure of a coiled-coil segment (residues 345-468 and 694-2 814) of pyrococcus yayanosii smc |
| 5  | <a href="#">c3wuqA_</a> | Alignment          |          | 97.7       | 8      | <b>PDB header:</b> motor protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> cytoplasmic dynein 1 heavy chain 1;<br><b>PDBTitle:</b> structure of the entire stalk region of the dynein motor domain   |
| 6  | <a href="#">c1ciiA_</a> | Alignment          |          | 97.6       | 10     | <b>PDB header:</b> transmembrane protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> colicin ia;<br><b>PDBTitle:</b> colicin ia  |
| 7  | <a href="#">c4cgkA_</a> | Alignment          |          | 97.5       | 11     | <b>PDB header:</b> cell cycle<br><b>Chain:</b> A; <b>PDB Molecule:</b> secreted 45 kda protein;<br><b>PDBTitle:</b> crystal structure of the essential protein pcsb from streptococcus2 pneumoniae  |
| 8  | <a href="#">c5nnvD_</a> | Alignment          |          | 97.5       | 14     | <b>PDB header:</b> cell cycle<br><b>Chain:</b> D; <b>PDB Molecule:</b> chromosome partition protein smc,chromosome partition<br><b>PDBTitle:</b> structure of a bacillus subtilis smc coiled coil middle fragment                             |
| 9  | <a href="#">c2oevA_</a> | Alignment          |          | 97.4       | 11     | <b>PDB header:</b> protein transport<br><b>Chain:</b> A; <b>PDB Molecule:</b> programmed cell death 6-interacting protein;<br><b>PDBTitle:</b> crystal structure of alix/aip1   |
| 10 | <a href="#">c5j1iA_</a> | Alignment          |          | 97.4       | 11     | <b>PDB header:</b> structural protein<br><b>Chain:</b> A; <b>PDB Molecule:</b> plectin;<br><b>PDBTitle:</b> structure of the spectrin repeats 7, 8, and 9 of the plakin domain of2 plectin  |
| 11 | <a href="#">c6gapB_</a> | Alignment          |          | 97.2       | 16     | <b>PDB header:</b> viral protein<br><b>Chain:</b> B; <b>PDB Molecule:</b> outer capsid protein sigma-1;<br><b>PDBTitle:</b> crystal structure of the t3d reovirus sigma1 coiled coil tail and body  |

**Fig. S1.** Structural modelling of the NUP-1 repeat. The sequence corresponding to a single repeat was submitted to the Phyre2 server (<http://www.sbg.bio.ic.ac.uk/phyre2>) for modeling (Kelley et al., 2015). The intensive algorithm was selected.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858. doi:10.1038/nprot.2015.053

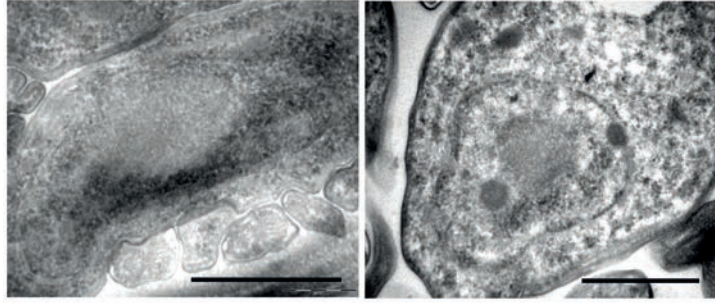


**Fig. S2.** Characterization of NUP-1 overexpression mutants. (A) Detection of NUP-1 mutants by immunoblotting using an anti-HA antibody. Bands for N-terminal (83 kDa), C-terminal (63 kDa) and N+C (144 kDa) are shown. An extra band for C-terminal mutant is detected of ~80 kDa (red star). Ponceau staining of membranes is shown as loading control (B) Growth curves were followed across six days to assess the growth rate and the doubling times were calculated. SMB cells (parental line) and cells expressing the N-terminal, C-terminal and N+C mutants were used. Growth curves were assessed at 0, 0.1, 0.3 and 1.0  $\mu\text{g/ml}$  of tetracycline for three different experiments. Cells were counted every 24 hours. The average doubling time and standard deviation were calculated. Error bar represents standard deviation. A gradual increment in the doubling time is observed at doses of 0.3 and 1.0  $\mu\text{g/ml}$  of tetracycline. (C) Bar chart representing the number of maxi and mini assemblies per nucleus. Overexpression of NUP-1 mutants was induced during 24 h with 0.1  $\mu\text{g/ml}$  Tet. Data for N-terminal, C-terminal and N+C fusion mutants is shown. The relative frequency of number of assemblies per nucleus is shown (n=50). For C-terminal mutant, the frequency of nucleoplasmic pattern (N) is also reported. (D) Effect of time in the formation of NUP-1 mutant assemblies. The relative frequency of number of assemblies per nucleus is shown (n=50). Overexpression of NUP-1 mutants was induced during 12 h, 24 h and 48 h with 1.0  $\mu\text{g/ml}$  Tet. Data for N-terminal, C-terminal and N+C fusion mutants is shown. For C-terminal mutant, the frequency of nucleoplasmic pattern (N) is also reported. (E) Immunofluorescence analysis in cells bearing the C-terminal mutant after 24 hours of induction with 1.0  $\mu\text{g/ml}$  of tetracycline. NUP-1 C-terminal mutants were probed with an anti-HA antibody (green). Mini-assemblies (white arrowhead) occurring at the same time with the nucleoplasmic pattern are shown. DAPI was used to visualize DNA. Scale bar 2  $\mu\text{m}$ . (F) Bar chart representing the frequency of phenotypes with nucleoplasmic distribution during the expression of the C-terminal mutant. Cells with nucleoplasmic distribution from C and D (24 h) were taken. Two different phenotypes were found: i) one showing only nucleoplasmic distribution and ii) nucleoplasmic distribution plus mini-assemblies ( $\leq 5$ ) as no maxi-assemblies were found to cooccur with the nucleoplasmic phenotype. Overexpression of C-terminal mutant was induced during 24 h with 0.1 and 1.0  $\mu\text{g/ml}$  Tet. The relative frequency of phenotypes per nucleus is shown.

