INVESTIGATION OF THE DISTAL APPENDAGE PROTEIN CEP164 IN *Trypanosoma brucei*

Madison Atkins

This thesis is submitted in partial fulfilment of the requirements of the award for the degree of Doctor of Philosophy

Awarded by Oxford Brookes University

August 2019

Abstract

Cilia and flagella are highly conserved organelles required for eukaryotic cells to perform sensory and motility functions. They are found in many cells of the human body and are essential for the survival of single celled protozoans. These organelles are nucleated from a basal body after docking to the plasma membrane and are structurally similar to centrioles. Docking of the basal body requires the presence of transitional fibres, which mature from distal appendages on the centriole. These structures rely on the presence of the Cep164 protein to carry out docking and ciliogenesis. Understanding of basal body docking and flagellum length regulation in the *Trypanosoma brucei* cell is limited and is thought to involve the transitional fibres and surrounding structures. Three diverse Cep164 orthologues in the *T. brucei* parasite have been identified, but knowledge on their function is limited as functional analysis of these proteins has not been carried out.

This project, in collaboration with the TrypTag project used a bioinformatic approach to identify basal body proteins in the *T. brucei* cell. Candidate proteins were confirmed as basal body components through endogenous tagging and co-localisation studies. The Cep164 orthologues (Cep164A, Cep164B and Cep164C) localised to the distal section of the mature basal body only, with Cep164C showing a cell cycle dependent localisation. Functional analysis of the Cep164 proteins was performed through the generation of inducible RNAi cell lines. Ablation of these proteins showed a functional role of flagellum length regulation, or formation of the transitional fibres and correct basal body docking.

This work provides further understanding on the three Cep164 proteins in the *T. brucei* cell and proposes how the cell regulates its flagellum length.

Acknowledgements

Firstly, I would like to thank my supervisor Professor Sue Vaughan in allowing me the opportunity to carry out my PhD project in her lab and for supervising me for the last four years. Also, a huge thank you to thank Dr Jiri Tyc for collaborating with me in this study and his never-ending patience and support during the project.

I would also like to thank all members of the Trypanosoma group at Oxford Brookes University and others in the Sinclair Annex for all the help provided during the years.

To my mom and dad, a huge thank you for helping and supporting me through my education while putting up with never quite knowing what to tell people when they ask what I am doing in my studies. To my close family and friends, thank you for listening and providing support when it was needed.

George, I cannot thank you enough for how patient, supportive and understanding you have been over the last few years, while never losing confidence that I could complete this achievement before our own personal deadline – Thank you.

This thesis is dedicated to my son Leo North, who gave me all the motivation I needed to complete my studies before his arrival and to hopefully show him when he is older to never give up when things get tough.

Contents

List of Figures

List of Tables

List of abbreviations

Abbreviation	Term
µm	Micrometer
1K1N	One kinetoplast and one nucleus
1K2N	One kinetoplast and two nuclei
2K1N	Two kinetoplasts and one nucleus
2K2N	Two kinetoplasts and two nuclei
A-tubule	A-microtubule
BB	Basal body
BB/C	Basal body/centriole
BLAST	Basic local alignment search tool
BBP	Basal body pairs
bp	Base pairs
BSF	Bloodstream form
C-	Carboxyl (terminus)
C2CD3	C2 calcium dependent domain containing 3
Cby	Chibby protein
CEP	Centrosomal protein
Cep164	Centrosomal protein 164
Cep164A	Centrosomal protein 164 A
Cep164B	Centrosomal protein 164 B
Cep164C	Centrosomal protein 164 C
CP110	Centriolar coiled-coil protein of 110
C-tubule	C-microtubule
CV	Ciliary vesicles
DAs	Distal appendages
DAPs	Distal appendage proteins
DAPI	4,6-diamidino-2-phenylindole
DC	Daughter centriole
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

dNTPs	Deoxyribonucleotide triphosphates
FAZ	Flagellum attachment zone
FCS	Fetal calf serum
FP	Flagellar pocket
FRAP	Fluorescence recovery after photobleaching
G1 phase	Gap 1 phase
G2 phase	Gap 2 phase
gDNA	Genome DNA
GFP	Green fluorescent protein
HAT	Human African trypanosomiasis
HMM	Hidden Markov model
hrs	Hours
IFT	Intraflagellar transport
K	Kinetoplast
Kb	Kilobase
kDa	Kilodalton
kDNA	Kinetoplast deoxyribonucleic acid
LB	Lysogeny broth
MBB	Mature basal body
MC	Mother centriole
MIB	Microscopy image browser
MNG	Monomeric neon green
M-phase	Mitosis phase
mRNA	Messenger ribonucleic acid
MTOC	Microtubule organising centre
N	Nucleus
NCBI	National Centre for Biotechnology Information
NF	New flagellum
nm	Nanometers
NMBB	New mature basal body
ns	Not significant
ODF1	Outer dense fibre of sperm tail 1

ODF2	Outer dense fibre of sperm tail 2
OF	Old flagellum
OMBB	Old mature basal body
PBB	Pro-basal body
PBS	Phosphate buffer saline
PCF	Procyclic form
PCM	Pericentriolar material
PCR	Polymerase chain reaction
PDVF	Polyvinylidene difluoride
PFAM	Protein families database
PFR	Paraflagellar rod
PLK	Polo-like kinase
pPOT	Plasmid for PCR only tagging
R	Pearson correlation coefficient
RNAi	Ribonucleic acid interference
RP2	Retinitis pigmentosa 2 protein
RPM	Rotation per minute
SAS	Spindle assembly abnormal
SBF-SEM	Serial block-face scanning electron microscopy
SDAs	Sub-distal appendages
SDM-79	Semi-defined media 1979
SDS	Sodium dodecyl sulphate
SEM	Scanning electron microscopy
siRNA	Small interfering ribonucleic acid
SmOx	Single marker Oxford
S-phase	Synthesis phase
STIL	SCL/TAL1-interupting locus
TEM	Transmission electron microscopy
TF	Transitional fibre
TRITC	Tetramethylrhodamine
tRNA	Transfer ribonucleic acid
TTBK2	Tau tubulin kinase 2

TZ	Transition zone
WW domain	Tryptophan/tryptophan domain
WHO	

1. Introduction

1.1 Tubulin

Tubulins are highly conserved globular proteins that perform a large number of functions in the eukaryotic cell. These proteins form or function at the microtubules and current members of this superfamily include alpha (α), beta (β), gamma (γ), delta (δ), epsilon (ε) and zeta (ζ) (McKean *et al;* 2001; Dutcher 2003; Turk *et al*, 2015). Studies have shown not all tubulins are present in each organism (Luduena, 2013), aside from α, β and γ which are present in all kingdoms (figure 1).

The six tubulins and their presence and absence. α, β and γ are present in all kingdoms

					Localisation		
Tubulin	Animalia	Plantae	Fungi	Protists	Basal Body	Centrosome	Other
α + β (alpha and beta)	+	+	+	+	+	+	Ubiquitous
γ (gamma)	+	+	+	+	+	+	Nucleus
δ (delta)	+	+		+	+	+	Perinuclear ring
ε (epsilon)	+	+		+	+	+	
ζ (zeta)	+			+	+		

Figure 1: Tubulin family

(+). Edited from Luduena (2013).

Both α and β are the most abundant tubulins and are structural homologues to each other. These form the subunit of the microtubule and form heterodimers once each monomer binds to a guanine nucleotide (Farr *et al,* 1989). Subunits then form and associate in a head-to-tail manner producing protofilaments (figure 2A) (Nogales *et al,* 1998; Brouhard and Rice, 2014). Whereas γ tubulin is present in all eukaryotic cells and functions in nucleation of the microtubule and microtubule assembly (Oakley *et al,* 1999; Oakley 2000).

1.2 Protofilaments and microtubule structures

In eukaryotic cells, 13 protofilaments produce one microtubule. This active structure switches between growth (figure 2B) and shrinkage (figure 2C) phases, allowing rapid remodelling of the microtubule during the cell cycle (Brouhard and Rice, 2014).

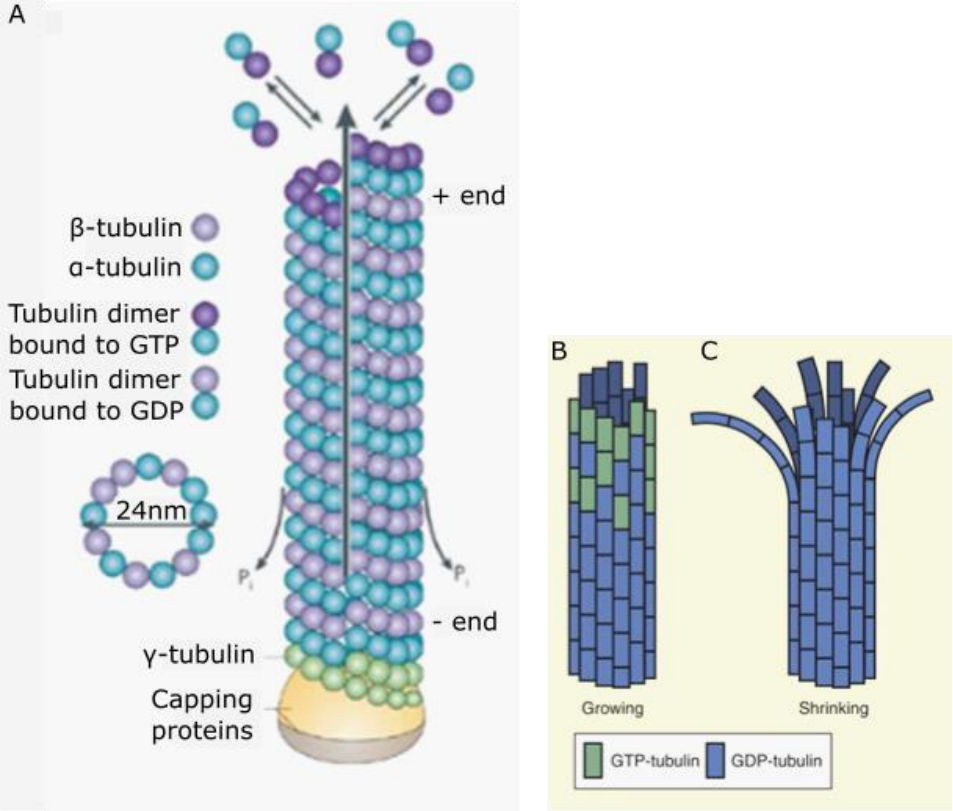


Figure 2: Protofilaments form microtubules

A) α and β tubulins joining in a head-to-tail manner to form 13 protofilaments, making a microtubule. Growth (B) and shrinkage (C) of a microtubule allows rapid remodelling of the cilium. A) Edited from Nature Education (2014). B and C taken from Sept *et al* (2007).

Microtubules are essential in providing support to the cell, determining cell morphology and for allowing motility and cellular division (Ganguly *et al,* 2012; Nogales, 2000). The active plus end of a microtubule grows rapidly, while the minus end is often anchored to a microtubule organizing centre (MTOC) (Wu and Akhmanova, 2017).

1.3 Basal bodies and centrioles

MTOCs are structures within a cell that organise, nucleate and anchor microtubules, allowing new microtubules to carry out functions including mitosis and assembly of cilia (Wu and Akhmanova, 2017; Paz and Luders, 2018). A well-studied MTOC is the centrosome and is considered the main MTOC in animal cells (Paz and Luders, 2018). Centrioles are organelles that form the MTOC and are two symmetrical structures surrounded by pericentriolar material (PCM) (Sanchez and Feldman *et al,* 2018) (figure 3; PCM).

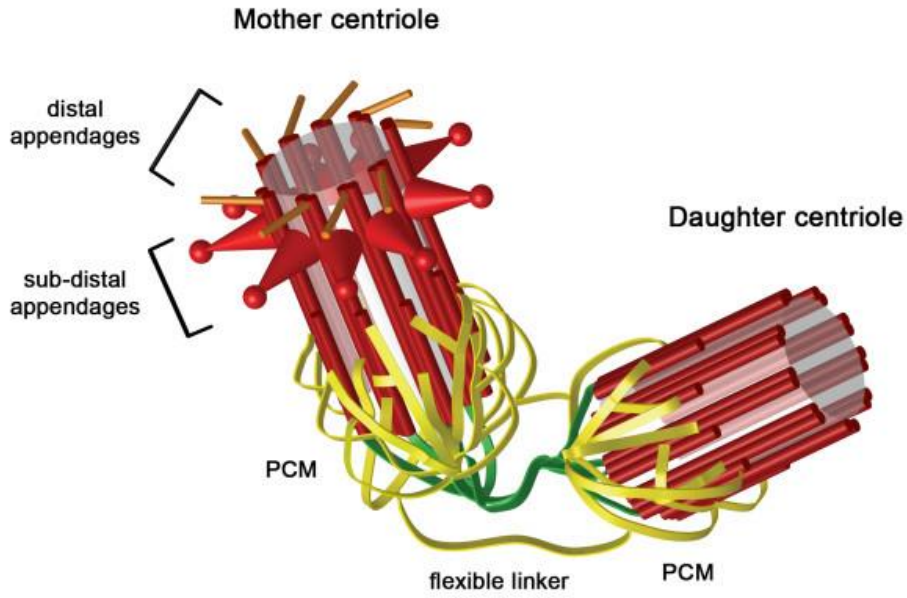


Figure 3: Centrioles and their components

The mother and daughter centriole with morphological differences at the distal end. The mother centriole contains distal and sub-distal appendages, which are lacking in the daughter centriole. Taken from Sillibourne *et al* (2010).

These self-reproducing organelles nucleate microtubules during cell division to capture chromosomes, which depolymerise to separate the duplicated chromosomes (Meunier and Vernos, 2012). Centrioles are formed by nine-triplet microtubules arranged in a cylindrical structure, which nucleate from the distal end. These structures contain a diameter of approximately 250nm, with a varying length ranging between 150-500nm depending on the cell type (Winey and O'Toole, 2014).

Each centriole pair contains a mother centriole (MC) and a daughter centriole (DC), with the MC being distinguished by the addition of its distal appendages (DAs) (figure 3; distal appendages), which are absent in the DC (Hatch and Stearns *et al,* 2014). During cell cycle quiescence, centrioles transform into the basal body (BB) organelle, dock to the plasma membrane and organise the microtubules (figure 4C) (Kobayashi and Dynlacht, 2011; Pearson, 2014). These organelles are similar in structure and function but have different termed components, with the centriole containing DAs (figure 4A; appendages) and the BB containing transitional fibres (TFs) (figure 4B; transitional fibres).

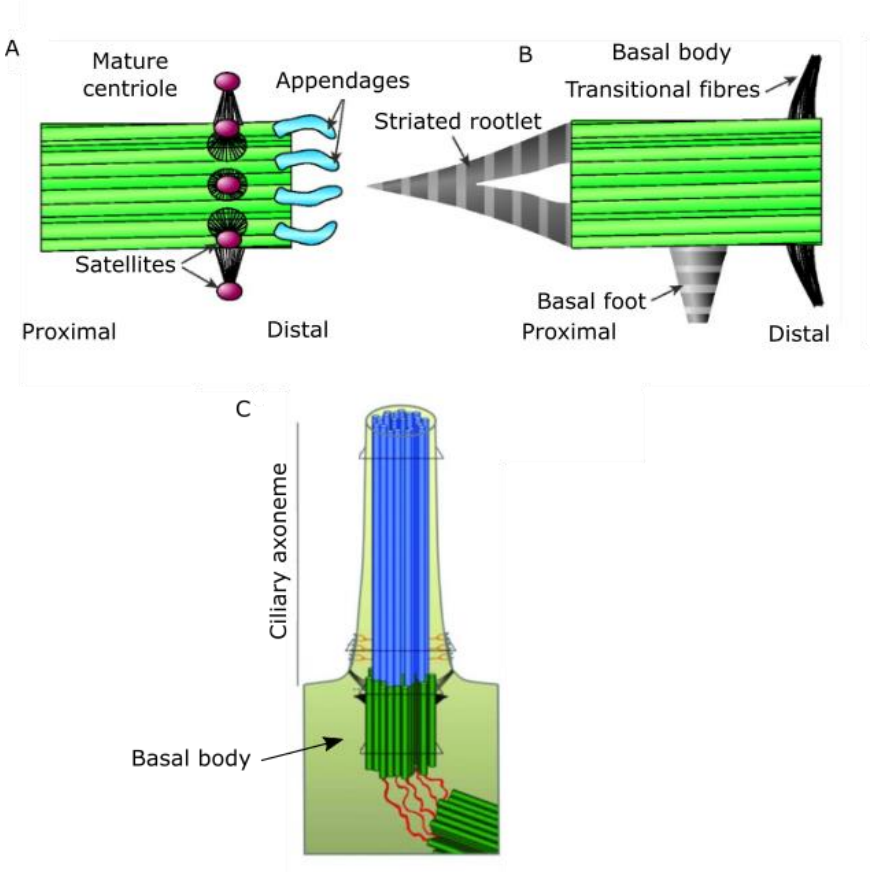


Figure 4: Composition of centriole and basal body distal sections

A and B) Differences between mature centriole (appendages) and basal body (transitional fibres) distal structures. C) During cell cycle quiescence, the basal body docks to the plasma membrane and nucleates a cilium. A and B taken from Dawe *et al* (2007). C edited from Pedersen *et al* (2012).

The proximal end of a BB is composed of a 9 + 0 arrangement of triplet microtubules (figure 5; B2), with the distal section (location of TFs) being composed of a 9 + 0 arrangement of doublet microtubules (figure 5; B4) (Vaughan and Gull, 2016). At the very proximal section of the BB is a layer of 5-6 cartwheel stacks (figure 5; B1) which form a central ring. Nine filaments radiate from this ring, with each filament connecting to the A-tubule of the triplets (Hirono, 2014).

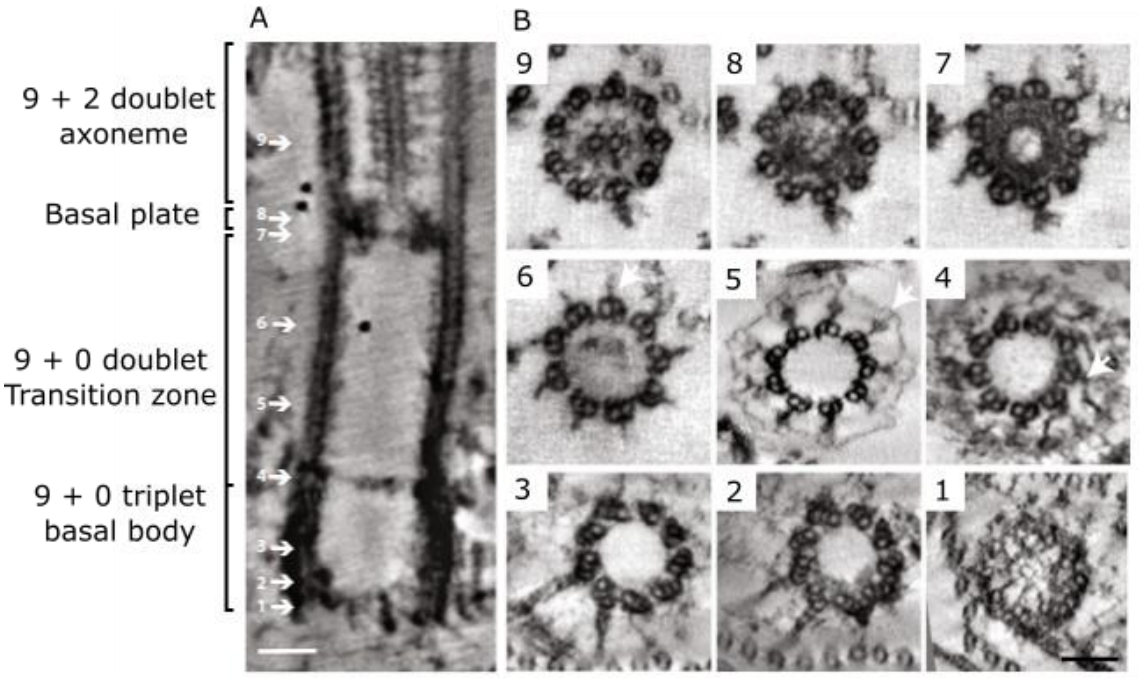


Figure 5: Ultrastructure of *T. brucei* basal body by TEM

A) Longitudinal TEM section of a basal body marking the different sections of a basal body and axoneme. B) Cross sections of a mature basal body from proximal (1) to distal (9). Taken from Vaughan and Gull (2016).

The cartwheel structure has an essential function in developing the nine-fold symmetry of the basal body/centriole (BB/C). Studies show that lack of the Bld10p protein in the cartwheel structure causes abnormalities in the numbers of microtubule triplets formed (Hiraki *et al,* 2007). This structure is also required for stability, as BB/Cs lacking the Bld12 (SAS-6) protein forming the central part of the cartwheel, caused centrioles to split into one to five triplet microtubules (Nakazawa *et al,* 2007).

1.4 The centriole duplication cycle

During early stages of the cell cycle, the BB is docked to the plasma membrane and has nucleated the cilium. This cilium extending from the BB is then reabsorbed, where the BB then returns to its previous role as a centriole required for cell division (Pearson, 2014). Both BBs and centrioles undergo duplication once in each cell cycle. At the beginning, the DC re-orientates from its orthogonal position to a parallel position to the MC (figure 6A) but is still attached through flexible connections. Both mature centrioles then develop their own DCs orthogonal to the MC (figure 6B) (Crasta and Surana, 2006).

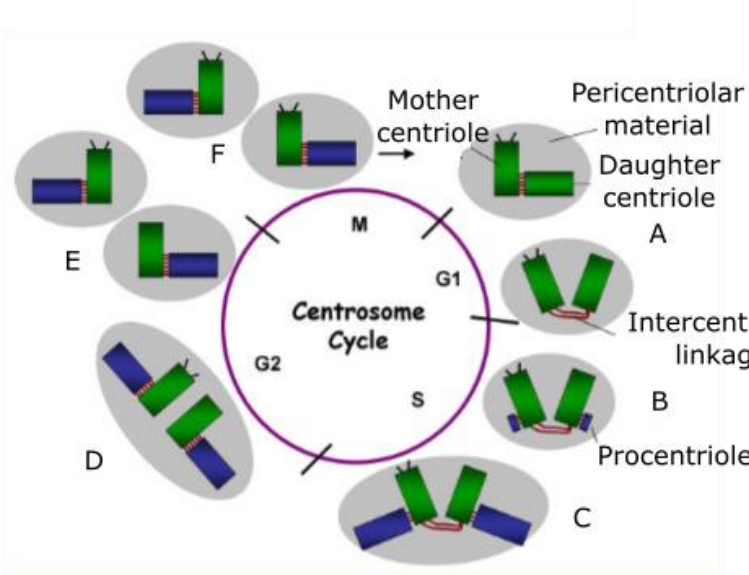


Figure 6: Centrosome cell cycle

The cycle begins with one centriole pair, with the daughter centriole maturing after entry into the cell cycle (A). Each mature centriole develops their own daughter centriole orthogonal to the parent (B-C). Once developed the tether holding the two pairs together dissolves (D), the two new centriole pairs then separate (E-F). Taken from Crasta and Surana (2006).

The developing DC continues to elongate until reaching the length of the MC (figure 6C). The tether between the two mature centrioles then dissolves, leaving two centriole pairs of one MC and one DC (figure 6D). Once separated the youngest parent centriole begins maturation by acquiring DAs and sub-distal appendages to the distal end (SDAs) (figure 6E and 6F).

Biogenesis of the DC/pro-basal body (PBB) is initiated and driven by three major proteins; PLK4, STIL and SAS-6. SAS-6 has previously been described (section 1.3) and is required for stability. PLK4 however drives centriole duplication (Bettencourt-dias *et al,* 2005) phosphorylating the STIL protein, allowing STIL to bind with SAS-6 and initiate centriole formation.

Maturation of the younger BB/C requires the acquisition of DAs (or transitional fibres in BBs) and SDAs (or basal feet in BBs) to the distal end (Wei *et al,* 2015). These appendages are essential for BB/C docking to the plasma membrane, for recruitment of the intraflagellar transport (IFT) particles and for ciliogenesis. DAs and SDAs require separate proteins for formation, with SDAs requiring the ODF2 protein and DAs requiring the ODF1 and C2CD3 proteins. Each protein then leads to recruitment of others for further development, with the DA proteins recruiting the centrosomal protein 89 (Cep89) and the centrosomal protein 164 (Cep164) (Loncarek and Bettencourt-Dias, 2018).

1.5 Cilia and flagella

Cilia and flagella are antennae-like structures projecting from the cell, anchored by the BB, and are present in most cells (Mitchison and Valente, 2016). These highly conserved organelles are essential in performing sensory or motility functions and are categorised into two classes based on their structure and function. Motile cilia or flagella contain dynein proteins (motor proteins) that allow movement, including removing debris or cell motility. Primary cilia are non-motile and are involved in sensory functions or transportation (Ishikawa, 2017).

Nucleation of the cilium only occurs after docking of the BB to the plasma membrane. The microtubules nucleated generate the axoneme of the structure, with the BB and axoneme being connected by the transition zone (TZ) (Ishikawa, 2017; Vaughan and Gull, 2016) (figure 5; B5). The axoneme is an external structure and consists of a 9 + 2 microtubule arrangement with the addition of two central microtubules required for beating (Nakazawa *et al,* 2014).

1.6 *Trypanosoma brucei* life cycle and disease

Trypanosoma brucei are protozoan parasites used in this study and rely on their flagellum for motility and survival. These organisms are parasitic disease-causing agents affecting both human and animal species, creating huge economic losses in sub-Saharan African regions through loss of animal productivity. *T. brucei* causes human African trypanosomiasis (HAT) in humans and nagana in domestic species that is transmitted through the tsetse fly and fatal if left untreated (Franco *et al,* 2014).

With a complex life cycle, this parasite requires an insect and mammalian host for development. The cycle begins when the tsetse fly takes a blood meal from the host, injecting metacyclic trypomastigotes into the mammalian host (figure 7;1). Metacyclic trypomastigotes then transform into bloodstream trypomastigotes (figure 7; 2) which divide through binary fission (figure 7; 3-4) into either the slender proliferative cell, or the non-dividing short stumpy cell (Matthews *et al,* 2005). A tsetse fly then takes a bloodmeal and ingests the bloodstream trypomastigotes which transform into the procyclic trypomastigotes in the insect vector (figure 7; 5-6). Procyclic trypomastigotes then transform into epimastigotes, which transform again into the metacyclic trypomastigotes in the salivary glands of the fly (figure 7; 7-8), where the life cycle is then repeated.

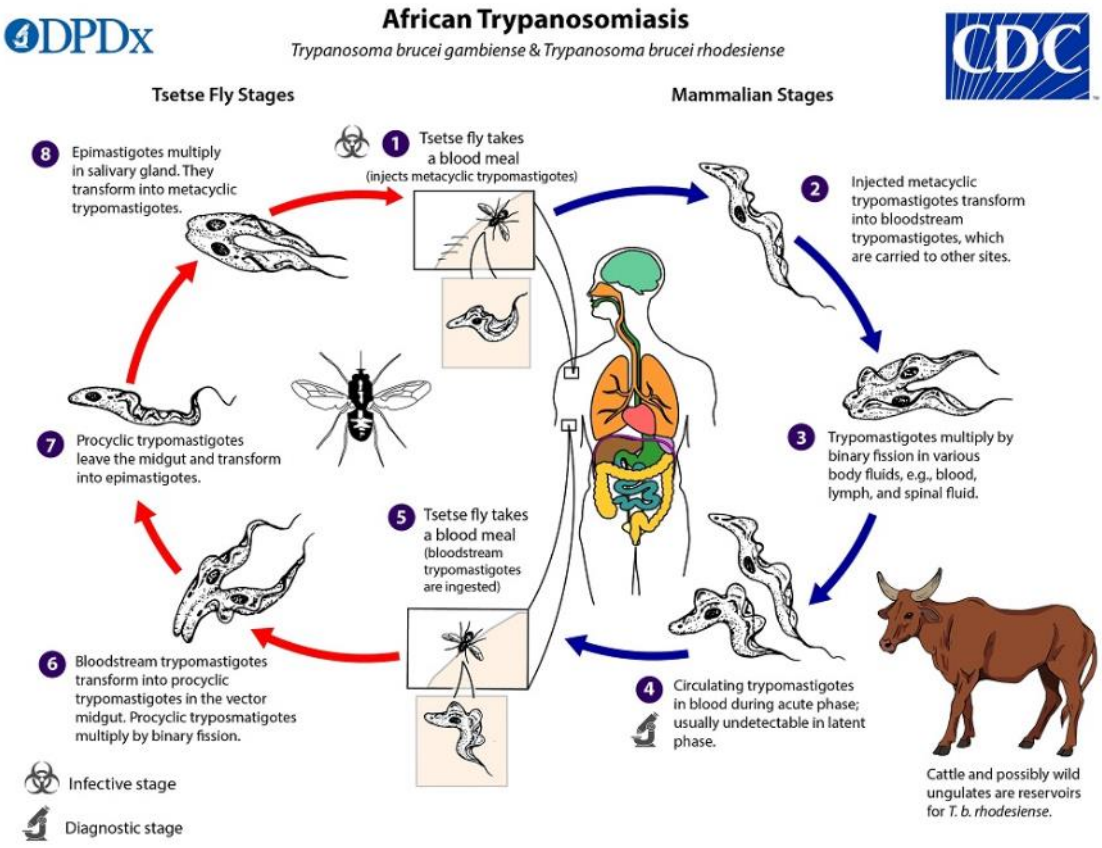


Figure 7: *Trypanosoma brucei* life cycle

This complex life cycle requires both an insect vector (tsetse fly) and mammalian host to complete its development. Taken from CDC (2019).

Only 1446 cases of HAT were reported in 2017 (WHO, 2019), however nagana still effects many domesticated and wild mammals, with an estimation of one quarter of Africa being unsuitable for livestock farming due to the parasite (Li *et al,* 2010; Troeberg *et al,* 2000). The first stage of the disease, when the parasites are present in the peripheral circulation of the mammalian host can cause fever, headaches, muscle and joint pains. The second stage is after the parasite invades the central nervous system causing deterioration, sleeping disturbances and other neurological symptoms leading to the death of the host (CDC, 2019).

Studies have shown that flagellum motility is essential in disease pathogenesis in the mammalian host and for parasite development in the insect vector. In the mammalian host, motility is required for migration purposes, to reach specific host

tissues and for traversing through the blood brain barrier. The flagellum is also required for immune-evasion, as motility directs all immunoglobulin-bound VSGs on the cell surface towards the flagellar pocket for clearance, protecting the cell from the hosts immune response (Engstler *et al,* 2007). The flagellum is required in the insect vector to attach to the salivary gland epithelium, allowing development into the mammalian infective form (Ralston *et al,* 2009).

1.7 *Trypanosoma brucei* are model organisms for flagellum studies

T. brucei are ideal models for eukaryotic flagella and cilia-based research. The single celled parasite contains one copy of each of its organelles, useful for tracking changes during investigative studies (Vaughan *et al,* 2003). The cell contains one constantly present and assembled flagellum (G1 cells), and a newly growing flagellum (dividing cells) not seen in mammalian cells, making this an ideal model for flagellum-based research (Langousis *et al,* 2014). The presence of a fully formed and a newly growing flagellum allows the identification of the age of BBs during cell division, which is useful when tracking BB proteins. *T. brucei* are also ideal models due to the availability of RNAi machinery, reverse genetics and an advanced genome project that encourages the use of this parasite in flagellum research.

1.8 *Trypanosoma brucei* cell cycle

The *T. brucei* cell cycle follows the same phases as other eukaryotic cells consisting of G0/G1, S, G2 and M phases (McKean, 2003; Hammarton, 2007). In the first stages of the cell cycle, the cell contains a single kinetoplast, nucleus (1K1N) and one basal body pair (BBP), the mature BB (BB1) and the pro-basal body (PBB) (BB2) (figure 9A) and a single flagellum (figure 8A). The first detectable event of entering the cell cycle is the maturation of the PBB, which docks to and invades the existing flagellar pocket (a small invagination of the plasma membrane), leading to the outgrowth of a new flagellum in S-phase (figure 8B; yellow arrow).

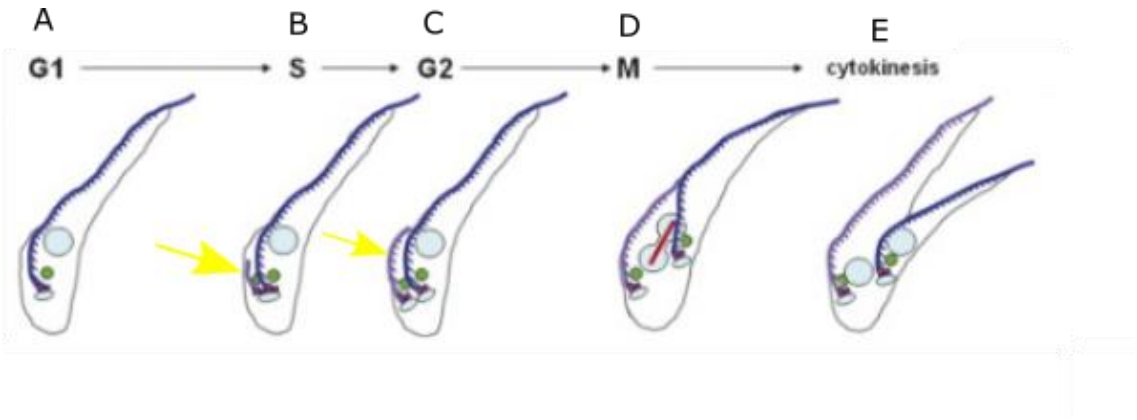


Figure 8: *Trypanosoma brucei* cell cycle

A) G1 (1K1N) cell with one kinetoplast, nucleus and basal body pair. B and C) The kinetoplast and basal body have begun duplication to form the 2K1N (early dividing) cells and has nucleated a new flagellum (yellow arrow). D) The nucleus divides, forming the late dividing (2K2N) cells. E) Cytokinesis after nuclear division. Edited from Hammarton (2007).

At S-phase the old mature BB (BB1) and newly mature BB (BB2) have formed PBBs orthogonal to themselves known as BB3 and BB4 (figure 9C). A flexible linker holding BB1 and BB2 together dissolves with two new flexible linkers attaching BB1 and BB3, and BB2 and BB4 together (figure 9D). During elongation of the new flagellum the BBPs perform an anti-clockwise rotation of the BB2 pair around the BB1 pair (figure 9E-F), required for correct morphological development of the cell, while the new flagellum continues to extend in length (figure 8B-C; yellow arrow). Prior to mitosis, the cell contains two kinetoplasts, two BBPs, two flagella and one nucleus (2K1N) (figure 8C). DNA division then begins and results in one cell containing two kinetoplasts, two BBPs, two flagella and two nuclei (2K2N) (figure 8D). Once mitosis is complete, cells undergo cytokinesis where two daughter cells are generated (figure 8E), with one cell inheriting the old BBP and one cell inheriting the new BBP.

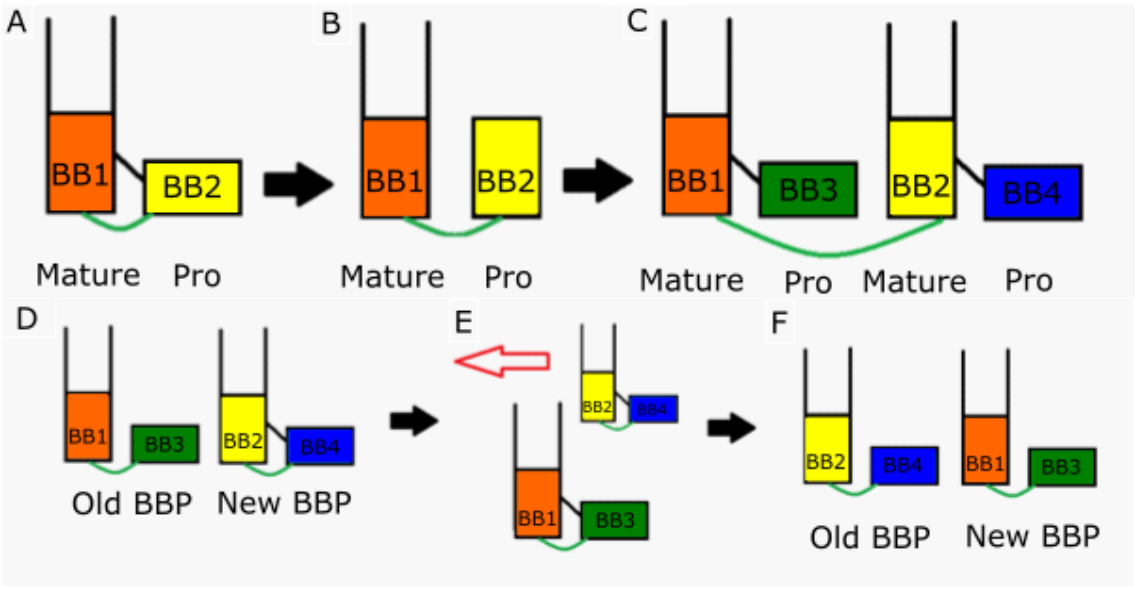


Figure 9: Basal body duplication cycle

A - B) Pro-basal body first orientates parallel to the mature basal body and docks to the plasma membrane, generating a new flagellum. C) Two pro-basal bodies form orthogonal to the mature basal bodies. D – F) The connection between BB1 and BB2 disintegrates and a rotation of the new basal body pair around the old basal body pair occurs.

In addition to nuclear division, the trypanosoma cell must also replicate and segregate its second genome, the kinetoplast DNA (kDNA). The organelle is unique to the kinetoplastid species and contains complex mitochondrial DNA. kDNA replication occurs prior to nuclear division and is thought to be essential for BB repositioning (Gluenz *et al,* 2010).

1.9 Basal bodies in the *Trypanosoma brucei* cell

Prior to entering the cell cycle, the cilium in mammalian cells is reabsorbed and the BB is removed from the plasma membrane, transforming into a centriole where it migrates near the nucleus to function in mitosis (Kobayashi and Dynlacht, 2011). This interchangeability between the BB/C does not occur in the *T. brucei* cell, instead they undergo a closed mitosis which does not require centrioles for nuclear division (Zhou *et al,* 2014).

Once the PBB matures, docks to the flagellar pocket and generates a new flagellum, the flagellum remains present and assembled at all times. As the flagellum is not reabsorbed, the BBs are constantly attached and are never dissociated from the flagellar pocket (Vaughan and Gull, 2016).

1.10 Docking of the basal body to the plasma membrane

BB docking is essential for the growth of the cilium or flagellum. For docking to occur, the BB must first mature and develop additional structures. Studies have shown that docking is mediated through pin-wheel structures locating to the distal ends of the BB (figure 5; B4) (Tanos *et al,* 2013). Previously mentioned (section 1.4), these additional appendages (DAs/TFs and SDAs/basal feet) are thought to form part of the ciliary gate, working with the ciliary TZ, allowing selected protein entry to and from the ciliary compartment (Tanos *et al,* 2013). Many proteins have been localised to the appendages, with research showing depletion of certain proteins lead to impairment of BB docking, suggesting they are essential structures for successful docking to the plasma membrane (Wei *et al,* 2015).

1.11 Cep164 localisation in mammalian cells

One protein thought to be essential for docking is the centrosomal protein 164 (Cep164). This protein localises to the distal section of each mature BB in mammalian cells and has been observed as a large ring resembling a pin-wheel pattern typical for distal appendage proteins (DAPs) (figure 10A; Cep164).

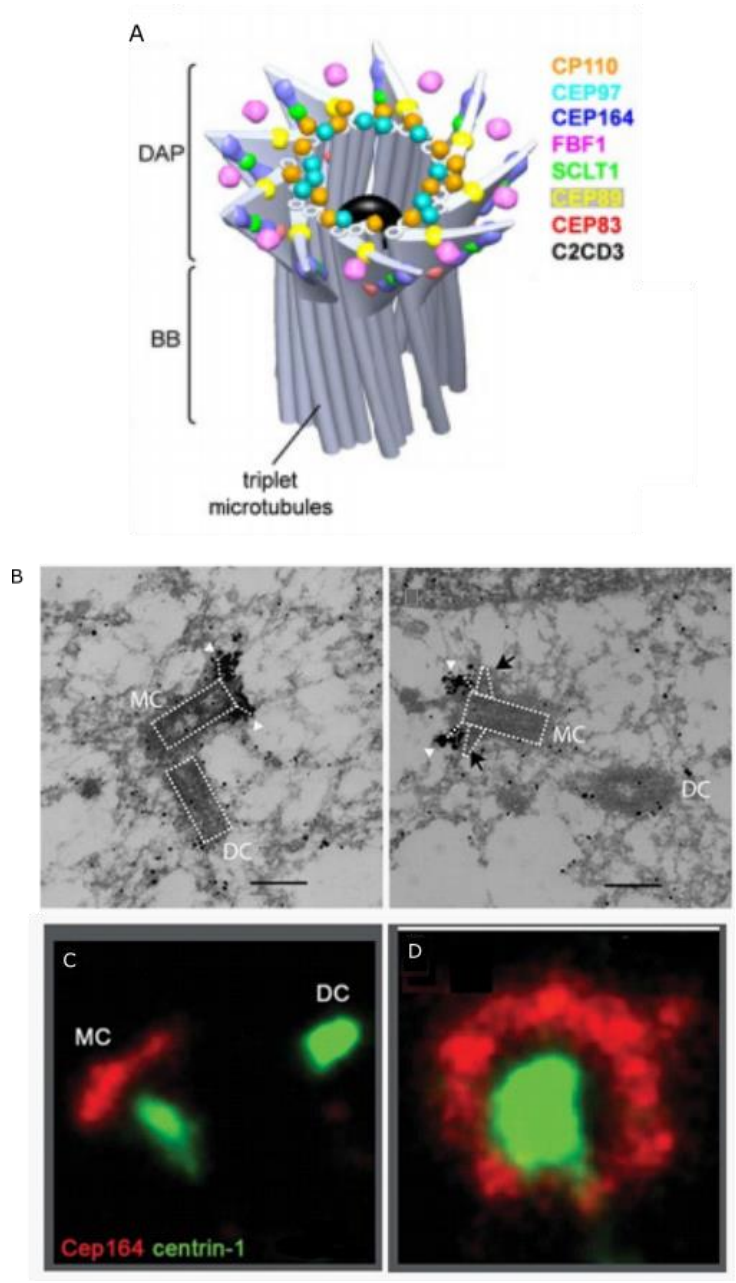


Figure 10: Cep164 localisation to the mother centriole

A) Diagrammatic representation of centriole proteins localising to the mother centriole. B) A micrograph of Cep164 localising to the distal section of the mother centriole through immunogold labelling. No Cep164 labelling was observed on the daughter centriole. C-D) Combined PALM/dSTORM imaging on HeLa cells expressing centrin-1 (green) and stained with an antibody for Cep164 (red). The Cep164 protein is seen to localise in a doughnut shape or as a bar-type pattern depending on the orientation of the centriole. A) Taken from Yang *et al* (2017). B) Taken from Graser *et al* (2007). C and D) Taken from Sillibourne *et al* (2011).

Cep164 has been localised to the DAs/TFs of the MC through immuno-gold labelling, with no labelling observed on the DCs (figure 10B) (Graser *et al,* 2007) and has been suggested to form the primary backbone of the DAPs (Yang *et al,* 2017). Studies using photo-activated localisation microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) demonstrated that the orientation of the centriole determines how the localisation of Cep164 is observed. Centrioles observed from the side showed Cep164 as a bar-like structure (figure 10C), whereas viewing from above produces a ring-like pattern, surrounding the triplet microtubules (figure 10D) (Sillibourne *et al,* 2011).

1.12 Cep164 function in mammalian cells

Cep164s localisation to the DAs makes this protein a candidate for anchoring the BB/C to the cell surface and an essential protein required for ciliogenesis. In mammalian cells, studies have shown that functions of the Cep164 protein involves the regulation of cilium formation through promoting trafficking during ciliogenesis, while aiding the docking of vesicles to the mother centrioles (Slaats *et al,* 2014).

Prior to docking to the plasma membrane, the BB first recruits vesicles, derived from the Golgi complex to the distal section of the BB (Joo *et al,* 2013). Once Cep164 is present it recruits the Chibby protein, tethering Golgi-derived vesicles to the distal section of the BB. This recruitment leads to a large merger of vesicles, forming a membranous cap known as the ciliary vesicle (CV) (Short, 2014) (figure 11; CV). The formation of the CV marks the transition from the centriole to the BB organelle (Sorokin*,* 1962; Schmidt *et al,* 2012; Kobayashi *et al,* 2011). Once the CV is formed, this allows fusion of the BB with the plasma membrane, allowing docking and resulting in ciliogenesis (Short, 2014).

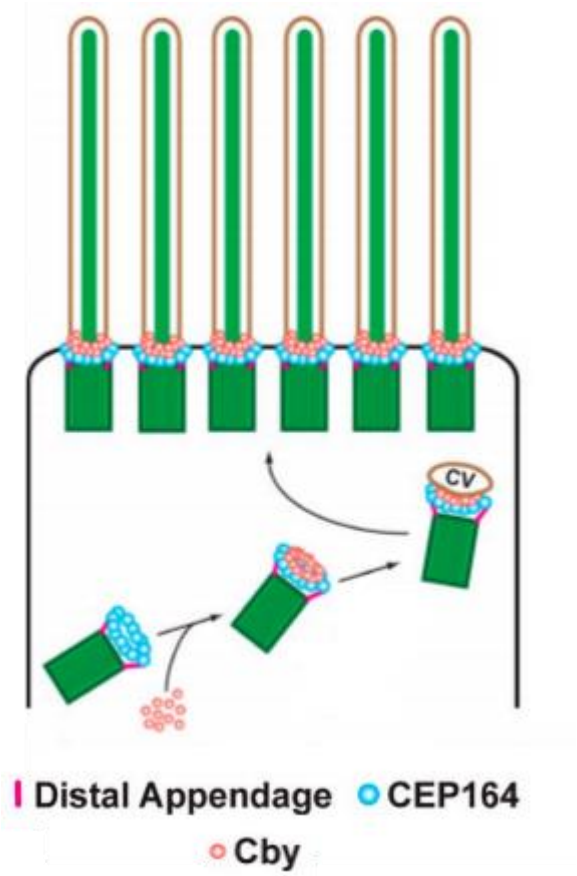


Figure 11: Basal body docking in mammalian cells

Cep164 at the distal end of the centriole recruits the Chibby protein (Cby), tethering Golgi-derived vesicles to the distal section and generating the ciliary vesicle (CV) needed for docking to the plasma membrane. Edited from Burke *et al* (2014).

While the BB is not docked to the plasma membrane a microtubule cap is present at the distal section, preventing extension of the axoneme (Cajanek *et al,* 2014). Cep164 in mammalian cells functions to recruit and form a complex with the Tau tubulin kinase 2 (TTBK2) protein, removing the microtubule cap (CP110 protein) and promoting axoneme formation while aiding in assembling distal appendages (Cajanek *et al,* 2014).

1.13 Consequences of loss of function of Cep164 in mammalian cells

Due to the importance of the Cep164 protein, studies have investigated the consequences in the loss of function in eukaryotic cells. RNAi techniques have suggested this protein is indispensable for ciliogenesis (Andersen *et al,* 2003; Chaki *et al,* 2012) and cells lacking this protein have a high failure rate of cilium formation (Graser *et al,* 2007).

In mammalian cells Cep164 is a cell cycle dependent protein, with protein being present at the MC during interphase but has a drastic reduction during mitosis (Graser *et al,* 2007). Loss of the Cep164 protein through siRNA accelerated the cell cycle, while slightly delaying the cell entering the S-phase (Slaats *et al,* 2014). Cells lacking Cep164 also had an increase in apoptosis and signs of DNA damage, although this topic is controversial. Sivasubramaniam *et al* (2008) stated that Cep164 located to the nucleus and only a fraction localised to the centrosome, whereas other studies have observed a consistent localisation of the Cep164 protein to the centrosome only (Daly *et al,* 2016).

1.14 The Cep164 protein and disease

Interest in the Cep164 protein is also related to the disease it causes in mammals. As cilia are present on almost all cell types in the human body, mutations in the cilium or flagellum lead to cilia-related disorders termed ciliopathies that affect many organ systems (Fliegauf *et al,* 2007). In humans, mutations in 19 different proteins cause the ciliopathy nephronophthisis, characterised by the progressive wasting of the filtering unit of the nephron (Badano *et al*, 2006). Studies have shown that four mutations in the Cep164 gene contributes towards the development of this disease (Stockman *et al,* 2016; Daly *et al,* 2016).

1.15 Conservation of the Cep164 protein

Cep164 is a conserved protein that is present and transcribed in organisms outside of mammals, including the zebrafish (*Danio rerio*), the African clawed frog (*Xenopus laevis*) and the fruit fly (*Drosophilia melanogaster*). Organisms that do

not possess a cilium such as *Arabidopsis thaliana* (mouse-ear cress) and *Saccharomyces cerevisiae* (brewers yeast*)* lack the Cep164 protein, potentially as no BB is required to dock to the membrane. Interestingly, not all organisms transcribe one orthologue of the protein. Some species including *T. brucei* and *Leishmania* transcribe three Cep164 proteins. However other organisms such as *Trichomonas* and the *Paramecium* species transcribe four Cep164 proteins (Hodges *et al,* 2010).

1.16 Building of the cilium or flagellum

Although the BB anchors the cilium or flagellum to the cell membrane, other processes and structures are required for ciliogenesis. Transitional fibres at the distal section of the BB act as the main attachment point to the ciliary membrane (Reiter *et al,* 2012) and is located between the internal and external cell environment (figure 12; TF). This allows the TFs to act as a physical block, as the spaces between the fibres are too small to allow entry of proteins or vesicles (Geimer and Molkonian, 2004). The transition zone is another essential structure for cilium development and is located at the proximal section of the axoneme. It is characterised by the Y-shape linkers that connect the microtubules to the ciliary membrane (figure 12: TZ) (Reiter *et al,* 2012).

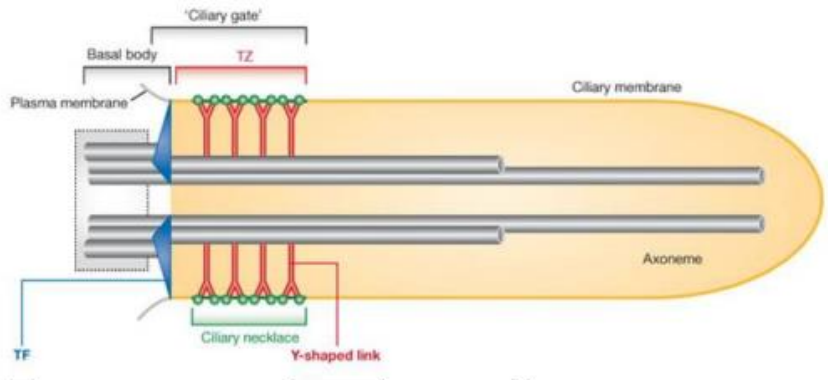


Figure 12: Structures forming the ciliary gate

The transitional fibres (TFs) and transition zone (TZ) form the ciliary gate, which regulates protein and vesicle entry into the developing cilium. Edited from Reiter *et al* (2012).

Both the TFs and TZ form the ciliary gate (figure 12), an important structure for ciliogenesis and for control of protein and vesicle entry into the cilium (Reiter *et al,* 2012). Although the ciliary gate regulates particle entry into the cilium, proteins cannot be synthesized *de novo* from within the cilium and must be transported into the organelle from the cell body (Reiter *et al,* 2012; Ishikawa and Marshall, 2017). Located by the ciliary gate is a docking site for intraflagellar transport (IFT) particles. This bidirectional protein transport system functions in transporting proteins (tubulins and others) along the microtubules of the cilium using molecular motors, building the cilium from the distal tip (Ishikawa and Marshall, 2017). The IFT system consists of two large complexes, termed IFT-A and IFT-B. IFT-B transports particles anterograde (towards the distal tip of the cilium), delivering particles used for cilium growth (figure 13; IFT complex B). IFT-A transports particles retrograde (towards the base of the cilium) and returns proteins to the cell body (figure 13; IFT complex A) (Ishikawa and Marshall, 2017).

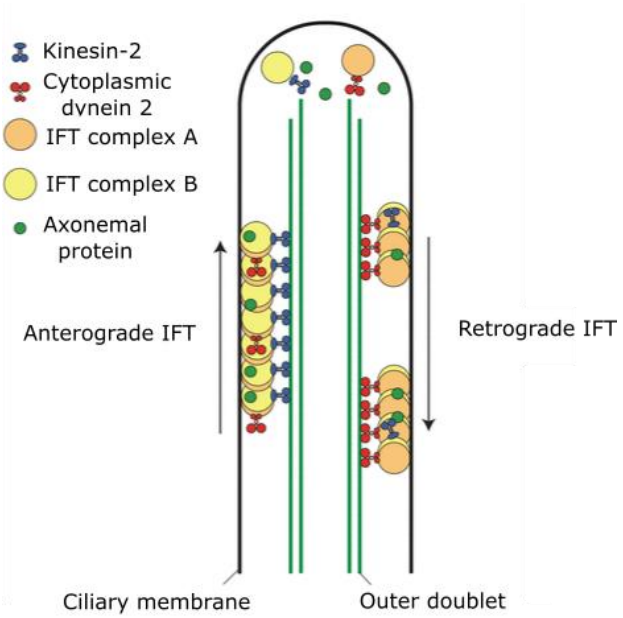


Figure 13: Intraflagellar transport system

The intraflagellar transport system (IFT) works to transport proteins (tubulin and others) to build the cilium. The IFT-B transports proteins anterograde to the distal tip, to build the cilium. Whereas IFT-A transports proteins retrograde, returning the proteins to the base of the cilium. Taken from Ishikawa and Marshall (2017).

The length of cilia and flagella differ in each cell type with cilia often being shorter and more numerous in cells than flagella. Cilia lengths differ not just in different organisms but also in each cell type, with some cells generating cilia between 4-9μm in length (Dummer *et al,* 2016). Flagella are often longer in length and are less abundant than cilia. The *T. brucei* new and old flagellum in a dividing cell differs in length, with the new flagellum averaging at 15μm, whereas the old flagellum measures at approximately 20μm. The new flagellum is shorter as development is not completed until after cellular division (Bertiaux *et al,* 2018; Farr and Gull, 2009).

1.17 Manipulation of the flagellum length in the *Trypanosoma brucei* cell

Research in *Chlamydomonas* and *T. brucei* cells have been carried out to investigate flagellum length regulation (Tuxhorn *et al,* 1998) and IFT entry into the flagellum (Dentler *et al,* 2005). Studies focusing on the IFT system have shown the importance of these particles in successful building of the flagellum. RNAi experiments of IFT-B complexes have led to shorter flagella by a reduced number of IFT particles entering the flagellum, whereas IFT-A generates shorter flagella with excess material collecting at the distal tip (Absalon *et al,* 2008; Fort *et al,* 2016).

Flagellum length influences morphology and viability in the *T. brucei* cell. In dividing cells, the length of the new flagellum and the flagellum attachment zone (FAZ) filament (cytoskeletal and membranous connection holding the flagellum and cell body in close proximity) determines where cytokinesis is initiated, influencing the length of the daughter cells (Rotureau *et al,* 2014). Cells which produce shorter flagella (IFT experiments) can carry out cellular division but generate smaller new daughter cells. However, those which are unable to produce a new flagellum are non-viable (Ralston and Hill, 2008). Bertiaux *et al* (2018) investigated the regulation behind flagellum extension in *T. brucei*. It was suggested that to regulate flagellum length, a locking event must take place once the flagellum has fully extended, to prevent the organelle from shortening or lengthening in following cell cycles. This “locked” stage is triggered prior to cell division and is controlled at the cell-cycle

level to ensure the new flagellum is in the correct location for cytokinesis (Bertiaux *et al,* 2018). The study suggested that alteration of the growth rate or timing of the locking event could lead to the development of flagella of different lengths.

Project Aims

The aim of this project is to screen proteins donated from the TrypTag project at Oxford University to identify basal body proteins in the *T. brucei* parasite. Proteins taken forward will be investigated to understand their role in basal body biogenesis or maturation.

The aims of researching the Cep164 proteins in *T. brucei* are to observe the function of this protein in relation to the mature basal body. Cep164 in mammalian cells is known to recruit Golgi-derived vesicles to the distal section of the BB to help dock the BB to the plasma membrane, and to recruit the TTBK2 protein for release of the microtubule cap, allowing ciliogenesis. Currently, little information is known about the three Cep164 proteins and their function in *T. brucei*. Characterisation will reveal whether these proteins function in docking the BB to the flagellar pocket, finalise maturation of the transitional fibres or whether the proteins act as a marker to identify the old mature basal body. The specific objectives of this work were to:

- Identify those proteins which localise to the basal bodies and are of interest in basal body biogenesis or maturation
- Characterise the three Cep164 proteins in *T. brucei* and uncover their function to the basal body

2. Methodology

2.1 Methodology for bioinformatics and the basal body proteome

2.1.1 Obtaining and identifying basal body proteins through TrypTag

All proteins suggested to localise to the basal body were donated through the TrypTag project; the *Trypanosoma brucei* genome-wide resource (Dean *et al,* 2017). An excel document was generated for the 267 proteins and information such as the localisation pattern, PFAM domains and orthologues in other species was collected.

2.1.2 Databases used for proteome analysis

Genomes of various organisms that were fully sequenced were used in this study to identify those basal body proteins conserved outside of the kinetoplastidae species, or whether the proteins are Trypanosoma specific.

Online databases used in this study include:

- TrypTag (Dean *et al,* 2016) World Wide Web URL: http://tryptag.org/
- TriTrypDB, Kinetoplastid genomics resource (Aslett *et al,* 2010) World Wide Web URL: http://tritrypdb.org/tritrypdb/
- National Centre for Biotechnology Information (NCBI) GenBank (Benson *et al,* 2013) World Wide Web URL: https://blast.ncbi.nlm.nih.gov/Blast.cgi
- InterPro protein sequence analysis and classification (Finn *et al,* 2017) World Wide Web URL: https://www.ebi.ac.uk/interpro/
- PFAM EMBL-EBI (Finn *et al,* 2016) World Wide Web URL: https://pfam.xfam.org/
- HMMER search sequence(s) against a profile database (Finn *et al,* 2011) World Wide Web URL: http://hmmer.org/

- PlantGDB resources for comparative plant genomics (Duvick *et al,* 2008) World Wide Web URL: http://www.plantgdb.org/
- Tetrahymena genome database wiki (Eisen *et al,* 2006) World Wide Web URL: http://ciliate.org/index.php/home/welcome
- TrichDB Trichomonas genomics resource (Aurrecoechea *et al,* 2009) World Wide Web URL: http://trichdb.org/trichdb/
- XenBase (Karimi *et al,* 2018) World Wide Web URL: http://www.xenbase.org/entry/
- ZFIN a zebrafish information network (Howe *et al,* 2013) World Wide Web URL: https://zfin.org/
- SGD Saccharomyces genome database (Cherry *et al*, 2012) World Wide Web URL: https://www.yeastgenome.org/
- TAIR the Arabidopsis information resource (Berardini *et al,* 2015) World Wide Web URL: https://www.arabidopsis.org/index.jsp
- WormBase facilitating insights into nematode biology (Stein *et al,* 2001) World Wide Web URL: https://www.wormbase.org/#012-34-5
- PlasmoDB plasmodium genomics resource (Aurrecoechea *et al,* 2008) World Wide Web URL: http://plasmodb.org/plasmo/
- FlyBase A database of Drosophila genes and genomes (Gramates *et al,* 2017) World Wide Web URL: http://flybase.org/

2.1.3 Identifying protein orthologues

Basic local alignment tool (BLAST) searches were conducted for each organism (table 1) and each of the 267 proteins suggested to localise to the *T. brucei* basal body. The amino acid sequences of the 267 proteins in *T. brucei* was taken and a BLAST search was run in the organisms database, that amino acid sequence was then taken and run in the *T. brucei* database to ensure protein identification. The

EMBOSS needle tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) was used to analyse the identity and similarity values of two protein sequences. The Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) multiple sequence alignment tool on EBI was used to analyse the identity and similarity values of multiple protein sequences.

Table 1: Genome references of organisms used in this study

12 genome databases were used for bioinformatic and conservation analysis in this study.

Organism	NCBI Taxid	Genome reference
Chlamydomonas reinhardtii	3055	Merchant, S.S *et al* (2007)
Tetrahymena thermophila	5911	Eisen *et al* (2006)
Trichomonas vaginalis	5722	Aurrecoechea, C *et al* (2009)
Danio rerio	7955	Howe *et al* (2013)
Xenopus laevis	8355	Karimi *et al* (2018)
Mus musculus	10090	Mouse Genome Sequencing *(*2002)
Homo sapiens	9606	Venter *et al* (2001)
Saccharomyces cerevisiae	4932	Cherry *et al* (2012)
Arabidopsis thaliana	3702	Arabidopsis Genome (2000)
Caenorhabditis elegans	6239	C. elegans Sequencing Consortium (1998)
Plasmodium falciparum	5833	Gardner *et al* (2002)
Drosophila melanogaster	7227	Adams *et al* (2000)

2.1.4 Identifying protein domains

Amino acid sequences for each of the 267 suggested basal body proteins were inserted into Interpro (Finn *et al,* 2017) and PFAM (Finn *et al,* 2016) databases, to identify any known domains of interest using the sequence search tool.

2.2 Methodology for molecular biology

2.2.1 Plasmid generation and primer design for endogenous tagging and RNA interference

To generate cell lines, a plasmid containing the monomeric neon green (mNG) fluorophore with the addition of TY1 tags was constructed. The original pPOTv4 plasmid containing the mNG fluorophore, blasticidin resistance marker and BamH1 restriction sites (figure 14) was amplified in *Escherichia coli* cells, where bacterial colonies were chosen and purified. The plasmid was then digested by BamH1 and dephosphorylated in preparation to insert the TY1 tags, with the tags being phosphorylated prior to ligation to either side of the plasmid. The plasmid transformed into *E. coli* cells developed cultures on agar plates containing ampicillin. Mini-preps were generated, digested with BamH1 and were run through gel electrophoresis. Those bacterial cultures confirmed to contain the modified plasmid (through sequencing) were put through a midi-prep and used to generate mutant cell lines.

Endogenous expression of each protein tagged in this project was generated using the modified pPOTv6 mNG blasticidin resistance plasmid containing TY1 tags. Fluorescent tagging of proteins was carried out through a polymerase chain reaction (PCR) based tagging system, where the template for the PCR is a plasmid instead of gDNA. This technique used in this project known as the PCR only tagging (pPOT) system was developed by Dean *et al* (2015).

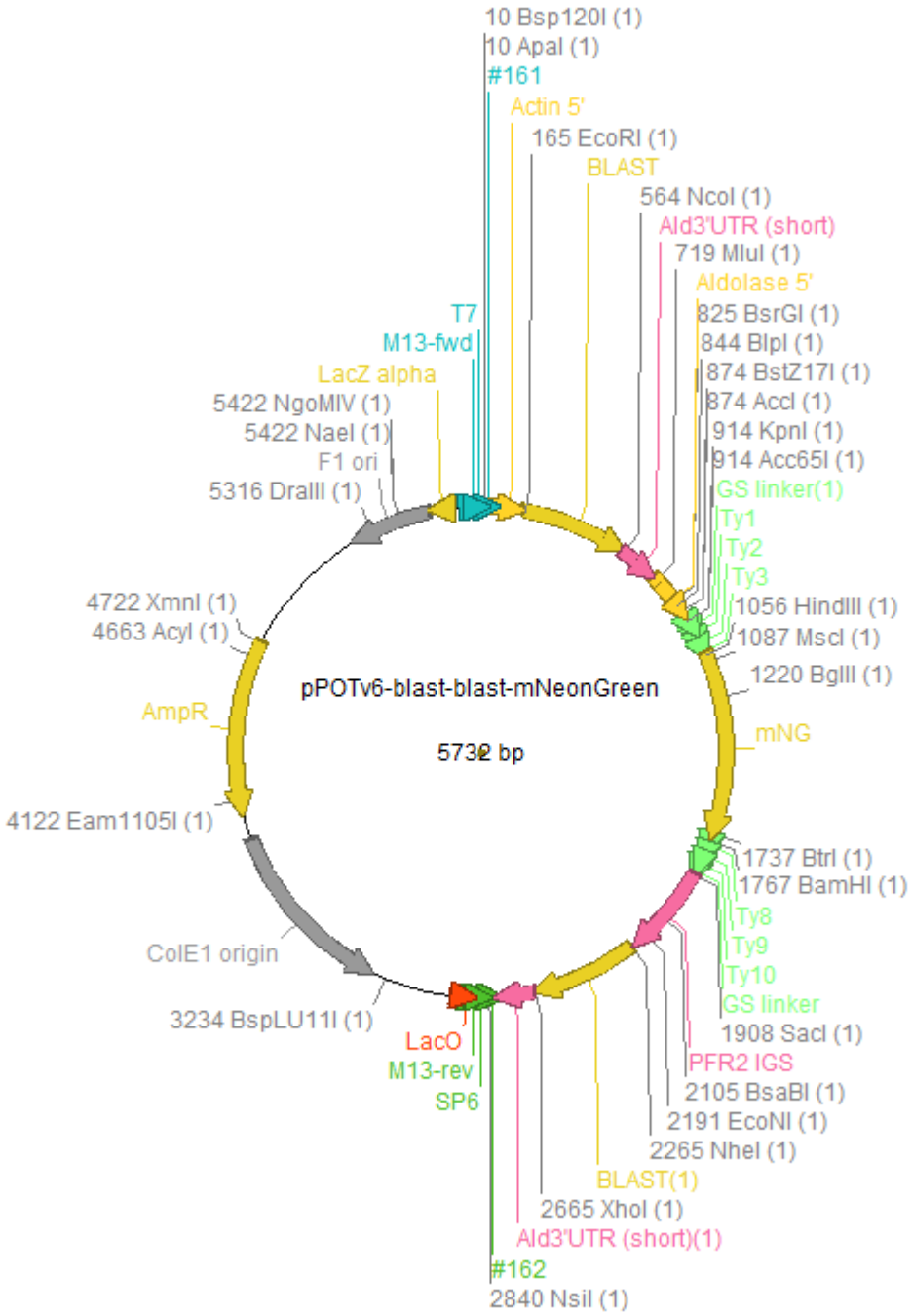


Figure 14: Modified pPOTv4 plasmid map

The plasmid map of the modified pPOTv4 plasmid, now referred to as the pPOTv6 plasmid containing the mNG fluorophore, blasticidin resistance marker and TY1 tags for long primer PCR.

