This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Do different African trypanosome species share quorum-sensing signal responses?

Eleanor Silvester
Declaration

I declare that this thesis was written by me, and the work presented was carried out by me, unless otherwise stated in the text. This work has not been submitted for any other degree previously.

Eleanor Silvester
Lay Summary

Animal African Trypanosomiasis is a veterinary disease affecting livestock in Africa that limits productivity and economic development. The disease is caused by single-celled parasites of the species Trypanosoma brucei brucei, Trypanosoma congolense or Trypanosoma vivax. These African trypanosomes replicate within the bloodstream of the mammalian host, and are transmitted between hosts by the blood-feeding tsetse fly.

To ensure its continued survival a parasite population must not overwhelm its host before it can be transmitted. *T. b. brucei* has developed a density-sensing mechanism to facilitate this. The parasite population is maintained in the host by proliferation of slender forms, these forms release a density-sensing signal known as SIF (stumpy induction factor). At high parasite density, SIF accumulates to a level that induces slender forms to develop into stumpy forms. Stumpy forms are growth arrested so limit parasite numbers, prolonging host survival. They have features that allow them to survive and continue development when taken up by feeding tsetse flies, enabling transmission.

*T. congolense* and *T. vivax* do not have slender and stumpy forms, but like *T. b. brucei* they must survive long enough in the host for transmission to occur. The aim of this thesis was to establish whether these parasites control their growth through a density-sensing mechanism comparable to *T. b. brucei*, and if so whether different parasite species could respond to each other’s signals.

In *T. congolense* infections high parasite numbers were associated with a degree of growth arrest. Features of parasites from low and high density *T. congolense* infections were compared to those characteristic of *T. b. brucei* slender and stumpy forms.

The finding that a *T. congolense* protein was able to compensate for the absence of a protein needed in the SIF-response in *T. b. brucei*, indicates that *T. congolense* may be able to detect the *T. b. brucei* density-sensing signal. Additional findings demonstrate that *T. b. brucei* responds to an unidentified signal from a dense *T. congolense* population by becoming more stumpy-like.
Inter-communication between these parasite species is of interest, as this could alter parasite virulence and transmission when co-infections occur.
Abstract

The protozoan parasites *Trypanosoma brucei brucei*, *Trypanosoma congoense* and *Trypanosoma vivax* cause Animal African Trypanosomiasis, a disease responsible for costly livestock pathology and economic losses in Africa. Each of these African trypanosomes are vector-borne and transmitted by the blood-feeding tsetse fly. Additional blood-feeding vectors can spread *T. vivax*, extending its range into South America.

*T. b. brucei* infection of the mammalian host progresses in waves, with periodic clearance of antigenic variants. Accumulation of slender parasites in the blood is accompanied by accumulation of the density-sensing factor, SIF (stumpy induction factor). SIF drives differentiation from the proliferative slender form to the growth arrested stumpy form at the peak of parasitaemia. This differentiation step aids host survival, and the stumpy form is pre-adapted for continuation of development in the tsetse fly, ensuring transmissibility.

Despite facing challenges comparable to *T. b. brucei* during their life cycle, *T. congoense* and *T. vivax* are not found to have morphologically distinguishable slender and stumpy forms. The growth control mechanisms used by these important veterinary pathogens have been investigated in this thesis. Particular focus has been placed on the conservation of quorum sensing pathways within the African Trypanosomes. The potential for cross-species communication has implications for co-infections.

*T. congoense* was found to undergo growth arrest at peak parasitaemia, and transcriptomic changes occurring between ascending and peak parasitaemia were identified and comparisons made to *T. b. brucei* slender and stumpy transcriptomes.

In an examination of the conservation of the SIF-responsive pathway, expression of a *T. congoense* orthologue was found to rescue stumpy formation in an otherwise SIF-resistant null mutant for the corresponding *T. b. brucei* gene.

The capacity for cross-talk between density-sensing signals in different trypanosome species was tested using conditioned medium from *T. congoense* bloodstream form cultures. This could activate the expression of a stumpy specific reporter protein in *T.*
*b. brucei.* A cell line deficient in a SIF-responsive gene showed resistance to the conditioned medium with a delay in reporter expression.

These results highlight the unanticipated capacity for different trypanosome species to exhibit intra and inter specific cell-cell communication in the mammalian bloodstream, with possible consequences for their virulence, transmission and evolution.
Acknowledgements

First and foremost I would like to thank Professor Keith Matthews who has mentored me with a good balance of patience, encouragement and humour throughout my PhD. I am also grateful to Keith for the opportunity to work on a project that has at times been perplexing, but has always been fascinating. I was supported financially during my PhD by a studentship from BBSRC. The Wellcome Trust funds work in Keith Matthews’ laboratory.

Thanks to Julie Young for being such a supportive and helpful lab manager. I am indebted to both Keith and Julie for undertaking the animal handling necessary for this project. I am particularly grateful to Dr Alasdair Ivens for collaborating with me on the RNA-seq analyses, and for taking time to advise me on many aspects of bioinformatics. I would also like to thank Dr Margo Chase-Topping for advice on statistical tests and Dr Martin Waterfall for assistance with flow cytometry. Thanks also to Dr Joanne Thompson, my second supervisor, for taking an interest in this project and for asking many thought-provoking questions.

Thanks to Dr Bill Wickstead and Dr Simon D’Archivio (University of Nottingham) for providing *T. vivax* epimastigotes and for showing me how to keep them alive. Thanks to Dr Annette MacLeod (University of Glasgow) and Dr Liam Morrison (University of Edinburgh) for providing *T. congolense* bloodstream forms. Thanks also to Dr Jane Munday (University of Glasgow) for sharing her knowledge on *T. congolense* bloodstream form culture, and for answering my many questions.

Thanks to all members of the Matthews lab past and present for helping to make my time in the lab both fun and productive. I would particularly like to thank Balázs, Binny, Eva, Federico, Lindsay, Maria, Paula and Rachel who gave me invaluable advice and helped me find my feet at the start of my PhD.

Finally, thanks to Mum, Dad, Chris and Loki for keeping me smiling.
Abbreviations

8-CPT-cAMP  8-(4-chlorophenylthio)-cAMP
AAT  Animal African Trypanosomiasis
ADSL  Adenylosuccinate lyase
AHL  Acyl homoserine lactones
AI-2  Autoinducer-2
AMP  Adenosine monophosphate
AMPK  AMP kinase
ANOVA  Analysis of variance
ATP  Adenosine triphosphate
BARP  Brucei alanine-rich protein
BLAST  Basic local alignment search tool
BRCA1  Breast cancer 1
BSA  Bovine serum albumin
BSF  Bloodstream form
cAMP  Cyclic AMP
CARP  cAMP responsive protein
CAT  Chloramphenicol acetyltransferase
CDD  Conserved domains database
CDK  Cyclin-dependent protein kinase
CESP  Congolense epimastigote specific protein
CF  Counting factor
cGMP  Cyclic guanosine monophosphate
CM  Conditioned medium
CMF  Conditioned medium factor
CRK  cdc2 related kinase
Da  dalton
DABCO  1, 4-diazabicyclo[2.2.2]octane
DAPI  4’, 6-diamidino-2-phenylindole
DIF  Differentiation inducing factor
DIG  Digoxigenin
dKO  double knockout
DNA: Deoxyribonucleic acid
EDTA: Ethylenediaminetetraacetic acid
EF1: Elongation factor 1-alpha
eIF4E: Eukaryotic translation initiation factor 4E
ELISA: Enzyme-linked immunosorbent assay
ELM: Eukaryotic linear motifs database
ES: Expression site
ESAG: Expression site associated gene
FACS: Fluorescence-activated cell sorting
FBS: Fetal bovine serum
FC: Fold change
G6PDH: Glucose-6-phosphate dehydrogenase
GARP: Glutamic acid/alanine-rich protein
gDNA: Genomic DNA
GO: Gene Ontology
GP63: Glycoprotein 63
HAT: Human African Trypanosomiasis
HEPES: 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HRP: Horseradish peroxidase
Hyp2: Hypothetical protein 2
IFA: Immunofluorescence assay
IPTG: Isopropyl β-D-1-thiogalactopyranoside
ISG: Invariant surface glycoprotein
I-TASSER: Iterative-Threading Assessment Refinement
Kb: Kilobase(s)
LB: Luria-Bertani
MAPK: Mitogen-activated protein kinase
MEK: Mitogen-activated protein kinase kinase
MOPS: 3-(N-Morpholino) propanesulfonic acid
MSP-B: Major Surface Protease B
NAD+: Nicotinamide adenine dinucleotide
NADH: NAD+ reduced
NEK  NimA related kinase
OE   Overexpression
PAD  Protein associated with Differentiation
PAG  Procyclin-associated gene
PBS  Phosphate buffered saline
PCR  Polymerase chain reaction
PCV  Packed cell volume
PDB  Protein Data Bank
PDEB1 cAMP-specific phosphodiesterase
PFR  Paraflagellar rod
PKA  Protein kinase A
PP1  Protein Phosphatase 1
PSF  Prestarvation factor
PSSA-2 Procyclic specific surface protein
RB   Retinoblastoma protein
RBP  RNA-binding protein
RNA  Ribonucleic acid
RNAi RNA interference
RNA-seq RNA-sequencing
rRNA ribosomal RNA
SCOP Structural classification of proteins database
SDS Sodium Dodecyl Sulphate
SEM  Standard error of the mean
SIF  Stumpy induction factor
SILAC Stable isotope labelling by amino acids in culture
siRNA small interfering RNA
sKO  single knockout
SSC  Saline sodium citrate
TAE  Tris-acetate-EDTA
TbPIP39 T. b. brucei PTP-interacting protein 39kDa
TbPTP1 T. b. brucei protein tyrosine phosphatase 1
TCA Tricarboxylic acid
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDB</td>
<td>Trypanosome dilution buffer</td>
</tr>
<tr>
<td>TEMED</td>
<td>N,N,N',N'-Tetramethylethylenediamine</td>
</tr>
<tr>
<td>TgIF2</td>
<td>Eukaryotic initiation factor-2 of <em>Toxoplasma gondii</em></td>
</tr>
<tr>
<td>TMHMM</td>
<td>Transmembrane hidden Markov model</td>
</tr>
<tr>
<td>TNFR</td>
<td>Tumour necrosis factor receptor</td>
</tr>
<tr>
<td>TOR4</td>
<td>Target of rapamycin</td>
</tr>
<tr>
<td>TORC1</td>
<td>Target of rapamycin complex 1</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>USD</td>
<td>United States Dollar</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VSG</td>
<td>Variant Surface Glycoprotein</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>WT</td>
<td>wild type</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>YFP</td>
<td>Yellow fluorescent protein</td>
</tr>
<tr>
<td>ZFK</td>
<td>Zinc-finger kinase</td>
</tr>
</tbody>
</table>
Table of contents

Declaration .............................................................................................................................................i
Lay Summary .......................................................................................................................................iii
Abstract ...............................................................................................................................................v
Acknowledgements ...........................................................................................................................vii
Abbreviations .....................................................................................................................................viii
Table of contents ...............................................................................................................................xii

Chapter 1: Introduction .....................................................................................................................1
1.1 Interactions in the field ..................................................................................................................2
  1.1.1 Impact of Animal African Trypanosomiasis .................................................................2
  1.1.2 Pathological aspects of AAT ...........................................................................................3
  1.1.3 Available treatments and drug efficacy ..........................................................................4
  1.1.4 Distribution of different trypanosome species and incidence of co-infection ............5
  1.1.5 Impact of co-infection on infection outcome .................................................................6
  1.1.6 Impact of co-infection on virulence and transmission of another unicellular parasite - Plasmodium .................................................................8
1.2 Pleomorphism and differentiation ............................................................................................10
  1.2.1 Life cycle ..........................................................................................................................10
  1.2.2 Antigenic variation ............................................................................................................13
  1.2.3 Growth dynamics in the mammalian host ....................................................................14
  1.2.4 Monomorphism ...............................................................................................................15
  1.2.5 Cell cycle arrest in stumpy forms ..................................................................................16
  1.2.6 Quiescence in diverse organisms ................................................................................17
  1.2.7 Markers associated with pleomorphism ......................................................................18
  1.2.8 Regulation of gene expression ......................................................................................18
  1.2.9 Quiescence in stumpy forms and gene expression changes during differentiation ....19
  1.2.10 Pleomorphism in T. congolense and T. vivax ............................................................21
1.3 Quorum sensing and environmental sensing ............................................................................24
  1.3.1 Environmental sensing in a simple eukaryote - yeast ..............................................24
  1.3.2 Environmental sensing in eukaryotic parasites ..........................................................25
  1.3.3 Prokaryotic quorum sensors .......................................................................................26
  1.3.4 Eukaryotic quorum sensors ..........................................................................................27
1.3.5 Density dependent control in eukaryotic parasites .................................. 28
1.3.6 Density dependent control of *T. b. brucei* growth: SIF ...................... 29
1.3.7 Regulators of stumpy formation .............................................................. 29
1.4 Aims .............................................................................................................. 32

**Chapter 2: Materials and Methods** .............................................................. 35

2.1 Trypanosome strains .................................................................................. 36
2.2 Trypanosome maintenance *in vitro* .......................................................... 36
2.3 Cloning and DNA manipulations ............................................................... 38
  2.3.1 Extraction of genomic DNA ................................................................. 38
  2.3.2 Polymerase chain reaction (PCR) ......................................................... 38
  2.3.3 Agarose gel electrophoresis ................................................................. 40
  2.3.4 Purification of PCR products ............................................................... 40
  2.3.5 Ligation .................................................................................................. 41
  2.3.6 Transformation of competent cells ....................................................... 41
  2.3.7 Small scale plasmid preparation (Miniprep) ......................................... 42
  2.3.8 Restriction digestions ........................................................................... 42
  2.3.9 Sequencing ............................................................................................ 43
  2.3.10 Preparation of DNA for transfection ................................................ 43
2.4 Transfection of bloodstream form trypanosomes ..................................... 44
2.5 Monitoring growth of cell lines *in vitro* .................................................. 45
  2.5.1 Testing induction of RNAi or overexpression *in vitro* ......................... 45
  2.5.2 Testing response to 8-CPT-cAMP *in vitro* ........................................ 45
  2.5.3 Cell cycle analysis *in vitro* .................................................................. 45
  2.5.4 Protein localisation *in vitro* ............................................................... 46
2.6 Conditioned medium generation .............................................................. 47
2.7 CAT reporter assay .................................................................................... 48
  2.7.1 Sample preparation ............................................................................. 48
  2.7.2 CAT ELISA .......................................................................................... 49
2.8 Trypanosome infections ............................................................................ 49
2.9 Monitoring growth of cell lines *in vivo* .................................................... 50
  2.9.1 Cell cycle analysis .............................................................................. 50
  2.9.2 Immunofluorescence analysis of protein expression ............................ 51
2.10 Induction of differentiation to procyclic forms *in vitro* ............................ 51
Chapter 3: Characterisation of *T. congolense* bloodstream infection .......... 61

3.1 Introduction ............................................................................................................. 62
3.2 *T. congolense* bloodstream forms display evidence of growth control at the peak of parasitaemia ........................................................................................................... 63
3.3 Preparation of RNA samples from ascending and peak *T. congolense* parasitaemia for RNA-seq analysis ................................................................................. 68
3.4 Analysis of *T. congolense* ascending and peak parasitaemia RNA-seq data 69
   3.4.1 Alignment to the *T. congolense* genome and analysis of differential expression ................................................................. 69
   3.4.2 Identification of *T. b. brucei* orthologues of the *T. congolense* genes ....... 74
   3.4.3 Predicted protein domains in hypothetical proteins that are more abundant at peak parasitaemia .......................................................... 76
3.5 Preparation of RNA from slender and stumpy populations of *T. b. brucei* for RNA-seq analysis ............................................................................................................. 78
3.6 Analysis of *T. b. brucei* slender and stumpy RNA-seq data with comparison to *T. congolense* RNA-seq data ................................................................................ 82
   3.6.1 Alignment to the *T. b. brucei* genome and analysis of differential expression ................................................................. 82
   3.6.2 Comparison to *T. congolense* RNA-seq data ............................................. 89
   3.6.3 Predicted protein domains of hypothetical proteins with transcripts more abundant in stumpy forms .......................................................... 90
   3.6.4 *T. congolense* orthologues of *T. b. brucei* proteins with transcripts increased in abundance in stumpy forms ............................................. 92
3.7 Analysis of RNA-seq data by orthologous group ................................................. 93
3.8 Conclusions and future directions ......................................................................... 102

Chapter 4: Conservation of components of the *T. b. brucei* SIF response pathway in *T. congolense* and *T. vivax* ......................................................... 105

4.1 Introduction ............................................................................................................. 106
4.2 Identification of *T. congolense* and *T. vivax* orthologues of SIF-responsive genes ................................................................................................................ 107
4.3 Strategy to test capacity of identified orthologues for functional complementation .......................................................... 114
4.4 Selection of identified orthologues for functional complementation .... 115
  4.4.1 Adenylosuccinate lyase (Tb927.9.7550) ................................. 115
  4.4.2 Hypothetical protein 2 (Tb927.9.4080) ................................. 116
4.5 Bioinformatics analysis of Hypothetical protein 2 (Tb927.9.4080/ TbHyp2) 118
4.6 Cloning strategy for overexpression ........................................... 126
4.7 Validation of functional complementation strategy .......................... 128
4.8 Characterization of Hyp2 overexpression in vitro ............................ 130
  4.8.1 Growth arrest caused by overexpression ................................. 130
  4.8.2 Localization ........................................................................ 132
  4.8.3 Overexpression associated with increased kinetoplast-posterior distance 136
  4.8.4 Effect of Overexpression on responsiveness to 8-CPT-cAMP ......... 141
  4.8.5 Knock out of TbHyp2 ................................................................ 143
4.9 Characterisation of Hyp2 overexpression in vivo ............................ 147
  4.9.1 TbHyp2 dKO does not become stumpy or control parasitaemia .... 147
  4.9.2 Is Hyp2 orthologue overexpression able to rescue stumpy formation in vivo? 149
  4.9.3 Effect of TbHyp2 overexpression in TbHyp2 dKO ...................... 152
  4.9.4 Effect of TvHyp2 overexpression in TbHyp2 dKO ...................... 157
  4.9.5 Effect of TcHyp2 overexpression in TbHyp2 dKO ...................... 160
4.10 Conclusions and future directions ............................................... 162

Chapter 5: How do T. congolense and T. b. brucei interact in co-infections? Is there potential for cross-talk? .................................................. 165
  5.1 Introduction .............................................................................. 166
  5.2 Validation of T. b. brucei conditioned medium effect ..................... 167
  5.3 Optimising T. congolense conditioned medium test conditions ....... 169
  5.4 Pleomorphic and monomorphic cell lines differ in their response to T. congolense conditioned medium ..................................................... 171
  5.5 Variable sensitivity of monomorphic cell lines to T. congolense conditioned medium ............................................................ 175
  5.6 Is sensitivity to T. congolense conditioned medium lessened in cells depleted for a SIF-responsive gene? ............................................. 179
5.7 Testing *T. congolense* conditioned medium effect on expression of a stumpy reporter ................................................................. 183

5.8 Does a cell line depleted of a SIF-response gene produce less of a stumpy reporter when treated with conditioned medium? ........................................ 192

5.8.1 Validation of the new reporter line *in vitro* ........................................ 193

5.8.2 Validation of the new reporter line *in vivo* ........................................ 195

5.8.3 Effect of *T. congolense* conditioned medium on the TbHyp2 RNAi CAT-PAD reporter line ............................................................ 196

5.9 Co-infection ......................................................................................... 203

5.9.1 Simultaneous infection with *T. b. brucei* and *T. congolense* ............ 203

5.9.2 *T. congolense* infection followed by *T. b. brucei* challenge ............ 210

5.9.3 *T. congolense* infection followed by *T. b. brucei* challenge in triplicate 215

5.10 Conclusions and future directions .................................................... 219

Chapter 6: Discussion ............................................................................. 225

Bibliography .............................................................................................. 245

Appendix 1: Supplementary information on *T. congolense* and *T. b. brucei* RNA-seq analysis ............................................................. 263

Appendix 2: Supplementary Information on Hypothetical Protein 2 and orthologues ............................................................... 279

Appendix 3: Supplementary information on *T. congolense* and *T. b. brucei* interactions *in vitro* and *in vivo* ........................................ 291

Appendix 4: Buffers and reagents ............................................................. 297

4.1 *T. congolense* TcBSF3 media .............................................................. 298

4.2 HMI-9 Media ...................................................................................... 299

4.3 TV3 media ......................................................................................... 299

4.4 TE solution ......................................................................................... 300

4.5 TELT (lysis buffer) ............................................................................. 300

4.6 6x DNA loading buffer ....................................................................... 300

4.7 TAE buffer ......................................................................................... 300

4.8 Miniprep solution I ........................................................................... 300

4.9 Miniprep solution II .......................................................................... 301

4.10 Miniprep solution III ......................................................................... 301

4.11 Trypanosome Dilution Buffer (TDB) ................................................ 301

4.12 Phosphate buffered saline (PBS) ....................................................... 301
4.13 SDM-79 Media ................................................................. 302
4.14 FACS fix ................................................................. 302
4.15 Laemmli buffer (2x) .................................................. 302
4.16 Solutions for protein gels ........................................... 303
  4.16.1 Separating solution (10ml) ..................................... 303
  4.16.2 Stacking solution (5ml) .......................................... 303
4.17 Tris-glycine electrophoresis buffer ............................... 303
4.18 Western Blot transfer buffer ....................................... 304
4.19 Western Blot stripping buffer ...................................... 304
4.20 10X MOPS ............................................................... 304
4.21 RNA gel loading buffer .............................................. 304
4.22 20xSSC ................................................................. 304
4.23 Hybridisation buffer .................................................. 305
4.24 Maleic Acid Buffer ................................................... 305
4.25 10% DIG block ........................................................ 305
4.26 Wash buffer (for Northern blot) ................................. 305
4.27 Detection buffer ........................................................ 305
Chapter 1: Introduction
1.1 Interactions in the field

1.1.1 Impact of Animal African Trypanosomiasis

Disease caused by salivarian trypanosomes has been a persistent burden for both health and economic development in afflicted regions. Alongside the direct impact of Human African trypanosomiasis (HAT) (caused by *T. b. gambiense* and *T. b. rhodesiense*) on human health, there is an additional cost associated with Animal African Trypanosomiasis (AAT) that affects a vast area. African trypanosomiasis infections often lead to anaemia and a progressive wasting disease in susceptible cattle and if untreated can be fatal. This has an enormous impact on agriculture, through reduced animal fertility, losses in meat and milk production, draught power (effect on arable farming), and less animal protein available to the workforce. Some areas of Sub-Saharan Africa are completely incompatible with keeping certain trypanosome susceptible livestock due to the prevalence of trypanosome-transmitting tsetse (Figure 1.1), holding back development in these areas (reviewed in (Shaw, 2004)). Trypanosomiasis causes losses of approximately 1.5 billion USD annually and afflicts 21 out of 25 of the world’s poorest countries (FAO, 2008).

![Figure 1.1 Tsetse fly infestation as a limitation to cattle distribution. A) Global distribution of cattle: Adapted from (Robinson et al., 2014). B) Predicted distribution of tsetse: Adapted from (Wint and Rogers, 2000). Predicted presence assumed when predicted probability of presence ≥50%. Large areas of tsetse infestation are mostly devoid of cattle.](image-url)
1.1.2 Pathological aspects of AAT

Anaemia is associated with the early acute phase of *T. congolense* or *T. vivax* infection, but may persist in the chronic phase in the case of susceptible cattle. Erythrocyte destruction by phagocytosis contributes to anaemia during the acute and chronic phases. This could result from binding of host autoantibodies or trypanosome antigen immune complexes to the erythrocyte surface or damage of erythrocytes by trypanosome enzymes. Ineffective erythropoiesis may contribute to anaemia in the chronic phase. East African isolates of *T. vivax* can cause a hyperacute disease with severe anaemia and haemorrhages. This has been linked to disseminated intravascular coagulation (DIC), which may result from products released by the destruction of platelets during infection. The mechanisms of platelet damage in trypanosome infection may be similar to those proposed for erythrocyte destruction (Taylor and Authie, 2004).

Products released by African trypanosomes into the bloodstream have been associated with pathology and can be immunogenic. Congopain, a cysteine protease of *T. congolense*, is likely to play a role in pathology as prominent IgG responses to congopain have been observed in trypanotolerant, but not trypanosusceptible, cattle (Authie, 1994). Trans-sialidases released by *T. vivax* were demonstrated to contribute to anaemia by desialylating erythrocytes resulting in erythrophagocytosis (Guegan et al., 2013). Trans-sialidases released by *T. congolense* were proposed to act in a similar manner, contributing to pathogenesis (Coustou et al., 2012).

The pathology observed in many instances may relate to the tissues that the parasites occupy. For example, *T. congolense* parasites were observed in the microvasculature alongside degenerative lesions in the pituitaries of cattle (Abebe et al., 1993). Reproductive disorders observed in AAT may in part result from damage to the pituitary gland. Recently, *T. b. brucei* mouse infections have revealed a significant reservoir of parasites in the adipose tissue (Trindade et al., 2016). These parasites had adapted their metabolism to the adipose tissue environment, and it is possible that parasites in the adipose tissue are contributing to the weight loss observed in AAT.
1.1.3 Available treatments and drug efficacy

Drug resistance is becoming an increasing problem in the treatment of AAT (reviewed in (Holmes et al., 2004)), with relief in the form of new compounds a distant prospect (reviewed in (Barrett et al., 2004)). Treatment of AAT is less regulated than treatment of the human disease, and there is concern that selection of resistance to Berenil (diminazene aceturate), used to treat livestock, could provide cross-resistance to melarsoprol used to treat late stage human disease (Barrett, 2001; Barrett and Fairlamb, 1999).

Treatment with frequently used drugs, isometamidium chloride and diminazene aceturate, was shown to be ineffective at preventing relapse in greater than 20% of cases when naïve Zebu cattle were treated with recommended doses of these trypanocidal drugs following an initial challenge with *T. vivax* (Dagnachew et al., 2015). This study used field isolates of *T. vivax* from North West Ethiopia, from areas of tsetse transmission or non-tsetse (mechanical) transmission. In both cases relapse was observed, indicative of problematic drug resistance occurring in this study area. Resistance to diminazene aceturate was also detected in *T. congolense* from cattle sampled in Ethiopia (Moti et al., 2012). Drug resistance has been detected in at least 17 African countries, with in some instances cases of multiple drug resistance (reviewed in (Delespaux et al., 2008)). Drug resistance is not new, with an early study showing that strains of *T. vivax* and *T. congolense* retained their drug resistance with cyclical transmission through tsetse flies (Gray and Roberts, 1971).

Even in the absence of drug resistance, recurrent reinfection means that keeping their animals healthy becomes costly for the livestock holder. A vaccination approach could offer a long-term solution. However, the elaborate immune evasion methods of the parasite, which include antigenic variation through variant surface glycoprotein (VSG) switching (reviewed in (McCulloch, 2004)) and rapid clearance of surface bound antibody (Engstler et al., 2007), make vaccine development a challenging prospect. Therefore, novel approaches must be taken to find aspects of the parasite’s biology that could be targeted therapeutically.
1.1.4 Distribution of different trypanosome species and incidence of co-infection

Animal African Trypanosomiasis is predominantly caused by *T. congolense* and *T. vivax*, and to a lesser extent *T. b. brucei*. For instance, in a meta-analysis of the prevalence of bovine trypanosomiasis in Ethiopia *T. congolense* and *T. vivax* were found to be the principal causes (Leta et al., 2016). Simplistically *T. vivax* infections are the most prevalent in Western Africa, whereas *T. congolense* prevails in Eastern and Central Africa, however severe outbreaks of haemorrhagic *T. vivax* are known to occur in East Africa. The progression of *T. congolense* infection is more typically chronic than the often acute *T. vivax*, although the consequences of both infections are frequently severe (reviewed in (Eisler et al., 2004)).

Mixed infections between *T. b. brucei*, *T. congolense* and *T. vivax* occur in the field. One study of Animal African trypanosomiasis spatial distribution in two districts of Western Kenya detected dual infections of *T. b. brucei/T. congolense*, *T. vivax/T. b. brucei* and *T. vivax/T. congolense* (Thumbi et al., 2010). *T. vivax* was the predominantly infecting trypanosome in this study, and was also detected in tsetse-free areas, likely as a result of mechanical transmission.

In another study (Cox et al., 2010) sampling indigenous African Zebu cattle in Uganda, 14 out of 35 cattle observed were found to have coinfections involving *T. vivax* and *T. congolense*, where there were either *T. vivax-T. congolense*, *T. vivax-T. brucei*, *T. congolense-T. brucei*, or a mix of all 3 species. There were also some coinfections involving *T. congolense* or *T. vivax* with *T. theileri* (a non-pathogenic species).

A recent field study of 8 sites in Tanzania sampled cattle, sheep and goats (Ruiz et al., 2015). Mostly infections comprised *T. b. brucei* only, with a low proportion testing positive for *T. b. rhodesiense*. Nevertheless, instances of *T. b. brucei* and *T. congolense* co-infection were detected in cattle.

In a study characterising the infected wildlife reservoir in the Luangwa Valley, Zambia, mixed infections (*T. b. brucei/T. congolense*, *T. b. brucei/T. vivax*) accounted for 6.9% of total infections (Anderson et al., 2011). The authors speculate
that the sensitivity of detection in this study was not particularly high, so could underestimate prevalence of mixed infections.

The establishment in a mammalian host of a mixed infection comprising two different species of African trypanosomes can occur by two mechanisms. The mammal may be infected by a single tsetse bite delivering metacyclic forms from two species. Alternatively, two separate tsetse bites deliver the mixed infection, with one species of metacyclic trypanosomes transmitted by the first bite followed by a second superinfecting bite with a different trypanosome species. There is experimental evidence for the ability of *T. b. brucei* and *T. congolense* from mixed cattle infections to concurrently establish mature metacyclic infections in the same fly (Van den Bossche et al., 2004). Furthermore, tsetse flies already infected by one trypanosome species, remain as susceptible to infection by a second trypanosome species as uninfected flies of the same age (Kubi et al., 2005). Thus, mixed mature tsetse fly infections may be established by one or two infective bloodmeals. Detection of mature mixed infections in tsetse flies has been recorded in field studies (Jamonneau et al., 2004; Majiwa and Otieno, 1990; Masiga et al., 1996; Morlais et al., 1998). Thus simultaneous infection by two trypanosome species from one infective tsetse fly bite is a possibility. However, considering the chronicity of African trypanosome infections, there is a greater probability of mixed infections arising from multiple tsetse fly bites on a chronically infected host resulting in superinfection.

1.1.5 Impact of co-infection on infection outcome

One study looked at the prevalence of trypanosome infection (grouped by species) in horses and donkeys, increasingly used for agriculture and transportation, in The Gambia (Pinchbeck et al., 2008). The overall detected trypanosome prevalence was 91% (infection rate 31% for *T. congolense* savannah, 87% for *T. vivax* and 18% for *T. b. brucei*). Multiple species were present in 43% of infections. *T. congolense* savannah appeared to cause the greatest negative effect on the clinical indicator packed cell volume (PCV). Notably, co-infection with *T. vivax* caused less of an effect on PCV than *T. congolense* infection alone, indicative of an interaction, possibly competition, affecting disease morbidity.
Van den Bossche (Van Den Bossche et al., 2005), found that upon challenge with *T. congolense*, cattle with low level *T. b. brucei* infection developed an increased *T. b. brucei* parasitaemia. This could be due to the immune response being diverted towards the newly invading pathogen, allowing the *T. b. brucei* population to expand. An alternative explanation could be a direct communication between the cohabiting parasites.

Further studies in cattle (Morrison et al., 1982) demonstrated that active infection with one serodeme of *T. congolense* could inhibit establishment of infection with a second unrelated serodeme. This interference did not seem to be a result of specific immunity, but could be the result of some form of competition between the trypanosome populations.

In another study of superinfection (Dwinger et al., 1989), goats infected with *T. congolense* were later challenged with either *T. b. brucei* or *T. vivax*. An established *T. congolense* infection delayed appearance of *T. b. brucei* parasites and specific antibodies in the blood. In contrast, no significant delay in establishment was observed for *T. vivax* superinfection in goats pre-infected with *T. congolense*.

Interactions between co-infecting strains of *T. b. brucei* have also been described. Multiple strain *T. b. brucei* infections occur across Africa and are thought to account for around 10% of *T. b. brucei* infections (Balmer and Caccone, 2008). One study found that a mixed infection comprising an avirulent strain and a virulent strain of *T. b. brucei* reduced the deleterious effects on the host relative to an infection with the virulent strain alone (Balmer et al., 2009). Additionally, there was a reduced total parasitaemia in the mixed infection relative to the single virulent strain infection indicating that the inhibition experienced by both strains was not a result of reaching maximum population size capacity. It was suggested that the observed interference could be immune mediated or the result of parasite-released factors having a direct effect on competitors (allelopathy).

Whatever the causative mechanism, different African trypanosome species and strains are capable of impacting each other’s growth dynamics and the pathogenic outcome when found together in co-infection. If interspecies competition in zoonotic
reservoirs was to select for more virulent strains of human infective *T. b. gambiense* or *T. b. rhodesiense* (discussed in (Matthews et al., 2015)), the consequences for human health could be severe.

The potential for interaction between trypanosome species does not stop at the mammalian host. A study of experimentally infected *Glossina morsitans morsitans* indicated that competition existed between isolates of *T. b. brucei* and *T. congolense* savannah in the fly midgut (Ahmed et al., 2015). Both species were capable of establishing similar numbers in the midgut individually. However, following simultaneous challenge, whereas *T. b. brucei* numbers were largely unaffected, *T. congolense* numbers were significantly lower. If flies were fed *T. b. brucei* first, then *T. congolense* was unable to establish. On the other hand, *T. b. brucei* could establish a small population when fed after *T. congolense*.

### 1.1.6 Impact of co-infection on virulence and transmission of another unicellular parasite - *Plasmodium*

Numerous investigations have probed the impact of co-infection on different strains and species of *Plasmodium* parasites, and these studies may be used to draw parallels with and inform on the interactions between African trypanosomes in co-infections.

As for *Trypanosoma* species, the life cycle of *Plasmodium* parasites comprises different specialised forms adapted to the different environments encountered in the vertebrate host and during transmission. During its erythrocytic cycle, the parasite undergoes asexual replication in the red blood cells of the host. During this phase, the parasite may choose to invest a proportion of its resources in transmission by production of gametocytes. The degree of investment in gametocytes is plastic and may vary when a parasite is placed under different degrees of stress. For instance, investment in gametocytes may be affected by the immune status of the host, the extent of host anaemia or the presence of co-infecting parasites of different species or strains (reviewed in (Reece et al., 2009)). When competition results in low levels of stress, parasites may respond by reducing costly investment in gametocyte production to better invest in asexual replication. Alternatively, when survival in the host is threatened, parasites may switch to a terminal investment strategy by increasing gametocyte production. Different *Plasmodium* species may switch to a
terminal investment strategy at different levels of stress depending on their life cycle characteristics, for instance *P. falciparum*, whose gametocytes require longer for maturation than mouse-infective species of *Plasmodium*, may switch to investment in transmission at lower levels of stress (Carter et al., 2013). Furthermore it was demonstrated that *P. malariae* co-infection could increase *P. falciparum* gametocyte production (Bousema et al., 2008).

One study of co-infection using the *P. chabaudi* model has demonstrated that competitive success of a parasite strain in a mixed infection could be related to the genotype of the parasite, the order in which the two strains were inoculated, and the length of time between the initial infection and superinfection with a second strain. The authors suggested that cross-strain immune responses, as well as competition for the limited resources within the blood could explain the competitive suppression observed in this study (de Roode et al., 2005).

Another *P. chabaudi* study, which featured two *P. chabaudi* strains of different virulences, indicated that more virulent parasites would have greater transmission success in co-infection than less virulent strains (Wargo et al., 2007). Furthermore, when genetically divergent *Plasmodium chabaudi* clones of different virulences were compared pair-wise, it was observed that the most virulent parasite had a competitive advantage during the acute phase of parasitaemia. Therefore, it was concluded that mixed infections favoured selection of more virulent parasites, and interventions aimed at decreasing the number of mixed infections could help to reduce virulence (Bell et al., 2006). Alternatively, it has been theorised that co-infection could result in reduced virulence as a result of unrelated parasites directly interfering with each other during a co-infection; such interference could have an associated energy cost, and this could reduce the extent to which both species are able to exploit the host, resulting in reduced virulence (Chao et al., 2000).
1.2 Pleomorphism and differentiation

1.2.1 Life cycle

The lifecycles of *T. b. brucei*, *T. vivax* and *T. congolense* are broadly similar, with mammalian bloodstream form stages, as well as insect form stages that develop in the shared tsetse fly vector. However, there are some important differences (Figure 1.2).
Figure 1.2 Life cycles of the African trypanosomes. Following a blood meal on an infected mammal, development of *T. b. brucei* and *T. congolense* proceeds through the tsetse midgut and proventriculus. Differentiation to epimastigote forms is followed by production of mammalian-infective metacyclic forms from the salivary gland in *T. b. brucei* and from the mouthparts in *T. congolense*. On transfer to a new mammalian host via a tsetse fly bite, metacyclic forms differentiate to bloodstream forms. In *T. b. brucei* infection, the parasitaemia is composed of slender and stumpy form parasites, and it is the stumpy forms that are ready for transmission to a feeding tsetse fly. Slender and stumpy forms are not found in *T. congolense* infections. *T. vivax* has a comparatively simpler life cycle. On uptake by a feeding tsetse fly, *T. vivax* develops exclusively in the tsetse mouthparts, with differentiation to epimastigote forms followed by production of mammalian infective metacyclic forms. On transfer to a new mammalian host, differentiation of metacyclic to bloodstream forms occurs and (as for *T. congolense*) no slender and stumpy forms are found.
*T. b. brucei* and *T. congolense* have distributions largely limited by their tsetse fly vector. In contrast, biting insects other than tsetse flies can mechanically transmit *T. vivax*. This has resulted in the spread of *T. vivax* to South America where transmission is maintained in the absence of tsetse flies (reviewed in (Osório et al., 2008)). Mechanical transmission is also thought to occur for *T. b. brucei* and *T. congolense*, although not to the same extent as *T. vivax* (Hoare, 1972).

There are further key differences between *T. vivax*, *T. congolense* and *T. b. brucei* life styles during infection of a mammalian host. Whereas, *T. congolense* resides entirely within blood vessels, *T. b. brucei* can invade beyond the bloodstream into tissues and in late stages of infection can be found in the brain parenchyma (Masocha et al., 2007). Recently, *T. vivax* has been detected in the brain of infected mice, although it is uncertain as to whether parasites are intra or extra vascular (D'Archivio et al., 2013). *T. b. brucei* has also been found to accumulate in the adipose tissue of mice (Trindade et al., 2016). *T. congolense* has a tendency to adhere to blood vessel walls and erythrocytes, in addition to swimming free in the blood (Banks, 1978, 1979). A recent study of the motility of different African trypanosome species revealed that differences in motility between *T. b. brucei*, *T. congolense* and *T. vivax* may be attributable to the different anatomical niches occupied by these parasites (Bargul et al., 2016). *T. vivax* was found to be specialised for fast swimming in the peripheral bloodstream, but not for invasion of tissue spaces. While *T. b. brucei* was found to swim most efficiently at higher viscosities than that of blood consistent with this species’ predilection for tissue invasion. On the other hand, *T. congolense* was found to be the least motile of the species examined, and it was considered that excessive motility would prevent the observed attachments to host blood vessels.

In addition to the variation in tissue tropism in the host, there is also variation in the dynamics of growth of the three parasite species within the blood, in terms of; antigenic variation, morphology, presence of growth-arrested forms and pre-adaptation for transmission (see sections 1.2.2, 1.2.3, 1.2.4, 1.2.5, 1.2.7, 1.2.9 and 1.2.10).
1.2.2 Antigenic variation

*T. b. brucei* is able to evade the host immune response by periodically changing its surface coat. The dense layer of variant surface glycoprotein (VSG) covering the parasite surface prevents adherence of host antibody to non-variant surface antigens (Cross, 1975; Vickerman, 1969b). The parasitaemia progresses in waves with one predominant VSG type succeeded by a new VSG (Vickerman, 1985), which is initially unrecognised by the host. The extensive repertoire of VSG is found in telomeric expression sites (ES) (De Lange and Borst, 1982), accompanied by expression site associated genes (ESAG) (Cully et al., 1985), in subtelomeric arrays on large and intermediate sized chromosomes, and on minichromosomes (Williams et al., 1982). Expression is limited to one ES per cell; transcription of this active ES by PolII is localised to the expression site body (Navarro and Gull, 2001). VSG switching may occur by changing the active ES, or by homologous recombination to insert a single complete or mosaic of stored VSG into the active ES (reviewed in (Horn, 2014)). A rich selection of mosaic gene sequences were detected during chronic infection of mice (Hall et al., 2013; Mugnier et al., 2015). This antigenic diversity contributes to the chronicity of infection.

*T. b. brucei*, *T. congolense* and *T. vivax* differ in their mechanisms of generating antigenic diversity. *T. congolense* uses VSG derived from multiple ancestral VSG lineages, whereas *T. b. brucei* VSG has recent origins, with ancestral VSG lineages repurposed to novel functions. There is a comparable recombination frequency in *T. b. brucei* and *T. congolense*, but this is much lower in *T. vivax* (Jackson et al., 2012). A study of the African trypanosome cell-surface phylome presents evidence that most ESAGs are *T. b. brucei* specific (Jackson et al., 2013).