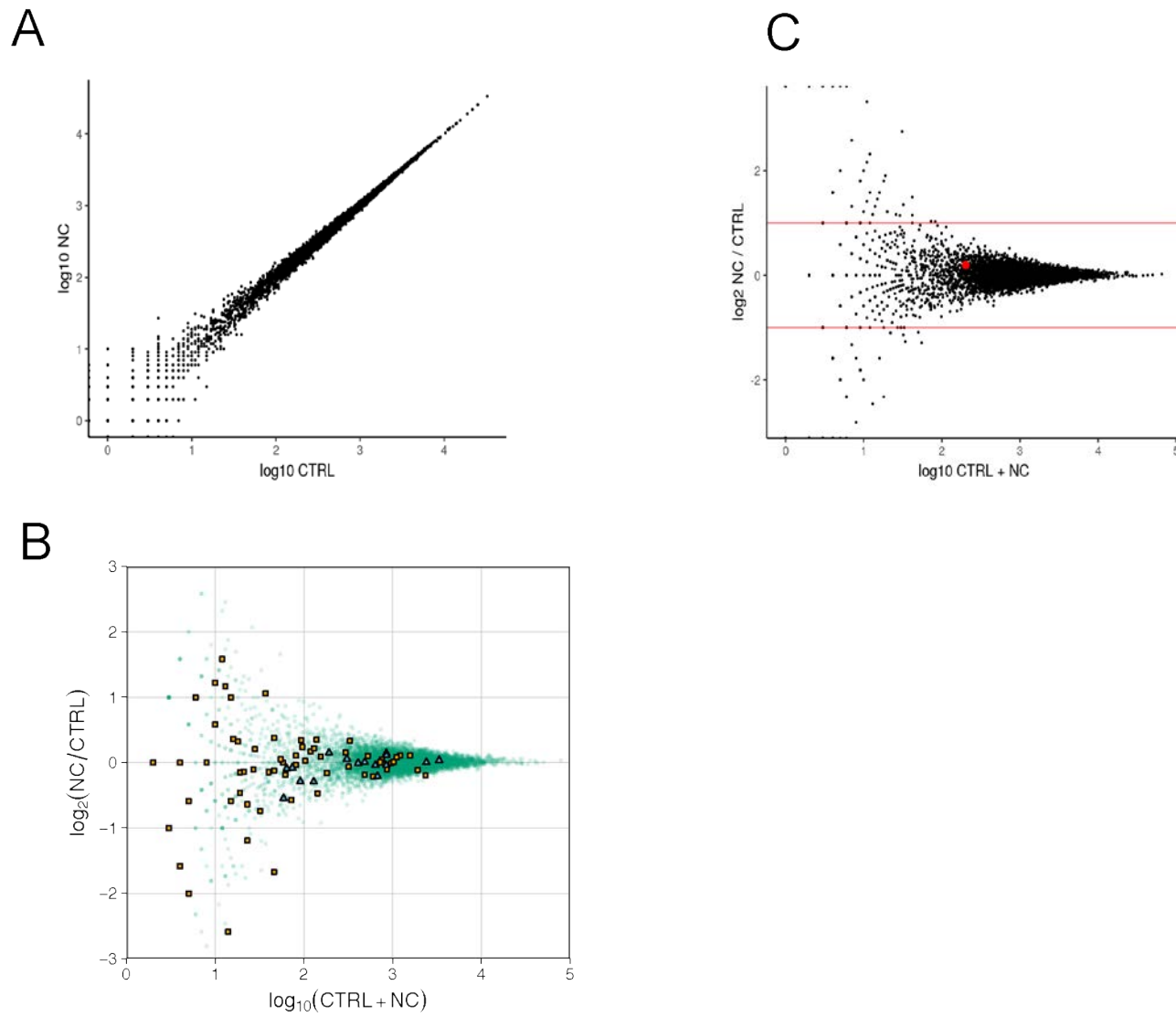


S3

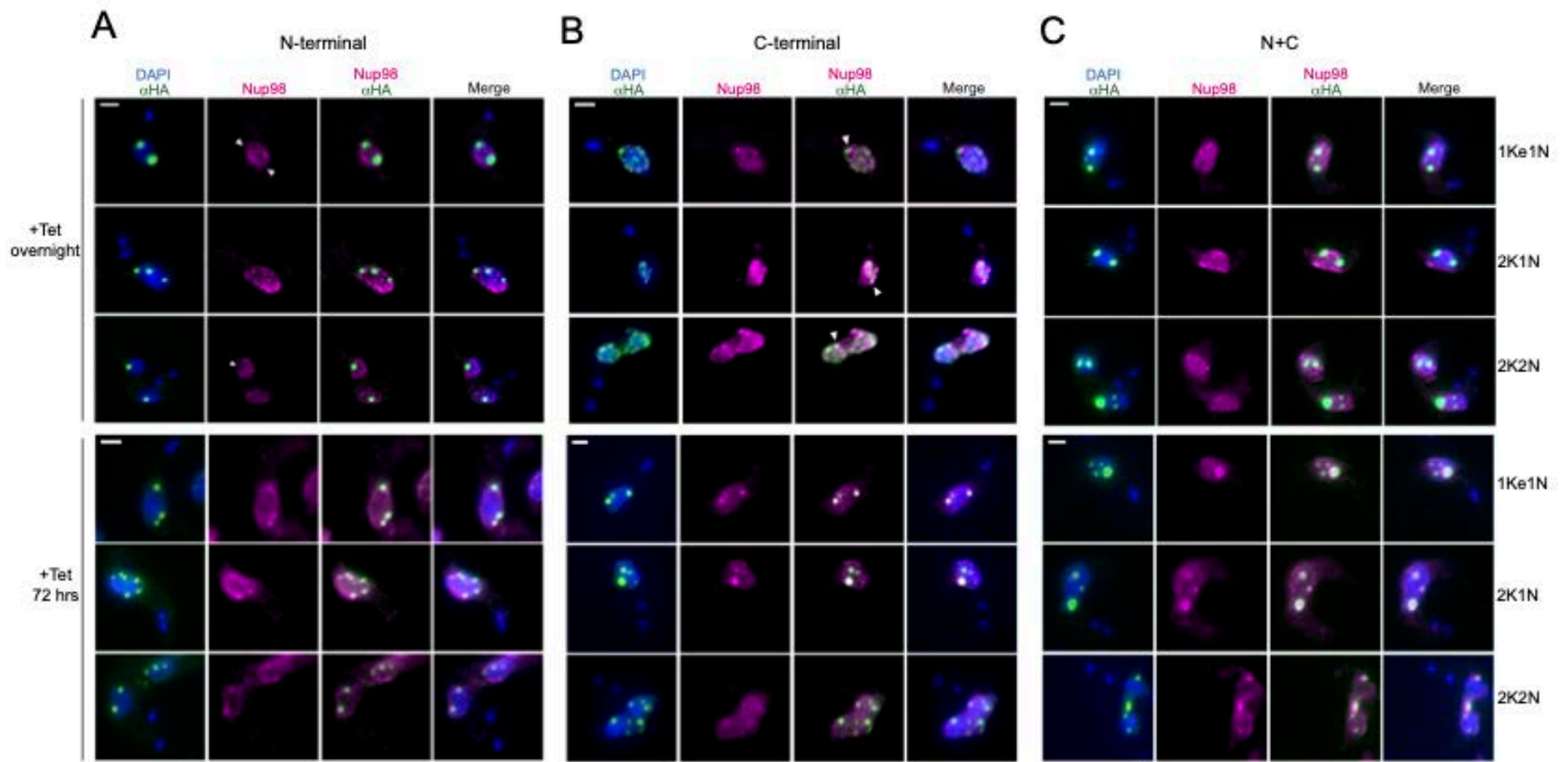
**Fig. S3. NUP-1 assemblies disrupt normal arrangement of endogenous lamina through the cell cycle.** The associations between the assemblies and NUP-1 repeats were unaltered across the cell cycle. BSF cells containing the HA-tagged NUP-1 variants were fixed, stained and visualized by confocal immunofluorescence microscopy. Cells in different stages of cell cycle are shown. Cells were co-stained with anti-HA (green), anti- NUP-1 repeats serum (red) and DAPI (blue), as indicated. Central z-stacks are presented. Scale bar 2 μm. (A) N-terminal mutant, (B) C-terminal mutant and (C) N+C mutant. In (A) maxi (white arrowheads) and mini-assemblies (white stars) are labelled in the first column (DAPI/anti-HA). (D) Effect of the cell cycle in the number of NUP-1 mutant maxi-assemblies. The relative frequency of number of assemblies per nucleus is shown (n=50). Overexpression of NUP-1 mutants was induced during 24 h with 0.1 μg/ml Tet. Data for N-terminal, C-terminal and N+C fusion mutants is shown. (E) NUP-1 α-helical repeats (NUP-1R) in SMB (parental cells) detected by Immunofluorescence (stained with anti-NUP-1 repeats serum, red). A cell in interphase and one in mitosis are shown. DAPI was used to visualize DNA. Scale bar 2 μm.