Generation of the long PCR primers for the pPOT technique were produced by the Perl script primer design tool accompanying the plasmid template for pPOT. Long PCR primer sequences for the N and C terminus was kindly donated by Jack Sunter (Oxford University) (table 2). These were required for long primer PCR amplification to allow homologous recombination of the amplified product at the loci of interest. All pPOT primers were ordered from Invitrogen (Paisley, UK).

Table 2: N and C primer sequences for endogenous tagging generated by the Perl script primer design tool.

Primer Name	Primer Sequence (5' – 3')	Gene ID	Template
3NTb927.5.1120F	TAAACACACACACACACACACGCACACACACAAAAGTGTATATTTATTTATATATTTAGGTCGCCACGGTACAATTAAgtataatgcagacctgctgc	Tb927.5.1120	pPOTv6
3NTb927.10.2860F	CTAAGGGAGAACGTCACATATAATATTCTGTACAGTCGCCTCACCCGACATAGGGAGTGAGAAGGGACACGGCCACGCCCgtataatgcagacctgctgc	Tb927.10.2860	pPOTv6
3NTb927.10.3130F	AAATACTGAAAAGAGGAACATATATACATTAACACTAACGAAGAAAGAAAGAGACGCTCAAAAGGGGGCAATATCAGGGGgtataatgcagacctgctgc	Tb927.10.3130	pPOTv6
3NTb927.7.4130F	TTCCTGTTTGTTTACAAGGTAGGGTTGTCGCATAGCTTGAGACAATAGCCGTGCTGTGAAGGGGACCCAAATAAGCAACGgtataatgcagacctgctgc	Tb927.7.4130	pPOTv6
3NTb927.8.4210F	CAAGGTTTAATCTCATTTACTTTATATATATTTATATCCGACTTCGTTTAACTCTTCCCCCACATCCACTTACTGAACCCgtataatgcagacctgctgc	Tb927.8.4210	pPOTv6
3NTb927.11.5030F	TGCATCGATTCCTATTACTGCTAGCTTTTGCTCGTTGTCGGTGCCATCACCGTTGCTGCCGCGCAACTTCGTTTTTGGCAgtataatgcagacctgctgc	Tb927.11.5030	pPOTv6
3NTb927.7.7200F	GTTTCCCCTAAAACCACTGATTGCTACATTGTTGCCTTTATCTTTTTCAATAATTCACACTATCACGCAGGGACATTTGTgtataatgcagacctgctgc	Tb927.7.7200	pPOTv6
3NTb927.9.7720F	GAAAAAAAGTGGAAGTTAAATATTGATAAGCACAACTACAAGTGAGATAAGGGTTCGAAGCTGATCGGTGTTGTGAGGCAgtataatgcagacctgctgc	Tb927.9.7720	pPOTv6

3NTb927.10.10090F	CGTGTGCTTGACGCGATAAGAAGGGTAGGGCACATATCAATCACAGATTTAGGAGGTTTGACACTGTTGGAGACAAAGTGgtataatgcagacctgctgc	Tb927.10.10090	pPOTv6
3NTb927.10.12590F	CTTCCTCACAATCATTCCCCCTCCCCTATTTGCATTCTCACACAAGGGTACTTAAAAAGGAAGACAAATAAAACGTAACTgtataatgcagacctgctgc	Tb927.10.12590	pPOTv6
3NTb927.11.13300F	ATTTTTAAAAAAAAAGGAGAAAGATCGTGGAAATCAAGAGAAATAAAACTTAAAAGAAAAAAGGAGAGGGAGAAGTGCGAgtataatgcagacctgctgc	Tb927.11.13300	pPOTv6
3NTb927.9.14300F	GGCAGCAGCTGTGCGGTGTCGGTGGCGACCACCTGAACAAACCTTGTAGGTTTCATTTGAGGAAGACGGCGTAGTTTGGCgtataatgcagacctgctgc	Tb927.9.14300	pPOTv6
3NTb927.10.14770F	GTAAGAAAGCAAAAAAAGGCAGCGTTGCTTCCCTTTTATATACCAGTGATAATCTCTCTGCGTTGGGGTTTCTTATTGAAgtataatgcagacctgctgc	Tb927.10.14770	pPOTv6
3NTb927.11.15450F	AGGAGAGCGTCGTAGCGGCGGTTCCTCCAAGGTGCCACCTCCCCCTACACTTATCGTCGCCGTCAAACAGAAAAAGGAAAgtataatgcagacctgctgc	Tb927.11.15450	pPOTv6
3NTb927.11.2700F	TGTGTGTGTCTTCTAGTTTTTGTCCTGCTGCTGAAGCAGTCATCAGTTGCTCCTCTACAGTTGTTGCCGGTTTCGCAGAAgtataatgcagacctgctgc	Tb927.11.2700	pPOTv6
3NTb927.7.5190F	GAAAGGAAGCATCTATAAAAGAAGCATACGGAAAATACATCCATAAATCAATTTATACATTTGCAACTGTTGTATCACTGgtataatgcagacctgctgc	Tb927.7.5190	pPOTv6
3NTb927.10.13610F	CGTGCAACTGCGGACTAACCCGCGTGAACACAGTTTAGTTTTGCGGCGCATCACGTGTATATTATTTAGATTGTCAGGGGgtataatgcagacctgctgc	Tb927.10.13610	pPOTv6
3NTb927.10.4990F	GGTTGCTACATAAAATTTTTATTTTTTGTCTCTGCTTTCCTCTTAAGTTCTTCAGGAAACGTTAGGTTGAAGGAGGAGATgtataatgcagacctgctgc	Tb927.10.4990	pPOTv6
3NTb927.11.3560F	CTTGGCTGCTCAGGGAATTTGTTCAACCTTTGGAAGCGGCTGTTAAACTAAGGTAAGAACCACCTTTAACGCTCCCCGGTgtataatgcagacctgctgc	Tb927.1.3560	pPOTv6

3NTb927.5.2440F	GGAAACAATGAAGACTTTTTTTTGTTGTGGTTGGTTTTTTAACGGAAGGTGTCGAGCGTTTAGGTTGATAGAAGGGCCACgtataatgcagacctgctgc	Tb927.5.2440	pPOTv6
3NTb927.11.11650F	GGCAAGCGTAGTTCCACCACGTAGTGCAGAGGGTTAGTAGCATCCCTCTATTCCTTACCGTTTTTGTTACTCGGCTGAGAgtataatgcagacctgctgc	Tb927.11.11650	pPOTv6

RNAi cell lines of Cep164A, Cep164B and Cep164C were generated using the pQuadra RNAi plasmid (Kalidas *et al,* 2011) which requires one plasmid to generate the stem loop (figure 15) and one plasmid being the vector backbone (figure 16).

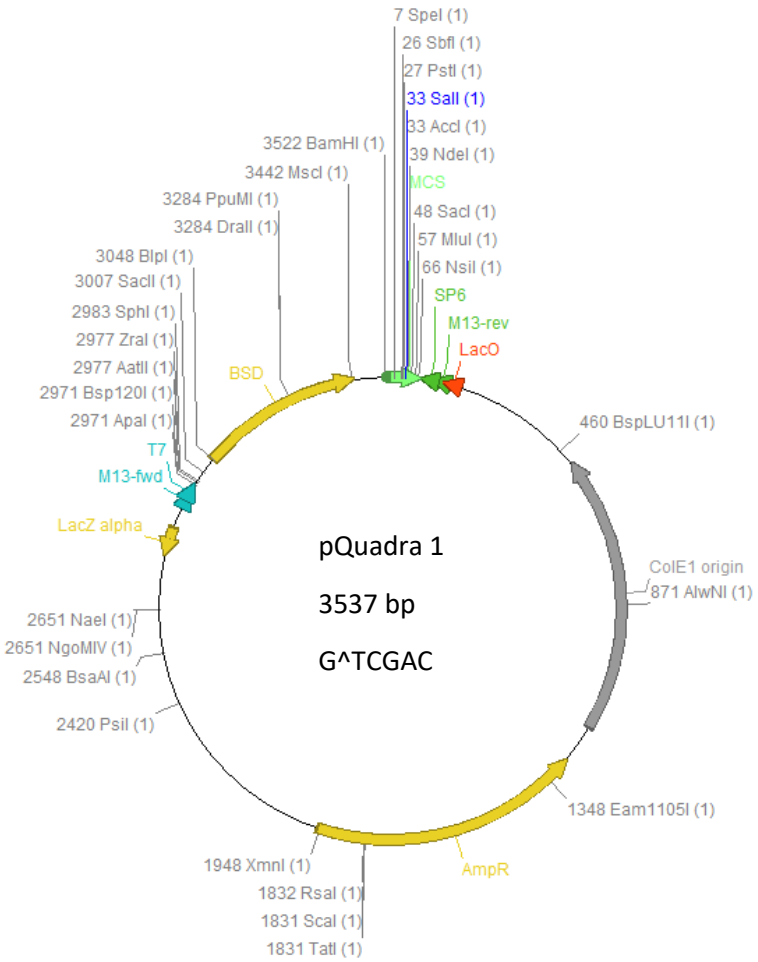


Figure 15: The plasmid map of the pQuadra stem-loop used for inducible RNAi

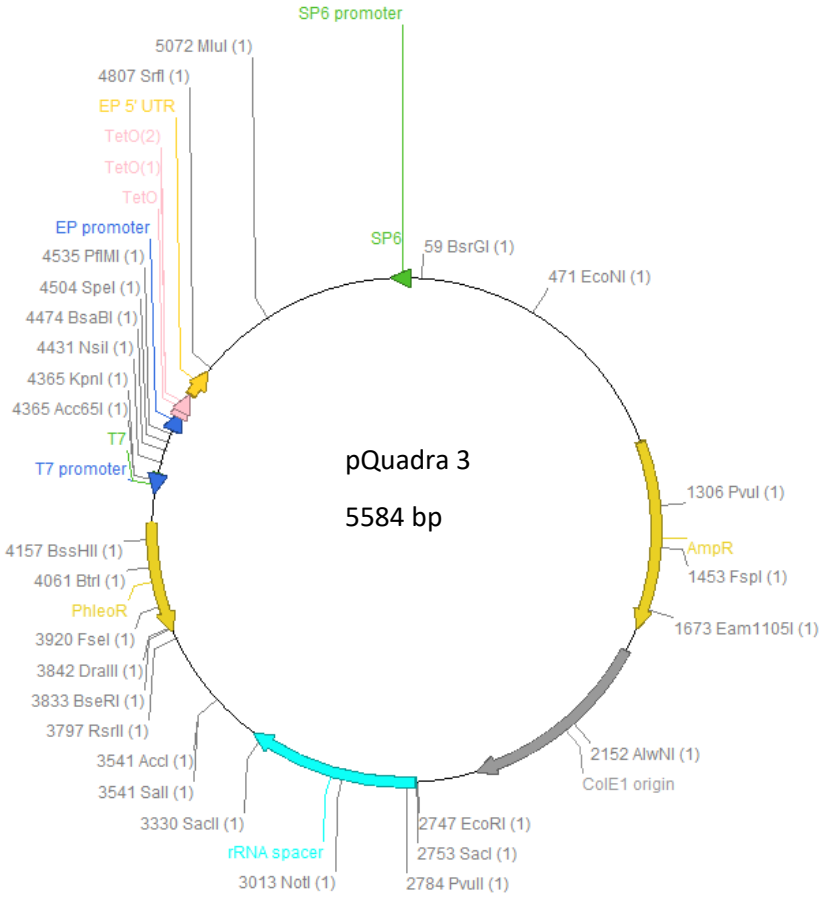


Figure 16: The plasmid map of the pQuadra backbone used for inducible RNAi

Generation of the pQuadra plasmid containing the PCR amplicons involves a four-way ligation assay, allowing construction of the vector in one cloning step (Park *et al,* 2014). Gene-specific primers for Cep164A, Cep164B and Cep164C were generated by taking the open reading frame sequences from TriTrypDB and inserting them into RNAit (Redmond *et al,* 2003). This generated target specific primers (table 3) which, through PCR, amplified a specific region of the open reading frame from gDNA. PCR amplicons were cloned into the pQuadra plasmid using the BstX1 and HindIII restriction sites. All RNAi primers were ordered from Eurofins Genomics (Wolverhampton).

Table 3: Inducible RNAi primer sequences for Cep164A, B and C generated by inserting the open reading frame sequences into RNAit

Primer Name	Primer Sequence (5' – 3')	Gene ID	Template
2440 RNAiQ FW	Forward: ataccaatgtgatggACCGCGCGAGAAGATTAAA	Tb927.5.2440 (Cep164A)	pQuadra
2440 RNAiQ RW	Reverse: ataccatagagttggGCTTCTCGCTACCGTTTGTC	Tb927.5.2440 (Cep164A)	pQuadra
11650 RNAiQ FW	ata**cca**atgtga**tgg**GCAACTTCTTGCTGAGGACC	Tb927.11.11650 (Cep164B)	pQuadra
11650 RNAiQ RW	ata**cca**tagagt**tgg**GCCTCACCAACTGCTTCTTC	Tb927.11.11650 (Cep164B)	pQuadra
3560 RNAiQ FW	Forward: ataccaatgtgatggTCTGAAACGCATTTGTCTGC	Tb927.1.3560 (Cep164C)	pQuadra
3560 RNAiQ RW	Reverse: ataccatagagttggGCACAGCACAGAGGTTGAAA	Tb927.1.3560 (Cep164C)	pQuadra

2.2.2 Polymerase chain reaction

To generate long primer PCR amplicons for direct transfections into the *T. brucei* cells, the expand high fidelity PCR system (11732650001, Roche) was used according to the published protocol (Dean *et al,* 2015). For amplification, the thermocycler was heated to 100^{o}C to prevent sample evaporation and PCR conditions were set as follows; template DNA was denatured at 95^{o}C for 3 minutes then 34 cycles of further denaturation for 30 seconds at 95^{o}C was set. Annealing of the primers was carried out at 60^{o}C for 30 seconds with an elongation step at 72^{o}C for 2 minutes. After the 34 cycles a final elongation step at 72^{o}C for 5 minutes was carried out.

Amplicons to clone into the pQuadra RNAi plasmid were generated using a PCR system with the Q5 HiFi DNA polymerase (M0491S; New England Biolabs, Hitchin, UK). 0.25µl of gDNA was added into a 25µl reaction containing 5µl of 5x Q5 reaction buffer, 1µl of forward and reverse primers, 0.5µl of 10mM dNTPs and 0.25µl of Q5 DNA polymerase, the rest of the volume was made up of distilled water. The PCR conditions were set as follows; initial denaturation was conducted

at 98°C for 2 minutes, with 1µl of polymerase added after the lid temperature reached 100°C. A further denaturation step was carried out at 98°C for 10 seconds, annealing of the primers was done at 60°C for 1 minute and an elongation step at 72°C for 1 minute was carried out for 5 cycles. Primers were then added to the solution and a further denaturation step at 98°C for 10 seconds was carried out, annealing was done at 60°C for 30 seconds and elongation at 72°C for 1 minute occurred for 35 cycles.

2.2.3 Agarose gel electrophoresis

Products amplified by PCR were analysed through agarose gel electrophoresis. Amplicons were run on a 1% agarose gel by melting electrophoresis grade agarose (Fisher scientific) in 1x TAE buffer (Tris-acetate-EDTA 40mM Tris (pH7.6), 20M acetic acid, 1mM EDTA) which was heated until boiling. Once cooled, SYBR Safe (0.5µl of a 10,000units/400µl) was added to allow visualisation of DNA using a blue light in a transilluminator (Syngene GBOX), with the image taken through the GeneSys software (California, USA). The agarose was cooled slightly before poured into a sealed gel tray with a multi-well comb prior to setting of the gel. Once set, the gel was placed into the electrophoresis tank (Bio-Rad, California, USA) flooded with 1 x TAE buffer until the surface of the gel was covered. Either a 1Kb DNA ladder (NEB N0468S) or 100bp DNA ladder (NEB N0467S) was loaded onto the gel next to the amplicons (1µl), where the amplicons were loaded with a coloured (NEB B7024S; NEB B7022S) loading dye. The gel was run at 50 volts for 5 minutes to ensure all products entered the gel, it was then run at 80 volts for 1-2 hours to allow separation of the DNA fragments.

Plasmid digests were also analysed through agarose gel electrophoresis at a 0.7% gel under similar conditions as previously described. Visualised bands were cut from the agarose gel using a sterile scalpel. DNA extracted from the gel was purified from the gel slice using the Monarch nucleic acid purification (NEB) kit according to manufacturer's instructions. DNA extracted was analysed by the NanoDrop spectrophotometer (ND-100, LabTech UK), calibrated with distilled water.

2.2.4 DNA purification and sequencing

Amplicons were purified using the Monarch PCR and DNA clean up kit (T1030L, NEB) according to manufacturer's instructions.

2.2.5 Restriction enzyme digest

Each DNA sequence digested had an excess of the appropriate restriction enzyme added to ensure efficient digestion, with all enzymes used with their supplied buffer (table 4). PCR products or plasmids were digested for 2-3 hours at 37^{o}C in either a water bath or 37^{o}C incubator. PCR products digested were purified through the PCR purification kit (Monarch), or extraction of either products or plasmids was carried out using the Monarch DNA gel extraction kit (T1020S, NEB) protocol after agarose gel electrophoresis.

Table 4: Restriction enzymes

Restriction enzymes used for digesting, ligating and linearising plasmids used in this study.

Enzyme	Concentration (units)	Catalogue number	Supplier
BamHI	20,000	R3136L	New England Biolabs
HindIII	20,000	R3104S	New England Biolabs
NotI	20,000	R3189L	New England Biolabs
Bstx1	10,000	R0113L	New England Biolabs

2.2.6 Ligation of PCR products into plasmids

PCR products were directly ligated into the pQuadra vector. A solution of approximately 40ng of the vector backbone and a minimum of 80ng of insert DNA was made. Ligations were made using the T4 DNA ligase (M180A, Promega) with the ligation buffer (C671A, Promega), and were carried out at 4^{o}C overnight.

2.2.7 Bacterial transformation

Amplification of the ligated PCR product and plasmid was carried out through bacterial transformation using *E.coli* SURE competent cells (200238, Agilent Technologies). The ligation mixture was first chilled on ice and 35µl of competent cells were added to the mixture once defrosted and placed back on ice. The cells then underwent heat shock at 42ºC for 40 seconds and were placed back on ice for 5 minutes before the addition of 100µl of defrosted SOC media (B90205, NEB). Cell suspensions were incubated in a 37ºC shaking incubator for 30 minutes before being spread on LB agar plates containing ampicillin at 100µg/ml (A40040-1000.0, Melford) for selection. Agar plates were inverted and incubated at 37ºC for 16 hours prior to bacterial colonies being counted and picked using a sterile pipette tip.

2.2.8 Bacterial culture

Transformed bacteria from the agar plate were grown in 3ml LB (BP1426-500, Miller) containing ampicillin at 100µg/ml. These were incubated in a shaking incubator at 120 RPM at 37ºC overnight in preparation for plasmid extraction.

2.2.9 Plasmid DNA extraction

Bacterial cultures were centrifuged at 13,000 RPM, with the plasmid DNA being extracted using the Monarch miniprep kit (T1010S, NEB) as instructed. Plasmid DNA was purified, eluted with the elution buffer provided and concentration of the plasmid was determined using the NanoDrop spectrophotometer (ND-100, Lab Tech UK).

2.2.10 Plasmid linearisation

Prior to transfection, plasmids were linearised with the Not1-HF (R3189L, NEB) restriction enzyme. Approximately 20µg of plasmid was incubated at 37ºC overnight with Not1-HF, buffer and water. To determine if the digestion was successful, the linearised plasmid was run through agarose gel electrophoresis as

previously described. The plasmid was then purified as instructed by the Monarch PCR and DNA clean up kit (T1030L, NEB) protocol.

2.3 Methodology of *Trypanosoma brucei* cell culture

2.3.1 Preparation of *Trypanosoma brucei* media

Procyclic trypomastigotes were grown in a semi-defined media (SDM-79) manufactured by Gibco (Thermofisher) which was provided in powdered form (074-90916N, Gibco). 5 litres of SDM-79 media was prepared by adding 10g sodium hydrogen carbonate (236527, Sigma-Aldrich) to the SDM-9 powder (Gibco) with 5 litres of distilled water. The media was stirred while adjusting the pH to 7.3, allowing the media to dissolve. A supplement of 15ml of 2.5mg/ml porcine haemin (51280-1G, Sigma) and a final concentration of 10% fetal calf serum (FCS) (10500-064, Gibco) was added to the media. The prepared SDM-79 media was filtered using a vacuum filter and a stericup system (Z660507-12EA, Millipore), which was stored at 4^{o}C until needed.

2.3.2 Procyclic *Trypanosoma brucei* cell culture

Procyclic trypomastigotes were maintained at 28^{o}C in SDM-79 culture media. The cell line used in this study was the Single Marker Oxford (SmOx) cells, which were maintained on 1µg/ml of puromycin (InvivoGen). Cell lines were cultured in plastic non-vented sterile culture flasks (690160, Greiner Bio-One) and sub-cultured using a 10ml sterile serological stripette (607160, Greiner Bio-One).

2.3.3 Electroporation of *Trypanosoma brucei* cells

Cells were electroporated with the Amaxa Nucleofector II Device (Lonza) using the programme X-001 with the Amaxa Human T-cell Nucleofector kit (VPA-1002, Lonza) or prepared roditi buffer (200mM Na_2HPO_4, 70mM NaH_2PO_4, 15mM KCl, 150mM HEPES pH 7.3 with a calcium chloride stock 1.5mM $CaCl_2$). For electroporation, cells were transfected with 10µg of purified linearised DNA construct (pQuadra, or purified PCR product). For each transfection 2 x 10^7 cells

were centrifuged at 800g for 5 minutes in a Heraeus Biofuge Pico (75003235, Labcare) centrifuge, where the supernatant was discarded and the cell pellet was re-suspended in 100µl of Amaxa transfection buffer or 35µl of 3 x roditi buffer and 10µl of calcium chloride. DNA was added to the cell suspension, mixed and the total volume was transferred to a cuvette provided by the Amaxa Human T-cell Nucleofector kit (VPA-1002, Lonza). The cuvette was placed into the Amaxa Nucleofector II Device and programme X-001 was selected. Cells were then transferred into 5mls of warm SDM-79 media, free from drugs at 28^{o}C. After an overnight recovery process, appropriate drugs were chosen for selection. These drugs were added to a reagent reservoir (15075, Pierce) along with 16ml fresh warm media and 4ml of electroporated cells. Cells were then distributed into a 96-well plate (655180, Greiner bio-one), and kept in the 28^{o}C incubator until cell lines grew.

2.3.4 Production of conditioned media for clonal population of cell lines

Cells were grown in 50mls of SDM-79 media (074-90916N, Gibco) without drugs and left to grow until reaching a density of 4 x 10^7 cells. These were then centrifuged at 800g for 10 minutes, where the supernatant was transferred into a sterile 50ml falcon tube (227261, Greiner bio-one) and the cell pellet was discarded. The supernatant was filtered using a 0.4 micron filter and 5ml of the media was incubated at 28^{o}C to test for cell growth. The conditioned media was kept at 4^{o}C until required.

2.3.5 Producing a clonal population using conditioned media

To generate clonal cell lines, cells in log phase were diluted to a density of 1 x 10^3/ml. 15µl of the diluted culture was placed into 21ml of SDM-79 containing 15% conditioned media and were plated into a 96 well plate (655180, Greiner bio-one) using a multi-channel pipette. The 96 well plate was placed into a humidified CO_2 incubator at 28^{o}C and left for 6-10 days for cell growth. Cells that had grown in the 96 well plate were placed into 5ml of fresh warm SDM-79 media and routinely cultured.

2.3.6 Preparation of *Trypanosoma brucei* cells for cryopreservation

For cryopreservation, cells were first grown in 5ml of SDM-79 media until reaching a density of between 2-6 x 10^6/ml. 500µl of culture was transferred from the flask into a cryovial (T311-2, Simport), where 500µl of freezing mix (14% glycerol in SDM-79) was added. Cryovials were transported using a screw-top BioJar (490, Air Sea Containers Ltd) to the -80^{o}C freezer (U570, New Brunswick Scientific) where cells were stored overnight. Cells were then transferred into liquid nitrogen (Statebourne biorack 3000) for long term storage.

2.3.7 Revival of cryopreserved cell lines

Cell lines were removed from liquid nitrogen, placed in a BioJar at room temperature for 5 minutes to defrost and placed into 10mls of SDM-79 drug free media at 28^{o}C overnight. The appropriate antibiotic was selected and added to the media (antibiotics used in this study seen in table 5).

Table 5: Antibiotics

Antibiotics used for *T. brucei* maintenance and RNAi induction in this study

Antibiotic	Working concentration (µg/ml)	Catalogue number	Supplier
Puromycin dihydrochloride	1	Ant-pr-1	InvivoGen
Phleomycin	5	10319787	Fischer Scientific
Blasticidine S hydrochloride	20	15202	Sigma-aldrich
Doxycycline hydrochloride	1	D9891-5G	Sigma
G418-Disulfate	5	G64000-1.0	Melford

2.3.8 Plasmid induction

Cell lines containing the inducible RNAi (pQuadra) plasmid were removed from antibiotic selection by placing cells into fresh SDM-79 media 24 hours before induction. Two flasks, each containing the same cell line at a density of between 2-8 x 10^6/ml were made, with one being induced with doxycycline (D9891-5G, Sigma) and one without doxycycline as a control.

2.3.9 Growth curves

For the accumulative growth curves, cells were routinely sub-cultured in fresh media with no antibiotic selection to 1 x 10^6/ml and incubated for 24 hours at 28^{o}C. After 24 hours cells were counted using the Beckman Coutler (Z2) and sub-cultured back to 1 x 10^6/ml. This was carried out for both induced and uninduced cell lines and repeated for 96 hours, or for 192 hours for the Cep164C cell line.

2.3.10 Cell cycle analysis of the Cep164 proteins

Cell cycle analysis for the induced and uninduced cell lines were conducted in cells at different points of the cell cycle, according to the number of DNA containing organelles in each individual cell. kDNA and nucleus numbers were recorded for each cell line to determine the timing of recruitment for the Cep164 proteins. 100 cells of each classification (1K1N, 2K1N etc) were imaged using the Zeiss Imager Z2 fluorescence microscope and post-processed on ImageJ.

2.3.11 Asynchronous and enriched SmOx and Cep164C cell line

SmOx cells were routinely sub-cultured between 2-8 x 10^6/ml in 10ml of fresh SDM-79 media without antibiotic selection. Cells were sub-cultured to a density of 5 x 10^4/ml in 6ml of SDM-79 media and allowed to cultivate for five days with no additional fresh media. Every 24 hours, flasks were gently swirled and cells were observed by light microscopy for bacterial contamination. After five days, cells were released at 1.2 x 10^7/ml in 6ml of fresh SDM-79 media, with a sample of the culture being taken each hour for eight hours to observe enrichment of cells through the cell cycle.

To determine the localisation of Cep164C, the Cep164C::mNG cell line was put through the enrichment protocol as described above. 100 cells were imaged for each point in the cell cycle and the presence or absence of the protein was noted.

2.4 Methodology of protein analysis

2.4.1 Whole cell protein samples for western blots

Protein samples were prepared by sub-culturing all cell lines to 1 x 10^9/ml using the Beckman Coulter (Z2) to determine cell density. Cells were pelleted using a centrifuge and the supernatant was discarded, where the pellet was then re-suspended in phosphate buffer saline (PBS) and pre-heated Laemmli buffer at 100^{o}C on a heat-block. The suspension was placed back onto the heat block at 100^{o}C for a further 15 minutes and vortexed, after which the protein lysates were stored at -20^{o}C.

2.4.2 Polyacrylamide gel electrophoresis

All protein work was done using the mini protean 12% polyacrylamide precast gels (456-1045, BioRad). Gels were loaded into the mini protean tank (1658004, BioRad), where the tank was flooded with running buffer (250 mM Tris, 1.92 M glycine, 1% SDS dissolved in in 1 litre of deionised water). Protein samples were loaded into the lanes at 10µl (1 x 10^7 per lane) and a pre-stained protein ladder (P7712S, NEB) was loaded for band size. The gels were run at 100 volts for 140 minutes.

2.4.3 Staining with Coomassie brilliant blue

Prior to transferring proteins from the polyacrylamide gels to the PDVF membrane, gels were stained with Coomassie brilliant blue (B5133, Sigma Aldrich). Gels were immersed in the stain for 15 minutes, before removing and using the de-staining solution (10% v/v methanol and 20% v/v acetic acid in distilled water) for 10 minutes. The gel was then removed and washed again in the de-staining solution.

2.4.4 Protein transfer and western blotting

Proteins were transferred from the polyacrylamide gel to the Immobilon PVDF membrane (IPVH00010, Millipore), with a trans-blot electrophoretic transfer cell (1703930, BioRad) carrying out the transfer. Prior to transfer, the membrane was

incubated in methanol for 10 minutes for activation. Prior to assembling the cassette, sponges and filter paper were soaked with transfer buffer (25 mM Tris, 1.92 mM glycine containing 20% methanol and 1% SDS dissolved in 1 litre of deionised water) and the mini-blot transfer cell was filled with transfer buffer after inserting the cassette. A cold pack and a magnetic stirrer were added to the cell, where the transfer was run for 2 hours at 100 volts. Transfer of protein was confirmed by visibility of the pre-stained protein ladder on the membrane.

The PVDF membrane was incubated in blocking buffer (5% non-fat w/v milk in PBS tween 20) overnight at 4^{o}C. This was washed with PBS tween 20 and incubated in the primary antibody in 10ml blocking solution for 2 hours, to prevent non-specific binding of antibody, on a rotator. The antibody solution was discarded and the membrane was washed in PBS tween 20 for 20 minutes on a rotator, changing the wash every 5 minutes. The membrane was incubated in the secondary antibody in 10ml blocking solution for 1 hour. The membrane was washed again in PBS tween 20 as previously carried out prior to detection using the Western Bright substrate kit (K-12042-D10, Advansta).

2.5 Methodology of light microscopy and confocal microscopy

2.5.1 Observing live and detergent-extracted cells by microscopy

Transfectants were grown in SDM-79 media to a density between 2-8 x 10^{6}/ml prior to microscopy observations. Cells imaged live were centrifuged at 800g for 5 minutes, the supernatant was removed and the pellet was washed twice with PBS supplemented with 20% sucrose. After a second re-suspension, 50µl of cells were transferred onto SuperFrost microscope slides. Cells were observed by the Zeiss Imager Z2 fluorescence microscope using the Zen blue (2012) software and the ORCA-R^{2} camera (Hamamatsu). mCherry (610nm), GFP (509nm) and 4′,6-diamidino-2-phenylindole (DAPI) (465nm) filters were used at 3.000, 3.000 and 0.500 second exposure times to observe fluorescence in the cells, with images being post-processed in the ImageJ software (FIJI).

Fixed whole cells were handled as mentioned previously until transferred onto the SuperFrost microscope slides, where they were plunged into -20^{o}C methanol and

kept for 30 minutes at -20°C. To generate cytoskeletal cells, the addition of PEME buffer (0.1 M PIPES, 2 mM EGTA, 1 mM $MgSO_4$, 0.1 mM EDTA) and NP40 was added to the slide for 2 minutes, checked by the light microscope and plunged into methanol at -20°C for 30 minutes. Fixed cells on the SuperFrost slides were rehydrated using PBS for 10 minutes prior to adding the DNA marker DAPI (1:250,000) or Hoechst 33342 (1:1000) for 5 minutes. Cells were washed for 3 x 5 minutes in PBS and checked by light microscopy as previously mentioned.

2.5.2 Immunofluorescence labelling

Cells on the SuperFrost slides were labelled with the primary antibody (table 6) for 1 hour at room temperature in a humid chamber. The slides were then washed with PBS for 3 x 5 minutes, before adding the secondary antibody (table 7) in 1% w/v bovine serum albumin (BSA) blocking buffer (A-947, Sigma) for 1 hour in the humid chamber. The slides were washed with PBS for 3 x 5 minutes.

Table 6: Primary antibodies used in this study

Primary antibodies	Class	Origin	Labels	Dilution	Supplier
BBA4	IgM	Mouse monoclonal	Basal body (Woodward *et al,* 1995)	1:50	Generated by Oxford University
L8C4	IgG	Mouse monoclonal	Paraflagellar Rod (PFR2) (Kohl *et al,* 1999)	1:50	Generated by Oxford University
YL1/2	IgG	Rat monoclonal	RP2 on transitional fibres, and acetylated tubulin (Andre *et al,* 2014)	1:50	Generated by University of Cambridge

Slides were then flooded with the DNA marker DAPI (1:250,000) or Hoechst 33342 (1:1000) for 10 minutes and mounted with the Vectashield mounting medium (H-1000 Vector laboratories). Coverslips (474030-9000-000, Zeiss) were applied to the slides and secured using clear nail varnish and allowed to dry for a minimum of 1 hour prior to observation by microscopy.

Table 7: Secondary antibodies used in this study

Primary antibodies	Secondary antibodies	Origin	Antigen	Dilution	Secondary antibody supplier
BBA4	Goat anti mouse IgM-TRITC	Goat polyclonal	Mouse IgM	1:50	Jackson Immuno-Research 715-025-140
L8C4	Goat anti mouse IgG-FITC	Goat polyclonal	Mouse IgG	1:50	Jackson Immuno-Research 115-095-062
YL1/2	Rhodamine Goat anti-rat IgG	Goat polyclonal	Rat IgG	1:1000	Jackson Immuno-Research AB_2338111

2.5.3 Measurements of flagellum lengths

200 dividing and G1 cells labelled with the PFR antibody were collected by fluorescence microscopy, where the flagella was traced using the ImageJ software and the segmented line tool. PFR labelling was used to measure either one flagellum (1K1N) (figure 17) or two flagella (2K2N) for the Cep164A, Cep164B, Cep164C and Cep164A+C induced and uninduced cell lines, all results were recorded in Microsoft excel.

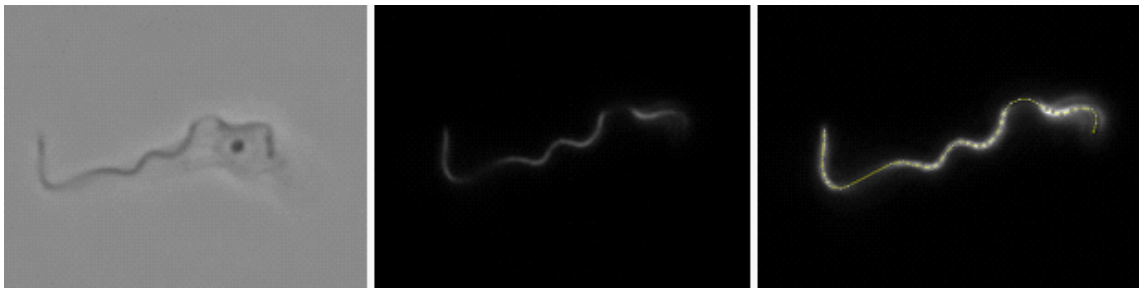


Figure 17: A *trypanosoma brucei* cell labelled with the flagellum marker antibody, L8C4. The segmented line tool in the ImageJ software is used for flagellum measurements.

2.5.4 Measurements of kinetoplast and nucleus distances

The 200 cells collected for flagellum length measurements were also investigated to collect the distance measurements between the kinetoplasts and nuclei. In G1 cells the kinetoplast and nucleus was measured. In dividing cells, distances between the new daughter kinetoplast and nuclei (figure 18A), old daughter kinetoplast and nuclei (figure 18B) and new and old kinetoplasts (figure 18C) were measured.

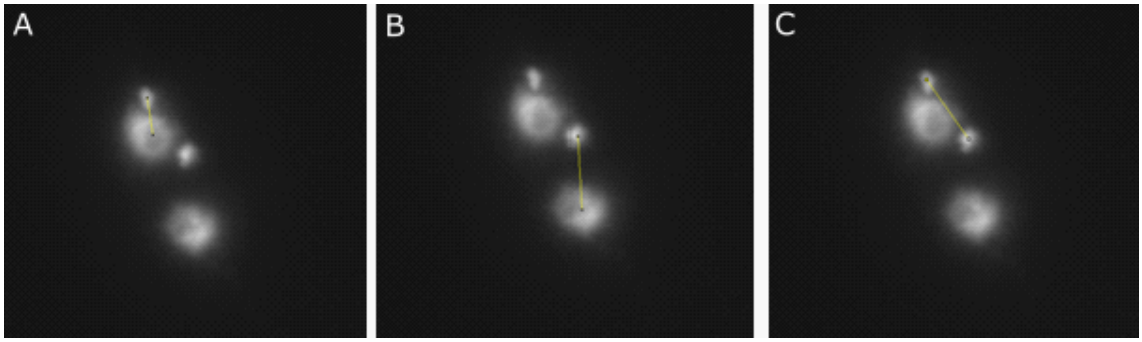


Figure 18: A *Trypanosome brucei* cell labelled with the DNA marker DAPI, labelling the kinetoplast and nuclei. Distances between the nucleus and kinetoplast was carried out using the segmented line tool in ImageJ.

2.5.5 Fluorescence recovery after photobleaching

The Cep164 inducible RNAi cell lines were cultured without antibiotics between a density of 2-8 x 10^6/ml prior to observation. Live cells were placed straight from culture onto the slide and were observed by the confocal microscope, using the velocity software. The region of interest (RIO) was set to the BB area, where the RIO was manually bleached 3-4 times to ensure there was no mNG expression. A recording was set prior to bleaching, which continued to record for 1 minute at 1 frame/second. The movies captured were post-processed and analysed in ImageJ.

2.6 Electron microscopy

2.6.1 Scanning electron microscopy (SEM)

For SEM imaging, between 2-8 x 10^6/ml cells were fixed in suspension with a final concentration of 2.5% v/v glutaraldehyde (G002, Taab). Cells were centrifuged, washed twice with PBS before being transferred onto Thermanox coverslips (174950, Thermo fisher) in 24-well plates, which were left to settle for 10 minutes. An ethanol dehydration series with distilled water was carried out at 30%, 60% and 90% ethanol concentrations for 5 minutes each, which were then dehydrated in 100% ethanol for 3 x 5 minutes. After dehydration, cells were put through critical point drying in an E3000 (Polaron, UK). Coverslips were mounted onto 15nm SEM stubs (G3313, Agar Scientific) using sticky carbon tabs (G3347N, Agar Scientific) and gold coated for 30 seconds using the Agar auto sputter coater (B7431, Agar Scientific). The stub was inserted into a Hitachi S-3400 scanning electron microscope, where the chamber was put under vacuum. Cells on the coverslip were first visualized using the “fast 1” setting. After focusing the beam, the “slow 3” setting was used for high resolution imaging.

2.6.2 Transmission electron microscopy (TEM)

For TEM imaging, between 2-8 x 10^6/ml cells were fixed in suspension with a final concentration of 2.5% v/v glutaraldehyde for (G002, Taab) for five minutes. Cells were then centrifuged for five minutes at 800g, where the supernatant was discarded as appropriate. The cell pellet was re-suspended into the primary fixative (2.5% v/v glutaraldehyde, 2% v/v paraformaldehyde in 0.1M phosphate buffer pH 7.0) and incubated for 1 hour at room temperature.

Once incubated, the suspension was transferred to a 1.5ml centrifuge tube where all centrifugation steps were carried out at a 90-degree angle to generate flat pellets. These pellets were centrifuged and washed 5 times with 0.1M phosphate buffer, then post-fixed in 1% v/v osmium tetroxide (O017/1, Taab) in a 0.1M phosphate buffer for 1:30 minutes at room temperature. Osmium was removed as appropriate and the pellet was washed 5 times in distilled water, prior to being stained *en bloc* overnight in 2% v/v uranyl acetate (R1260A, Agar Scientific Ltd) in

dark conditions. Cell pellets were washed 5 times in distilled water before beginning the acetone dehydration series with distilled water from 10-90% acetone concentration at 20 minutes each. The pellets were then dehydrated 3 times in 100% acetone at 20 minutes each.

To embed the samples into 812 embedding resin (T030, Taab), the final 100% dehydration in acetone was removed from the pellet and a mixture of 25% v/v resin in acetone was added for 2 hours on a rotator in room temperature. The pellet was then centrifuged at 13,000 RPM, where the 25% v/v resin was removed and replaced with 50% resin and 50% acetone for 1 hour on rotation. After another centrifugation step the 50% resin and 50% acetone was removed and replaced with 75% resin in 25% acetone to the pellet overnight on rotation. Samples were again centrifuged and the 75% v/v resin was replaced with 100% resin overnight on rotation. Finally, the 100% resin was changed 3 times at 2 hours incubation each. The pellets in 100% resin were then polymerised in a resin oven at 70^{o}C for 24 hours.

2.6.3 Sectioning samples for TEM

Resin blocks were removed from the 1.5ml centrifuge tubes after polymerisation. To section the samples, the resin block was mounted in a PowerTome (RMC) and trimmed with a glass knife to create a cutting edge. Once the surface of the block was straight, a diamond knife was used to cut sections at 70nm into the sectioning boat filled with distilled water. Each section cut floated onto the water and was collected onto 3.05nm copper grids (GG005/C, Taab). Excess water was removed from the grid using filter paper and was left to dry in a suitable grid container.

2.6.4 Post-staining for TEM

Once dried the sections were post-stained using filtered uranyl acetate, which were incubated in a Petri dish for 15 minutes in dark conditions. The sections were dipped in distilled water 10 times before being incubated with filtered lead citrate, surrounded by sodium hydroxide (S/4920/60, Fischer Scientific) for 15 minutes. To

wash the sections, these were dipped again in fresh distilled water 10 times prior to drying in a sealed suitable container.

Micrographs of the sections were taken using a Hitachi H-7650 transmission electron microscope, operated at 100kV using an AMT 2kx2k CCD camera (Advanced Microscopy Technologies, Suffolk, UK). Images of the samples were captured at either 10,000 or 20,000 x original magnification.

2.6.5 Serial block-face scanning electron microscopy (SBF-SEM)

Samples for SBF-SEM were prepared in a similar manner as the TEM (section 2.6.2), however after trimming the polymerised block on the PowerTome (RMC), blocks were mounted onto stubs suitable for SBF-SEM (Gatan, UK) using superglue. These mounted blocks were loaded into a Merlin compact field emission SEM (Ziess) fitted with a 3View serial block-face imaging system (Gatan). Images were taken at an accelerating voltage of 4kV and at a chamber pressure of 45.4 pascals. Magnification was 7,000 x original size and the section slice thickness was 75nm. Images were recorded using the Digital Micrograph (Gatan).

2.6.6 Serial block-face scanning electron microscopy data analysis

Data collected from the Digital Micrograph (Gatan) was converted from .dm3 or .dm4 files to a .mrc stack for image assembly and analysis in the IMOD software (Kremer *et al,* 1996) via Cygwin (Cygnus solutions). Once generated, the axes and pixel spacing were corrected and images aligned using Etomo. Segmentation of the flagellum, flagellar pockets, basal bodies and whole cells were carried out using the manual brush tool on 3dmod. Each component modelled was saved as a different material within 3dmod. Measurements of distances between organelles was carried out by creating an open contour, where a line was drawn and measured. Closed contours were used for modelling the flagellum, flagellar pockets and basal bodies. Generated models were observed using the “Model View” section of 3dmod.

2.6.7 Statistical analysis

Microsoft Office Excel and the SPSS software (IBM, 1968) were used to perform the statistical analysis on the data. The p-value of independent *t*-tests were calculated with a 99% confidence interval. The relationship between cell body length and flagellum length was determined using the Pearson correlation coefficient.

3. Bio-informatic and basal body screen results

This PhD study began at the beginning of The Wellcome Trust TrypTag project, a genome wide localisation screen aimed at tagging all proteins in the *Trypanosoma brucei* parasite (Dean *et al,* 2017). To begin, 20 putative BB proteins were chosen from one of the pilot projects. These proteins were shown to localise close to the BB area in *T. brucei* (in collaboration with the TrypTag Project). The aim of the first set of experiments were to confirm localisation using the well characterised BB protein SAS-6 (Hu *et al,* 2015) and co-localise the 20 putative proteins with this marker.

3.1 Plasmid construction and PCR-only tagging for the 20 putative basal body proteins

To confirm the 20 proteins located to the BB, new cell lines were made to co-localise with the well characterised BB protein SAS-6. This protein localises to the mature basal body and pro-basal body of all cells (Nakazawa *et al,* 2007). In G1 cells with a single flagellum, nucleus and kinetoplast there are 2 foci close to the kinetoplast. When the cell cycle begins, the BBs duplicate and 4 foci are seen in each cell. This is then a useful marker to identify BB proteins, discover their proximity to SAS-6 (proximal, co-localised or distal) and whether they are present in dividing cells. An mCherry cell line endogenously expressing the SAS-6 protein was used in this study (gifted from Jack Sunter at Oxford University). The mCherry::SAS-6 cell line was used as a marker to investigate the co-localisation of the 20 genes of interest to the BB and those expressed during BB maturation. Cell lines were generated using the original pPOTv4 plasmid containing the mNG fluorophore and the Blasticidin resistance marker with BamH1 restrictions sites either side of the fluorophore (Dean *et al,* 2015). This plasmid underwent modification to include the TY1 epitope tag recognised by the BB2 antibody (Bastin *et al,* 1996), so the protein could be identified by western blotting if required (figure 19).

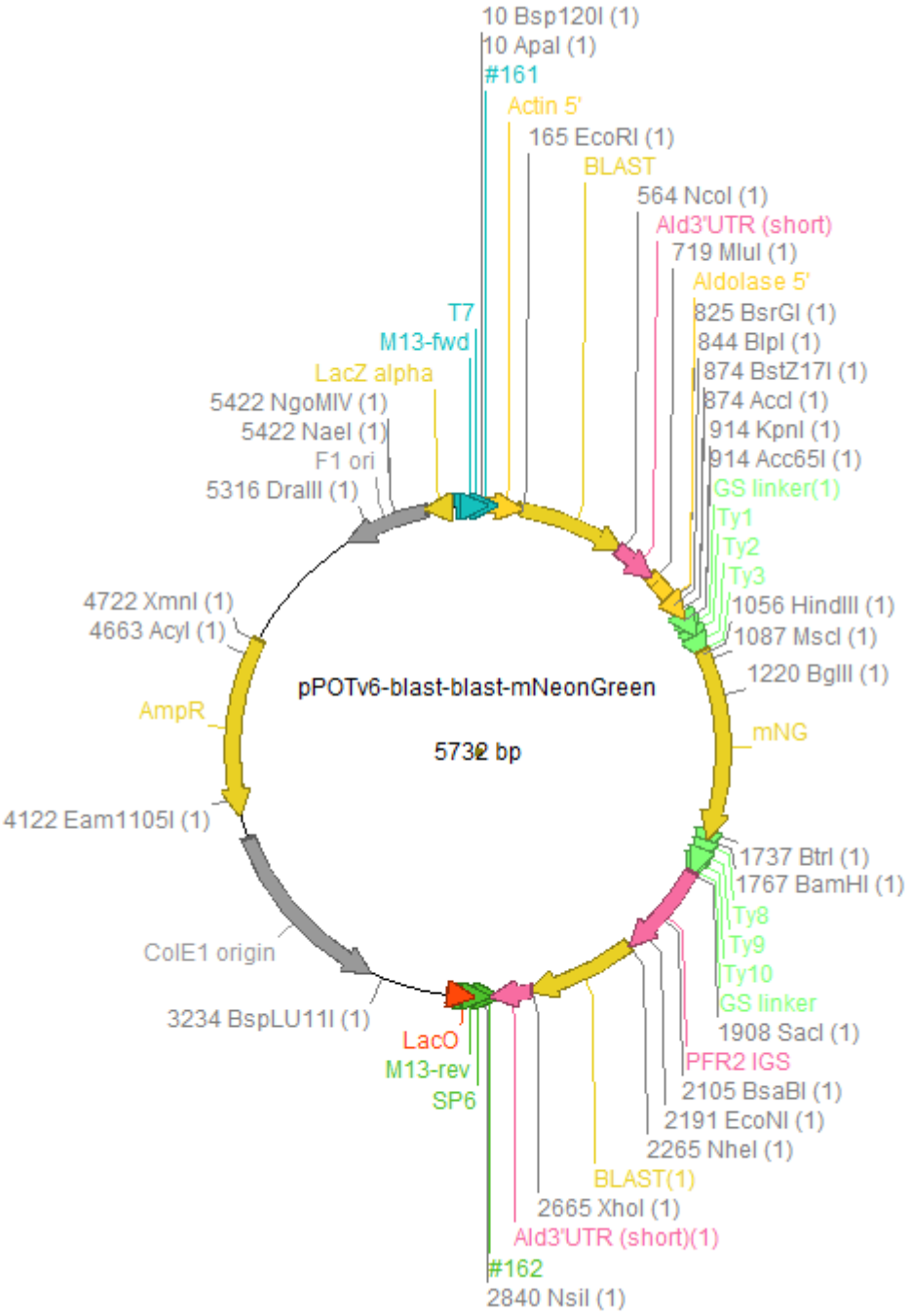


Figure 19: Plasmid map of the pPOTv4 plasmid

The plasmid map of the pPOTv4 plasmid prior to modification to include the TY1 tags

The original pPOTv4 plasmid was dephosphorylated and TY1 tags were inserted into the plasmid, these were amplified in *E.coli* cells. As BamH1 was used to cut the TY1 tags and the plasmid, the TY1 tags could have ligated with the plasmid in either orientation. To determine which bacterial colonies contained the plasmid in the correct orientation, cells were prepared and run through gel electrophoresis (figure 20).

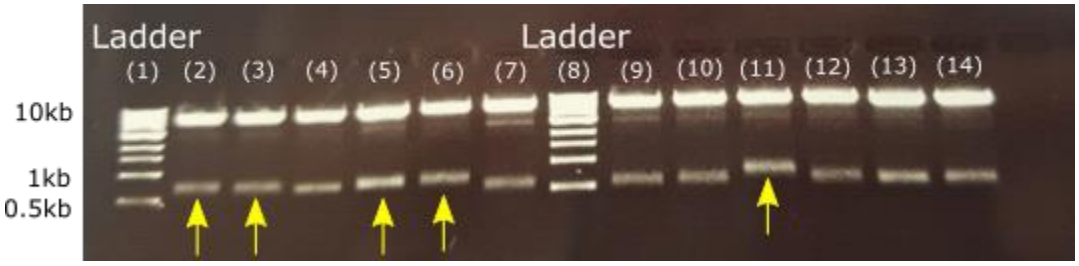


Figure 20: Modified plasmids with either the incorrect or correct TY1 tag orientation. Lanes 1 and 8 – Ladder. Lanes 2, 3, 5, 6, 11 – modified pPOTv4 plasmid with the correct orientation of TY1 tags. Lanes 4, 7, 9, 10, 12, 13, 14 – modified pPOTv4 plasmid with the incorrect orientation of TY1 tags.

Those bacterial cells containing larger bands for the plasmid backbone and shorter bands for the insert identified bacterial colonies containing the mNG pPOTv4 plasmid with Blasticidin resistance and the inserted TY1 tags in the correct orientation (figure 20; yellow arrows). Using the modified plasmid, 20 of the BB genes were amplified by the pPOT PCR technique and transfected into the mCherry::SAS-6 cell line, where the localisation of the mNG fluorophore to the BB was assessed.

3.2 Analysis of the 20 putative basal body proteins

Detergent extracted *T. brucei* cells of the 20 generated cell lines were observed through fluorescent microscopy to confirm the localisation of the proteins to the BB. Detergent extracted cells were generated to identify whether any of the 20 proteins were present after membrane removal, confirming whether they were tightly bound

to the cytoskeleton. Live cells were imaged when no protein localisation was seen in detergent extracted cells. 19 cell lines were generated successfully with clear signal expression, with one cell line failing to grow. All images were assessed for their relative position to SAS-6, whether the proteins localised to the MBB or PBB and what happened to the signal through the cell cycle. A summary of the screen is seen in table 8 (figure 21-39), with protein present in green and protein absent in red.

Six proteins co-localised to the SAS-6 protein on the MBB and PBB in 1K1N and 2K2N cells (table 8; F (figure 26A-L), G (figure 27A-L), H (figure 28A-L) and I (figure 29A-L)). Tb927.10.14770 had signal expression in un-fixed areas (including the flagellum) with similar intensity to those observed in the BB (figure 29C and I). Two proteins localised with a lower expression level on the MBB in comparison to the PBB (table 8; A (figure 21A-L) and C (figure 23A-L)).

Four proteins had no localisation in 1K1N or 2K2N detergent-extracted cells, suggesting these proteins are not tightly bound to the cytoskeleton (table 8; J (figure 30A-H), K (figure 31A-H), L (figure 32A-L) and M (figure 33A-L)). Two of these proteins contained a localisation distal to the MBB in live cells, with no localisation to the PBB (table 8; J (figure 30A-H) and K (figure 21A-H)).

All other cell lines co-localised to SAS-6 at either the MBB or PBB in 1K1N and 2K2N cells. Three proteins displayed a differing BB localisation dependent on the cell cycle stage of the cell. From these, one protein localised to the PBB only in 1K1N cells and to the MBB and PBB in 2K2N cells (table 8; B (figure 22A-L)). One protein localised in-between the MBB and PBB, distal to the PBB in 1K1N cells and distal to the MBB and PBB in 2K2N cells (table 8; E (figure 25A-L)). From these, the remaining protein localised to the distal section of the MBB and PBB and in-between both BBPs creating a band-like structure in 1K1N cells. Whereas the protein localised to the distal section of the MBB and in-between the BBPs in 2K2N cells, with no localisation present at the PBB (table 8; D (figure 24A-L)).

Table 8: Localisation of the 20 basal body proteins

Localisation of the basal body proteins tagged with the mNG fluorophore in the background of the SAS-6 protein. Green = protein present, Red = protein absent. 1K1N = G1 cell, 2K2N = dividing cell.

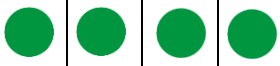
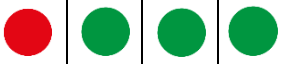
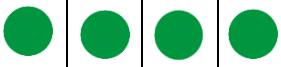
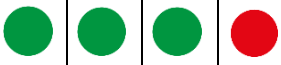
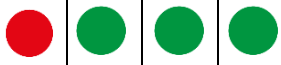
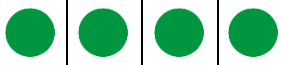
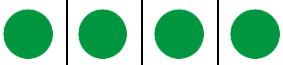
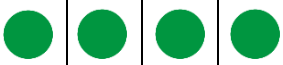
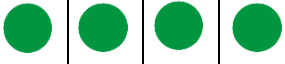
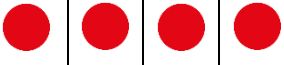
Accession number	Localisation				Comments	Distal to SAS-6	Co-localised to SAS-6	No localisation
	1K1N MMB/ PBB		2K2N MMB/ PBB					
(A) Tb927.5.1120 (Figure 21)					Distal to PBB and lower expression on MBB in 1K1N cells. MBB and PBB in 2K2N cells.	✓		
(B) Tb927.10.3130 (Figure 22)					PBB only in 1K1N cells, MBB and PBB in 2K2N cells		✓	
(C) Tb927.10.10090 (Figure 23)					PBB and lower expression on MBB in 1K1N cells, MBB and PBB in 2K2N cells	✓		
(D) Tb927.9.7720 (Figure 24)					Band across MBB and PBB in 1K1N cells	✓		
(E) Tb927.9.14300 (Figure 25)						✓		
(F) Tb927.10.2860 (Figure 26)							✓	
(G) Tb927.11.5030 (Figure 27)							✓	
(H) Tb927.7.7200 (Figure 28)						✓		
(I) Tb927.10.14770 (Figure 29)							✓	
(J) Tb927.8.4210* (Figure 30)					No localisation in extracted cells and BB area in whole cells		✓	

Table 8: Localisation of the 20 basal body proteins

Accession number	Localisation				Comments	Distal to SAS-6	Co-localised to SAS-6	No localisation
	1K1N MMB/ PBB		2K2N MMB/ PBB					
(K) Tb927.11.13300 * (Figure 31)	🔴	🔴	🔴	🔴	No localisation in extracted cells	✓		
(L) Tb927.10.4990* (Figure 32)	🔴	🔴	🔴	🔴	No localisation in extracted cells			✓
(M) Tb927.10.13610 * (Figure 33)	🔴	🔴	🔴	🔴	No localisation in extracted cells			✓
(N) Tb927.10.12590 (Figure 34)	🟢	🔴	🟢	🔴		✓		
(O) Tb927.1.3560 (Figure 35)	🟢	🔴	🟢	🔴	Cell cycle regulated. Observed only on BB1	✓		
(P) Tb927.7.5190 (Figure 36)	🔴	🟢	🔴	🟢	Distal to PBB in 1K1N and 2K2N cells	✓		
(Q) Tb927.7.4130 (Figure 37)	🔴	🔴	🔴	🔴	In-between MBB and PBB			✓
(R) Tb927.11.2700 (Figure 38)	🔴	🔴	🔴	🔴	In-between MBB and PBB			✓
(S) Tb927.11.15450 (Figure 39)	🔴	🔴	🔴	🔴	No co-localisation with SAS-6	✓		
(T) Tb927.4.4150	---	---	---	---	Did not generate a cell line	---	---	---

* No localisation in detergent-extracted cells

One protein localised to the MBB only in 1K1N cells, and to the MBB on both BBPs in 2K2N cells (table 8; N (figure 34A-L)). Whereas one protein localised to the MBB only in 1K1N cells, but only localised to the MBB of the old BBP in 2K2N cells (table

8; O (figure 35A-L)). Another protein only localised to the distal section of the PBB for both the new and old BBPs throughout the cell cycle (table 8; P (figure 36A-L)).

Two proteins had a localisation in-between the MBB and PBB in 1K1N and 2K2N cell cycle stages (table 8; Q (figure 37A-L) and R (figure 38A-L)). Finally, the last protein did not co-localise to SAS-6 or localise to the BBs. Tb927.11.15450 instead localised further distal to the BBs and was not considered a BB related protein for the MBB or PBB (table 8; S (figure 39A-L). Tb927.4.4150 did not generate a cell line and no localisation to the BBs was determined (table 8; T).

In those proteins which localised to both the MBB and PBB during the cell cycle, six proteins were identified to have a brighter localisation on the MBB (figure 24K, figure 26K, figure 27K, figure 28K, figure 29K and figure 34K; white arrow head) in comparison to the PBB. Only one protein had a brighter localisation to the PBB (figure 21; white arrow) in comparison to the MBB (figure 21; white arrow head). These results suggest a signal strength variation in some of the 20 generated proteins. It is noted that the localisations observed in this screen were not taken from clonal cell lines.