Recent RNA-seq analysis of *T. vivax* developmental stages (Jackson et al., 2015) found a comparable VSG-bloodstream expression pattern to *T. b. brucei* with 1 or 2 dominant VSGs at a time. This study also found unique *T. vivax* surface proteins differentially expressed through its life cycle, which could explain the less dense VSG coat than *T. b. brucei*. 
1.2.3 Growth dynamics in the mammalian host

In the blood of a mammalian host, *T. b. brucei* metacyclic forms develop into long slender forms. As the parasite population density increases, slender parasites transition through intermediate forms into short stumpy forms, which are poised for transmission to the tsetse fly vector (Figure 1.3).

![Diagram showing growth dynamics](image)

**Figure 1.3** Stumpy forms accumulate with increasing parasitaemia, and are prepared for transmission to the tsetse fly vector (Rico et al., 2013)

Stumpy forms are arrested in the G0/G1 stage of the cell cycle (Shapiro et al., 1984) and the slender-stumpy differentiation step forms an important component in the control of parasite growth (reviewed in (MacGregor et al., 2012)).

Differentiation to stumpy forms does not require an intact immune response from the host (Balber, 1972; Seed and Sechelski, 1988). However, both density-dependent differentiation and the host immune response give the best predictive capacity to model the growth dynamics of *T. b. brucei* in the mammalian host (Tyler et al., 2001). When chronic mouse infections were monitored for the proportions of the different cell types, early stages of infection were predominated by slender forms, and in later stages stumpy and intermediate forms were shown to dominate (MacGregor et al., 2011). This profile of parasite differentiation means that, on initiating an infection in the field, the high level of antigen switching by slender forms of the parasite (Turner and Barry, 1989) would ‘probe’ the host immune status.
in order to overcome any pre-existing herd immunity to the circulating parasite populations. Once the infection has established, high levels of stumpy forms (which are quiescent and do not undergo antigen switching) prevents the parasitaemia from becoming too high and killing the host, and restraints use of the parasite’s antigen repertoire, preventing the host from being overwhelmed by too many antigen types. High levels of transmissible stumpy forms may also be necessary in the field because of the poor vectoral capacity of the tsetse fly vector. On uptake to the vector, slender cells are killed and only stumpy forms survive to differentiate to the next life cycle stage (the procyclic form) (Turner et al., 1988). Adaptations that aid survival of the stumpy form in the tsetse fly midgut include increased resistance to acidic and proteolytic stress (Nolan et al., 2000).

Mathematical models have been applied to slender to stumpy differentiation dynamics. A short stumpy survival time of 24 hours was proposed (Seed and Black, 1999), this was predicted to be roughly twice as long based on a study of chronic infection (MacGregor et al., 2011). It has been suggested that the stumpy population consists of 2 parts, one competent for transmission, and the other part an apoptotic-like, altruistic component, with such cells helping to stimulate an immune response against the dominant antigen on slender and stumpy forms to aid in chronic relapsing parasitaemia, and maintain infection without overwhelming the host (Seed and Wenck, 2003). This suggestion is controversial however, and it seems unlikely that the dying stumpy forms have evolved to be beneficial for the parasite's overall survival.

1.2.4 Monomorphism

Multiple syringe-passage of *T. b. brucei* results in monomorphism, while tsetse-passaged parasites retain pleomorphism (Ashcroft, 1960). Monomorphic species can also be found in the field. *T. evansi* and *T. equiperdum* are morphologically comparable to slender forms of *T. b. brucei* and both cause diseases of veterinary importance. *T. evansi* is transmitted mechanically by biting insects to numerous working animals including donkeys and camels. *T. equiperdum* can be transmitted during coitus between horses or mechanically by a vector. Both infections are widespread, impacting Asia, Africa and South America (Brun et al., 1998).
1.2.5 Cell cycle arrest in stumpy forms

The cell cycle of *T. b. brucei* involves faithful duplication of the flagellum, nucleus and mitochondrial genome (kinetoplast), as well as other cell components. For a summary of the major events see Figure 1.4.

![Figure 1.4 Cell cycle of *T. b. brucei* (Sherwin and Gull, 1989). Initial elongation of a daughter flagellum begins from the posterior basal body. Generation of two new probasal bodies is followed by extension of the daughter flagellum outside of the flagellar pocket, including formation of the paraflagellar rod. Division of the kinetoplast and separation of the kinetoplast and basal bodies follows. Mitosis is finally followed by cytokinesis to generate two identical daughter cells.](image)

Evidence from analysis of the capacity to differentiate to procyclic forms inferred that there was a point in the cell cycle (G1/G0) when cells were receptive of the signal to differentiate and re-enter the cell cycle (Matthews and Gull, 1994). Stumpy forms already arrested at this point were therefore able to differentiate synchronously, unlike asynchronous slender forms (Ziegelbauer et al., 1990). The G1 phase is thus an important cross-roads where cells may decide to continue progression through the cell cycle, or to quiesce and differentiate. *T. b. brucei* has numerous cyclins and at least two cdc2 related kinases (CRKs) regulating the G1/S phase transition, this is more comparable to the complexity of control seen in animal cell cycles than to the relatively simpler yeast model (reviewed in (Li, 2012)). The
generation of a kinome-wide RNAi library in *T. b. brucei* resulted in identification of novel cell cycle regulators, as well as two kinases involved in repressing bloodstream form to procyclic form differentiation (Jones et al., 2014).

**1.2.6 Quiescence in diverse organisms**

Yeast cells are known to survive environmental stresses by entering into a quiescent state which is characterised by arrest in G1 of the cell cycle, suppressed translational activity and changes in metabolism (De Virgilio, 2012). Mammalian cells such as lymphocytes also have a quiescent state that features arrest in G1 of the cell cycle, and transcriptional changes that include expression of inhibitors of cyclin dependent kinases, and programmes to inhibit senescence (Coller, 2011).

*T. b. brucei* is not unique in the presence of quiescent forms at life cycle transition stages. *Plasmodium* gametocytes, like stumpy forms, are growth-arrested forms responsible for transmission to the insect vector (Peatey et al., 2013). Indeed, the resistance of *Plasmodium falciparum* to the antimalarial artemisinin was even attributed to cell cycle arrest of ring stage parasites, with such parasites able to resume development following removal of drug pressure (Witkowski et al., 2010).

Similarly, *Toxoplasma gondii* switches from a proliferative form (tachyzoite) to a comparatively quiescent form (bradyzoite); the transition can be triggered by various stressful stimuli. Phosphorylation of eukaryotic initiation factor-2 (TgIF2) results in global translational repression, and this has been linked to conversion to the bradyzoite form and maintenance of latency (Narasimhan et al., 2008). The bradyzoite differentiation programme was thought to initiate from late S/G2 phase of the cell cycle, with differentiating parasites proceeding through M phase before arresting in G0/G1. Mature bradyzoites had been observed to have a uniform 1N DNA content (Radke et al., 2003). However, a recent *in vivo* study has observed episodic replication of bradyzoites within tissue cysts (Watts et al., 2015), calling into question the nature of bradyzoites in a chronic infection (Kim, 2015). Nevertheless, the relative dormancy of the bradyzoite cyst form can prevent the parasite from overwhelming and killing the host before transmission. Furthermore,
the bradyzoite tissue cysts remain viable when the current host is consumed by a predator enabling transmission (reviewed in (Sullivan et al., 2009)).

1.2.7 Markers associated with pleomorphism

The slender forms of *T. b. brucei* rely simply on glycolysis, with a limited mitochondrial structure. Stumpy forms, on the other hand, have an elaborated mitochondrion ready for transition to the amino acid-based metabolism of the insect stage in which oxidative phosphorylation takes place. NAD diaphorase activity can be used to report on mitochondrial activity. Monomorphic or slender forms of *T. b. brucei* or *T. evansi* have no diaphorase activity in their mitochondria, while stumpy forms have high activity. In contrast, *T. congolense* and *T. vivax* bloodstream forms described as monomorphic displayed diaphorase activity in all parasites (Vickerman, 1965). Additionally, these parasites were described as having elaborated cristae in the mitochondrion as well as the capacity to metabolize pyruvate (Ryley, 1956). This could indicate that the majority of *T. congolense* and *T. vivax* parasites are pre-adapted for uptake into the tsetse fly.

More recently the protein PAD1 (protein associated with differentiation) was shown to be specifically detectable on stumpy, transmissible parasites, but not on slender forms (Dean et al., 2009). Use of this surface marker is now widespread in the categorisation of stumpy forms, and a greater understanding of the regulatory elements controlling expression of PAD1 has led to the generation of various reporter lines able to report on stumpy formation (MacGregor and Matthews, 2012).

Additional transcripts found to be upregulated in stumpy forms are the ESAG9 family of expression site associated genes (Barnwell et al., 2010).

1.2.8 Regulation of gene expression

Control of gene expression in *T. b. brucei* is divergent from other eukaryotes. Genes are transcribed as polycistronic units with few promoters identified. These polycistronic units are processed into their individual components by 5’ *trans* splicing and 3’ polyadenylation. The majority of regulation acts at the post-transcriptional level through binding of regulators to the 3’UTR of a particular gene, affecting transcript stability and translational efficiency (Clayton, 2002). Indeed a
number of 3’UTRs have been identified as important in controlling expression of
developmentally regulated genes, such as PAD1 (MacGregor and Matthews, 2012)
and ESAG9 (Monk et al., 2013). Much post-transcriptional regulation may be
through RNA-binding proteins (Clayton, 2013). Various RNA-binding proteins
(RBP) have been implicated in life cycle transitions in T. b. brucei, for example
RBP7 in the slender to stumpy transition (Mony et al., 2014), or RBP6
(overexpression of which triggered metacyclogenesis from cultured procyclic forms
(Kolev et al., 2012)).

1.2.9 Quiescence in stumpy forms and gene expression
changes during differentiation

Cell cycle arrest is accompanied by reduced protein synthesis in the stumpy form of
T. b. brucei, with control suggested to be at the level of translational initiation
(Brecht and Parsons, 1998). When transcriptional changes were followed through
slender to stumpy to procyclic differentiation, transcripts associated with an active
cell cycle (e.g. histones) were found to be down-regulated in stumpy cells, with up-
regulation of transcripts involved in protein synthesis found on differentiation to
procyclic forms (Kabani et al., 2009). Another study (Capewell et al., 2013) analysed
transcripts associated with polysomal material in stumpy forms, so potentially
escaping the general translational repression of this stage. These results found
components of oxidative phosphorylation and amino acid metabolism, amino acid
transporters, as well as mitochondrial transcripts to be enriched in stumpy forms.
This matched expectation for the life cycle stage primed for transmission to the
insect vector. In other transcriptomic studies, Jensen et al (Jensen et al., 2009) and
Nilsson et al (Nilsson et al., 2010) found glycolytic pathway components and
cytoskeletal transcripts down-regulated in this non-dividing, quiescent life cycle
stage. Transcripts encoding enzymes involved in oxidative phosphorylation and
amino acid metabolism were found to be more abundant in procyclic and
epimastigote forms of T. congolense than in the bloodstream and metacyclic forms
(Helm et al., 2009). Helm et al did not investigate transcriptional changes at different
points of the parasitaemia during T. congolense bloodstream infection.
Regulation at the level of translation or protein stability may result in differences between transcript abundances and the levels of expressed protein. Differential protein expression during the *T. b. brucei* life cycle has been investigated in numerous studies. Some studies (Butter et al., 2013; Urbaniak et al., 2012) have compared protein expression between cultured bloodstream and procyclic forms of *T. b. brucei* often using a methodology known as SILAC (stable isotope labeling by amino acids in cell culture). These studies identified changes in abundance of RNA-binding proteins, cell surface proteins and proteins involved in mitochondrial and glycosomal energy metabolism. One study, which used SILAC to study differences in protein abundance between slender, stumpy and procyclic forms of the pleomorphic *T. b. brucei* cell line AnTat 1.1 (Gunasekera et al., 2012), found that mitochondrial proteins were upregulated in stumpy and procyclic forms relative to slender forms, with the majority of changes occurring between the slender and stumpy stages. The increase in mitochondrial proteins in procyclic forms was of a greater magnitude in this study than in a previous study that did not use a pleomorphic strain of *T. b. brucei*. In the Gunasekera *et al* study, a number of proteins with the GO term plasma membrane were also increased in abundance in stumpy relative to slender forms. These data were also compared to a *T. congolense* proteomic data set (Eyford et al., 2011). There was some correlation in the changes in expression between slender bloodstream and procyclic forms in *T. b. brucei* and the comparable life cycle stages in *T. congolense*, however some proteins showed opposite patterns of expression in the two species. Proteomic analysis of *T. congolense* IL3000 bloodstream, procyclic and epimastigote forms (Eyford et al., 2011) revealed that many of the proteins differentially regulated through the life cycle were cell surface associated, such as VSG, ISG65, ISG75, congolense epimastigote specific protein (CESP), the surface protease GP63, an amino acid transporter, and a haptoglobin-hemoglobin receptor.

More recent proteomic studies have analysed not only distinct life cycle stages of *T. b. brucei* (i.e. slender, stumpy and procyclic forms), but have also investigated proteomic changes occurring early during differentiation from stumpy to procyclic forms. One study followed changes in protein abundance and phosphorylation before and after the point of commitment to differentiation from bloodstream to procyclic
forms (Domingo-Sananes et al., 2015). Proteins associated with the procyclic form cell surface were upregulated just 3 hours after stimulation of differentiation with cis-aconitate, as well as proteins implicated in developmental changes in metabolism. Additionally, the phosphorylation status of eukaryotic initiation factors changed, consistent with reinitiation of protein synthesis early in differentiation (Kabani et al., 2009). Similarly, a very recent proteomic study (Dejung et al., 2016) found increases in protein abundance from 2 hours following induction of differentiation to procyclic forms that were consistent with a changing metabolism and cell surface (e.g. a putative major surface protease GP63 (MSP-B) and an NADH-ubiquinone oxidoreductase complex I subunit). A number of proteins with increased expression in procyclic forms were already enriched in stumpy forms relative to slender forms, these included a procyclic form surface phosphoprotein (PSSA-2), Protein Associated with Differentiation 2 (PAD2), as well as the PTP-interacting protein, TbPIP39.

1.2.10 Pleomorphism in T. congolense and T. vivax

Early studies on the transmissibility to tsetse flies of T. b. brucei and T. b. gambiense established that increased capacity to be transmitted correlated with a greater proportion of stumpy forms, and not with parasitaemia (Wijers and Willett, 1960). The method of growth control in T. vivax and T. congolense remains unknown. Experimental infections to characterise T. congolense and T. vivax bloodstream forms at different points during parasitaemia have been carried out. In a study using FACs to analyse DNA content as an indicator of cell cycle position (Shapiro et al., 1984), an initially dividing population of T. vivax was superseded by a non-dividing population. However, there was no significant variation in size between the early and late populations of the parasite. Similarly, accumulation of non-dividing, non-infective forms at peak parasitaemia was observed in another study (Mahan and Black, 1989). However, again there did not appear to be a particular morphological form associated with reduced infectivity to mice (although a longer form seemed to be linked to the remission phase, compared to forms in the exponential or plateau phase of growth). Other authors have described polymorphism in T. vivax, postulating that a longer form may be equivalent to the stumpy form of T. b. brucei.
(Gardiner and Wilson, 1987). Hoare (Hoare, 1972) described *T. vivax* as dimorphic, losing its ‘tadpole’ forms and retaining only slender ones when transmitted mechanically or as rodent-adapted laboratory strains. This is comparable to *T. b. brucei* that become monomorphic in the laboratory. One early study attributed certain morphological forms to the degree of virulence; strains causing acute disease in cattle of West Africa had characteristically short forms, whereas long forms were chiefly associated with strains causing chronic infection in East Africa (Fairbairn, 1953).

Studies on *T. congolense* infection dynamics also come to conflicting conclusions on pleomorphism in this parasite. An early paper by Nantulya (Nantulya et al., 1978b) found that short forms dominate in the ascending parasitaemia, whereas long forms predominate at peak parasitaemia (days 7-8 of infection). On challenge of fresh mice, the longer forms were found to be less infective than the short forms. Furthermore, in a preliminary study (Nantulya et al., 1978a), parasites were found to be more successful in establishing mature infections in tsetse flies when flies were fed on mice at peak parasitaemia when compared to ascending or descending parasitaemia (although the number of flies used in this study was low). This work suggests that *T. congolense* longer forms are equivalent to *T. b. brucei* stumpy form parasites, while *T. congolense* shorter forms are equivalent to *T. b. brucei* slender forms. In opposition to this view, Vickerman (Vickerman, 1969a) had previously described *T. congolense* as monomorphic. More recently, in a study looking at transmissibility to tsetse flies at various points during parasitaemia, Akoda et al (Akoda et al., 2008) found that high parasitaemia of *T. congolense* did not correlate with increased transmissibility to tsetse flies. However, it was found that the proportion of infected flies was significantly higher for mice on day 5 or 10 post-infection compared to the first day parasites were observed in the blood (day 4). Moreover, it has been observed that the proportion of procyclic and metacyclic infections established in *Glossina morsitans morsitans* was significantly higher when flies were fed on mice at the peak of parasitaemia during acute infection, relative to those fed at peak parasitaemia during the chronic phase of infection (Masumu et al., 2010). Since cloned isolates were used in this study and flies were fed at the same level of parasitaemia, the changes in transmissibility are likely to be a result of changes in the
trypanosome during infection. Hence, *T. congolense* may have transmissible forms present at different points during the infection analogous to the situation in *T. b. brucei*. When analysing transmission stages, a potential confounding factor is that *T. congolense* is more infective to tsetse than *T. b. brucei* and this may mean that the proportion of transmissible forms maintained throughout mammalian infection does not need to be as high as for *T. b. brucei*. *T. vivax* has a higher infectivity for tsetse than both *T. congolense* and *T. b. brucei* (ILRAD, 1981).

Further changes that may occur during *T. congolense* chronic infections include changes in virulence and drug resistance. For example, during the course of a goat infection with a cloned *T. congolense* strain, attempts were made to infect naïve mice with a single trypanosome (Joshua, 1990). Parasites collected towards the end of the chronic infection were found to have higher virulence in mice compared to those collected earlier in infection. Virulence in mice has in some cases been shown to correlate with that in cattle, for instance Savannah type *T. congolense* were found to be more virulent in both zebu cattle and mice than Forest-type and Kilifi-type *T. congolense* (Bengaly et al., 2002a; Bengaly et al., 2002b). Additionally, a Savannah-type *T. congolense* strain was demonstrated to have higher transmissibility by tsetse flies than a strain of the Riverine-forest type (Reifenberg et al., 1997). Variations in virulence have been detected in *T. congolense* strains isolated from cattle in eastern Zambia, including identification of extremely virulent strains (Masumu et al., 2006a). The maintenance of virulent strains in the field may be aided by an association between increased virulence and increased transmissibility, which was reported in a study of the transmissibility of *T. congolense* strains isolated from cattle from a single geographical area (Masumu et al., 2006b).

In a study of drug sensitivity, *T. congolense* became refractory to diminazene treatment when treated on day 19 of infection rather than 24 hours post-infection. When these previously drug-resistant parasites were used to infect naïve goats the same pattern of drug-sensitive parasites on day 1 and drug-resistant parasites on day 19 was observed (Mamman et al., 1993), indicating that the alteration in resistance phenotype was associated with the chronicity of infection. Another study has found that of three isogenic clones of *T. congolense* with differing resistance to
isometamidium chloride, the clone with the highest level of resistance had the
highest tsetse fly infection rate (van den Bossche et al., 2006).

The role of stumpy forms of *T. b. brucei* as part of the parasite life cycle was
previously a subject of controversy, with some convinced that the stumpy form was
important for transmission, and others suspecting that it was simply a degenerating
parasite. The strong evidence that stumpy forms are able to survive on uptake into
the tsetse midgut and resume proliferation on differentiation to the procyclic form
has led to the wide acceptance of the role of stumpy forms in the parasite life cycle
((Vickerman, 1965), reviewed in (MacGregor and Matthews, 2010)). Similar
clarification of the presence or absence of pleomorphism in *T. congolense* and *T.
vivax* is required.

### 1.3 Quorum sensing and environmental sensing

#### 1.3.1 Environmental sensing in a simple eukaryote - yeast

Single celled organisms may face a barrage of challenges in the diverse extracellular
environments that they inhabit. Survival depends on plasticity, the ability to optimise
gene expression to the particular conditions faced.

One key aspect of survival relies on the ability to detect nutrient availability and to
adapt metabolism to make the most of nutrients that are available. For example, yeast
cells may respond to nutrient starvation by entering a quiescent phase. The presence
of glucose is sensed through a pathway involving protein kinase A (PKA), activity of
which promotes expression of the translational machinery and growth. Cyclic AMP
(cAMP) interacts with regulatory domains of PKA to feed into this pathway. Yak1 is
a proquiescence kinase inhibited by the action of PKA. On the other hand, low
glucose may be signalled through Snf1 kinase, leading to a shift in metabolism to use
of alternative carbon sources. Snf1 kinase is an orthologue of the mammalian AMP
kinase (AMPK), whose activity is stimulated by the accumulation of AMP in
nutritionally stressed cells. One important regulator in maintaining active growth in
the presence of nutrients is target of rapamycin complex 1 (TORC1). TORC1,
through multiple effectors, inhibits stress responses and promotes translational
activity. These pathways do not act in isolation; instead they act to integrate different
signals in order to coordinate the appropriate response to nutrient limitation (reviewed in (Broach, 2012; De Virgilio, 2012)).

A specific example of cross-talk between environmental sensing pathways involves *Schizosaccharomyces japonicas*, which switches from a yeast to hyphal lifestyle in response to nutrient limitation or DNA damage, in both cases the hyphal response is inhibited by cAMP (Furuya and Niki, 2012).

TORC1 is not just important in yeast but also plays a role in mammalian cell nutrient signalling, and defects in the TORC1 network may contribute to cancerous growth, as may abnormalities in other nutrient sensing pathways (Menon and Manning, 2013).

### 1.3.2 Environmental sensing in eukaryotic parasites

Eukaryotic parasites may face diverse environments in progression through their life cycle, and they must detect environmental triggers in order to respond appropriately.

One example of host cues influencing parasite fate decisions is the decreased investment in gametocytes observed following treatment of *Plasmodium chabaudi* with lysed parasitized red blood cells. The molecular mechanisms that regulate such fate decisions require further elucidation (Carter et al., 2014).

Nevertheless, differentiation through the *Plasmodium* life cycle in response to various environmental triggers relies upon, in several instances, signalling through the cyclic nucleotides cAMP or cGMP. Examples include gametocytogenesis or exflagellation of the activated male gametocyte in the mosquito vector (reviewed in (Baker, 2011)).

Recently a role of cAMP in environmental sensing and coordinated behaviour has been demonstrated in *T. b. brucei* procyclic forms that demonstrate social motility when plated. Knock down of the cAMP-specific phosphodiesterase PDEB1 was found to inhibit social motility (Oberholzer et al., 2015). This is in line with the finding that knock down of certain adenylate cyclases results in a hypersocial phenotype (Lopez et al., 2015). The role of cyclic nucleotides in environmental responses is not limited to eukaryotic parasites. For instance, cyclic-di-GMP can
function in swarming motility in various bacterial species (Simm et al., 2004). Similarities to well-studied cyclic nucleotide based signalling in prokaryotes or other eukaryotes may help in investigating such signalling pathways in trypanosomes. However, care needs to be taken as there are clear eccentricities in the cAMP pathways in trypanosomes (Tagoe et al., 2015); for example there are differences in cAMP responsive proteins, with PKA unresponsive to cAMP, but new cAMP responsive proteins (CARPs) identified (Gould et al., 2013).

One well-studied example of environmental sensing in T. b. brucei is the control of differentiation to the procyclic form on entering the tsetse midgut. Initial study showed that cold shock was able to sensitise stumpy (but not slender) parasites for perception of the differentiation signals cis-aconitate and citrate (Engstler and Boshart, 2004). Later the increased sensitivity of stumpy forms to the differentiation trigger was shown to be due to PAD proteins (a carboxylate transporter family), expression of which was temperature sensitive (Dean et al., 2009). Perception of either the citrate/cis-aconitate signal or mild acid treatment leads to differentiation via a protein phosphorylation cascade (Szöör et al., 2013). The protein tyrosine phosphatase TbPTP1 (Szöör et al., 2006) dephosphorylates and inactivates the DxDxT phosphatase TbPIP39 (TbPTP1-interacting protein, 39kDa) (Szöör et al., 2010) to prevent inappropriate differentiation of stumpy cells. On perception of the differentiation signal, TbPIP39 is phosphorylated and is translocated to the glycosome where it has activity, and is protected from TbPTP1.

1.3.3 Prokaryotic quorum sensors

In contrast to the relatively limited information relating to eukaryotic quorum sensing, there is a wealth of information relating to bacterial quorum sensing (Atkinson and Williams, 2009). Molecules used to communicate population status include the acyl homoserine lactones (AHLs); receiving such a signal may alter expression of bacterial virulence factors, including genes involved in biofilm formation or host colonisation. Perception of quorum sensing molecules is not restricted to the species producing them. For example, autoinducer-2 (AI-2) signalling is shared between gram negative and gram positive bacteria, with the result that certain species may be able to interfere with another’s signalling.
*Pseudomonas aeruginosa* does not produce AI-2 but may still respond to it. Such interactions could result in alterations in pathology in co-infection. Sibley *et al* (Sibley et al., 2008) observed that in a *Drosophila* infection model increased virulence was observed when *Pseudomonas* infection was accompanied by a normally non-pathogenic *Streptococcus* species. Bacterial interactions may be cooperative, for instance in establishment of a biofilm, or competitive. It has even been proposed that bacterial quorum-sensing peptides, extracellular death factors, that can mediate interspecies death could be used as a basis for novel antimicrobials (Kumar and Engelberg-Kulka, 2014).

Bacterial quorum sensing signals may not just cause effects within their kingdom but may interact with their eukaryotic neighbours. For instance, directed locomotion of *Caenorhabditis elegans* (*C. elegans*) was found to be preferentially targeted towards AHL producing bacteria (Beale et al., 2006). Another such example is the attraction of microalgae zoospores by AHL producing bacteria found in tidal pools (Joint et al., 2002). There is also evidence of interspecies signalling between *C. albicans* and *P. aeruginosa*, with the yeast pathogen able to prevent production of a *P. aeruginosa* product (Cugini et al., 2007), and a *P. aeruginosa* quorum sensing molecule able to prevent mycelium production by the yeast (Hogan et al., 2004). Thus, each competes to establish its own foothold in a shared environment.

### 1.3.4 Eukaryotic quorum sensors

In addition to the ability of certain eukaryotes to detect prokaryotic quorum sensing molecules, certain eukaryotes are able to produce and respond to their own signals. A study of *Candida albicans* described a role for a quorum sensing molecule, identified as farnesol, in controlling the density-dependent switch from yeast to hyphal growth. Farnesol was able to inhibit the yeast-to-mycelium transition, as was supernatant from *Candida albicans* (such supernatants were active between strains) (Hornby et al., 2001).

*Dictyostelium discoideum* is a social amoebae that develops from a unicellular to multicellular aggregated body, which displays coordinated behaviour in migration towards optimal conditions. The aggregates form fruiting bodies that consist of a
stalk with multiple spores. Various small molecules are involved in coordinating this developmental behaviour, including the glycoprotein CMF (Yuen et al., 1995), the protein complex counting factor (CF) (Tang et al., 2001), the chlorinated hydrocarbon DIFs (Kay et al., 1999), and, as with other cases of density sensing, cAMP plays a key role (reviewed in (Firtel, 1996)). Density sensing may be coordinated with the detection of available nutrients, for example prestarvation factor (PSF) accumulates with increasing density and signals the transition from proliferation to development. However, this response is inhibited by the detection of food bacteria (Clarke et al., 1988). Another factor, CMF450, is involved in coordinating the decision of feeding cells to arrest and begin the transition through development (Iijima et al., 1995). The decision to differentiate into a particular cell type may be related to the cell cycle position of a given cell (Weijer et al., 1984).

Density dependent signalling has been suggested as a reason for resurgence from remaining metastases when a primary mammalian tumour is removed. It has been suggested that released tumour factors may slow growth when there is a high density of tumour cells (Guba et al., 2001; Peeters et al., 2008). In addition, it has been proposed that metastatic colonisation by tumour cells may involve social interaction in an analogous way to bacterial biofilm formation (Hickson et al., 2009).

### 1.3.5 Density dependent control in eukaryotic parasites

A study on co-infecting plasmodium species carried out by Bruce (Bruce et al., 2000) suggests that a particular pattern of infection may occur because the dominant species in an infection reaches a threshold leading to density dependent regulation, which limits the growth of other co-infecting species. Once a species-specific response clears the one majority species, the density falls below the threshold and other co-infecting species can expand leading to the sequential pattern of infection observed. The implication for a species-specific vaccine is that removal of one species in endemic regions may lead to expansion of other species through removal of density dependent regulation.
1.3.6 Density dependent control of *T. b. brucei* growth: SIF

Stumpy forms are not predicted to contribute to the change in environment that triggers slender to stumpy differentiation, so that the stumpy population never reaches 100% (MacGregor et al., 2011; Seed and Black, 1997).

The differentiation of *T. b. brucei* to stumpy forms is driven by accumulation of a soluble parasite-derived factor called SIF as parasite population density increases. To date the identity of SIF has remained elusive. However, it has been found that cell-permeable analogues of cAMP are able to drive slender forms to differentiate to stumpy-like forms in culture, leading to a proposed role of cAMP in the SIF signalling pathway (Breidbach et al., 2002; Reuner et al., 1997; Vassella et al., 1997a). Indeed intracellular levels of cAMP had been shown to increase during ascending parasitaemia, but decrease as the transition to stumpy forms began (Mancini and Patton, 1981). Further to this, it was discovered that cell-permeable analogues of the hydrolysis products of cAMP, such as 5’AMP and adenosine are able to trigger the conversion of monomorphic bloodstream form *T. b. brucei* into stumpy-like forms, and hydrolysis resistant cAMP analogues are not able to trigger the stumpy-like response (Laxman et al., 2006). Therefore, it has been proposed that it is the hydrolysis products of cAMP that feed into the SIF signalling pathway. If and where the cAMP response pathway intersects with the natural SIF signalling pathway is currently undefined.

1.3.7 Regulators of stumpy formation

Previous studies have identified genes that when knocked out lead to premature stumpy formation. For example, TbMAPK5 gene deletion in pleomorphs prevented infections reaching high parasitaemias in mice. MAPK5 knock out was associated with premature differentiation to stumpy forms both *in vitro* and *in vivo*. Consequently, it was proposed that MAPK5 acts as a negative regulator of slender-stumpy differentiation. The action of MAPK5 is likely to occur at the level of downstream signalling from the SIF response, since MAPK5 null mutants produce a similar amount of SIF as wild type cells in conditioned media experiments (Domenicali Pfister et al., 2006). A Zinc-finger kinase was also postulated to be a negative regulator of SIF-mediated differentiation to stumpy forms (Vassella et al.,
2001), as null mutants showed reduced growth and increased expression of markers of stumpy differentiation in vitro. However, there was no difference between the parasitaemia of wild-type or null-mutant cells in vivo. The authors proposed that there may be multiple redundant mechanisms antagonizing the effect of SIF in vivo, leading to the lack of a phenotype for the ZFK null mutant. They suggest that suboptimal growth conditions in vitro have allowed an effect of ZFK deletion to be observed, and this may reflect the situation in natural infections. More recently, TbTOR4 was identified as another negative regulator of stumpy differentiation (Barquilla et al., 2012). Unlike the MAPK5 and ZFK, where knock down only showed a phenotype in pleomorphs, monomorphs also showed a phenotype as a result of TOR4 knock down. Cell lines depleted for TOR4 also showed an increased sensitivity to the experimental trigger of procyclic form differentiation, cis-aconitate. Furthermore, when treated with AMP or hydrolysable analogues of cAMP, monomorphs showed a reduction in TbTOR4 expression associated with differentiation to a stumpy-like form. The authors suggested that AMP rather than cAMP is involved in the slender-stumpy differentiation process, and that reduced cellular energy (high AMP:ATP ratio) inactivates TORC4 (the complex that includes TOR4) to trigger differentiation to the stumpy form.

More recently in a genome-wide screen, Mony et al (Mony et al., 2014) identified many drivers of stumpy formation. These genes could plausibly form signalling cascades and comprise both suspected upstream (e.g. identified phosphatases and kinases) and downstream components (e.g. Hypothetical protein 2 (identified as a post-transcriptional regulator in another screen (Erben et al., 2014), or RBP7). Comparisons can be made to nutrient sensing pathways and growth control in other single-celled organisms (as discussed in (Mony and Matthews, 2015)). Many of the identified stumpy inducers had orthologues in yeast known to be involved in environmental sensing and pathways to quiescence. For example, the T. b. brucei regulatory subunit of PKA was identified as a stumpy inducer, and the PKA pathway in yeast is known to play a role in the transcriptional changes leading to quiescence in response to changing glucose availability (Zaman et al., 2008). Another example is the identification of an AMPK catalytic subunit in the screen for stumpy inducers;
AMPK in yeast (known as Snf) functions in driving a cell towards a state of reduced activity when its ATP levels are low (Carling et al., 2011).

If these stumpy inducer genes are conserved in *T. vivax* and *T. congolense*, then it is possible that mechanisms of SIF signalling are also conserved in these species, raising the potential for interaction between the species in a co-infection. This has relevance for understanding the potential for cross talk between species, and consequences of using therapeutics developed to target stumpy formation in *T. b. brucei*. For instance if a therapeutic could drive *T. b. brucei* in an infected host to turn stumpy leading to eventual clearance of the parasite, then this could result in an increased prevalence of co-infecting *T. vivax* and *T. congolense*. Alternatively, if shared signals were used, therapeutics targeting one parasite might also target other species.

Moreover, the effect of mixed infections on virulence may depend on whether or not co-infecting strains or species are able to coordinate behaviours in order to cooperate for the good of the total parasite population, for example in the case of quorum sensing between bacterial species ((Brown et al., 2002) and see section 1.3.3). If the individual species comprising a mixed infection behave as separate entities then the most advantageous strategy for each parasite species would be to deplete the hosts resources before the competitor species could, at the expense of host fitness, thus favouring evolution of more virulent strains and selection against more prudent parasites. However, if parasites were able to respond to a conserved signal that resulted in less host exploitation by individual species (e.g. as a result of growth arrest) then the presence of a co-infecting species could actually reduce the virulence of a given species. For instance, *T. b. brucei* bloodstream forms have quorum-sensing controlled system for density control (see section 1.3.6). These parasites are prudent in their use of host resources enabling establishment of chronic infections. In field situations, co-infection with other species of trypanosome such as *T. congolense* and *T. vivax* occur, and if these parasites did not respond to density control, they could recover a greater share of host resources at the expense of *T. b. brucei* and potentially the host. Therefore, it is important to understand the conservation of
density sensing mechanisms in African trypanosomes in order to understand the dynamics of coinfections and the likely effect of mixed infections on virulence.

1.4 Aims

The aim of this thesis is to investigate the methods of growth control used by *T. vivax* and *T. congolense*, as well as examining the potential for interaction with *T. b. brucei* in co-infection. Specifically I aim to address the following questions:

1. Does *T. congolense* generate arrested forms that resemble *T. b. brucei* stumpy forms *in vivo*?
   - Is there cell cycle arrest at peak parasitaemia?
   - What transcriptomic changes accompany the transition to peak parasitaemia?
   - How do *T. congolense* peak enriched genes compare to *T. b. brucei* stumpy enriched genes?

2. Are components of the *T. b. brucei* SIF-responsive pathway conserved in *T. congolense* and *T. vivax*?
   - Can *T. congolense* and *T. vivax* orthologues of *T. b. brucei* SIF-responsive genes be identified?
   - Does orthologue overexpression *in vitro* drive cell cycle arrest?
   - Can orthologue overexpression *in vivo* rescue stumpy formation in a SIF resistant *T. b. brucei* KO line?

3. Does *T. congolense* produce a SIF-like molecule?
   - Is *T. congolense* conditioned media (CM) able to inhibit growth of low-density *T. brucei*? Is inhibition associated with accumulation in 1K1N or increased expression of a stumpy-specific reporter?
   - Do monomorphic and pleomorphic *T. b. brucei* differentially respond to *T. congolense* CM? Are SIF-resistant *T. b. brucei* RNAi or KO lines resistant to *T. congolense* CM?
- How do *T. b. brucei* and *T. congolense* impact each other’s growth in a co-infection?
Chapter 2: Materials and Methods
2.1 Trypanosome strains

*T. b. brucei* parasites of the pleomorphic cell line AnTat 1.1 90.13 (Engstler and Boshart, 2004) were transfected to generate all of the cell lines produced in thesis. This cell line co-expresses T7 RNA polymerase and the tet repressor. Additionally, all infections were carried out with this pleomorphic cell line or modified versions of this parental line.

*T. b. brucei* parasites of monomorphic strains were used only in conditioned medium experiments. The monomorphic strain most commonly used was Lister 427, a potential history of which was recorded by George Cross at Rockefeller University (http://tryps.rockefeller.edu/trypsru2_pedigrees.html). This included use of the Lister 427 90.13 strain, a cell line co-expressing T7 RNA polymerase and the tet repressor (Wirtz et al., 1999). An additional derivation of the Lister 427 strain that was used was the ‘2T1’ strain (Lister 427, MITat1.2, clone 221a, (Alsford and Horn, 2008)).

*T. congolense* parasites of the IL3000 strain were used both for infections and *in vitro* experiments. This strain was derived from the ILC-49 strain that was isolated from a cow in the Trans Mara, Kenya (Wellde et al., 1974). The *T. congolense* IL3000 parasites used for *in vivo* experiments were provided by Dr Annette MacLeod in a blood straw. The *T. congolense* IL3000 parasites used for *in vitro* experiments were supplied as culture-adapted bloodstream forms by Dr Liam Morrison, who had received them from Professor Théo Baltz.

*T. vivax* parasites of the IL1392 strain were cultured as epimastigote forms and used to generate genomic DNA. The IL1392 strain was derived from the Zaria Y486 Nigerian isolate (Chamond et al., 2010; Leeflang et al., 1976).

2.2 Trypanosome maintenance *in vitro*

Pleomorphic bloodstream form *T. b. brucei* were cultured in HMI-9 supplemented with 10% FBS and 1% Penicillin/ Streptomycin at 37°C with 5% CO₂ (Hirumi and Hirumi, 1989). *T. b. brucei* parasites were cultured in suspension, and were passaged at least every 2-3 days by dilution into fresh culture medium. Cultures were not allowed to reach a density greater than 1x10⁶ cells/ml. Freezer stocks were prepared by pelleting 2-3 x10⁶ cells (1400 x g, 5 minutes) and resuspending these in 1ml HMI-
containing 14% glycerol. Freezer stocks were stored at -80°C. When recovering parasites from frozen, thawed parasites were first pelleted and resuspended in 10ml fresh HMI-9 to remove glycerol. Monomorphic bloodstream form T. b. brucei were treated in the same way as pleomorphic T. b. brucei parasites, but were allowed to reach higher densities in culture (up to 2-3x10^6 cells/ml).

*T. congolense* bloodstream forms were cultured in TcBSF3 supplemented with 25% goat serum, 5% serum plus and 1% Penicillin/ Streptomycin at 34°C with 5% CO₂ (Coustou et al., 2010). *T. congolense* bloodstream forms are adherent and were passaged at least every 2-3 days by flushing adherent parasites with a pipette followed by dilution of resuspended parasites into fresh culture medium. Parasites were cultured up to a density of 1x10^7 cells/ml, and were generally not maintained at a density lower than 5x10^4 cells/ml, as parasites did not respond well to low density. Parasites were frozen at a density of 3-6x10^6 parasites in 1ml of freezing mix (TcBSF3 with 30% glycerol). Frozen stocks were stored at -80°C. When recovering parasites from frozen, thawed parasites were first pelleted and resuspended in 2ml fresh TcBSF3 to remove glycerol.

*T. vivax* epimastigotes were cultured in TV3 media (D'Archivio et al., 2011). Epimastigotes are adherent and were not passaged until confluent on the base of the culture flask. Until confluence was reached the medium was replaced every 2-3 days by first washing the parasites with 1x PBS and then adding TV3 media (when parasite density was low this was supplemented with 100µl of 1M HEPES per 3ml culture to maintain an optimal pH). Once a culture had reached confluence it could be passaged, so adherent cells were dislodged by firmly tapping the flask and these were transferred to a new flask. Once parasites had adhered to the new flask, these were washed with 1xPBS and TV3 media was added. Parasites were frozen by tapping the flask to release adherent cells into the supernatant, the cells were then pelleted (800 x g, 5 minutes) and were resuspended in TV3 media containing 10% glycerol. Only confluent cultures were used to make freezer stocks. To recover parasites from frozen, a thawed aliquot was added directly to fresh TV3 media, and after parasites had adhered to the flask the cells were washed with 1x PBS and fresh TV3 media was added.
2.3 Cloning and DNA manipulations

2.3.1 Extraction of genomic DNA

5x10^7 parasites were pelleted, washed in 1xPBS (when starting from cultured parasites), and were resuspended in 150µl of lysis buffer (TELT) (Medina-Acosta and Cross, 1993). The lysis reaction was incubated for 5 minutes at room temperature, before addition of 150µl water equilibrated phenol/chloroform (1:1 v/v). This was mixed by inverting for 5 minutes, and the mixture was then centrifuged at 11330 x g for 10 minutes. The upper phase was transferred to a new eppendorf tube and the phenol/chloroform extraction was repeated to clean up the sample. Again the upper phase was transferred to a new eppendorf tube and 300µl 100% ethanol was added, the mixture was inverted for 15 seconds and was then incubated for 5 minutes at room temperature. The mixture was centrifuged (11330 x g, 10 minutes) to pellet the DNA. The pellet was washed in 100% ethanol, and resuspended in 50µl TE buffer (containing 20µg/ml pancreatic ribonuclease).