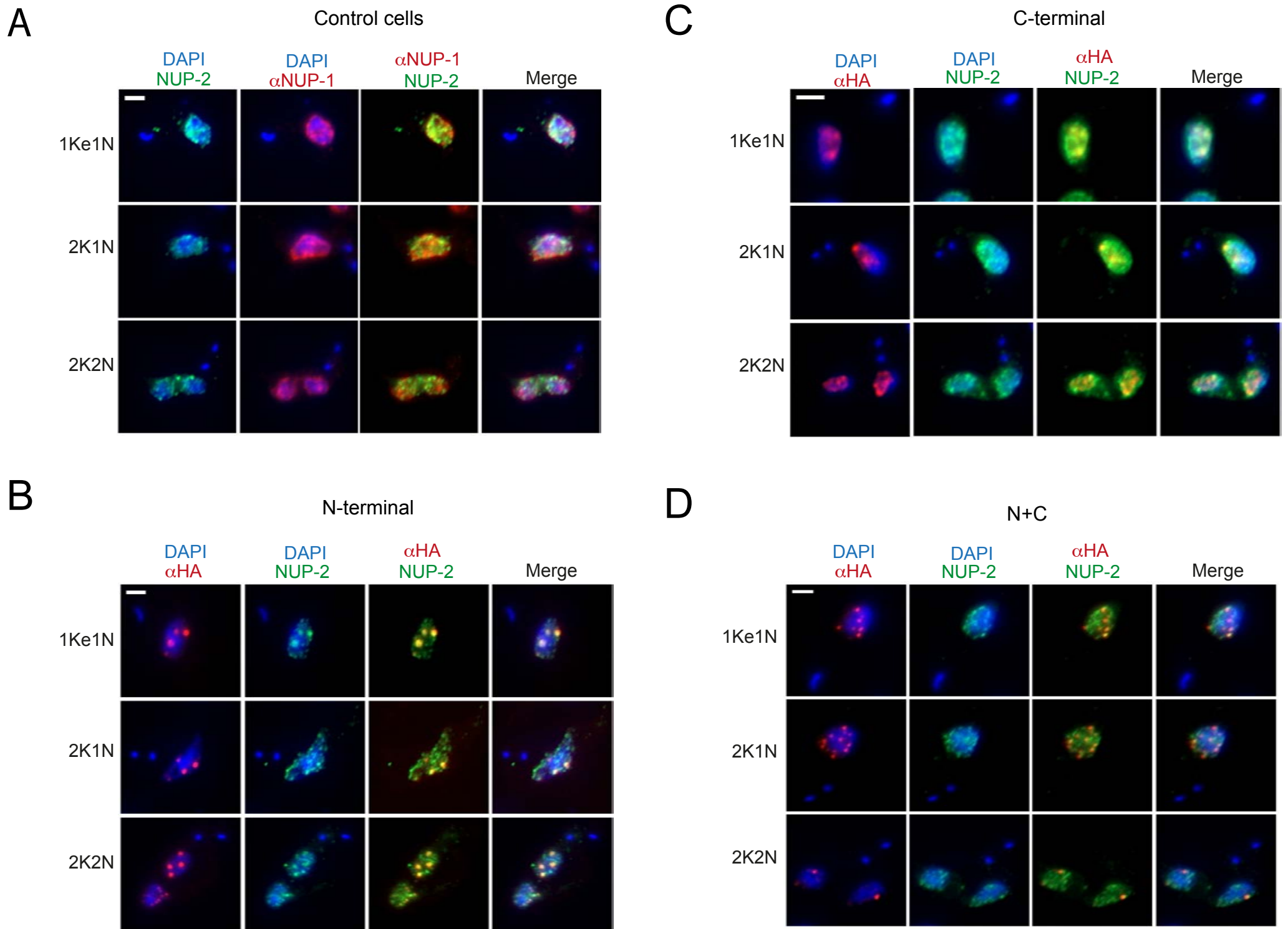
**Fig. S4. Negative controls for immunogold localization.** Left: SMB parental cells, Right: cells expressing N-terminal mutant, stained only with secondary antibodies. Scale bar 1  $\mu$ m.



**Fig. S5. Transcriptomics. A pilot RNAseq experiment was conducted using cells expressing the N+C fusion mutant and SMB cells as control.** (A) Scatter plot of normalized RNA-seq read counts, showing no differentially expressed genes between the two conditions. Intensity-ratio plot of RNA-seq data. Plots based on raw counts per gene. There is no evidence for changes in the expression of the (B) intrinsic VSG-2 (VSG221 or Tb427.BES40.22), represented by a red dot nor (C) other VSG genes (orange squares) or procyclin genes (blue triangles). Green small points represent the rest of the transcripts across the genome. No evidence is found for any of the genes changing between conditions.



**Fig. S6. TbNup98 interacts with both N-terminal and C-terminal domains of NUP-1 across the cell cycle.** TbNup98::myc is visualised together with NUP-1 variants after two induction times with tetracycline: 16 (overnight) and 72 hours. Cells were fixed, stained as indicated and visualized by confocal immunofluorescence microscopy. Overexpressed NUP-1 domains are shown for the (A) N-terminal, (B) C-terminal and (C) N+C fusion variant. Cells are co-stained with anti-HA (green) and anti-myc antibodies (magenta). DAPI was used to visualize DNA (blue). After 16 hours induction, there are regions where TbNup98 is absent from the normal distribution (arrowheads) and do not interact with the NUP-1 assemblies. After 72 hours, NUP-1 mutants colocalize with TbNup98. Central z-stacks are shown. Scale bar 2  $\mu$ m.