Tb927.5.1120

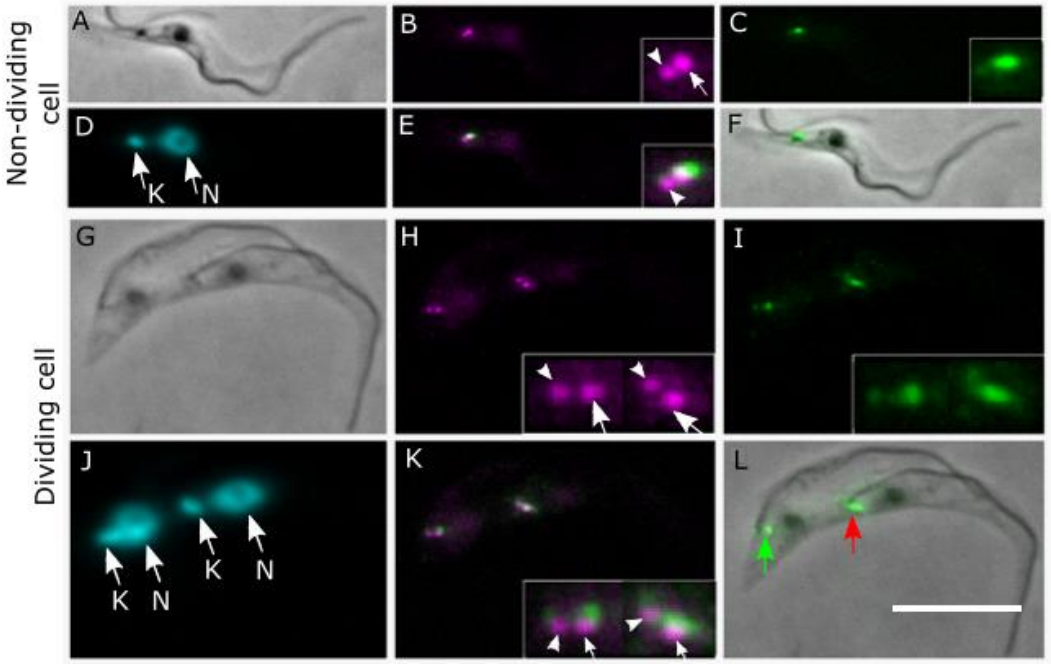


Figure 21: Localisation of Tb927.5.1120

Co-localisation of Tb927.5.1120 (mNG) with the SAS-6 mCherry (magenta) cell line to the pro-basal body (white arrow) and lower expression on the mature basal body (arrowhead) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.3130

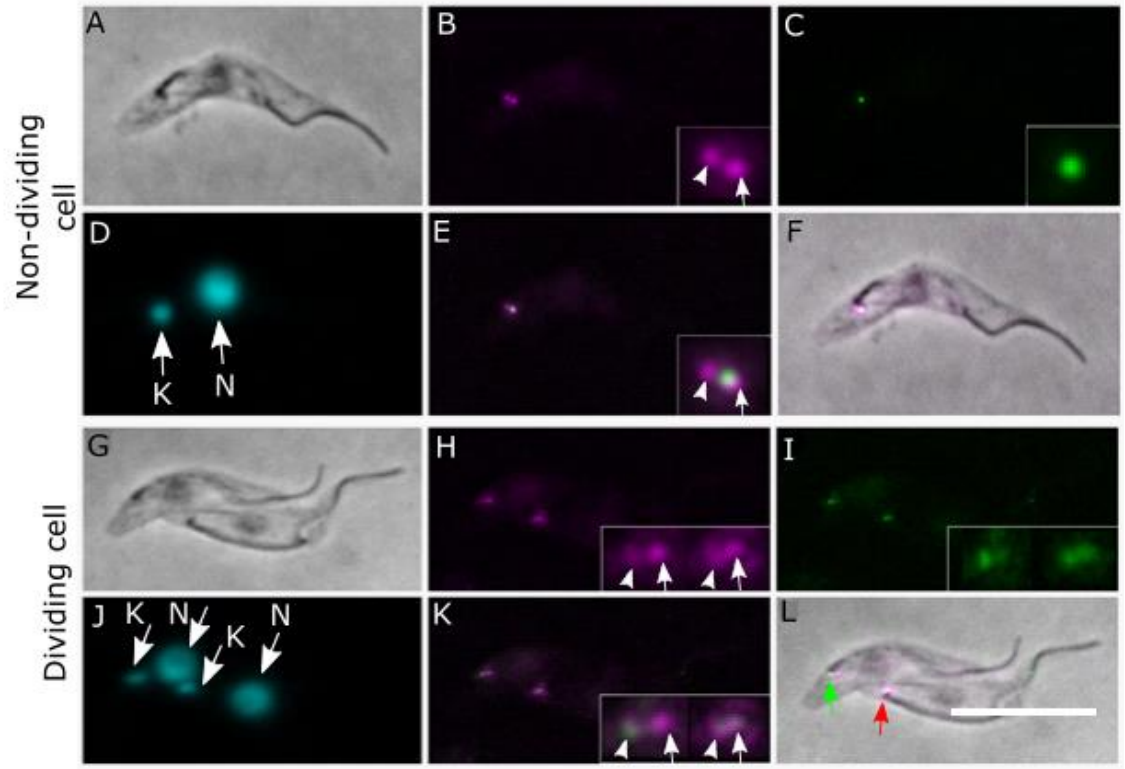


Figure 22: Localisation of Tb927.10.3130

Co-localisation of Tb927.10.3130 (mNG) with the SAS-6 mCherry (magenta) cell line to the pro-basal body only in 1K1N cells (A-F) (E; arrow) and to the mature (arrow head) and pro-basal body (arrow) in 2K2N cells (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.10090

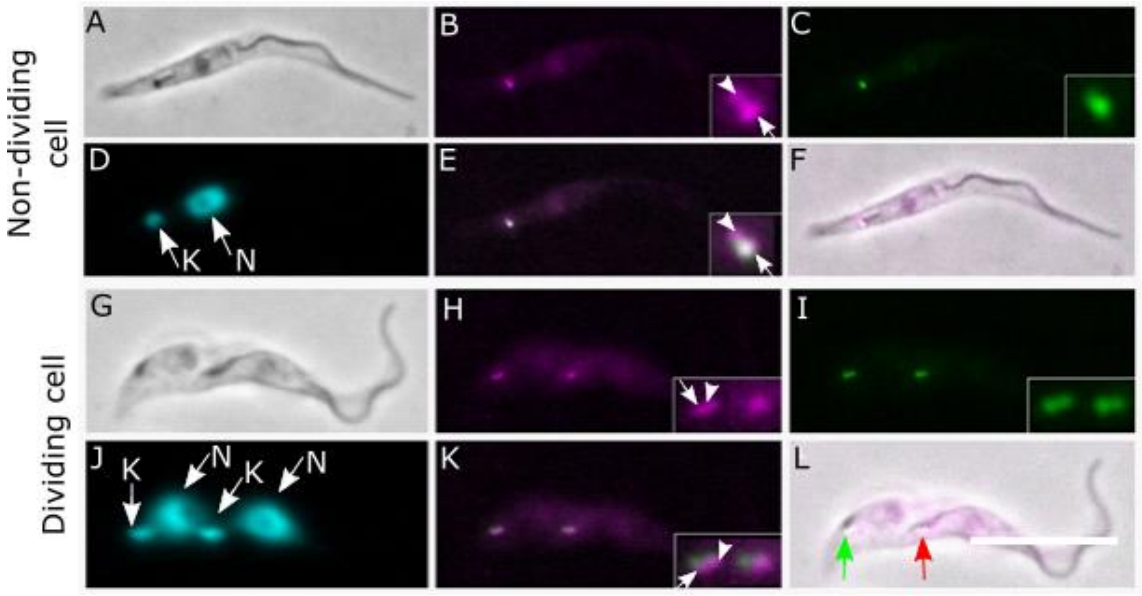


Figure 23: Localisation of Tb927.10.10090

Co-localisation of Tb927.10.10090 (mNG) with the SAS-6 mCherry (magenta) cell line to the pro-basal body (white arrow) and low expression on the mature basal body (white arrowhead) in 1K1N cells (A-F), and to both the mature basal body and pro-basal body in 2K2N cells (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.9.7720

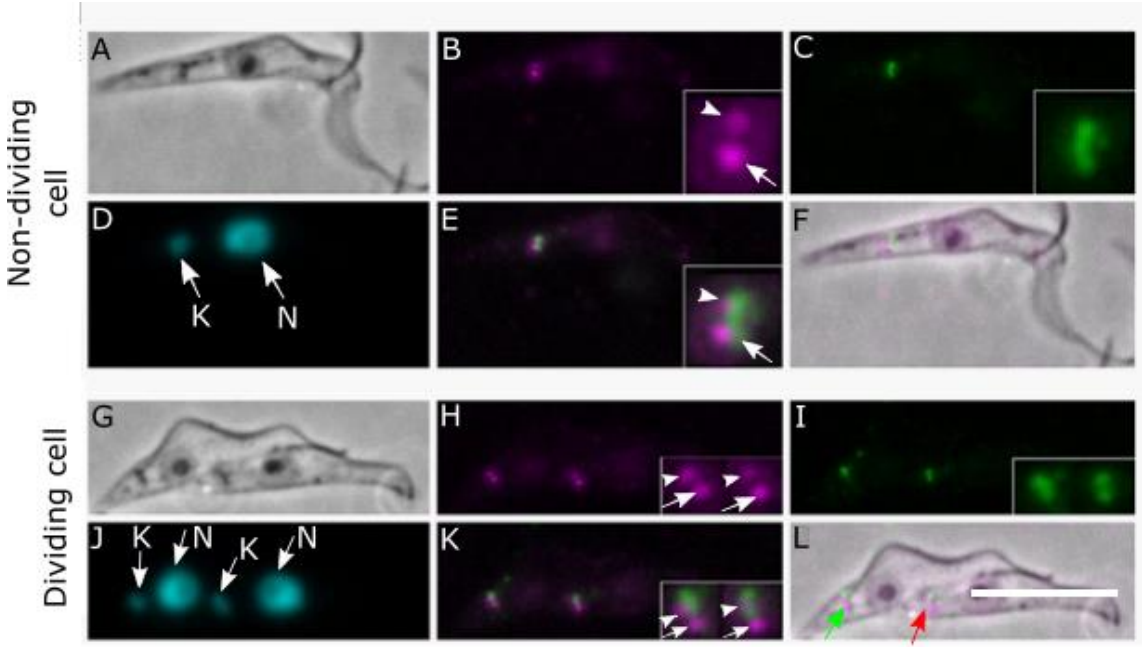


Figure 24: Localisation of Tb927.9.7720

Co-localisation of Tb927.9.7720 (mNG) with the SAS-6 mCherry (magenta) cell line to the distal end of the mature (arrow head) and pro-basal body (arrow) and in-between the organelles in 1K1N cells (A-F) and localising to the distal end of the mature basal body in 2K2N cells (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.9.14300

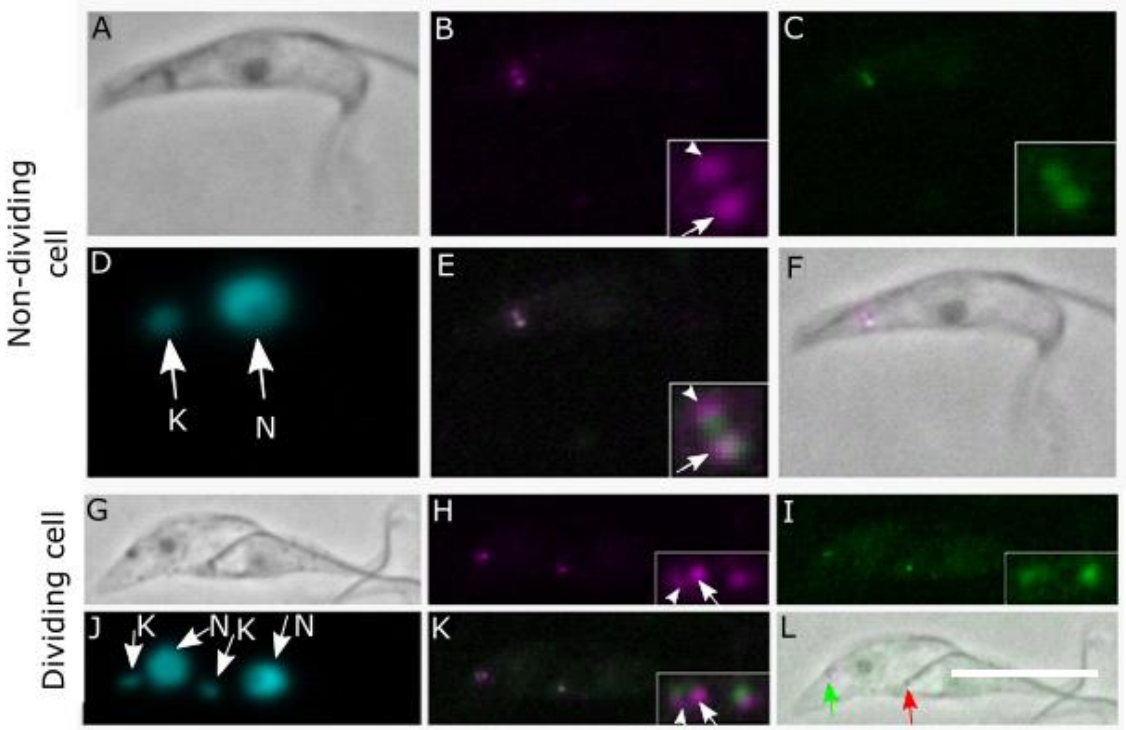


Figure 25: Localisation of Tb927.9.14300

Co-localisation of Tb927.9.14300 (mNG) with the SAS-6 mCherry (magenta) cell line to in-between the mature basal body (arrowhead) and pro-basal body (arrow) and distal to the pro-basal body in 1K1N cells (A-F), and distal to the mature basal body and pro-basal body in 2K2N cells (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.2860

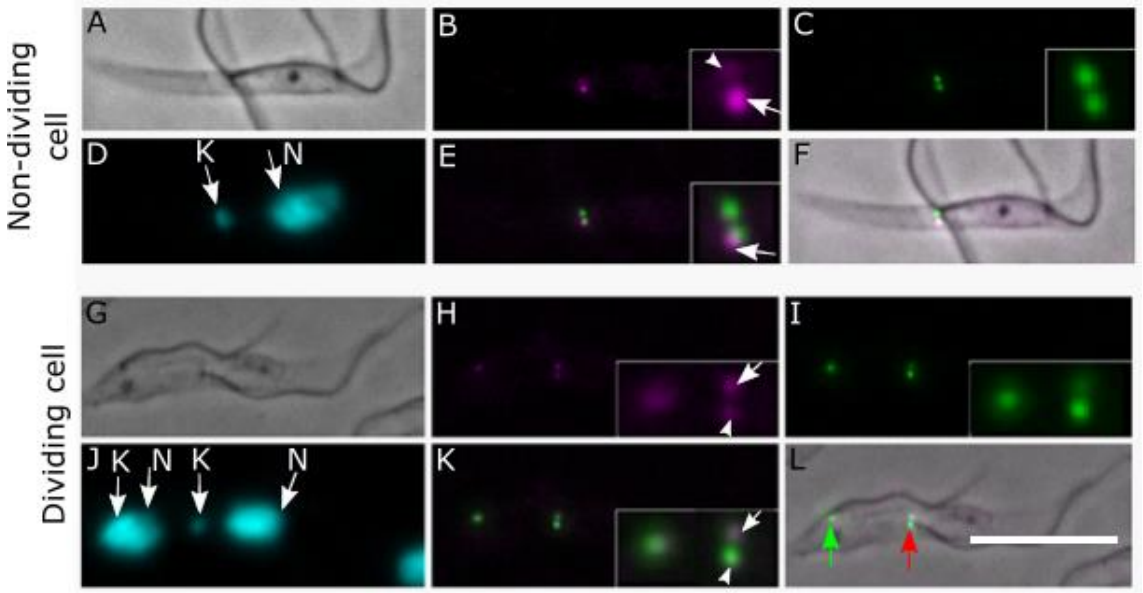


Figure 26: Localisation of Tb927.10.2860

Co-localisation of Tb927.10.2860 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature basal body (arrow head) and pro-basal body (arrow) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.11.5030

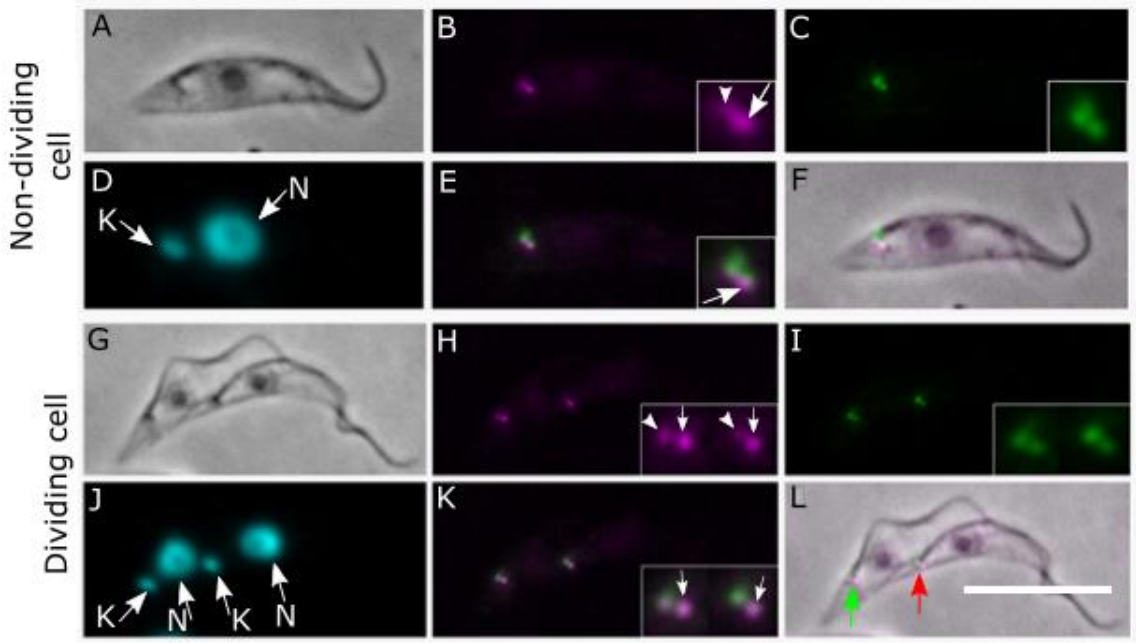


Figure 27: Localisation of Tb927.11.5030

Co-localisation of Tb927.11.5030 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature basal body (arrow head) and pro-basal body (arrow) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.7.7200

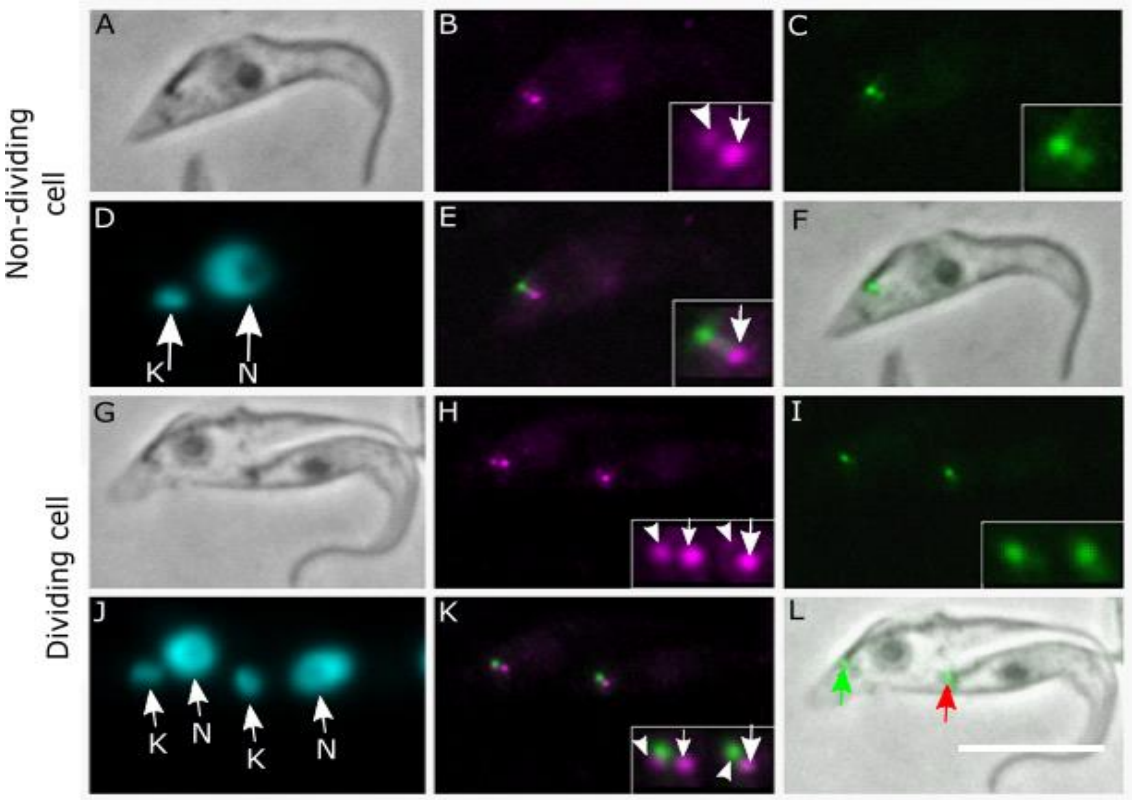


Figure 28: Localisation of Tb927.7.7200

Co-localisation of Tb927.7.7200 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature basal body (arrow head) and pro-basal body (arrow) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.14770

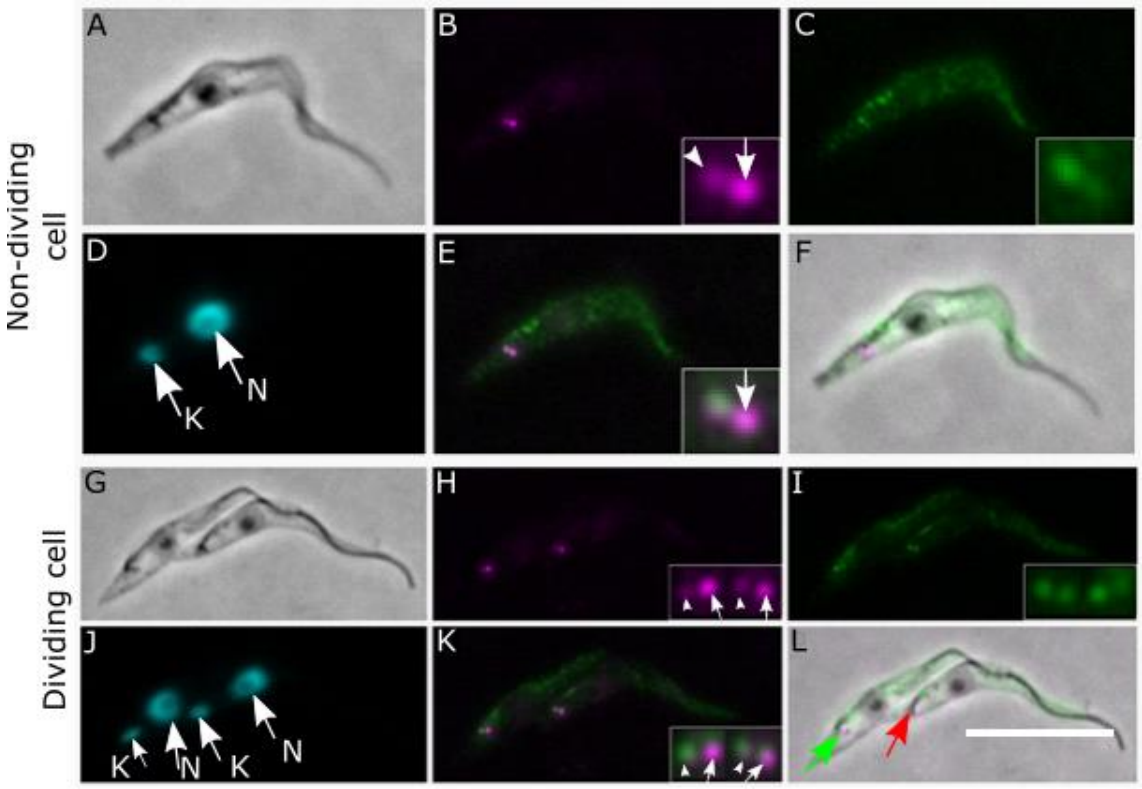


Figure 29: Localisation of Tb927.10.14770

Co-localisation of Tb927.10.14770 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature (arrow head), pro-basal body (arrow) and flagellum in 1K1N cells (A-F) and 2K2N cells (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.8.4210

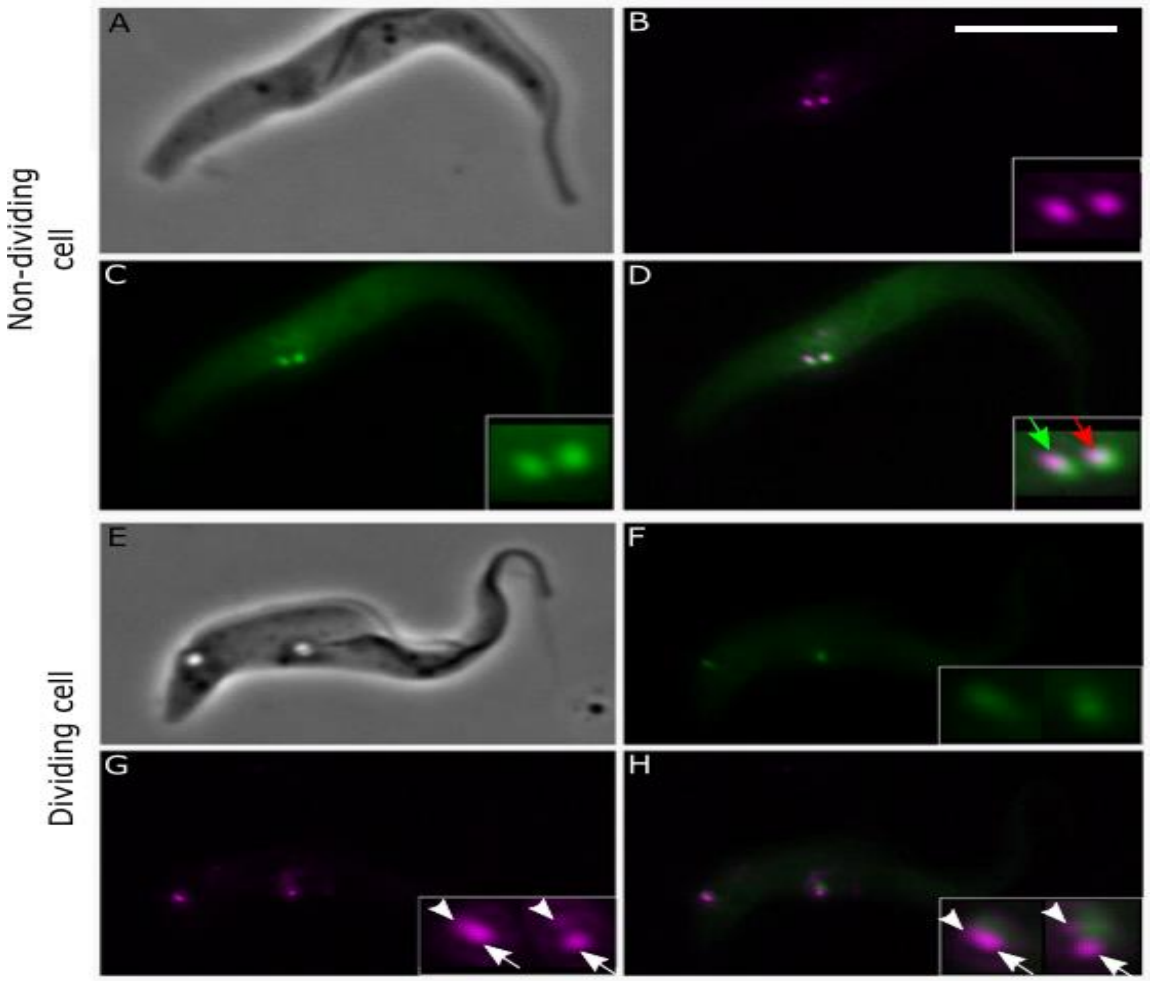


Figure 30: Localisation of Tb927.8.4210

Co-localisation of Tb927.8.4210 (mNG) with the SAS-6 mCherry (magenta) cell line to both the mature and pro-basal body in whole cells. No localisation was observed in detergent extracted cells. Localisation is seen at the old basal body pair (green arrow) and the new basal body pair (red arrow). Scale bar = 10μm.

Tb927.11.13300

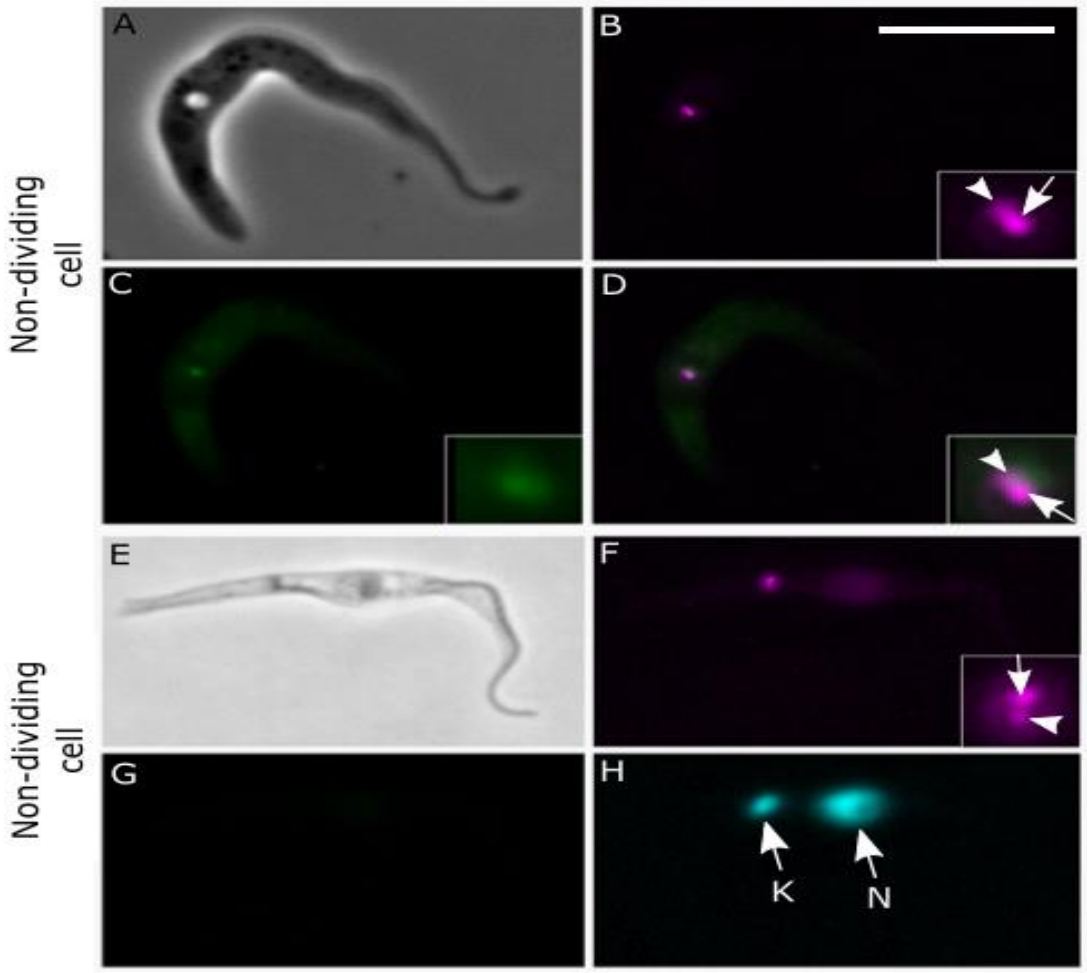


Figure 31: Localisation of Tb927.11.13300

Co-localisation of Tb927.11.13300 (mNG) with the SAS-6 mCherry (magenta) cell line with a faint localisation to the distal section of the mature (arrow head) and pro-basal body (arrow) in live 1K1N cells (A-D). No localisation was observed in detergent-extracted cells (E-H). Scale bar = 10μm.

Tb927.10.4990

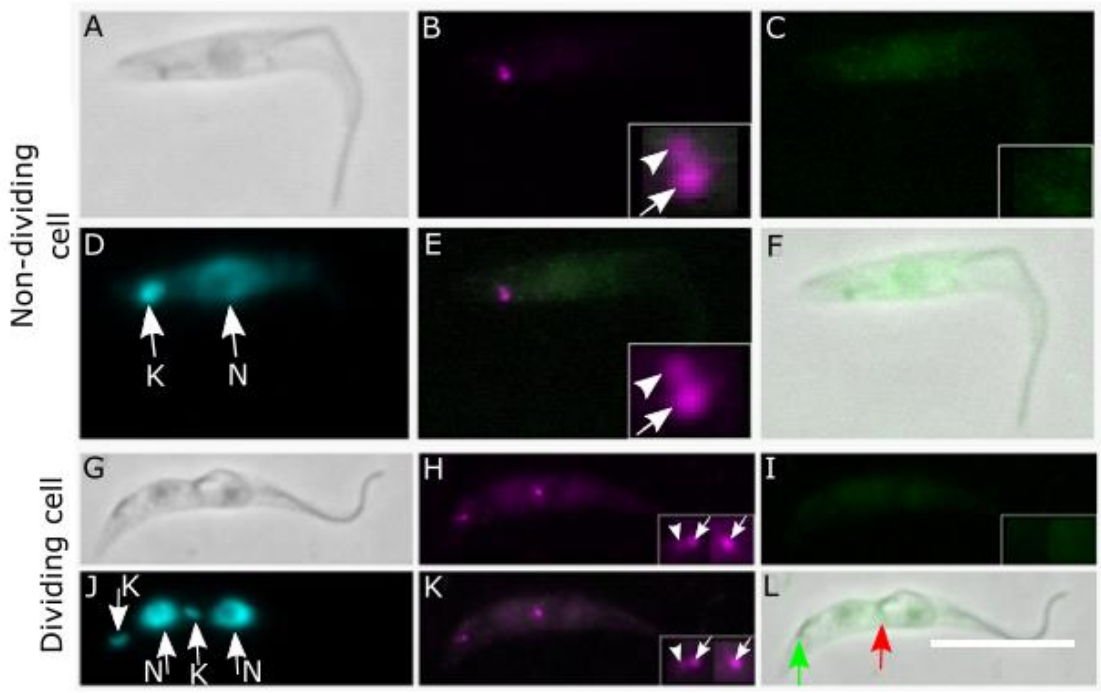


Figure 32: Localisation of Tb927.10.4990

No co-localisation of Tb927.10.4990 with the SAS-6 mCherry (magenta) cell line was observed in detergent-extracted cells in either 1K1N (A-F) or 2K2N (G-L) cells, L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.13610

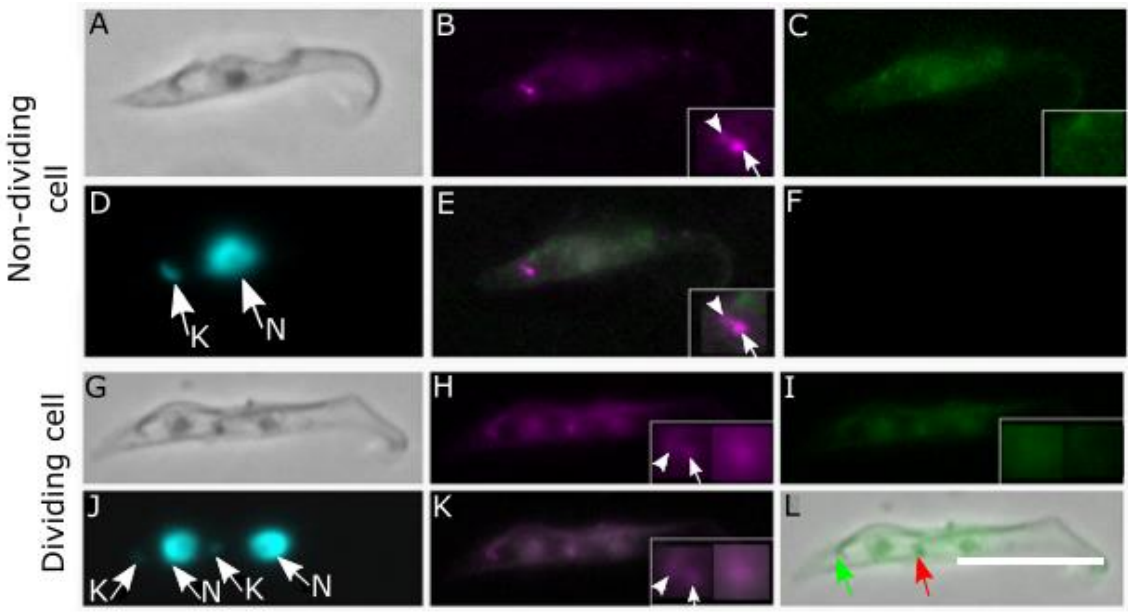


Figure 33: Localisation of Tb927.10.13610

No co-localisation of Tb927.10.13610 with the SAS-6 mCherry (magenta) cell line was observed in detergent-extracted cells in either 1K1N (A-F) or 2K2N (G-L) cells, L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.10.12590

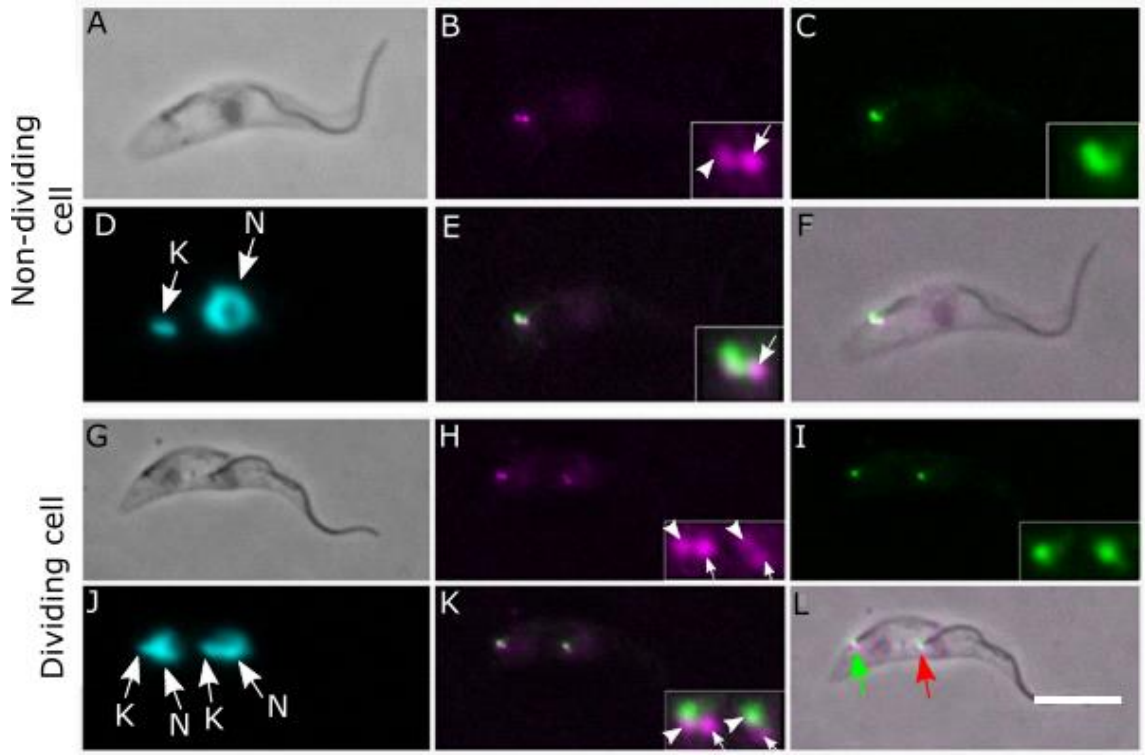


Figure 34: Localisation of Tb927.10.12590

Co-localisation of Tb927.10.12590 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature basal body (arrow head) only in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 5μm.

Tb927.1.3560

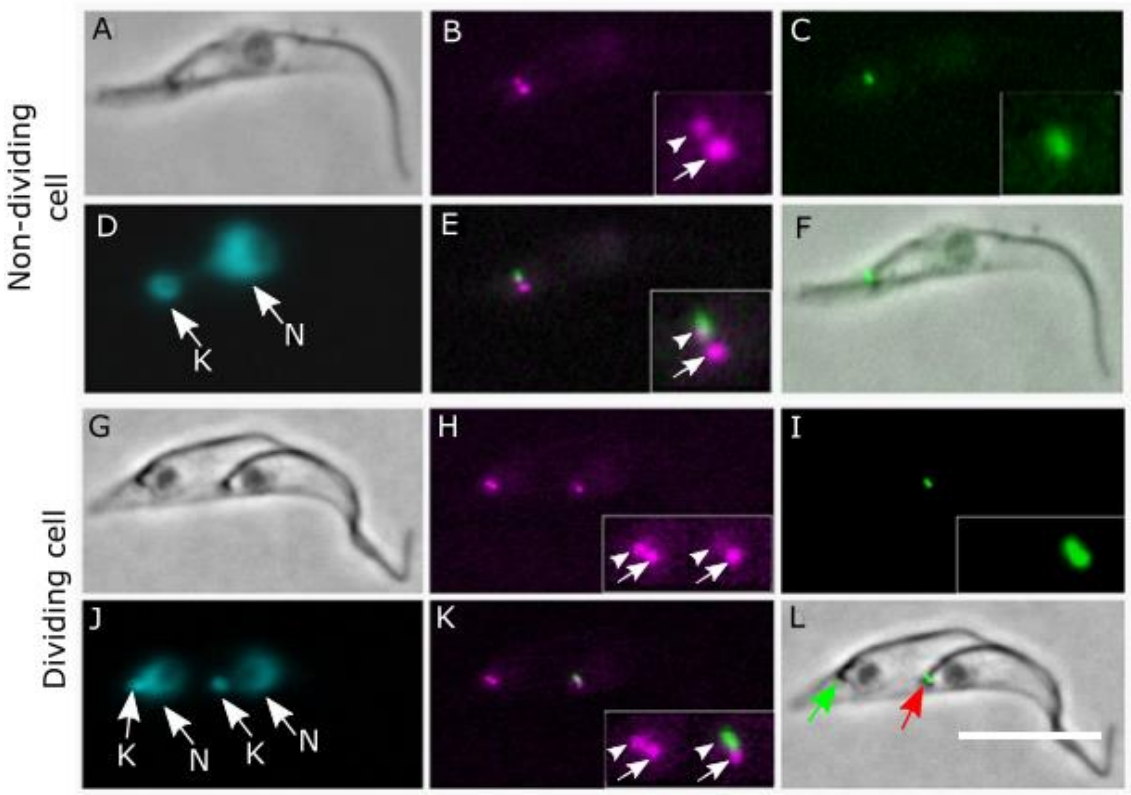


Figure 35: Localisation of Tb927.1.3560

Co-localisation of Tb927.1.3560 (mNG) with the SAS-6 mCherry (magenta) cell line to the mature basal body (arrow head) in 1K1N cells (A-F) and to the old mature basal body only in 2K2N cells (arrow head) (G-L). L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.7.5190

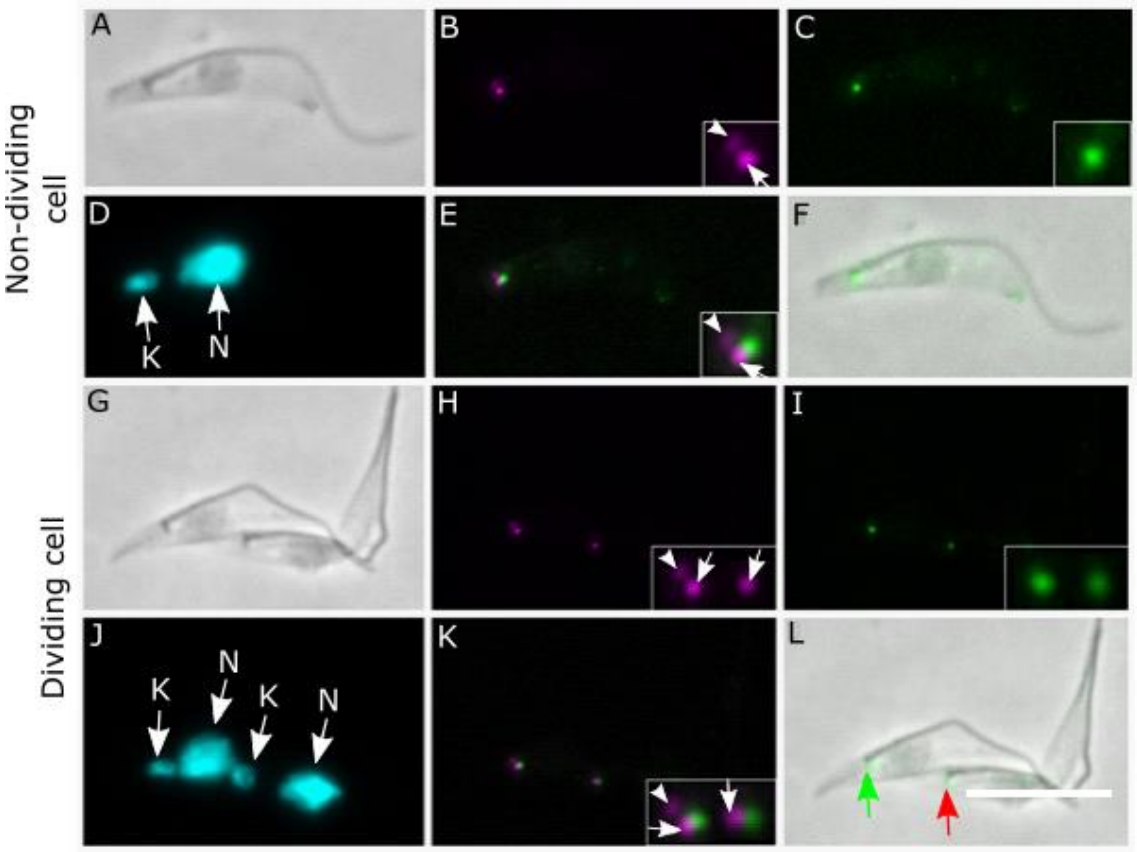


Figure 36: Localisation of Tb927.7.5190

Co-localisation of Tb927.7.5190 (mNG) with the SAS-6 mCherry (magenta) cell line to the distal section of the pro-basal body (arrow) only in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.7.4130

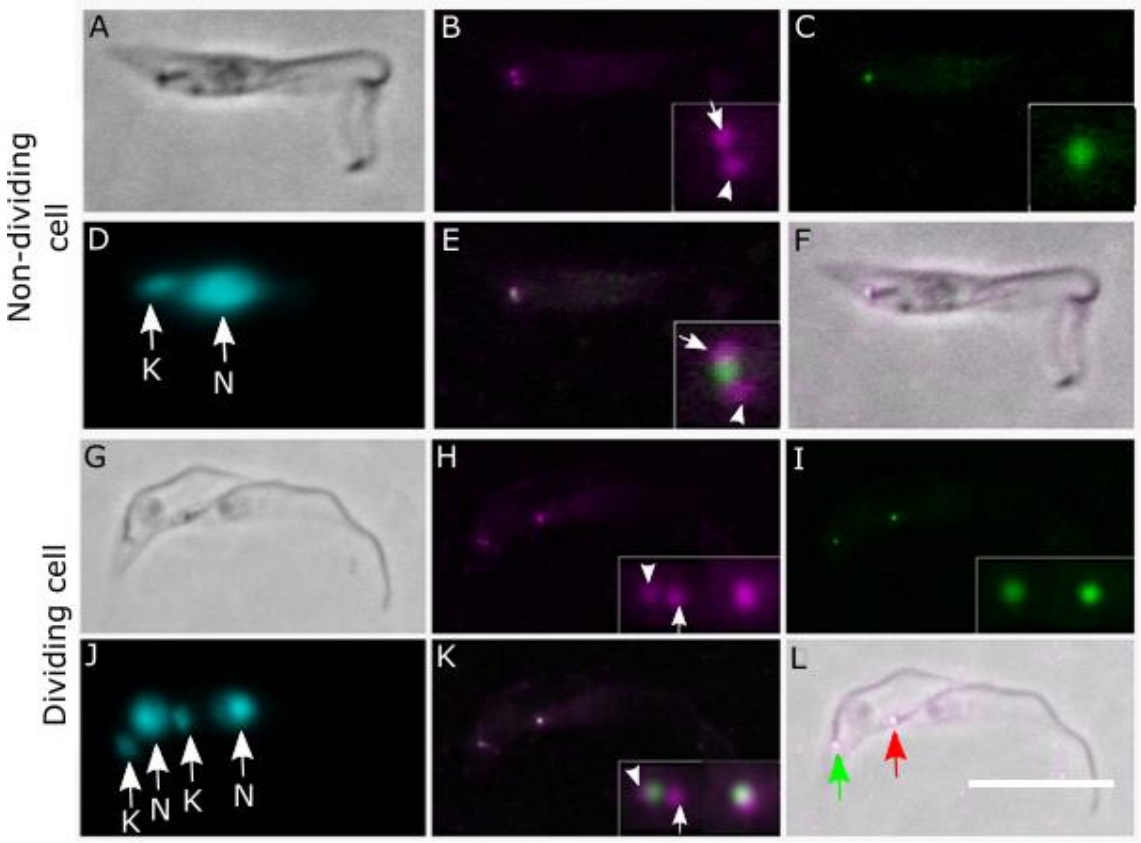


Figure 37: Localisation of Tb927.7.4130

Co-localisation of Tb927.7.4130 (mNG) with the SAS-6 mCherry (magenta) cell line to in-between the mature (arrow head) and pro-basal body (arrow) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.11.2700

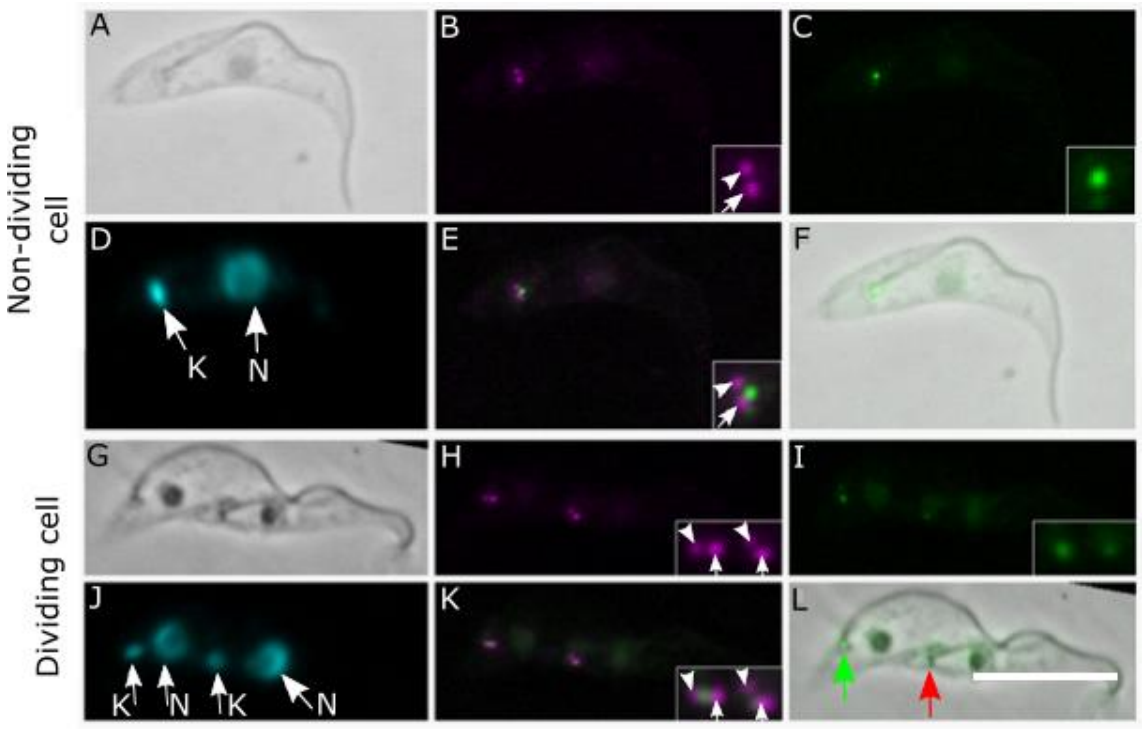


Figure 38: Localisation of Tb927.11.2700

Co-localisation of Tb927.11.2700 (mNG) with the SAS-6 mCherry (magenta) cell line to in-between the mature (arrow head) and pro-basal body (arrow) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

Tb927.11.15450

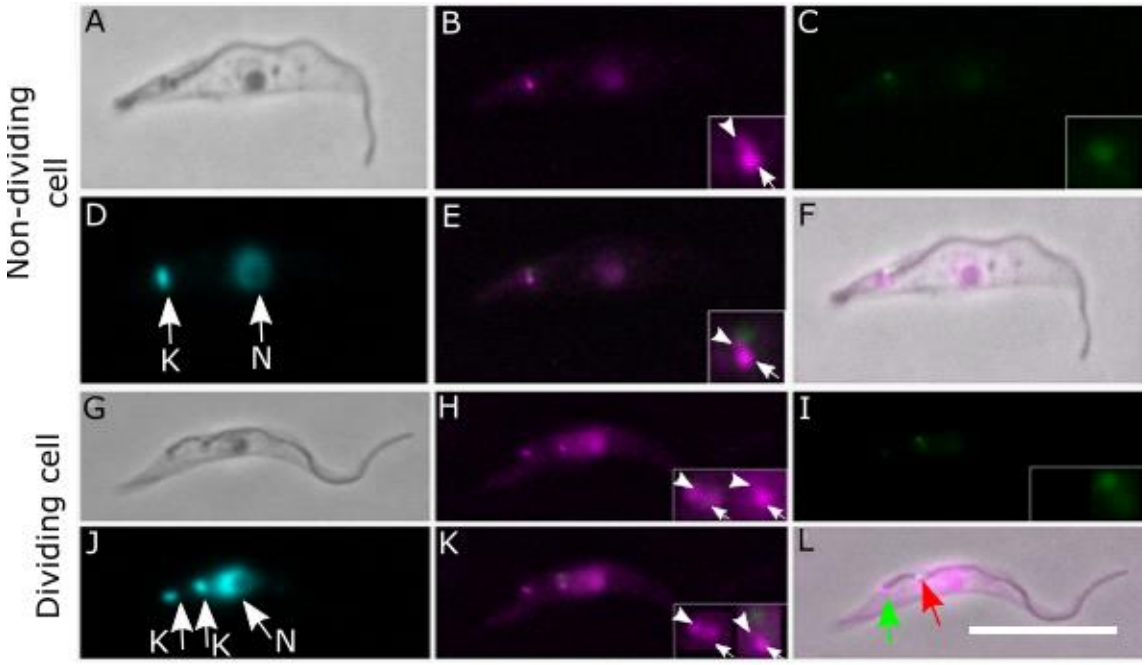


Figure 39: Localisation of Tb927.11.15450

No co-localisation of Tb927.11.15450 (mNG) with the SAS-6 mCherry (magenta) cell line was observed in detergent-extracted cells. Localisation was observed surrounding the areas of the mature basal body (arrow head) in 1K1N (A-F) and 2K2N (G-L) cells. L: new basal body pair (green arrow) and old basal body pair (red arrow). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

3.3 Summary of screened proteins

Basal body localisation differs throughout the cell cycle

Three proteins localised differently throughout the cell cycle, with proteins localising to the PBB only in 1K1N cells and to both the MBB and PBB in 2K2N cells, or in the reverse order.

Localisation to both the mature (BB1) and pro-basal body (BB2)

Six proteins co-localised with the MBB and PBB in G1 (1K1N) cells and remained present at both the MBB and PBB in dividing (2K2N) cells.

Absent localisation in extracted cells

Four proteins did not localise to the BBs or elsewhere in detergent-extracted cells, suggesting these proteins are not tightly bound to the BB.

Localisation observed in-between mature (BB1) and pro basal body (BB2)

Two proteins localised between the MBB and PBB in 1K1N and 2K2N cells.

Localisation to the pro-basal body only throughout the cell cycle

One protein localised to the distal section of the PBB only throughout the cell cycle.

Localisation to the mature basal body (BB1 and BB3) throughout the cell cycle

One protein localised to the distal section of the MBB in 1K1N cells and localised to both MBBs in 2K2N cells.

No localisation to the basal bodies

One protein did not co-localise to SAS-6 and instead located distal to the BBs. This was not considered a BB protein.

Localisation to the mature basal body (BB1) only throughout the cell cycle

One protein localised to the distal section of the MBB in 1K1N cells, however was only localised to the MBB of the old BBP and was not observed on the new BBP. This protein had a unique localisation with no other BB protein in TrypTag showing a similar localisation.

Overall, 19 proteins were tagged with the mNG fluorophore and were assessed for their localisation to the BBs. Localisation of each protein was analysed on the wide-field fluorescent microscope and uncovered characteristics were noted. The screen highlighted one protein, Cep164C (Tb927.1.3560) which demonstrated an unusual localisation pattern (figure 35). No other BB protein has shown a similar localisation pattern, regarding a protein which is located to the old MBB, but not located on the younger MBB.

3.4 Analysing the dataset generated in this project

Whilst generating and analysing cell lines, bio-informatic analysis on all the 267 putative BB proteins identified by The Wellcome Trust Tryptag project (Dean *et al,* 2017) was conducted. Each localisation and accession number identified by the TrypTag project can be accessed through http://tryptag.org/.

Table 9: Organisms investigated in the Orthofinder against the *T. brucei* basal body proteins

Group of Organism	Name of Organism	Organisms that contain cilia
Plantae Group	*Arabidopsis thaliana, Populus trichocarpa, Oryza sativa, Physcomitrella patens, Ostreococcus tauri, Chlamydomonas reinhardtii, Cyanidioschyzon merolae*	*Physcomitrella patens, Chlamydomonas reinhardtii,*
Excavata Group	*Trypanosoma brucei, Leishmania mexicana, Naegleria gruberi, Giardia intestinalis, Trichomonas vaginalis*	All contain cilia
Chromalveolata Group	*Plasmodium berghei, Plasmodium falciparum, Theileria annulata, Toxoplasma gondii, Cryptosporidium Parvum, Tetrahymena thermophila, Paramecium tetraurelia, Phaeodactylum tricornutum, Thalassiosira pseudonana, Phytophthora sojae, Aureococcus anophagefferens*	*Plasmodium berghei, Plasmodium falciparum, Toxoplasma gondii, Tetrahymena thermophila, Paramecium tetraurelia, Thalassiosira pseudonana, Phytophthora sojae, Aureococcus anophagefferens*
Holozoa Group	*Homo sapiens, Gallus gallus, Takifugu rubripes, Danio rerio, Ciona intestinalis, Strongylocentrotus purpuratus, Drosophila melanogaster, Apis mellifera, Caenorhabditis elegans, Capitella sp, Schistosoma mansoni, Lottia gigantea, Nematostella vectensis, Trichoplax adhaerens, Monosiga brevicollis*	All contain cilia
Fungi Group	*Neurospora crassa, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus nidulans, Ustilago maydis, Rhizopus delemar, Batrachochytrium dendrobatidis, Encephalitozoon cuniculi*	Only *Batrachochytrium dendrobatidis*

Table 9: Organisms investigated in the Orthofinder against the *T. brucei* basal body proteins

Amebozoa Group	*Entamoeba histolytica, Dictyostelium discoideum*	All contain cilia

Additional information on the orthologues was provided through a previously run OrthoFinder database by Richard Wheeler (Oxford University, 2017 personal communication). This database compared *T. brucei* proteins against 48 different organisms included in table 9. Organisms included either possessed a cilium *Homo sapiens, Giardia intestinalis* etc) or lacked a cilium (*Arabidopsis thaliana, Saccharomyces cerevisiae* etc).

3.5 Localisation pattern and differences in fluorophore intensity

Detection of the mNG fluorophore at the BBs by fluorescence microscopy identifies proteins that locate to the MBB, PBB or to a structure associated with the BBs. Proteins containing one focus only localises to one of the BB organelles (figure 40; A-C), whereas proteins with two foci localise to both BBs (figure 40; D-F).

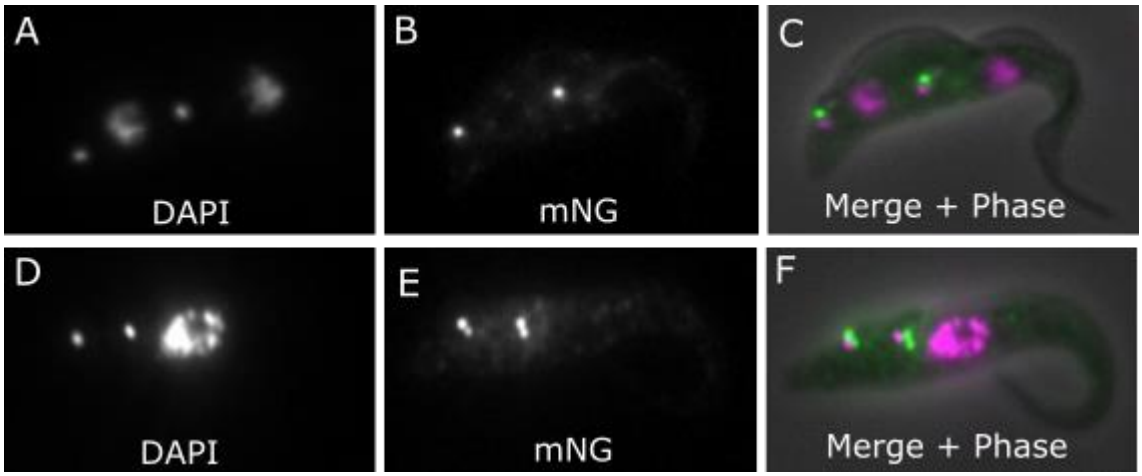


Figure 40: Localisation of proteins with a 1-2 or 2-4 foci.

A dividing *Trypanosoma brucei* cell expressing a basal body protein localising to either the mature or pro-basal body on each basal body pair (A-C) with a 1-2 foci, or to both the mature and pro-basal bodies on each basal body pair (D-F) with a 2-4 foci. Image taken from the TrypTag project - Dean *et al* (2017)

Using the criteria from above, all annotated BB proteins on TrypTag were re-analysed to determine in more detail the localisation pattern for each of the 267 proteins. The analysis took into consideration the localisation pattern in both G1 and dividing cells. In G1 cells, only one BBP is present. During cell division the PBB matures after which two new PBBs are generated to create two BBPs. Proteins which contain a localisation pattern of one focus at the BB in G1 cells, could then divide to contain two foci (1-2) in a single cell (figure 40; A-C) Whereas those containing two foci in G1 cells could divide to contain four foci (2-4) in a single cell (figure 40; D-F).

The most frequent localisation pattern observed in the 267 BB proteins was those containing a 2-4 foci, localising at both the MBB and PBB organelle (figure 41). Less frequent was the 1-2 localisation pattern at 27%. For 8% of proteins, identification of the localisation pattern was not possible due to the image quality and availability of images on TrypTag (figure 41).

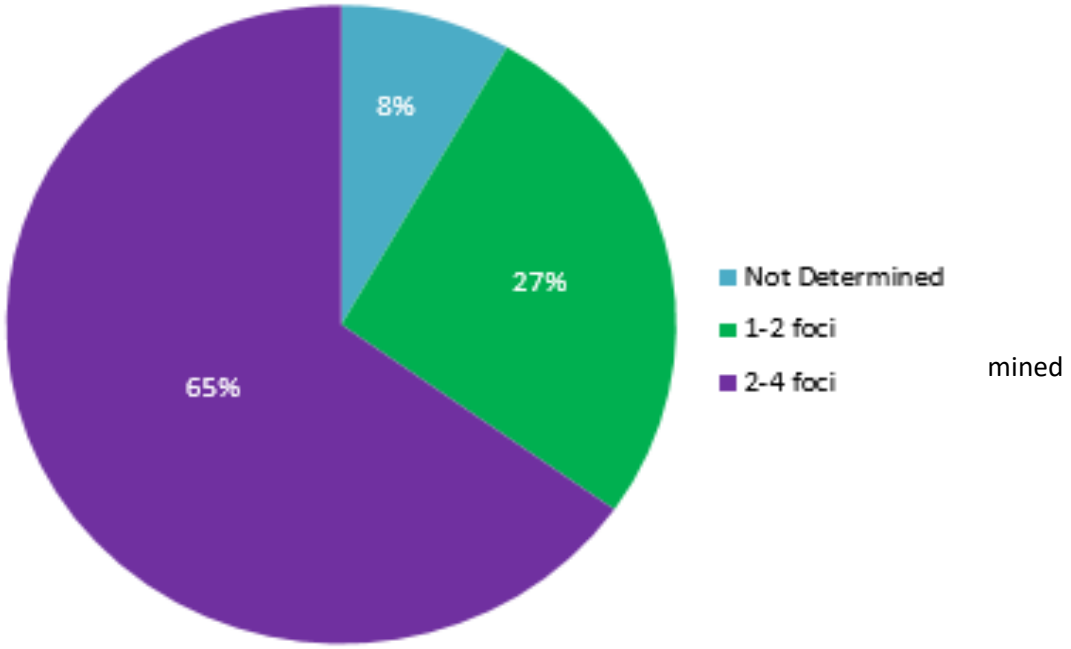


Figure 41: Localisation pattern of the 267 basal body proteins

The majority of basal body proteins (65%) had a 2-4 foci pattern, 27% had a 1-2 foci pattern with the remaining 8% of proteins not determined.

Analysis of the 267 proteins led to the identification of observable differences in intensity between the MBB and the PBB, or intensity differences were observed in a cell cycle manner. An example is shown in figure 42, where the MBB had a higher intensity than the PBB.

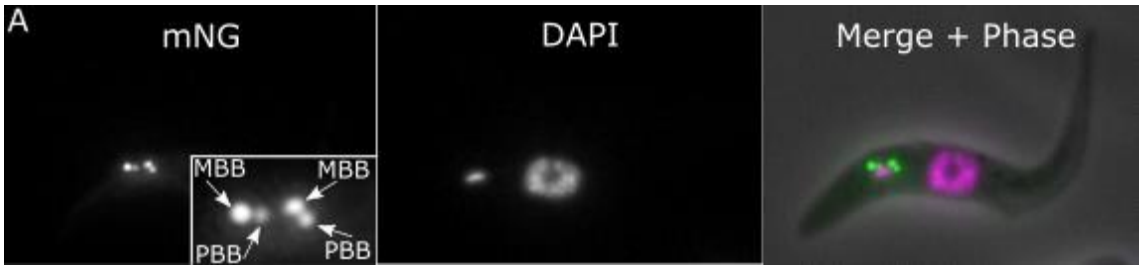


Figure 42: Intensity variation in the mature and pro-basal body.

Intensity differences were observed between the mature and pro-basal body in the 267 suggested basal body proteins in the TrypTag project. A higher level of intensity is observed on the mature basal body than the pro-basal body (A; inset). Image taken from the TrypTag project - Dean *et al* (2017).

40% of suggested BB proteins did not show any fluorophore intensity variations between the MBB and the PBB (figure 43; B-C). 20% of proteins in the dataset had an increased fluorophore intensity to the MBB, in comparison to the PBB (figure 43; D-E). Whereas only 3% of proteins had an increased fluorophore intensity to the PBB, in comparison to the MBB (figure 43; F-G). 10% of BB proteins were not determined due to the quality and availability of images on TrypTag (figure 43;A).

27% of BB proteins had a 1-2 foci only, with 69% of these proteins failing to show a fluorophore intensity difference between the old and new BBPs. 22% of these proteins had an increased fluorophore intensity to the MBB, in comparison to the newly MBB prior to BB duplication. This differed greatly to the 1.5% of BB proteins which were seen to have an increased fluorophore intensity on the newly MBB, in comparison to the old MBB. 2 proteins had a varied fluorophore intensity, dependent on the stage of the cell cycle. Only one protein had a localisation pattern at the old MBB only, with no localisation to the new BBP (Tb927.1.3560).

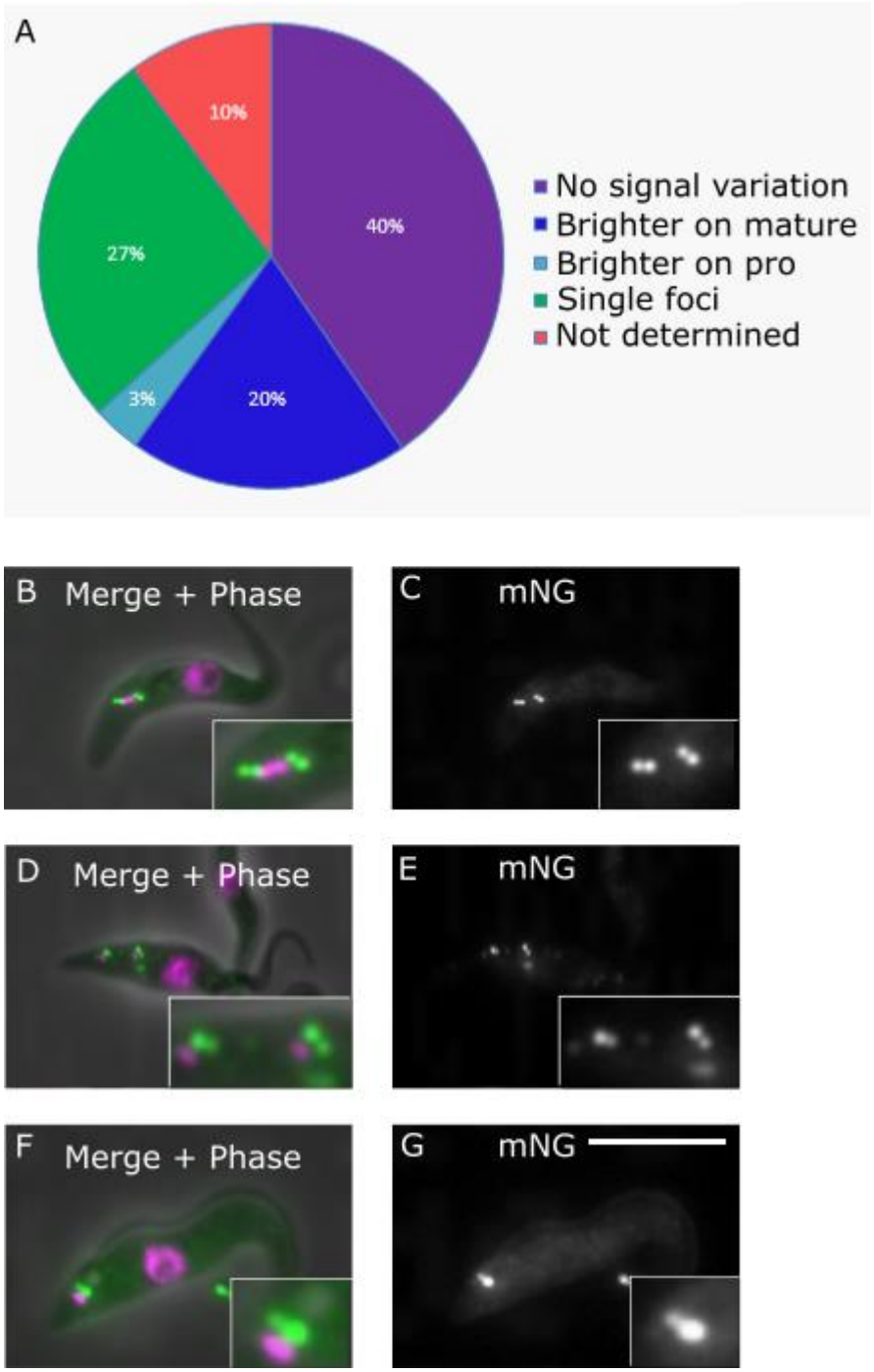


Figure 43: Protein intensity variation of the basal body proteins.

A) A break-down of the signal variation in the basal body proteins, with 40% of proteins having no signal variation (B-C), 20% of proteins having a brighter signal intensity on the mature basal body (D-E) and 3% of proteins having a brighter signal intensity on the pro-basal body (F-G).

3.6 Additional protein localisations

Analysing the localisation of the BB proteins on TrypTag allowed the identification of those proteins which localise to organelles other than the BBs. 84 (31%) of proteins exclusively localised to the BBs or surrounding structures only (table 10), suggesting that these proteins have a functional role to the BB. The remaining 69% of proteins localised to the BBs and other structures in the cell. 32% of proteins shared a localisation pattern to at least one other structure, with 62/267 proteins localising to the cytoplasm and the BBs. The TZ, flagellar tip and the axoneme are other examples of where the proteins were observed in those which shared two localisations. The remaining proteins contained multiple localisations, but at a much lower frequency (table 10).

19% of proteins were present in three different organelles (including the BBs), with a variety of organelles or structures identified, with the most frequent being at the cytoplasm and flagellar cytoplasm (9) (table 10). 12% of proteins were present in four different localisations (including the BBs) with the cytoplasm, nuclear lumen and flagellar cytoplasm being the most common (8). Only 4% of proteins shared a localisation in five different organelles or structures, with only one protein seen for each category (table 10). Finally, one protein was seen to be positive for six localisations, and one protein positive for seven localisations.