Alternatively DNA was extracted using the Qiagen DNeasy Blood and Tissue kit (Qiagen, 69506) using the protocol ‘Purification of Total DNA from Animal Blood or Cells’ as per the manufacturer’s instructions.

2.3.2 Polymerase chain reaction (PCR)

PCRs were performed in 0.2ml PCR tubes (Axygen, 321-10-051) using a Biometra TProfessional Basic Gradient Thermocycler. Either the Roche Expand High Fidelity PCR system (11732650001) or the Q5 High Fidelity DNA polymerase system (NEB, M0491S) was used. Below are example reactions for each system:

For the Expand High fidelity PCR system, each reaction typically contained:

31.5µl dH2O
10µl Expand hifi plus 5x reaction buffer
1µl PCR grade nucleotide mix
2.5µl 10µM forward primer
2.5µl 10µM reverse primer
2µl gDNA or dH2O (for negative control)
0.5µl Expand hifi plus enzyme blend
Expand high fidelity PCR reaction programme:

Preheat lid (94°C)

94°C for 2 minutes

94°C for 30 seconds

(5°C for 30 seconds)

72°C for 1 minute/kb

72°C for 5 minutes

10°C hold

For the Q5 High fidelity PCR system, each reaction typically contained:

5µl 5x Q5 reaction buffer

0.5µl dNTPs (10mM)

1.25µl 10µM Forward primer

1.25µl 10µM Reverse primer

1µl gDNA or dH₂O (for negative control)

0.25µl Q5 high fidelity DNA polymerase

15.75µl dH₂O

Q5 high fidelity PCR reaction programme:

Annealing temperatures were calculated using the NEB Tₘ calculator (http://tmcalculator.neb.com). As Q5 polymerase produces blunt ended products, 1µl GoTaq polymerase (Promega) was added at the end of the reaction and this was incubated for 20 minutes at 72°C for addition of overhangs for TA cloning.
Preheat lid (98°C)

98°C for 30 seconds

98°C for 10 seconds

Annealing temperature for 30 seconds

72°C for 2 minutes (20-30 seconds/kb)

72°C for 2 minutes

Hold at 10°C

Primers used are detailed in Appendix 2.4. Where primers were designed to incorporate restriction sites, a 6 base sequence of Gs and Ts was incorporated prior to the restriction site. When calculating the annealing temperature only complementary regions of sequence were taken into account. Primers were checked for the potential to form hairpins, self-dimers and heterodimers using the Oligo Analyser 3.1 tool (integrated DNA technologies).

2.3.3 Agarose gel electrophoresis

DNA samples were visualised by agarose gel electrophoresis. Generally, 1% agarose gels were prepared (for 0.4-10kb products), but in some cases different percentage gels were prepared when products were expected to be particularly small or large. To prepare a 1% gel, 1g agarose (Biolab, BIO-41025) was dissolved in 100ml Tris-acetate-EDTA (TAE) by heating in a microwave on full power for about 90 seconds. The solution was then cooled to approximately 65°C before addition of 3.75µl Ethidium bromide (10mg/ml) for a 100ml gel (5µl for a 150ml gel). The gel was then cast in a Fisher Brand gel apparatus. DNA samples were mixed with 6x DNA loading buffer and were run alongside 5µl DNA Smart Ladder (Eurogentec, MW-1700-10) at 100-130V for approximately 60 minutes. Images of bands were captured using a G:Box (Syngene).

2.3.4 Purification of PCR products

DNA was purified using the NucleoSpin Gel and PCR Clean-up kit (Machery-Nagel) according to the manufacturer’s instructions. When multiple bands were obtained the
band of expected size was excised from the agarose gel prior to purification. Occasionally, instead of using the kit, DNA was purified by squeezing the DNA from the gel between two sheets of Parafilm® (Sigma, P7793-1EA), and making the resultant liquid up to a volume of 200µl with water. The DNA was then precipitated by addition of 20µl 3M sodium acetate and 550µl 100% ethanol, the solutions were mixed by vortexing and the precipitation reaction was stored at -80°C for at least 1 hour. The precipitated DNA was then pelleted at 4°C for 10 minutes at 18,000 x g, and the pellet was resuspended in dH$_2$O.

### 2.3.5 Ligation

A typical ligation reaction is shown below:

- 1µl 10x ligation buffer (NEB)
- 1µl T4 ligase (NEB)
- _µl Insert *
- _µl Vector*

Volume of dH$_2$O needed to complete a 10µl ligation reaction.

*The volume of insert required to generate a 3:1 insert:vector ratio was calculated using the equation: (kb of insert/kb of vector) x ng of vector = ng of insert needed for a 1:1 ratio. Although the 3:1 ratio was most commonly used, occasionally ligation reactions were set up to maximise the ratio of the insert relative to the vector. Ligation reactions were incubated at 4°C overnight.

### 2.3.6 Transformation of competent cells

An entire 10µl ligation reaction was used to transform 90µl of competent XL1 blue *E. coli* cells. This reaction was then incubated on ice for 30 minutes, prior to heat shock in a 42°C water bath for 30 seconds. 800µl LB broth was added to the reaction and this was incubated for 1 hour, shaking at 37°C. The cells were then gently pelleted (2400 x g, 90 seconds), 800µl of the supernatant was removed, and the cells were resuspended in the remaining 100µl volume. This was spread onto LB agar plates containing Ampicillin (100µg/ml). When the pGEM®-T Easy vector
(Promega, A1360) had been used in the ligation, IPTG (Isopropyl β-D-1-thiogalactopyranoside) and X-gal were applied to the plates prior to spreading of the cells, to enable blue:white screening for identification of recombinant colonies which appeared white due to insertional inactivation of the enzyme β-galactosidase.

2.3.7 Small scale plasmid preparation (Miniprep)

Single colonies were picked from LB Amp plates and were incubated overnight in a shaking incubator at 37°C in 5ml LB broth with 100µg/ml Ampicillin. The following day 5ml cultures were pelleted at 3000 x g at 4°C, and pellets were resuspended in 300µl of miniprep solution I and were transferred to eppendorf tubes. 300µl of miniprep solution II was added to each tube, and the solution was mixed by inversion. 350µl of miniprep solution III was added to the mixture and was incorporated by inversion. The mixture was centrifuged at 15,000 x g at 4°C for 15 minutes, and the supernatant was then carefully transferred to a new eppendorf tube. The supernatant was then centrifuged (15,000 x g) for a further 10 minutes at 4°C. The supernatant was again removed to a new eppendorf tube, 650µl of isopropanol was added, and the solutions were vortexed for mixing. The mixture was centrifuged (15,000 x g) for 30 minutes at 4°C, the isopropanol was then removed and 1ml 70% ethanol was added to the DNA pellet. Samples were centrifuged (15,000 x g) for 5 minutes, the ethanol was removed, and the DNA pellets were air-dried. Each DNA pellet was finally resuspended in 30µl dH2O containing 330µg/ml RNase. The DNA concentration was determined on a Nanodrop spectrophotometer.

2.3.8 Restriction digestions

Restriction digests were carried out using enzymes from NEB or Promega. Up to 40 units of restriction enzyme were used to digest up to 30µg of DNA in 25µl or 50µl reactions. Reactions were usually incubated for 1 hour or overnight. The most appropriate buffer and reaction temperature for each double digest reaction was identified using the NEB double digest tool (https://www.neb.com/tools-and-resources/interactive-tools/double-digest-finder). When SpeI and BclI were used in a double digest, the reaction was incubated for 1 hour at 37°C after addition of SpeI, before addition of BclI and incubation for 2 hours at 50°C. The digested DNA fragments were separated by agarose gel electrophoresis. Where the insert or vector
backbone of a digested plasmid required purification the appropriately sized DNA band was excised from the gel and purified as for PCR product purification (see section 2.3.4).

2.3.9 Sequencing

1µl of DNA (1-10 µg/µl) from a small-scale plasmid preparation was diluted with 4µl dH₂O, and was mixed with 1µl 3.2µM forward or reverse sequencing primer in a 0.2ml PCR tube. Edinburgh Genomics (University of Edinburgh) performed DNA sequencing. Sequencing data was viewed using the 4Peaks tool (http://nucleobytes.com/4peaks/index.html) and sequence alignments were carried out using the EMBL-EBI EMBOSS Needle tool for Pairwise nucleotide sequence alignments (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html) or the EMBL-EBI Clustal Omega tool for Multiple nucleotide sequence alignments (http://www.ebi.ac.uk/Tools/msa/clustalo/).

2.3.10 Preparation of DNA for transfection

To prepare a plasmid for transfection 60µg of plasmid DNA was split between two 25µl restriction digest reactions to linearise the plasmid. Digests for linearization were carried out at the temperature recommended by the enzyme manufacturer overnight. 1µl of each digest reaction was run on an agarose gel to check linearization had occurred. The linearised DNA was then purified from the digest reaction. Each digest reaction was divided between two nucleospin columns (Machery-Nagel) and DNA was purified according to the manufacturer’s instructions. In the final step DNA was eluted four times by addition of 25µl of dH₂O to the nucleospin columns. The two resultant 100µl volumes for each restriction digest were combined together and the DNA was precipitated by addition of 20µl 3M sodium acetate and 550µl 100% ethanol. The precipitation mixture was mixed well and stored at -80°C for at least 1 hour or overnight. The precipitated DNA was pelleted (18,000 x g) for 10 minutes at 4°C, and the air-dried pellet was then resuspended in 10µl dH₂O and was stored at -20°C prior to transfection.
2.4 Transfection of bloodstream form trypanosomes

2-3x10^7 parasites were used per transfection. The cells were pelleted (1400 x g, 5 minutes) and resuspended in 1ml trypanosome dilution buffer (TDB), and were transferred to an eppendorf tube. The cells were again pelleted (2370 x g, 5 minutes), and were then resuspended in 100µl Amaxa transfection buffer (Amaxa Basic Parasite Nucleofector Kit II, Lonza, VMI-1021) with 10µl linearised plasmid. The cells and DNA were transferred to an electroporation cuvette and the electroporation was carried out in an Amaxa Nucleofector II using the programme ‘Z001: free program choice’. The transfected cells were added to 30ml pre-warmed HMI-9, a 1:10 dilution was made from this undiluted flask, and a 1:100 dilution was also prepared, each in 30ml volumes. Drug treatment of the transfected cells was applied 6 hours later or the following day. Cultures of transfected parasites were mixed with an equal volume of HMI-9 containing a 2x concentration of the selection drug(s). 2ml volumes of the drugged transfectants were then plated into a 24 well plate (Corning® Costar®, CLS3524). One plate was prepared for each dilution (undiluted, 1:10, 1:100). A well of untransfected cells was included as a control for the efficacy of the drug treatment. Once all the cells in the control well were dead, and parasite numbers in some of the 2ml wells had begun to increase (3-8 days post-transfection), parasites from these positive wells were passaged into fresh HMI-9 containing the selective antibiotics. After 2 successful passages, drug-resistant parasites were frozen in HMI-9 containing 14% glycerol, and the resultant clones were tested for their phenotype.

Drug concentrations used to select and maintain transfectants:

1.5µg/ml phleomycin
0.5µg/ml puromycin
10µg/ml blasticidin
0.5µg/ml hygromycin
2.5µg/ml G418
2.5 Monitoring growth of cell lines *in vitro*

2.5.1 Testing induction of RNAi or overexpression *in vitro*

In most cases induction of RNAi or overexpression was triggered by treatment with 1µg/ml doxycycline. For some experiments, where the effect of overexpression was severe, the doxycycline was titrated to lower doses. The growth of induced cultures was compared to that of uninduced cultures for 3-6 days post-induction. Cell numbers were monitored every 24 hours using a Beckman Coulter Z2 Coulter Particle count and size analyser (or haemocytometer if there was a number of dead cells). The effectiveness of RNAi was monitored by preparation of RNA for analysis by Northern Blot (see section 2.13) from 5x10⁷ cells collected after 48 hours of induction with doxycycline, as well as uninduced control cells. The overexpression of a tagged protein was monitored by collection of at least 3x10⁶ cells and preparation of protein samples for Western Blot (see section 2.12) 1, 2 and 3 days post-induction, as well as uninduced control samples.

2.5.2 Testing response to 8-CPT-cAMP *in vitro*

In order to test the response of newly generated cell lines to 8-CPT-cAMP, cultures were first induced with doxycycline (or remained uninduced) for 24 hours prior to addition of 100µM 8-CPT-cAMP. The growth of treated cultures was compared to that of untreated cultures for 4-5 days after treatment with 8-CPT-cAMP. Cell numbers were monitored every 24 hours using a Beckman Coulter Z2 Coulter Particle count and size analyser (or haemocytometer if there was a number of dead cells).

2.5.3 Cell cycle analysis *in vitro*

Cells were smeared onto slides using the edge of a cover slip and left to dry, and these slides were used for cell cycle analysis. Slides were initially rehydrated for 5 minutes in 1x PBS. The edges of the slides were then dried and 30µl of a DAPI working dilution (10µg/ml in PBS) was applied to the smears. Slides were incubated in a humidity chamber for 2 minutes and were then washed for 5 minutes in 1x PBS. Slides were then mounted with 40µl Mowiol containing 2.5% DABCO (1, 4-diazabicyclo[2.2.2]octane) and were left to dry overnight at room temperature in the
dark, before storage at 4°C. Slides were analysed on a Zeiss Axioskop 2 plus for KN configuration counts, and QCapture software (QImaging) was used for image capture.

### 2.5.4 Protein localisation in vitro

Methanol fixed slides were rehydrated in 1xPBS for 5 minutes, and excess PBS was removed before blocking for 30 minutes in 2% BSA:PBS in a humidity chamber. Slides were washed once in PBS, and were then incubated in 50µl of primary antibody (diluted in 2% BSA:PBS, αPAD1 1:1000, BB2 1:2, YL1/2 1:5000 (Abcam)) for 1-2 hours at 37°C in a humidity chamber. Slides were then washed three times in PBS for 5 minutes, before incubation with secondary antibody (diluted in 2% BSA:PBS, α-rabbit Alexa fluor 488 1:500 (Life technologies), α-mouse Alexa fluor 568 1:500 (Abcam), α-rat Alexa fluor 568 1:500 (Life technologies)) for 1 hour at 37°C in a humidity chamber. Slides were again washed three times in PBS for 5 minutes, before incubation in the humidity chamber for 2 minutes with 30µl of a DAPI working dilution (10µg/ml in PBS). Slides were finally washed for 5 minutes in PBS. Slides were then mounted with 40µl Mowiol containing 2.5% DABCO (1, 4-diazabicyclo[2.2.2]octane) and were left to dry overnight at room temperature in the dark, before storage at 4°C. Slides were analysed on a Zeiss Axioskop 2 plus or Zeiss Axio Imager Z2, and QCapture software was used for image capture.

Alternatively, cells that had been paraformaldehyde fixed were used for immunolocalisation analysis. In this case 2x10^6 cells were pelleted at 2450 x g for 5 minutes and washed with cold PBS, and these were then resuspended in 125µl of cold PBS and 125µl of 8% paraformaldehyde. Cells were fixed on ice for 10 minutes. The paraformaldehyde was then removed and fixed cells were resuspended in 130µl 0.1M glycine in PBS. The cells were incubated at least overnight at 4°C before resuspension in 10µl 1xPBS per 1x10^5 cells. Fixed cells were incubated for 1 hour at room temperature on slides (8 well 6mm microscope slides Thermo-scientific, ER-201B-CE24) pre-coated with poly-L-lysine (15 minute incubation, 0.1mg/ml BiochromAG, 7240) in a humidity chamber. Alternatively, cells were adhered to Polysine® slides (VWR, 631-0107) in areas demarcated using an ImmEdge hydrophobic barrier pen (Vector laboratories, H-4000). Excess PBS was
removed from the wells using an aspirator. 20µl 0.1% triton in PBS was applied to each well for 2 minutes, this was then aspirated and wells were washed with a large drop of PBS. Wells were blocked with 2% BSA:PBS for 45 minutes at 37°C in a humidity chamber, before removal of excess blocking solution and application of 20µl primary antibody. Wells were incubated with primary antibody (diluted in 2% BSA:PBS, αPAD1 1:1000, BB2 1:2, YL1/2 1:5000) for 45 minutes at 37°C in a humidity chamber. Positive control wells and secondary antibody only wells were included for each experiment. Wells were each washed 5 times by repeatedly applying and aspirating 1x PBS. Wells were incubated with 20µl secondary antibody (diluted in 2% BSA:PBS, α-rabbit Alexa fluor 488 1:500, α-mouse Alexa fluor 568 1:500, α-rat Alexa fluor 568 1:500) for 45 minutes at 37°C in a humidity chamber. The secondary antibody was then aspirated and 20µl of a DAPI working dilution (10µg/ml in PBS) was applied to each well for 1 minute, followed by 5 washes with PBS. The slides were allowed to dry for 5 minutes at 37°C. Slides were mounted with a cover slip by application of 30µl drops of Mowiol containing 2.5% DABCO between wells. The cover slip was applied firmly to minimise the number of bubbles trapped, and the slides were stored in darkness to dry before visualisation. Slides were analysed on a Zeiss Axioskop 2 plus or Zeiss Axio Imager Z2, and QCapture software was used for image capture.

2.6 Conditioned medium generation

For generation of conditioned medium, cultures of *T. b. brucei* AnTat 1.1 90.13 were established at 1x10^5 cells/ml in HMI-9 (Hirumi and Hirumi, 1989) and were incubated for 2 days at 37°C (and 5%CO₂). Conditioned medium was harvested after 2 days of growth when cells had reached a density of 2-3x10^6/ml, by pelleting the cells (1400 x g, 5 minutes) and passing the supernatant through a 0.22µm filter. Filtered supernatant was stored at 4°C. Cultures of *T. congoense* for generating conditioned medium were in most cases established at 1x10^5 cells/ml in TcBSF3 and were incubated for 3 days at 34°C (and 5%CO₂). Conditioned medium was collected at a range of densities from 2x10^6 cells/ml to 1x10^7 cells/ml, and prepared as for *T. b. brucei* conditioned medium by filtering of conditioned supernatant through a 0.22µm filter. Conditioned medium was stored for a maximum of 5 days at 4°C before use.
Alongside flasks for conditioned medium generation, flasks containing either HMI-9 or TcBSF3 without parasites were prepared, and this control medium was treated in the same way as the conditioned medium and used as a negative control in all conditioned medium experiments.

2.7 CAT reporter assay

Two cell lines had been generated by Paula MacGregor (MacGregor and Matthews, 2012), one monomorphic Lister 427 90.13 cell line and one pleomorphic AnTat 1.1 90.13 cell line, each with a CAT reporter under the control of the PAD1 3’UTR (see section 5.7).

2.7.1 Sample preparation

Either the AnTat CAT-PAD or the 427 CAT-PAD reporter cell line was pelleted, washed once in HMI-9, and was resuspended in a mixture of conditioned medium or control medium and HMI-9 at a density of 2x10^5 cells/ml. Cultures were incubated for 3 days at 37°C (and 5%CO₂) without passage. Each day cell number was estimated using a Beckman Coulter Z2 Coulter Particle count and size analyser (or haemocytometer if there was a number of dead cells), and CAT ELISA samples were collected. For some experiments methanol-fixed slides for KN counts were prepared by spreading 30µl of culture on to a slide, which was air-dried before fixation in methanol. For experiments involving Hyp2 RNAi, induction of the RNAi with doxycycline was initiated one day before addition of conditioned medium, and was maintained throughout the experiment.

To prepare the day 0 samples, 1 or 2x10⁶ parasites were resuspended to a volume of 5ml with HMI-9. Two replicates were prepared for each cell line. Parasites were pelleted (1400 x g, 5 minutes) and washed three times in 1x PBS. After the last wash cells were resuspended in 1ml CAT lysis buffer (Roche, made up from 5x stock by diluting with dH₂O), and were incubated at room temperature for 25 minutes. The lysis reaction was then centrifuged (16,600 x g, 5 minutes) to pellet debris and the supernatant was transferred to 1.5ml eppendorf tubes as 500µl aliquots. Samples were snap frozen in liquid nitrogen and stored at -80°C.
To prepare samples on days 1-3 of the experiment 5ml of culture was pelleted (1400 x g, 5 minutes) and washed three times in 1x PBS. The cell count of each culture was recorded when the sample was taken in order to normalise CAT concentration measurements to cell number. Washed cells were resuspended in 1ml CAT lysis buffer (Roche, made up from 5x stock by diluting with dH₂O), and were incubated at room temperature for 25 minutes. The lysis reaction was then centrifuged (16,600 x g, 5 minutes) to pellet debris and the supernatant was transferred to 1.5ml eppendorf tubes as 500µl aliquots. Samples were snap frozen in liquid nitrogen and stored at -80°C.

2.7.2 CAT ELISA
Samples were analysed by CAT ELISA (Roche, 11363727001) to determine their CAT concentration according to the manufacturer’s instructions. Each sample was loaded into two wells of a 96 well plate. The CAT concentration of the samples was estimated by comparing the absorbance at 405nm to that of a CAT standard curve (provided by the manufacturer). The standard curve included a range of CAT concentrations of 0.0625ng/ml CAT to 2ng/ml CAT, as well as a blank of 0ng/ml. Absorbance was measured using a BioTek ELx808 Absorbance Microplate reader with Gen5 data analysis software (BioTek). Reads were taken at various times after addition of POD substrate and the read out chosen for analysis was that for which the standard curve had an R² value closest to 1, and for which the reads fitted within the standard curve. If reads did not fit within the standard curve then alternative dilutions of the samples were prepared and the ELISA was repeated. Reads of duplicate wells at 405nm were averaged and converted to CAT concentrations using the standard curve. CAT concentration/ cell was calculated using the number of cells in each 5ml sample collected during the experiment. Fold change in CAT concentration/ cell was calculated for each condition relative to day 0.

2.8 Trypanosome infections
All trypanosome infections were carried out in MF1 mice. Mouse husbandry was carried out by staff of the MARCH Building, University of Edinburgh. Mouse
handling was carried out by Julie Young and Keith Matthews according to the
conditions of Home-Office personal and project licenses.

In experiments that followed differentiation of *T. b. brucei* from slender to stumpy
forms, for generation of *T. b. brucei* RNA-seq samples, and for co-infection studies
mice were treated with 25mg/ml cyclophosphamide at least two hours prior to
infection. When only *T. b. brucei* slender forms were required, for *T. congolense*
chronic infection studies, or for generation of *T. congolense* RNA-seq samples
cyclophosphamide was not used.

Infections were usually monitored daily from day 3 post-infection by tail snip. On
the final day of an experiment, total blood was harvested from mice by cardiac
puncture using a MicroLance 0.6 x 25mm needle and a 2ml syringe containing 250µl
of 2% sodium citrate. Trypanosomes were purified from whole blood by passage
through a DE52 column (Whatman® anion exchange cellulose, Z742600) at pH 7.8.

### 2.9 Monitoring growth of cell lines in vivo

Infections were generally monitored from day 3 post-infection. Each day
parasitaemia was monitored by tail snip, and a wet blood smear was prepared as well
as two dry blood smears. The wet blood smear was used to estimate parasitaemia
based on the Herbert and Lumsden rapid matching method (Herbert and Lumsden,
1976). The dry blood smears were fixed in ice-cold methanol and stored at -20°C.
Occasionally, 10µl of blood was collected on each day and stored on ice. Cells were
then pelletted at 2450 x g for 5 minutes and washed in 200µl cold PBS. The cells
were resuspended in 125µl cold PBS and 125µl 8% paraformaldehyde, and were
fixed on ice for 10 minutes. The paraformaldehyde was then removed and fixed cells
were resuspended in 130µl 0.1M glycine in PBS. The cells were incubated at least
overnight at 4°C before resuspension in 1xPBS of a volume dependent on the
number of cells.

#### 2.9.1 Cell cycle analysis

Blood smears that had been fixed in methanol were processed as in section 2.5.3.
2.9.2 Immunofluorescence analysis of protein expression
Methanol fixed slides were processed as in section 2.5.4.
Paraformaldehyde fixed cells were processed as in section 2.5.4.

2.10 Induction of differentiation to procyclic forms \textit{in vitro}
Parasites were purified from blood by passage through a DE52 column (Whatman® anion exchange cellulose, Z742600). Purified cells were resuspended at 2x10^6/ml in SDM79 media (GIBCO by Life technologies) containing 6mM cis-aconitate (Sigma, A3412) and were incubated at 27°C. Samples were collected for flow cytometry at 0, 4 and 24 hours.

2.11 Flow cytometry
2-5x10^6 cells were pelleted in FACS tubes (7 minutes, 1350 x g) and washed twice in 1xPBS before resuspension in 500µl FACS fix, and storage at 4°C overnight. The cells were then washed three times in 1xPBS and were resuspended in 200µl 2% BSA:PBS for 30 minutes. The cells were then resuspended in 200µl primary antibody diluted in 2% BSA:PBS (αPAD1 was diluted 1:200, αEP procyclin (Cedar Lane laboratories) was diluted 1:500) and were incubated overnight at 4°C. The cells were washed twice in 1x PBS and were resuspended in 200µl secondary antibody diluted in 2% BSA:PBS (α-rabbit CY5 and α-mouse FITC were each diluted 1:1000). The cells were washed twice in 1x PBS and were finally resuspended in 500µl PBS containing 0.02µg/ml DAPI. Samples were then processed on an LSRII Flow Cytometer (BD Biosciences). Positive controls and secondary antibody only controls were included for all experiments. Analysis was performed using FlowJo software (Tree Star).

2.12 Western Blot
Protein samples were prepared as follows: cells were pelleted and washed once in cold 1xPBS and were then resuspended in 2x laemmli buffer with β-mercaptoethanol. 3x10^6 cells were resuspended in 15µl laemmli buffer. Samples for detection of Hyp2 were boiled at 95°C for 5 minutes. Samples were stored at -20°C.
The components of the separating solution were combined and poured between two glass gel plates (1.5mm spacer and short plate), and overlaid with isopropanol to polymerise. An 8%, 10% or 12% separating gel was prepared depending on the size of the protein of interest. After removal of the isopropanol, the TEMED was added to the stacking solution and this was poured on top of the polymerized resolving gel. A 10 or 15-toothed comb was inserted into the stacking gel, which was then left to polymerise.

Polymerized gels were assembled in the BioRad mini Protean II system. Tris-glycine electrophoresis buffer was used to fill the central reservoir, and the surrounding tank was filled to 50%. 15µl of sample was loaded per well, alongside 5µl Novex® Sharp pre-stained protein standard (ThermoFisher Scientific LC5800). Gels were run at 125V for 1-2.5 hours, until the bromophenol blue just ran off the bottom of the gel.

Gels were extracted from the electrophoresis apparatus and assembled in a transfer cassette for transfer to an Amersham™ Protran™ Premium 0.45µm nitrocellulose membrane (GE healthcare, 10600003). The transfer cassette was assembled in the Mini-Protean Tetra blotting module (BioRad, 1660827EDU), which was then filled with transfer buffer and an ice pack. Transfer was carried out for 1 hour at 80V.

Following extraction of the nitrocellulose membrane from the transfer apparatus Ponceau stain was applied to monitor even loading and transfer. Ponceau stain was removed by washing in 1xPBS with 0.05% Tween. The membrane was then blocked in 5% non-fat milk solution (prepared in 1xPBS 0.05% Tween) for 30 minutes to 1 hour. The blocking solution was removed and the primary antibody added, and this was incubated at 4°C overnight. Primary antibody was made up in 5% milk solution. An initial stock of BB2 antibody (Bastin et al., 1996) was used at a 1:4 dilution, but when a fresh stock of αBB2 antibody was prepared from hybridoma cells this was used at 1:20. αPAD1 antibody (Dean et al., 2009) was used at 1:1000. αG6PDH (glucose-6-phosphate dehydrogenase, kind gift of Paul Michels) was used for loading controls at 1:10,000. αEF1 (elongation factor 1-alpha, Merck Millipore 05-235) was used for loading controls at 1:7000.
Primary antibody solution was removed, and was stored at -20°C and was frequently re-used. The membrane was washed three times in 1xPBS 0.05% Tween, before incubation for 1 hour with the secondary antibody prepared in 5% milk. Anti-mouse HRP and anti-rabbit HRP were each used at 1:8000. The membrane was again washed three times in 1xPBS 0.05% Tween before detection. Pierce ECL Western blotting substrates (ThermoScientific 11884584) were mixed together according to the manufacturer’s instructions before applying the substrate to the membrane. After 5 minutes incubation, excess substrate was removed and the membrane was placed into a plastic wallet in a developing cassette. In a darkroom, photographic films (Fujifilm, Fisher 12705325) were exposed to the membrane for various lengths of time (seconds to minutes), and the films were developed using a Konica Minolta SRX-101A developer.

Alternatively the Li-cor detection system was used. In this case the method detailed above was followed up until addition of secondary antibody. Secondary antibody was diluted in 50% 1xPBS and 50% Li-cor blocking buffer. Both anti-mouse (IRDye® 680 goat anti-mouse, Li-cor) and anti-rabbit (goat anti-rabbit IgG (H+L) Dylight 800, ThermoScientific) secondary antibodies were diluted 1:8000. Immunofluorescence was detected on the Li-cor Odyssey imaging system.

In order to reprobe a blot with a different primary antibody, the membrane was first stripped by incubation in stripping buffer for 10 minutes, before reapplying blocking solution and proceeding as above through to detection.

### 2.13 Northern Blot

RNA samples were prepared using the Qiagen RNeasy kit (Qiagen, 74106) according to the manufacturer’s instructions. Briefly no more than 5x10⁷ cells were pelleted and resuspended in 594μl RLT buffer with 6μl β-mercaptoethanol. Samples were stored at -80°C before processing according to the protocol ‘Purification of Total RNA from Animal cells using spin technology’. RNA samples were resuspended in 10μl RNase free water per 10⁷ cells used to prepare the sample. The concentration and purity of the resultant RNA was measured on a Nanodrop spectrophotometer. RNA samples were stored at -80°C.
Prior to use, all equipment that would come into contact with the RNA gel or membrane were thoroughly washed with hot water and detergent and then 70% ethanol.

Agarose gels were prepared by dissolving 1.8g of agarose (Bioline BIO-41026) in 130.5ml of distilled water, before addition of 15ml 10x MOPS. Once this had cooled to approximately 60°C, 4.5ml of 37% formaldehyde was added and the gel left to set in a chemical hood.

A master mix was prepared consisting of 9µl formamide, 3µl formaldehyde (37%), 2µl 10x MOPS and 2µl RNA gel loading buffer per sample. 16µl volumes of the master mix were taken into eppendorf tubes and 1µg of RNA sample was added to each tube, followed by heating at 65°C for 5 minutes. RNA samples were then loaded into the prepared agarose gel, which was run at 150V in 1x MOPS running buffer for 90 minutes. The gel was then stained with 20µl Ethidium bromide (10mg/ml) in 200ml 1x MOPS for 15 minutes on a platform rocker. The gel was destained by three 30 minute incubations with water on the platform rocker. The quality and equal loading of the RNA samples was determined by visualising the rRNA bands on a UV transilluminator.

The transfer was assembled as illustrated and left overnight:
The following day the nylon membrane (Roche, 000000011209299001) was allowed to dry, before UV cross-linking in a Stratalinker® UV crosslinker at 0.12 joules. The blot was placed into a hybridisation tube along with 10ml hybridisation buffer and was incubated for 1 hour at 68°C in a hybridisation oven. 1.5µl of probe was added to 100µl of hybridisation buffer and was boiled for 7 minutes. The boiled probe was added to 7ml prewarmed (68°C) hybridisation buffer, and this solution was used to replace the prehybridisation solution already on the blot. The blot was left overnight at 68°C for hybridisation. Probes used in this thesis included one to detect TbHyp2 (prepared by Paula MacGregor) and one to detect PAD1 (prepared by Eva Rico Vidal).

The following day the blot was washed twice for 30 minutes at 68°C in 2xSSC/0.1%SDS, and once for 30 minutes at 68°C in 0.5xSSC/0.1%SDS. The blot was then washed for 1 minute in wash buffer at room temperature, before being blocked for 1 hour in 50ml Maleic acid buffer with 1% DIG block. The blot was then incubated for 30 minutes in 50ml of Maleic acid buffer with 1% DIG block and 2µl anti-DIG (Roche, 000000011093274910, Anti-DIG was centrifuged for 1 minute at 11,300 x g before use). The blot was then washed three times in Maleic acid buffer with 0.3% Tween 20 for 10 minutes each wash, before being soaked for 2 minutes in
Detection buffer. Excess detection buffer was removed and the blot was placed in Bag W (a polyethylene bag), and CDP-Star (Roche, 000000011685627001) diluted 1:100 in detection buffer was applied to the blot and incubated for 2 minutes. Excess CDP-Star detection reagent was removed and the blot was sealed in the Bag W using a heat sealer. The sealed blot was incubated at 37°C for 15 minutes, before being placed in a developing cassette. In a darkroom, photographic films (Fujifilm, Fisher 12705325) were exposed to the membrane for various lengths of time (seconds to minutes), and the films were developed using a Konica Minolta SRX-101A developer.

2.14 Southern Blot
A gene probe was produced to detect the presence of the TbHyp2 gene, and a 5’UTR probe was designed to detect the correct integration of the knock out construct. To generate the 5’UTR probe Hyp2 5UExtF and 5UIntR primers (Appendix 2.4) were used in a PCR reaction with Q5 high fidelity DNA polymerase (NEB, M0491S). To generate the gene probe TbHyp2 gene primers were used (Forward primer: GGG TTT ACT AGT ATG GCA TCG GAG GCA GCG, Reverse primer: GGG TTT GGA TCC TTA TTC GCC CCT AAC TG). Multiple PCRs were combined to generate enough material for the probes. Where only one band was visualised after running the PCRs on a gel, the reactions were combined and made up to a volume of 200µl with water. When multiple bands were visualised after running the PCRs on a gel, the bands of expected size were first excised and purified by squeezing the DNA from the gel between two sheets of Parafilm® (Sigma, P7793-1EA), and making the resultant liquid up to a volume of 200µl with water. The DNA was then precipitated by addition of 20µl 3M sodium acetate and 550µl 100% ethanol, vortexing, and storage at -80°C for at least 1 hour. The precipitated DNA was then pelleted at 4°C for 10 minutes at 18,000 x g, and the pellet was resuspended in 16µl dH2O. The DNA was quantified using a Nanodrop spectrophotometer and 1µg of DNA was made up to 16µl with dH2O. The DNA was denatured by heating to 100°C for 10 minutes and was then quickly chilled on ice. 4µl DIG High Prime (vial 1 of DIG High Prime labelling and detection starter kit II, Roche, 000000011585614910) was added to the denatured DNA and the reaction was incubated at 37°C for 3 hours. The DIG-
labelling reaction was stopped by addition of 2µl 0.2M EDTA (pH8) and by heating to 65°C for 10 minutes. DIG-labelled probes were then stored at -20°C.

Genomic DNA was extracted from parental AnTat 1.1 90.13 parasites, from suspected single knock out parasites for TbHyp2, and from suspected double knock out parasites for TbHyp2. 1µg of DNA was digested overnight at 37°C by PstI (Promega, R6111) according to the manufacturer’s instructions.

Digested DNA was loaded on to a 0.8% 150ml agarose gel (containing 5µl Ethidium bromide). Each 50µl digest was combined with 10µl 6x DNA loading buffer and was divided equally between duplicate wells. Additionally 1µg of undigested DNA was loaded as a control. The gel was run for 3 hours at 100V and an image of the gel was captured alongside a ruler. The gel was soaked for 15 minutes in 250ml depurination solution (0.25M HCl); the gel was then soaked twice for 15 minutes in 250ml denaturation solution (1.5M NaCl/ 0.5M NaOH). The gel was next rinsed in deionized water before being soaked twice for 15 minutes in neutralisation solution (1M Tris/ 1.5M NaCl/ pH7.4). The gel was then assembled for transfer of the DNA to a nylon membrane overnight using the same method of assembly as for the Northern Blot (section 2.13).

The following day the Southern blot was UV cross-linked in a Stratalinker® UV crosslinker at 0.12 joules. Pre-hybridisation of the membrane was carried out with pre-warmed DIG Easy Hyb Buffer (Roche, 000000011603558001) in a hybridisation oven at 42°C for 30 minutes. 1µl of the DIG-labelled DNA probe was added to 50µl Hyb buffer and this was boiled for 5 minutes and then rapidly cooled on ice before addition of the denatured probe to 7ml pre-warmed Hyb buffer. The probe was then used to replace the pre-hybridisation solution, and hybridisation was carried out overnight in a hybridisation oven at 42°C.

The following day the membrane was washed twice for 5 minutes in 2x SSC/0.1% SDS at room temperature. The membrane was then returned to a hybridisation tube and washed twice for 15 minutes in 0.5x SSC/0.1% SDS (pre-warmed) at 68°C. Boxes that had been cleaned with detergent and ethanol were used for later steps. The membrane was rinsed in wash buffer (1x Maleic acid buffer/0.3% Tween 20) for
1 minute at room temperature, before being blocked for 1 hour in Maleic acid buffer containing 1% DIG Block. The membrane was then incubated for 30 minutes in Maleic acid buffer containing 1% DIG Block and 2µl Anti-DIG (Roche, 000000011093274910, Anti-DIG was centrifuged for 1 minute at 11,300 x g before use). The blot was washed three times for 10 minutes in wash buffer, and was then soaked for 2 minutes in detection buffer. Excess detection buffer was removed and the blot placed inside a Bag W, followed by application of 1ml CDP-Star (Roche, 000000011685627001, diluted 1:100 in detection buffer) and incubation for 2 minutes. Excess CDP-Star was removed and the blot sealed inside the Bag W using the heat sealer, before incubation at 37°C for 15 minutes. Films were developed as for a Northern Blot (section 2.13), with 5-8 minutes exposure.

2.15 RNA-seq analysis
RNA samples were shipped on dry ice to BGI Tech solutions in Hong Kong for RNA-seq analysis. The quality of the transcriptome data produced by BGI Hong Kong was evaluated by Dr Alasdair Ivens using the Fast QC report programme (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The reads were aligned to either the *T. b. brucei* TREU927/4 or the *T. congolense* IL3000 genome by Dr Alasdair Ivens using the Bowtie 2 programme (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml, (Langmead and Salzberg, 2012)). A group-wise comparison using the R/Bioconductor software package limma (Ritchie et al., 2015) was performed between the ascending parasitaemia and peak parasitaemia replicates.

2.16 Bioinformatic analysis
The BLASTP tool on TritrypDB (Aslett et al., 2010) was used for identification of orthologues. Conserved domains in proteins of interest were identified using various tools: InterPro (Mitchell et al., 2015), ScanProsite (de Castro et al., 2006), Gene3D/CATH database (Sillitoe et al., 2015), Superfamily/SCOP database (Andreeva et al., 2007), PFAM (Finn et al., 2014), Conserved Domain Database (Marchler-Bauer et al., 2015), Eukaryotic Linear Motif database (Dinkel et al., 2012), I-TASSER (Yang et al., 2015), Phobius (Kall et al., 2004), TMHMM (Krogh et al., 2001), Panther (Mi et al., 2013) and PRINT (Attwood et al., 1994).
2.17 Statistical analysis and production of figures

All statistical analyses were carried out in GraphPad Prism version 6 (GraphPad Software, La Jolla, California, USA, www.graphpad.com), except for the general linear models used to analyse the T. congolense chronic infection data which were performed using Minitab® Statistical Software. A number of the analyses used work on the assumption of a normal distribution; therefore the distribution of the data was checked before performing the analyses, and if required the data was transformed (e.g. by logarithm) prior to testing. In all analyses a p-value <0.05 was considered significant. Guidance on statistical testing was received from Dr Margo Chase-Topping.

All graphs were produced using GraphPad Prism version 6. Images were manipulated (i.e. cropping, brightness, contrast, overlay, scale bar application) using Image J 64 (Rasband, 1997-2015). When cell counts were performed using Image J 64 the Cell Counter Plugin was used.
Chapter 3: Characterisation of *T. congolense* bloodstream infection
3.1 Introduction

*T. b. brucei* parasites have evolved a life cycle strategy that promotes survival and transmission. These cells respond to the unidentified quorum-sensing signal(s), SIF, which accumulates at high parasitaemia, by differentiating from the proliferative slender form to the growth-arrested, transmissible, stumpy form. This differentiation strategy in combination with the periodic clearance of expressors of the predominant variant surface glycoprotein (VSG) results in the characteristic wave-like parasitaemia profile of a *T. b. brucei* infection. Differentiation of *T. b. brucei* parasites to the stumpy forms assists in maintenance of a chronic infection by preventing high parasitaemia from overwhelming the host. The advantages associated with *T. b. brucei* pleomorphism in enabling parasite survival and transmission would be expected to be similarly advantageous for *T. congolense* parasites that share hosts and a vector with *T. b. brucei*. However, there is conflicting evidence for and against the presence of pleomorphism in *T. congolense* bloodstream infections. Therefore, we have investigated whether there are changes in the morphology, or cell cycle arrest, during infection with *T. congolense* bloodstream forms of the IL3000 strain, and whether any observed changes were associated with particular points in the parasitaemia.