**Fig. S7. NUP-2 interactions with NUP-1 across the cell cycle.** NUP-2::TY1 is visualised together with endogenous NUP-1 and NUP-1 mutant variants. Cells were fixed, stained as indicated and visualised by confocal immunofluorescence microscopy. DAPI is used to visualise DNA (blue). (A) Parental SMB cells bearing NUP-2::TY1 were costained with anti-TY1 antibodies (green) and  $\alpha$ -NUP-1 repeats serum (red). Cells expressing the NUP-1::HA variants were tet-induced and co-stained with anti-HA (red) and anti-TY1 (green) antibodies: (B) N-terminal, (C) C-terminal and (D) N+C fusion variant. Central z-stacks are shown. Scale bar 2  $\mu$ m.



```

ATGTTTTCTG CTGGGGACGC ACGGCGGTAC CCCGTTTTCT TCACACGAAC GTGGACTCCA 60
CCACCCGAAA ATATTGGGCA TGTGAGAAGC AACCGAAGCG CCAGCAGCAT CCAAGGAGGT 120
TTGACACATG AGACACCGCC GCTGCTCACA CCGCGGCTGG CCGCACCAAT TAATGTGCGA 180
GGATTGGCCG CAACGGATTC CATACTCGT CTCAACAATC CTTTACCATC TCCTACGGGG 240
TTGCTGACTA ATTCTGCAC TCGTCGGCATG TCACGTGAGG AACCCAGCAT AATGCCGCTT 300
CTCAGCTCCA AACAACTGC TCCTATGGGA ATGCACCCCA CCATCTTCGC TGGGCGTCGG 360
GGCTTGGCCG ACATGTCAGA GGAGGAGCGC ATGGAGTACA CAAATCGCTT GGAGGGAGAT 420
ATGACTCATG TGCATAATAC CCTTAGCCGT GCGTATCAGT TGGGTGATGA CTACAAAAAT 480
GAGGCTGCAC GCCTCCATCG TGAGTTGCAG GACAAAAATC ATCGTTTTGA TTGCCTTTTG 540
CGCGAGCATA GTGCATGCAA CGACGTAATA TACCGCTGCA AGCGAGAAAA CGAAGAAGT 600
CGTCAAAAAC TTGACGAATC GGAAGGGGAG GTGCGACAGC TTAGGGATAA GCTAGTGTCC 660
GTAAATTCAC AAGGCAATA TGTCCCATCT GGAGGTGAAC GACATGTTGG CCGTCAAGAG 720
ATAAGTGCGC TTGAAGAGAA GAACAAAAAA CTTGAGGAAG AACTTCTGGA GCTAACTAAA 780
GAGCTTGAAA GGGAACGTGA GTGTATTAGG CACCACGCTG TCGCTGCAGA AATGGGGAAA 840
AGTGAAAAACA CGAGTCATGA AGAGGAATTG GCGCAATCAA GGTACCTGTT ACAGGTGACA 900
CGTACAGAGA TAACGGATCT ACAGCAGCTT CTTGAAAAG AACGTGAAGA CTACGAAGAA 960
AGCCTTAGGG AAGCAATACA AGCGCGAAC AACCTTCACC AACAAAACAC CGCACTGCAG 1020
GAGCAAAAGG AACAAATTGCA GGAAATGTGT GACGAACAGC ACAGGACGAT TGAGGATCTT 1080
ACTTCACAAC TACTACAGCG CAAAACAGG CAAGCAGTAC AACGTGGGGC ACCTGCACATA 1140
CAAATGGAGA CAACAGACGA GAACAAAAACA GACACAAAATA CAAACAACGA TGATGAAGTG 1200
TACAGAATGC TGGAACTACA ACAGCACACA CTTCAACAGC AGTTCTTCTT GCTCCGTAGG 1260
GAAGGAGAAG CGAAGGACAT CCTGCTCCAA AAAGCGAGCG AAGAAATATT TAACCTCCAA 1320
AATCTTCAGC AGCAACTAGA GGCCGCCCTT CAGAAATCGA GAGAACACGC TGCAGAACTA 1380
ACGAAAAGCC TTTCCACAC ACAAAACCAA CTGCAGACTG CTCAGAAAAG GATCACCGAA 1440
GATAGCTATG TGATCAACAA CTTTCATCAT CAACTCAGAG AAAAAATACA AATATCAGGC 1500
TCCATAAGCG GTGAAAAGAA CATTCCACAA GGCGGGAATA AAGAGGAATC AATAGAGTTA 1560
GTGACAAGAG AGACACAGAT GCCATCAAGG TCAGGAAATG ATTGCAATA TATTACTGCA 1620