Table 10: Analysis of the additional localisations to where the 267 basal body proteins localise to in the T. brucei cell. Taken from TrypTag – Dean et al (2017)

Localisation	Additional localisation	Number of proteins
One localisation only		
Basal body only	-	84
Two localisations		
Basal body	Cytoplasm	62
Basal body	Transition zone	7
Basal body	IFT	1
Basal body	Mitochondrion	2
Basal body	Flagellar tip	1
Basal body	Microtubule quartet	1
Basal body	Axoneme	3
Basal body	Paraflagellar rod	2
Basal body	Cell tip	2
Basal body	Tripartite attachment complex	1
Basal body	Hook complex	1
Basal body	Pellicular membrane	1
Basal body	Cortical cytoskeleton	2
Three localisations		
Basal body	Cytoplasm + flagellum	7
Basal body	Cytoplasm + spindle pole	1
Basal body	Cytoplasm + cell tip	3
Basal body	Cytoplasm + transition zone	2
Basal body	Cytoplasm + endoplasmic reticulum	1
Basal body	Cytoplasm + flagellar pocket	2
Basal body	Cytoplasm + endocytic	8
Basal body	Cytoplasm + flagellar cytoplasm	9
Basal body	Cytoplasm + cortical cytoskeleton	2
Basal body	Cytoplasm + axoneme	4
Basal body	Cytoplasm + IFT particle	2
Basal body	Hook complex + cell tip	1
Basal body	Cytoplasm + hook complex	1
Basal body	Cell tip + flagellum attachment zone	1
Basal body	Cytoplasm + flagellar attachment zone	1
Basal body	Cytoplasm + mitochondrion	1
Basal body	Cortical cytoskeleton + cell tip	1
Basal body	Cytoplasm + flagellar tip	1
Basal body	Cytoplasm + flagellar connector	1
Basal body	Flagellar pocket + transition zone	1
Basal body	Cytoplasm + nuclear lumen	1

Table 10: Analysis of the additional localisations to where the 267 basal body proteins localise to in the T. brucei cell. Taken from TrypTag – Dean et al (2017)

Localisation	Additional localisation	Number of proteins
Four localisations		
Basal body	Cytoplasm + flagellum + cell tip	1
Basal body	Cytoplasm + cytoskeleton + flagellum	1
Basal body	Cytoplasm + axoneme + nucleus	1
Basal body	Cytoplasm + flagellum + IFT	1
Basal body	Cytoplasm + flagellum + flagellar tip	1
Basal body	Cytoplasm + cell tip + hook complex	1
Basal body	Cytoplasm + flagellum + axoneme	1
Basal body	Cytoplasm + nuclear lumen + flagellar cytoplasm	8
Basal body	Cytoplasm + axoneme + nucleoplasm	1
Basal body	Cytoplasm + flagellum + hook complex	1
Basal body	Cytoplasm + nuclear lumen + flagellum	1
Basal body	Cytoplasm + flagellar tip + cell tip	1
Basal body	Cytoplasm + flagellar cytoplasm + nucleoplasm	2
Basal body	Cytoplasm + axoneme + flagellar tip	1
Basal body	Cell tip + cleavage furrow + flagellar attachment zone	1
Basal body	Flagellum + cleavage furrow + flagellum attachment zone	1
Basal body	Cytoplasm + flagellum + endocytic	1
Basal body	Cytoplasm + hook complex + plasma membrane	1
Basal body	Cytoplasm + transition zone + endocytic	1
Basal body	Cytoplasm + flagellar cytoplasm + flagellar pocket	1
Basal body	Cytoplasm + nucleus + flagellar cytoplasm	1
Basal body	Cytoplasm + flagellar cytoplasm + spindle pole	1
Basal body	Cytoplasm + flagellar cytoplasm + mitochondrion	1
Five localisations		
Basal body	Cytoplasm + nucleus + nucleolus + kinetoplast	1
Basal body	Cytoplasm + nuclear lumen + flagellar cytoplasm + cell tip	1
Basal body	Cytoplasm + cell tip + axoneme + flagellar attachment zone	1
Basal body	Cell tip + nucleus + axoneme + flagellar attachment zone	1
Basal body	Cytoplasm + nucleoplasm + nucleolus + flagellar cytoplasm	1
Basal body	Cytoplasm + cell tip + cytoskeleton + flagellar attachment zone	1
Basal body	Cytoplasm + cell tip + hook complex + flagellar tip	1
Basal body	Cytoplasm + flagellar cytoplasm + kinetoplast + mitochondrion	1
Basal body	Cytoplasm + flagellar cytoplasm + nuclear lumen + endocytic	1
Basal body	Cell tip + nucleus + flagellar connector + axoneme	1
Basal body	Cytoplasm + flagellum + flagellar cytoplasm + nucleus	1
Six localisations		
Basal body	Cytoplasm + flagellar cytoplasm + nuclear lumen + cell tip + nucleoplasm	1
Seven localisations		
Basal body	Cytoplasm + nucleoplasm + nucleolus + cell tip + flagellar cytoplasm + endocytic	1
Not determined	-	2

3.7 Hypothetical proteins and PFAM domains

After observing the localisation of the proteins, it was then questioned how many had been previously identified (from previous experiments) and how many were hypothetical proteins. Information was collected from the TriTrypDB website, which states whether a protein is hypothetical. 53% of proteins contained a descriptive annotation which included proteins such as centrosomal proteins, BB proteins, tubulin and kinesins etc. The remaining 45% were classed as hypothetical with little information available (figure 44). Only one protein was labelled as not determined due to the lack of available information.

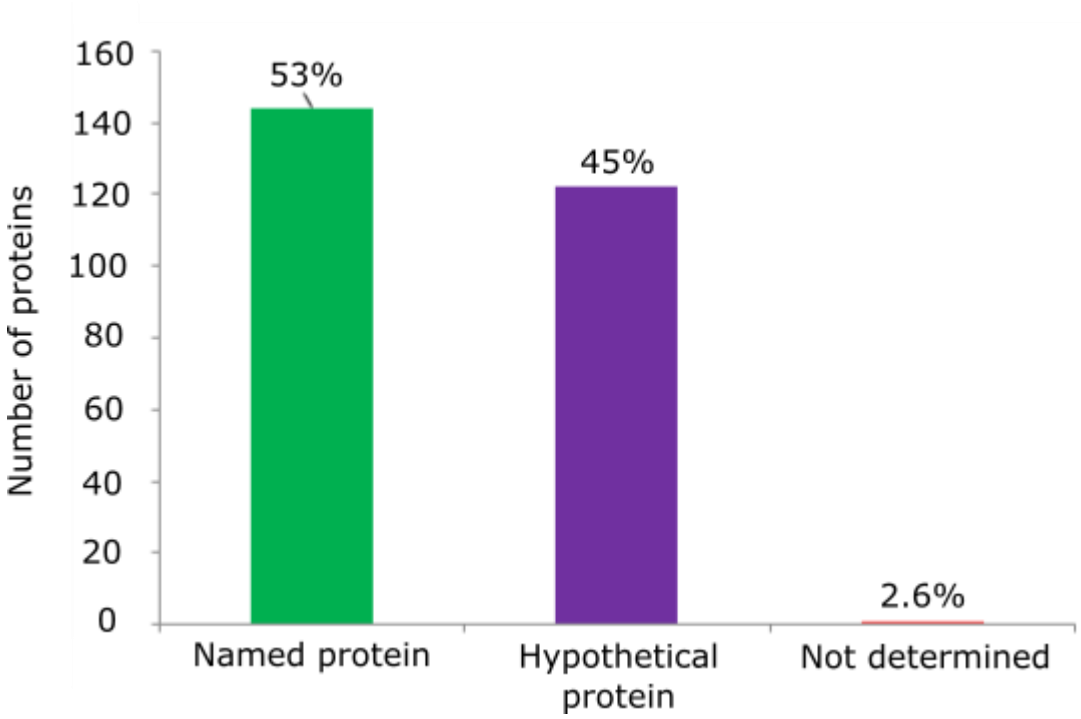


Figure 44: Hypothetical or known proteins

TriTrypDB description of each of the 267 basal body proteins identifying either named or hypothetical proteins

For more information on the BB proteins, a bio-informatic search was conducted through the PFAM database, using the Hidden Markov Modelling technique, to analyse protein sequences and identify known domains. From the 267 BB proteins, 59% did not identify any known domains, whereas 41% contained sections in the sequence which were recognised as specific domains (figure 45).

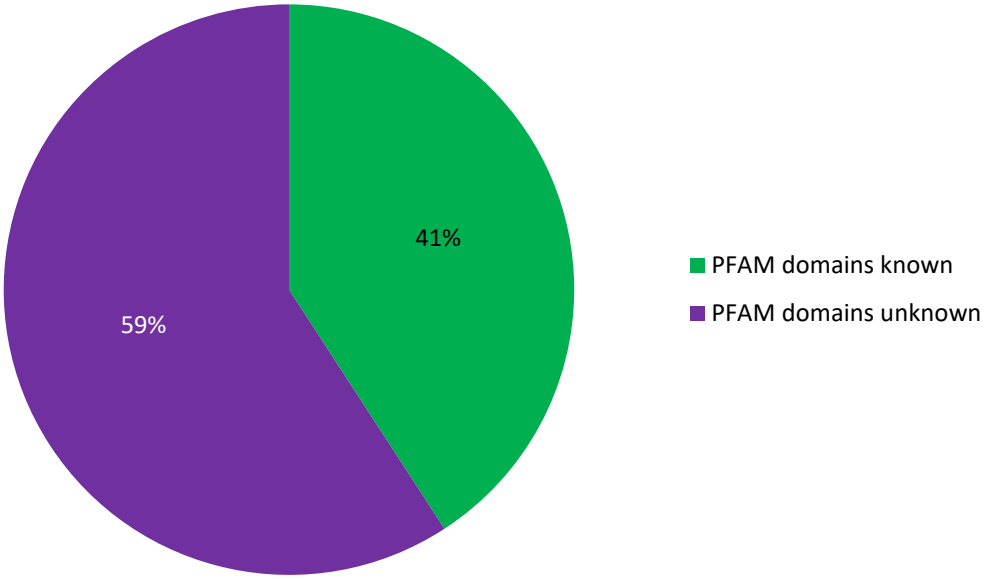


Figure 45: Basal body proteins with known or unknown PFAM domains

Although 41% of proteins identified a known domain, no domain was overly frequent (table 11), with the most frequent domain being the protein kinase domain at 6.5%. Other domains containing more than one protein included leucine-rich repeat (5%), zinc-finger (5%), kinase (4.5%) and tetratricopeptide (4.5%) domains (table 11). The remaining proteins contained three or less proteins in a group, with the majority of known domains containing only one protein.

Table 11: PFAM domains in the 267 basal body proteins, as determined through the Protein Families Database (PFAM)

PFAM domains	Number of proteins with domain
Protein kinase domain	7
Leucine-rich repeat	6
Zinc_Finger C2H2	6
Kinesin	5
Tetratricopeptide repeat, DNAj domain	5
EF-hand domain pair	3
ADP-ribosylation factor family	3
Domain of unknown function	3
WW domain	3
2-oxoacid dehydrogenases, dehydrogenase E1 component, pyrimidine binding domain	2
FOP N terminal dimerisation domain	2
MORN repeat	2
WD domain, G-beta repeat	2
WD40	2
Tubulin/FtsZ family, GTPase domain	2
Dynien light chain type 1/2 domain	1
Ras of Complex, Roc, domain of DAPkinase	1
Cyclin-dependant kinase regulatory subunit	1
SpoU methylase	1
Ras family	1
Spc97/Spc98 family	1
Casein kinase II regulatory subunit	1
Glycosyltransferase family 25 (LPS biosynthesis protein)	1
CEP76 C2 domain	1
Serine-threonine protein phosphatase, Metallo-dependent phosphatase	1
Cilia and flagella associated protein, WD40/YVTN	1
Iguana/Dzip1-like DAZ-interacting protein	1
ATPase family associated with various cellular activities (AAA). Vps4 C terminal oligomerisation domain	1
Carboxypeptidase Taq (M32) metallopeptidase	1
Dehydrogenase E1 component, Transketolase pyrimidine binding domain, 2-oxoglutarate dehydrogenase C terminal	1
Calpain family cysteine protease	1
Ciliary basal body-associated, B9 protein	1
CEP19-like protein	1
Tir chaperone protein (CesT) family	1
Acyltransferase	1
Dual specificity phosphatase, catalytic domain	1
tRNA anti-codon, tRNA-synt 2	1
PB1 domain	1
Centrosomal protein 95kDa	1
microtubule-binding protein MIP-T3	1
5'-3' exonuclease, N-terminal resolvase-like domain	1

Table 11: PFAM domains in the 267 basal body proteins, as determined through the Protein Families Database (PFAM)

PFAM domain	Number of proteins with domain
Growth-arrest specific micro-tubule binding	1
Roadblock/LC7 domain	1
KMP11 domain	1
Kinetoplast membrane protein 1	1
Centriolar protein SAS N-terminal	1
Glyceraldehyde-3-phosphate, C-terminal domain	1
MMR HSR1, TGS	1
Calmodulin-binding	1
First C2 domain of RPGR-interacting protein 1	1
AhpC/TSA family, C-terminal domain of 1-Cys peroxiredoxin	1
MaoC like domain	1
Centrosomal spindle body, CEP44	1
SKP1, BTB/POZ, Potassium channel tetramerisation, Regulator of chromosome condensation 1	1
Hydroxymethylglutaryl-coenzyme A reductase	1
IFT complex B, subunit 20	1
NAD binding domain of 6-phosphogluconate dehydrogenase	1
Plus-3, zf-C3HC4 3	1
Biotin lipoyl, E3 binding, 2-oxoacid_dh	1
Anaphase-promoting complex WD40 domain	1
Sec3 C	1
FAD binding domain, Oxidoreducatase NAD-binding domain	1
3'5'-cyclic nucleotide phosphodiesterase	1
STPPase N, Metallophos	1
Rhodanese-like domain	1
RNA recognition motif (RRM, RBD)	1
NLI interacting factor-like phosphatase	1
Cullin family, neddylation domain	1
CTP_synth_N, GATase	1

3.8 Conserved and kinetoplastid specific proteins

After identification of the PFAM domains, it was investigated into which proteins were conserved and which were kinetoplastid specific. The data collected from the OrthoFinder and from reciprocal best hit BLASTs (Benson *et al,* 2013), showed that 47% of proteins in the BB dataset were conserved outside of kinetoplastids. The remaining 53% of proteins were kinetoplastid specific (figure 46) and were only identified in other kinetoplastid species.

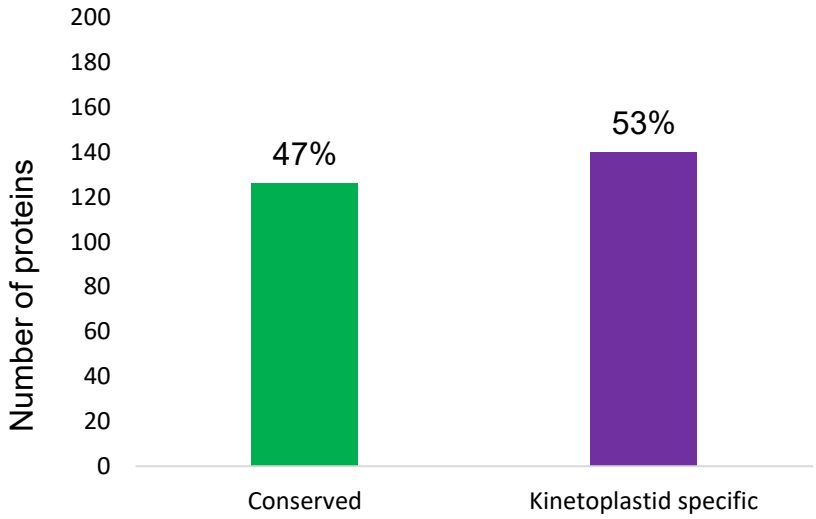


Figure 46: Conserved or kinetoplastid specific proteins

Based on the results from OrthoFinder, the 267 basal body proteins were analysed to determine how many proteins were conserved outside of kinetoplastids (47%) and how many were kinetoplastid specific (53%)

Further investigation identified that 78 proteins were conserved within the kinetoplastid species, however 59 proteins were identified to *T. brucei* species only (figure 47).

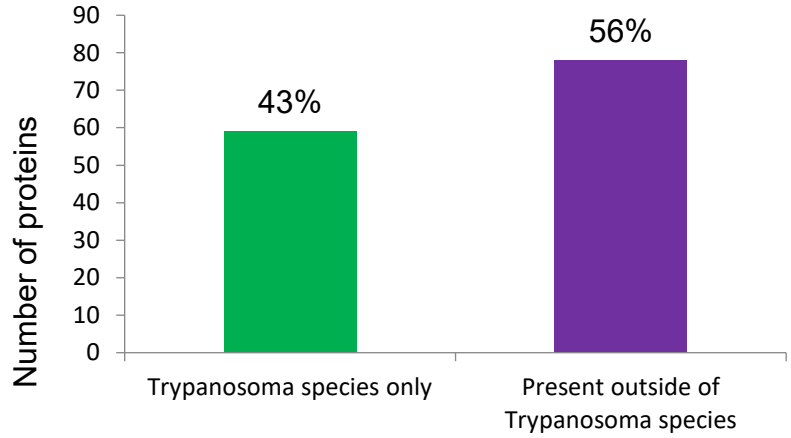


Figure 47: Percentage of BB proteins present in the Trypanosoma species only

The basal body proteins which were kinetoplastid specific were further analysed to determine how many were present in Trypanosoma species only (43%), and how many were present outside of Trypanosoma species (56%)

Analysing the presence and absence of orthologues in other organisms for the BB proteins in *T. brucei* showed that the highest frequency of orthologues was observed in the *Homo sapiens* group. Percentages were calculated by taking the number of proteins present, dividing them by the total number of proteins (267) and multiplying by 100. In the *Homo sapiens* group, only 104 (38%) proteins out of the total 267 were present (table 12).

Table 12: Conservation of basal body proteins in investigated organisms

Organism	**Present**	**Not detected**	**Not Determined**
Trypanosoma brucei	267	0	0
Homo sapiens	104 (38%)	162	1
Tetrahymena thermophila	99 (37%)	167	1
Danio rerio	98 (36%)	168	1
Mus musculus	95 (35%)	171	1
Xenopus laevis	93 (34%)	173	1
Chlamydomonas reinhardtii	86 (32%)	180	1
Drosophila melanogaster	82 (30%)	184	1
Caenorhabditis elegans	74 (27%)	192	1
Trichomonas vaginalis	71 (26%)	195	1
Arabidopsis thaliana	64 (23%)	202	1
Plasmodium falciparum	62 (23%)	204	1
Saccharomyces cerevisiae	49 (18%)	217	1

This highlighted that the 162 of BB proteins (62%) present in *T. brucei* are not present within this group. *Tetrahymena thermophila* had the next highest number of BB orthologues, with 99 proteins present (37%) (table 12). Analysis showed that other than *Tetrahymena thermophila,* all other organisms containing a higher level of conservation were multicellular eukaryotes. Organisms that have a lower level of conservation between BB proteins include *Arabidopsis thaliana, Plasmodium falciparum* and *Saccharomyces cerevisiae* (table 12), which show a 23%, 23% and 18% conservation respectively.

3.9 36 proteins are present in all investigated organisms

Analysis of the proteins identified that 47% of proteins were conserved outside of the kinetoplastid species. Bio-informatic results showed that 36 (13%) suggested BB proteins were present in each organism analysed in this project (table 13).

Table 13: Information collected on the 36 basal body proteins that were present in each organism investigated in this study.

Accession number	Known as	Foci	PFAM domain
Tb927.3.690	-	2-4	Protein kinase
Tb927.5.800	Casein kinase I	2-4	Protein kinase
Tb927.6.1780	-	2-4	Protein kinase
Tb927.6.5100	Polo-like kinase	2-4	Protein kinase
Tb927.9.14430	Casein kinase II	2-4	Protein kinase
Tb927.10.460	-	2-4	Protein kinase
Tb927.10.1940	-	Not determined	Protein kinase
Tb927.6.2880	-	1-2	Kinesin
Tb927.7.3000	-	2-4	Kinesin
Tb927.7.7260	-	2-4	Kinesin
Tb927.11.3280	-	2-4	Kinesin
Tb927.11.5300	-	2-4	Kinesin
Tb927.3.3450	-	1-2	ADP-ribosylation
Tb927.6.3650	-	1-2	ADP-ribosylation
Tb927.8.5060	-	1-2	ADP-ribosylation
Tb927.7.3410	Centrin-4	2-4	EF-hand
Tb927.8.1080	Centrin	2-4	EF-hand
Tb927.9.11230	Calmodulin-like	2-4	EF hand
Tb927.1.2340	Alpha tubulin	2-4	Tubulin
Tb927.3.910	Gamma tubulin	2-4	Tubulin
Tb927.4.3630	-	2-4	Serine/threonine
Tb927.11.8090	-	1-2	Serine/threonine
Tb927.3.2110	-	2-4	NLI interacting factor
Tb927.4.2220	-	2-4	TPR repeat
Tb927.8.1990	-	2-4	AhPC/TSA
Tb927.8.4210	TbBBP46	2-4	WD40
Tb927.10.2290	-	1-2	DNAj
Tb927.10.2860	TbPOC1	2-4	WD domain
Tb927.10.10350	-	2-4	MORN repeat
Tb927.10.12**6**0	-	2-4	tRNA anti-codon
Tb927.9.9020	-	Not determined	MMR HSRI
Tb927.11.3870	-	2-4	ATPase
Tb927.11.9980	-	1-2	2-oxoacid dehydrogenase
Tb927.11.16390	-	2-4	Cyclin dependant kinase
Tb927.11.18680	-	2-4	Dynein light chain
Tb927.1.1380	-	2-4	None

From these proteins, 35/36 contained a known PFAM domain, with only one (Tb927.1.1380) failing to identify any recognised domains. BB proteins containing a known PFAM domain of protein kinase (7/7), kinesin (5/5) ADP-ribosylation (3/3) EF-hand (3/3) and tubulin (2/2) were present in all organisms investigated (table 13).

3.10 RNAi phenotypes in procyclic and bloodstream *Trypanosoma brucei* proteins

Previous experiments on RNAi phenotypes for each of the *T. brucei* proteins are available on the TriTrypDB database (Alsford *et al,* 2011) and were collected for both the procyclic and bloodstream forms of the parasite. RNAi of 85 (31%) of procyclic form and 102 (38%) of bloodstream form BB proteins had no effect on the growth rates of the cells after induction (figure 48 and 49).

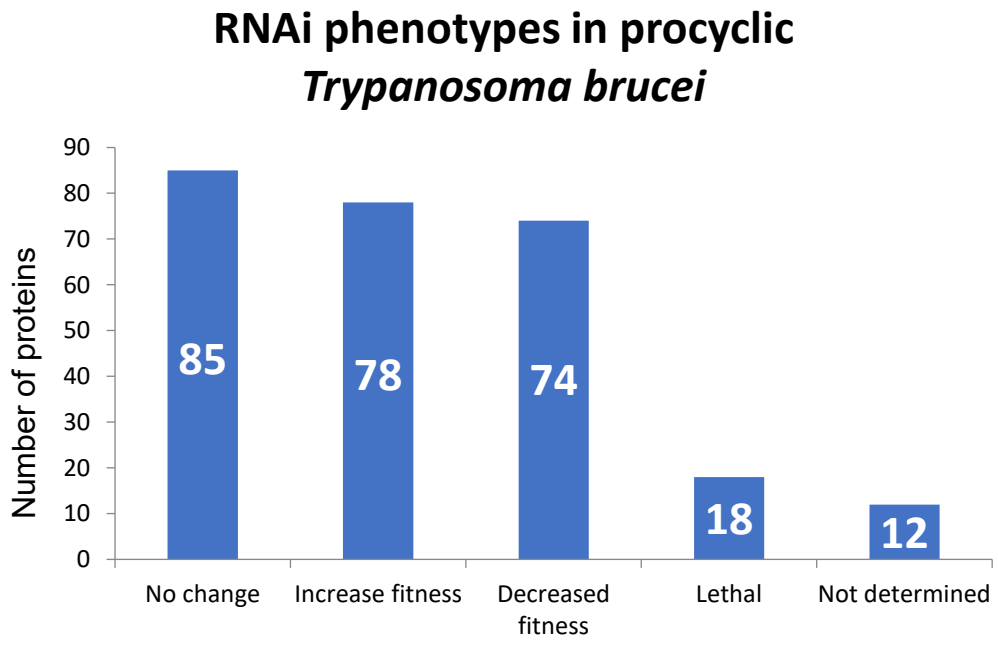


Figure 48: RNAi phenotypes of basal body proteins in procyclic cells

Information collected from TriTrypDB (Aslett *et al,* 2010) on the 267 basal body proteins and the effect that RNAi induction of the proteins had on the fitness of a *T. brucei* procyclic cell.

Absence of other proteins through RNAi induction showed that 78 (29%) and 75 (28%) of procyclic and bloodstream form BB proteins increased the growth rate of cells, in comparison to the uninduced cell line. Whereas 74 (27%) and 66 (24%) of BB proteins presented a decreased growth rate after RNAi induction.

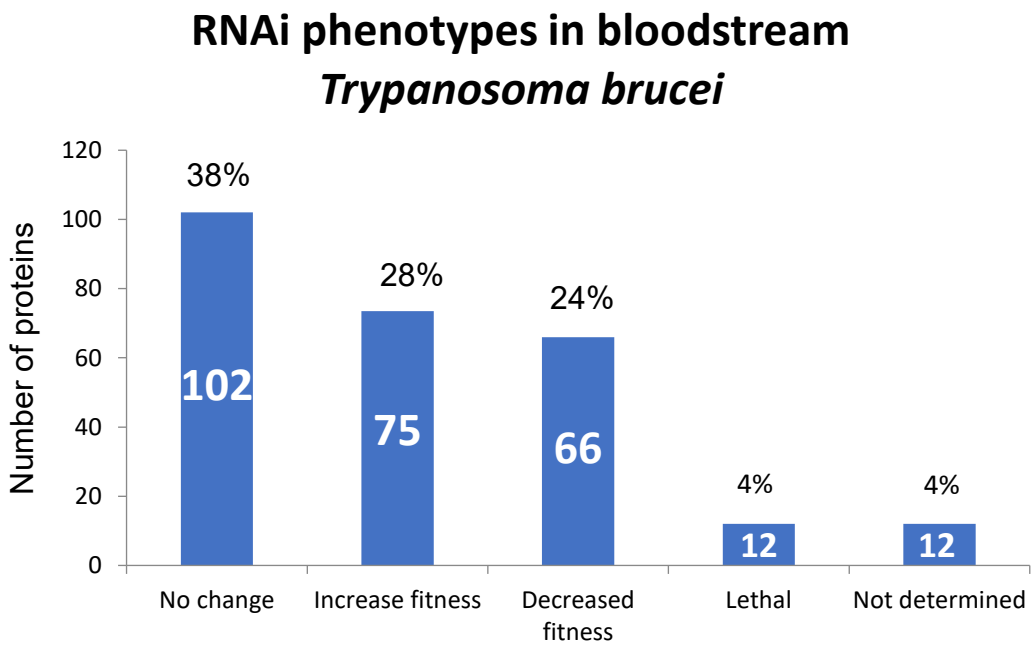


Figure 49: RNAi phenotypes of basal body proteins in bloodstream cells

Information collected from TriTrypDB (Aslett *et al,* 2010) on the 267 basal body proteins and the effect that RNAi induction of the proteins had on the fitness of a *T. brucei* bloodstream cell.

Very few BB proteins of 18 (6%) for procyclic cells and 12 (4%) for bloodstream cells showed a substantial decrease in the fitness of cells and were classified as potentially lethal. 12 (4%) proteins from each life cycle stage was classed as “not determined”, as proteins failed to provide data during the experiment.

3.11 Conservation of the Cep164 proteins in *Trypanosoma brucei* and other investigated organisms

Three proteins from the 267 suggested BB proteins contained a WW domain and had previously been uncovered as Cep164A (Tb927.5.2440), Cep164B (Tb927.11.11650) and Cep164C (Tb927.1.3560). These proteins are orthologues to the human Cep164 protein, with this study showing that Cep164C is the main conserved protein, being present in nine of the organisms investigated (table 14). The least conserved was Cep164B, identified in four organisms. Whereas Cep164A was identified in 7 of the organisms investigated (table 14).

Table 14: Conservation of Cep164 proteins in the basal body database

Cep164A (Tb927.5.2440)	
Conserved in:	*Trypanosoma brucei, Chlamydomonas reinhardtii, Tetrahymena thermophila, Trichomonas vaginalis, Xenopus laevis, Mus musculus, Homo sapiens*
Absent in:	*Danio rerio, Saccharomyces cerevisiae, Arabidopsis thaliana, Caenorhabditis elegans, Plasmodium falciparum, Drosophila melanogaster*
Cep164B (Tb927.11.11650)	
Conserved in:	*Trypanosoma brucei, Danio rerio, Mus musculus, Drosophila melanogaster*
Absent in:	*Chlamydomonas reinhardtii, Tetrahymena thermophila, Trichomonas vaginalis, Xenopus laevis, Saccharomyces cerevisiae, Arabidopsis thaliana, Caenorhabditis elegans, Plasmodium falciparum, Homo sapiens*
Cep164C (Tb927.1.3560)	
Conserved in:	*Trypanosoma brucei, Chlamydomonas reinhardtii, Tetrahymena thermophile, Trichomonas vaginalis, Danio rerio, Xenopus laevis, Mus musculus, Homo sapiens, Drosophila melanogaster*
Absent in:	*Saccharomyces cerevisiae, Arabidopsis thaliana, Caenorhabditis elegans, Plasmodium falciparum*

No orthologues were identified in *Arabidopsis thaliana, Caenorhabditis elegans, Plasmodium falciparum* and *Saccharomyces cerevisiae.* It was also noted that based on the bio-informatics conducted in this study, *Homo sapiens* did not have an orthologue to Cep164B but contained orthologues for Cep164A and Cep164C.

4. Investigation of Cep164C results

In results chapter three, one of the 20 proteins screened (Tb927.1.3560) was identified as an orthologue to the highly conserved basal body/centriole protein Cep164 (Graser *et al,* 2007). Three Cep164 proteins have been previously identified by Hodges *et al* (2010) in *T. brucei* and termed Cep164A, Cep164B and Cep164C. A publication by Dang *et al* (2017) demonstrated that Cep164A and Cep164B localised to the BB area in *T. brucei* but no localisation for Cep164C was observed. As outlined in the previous chapter, this protein had an unusual cell cycle-dependent localisation and was chosen to take forward to discover the role of this orthologue in *T. brucei*.

4.1 Cep164C is a cell cycle-dependent protein

Localisation of Cep164C was analysed in detergent-extracted cells expressing the mNG::Cep164C protein, which was transfected into the background of an mCherry::SAS-6 cell line. The SAS-6 protein is ideal for co-localisation studies to the BBs, as it locates to the very proximal section of the organelles (Hu *et al,* 2015) and labels the mature basal body (figure 50A-F; white arrow head) and the pro-basal body (figure 50A-F; white arrow).

At the G1 cell cycle stage (cells with one BBP, one kinetoplast and one nucleus) cells contain only a single flagellum (figure 50A-B; orange arrow), nucleated from the MBB. In dividing cells, the single flagellum remains assembled, termed the old flagellum and is located towards the anterior section of the cell (figure 50E; blue arrow head). Following BB duplication, a new flagellum is nucleated from a newly matured BB and is located towards the posterior section of the cell (figure 50E; green arrow head). This positioning of the two flagella is well characterised and remains throughout the cell cycle, making it possible to determine the old and new flagellum and BBPs in dividing cells.

Analysis of the mNG::Cep164C cell line showed that this cytoskeletal protein was a cell cycle-dependent protein. It localised distal to the mCherry::SAS-6 protein (figure 50B; red arrow head) and to the MBB of the old flagellum only (figure 50B-E; red arrow head). An important question is when localisation to the MBB of the

old flagellum appears. In some cells with a single flagellum no Cep164C labelling was present (figure 50A; yellow arrow head), however a proportion of these cells were positive for Cep164C labelling (figure 50B; red arrow). The protein then continues to localise to the MBB of the old flagellum throughout the remainder of the cell cycle (figure 50C-E; red arrow), before becoming absent prior to cytokinesis (figure 50F; yellow arrow). No Cep164C localisation was observed on the PBBs at any point in the cell cycle. This protein also localised to the MBB only in whole cells (figure 50G) and showed the same localisation pattern.

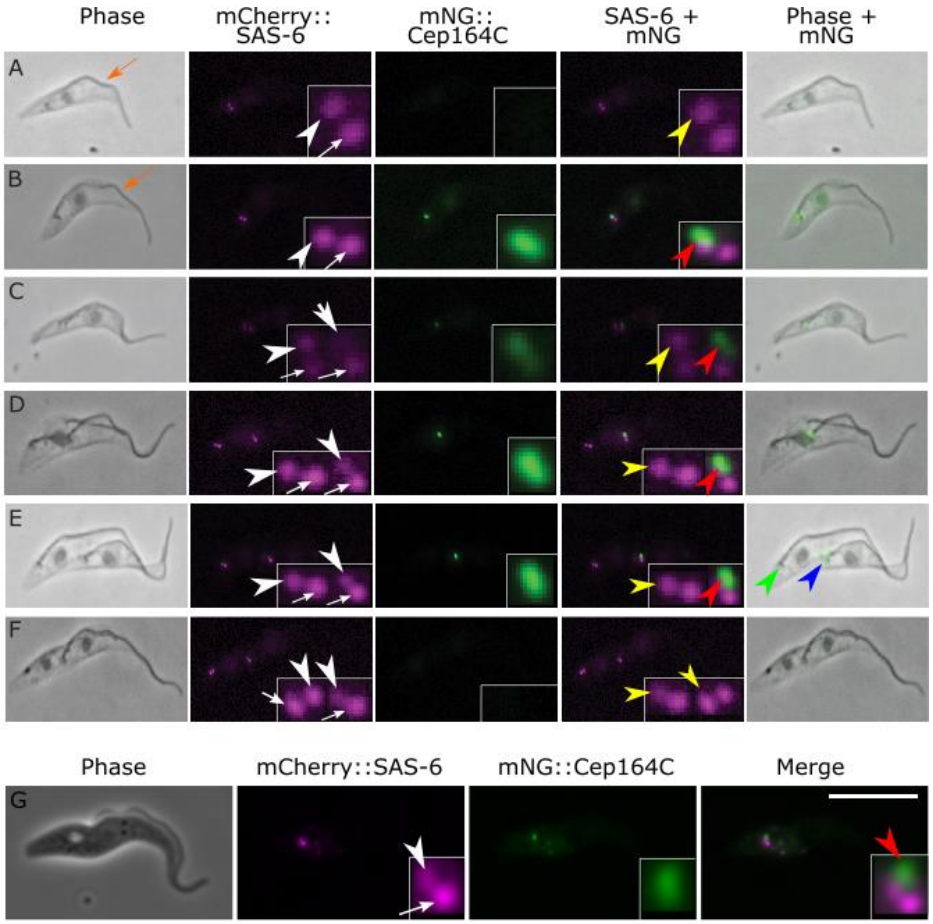


Figure 50: Cep164C localises distal to the SAS-6 protein.

Cells expressing the mNG::Cep164C (green) protein in the background of the basal body marker SAS-6 (mCherry). Cep164C is a cell cycle-dependent protein that recruits to the distal section of the old mature basal body only and is absent prior to basal body duplication (A). The protein is recruited distal to the SAS-6 protein throughout the cell cycle (B-E), becoming absent prior to cytokinesis (F). Localisation in whole cells was also seen at the old mature basal body (G). Scale bar = 10μm.

In summary, Cep164C localises in a cell cycle-dependent manner to the oldest mature BB only in a cell. Since it is not present in all cells with one flagellum (figure 50A), this suggests a recruitment event at the start of the cell cycle. This is an intriguing localisation pattern and essentially represents the 'age' of a basal body (figure 51).

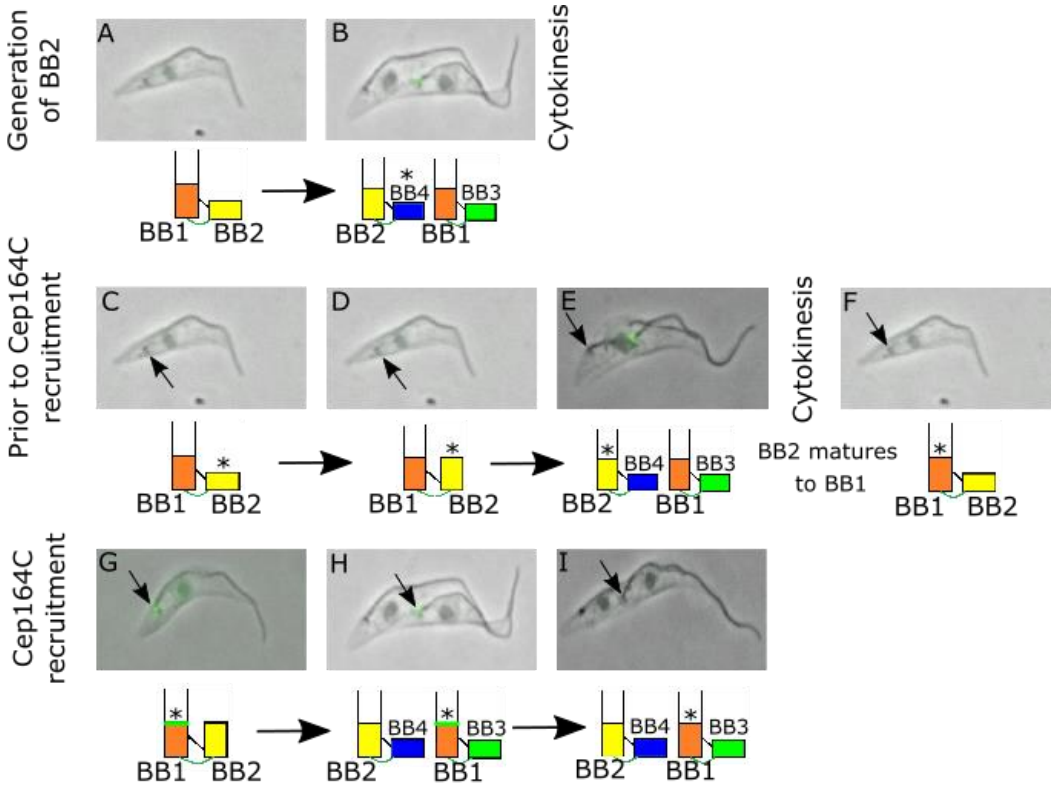


Figure 51: Cep164C localises to the old mature basal body only

Cells expressing the mNG::Cep164C protein (green). The pro-basal body does not contain Cep164C (A-B) but is able to dock to the flagellar pocket and produce a new flagellum (C-E). Prior to basal body duplication, the now matured basal body then recruits Cep164C (G), where it remains throughout the cell cycle (H), before becoming absent prior to cytokinesis (I).

The MBB with the old flagellum (BB1) is connected to its pro-basal body (BB2) (figure 51A-B). During the first cell cycle, BB2 matures and two PBBs known as BB3 and BB4 form next to each MBBs (figure 51B).

After cytokinesis, the daughter cells contain a BPP (figure 51C), where the PBB matures and grows a new flagellum (figure 51D-E) in cell cycle two. Only during the third cell cycle after it was assembled does the BB (not containing the old flagellum) recruit Cep164C (figure 51G-H).

4.2 Development of an enrichment protocol to identify the recruitment pattern of Cep164C

After observation of Cep164Cs recruitment to the MBB of the old flagellum in a cell cycle-dependent manner, the recruitment process was further identified through cell enrichment methodology. The enrichment protocol was carried out in collaboration with Dr Vladimir Varga from Oxford University (V. Varga 2018, personal communication). Dr Varga kindly developed the enrichment protocol used in this study by altering a starvation-synchronization protocol for the *T. brucei* cells originally created by Archer *et al* (2011). Cells in stationary phase are predominantly in the G1 phase, once released into fresh media they begin division and quickly enter the exponential phase showing a rapid increase in cell numbers. Once the level of nutrients in the media begin to decline the population then enters the stationary phase, where a plateau is reached and the number of dying cells equals the number of dividing cells. The starvation methodology was developed after the expectation that starved cells at the stationary phase will be unable to continue through the cell cycle, with most cells stopping at the same stage. Cells which are not released into fresh media at this point will then enter the death phase (Creek *et al,* 2013). The aim was to investigate the recruitment of Cep164C to cells as they start to divide.

The SmOx parental cell line (Poon *et al,* 2012) were used for carrying out the starvation protocol. In this protocol, five different densities were piloted for observation of the maximum cell growth rate per/ml. These densities ranged from 1×10^4 and 1×10^6 and interestingly each cell density after five days grew to a similar maximum density of $4.5\text{-}5 \times 10^7$ (figure 52).

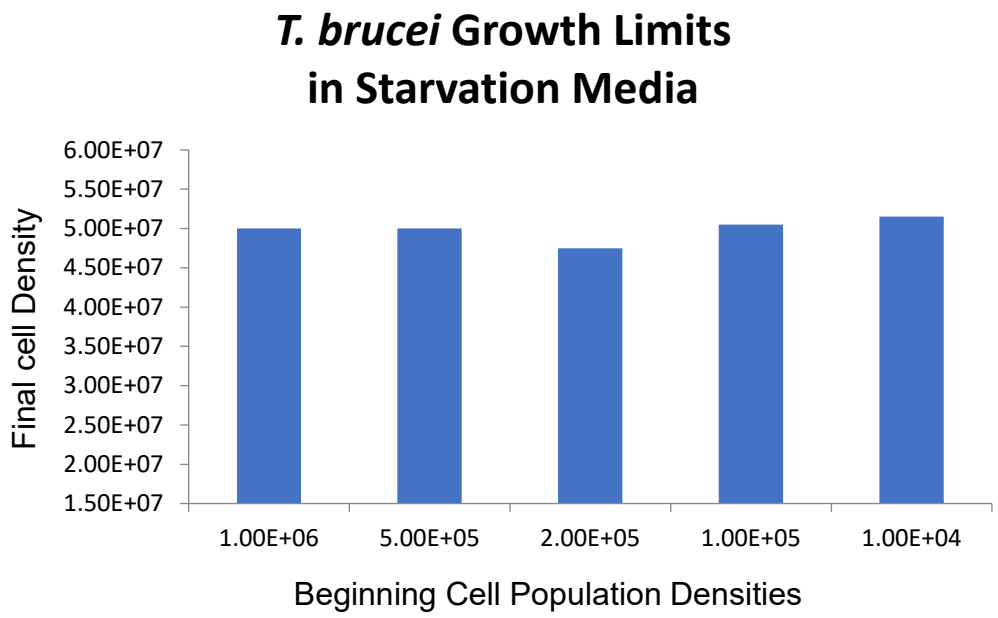


Figure 52: SmOx growth rate in starvation media

SmOx cell density after continuous growth of cells in 6ml of media for five days without additional fresh media for the enrichment protocol. Five starting cell densities were used (ranging from 1 x 10^4 – 1 x 10^6), with each reaching a similar maximum density of 4.5-5x10^7.

A cell density at 5 x 10^4 ml was chosen, cells were released into 6mls of drug-free media for five days without additional culture. This density was chosen due to the similar maximum densities observed between the range of 1 x 10^4 to 1 x 10^7 and that it was shown to produce enriched cells in a previously written protocol from Vladimir Varga from Oxford University (V. Varga 2018, personal communication).

After release, cells had an elongated cell body with a longer posterior, potentially due to media starvation (figure 53). Cells were labelled with the BB marker BBA4 antibody, which is ideal for use in cell cycle studies as it detects a protein located to both the MBB and PBB in *T. brucei* (Woods *et al,* 1989; Woodward *et al,* 1995).

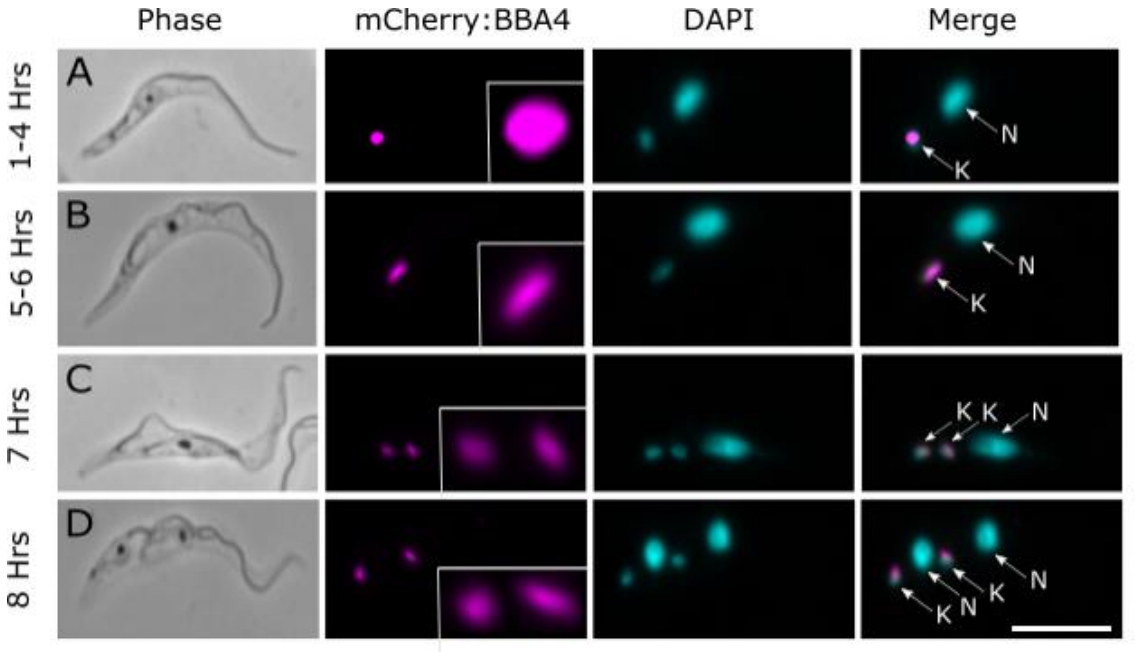


Figure 53: Enrichment halts cell cycle progression prior to basal body duplication

SmOx cells labelled with the basal body marker BBA4 after release from media starvation. The majority of cells between 0-4 hours were in the G1 cell cycle stage, prior to basal body duplication (A). At 5-6 hours the cells had begun duplication, with the kinetoplast beginning to elongate (B). At 7 hours, the kinetoplast has divided and two basal body pairs are present (C). Finally, at 8 hours post-release, most cells are in the early or late stages of cell division (D). K = Kinetoplast, N = Nucleus. Scale bar = 10μm.

From immediate release to 4 hours post-release, most cells had arrested prior to BB duplication, with the BBA4 antibody labelling showing only one BBP (figure 53A). Those cells released between 5-7 hours were beginning duplication and segregation of their BBs (figure 53B), with division of one kinetoplast to two and one BBP to two BBPs (each containing one MBB and one PBB). Cells at 8 hours post-release were in the later stages of the cell cycle and contained fully segregated kinetoplasts, a divided nucleus and two BBPs (figure 53D).

These results suggest that the enrichment protocol was effective at arresting cells prior to BB duplication, and that those released into fresh media were able to recover and re-start cell division. 100 cells were analysed for each hour of release from immediate to 8 hour post release. 90-95% of cells between immediate to two

hours release had either not entered the cell cycle or were prior to BB duplication (figure 54A and B).

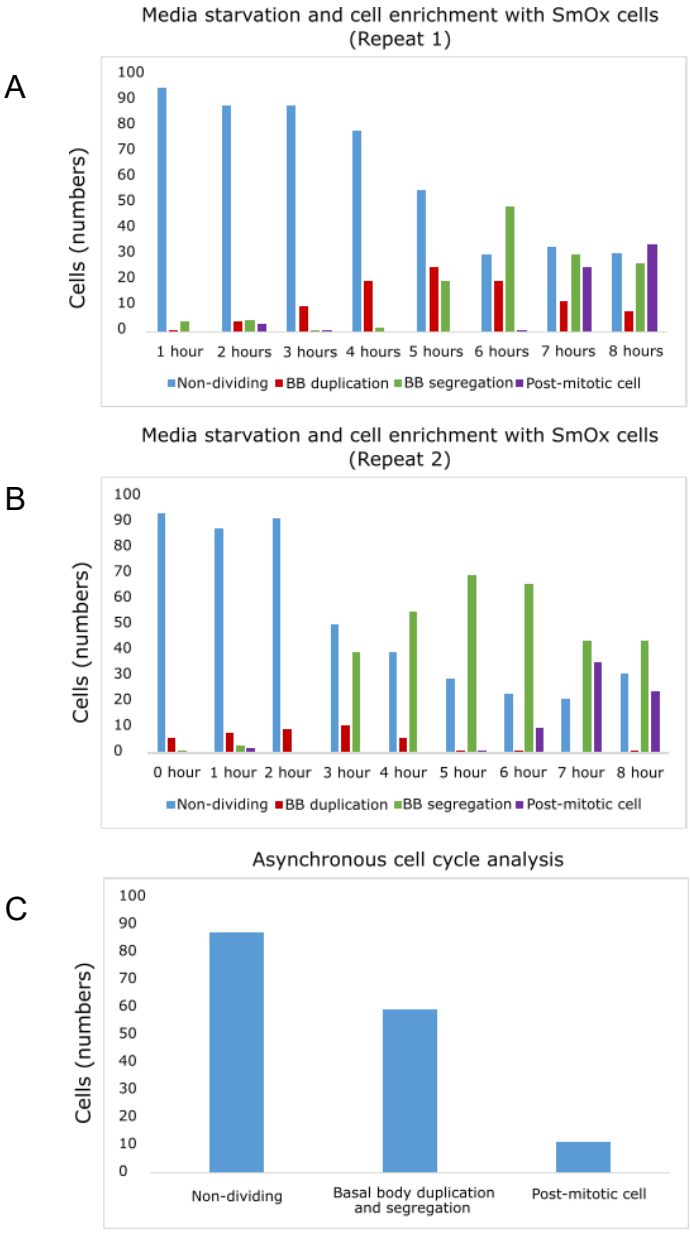


Figure 54: Progression through the cell cycle causes loss of enrichment

A and B) First and second repeat of SmOx p927 *T. brucei* cells undergoing media starvation. The majority of cells after immediate release were prior to basal body duplication. Cells progressed through the cell cycle in a less enriched manner. C) 157 Cep164C cells were analysed to show a typical asynchronous culture. BB = basal body.

By the 5-6th hour of release, cells had begun to duplicate and segregate their basal bodies (BB segregation). However, by this time point cells had lost synchronization but an enrichment of cells in certain stages were still present (figure 54A and B). During the 5-6th hour, a mixture of G1 cells, early dividing and later dividing cells were present. By the 7-8th hour of release, over half the population were progressing through the cell cycle. The cell cycle analysis results for the enrichment protocol showed differences in those cell cycle counts for the asynchronous population.

In order to compare the cell cycle analysis of enriched SmOx cells to an asynchronous population, 157 asynchronous cells of the Cep164C cell line were labelled with DAPI and counted. High numbers of 1K1N cells (55%) and lower numbers of 2K1N (38%) and 2K2N (7%) cells were seen, showing a typical asynchronous cell cycle analysis (Siegel *et al,* 2008). In comparison to the asynchronous cells, the enrichment protocol showed an increase of up to 30-40 cells in the late stage of division (2K2N), and a higher number of early dividing cells between the 5-6th hour of release. These results show that the starvation protocol arrests cells prior to BB duplication and that this methodology can be used to further understand the recruitment process of Cep164C as this protein is recruited prior to BB duplication.

4.3 Recruitment of Cep164C occurs prior to basal body duplication.

To further understand the timing of recruitment of the Cep164C protein to the MBB of the old flagellum, the starvation method was carried out on clonal populations of the mNG::Cep164C cells. Cells expressing the mNG::Cep164C and mCherry::SAS-6 pPOTv6 plasmid were grown in conditioned media, put through the starvation protocol and analysed. Two clonal cell lines were fixed and labelled with BBA4 to indicate whether the BBs had duplicated. In cells where BB duplication had not occurred 40% were positive and 60% were negative for Cep164C recruitment (figure 55A and 55B; prior to BB duplication), confirming the

observations where some cells lacked (figure 50A) or contained (figure 50B) Cep164C.

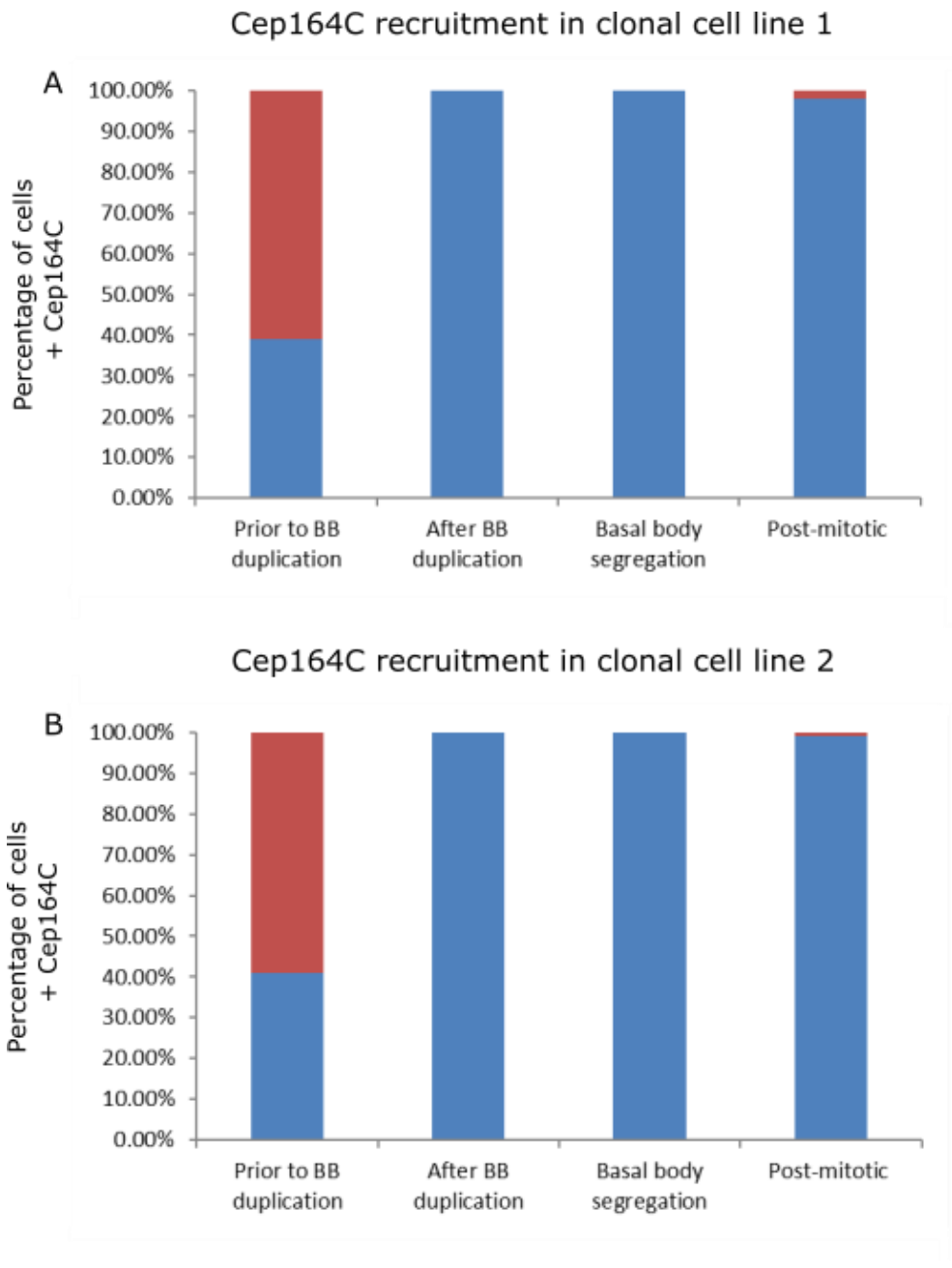


Figure 55: Cep164C is recruited prior to basal body duplication

A and B) 100 cells from two clonal cell lines expressing mNG::Cep164C and mCherry::SAS-6. After carrying out the enrichment studies, Cep164C was seen to recruit to the basal body prior to basal body duplication which remained present until prior to cytokinesis. BB = basal body.

Since duplication of BBs is the first morphological marker of cells entering the cell cycle, these results demonstrate that Cep164C recruitment occurs prior to BB duplication on the MBB. Results show that after recruitment, Cep164C continues to localise throughout the cell cycle, before becoming absent prior to cytokinesis or immediately after abscission (figure 51; I). The protein is then recruited again at the start of the next cell cycle (figure 51; G).

4.4 Cep164C localises immediately distal to the transitional fibre protein TbRP2

In mammalian cells, the Cep164 protein is a component of the distal appendages/transitional fibres in mature centrioles/basal bodies (Graser *et al,* 2007). Therefore, it is possible that the Cep164C protein may also localise to the transitional fibres (TFs) in *T. brucei.* TFs are situated at the distal section of the BBs and are important for the docking process in many eukaryotic cells (Wei *et al,* 2015; Vaughan and Gull, 2016; Stephan *et al,* 2007; Trépout *et al,* 2018).

To investigate further the timing of recruitment and localisation, a co-localisation study using the established transitional fibre antibody, YL1/2 (Andre *et al,* 2014) was carried out. The antibody generated by Kilmartin *et al* (1982) labels the carboxyl-tyrosinated alpha tubulin, composing the new microtubules building the sub-pellicular corset. Recently it was discovered that the YL1/2 antibody cross reacts with the retinitis pigmentosa 2 (RP2) protein localising to the TFs of the MBB throughout the cell cycle (Stephan *et al,* 2007; Andre *et al,* 2014).

Analysis of the Cep164C cell line labelled with the YL1/2 antibody showed that in accordance with the cell cycle-dependency of this protein, Cep164C was not present in the G1 cells or early dividing cells as expected (figure 56A). However, Cep164C was shown to localise immediately distal to the TbRP2 protein in cells which had entered the cell cycle and were just prior to duplicating their BBs (figure 56C; white arrow). The localisation was also observed for those cells in the later cell cycle stages (56D-E; white arrow).

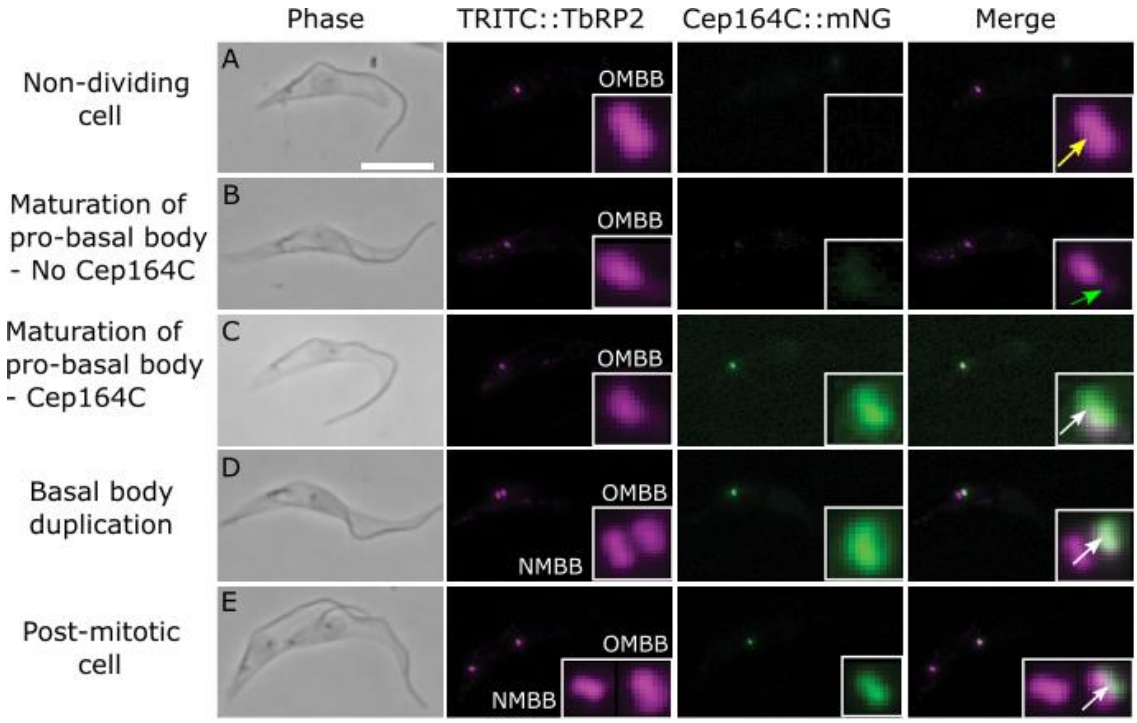


Figure 56: Cep164C localises distal to the TbRP2 protein

Cells expressing mNG::Cep164C, labelled with the YL1/2 antibody shows Cep164C to localise immediately distal to the TbRP2 protein at the transitional fibres through the cell cycle (C-E). As expected Cep164C is not recruited until prior to basal body duplication (A-B). OMBB – Old mature basal body, NMBB – New mature basal body. Scale bar = 10μm

TFs are only present when MBBs dock to the flagellar pocket, with PBBs failing to show labelling with the YL1/2 antibody (Andre *et al,* 2014). Significantly, although both MBBs are docked to their respective flagellar pockets, only the old MBB of the old flagellum is positive for Cep164C. Therefore, TFs are assembled on the new MBB of the new flagellum in the absence of Cep164C.

In summary, Cep164C localised to the TF area of the old MBB of the old flagellum, appearing more distal with no full co-localisation to the TbRP2 protein. This raises important questions such as whether the Cep164C protein localises to the TFs or to other surrounding structures of the basal body.

4.5 Cell line construction for analysis of Cep164 proteins in *Trypanosoma brucei*

In *T. brucei* two other Cep164 proteins (A and B) are present and are also shown to localise to the BB area (Dang *et al,* 2017). To understand how these proteins co-localise and function, a number of cell lines were generated in collaboration with Dr Jiri Tyc (Vaughan Lab), outlined in table 15 (A-M). Dr Jiri Tyc and myself collaborated to generate these cell lines, with some of the cell lines being generated by Dr Jiri Tyc and some by myself.

Table 15: Cell lines generated in this study

Cell lines were generated with either the mNG or mCherry fluorophore and either lacking or containing the RNAi pQuadra plasmid. Double tagged and double RNAi constructs were also generated.

Cell line	Cep164A (fluorophore)	Cep164B (fluorophore)	Cep164C (fluorophore)	RNAi construct
(A) Parental SmOx p927	-	-	-	No
(B) Cep164C cell line + Cep164C RNAi construct	-	-	mNG	Yes
(C) Cep164A+C double tagged cell line	mCherry	-	mNG	No
(D) Cep164B+C double tagged cell line	-	mNG	mCherry	No
(E) Cep164A+B double tagged cell line	mCherry	mNG	-	No
(F) Cep164A+C double tagged + Cep164A RNAi construct	mCherry	-	mNG	Yes
(G) Cep164A+C double tagged + Cep164C RNAi construct	mCherry	-	mNG	Yes
(H) Cep164A+C double tagged + Cep164A+C RNAi construct	mCherry	-	mNG	Yes

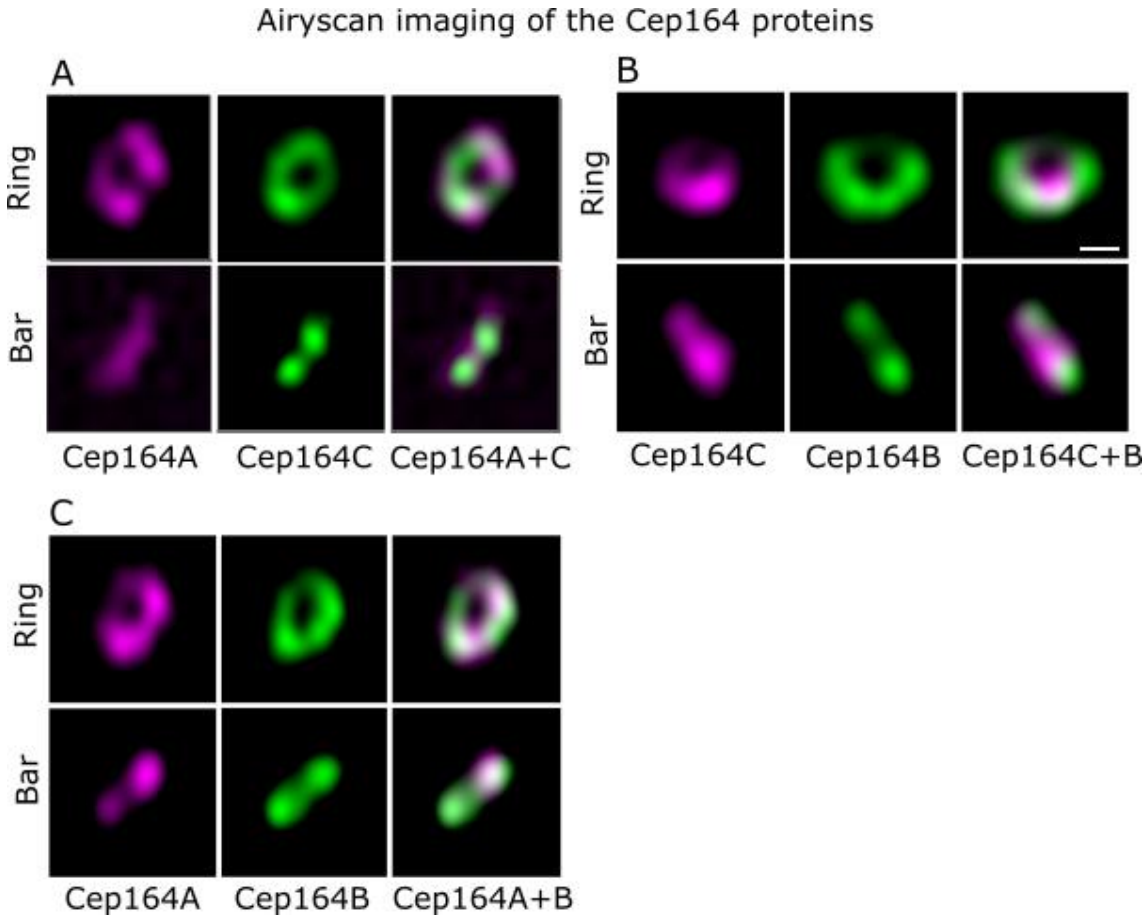
Table 15: Cell lines generated in this study

(I) Cep164B+C double tagged + Cep164B RNAi construct	**-**	**mNG**	**mCherry**	**Yes**
(J) Cep164B+C double tagged + Cep164C RNAi construct	**-**	**mNG**	**mCherry**	**Yes**
(K) Cep164B+C double tagged + Cep164B+C RNAi construct	**-**	**mNG**	**mCherry**	**Yes**
(L) Cep164A+B double tagged + Cep164A RNAi construct	**mCherry**	**mNG**	**-**	**Yes**
(M) Cep164A+B double tagged + Cep164B RNAi construct	**mCherry**	**mNG**	**-**	**Yes**

Cell lines were generated using the mNG or mCherry fluorophore, with some cell lines having two Cep164 proteins tagged e.g. Cep164A + B, Cep164A + C. For some of these double tagged cell lines there was the addition of an inducible RNAi plasmid.