Transcriptomic studies in *T. b. brucei* have compared the abundance of transcripts from different points of the parasite life cycle. Kabani *et al* compared mRNA abundance in slender, stumpy and procyclic forms using whole-genome microarray analysis (Kabani et al., 2009). Another study used Nimblegen array analysis to investigate changes in mRNA abundance during these life cycle stages (Jensen *et al*., 2009). Nilsson *et al* developed a spliced leader trapping method to identify variations in transcript abundance between slender, stumpy and procyclic forms, as well as variability in alternative splicing patterns between these life cycle stages (Nilsson *et al*., 2010). Here, I have used RNA-seq analysis to identify transcripts that vary in abundance between *T. congolense* parasites from ascending parasitaemia as compared to those from peak parasitaemia. Additionally, this *T. congolense* data set has been compared to a new *T. b. brucei* slender and stumpy data set that we have generated using the same methodology. The goal was to establish the extent to which
*T. congolense* parasites show growth control in their mammalian host and their molecular characteristics with respect to *T. b. brucei* stumpy forms.

### 3.2 *T. congolense* bloodstream forms display evidence of growth control at the peak of parasitaemia

Infections with *T. congolense* IL3000 were initiated in 6 MF1 mice and were monitored over 21 days. An initial peak of parasitaemia was typically followed by clearance of parasites and emergence of a second peak (Figure 3.1). The cell cycle status of the parasites was monitored at various points during the infections.

The percentage of 2K1N, 2K2N cells was markedly reduced at peak parasitaemia relative to the ascending or descending parasitaemia in several instances, but this was not always observed. When the percentage of 2K1N, 2K2N cells is plotted against the parasitaemia, scatter plots in some cases showed a negative association between the two measures, but this was not always the case and data points were dispersed (Figure 3.2).

The effect of parasitaemia on the percentage of proliferating (2K1N, 2K2N) cells was investigated with a general linear model using Minitab® Statistical Software. Mouse was included as a random factor and the effect was found not to be significant (p=0.119). Parasitaemia had a significant effect on the percentage of proliferating (2K1N, 2K2N) cells (p=0.013) when all 21 days of the infections were taken into account. Additionally, the effect of parasitaemia on the percentage of 2K1N, 2K2N cells was investigated for just the first 14 days of infection, in order to focus on the first wave of parasitaemia.
Figure 3.1 The parasitaemia profiles for each *T. congolense* infection, including the proportion of proliferating cells. A) Representative images illustrating the appearance of cells with 1 kinetoplast (K) and 1 nucleus (N) (G0/G1, plus S-phase cells), 2K1N (G2-phase cells) or 2K2N (post-mitotic cells). Scale bar represents 10µm. B) Each profile represents the parasitaemia for one infection of an MF1 mouse (black line). In most cases an initial peak of parasitaemia, was followed by clearance of parasites and then a second peak. Blue bars represent the percentage of cells on a given day with the configuration 2K1N or 2K2N. On each day 500 cells were assessed for KN configuration, abnormal 0K1N cells were occasionally observed (on average 15 per day, range 0-56), but these were not included in the 500 cells counted. In some cases (marked *) when parasites were of low density only 200 cells were counted.
During the first 14 days of infection a reduction in proliferating (2K1N, 2K2N) cells was observed at the peak of parasitaemia, and there was a negative correlation between the two measures when multiple days were taken into account (Figure 3.3). The significance of the effect of parasitaemia on the percentage of proliferating (2K1N, 2K2N) cells was investigated using a general linear model that incorporated mouse as a random factor. As with the entire time course mouse was not a significant factor (p=0.553). However, parasitaemia was found to have a significant effect on the percentage of proliferating (2K1N, 2K2N) cells (p=0.001) when the first 14 days of infection were taken into account.

Figure 3.2 Scatter plots of the percentage of cells that have a 2K1N or 2K2N configuration plotted against parasitaemia for the entire 21 days of infection. In some infections (2, 3, 4 and 5) the percentage of cells that are 2K1N or 2K2N correlated negatively with the parasitaemia, but this was not always the case. Lines of best-fit were applied in GraphPad Prism.
Figure 3.3 Scatter plots of the percentage of cells that have a 2K1N or 2K2N configuration plotted against parasitaemia for the first 14 days of infection. In all infections the percentage of cells that are 2K1N or 2K2N correlated negatively with the parasitaemia. Lines of best-fit were applied in GraphPad Prism.

In addition to cell cycle analysis, variations in parasite morphology were considered when observing hundreds of cells on the various days of parasitaemia for which KN configurations were assessed. Representative images were captured for numerous days of the infections as a record of parasite morphology (Figure 3.4).
Figure 3.4 Images captured during the *T. congolense* infections. The scale bar represents 10µm. The images on the upper row were captured during ascending parasitaemia, those on the middle row were captured at peak parasitaemia, and those on the lower row were captured during descending parasitaemia. Despite some variation in size and morphology on any given day, this did not appear to be associated with the parasitaemia.

Pronounced variations in parasite size and morphology were not observed at any particular level of parasitaemia (measurements were not made), indicating that in terms of morphology, these parasites appeared monomorphic. Nevertheless the significant correlation between parasitaemia and the percentage of proliferating cells, and the apparent reduction in percentage of proliferating cells coinciding with peak parasitaemia may be indicative of a growth control strategy driven by the parasite itself. This is comparable to the accumulation of growth-arrested stumpy forms at the peak of *T. b. brucei* parasitaemia. The proportion of *T. b. brucei* cells that are growth arrested at the peak of parasitaemia is ordinarily higher than that observed in these *T. congolense* infections. However, this may not be the case for all *T. congolense* strains. It is possible that the strain we have used is more monomorphic than field isolates, due to long-term maintenance in the laboratory.
3.3 Preparation of RNA samples from ascending and peak *T. congolense* parasitaemia for RNA-seq analysis

Growth arrest at peak *T. congolense* parasitaemia could be accompanied by further differences between parasite populations in ascending or at peak parasitaemia. Therefore RNA samples for use in RNA-seq analysis were prepared from populations of parasites taken from ascending parasitaemia, as well as parasites collected at peak parasitaemia.

RNA samples were collected from three *T. congolense* infections at peak parasitaemia and three replicate samples were also generated for the ascending parasitaemia. In order to generate enough RNA, parasites from more than one infection had to be combined for ascending parasitaemia samples. Before sending RNA samples for RNA-seq analysis, 1µg of each RNA sample was loaded on a denaturing agarose gel, and ethidium bromide staining revealed the 5 bands expected for *T. congolense* rRNA in all 6 samples (3 ascending, 3 peak), indicating that the samples were not degraded (Figure 3.5A). Additionally the proportion of proliferating cells was investigated for each of the infections used to generate the RNA samples. As expected there was a reduction in the proportion of proliferating cells (2K1N, 2K2N) at the time point when the peak parasitaemia RNA samples were collected relative to earlier time points in these infections (Figure 3.5B). Additionally, *T. congolense* cells collected from ascending parasitaemia had greater proportions of 2K1N or 2K2N cells than cells collected from peak parasitaemia (except for peak sample 3) (Figure 3.5C). RNA samples were quantified by nanodrop (Figure 3.5D) and were deemed to be of sufficient quantity and quality for RNA-seq analysis by BGI Tech Solutions Hong Kong. Quality was verified by BGI, who provided a quality report on receipt of the RNA samples (not shown).
control

3.4.1 Analysis of T. congolense ascending and peak parasitaemia RNA-seq data

Once RNA-seq reads were received from BGI, Dr Alasdair Ivens performed quality control before aligning reads to the T. congolense genome. Quality control was
performed using the Fast QC report programme (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Paired end reads of 90 bases in length met quality control criteria for sequence quality scores and GC content except for the first 10 bases of each read. Therefore the first 10 bases from each read were trimmed before alignment of the reads with the most up-to-date *T. congolense* genome available from the Sanger institute (ftp://ftp.sanger.ac.uk/pub/project/pathogens/gff3/CURRENT/). The alignment was carried out using the Bowtie 2 programme (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml, (Langmead and Salzberg, 2012)), and was not restricted to annotated genes. Read counts were normalised to reads per kb/ map (rpkm), in order to account for size of gene and the different read depths of the replicates. A joint quantile threshold cut-off of 10% was applied across the relevant replicate rpkm values, in order to remove the lowest numbers of reads. A group-wise comparison using the R/Bioconductor software package limma (Ritchie et al., 2015) was performed between the ascending parasitaemia and peak parasitaemia replicates. The output table reports for each transcript detected the fold change (FC) between peak and ascending replicates as Log2FC, so that a value of 1 or greater indicates a transcript at least 2 fold more abundant in one category than the other, and a value of -1 or less indicates a transcript at least 2 fold less abundant in one category relative to the other. The comparison performed was ‘peak’ relative to ‘ascending’, so that a positive Log2FC indicates a transcript that is more abundant at peak parasitaemia than during ascending parasitaemia, and a negative Log2FC indicates a transcript that is less abundant at peak parasitaemia than during ascending parasitaemia. When an adjusted p-value cut-off of p<0.05 was applied to the data, 804 transcripts were found to have significant changes in abundance in peak relative to ascending data sets. Average transcript levels at peak versus ascending parasitaemia for each of the 804 transcripts with significant changes were plotted (Figure 3.6). Notably, the magnitude of the fold changes were larger for transcripts that were more abundant at peak parasitaemia, compared to those that were more abundant in ascending parasitaemia, for which only smaller (<2 fold) significant changes were observed. The 804 transcripts were further sorted to exclude transcripts where the magnitude of change was less than 2 fold, and this generated a short-list of 372 genes. When
transcripts for which no protein product was annotated were excluded 322 genes remained. This short-list was reduced to 170 genes when transcripts described as VSG were excluded; of these 170 genes, 118 were ≥3 fold more abundant in peak relative to ascending parasitaemia, and 65 were ≥4 fold more abundant in peak relative to ascending parasitaemia.

Figure 3.6 Schematic workflow of the analysis of RNA-seq data following alignment to the *T. congolense* genome. Once reads had been mapped to the *T. congolense* genome, fold changes in transcript abundance for peak relative to ascending parasitaemia were calculated. Transcripts for which the fold change had an adjusted p-value of <0.05 were considered significant. When these transcripts were short-listed by removal of transcripts described as not having a protein product or as VSG, 170 genes were ≥2 fold more abundant in peak relative to ascending parasitaemia. The 804 transcripts that were significantly changed in abundance are plotted in the graph on the right, with ‘peak’ abundance on the x-axis and ‘ascending’ abundance on the y-axis. Transcripts described as VSG are in orange and all other transcripts are in blue.

When genes with significant changes in abundance were ordered by the magnitude of the fold change, the gene with the largest increase in abundance in the peak relative to ascending population was described as a hypothetical protein (Table 3.1).
The genes with the greatest increase in abundance in parasites from peak relative to ascending parasitaemia.

The 804 genes that showed significant changes between peak and ascending parasitaemia were grouped by the description of their protein product. The category that was most represented for transcripts more abundant at peak parasitaemia were ‘VSG’, followed by ‘hypothetical protein’ (Figure 3.7A). The predominant categories for transcripts with reduced abundance at peak parasitaemia were ‘hypothetical protein’, followed by ‘other’ (Figure 3.7B). Some of the transcripts with reduced abundance at peak parasitaemia that fell into the ‘other’ category (including paraflagellar rod protein, DNA polymerase catalytic subunits, and a putative cell division protein kinase) were consistent with the reduced proliferation observed at peak parasitaemia. Three transcripts described as putative amino acid transporters were also found to have reduced abundance at peak parasitaemia (Appendix 1.4). Transcripts that were increased in abundance at peak parasitaemia and were classified as ‘other’ are listed in Appendix 1.5.

Further analysis was required to investigate potential roles for the majority of transcripts that varied between peak and ascending parasitaemia that were described as hypothetical proteins.

<table>
<thead>
<tr>
<th>SysID</th>
<th>T. congolense Annotation</th>
<th>Fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td>TcIL3000_0_23590</td>
<td>hypothetical protein, conserved</td>
<td>20.18</td>
</tr>
<tr>
<td>TcIL3000_0_34550</td>
<td>ESAG protein, putative</td>
<td>12.41</td>
</tr>
<tr>
<td>TcIL3000_0_41620</td>
<td>hypothetical protein, conserved</td>
<td>11.53</td>
</tr>
<tr>
<td>TcIL3000_0_43500</td>
<td>hypothetical protein, conserved</td>
<td>10.66</td>
</tr>
<tr>
<td>TcIL3000_0_13130</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>10.47</td>
</tr>
<tr>
<td>TcIL3000_0_35150</td>
<td>invariant surface glycoprotein, putative</td>
<td>10.18</td>
</tr>
<tr>
<td>TcIL3000_0_01540</td>
<td>ESAG protein, putative</td>
<td>9.89</td>
</tr>
</tbody>
</table>

Table 3.1 The genes with the greatest increase in abundance in parasites from peak relative to ascending parasitaemia.
Figure 3.7 Transcripts with significantly changed abundance between ascending and peak parasitaemia that have been grouped by product description. A) Protein description categories for transcripts significantly more abundant at peak parasitaemia than ascending parasitaemia. B) Protein description categories for transcripts with significantly reduced abundance at peak parasitaemia relative to ascending parasitaemia.
3.4.2 Identification of *T. b. brucei* orthologues of the *T. congolense* genes

A BLASTP search against *T. b. brucei* 927 proteins for all *T. congolense* proteins whose gene IDs were listed in the RNA-seq results table was carried out by Dr Alasdair Ivens. The top hit *T. b. brucei* orthologue was incorporated into the RNA-seq results table to aid in interpretation of the function of genes with significantly altered transcript abundance where the *T. congolense* annotation was limited.

A good orthologue of the *T. b. brucei* stumpy enriched protein PAD1 was identified in *T. congolense*, but not in *T. vivax* (Appendix 1.1). If the *T. congolense* PAD1 orthologue, TcIL3000.0.18180, had the same role as *T. b. brucei* PAD1 it might be expected that the *T. congolense* transcript would be more abundant at the peak of parasitaemia if there was an accumulation of transmissible forms at the peak of *T. congolense* parasitaemia. However, this *T. congolense* PAD1 orthologue was not included in the list of transcripts significantly more abundant in peak relative to ascending parasitaemia. This does not rule out a pattern of differential expression for the *T. congolense* PAD1 orthologue in the bloodstream. It is possible that parasites collected from ascending parasitaemia samples were already expressing the PAD1 orthologue transcript, as an intermediate *T. b. brucei* population would express the PAD1 transcript but not the protein. This would be consistent with the relatively high average expression of the PAD1 orthologue detected in the *T. congolense* RNA-seq data. Additionally, a BLASTP search of the TcIL3000.0.18180 amino acid sequence identified a number of proteins in *T. congolense* of very similar sequence, some of which had not been assigned chromosomes, however some were found on chromosome 7 (the location of the *T. b. brucei* PAD array). If the *T. congolense* PAD1 orthologue is part of a closely related family, as for the *T. b. brucei* PAD family, this could mask the expression profile of individual members of that family.

Genes with significant changes in abundance of the greatest magnitude were annotated with information from GeneDB (Logan-Klumpler et al., 2012). These annotations included whether the *T. congolense* protein or *T. b. brucei* orthologue were part of the cell surface phylome, or, in the case of the *T. b. brucei* orthologue, whether loss-of-fitness was observed following RNAi at any life cycle stage, but
particularly differentiation. The annotation was applied to all transcripts at least 4-fold more abundant in peak relative to ascending parasitaemia (Appendix 1.2), as well as the eight transcripts showing the greatest reduction in abundance in peak relative to ascending parasitaemia (Appendix 1.3).

In the case of the *T. congolense* transcripts more abundant at peak parasitaemia, RNAi targeting some of the identified *T. b. brucei* BLASTP orthologues were found to cause a loss-of-fitness in the transition from bloodstream to procyclic forms (Alsford et al., 2011). This could be the result of a loss-of-fitness in either life cycle stage, or an important role for the protein in differentiation itself. Of the *T. b. brucei* orthologues with a possible role in differentiation some did not have a low e-value, indicating that these may not be good orthologues of the *T. congolense* transcripts. More convincing *T. b. brucei* orthologues were procyclin-associated gene 5 (PAG5), and expression-site associated gene 2 (ESAG2), but RNAi targeting these transcripts was also associated with defects in bloodstream and/or procyclic form proliferation, so these may not be important for the process of differentiation itself. Nevertheless, the identification of procyclin-associated genes as orthologues of *T. congolense* transcripts enriched at peak parasitaemia could indicate that these parasites were in preparation for transmission. One *T. congolense* transcript with a good orthologue was TcIL3000.1.10, the identified orthologue was described as a putative UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase (Tb927.8.7140), and RNAi targeting this orthologue resulted in a loss-of-fitness during bloodstream to procyclic form differentiation, but not proliferation of the individual life cycle stages. Additionally, Tb927.8.7140 was associated with GO terms for membrane, protein glycosylation and GPI-anchor biosynthetic process, indicating that the *T. congolense* transcript could be up-regulated at peak parasitaemia in preparation for cell surface modifications required for transmission.

In many cases, the *T. congolense* transcripts that were at least 4-fold more abundant at peak parasitaemia were predicted to form part of the *T. congolense* cell surface phylome. In other cases, the *T. b. brucei* orthologues of these *T. congolense* transcripts were predicted to form part of the *T. b. brucei* cell surface phylome (in most cases these orthologues had relatively low e-values). This could indicate that
the *T. congolense* cell surface had undergone a change in composition between ascending and peak parasitaemia, or was preparing to undertake modification of the cell surface necessary for transmission.

While identification of the *T. b. brucei* orthologues of the *T. congolense* transcripts has allowed some useful comparisons to be made, it is of limited use in the case of *T. congolense* hypothetical proteins for which a good orthologue could not be found. There were a considerable number of hypothetical proteins of differing abundance between ascending and peak parasitaemia, so these were investigated for any domains indicative of function.

### 3.4.3 Predicted protein domains in hypothetical proteins that are more abundant at peak parasitaemia

Amongst the *T. congolense* transcripts at least 4-fold more abundant at peak relative to ascending parasitaemia there were 23 proteins described as ‘hypothetical’. The amino acid sequences for these 23 hypothetical proteins were searched using the InterPro search tool to identify potential domains of interest (Figure 3.8). InterPro (Mitchell et al., 2015) draws on information from several member databases for output, providing a useful overview. There was not much information available for the majority of these hypothetical proteins. However, 7 of these proteins were predicted to have one transmembrane domain by the Phobius tool (Kall et al., 2004), 3 of these predictions were also supported by TMHMM (Krogh et al., 2001). Additionally, 3 of the hypothetical proteins were predicted to have more than one transmembrane domain by Phobius, and in some cases this was supported by predictions by TMHMM. Although these hypothetical proteins were not described as part of the cell surface phylome (Jackson et al., 2013), there is a possibility that these could be cell surface expressed proteins for which expression is up-regulated at peak parasitaemia. Alternatively these hypothetical proteins may localize to internal membranes. One of the ‘unlikely hypothetical proteins’ had a predicted Trypanosome VSG, B-type, N-terminal domain, and may represent simply a VSG fragment.
<table>
<thead>
<tr>
<th>SystemID</th>
<th>Updated Description</th>
<th>logFC</th>
<th>adj.P</th>
<th>Comments from Gene DB</th>
<th>Phobius</th>
<th>HMM</th>
<th>PFAM</th>
<th>SignalP- euk</th>
<th>Panther</th>
<th>GO terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>TcIL3000_0_23050</td>
<td>hypothetical protein, conserved</td>
<td>4.335</td>
<td>0.0423</td>
<td>T.con orthologues reverse transcriptase</td>
<td>1-40 GY, 41-61 TM, 62-310 NON-CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_41620</td>
<td>hypothetical protein, conserved</td>
<td>3.5275</td>
<td>0.0235</td>
<td>T.con orthologues VSG</td>
<td>1-19 NON-CYT, 20-44 TM, 45-240 CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_43300</td>
<td>hypothetical protein, conserved</td>
<td>3.4135</td>
<td>0.0133</td>
<td>T.con orthologues hypothetical</td>
<td>1-31 NON-CYT, 32-328 TM, 330-468 CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_50570</td>
<td>hypothetical protein, conserved</td>
<td>3.0385</td>
<td>0.0133</td>
<td>T.con orthologues hypothetical</td>
<td>1-21 NON-CYT, 22-45 TM, 46-153 CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_60600</td>
<td>hypothetical protein, conserved</td>
<td>2.6998</td>
<td>0.0288</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_20630</td>
<td>hypothetical protein</td>
<td>2.6655</td>
<td>0.0201</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26040</td>
<td>hypothetical protein, conserved</td>
<td>2.6591</td>
<td>0.0133</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_24580</td>
<td>hypothetical protein, unlikely</td>
<td>2.6138</td>
<td>0.0469</td>
<td>T.con orthologues VSG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26360</td>
<td>hypothetical protein, conserved</td>
<td>2.5708</td>
<td>0.0370</td>
<td>T.con orthologues hypothetical</td>
<td>1-96 CYT, 97-115 TM, 116-134 NON-CYT, 135-151 TM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_32470</td>
<td>hypothetical protein, conserved</td>
<td>2.5665</td>
<td>0.0090</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_53060</td>
<td>hypothetical protein, conserved</td>
<td>2.5648</td>
<td>0.0238</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_03960</td>
<td>hypothetical protein, conserved</td>
<td>2.4150</td>
<td>0.0118</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_12570</td>
<td>hypothetical protein, conserved</td>
<td>2.3877</td>
<td>0.0035</td>
<td>T.con orthologues hypothetical</td>
<td>1-172 CYT, 173-192 TM, 193-198 NON-CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_60540</td>
<td>hypothetical protein, conserved</td>
<td>2.3346</td>
<td>0.0312</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_09900</td>
<td>hypothetical protein, conserved</td>
<td>2.3200</td>
<td>0.0119</td>
<td>T.con orthologues VSG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_10120</td>
<td>hypothetical protein, unlikely</td>
<td>2.2381</td>
<td>0.0477</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_29730</td>
<td>hypothetical protein, conserved</td>
<td>2.2295</td>
<td>0.0063</td>
<td>T.con orthologues VSG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_5_5260</td>
<td>hypothetical protein, unlikely</td>
<td>2.1863</td>
<td>0.0062</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_22950</td>
<td>hypothetical protein, conserved</td>
<td>2.1814</td>
<td>0.0033</td>
<td>T.con orthologues VSG</td>
<td>1-24 SIGNAL PEPTIDE, 25-145 NON-CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_60190</td>
<td>hypothetical protein, conserved</td>
<td>2.1568</td>
<td>0.0331</td>
<td>T.con orthologues hypothetical</td>
<td>1-17 NON-CYT, 18-40 TM, 41-309 CYT, 310-331 TM, 332-342 NON-CYT, 342-363 TM, 364-599 CYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Dicer-like - ribonuclease III activity, RNA-binding**
Figure 3.8 *T. congolense* transcripts that were at least 4-fold more abundant at peak relative to ascending parasitaemia and were described as hypothetical proteins. The InterPro search tool was used to investigate predicted domains for these hypothetical proteins.

Two of the *T. congolense* hypothetical proteins with one predicted transmembrane domains were predicted to have a beta-1,3-N-acetylglucosaminytransferase domain (PTHR11214) by the Panther tool (Mi et al., 2013). This was accompanied by the GO terms for protein glycosylation (GO:0006486), galactosyltransferase activity (GO:0008378), and membrane (GO:0016020) on the InterPro output. It is possible that these transcripts involved in protein glycosylation were up regulated at peak parasitaemia in preparation for transmission. On transmission to the tsetse fly *T. congolense* parasites change their surface coat to incorporate *T. congolense* procyclins, which are highly glycosylated parasite surface molecules (Utz et al., 2006).

### 3.5 Preparation of RNA from slender and stumpy populations of *T. b. brucei* for RNA-seq analysis

*T. b. brucei* slender and stumpy RNA samples were prepared for RNA-seq analysis in order to allow for comparison of changes in transcript abundance between these *T. b. brucei* life cycle stages, and those changes observed between *T. congolense* ascending and peak parasitaemia.

Parasites from infections with *T. b. brucei* AnTat 1.1 90.13 were collected at either ascending parasitaemia when parasites were of slender morphology, or at peak parasitaemia when parasites were of stumpy morphology. The stumpy RNA samples,

| TcSIL3000_1 | hypothetical protein, conserved | 2.1090 | 0.0312 | 1-19 NON-CYT, 20-37 TM, 38-213 CYT |  |  | 11-110, 128-203 BETA-1,3-N-ACETYLGLUCOSAMINYLTRANSFÉRASE (PTHR11214), GO:0006486 protein glycosylation, GO:0008378 galactosyltransferase activity, GO:0016020 membrane |
| TcSIL3000_0 | hypothetical protein | 2.0913 | 0.0400 | 1-91 CYT, 92-108 TM, 109-144 NON-CYT, 145-165 TM, 166-223 CYT | 91-108 TM helix, 144-166 TM helix |  |  |
and slender RNA samples 1 and 2 were each derived from an individual AnTat 1.1 90.13 infection. The third slender RNA sample was derived from a combination of 4 infections. When 1µg of RNA was run on a denaturing agarose gel, the three rRNA bands expected for *T. b. brucei* RNA were visualised for each sample. When these RNA samples were analysed by Northern blotting using a probe targeting PAD1, the levels of the PAD1 transcript were shown to be higher in the stumpy RNA samples than in the slender RNA samples (Figure 3.9A). Slender and stumpy RNA samples were also analysed by nanodrop and the quantity and quality of the samples were deemed to be sufficient for RNA-seq analysis by BGI (Figure 3.9B). Slender forms are the proliferative bloodstream form, whereas stumpy forms are growth-arrested with an accumulation of cells with the 1K1N configuration. Therefore, the KN configurations of parasites before and at the time of RNA collection were determined. This confirmed that at the time of collection the parasites used to prepare the stumpy RNA samples were predominantly accumulated in 1K1N, whereas those parasites used to prepare slender RNA samples were more proliferative with a greater proportion of cells that had either a 2K1N or 2K2N configuration (Figure 3.9C).

![A](image1)

![B](image2)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration (ng/µl)</th>
<th>260/280</th>
<th>260/230</th>
<th>Amount for RNA-seq (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slender 1</td>
<td>386.8</td>
<td>2.12</td>
<td>2.19</td>
<td>10.1</td>
</tr>
<tr>
<td>Slender 2</td>
<td>309.2</td>
<td>2.15</td>
<td>2.41</td>
<td>6.8</td>
</tr>
<tr>
<td>Slender 3</td>
<td>266.9</td>
<td>2.15</td>
<td>2.25</td>
<td>6.1</td>
</tr>
<tr>
<td>Stumpy 1</td>
<td>663.3</td>
<td>2.17</td>
<td>2.11</td>
<td>11.9</td>
</tr>
<tr>
<td>Stumpy 2</td>
<td>549.8</td>
<td>2.21</td>
<td>2.42</td>
<td>12.1</td>
</tr>
<tr>
<td>Stumpy 3</td>
<td>771.5</td>
<td>2.17</td>
<td>2.37</td>
<td>12.3</td>
</tr>
</tbody>
</table>
Figure 3.9 Preparation of T. b. brucei slender and stumpy RNA samples for RNA-seq analysis. A) 1µg of RNA was loaded on a denaturing agarose gel for each slender and stumpy RNA-seq sample, as well as an AnTat 1.1 90.13 control. The 3 bands expected for T. b. brucei rRNA were visualised by ethidium bromide staining. When this RNA was analysed by Northern blotting using a probe directed against PAD1, the PAD1 transcript was detected in greater abundance in the stumpy samples than in the slender samples. B) The slender and stumpy T. b. brucei RNA samples were analysed by nanodrop and met the concentration, quantity and quality criteria for RNA-seq analysis by BGI. C) A comparison of the parasitaemia values (black bars) and the percentage of cells that were 2K1N or 2K2N (blue bars) was made for slender and stumpy samples. The slender 1 and 2 samples were each comprised of 1 infection, the slender 3 sample was prepared from 4 infections. Additionally, the profile of each parasitaemia for the 3 stumpy samples (black lines) was plotted against the percentage of 2K1N or 2K2N cells (blue bars). An asterisk (*) indicates at what point the stumpy sample was collected. Slender samples were collected at 72 hours post-infection.

Further confirmation that the parasites used to generate the stumpy RNA-seq samples were indeed stumpy was obtained by western blotting using a PAD1 specific antibody. The levels of PAD1 protein expression were found to be considerably higher in the protein samples from the parasites used for the stumpy RNA, than in control protein samples from slender AnTat 1.1 90.13 parasites cultured in vitro (Figure 3.10A). Additionally, at the time of RNA collection parasites used to generate stumpy RNA were of stumpy morphology (Figure 3.10B), while those used to generate slender RNA were of slender morphology (Figure 3.10C).
These slender and stumpy RNA samples were then sent for RNA-seq analysis by BGI Tech Solutions Hong Kong. RNA quality was verified by BGI, who provided a quality report on receipt of the RNA samples (not shown).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration (ng/µl)</th>
<th>260/280</th>
<th>260/230</th>
<th>Amount for RNA-seq (µg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slender 1</td>
<td>386.8</td>
<td>2.12</td>
<td>2.19</td>
<td>10.1</td>
</tr>
<tr>
<td>Slender 2</td>
<td>309.2</td>
<td>2.15</td>
<td>2.41</td>
<td>6.8</td>
</tr>
<tr>
<td>Slender 3</td>
<td>266.9</td>
<td>2.15</td>
<td>2.25</td>
<td>6.1</td>
</tr>
<tr>
<td>Stumpy 1</td>
<td>663.3</td>
<td>2.17</td>
<td>2.11</td>
<td>11.9</td>
</tr>
<tr>
<td>Stumpy 2</td>
<td>549.8</td>
<td>2.21</td>
<td>2.42</td>
<td>12.1</td>
</tr>
<tr>
<td>Stumpy 3</td>
<td>771.5</td>
<td>2.17</td>
<td>2.37</td>
<td>12.3</td>
</tr>
</tbody>
</table>

A

B

Stumpy 1 day 6  Stumpy 2 day 6  Stumpy 3 day 6
Figure 3.10 The stumpy RNA-seq samples were collected from PAD1 positive populations of predominantly stumpy morphology. A) A western blot using a primary antibody directed against PAD1 revealed a high level of PAD1 expression in all 3 stumpy protein samples, as compared to slender AnTat 1.1 90.13 control parasites collected in vitro. G6PDH was used as a loading control. B) Phase contrast images revealed that the morphology of the parasites used for stumpy RNA-seq samples was principally stumpy. C) Phase contrast images revealed that the morphology of the parasites used for slender RNA-seq samples was predominantly slender. The scale bar for each image represents 10 µm.

3.6 Analysis of *T. b. brucei* slender and stumpy RNA-seq data with comparison to *T. congolense* RNA-seq data

3.6.1 Alignment to the *T. b. brucei* genome and analysis of differential expression

RNA-seq reads were converted to fold changes in abundance between stumpy and slender RNA samples by Dr Alasdair Ivens. Processing of the reads was carried out in the same manner as for the *T. congolense* RNA-seq data (see section 3.4.1). Once again fold changes (FC) were reported as Log2FC, so that values greater than 1 referred to increases in abundance of more than 2-fold, and values less than -1
referred to decreases in abundance of more than 2-fold. 5028 genes demonstrated significant fold changes in abundance (adj. P-value <0.05) (Figure 3.11). Of these significant changes in abundance, 3096 transcripts were increased in stumpy relative to slender parasites, and 1932 transcripts were decreased in stumpy relative to slender parasites. When changes of less than 2 fold magnitude were excluded, 421 transcripts were at least 2 fold more abundant in stumpy relative to slender parasites, and 501 transcripts were at least 2 fold less abundant in stumpy relative to slender parasites.
Figure 3.11 Schematic workflow summarising the analysis of the T. b. brucei slender and stumpy RNA-seq data following alignment to the T. b. brucei genome. A) Once the T. b. brucei reads had been mapped to the T. b. brucei TREU 927/4 genome, fold changes in abundance in stumpy relative to slender parasites were reported. 5028 transcripts demonstrated significant changes (adj. p-value <0.05) in abundance. Of these, 3096 transcripts had increased abundance in stumpy parasites, and 1932 transcripts had decreased abundance in stumpy parasites. When transcripts were sorted to exclude changes <2 fold, transcripts with no protein product or those described as VSG, then there remained 397 transcripts with increased abundance in stumpy parasites, and 409 transcripts with decreased abundance in stumpy parasites. Of these transcripts, 31 were increased at least 4 fold in stumpy parasites, and 103 were decreased at least 4 fold in stumpy parasites. B) The 5028 transcripts that were significantly changed in abundance were plotted, with ‘slender’ abundance on the x-axis and ‘stumpy’ abundance on the y-axis. Transcripts described as VSG are in orange and all other transcripts are in blue.

The large number of transcripts significantly decreased by at least 2-fold in stumpy relative to slender parasites contrasts with the transcript abundance changes observed for T. congolense. The only significant changes that were of at least 2-fold magnitude were increases in abundance in peak relative to ascending T. congolense parasites. There were no T. congolense transcripts that significantly decreased by at least 2-fold in peak relative to ascending parasites.
Transcripts were further sorted to exclude those described as VSG or with no protein product assigned. In the case of the *T. b. brucei* RNA-seq, few of the transcripts with significant changes in abundance were described as VSG. Furthermore, those transcripts that were described as VSG were all decreased in abundance in stumpy relative to slender parasites. This is in contrast to the *T. congoense* data where a large number of transcripts with increased abundance in peak relative to ascending parasites were described as VSG.

When genes with significant changes in abundance were ordered by the magnitude of the fold change, the gene with the largest increase in abundance in the stumpy relative to slender population was described as a hypothetical protein (Table 3.2), accompanied by procyclin transcripts, which may be held ready for transmission.

<table>
<thead>
<tr>
<th>SysID</th>
<th><em>T. brucei</em> Annotation</th>
<th>Fold change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.10.8340</td>
<td>hypothetical protein</td>
<td>39.81</td>
</tr>
<tr>
<td>Tb927.4.2090</td>
<td>hypothetical protein</td>
<td>22.69</td>
</tr>
<tr>
<td>Tb927.11.11070</td>
<td>hypothetical protein, unlikely</td>
<td>17.15</td>
</tr>
<tr>
<td>Tb927.6.510</td>
<td>GPEET2 procyclin precursor</td>
<td>12.95</td>
</tr>
<tr>
<td>Tb927.10.10260</td>
<td>EP1 procyclin</td>
<td>11.89</td>
</tr>
<tr>
<td>Tb927.9.6820</td>
<td>hypothetical protein, unlikely</td>
<td>10.97</td>
</tr>
<tr>
<td>Tb927.7.3480</td>
<td>hypothetical protein</td>
<td>8.94</td>
</tr>
</tbody>
</table>

Table 3.2 The genes with the greatest increase in abundance in stumpy relative to slender parasites.

The 3096 transcripts found to be significantly increased in stumpy relative to slender form parasites were sorted by their protein product description (Figure 3.12A). The 1932 transcripts found to be significantly decreased in stumpy relative to slender form parasites were also sorted by their protein product description (Figure 3.12B). The most common category of differing transcript abundance between stumpy and slender forms was ‘hypothetical proteins’. Other well-represented categories included protein kinases, phosphatases, and RNA-binding proteins. Changes in abundance of such transcripts is consistent with the different cell signalling requirements accompanying progression from one life cycle stage to another. A change in abundance of transcripts encoding transporter proteins was also observed. This could be associated with preparation for the change of environment faced by parasites on transmission to the tsetse fly, where different nutrients would be
available. For instance, a number of transcripts with reduced abundance in stumpy forms were described as glucose transporters, which are used to uptake the abundant glucose in the mammalian bloodstream. Many transcripts that increased in abundance in the stumpy forms were described as amino acid transporters, which might be required on parasite uptake by the tsetse fly when glucose is limiting. Additionally, consistent with changing metabolic requirements on transmission, transcripts described as mitochondrial carrier proteins, or as NADH-ubiquinone oxidoreductase (part of the mitochondrial electron transfer chain in procyclic forms, (van Hellemond et al., 2005)) were also increased in abundance in the stumpy form, as would be expected for this life cycle stage with a more elaborated mitochondrion than slender forms (Vickerman, 1965).

Further transcripts increased in the stumpy form that were consistent with preparation for transmission to the insect vector included EP Procyclin, a protein expressed on the surface of procyclic forms (Richardson et al., 1988) and BARP, a protein expressed on the surface of epimastigote forms (Urwyler et al., 2007).
Increased abundance in stumpy

- Hypothetical Proteins (60.0%)
- No protein product (3.0%)
- Other (20.8%)
- Transporter (2.4%)
- Protein kinase (1.7%)
- Protein phosphatase (0.8%)
- Palmitoyl acyltransferase (0.2%)
- ESAG (0.7%)
- Receptor-type adenylate cyclase GRESAG 4, putative (0.7%)
- EP Procyclin (0.2%)
- BARP protein (0.4%)
- RNA binding protein, putative (1.1%)
- Zinc finger protein, putative (0.8%)
- Leucine-rich repeat protein (LRRP), putative (0.4%)
- Small nuclear ribonucleoprotein (snRNP), putative (0.2%)
- Exonuclease (0.2%)
- Ribosomal protein (0.6%)
- 40S ribosomal protein (1.7%)
- 60S ribosomal protein (2.7%)
- Eukaryotic initiation factor (0.5%)
- ATP-dependent RNA helicase (0.7%)
- Mitochondrial carrier protein (0.4%)
- NADH-ubiquinone oxidoreductase (0.2%)
- DNA repair (0.3%)
- Chaperone protein DNAJ, putative (0.7%)
- Retrotransposon hot spot (RHS) protein, putative (0.5%)
- DNA-directed RNA polymerase (0.6%)
Transcripts within the expression site-associated gene (ESAG) category were found to have both significant increases and decreases in abundance in stumpy forms. Transcripts falling into one ESAG family, the ESAG9 family, were only significantly increased in abundance in stumpy forms. This was consistent with the published literature (Barnwell et al., 2010; Jensen et al., 2009). Another transcript that was expected to have increased abundance in stumpy forms was the protein-associated with differentiation 1 (PAD1) transcript (Dean et al., 2009). As expected, this transcript, alongside those of PAD2 and 3, were found to be of increased abundance in stumpy forms. The fold change in PAD1 was not as high as expected based on the northern blot in Figure 3.9. However, this may have been caused by reads mapping to conserved sequences in other members of the PAD array that are not

---

**Figure 3.12 Transcripts with significant differences in abundance between slender and stumpy parasites.** A) Transcripts with increased abundance in stumpy relative to slender parasites, grouped by product description. B) Transcripts with decreased abundance in stumpy relative to slender parasites, grouped by product description.
developmentally regulated, so reducing the number of reads that should map to PAD1. Furthermore increased abundance of the PTP1-interacting protein, TbPIP39, transcript in this stumpy data set is consistent with expectation from the literature (Szöör et al., 2010), and preparation for onward life cycle progression. In a recent study of early proteomic changes made during commitment to differentiation to the procyclic form (Domingo-Sananes et al., 2015) TbPIP39 protein was found to increase in abundance 3 hours after induction of differentiation, alongside a trans-sialidase and eukaryotic initiation factors. The increased abundance of these transcripts in stumpy forms is likely necessary in preparation for the rapid proteomic changes required after transmission.

Finally, reduced abundance of flagellar protein, histone, and beta-tubulin transcripts is consistent with the growth-arrested status of the stumpy form. Additionally transcripts for the haptoglobin-haemoglobin receptor were reduced in abundance in the stumpy form as expected (Vanhollebeke et al., 2010). In conclusion, the slender and stumpy RNA-seq data obtained matched expectation from the literature and the expected pre-adaptation of the stumpy form for transmission. This data set thus formed a good base for comparison with the T. congolense RNA-seq data.

### 3.6.2 Comparison to T. congolense RNA-seq data

In both T. congolense peak parasites and T. b. brucei stumpy forms, transcripts that would be useful for cell growth and replication were decreased in abundance (Figure 3.12B, Appendix 1.4), consistent with the increased growth arrest observed in these life cycle stages. Additionally, changes in transcript abundance for components of signalling pathways, such as kinases, were observed both for stumpy compared to slender, and peak compared to ascending parasites (Figure 3.7, Figure 3.12). This could be associated with the need for the transmissible life cycle stage to respond to a different environment than the proliferative bloodstream stage. ESAG transcripts also had differing abundance between peak and ascending, and stumpy and slender parasites (Figure 3.7, Figure 3.12).

Transcripts for putative T. congolense specific surface proteins varied in abundance between ascending and peak parasitaemia (Figure 3.7). This could have been in
preparation for alteration of the cell surface as required for the new environment encountered on transmission. Similarly, many *T. b. brucei* transcripts for cell surface proteins required in the insect vector were of increased abundance in the stumpy form (Figure 3.12).

### 3.6.3 Predicted protein domains of hypothetical proteins with transcripts more abundant in stumpy forms

As for the *T. congolense* RNA-seq data, many transcripts of increased or decreased abundance in stumpy forms were described as hypothetical proteins. The amino acid sequences for these hypothetical proteins were searched for predicted domains using the InterPro tool (Mitchell et al., 2015) (Figure 3.13).