AATGTGCAAC ATGAGAAACT AAACCAACCA CAAAAAGCCG ATAGCGGTCA CAATGCCACA 1680
GGGAATAACA AGGAATTGTC GTCAGCTCAA AATGACGAGT ACGAACAAGC CATAATCAAG 1740
CACAAAGATG CAGAAGAAGG TTTAACAGAA GTCATAGAAG CACTGAAAAC GGAAGTGCAG 1800
CACACACAGA AGTGCCTTCG CGAGGCAGGG GAAGAGAAGC TGCAACTAAC GAATAAGCTG 1860
AATGCGGCTG GAGCTCGGGG CCGCTCAACA AGTACAACAA GAAGTGGGAG TTTAACACCA 1920
AATGATACAG AAGGATCACT GAGAACATAT AATGCCGGAT TGAAAACACA GTTATCCTCC 1980
GCCTTGGCAG CCTTAACACA GCTAGCAGAA CAGCAGATG CTACACTAGC AAGAGCAACT 2040
GAAATGGAGG AACGTGTTTC TACTACTGAG GAGGAACTCC GTACAGCGCA CTCAACCACG 2100
//
GATCTTGGTA CACAATTGGC ATCTGCTTTA GTAGCACTTG AGCGACTTGC TGAGGAACGG 9360
GAAGCTGCTT TAGAAAAGGC AACTGAGATG GAGGAACGTG TTTCTACACT TGAGGAGGAA 9420
CTCCGTACAG CAAAGGAGAA GCTGGAGAGG AGTGTGAGG AATATCTTTT TTTAAAAGAT 9480
GAAGTTTTGG TTAGTAATCG TTTGCTTGTG GATAGTGTTT CTTCTTTGAA TGGTAAAGTG 9540
GGGGATAGTG ATGGTGTCTG TGGTGCAGAT GTTGAGAGGT TGTCTCGGGT TGTGGATGAA 9600
CTTCATGCTC AAGTTTCCCG TACGAAGCGT GGTTTTGAAG AATTTTATGA CCGTAGGAGT 9660
GAGGGTTGCG TGACGGAACT TATTGTGCGA AGAAGATCCG TTGATCGCTC TAATGATGCG 9720
AGGAGGAGGT TGGAGGAGCG CAATGTTCCG CTAGAGCAAG ATTTGGAAAAG AAAATGTCTG 9780
GAAGTAGTTA AATTGCAGAA GGAGTGCCAG CGGTTGGAAC AATTTGTTCC GGCAAAGGAC 9840
GTTCTGTCG CACACAGTGT ATTAGGTGTG GATGGTTCCG TTGATGTGAG TTCTGTGGGG 9900
GCGGAACCGG TGGACTTAGA GGCTGTAGAT CTCGCCAGT TTCTACAAAT ATCAAGCCTT 9960
CACGCAGATC TAATGCTCTG CCGAAAAACT TGCCGTCAGT TGGAAATCAA TCAAGAGGAA 10020
CTTCTCTTGT CATTGGAACA GAATTCGTCG CAGTCAAATG CGTATCTGGA AGATTTAGAT 10080
GAAATTCGTC AACAGTTGGT GGAAATGCGT CAGCAACGTG AAGAACTCAT AGCTGAGCGT 10140
CGTACTCTCA CCGAGAGGGT CGATGAACTT GGTGCTGAGC GAGGTGAGGA AGTTAGTCGG 10200
TTGAAAACAGC AGAACAACCT GCTCTCCGCA CAGTTGCAAG CGAGCCGCAA TAAACTCTCC 10260
GCACTGGAGG CATCGAAACG TGAGGGTGAA CTTGCGGGCA GGCAGCAAGC TGAGGAACTG 10320
GCGAAAAGCAT TCAGTCTGAT GGAGGCTCAG GTGCAAAACAC TCCCGGAGGA GGTTCATCG 10380
ACGAGTGGTT CCCCAAAACG GCAAAGCGGT TCCTCTCGTC AAAAGGCCGT TGTGGAGGGG 10440
GATGAAGCGC GTATTGCGAT GTCGCAGGCG CGCGTTACGT TCCTTGAAAA AGCTCTACAA 10500
CGAAAGGATG AGGAGGTTCA ACGGCTGCAG GATGAGCTTG TACAGAAGGA CGAACAACCT 10560
GACCAGTATG AACAAAGATG GGCCAAGGCG GCACAAGATG CCGAGAATGC ATCAAGGAAG 10620
ACCTTACAAC TTGAAAGCGC AGTTTCAGAAG TTGCAGGGCG ATAAGAAAAG TCTGGAGGAC 10680
GAGCTTCGAT ATGCCAAGAC AAGGGTTGTA ACCTATGGTG GTCGTGTGTC ATCAGAAGTG 10740
GCACAACACA GCAGCCACC GGAACAGCAA ATTGCGGGGT CACCTGTGCT AGGTGCAGGA 10800
AGAACCACCA GAGAGAGGGT GAGCTTGTCC GTTGAGTATC CACATCATTC CAGAATCACT 10860
GAACAAACAC AGCGACAGGT ACGGCAAGTC ATGGACATAC GTAGCACAAG GAAAAGGTCT 10920
CGTTCAGCCA ATGCGGTCTC GTGA 10944