4.6 Cep164C co-localises with Cep164A and Cep164B

Confocal imaging of double tagged cell lines of Cep164A + C (table 15C), Cep164B + C (table 15D) and Cep164A + B (table 15E) revealed a doughnut shaped localisation pattern at the BB area (figure 57A, B and C). Both the side view (observed as a bar structure) and the above view (observed as a circular doughnut structure) were collected. In each case, each Cep164 protein co-localised with the other well (figure 57A, B and C), suggesting these proteins are recruited to the same location on the TFs. The full analysis of Cep164A and B will be shown in chapter 5.



Double tagged cell lines expressing two Cep164 proteins in either the mNG or mCherry fluorophore in whole cells, analysed by Airyscan technology. mNG::Cep164C and mCherry::Cep164A (A), mNG::Cep164B and mCherry::Cep164C (B) and mNG::Cep164B and mCherry::Cep164A (C) co-localised to each other either from the side (bar pattern) or from above (doughnut pattern). Scale bar = 100nm

Figure 57: All Cep164 proteins co-localise to each other

4.7 The Cep164C uninduced double tagged and RNAi cell lines show minimal growth defects

To investigate the function of Cep164C and whether it is required for localisation of both Cep164A and Cep164B, double tagged cell lines with inducible RNAi plasmids were analysed. Growth curve controls were conducted to assess the growth rate of two double tagged cell lines; mCherry::Cep164A + mNG::Cep164C (table 15G) and mNG::Cep164B + mCherry::Cep164C (table 15J) with inducible RNAi to Cep164C.

To demonstrate any fitness cost of the constructs, the growth of the parental SmOx p927 without the pPOTv6 tagging plasmid or RNAi construct were grown for 4 days in the same conditions, alongside the double tagged cell line and cell lines containing the RNAi construct to the Cep164C protein. Growth rates of the RNAi cell lines were compared against the SmOx p927 cell line to determine whether there was a growth phenotype. Cell lysates were collected to identify the amount of Cep164C protein that was present.

In growth rate experiments, cells were sub-cultured to 1 x 10^6 each day and over 4 days an accumulated growth curve was generated. In both experiments the parental SmOx p927 cell line grew to an accumulated growth rate of between 1 x 10^9 – 1 x 10^{10} (figure 58A and 58C; dark green line). The double tagged cell lines lacking the RNAi construct showed a slightly lower accumulative growth rate, but still between 1 x 10^9 – 1 x 10^{10} (figure 58A and 58C; red line). Cell lines containing the RNAi construct against Cep164C generated an accumulative growth rate between 1 x 10^8 – 1 x 10^9, showing a slightly lower growth rate than the parental or double tagged cell lines (figure 58A; light green line; and 58C; purple line). These results suggest that the addition of the RNAi construct only caused a minimum effect on the growth rate of the cells.

Cell lines were generated with a plasmid containing the TY1 epitope tag, recognised by the BB2 antibody (Bastin *et al,* 1996), for use with western blots. No BB2 antibody was detected in the parental SmOx p927 cell line as expected (figure 58B; lane 1 and 58D; lane 1), while the double tagged cell line (figure 58B; lane 2 and 58D; lane 2) and the Cep164C inducible RNAi cell lines (figure 58B; lane 3 and 58D; lane 3) generated positive detection of the BB2 antibody.

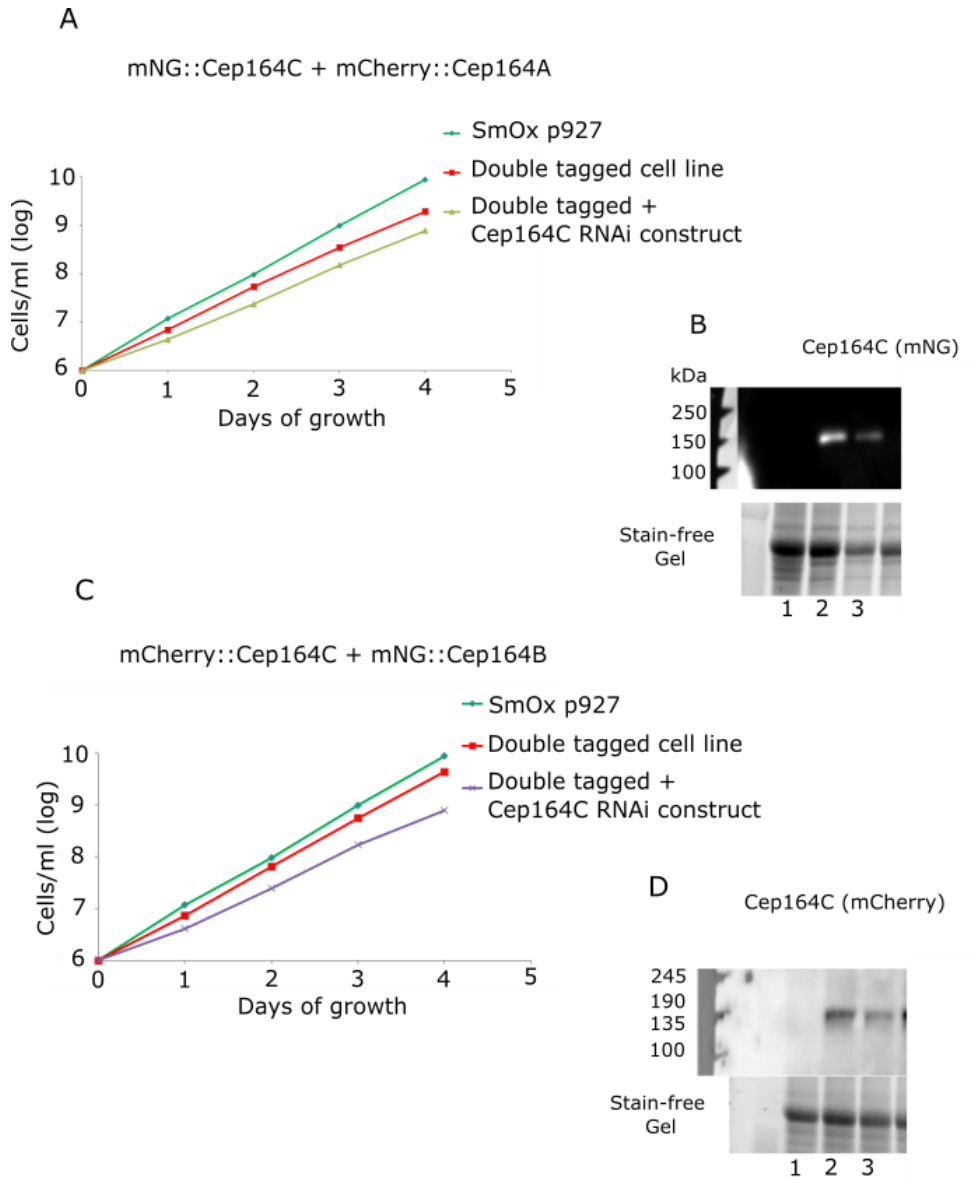


Figure 58: Uninduced Cep164C RNAi cell line shows minimal growth defects

A+C) Growth curves for SmOx cells, Cep164 double tagged cell lines (Cep164C/Cep164A and Cep164C/Cep164B) and uninduced double tagged cell lines with the additional Cep164C RNAi construct were carried out. Those containing the RNAi construct show a slower growth rate. B+D) Western blots were run for all five cell lines. No TY1 tag was detected by the BB2 antibody for the SmOx cells as expected (B; lane 1 and D; lane 1). TY1 tags were detected by western blot for both the double tagged cell line without RNAi constructs: Cep164C/A (B; lane 2) and Cep164C/B (D; lane 2), and those with RNAi constructs: Cep164C/A + Cep164C RNAi construct (B; lane 3) and Cep164C/B + Cep164C RNAi construct (D; lane 3).

The molecular weight of Cep164C without additional plasmids is approximately 98kDa, and approximately 128kDa containing the fluorescent tag. The double tagged cell line however was detected at approximately 160/170 kDa (figure 58B and 58D). These experiments were conducted in collaboration with Dr Jiri Tyc. Both myself and Dr Jiri Tyc worked together in all aspects to conduct the experiment.

4.8 Silencing of the Cep164C protein does not affect growth of the population

Once it was determined that the introduced constructs had little effect on the growth rate, cells were induced using doxycycline for a period of 4 days. Growth rates of induced and uninduced cultures were measured daily on a Beckman Coulter counter and plotted. Double tagged cell lines containing the RNAi construct targeting the Cep164C protein were analysed in the background of either Cep164A or Cep164B. Growth rates of the induced cell line grew to similar densities to the uninduced population (figure 59A, D).

In order to assess knockdown of Cep164C in these double tagged lines western blotting was carried out. The cell lysates for the uninduced cell line produced a visible band as expected (figure 59C; lane 1 and 59F; lane 1), however the induced cell lines failed to produce visible bands, suggesting a successful knockdown of the Cep164C protein (figure 59C; lane 2-3 and 59F; lane 2-3). Observation of these cell lines showed the lack of detectable Cep164C (mNG or mCherry) protein after induction (figure 59B, E; arrow).

Overall, reducing the expression of Cep164C only caused minimal fitness cost to cell growth, with cells growing only slightly slower than the uninduced cell line (figure 59A and 59D).

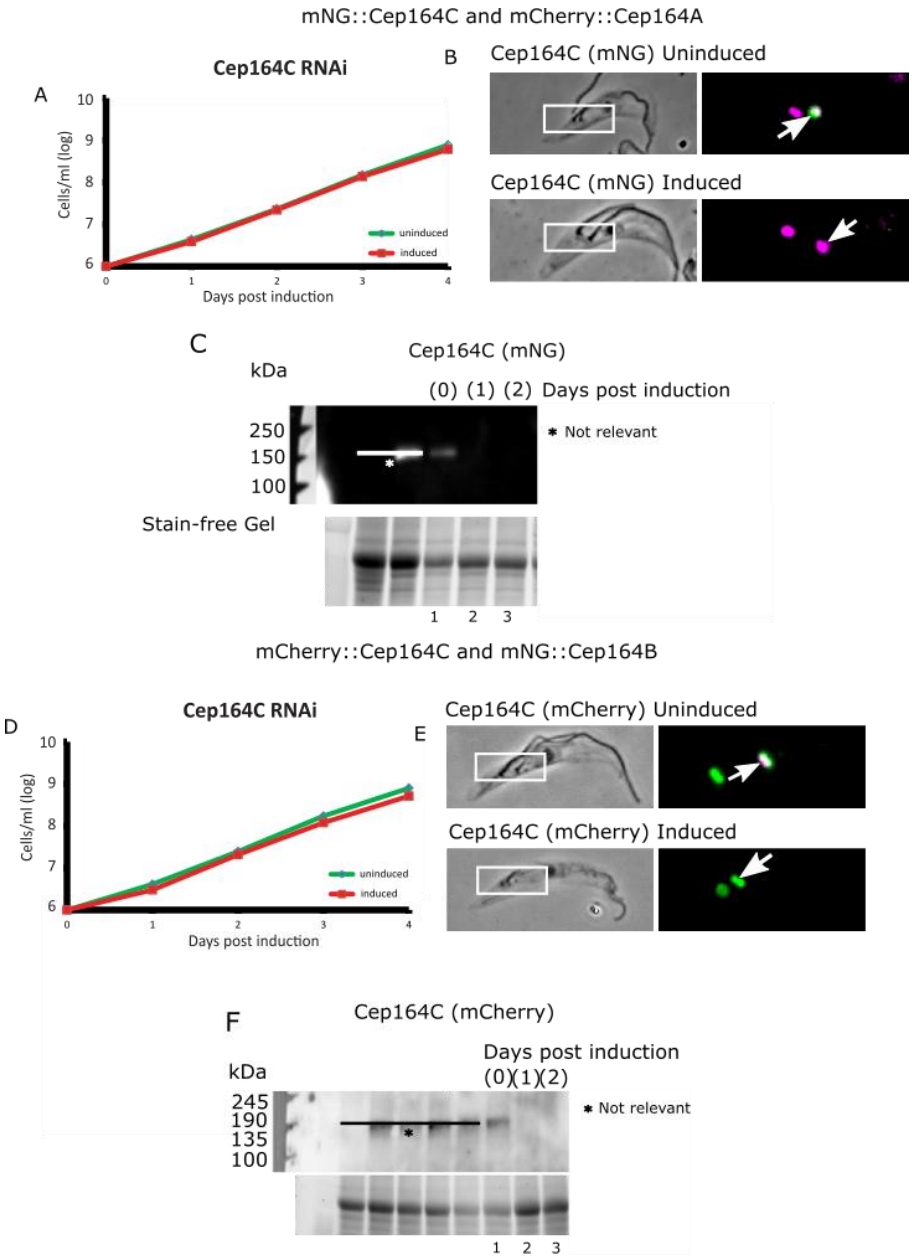


Figure 59: Induced Cep164C RNAi cell line shows minimal growth defects.

A+D) Growth curves for the uninduced and inducible cell lines containing the Cep164C RNAi construct, showing only a minimal fitness cost to cell growth to the inducible cell lines (A + D; red line). Western blots were carried out to determine whether the TY1 tag could be detected after RNAi induction. In uninduced cell lines (C; lane 1 and F; lane 1) TY1 tags were detected as expected. After RNAi induction no TY1 tags were detected at day 1 or day 2 of induction (C; lane 2-3 and F; lane 2-3), suggesting a successful knockdown. Fluorescent imaging of the Cep164C/A or Cep164C/B cell lines showed a lack of detectable mNG::Cep164C (B) or mCherry::Cep164C (E) after induction.

4.9 Induction of Cep164C cell line generates normal counts for the cell cycle analysis

By labelling cells with the DNA marker DAPI, the number of cells in a specific stage of the cell cycle can be counted based on the number of kinetoplasts and nuclei present in each cell. As growth rates for the induced Cep164C population did not differ from the uninduced cell line, the numbers of dividing and G1 cells are likely to be normal.

A kinetoplast and nuclei count in 500 cells for the uninduced, 24 hour induced or 96 hour induced populations was conducted. 75% of the induced and uninduced cells were shown to be in the G1 cell cycle stage (figure 60).

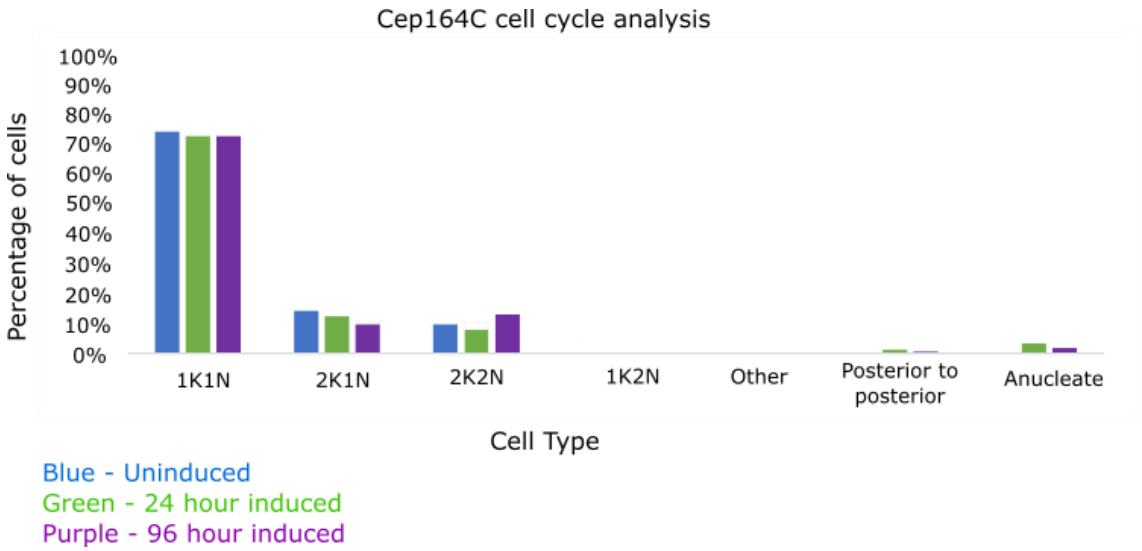


Figure 60: Cep164C cell line produces a normal cell cycle analysis

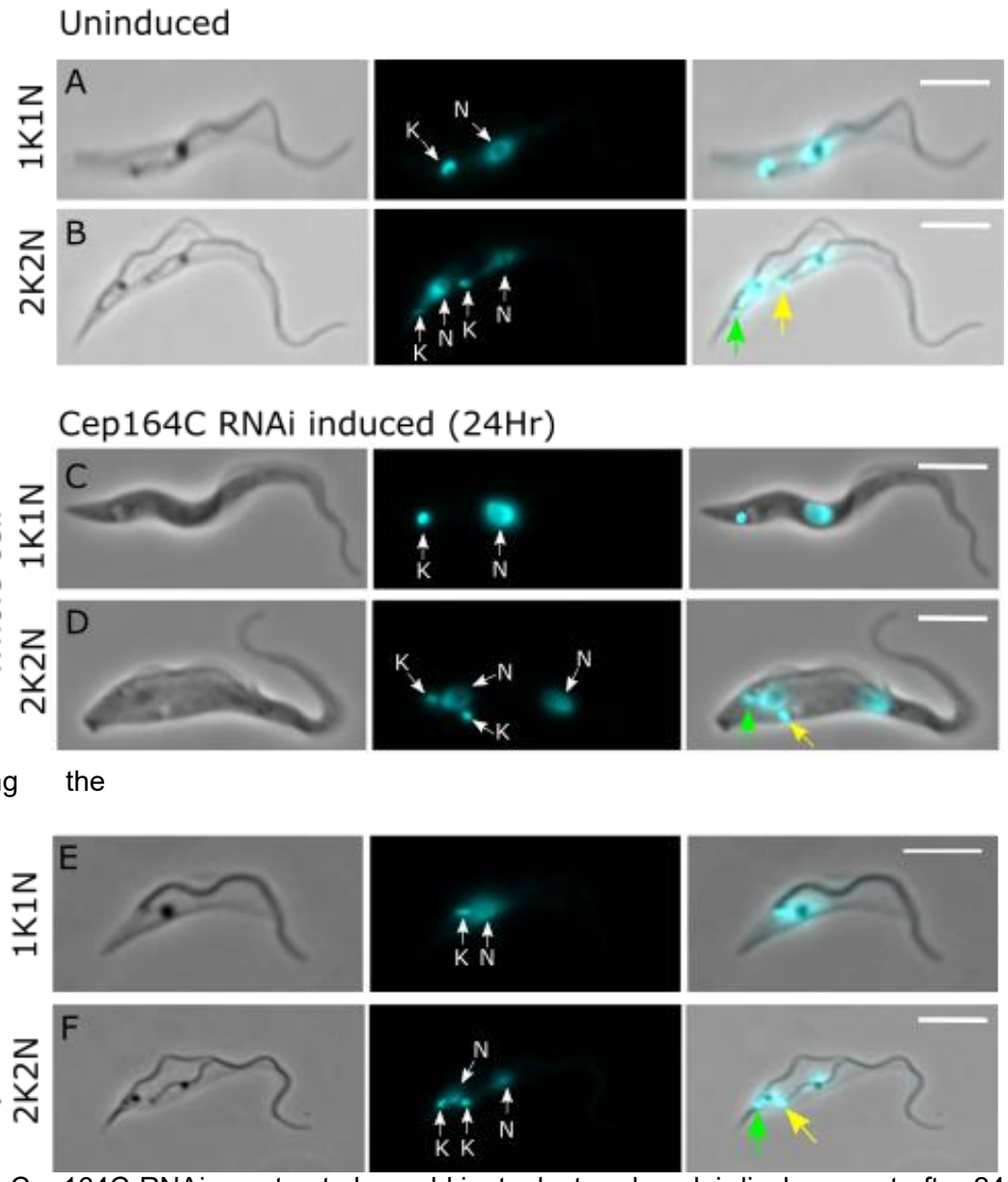
500 uninduced and induced cells expressing the mNG::Cep164C construct were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope to identify the cell cycle stage of the cell. Cells induced for 24 hours and 96 hours were imaged. Induction of the Cep164C RNAi construct showed a normal cell cycle analysis in comparison to the uninduced cell lines, with only a small amount of cells lacking a nucleus (anucleate cells). 1K1N – G1 cells. 2K1N – Early dividing cells. 2K2N – Late dividing cells.

Between 10-15% of induced and uninduced cells were in the early stages of cell division (2K1N), whereas between 8-13% were in the later stages (2K2N). A small number of cells were aberrant, containing anucleate cells (2-3%) or cells with more nuclei/kinetoplasts than expected (less than 1%). Overall, the cell cycle analysis results from the induced and uninduced cell lines were shown to be similar (figure 60).

4.10 RNAi induction of the Cep164C cell line generates asymmetrical daughter cells and morphological phenotypes

After the growth rates and cell cycle counts showed the induced Cep164C cell line generates similar results to the uninduced population, next was to analyse the morphology of the induced cells. Careful morphological analysis of the induced cells revealed a significant number of abnormalities. Initial observations of 1K1N cells showed a mixed population of larger (figure 61C) and smaller (figure 61E) cells in comparison to the uninduced cell line (figure 61A), suggesting an asymmetrical cell division producing daughter cells of different sizes. The two daughters in induced dividing cells also appeared different sizes (figure 61F) compared to the uninduced (figure 61B), initiating further studies.

Analysis also revealed evidence of the kinetoplasts and nuclei being in an incorrect position in the cell. Kinetoplasts of some induced cells appeared to be further posterior (compare figure 61C to 61A) or anterior (figure 61E to 61A) when compared to the uninduced cell line. This morphological abnormality was also observed in dividing cells, with kinetoplasts closer together post-mitosis (figure 61D and 61F; yellow arrows) in comparison to the uninduced cells (figure 61B; yellow arrow). New daughter cells contained anterior kinetoplasts closer to the nucleus in induced cells (figure 61D and 61F; green arrow) in comparison to the uninduced cells (figure 61A; green arrow).



Cells containing the inducible Cep164C RNAi construct showed kinetoplast and nuclei displacement after 24 hours induced, in comparison to the uninduced cell line. Cells in the G1 stage demonstrated having either a far posterior (C) or anterior (E) kinetoplast in comparison to the uninduced (A). The kinetoplasts of the new and old daughter cell was closer to each other (D and F; green and yellow arrow), in comparison to the uninduced cell (B; green and yellow arrow). K- Kinetoplast, N- Nuclei. Scale bars = 5μm.

Figure 61: Inducible Cep164C cell line shows organelle displacement

Morphometric analysis was carried out in order to quantify these positioning abnormalities. Measurements in (N= 200) 1K1N induced and uninduced cells was carried out. After 24 hours and 96 hours of Cep164C RNAi induction, the population contained cells with kinetoplasts closer to the nucleus in comparison to the uninduced cells (figure 62A). An independent t-test showed there was a statistically significant decrease in distance ($p<0.0001$) between the kinetoplast and nucleus in 1K1N uninduced and 24 hour induced cells (figure 62A), and no statistically significant difference ($p<0.130$) between uninduced and 96 hour induced 1K1N cells. However, the wide variation in error bars indicate there were cells with kinetoplasts further away as well as closer together.

Measurements of the late dividing cells containing 2 flagella showed a difference in organelle displacement, between the new daughter cell and the old daughter cell. The kinetoplast of the new flagellum daughter cell in 24 hour and 96 hour were closer to the nucleus than the uninduced (figure 62B; green arrow), while the kinetoplast of the old flagellum daughter cell is further away from the nucleus (figure 62C; yellow arrow). Both measurements were statistically significant at p-value <0.0001.

In dividing cells, measurements between the two kinetoplasts showed the distances were closer together for the 24 hour and 96 hour induced cells, in comparison to the uninduced kinetoplasts. With most distances ranging from 0.5µm-3µm, whereas the uninduced ranged from 3µm-6µm. The data shows there is a statistically significant decrease in distance ($p<0.0001$) between the new (figure 62D; green arrow) and old (figure 62D; yellow arrow) daughter kinetoplasts of dividing cells for the uninduced, 24 hour and 96 hour cells.

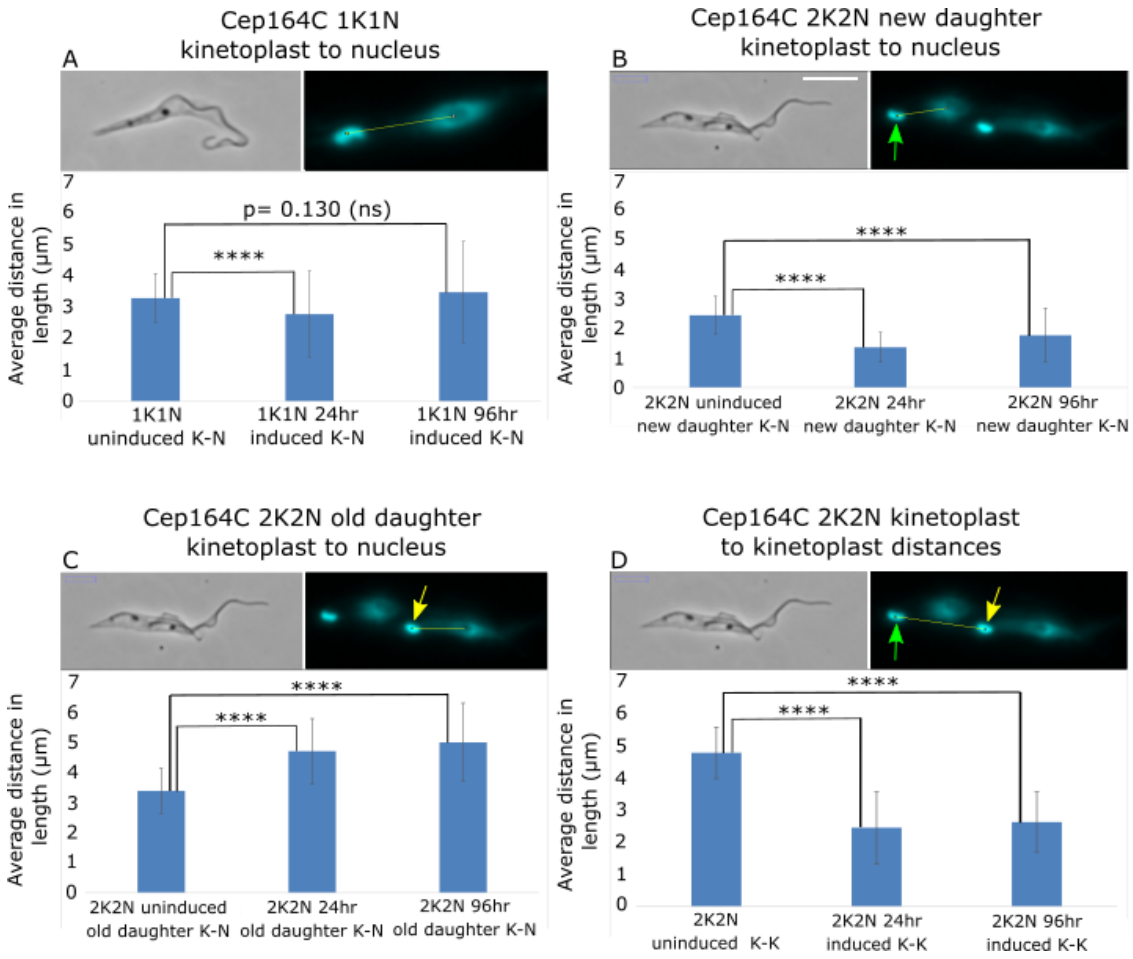


Figure 62: Inducible Cep164C shows kinetoplast and nucleus phenotypes

Cells containing the Cep164C RNAi construct were collected for either the uninduced or induced cell lines. 200 cells were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope. Induced cells were induced for either 24 hours or 96 hours. Statistically significant differences between the distances of the kinetoplast and nuclei were observed for the G1 cells (A) and dividing cells. B) shows the K-N distance for the new daughter cell, C) shows the K-N distance for the old daughter cell and D) shows the K-K distance between the old and new kinetoplast. K = Kinetoplast, N = Nucleus. Independent t-tests were used, **** = $p \leq 0.0001$. Scale bar = 10µm.

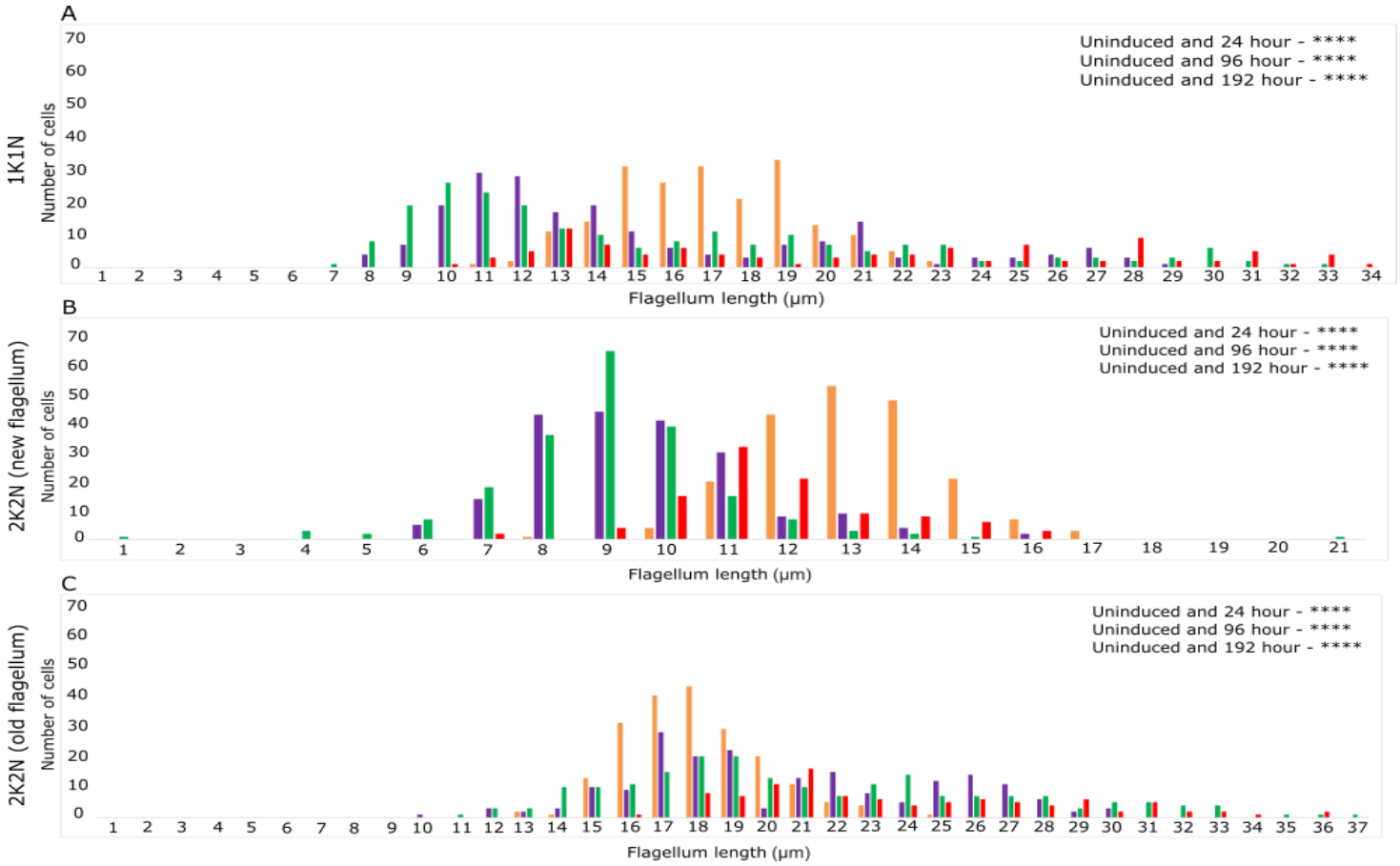
In summary, these results suggest defects in the correct placement of kinetoplasts and nuclei as well a variation in cell sizes resulting from an asymmetrical division of daughter cells.

4.11 Knockdown of Cep164C generates shorter and longer flagella in G1 cells

As RNAi induction of Cep164C generated cells of different sizes it may indicate that the length of the flagellum is also altered, so this was investigated through measurements of the flagellum. Cells were labelled with the DNA marker DAPI and the L8C4 antibody (labelling the PFR structure in the flagellum (Kohl *et al,* 1999)). 200 cells were imaged at uninduced, 24 hour and 96 hour induction time points, with 100 cells being collected at the 192 hour induced time point. Flagellum lengths were measured using the PFR labelling.

In the uninduced population of cells, the 1K1N flagellum lengths ranged from 11μm-23μm (figure 63A; orange). Whereas 24 hour post-induction flagellum lengths ranged from 8μm-29μm (figure 64A and 64B), the 96 hour flagellum measurements ranged from 7μm-33μm (figure 64E and 64F) and the 192 hour flagellum lengths ranged from 10μm-34μm (figure 63A; red). Whilst RNAi studies of flagellum components have yielded cells with shorter flagella (Kohl *et al,* 2003; Absalon *et al,* 2008), this is the first time to our knowledge cells with longer flagella as well as shorter flagella have been observed in *T. brucei.*

The average between uninduced and 24 hour 1K1N cells was 17.6μm and 15.6μm respectively, with results showing a statistically significant difference at $p<0.0001$ (figure 63A). The average for the 96 hour and 192 hour induced cells was 21.9μm and 21.3μm respectively, both showing statistically significant differences in flagellum lengths at $p<0.0001$ (figure 63A). These results show that there is a wide range in flagellum lengths in the induced cells versus the uninduced cells.



Orange: Un-induced, Purple: 24 hour induced, Green: 96 hour induced, Red: 192 hour induced

Figure 63: Inducible Cep164C shows a range in flagellum lengths

200 induced and uninduced Cep164C cells were labelled with the DAPI and L8C4 antibody and flagellum lengths were measured. Measurements showed a wide range of flagellum lengths in G1 cells (A), with a shorter new flagellum (B) and a longer old flagellum (C) in dividing cells. Independent t-test was used. **** = $P \leq 0.0001$

4.12 Longer flagella are only observed in the old flagellum of dividing cells

Post-mitotic cells were used to compare the new and old flagellum lengths, as the new flagellum at this stage has almost completed its growth (Farr *et al,* 2009). Measurements of the old flagellum showed an interesting difference in comparison to the uninduced cells. Flagellum lengths for the uninduced cell line ranged between 13μm-25μm (figure 63C; orange) with an average of 18.5μm. 24 hour induced cells ranged from 10μm-30μm for their flagellum lengths (figure 64C and 64D) with an average of 21.1μm, results showed there was a statistically significant difference $p<0.0001$ (figure 63C). The flagellum lengths for 96 hour and 192 hour induced cells ranged from 11μm-37μm (figure 64G and 64H) and from 16μm-36μm respectively. The averages were 21.9μm and 24.4μm, with both being longer than the uninduced flagellum length average. There was a statistically significant difference at $p<0.0001$ (figure 63C). These results compare with those taken for 1K1N cells and shows that those with a very long flagellum are the old flagellum.

Measurements of the new flagella in post-mitotic dividing cells also showed an interesting difference in flagellum lengths. The uninduced new flagellum ranged from 10μm-17μm (figure 63B), with a mean of 13.6μm. Measurements of the 24 hour induced new flagellum ranged from 6μm-15μm and had an average of 10.0μm (figure 63B). The range of new flagellum lengths are smaller than the uninduced, with the data showing statistically significant differences between the uninduced and 24 hour induced cell lines at $p<0.0001$ (figure 63B). 96 hour and 192 hour cells ranged from 4μm-17μm and 7μm-16μm respectively (shorter than the uninduced population), with an average of 9.6μm and 12.2μm, with both showing a statistically significant difference at $p<0.0001$ (figure 63B). These results show a wide range in flagellum lengths in the induced cells, with a shorter new flagellum to the uninduced population. This suggests an issue in the growth of the new flagellum.

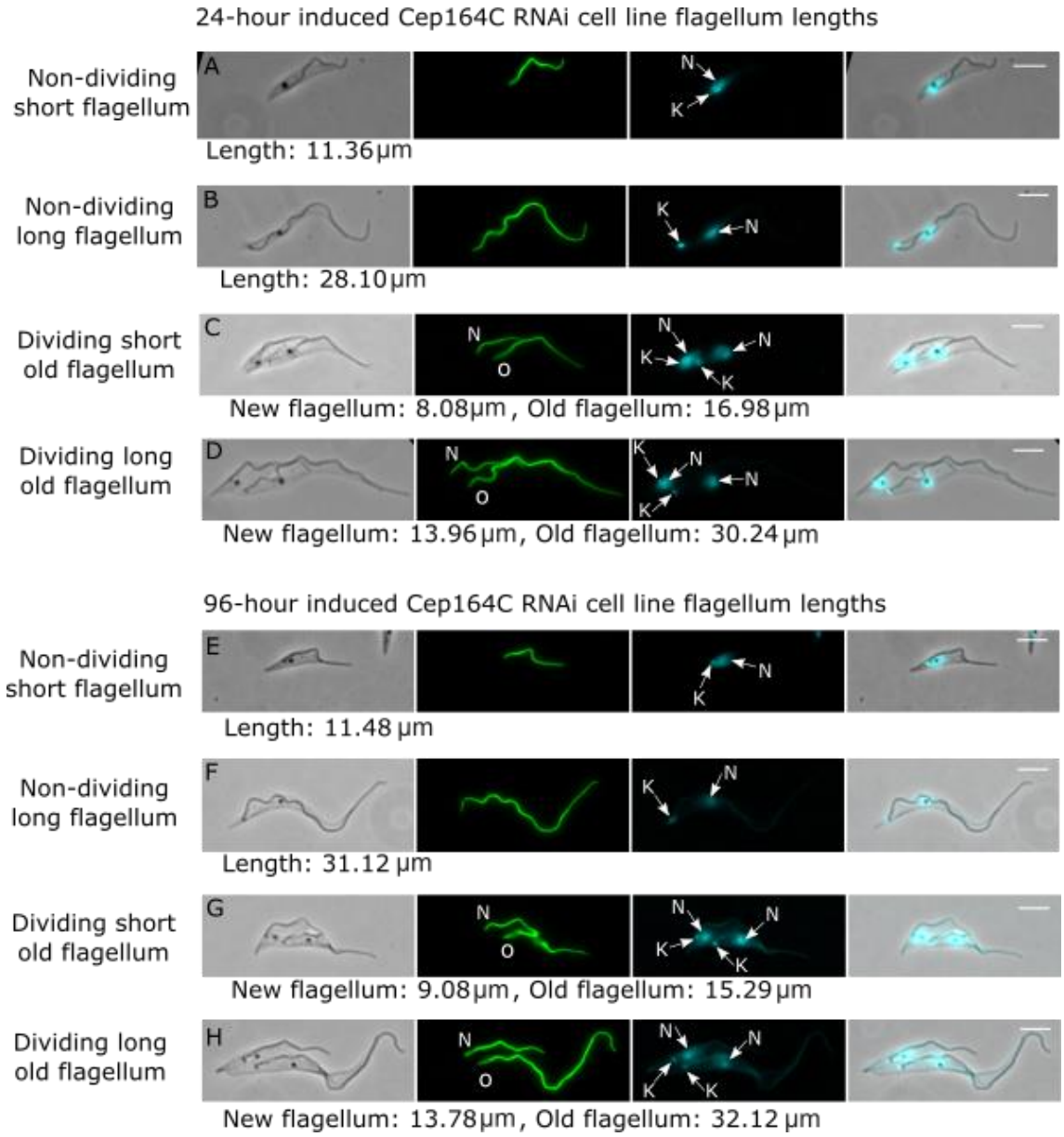


Figure 64: Knockdown of Cep164C causes extreme flagellum lengths

Induced cells containing the Cep164C RNAi construct were induced for 24 hour and 96 hours, collected on SuperFrost microscope slides, labelled with the PFR (flagellum) antibody and DAPI, and measured using the segmented tool on ImageJ. A wide variation in flagellum and cell lengths was observed in both the 24 hour induced 1K1N (A-B) and 2K2N (C-D) cells and the 96 hour 1K1N (E-F) and 2K2N (G-H) cells. O – Old flagellum, N – New flagellum. Scale bar – 5µm.

In summary, RNAi induction of Cep164C results in a wide variation of flagellum lengths. Shorter flagella could be formed by a defect in flagellum growth, however this does not explain how longer flagella are formed. One hypothesis is that there is further growth to the old flagellum after the cell has exited the cell cycle. Recent work by Bertiaux *et al* (2018) demonstrated that the new flagellum does not continue to grow after the end of the cell cycle and there is a 'lock' to growth. The results here suggest that knockdown of Cep164C could release this lock, allowing further growth of the flagellum after cytokinesis, or no lock was applied.

4.13 Longer flagellum correlate with a longer cell body and a longer unattached flagellum

After uncovering the longer flagellum phenotype, it was then questioned how cells were growing and incorporating the additional flagellum length of up to 37μm. To investigate this, cell body length and flagellum length of the *T. brucei* cell were plotted. Results showed that a shorter flagellum correlated with a shorter cell body in induced and uninduced cell lines (figure 65). The shortest flagellum length was 7μm, which had a cell body measurement of 8μm. The largest uninduced flagellum length was 23μm with a cell body length at 19μm (figure 65; orange). The induced cell lines with a longer flagellum of 33μm had a longer cell body length at 28μm (figure 65; green). This data confirmed that the cell body length and flagellum length are correlated, even in cells with a very long flagellum.

At the anterior end of normal cells, there is always a small portion of the flagellum which is unattached. It was investigated whether the increase in flagellum length was correlated with an increase of the unattached flagellum. 200 cells were measured for the length of the flagellum and the length of the unattached flagellum in G1 cells and the old flagellum of dividing cells in uninduced and induced cells.

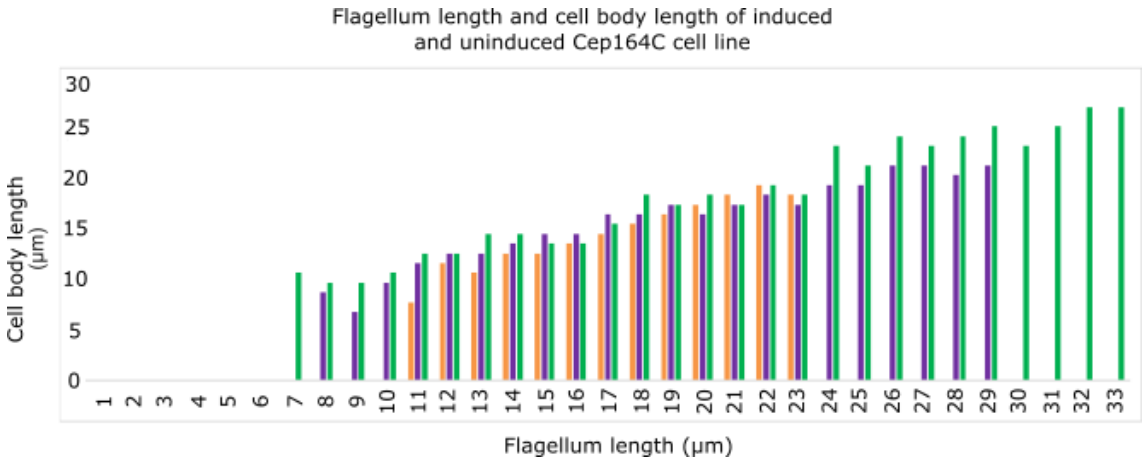


Figure 65: Cell body increases as flagellum length increases

Induced and uninduced cells containing the Cep164C RNAi construct were measured for their flagellum length and their cell body length through the segmentation tool in ImageJ. Measurements demonstrate a correlation, where cells with a longer flagellum have a longer cell body, and those with a shorter flagellum length have a shorter cell body. Orange – uninduced, Purple – 24 hour induced and Green – 96 hour induced.

The G1 uninduced population showed flagellum lengths between 11µm (2µm unattached flagellum), and 23µm (4-5µm unattached flagellum). A Pearson's correlation showed a strong correlation and statistically significant result at 0.891 ($p<0.0001$) (figure 66C; orange). The induced cells ranged from 7µm (with a shorter unattached flagellum at 1-2µm) to 33µm (with an unattached flagellum at 7-8µm). The Pearson's correlation showed a strong and statistically significant result at 0.979 ($p<0.0001$) (figure 66C; blue).

Measurements for the old flagellum of the dividing cells showed a similar result to the G1 cells. Flagellum lengths of the uninduced cells ranged from 15µm-25µm with the unattached flagellum length increasing from 3µm-5µm (figure 66D; orange), with examples shown in figure 66(A). The Pearson's correlation showed a statistically significant result at 0.751 ($p<0.0001$). The induced population ranged from 13µm (with an unattached flagellum at 2-3µm) to 37µm (with an unattached flagellum at 8-9µm). A good Perasons' correlation and a statistically significant results was seen at 0.789, ($p<0.0001$) (figure 66D; blue), with examples shown in figure 66(B).

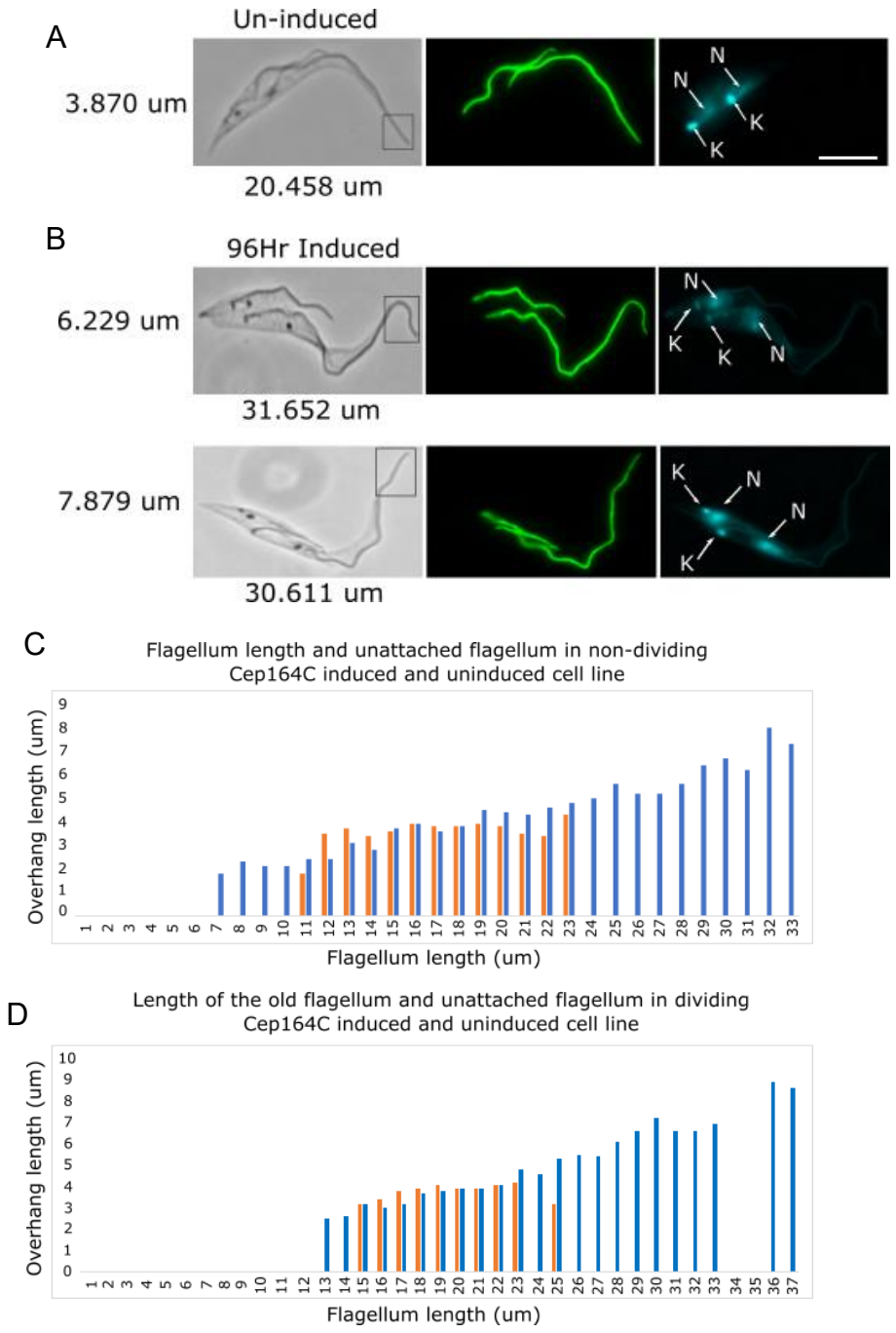


Figure 66: Unattached flagellum lengths increase as total flagellum length increases.

Uninduced (A) and induced (B) cells containing the Cep164C RNAi construct were measured for their unattached flagellum lengths through the segmentation tool in ImageJ. 200 cells of either G1 cells (C) or dividing cells (D) were measured and analysed. Analysis showed that cells with longer cell bodies have longer unattached flagella. K = Kinetoplast, N = Nucleus. Pearson's correlation was used. Scale bar = 10μm.

These results show that the induced and uninduced cell lines both have an increase in the length of their unattached flagellum, however those induced cells that generate a longer flagellum (due to lack of Cep164C expression) also generate a longer unattached flagellum. Overall, these results demonstrate that the increase in flagellum length is accomplished by cells increasing the length of their cell body AND increasing the length of the unattached portion of the flagellum.

4.14 The FRAP technique shows the Cep164C protein does not have a high turnover rate

It was questioned whether the Cep164C protein had a high turn-over rate and be actively maintained at the TFs to continue the “lock” and prevent growth of the old flagellum. For this, the Fluorescence Recovery After Photobleaching (FRAP) technique was carried out using live mNG::Cep164C cells. TFs were bleached using a strong excitation laser and recordings were captured for 54 seconds. Prior to bleaching, both mNG::Cep164C fluorophores were seen (figure 67; yellow and white arrows). After 5 seconds, one of the two mNG::Cep164C regions of interest was bleached (figure 67C; white arrow) and was not seen to recover to the TFs for the remainder of the recording (figure 67D-F; white arrow), suggesting this protein does not have a high turn-over rate.

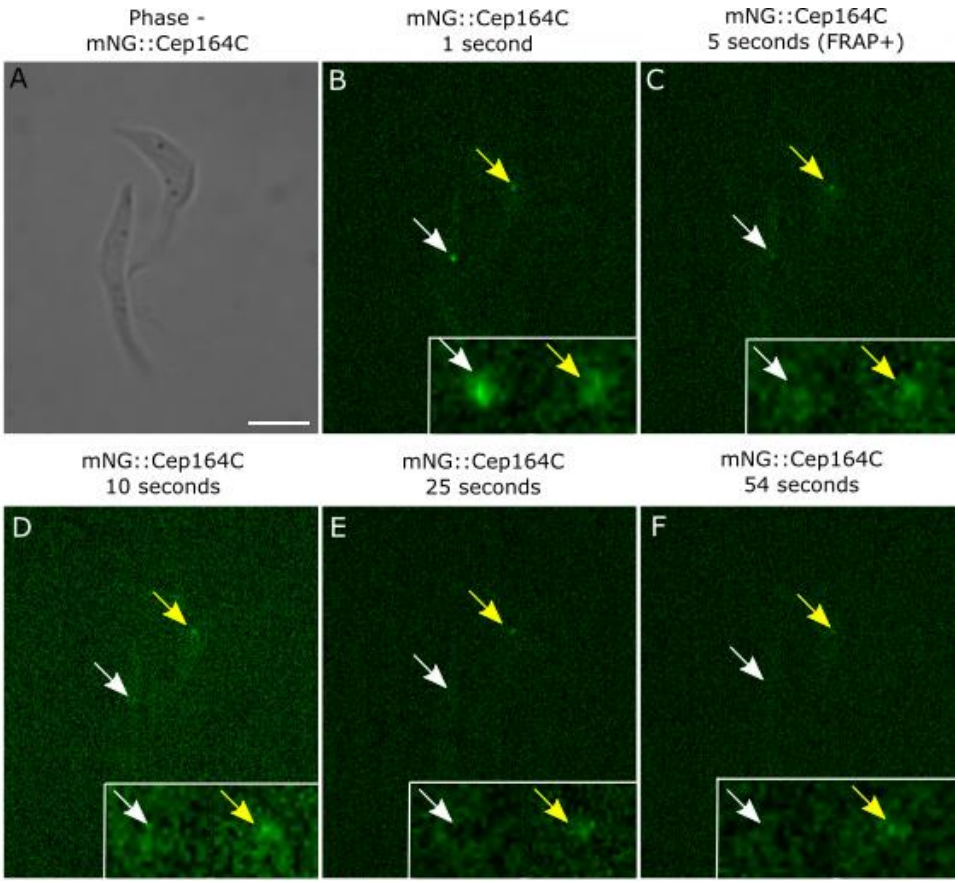


Figure 67: Cep164C does not have a high turn-over rate.

Live cells expressing the mNG::Cep164C were analysed on the fluorescence microscope using the FRAP technique (A-B). The mNG::Cep164C of one cell was bleached after 5 seconds of recording (C) and was not seen to recover throughout the remainder of the movie (D-F; white arrow). A nearby cell was not bleached, with the mNG::Cep164C being continuously present throughout the recording. (B-F; yellow arrow). Scale bar = 10μm.

4.15 SEM imaging confirms the Cep164C phenotype and shows longer and shorter cells.

Scanning electron microscopy of the uninduced and induced Cep164C cells confirmed that there were smaller cells with shorter flagella (figure 68B) and larger cells with longer flagella (figure 68C) in comparison to the uninduced cells (figure 68A). Those cells with a longer cell body showed a longer unattached flagellum, than is typically seen in uninduced cells (figure 68C-D; white arrow) and a longer posterior (figure 68C; yellow arrow), whereas in dividing cells, some had a smaller new daughter cell than expected (figure 68E; yellow arrow).

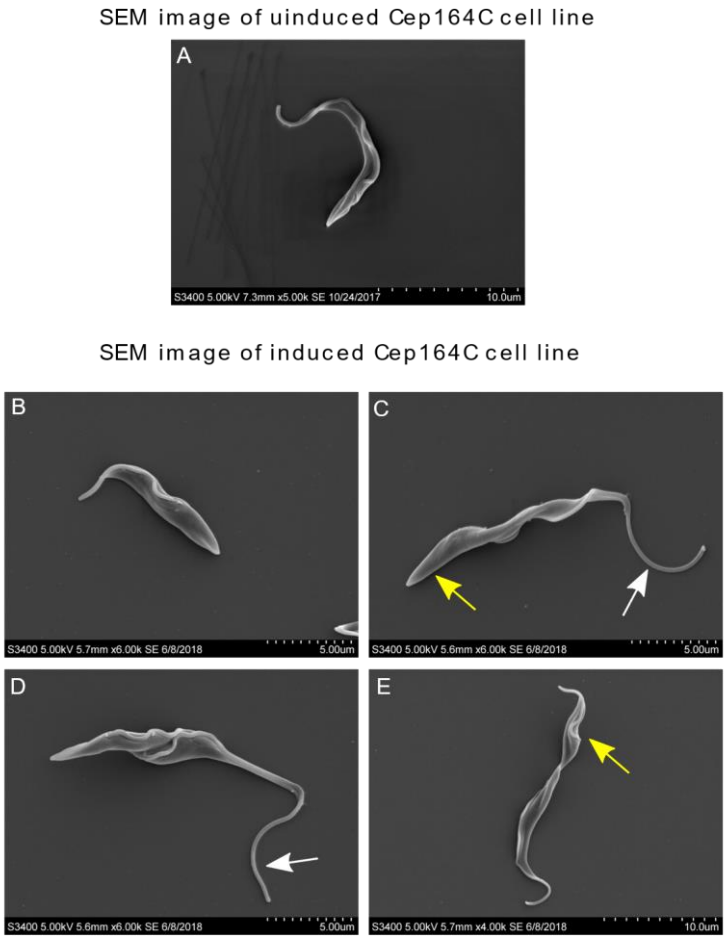


Figure 68: SEM shows extreme flagellum lengths in Cep164C cell line

SEM imaging of the Cep164C induced cell line showing smaller (B) and longer (C) cells than the uninduced (A), dividing cells with longer unattached flagella (D; white arrow) and dividing cells with smaller new daughter cells (E; yellow arrow).

Cells prior to abscission had one longer daughter cell (likely to be the old daughter cell) and one smaller daughter cell (new daughter cell), with the size differences between the two being apparent (figure 68E). Images were taken between 5.6-5.7 mm x 6.00k on the SEM (apart from figure 68A and E), showing the difference in size between figure 68B and 68C.

4.16 SBF-SEM imaging further confirms longer and shorter flagellum lengths and misplaced organelles

The serial block-face scanning electron microscopy technique was employed to further demonstrate the Cep164C phenotypes. SBF-SEM is a method to obtain automated sequential serial sections of cells and tissues, allowing three-dimensional reconstruction (Cocks *et al,* 2018; Hughes *et al,* 2013). Cell pellets of induced (24hrs) and uninduced cells were fixed for transmission electron microscopy (TEM). 456 (for uninduced) and 521 sequential serial sections (for induced) (slice thickness was 75nm) were obtained and reconstructed using iMOD and modelled using MIB (Belevich *et al,* 2016). For the induced cell line, one whole dividing cell was modelled. This model shows a short new daughter cell and flagellum length (figure 69A and 69B; pink) and a longer old daughter cell and flagellum length (figure 69A and 69B; yellow), with a longer unattached flagellum (figure 69B; white arrow). The model also confirmed the misplaced organelles, with the nuclei being too close together (figure 69A; white arrows) and incorrect positioning of the BBs.

SBF-SEM images of the induced Cep164C cell line with longer flagella and misplaced organelles

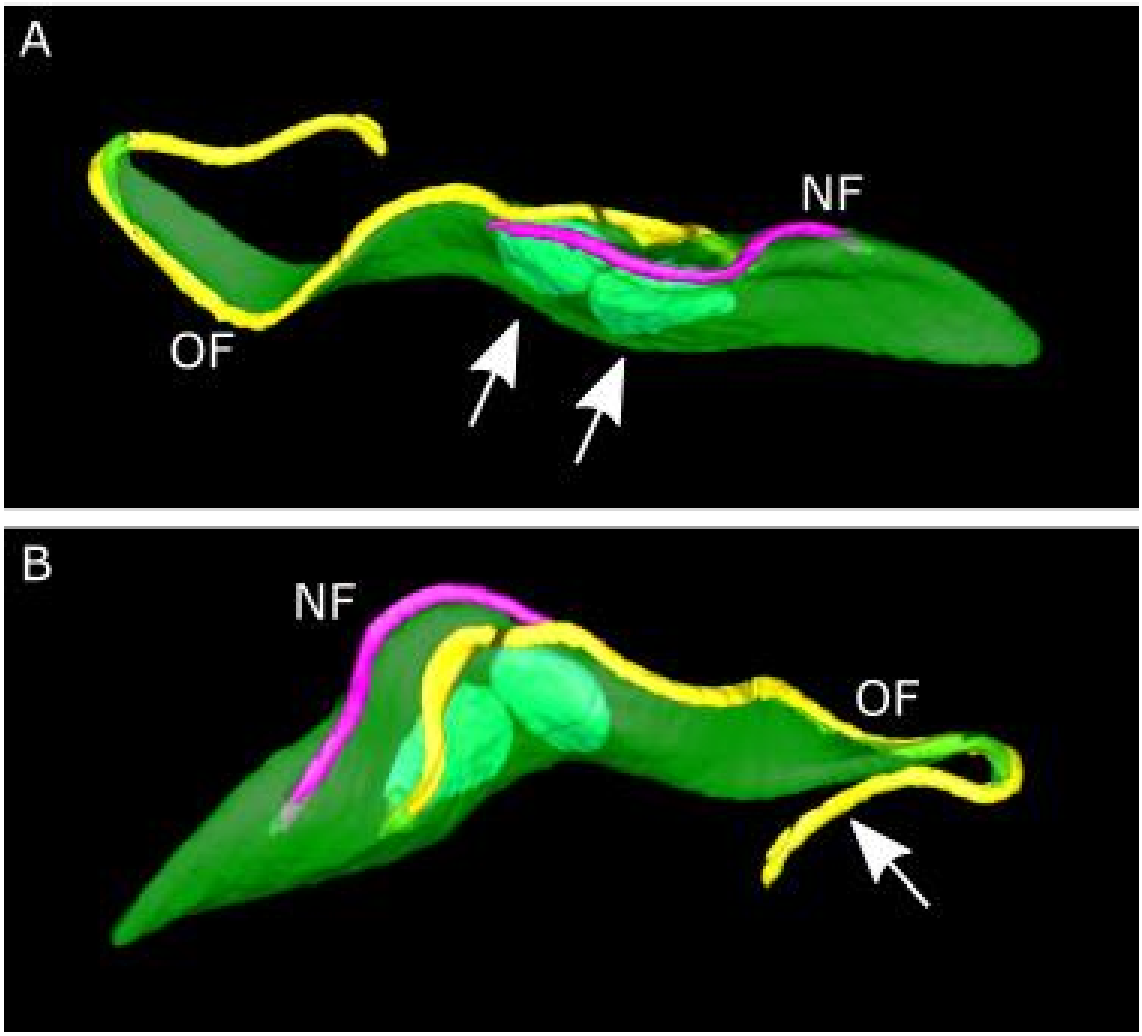


Figure 69: SBF-SEM shows longer flagella and organelle displacement

One dividing whole cell of the induced Cep164C cell line was modelled after induction, showing a shorter new flagellum (A-B; pink) and a longer old flagellum (A-B; yellow) with a longer unattached flagellum (B; white arrow).

4.17 SBF-SEM analysis shows an increase in flagellar pocket volume in induced Cep164C cells.

The Cep164C protein is localised to the TF area of the BB which docks the BB to the flagellar pocket. Both the TFs and TZ are important in controlling trafficking into the flagellum. It was therefore investigated whether there was an increase or decrease in flagellar pocket volume in the Cep164C induced cell line.

Firstly, G1 cells containing only one flagellar pocket were analysed (figure 70D) by modelling the flagellar pocket in each of the sections to work out the volume (figure 70E). 20 uninduced (figure 70A) and 19 induced (figure 70B) flagellar pockets from the Cep164C cell line were modelled and measured using MIB software. Flagellar pockets were modelled with the MBB and PBB (figure 70A and 70B; white arrows). The Cep164C uninduced pocket volumes ranged from 0.18μm^3 to 0.85μm^3 whereas the Cep164C induced population generated larger pocket volumes ranging from 0.44μm^3 to 1.5μm^3 (figure 70C). On average, the induced cell line generated a flagellar pocket volume of 0.78μm^3, whereas the uninduced generated a volume of 0.55μm^3, there was a statistically significant difference of p-value <0.005 between these values.

Using the SBF-SEM technique whole individual cells were obtained, allowing the identification of dividing cells (figure 70G). 21 induced and 20 uninduced dividing cells were extracted from the serial section dataset. The new and old pockets from each of these cells were modelled and the volumes calculated.

New Flagellar pockets: The volume of the uninduced new flagellar pockets ranged from 0.07μm^3 to 0.43μm^3, with an average of 0.25μm^3. The induced new flagellar pocket ranged from 0.06μm^3 to 1.71μm^3, with an average of 0.31μm^3 (figure 70F). This was not statistically significantly different (p-value = 0.548).

Old flagellar pockets: The volume of the uninduced old flagellar pockets ranged from 0.04μm^3 to 0.50μm^3, with an average of 0.25μm^3. The induced old flagellar pocket ranged from 0.08μm^3 to 2.13μm^3, with an average of 0.46μm^3 (figure 70F). This result showed the induced volume was statistically significantly larger than the uninduced volume (p-value <0.05).