<table>
<thead>
<tr>
<th>SystID</th>
<th>Description</th>
<th>logFC</th>
<th>adj.P.Val</th>
<th>Phobius</th>
<th>SignalP</th>
<th>InterPro Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.10.8340</td>
<td>hypothetical protein</td>
<td>5.315</td>
<td>7.1E-05</td>
<td>1,426 CYT, 143,161 TM, 162,171 NON-CYT</td>
<td>41,63 TM helix, 145,167 TM helix</td>
<td></td>
</tr>
<tr>
<td>Tb927.4.2090</td>
<td>hypothetical protein</td>
<td>4.504</td>
<td>1.3E-03</td>
<td>1,71 CYT, 72,94 TM, 95,111 NON-CYT</td>
<td>72,94 TM helix</td>
<td></td>
</tr>
<tr>
<td>Tb927.11.11070</td>
<td>hypothetical protein, unlikely</td>
<td>4.100</td>
<td>7.3E-04</td>
<td>1,19 NON-CYT, 20,26 TM, 37,110 CYT, 111,129 TM, 130,175 NON-CYT</td>
<td>176,200 TM, 201,216 CYT</td>
<td>174,186 TM helix</td>
</tr>
<tr>
<td>Tb927.9.6820</td>
<td>hypothetical protein, unlikely</td>
<td>3.455</td>
<td>9.7E-03</td>
<td>1,19 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
</tr>
<tr>
<td>Tb927.7.3480</td>
<td>hypothetical protein</td>
<td>3.160</td>
<td>1.2E-02</td>
<td>1,19 SIGNAL PEPTIDE, 20,159 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
</tr>
<tr>
<td>Tb927.9.3210</td>
<td>hypothetical protein, unlikely</td>
<td>3.153</td>
<td>2.2E-02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.1.2950</td>
<td>hypothetical protein, unlikely</td>
<td>2.716</td>
<td>4.0E-04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.14850</td>
<td>hypothetical protein, unlikely</td>
<td>2.573</td>
<td>2.9E-03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.3060</td>
<td>hypothetical protein</td>
<td>2.443</td>
<td>7.0E-05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.5.150</td>
<td>hypothetical protein, conserved</td>
<td>2.421</td>
<td>3.6E-05</td>
<td>1,19 SIGNAL PEPTIDE, 20,159 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
</tr>
<tr>
<td>Tb927.10.3050</td>
<td>hypothetical protein</td>
<td>2.318</td>
<td>1.2E-02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.9.720</td>
<td>hypothetical protein, conserved</td>
<td>2.291</td>
<td>4.0E-05</td>
<td>1,17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,136 NON-CYT</td>
</tr>
<tr>
<td>Tb11.v5.0984</td>
<td>hypothetical protein, conserved hypothetical protein, unlikely</td>
<td>2.280</td>
<td>2.3E-04</td>
<td>1,17 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
<td>1-17 SIGNAL PEPTIDE, 18,271 NON-CYT</td>
</tr>
<tr>
<td>Tb927.1.900</td>
<td>hypothetical protein, unlikely</td>
<td>2.265</td>
<td>4.7E-05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.13 *T. b. brucei* transcripts that were at least 4-fold more abundant in stumpy relative to slender parasites and were described as hypothetical proteins. The InterPro search tool was used to investigate predicted domains for these hypothetical proteins.

Many of the transcripts that increased in abundance in stumpy relative to slender parasites encoded proteins that were predicted to have one or more transmembrane domains. This is comparable to transcripts with increased abundance at peak *T. congoense* parasitaemia, for which the encoded proteins also had predicted transmembrane domains. A number of *T. b. brucei* transcripts encoded proteins that had predicted N-terminal signal peptides without accompanying transmembrane domains, and there is a possibility that such proteins could be secreted. One such transcript was also found to have increased abundance at *T. congoense* peak parasitaemia (Figure 3.8). Some transcripts encoding proteins with predicted transmembrane domains or signal peptides were found to be of significantly reduced abundance in stumpy forms (Appendix 1.6).

Furthermore, one of the hypothetical *T. b. brucei* transcripts of increased abundance in stumpy forms, Tb927.5.2260, was predicted by PFAM to belong to a family of proteins of unknown function (IPR005184). Some proteins in this family were described as heat shock proteins (HsJ), one member was described as a secreted

<table>
<thead>
<tr>
<th>Transcript</th>
<th>Description</th>
<th>InterPro</th>
<th>Pfam</th>
<th>Gene3D</th>
<th>GO Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.2.5720</td>
<td>hypothetical protein</td>
<td>1:23 SIGNAL PEPTIDE, 24-121 NON-CYT</td>
<td>10-184 Domain of unknown function (DUF306), Meta/HsJ (IPR005184), META (PF03724)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.9.730</td>
<td>hypothetical protein</td>
<td>1:57 NON-CYT, 58-79 TM, 80-179 CYT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.5.2260</td>
<td>hypothetical protein</td>
<td>712-735 TM, 735-775 CYT</td>
<td>Sialidase (PF13859) (BNR_3) 82 - 406. Superfamily (SSF50939) 44-453. Gene3D (G3DSA:2.120.10.10) 42-436. Concanavalin A-like lectin/glucanase domain. Superfamily (SSF49699) and Gene3D (G3DSA:2.60.120.200)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.9550</td>
<td>hypothetical protein</td>
<td>1:42 SIGNAL PEPTIDE, 43-711 NON-CYT</td>
<td>Glutamic acid/alanine-rich protein of Trypanosoma (GARP) (PF16731) (IPR031987)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb11.v5.0223</td>
<td>hypothetical protein</td>
<td>1:24 SIGNAL PEPTIDE, 25-272 NON-CYT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Tb927.5.4020 | hypothetical protein | 1-24 SIGNAL PEPTIDE, 25-272 NON-CYT | 1-24 |}

| SignalP | TMhelix | Glutamic acid/alanine-rich protein of Trypanosoma (GARP) (PF16731) (IPR031987) | 25-45 | 186-216 | 1-24 |
| no TM | Coil | 25-45 | Coil | 186-216 | 1-24 | 25-45 | 186-216 | 1-24 |
bacterial protein implicated in motility, and in *Leishmania amazonensis* one member was implicated in virulence.

Another hypothetical protein, Tb11.v5.0223, was predicted to be a trypanosome sialidase (PR01803) by the PRINT predictive tool (Attwood et al., 1994). Members of this family in *T. cruzi* are involved in transferring sialic acid from host cells onto the parasite surface to assist in immune evasion. *T. congolense* trans-sialidases have been detected in infected mouse blood and have been implicated in anaemia (Coustou et al., 2012). *T. vivax* trans-sialidase enzymes released during the early phase of infection have also been implicated in anaemia (Guegan et al., 2013). Additionally, *T. congolense*, *T. vivax* and *T. b. gambiense* (but not *T. b. brucei*) trans-sialidases have been shown to activate the pro-inflammatory endothelial NF-κB pathway (Ammar et al., 2013). Thus trypanosome sialidases are implicated in pathogenesis. *T. congolense* trans-sialidases, however, were not found to have significantly increased abundance at peak parasitaemia.

Finally, one transcript annotated as a hypothetical protein was predicted to belong to the Glutamic acid/alanine-rich protein of Trypanosoma family (GARP). GARP is expressed on *T. congolense* epimastigote forms in the tsetse mouthparts (Butikofer et al., 2002). The *T. b. brucei* Brucei Alanine-Rich Protein (BARP) is described as related to GARP (Jackson et al., 2013), and both are expressed on epimastigote forms. Therefore, it is possible that such a transcript could have increased abundance in stumpy forms in preparation for transmission.

In summary, a number of transcripts, whose products were described as hypothetical proteins, that had increased abundance in *T. b. brucei* stumpy forms or *T. congolense* parasites at peak parasitaemia, may be held in preparation for alteration of the parasite surface coat on continuation of the life cycle in the insect vector.

### 3.6.4 *T. congolense* orthologues of *T. b. brucei* proteins with transcripts increased in abundance in stumpy forms

*T. congolense* orthologues for all of the *T. b. brucei* transcripts recovered by RNA-seq were identified by Dr Alasdair Ivens using BLASTP, hit quality was given a score and analysis was focused on hits scoring better than ‘poor’. Of 260 *T. b. brucei*
genes that were at least 2-fold more abundant in stumpy forms than slender forms, where a *T. congolense* orthologue could be identified (hit quality better than ‘poor’), only 4 were also significantly increased in abundance at *T. congolense* peak relative to ascending parasitaemia (Table 3.3). Two of these transcripts encoded hypothetical proteins, and an InterPro search did not reveal any predicted domains for Tb927.11.12000, or its *T. congolense* orthologue. The other hypothetical protein, Tb927.8.1190 has a predicted coil, but no other predicted domains. The *T. congolense* orthologue for this protein also has a predicted coil, and moreover has a predicted N-terminal signal peptide, and one predicted transmembrane domain (predicted by Phobius and TMHMM). The increased abundance of transcripts for a putative succinyl-CoA synthetase subunit in *T. b. brucei* stumpy forms, and its orthologue in *T. congolense* parasites at peak parasitaemia, is indicative of preparation for the change in metabolism required on transition to the tsetse fly vector. Likewise, the increased abundance of a transcript for an DNA-directed RNA polymerase I subunit, which is involved in transcription of procyclin as well as VSG (Gunzl et al., 2003), may be indicative of preparation for changing to the procyclic surface coat on transmission.

<table>
<thead>
<tr>
<th>SysID</th>
<th><em>T. brucei</em> Annotation</th>
<th>Fold change</th>
<th>Adj. p value</th>
<th>TcIL3000_BestHit</th>
<th>Hit quality</th>
<th>T.con orthologue</th>
<th>Up or Down?</th>
<th>Fold change</th>
<th>Adj. p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.10.13310</td>
<td>DNA-directed RNA polymerase I subunit RPB5z, putative</td>
<td>2.89</td>
<td>0.00053</td>
<td>TcIL3000_10_11410.1</td>
<td>Reasonable</td>
<td>Up</td>
<td>1.33</td>
<td>0.04945</td>
<td></td>
</tr>
<tr>
<td>Tb927.11.12000</td>
<td>hypothetical protein, conserved</td>
<td>2.48</td>
<td>0.02487</td>
<td>TcIL3000.11.12650.1</td>
<td>Reasonable</td>
<td>Up</td>
<td>3.03</td>
<td>0.02410</td>
<td></td>
</tr>
<tr>
<td>Tb927.3.2230</td>
<td>succinyl-CoA synthetase alpha subunit, putative</td>
<td>2.09</td>
<td>0.00417</td>
<td>TcIL3000_3_1350.1</td>
<td>Good</td>
<td>Up</td>
<td>1.30</td>
<td>0.03088</td>
<td></td>
</tr>
<tr>
<td>Tb927.8.1190</td>
<td>hypothetical protein, conserved</td>
<td>2.08</td>
<td>0.00037</td>
<td>TcIL3000_8_790.1</td>
<td>Reasonable</td>
<td>Up</td>
<td>1.37</td>
<td>0.04162</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.3 *T. b. brucei* transcripts and their *T. congolense* orthologues that were significantly increased in abundance in both *T. b. brucei* stumpy parasites and *T. congolense* parasites from peak parasitaemia.

### 3.7 Analysis of RNA-seq data by orthologous group

Many of the *T. b. brucei* transcripts identified by RNA-seq shared the same ‘best hit’ *T. congolense* orthologue. Therefore, in order to get a better comparison of the *T. b.*
*T. brucei* and *T. congoense* data sets, the transcripts identified by RNA-seq were placed into orthologues groups. The orthologous groups were obtained from GeneDB and the trypanosome cluster of orthologous was used, which included *T. b. brucei*, *T. congoense*, *T. vivax*, *T. b. gambiense* and Leishmania species.

When the Log$_2$FC values for *T. congoense* peak relative to ascending parasitaemia were plotted against orthologue cluster, certain clusters had multiple members with pronounced increases in abundance in peak relative to ascending parasitaemia. These orthologous groups did not display the same trend in the *T. b. brucei* stumpy relative to slender plot (Figure 3.14A). For the orthologous groups with a *T. congoense* member at least 4-fold more abundant (log$_2$FC>2) at peak parasitaemia, the size of each group, the main product description associated with the group, and which species made up the group were investigated using GeneDB (Figure 3.15). The majority of these orthologous groups that contained hypothetical proteins were *T. congoense* specific groups. Additionally, one orthologous group contained members of the FAM22 cell surface family (Jackson et al., 2013), which is a *T. congoense*-specific family. Where orthologous groups contained members from other species, the product description was predominantly VSG or ESAG (Figure 3.14B). The *T. b. brucei* transcripts that had greater than 4-fold increases in abundance in stumpy relative to slender parasites belonged to orthologous groups that were predominantly described as ESAG9, procyclin or hypothetical proteins. These hypothetical proteins frequently belonged to *T. brucei* specific clusters containing just two members, one *T. b. brucei* member and one *T. b. gambiense* member (Figure 3.14C). Product descriptions for orthologue clusters of reduced abundance in *T. b. brucei* stumpy forms relative to slender forms are listed in Appendix 1.7.

Overall, it seems that many of the changes observed between peak and ascending *T. congoense* parasitaemia feature transcripts whose products belong to *T. congoense* specific orthologous groups. This indicates that the changes in transcript abundance that are observed between ascending and peak *T. congoense* parasites, differ considerably from those observed between slender and stumpy *T. b. brucei* parasites.
Trypanosome clusters for *T. congolense* transcripts elevated at peak parasitaemia

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Description</th>
<th>ID</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>VSG/ESAG</td>
<td>120</td>
<td>Invariant surface glycoprotein</td>
</tr>
<tr>
<td>11</td>
<td>VSG/ESAG</td>
<td>153</td>
<td>SecA/DEAD-like domain containing protein</td>
</tr>
<tr>
<td>15</td>
<td>VSG/ESAG</td>
<td>306</td>
<td>Cysteine-rich integral membrane precursor</td>
</tr>
<tr>
<td>20</td>
<td>VSG/ESAG</td>
<td>357</td>
<td>Hypothetical/reverse transcriptase</td>
</tr>
<tr>
<td>28</td>
<td>VSG/ESAG</td>
<td>630</td>
<td>Zinc finger C2H2 type/ hypothetical</td>
</tr>
<tr>
<td>48</td>
<td>VSG/ESAG</td>
<td>853</td>
<td>Hypothetical</td>
</tr>
<tr>
<td>56</td>
<td>VSG/ESAG</td>
<td>1736</td>
<td>Hypothetical</td>
</tr>
<tr>
<td>74</td>
<td>VSG/ESAG</td>
<td>4428</td>
<td>Hypothetical</td>
</tr>
<tr>
<td>75</td>
<td>Hypothetical</td>
<td>5748</td>
<td>Invariant surface glycoprotein/hypothetical</td>
</tr>
<tr>
<td>83</td>
<td>VSG/ESAG</td>
<td>6038</td>
<td>Hypothetical</td>
</tr>
<tr>
<td>96</td>
<td>VSG/ESAG</td>
<td>6046</td>
<td>Putative cell-surface expressed/hypothetical</td>
</tr>
</tbody>
</table>
Figure 3.14 Transcripts of increased abundance at *T. congolense* peak parasitaemia do not fall into the same orthologous group as transcripts increased in abundance in *T. b. brucei* stumpy forms. A) Fold changes in abundance for *T. congolense* transcripts between peak and ascending parasitaemia or *T. b. brucei* transcripts between stumpy and slender populations were plotted against Trypanosome orthologous group. This revealed enrichment of certain orthologous groups at *T. congolense* peak parasitaemia (B), which were distinct from those clusters enriched in *T. b. brucei* stumpy forms (C).

| Trypanosome clusters for *T. b. brucei* transcripts elevated in stumpy forms |
|-----------------|-----------------|-----------------|
| 263 | hypothetical/ META domain containing protein | 9167 | hypothetical |
| 865 | trans-sialidase | 9331 | ESAG9 |
| 991 | zinc-finger CCCH domain containing protein II | 9899 | EP1 procyclin |
| 5635 | RNA-binding protein (5) | 10940 | hypothetical |
| 5764 | hypothetical/ ESAG9 | 10954 | hypothetical |
| 6069 | hypothetical/ putative cell-surface expressed gene family | 11073 | hypothetical, T. b. specific* |
| 6872 | procyclin/ procyclin precursor | 11081 | GPEET procyclin precursor/ GPEET procyclin |
| 7154 | hypothetical | 11140 | hypothetical, T. b. specific* |
| 7183 | ESAG9 | 11173 | small nuclear RNA, U2 snRNA |

*Just 2 members in cluster. *T. b. brucei* and *T. b. gambiense*
Trypanosome orthologous groups for which a *T. congolense* member is at least 4-fold more abundant at peak parasitaemia than during ascending parasitaemia. The number of *T. congolense* transcripts that had significantly altered abundance were listed alongside the total number of *T. congolense* transcripts in that cluster. If any genes in the cluster belonged to cell surface phylome families (Jackson et al., 2013) then this was also recorded.

<table>
<thead>
<tr>
<th>Trypanosome_ORTHOMCL10114</th>
<th>Log FC&gt;2 (and adj p&lt;0.05)</th>
<th>Gene type</th>
<th>Genes in cluster [Brucei=B, Congo=C, Vivax=V, Leishmania=L, G=Gambienne, CR-Cruzi]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trypanosome_ORTHOMCL11</td>
<td>52/213 (19/213)</td>
<td>VSG</td>
<td>213C - FAM13 Cell surface phylome (T.congo/T.br shared b-VSG family (ESAG2-like), 317 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL120</td>
<td>3/32 (3/32)</td>
<td>VSG</td>
<td>3C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL15</td>
<td>3/186 (1/186)</td>
<td>VSG</td>
<td>186C - FAM22 Cell surface phylome (T.congo specific)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL1738</td>
<td>4/11 (2/11)</td>
<td>Hypothetical conserved</td>
<td>11C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL20</td>
<td>34/146 (29/146)</td>
<td>VSG</td>
<td>146C - FAM16 Cell surface phylome (T.congo only b-VSG family, 512 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL28</td>
<td>35/110 (29/110)</td>
<td>Transferrin-receptor ESAG 6-like</td>
<td>110C, 2B (ESAG6 and ESAG7), 1G</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL306</td>
<td>4/15 (0/15)</td>
<td>Hypothetical protein, Reverse transcriptase (RNA-dependent DNA polymerase)</td>
<td>15C, 1B, 1G, 1CR</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL357</td>
<td>6/14 (1/14)</td>
<td>Zinc finger, C2H2 type, hypothetical conserved</td>
<td>14C, 1B, 1G, 1L</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL428</td>
<td>1/10 (0/10)</td>
<td>Hypothetical, conserved</td>
<td>10C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL46</td>
<td>10/70 (10/70)</td>
<td>VSG</td>
<td>70C - FAM16 Cell surface phylome (T.congo only b-VSG family, 512 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL5</td>
<td>42/287 (33/287)</td>
<td>VSG</td>
<td>287C - FAM16 Cell surface phylome (T.congo only b-VSG family, 512 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL56</td>
<td>6/59 (6/59)</td>
<td>VSG</td>
<td>59C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL630</td>
<td>1/13 (0/13)</td>
<td>SecA DEAD-like domain containing protein, putative</td>
<td>13C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL6700</td>
<td>2/6 (0/6)</td>
<td>VSG</td>
<td>6C - FAM13 Cell surface phylome (T.congo/T.br shared b-VSG family (ESAG2-like), 317 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL7348</td>
<td>1/5 (0/5)</td>
<td>VSG</td>
<td>5C - FAM13 Cell surface phylome (T.congo/T.br shared b-VSG family (ESAG2-like), 317 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL7355</td>
<td>1/5 (1/5)</td>
<td>Hypothetical, conserved</td>
<td>5C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL7357</td>
<td>2/5 (2/5)</td>
<td>Hypothetical conserved</td>
<td>5C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL74</td>
<td>15/51 (13/51)</td>
<td>VSG</td>
<td>51C - FAM16 Cell surface phylome (T.congo only b-VSG family, 512 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL75</td>
<td>4/51 (4/51)</td>
<td>Hypothetical</td>
<td>51C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL83</td>
<td>1/23 (1/23)</td>
<td>VSG, ESAG2, GRESAG2</td>
<td>23C, 18B, 6G - FAM13 Cell surface phylome (T.congo/T.br shared b-VSG family (ESAG2-like), 317 genes total)</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL853</td>
<td>1/12 (0/12)</td>
<td>Hypothetical, conserved</td>
<td>12C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL9103</td>
<td>2/4 (2/4)</td>
<td>Hypothetical, conserved</td>
<td>4C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL9222</td>
<td>1/3 (1/3)</td>
<td>Hypothetical, conserved</td>
<td>3C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL9540</td>
<td>1/3 (1/3)</td>
<td>Hypothetical, conserved</td>
<td>3C</td>
</tr>
<tr>
<td>Trypanosome_ORTHOMCL96</td>
<td>4/40 (3/40)</td>
<td>VSG, ESAG2</td>
<td>40C - FAM13 Cell surface phylome (T.congo/T.br shared b-VSG family (ESAG2-like), 317 genes total)</td>
</tr>
</tbody>
</table>

Figure 3.15 Trypanosome orthologous groups for which a *T. congolense* member is at least 4-fold more abundant at peak parasitaemia than during ascending parasitaemia. The number of *T. congolense* transcripts that had significantly altered abundance were listed alongside the total number of *T. congolense* transcripts in that cluster. If any genes in the cluster belonged to cell surface phylome families (Jackson et al., 2013) then this was also recorded.
<table>
<thead>
<tr>
<th>TrypanosomE_Cluster</th>
<th>PvA <a href="mean">&amp;</a></th>
<th>SttvB (mean)</th>
<th>Genes in cluster (Brucei=B, Congo=C, Vivax=V, Leishmanial=L, G=Gambiense, CR-Cruzi)</th>
<th>Type of genes</th>
<th>RIT-seq phenotype (N=None, abnormal proliferation in: D=day 3 baf, B6 = day 6 baf, P=pcf)</th>
<th>Cell Surface Phylome (Y/N)</th>
<th>Fam No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>357</td>
<td>2.74</td>
<td>0.95</td>
<td>1B, 14C, 1L, 1G Hypothetical proteins or reverse transcriptase (RNA-dependent DNA polymerase) putative</td>
<td>D</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>1.82</td>
<td>-1.14</td>
<td>2B, 110C, 1G ESAG6 putative, ESAG 7, transferrin-receptor like/ ESAG6 like</td>
<td>N</td>
<td>Y - 0a, 14, 15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>1.82</td>
<td>0.12</td>
<td>1B,4C,44V,3G Hypothetical proteins conserved or unlikely, 1 retrotransposon hot spot VSG, VSG-related putative, VSG-domain containing protein</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.74</td>
<td>0.65</td>
<td>1B, 211C, 1G Hypothetical protein conserved G6, D</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3122</td>
<td>1.60</td>
<td>1.31</td>
<td>1B, 1C, 1V, 5L, 1G, 2CR Hypothetical protein conserved D, 3B, 6B</td>
<td>N</td>
<td>Y - Fam6b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>306</td>
<td>1.53</td>
<td>0.77</td>
<td>1B, 15C, 1G, 1CR Cysteine-rich, acidic integral membrane protein precursor CRAM, hypothetical protein conserved</td>
<td>N (but up in pcf by microarray)</td>
<td>N (but N-terminal signal peptide)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6738</td>
<td>1.31</td>
<td>-0.22</td>
<td>1B, 1C, 1V, 1G, 2CR Hypothetical protein conserved D, 3B, 6B</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9746</td>
<td>1.22</td>
<td>-0.52</td>
<td>1B, 1C, 1G Hypothetical protein, Protein of unknown function</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>996</td>
<td>1.11</td>
<td>0.09</td>
<td>1B, 1C, 2V, 1G, 2CR, 5L Aspartate aminotransferase, putative</td>
<td>N</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1165</td>
<td>1.00</td>
<td>0.21</td>
<td>2B, 1C, 1G, 3CR, 5L Triglyceride lipase, putative B6, D, P</td>
<td>Y - Fam63</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.16 Trypanosome orthologue clusters for which the mean fold change in abundance between peak and ascending *T. congolense* parasitaemia is greater than 2-fold (Log₂FC>1). For each orthologous cluster the mean Log₂FC for *T. congolense* peak vs. ascending and *T. b. brucei* stumpy vs. slender is reported alongside a list of the species with genes in the cluster, and the type of genes found in the cluster. Additionally, if RNAi targeting the *T. b. brucei* cluster member resulted in loss of fitness in any life cycle stage then this life cycle stage was listed. If members of the cluster formed part of the Trypanosome cell surface phylome, this was also recorded. Two hypothetical proteins, for which knock down resulted in loss of fitness through differentiation are highlighted in dark green.

In order to focus on those orthologous groups for which members consistently changed abundance between peak and ascending *T. congolense* parasites, the mean log fold change (Log₂FC) from peak to ascending parasitaemia was calculated for each orthologous group by Dr Alasdair Ivens. Those orthologue groups for which the mean fold change in abundance was greater than 2 (Log₂FC>1) for *T. congolense* peak relative to ascending parasitaemia are listed in Figure 3.16. In most cases the mean change in abundance for stumpy relative to slender forms was also positive for these orthologue groups, but in only one case (cluster 3122) was the mean change in abundance of comparable magnitude between the *T. b. brucei* and *T. congolense* data sets. Cluster 3122 members were described as conserved hypothetical proteins, and RNAi-mediated knock down of the *T. b. brucei* orthologue resulted in loss-of-fitness on differentiation from bloodstream to procyclic forms. An InterPro search using the *T. congolense* and *T. b. brucei* sequences from cluster 3122 did not reveal any predicted domains. In contrast, some clusters that had on average increased abundance in *T. congolense* peak parasitaemia were of reduced abundance in *T. b. brucei* stumpy forms. For instance, in cluster 6738 there was enrichment in *T. congolense* peak parasites, but not in *T. b. brucei* stumpy forms. Members of cluster...
6738 were described as conserved hypothetical proteins, and RNAi-targeting the *T. b. brucei* orthologue resulted in a loss-of-fitnes in differentiation from bloodstream to procyclic forms. Both the *T. b. brucei* and the *T. congolense* orthologues are predicted by Phobius to have one transmembrane domain. However, neither the *T. b. brucei* nor *T. congolense* transcripts falling into cluster 6738 underwent significant fold changes (adjusted p-value<0.05), so differences between the *T. congolense* and *T. b. brucei* data in this case must be interpreted with care. Nevertheless, the increased abundances of *T. congolense* transcripts belonging to particular trypanosome orthologue groups at peak parasitaemia were not always mirrored by comparable increases in the *T. b. brucei* transcripts belonging to the same orthologous groups in stumpy forms.

Finally for the orthologue clusters common between *T. b. brucei* and *T. congolense*, Dr Alasdair Ivens plotted the differences between the mean Log$_2$FC values for *T. congolense* peak versus ascending data and *T. b. brucei* stumpy versus slender data (Figure 3.17A). Clusters with the largest differences between *T. congolense* and *T. b. brucei* were labelled with their cluster ID, but only those where at least one transcript had a significant (adjusted p-value <0.05) difference at peak compared to ascending parasitaemia were described in detail (Figure 3.17B). In many cases differences between *T. b. brucei* and *T. congolense* occurred because transcripts that usually had a pronounced reduction in stumpy relative to slender forms, were slightly increased in peak relative to ascending *T. congolense* parasites. Examples of such clusters include cluster 76 described as histone 2B, and cluster 65 described as paraflagellar rod protein. This could be indicative of the less complete growth arrest observed at *T. congolense* peak parasitaemia, as compared to a stumpy population at *T. b. brucei* peak parasitaemia. One cluster that appears to be particularly differentially regulated between *T. b. brucei* and *T. congolense* is cluster 28 described as Transferrin-receptor-like. Transferrin receptor components (ESAG6 and 7) are encoded in polycistronic telomeric expression sites along with VSG (Salmon et al., 1997; Steverding et al., 1995). It has been demonstrated in *T. b. brucei* that expression of different transferrin receptor variants (from different expression sites) affect the affinity for different host transferrins and the ability of trypanosomes to grow in serum from various hosts. It has also been shown that *T. b. brucei* parasites growing
in serum for which they cannot uptake transferrin arrest growth, but parasites that are able to grow out have switched transferrin receptor expression (Bitter et al., 1998). It is unclear why so many *T. congo**lense* transcripts belonging to cluster 28 are increased in abundance at peak parasitaemia, but it may be related to a high level of expression site switching at *T. congo**lense* peak parasitaemia, as many VSG and ESAG clusters were also increased in abundance at peak relative to ascending parasitaemia (Figure 3.14B). *T. b. brucei* transcripts belonging to cluster 5171 described as cyclophilin a, were of reduced abundance in stumpy relative to slender parasites in agreement with published data (Pelle et al., 2002).

In the case of one orthologue cluster, 6069, the *T. b. brucei* transcript was elevated in stumpy forms, but not at *T. congo**lense* peak parasitaemia. The *T. b. brucei* orthologue had been shown to decrease mRNA stability or translation in a tethering screen (Erben et al., 2014), whereas *T. congo**lense* orthologues were proposed as putative cell surface expressed proteins.

In summary, for many transcripts that were significantly reduced in abundance in *T. b. brucei* stumpy relative to slender forms, *T. congo**lense* transcripts belonging to the same cluster were not reduced in abundance at peak relative to ascending parasitaemia. This, taken together with the analysis of cell cycle arrest in *T. congo**lense* (section 3.2), indicates a higher degree of quiescence in *T. b. brucei* pleomorphic AnTat 1.1 90.13 parasites at peak parasitaemia than in the *T. congo**lense* IL3000 parasites at peak parasitaemia.
Figure 3.17 Differences between mean LogFC values for *T. congolense* peak vs. ascending data and *T. b. brucei* stumpy vs. slender data. A) A positive LogFC value indicates an orthologous group for which the fold change in abundance was greater for *T. congolense* ascending to peak parasitaemia, than for *T. b. brucei* slender to stumpy differentiation. A negative LogFC value indicates an orthologue group for which the fold change in abundance was greater for *T. b. brucei* slender to stumpy differentiation, than for *T. congolense* ascending to peak parasitaemia. Clusters which had LogFC differences of $>2$ or $<-2$ were coloured red and labelled with the cluster ID. Other clusters were coloured light blue. B) Description of clusters (green) for which the LogFC difference was $>2$ and at least one *T. congolense* transcript had a significant fold change (adjusted p-value <0.05) in the RNA-seq data, and clusters (red) for which LogFC difference was $<-2$ and at least one *T. b. brucei* transcript had a significant fold change (adjusted p-value <0.05) in the RNA-seq data.
3.8 Conclusions and future directions

At peak parasitaemia, the increased proportion of *T. congolense* parasites that have a 1K1N configuration is indicative of a degree of growth arrest, and possibly a parasite intrinsic mechanism of growth control. The level of growth arrest is not as severe as that observed at *T. b. brucei* peak parasitaemia, but this may be related to the *T. congolense* strain used, rather than being a common phenomenon for *T. congolense* parasites. For instance, some elements of growth control may have been lost through multiple rodent passages, as for *T. b. brucei* monomorphs. This could be addressed in the future by assessing the degree of growth arrest of newly isolated field strains, with limited passage history. Nevertheless, the increased arrest observed at peak parasitaemia relative to ascending parasitaemia encouraged investigation of ascending and peak parasites to determine whether there were further differences in these parasite populations.

Consistent with the increased cell cycle arrest coinciding with peak *T. congolense* parasitaemia, some of the transcripts found to have reduced abundance at peak parasitaemia included paraflagellar rod protein, DNA polymerase catalytic subunits, and a putative cell division protein kinase. Similarly, *T. b. brucei* stumpy forms had less abundant transcripts for histones and paraflagellar rod proteins than slender forms, as expected for their growth-arrested state.

While the replicative slender form of *T. b. brucei* can be distinguished from the transmission-ready stumpy form on the basis of morphology, no such distinction could be made between the ascending parasites and peak parasites of a *T. congolense* infection. This is consistent with observations describing *T. congolense* as monomorphic made by Vickerman (Vickerman, 1969a) and Akoda et al (Akoda et al., 2008), but not those made by Nantulya et al (Nantulya et al., 1978b) who described predominantly long forms at peak parasitaemia and short forms during ascending parasitaemia. Differences in the extent of pleomorphism observed might have been a result of the different strains of *T. congolense* used in each of these studies, which had different passage histories.
A greater number of transcripts significantly differed in abundance between *T. b. brucei* slender and stumpy forms than between *T. congolense* ascending and peak parasites. For *T. b. brucei* changes in transcript abundance were spread across many orthologous groups, with particular enrichment for groups containing ESAG9 and Procyclin in stumpy forms. In contrast, *T. congolense* transcripts enriched at peak parasitaemia were focused into a few prominent orthologous groups that were mainly described as hypothetical proteins or cell surface proteins (including VSG).

A number of transcripts that differed in abundance between *T. b. brucei* stumpy and slender forms were linked to changes in metabolism that accompany the changing nutrient environments on transitioning from the mammalian bloodstream to the tsetse fly midgut. These changes included a decrease in transcripts for glucose transporters, and an increase in transcripts for amino acid transporters and mitochondrial enzymes. While some mitochondrial enzymes such as the succinyl-CoA synthetase alpha-subunit were found to have increased transcript abundance at peak relative to ascending *T. congolense* parasitaemia, other mitochondrial enzymes were found to have a decreased transcript abundance, such as mitochondrial malate dehydrogenase. Furthermore amino acid transporters were not found to have increased transcript abundance at peak *T. congolense* parasitaemia, unlike in the *T. b. brucei* stumpy forms.

There is a possibility that transmissible forms are present throughout the course of *T. congolense* infection without a focused accumulation at peak parasitaemia. This would be consistent with the observations by Vickerman (Vickerman, 1965) that describe mitochondrial activity in all bloodstream form *T. congolense* and *T. vivax*, but not in *T. b. brucei* where only stumpy forms have active mitochondria and slender forms (as well as monomorphic *T. evansi*) have inactive mitochondria. This may explain the fewer significant changes in transcript abundance between *T. congolense* peak and ascending parasitaemia, as compared to changes in transcript abundance between *T. b. brucei* slender and stumpy parasites uncovered by RNA-seq. Further investigation of the *T. congolense* mitochondrial structure through infection could be made using the mitotracker stain, using the method described in Vassella *et al* (Vassella et al., 1997b).
Nevertheless, transcripts of increased abundance at peak *T. congolense* parasitaemia included those for *T. congolense* cell surface proteins, and a hypothetical protein implicated in protein glycosylation, which may be held in readiness for the transformation of the parasite surface landscape required on transition to the insect vector. Enrichment of such transcripts at peak parasitaemia was comparable with the enrichment of procyclin transcripts observed in *T. b. brucei* stumpy forms. It appears to be important for both parasites to be prepared to change their surface to face the new environment encountered following transmission.

In conclusion, there were some similarities between *T. congolense* and *T. b. brucei* ascending and peak parasitaemia during infections, such as an increased growth arrest at the first peak of parasitaemia, and variation in abundance of transcripts for surface proteins. However, when the *T. b. brucei* and *T. congolense* RNA-seq data were compared it became clear that the changes in transcript abundance occurring from slender to stumpy forms were more extensive than those occurring from *T. congolense* ascending to peak parasitaemia. Therefore, these parasites appear to have evolved distinct methods for ensuring infection longevity and transmission.
Chapter 4: Conservation of components of the *T. b. brucei* SIF response pathway in *T. congolense* and *T. vivax*
4.1 Introduction

The generation of an inducible RNAi library in bloodstream form trypanosomes has enabled high throughput screening for essential genes (Alsford et al., 2011), as well as effectors of HAT drug action (Alsford et al., 2012). In the latter screen, clones able to grow in the presence of HAT drugs were analysed to identify the genes knocked down. In this way genes involved in susceptibility to these pharmacological agents were identified.

Mony et al (Mony et al., 2014) applied a similar principal to screen for genes that facilitate the response to SIF. As the library is in a monomorphic cell line it was unlikely to be inhibited by the SIF that accumulates in high-density *T. b. brucei* cultures or *in vivo*, therefore the screen could not be carried out using conditioned media. However, it was known that cell permeable analogues of cAMP and AMP (Breidbach et al., 2002; Laxman et al., 2006; Vassella et al., 1997a) can act as SIF-mimics *in vitro*, inhibiting growth of both pleomorphic and monomorphic cell lines. Consequently, a cell permeable cAMP analogue, 8-CPT-cAMP, was applied to the library and ion-torrent sequencing identified RNAi inserts overrepresented in the outgrowing parasite populations. A list of potential components of the stumpy induction pathway thereby was generated. Thereafter, pleomorphic cell lines were created with inducible RNAi targeted against individual hits from the screen. For those hits that were validated, RNAi-mediated knock down resulted in resistance to SIF and reduced stumpy formation at high parasitaemia *in vivo*.

The validated list of genes provides unprecedented insight into the molecular machinery of quorum sensing in the *T. b. brucei* bloodstream form. If these SIF-responsive genes are found to be conserved in other African trypanosome species, cross talk between cohabiting parasites becomes an interesting possibility in co-infection. Therefore a bioinformatics analysis was carried out to identify potential *T. congolense* and *T. vivax* orthologues of the SIF-responsive genes from *T. b. brucei*. This was followed up with complementation experiments to analyse whether one of the *T. congolense* or *T. vivax* orthologues was able to provide a rescue in cells deficient in the corresponding *T. b. brucei* gene.
# 4.2 Identification of *T. congolense* and *T. vivax* orthologues of SIF-responsive genes

Genes which when targeted by RNAi led to resistance to SIF-mediated growth control in *T. b. brucei* were subjected to a BLAST search against the *T. congolense* and *T. vivax* genome databases. The *T. b. brucei* protein sequences of interest were entered into the BLASTP tool on TritrypDB (Aslett et al., 2010) (default parameters were used). A table was compiled listing the top hit of each BLAST search (Table 4.1 and Table 4.2). Hits were considered to be more convincing orthologues if the reciprocal BLAST search using the *T. congolense*/*T. vivax* protein generated the original *T. b. brucei* protein of interest as the top hit.

<table>
<thead>
<tr>
<th>Predicted gene function</th>
<th><em>T. brucei</em> ID</th>
<th><em>T. vivax</em> hit 1</th>
<th>Result from input of <em>T. vivax</em> hit 1 into BLAST (E value, hit number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3620</td>
<td>TvY46_0403330</td>
<td>1.20E-157 97% 98% Tb927.4.3620 (1.0e-157, 1)</td>
</tr>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3640</td>
<td>TvY46_0403330</td>
<td>2.40E-154 94% 97% Tb927.4.3640 (2.0e-154, 2)</td>
</tr>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3630</td>
<td>TvY46_0403330</td>
<td>7.30E-153 93% 96% Tb927.4.3630 (6.0e-153, 3)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5940</td>
<td>TvY46_1005880</td>
<td>3.00E-234 86% 94% Tb927.10.5940 (2.5e-234, 1)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5950</td>
<td>TvY46_1006880</td>
<td>3.50E-233 86% 94% Tb927.10.5950 (2.9e-233, 2)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5950</td>
<td>TvY46_1005880</td>
<td>4.00E-232 85% 93% Tb927.10.5950 (3.3e-232, 3)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.2.2720</td>
<td>TvY46_0200560</td>
<td>8.20E-149 35% 52% Tb927.2.2720 (8.7e-149, 1)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.15020</td>
<td>TvY46_1011760</td>
<td>9.90E-168 46% 62% Tb927.10.15020 (8.2e-168, 1)</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 1)</td>
<td>Tb927.11.6600 (previously known as Tb11.02.4450)</td>
<td>TvY46_11014430</td>
<td>9.0E-168 46% 62% Tb927.10.15020 (8.2e-168, 1)</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 2)</td>
<td>Tb927.9.4080 (previously known as Tb09.160.2650)</td>
<td>TvY46_0901410</td>
<td>1.90E-228 49% 65% Tb927.10.15020 (8.2e-228, 1)</td>
</tr>
<tr>
<td>RNA-binding protein</td>
<td>Tb927.10.12100</td>
<td>TvY46_1011760</td>
<td>1.80E-37 69% 79% Tb927.10.12100 (1.5e-37, 2)</td>
</tr>
<tr>
<td>Adenylosuccinate synthetase (Purine pathway)</td>
<td>Tb11.02.4450 (segment of sequence similarity from aa 214-493)</td>
<td>2.00E-40 32% 50% Tb927.10.15020 (8.2e-40, 1)</td>
<td></td>
</tr>
<tr>
<td>Adenylosuccinate Lyase (Purine pathway)</td>
<td>Tb09.160.5560</td>
<td>TvY46_0902900</td>
<td>2.50E-239 77% 88% Tb927.10.15020 (8.2e-239, 1)</td>
</tr>
<tr>
<td>GMP synthase</td>
<td>Tb927.7.2100</td>
<td>TvY46_0701800</td>
<td>2.40E-232 71% 83% Tb927.7.2100 (2.0e-232, 1)</td>
</tr>
<tr>
<td>ubiquitin activating enzyme, putative</td>
<td>Tb927.2.4020</td>
<td>TvY46_0301910</td>
<td>4.40E-137 44% 56% Tb927.2.4020 (3.6e-137, 1)</td>
</tr>
<tr>
<td>serine/threonine protein kinase, putative</td>
<td>Tb927.3.4560</td>
<td>TvY46_0303330 (segment of sequence similarity covers aa 68-630 of the query sequence)</td>
<td>6.40E-145 51% 67% Tb927.3.4560 (5.3e-145, 1)</td>
</tr>
<tr>
<td>hypothetical protein, conserved (Hyp 3)</td>
<td>Tb927.4.670</td>
<td>TvY46_040390</td>
<td>5.50E-187 56% 70% Tb927.4.670 (4.5e-187, 1)</td>
</tr>
</tbody>
</table>
Table 4.1: Top hit of each BLAST search of *T. b. brucei* SIF-responsive proteins against *T. vivax*. *T. b. brucei* amino acid sequences were entered into the BLASTP tool of TritrypDB. *T. vivax* was selected as the target organism for each search. Default parameters were used (Expectation value = 10, Maximum descriptions/alignments (V=B) = 50, Low complexity filter = no). As of 3 Feb 2014, NCBI-BLAST was used to determine sequence similarity on TritrypDB. Prior versions of the search used WU-BLAST. Boxes in grey represent searches that used NCBI-BLAST, the previous searches used WU-BLAST. The final column reports the result of inputting the top *T. vivax* hit into a BLAST search against *T. b. brucei*. Only proteins where the top *T. vivax* hit returned the original *T. b. brucei* protein as the top hit are highlighted. *T. b. brucei* SIF-responsive proteins with strong *T. vivax* hits are highlighted in dark green, while those with weaker hits are highlighted in light green.