```

**Fig. S8. NUP-1 sequences used to create NUP-1 variants.** NUP-1 (Tb927.2.4230) gene sequence is shown above with a gap in the central region of the gene. The N-terminal region is shown in blue, the initial start codon is underlined. The C-terminal region is shown in green with the endogenous NLS highlighted in yellow. The stop codon is shown in red. The N-terminal variant was built cloning the N-terminal sequence shown here followed by the native NLS. The C-terminal variant was built using the sequence in green, avoiding the stop codon. For the N+C variant a fusion of both sequences was made. In all mutants, the HA tag sequence was added at the end.

**Table S1.** Excel sheet. Data from label-free quantitative mass spectrometry of whole lysates of cells expressing the NUP-1 constructs.

[Click here to download Table S1](#)

**Table S2.** List of differentially upregulated proteins in cells expressing NUP-1 constructs

| Gene ID        | Exclusive N-terminal<br>Product Description                                       |
|----------------|---|
| Tb927.10.5250  | zinc finger protein family member, putative                                       |
| Tb927.10.11300 | paraflagellar rod component, putative   |
| Tb927.10.11760 | pumilio/PUF RNA binding protein 6   |
| Tb927.10.12820 | hypothetical protein, conserved   |
| Tb927.10.2610  | Domain of unknown function (DUF1935), putative                                    |
| Tb927.10.3230  | CMGC/MAPK protein kinase, putative  |
| Tb927.10.3810  | Nucleoporin NUP65   |
| Tb927.11.13870 | uncharacterized protein, PH0010 family/AmmeMemoRadiSam system protein A, putative |
| Tb927.11.15030 | small GTPase  |
| Tb927.11.3360  | Component of motile flagella 22   |
| Tb927.11.3500  | Dpy-30 motif containing protein, putative   |
| Tb927.11.5650  | replication factor C, subunit 1, putative   |
| Tb927.2.3580   | transcription elongation factor s-II, putative                                    |
| Tb927.2.4330   | paraflagellar rod protein 5, putative   |
| Tb927.3.1040   | cAMP Response Protein 4   |
| Tb927.3.1800   | hypothetical protein, conserved   |
| Tb927.3.4190   | endosomal integral membrane protein, putative                                     |
| Tb927.3.5020   | Flagellar Member 6  |
| Tb927.6.1080   | hydroxyacylglutathione hydrolase, putative  |
| Tb927.6.3100   | Intraflagellar transport complex B protein 46 C terminal, putative                |
| Tb927.7.3630   | TPR-repeat-containing chaperone protein DNAJ, putative                            |
| Tb927.7.4750   | hypothetical protein, conserved   |
| Tb927.9.11540  | hypothetical protein, conserved   |
| Tb927.9.2390   | hypothetical protein, conserved   |
| Gene ID        | Exclusive C-terminal<br>Product description                                       |
| Tb927.1.1540   | Tubulin/FtsZ family, putative   |
| Tb927.10.10140 | paraflagellar rod component, putative   |
| Tb927.10.11800 | Axonemal inner arm dynein light chain, putative                                   |
| Tb927.10.12360 | hypothetical protein, conserved   |
| Tb927.10.15850 | Peroxisome biogenesis factor 12   |
| Tb927.10.350   | protein kinase PK4, putative  |
| Tb927.10.5630  | hypothetical protein, conserved   |
| Tb927.10.5880  | Proteophosphoglycan, putative   |
| Tb927.10.6670  | dynein light chain, putative  |
| Tb927.10.7880  | Sperm tail/Sperm tail C-terminal domain containing protein, putative              |
| Tb927.10.8650  | ran binding protein, putative   |
| Tb927.10.8930  | paraflagellar rod component, putative   |
| Tb927.11.10540 | hypothetical protein, conserved   |
| Tb927.11.15480 | heat shock protein Hsp20, putative  |
| Tb927.11.15910 | iron superoxide dismutase   |
| Tb927.11.16090 | Outer dynein arm docking complex protein 2, putative                              |
| Tb927.11.4450  | ALBA-Domain Protein   |
| Tb927.11.4920  | hypothetical protein, conserved   |
| Tb927.11.510   | RNA-binding protein, UBP2, UBP1   |
| Tb927.11.6280  | pyruvate phosphate dikinase   |
| Tb927.11.6370  | leucine-rich repeat protein (LRRP), putative                                      |
| Tb927.11.8440  | haloacid dehalogenase-like hydrolase, putative                                    |
| Tb927.11.9470  | ADP-ribosylation factor GTPase activating protein, putative                       |
| Tb927.2.1890   | E2-like ubiquitin-conjugation enzyme  |
| Tb927.2.5810   | Holliday-junction resolvase-like of SPT6/SH2 domain containing protein, putative  |
| Tb927.3.1010   | hypothetical protein, conserved   |
| Tb927.3.3690   | flagellar radial spoke protein-like, putative                                     |
| Tb927.3.5010   | hypothetical protein, conserved   |
| Tb927.4.750    | 50S ribosomal protein L7Ae, putative  |
| Tb927.5.2620   | hypothetical protein, conserved   |
| Tb927.5.4150   | hypothetical protein, conserved   |
| Tb927.6.1730   | hypothetical protein, conserved   |
| Tb927.6.3160   | splicing factor 3a, putative  |
| Tb927.6.3720   | hypothetical protein, conserved   |
| Tb927.7.1120   | trypanothione/trypanothione dependent peroxidase 1, 3                             |
| Tb927.7.1310   | hypothetical protein, conserved   |
| Tb927.8.1340   | 60S ribosomal protein L7a, putative   |
| Tb927.9.3770   | hypothetical protein, conserved   |
| Tb927.9.6560   | NAK family pseudokinase, putative   |
| Tb927.9.9940   | parkin coregulated gene protein   |
| Gene ID        | Exclusive N+C<br>Product Description  |
| Tb927.11.3510  | hypothetical protein, conserved   |
| Tb927.11.6510  | 40S ribosomal protein S21, putative   |
| Tb927.3.4920   | LETM1 and EF-hand domain-containing protein 1, putative                           |
| Tb927.5.1250   | GAF domain/TIP41-like family, putative  |
| Tb927.5.2900   | histone deacetylase 4   |
| Tb927.9.10790  | hypothetical protein  |
| Tb927.9.13320  | hypothetical protein, conserved   |
| Tb927.9.1600   | hypothetical protein, conserved   |

[Click here to download Table S2](#)