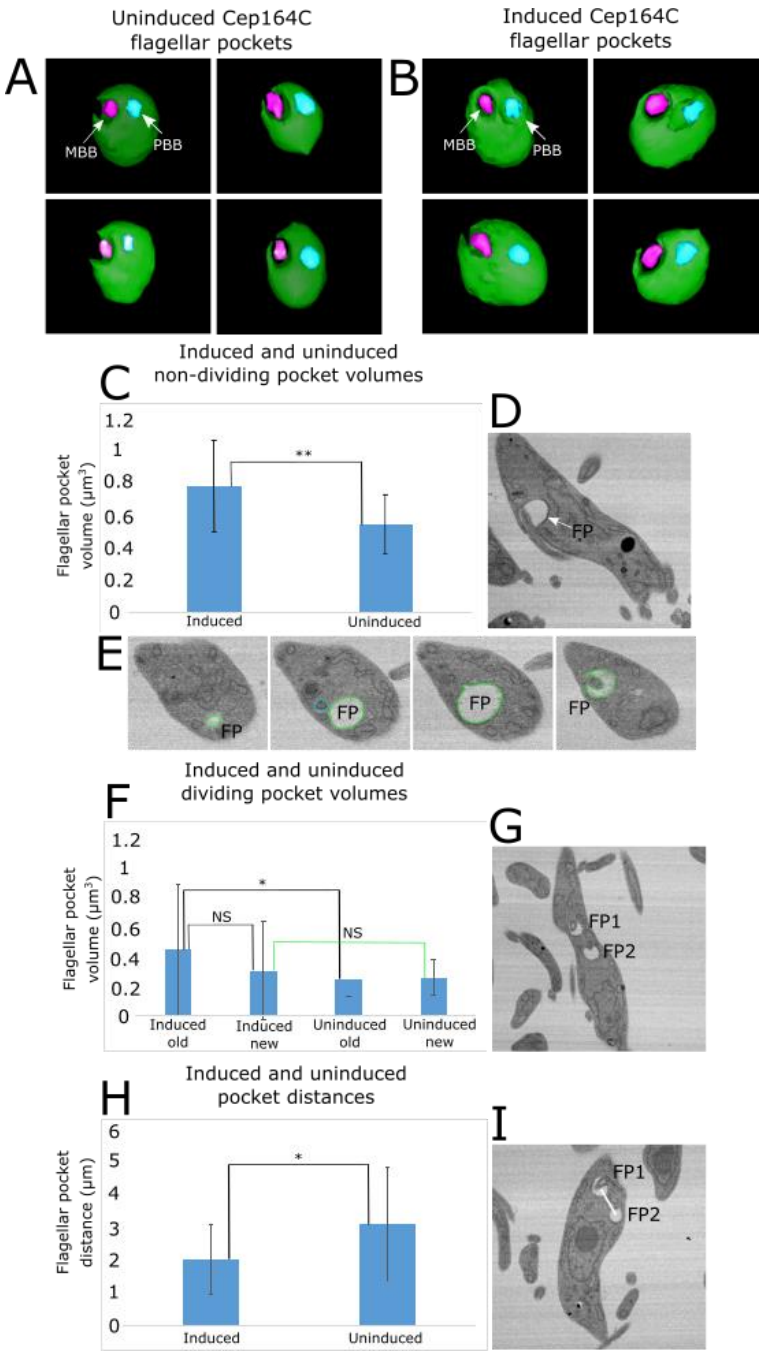


Figure 70: Knockdown of Cep164C causes old pocket volume increase

Uninduced and induced cells of the inducible Cep164C RNAi cell line were prepared for the SBF-SEM, where sections were collected from the samples and analysed through iMOD and MIB. From the sections collected, the pocket volumes from the Cep164C uninduced (A) and induced (B) cells were modelled and analysed. Pocket volumes of the G1 cells (C-E) and dividing cells (F-G) were analysed, and the pocket distances between the induced and uninduced were also compared (H-I). Independent t-test were used. NS = not significant, * = $p \leq 0.05$ and ** = $p \leq 0.01$.

Induced old and new flagellar pockets: Volume measurement data collected for both the induced old flagellar pocket and the induced new flagellar pocket were then analysed. Results of the induced old flagellar pocket volumes (0.08μm^3 to 2.13μm^3) and induced new flagellar pocket volumes (0.06μm^3 to 1.71μm^3) showed that there was not a statistically significant difference between these values (p-value = 0.234) (figure 70F).

This data supports that the old flagellar pocket volumes in the induced cell line is statistically significantly larger than the old flagellar pocket volume in the uninduced cell line, with no significant volume increase seen for either new flagellar pocket. However no statistically significant volume differences were seen between the new and old induced flagellar pockets. Volumes were also increased in the induced Cep164C cell line for the G1 cells with a single flagellar pocket, in comparison to the uninduced.

Previous fluorescence microscopy in the Cep164C RNAi cell line indicated that the distances between the kinetoplasts and BBs in dividing cells were closer in the induced cells, therefore the distance between both flagellar pockets was investigated (figure 70I). The average distance between the flagellar pockets in dividing cells was 2.1μm for induced cells and 3.1μm for the uninduced population. These results show that the induced pockets of the Cep164C cell line are statistically significantly closer in distance in comparison to the uninduced (p-value <0.05) (figure 70H).

4.18 Flagellar pocket volumes in the uninduced cell line increases in size as the dividing cells progress through the cell cycle

The modelled uninduced Cep164C cell line raised questions on whether the flagellar pockets have a different volume, dependent on the stage of the cell cycle the dividing cell is in. In those cells which contained the Cep164C RNAi plasmid but were not induced, 27 new flagellar pockets and 27 old flagellar pockets (of the same cell) were modelled.

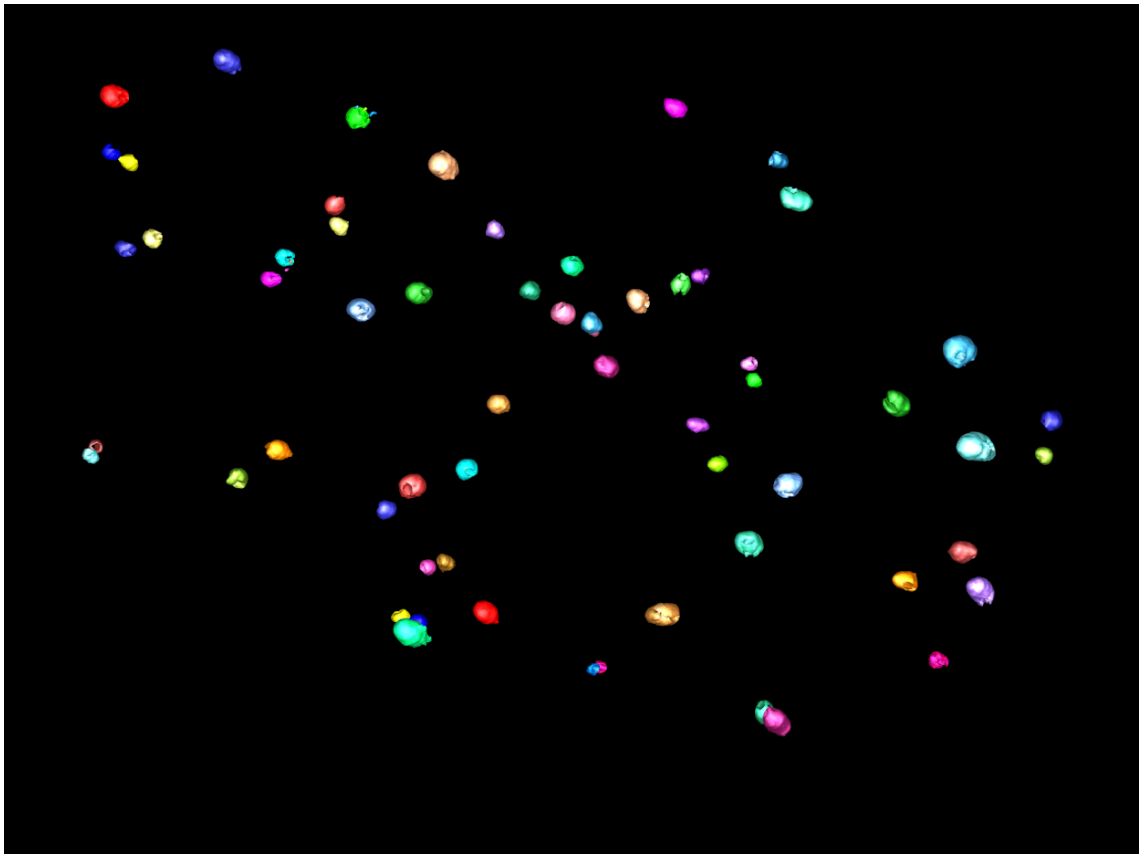


Figure 71: SBF-SEM modelling of flagella pocket volumes

Uninduced cells from the inducible Cep164C RNAi cell line were prepared for the SBF-SEM, where sections were collected from the sample and analysed through iMOD and MIB. From the data collected, cells containing whole pockets were measured and modelled. Cells were modelled as either early dividing cells or late dividing cells and each flagellar pocket is modelled above. From this data 27 uninduced new and old pocket volumes were measured.

Uninduced cells ranging from early dividing cells to late dividing cells based on the distances from the pockets were modelled (figure 71). Results showed that those in early stages of division had a much smaller pocket volume for the old and new flagellar pocket, in comparison to those which were in the later stages of the cell cycle (figure 72). The old and new flagellar pocket volumes also appeared to increase in volume at the same time.

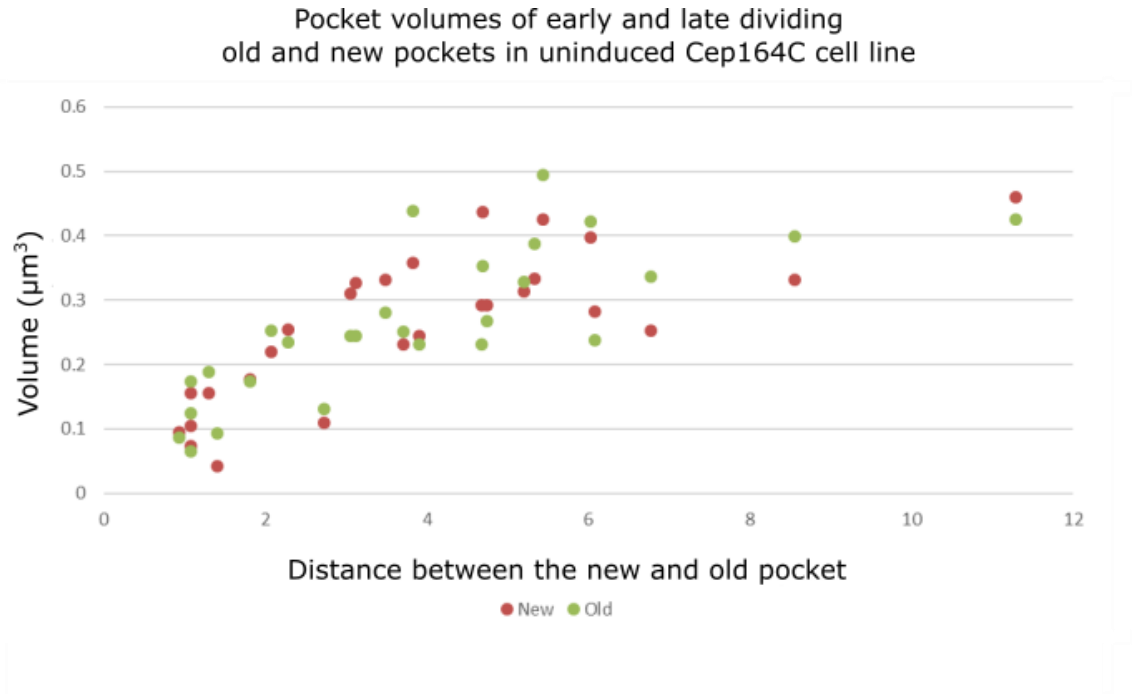


Figure 72: Cell cycle progression leads to pocket volume increase

The sections collected from the uninduced cells from the inducible Cep164C RNAi cell line on the SBF-SEM were analysed and modelled. Pockets modelled were from cells ranging from early dividing cells to late dividing cells and showed that early dividing cells had a much smaller flagellar pocket volume than those in later stages of division.

4.19 Axoneme and basal body structures develop as expected in the Cep164C induced cell line

After observing the phenotype of the RNAi induced Cep164C cell line, cells were then observed using the TEM to discover any observable ultrastructural changes in the flagellum and the TZ. As the Cep164C protein localises immediately distal to the TF protein TbRP2, it was investigated as to whether the Cep164C protein functions in regulating proteins entering the cilium. As the RNAi Cep164C cell line was shown to generate cells with longer or shorter flagella than expected for wild-type cells, the morphology of the induced axoneme was investigated. 32 uninduced axonemes were observed in cross-section on the TEM, with all but one axoneme showing a 9+2 microtubule arrangement (figure 73A and 73B).

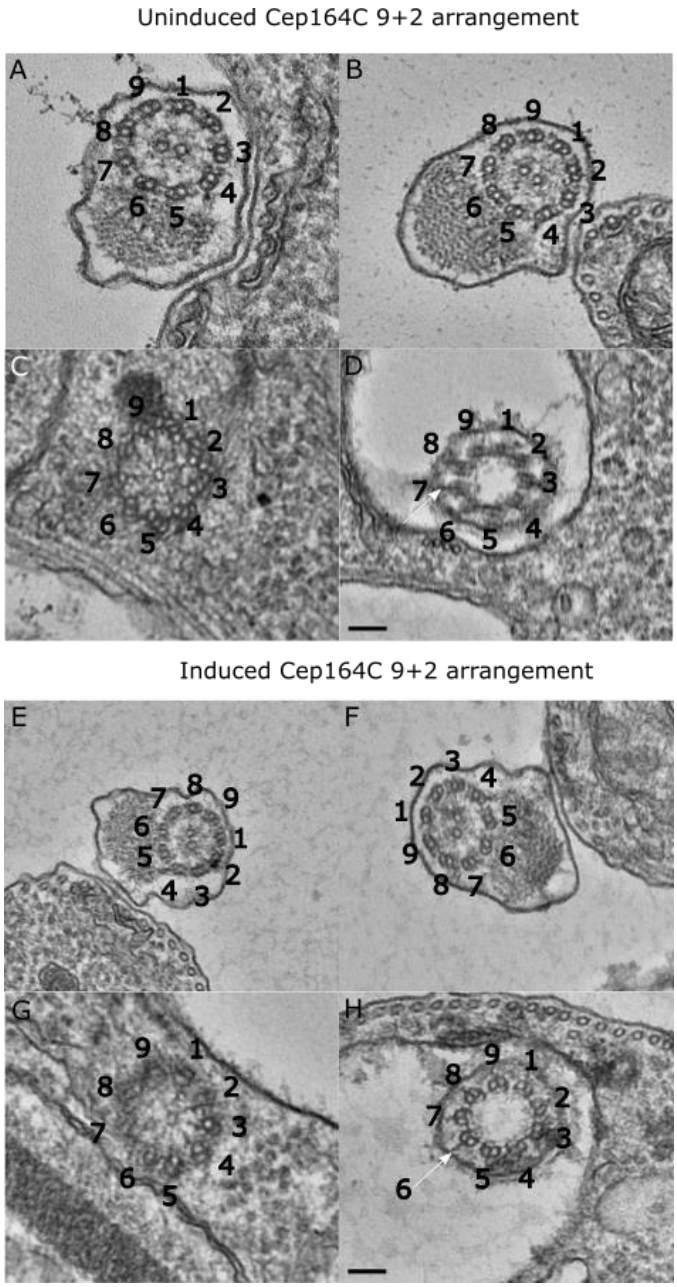


Figure 73: Knockdown of Cep164C shows no morphological differences at basal body structures

TEM imaging of uninduced (A-D) and induced (E-H) ultrastructural components develop normally in the Cep164C induced cell line. A, B, E and F – Axonemes. C and G – Cartwheel structure. D and H – Transition zone. Scale bars 100nm

In comparison to the uninduced axonemes, the induced axonemes (45) showed a similar outcome, with the majority having a 9+2 configuration (figure 73E and 73F). Four showed a lower number of triplet microtubules, however these lacked a PFR, suggesting it is a growing flagellum and unlikely to be due to the lack of the Cep164C protein.

TEM imaging also showed one uninduced cell that had a typical 9+0 microtubule arrangement of the proximal section of the BB (figure 73C). Two proximal BBs of the induced cell line showed a normal 9+0 microtubule arrangement (figure 73G). The cross-section of the TZ was also imaged for comparison.

Five uninduced cells of the Cep164C cell line showed a TZ containing a typical 9+0 microtubule arrangement, with Y-linkers protruding from the microtubules (figure 73D; white arrow). In comparison, four TZs from the induced cell line also showed a 9+0 microtubule arrangement and Y-linkers (figure 73H). Observation of these structures suggests that no difference is seen in the development of the axoneme, proximal section of the BB or the TZ for the induced Cep164C cell population.

4.20 Transition zone develops normally in the Cep164C RNAi cell line

Next, the length of the transition zone between the uninduced and induced cell lines were analysed. 14 longitudinal sections of the TZ were captured from the uninduced cell line (figure 74A and 74B), showing the lengths of the TZs ranging from 0.225μm to 0.456μm. In comparison 9 longitudinal sections of the TZ from the induced Cep164C cell line was captured (figure 74C and 74D), ranging 0.370μm to 0.500μm. The average length was 0.385μm for the uninduced TZ and 0.428μm for the induced TZ. Although on average the induced TZ lengths were slightly longer than the uninduced TZ lengths, there was no statistical significant difference between the two (p-value = 0.120) (figure 74E).

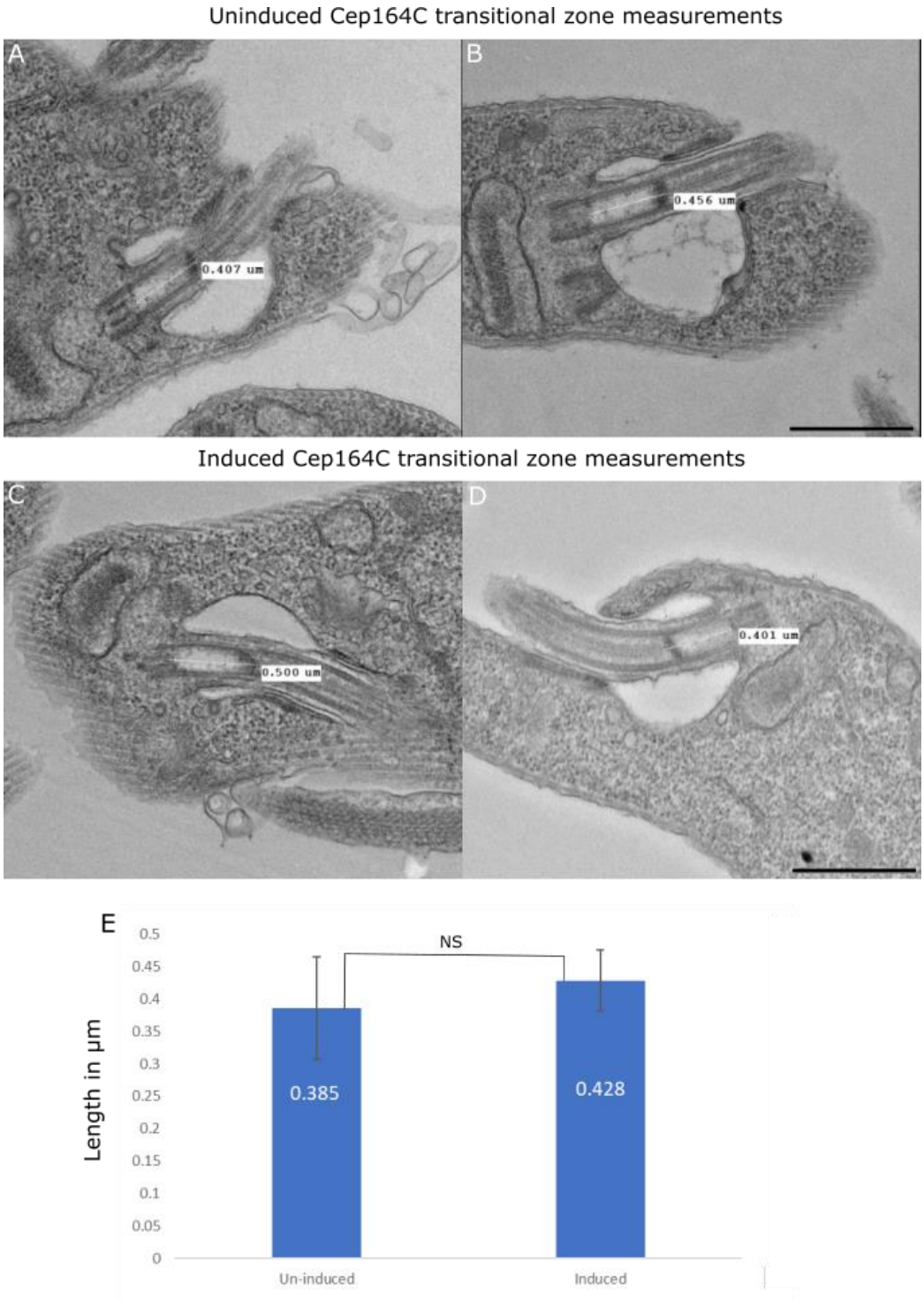


Figure 74: TEM imaging shows no transition zone phenotypes

The transition zones from the uninduced (A-B) and induced (C-D) Cep164C RNAi cell lines were imaged and measured. No statistically significant differences were observed between the uninduced and induced for their transition zone lengths (E). Independent t-test was used. NS = not significant. Scale bars 500nm.

5. Investigation of Cep164A and Cep164B results

In results chapter four, one of the identified Cep164 proteins was investigated to further understand the function in the *T. brucei* cell. After investigation the Cep164C protein was hypothesized to 'lock' the old flagellum and prevent extension after the end of the cell cycle, with the RNAi cell line thought to break this mechanism allowing the old flagellum to extend further. Hodges *et al (2010*) identified three Cep164 proteins in *T. brucei,* with the two remaining proteins Cep164A (Tb927.5.2440) and Cep164B (Tb927.11.11650) being investigated in this chapter. Dang *et al* (2017) localised the Cep164A and Cep164B proteins to the basal body area of the cell, however further investigation has been conducted to discover the functional role of these proteins.

5.1 Cep164A and B localise to both the mature and newly matured basal bodies

The localisation of both proteins was carried out in detergent-extracted cells expressing either the mNG::Cep164A and mNG::Cep164B, in the background of an mCherry::SAS-6 cell line for co-localisation. SAS-6 was used to confirm whether either protein co-localised to the BBs. As previously mentioned, the SAS-6 protein labels both the MBB (figure 75A-H; white arrow head) and the PBB (figure 75A-H; white arrow). This work was carried out in collaboration with Dr Jiri Tyc (Vaughan lab). Dr Jiri Tyc generated and observed the localisation for mNG::Cep164B cell line, while I generated and observed the localisation for mNG::Cep164A cell line.

Analysis of the G1 cells, containing one BB and one flagellum (figure 75A and 75E; orange arrow) showed that both Cep164A and Cep164B localised to the distal section of the MBB only (figure 75A-H; white arrow head). Interestingly, cells which had duplicated their BBs and were building a new flagellum (figure 75B and 75F; orange arrow) were positive for Cep164A and Cep164B at the distal section of the newly matured basal body. Both proteins continued to localise to the newly MBB of the new flagellum (figure 75C and 75G; green arrow) and the old MBB of the old flagellum (figure 75C and 75G; blue arrow) in post-mitotic cells. No localisation to the PBBs was observed at any point throughout the cell cycle (figure 75A-H; white

arrow). A similar localisation to the MBB was also obtained in whole cells (figure 75D and 75H; white arrow head).

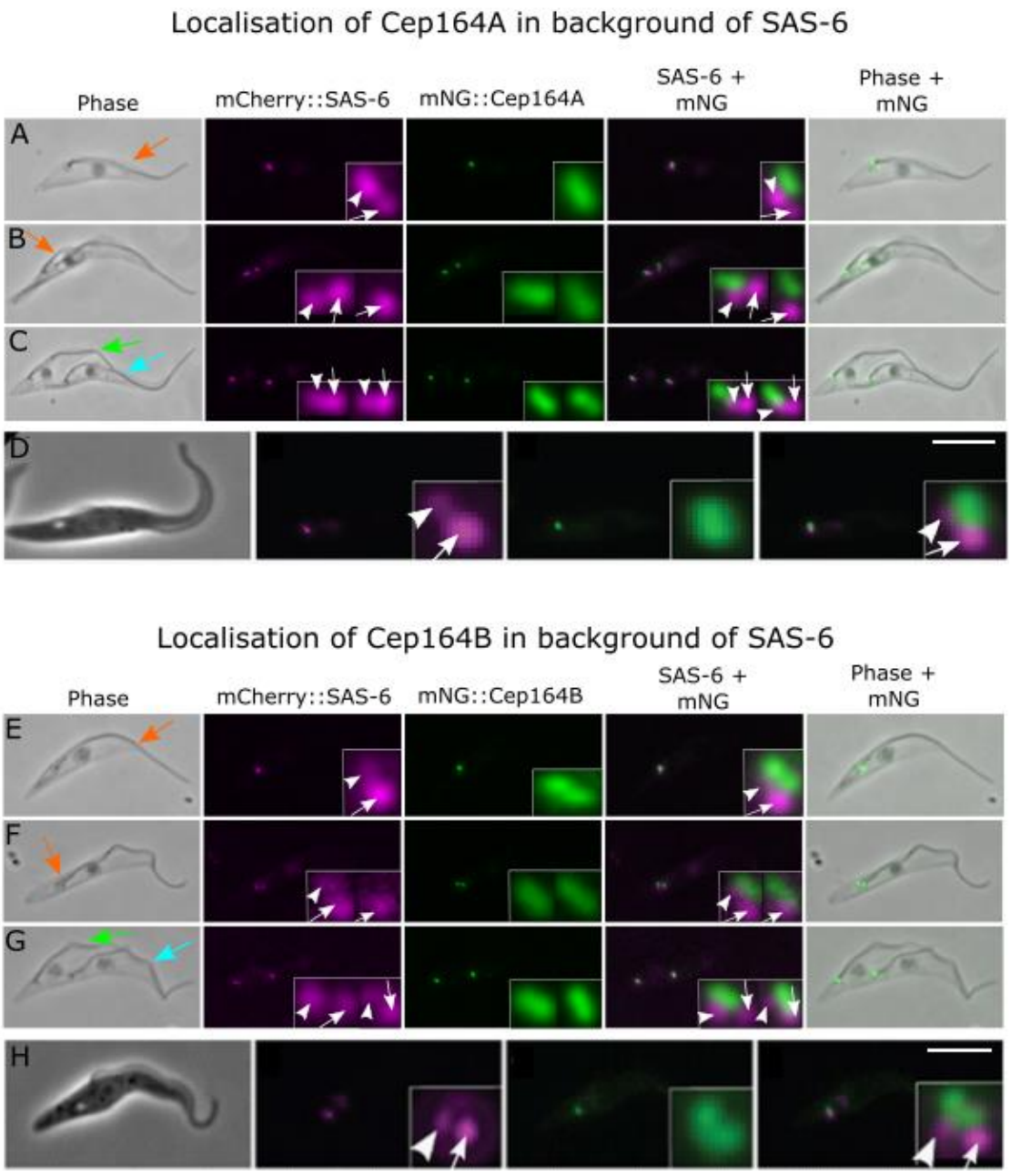


Figure 75: Cep164A and B localise to the mature and newly mature basal body.

mNG::Cep164A (A-D) and mNG::Cep164B (E-H) in the background of mCherry::SAS-6 localised distal to the mature basal body in 1K1N (A and E) and to the mature and newly mature basal body in dividing cells (B-C and F-G) throughout the cell cycle. Localisation in whole cells was also seen at both the newly mature and old mature basal body for Cep164A (D) and Cep164B (H). Scale bar = 10μm.

These results show an interesting difference between the three Cep164 proteins. Although all proteins localised distal to the SAS-6 protein, Cep164A and Cep164B localised to the mature and newly mature BBs, whereas Cep164C only localised to the old MBB only (section 4.1). This localisation to both mature BBs suggests that these proteins have a different role to the Cep164C protein. As both proteins were present throughout the cell cycle, it is hypothesized that they were not cell cycle-dependent. Further analysis was carried out to confirm or reject this hypothesis.

5.2 Fluorescent microscopy of Cep164A and B shows constant recruitment to the mature basal body

It was previously identified that Cep164C localised to the MBB in a cell cycle dependent manner, with recruitment prior to BB duplication and absence prior to cytokinesis (section 4.1). As Cep164A and Cep164B are part of the same protein family, it was investigated whether these proteins had a similar localisation pattern. Analysis of 100 mNG::Cep164A and mNG::Cep164B cells showed these proteins were constantly present at the MBBs throughout the cell cycle (figure 76A and 76B). These results suggest that neither protein is cell cycle dependent. Instead it is hypothesised that they are recruited after maturation of the BB (figure 75B and 75E) and remain localised throughout the life of the cell.

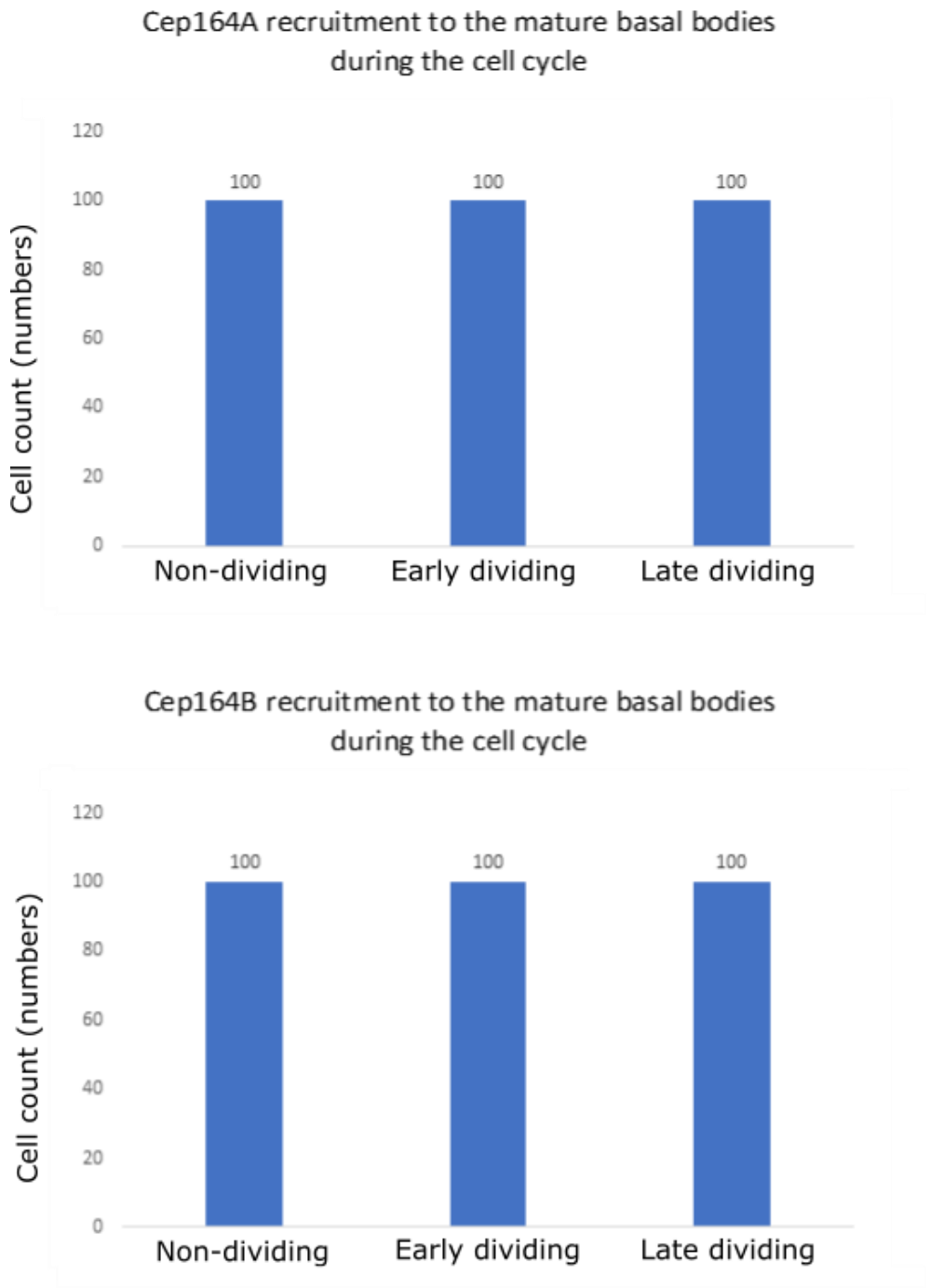


Figure 76: Cep164A and B are constantly recruited to the mature basal bodies

100 mNG::Cep164A and mNG::Cep164B cells were imaged and analysed on the fluorescence microscope for the presence or absence of the Cep164A or Cep164B protein throughout the cell cycle. Analysis showed that both proteins were present at all stages through the cell cycle.

5.3 Cep164A and Cep164B localises immediately distal to TbRP2

As both proteins localised distal to the SAS-6 protein, it was hypothesized that they may localise to the TFs, similar to the Cep164C protein (section 4.4). Transitional fibres, as previously described are situated at the distal section of the BBs and are essential for BB/C docking to the flagellar pocket/plasma membrane (Wei *et al,* 2015; Vaughan and Gull, 2016; Stephan *et al,* 2007). To identify further where the proteins localised, cells expressing the mNG::Cep164A and mNG::Cep164B proteins were labelled with the YL1/2 antibody (labelling TbRP2 at the TFs as described in section 4.4).

Cells in the G1 stage of the cell cycle, with one flagellum had a localisation on the MBB, immediately distal to the YL1/2 (TbRP2) labelling (figure 77A and 77E; white arrow). Both proteins continued to localise distal to TFs on the MBB and newly MBB for the remainder of the cell cycle (figure 77C-D and 77G-H; white arrow). No localisation was observed on the PBBs for either protein (figure 77A and 77H; white arrow head).

In summary, Cep164A and Cep164B localised immediately distal to the TbRP2 protein, but no full co-localisation was observed. Both proteins appeared to localise to the same area in which the Cep164C protein was located (co-localisation was shown with Airyscan imaging in section 4.6), suggesting all three proteins localise to the same structure.

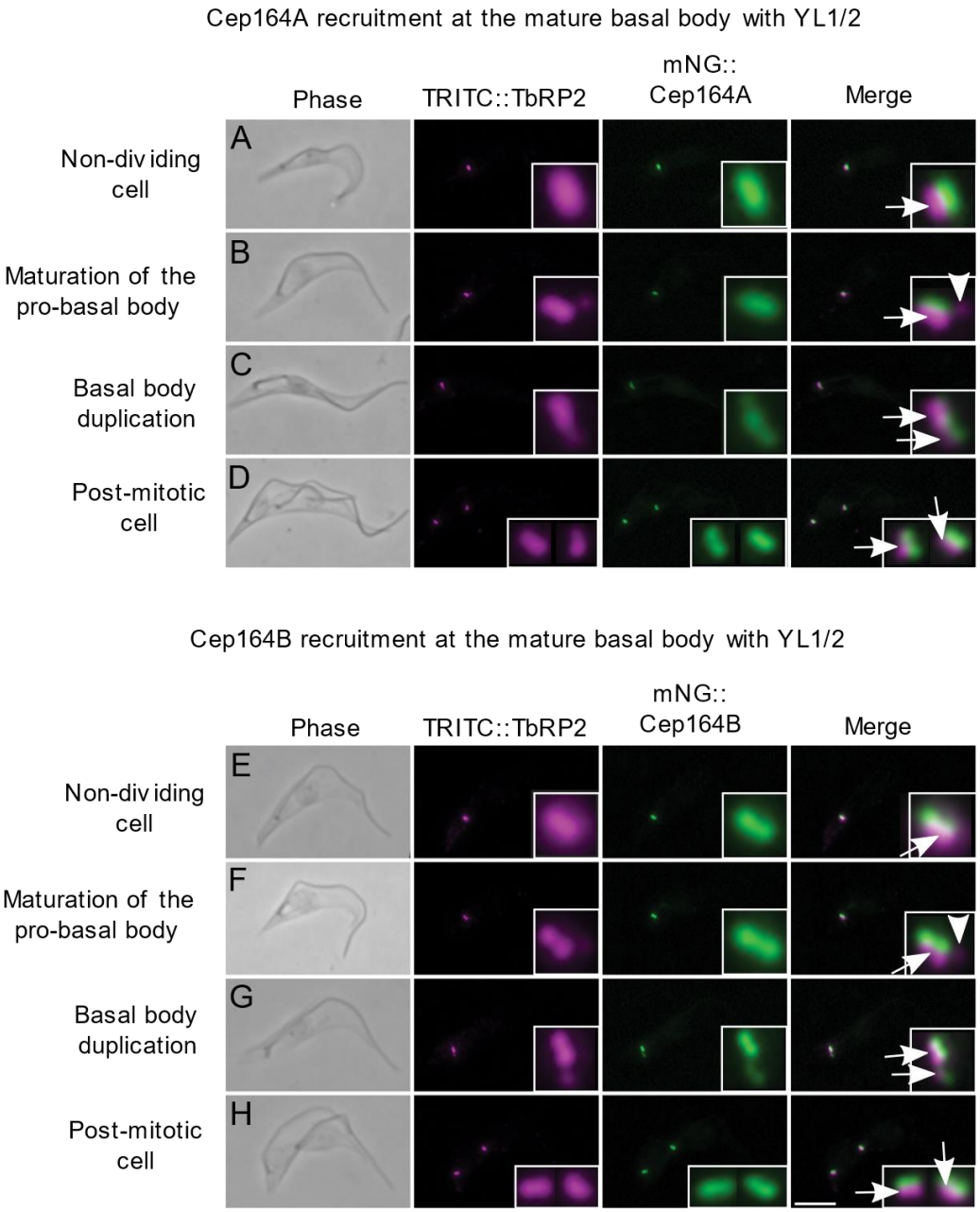


Figure 77: Cep164A and Cep164B localise distal to the TbRP2 protein

Analysis of detergent-extracted cells of mNG::Cep164A and mNG::Cep164B show a localisation of the proteins distal to the TbRP2 protein (magenta) on the mature basal body and newly matured basal body in G1 (A and E) and dividing (B-D and F-H) cells. Scale bar = 10μm.

5.4 Addition of plasmids causes minimal growth defects

After confirming Cep164A and Cep164B localised distal to the TFs, double tagged cell lines of Cep164A and Cep164B (in the background of Cep164C) with inducible RNAi plasmids were generated and analysed. Prior to assessing the induced cell lines, growth curves were carried out on the uninduced mCherry::Cep164A + mNG::Cep164B (table 15M and L), mCherry::Cep164A + mNG::Cep164C (table 15F) and mNG::Cep164B + mCherry::Cep164C (table 15I) cells.

Uninduced growth curves in the double tagged cell line and the cell lines containing the RNAi construct to either Cep164A or Cep164B was carried out as described in section 4.7 and compared against the SmOx p927 parental cell line. The SmOx p927 cells lack the pPOTv6 or RNAi constructs and can be compared to demonstrate fitness cost to the cell line. Cell lysates were collected to identify the amount of Cep164A or Cep164B protein present.

In all growth curve experiments, cells were sub-cultured to 1 x 10^6 each day and over 4 days an accumulated growth curve was generated. In each experiment, the parental SmOx p927 cells grew to an accumulated growth rate between 1 x 10^9 – 1 x 10^{10} (figure 78A, 79A and C; dark green line). The cell line containing the double tagged Cep164 proteins but lacking the RNAi construct had a slightly lower accumulative growth rate, but still between 1 x 10^9 – 1 x 10^{10} (figure 78A, 79A and C; red line). The cell lines containing the RNAi construct to either Cep164A or Cep164B showed a slightly lower accumulative growth rate to the double tagged cell lines, between 1 x 10^8 – 1 x 10^9 (figure 78A; light green, 79A; purple line and 79C; light green line). Apart from the mCherry::Cep164A and mNG::Cep164C (construct against Cep164A) which had an accumulate growth rate between 1 x 10^9 – 1 x 10^{10} (figure 79A; purple line). These results indicate that the addition of either construct (pPOTv6 or pQuadra) creates only a minimal effect on the growth rate of the cells.

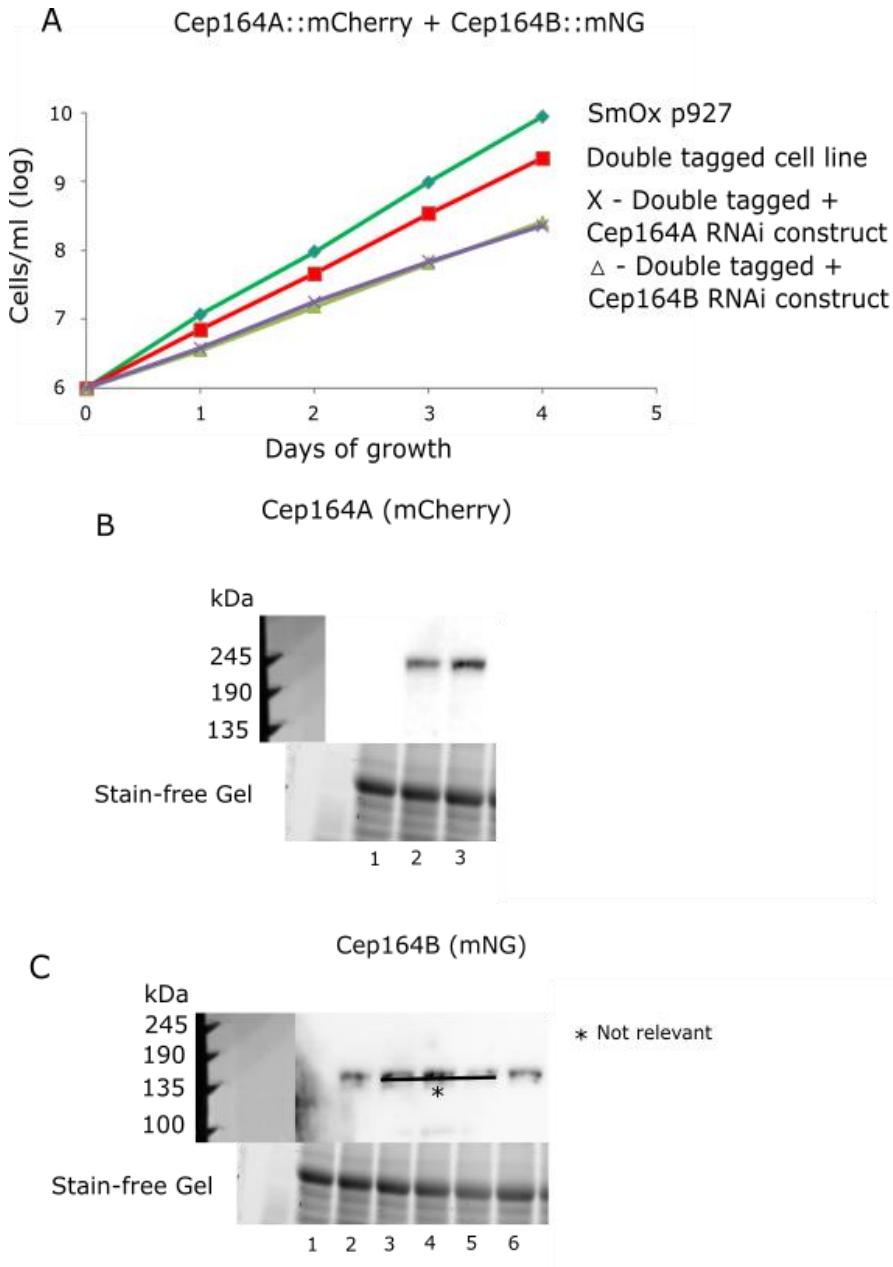


Figure 78: Addition of constructs in uninduced Cep164A+B cell lines show minimal growth defects

A) Growth curves for SmOx cells, the Cep164A/B double tagged cell line and the cell lines containing the Cep164A or Cep164B RNAi construct. Those containing the RNAi construct had the slowest growth rate, with the double tagged cell line also showing a lower growth rate to the SmOx. B-C) Western blots were run for all cells lines shown in A. No TY1 tags were detected by the BB2 antibody for the SmOx cells as expected (B; lane 1 and C; lane 1). TY1 tags were detected by western blot for the uninduced double tagged cell line (B; lane 2 and C; lane 2) and the cell lines containing the RNAi constructs of Cep164A (B; lane 3) and Cep164B (C; lane 6).

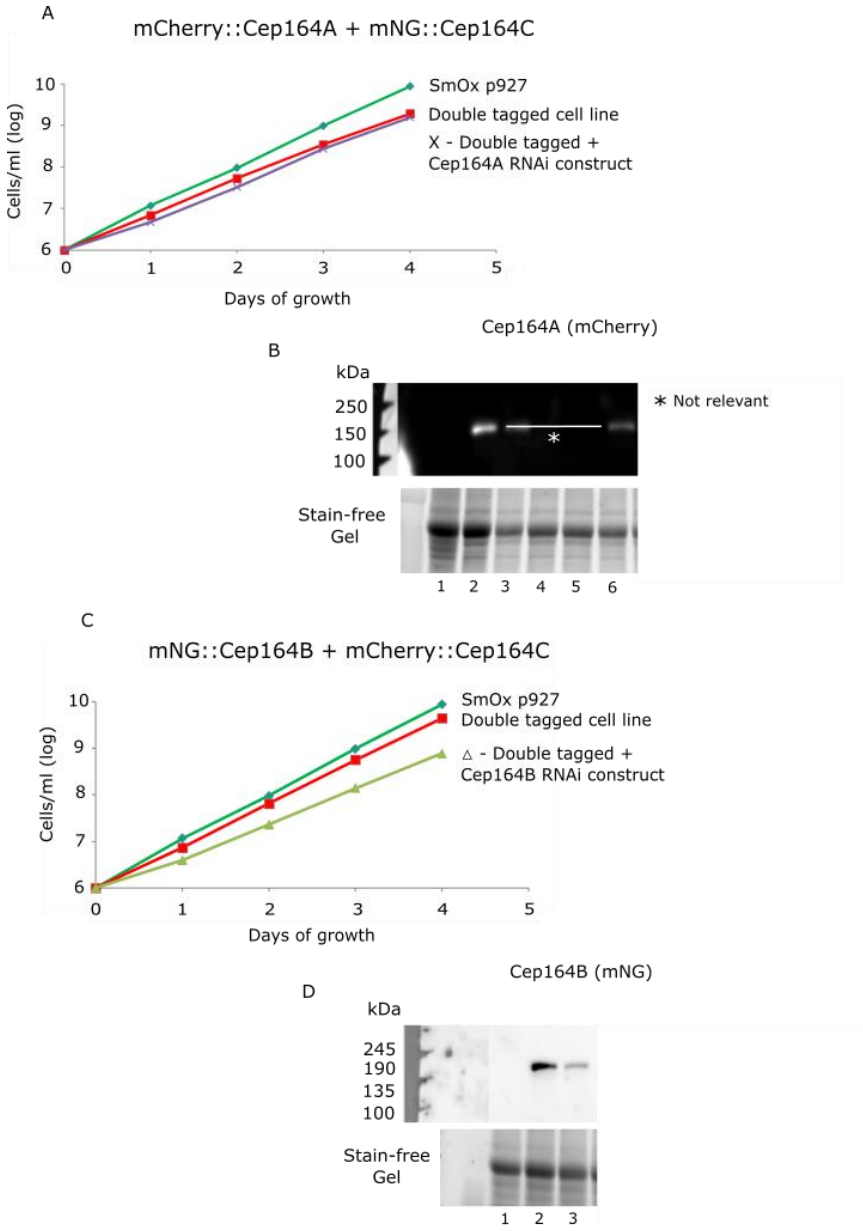


Figure 79: Addition of pPOTv6 or RNAi constructs in uninduced cell lines show minimal growth defects

A + C) Growth curves of the SmOx cells, the Cep164A/C (A) and Cep164B/C (B) double tagged cell lines and the cell lines containing either the Cep164A or Cep164B RNAi construct. Those containing the RNAi construct had the slowest growth rate, with the double tagged cell line also growing to a lower density. B-D) Western blots were run for cell lines shown in A and C. No TY1 tags were detected by the BB2 antibody for the SmOx cells as expected (B; lane 1 and D; lane 1). TY1 tags were detected for the uninduced double tagged cell lines (B; lane 2 and D; lane 2) and the cell lines containing the RNAi constructs of Cep146A (B; lane 6) and Cep164B (D; lane 3).

As described in section 4.7, plasmids contain the TY1 epitope tag which is recognised by the BB2 antibody and was collected in the cell lysates to conduct western blots. The western blots showed that no BB2 antibody was detected for the SmOx p927 parental cell line as expected (figure 78B; lane 1 and C; lane 1 and 79B; lane 1 and D; lane 1). The double tagged cell lines (figure 78B; lane 2 and C; lane 2 and 79B; lane 2 and D; lane 2) and the double tagged cell lines containing the Cep164A/B RNAi construct (figure 78B; lane 3, 78C; lane 6 and 79B; lane 6 and 79D; lane 3) were all positive for the BB2 antibody. The molecular weight for the Cep164A and Cep164B protein without additional plasmids is approximately 131kDa and 111kDa respectively, and approximately 161kDa and 141kDa containing the fluorescent tag. However, the double tagged cell lines were detected on the western blot between 150-240kDa (figure 78B-C and 79B and D). The above experiments were conducted in collaboration with Dr Jiri Tyc (Vaughan lab). Both myself and Dr Jiri Tyc worked together in all aspects to conduct the experiment.

5.5 Silencing of the Cep164A or B protein does not create further growth defects

After discovering that the introduced constructs (pPOTv6 or pQuadra) had little effect on the growth rates, the induced cell lines were then analysed. Cells were induced using doxycycline for a period of 4 days, sub-cultured to 1 x 10^6 each day where the cell population was measured on a Beckman Coulter and plotted. Growth rates were carried out on the induced cell lines of mCherry::Cep164A + mNG::Cep164B (table 15M and L), mCherry::Cep164A + mNG::Cep164C (table 15F) and mNG::Cep164B + mCherry::Cep164C (table 15I), with an RNAi construct against one of the proteins, to observe if any growth phenotype is seen against the uninduced cell lines. Growth rates of the induced cells in the background of either Cep164A, Cep164B or Cep164C (figure 80A, D; red line and figure 81A, D; red line) had a similar accumulative growth rate to the uninduced cell line (figure 80A, D; green line and 81A, D; green line). To assess the knockdown of Cep164A or Cep164B in the double tagged cell lines, a western blot was carried out. Both

proteins were identified by the TY1 tag (as explained in section 4.8). The cell lysates for the uninduced cell lines produced a visible band as expected (figure 80C; lane 3, 80F; lane 6 and figure 81C; lane 6 and 81F; lane 3).

All induced cell lines failed to produce a visible band, suggesting a successful knockdown of the Cep164A or Cep164B proteins (figure 80C; lane 4-5, 80F; lane 7-8, 81C; lane 7-8 and 81F; lane 4-5). Results showed a lack of detectable Cep164A (mCherry) or Cep164B (mNG) protein after induction and that reducing the protein only causes a minimal effect to growth. Cells had an accumulative growth, growing either similar to or only slightly slower than the uninduced cell line (figure 80A, D and 81A, D)

5.6 Cep164A and B are dependent on each other for their localisation to the basal body

Fluorescent microscopy investigations to further confirm whether the Cep164A or Cep164B RNAi cell lines were successful and did not have a positive detection of the knocked-down protein showed interesting results for the mCherry::Cep164A + mNG::Cep164B cell lines (table 15M and L).

mCherry::Cep164A + mNG::Cep164B (with inducible RNAi construct for Cep164A)

Induction showed that Cep164A was no longer present in the cell as indicated by the western blot (figure 80C; lane 4-5) and was confirmed by the lack of the mCherry::Cep164A fluorophore (figure 80B; white arrows). However, the western blot was still detecting TY1 tags in the mNG::Cep164B protein (figure 80C; lane 7-8) but showed no localisation of the mNG::Cep164B protein to the basal body (figure 80B; white arrows).

mCherry::Cep164A + mNG::Cep164B (with inducible RNAi construct for Cep164B)

Investigation showed Cep164B produced similar results to the Cep164A RNAi cell line. After induction, western blots showed that no mNG::Cep164B protein was present in the cell (figure 80F; lane 7-8), with confirmation through fluorescence microscopy showing no detectable mNG fluorophore (figure 80E; white arrows). However, the western blot was detecting the TY1 tags in the mCherry::Cep164A protein (figure 80F; lane 4-5), however no localisation of the mCherry::Cep164A protein was seen at the BB (figure 80E; white arrows).

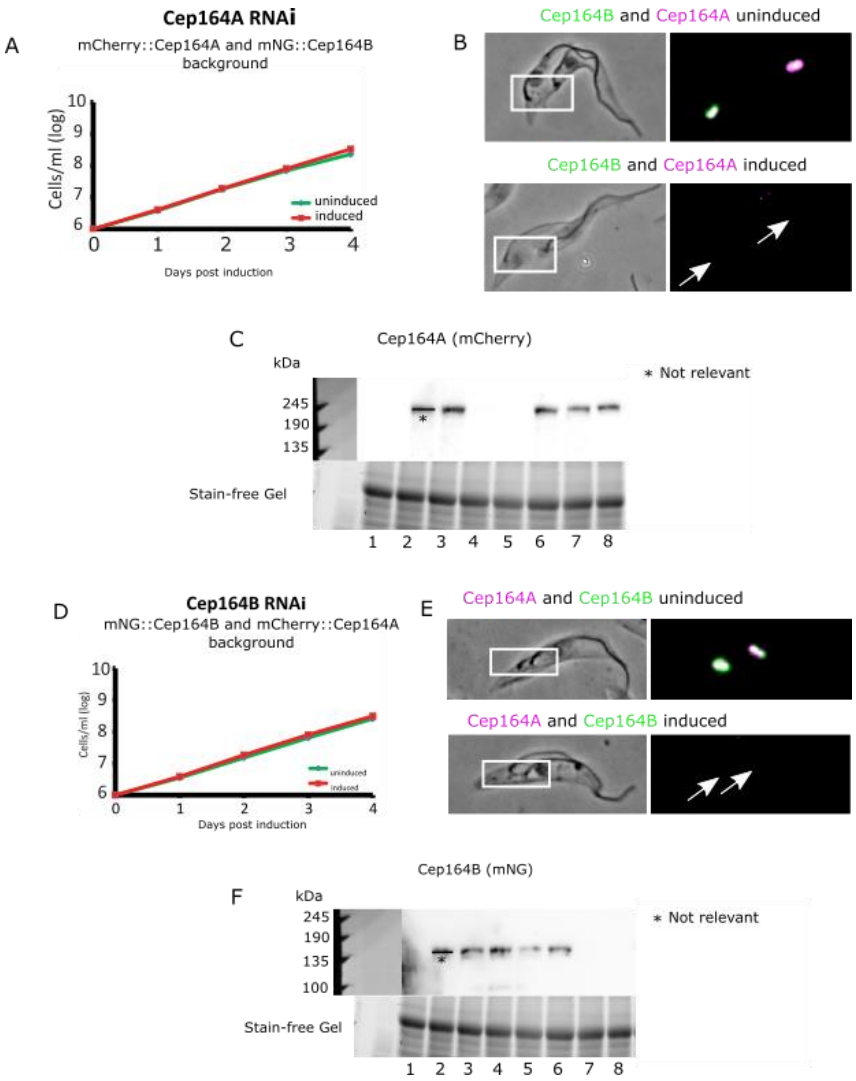


Figure 80: Cep164A and B require each other for basal body location.

A + D) Growth curves of the cell lines containing the Cep164A RNAi construct (A) and the Cep164B RNAi construct (D) in comparison to the uninduced cell lines. Those induced cells contaning the RNAi constructs grew to similar densities to the uninduced cell lines. B + E) The induced and uninduced cell lines containing the Cep164A RNAi construct (B) and the Cep164B RNAi construct (E) were observed on the flourescent microscope, no expression was observed for either mNG::Cep164 proteins after induction of the RNAi constructs. C + F) Western blots showed no detection of TY1 tags in the SmOx as expected (C; lane 1 and F; lane 1). TY1 tags were detected for the uninduced Cep164A (C; lane 3) and Cep164B (F; lane 6) cell line, however no TY1 tags were detected for the induced Cep164A (C; lane 4-5) and Cep164B (F; lane 7-8). Interestingly, the opposite tagged Cep164 protein was present in the western blot for Cep164A (C; lane 6-8) and Cep164B (F; lane 3-5).

In summary, induction of the Cep164A or Cep164B protein leads to loss of detection of the fluorophore using fluorescent microscopy (as expected), but also leads to the loss of detection of the opposite tagged Cep164 protein (A or B). Although the tagged proteins without the RNAi construct are not localised to the BB, they are still present in the cell. This suggests the Cep164A and Cep164B proteins are dependent on each other for correct localisation to the BB organelle and may require each other to correctly carry out their function. The same experiment was carried out for the Cep164C protein (section 4.8) which showed a successful knockdown, however both Cep164A and Cep164B were still present at the BBs.

5.7 Induction of Cep164A or B causes the Cep164C protein to localise to the newly matured basal body

After discovering that Cep164A and Cep164B may require each other for localisation to the BB, the mCherry::Cep164A + mNG::Cep164C and mNG::Cep164B and mCherry::Cep164C cell lines were investigated to uncover whether a similar result was produced. Interestingly, localisation of the Cep164C protein in these cell lines produced a different result to the induced cell lines containing mCherry::Cep164A + mNG::Cep164B (either Cep164A or Cep164B RNAi construct).

mCherry::Cep164A + mNG::Cep164C (with inducible RNAi construct for Cep164A)

The western blot showed that RNAi induction of the mCherry::Cep164A protein failed to detect any TY1 tags (figure 81C; lane 7-8), with fluorescent microscopy confirming no detection of the mCherry fluorophore (figure 81B; white arrow). Analysis of the cell line also showed that the Cep164C protein was detected on both the western blot (figure 81C; lane 4-5) and through fluorescent microscopy. It was also noted that without the Cep164A or Cep164B protein present on the BB, the Cep164C protein was now observed on the new MBB at a lower intensity to the old MBB (figure 81B; compare white arrow to yellow arrow).

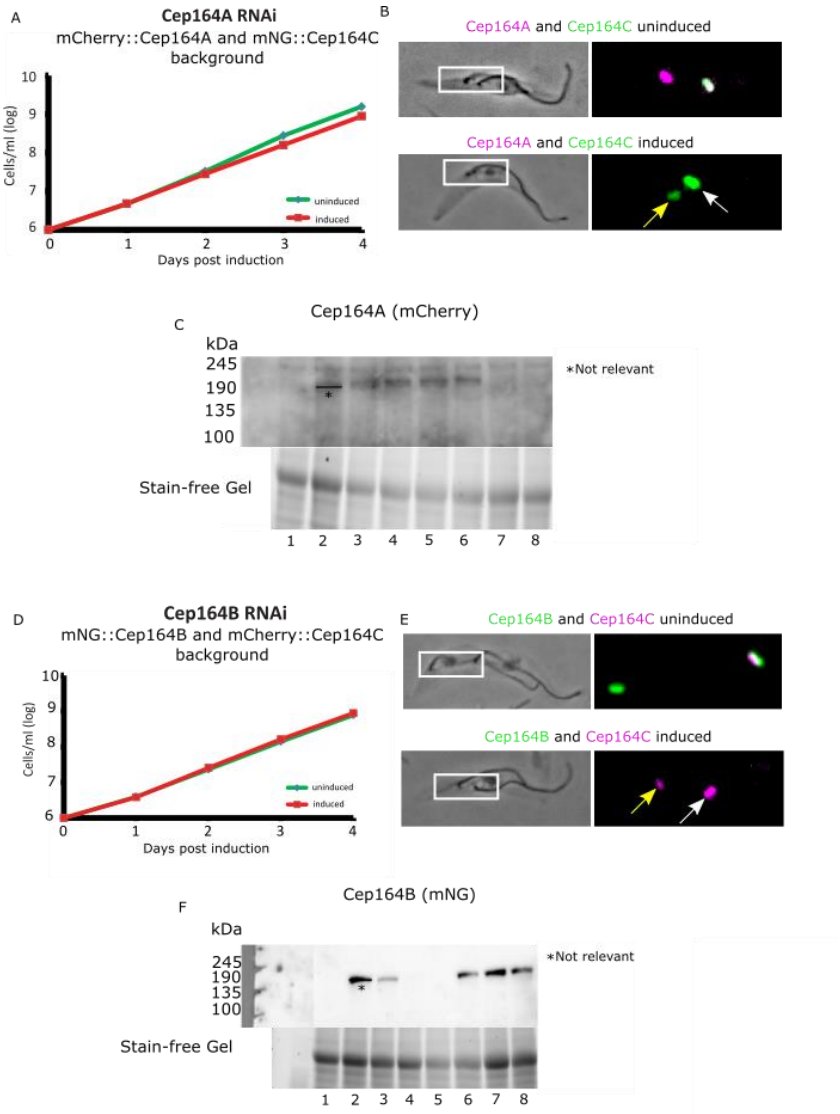


Figure 81: Loss of Cep164A and B forces Cep164C to the newly matured basal body.

A + D) Growth curves of the cell lines containing the Cep164A RNAi construct (A) and the Cep164B RNAi construct (D) with mNG::Cep164C in the background, in comparison to the uninduced cell lines. Those induced cells contaning the RNAi constructs grew to similar densities to the uninduced cell lines. B + E) After induction of the cell lines contaning the RNAi constructs, depletion of the Cep164A (B) and Cep164B (E) proteins caused the Cep164C protein to also localise to the newly mature basal body (B and E; yellow arrow), which is not observed under normal circumstances. C + F) Western blots showed no detection of TY1 tags in the SmOx cell line as expected (C; lane 1 and F; lane 1). TY1 tags were detected for the uninduced Cep164A (C; lane 6) and Cep164B (F; lane 3), however no TY1 tags were detected for the induced Cep164A (C; lane 7-8) or the Cep164B (F; lane 4-5) proteins. Scale bar = 10μm.

mNG::Cep164B and mCherry::Cep164C (with inducible RNAi construct for Cep164B)

A similar result was seen for the RNAi inducible cell line for Cep164B. Induction of this cell line showed no positive detection of the TY1 tags in the western blot for Cep164B (figure 81F; lane 4-5), with fluorescent microscopy confirming no detection of the mNG fluorophore (figure 81E; white arrow). However, TY1 tags were detected for the tagged Cep164C protein on the western blot (figure 81F; lane 7-8) and the mNG fluorophore was detected under fluorescent microscopy (figure 81E; white arrow). However, without Cep164B or Cep164A at the BB, Cep164C was seen to localise to the newly MBB at a lower intensity to the old MBB (figure 81E; compare white arrow to yellow arrow).

In summary, successful knockdown of either the Cep164A or Cep164B proteins produce a cell containing neither protein at the BB. It is then hypothesized that due to the lack of either protein, the cell re-locates some of the Cep164C protein from the old MBB to the newly MBB, at a lower intensity. It is possible to hypothesize that at least one of the Cep164 proteins has to be present on the BBs at all times after maturation. It is interesting to note that although Cep164A and Cep164B depend on each other for localisation to the BB, knockdown of Cep164C does not affect the localisation of either Cep164A or Cep164B.

5.8 Induction of Cep164A, B or A+C cell lines generate normal cell counts

As described in section 4.9, cells can be labelled with the DNA marker DAPI and the cell cycle stage can be determined through counting the number of kinetoplasts or nuclei in each cell. As the addition of constructs to the cells showed a minimal growth defect in comparison to the uninduced cell line, it was investigated into whether the number of G1 and dividing cells are normal. The mCherry::Cep164A + mNG::Cep164B (either Cep164A or Cep164B RNAi construct) and the mCherry::Cep164A and mNG::Cep164C double tagged and double RNAi cell line was investigated.

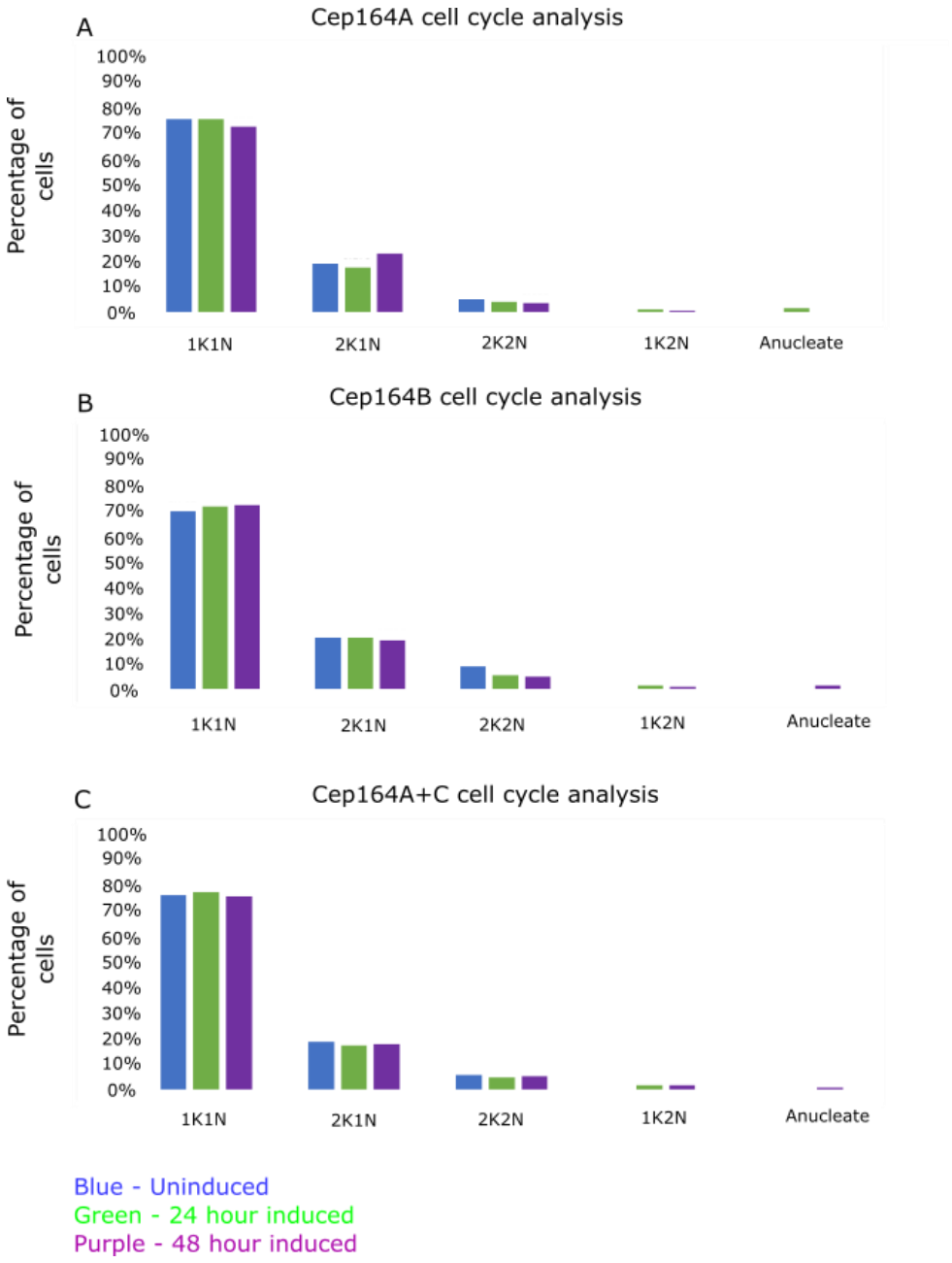


Figure 82: Induction of Cep164 RNAi cell lines produce a normal cell cycle analysis

200 cells containing either the Cep164A (A), Cep164B (B) or Cep164A+C (C) RNAi construct were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope to identify the cell cycle stage of the cell. Both uninduced and induced cells were imaged. Induction of all cell lines showed a normal cell cycle analysis in comparison to the uninduced cell lines, with only a small number of cells showing an unusual 1K2N configuration or anucleate cells. 1K1N – G1 cells. 2K1N – Early dividing cells. 2K2N – Late dividing cells.