<table>
<thead>
<tr>
<th>Hypothetical protein (Hyp 4)</th>
<th>TvY486_0701350</th>
<th>Tb927.4.3650</th>
<th>0.68</th>
<th>27%</th>
<th>44%</th>
<th>Tb927.4.3650 (0.56, 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenosine kinase, putative</td>
<td>Tb927.6.2300</td>
<td>TvY486_001850</td>
<td>1.90E-120</td>
<td>63%</td>
<td>76%</td>
<td>Tb927.6.2300 (1.6e-120, 1)</td>
</tr>
<tr>
<td>Adenosine kinase, putative</td>
<td>Tb927.6.2360</td>
<td>TvY486_001850</td>
<td>5.10E-120</td>
<td>63%</td>
<td>76%</td>
<td>Tb927.6.2360 (4.3e-120, 2)</td>
</tr>
<tr>
<td>Kinetoplastid-specific dual specificity phosphatase, putative</td>
<td>Tb927.7.7160</td>
<td>TvY486_0709784</td>
<td>7.80E-27</td>
<td>39%</td>
<td>56%</td>
<td>Tb927.7.7160 (6.5e-27, 1)</td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 5)</td>
<td>Tb09.3.10530 (previously known as : Tb09.211.4370)</td>
<td>TvY486_0906760</td>
<td>9.60E-32</td>
<td>44%</td>
<td>61%</td>
<td>Tb09.211.4370 (7.9e-32, 1)</td>
</tr>
<tr>
<td>Inosine-5'-monophosphate dehydrogenase</td>
<td>Tb927.10.16120</td>
<td>TvY486_0027160</td>
<td>1.30E-201</td>
<td>83%</td>
<td>92%</td>
<td>Tb927.10.16120 (1.1e-201, 1)</td>
</tr>
<tr>
<td>Phosphatase and tensin homologue, putative</td>
<td>Tb927.11.11470</td>
<td>TvY486_1100710 (area of sequence similarity covers aa 16-247 of query sequence)</td>
<td>1.00E-73</td>
<td>59%</td>
<td>75%</td>
<td>Tb11.03.0850 (8.6e-74, 1)</td>
</tr>
<tr>
<td>Protein kinase A regulatory subunit</td>
<td>Tb927.11.4410 (previously known as : Tb11.02.2210)</td>
<td>TvY486_1104700</td>
<td>5.00E-209</td>
<td>76%</td>
<td>87%</td>
<td>Tb11.02.2210 (4.1e-209, 1)</td>
</tr>
<tr>
<td>Transcription silencer (ISWI)</td>
<td>Tb22.2.1810</td>
<td>TvY486_0200200</td>
<td>0</td>
<td>67%</td>
<td>79%</td>
<td>Tb22.2.1810 (0.0, 1)</td>
</tr>
<tr>
<td>Phosphoglycerate mutase</td>
<td>Tb927.5.3580</td>
<td>TvY486_0502950</td>
<td>3.00E-103</td>
<td>76%</td>
<td>85%</td>
<td>Tb927.5.3580 (3e-103, 1)</td>
</tr>
<tr>
<td>Hypothetical protein (Hyp 5)</td>
<td>Tb927.8.2860</td>
<td>TvY486_0802390</td>
<td>2.00E-51</td>
<td>40%</td>
<td>56%</td>
<td>Tb927.8.2860 (3e-52, 1)</td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 5)</td>
<td>Tb927.11.300</td>
<td>TvY486_0004610 (small region of similarity)</td>
<td>1.1</td>
<td>29%</td>
<td>38%</td>
<td>Tb927.11.300 not returned as a hit when BLAST with top <em>T. vivax</em> hit</td>
</tr>
<tr>
<td>DNA repair protein, putative (Hyp 9)</td>
<td>Tb927.11.750</td>
<td>TvY486_1100600 (2 small regions of similarity)</td>
<td>1.00E-81</td>
<td>72%</td>
<td>84%</td>
<td>Tb927.11.750 (6e-82, 1)</td>
</tr>
<tr>
<td>Product: protein phosphatase 2C, putative</td>
<td>Tb927.11.760</td>
<td>TvY486_1100610</td>
<td>0</td>
<td>80%</td>
<td>90%</td>
<td>Tb927.11.760 (0.0, 1)</td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 10)</td>
<td>Tb927.11.1640</td>
<td>TvY486_1101570</td>
<td>1.00E-44</td>
<td>36%</td>
<td>52%</td>
<td>Tb927.11.1640 (7e-33, 1)</td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 11)</td>
<td>Tb927.11.2250</td>
<td>TvY486_0802140</td>
<td>0.91</td>
<td>34%</td>
<td>45%</td>
<td>Tb927.11.2250 not returned as a hit when BLAST with top <em>T. vivax</em> hit</td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 13)</td>
<td>Tb927.11.1470</td>
<td>TvY486_1112280</td>
<td>2.00E-120</td>
<td>60%</td>
<td>74%</td>
<td>Tb927.11.1470 (1e-112, 1)</td>
</tr>
<tr>
<td>Trichohyalin, putative</td>
<td>Tb927.11.11480</td>
<td>TvY486_1112510 (results here are for best one of 3 alignments shown)</td>
<td>2.00E-67</td>
<td>40%</td>
<td>66%</td>
<td>Tb927.11.11480 (4e-48, 1)</td>
</tr>
</tbody>
</table>

**Key**

<table>
<thead>
<tr>
<th>Promising hits</th>
<th>Proportion of hits</th>
<th>Percentage of hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hits with low E value, % identity over 50%, and hits for which the top <em>T. vivax</em> hit completely amino acid sequences returned the original <em>T. brucei</em> protein as the top hit when put through BLAST</td>
<td>10/35</td>
<td>29%</td>
</tr>
<tr>
<td>Hits as above but including % identity over 30%</td>
<td>26/35</td>
<td>74.29</td>
</tr>
</tbody>
</table>

---

1. **Maximum descriptions/alignments (V=B)** = 50, Low complexity filter = no.
2. As of 3 Feb 2014, NCBI-BLAST was used to determine sequence similarity on TritrypDB. Prior versions of the search used WU-BLAST. Boxes in grey represent searches that used NCBI-BLAST, the previous searches used WU-BLAST. The final column reports the result of inputting the top *T. vivax* hit into a BLAST search against *T. b. brucei*. Only proteins where the top *T. vivax* hit returned the original *T. b. brucei* protein as the top hit are highlighted. *T. b. brucei* SIF-responsive proteins with strong *T. vivax* hits are highlighted in dark green, while those with weaker hits are highlighted in light green.
<table>
<thead>
<tr>
<th>Predicted gene function</th>
<th>T. brucei ID</th>
<th>T. congolense hit 1</th>
<th>% identity</th>
<th>% similarity</th>
<th>Result from input of T. congolense hit 1 into BLAST (E value, hit number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3620</td>
<td>TcIL3000.11.8600</td>
<td>58%</td>
<td>78%</td>
<td>Tb927.4.3620 (2.2e-100, 4)</td>
</tr>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3640</td>
<td>TcIL3000.11.8600</td>
<td>57%</td>
<td>78%</td>
<td>Tb927.4.3640 (7.4e-100, 5)</td>
</tr>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3630</td>
<td>TcIL3000.11.8600</td>
<td>57%</td>
<td>78%</td>
<td>Tb927.4.3630 (4.1e-99, 6)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5940</td>
<td>TcIL3000.10.4970</td>
<td>89%</td>
<td>95%</td>
<td>Tb927.10.5940 (5.1e-243, 1)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5930</td>
<td>TcIL3000.10.4970</td>
<td>50%</td>
<td>85%</td>
<td>Tb927.10.5930 (1.1e-242, 2)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5920</td>
<td>TcIL3000.10.4970</td>
<td>88%</td>
<td>93%</td>
<td>Tb927.10.5920 (5.2e-235, 1)</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.2.2720</td>
<td>TcIL3000.0.44450</td>
<td>39%</td>
<td>54%</td>
<td>Tb927.2.2720 (1.1e-220, 1)</td>
</tr>
<tr>
<td>Hypothetical Protein</td>
<td>Tb927.10.15020</td>
<td>TcIL3000.10.12870</td>
<td>49%</td>
<td>63%</td>
<td>Tb927.10.15020 (3.4e-262, 1)</td>
</tr>
<tr>
<td>Hypothetical Protein</td>
<td>Tb927.9.4080</td>
<td>TcIL3000.10.12870</td>
<td>45%</td>
<td>58%</td>
<td>Tb927.9.4080 (3.3e-262, 1)</td>
</tr>
<tr>
<td>Hypothetical Protein</td>
<td>Tb927.10.12100</td>
<td>TcIL3000.0.55500</td>
<td>77%</td>
<td>87%</td>
<td>Tb927.10.12100 (3.2e-100, 1)</td>
</tr>
<tr>
<td>Adenylosuccinate</td>
<td>TcIL3000.10.12870</td>
<td>TcIL3000.11.7200</td>
<td>86%</td>
<td>93%</td>
<td>Tb927.10.12100 (3.4e-205, 1)</td>
</tr>
<tr>
<td>Lyase (Purine pathway)</td>
<td>TcIL3000.10.12870</td>
<td>TcIL3000.11.7200</td>
<td>86%</td>
<td>93%</td>
<td>Tb927.10.12100 (3.4e-205, 1)</td>
</tr>
<tr>
<td>GMP-synthase</td>
<td>TcIL3000.10.12870</td>
<td>TcIL3000.11.7200</td>
<td>86%</td>
<td>93%</td>
<td>Tb927.10.12100 (3.4e-205, 1)</td>
</tr>
<tr>
<td>ubiquitin</td>
<td>TcIL3000.2.560</td>
<td>TcIL3000.2.560</td>
<td>54%</td>
<td>70%</td>
<td>Tb927.2.4020 (1.9e-179, 1)</td>
</tr>
<tr>
<td>serine/threonine</td>
<td>Tb927.2.4020</td>
<td>TcIL3000.2.560</td>
<td>51%</td>
<td>69%</td>
<td>Tb927.2.4020 (8.5e-81, 2)</td>
</tr>
<tr>
<td>protein kinase,</td>
<td>Tb927.3.4560</td>
<td>TcIL3000.0.4480</td>
<td>51%</td>
<td>69%</td>
<td>Tb927.3.4560 (8.5e-81, 2)</td>
</tr>
<tr>
<td>conserved (Hyp 3)</td>
<td>Tb927.4.670</td>
<td>TcIL3000.4.290</td>
<td>59%</td>
<td>74%</td>
<td>Tb927.4.670 (3.2e-202, 1)</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>Tb927.4.3650</td>
<td>TcIL3000.10.810</td>
<td>25%</td>
<td>44%</td>
<td>Tb927.4.3650 (was not even on the list of hits for TcIL3000.10.810)</td>
</tr>
<tr>
<td>adenosine kinase,</td>
<td>Tb927.6.2300</td>
<td>TcIL3000.6.1860</td>
<td>72%</td>
<td>86%</td>
<td>Tb927.6.2300 (1.9e-140, 2)</td>
</tr>
<tr>
<td>putative</td>
<td>Tb927.6.2300</td>
<td>TcIL3000.6.1860</td>
<td>73%</td>
<td>86%</td>
<td>Tb927.6.2300 (2.7e-141, 1)</td>
</tr>
<tr>
<td>kinetoplastid-</td>
<td>Tb927.7.7160</td>
<td>TcIL3000.5.52020</td>
<td>41%</td>
<td>52%</td>
<td>Tb927.7.7160 (1.4e-52, 1)</td>
</tr>
<tr>
<td>specific dual</td>
<td>Tb927.9.13530</td>
<td>TcIL3000.9.5770</td>
<td>48%</td>
<td>66%</td>
<td>Tb927.9.13530 (7.2e-38, 1)</td>
</tr>
<tr>
<td>specificity phosphatase</td>
<td>Tb927.10.16120</td>
<td>TcIL3000.5.1940</td>
<td>37%</td>
<td>55%</td>
<td>Tb927.10.16120 (6.7e-67, 2)</td>
</tr>
<tr>
<td>homologue, putative</td>
<td>Tb927.11.200</td>
<td>TcIL3000.11.210</td>
<td>67%</td>
<td>80%</td>
<td>Tb927.11.200 (2.3e-80, 1)</td>
</tr>
<tr>
<td>protein kinase A</td>
<td>Tb927.7.4670</td>
<td>TcIL3000.0.15340</td>
<td>76%</td>
<td>88%</td>
<td>Tb927.7.4670 (8.1e-213, 1)</td>
</tr>
<tr>
<td>regulatory subunit</td>
<td>Tb927.2.1810</td>
<td>TcIL3000.2.500</td>
<td>62%</td>
<td>76%</td>
<td>Tb927.2.1810 (0.0, 1)</td>
</tr>
<tr>
<td>transcripion silencer</td>
<td>Tb927.5.3580</td>
<td>TcIL3000.0.40660</td>
<td>79%</td>
<td>90%</td>
<td>Tb927.5.3580 (2.0e-107, 1)</td>
</tr>
</tbody>
</table>
Table 4.2: Top hit of each BLAST search of T. b. brucei SIF-responsive proteins against T. congolense. T. b. brucei amino acid sequences were entered into the BLASTP tool of TritrypDB. T. congolense was selected as the target organism for each search. Default parameters were used (Expectation value = 10, Maximum descriptions/alignments (V=B) = 50, Low complexity filter = no). As of 3 Feb 2014, NCBI-BLAST was used to determine sequence similarity on TritrypDB. Prior versions of the search used WU-BLAST. Boxes in grey represent searches that used NCBI-BLAST, the previous searches used WU-BLAST. The final column reports the result of inputting the top T. congolense hit into a BLAST search against T. b. brucei. Only proteins where the top T. congolense hit returned the original T. brucei protein as the top hit are highlighted. T. b. brucei SIF-responsive proteins with strong T. congolense hits are highlighted in dark green, while those with weaker hits are highlighted in light green.

BLAST analysis was carried out for 35 genes identified from the T. brucei RNAi screen (Mony et al., 2014). To evaluate degree of conservation of the T. b. brucei SIF response in T. congolense and T. vivax, genes satisfying arbitrarily selected criteria for ‘good orthologues’ were short-listed (Table 4.3). Hits were only included if the T. vivax/ T. congolense amino acid sequence returned the original T. brucei protein as the top hit upon reciprocal BLAST. A low E value was considered important; indicating close sequence similarity to the query, but no strict cut off was used. When a cut off of ≥50% identity was used 16 out of 35 (46%) T. vivax genes were considered to be ‘good orthologues’, along with 16 out of 35 (46%) T. congolense genes. When a less stringent threshold of ≥30% identity was applied 26 out of 35...
(74%) *T. vivax* genes were considered to be ‘good orthologues’, with 25 out of 35 (71%) *T. congolense* genes.

A number of orthologues were excluded from the ‘good orthologue’ shortlist for not being reciprocal BLAST top hits. In some cases this is likely a result of some of *T. b. brucei* genes being part of arrays with very similar sequence. For example, 3 hits from the screen are Tb927.10.5930, Tb927.10.5940 and Tb927.10.5950. These protein kinases are part of an array of very similar genes such that RNAi targeting Tb927.10.5930 is expected to also target Tb927.10.5950 and Tb927.10.5940 (97% & 99% identical to Tb927.10.5930 respectively). All of these genes have TcIL3000.10.4970 as the top *T. congolense* hit, but only one can be a best hit by reciprocal BLAST. As a result, two of the genes are excluded from the ‘good orthologue’ short-list. However, the 3 arrayed *T. b. brucei* genes show amino acid sequence similarity to 3 closely related *T. congolense* genes - TcIL3000.10.4970, TcIL3000.10.4990, TcIL3000.10.4980, and occur at similar locations along chromosome 10. All of these genes occur in an orthologous group and are listed as syntenic.

Similarly the adjacent RNA-binding proteins Tb927.10.12100 and Tb927.10.12090 shared the same top hit, TcIL3000.10.10320, a BLAST search with which resulted in Tb927.10.12090 as the reciprocal top hit. Thus, the Tb927.10.12100 orthologue does not count towards the ‘good orthologues’ in the shortlist, and can be considered a false negative.

Therefore, when classifying 74% and 71% of the *T. b. brucei* genes of interest with good *T. vivax* and *T. congolense* orthologues respectively, care needs to be taken with interpretation. This does not necessarily indicate that other *T. b. brucei* genes from the screen are absent in *T. congolense* and *T. vivax*. Genes may have been excluded for comprising an array of similar sequence, sharing a common orthologue. Additionally, lack of a convincing orthologue for some genes may not infer lack of a complete SIF-response pathway in these other trypanosome species. To date, not all *T. b. brucei* genes from the screen have been targeted individually by RNAi and validated for SIF resistance *in vivo*, so some may not be part of the physiological SIF-response.
In summary, BLAST analysis supports conservation of the SIF-response genes between different trypanosome species, although sequence identity can be low. Identified orthologues can then be taken forward to test functional conservation. Our initial focus was placed on orthologues of *T. b. brucei* SIF-response genes validated *in vivo* (Table 4.3 and (Mony et al., 2014)). These genes have been individually targeted by RNAi in monomorphic cell lines resulting in resistance to growth inhibition by 8-CPT-cAMP *in vitro*, and in pleomorphic cell lines RNAi resulted in cells refractory to stumpy induction *in vivo*.
Table 4.3: Promising orthologues of SIF-responsive proteins identified in *T. vivax* and *T. congolense*. Hits with low E value, % identity over 30%, and hits for which the *T. vivax*/*T. congolense* amino acid sequence returned the original *T. b. brucei* protein as the top hit when put through BLAST.

<table>
<thead>
<tr>
<th>Predicted gene function</th>
<th><em>T. b. brucei</em> gene ID</th>
<th><em>T. vivax</em> gene ID</th>
<th><em>T. congolense</em> gene ID</th>
<th><em>T. brucei</em> RNAi SIF resistant in vivo?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3620</td>
<td>TyY486_0403330</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5940</td>
<td>TyY486_1005880</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.2.2720</td>
<td>TyY486_0200560</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.15020</td>
<td>TyY486_1014430</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 1)</td>
<td>Tb927.11.6600</td>
<td>TyY486_1106660</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 2)</td>
<td>Tb927.9.4090</td>
<td>TyY486_0901410</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Adenylosuccinate synthetase (Purine pathway)</td>
<td>Tb927.11.3650</td>
<td>TyY486_1103820</td>
<td>TcIL3000.11.3610</td>
<td></td>
</tr>
<tr>
<td>Adenylosuccinate Lyase (Purine pathway)</td>
<td>Tb927.9.7550</td>
<td>TyY486_0902900</td>
<td>TcIL3000.0.18960</td>
<td></td>
</tr>
<tr>
<td>GMP synthase</td>
<td>Tb927.7.2100</td>
<td>TyY486_0701820</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>ubiquitin activating enzyme, putative</td>
<td>Tb927.2.4020</td>
<td>TyY486_0201010</td>
<td>TcIL3000.2.550</td>
<td></td>
</tr>
<tr>
<td>serine/threonine protein kinase, putative</td>
<td>Tb927.3.4560</td>
<td>TyY486_0303830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein, conserved (Hyp 3)</td>
<td>Tb927.4.670</td>
<td>TyY486_0400390</td>
<td>TcIL3000.4.290</td>
<td></td>
</tr>
<tr>
<td>adenosine kinase, putative</td>
<td>Tb927.6.2300</td>
<td>TyY486_0601850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>adenosine kinase, putative</td>
<td>Tb927.6.2360</td>
<td></td>
<td>TcIL3000.6.1860</td>
<td>✓</td>
</tr>
<tr>
<td>kinetoplastid-specific dual specificity phosphatase, putative</td>
<td>Tb927.7.7160</td>
<td>TyY486_0706794</td>
<td>TcIL3000.0.52520</td>
<td></td>
</tr>
<tr>
<td>hypothetical protein, conserved (Hyp 6)</td>
<td>Tb927.9.13530</td>
<td>TyY486_0906760</td>
<td>TcIL3000.9.5770</td>
<td></td>
</tr>
<tr>
<td>inosine-5'-monophosphate dehydrogenase</td>
<td>Tb927.10.16120</td>
<td>TyY486_0201010</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>phosphatase and tensin homologue, putative</td>
<td>Tb927.11.290</td>
<td>TyY486_1100170</td>
<td>TcIL3000.11.210</td>
<td></td>
</tr>
<tr>
<td>protein kinase A regulatory subunit</td>
<td>Tb927.11.4610</td>
<td>TyY486_1104700</td>
<td>TcIL3000.0.15340</td>
<td></td>
</tr>
<tr>
<td>transcription silencer (ISWI)</td>
<td>Tb927.2.1610</td>
<td>TyY486_0202000</td>
<td>TcIL3000.2_50</td>
<td></td>
</tr>
<tr>
<td>Phosphoglycerate mutase</td>
<td>Tb927.5.3580</td>
<td>TyY486_0502950</td>
<td>TcIL3000.0_40660</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein (Hyp 5)</td>
<td>Tb927.8.2860</td>
<td>TyY486_0802390</td>
<td>TcIL3000.0_41460</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 8)</td>
<td>Tb927.11.300</td>
<td></td>
<td>TcIL3000.11.220</td>
<td></td>
</tr>
<tr>
<td>DNA repair protein, putative (Hyp 9)</td>
<td>Tb927.11.750</td>
<td>TyY486_1100600</td>
<td>TcIL3000.0_15260</td>
<td></td>
</tr>
<tr>
<td>Product: protein phosphatase 2C, putative</td>
<td>Tb927.11.760</td>
<td>TyY486_1100610</td>
<td>TcIL3000.0_15250</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 10)</td>
<td>Tb927.11.1640</td>
<td>TyY486_1101570</td>
<td>TcIL3000.11.1420</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 11)</td>
<td>Tb927.11.2250</td>
<td></td>
<td>TcIL3000.11.2000</td>
<td></td>
</tr>
<tr>
<td>Hypothetical protein, conserved (Hyp 13)</td>
<td>Tb927.11.11470</td>
<td>TyY486_1112490</td>
<td>TcIL3000.0_22050</td>
<td></td>
</tr>
<tr>
<td>Trichohyalin, putative</td>
<td>Tb927.11.11480</td>
<td>TyY486_1112510</td>
<td>TcIL3000.0_28560</td>
<td></td>
</tr>
</tbody>
</table>

Key: % Identity > 50% % Identity > 30%
4.3 Strategy to test capacity of identified orthologues for functional complementation

To assess functional conservation, a selected *T. congolense* or *T. vivax* protein was overexpressed in a *T. b. brucei* cell line in which the orthologous *T. b. brucei* gene was knocked down by RNAi (Figure 4.1). Pleomorphic *T. b. brucei* AnTat 1.1 90.13 cell lines with RNAi targeting individual SIF responsive genes had previously been generated in the laboratory (Mony et al., 2014). It was hoped that the capacity for rescue could be assessed through transfection of these cell lines with constructs engineered for inducible overexpression of a *T. vivax* or *T. congolense* orthologue.

*Figure 4.1: Experimental strategy to assess functional complementation of SIF signal components in *T. b. brucei*, *T. congolense* and *T. vivax*

One caveat of this approach is that the foreign gene selected for expression needs to be sufficiently different from the RNAi target, so that it is not knocked down in parallel. In trypanosomes RNAi works by cleavage of dsRNA into 21-25nt siRNA, where transcripts that are targeted by the siRNA are degraded. Therefore if there were stretches of sequence in a *T. vivax* or *T. congolense* gene with 21+ nucleotides in a row identical to the target *T. b. brucei* RNAi sequence, then there was potential for that gene to be co-silenced (Durand-Dubief et al., 2003). Consequently, a BLAST search against the *T. vivax*/ *T. congolense* genome using the *T. b. brucei* RNAi targeting sequence was implemented to assess whether there were regions where co-silencing was possible (Table 4.4). In most cases, the analysis could not rule out the possibility of cross RNAi since, although there were not stretches with 20+ nucleotides in a row identical, in some cases stretches of identity were only broken up by single nucleotide differences potentially allowing RNAi. However, the knock down of the *T. vivax* and *T. congolense* genes could be inefficient compared to the
knock down of the *T. b. brucei* transcripts, particularly if the *T. vivax* or *T. congoense* gene is overexpressed to a sufficient extent.

<table>
<thead>
<tr>
<th>Gene function</th>
<th>T. b. brucei gene ID</th>
<th>T. vivax gene ID</th>
<th>T. congoense gene ID</th>
<th>Will the RNAi used in <em>T. brucei</em> also knock down the <em>T. vivax</em> gene?</th>
<th>Will the RNAi used in <em>T. brucei</em> also knock down the <em>T. congoense</em> gene?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein phosphatase</td>
<td>Tb927.4.3620</td>
<td>TvY486_0403330</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>Protein Kinase</td>
<td>Tb927.10.5940</td>
<td>TvY486_1005880</td>
<td>TcIL3000.10.4970</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 1)</td>
<td>Tb927.11.6600</td>
<td>TvY486_1106960</td>
<td>TcIL3000.11.7200</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hypothetical Protein (Hyp 2)</td>
<td>Tb927.9.4080</td>
<td>TvY486_0901410</td>
<td>TcIL3000.0.19510</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>RNA-binding protein</td>
<td>Tb927.10.12100</td>
<td>TvY486_1011760</td>
<td>TcIL3000.10.10320</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Adenylosuccinate synthetase (Purine pathway)</td>
<td>Tb927.11.3650</td>
<td>TvY486_1103820</td>
<td>TcIL3000.11.3610</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Adenylosuccinate Lyase (Purine pathway)</td>
<td>Tb927.9.7550</td>
<td>TvY486_0902900</td>
<td>TcIL3000.0.18960</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Key: ✓ Probably – there are 20 bases in a row identical and there are stretches where 20+ identical bases are only broken up by single nucleotide differences. 
? Possibly - there are not 20+ bases in a row identical, but there are stretches where 20+ identical bases are only broken up by single nucleotide differences. 
✗ Probably not – no BLAST hits were returned for the RNAi sequence so it is unlikely it would target anything in *T. congoense*/*T. vivax*.

Table 4.4: Likelihood of cross-RNAi targeting *T. congoense* or *T. vivax* orthologue

### 4.4 Selection of identified orthologues for functional complementation

#### 4.4.1 Adenylosuccinate lyase (Tb927.9.7550)

One gene to emerge from the initial RNAi screen in *T. b. brucei* was an Adenylosuccinate lyase (Tb927.9.7550/ previously known as Tb9.160.5560). However, when a pleomorphic RNAi line was tested *in vivo* to validate the hit, it was found that knock down did not cause resistance to SIF in mice. One possible explanation is that trypanosomes cannot synthesise purines *de novo* and rely on a salvage pathway for production (Vertommen et al., 2008), such that Adenylosuccinate synthetase and Adenylosuccinate lyase are important genes for which knock down has a growth effect *in vitro*. Therefore it was possible that the cell permeable cAMP used in the screen was channeled into another pathway to compensate for the Adenylosuccinate lyase knock down, rather than the pathway for stumpy induction, allowing such cells to outgrow. Nevertheless, I chose to overexpress *T. congoense* and *T. vivax* Adenylosuccinate lyase orthologues, in the
context of a *T. b. brucei* RNAi knock down line, to determine whether functional complementation with a *T. congo* or *T. vivax* gene is possible in *T. b. brucei*. Functional complementation by Adenylosuccinate lyase is expected to occur independently of the presence or absence of a conserved density dependent signaling pathway in *T. b. brucei*, *T. congo* and *T. vivax*. However, this gene was expected to provide a proof of principle for the ability to compensate an RNAi-mediated phenotype by complementation.

### 4.4.2 Hypothetical protein 2 (Tb927.9.4080)

A major hit from the initial RNAi screen was a hypothetical protein (Tb927.9.4080, (previously known as Tb09.160.2650)). This hit had been successfully validated both *in vitro*, where the knock down line showed resistance to 8-CPT-cAMP induced growth arrest, and *in vivo*, where parasites showed resistance to SIF-induced stumpy formation (Mony et al., 2014). In addition, the knock down did not cause a growth effect *in vitro*. The most similar *T. vivax* orthologue to the *T. b. brucei* gene (in terms of amino acid sequence) is TvY486_0901410 (E value: 1.9e-228, % identity: 49%, % similarity: 65%). When a reciprocal BLAST search against *T. b. brucei* was carried out, the *T. b. brucei* gene Tb09.160.2650 was returned as the top hit (Table 4.1). The most similar *T. congo* orthologue to the *T. b. brucei* gene (in terms of amino acid sequence) is TcIL3000.0.19510 (E value: 5.9e-262, % identity: 45%, % similarity: 58%), which also returned Tb09.160.2650 by reciprocal BLAST. A pairwise comparison based on amino acid sequence was constructed for the predicted orthologous group (Table 4.5). For the entire alignment see Appendix 2.1.

The TcIL3000.9.1290 gene, but not the TcIL3000.0.19510 gene is listed as syntenic to Tb927.9.4080. However, this is probably because TcIL3000.0.19510 does not have a chromosome assigned. TcIL3000.0.19510 is more similar in length to Tb927.9.4080, and it is possible that TcIL3000.9.1290 is a truncated version of the gene.
Orthologues of Tb927.9.4080 were also identified in other species with sequence available on TritrypDB, including *T. cruzi* and various Leishmania species (Figure 4.2).

**Table 4.5: Pairwise comparison of Hypothetical protein 2 orthologous group.** Analysis was carried out using CLC genomics workbench for the entire amino acid sequence of each protein.

**Figure 4.2: TritrypDB WU-BLAST results for Tb927.9.4080 in Tritryp species.** The *T. b. brucei* gene used for the BLAST analysis is highlighted in yellow, identified *T. congoense* orthologues are highlighted in blue and a *T. vivax* orthologue is highlighted in green.

Overall, the strongest orthologues of Tb927.9.4080 are found within the African trypanosomes. However, one *T. cruzi* orthologue has a lower E value than the top *T. vivax* orthologue. Interestingly, the scores and E values for the orthologues in the more distantly related *Leishmania* species and *T. cruzi* compare favourably with the orthologues within African trypanosomes. Thus there could be a conserved function in these more distantly related kinetoplastid parasites.
4.5 Bioinformatics analysis of Hypothetical protein 2 (Tb927.9.4080/ TbHyp2)

Tb927.9.4080 (TbHyp2) was designated as a positive regulator of gene expression in a genome-wide tethering screen to identify post-transcriptional regulators in *T. b. brucei* (Erben et al., 2014). Although it may not bind to RNA, other publications show it to interact with DRBD3 and MKT1 (Singh et al., 2014). The interaction with MKT1 may implicate TbHyp2 in post-transcriptional regulation. DRBD3 is thought to be involved in stabilization of a subset of mRNAs encoding membrane proteins that are up-regulated in procyclic forms (Estévez, 2008). Hence, interaction of the stumpy inducer TbHyp2 with this regulator could play a role in preparing stumpy forms for transmission to the insect vector.

Predicted domains in TbHyp2 and orthologues were then identified using various bioinformatics tools. In the following sections, numbers in brackets correspond to E values provided by the various databases if available (unless specified).

Initially, an InterPro search was used to identify conserved domains that could hint at function for this hypothetical protein and it’s orthologues. InterPro (Mitchell et al., 2015) draws on information from several member databases (including PFAM and Prosite) for output, providing a useful overview. An InterPro search revealed that TbHyp2 is predicted to have a zinc-finger domain (DksA/TraR C4 type, IPR000962), based on information from Prosite, and a separate search using ScanProsite (de Castro et al., 2006) revealed the same domain (score - 9.605) (Figure 4.3). This domain is particularly found in prokaryotic cells and is part of a bacterial response to amino acid starvation. The DksA protein in *E. coli* is involved in limiting transcription from certain stable promoters by binding RNA polymerase. The protein works in concert with the internal signal ppGpp (guanosine tetraphosphate or alarmone) to create a synergistic effect. Additionally, transcripts involved in amino acid biosynthesis are upregulated by both the direct and indirect action of ppGpp and DksA (Paul et al., 2005). The presence of this domain, found in a bacterial stress response protein, in TbHyp2 is of interest as it could implicate a role for this *T. b. brucei* hypothetical protein in the regulation of transcription as part of the SIF signaling response, particularly since DksA has also been proposed to have an effect...
on quorum sensing in bacteria. This domain is not found to be present in TcIL3000.0.19510 (TcHyp2) or TvY486_0901410 (TvHyp2). However, an aminoacyl tRNA ligase I signature is found in TcHyp2, and a regulator of chromosome condensation I signature is found in TvHyp2.

Figure 4.3 Alignment of Hyp2 orthologues showing predicted dksA C4-type zinc finger domain of TbHyp2. Highlighted in yellow is the predicted zinc-finger, highlighted in green are the surrounding amino acids used by Prosite to predict presence of the domain. Below is a schematic of the location of the predicted domain in the T. b. brucei protein, accompanied by a 1000 amino acid ruler. The schematic was taken from Prosite.

A further domain identified through the InterPro search with TbHyp2 was a signature of P-loop containing nucleoside triphosphate hydrolases (IPR027417). This was identified through the Gene3D and Superfamily databases, and the same result was obtained when searching these databases individually (Gene3D - 2.4e-14, Superfamily - 1.76e-13). TcHyp2 (Gene3D – 9.4e-15, Superfamily – 1.03e-13) and TvHyp2 (Gene3D – 3.7 e-13, Superfamily – 1.11e-11) were also predicted to have the domain. Both databases use Hidden Markov models based on particular protein datasets to infer protein domain type from amino acid sequences of uncharacterized proteins. Gene3D draws protein information from the CATH database (Sillitoe et al., 2015) which groups protein domains into superfamilies when they are suspected to be derived from a common ancestor. Superfamily uses the Structural Classification of Proteins database (SCOP (Andreeva et al., 2007)) which orders proteins based on structure and evolutionary relationships. Furthermore, some of the Leishmania orthologues identified (Figure 4.2) are also predicted to have the P-loop containing
nucleoside triphosphate hydrolase domain by Superfamily; LinJ.01.0480 (3.9e-18), LmjF.01.0460 (1e-18), LbrM.01.0510 (2e-19).

A PFAM (Finn et al., 2014) search revealed an AAA+ domain as an insignificant match for the TbHyp2 (AAA 18 - 330, AAA 16 - 0.00012, Thymidylate kinase – 18), TcHyp2 (AAA 18 - 1.2e-5), or TvHyp2 (AAA 18 - 2.4e-05) amino acid sequence. Although not a significant match in PFAM, AAA+ domains form part of the P-loop containing nucleoside triphosphate hydrolase superfamily, so PFAM supports the prediction by Gene3D and Superfamily. AAA stands for ATPases associated with a variety of cellular activities. AAA+ domain proteins often form oligomers with pore structures. AAA+ domain proteins have diverse cellular functions including housekeeping, cell division and differentiation (Patel and Latterich, 1998), and often perform chaperone-type roles in assembly, disassembly or functioning of protein complexes (Neuwald et al., 1999).

A conserved domain search using the NCBI tool CDD (Marchler-Bauer et al., 2015) also revealed two AAA domains in the TbHyp2 sequence (9.11e-03, 3.16e-03), alongside an adenylation domain of proteins similar to ATP-dependent polynucleotide ligases (1.24e-3). These domains are also identified in TcHyp2 (AAA - 2.59e-4, adenylation - 4.85e-05) and TvHyp2 (AAA 18 - 6.73e-07, adenylation - 6.46e-04). CDD pulls in information from PFAM and other databases, and uses RPS BLAST with well annotated multiple sequence alignment models to identify conserved domains.

Both Gene3D (4.1e-09) and Superfamily (5.28e-13) also predict a domain belonging to the DNA ligase/mRNA capping superfamily in TbHyp2. This domain is also predicted to be present in TcHyp2 (Gene3D - 5.28e-11, Superfamily – 1.3e-6) and TvHyp2 (Gene3D - 2.9e-5, Superfamily - 4.27e-08). PFAM also predicts this domain as an insignificant match: TvHyp2 (0.17), TcHyp2 (0.0075). Two putative transposase DNA-binding domains are also weakly predicted for TcHyp2 (6.4, 8.2).

Finally the Eukaryotic Linear Motifs (ELM) database was searched (Dinkel et al., 2012). This identifies short linear motifs in the query sequence; such motifs can inform on cell signalling, protein localization or degradation. Numerous motifs were
identified in the TbHyp2 sequence (Table 4.6). Some motifs indicate that this hypothetical protein may be trafficked into and out of the nucleus: 3 C-terminal nuclear localization signals (2.588e^-4, 7.252e^-4, 1.276e^-3), 1 nuclear export signal (7.626e^-4). Further trafficking signals include an endosome-lysosome basolateral sorting signal (2.758e^-03) used to signal trafficking of transmembrane proteins from the trans-golgi network to the endosomal-lysosomal compartments, as well as an ER retention/retrieval motif (5.369e^-03) found in transmembrane proteins. Potential N-glycosylation sites were also identified (5.018e^-3), as well as numerous phosphorylation sites, including ones favoured by PKA-type AGC kinase (2.315e^-3), NEK2 (9.798e^-3) or cyclin-dependent protein kinase (CDK, 1.929e^-3). The potential interaction with CDK could link TbHyp2 with cell cycle control. Additionally, if the PKA-type and NEK2 phosphorylation sites can be used by the PKA and NEK proteins that are also predicted to form part of the stumpy-induction pathway in *T. b. brucei* ((Mony et al., 2014) and Table 4.3), this would identify a link between these pathway components and TbHyp2. The presence of a Protein Phosphatase 1 (PP1) catalytic subunit interaction motif (8.301e^-4), which can help target docked proteins for dephosphorylation, also provides another potential link into the known stumpy-induction pathway through PP1. A docking motif that assists in regulating specific interactions with the MAPK cascade is also predicted (4.324e^-3), possibly linking into the stumpy-induction pathway through MEK kinase.

<table>
<thead>
<tr>
<th>ELM Name</th>
<th>ELM description</th>
<th>Probability</th>
<th>Present in <em>T. congolense</em></th>
<th>Present in <em>T. vivax</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>TRG_NES_CRM1_1</td>
<td>Some proteins re-exported from the nucleus contain a Leucine-rich nuclear export signal (NES) binding to the CRM1 exportin protein</td>
<td>7.626e^-04</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>TRG_NLS_Bipartite_1</td>
<td>Bipartite variant of the classical basically charged NLS.</td>
<td>2.588e^-04</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>TRG_NLS_MonoExtC_3</td>
<td>Monopartite variant of the classical basically charged NLS. C-extended version.</td>
<td>7.252e^-04</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td><strong>TRG_NLS_Mono</strong></td>
<td>Monopartite variant of the classical basically charged NLS. N-extended version.</td>
<td>1.276e^−3</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td><strong>TRG_ENDO CYTIC</strong></td>
<td>Tyrosine-based sorting signal responsible for the interaction with μ subunit of AP (Adaptor Protein) complex</td>
<td>2.587e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>TRG_ER_di Arg_1</strong></td>
<td>The di-Arg ER retention motif is defined by two consecutive arginine residues (RR) or with a single residue insertion (RXR). The motif is completed by an adjacent hydrophobic/arginine residue which may be on either side of the Arg pair.</td>
<td>5.369e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>TRG_LysEn d_AP s AcLL_1</strong></td>
<td>Sorting and internalisation signal found in the cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes</td>
<td>2.758e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_N-GLC_1</strong></td>
<td>Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.</td>
<td>5.018e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_N-GLC_2</strong></td>
<td>Atypical motif for N-glycosylation site. Examples are Human CD69, which is uniquely glycosylated at typical (Asn-X-Ser/Thr) and atypical (Asn-X-Cys) motifs, beta protein C</td>
<td>2.973e^−4</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_PKA_1</strong></td>
<td>Main preference for PKA-type AGC kinase phosphorylation.</td>
<td>2.315e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_PKA_2</strong></td>
<td>Secondary preference for PKA-type AGC kinase phosphorylation</td>
<td>9.458e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_NEK2_1</strong></td>
<td>NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.</td>
<td>9.798e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>MOD_NEK2_2</strong></td>
<td>NEK2 phosphorylation motif with specific set of residues in the +1 and +2 position to compensate for less favorable residues in the -3 position.</td>
<td>1.295e^−3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MOD_CDK_1</td>
<td>Substrate motif for phosphorylation by CDK</td>
<td>1.929e-03</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DOC_CYCLIN_N_1</td>
<td>Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site. Also used by cyclin/cdk inhibitors</td>
<td>5.324e-03</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DOC_PP1_R_VXF_1</td>
<td>Protein phosphatase 1 catalytic subunit (PP1c) interacting motif binds targeting proteins that dock to the substrate for dephosphorylation.</td>
<td>8.301e-04</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DOC_MAPK_1</td>
<td>MAPK interacting molecules (e.g. MAPKKs, substrates, phosphatases) carry docking motif that help to regulate specific interaction in the MAPK cascade. The classic motif approximates (R/K)xxxx#x# where # is a hydrophobic residue.</td>
<td>4.324e-03</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>MOD_ProDK_in_1</td>
<td>Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes</td>
<td>1.543e-02</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DOC_WD40_RPT_OR_TOS_1</td>
<td>The TOR pathway adaptor protein Raptor links the mTOR kinase to the TOS motif containing substrates 4E-BP1 and S6-beta kinases. Proteins with TOR motif (e.g. 4E-BP1, S6KB1) participate in the transcription mechanism.</td>
<td>2.075e-05</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>MOD_LATS_1</td>
<td>The LATS phosphorylation motif is recognised by the LATS kinases for Ser/Thr phosphorylation. Substrates are often found toward the end of the Hippo signalling pathway.</td>
<td>4.776e-04</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>LIG_T_RAF2_1</td>
<td>Major TRAF2-binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails.</td>
<td>4.300e-03</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LIG_R_b_pA_Bgroove_1</td>
<td>The LxxLFD motif binds in a deep groove between pocket A and pocket B of the Retinoblastoma protein</td>
<td>2.475e-03</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>LIG_R_b_LxC_xE_1</td>
<td>Interacts with the Retinoblastoma protein</td>
<td>1.350e-04</td>
<td>✓</td>
<td>x</td>
</tr>
</tbody>
</table>
Table 4.6 Summary of Eukaryotic Linear Motifs found in TbHyp2 using the ELM database. Presence ✓ or absence ✗ in *T. congolense* and *T. vivax* is marked.