**Table S3.** List of differentially downregulated proteins in cells expressing NUP-1 constructs.

| Gene ID                       | Product description   |
|-------------------------------|---|
| Tb11.02.5380                  | exosome complex exonuclease RRP44p homologue  |
| Tb11.v5.0480                  | DNA-directed RNA polymerase, alpha subunit, putative                                      |
| Tb927.10.12630                | hypothetical protein, conserved   |
| Tb927.10.1320                 | hypothetical protein, conserved   |
| Tb927.10.14840                | Mitochondrial ADP/ATP carrier protein 5a  |
| Tb927.10.3150                 | N-acetyltransferase, putative   |
| Tb927.10.4000                 | methylglutaconyl-CoA hydratase, mitochondrial precursor, putative                         |
| Tb927.10.4740                 | nucleolar RNA-binding protein, putative   |
| Tb927.10.5480                 | 60S ribosomal protein L24, putative   |
| Tb927.10.6970                 | serine peptidase, Clan SC, Family S9B   |
| Tb927.10.730                  | ATP synthase, putative  |
| Tb927.10.8830                 | Flagellum attachment zone protein 5   |
| Tb927.10.9430                 | phosphoribosylpyrophosphate synthetase, putative  |
| Tb927.11.13010                | hypothetical protein  |
| Tb927.11.14430                | proteasome regulatory non-ATP-ase subunit   |
| Tb927.11.14780                | phosphomannose isomerase  |
| Tb927.11.2260                 | Eukaryotic translation initiation factor 4E-1   |
| Tb927.11.2400                 | Flabarin-like protein   |
| Tb927.11.2730                 | UDP-galactose 4-epimerase   |
| Tb927.11.800                  | prefoldin subunit, putative   |
| Tb927.2.6150                  | adenosine transporter 2   |
| Tb927.3.2230                  | succinyl-CoA synthetase alpha subunit, putative   |
| Tb927.5.1160                  | Degradation arginine-rich protein for mis-folding, putative                               |
| Tb927.5.590                   | protein phosphatase 1, regulatory subunit, putative                                       |
| Tb927.7.1670                  | Eukaryotic translation initiation factor 4E type 6  |
| Tb927.7.2790                  | Component of motile flagella 10   |
| Tb927.7.3080                  | Kinetochore interacting protein 4   |
| Tb927.7.4440                  | NAD dependent epimerase/dehydratase family, putative                                      |
| Tb927.7.5300                  | C2 domain/Ankyrin repeats (3 copies), putative  |
| Tb927.8.730                   | nucleolar RNA-binding protein, putative   |
| Tb927.9.10690                 | Protein of unknown function (DUF2009), putative   |
| Tb927.9.11410;Tb927.9.11380   | 60S ribosomal protein L23, putative   |
| Tb927.9.3990;Tb927.9.3920     | ribosomal protein S7, putative  |
| Tb927.9.8720                  | fructose-1,6-bisphosphatase   |
| <b>Exclusive C-terminal</b>   |   |
| Gene ID                       | Product description   |
| Tb927.1.1420                  | conserved protein, unknown function   |
| Tb927.10.12510                | P-type H <sup>+</sup> -ATPase, putative   |
| Tb927.11.10030                | 60S ribosomal protein L29, putative   |
| Tb927.11.1090                 | calpain-like protein, putative  |
| Tb927.11.11830;Tb927.11.11820 | 40S ribosomal protein S17, putative   |
| Tb927.11.11980                | cytoskeleton-associated protein 15  |
| Tb927.11.13970                | Pab1p-dependent poly(A) ribonuclease subunit, putative                                    |
| Tb927.11.14560                | Cleavage and polyadenylation specificity factor CPSF160, putative                         |
| Tb927.11.14690                | Microtubule-binding protein MIP-T3, putative  |
| Tb927.11.16760                | T-complex protein 1, alpha subunit, putative  |
| Tb927.11.2250                 | conserved protein, unknown function   |
| Tb927.11.230                  | cleavage and polyadenylation specificity factor subunit 2                                 |
| Tb927.11.4820;Tb927.10.14580  | 60S ribosomal protein L17, putative   |
| Tb927.11.740                  | eukaryotic translation initiation factor 5A   |
| Tb927.2.1330                  | retrotransposon hot spot protein 6 (RHS6), degenerate                                     |
| Tb927.2.2230                  | hypothetical protein, conserved   |
| Tb927.3.3330                  | heat shock protein 20, putative   |
| Tb927.3.4260                  | Acetyl-CoA hydrolase  |
| Tb927.3.5580                  | tryptophanyl-tRNA synthetase  |
| Tb927.4.2170                  | hypothetical protein, conserved   |
| Tb927.4.2530                  | hypothetical protein, conserved   |
| Tb927.4.4360;Tb927.8.8020     | monoglyceride lipase, putative  |
| Tb927.5.1260                  | Sulfate transporter N-terminal domain with GLY motif/Sulfate transporter family, putative |
| Tb927.6.1140                  | dolichyl-P-Man:GDP-Man5GlcNAc2-PP-dolichyl alpha-1,2-mannosyltransferase, putative        |
| Tb927.6.1650;Tb927.6.1640     | single strand-specific nuclease, putative   |
| Tb927.6.2170                  | co-chaperone GrpE, putative   |
| Tb927.6.3840                  | reticulon domain protein  |
| Tb927.6.4090                  | chaperonin HSP60, mitochondrial precursor, putative                                       |
| Tb927.7.240;Tb927.7.230       | 40S ribosomal protein S33, putative   |
| Tb927.7.4220                  | WD domain, G-beta repeat/Djp2/Utp12 Family, putative                                      |
| Tb927.7.7460                  | hypothetical protein, conserved   |
| Tb927.8.4640                  | Component of motile flagella 19   |
| Tb927.8.6070                  | Trypanosome basal body component protein  |
| Tb927.8.8230;Tb927.8.8220     | amino acid transporter, putative  |
| Tb11.v5.0621                  | hypothetical protein, conserved   |
| Tb927.9.1780                  | sec1 family transport protein, putative   |
| Tb927.9.5410                  | hypothetical protein, conserved   |
| Tb927.9.9670                  | proteasome alpha 1 subunit, putative  |
| <b>Exclusive N+C</b>          |   |
| Gene ID                       | Product description   |
| Tb927.10.5520                 | AmmeMemoRadiSam system protein B, putative  |
| Tb927.11.5590                 | Anaphase-promoting complex-associated protein AP1   |
| Tb927.4.2380                  | sarcoplasmic reticulum glycoprotein, putative   |
| Tb927.11.9810                 | NUDIX hydrolase 3, putative   |
| Tb927.10.260                  | BoIA-like protein, putative   |
| Tb927.11.16610                | zinc-finger of a C2HC-type, putative  |

[Click here to download Table S3](#)