A count of the number of kinetoplasts and nuclei in 200 cells was carried out in 24 hour, 96 hour and uninduced cell lines. In all three experiments, the uninduced and induced cells showed 70-75% of cells were observed in the G1 cell cycle stage (1K1N) (figure 82A, B and C). Both the induced and uninduced had between 10-20% of cells in the early dividing stages (2K1N) (figure 82A, B and C). A smaller number (10%) of cells were in the later stages of the cell cycle (2K2N) for both the induced and uninduced for all cell lines (figure 82A, B and C). Analysis of the induced cell line identified a small amount of cells with an unusual 1 kinetoplast and 2 nuclei (1K2N) morphology (figure 82A, B and C; 1K2N), with no cells showing this phenotype in the uninduced cell line. Another small amount (1-2%) of cells were aberrant, containing anucleate cells in the induced cell line. Overall, the induced cells were seen to have a similar cell cycle analyses to the uninduced cells, with the addition of a few cells showing an unusual 1K2N phenotype.

5.9 RNAi induction of the Cep164A, B and A+C cell lines generates asymmetrical daughter cells and abnormal morphological phenotypes

As observation of the cells showed a normal cell cycle analysis, the next step was to observe the morphology of the cells. Careful morphological analysis of the induced cells versus the uninduced cells showed a significant number of abnormalities. Induction of the mCherry::Cep164A + mNG::Cep164B (either Cep164A or Cep164B RNAi construct) and the mCherry::Cep164A and mNG::Cep164C double tagged and double RNAi construct cells lines generated a similar result to the induced Cep164C cell line. Analysis showed that cells with a single flagellum had a mixed variation of smaller (figure 83C, 84C and 85C) and larger cells (figure 83D, 84D and 85D) in comparison to the uninduced cells (figure 83A, 84A and 85A), suggesting an asymmetrical cell division generating daughters of different sizes.

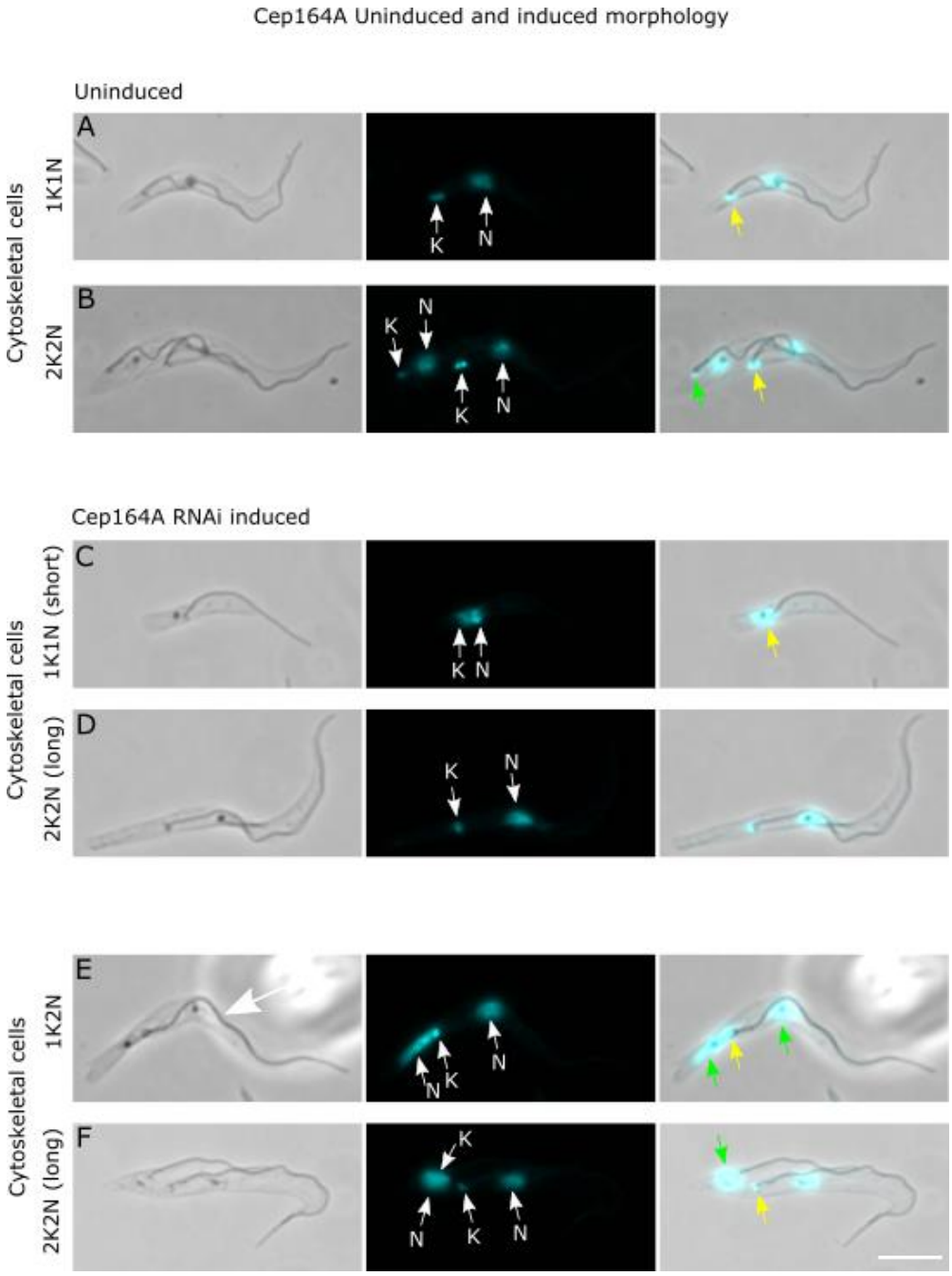


Figure 83: Knockdown of Cep164A shows organelle displacement.

The cell line containing the Cep164A RNAi construct was induced for 24 hours, collected on SuperFrost slides and labelled with DAPI. Induction of the RNAi construct lead to morphology changes in G1 (C-D) and dividing (E-F) cells, in comparison to the uninduced cells (A-B). Cells show asymmetrical division leading to shorter (C) or longer (F) flagella, with some cells lacking a new flagellum (E). Cells also generate mis-placed kinetoplasts (yellow arrow) and nuclei (green arrow). Scale bar = 10μm.

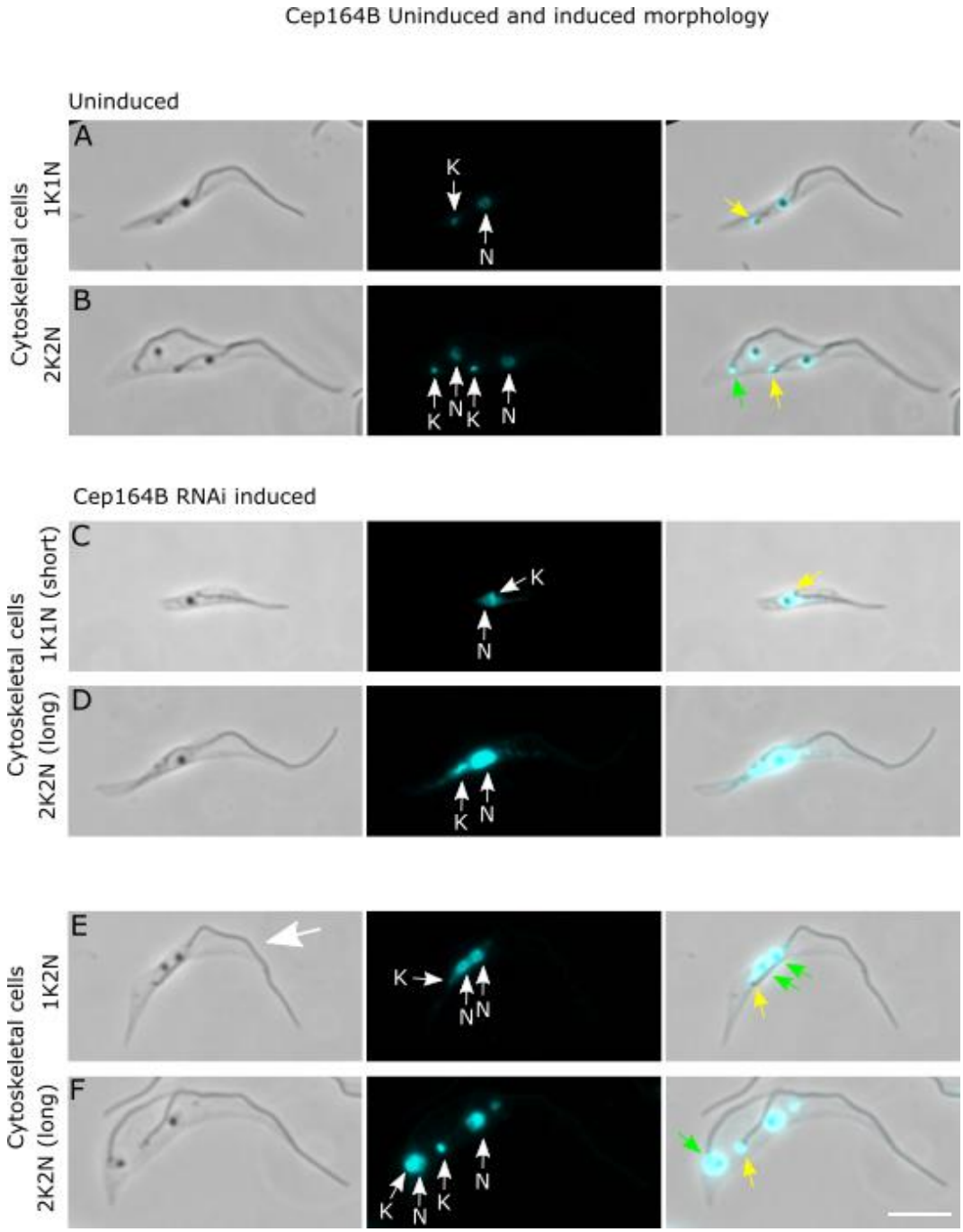


Figure 84: Knockdown of Cep164B shows organelle displacement.

The cell line containing the Cep164B RNAi construct was induced for 24 hours, collected on SuperFrost slides and labelled with DAPI. Induction of the RNAi construct lead to morphology changes in G1 (C-D) and dividing (E-F) cells, in comparison to the uninduced cells (A-B). Cells show an asymmetrical division, leading to shorter (C) or longer (D) flagella, with some cells lacking a new flagellum (E). Cells also generate mis-placed kinetoplasts (yellow arrow) and nuclei (green arrow). Scale bar = 10μm.

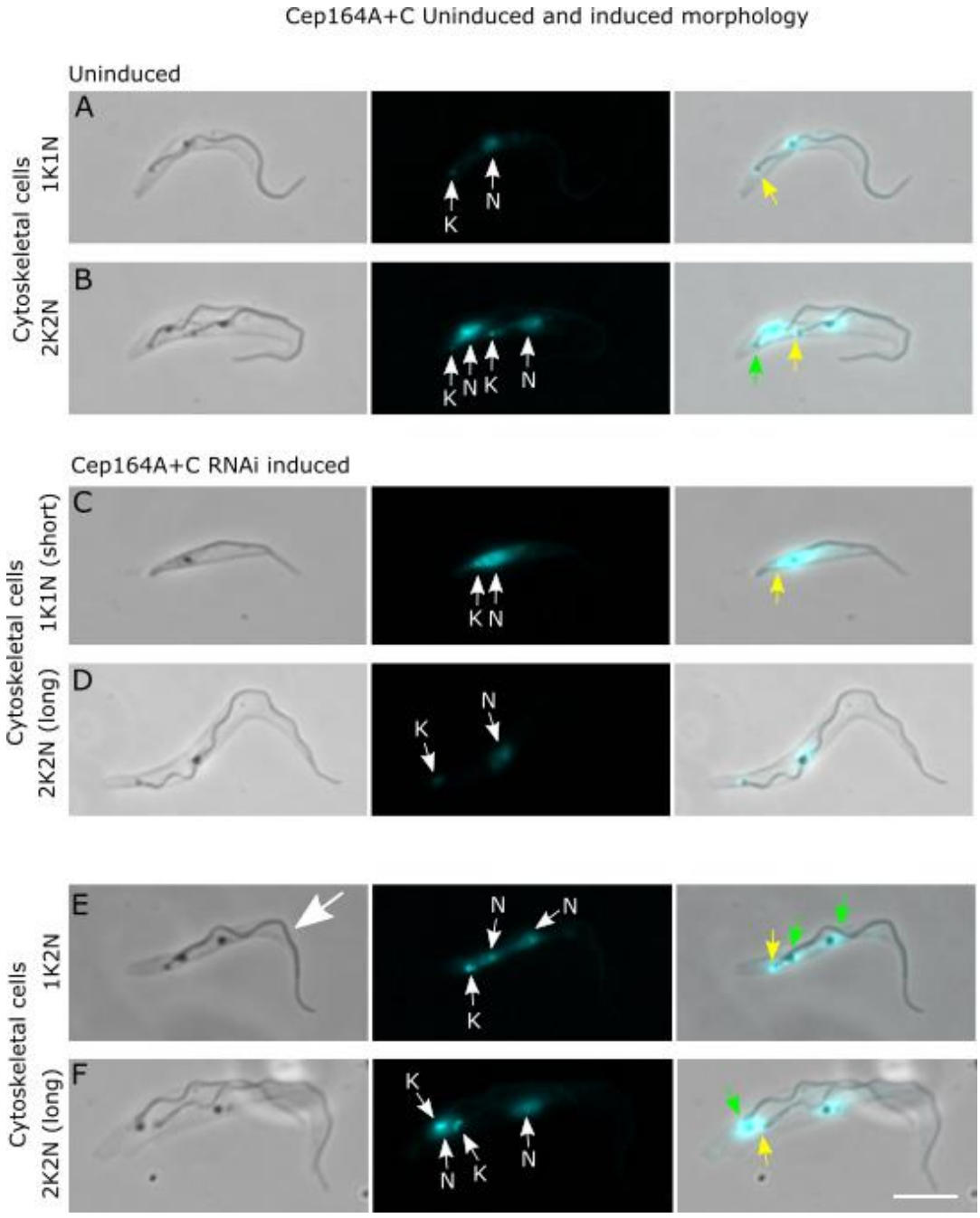


Figure 85: Knockdown of Cep164A+C shows organelle displacement.

The cell line containing the Cep164A+C RNAi construct was induced for 24 hours, collected on SuperFrost slides and labelled with DAPI. Induction of the RNAi construct lead to morphology changes in G1 (C-D) and dividing (E-F) cells, in comparison to the uninduced cells (A-B). Cells show an asymmetrical division, leading to shorter (C) or longer (D) flagella, and show a mixture of cells with Cep164A/B phenotypes (1K2N cells (E)) and Cep164C phenotypes (longer flagella). Cells also generate mis-placed kinetoplasts (yellow arrow) and nuclei (green arrow). Scale bar = 10μm.

In cells containing two flagella (2K2N), the two daughter cells appeared asymmetric in cell length (figure 83F, 84F and 85F) in comparison to the uninduced (figure 83B, 84B and 85B), with some dividing cells appearing larger than the uninduced (compare figure 84F to 84B and 85F to 85B). Cells containing the unusual 1 flagellum (figure 83E, 84E and 85E; white arrow), 1 kinetoplast and 2 nuclei (1K2N) morphology (figure 83E, 84E and 85E; yellow and green arrows respectively) were also observed in each cell line, suggesting a cell division issue, the BBs did not duplicate/segregate correctly or the axoneme did not develop a new flagellum.

In the induced cell lines, there was also evidence that organelles were in an incorrect position. In those cells smaller than the uninduced cell line, the kinetoplasts were more anterior than expected and were sometimes seen anterior to the nuclei (figure 83C, 84C and 85C; yellow arrow) in comparison to the uninduced cells (figure 83A, 84A and 85A; yellow arrow). This morphological abnormality was also seen in dividing cells, with kinetoplasts closer together post-mitosis (figure 83F, 84F and 85F; green and yellow arrows) in comparison to the uninduced post-mitotic cells (figure 83B, 84B and 85B; green and yellow arrows).

In order to quantify the organelle positioning abnormalities, analysis of the three induced cell lines was carried out. Measurements between the kinetoplast and nucleus, or kinetoplast and kinetoplast in induced and uninduced G1 and dividing cells was carried out (N=100 cells). Fluorescence microscopy indicated that 24 hour and 96 hour induction generated a population (1K1N) containing a mixture of cells that had either a closer or further away nucleus in comparison to the uninduced cells (compare figure 83C and D with 83A). Statistical analysis of this data showed a mixture of significant and non-significant results for each cell line.

Analysis of the RNAi induced Cep164A cell line

In the 1K1N cell population, results showed there was a statistically significant shorter distance between the kinetoplast and the nucleus after 96 hour induction in comparison to the uninduced cell line (p-value <0.0001). However, a statistically non-significant result was observed between the 24 hour and the uninduced cell line (p-value = 0.440) (figure 86A). In the dividing population, measurements between the new daughter kinetoplast and nucleus was measured, showing that there was a statistically significant decrease in distance between the 96 hour and uninduced (p-value <0.0001) cells, but not for the 24 hour cells (p-value = 0.087) (figure 86B). Measurements between the old daughter kinetoplast and nucleus showed no statistically significant results (p-value = 0.148 and 0.272 respectively) (figure 86C). However, a statistically significant decrease in distance was observed for the kinetoplast to kinetoplast measurements between the 24 hour and uninduced (p-value <0.0001) and the 96 hour and uninduced (p-value <0.0001) cells (figure 86D).

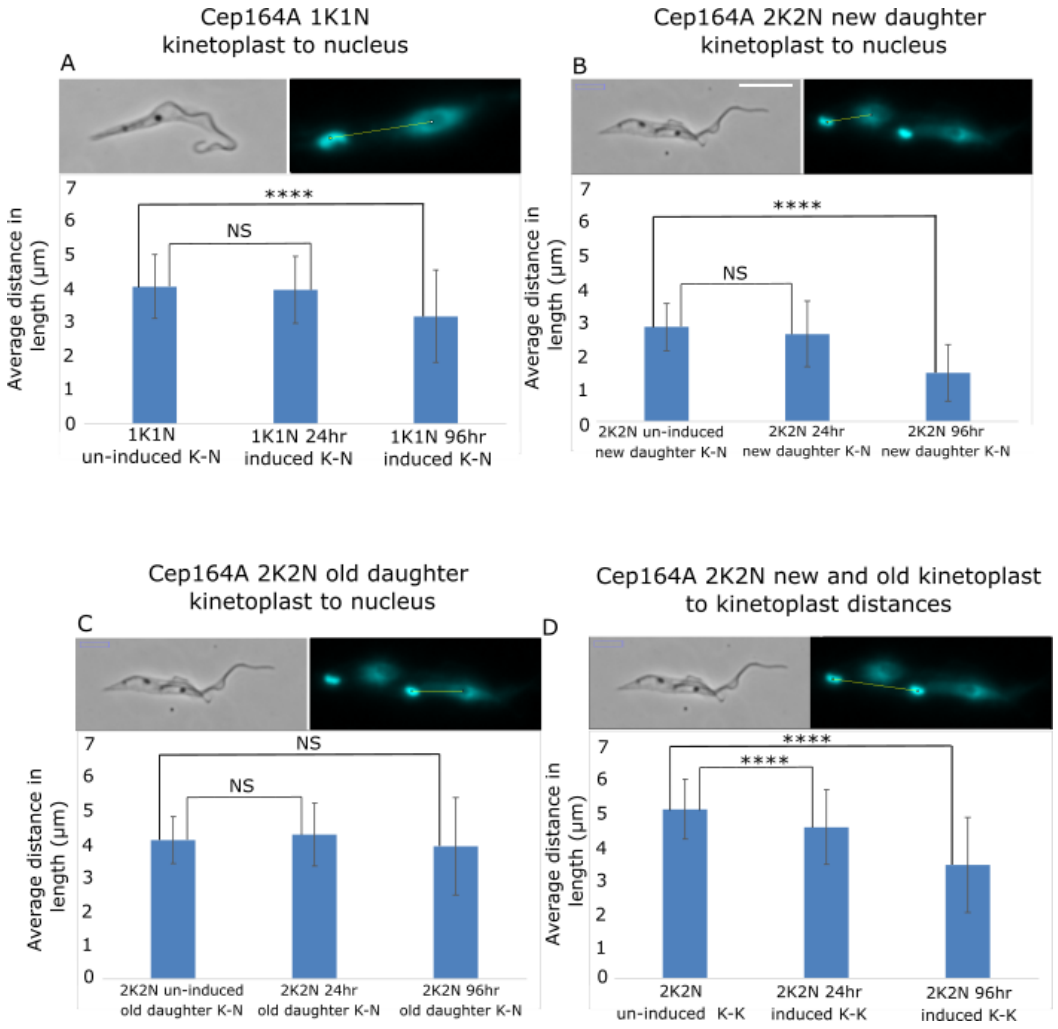


Figure 86: Induction of Cep164A shows misplacement of organelles.

Cells containing the Cep164A RNAi construct were collected from either the uninduced or induced cell lines. 100 cells were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope. Kinetoplasts/nuclei were measured in the ImageJ software using the segmented tool. Statistically significant changes in distances between the kinetoplast and nuclei were observed for G1 cells (A) and dividing cells (B-D). Measurements between the new kinetoplast and new nucleus (B), old kinetoplast and old nucleus (C) and between the old and new kinetoplast (D) were collected. K= Kinetoplast, N= Nuclei. Independent t-test was used. NS = not significant, **** = $p \leq 0.0001$ Scale bar = 10µm.

Analysis of the RNAi induced Cep164B cell line

Measurements in the G1 population between the kinetoplast and nucleus for the induced RNAi Cep164B cell line showed a statistically significant decrease in distance between the 24 hour and uninduced (p-value <0.0001) and 96 hour and uninduced (p-value <0.0001) cells (figure 87A). A statistically significant decrease in distance was also seen between the new daughter kinetoplast and nucleus in dividing cells (2K2N) for both induced time points at p-value <0.0001 and p-value <0.0001 respectively (figure 87B). A similar statistically significant decrease was observed in the measurements in the induced and uninduced cell line between the kinetoplast and nucleus of the old daughter cell at p-value <0.0001 and p-value <0.0001 respectively (figure 87C). Measurements between the kinetoplast and kinetoplast in dividing cells also showed there was a statistically significant decrease in distance, with both cells lines showing a result of p-value <0.0001 (figure 87D).

Analysis of the RNAi induced Cep164A+C cell line

Analysis was carried out for the inducible RNAi A+C cell line. Statistical analysis showed no significant results between the measurements of the kinetoplast and nucleus for the induced and uninduced cells (p-value = 0.155 and p-value = 0.779 respectively) (figure 88A). Measurements between the new daughter kinetoplast and nucleus showed there was a statistically significant decrease in distance between the uninduced and induced 24 hour cells (p-value <0.0001), but no statistically significant result for the 96 hour and uninduced cells (p-value = 0.085) (figure 88B). Results between the old daughter kinetoplast and nucleus in dividing cells (figure 88C), and between the kinetoplast and kinetoplast in dividing cells (figure 88D) all showed a statistically significant result between both induced time points and the uninduced cell line measurements (p-value <0.0001).

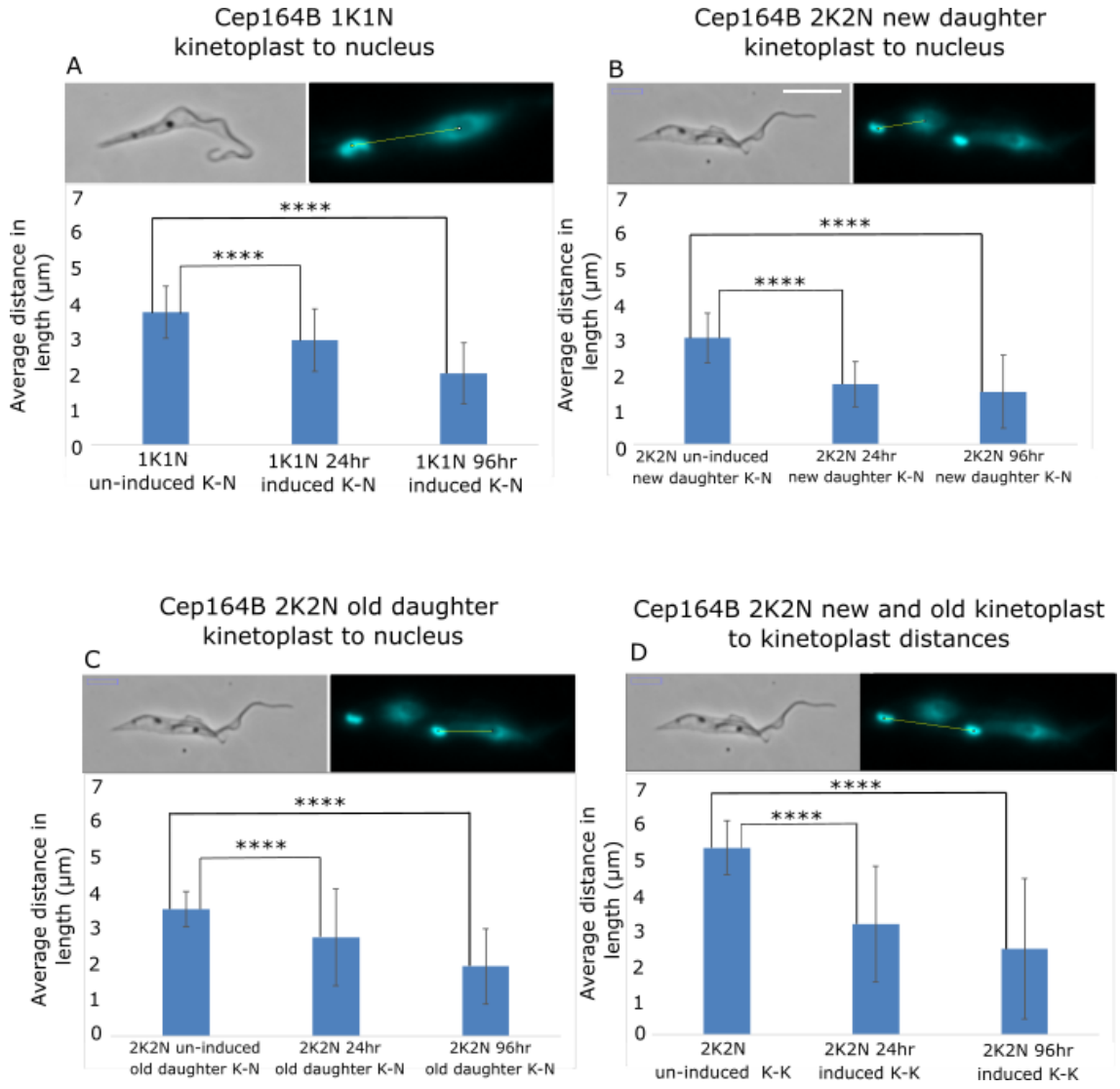


Figure 87: Induction of Cep164B shows misplacement of organelles. Scale bar = 10µm.

Cells containing the Cep164B RNAi construct were collected from either the uninduced or induced cell lines. 100 cells were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope. Kinetoplasts/nuclei were measured in the ImageJ software using the segmented tool. Statistically significant changes were observed between all measurements. Measurements of the kinetoplast and nucleus in G1 cells (A), new kinetoplast and new nucleus (B), old kinetoplast and old nucleus (C) and between the old and new kinetoplast (D) were collected. K= Kinetoplast, N= Nuclei. Independent t-tests were used. **** = $p \leq 0.0001$. Scale bar = 10µm.

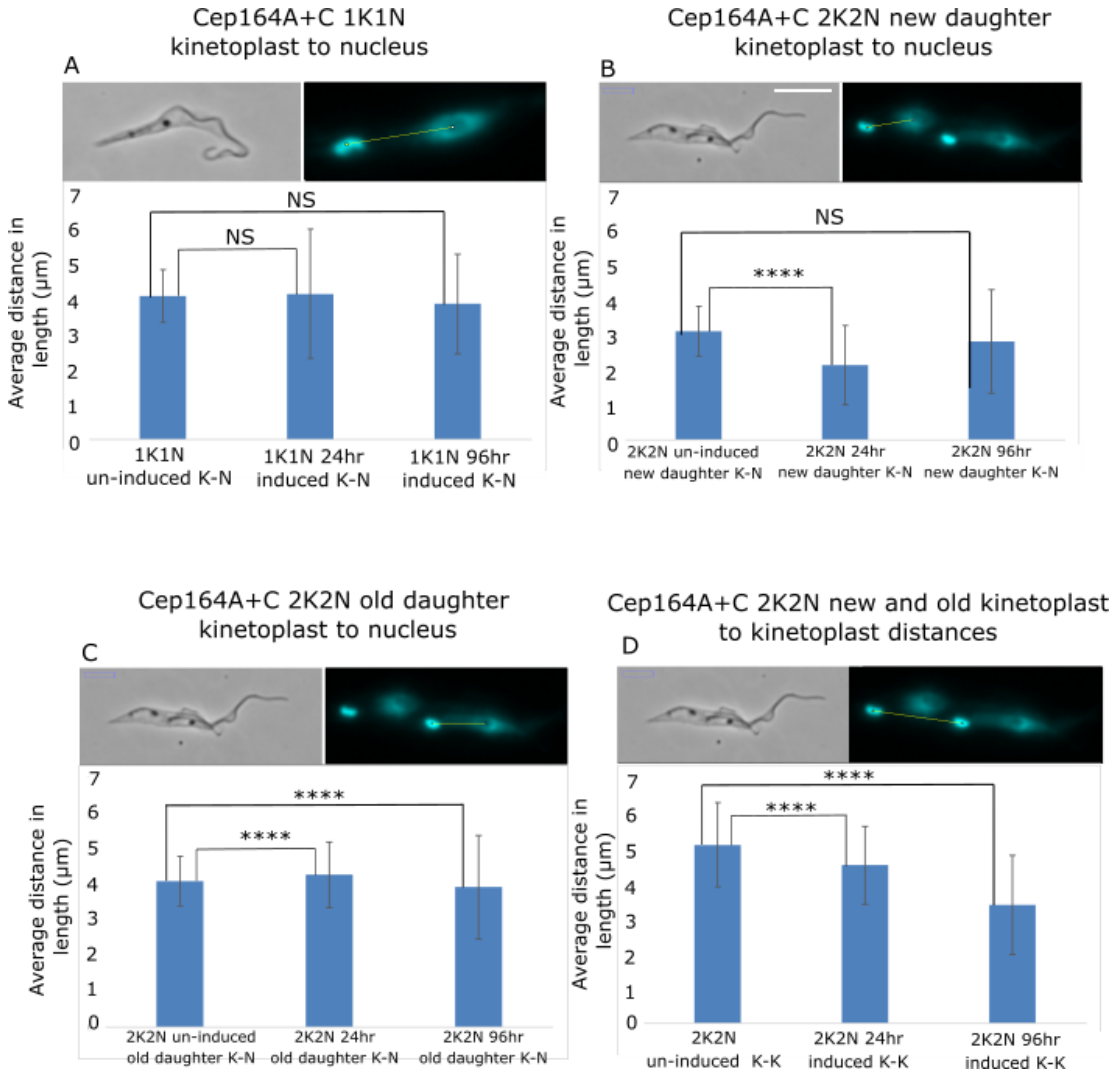


Figure 88: Induction of Cep164A+C shows misplacement of organelles.

Cells containing the Cep164A+C RNAi construct were collected from either the uninduced or induced cell lines. 100 cells were collected on SuperFrost microscope slides, labelled with DAPI and imaged on the fluorescent microscope. Kinetoplasts/nuclei were measured in the ImageJ software using the segmented tool. No statistically significant changes to the distance of the kinetoplast and nucleus were observed for G1 cells (A). Significant changes to the distance between the kinetoplast and nucleus was observed for the new kinetoplast and new nucleus (B), old kinetoplast and old nucleus (C) and between the old and new kinetoplast (D). K= Kinetoplast, N= Nuclei. Independent t-tests used. NS = not significant, **** = $p \leq 0.0001$. Scale bar = 10µm.

In summary, the morphological analysis showed defects in the correct placements of both the kinetoplast and nucleus in the Cep164A, Cep164B and Cep164A+C inducible cell lines resulting from an asymmetric division of daughter cells.

5.10 Knockdown of Cep164A or B by inducible RNAi generates shorter new flagella and dividing cells with no new flagellum growth

In previous experiments it was indicated that induction of the Cep164A or Cep164B cell line generated daughter cells of different sizes, and the asymmetrical division indicates that the length of the flagellum could also be affected. The methodology as described in section 4.11 was carried out for the cell lines containing the RNAi construct for Cep164A and for Cep164B. Cells were collected at uninduced, 24 hour and 96 hour time points, with the flagellum measurements analysed through the FIJI (Image J) software.

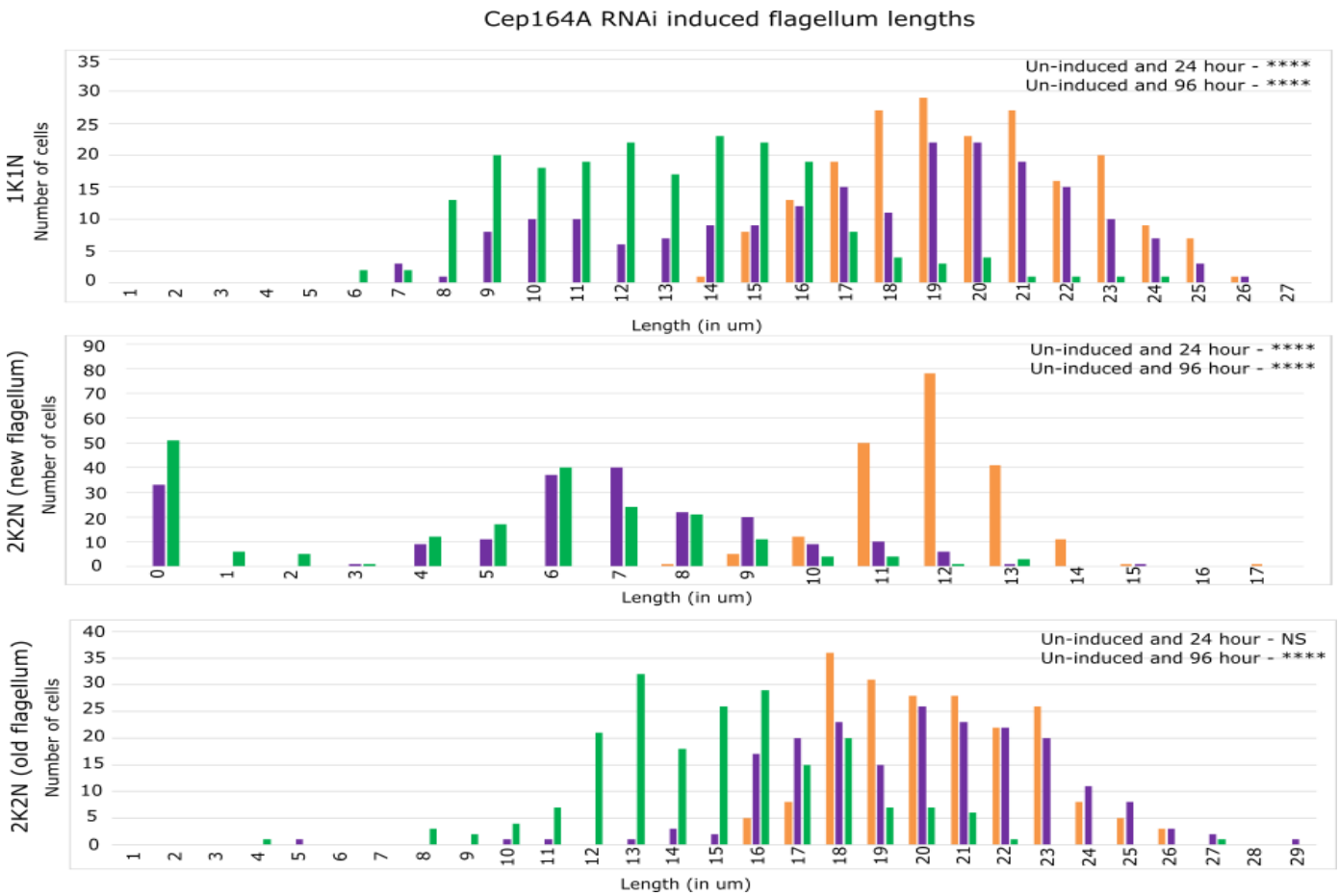
Flagellum length measurements for the RNAi Cep164A cell line

For the uninduced G1 (1K1N) population of cells with a single flagellum, one kinetoplast and one nucleus (figure 91A), the flagellum lengths ranged from 14μm-26μm in length (figure 89A; orange). The 24 hour induced cells showed much shorter flagellum lengths starting from 7μm (figure 91C), with the longer flagellum length at 26μm (figure 91D) (figure 89A; purple), similar to the uninduced. The 96 hour induced cells had similar flagellum lengths to the 24 hour induced cells, ranging between 6μm-24μm (figure 89A; green). The average between the uninduced, the 24 hour and 96 hour induced G1 (1K1N) cells was 19.3μm, 16.8μm and 13.4μm respectively, with both induced time points showing there was a statistically significant difference at p-value <0.0001 (figure 89A). These results indicate that the induced cells are generated with a much shorter flagellum in comparison to the uninduced cell line, but not extended flagella as observed in the Cep164C cell line.

Measurements of the old and new flagellum were also carried out in the dividing cells, with the new flagellum lengths showing an interesting difference in comparison to the new flagellum lengths of the uninduced cell line. The flagellum

length in the uninduced cell line ranged from 8µm-17µm (figure 89B; orange) (figure 91B; NF). In comparison, the flagellum length in the 24 hour and 96 hour ranged between 0µm-13µm (figure 91E; NF), with 33 and 50 cells respectively having no new flagellum growth (figure 89B; purple and 89B; green). The average length of the new flagellum in the uninduced cell line was 12.4µm. In comparison, the 24 hour and the 96 hour flagellum length averages were 6.7µm and 5.2µm respectively, there was a statistically significant difference between these two values at p-value <0.0001 (figure 89B). These results suggest that the induced cell line generates daughter cells that undergo mitosis, but do not generate a new flagellum.

Finally, measurements of the old flagellum in the dividing cells were measured for the Cep164A RNAi cell line. The uninduced old flagellum lengths ranged from 16µm-26µm (figure 91B; OF), with an average of 19.6µm (figure 89C; orange). The flagellum length measurements for the 24 hour induced cell line ranged between 5µm-29µm (figure 91E; OF), but had a similar average of 19.5µm, with a statistically non-significant result at p-value = 0.144 (figure 89C; purple). The flagellum length measurements for the 96 hour cell line ranged from 4µm-22µm (figure 89C; green), with an average at 14.5µm, results showed there was a statistically significant difference at p-value <0.0001 (figure 89C). Overall, the results from this cell line suggest that lack of the Cep164A protein generates cells which have a shorter flagellum or fail to produce a new flagellum.



Orange: Un-induced, Purple: 24-hour induced, Green: 96-hour induced

Figure 89: Knockdown of Cep164A produces 1K2N cells

200 induced and uninduced Cep164A cells were labelled with DAPI and the L8C4 antibody and flagellum lengths were measured. Measurements showed shorter flagella in G1 cells and in new flagella of dividing cells, with some not generating a new flagellum. Independent t-tests were used. NS = not significant, **** = $p \leq 0.0001$

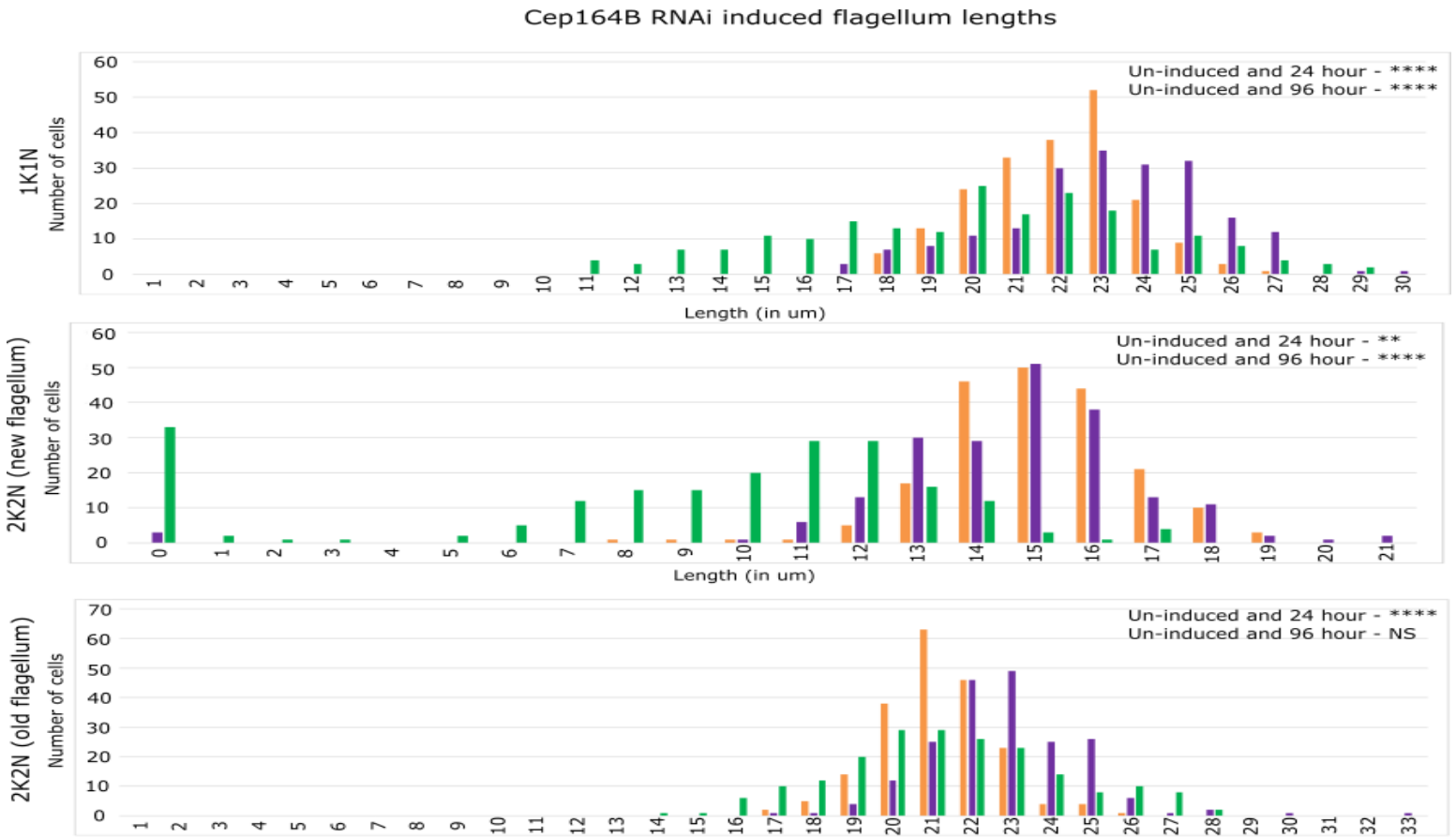
Flagellum length measurements for the RNAi Cep164B cell line

After the flagellum lengths were measured for the Cep164A cell line, the induced and uninduced Cep164B RNAi cell line was investigated. In the uninduced G1 population, the flagellum lengths ranged from 18μm-27μm (figure 90A; orange) (figure 91F). In comparison to the uninduced, the 24 hour flagellum lengths ranged from 17μm-30μm (figure 90A; purple) (figure 91H-I), and the 96 hour flagellum lengths ranged from 11μm-29μm (figure 90A; green). The average length of the uninduced G1 flagellum was 21.4μm, in comparison to the 24 hour and 96 hour having a similar average at 22.7μm and 19.4μm respectively, with both showing a statistically significant difference at p-value <0.0001 (figure 90A). Similar to the Cep164A RNAi cell line, these results show that the G1 daughter cells produce shorter flagellum lengths in comparison to the uninduced cell line, with only a small number of cells producing a flagellum length up to 30μm.

Measurements for the new and old flagellum was also carried out in dividing cells. For the uninduced cell line, flagellum lengths ranged between 8μm-19μm (figure 90B; orange) (figure 91G; NF). In comparison, the 24 hour flagellum lengths ranged from 0μm-21μm (figure 90B; purple) (figure 91J; NF) and the 96 hour flagellum lengths ranged from 0μm-17μm (figure 90B; green). The average length for the uninduced new flagellum was 15.6μm, with the 24 hour average new flagellum length at 15.1μm. There was a statistically significant result between these two values at p-value <0.05. The 96 hour induced cell line had an average new flagellum length of 9.2μm and a statistically significant difference at p-value <0.0001 (figure 90B) was seen. The results showed that induced daughter cells were produced with no new flagellum growth, which was not observed for the uninduced cell line.

Finally, measurements for the old flagellum in dividing cells was carried out. Results showed that the uninduced, 24 hour and 96 hour produced similar flagellum lengths. The uninduced old flagellum lengths ranged between 17μm-26μm (figure 90C; orange) (figure 91G; OF), with an average flagellum length of 20.9μm. In comparison the 24 hour old flagellum lengths ranged from 19μm-33μm (figure 90C; purple) (figure 91J; OF), with an average of 22.3μm and a statistically

significant result at p-value <0.0001 (figure 90C). The 96 hour old flagellum lengths however had a range from 14µm-28µm (figure 90C; green) and no statistically significant result at p-value = 0.448 (figure 90C).



Orange: Un-induced, Purple: 24-hour induced, Green: 96-hour induced

Figure 90: Knockdown of Cep164B produces 1K2N cells

200 induced and uninduced Cep164B cells were labelled with the DAPI and L8C4 antibody and flagellum lengths were measured. Measurements showed shorter flagella in G1 cells and in new flagella of dividing cells, with some not generating a new flagellum. Independent t-test were used. NS = not significant, * = $p \leq 0.05$ and **** = $p \leq 0.0001$

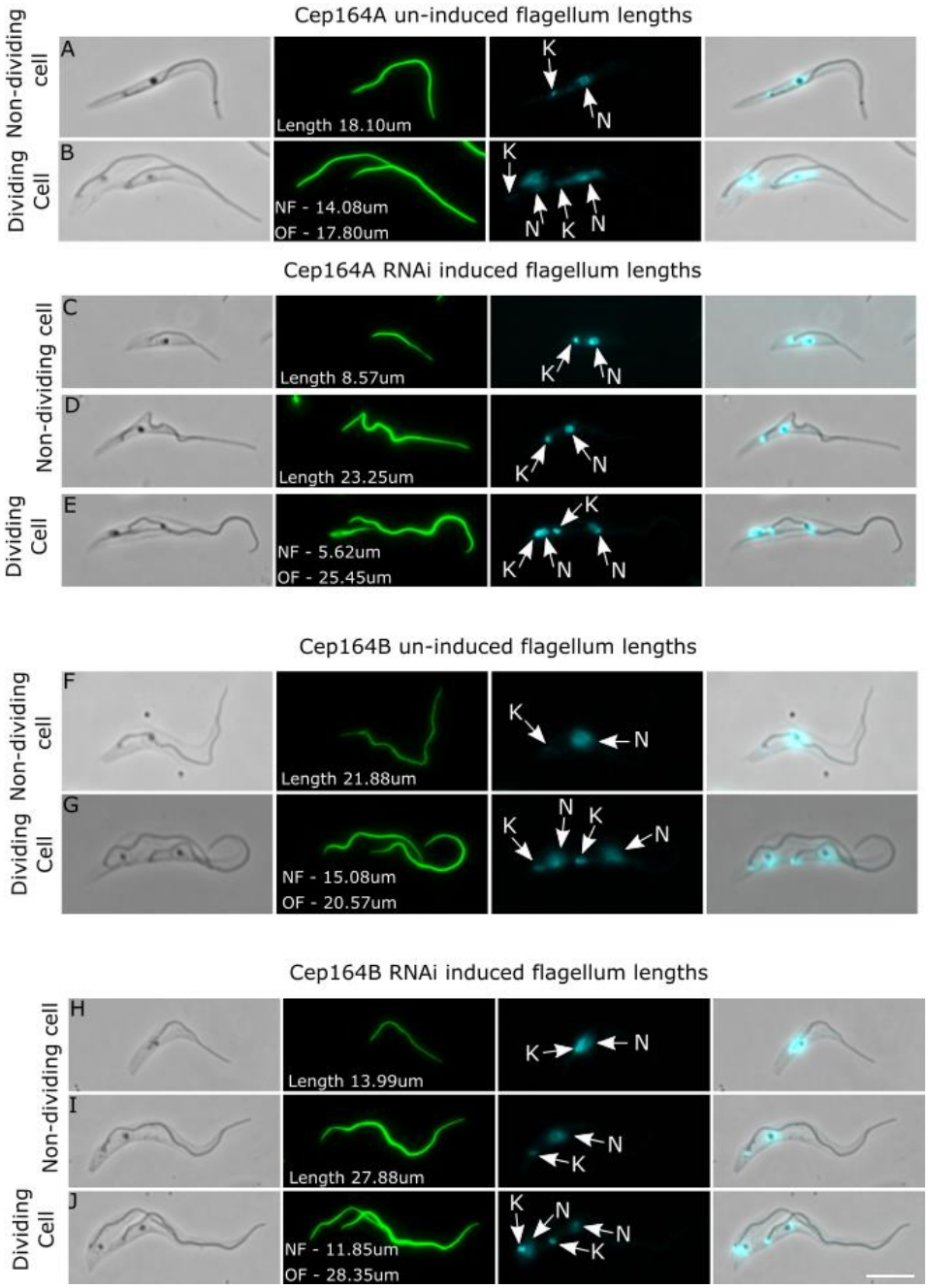


Figure 91: Induced Cep164A and Cep164B show shorter new flagella.

Induced cells containing the Cep164A (A-E) and the Cep164B (F-J) RNAi construct were induced for 24 hours, collected on SuperFrost microscope slides and labelled with the PFR (flagellum) antibody and DAPI. Measurements were carried out in ImageJ using the segmented tool. Induced cell lines had a wider variation in flagellum length, with cells showing shorter flagellum lengths (C and H) in comparison to the uninduced Cep164A (A-B) and Cep164B (F-G) cell lines. O – Old flagellum, N – New flagellum. Scale bar – 5µm.

These results show that on average the majority of daughter cells are produced with a shorter new flagellum and with some not generating a new flagellum overall. Whereas fewer cells are generating flagellum lengths exceeding 26-27μm and shows a different phenotype to the Cep164C induced cell line.

5.11 Induction of Cep164A and B generates dividing cells with one flagellum

During cell division, cells develop from having one BB and one flagellum, to containing two BBPs, two kinetoplasts and two flagella after mitosis. Investigating the phenotype of the induced Cep164A and Cep164B RNAi cell lines resulted in the observation of dividing cells which either contained a shorter flagellum then expected (figure 92B; NF), or aberrant post-mitotic cells containing only a single flagellum (1K1N). 142 cells of the Cep164A cell line were labelled with DAPI, the flagellum antibody L8C4 and the basal body antibody BBA4.

These cells were analysed to understand whether the new flagellum had not grown due to the lack of a new BB (BB duplication failure), or the BB was present, but no new flagellum had nucleated. Analysis showed that 42% of cells had successfully duplicated and segregated their BBs (figure 92A; BB1 and BB2) and had developed a new flagellum (figure 92A; NF). 23% of cells had successfully duplicated and segregated their BBs (figure 92B; BB1 and BB2) but had generated a shorter new flagellum than expected (figure 92B; NF). A lower number of 14% lacked BB duplication (figure 92C; BB1) and did not generate a new flagellum, with only the old flagellum present (figure 92C; OF). Fewer than 10% of cells were seen to successfully duplicate their BBs, but were unable to segregate the organelles correctly, with the BBs seen close together in post-mitotic cells (figure 92D; BB1 and BB2). These cells were able to generate a short new flagellum (figure 92D; NF).

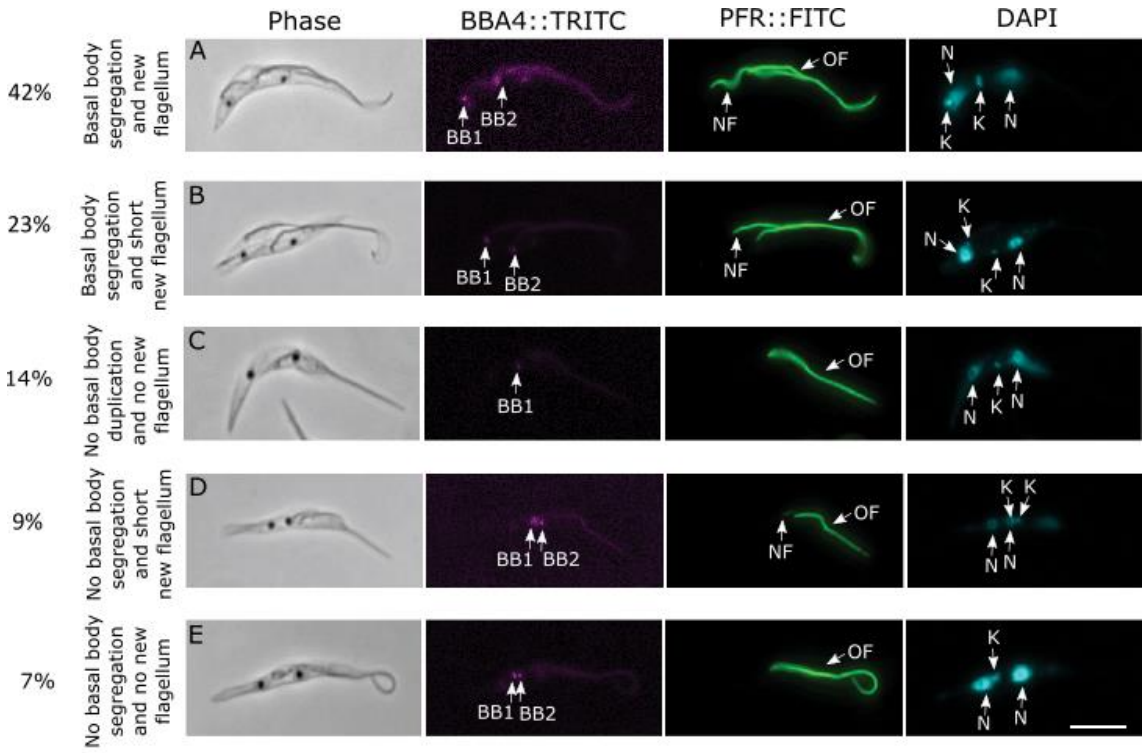


Figure 92: A percentage of Cep164A cells do not divide their basal bodies.

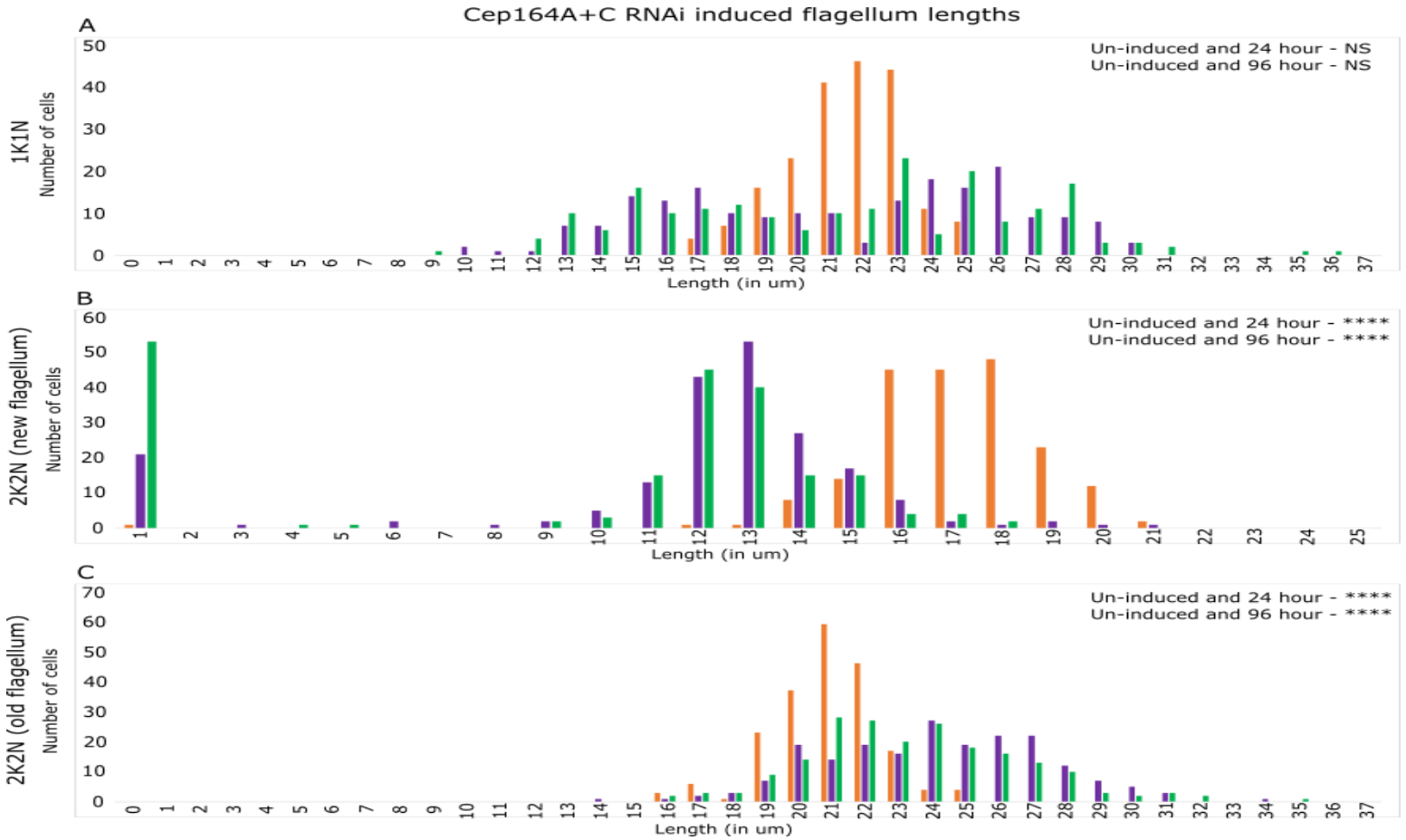
142 cells containing the mNG::Cep164A RNAi construct were analysed after 24 hours of induction. These cells were collected on SuperFrost slides, labelled with the PFR (flagellum) antibody, the BBA4 (basal body) antibody, the DNA marker DAPI and analysed. Induced cells showed a wide range of basal body/flagellum nucleation issues. With many generating shorter flagella than expected (A, B and D), or generate no new flagellum at all (C and E). Basal bodies had either duplicated and segregated (A and B), or failed to segregate (D, E) or divide (C). Scale bar - 5μm.

Finally, 7% of cells had successfully duplicated their BB, but failed to segregate these organelles (figure 92E; BB1 and BB2) and failed to generate a new flagellum (figure 92; OF). No cells in the uninduced Cep164A cell line lacked a new flagellum. In summary, a variety of phenotypes were observed in the induced cell line, with some failing to duplicate or segregate BBs, or generate a new flagellum. It is possible that the lack of this Cep164A causes complications in regulating BB duplication and segregation and could possibly be involved in BB docking, which leads to some cells failing to generate a new flagellum.

5.12 Induction of the Cep164A+C population generates cells with shorter and longer flagella and cells which do not build a new flagellum.

In previous experiments, induction of the Cep164C cell line showed a phenotype of extended flagellum lengths (section 4), while induction of the Cep164A and Cep164B showed a mixture of shorter flagellum lengths with a proportion of cells failing to grow a new flagellum. It was questioned whether generating an inducible RNAi cell line against both Cep164A and Cep164C proteins would produce daughter cells showing a mixture of both phenotypes.

For the uninduced G1 population of cells, the flagellum ranged from 17-25μm (figure 93A; orange). The 24 hour and 96 hour induced G1 cells ranged from 10-30μm (figure 93A; purple) and 9-31μm (figure 93A; green) respectively. Examples of induced cells containing a short flagellum are seen in figure 94(C), and those with a longer flagellum seen in figure 94(D). Induced cells show a wider range of shorter and longer flagellum lengths, while also showing much smaller flagellum lengths than the uninduced cell line. The average between the uninduced, 24 hour and 96 hour was 21.3μm and 21.7μm and 21.7μm respectively with both induced time points showing a statistically non-significant difference (p-value = 0.432 and p-value = 0.414) (figure 93A). These results indicate that there is a wide range of flagellum lengths in the induced cell line (ranging from 9-31μm), showing that cells are generating both shorter and longer flagellum lengths in comparison to the uninduced cell line.



Orange: Un-induced, Purple: 24-hour induced, Green: 96-hour induced

Figure 93: Knockdown of Cep164A+C produce 1K2N and longer old flagellum lengths

200 induced and uninduced Cep164A+C cells were labelled with the DAPI and L8C4 antibody and flagellum lengths were measured. Measurements showed a mixture of shorter flagella in G1 and dividing cells, as well as longer flagellum lengths seen in the Cep164C induced cell line. Independent t-test were used. NS = not significant, **** = $p \leq 0.0001$

Measurements of the old and new flagellum were carried out in dividing cells. Firstly, the new flagellum lengths in the induced cell line were compared against the new in the uninduced cell line. The uninduced new flagellum measured between 12-21μm (figure 93B; orange) (figure 94; NF). As expected, induced cells were seen to contain dividing daughter cells that lack the development of a new flagellum (figure 93B; purple and green) (figure 94G) but also range from 0-21μm for 24 hour and 0-18μm for 96 hour (figure 94E and F; NF) with the majority being between 10-16μm. The average of the new flagellum length in the uninduced cell line was 16.5μm, in comparison the 24 hour new flagellum average was 11.3μm and was statistically significant at p-value <0.0001. The 96 hour however had a shorter length average than the 24 hour at 9μm and results showed there was a statistically significant difference at p-value <0.0001. These results suggest that the new flagellum in the uninduced cell line is longer on average than the induced cell line, and RNAi induction of Cep164A and Cep164C generates cells lacking a new flagellum.

Finally, measurements of the old flagellum, in the dividing cells was carried out. The uninduced cells ranged between 16-25μm in flagellum length (figure 93C; orange) (figure 94B; OF), with the 24 hour and 96 hour showing a much longer flagellum range between 16-34μm (figure 93C; purple) and 16-35μm (figure 93C; green) respectively (figure 94E, F and G; OF). The average between the uninduced and 24 hour was 21.3μm and 24.3μm, with a statistically significant difference at p-value <0.0001. The 96 hour induced cells had an average of 23.8μm and showed there was a statistically significant difference at p-value <0.0001. Overall, these results show that the induced Cep164A and Cep164C RNAi cell line generate dividing cells with shorter or no new flagellum growth (Cep164A phenotype), but can also generate cells with much longer flagellum lengths than the uninduced (up to 35μm) which is typically seen in the Cep164C phenotype.

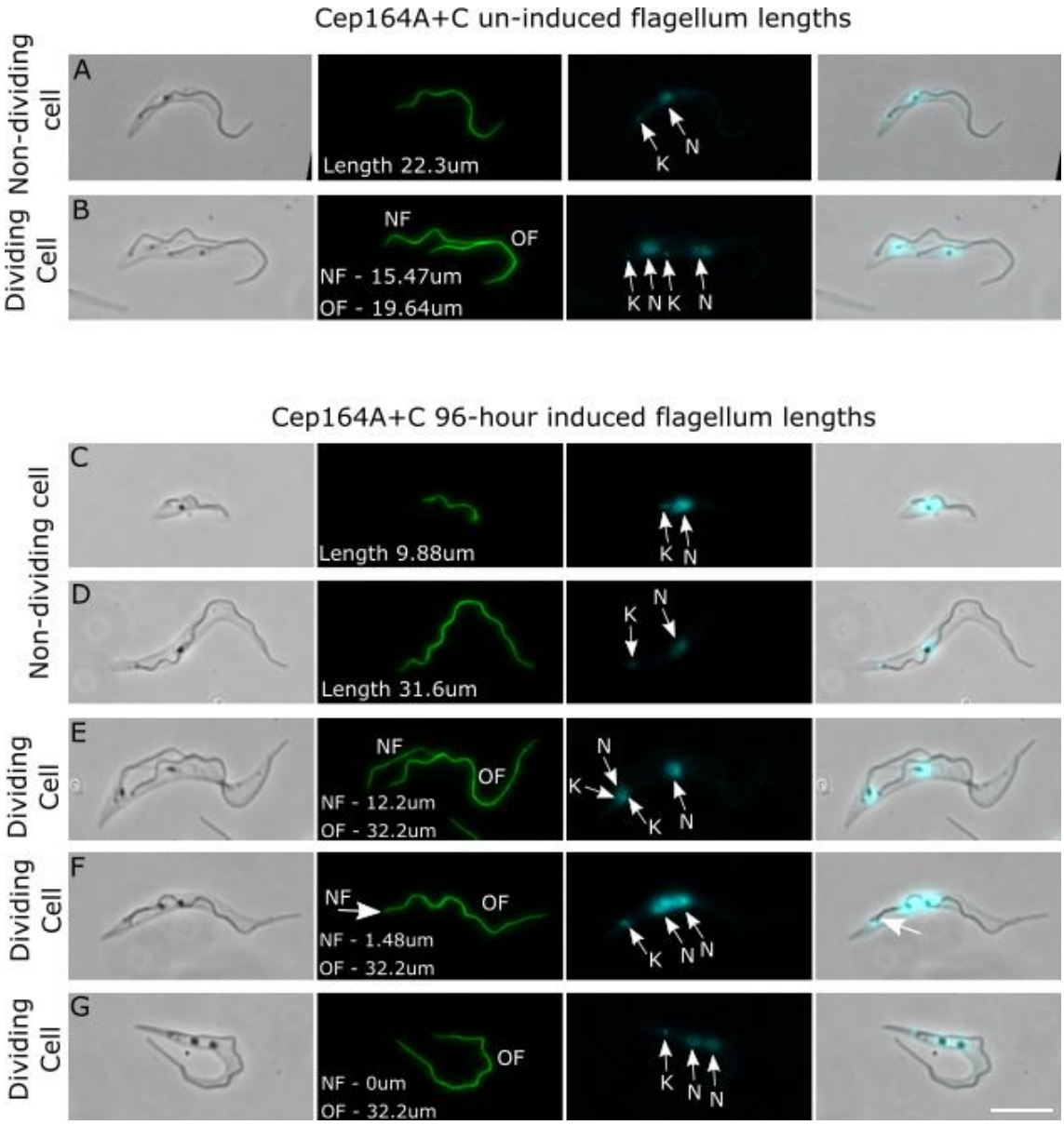


Figure 94: Induced Cep164A+C shows no new flagellum growth and longer old flagellum lengths.