<table>
<thead>
<tr>
<th>Motif (LIG)</th>
<th>Description</th>
<th>P-value</th>
<th>Presence</th>
<th>Absence</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIG_BRCT_BRCA_1_1</td>
<td>Phosphopeptide motif which directly interacts with the BRCT (carboxy-terminal) domain of the Breast Cancer Gene BRCA1 with low affinity</td>
<td>1.912e-03</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>LIG_eIF4E_1</td>
<td>Motif binding to the dorsal surface of eIF4E.</td>
<td>1.891e-04</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

eIF4E binds the RNA cap for the progression of translation in *T. brucei*, inhibitors of protein synthesis can act by competing with eIF4G for binding to eIF4E (Clayton and Shapira, 2007). The presence of a motif for eIF4E binding (1.891e-04) in TbHyp2 supports a role for this protein in post-transcriptional regulation as found in the screen for regulators mentioned above (Erben et al., 2014).

Interestingly, a number of motifs suggest interactions with pathways known to be involved in the control of mammalian cell proliferation, which are commonly disrupted in cancer. Two domains are predicted to be involved in interaction with the Retinoblastoma protein (RB) (2.475e-05, 1.350e-04). RB is classified as a tumour suppressor gene, with a role in inhibiting entry into S-phase through interaction with members of the E2F family of transcription factors. Loss of function of RB is found in a number of cancers (Sherr and McCormick, 2002). An interaction with another tumour suppressor gene, BRCA1, is also predicted (1.912e-03). BRCA1 is involved in the DNA damage response and also feeds into cell cycle checkpoint control (Roy et al., 2012). Another motif, LIG_TRAF2_1 (4.300e-03) indicates potential interaction with the tumour necrosis factor receptor (TNFR) signalling pathway, which can be linked to cell death, proliferation and differentiation (Dempsey et al., 2003). Additionally, a motif possibly targeted for phosphorylation by components of the Hippo signalling pathway, a pathway frequently disrupted in human cancer (Harvey et al., 2013), was predicted. Signalling through the Hippo pathway promotes cell death and differentiation and inhibits cell proliferation (Yu and Guan, 2013). Thus, TbHyp2 is predicted to link in to a number of pathways that inhibit cell proliferation, an observation of interest considering that TbHyp2 is shown to have a role in signalling the transition to the growth-arrested stumpy form of *T. b. brucei*. 
A summary of interesting short linear motifs was overlaid onto the amino acid sequence using Protter (Omasits et al., 2014) (Appendix 2.2).

The ELM database was then searched with the *T. congo**lo**nese* and *T. vivax* orthologues of TbHyp2, to compare their potential interactions. While all nuclear trafficking signals found for *T. b. brucei* are also found for *T. congo**lo**nese*, only the nuclear export signal is found for *T. vivax*. The other signals relating to trafficking are identified in all three species. Potential N-glycosylation sites were conserved. Phosphorylation sites that could link TbHyp2 with other components of the SIF response pathway were also found in *T. congo**lo**nese* and *T. vivax* orthologues. Motifs for interaction with PP1 and MAPK cascade also appear conserved in the African trypanosomes. The motif for eIF4E binding is conserved, as are a number of motifs linking TbHyp2 and orthologues to identified eukaryotic tumour suppressor pathways and cell cycle control. Evidence from the ELM database indicates that the three orthologues may have similar molecular interactions, with potential for conserved roles in pathways relevant to parasite growth control.

Finally, I-TASSER (Yang et al., 2015) was used to make structural predictions about the three orthologues. The top template used was the same for all three species, Human complement factor H. Only one model each had a C-score > -1.5, which indicates a model of correct global topology. However, none of the models appeared a very good fit. The top GO template in the Protein Data Bank (PDB) for TbHyp2 (C-score=0.08), TcHyp2 (C-score=0.08) and TvHyp2 (C-score=0.08) is the cytoplasmic dynein heavy chain motor domain. The C-scores indicate low confidence in these predictions (scores range from 0-1 where higher scores are better). Such domains help power ciliary beating, transport intracellular cargos, and construct mitotic spindles, and evolved from ring-shaped hexameric AAA-family ATPases, perhaps consistent with the AAA+ domain predicted weakly by PFAM for TcHyp2 (see above). The top molecular function GO terms, based on PDB templates, associated with Hyp2 orthologues are ATP-binding (GO score: TbHyp2 and TvHyp2 = 0.14, TcHyp2 = 0.07) and microtubule motor activity (GO score: TbHyp2 and TvHyp2 = 0.08, TcHyp2 = 0.07). Once again the GO scores associated with the
predictions from I-TASSER do not indicate high confidence, so should be considered in line with other predictive tools, and interpreted with caution.

Bioinformatic analysis of TbHyp2 and orthologues indicates that these proteins may perform multiple roles in the cell, with roles in cell cycle control implied.

**4.6 Cloning strategy for overexpression**

Primers were designed that would result in amplification of the gene of interest with appropriate terminal restriction sites for insertion into the pDex577-Y vector (Appendix 2.3). For primer sequences used see Appendix 2.4. The pDex577-Y vector can be used for tetracycline-inducible overexpression of a gene of interest with an N- or C-terminal TY epitope tag and/or a YFP tag (Kelly et al., 2007). The TY tag is small, and use of an anti-TY antibody (BB2, (Bastin et al., 1996)) allows detection of protein expression when a protein-specific antibody is not available. Therefore, the *T. b. brucei, T. congolense* and *T. vivax* orthologue genes of interest were expressed with either an N- or C-terminal TY tag. For expression with an N-terminal TY tag, the gene of interest and vector must both be digested with *SpeI* and *BamHI*. For expression with a C-terminal TY-tag, both must be digested with *HindIII* and *XbaI*.

*T. congolense* IL3000 genomic DNA was isolated from bloodstream form parasites collected from an infected MF1 mouse and purified over a DE52 column. *T. vivax* genomic DNA was isolated from cultured IL1392 epimastigote forms. *T. b. brucei* genomic DNA was isolated from AnTat 1.1 90.13 bloodstream forms. Genes of interest were amplified from genomic DNA and cloned initially into pGEM-T Easy. Once the expected sequence was verified, each gene of interest was excised and cloned into pDex577-Y vector. Diagnostic restriction digests were carried out and positive clones were verified by DNA sequencing to confirm the presence of the insert with an appropriately positioned TY tag.

TcHyp2 proved difficult to ligate into pGEM-T Easy. The TcHyp2 gene is 3852bp in length, whereas the pGEM-T Easy vector itself is only ~3000bp. Therefore, new primers were designed to incorporate 6 extra bases adjacent to the restriction site (see Appendix 2.4), to allow direct digest of the PCR product and insertion immediately
into pDex577-Y, circumventing the use of pGEM Teasy. Following this, clones with an insert of expected size were identified by diagnostic restriction digest. Sequencing confirmed that the insert sequence was intact and in frame with the TY tag. Additional sequencing primers were also designed to bind within pDex577-Y to confirm the sequence at the start and end of the insert, plus sequencing primers were designed to confirm the sequence of the centre of this long gene. The amino acid sequence was determined to be complete and as predicted, with the exception of 12 amino acid changes (~1% of the overall protein) that occurred in multiple clones. These changes could reflect differences between the genome strain and the IL3000 T. congolense parasites used to isolate the gDNA or, alternatively, could have been a consequence of errors by the DNA polymerase. For this reason, the PCR was repeated and the purified product sent for sequencing. This confirmed that all amino acid changes were conserved and particular to the parasite cells used, except for one (P>Q). Given the size of the overall protein, it was decided to proceed to overexpress the protein with this one incorrect amino acid (glutamine (uncharged polar side chain) instead of proline (nonpolar side chain)). This change did not effect any predicted linear motifs. Both TvHyp2 and TbHyp2 were cloned directly into pDex577-Y, and their sequence verified. The TvHyp2 sequence was intact and as predicted. However, the TbHyp2 sequence had 6 amino acid changes from those predicted. To test whether these changes were due to PCR errors or differences between the AnTat 1.1 90.13 genomic DNA used and the TREU927 genome strain, an independent PCR reaction was sequenced. This verified that the amino acid changes detected were differences between the strain used and the genome strain.

Completed overexpression constructs were transfected into AnTat 1.1 90.13 bloodstream forms with or without capacity for inducible gene knock down of the corresponding T. b. brucei SIF-response pathway component. The pDex577-Y (Kelly et al., 2007) construct integrates into the 177bp repeats of minichromosomes or intermediate chromosomes, and is thus designed to have tighter regulation of inducible gene expression as these regions are usually transcriptionally silent (Wickstead et al., 2002). A tetracycline-inducible T7 promoter controls transgene expression, and successfully transfected parasites are selected in the presence of 1.5µg/ml phleomycin (a bleR gene is under control of a rRNA promoter).
4.7 Validation of functional complementation strategy

As mentioned previously, functional complementation with Adenylosuccinate lyase (ADSL) is expected to occur independently of the presence or absence of a conserved density dependent signaling pathway and so was used to validate the experimental strategy.

Knock down of *T. b. brucei* ADSL resulted in a growth defect *in vitro* (Figure 4.4A, Figure 4.5A, black lines). Growth was rescued when knock down of *T. b. brucei* ADSL was accompanied by overexpression of the *T. congolense* or *T. vivax* ADSL orthologue (Figure 4.4A, Figure 4.5A, red lines). This result indicates that the complementation strategy used in this analysis is capable of detecting when two orthologues are able to carry out the same function. In other words, the *T. congolense* or *T. vivax* ADSL protein is able to carry out the equivalent function in *T. b. brucei*. Inducible overexpression of the TY-tagged *T. congolense* ADSL (Figure 4.4B) and *T. vivax* ADSL (Figure 4.5B) proteins was confirmed by detection with an anti-TY antibody. Inducible knock down of the endogenous TbADSL transcript was confirmed by Northern Blotting, and the TbADSL transcript was found to be depleted in TbADSL RNAi cell lines with concomitant TcADSL or TvADSL overexpression (Figure 4.4C, Figure 4.5C). This confirmed that the observed rescue was the result of overexpression of either the *T. congolense* or *T. vivax* protein, and not reduced efficacy of the RNAi targeting the *T. b. brucei* transcript in these newly generated cell lines.
Figure 4.4: Functional complementation of T. b. brucei Adenylosuccinate lyase with T. congolense Adenylosuccinate lyase. A) Overexpression of the T. congolense ADSL orthologue rescues growth in cells depleted for the T. b. brucei transcript. Each data point represents the mean and SEM for triplicate wells. B) Western blot showing inducible expression of TY-tagged TcADSL (upper) with a G6PDH loading control (lower). C) Northern blot showing inducible knock down of TbADSL in the TbADSL RNAi cell line (RNAi) and TbADSL RNAi cell line with concomitant TcADSL overexpression (OE + RNAi).

Figure 4.5 Functional complementation of T. b. brucei Adenylosuccinate lyase with T. vivax Adenylosuccinate lyase A) Overexpression of the T. vivax ADSL orthologue rescues growth in cells depleted for the T. b. brucei transcript. Each data point represents the mean and SEM for triplicate wells. B) Western blot showing inducible expression of TY-tagged TvADSL (upper) with EF1 loading control (lower). C) Northern blot showing inducible knock down of TbADSL in TbADSL RNAi cell line (RNAi) and TbADSL RNAi cell line with concomitant TvADSL overexpression (OE + RNAi).
4.8 Characterization of Hyp2 overexpression *in vitro*

Following validation of the experimental strategy, orthologues of *T. b. brucei* Hyp2 were investigated as potentially conserved regulators of differentiation. Initially, the consequence of overexpression of each orthologue was characterised *in vitro*.

4.8.1 Growth arrest caused by overexpression

Overexpression of TbHyp2 and orthologues in an AnTat 1.1 90.13 background with or without knock out of endogenous TbHyp2 resulted in severe growth inhibition when a high concentration of doxycycline was used (1µg/ml). The extent of growth inhibition was lessened when a lower doxycycline concentration (10^-4 µg/ml) was used. This was the lowest concentration with detectable protein (Figure 4.6).
Overexpression of TbHyp2, TcHyp2 or TvHyp2 results in growth inhibition in vitro when high and low concentrations of doxycycline are used for induction. Parasites/ml measurements were taken for triplicate wells of each condition, with the mean and SEM displayed. All overexpressed Hyp2 orthologues were N-terminally TY-tagged, and were detected by western blot using BB2 antibody. Overexpression was induced in either an AnTat 1.1 90.13 background (A), or in the context of a TbHyp2 null mutant (B, the generation of which is discussed in section 4.8.5). Either a high (1µg/ml) or a low (10^{-4}µg/ml) concentration of doxycycline was used for induction.

Differentiation to the stumpy form is associated with accumulation in 1K1N (Shapiro et al., 1984). In order to determine if cell cycle arrest was linked to Hyp2 overexpression the cell cycle position of each population was analysed. However, induction of a high level of overexpression using 1µg/ml doxycycline was associated in some cases with a considerable proportion of cells with abnormal KN configurations (for examples see Appendix 2.6). This made it difficult to interpret the degree of cell cycle arrest under these conditions. Therefore, KN counts were carried out to investigate the degree of arrest following induction of Hyp2 overexpression by 10^{-4}µg/ml doxycycline, as compared to uninduced cells. Cells were categorised as 1K1N, 2K1N and 2K2N, with the latter two categories considered to be representative of proliferating cells. At least 250 cells were counted under each condition, including abnormal KN configurations. In all cases, Hyp2 overexpression resulted in a clear reduction in the proportion of proliferating cells (Figure 4.7). This
was mostly a result of accumulation in 1K1N, as abnormal configurations were minimal at this lower concentration of doxycycline (see Appendix 2.7).

Figure 4.7 Effect of overexpression of Hyp2 orthologues on cell cycle in vitro. For all orthologues Hyp2 overexpression using a low doxycycline concentration (10^{-4}µg/ml) for induction resulted in a significant reduction in proliferating cells (P values: ★ P < 0.05, ★★ P<0.005, ★★★★ P<0.0001). Unpaired t tests were performed using GraphPad Prism. Columns display mean and standard deviation for each condition based on three technical replicates. More than 250 cells were counted for each replicate. It was assumed that any variation within the cell line was minimal in comparison to the effect of induction.

4.8.2 Localization

Each orthologue was overexpressed with an N-terminal TY-tag, so was detectable using BB2 antibody (Bastin et al., 1996). Thus, the localization of the respective orthologues could be visualized and compared (Figure 4.8). The most frequently observed staining consisted of a few bright dots. This was not observed in every cell, but often comprised two bright dots either side of the nucleus. Occasionally three or more dots were observed, sometimes close to the kinetoplast or at the posterior end of the cell (which can be elongated with Hyp2 overexpression, see section 4.8.3).
### A  TbHyp2 overexpression

<table>
<thead>
<tr>
<th></th>
<th>BB2</th>
<th>DAPI</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dox (µg/ml)</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
</tr>
</tbody>
</table>

### B  TcHyp2 overexpression

<table>
<thead>
<tr>
<th></th>
<th>BB2</th>
<th>DAPI</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dox (µg/ml)</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
</tr>
</tbody>
</table>

### C  TnHyp2 overexpression

<table>
<thead>
<tr>
<th></th>
<th>BB2</th>
<th>DAPI</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dox (µg/ml)</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
<td>10^{-4}</td>
</tr>
</tbody>
</table>
D  TbHyp2 overexpression in TbHyp2 knockout – Clone 1

<table>
<thead>
<tr>
<th>Dox(µg/ml)</th>
<th>1</th>
<th>10^-4</th>
<th>1</th>
<th>10^-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

BB2

DAPI

Phase

---

E  TbHyp2 overexpression in TbHyp2 knockout – Clone 2

<table>
<thead>
<tr>
<th>Dox(µg/ml)</th>
<th>1</th>
<th>10^-4</th>
<th>1</th>
<th>10^-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

BB2

DAPI

Phase

---

F  TbHyp2 overexpression in TbHyp2 knockout

<table>
<thead>
<tr>
<th>Dox(µg/ml)</th>
<th>1</th>
<th>10^-4</th>
<th>1</th>
<th>10^-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

BB2

DAPI

Phase
Figure 4.8 Localisation of TY-tagged Hyp2 orthologues. Cells were incubated with high (1µg/ml) or low (10^{-4}µg/ml) concentrations of doxycycline for 1 or 2 days to induce overexpression of the Hyp2 orthologues. Cells were paraformaldehyde fixed prior to detection of TY-tagged Hyp2 with BB2 antibody (red). The nucleus and kinetoplast were visualised with DAPI (blue). In all images the scale bar represents 10µm. A) TbHyp2 overexpression was induced in an AnTat 1.1 90.13 background, and the overexpressed TbHyp2 was occasionally detected in a punctate pattern of 2 or more dots, or at the elongated posterior tip of the cell. B) TcHyp2 overexpression was induced in an AnTat 1.1 90.13 background, and the punctate patterns observed were comparable to those observed for TbHyp2 localisation. C) TvHyp2 overexpression was induced in an AnTat 1.1 90.13 background, and a punctate localisation pattern was occasionally observed. D-G) Localisation of overexpressed Hyp2 orthologues was also examined in the context of a TbHyp2 null mutant background (the generation of which is discussed in section 4.8.5). Comparable localisation patterns were observed for TbHyp2 (in 2 clones, D and E), TcHyp2 (F) and TvHyp2 (G) overexpression in the null mutant background.

The staining pattern observed with paraformaldehyde fixed cells was to some extent consistent with that observed with methanol fixation (for examples see Appendix 2.5). However, methanol fixation in some cases produced a greater number of dots. As mentioned previously (section 4.5) TbHyp2 is predicted to interact with MKT1. In procyclic forms stressed by starvation, the normally cytosolic MKT1 was shown to localize to granules (Singh et al., 2014), this localisation was comparable to that observed for Hyp2 overexpression in bloodstream forms. These comparable localisation patterns were compatible with a potential Hyp2 interaction with MKT1. MKT1 had been shown to partially associate with the starvation-stress granule marker SCD6, therefore Hyp2 localisation was tested with regard to SCD6. However, overexpressed Hyp2 did not colocalise with the starvation-stress granule marker SCD6 (Appendix 2.5) (SCD6 antibody kind gift of Mark Carrington, (Kramer et al., 2008)).
4.8.3 Overexpression associated with increased kinetoplast-posterior distance

An increased kinetoplast-posterior distance was observed in some cells overexpressing TbHyp2 and orthologues. Therefore, *in vitro* methanol fixed cells were collected 24 hours after inducing Hyp2 overexpression with doxycycline, and images were captured. Measurements of kinetoplast-posterior distance were carried out in Image J64 (Rasband, 1997-2015). A graticule image captured at the same magnification as the acquired images was used to ‘set scale’ in Image J before commencing measurements, providing quantitative measurements of the kinetoplast-posterior dimension.

<table>
<thead>
<tr>
<th>Dox concentration</th>
<th>Kinetoplast posterior distance (µm)</th>
<th>Log Kinetoplast posterior distance (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dox 1µg/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dox 10^-4 µg/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dox -0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A

![Graphs showing TbHyp2 overexpression and its transform](image-url)
B

**TcHyp2 overexpression**

![Graph showing TcHyp2 overexpression with various conditions.]

**TcHyp2 overexpression in TbHyp2 dKO**

![Graph showing TcHyp2 overexpression in TbHyp2 dKO with various conditions.]

**Transform of TcHyp2 overexpression**

![Graph showing the transformed TcHyp2 overexpression with various conditions.]

**Transform of TcHyp2 overexpression in TbHyp2 dKO**

![Graph showing the transformed TcHyp2 overexpression in TbHyp2 dKO with various conditions.]

- **Kinetoplast posterior distance (µm)**
- **Log Kinetoplast posterior distance (µm)**

**Conditions:**
- Dox 1µg/ml
- Dox 10⁻⁴ µg/ml
- Dox -

**Transform of TcHyp2 overexpression:**

- Adj. p = 0.0002
- Adj. p = 0.0596

**Note:**
- Adj. p < 0.0001
Figure 4.9 Kinetoplast-posterior distance measurements following 24h induction of Hyp2 overexpression. Measurements of kinetoplast-posterior distance were carried out in Image J64 following overexpression of TbHyp2 (A), TcHyp2 (B) or TvHyp2 (C). Overexpression was induced in either an AnTat 1.1 90.13 background, or in the context of a TbHyp2 null mutant (the generation of which is discussed in section 4.8.5). Either a high (1μg/ml) or a low (10^{-4}μg/ml) concentration of doxycycline was used for induction. Graphical representations of measurements showing mean and standard deviation are accompanied by log transformed graphs indicating significance of difference in kinetoplast-posterior distance. One-way ANOVA followed by Tukey’s multiple comparisons test was carried out in GraphPad Prism. Samples for measurement were collected from a single well for each condition. It was assumed that any variation within the cell line was minimal in comparison to the effect of induction.

The most pronounced increases in mean kinetoplast-posterior distance were observed for TbHyp2 or TvHyp2 overexpression with 1μg/ml doxycycline, or TcHyp2 overexpression alone or in TbHyp2 dKO with 1μg/ml or 10^{-4}μg/ml doxycycline. For statistical analysis, as standard deviations were not equal, data was first log transformed. Comparisons were made between different doxycycline concentrations using a one-way ANOVA with multiple comparisons. In all cases, induction of Hyp2
overexpression by high (1µg/ml) or low (10^-4µg/ml) concentrations of doxycycline significantly increased the mean kinetoplast-posterior distance relative to in the uninduced cells.

The observed elongated kinetoplast-posterior distance following Hyp2 overexpression potentially reflected microtubule extension at the posterior end of the cell. Therefore microtubule extension at the posterior end of the cell was investigated using the antibody YL1/2. The YL1/2 antibody (Kilmartin et al., 1982) detects a post-translational tyrosination of α-tubulin. In trypanosomes newly laid down microtubules are tyrosinated and subsequently detyrosinated. Thus, YL1/2 antibody can be used to detect newly laid down microtubules in trypanosomes. This antibody detects posterior fluorescence, bright basal bodies and a bright daughter flagellum in procyclic forms preparing for cell division (Matthews and Gull, 1994).

<table>
<thead>
<tr>
<th>Cell type</th>
<th>TbHyp2 overexpression</th>
<th>TbHyp2 overexpression in TbHyp2 dKO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dox</td>
<td>1µg/ml</td>
<td>10^-4µg/ml</td>
</tr>
<tr>
<td>Day</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

A
Figure 4.10 YL1/2 antibody staining of newly laid down α-tubulin in Hyp2 overexpressing cells. Staining with YL1/2 (red) is mostly focused at the posterior end of the cell following TbHyp2 (A), TcHyp2 (B) or TvHyp2 (C) overexpression. Overexpression was induced in either an AnTat 1.1 90.13 background, or in the context of a TbHyp2 null mutant (the generation of which is discussed in section 4.8.5). Either a high (1µg/ml) or a low (10^{-4}µg/ml) concentration of doxycycline was used for induction. In all images the scale bar represents 10µm.
Often an elongated kinetoplast-posterior distance following Hyp2 overexpression in bloodstream forms was associated with bright posterior YL1/2 staining (Figure 4.10) comparable to that seen in procyclic forms. In uninduced controls, occasionally basal body and newly formed flagella were stained in proliferating cells, but bright posterior end staining was mostly not observed. This indicated as expected that Hyp2 overexpression induced a posterior extension in the cells mediated by microtubule polymerisation.

**4.8.4 Effect of Overexpression on responsiveness to 8-CPT-cAMP**

Knock down of TbHyp2 with RNAi causes resistance to 8-CPT-cAMP *in vitro* and to SIF *in vivo* (Mony et al., 2014). Therefore, initial tests for functional complementation with the identified Hyp2 orthologues involved expressing each orthologue in the context of TbHyp2 RNAi to test whether these could restore sensitivity to 8-CPT-cAMP. Consistent with Table 4.4, the RNAi template targeting TbHyp2 did not appear to silence TcHyp2 expression, with inducible overexpression of the protein detectable from days 1-3 following induction (Figure 4.11A). However, the growth inhibition associated with overexpression was greater than the effect of 8-CPT-cAMP so it was not possible to determine whether there was a rescue (Figure 4.11B).
The severe growth effect associated with overexpression of TcHyp2 (induced with 1 µg/ml doxycycline) prevents analysis of sensitivity to 8-CPT-cAMP. A) TbHyp2 RNAi does not prevent expression of TY-tagged TcHyp2, as detected by Western Blot. B) The graph on the left demonstrates resistance to 8-CPT-cAMP when TbHyp2 RNAi is induced. The graph on the right demonstrates that the growth effect associated with Hyp2 overexpression is more severe than the effect of 8-CPT-cAMP. Each data point represents a single well.

In order to establish a concentration at which TcHyp2 overexpression did not have such an effect on growth, a titration of doxycycline was carried out. It was observed that a concentration of 2x10^{-5} µg/ml had only a minimal effect on growth and TY-tagged TcHyp2 was detectable at this concentration from days 1-3 (see Appendix 2.8). Unfortunately when 2x10^{-5} µg/ml doxycycline was tested on the TbHyp2 RNAi line, it was shown that this was not sufficient to induce resistance to 8-CPT-cAMP (not shown). Therefore doxycycline was again titrated, and it was found that 8x10^{-5} µg/ml doxycycline was the lowest concentration able to induce detectable resistance to 8-CPT-cAMP. If TcHyp2 was able to rescue TbHyp2 function then under inducing conditions this resistance to 8-CPT-cAMP should be lost. However, when 8x10^{-5} µg/ml doxycycline was used to induce TcHyp2 overexpression alongside TbHyp2 RNAi, the growth inhibitory effect was additive, with the induced cells growing worse in 8-CPT-cAMP than the uninduced cells (Figure 4.12). This could be consistent with rescue, however at this induction concentration, the growth inhibition caused by TcHyp2 overexpression was still significant and it was possible that slight variations in TcHyp2 expression levels in the +/-cAMP conditions could be responsible for the differences in growth observed. Hence, the main difficulty
encountered in these initial rescue experiments was the inability to individually control the knock-down of the \textit{T. b. brucei} transcript and overexpression of the \textit{T. congoense} protein; both were inducible with doxycycline. Therefore, a TbHyp2 knock out (dKO) cell line was generated.

Figure 4.12 The growth inhibitory effects of TcHyp2 overexpression and 8-CPT-cAMP were additive when $8 \times 10^5 \mu g/ml$ doxycycline was used for induction. The graph on the left demonstrates resistance to 8-CPT-cAMP when TbHyp2 RNAi is induced. The graph on the right demonstrates that Hyp2 overexpression and 8-CPT-cAMP treatment had additive inhibitory effects on growth. Each data point represents a single well.

4.8.5 Knock out of TbHyp2

The strategy to create a TbHyp2 null mutant involved the use of pEnT6B-Y and pEnT6P-Y vectors designed for endogenous tagging (Kelly et al., 2007). Primers were designed to amplify regions of the 5’UTR and 3’UTR of TbHyp2 in order to replace the endogenous gene with a cassette containing either a puromycin or blasticidin resistance marker (Appendix 2.4). Primers were designed such that the first transfection replaces a larger region external to the target gene, and the second transfection targets a region internal to this, so that the second resistance cassette replaces the remaining endogenous allele rather than the first integrated construct (Appendix 2.9).
Figure 4.13 Generation of TbHyp2 dKO in T. b. brucei AnTat 1.1 90.13 background. Generation of TbHyp2 dKO was confirmed by the restriction digest with PstI followed by Southern blotting. Digest with PstI results in a smeary pattern of genomic DNA, but by using probes targeted to either the 5’UTR or coding region of the targeted gene, a characteristic digest pattern is obtained (A). The digest pattern observed was consistent with expectation confirming generation of Hyp2 null mutants (B). Predicted band sizes obtained with the 5’UTR probe are represented by an orange line, those expected with the gene probe are represented by a blue line. The total region detectable by the gene probe is represented by a thick green line. The total region detectable by the 5’UTR probe is represented by a thick purple line.
The plasmids were linearised with SpeI prior to transfection. Although the internal primer design means that construct integration would also affect Tb927.9.4060, Tb927.9.4070 and Tb927.9.4090, these genes were annotated as ‘unlikely’ so should not contribute to off-target effects. Once the integration constructs were generated, the transfections themselves, for both allele replacements, were carried out by Lindsay McDonald. I then validated the two resultant clones as null mutants by restriction digest with PsiI followed by analysis of the resultant digest pattern by Southern Blot (Figure 4.13). As expected the Hyp2 gene probe detected two bands of 2161bp and 2903bp in the wild type and single knock out (sKO) lanes but not in the double knock out (dKO) lanes. Additionally, the 5’UTR probe detected the expected wild type band sizes (763bp and 2161bp) in the parental AnTat 1.1 90.13 line, these bands were also detectable in the sKO with an additional band of expected size for the first integration (472bp). The 763bp and 472bp bands expected for the sKO were also detectable in the dKO lanes alongside a predicted 1001bp band, indicating correct integration of the second replacement construct. Therefore generation of the Hyp2 dKO was considered to be successful. Hyp2 dKO clone C was then selected for transfection using overexpression constructs for TbHyp2, TcHyp2 and TvHyp2.

Once again, the doxycycline concentration was titrated to determine the level that did not result in an inhibition of growth, in order to test the ability of each Hyp2 orthologue to restore sensitivity to 8-CPT-cAMP. Although $1 \times 10^3 \mu g/ml$ doxycycline caused full growth inhibition, $1 \times 10^4 \mu g/ml$ doxycycline caused a reduced, but intermediate effect on growth, and $1 \times 10^5 \mu g/ml$ doxycycline was shown to have no growth effect on any of the clones tested (see Appendix 2.10). Therefore, $1 \times 10^5 \mu g/ml$ doxycycline was used to induce TbHyp2 overexpression in a TbHyp2 dKO background and the responsiveness of the cells to 8-CPT-cAMP was examined and compared with the dKO line. As expected the TbHyp2 dKO alone demonstrated resistance to 8-CPT-cAMP relative to parental AnTat 1.1 90.13 (Figure 4.14).
Figure 4.14 TbHyp2 dKO reduces sensitivity to 8-CPT-cAMP. A) The inhibition of growth caused by 8-CPT-cAMP was reduced in the TbHyp2 null mutant line relative to the parental AnTat 1.1 90.13 cell line. Each data point represents the mean and SEM for triplicate wells. B) Percentage reduction in growth was calculated by dividing the average reduction in growth caused by adding 8-CPT-cAMP to each cell line by the average growth -8-CPT-cAMP and multiplying by 100.

Overexpression of TbHyp2 resulted in an increased responsiveness to 8-CPT-cAMP in the dKO line. However, once again interpretation was complicated by a slight growth effect associated with overexpression itself. In an attempt to bypass this complication, the percentage reduction in growth caused by 8-CPT-cAMP for both the induced and uninduced conditions was compared to each of their respective untreated growth levels. However, this demonstrated no significant change in growth reduction caused by 8-CPT-cAMP when TbHyp2 overexpression was induced or not (Figure 4.15). TbHyp2 overexpression would be expected to restore functionality, so it was determined that in vivo experiments to analyse sensitivity to SIF would be required in order to get a clearer output for functional rescue.
Figure 4.15 Growth effect of 8-CPT-cAMP on the TbHyp2 dKO line with or without induction of TbHyp2 overexpression. A) The inhibitory effect of 8-CPT-cAMP on growth is increased by concomitant TbHyp2 overexpression in a TbHyp2 null mutant background. Each data point represents the mean and SEM for triplicate wells. B) Percentage reduction in growth with or without Hyp2 overexpression was calculated by dividing the average reduction in growth caused by adding 8-CPT-cAMP by the average growth -8-CPT-cAMP and multiplying by 100).

4.9 Characterisation of Hyp2 overexpression in vivo

The capacity for complementation of TbHyp2 by *T. congolense* or *T. vivax* orthologues was assessed in vivo by analysing the capacity of various overexpression and knock-out lines to differentiate to stumpy forms.

4.9.1 TbHyp2 dKO does not become stumpy or control parasitaemia

Before attempting rescue with overexpression of Hyp2 orthologues, the characteristics of Hyp2 dKO parasitaemia were examined.
Figure 4.16 Lack of growth control with TbHyp2 dKO. Parasitaemia in TbHyp2 dKO infections reached high levels (A) with a considerable proportion of proliferating cells (columns represent the mean and standard deviation) (B), and without appearance of cells with stumpy morphology (C). Scale bar represents 10 µm. For slender and stumpy n=2, for TbHyp2 dKO n=3.

TbHyp2 dKO infections reached high parasitaemia (>2x10^8/ml) by day 4 post infection without displaying evidence of growth control; the proportion of proliferating cells remained above 25%, and the morphology did not appear stumpy (Figure 4.16). Parasites were harvested on the final day of each infection (day 3 for AnTat 1.1 90.13 slender, day 6 for AnTat 1.1 90.13 stumpy and day 4 for TbHyp2 dKO). Protein samples were prepared to compare expression of the stumpy marker PAD1 in the Hyp2 dKO line relative to AnTat 1.1 90.13 cells. Consistent with their slender morphology, the Hyp2 dKO cells did not express PAD1 despite high parasitaemia (Figure 4.17A). Additionally, once harvested and purified parasites were incubated in SDM-79 containing cis-aconitate at 27°C to stimulate differentiation to procyclic forms. The TbHyp2 dKO line demonstrated less capacity to differentiate to procyclic forms compared to stumpy cells, as assessed by flow cytometry analysis of EP procyclin expression (Figure 4.17B). TbHyp2 dKO cells
were also less proliferative following 24h in SDM-79 and cis-acconitate (Figure 4.17C), indicating a failure to grow out as a differentiated procyclic form population.

Figure 4.17 Effect of TbHyp2 dKO on the capacity to differentiate. TbHyp2 dKO cells do not express PAD1 despite high parasitaemia as visualised by western blot (A). EF1 was used to demonstrate that loading was consistent between TbHyp2 dKO samples and slender and stumpy control samples. TbHyp2 dKO cells have reduced capacity to differentiate to procyclic forms following induction with cis-acconitate, as determined by EP procyclin expression at 4h and 24h after induction (B). Additionally TbHyp2 dKO cells have reduced proliferation after 24h of induction with cis-acconitate, compared to both slender and stumpy control parasites (C).

4.9.2 Is Hyp2 orthologue overexpression able to rescue stumpy formation in vivo?

Before proceeding with larger scale experiments to test the effect of Hyp2 orthologue expression in the Hyp2 dKO line, a trial experiment was carried out with a single infection per test condition. All infections were initiated at the same dose (~6x10⁵ cells). The absence of endogenous TbHyp2 was confirmed by Northern Blot in all samples except TcHyp2 OE in KO +Dox, of which there were not enough cells for
an RNA sample to be prepared (Figure 4.18H). TbHyp2 overexpression was included as a positive control as this was expected to compensate for the absence of Hyp2 in the TbHyp2 dKO line. Despite this, TbHyp2 overexpression was not effective at limiting parasitaemia, and the proportion of proliferating cells was comparable between TbHyp2 expressing (induced) and non-expressing (uninduced) cells (Figure 4.18A). However, there was higher PAD1 expression observed with ‘TbHyp2 induced’ than in the ‘uninduced’ control (Figure 4.18B), and the TY-tagged TbHyp2 was not detectable in the endpoint sample (Figure 4.18C). TcHyp2 and TvHyp2 overexpression both limited parasitaemia (Figure 4.18A), with a few stumpy cells observable on days 3 and 4, increasing in number on day 5 (and 6 in the case of TvHyp2). Indeed, the TcHyp2 overexpression effect on the parasitaemia was so severe that there were not enough cells for protein samples to be prepared on the final day of the experiment. Nonetheless, there was a clear reduction in the proportion of proliferating cells by day 5, with TcHyp2 overexpression induced (Figure 4.18F). Moreover, a few PAD1 positive cells were visible by IFA on day 5 following TcHyp2 induction. These results were suggestive of a capacity for TcHyp2 overexpression to rescue stumpy formation. The most convincing evidence for rescue of stumpy formation was obtained for TvHyp2. TY-tagged TvHyp2 was detectable on the final day of infection (Figure 4.18C) and was associated with increased expression of PAD1 relative to uninduced cells (Figure 4.18B). PAD1 expressing cells with stumpy morphology were also visualised by IFA on day 6 post-infection (Figure 4.18I). This was coincident with a reduced percentage of proliferating cells on day 6 following induction (Figure 4.18G).
TvHyp2 OE in TbHyp2 dKO +DOX

Parasitaemia % 2K1N, 2K2N % PAD +ve

TvHyp2 OE in TbHyp2 dKO -DOX

Parasitaemia % 2K1N, 2K2N % PAD +ve

TvHyp2 OE in TbHyp2 dKO clone 1 -DOX

Parasitaemia % 2K1N, 2K2N % PAD +ve

TvHyp2 OE in TbHyp2 dKO clone 1 +DOX

Parasitaemia % 2K1N, 2K2N % PAD +ve

* Note: 173 cells counted for % 2K1N, 2K2N and 150 counted for % PAD +ve for this sample. For all other samples 250 cells counted.
**Figure 4.18 Trial of Tb/Tc/TvHyp2 overexpression rescue of TbHyp2 dKO in vivo.** A) Effect of Tb/Tc/TvHyp2 overexpression on parasitaemia in the context of TbHyp2 dKO. Each data point represents the parasitaemia from a single infection. B) PAD1 expression with or without Hyp2 overexpression. EF1 was used as a loading control. C) TY-tagged TvHyp2 was faintly detectable with BB2 antibody. EF1 was used as a loading control. D-G) Percentage 2K1N, 2K2N cells (blue) and PAD1 positive cells (green) displayed alongside parasitaemia for each cell line tested. Each data point represents a measurement from a single infection. H) Northern blot confirming absence of endogenous TbHyp2 in the dKO cell lines tested. I) Image of TvHyp2 overexpression in KO +Dox (day 6). PAD1 positive cells are stained green. Scale bar represents 10µm.

**4.9.3 Effect of TbHyp2 overexpression in TbHyp2 dKO**

A lack of growth control was observed when TbHyp2 overexpression was induced in the dKO clone 1 (Figure 4.18). This clone was observed to undergo less growth limitation *in vitro* than other TbHyp2 overexpressing cell lines (not shown). Therefore, other TbHyp2 OE clones were tested and a new clone, clone 2, that showed stronger growth inhibition upon induction *in vitro* was selected for testing *in vivo* (Figure 4.19).
Figure 4.19 Analysis of TbHyp2 overexpression in a second TbHyp2 dKO clone. A) TbHyp2 overexpression induced by a high doxycycline concentration causes severe growth inhibition in vitro, the growth inhibition observed was less when a lower concentration of doxycycline was used. Each data point represents the mean and SEM for triplicate wells. B) Detection of TY-tagged TbHyp2 in vitro. TbHyp2 overexpression was induced by either a high (1 µg/ml) or low (10^{-4} µg/ml) concentration of doxycycline. C) TbHyp2 overexpression in clone 2 caused significantly reduced %2KN, 2K2N (★★★ p<0.001). An unpaired t test was performed using GraphPad Prism. The columns display mean and standard deviation for each condition based on three technical replicates. More than 250 cells were counted for each replicate. It was assumed that any variation within the cell line was minimal in comparison to the effect of induction.