Induced and uninduced cells containing the Cep164A+C RNAi construct were collected on SuperFrost microscope slides and labelled with the PFR (flagellum) antibody and DAPI. Cells were induced for 96 hours. Measurements were carried out in ImageJ using the segmented tool. Induced cell lines had a wider variation in flagellum lengths. With shorter new flagella (Cep164A phenotype) and longer old flagellum lengths (Cep164C phenotype) being observed in comparison to the uninduced (A-B). O – Old flagellum, N – New flagellum. Scale bar – 5μm.

5.13 The FRAP technique for Cep164A and B shows the proteins do not have a high turnover rate

One hypothesis for the Cep164A and Cep164B proteins is that they may help regulate flagellum length or have an impact on the development of the new flagellum. Similar to the Cep164C protein (section 4) it was questioned as to whether these proteins have a high turn-over rate and whether they are actively maintained at the TFs to continue growth of the new flagellum. The FRAP technique was employed with the TF region selected (as explained in section 4) for investigation into these proteins. Recordings of live cells was captured before bleaching the region of interest (figure 95A and 95G). Recording were carried out for 72 seconds for Cep164A and 60 seconds for Cep164B. Immediately after recording, two cells positive for mNG::Cep164A or mNG::Cep164B prior to bleaching was captured (figure 95B and 95H; yellow and white arrows). After 6 seconds for Cep164A and 8 seconds for Cep164B, bleaching on the region of interest was carried out on a targeted cell (figure 95C and 95I; yellow arrows) the remaining cell was not put under the FRAP technique and was still positive for the Cep164A/Cep164B protein (figure 95C and 95I; white arrow).

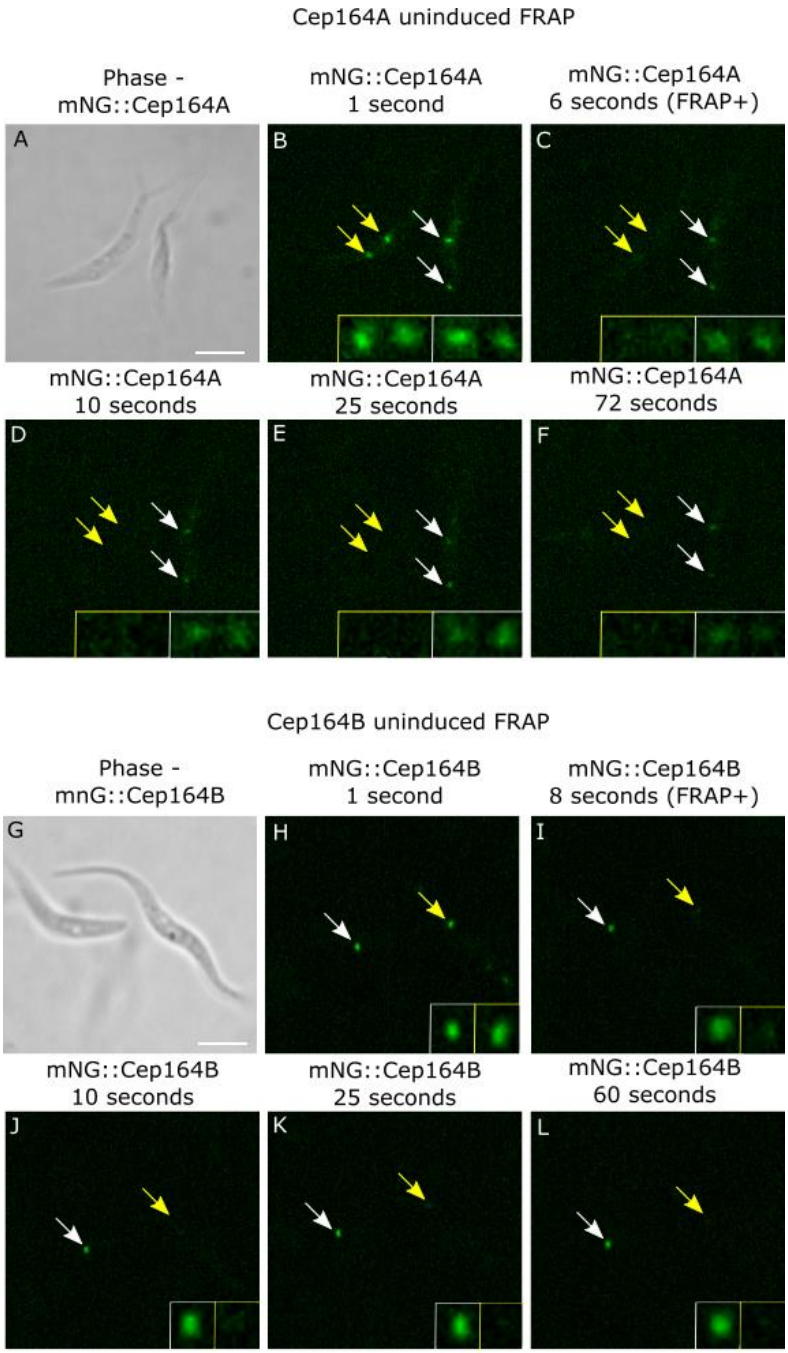


Figure 95: Cep164A and B do not have a high turn-over rate.

Live cells expressing either the mNG::Cep164A (A-F) or the Cep164B (G-L) protein were analysed using the FRAP technique. The mNG::Cep164A or mNG::Cep164B of one of the cells was bleached after a few seconds of recording and was not seen to recover throughout the remainder of the movie (B-F; yellow arrow and H-L; yellow arrow). A nearby cell was not bleached, with the mNG::Cep164A or mNG::Cep164B being continuously present throughout the recording (B-F; white arrow and H-L; white arrow). Scale bar = 10μm.

Throughout the remainder of the experiment both Cep164A and Cep164B proteins failed to return a positive expression of the mNG fluorophore (figure 95D-F and 95J-L; yellow arrows), which was not due to over bleaching of the sample as the non-targeted cell was still positive for the Cep164A/Cep164B protein (figure 95D-F and 95J-L; white arrows). These results suggest that Cep164A and Cep164B are not highly-turned over proteins and do not require this for their function.

Chapter 6. Discussion

The previous chapters have discussed in detail the results observed throughout this PhD study. Focusing on results from the screen and bio-informatic analyses for the 267 basal body proteins, and characterising the Cep164A, Cep164B and Cep164C proteins. This section will summarise the results of the previous three chapters and discuss how these results add to the knowledge of basal body/flagellum biology in both *T. brucei* and eukaryotic cells.

6.1 Basal body screen and bio-informatics of the basal body proteins

The screen identified proteins which were thought to localise to/near the SAS-6 protein. However fluorescent microscopy only determines the approximate localisation of a protein suggesting these proteins could localise to the BB, or to structures around the BB (Tahir *et al,* 2012). Proteins that localised to the PBB only suggests they could function in TF biogenesis but are not components in TF maintenance. Whereas proteins that localise to the distal section of the MBBs suggest they could function in maintenance or stability of the TFs, or involved in TZ building. Some proteins localised between the BBs, which could be the filament tethering the MBB and PBB together (figure 96A; yellow arrow) or could remove the tether in later stages of the cell cycle (figure 96B; yellow arrow) (Loncarek and Bettencourt-Dias, 2018).

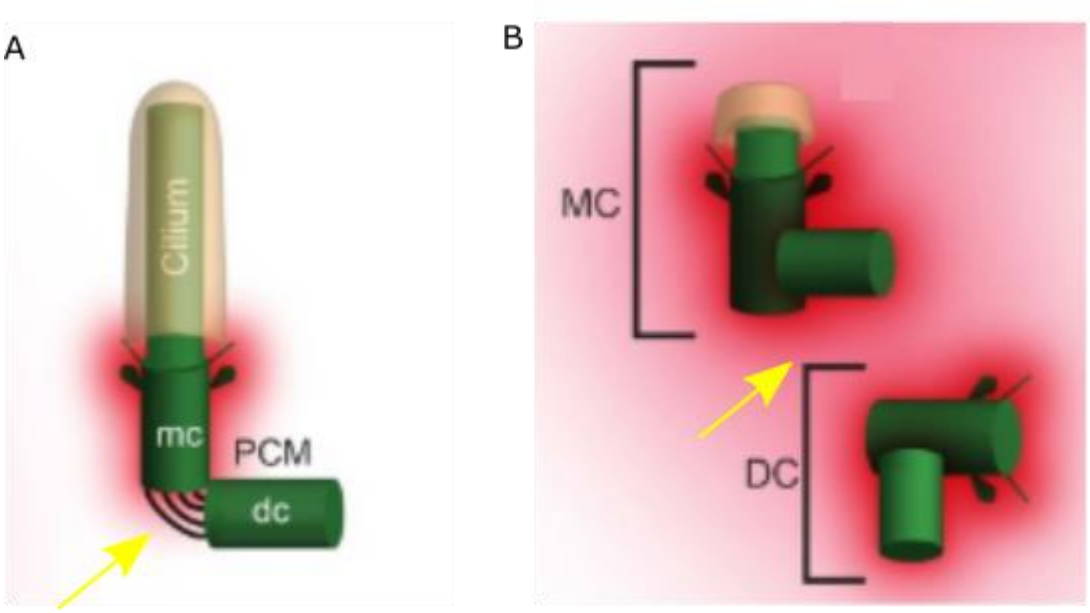


Figure 96: Presence and absence of the tether structure

A) The tether connects the mother and daughter centriole and is present prior to disassembly (yellow arrow) B) The tether is removed after development of the pro-centrioles to allow separation of the two centriole pairs (yellow arrow). Edited from Avidor-Reiss *et al* (2015).

Basal body proteins showed either a 1-2 or a 2-4 localisation pattern, with 2-4 foci being the most frequent (figure 97; D-F). As the pattern changes from 2 to 4 foci, the proteins then localise to the MBB and PBB on both BBPs in dividing cells, potentially recruited for BB duplication or development. A lower frequent of proteins showed a 1-2 localisation pattern (figure 97; A-C), showing only one BB in the BBP recruits this protein. These proteins are likely required for structure-associated functions with those on the MBB being related to stability/docking or flagellum growth, whereas those on the PBB could function in development of the TF/TZ structures.

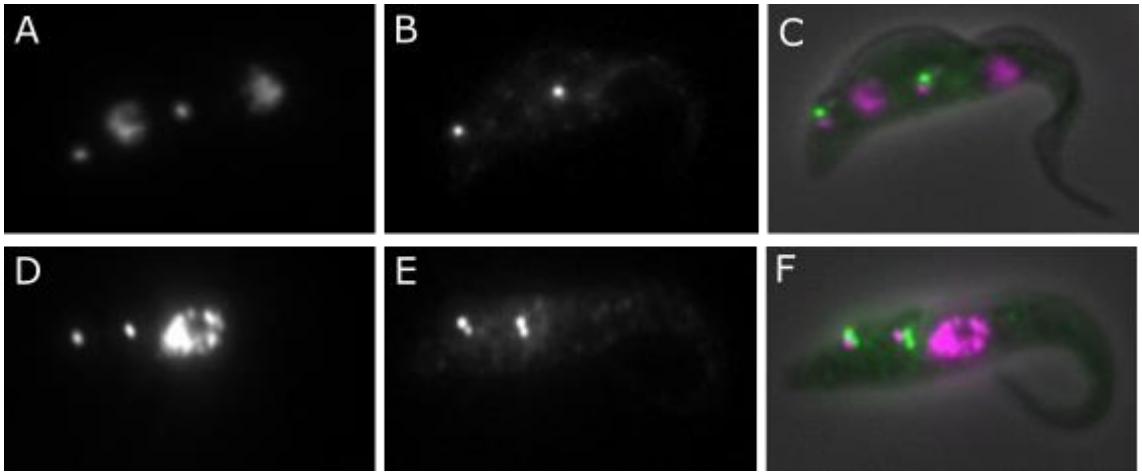


Figure 97: Localisation of proteins with a 1-2 or 2-4 foci.

A dividing *Trypanosoma brucei* cell expressing a basal body protein localising to either the mature or pro-basal body on each basal body pair (A-C) with a 1-2 foci, or to both the mature and pro-basal bodies on each basal body pair (D-F) with a 2-4 foci. Image taken from the TrypTag project - Dean *et al* (2017)

Results from the BB proteins showed 31% identified to the BB organelle or surrounding structures only, suggesting a functional role to the BBs. Constant recruitment of a protein to an organelle provides a good indication to its function (Dean *et al,* 2017), however 69% of proteins contained a localisation to the BB and to at least one other structure in the cell (table 10). From these, 32% localised to the BBs and cytoplasm, suggesting the presence of extra protein in the cytoplasm (Albert *et al,* 2002).

In the BB dataset only 41% of proteins contained a known domain including tubulin, WW and protein kinase domains. Protein kinases and kinesin domains were the most frequent domains observed. Possibly due to kinases phosphorylating cellular proteins and co-ordinating cell cycle control (Krupa *et al,* 2004), and kinesin domains required for kinesin proteins to transport proteins for growth of the flagellum (Kull *et al,* 1996).

Based on the BB dataset, the highest frequency of orthologues to the *T. brucei* was the *Homo sapiens* group at 38%. *Arabidopsis thaliana, Plasmodium falciparum* and *Saccharomyces cerevisiae* had the lowest level of BB conservation. This could be expected as *Arabidopsis thaliana* and *Saccharomyces cerevisiae* have lost the ability to develop BB organelles (Debec *et al,* 2010; Wasteneys, 2002). However some BB proteins appear to still be conserved in these species (table 12). Studies suggest that ciliary proteins in non-ciliated organisms could have gained novel functions to other roles in biology (Hodges *et al,* 2010). *Trichomonas vaginalis* had a low level of BB protein conservation to *T. brucei,* potentially indicating the level of evolutionary divergence which occurred between these two species.

6.2 Cep164C, the cell cycle dependent protein that identifies basal body age and controls flagellum length

This study identified Cep164C to localise distal to both SAS-6 and TbRP2 (TF) proteins (section 4.4). This localisation suggests that Cep164C could be a component of forming the distal section of the TFs, or could localise to structures distal to the TFs including the collarette (figure 98; B5), encircling the proximal section of the TZ (Vaughan and Gull, 2016).

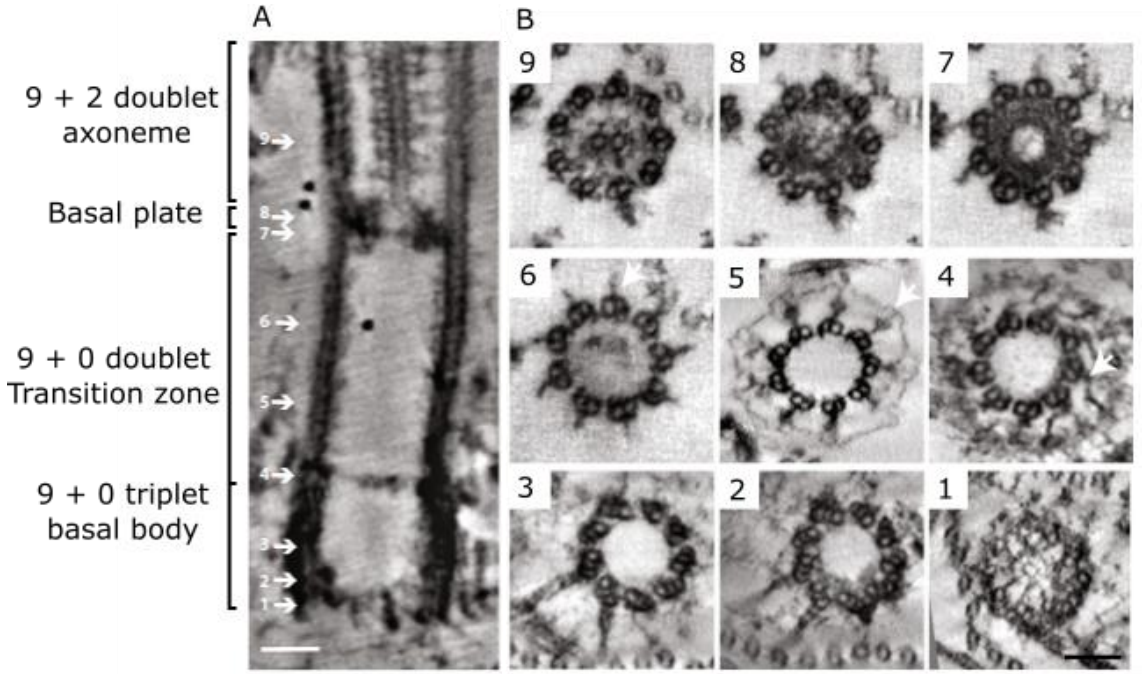


Figure 98: TEM ultrastructure of the *T. brucei* basal body

A) Longitudinal section of the *T. brucei* basal body and flagellum on TEM, with different sections of the structure seen by cross-section (B1-9). Taken from Vaughan and Gull, (2016).

Analysis of the mNG::Cep164C cell line identified Cep164C to localise to the old MBB only, with recruitment occurring prior to BB duplication and becoming absent prior to cytokinesis (figure 50). The localisation to BB1 only indicates that the protein is acquired to those BBs which have undergone multiple rounds of division, with figure 99 explaining in more detail. The model illustrates Cep164C identifies the age of BBs and is not recruited until full maturation of the MBB, at third round of BB duplication.

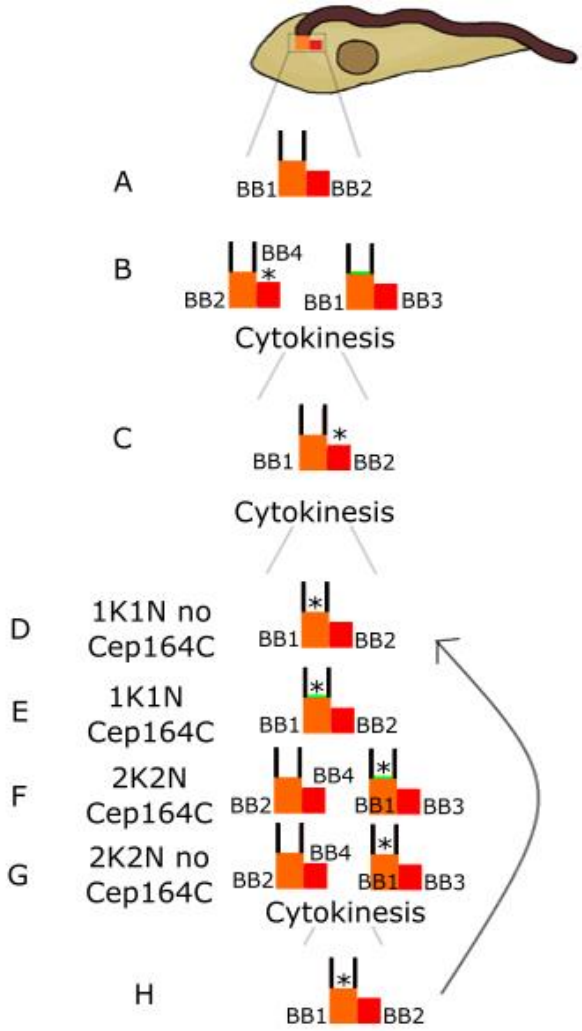


Figure 99: Model of Cep164C recruitment

Model of the recruitment of the Cep164C protein to the basal bodies. After cytokinesis (A), a new pro-basal body is formed, lacking the Cep164C protein (B). After another round of cytokinesis (C), the pro-basal body becomes the mature basal body, but still lacks the Cep164C protein (D). Just prior to basal body duplication the mature basal body then recruits the Cep164C protein (E), which continues to localise to the old mature basal body throughout the cell cycle (F). Finally, Cep164C becomes absent just prior to (G) and after (H) cytokinesis.

RNAi of the Cep164 proteins in the *T. brucei* cell lead to misplaced organelles (kinetoplasts and nuclei). Studies have previously identified organelle displacement as a phenotype and a consequence of RNAi studies leading to

flagellum shortening and generation of epimastigote-like cells (Moreira *et al,* 2017; Hayes *et al,* 2014). Studies have also shown that in natural conditions during the differentiation between the trypomastigote and the epimastigote morphology, flagellum shortening occurs causing the kinetoplast to reposition anterior to the nucleus (Sharma *et al,* 2008). Although all three induced cell lines contained phenotypes as described above, analysis showed a normal cell cycle count. Studies have shown that if cells generate a motile new flagellum, they are still capable of completing cell division (Ralston *et al,* 2006). Currently it is unknown as to whether the dividing cells with one flagellum (1K2N) in the Cep164A and Cep164B RNAi induced cell lines are capable of undergoing cell division.

Enrichment studies in this project showed that Cep164C is not required during PBB maturation, as Cep164C is absent during docking of the NMBB to the flagellar pocket. This suggests that Cep164C is not involved with initial TF formation or maintenance. Based on the Cep164C results from this study, it can be hypothesized that this protein regulates the flagellum length in the *T. brucei* cell. This recruitment is only recruited to the old MBB and is thought to “lock” flagellum length, preventing further development in following cell cycles after reaching its full length. If Cep164C “locks” the growth of the flagellum, this could explain why it is only present on the oldest MBB and not on the NMBB, as the NMBB is still to extend its flagellum length and the addition of Cep164C would prevent this (figure 100A; red arrow). Without the Cep164C present on the oldest MBB, the old flagellum could still extend further than required (figure 100B; red arrow). Analysis showed that Cep164C is lost prior to cytokinesis (figure 50) potentially because the protein is not required for regulation at this stage of the cell cycle, becoming present again prior to BB duplication to prevent additional flagellum growth. RNAi induction of the Cep164C protein produced a phenotype that displayed a wide range of flagellum lengths (figure 63). With the protein failing to “lock” the flagellum length at the appropriate time, the cell cannot regulate entry of IFT and other flagellum building proteins through the ciliary gate.

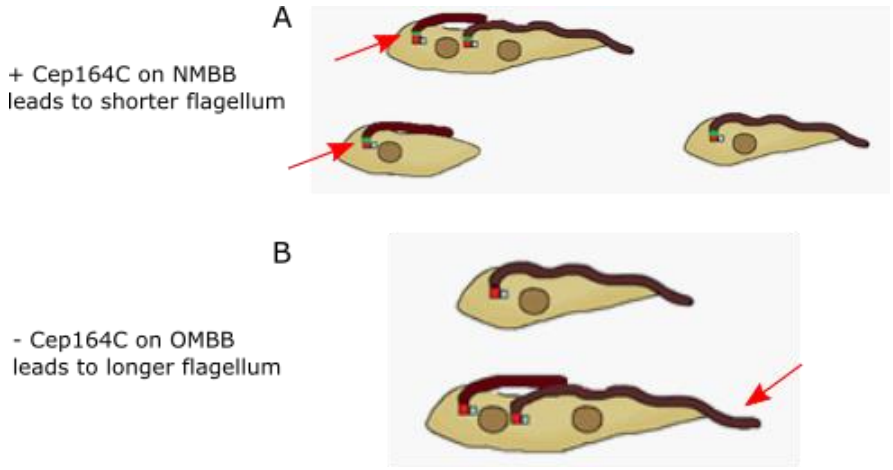


Figure 100: Consequences of + or - Cep164C in the *T. brucei* cell

Forcing the Cep164C protein onto the newly matured basal body after removing the Cep164A and Cep164B protein causes the new flagellum to be shorter in length (A), as Cep164C is preventing the full growth of the flagellum. Induction of the Cep164C RNAi construct also leads to the loss of Cep164C at the old mature basal body and causes a longer old flagellum (B), as Cep164C cannot regulate flagellum length.

Results demonstrated cells with a much smaller cell body and flagellum. Generation of these smaller cells is thought to be due to the reduced IFT entry into the new flagellum, as the old flagellum is also extending at similar rates. This could prevent correct growth of the new flagellum either due to a limited growth time prior to cytokinesis being initiated, or there is a biological limit of IFT particles for growth (Engel *et al,* 2009).

Although flagellum length regulation is considered lost in the induced Cep164C RNAi cell line, no flagellum had a measurement higher than 37μm. It is possible that after this length the flagellum becomes too long and cell motility is compromised, or there is a biological limit to flagellum length. A study demonstrated that trypomastigotes during asymmetrical division naturally develop long flagella, but with none extending past 38μm (Sharma *et al,* 2008). No previous studies have generated longer flagellum in the *T. brucei* cell, however longer flagella have been generated in the *Chlamydomonas* species. LF4 mutants were

unable to regulate their flagellum lengths and produced flagella twice the length of wild-type cells (Rosenbaum, 2003) (figure 101).

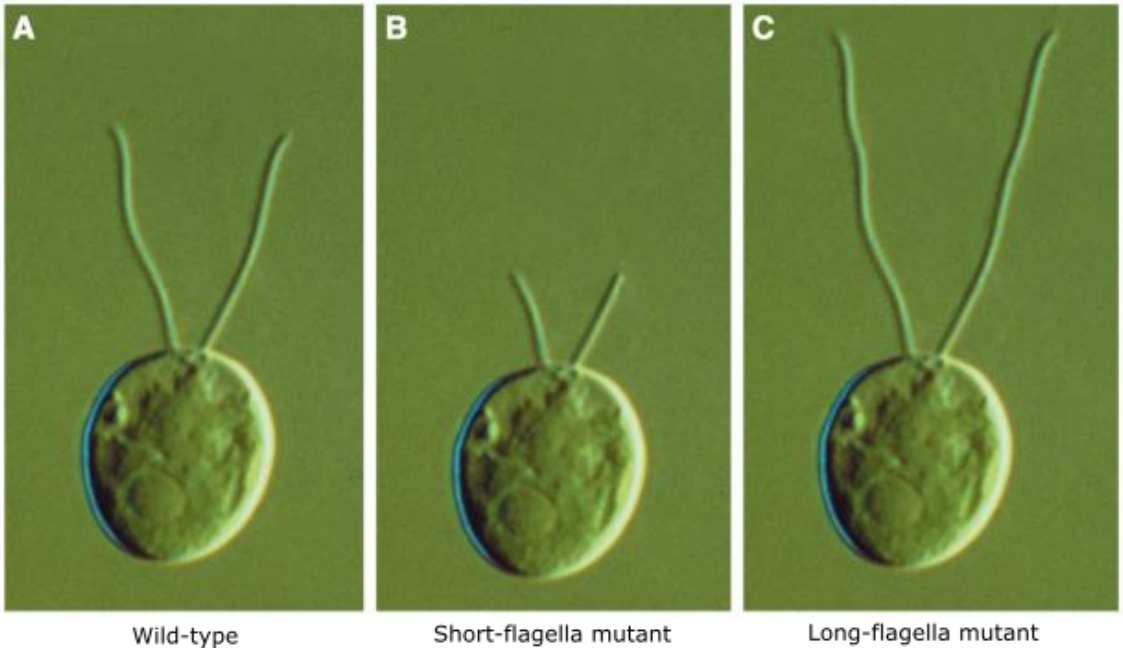


Figure 101: Consequence of the LF4 mutant in Chlamydomonas

Rosenbaum (2003) showed Chlamydomonas mutants that were also able to generate various flagellum lengths when targeting the LF4 protein. These mutants were able to generate either shorter flagellum lengths (B) or longer flagellum lengths (C) in comparison to the uninduced (A).

Investigations uncovered this protein was a protein kinase localising to the flagellum, resulting in lengths between 0-38μm, with wild-type cells typically producing lengths of 12μm (Berman *et al,* 2003). To our knowledge no other studies have demonstrated longer flagellum phenotypes.

In this study the flagellar pocket volumes were analysed. No studies have previously investigated flagellar pocket volumes in different stages of the cell cycle and it was hypothesised that the flagellar pocket volumes increase in size as the cell progresses through the cell cycle. Results confirmed that measurements of the uninduced Cep164C flagellar pockets were smaller in the early stages of the cell cycle (new and old) (figure 72), increasing in size in later cell cycle stages with the new and old pockets increasing simultaneously. This could be explained by an

increase in protein trafficking to the flagellar pockets for biological purposes, or to expand in size to accommodate the newly invading flagellum in the next cell cycle.

In the induced RNAi Cep164C cell line the single pockets were larger than the uninduced, whereas in dividing cells only the old flagellar pocket was statistically significantly larger to the uninduced. As lack of the Cep164C prevents the flagellum from “locking” this volume increase could be due to the increase in IFT particles and other required proteins to continuously build a longer flagellum, or potentially from an imbalance in membrane trafficking pathways as observed in a study conducted by Demmel *et al* (2014).

6.3 Cep164A and Cep164B has a different phenotype to Cep164C

As all three proteins co-localised to each other through Airyscan imaging, this suggests that Cep164A and Cep164B also localise distal to the TbRP2 protein at either the TFs or to surrounding structures including the collarette. Cep164A and Cep164B however are not cell cycle dependent proteins and remain recruited to both BB1 and BB2 at all times (figure 75). Localisation to the MBB only suggests this protein could form part of the TFs to allow docking of the BB to the flagellar pocket, explaining why these proteins are constantly present. As these proteins are recruited prior to new flagellum growth (figure 77C and 77G), they could also function in BB docking and are constantly present to keep the BBs docked to the flagellar pocket.

RNAi induction results showed Cep164A and Cep164B generate a different phenotype to Cep164C. Although both produced daughter cells with varied flagellum lengths, they did not generate extreme lengths similar to Cep164C with a maximum of 29μm (Cep164A) and 33μm (Cep164B) measured. A difference was in the new flagellum lengths which were much shorter than those in the Cep164C RNAi cell line, with some cells generating no new flagellum growth (figure 89 and 90). Previous studies have shown that producing dividing cells with one flagellum is rare, however 1K2N cells can be formed by RNAi induction of proteins which are required for kinetoplast and BB migration (Absalon *et al,* 2007). If the function of

Cep164A and Cep164B was related to BB docking, this could also explain the increase in 1K2N cells, as TFs formation or docking would not be effective leading to a reduced amount of MBBs successfully docking and producing a new flagellum.

Cep164A and Cep164B are not hypothesised to regulate flagellum length, however lengths of the induced cells have measured up to 29µm and 33µm respectively. This could be due to loss of both Cep164A and Cep164B at the BBs, forcing Cep164C onto the newly matured BB (figure 102).

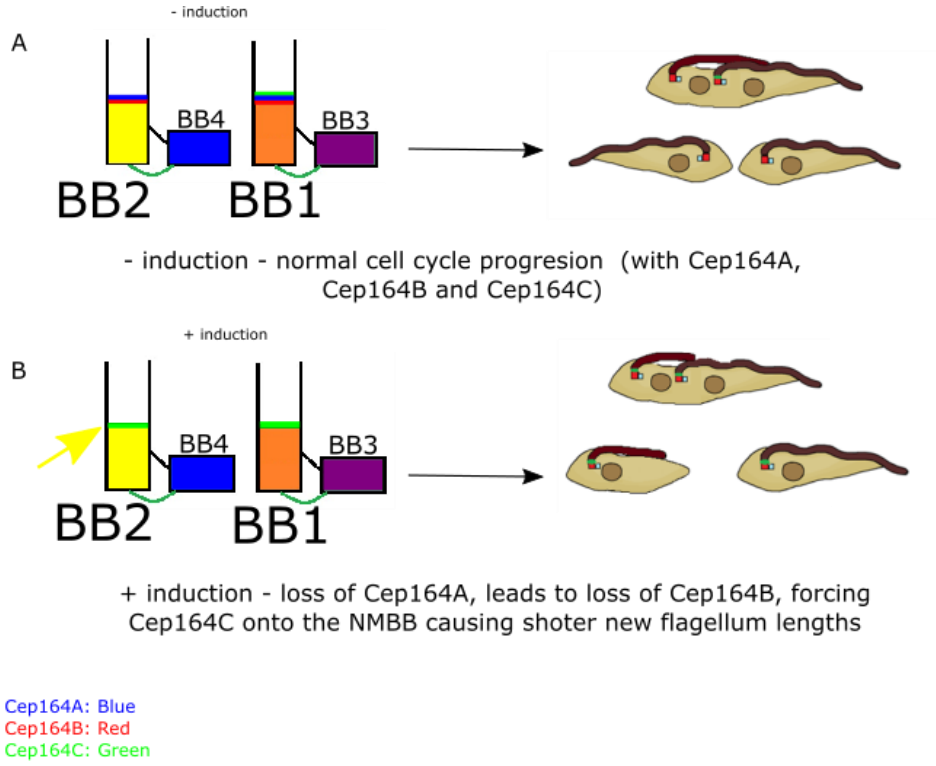


Figure 102: Loss of Cep164A and Cep164B forces Cep164C onto the newly matured basal body

T. brucei cells contain the Cep164A, Cep164B and Cep164C proteins to the basal bodies (A), when Cep164A and Cep164B are lost, Cep164C is forced onto the newly matured basal body, potentially causing the shorter new flagellum in RNAi induced cell lines.

In this study, RNAi induction of Cep164A and Cep164B failed to localise either protein to the BBs, although the proteins were still present in the cell. This suggests these proteins may form a complex and rely on each other for recruitment or stability. However, Cep164C is not dependent on either protein for BB recruitment. Instead, without Cep164A or Cep164B, Cep164C was forced onto the NMBB (figure 102B), not occurring in normal conditions and may explain why the new flagellum is much shorter in induced Cep164A and Cep164B cell lines. With both of these proteins absent, Cep164C is distributed onto both MBBs (less intensity on the NMBB), blocking particle entry into the new flagellum and preventing its extension (figure 102B).

In both proteins' absence, Cep164C may carry out the role of Cep164A and Cep164B, causing a compromise to Cep164Cs original function. This could then explain why Cep164A and Cep164Bs flagellum lengths are slightly longer than the uninduced cell line.

6.4 Impact of Cep164 protein knowledge to the *T. brucei* are eukaryotic field.

The three Cep164 proteins (A, B and C) from the same protein family have been investigated in the *T. brucei* cell line to identify the novel roles these proteins play in basal body docking and flagellum length control. Investigating the Cep164C protein has shown the mechanism of flagellum length control, and that removing this protein breaks this mechanism. Investigation of the Cep164 proteins has aided in understanding why the *T. brucei* requires three Cep164 proteins, whereas other species such as humans and other mammals only contain one Cep164 protein. With the Cep164A and Cep164B being required for basal body docking/TF formation, and Cep164C required for flagellum length control. Although the localisation to the Cep164 protein in humans has been identified to the distal appendages, research of the three proteins in *T. brucei* has shown that all three localise to the transitional fibres, helping to further understand the transitional fibre composition. Discovering that the Cep164 proteins in the *T. brucei* cell are located

very distal to the transitional fibres could also aid others in future experiments and use these proteins as a marker.

For the first time to our knowledge, it has been noted that the protein composition of the new and old mature basal bodies in the *T. brucei* cell are not identical. It has been shown that the old mature basal body contains an additional protein (Cep164C) in comparison to the newly matured basal body, and that this protein is recruited after the development of the transitional fibres whereas the Cep164A and Cep164B are present before transitional fibre development is complete. Induction of the Cep164C RNAi cell line has also shown that the longest a cell may generate a flagellum could be around 37μm, similar to the trypomastigote cells during asymmetrical division, which reach flagellum lengths of up to 38μm. This may indicate the maximum length a flagellum can extend, either for biological or viability purposes.

Studies on the Cep164C protein has highlighted a limiting pool of IFT proteins in the *T. brucei* cells. Removing the mechanism for flagellum length control leads to the old flagellum continuously extending, however as the old flagellum is continuously growing the new flagellum cannot grow to the correct length, suggesting a limited IFT pool. Interestingly, these studies have shown that the Cep164A and Cep164B proteins are reliant on each other to localise to the basal bodies, whereas the Cep164C protein is not dependant on either for recruitment. Investigation has also shown that the Cep164A and Cep164B proteins have a very different function to Cep164C, and that Cep164A/B may have a role in basal body docking/segregation, explaining why some cells during induction of the Cep164A/B RNAi cell line produce unusual 1K2N cells with one flagellum.

Finally, this study has highlighted that there is a mechanism for flagellum control in the *T. brucei* cell. As this protein has not been studied in many species, further studies could be carried out to understand whether this mechanism is something specific to trypanosomes, or whether other species outside the kinetoplastids can also regulate their flagellum/cilium lengths through the Cep164C protein.

6.5 Future perspectives

This study improved the understanding in the function of the Cep164 proteins in the *T. brucei* parasite. Due to time constraints, not all experiments could be carried out. The basal body screen identified a few proteins with interesting localisations including the Tb927.7.5190 protein, which localised to the PBB only through the cell cycle. Experiments could have been conducted to observe whether this protein functioned in maturation or development of the TFs on the PBB, or whether it functioned in recruitment of proteins for maturation.

The human Cep164C protein has an important function to recruit TTBK2 to remove the ciliary cap, allowing flagellum growth. TTBK2 has not been confirmed in *T. brucei* and additional experiments could have investigated this protein and whether the Cep164 proteins recruit TTBK2, allowing ciliogenesis.

Additional investigations to the Cep164 proteins could have been carried out for further understanding. Immuno-gold labelling could have been used for a more accurate localisation to the distal section of the MBBs as this uses antibodies that bind to specific antigens, with the localisation being detected through gold-particles (Murtey, 2015). This could have further identified whether the Cep164 proteins localise to the TFs, or the collarette structure.

Tracking tagged IFT particles could have been carried out to determine whether there is a reduced IFT entry in the new flagellum or increase IFT entry in the old flagellum. Further explaining the generation of the shorter new flagella and longer old flagella in the induced cell line. Kymographs could have been used to determine the rate of entry and compared against the uninduced cell line.

Further studies with fluorescent microscopy and TEM could have been conducted to investigate the recruitment of Cep164C to the NMBB, in the absence of Cep164A or Cep164B. These experiments may have uncovered whether the addition of Cep164C to the NMBB causes the shorter flagellum which would further support the hypothesised function. Generation of an overexpression cell line could also have been used to further force Cep164C onto the new flagellum, possibly generating shorter new flagella in dividing cells and supporting Cep164Cs function.

Dual tomography could be carried out for the induced RNAi cell lines of Cep164A and Cep164B to bring insight into whether these proteins function in docking of the BBs or generation of the TFs and whether BBs are able to dock correctly in their absence. TEM studies could have been carried out on the TFs to determine whether lack of these proteins lead to morphology changes, indicating their function. SBF-SEM modelling could also have investigated whether those 1K2N cells have successfully duplicated or docked their BBs. It could also investigate whether a very short flagellum had grown in those which appear to lack a flagellum, but no PFR was formed in the cell so no detection of a flagellum was found during fluorescent microscopy.

References

Absalon, S., Blisnick, T., Kohl, L. *et al* (2008) Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes. *Molecular biology of the cell.* 19(3), 929-944

Absalon, S., Kohl, L., Branche, C. *et al* (2007) Basal body positioning is controlled by flagellum formation in *Trypanosoma brucei. PLoS One*. 2(5), 437

Albert, B., Johnson, A., Lewis, J. *et al* (2002) From RNA to Protein. *Molecular Biology of the Cell.* 4th edition. USA: Garland. 335-364

Alsford, S., Turner, D. J., Obado, S. *et al* (2011) High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Research.* 21(6), 915-924

Andersen, J. S., Wilkinson, C. J., Mayor, T. *et al* (2003) Proteomic characterization of the human centrosome by protein correlation profiling. *Nature.* 426(6966), 570-574

Andre, J., Kerry, L., Qi, X. *et al* (2014) An alternative model for the role of RP2 protein in flagellum assembly in the African trypanosome. *The Journal of Biological Chemistry.* 289(1), 464-475

Archer, S. K., Inchaustegui, D., Queiroz, R. *et al* (2011) The cell cycle regulated transcriptome of *Trypanosoma brucei. PLoS one.* 6(3), 18425

Aslett, M., Aurrecoechea, C., Berrimen, M. *et al* (2010) TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Research.* 38(1), 457-462

Aurrecoechea, C., Brestelli, J., Brunk, B. P. *et al* (2008) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research.* 37, 539-543

Aurrecoechea, C., Brestelli, J., Brunk, B. P. *et al* (2009) GiardiaDB and TrichDB: integrated genomic resources of the eukaryotic protist pathogens *Giardia lambia* and *Trichomonas vaginalis*. *Nucleic Acids Research.* 37, 526-530

Avidor-Reiss, T., Khire, A., Fishman, E. L. *et al* (2015) Atypical centrioles during sexual reproduction. *Frontiers in Cell and Developmental Biology* [online]. 3(21) [viewed 20th June, 2019]. Available from: doi: 10.3389/fcell.2015.00021

Badano, J. L., Mitsuma, N., Beales, P.L. *et al* (2006) The ciliopathies: An emerging class of human genetic disorders. *Annual Review of Genomics and Human Genetics.* 7, 125-148

Bastin, P., Bagherzadeh, Z. *et al* (1996) A novel epitope tag system protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*. 77(2), 235-239

Belevich, I., Joensuu, M. *et al* (2016) Microscopy image browser: A platform for segmentation and analysis of multidimensional datasets. *PLoS Biology* [online]. 14(1) [Viewed 16th June 2019]. Available from: doi: 10.1371/journal.pbio.1002340.

Benson, D. A., Cavanaugh, M., Clark, K. (2013) GenBank. *Nucleic Acids Research.* 41, 36-42

Berardini, T. Z., Resier, L. and Li, D. (2015) The Arabidopsis information resource: making and mining the gold standard annotated reference plant genome. *Genesis*. 53(8), 474-485

Bettencourt-Dias, M., Rodrigues-Martins, L. and Carpenter, M. *et al* (2005) SAK/PLK4 is required for centriole duplication and flagella development. *Current Biology*. 15(24), 2199-2207

Berman, S. A., Wilson, N. F., Haas, N. A. *et al* (2003) A novel MAP kinase regulates flagellar length in Chlamydomonas. *Current Biology.* 13(13), 1145-1149

Bertiaux, E., Morga, B., Blisnick, T. *et al* (2018) Control of flagellum length by a grow-and-lock model. *Current Biology.* 28(23), 3802-3814

Brouhard, G. J and Rice, L. M (2014) The contribution of αβ-tubulin curvature to microtubule dynamics. *Journal of Cell Biology.* 207(3), 323-334

Burke, M. C., Li, F., Cyge, B. (2014) Chibby promotes ciliary vesicle formation and basal body docking during airway cell differentiation. *Journal of Cell Biology.* 207(1), 123-137

Cajanek, L. and Nigg, E. A. (2014) Cep164 triggers ciliogenesis by recruiting Tau tubulin kinase 2 to the mother centriole. *Proceedings of the National Academy of Sciences of the United States of America.* 111(28), 2841-2850

CDC (Centers for disease control and prevention) (2019) Parasites: African Trypanosomiasis. [online] CDC [08th July, 2019] available at:https://www.cdc.gov/parasites/sleepingsickness/disease.html

CDC (Centres for disease control and prevention) (2019) Trypanosomiasis, African. [online] CDC [28th July, 2019] available at: https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html

Chaki, M., Airik, R., Ghosh, A. K. *et al* (2012) Exome capture reveals ZNF423 and Cep164 mutations, linking renal ciliopathies to DNA damage response signalling. *Cell.* 150(3), 533-548

Cherry, J. M., Hong, E. L., Amundsen, C. *et al* (2012) Saccharomyces genome database: the genomics resource of budding yeast. *Nucleic Acids Research.* 40, 700-705

Cocks, E., Taggart, M. *et al* (2018) A guide to analysis and reconstruction of serial block face scanning electron microscopy data. *Journal of Microscopy.* 270(2), 217-234

Crasta, K and Surana, U. (2006) Disjunction of conjoined twins: Cdk1, Cdh1 and separation of centrosomes. *Cell Division* [online]. 1(12) [viewed 20th July 2019]. Available from: doi: 10.1186/1747-1028-1-12

Creek, D. J., Nijagal, B. *et al* (2013) Metabolomics guide rational development of a simplified cell culture medium for drug screening against *Trypanosoma brucei. Antimicrobial agents and chemotherapy.* 57(6), 2768-2779

Daly, O. M., Gaboriau, D., Karakaya, K. *et al* (2016) Gene-targeted Cep164-deficient cells show a ciliation defect with intact DNA repair capacity. *Journal of Cell Science.* 129(9), 1769-1774

Dang, H. Q,. Zhou, Q. *et al* (2017) Proximity interactions among basal body components in *Trypanosoma brucei* identify novel regulators of basal body biogenesis and inheritance. *mBio* [online]. 8(1) [viewed 11th March 2019]. Available from: doi: 10.1128/mBio.02120-16

Dawe, H. R., Farr, H. and Gull, K. (2007) Centriole/basal body morphogenesis and mitigation during ciliogenesis in animal cells. *Journal of Cell Science.* 120(1), 7-15

Dean, S., Moreira-Leite, F. *et al* (2016) Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes. *Proceedings of the National Academy of Sciences to the United States of America.* 113(35), 135-143

Dean, S., Sunter, J., Wheeler, R. J. *et al* (2015) A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biology.* 5(1), 140197

Dean, S., Sunter, J. D and Wheeler, R. J. (2017) TrypTag.org: A trypanosome genome-wide protein localisation resource. *Trends in Parasitology.* 33(2), 80-82

Debec, A., Sullivan, W. and Bettencourt-Dias, M. (2010) Centrioles: Active players or passengers during mitosis? *Cellular and Molecular Life Sciences.* 67(13), 2173-2194

Demmel, L., Schmidt, K., Lucast, L. *et al* (2014) The endocytic activity of the flagellar pocket in *Trypanosoma brucei* is regulated by an adjacent

phosphatidylinositol phosphate kinase. *Journal of Cell Science.* 127(10), 2351-2364

Dentler, W. (2005) Intraflagellar transport (IFT) during assembly and disassembly of Chlamydomonas flagella. *Journal of Cell Biology.* 170(4), 649-659

Dummer, A., Poelma, C., DeRuiter, M. C. (2016) Measuring the primary cilium length: Improved method for unbiased high-throughput analysis. *Cilia.* [online]. 5(7). [viewed 20th March 2019]. Available from: doi: 10.1186/s13630-016-0028-2

Dutcher, S. K. (2003) Long-lost relative reappear: Identification of new members of the tubulin superfamily. *Current Opinion in Microbiology.* 6(6), 634-640

Duvick, J., Fu, A., Muppirala, U. *et al* (2008) PlantGDB: a resource for comparative plant genomics. *Nucleic Acids Research.* 36, 959-965

Eisen, J .A., Coyne, R. S., Wu, M. *et al* (2006) Macronuclear genome sequence of the ciliate *Tetrahymena thermophila,* a model eukaryote. *PLoS.* 4(9), 286

Engel, B. D., Ludington, W. B. and Marshall, W. F. (2009) Intraflagellar transport particle size scales inversely with flagellar length: revisiting the balance-point length control mode. *Journal of Cell Biology.* 187(1), 81-89

Engstler, M., Pfohl, T., Herminghaus, S. *et al* (2007) Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell.* 131(3), 505-515

Farr, G. W., Yaffe, M, B. and Sternlicht, H. (1989) α-tubulin influences nucleotide binding to β-tubulin: An assay using picomoles of unpurified protein. *Proceedings of the National Academy of Sciences to the United States of America.* 87(13), 5041-5045

Farr, H and Gull, K. (2009) Functional studies of an evolutionary conserved, cytochrome b5 domain protein reveal a species role in axonemal organisation and the general phenomenon of post-division axonemal growth in trypanosomes. *Cell Motility and the Cytoskeleton.* 66(1), 24-35

Finn, R. D., Attwood, T. K., Babbitt. P. C. (2017) Interpro in 2017-beyond protein family and domain annotations. *Nucleic Acids Research.* 4(45), 190-199

Finn, R. D., Clements, J. and Eddy, S. R. (2011) HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research.* 1(39), 29-37

Finn, R. D., Coggill, P., Eberhardt. R. Y. (2016) The Pfam protein families database: toward a more sustainable future. *Nucleic Acids Research.* 4(44), 279-285

Fliegauf, M., Benzing, T. and Omran, H. (2007) When cilia go bad: Cilia defects and ciliopathies. *Nature Reviews. Molecular Cell Biology.* 8(11), 880-893

Fort, C., Bonnefoy, S., Kohl, L. *et al* (2016) Intraflagellar transport is required for the maintenance of the trypanosome flagellum composition but not its length. *Journal of Cell Science.* 129(15), 3026-3041

Foster, L. J., Hoog, C. L., Zhang, Y. *et al* (2006) A mammalian organelle map by protein correlation profiling. *Cell.* 125(1), 187-199

Franco, J. R., Simarro, P. P., Diarra, A. *et al* (2014) Epidemiology of human African trypanosomiasis. *Clinical Epidemiology.* 6(6), 257-275

Ganguly, A., Yang, H., Sharma, R. *et al* (2012) The role of microtubules and their dynamics in cell migration. *Journal of Biological Chemistry.* 287(52), 43359-43369

Geimer, S. and Melkonian, M. (2004) The ultrastructure of the *Chlamydomonas reinhardtii* basal apparatus: identification of an early marker of radial asymmetry inherent in the basal body. *Journal of Cell Science.* 117(13), 2663-2674

Gluenz, E., Hoog, J. L., Smith, A. E. *et al* (2010) Beyond 9+0: noncanonical axoneme structure characterize sensory cilia from protists to humans. *Federation of American Societies for Experimental Biology.* 24(9), 3117-3121

Gluenz, E., Povelones, M. L., Englund, P. T. *et al* (2011) The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. *Molecular and Cellular Biology. 3*1(5), 1012-1021

Gramates, L. S., Marygold, S. J., Santos, G. *et al* (2017) FlyBase at 25: looking to the future. *Nucleic Acids Research.* 4(45), 663-671

Hammarton, T. C. (2007) Cell cycle regulation in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology.* 153(1-4), 1-8

Hatch, E. and Stearns, T. (2014) The life cycle of centrioles. *Cold Spring Harbor Symposia on Quantitative Biology. 7*5, 425-431

Hayes, P., Varga, V., Olego-Fernandez, S. *et al* (2014) Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosoma morphology. *Journal of Cell Biology.* 206(3), 377-384

Hiraki, M., Nakazawa, Y., Kamiya, R. *et al* (2007) Bld10p constitutes the cartwheel-spoke tip and stabilizes the 9-fold symmetry of the centriole. *Current Biology.* 17(2), 1778-1783

Hirono, M. (2014) Cartwheel assembly. *Philosophical Transactions of the Royal Society B: Biological Science.* 369(1650), 20130458

Hodges, M. E., Scheumann, N. (*et al*) (2010) Reconstructing the evolutionary history of the centriole from protein components. *Journal of Cell Science.* 1(123), 1407-1413

Howe, K., Clark, M. D. and Torroja, C. F. (2013) The zebrafish reference genome sequence and its relationship to the human genome. *Nature.* 496(7446), 498-503

Hu, H., Liu, Y. *et al* (2015) The centriole cartwheel protein SAS-6 in *Trypanosoma brucei* is required for probasal body biogenesis and flagellum assembly. *Eukaryotic Cell.* 14(9), 898-907

Hughes, L., Towers, K. *et al* (2013) A cell-body groove housing the new flagellum tip suggests an adaptation of cellular morphogenesis for parasitism in the bloodstream form of *Trypanosoma brucei. Journal of Cell Science.* 126(24), 5748-5757

Ishikawa, H. and Marshall, W. F. (2017) Intraflagellar transport and ciliary dynamics. *Cold Spring Harbour Perspectives in Biology*. [online]. 9(3). [viewed 20th March 2019]. Available from: doi: 10.1101/cshperspect.a021998.

Ishikawa, T. (2017) Axoneme structure from motile cilia. *Cold Spring Harbour Perspectives in Biology.* [online]. 9(1). [viewed 15th March 2019]. Available from: doi: 10.1101/cshperspect.a028076.

Joo, K., Kim, C. G., Lee, M. *et al* (2013) CCDC41 is required for ciliary vesicle docking to the mother centriole. *Proceedings of the National Academy of Sciences of the United States of America.* 110(15), 5987-5992

Kalidas, S., Li, Q and Phillips, M. A. (2011) A gateway compatible vector for gene silencing in bloodstream for *Trypanosoma brucei. Molecular and Biochemical Parasitology.* 178(1-2), 51-55

Karimi, K., Fortriede, J. D., Lotay, V. S. *et al* (2018) Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Research.* 4(46), 861-868

Kilmartin, J. V., Wright, B. and Milstein, C. (1982) Rat monoclonal antitubulin antibodies derived by using a new nonsecreting rat cell line. *Journal of Cell Biology.* 93(3), 576-582

Kobayashi, T. and Dynlacht, B. D. (2011) Regulating the transition from centriole to basal body. *Journal of Cell Biology.* 193(3), 435-444

Kohl, L., Robinson, D. *(et al*) (2003) Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes. *The European Molecular Biology Organization Journal.* 22(20), 5336-5346

Kohl, L., Sherwin, T, Gull, K. (1999) Assembly of the paraflagellar rod and the flagellum attachment zone complex during the *Trypanosoma brucei* cell cycle. *The Journal of Eukaryotic Microbiology.* 46(2), 105-109

Kremer, J. R., Mastronarde, D. N. and McIntosh, J. R. (1996) Computer visualisation of three-dimensional image data using IMOD. *Journal of Structural Biology.* 116(1), 71-76

Krupa, A., Abhinandan, K. R and Srinivasan, N. (2004) KinG: A database of protein kinases in genomes. *Nucleic Acids Research.* 1(32), 153-155

Kull, F. J., Sablin, E. P., Lau, R. *et al* (1996) Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature.* 380(6574), 550-555

Langousis, G. and Hill, K. L. (2014) Motility and more: The flagellum of *Typanosoma brucei. Nature Reviews Microbiology.* 12(7), 505-508

Li, S., Luckins, A and Lun, Z. (2010) *Trypanosoma brucei brucei:* A comparison of gene expression in the liver and spleen of infected mice utilizing cDNA microarray technology. *Experimental Parasitology.* 128(3), 256-264

Loncarek, J and Bettencourt-Dias, M. (2018) Building the right centriole for each cell type. *Journal of Cell Biology.* 217(3), 823-835

Luduena, R. F. (2013) A hypothesis on the origin and evolution of tubulin. *International Review of Cell and Molecular Biology.* 302, 41-185

Matthews, K. R. (2005) The developmental cell biology of *Trypanosoma brucei. Journal of Cell Science.* 118(2), 283-290

McKean, P. G. (2003) Coordination of cell cycle and cytokinesis in *Trypanosoma brucei. Current Opinion in Microbiology.* 6(6), 600-607

McKean, P. G., Vaughan, S and Gull, K. (2001) The extended tubulin superfamily. *Journal of Cell Science.* 114(15), 2723-2733

Murtey, M. D (2015) *Immunogold Techniques in Electron Microscopy* [online]. IntechOpen. Available from: doi: 10.5772/61719

Meunier, S and Vernos, I. (2012) Microtubule assembly during mitosis – from distinct origins to distinct functions? *Journal of Cell Science.* 125(12), 2805-2814

Mitchison, H. M and Valente, E. M. (2016) Motile and non-motile cilia in human pathology: From function to phenotypes. *The Journal of Pathology.* 241(2), 294-309

Moreira, B., Fonseca, C., Hammarton, T. *et al* (2017) Giant FAZ10 is required for flagellum attachment zone stabilization and furrow positioning in *Trypanosoma brucei*. *Journal of Cell Science.* 130(6), 1179-1193

Nakazawa, Y., Ariyoshi, T., Noga, A. *et al* (2014) Space-dependent formation of central pair microtubules and their interactions with radial spokes. *PLoS One*. [online]. 9(10), e110513 [9th July 2019]. Available from: doi: 10.1371/journal.pone.0110513

Nakazawa, Y., Hiraki, M., Kamiya, R. *et al* (2007) SAS-6 is a cartwheel protein that establishes the 9-fold symmetry of the centriole. *Current Biology.* 17(24), 2169-2174

Nature Education (2014) Cytoskeletal networks provide spatial organization and mechanical support to eukaryotic cells. [online]. *Nature Education.* [10th July 2019]. Available from: https://www.nature.com/scitable/ebooks/cell-biology-for-seminars-14760004/118244414

Nogales, E. (2000) Structural insights into microtubule function. *Annual Review of Biochemistry.* 69, 277-302

Nogales, E., Wolf, S. G and Downing, K. H. (1998) Structure of the alpha beta tubulin dimer by electron crystallography. *Nature.* 391(6663), 199-203

Oakley, B. R. (1999) γ-tubulin. *Current Topics in Developmental Biology.* 49, 27-54

Oakley, B. R. (2000) An abundance of tubulins. *Trends in Cell Biology.* 10(12), 537-542

Otte, L., Wiedemann, U., Schlegel, B. *et al* (2003) WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Science.* 12(3), 491-500

Park, S. H., Nguyen, B. N., Kirkham, J. K. *et al* (2014) A new strategy of RNA interference that targets heterologous sequences reveals CITFA1 as an essential component of class 1 transcription factor A in *Trypanosoma brucei. Eukaryotic Cell.* 13(6), 785-795

Paz, J. and Luders, J. (2018) Microtubule-organizing centers: Towards a minimal parts list. *Trends in Cell Biology.* 28(3), 176-187

Pearson, C. G (2014) Choosing sides – Asymmetric centriole and basal body assembly. *Cell Science.* 127, 2803-2810

Pedersen, L. B., Schroder, J. M., Christensen, S. T. *et al* (2012) The ciliary cytoskeleton. *Comprehensive Physiology.* 2(1), 779-803

Poon, S. K., Peacock, L. *et al* (2012) A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open biology.* [online]. Vol. 2(2), e110037 [13th May 2019]. Available from: doi: 10.1098/rsob.110037.

Ralston, K. S. and Hill, K. L. (2008) The flagellum of *Trypanosoma brucei*: New tricks from an old dog. *International Journal for Parasitology.* 38(8-9), 869-884

Ralston, K. S., Kabututu, Z. P., Melehani, J. H. *et al* (2009) The *Trypanosoma brucei* flagellum: Moving parasites in new directions. *Annual Review of Microbiology.* 63, 335-362

Ralston, K., Lerner, A., Diener, D. *et al* (2006) Flagellar motility contributes to cytokinesis in *Trypanosoma brucei* and is modulated by an evolutionary conserved dynein regulatory system. *Eukaryotic Cell.* 5(4), 696-711

Redmond, S., Vadivelu, J. and Field, M. C. (2003) RNAit: An automated web-based tool for the selection of RNAi targets in *Typanosoma brucei.* 128(1), 115-118

Reiter, J. F., Blacque, O. E and Leroux, M. R. (2012) The base of the cilium: Roles for transition fibres and the transition zone in ciliary formation, maintenance and compartmentalization. *The European Molecular Biology Organization Journal.* 13(7), 608-618

Rosenbaum, J. (2003) Organelle size regulation: Length matters. *Current Biology.* 13(13), 506-507

Rotureau, B., Blisnick, T., Subota, I. *et al* (2014) Flagellar adhesion in *Trypanosoma brucei* relies on interactions between different skeletal structures in the flagellum and cell body. *Journal of Cell Science.* 127(1), 204-215

Sanchez, A. D. and Feldman, J. L. (2018) Microtubule-organizing centers: From the centrosome to non-centrosomal sites. *Current Opinion in Cell Biology.* 44, 93-101

Schmidt, K. N., Kuhns, S., Neuner, A., *et al* (2012) Cep164 mediates vesicular docking to the mother centriole during early steps of ciliogenesis. *The Journal of Cell Biology.* 199(7), 1083-1101

Sept, D. (2007) Microtubule polymerizations: One step at a time. *Current Biology.* 17(17), 764-766

Sharma, R., Peacock, L., Gluenz, E. *et al* (2008) Asymmetric cell division as a route to reduction in cell length and change in cell morphology in Trypanosomes. *Protist.* 159(1), 137-151

Short, B. (2014) Chibby's function in the dock. *Journal of Cell Biology.* 207(1), 2

Siegel, N. T., Hekstra, R. D. *et al* (2008) Analysis of the *Trypanosoma brucei* cell cycle by quantitative DAPI imaging. *Molecular and biochemical parasitology.* 160(2), 171-174

Sillibourne, J. E and Bornens, M. (2010) Polo-like kinase 4: The odd one out of the family. *Cell Division* [online]. Vol. 5(1) [9th July 2019]. Available from: doi: 10.1186/1747-1028-5-25.

Sillibourne, J. E., Specht, C. G., Izeddin, I. *et al* (2011) Assessing the localisation of centrosomal proteins by PALM/STORM nanoscopy. *Cytoskeleton.* 68(11), 619-627

Sivasubramaniam, S., Sun, X., Pan, Y. *et al* (2008) Cep164 is a mediator protein required for the maintenance of genomic stability through modulation of MDC1, RPA and CHK1. *Genes and Development*. 22(5), 587-600

Slaats, G. G., Ghosh, A. K., Falke, L. L. *et al* (2014) Nephronophthisis-associated Cep164 regulates cell cycle progression, apoptosis and epithelial-to-mesenchymal transition. *PLoS Genetics.* [online] Vol. 10(10) e1004594 [8th July 2019]. Available from: doi: 10.1371/journal.pgen.1004594

Sorokin, S (1962) Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells. *The Journal of Cell Biology.* 15(2), 363-377

Stein, L., Sternberg, P., Durbin, R. *et al* (2001) WormBase: network access to the genome and biology of Caenorhabditis elegans. *Nucleic Acids Research.* 1(29), 82-86

Stephan, A., Vaughan, Shaw, M. K. *et al* (2007) An essential quality control mechanism at the eukaryotic basal body prior to intraflagellar transport. *Traffic.* 8(10), 1323-1330

Stockman, M., Lillen, M. and Knoers, N (2016) Nephronophthisis. [online]. *Gene Reviews.* [14th November 2018]: Available from: http://www.ncbi.nlm.nih.gov/books/NBK368475/

Tahir, M., Khan, A and Majid, A. (2012) Protein subcellular localisation of fluorescence imagery using spatial and transform domain features. *Bioinformatics.* 28(1), 91-97

Tanos, B. E., Yang, H., Soni, R. *et al* (2013) Centriole distal appendages promote membrane docking, leading to cilia initiation. *Genes and Development.* 27(2), 163-168

Trépout, S., Tassin, A. M. *et al (*2018) STEM tomography analysis of the trypanosome transition zone. *Journal of Structural Biology.* 202(1), 51-60

Troeberg, T., Chen, X., Flaherty, T. M. *et al* (2000) Chalcone acyl hydrazide, and related amides kill cultured *Trypanosoma brucei brucei. Molecular Medicine.* 6(8), 660-669

Turk, E., Wills, A. A., Kwon, T. *et al* (2015) Zeta-tubulin is a member of a conserved tubulin module and is a component of the centriolar basal foot in multiciliated cells. *Current Biology.* 25(16), 2177-2183

Tuxhorn, J., Daise, T. and Dentler, W. L. (1998) Regulation of the flagellar length in Chlamydomonas. *Cell Motility and the Cytoskeleton.* 40(2), 133-146

Vaughan, S. and Gull, K. (2003) The trypanosome flagellum. *Journal of Cell Science.* 1(116), 757-759

Vaughan, S. and Gull, K. (2016) Basal body structure and cell cycle-dependent biogenesis in *Trypanosoma brucei. Cilia.* [online]. 5(5) [20th June 2019]. Available from: doi.org/10.1186/s13630-016-0023-7

Wasteneys, G. O. (2002) Microtubule organization in the green kingdom: Chaos or self-order? *Journal of Cell Science.* 115(7), 1345-1354

Wei, Q., Ling, K and Hu, J. (2015) The essential roles of transition fibres in the context of cilia. *Current Opinion in Cell Biology.* 35, 98-105

WHO (World health organization) (2019) Trypanosomiasis, human African (sleeping sickness). [online] World Health Organization [8th July 2019]. Available from: https://www.who.int/en/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)

Winey, M. and O'Toole, E. (2014) Centriole structure. *Philosophical Transactions of the Royal Society B: Biological Sciences.* [online]. 369(1650) [20th July 2019]. Available from: doi: 10.1098/rstb.2013.0457

Woods, A., Sherwin, T. *et al* (1989) Definition of individual components within the cytoskeleton of *Trypanosoma brucei* by a library of monoclonal antibodies. *Journal of cell science.* 93(3), 491-500

Woodward, R., Carden, M. J. *et al* (1995) Immunological characterization of cytoskeletal proteins associated with the basal body, axoneme and flagellum attachment zone of *Trypanosoma brucei. Parasitology.* 111(1), 77-85

Wu, J and Akhmanova, A. (2017) Microtubule-organizing centres. *Annual Review of Cell and Developmental Biology.* 33, 51-75

Yang, T. T., Chong, W. M., Wang, W. *et al* (2017) Architecture of mammalian centriole distal appendages accommodates distinct blade and matrix functional elements. *Nature Communications.* [online]. 9(1) e2023 [13th June 2019]. Available from: doi: 10.1038/s41467-018-04469-1

Yu, L., Tanwar, D. K., Emanuel, D. S. *et al* (2019) Grammar of protein domain architectures. *Proceedings of the National Academy of Sciences of the United States of America.* 226(9), 3633-3645

Zhou, Q., Hu, H and Li, Z. (2014) New insights into the molecular mechanisms of mitosis and cytokinesis in Trypanosomes. *International Review of Cell and Molecular Biology.* 308, 127-166