A trial in vivo experiment with the new TbHyp2 OE in KO clone demonstrated a clear reduction in parasitaemia when TbHyp2 overexpression was induced, with stumpy cells visible on several days of the infection. There was also a slight reduction in proliferating cells in the induced as compared to the uninduced cells (Figure 4.20).
Figure 4.20 *TbHyp2 overexpression in TbHyp2 dKO trial*. A) Overexpression of *TbHyp2* (+Dox) reduced parasitaemia relative to an uninduced infection (-Dox). B) Proportion of proliferating (2K1N, 2K2N) cells on each day of infection (blue columns). Each data point represents a measurement from a single infection.

This trial experiment was then followed up in triplicate. Once again, a reduced parasitaemia was observed for the cells overexpressing *TbHyp2* relative to the uninduced cells (Figure 4.21A). This was accompanied by a reduction in proliferating cells on day 3 in the induced infections, but the level of 2K1N, 2K2N was comparable between the induced and uninduced infections on day 4 (Figure 4.21B).

Detection of *TbHyp2* overexpression only in the induced cells (Figure 4.22A), alongside confirmation of the absence of endogenous *TbHyp2* in all knock out lines (Figure 4.22B), confirmed that the cells were behaving as expected in the infections. Fluorescence analysis detected the presence of TY-tagged *TbHyp2* in some but not all induced cells. Additionally, there was detection of some PAD1 positive stumpy cells in the induced infections, not observed in the uninduced (Figure 4.21C).
Figure 4.21 TbHyp2 overexpression in the TbHyp2 dKO clone 2 in vivo. A) Reduced parasitaemia was observed in induced relative to uninduced infections. B) A reduced proportion of proliferating cells (2K1N, 2K2N) was observed on day 3 in induced relative to uninduced infections (★ p<0.05), but not on day 4 (p=ns). An unpaired t test was performed using GraphPad Prism. Columns display the mean and standard deviation for each condition based on three biological replicates. More than 250 cells were counted for each replicate. C) Some stumpy cells were observed following induction of Hyp2 overexpression. TY-tagged TbHyp2 was detected with the BB2 antibody (red), occasionally coincident with PAD1 expression (green).

On the last day of each infection, parasites were collected and induced to differentiate to procyclic forms in SDM-79 containing cis-aconitate at 27°C. The TbHyp2 overexpressing cells did not display an improved capacity to differentiate to procyclic forms relative to the uninduced cells, and in both cases the differentiation level (assessed by EP procyclin expression) was more comparable to a slender population than a stumpy population (Figure 4.22C). This was in agreement with the low level of PAD1 expression in both induced and uninduced samples (Figure 4.22D).
Figure 4.22 TbHyp2 overexpression in TbHyp2 dKO clone 2 in vivo. A) TY-tagged TbHyp2 protein was detected using BB2 antibody in +Dox samples only. B) A Northern blot confirming absence of endogenous Hyp2 in all dKO lines; exogenous TbHyp2 was detected in the one +Dox replicate with enough cells for an RNA sample to be prepared on the final day of infection. C) A comparable efficiency of differentiation to procyclic forms in induced and uninduced cell lines was observed (p=ns). An unpaired t test was performed using GraphPad Prism. Columns display the mean and standard deviation for each condition based on three biological replicates. D) PAD1 expression with (+Dox) or without (-Dox) induction of TbHyp2 overexpression on the final day of each infection.

In this experiment, the uninduced and control TbHyp2 dKO cells behaved as expected for a non-differentiated, mostly slender population. In contrast, overexpression of exogenous TbHyp2 had been expected to restore differentiation to stumpy forms at high parasitaemia. However it was observed that this population also demonstrated low PAD1 expression and inefficient differentiation capacity (Figure 4.22). Nonetheless, stumpy morphology cells were observed on several days in the ‘induced’ infections, and the parasitaemia in these infections was considerably lower on days 3 and 4 than for wild-type pleomorphic cells, despite all infections having been initiated by a comparable infective dose. The exogenous overexpression appears to be inducing early differentiation in some cells despite low parasitaemia, consistent with the expected role of TbHyp2 in signalling differentiation to stumpy forms. Although strong growth control was not seen, the detection of stumpy cells in
the induced population indicates rescue. Selection against growth inhibition may then have allowed continued growth of non-expressors. Overall, evidence for the ability of TbHyp2 to rescue the Hyp2 KO phenotype was observed, however.

4.9.4 Effect of TvHyp2 overexpression in TbHyp2 dKO

In the trial experiment (section 4.9.2), TvHyp2 overexpression had shown more ability to rescue stumpy formation than the other orthologues tested. This preliminary result was then followed up in triplicate to further investigate rescue of the TbHyp2 dKO with TvHyp2 overexpression.

In infections where TvHyp2 overexpression was induced, the parasitaemia was limited relative to both the uninduced and AnTat 1.1 90.13 parental control infections (Figure 4.23A). The parasitaemia in these induced infections reached equivalent levels to AnTat 1.1 90.13 cells by day 6 in all except one replicate, which plateaued at a lower parasitaemia between days 5 and 6. In the uninduced infections, growth was unlimited, and infections had to be terminated on day 4. In the induced infections, stumpy cells were detectable from day 3, whereas uninduced parasites reached high parasitaemia without showing stumpy morphology (Figure 4.23D, and not shown). A reduction in the proportion of 2K1N, 2K2N cells was also observable in the induced populations on days 5 and 6, consistent with some restoration of G1/G0 growth arrest (Figure 4.23B). However on day 4, after reaching high parasitaemia without displaying stumpy morphology, the uninduced populations also appeared to be arresting in G1/G0. This data, based on KN counts, was supported by flow cytometry analysis of DNA content by DAPI staining (Figure 4.23C).

Furthermore, when PAD1 expression was analysed by western blot, the PAD1 expression levels were found to be higher in the uninduced samples than in the induced samples (Figure 4.24A). This finding is inconsistent with the previous observation that TbHyp2 dKO cells did not express PAD1 despite reaching high parasitaemia. However, when TY-tagged TvHyp2 protein expression was assessed by western blot, it was clear that expression had become leaky by the final stages of infection, and in fact TvHyp2 expression was detectable in the uninduced but not the induced samples (Figure 4.24B). This would indicate that it is the leaky TvHyp2 expression in the uninduced population that is responsible for the PAD1 expression
not usually detectable in a TbHyp2 dKO cells. PAD1 protein levels for all test samples were considerably lower than those of a stumpy control (Figure 4.24A).

**Figure 4.23 TvHyp2 overexpression in TbHyp2 dKO line in vivo.** A) Parasite population growth was limited when TvHyp2 overexpression was induced (+Dox) relative to uninduced controls. B) The proportion of proliferating cells (2K1N, 2K2N) was assessed by manual cell counting as an indicator of the degree of growth arrest (*** p<0.005). An unpaired t test was performed using GraphPad Prism. Columns display the mean and standard deviation for each condition based on three biological replicates. More than 250 cells were counted for each replicate. C) Manual counts were supported by flow cytometry analysis of DAPI stained cells. D) Occasional stumpy morphology cells with PAD1 staining were observed on days 4, 5 and 6 of the induced infections, but uninduced populations were predominantly slender in morphology. Also, BB2 antibody detected TY-tagged TvHyp2 in some cells of the induced population on days 4 and 5.

On the last day of each infection parasites were collected, and induced to differentiate to procyclic forms in SDM-79 containing cis-aconitate at 27°C. The controls in this experiment behaved as expected, with a stumpy population showing a better capacity to differentiate to procyclic forms (assessed by EP procyclin expression) compared to slender (*in vitro*) bloodstream forms. Both the TvHyp2 uninduced and induced populations showed comparable ability to differentiate to procyclic forms; both were able to differentiate better than slender controls, but not as efficiently as a stumpy population (Figure 4.24C). This result, alongside the PAD1 data, is consistent with some degree of restoration of stumpy formation when
TvHyp2 was overexpressed. While TY-tagged TvHyp2 was not detected by western on day 6 in the induced population, it was detectable by BB2 antibody staining of individual cells from day 4 and 5 of the infections (Figure 4.23D). Knock out of endogenous TbHyp2 in all 3 uninduced and 2 of the induced samples was confirmed by Northern blot (Figure 4.24D).

Figure 4.24 TvHyp2 overexpression in TbHyp2 dKO in vivo. A) Detection of PAD1 expression with (+Dox) or without (-Dox) the induction of TvHyp2 overexpression. B) TY-tagged TvHyp2 protein was detected using BB2 antibody in -Dox samples more strongly than in +Dox samples indicating leaky expression. EF1 was used as a loading control. C) Induced and uninduced TvHyp2 overexpression cell lines showed a comparable efficiency of differentiation to procyclic forms (p=ns). An unpaired t test was performed using GraphPad Prism. Columns display mean and standard deviation for each condition based on three biological replicates. D) The absence of endogenous Hyp2 in all dKO lines was confirmed by Northern blot.

Despite, the leakiness of TvHyp2 expression in the uninduced control, these data support the ability of TvHyp2 to restore stumpy formation to some extent. In both the induced and uninduced infections, TY-tagged TvHyp2 overexpression was detectable at some point, and this was concurrent with increased PAD1 expression and improved capacity to differentiate to procyclic forms compared to the parental TbHyp2 dKO cell line.
4.9.5 Effect of TcHyp2 overexpression in TbHyp2 dKO

In the trial experiment, the effect of TcHyp2 overexpression was so severe that there were few cells remaining for analysis when the experiment was ended on day 6. Therefore, when the experiment was repeated in triplicate, the induced infections were followed up to day 5 only.

Figure 4.25 TcHyp2 overexpression in TbHyp2 dKO line in vivo. A) Parasite population growth was limited when TcHyp2 overexpression was induced (+Dox) relative to uninduced controls. B) The proportion of proliferating cells (2K1N, 2K2N) was assessed by manual cell counting as an indicator of the degree of growth arrest (★★ p<0.05). An unpaired t test was performed using GraphPad Prism. Columns display mean and standard deviation for each condition based on three biological replicates. C) Staining with BB2 antibody detected TY-tagged TcHyp2 (red) in many cells of the induced population on days 3-5. PAD1 positive (green), stumpy morphology cells were also detectable in the induced population.

As observed previously, the uninduced populations and TbHyp2 dKO parental control reached high parasitaemias by day 4 without the appearance of stumpy morphology parasites. These infections had to be terminated on day 4. In contrast, while the parasitaemias in the induced infections did increase over time, the levels
remained lower than the AnTat 1.1 90.13 control (Figure 4.25A). There was also a reduction in proliferating cells on days 3 and 4 in the induced relative to uninduced populations, though this difference was only significant on day 3. By day 5 of the infections, the proportion of proliferating cells was greatly reduced in the AnTat 1.1 90.13 control, but remained at a higher level in the induced TcHyp2 overexpression line (Figure 4.25B). Nevertheless, PAD1 expression was clearly detectable by immunofluorescence on stumpy cells observed following TcHyp2 overexpression, particularly on days 4 and 5. TY-tagged TcHyp2 was also detectable in a number of cells in the induced population from days 3 to 5 of infection (Figure 4.25C). Consistent with the immunofluorescence analysis, western blotting confirmed that TcHyp2 expression was only detectable in the induced infections, and this was associated with increased PAD1 expression relative to uninduced samples (Figure 4.26A). PAD1 levels were less than in the stumpy control population, but were consistent with a partial rescue of stumpy formation. Northern blotting confirmed the absence of endogenous TbHyp2 in all uninduced infections, as well as the one induced population where enough parasites remained to extract RNA (Figure 4.26B).

At the end of each infection, parasites were induced to differentiate to procyclic forms in SDM-79 containing cis-aconitate at 27°C. The populations in which TcHyp2 overexpression was induced showed significantly better capacity to differentiate to procyclic forms than uninduced or slender controls, as judged by EP procyclin expression after 4 hours induction (Figure 4.26C). Differentiation in the induced populations was not as efficient or synchronous as the stumpy control. This is consistent with the differentiation of an ‘intermediate’ population, and a partial rescue of stumpy formation in the TbHyp2 dKO.
Figure 4.26 TcHyp2 overexpression in TbHyp2 dKO line in vivo. A) PAD1 expression was higher with (+Dox) induction of TcHyp2 overexpression than in uninduced samples. TY-tagged TcHyp2 protein was detected using BB2 antibody in +Dox samples but not in –Dox samples. B) Northern blot confirming absence of endogenous Hyp2 in all dKO lines. C) Efficiency of differentiation to procyclic forms was improved in induced cells relative to uninduced cells (★★★★ p<0.0001, ★★ p<0.01). An unpaired t test was performed using GraphPad Prism. Columns display the mean and standard deviation for each condition based on three biological replicates.

4.10 Conclusions and future directions

Potential *T. congolense* and *T. vivax* orthologues were identified for multiple components of the *T. b. brucei* SIF-responsive pathway. One SIF pathway component, TbHyp2, and its orthologues were chosen to investigate whether sequence conservation would follow through to functional conservation. A bioinformatics analysis indicated conservation of some but not all predicted domains and short linear motifs. Amongst those conserved are short motifs indicative of involvement in cell cycle control, post-transcriptional regulation and interaction with other SIF-responsive pathway components. The predicted interactions of the Hyp2 orthologues were also consistent with a role in a pathway to quiescence. This bioinformatics identification of potential interactions could be further informed by carrying out pull-down experiments to identify interaction partners of the individual Hyp2 orthologues.
Activity of the individual orthologues was assayed using overexpression lines, with the Hyp2 orthologue expressed either in a pleomorphic *T. b. brucei* background or in a cell line where endogenous TbHyp2 had been knocked out. All overexpressed proteins had a small TY tag at the N-terminal. Localization patterns were found to be consistent between the orthologues, often consisting of a punctate pattern of two or more dots particularly near to the nucleus or kinetoplast. Further work is required to identify these compartments, through colocalisation with known cell markers. An extended kinetoplast-posterior distance was occasionally observed following overexpression of TbHyp2 and its orthologues. This was found to coincide with newly laid down α-tubulin. Such cytoskeletal remodelling is usually observed in procyclic forms as they transit through their cell cycle (Matthews and Gull, 1994) and in cells differentiating to procyclic forms (Matthews and Gull, 1997). Additionally, all orthologues caused growth inhibition when overexpressed, and in all cases this was concurrent with an accumulation in 1K1N. This G1/G0 arrest is consistent with the expected activity of proteins that drive differentiation to the quiescent stumpy form.

Elimination of endogenous TbHyp2 expression *in vivo* was found to prevent parasite growth control through differentiation to the stumpy form. Overexpression of TbHyp2 in the context of the TbHyp2 dKO was expected to restore differentiation kinetics. However, lack of control over the level of overexpression made it impossible to observe a coordinated population-wide response, as is observed in naturally pleomorphic parasites. Nevertheless, occasional stumpy parasites were observed throughout infection following induction of TbHyp2 overexpression. SIF is predicted to be produced only by slender forms (Seed and Black, 1997). Therefore, the constant differentiation of just a small portion of the induced population to stumpy forms, could have prevented or delayed the accumulation of SIF to the levels necessary to induce differentiation, thus preventing the coordinated population-wide differentiation observed in wild-type infections. Additionally, regulatory mechanisms may be in place to restrict prolonged TbHyp2 expression, such that low levels were detectable by the experiment end point. This may have been particularly effective in the case of TvHyp2 overexpression. Despite a reduction in parasitaemia early on in the overexpressing population, by the endpoint of infection the protein was
undetectable in induced samples. TvHyp2 appeared to be responsible for a partial restoration of stumpy formation, irrespective of the observed leaky expression in the uninduced control. Finally, the strongest evidence for effective rescue of stumpy differentiation was observed with TcHyp2 overexpression, supported by increased characteristics of a stumpy population.

One component required for cross-talk between single-celled parasite species is a mechanism to receive and translate another’s signal. Evidence supporting functional conservation of a SIF-responsive protein in other African trypanosome species favours a model where such parasites are able to perceive growth-controlling signals from T. b. brucei.
Chapter 5: How do *T. congoense* and *T. b. brucei* interact in co-infections? Is there potential for cross-talk?
5.1 Introduction

Potential interactions between cohabiting trypanosome species need not be restricted to perception of *T. b. brucei* generated SIF. Other trypanosome species such as *T. congolense* may also produce quorum-sensing molecules; such a system would be in line with the milieu of signals produced in a mixed population of bacteria.

The description of SIF began with analysis of conditioned media from dense cultures of *T. b. brucei*. This medium was found to trigger expression of stumpy-like characteristics *in vitro* during culture on agarose plates. Filtration identified the component(s) responsible as smaller than 500Da (Vassella et al., 1997a), beyond this the identity of factor(s) termed SIF remain a mystery. In addition to conditioned culture media, plasma from *in vivo* infections at peak parasitaemia was found to result in growth inhibition in cultured parasites (Seed and Sechelski, 1989). However, modelling using chronic infection data indicated that SIF was only produced by slender forms (MacGregor et al., 2011). Additionally, monomorphic cell lines produced SIF, though they no longer responded to it (Reuner et al., 1997; Vassella et al., 1997a).

The identification of the regulatory elements responsible for the stumpy-restricted expression of PAD1 allowed construction of a cell line with a CAT (Chloramphenicol Acetyltransferase) reporter under the control of the PAD1 3’UTR to report on stumpy formation (MacGregor and Matthews, 2012). This cell line was shown to express the CAT reporter in stumpy parasites extracted from an infection, as well as in response to the SIF-mimic 8-CPT-cAMP used *in vitro*. Additionally, conditioned media collected from pleomorphic AnTat 1.1 90.13 parasites that had reached high density *in vitro* was demonstrated to increase CAT expression in the reporter line.

The established *T. b. brucei* reporter system provided a clear read out to determine whether conditioned medium from *T. congolense* could affect the stumpy formation of a *T. b. brucei* population. Combined with analysis of cell cycle arrest, this should assist in distinguishing a differentiation response from simply dying and sick cells, and thereby establish the potential for cross talk between the species.
Conditioned medium experiments provide a useful method to investigate cross-talk in vitro, and offer a simplified system separated from host effects. Nevertheless, it is important to consider interactions in vivo, as a closer representation of field populations. Thus, impacts of co-infection on the individual species involved have also been investigated.

5.2 Validation of *T. b. brucei* conditioned medium effect

Conditioned medium was collected from pleomorphic *T. b. brucei* (AnTat 1.1 90.13) that had been allowed to reach a high density (2x10^6 cells/ml) over 2 days. 8-CPT-cAMP was used as a positive control in this experiment. Both conditions were tested on pleomorphic (AnTat 1.1 90.13) and monomorphic (Lister 427) reporter lines (Figure 5.1), where CAT reporter expression was under the control of the PAD1 3’UTR. Importantly, control medium that had been incubated for the same length of time as the conditioned medium was included as a negative control.
Figure 5.1 The effect of *T. b. brucei* conditioned medium (CM) and 8-CPT-cAMP on pleomorphs (graphs on left) and monomorphs (graphs on right). A) The effect of AnTat CM on pleomorph (left) and monomorph (right) population growth. For each condition tested n=1. B) CAT concentration/cell following treatment of a pleomorphic (left) or monomorphic (right) reporter line with AnTat CM or 8-CPT-cAMP. C) The fold change in CAT concentration/cell relative to day 0 for the pleomorphic (left) or monomorphic (right) reporter lines. CAT concentration was estimated by ELISA, with test samples compared to CAT standard concentrations.
Treatment with *T. b. brucei* conditioned medium increased CAT reporter expression in the pleomorphic, but not the monomorphic, cell line, despite some effect of conditioned medium on monomorph growth (Figure 5.1A and B). CAT levels in the negative control treatment remained lower (Figure 5.1B and C). 8-CPT-cAMP increased CAT concentration in both reporter lines, though to a greater extent in monomorphs. While components necessary for responding to the SIF signal may be disrupted in monomorphs, they still retain capacity to respond to 8-CPT-cAMP. Hence, these results met expectation and the established experimental protocol could then be adapted to test *T. congolense* conditioned medium.

### 5.3 Optimising *T. congolense* conditioned medium test conditions

In order to determine the density of *T. congolense* cells necessary to obtain conditioned medium with an observable effect on *T. b. brucei* growth, medium from different lengths of culture (without passage) were collected. Conditioned media was collected from a *T. congolense* culture that had been passaged 2, 3 or 4 days previously. The densities of culture were; 2.7x10⁶/ml on day 2, 7.0x10⁶/ml on day 3, and 6.2x10⁶/ml on day 4. Additionally, *T. congolense* conditioned medium was diluted to varying degrees (100%, 75% or 50% conditioned medium) with HMI-9 (*T. b. brucei* culture medium). Responses were compared with negative control *T. congolense* culture medium (TcBSF3 medium incubated at 34°C alongside the conditioned medium). Conditions were tested on both pleomorph and monomorph cell lines for comparison (Figure 5.2).

Even 50% *T. congolense* conditioned medium (TcCM) from day 3 of culture was found to inhibit growth of pleomorphic *T. b. brucei*. However, this effect was more exaggerated with 75% TcCM from day 3, where a clear difference between the test and negative control was observed, without too severe a growth effect. There was no apparent difference between day 3 and day 4 conditioned medium. This was supportive of an effect distinct from simple nutrient depletion or toxic metabolite build up, as these would have been expected to accumulate to a greater extent between days 3 and 4.
Figure 5.2 Titration of *T. congolense* conditioned medium (T. con CM) effect on pleomorphs (AnTat 1.1 90.13) and monomorphs (WT 427). Conditioned medium was collected on days 2, 3 or 4 of culture, and was diluted to varying extents with HMI-9 (100%, 75% or 50% conditioned medium). Each data point represents the cell number from a singlet well. A) The effect of 100% *T. congolense* CM on pleomorphs (left) and monomorphs (right). B) The effect of 75% *T. congolense* CM on pleomorphs. C) The effect of 50% *T. congolense* CM on pleomorphs (left) and monomorphs (right).
In contrast to the observed effect on pleomorphs, 50% TcCM from day 3 was not effective at inhibiting monomorph growth. This could be due to a difference in ability to cope with nutrient stress conditions between pleomorphs or monomorphs. Alternatively, monomorphs may have lost responsiveness to a growth inhibitory signal present in TcCM. 100% TcCM from day 3 caused a comparably severe growth effect on both pleomorphs and monomorphs, indicating that testing conditioned medium without any nutrient replacement with HMI-9 may overshadow any effects independent of nutrient depletion. Therefore, in later conditioned medium experiments 75% TcCM from day 3 of culture and 25% HMI-9 was used. Attempts were made to ensure that such conditioned medium was always collected at a comparable density.

5.4 Pleomorphic and monomorphic cell lines differ in their response to *T. congolense* conditioned medium

The observation that a monomorphic cell line could be less responsive to *T. congolense* conditioned medium than a pleomorphic one merited further investigation considering that monomorphic cell lines have become resistant to *T. b. brucei* SIF. Therefore, 75% *T. congolense* conditioned medium collected on day 3 was tested on both AnTat 1.1 90.13 (pleomorphic) and 427 (monomorphic) cell lines. The parasite number following treatment was observed for triplicate wells (Figure 5.3A), and demonstrated that the inhibition of pleomorph growth following TcCM treatment was more pronounced than that on monomorph growth. This was coincident with a small accumulation in 1K1N, perhaps indicative of some cell cycle arrest on day 2 of the experiment (Figure 5.3B), although this was also observed for the monomorphs. On day 3 following treatment, no morphological difference was observed between 427 cells treated with negative control or conditioned medium. In contrast, AnTat 1.1 90.13 cells appeared somewhat ‘stumpier’ when treated with TcCM compared to the negative control, however the difference was not pronounced (Appendix 3.1). Indeed, even with treatment with 8-CPT-cAMP or *T. b. brucei* conditioned medium there was not a clear morphological change to stumpy-like forms in liquid culture by the pleomorphs.
Figure 5.3 The effect of *T. congolense* conditioned medium on *T. b. brucei* pleomorphs (AnTat 1.1 90.13) and monomorphs (427). A) The effect of treatment with conditioned medium on pleomorph and monomorph cell number. Each data point represents the mean and SEM for triplicate wells. Conditioned medium was collected on day 3 of *T. congolense* culture at a density of 7.6x10^6 cells/ml. B) The proportion of cells in 1K1N, 2K1N or 2K2N, following 2 or 3 days of treatment of pleomorphs (left) or monomorphs (right) with conditioned medium. KN proportions were calculated by counting 250 cells taken from a single well of each condition.

The differential effect of conditioned medium on pleomorphs and monomorphs was reproducibly observed. For instance, when TcCM was diluted with various amounts
of HMI-9, the growth inhibition was always most pronounced in the case of the pleomorphic cell line. This was the case whether 75%, 50% or 25% conditioned medium was used (Figure 5.4)

**Figure 5.4** The effect of *T. congolense* conditioned medium on pleomorphs (AnTat 1.1 90.13) and monomorphs (427). A) The effect of various dilutions of conditioned medium or negative control medium on AnTat 1.1 90.13 cells. B) The effect of various dilutions of conditioned medium or negative control medium on 427 cells. Each data point represents the mean and SEM of triplicate wells.

Finally, while the TcCM effect was greatest on pleomorphs and the effect was lesser on monomorphs, there was almost no effect on growth of a low density *T. congolense* culture (Figure 5.5). This could again be due to variability in sensitivity to nutrient depletion or different nutritional requirements. Alternatively, *T.
congolense may produce signal(s) that only *T. b. brucei* is capable of responding to. This could be a case similar to *T. b. brucei* monomorphism, with the *T. congolense* tested having reduced ability to respond to a signal, but still being able to produce it. Though this may be the case for the cell line tested, it is not necessarily indicative of the response of other isolates.
Figure 5.5 The effect of *T. congoense* conditioned medium on *T. b. brucei* pleomorphic and monomorphic cell lines and *T. congoense* itself. A) The effect of conditioned medium on AnTat 1.1 90.13 *T. b. brucei*. B) The effect of conditioned medium on 427 *T. b. brucei*. C) The effect of conditioned medium on *T. congoense*. Each data point represents the mean and SEM of triplicate wells.

### 5.5 Variable sensitivity of monomorphic cell lines to *T. congoense* conditioned medium

Different monomorphic cell lines might be expected to have become non-responsive to SIF through different genetic changes. Their different passage histories may also have resulted in differing sensitivities to non-optimal growth conditions (i.e. nutrient depletion and toxic waste accumulation). To assess this, conditioned medium was collected on day 2 of AnTat 1.1 90.13 culture at a density of 3.6 x10^6/ml. Conditioned medium was also collected on days 3 and 4 of *T. congoense* culture at densities of 4.1x10^6/ml and 9.2x10^6/ml respectively. Thereafter the growth of one pleomorphic and 3 monomorphic *T. b. brucei* strains, as well as one *T. congoense* strain was compared under the various conditions (Figure 5.6).
Figure 5.6 The effect of *T. congoense* conditioned medium on *T. congoense*, and one pleomorphic and three monomorphic *T. b. brucei* cell lines. The effect of AnTat 1.1 90.13 conditioned medium (AnD2CM) and *T. congoense* conditioned medium collected after 3 or 4 days (TcD3CM, TcD4CM) on growth of one pleomorphic (AnTat) (A), and 3 monomorphic *T. b. brucei* cell lines (WT427, 427 90.13, 2T1) (B-D), as well as *T. congoense* (E). Each data point represents the mean and SEM of triplicate wells.
The effect of conditioned medium on KN counts. The effect of the conditions tested on AnTat 1.1 90.13 were compared using unpaired t tests, *** P<0.001, * P<0.05. The effect of AnTat CM compared to HMI-9 control media on 427 90.13 cells was compared using an unpaired t test (P=ns). The effects of TcCM from day 3 or 4 were compared to that of the TcBSF3 negative control using a one-way ANOVA followed by Dunnett’s multiple comparisons test (P=ns). All analyses were carried out in GraphPad Prism. The columns display mean and standard deviation for each condition based on three technical replicates. Unless otherwise stated, 250 cells were counted for each replicate. It was assumed that any variation within the cell line was minimal in comparison to the effect of treatment.

It was observed that TcCM causes less growth inhibition than AnTat CM when both were collected at comparable densities. TcCM collected on day 4 at a high density caused a particularly severe growth inhibition to AnTat 1.1 90.13. In all three monomorphic cell lines the inhibitory effect of TcCM was less than in the pleomorphic line. However, there was variation between the responses of the monomorph lines, with 427 90.13 showing greater sensitivity to TcCM than WT 427 or 2T1. The effect of AnTat CM was less severe on monomorph lines than on the pleomorphic line, with all monomorphs affected to a similar extent. As observed previously, the TcCM did not limit T. congolense growth to any considerable extent, although by day 3 of the experiment the plateauing of growth could result from nutrient exhaustion. Furthermore, AnTat CM did not affect T. congolense growth. Therefore, the T. congolense IL3000 strain used in this experiment did not appear to respond to T. b. brucei SIF, whereas pleomorphic T. b. brucei responded to T. congolense conditioned medium.
Cell cycle counts were carried out for AnTat 1.1 90.13 and 427 90.13 in conditions where there were enough cells present (Figure 5.7). This was not possible for AnTat 1.1 90.13 cells in day 4 TcCM as the growth effect was too severe. This analysis demonstrated that the reduction in the proportion of proliferating cells (%2K1N, 2K2N) was greatest when pleomorphic cells were treated with AnTat CM. Nevertheless, the slight reduction in AnTat 1.1 90.13 growth following treatment with day 3 TcCM was accompanied by a slight but significant reduction in the proportion of proliferating cells. Although, 427 90.13 cells responded to day 4 TcCM with a reduction in growth this was not accompanied by a significant reduction in proportion of proliferating cells. Once again these results support differential responses to TcCM in cell lines with different histories.

One caveat of these analyses to investigate growth effects of conditioned medium on different cell lines, is that observed growth differences sometimes varied, and different experiments sometimes provided contradictory results that are not easily explained. For example, in a follow up experiment (Appendix 3.2), the response of all T. b. brucei cell lines to T. congolense CM was more severe than in Figure 5.6 though the pattern of sensitivity was consistent, with monomorphs showing greatest resilience. However, in this case TcCM had an unexpectedly severe effect on T. congolense growth, while again AnTat CM had no effect. When this experiment was repeated, the effect of the TcCM collected was so severe that all T. b. brucei cell lines were killed, but T. congolense was able to continue growth at a level similar to the negative control (Appendix 3.3). This once again supported the observation that T. congolense was neither responding to AnTat nor TcCM. This was further supported in a third repeat of the experiment, with T. congolense not responding to either conditioned medium (Appendix 3.4). Nevertheless, another aspect of this experiment did not match expectation: all T. b. brucei cell lines responded comparably to both AnTat CM and TcCM. One explanation for this could be that the AnTat 1.1 90.13 used in that experiment had been over-passaged and may, therefore, have progressed towards monomorphism. Another factor to consider is that for these three more recent experiments, although the TcCM was collected at a comparable density to the earlier experiments, the conditioned medium was collected after only 2 days (rather than the usual 3) of culture. Since this could have affected how any
factor may have accumulated in the medium, the experiment was repeated a fourth time with conditioned medium collected on day 3. In this experiment, the responses to AnTat CM were as expected, with AnTat 1.1 90.13 growth more severely inhibited than any of the monomorph cell lines (Appendix 3.5). However, the effect of TcCM was comparatively small for all cell lines tested, and effected *T. congolense* growth on days 2 and 3 of the experiment. Another anomaly was the small effect of AnTat CM on *T. congolense* observed in this experiment. This effect could be explained by an initial growth inhibition that is overcome on days 2 and 3. Alternatively the cause could be an error in setting up the initial dilutions for this condition, as thereafter the AnTat CM and negative control cultures grew with similar trajectories.

In summary, conditioned medium, unlike a drug that may be accurately prepared to a desired dose, is difficult to prepare consistently. Hence, any potential secreted signalling factors would be accompanied by varying levels of nutrients and toxic waste products, any one of which could impact the growth of a test culture. Therefore, while it was possible to observe growth effect trends, a more refined analysis of the action of conditioned medium was required.

**5.6 Is sensitivity to *T. congolense* conditioned medium lessened in cells depleted for a SIF-responsive gene?**

Multiple factors could be responsible for changes in growth following treatment with conditioned medium; these include nutrient depletion and accumulation of toxic waste products. To address this issue, and determine whether the TcCM effect could be partly attributed to a differentiation response, an experiment was designed to test whether TcCM could be acting through the SIF-response pathway. Components of the SIF–response pathway have been identified by Mony et al (Mony et al., 2014), and RNAi lines for these components were available. Therefore, TcCM was tested in triplicate on an AnTat 1.1 90.13 cell line with inducible RNAi targeting Hypothetical protein 2 (Tb927.9.4080).
Hyp2 RNAi induction led to resistance to 8-CPT-cAMP as expected. However, there was no apparent resistance to TcCM (Figure 5.8A) based on the growth phenotype. Therefore, this experiment did not suggest that TcCM was having its effect through the SIF signalling pathway. Nevertheless, the possibility was not ruled out, since it was possible that the very strong growth inhibition by TcCM observed in this experiment was mainly a result of nutrient depletion, which may not always be the case. Nonetheless, to observe any potential for cell cycle arrest, one replicate from each condition was examined for their KN ratios. It was observed that there were less cells with a 1K1N configuration with TcCM (and 8-CPT-cAMP) when Hyp2 RNAi was induced compared to the uninduced cells on day 1 and 2 of the experiment (Figure 5.8B). A difference with TcBSF3 negative control medium was also observed on day 2 of the experiment between induced and uninduced cultures. However, this may be expected considering that by day 2, uninduced cultures had reached a high density and were likely to be responding to accumulation of their own SIF with growth arrest and accumulation in 1K1N. The accumulation in 1K1N of uninduced cells treated with TcCM, despite remaining at low density, is perhaps indicative of a differentiation effect, as is the difference in cell cycle between cells induced or uninduced for Hyp2 RNAi. Therefore, this experiment provided some evidence that inhibiting a component of the SIF-response signalling pathway resulted in less cell cycle arrest when exposed to T. congolense conditioned medium.
Effect of *T. congolense* conditioned medium on *T. b. brucei* Hyp2 RNAi

A

**Effect of *T. congolense* conditioned medium on *T. b. brucei* Hyp2 RNAi**

- **Dox 100% HMI-9**
- **-Dox 75% T.con CM + 25% HMI-9**
- **-Dox 75% TcBSF3 -ve control media +25% HMI-9**
- **-Dox 100% HMI-9 + 100µM cAMP**
- **+Dox 100% HMI-9**
- **+Dox 75% T.con CM + 25% HMI-9**
- **+Dox 75% TcBSF3 -ve control media +25% HMI-9**
- **+Dox 100% HMI-9 + 100µM cAMP**

**Day**

0 1 2 3 4

**Parasites/ml**

0 1×10⁶ 2×10⁶ 3×10⁶ 4×10⁶ 5×10⁶
Figure 5.8 The response of a Hyp2 RNAi line to treatment with *T. congolense* conditioned medium and 8-CPT-cAMP with or without induction of RNAi by doxycycline. A) The effect of conditioned medium or 8-CPT-cAMP on parasite growth with or without the induction of RNAi by doxycycline. Each data point represents the mean and SEM of triplicate wells. B) The effect of conditioned medium or 8-CPT-cAMP on the cell cycle (KN proportions) on days 1-3 of the experiment, with or without the induction of Hyp2 RNAi by doxycycline. KN proportions were calculated by counting 250 cells taken from a single well of each condition.

Further to these initial results that indicated a possible role for the SIF-response pathway in the response of *T. b. brucei* pleomorphs to *T. congolense* conditioned medium, a new method was performed to investigate the link between conditioned medium and the characterized SIF-response pathway.
5.7 Testing *T. congoense* conditioned medium effect on expression of a stumpy reporter

In order to more satisfactorily test if the TcCM effect was linked to a differentiation effect, TcCM was tested on the stumpy reporter line with CAT expression under the control of the PAD1 3’UTR. *T. b. brucei* conditioned medium was known to increase CAT expression from the reporter line (Figure 5.9), and so experiments were carried out to determine whether TcCM could evoke a similar response.

Figure 5.9 Schematic representation of a pleomorphic reporter line’s response to *T. b. brucei* conditioned medium. Addition of *T. b. brucei* conditioned medium containing SIF would trigger expression of the regulated CAT reporter gene placed under the control of the PAD1 3’UTR. Hence, the CAT reporter expression acts as a quantifiable indicator of stumpy formation.
Growth of pleomorphic AnTat 1.1 90.13 and monomorphic 427 cell lines is restricted by T.congolense conditioned medium

CAT concentration/cell is elevated following treatment of pleomorphic T. b. brucei with T.congolense conditioned media (CM)

Fold change in CAT concentration/cell

AnTat +75% T.con CM + 25% HMI-9
AnTat +75% TcBSF3-ve control media + 25% HMI-9
AnTat + 100% HMI-9 + 100 µM 8-CPT-cAMP
427 + 75% T.con CM + 25% HMI-9
427 + 75% TcBSF3-ve control medium + 25% HMI-9
427 + 100% HMI-9 + 100 µM 8-CPT-cAMP

CAT concentration/cell is elevated following treatment of monomorphic T. b. brucei with T.congolense conditioned media (CM)

Fold change in CAT concentration/cell

AnTat +75% T.con CM + 25% HMI-9
AnTat +75% TcBSF3-ve control media + 25% HMI-9
AnTat + 100% HMI-9 + 100 µM 8-CPT-cAMP
427 + 75% T.con CM + 25% HMI-9
427 + 75% TcBSF3-ve control medium + 25% HMI-9
427 + 100% HMI-9 + 100 µM 8-CPT-cAMP

CAT concentration/cell is elevated following treatment with T.congolense conditioned media or 8-CPT-cAMP

****

***

****

****

ns

****

****

****
In this experiment, unexpectedly, the positive control 8-CPT-cAMP had no growth effect on the monomorphic 427 reporter line, and a minimal effect on the pleomorphic AnTat 1.1 90.13 reporter line. Nevertheless, *T. congolense* conditioned medium had a severe growth effect on both pleomorphic and monomorphic reporter lines (Figure 5.10A). This effect on the monomorphic line was more severe than previously observed but is consistent with the observation that different monomorph strains may vary in their responsiveness. When CAT reporter expression was analysed by ELISA, it was observed that both 8-CPT-cAMP and TcCM increased CAT expression on days 1 and 2 of the experiment in the pleomorphic and monomorph reporter lines (Figure 5.10B, C and D). The large increase in mean CAT expression was significant, so data from individual days were considered separately. Analyses were carried out in GraphPad Prism. C) The fold change in CAT concentration/cell when the pleomorphic reporter cell line was treated with TcCM, relative to day 0. D) The fold change in CAT concentration/cell when the monomorphic reporter cell line was treated with TcCM, relative to day 0. Each bar represents the mean and standard deviation of samples from the triplicate flasks.
concentration/cell on day 2 of treatment of the monomorph cell line with *T. congoense* conditioned medium was accompanied by considerable variation in this instance, but warrants further investigation. This is because monomorphic cell lines respond to 8-CPT-cAMP with CAT reporter expression (MacGregor and Matthews, 2012), but they do not express the reporter in response to *T. b. brucei* conditioned medium (Figure 5.1).

In support of this initial data, when the experiment was again repeated in triplicate, TcCM treatment was once more shown to result in increased CAT reporter expression in the pleomorphic and monomorphic line (Figure 5.11). The effect on the monomorphic line on day 2 was particularly pronounced, and this time the variation between replicates was acceptable. It is important to note that in this experiment the cultures were not passaged once the experiment had begun. Therefore, SIF begins to accumulate in the negative control culture between days 2 and 3 with the result that, by day 3, the pleomorphic cell line was expressing the CAT reporter to a high degree. Comparisons between the TcCM and negative control conditions are thus more meaningful on days 1 and 2 of the experiment. Since the monomorphic cell line does not respond to its own SIF, the CAT expression in the negative control remained low for this cell line throughout the experiment.

Despite showing the same pattern as the initial experiment, the magnitude of differences between TcCM and negative control conditions were minor in this second experiment. Therefore, the experiment was repeated for a third time.
T. congolense day3 conditioned medium (CM) inhibits growth of pleomorphic AnTat 1.1 90.13 and monomorphic 427 cells.

**A**

- Parasites/ml
- Day
- $T. congolense$ CM D3 collected at $\approx 8.6-9.4 \times 10^6$/ml
- $T. con$ CM D3 ~pH 7.5
- TcBSF3 D3 ~pH 8.0

**B**

- CAT concentration/cell
- Day
- $T. congolense$ conditioned media (CM) inhibits growth of pleomorphic AnTat 1.1 90.13 and monomorphic 427 cells.

**Fold change in CAT concentration/cell**

- Day 0
- Day 1
- Day 2
- Day 3

**Legend**

- AnTat +75% T.con CM +25%HMI-9
- AnTat +75% TcBSF3 -ve control media +25%HMI-9
- 427 +75% T.con CM +25%HMI-9
- 427 +75% TcBSF3 -ve control medium +25%HMI-9

CAT concentration/cell (ng/ml per cell)
Figure 5.11 The effect of *T. congolense* conditioned medium on CAT-PAD reporter lines – Experiment 2. A) The growth effect of TcCM on pleomorphic (AnTat) and monomorphic (427) reporter lines. *T. congolense* conditioned medium was collected on day 3 at 8.6-9.4x10^6/ml at pH7.5, negative control medium had a pH of 8. Each data point represents the mean and SEM of triplicate flasks. B) CAT concentration per cell following treatment with TcCM with comparisons between categories made using unpaired t tests (* P<0.05, *** P<0.001, **** P<0.0001). Analyses were carried out in GraphPad Prism. C) The fold change in CAT concentration/cell when the pleomorphic cell line was treated with TcCM. D) The fold change in CAT concentration/cell when the monomorphic cell line was treated with TcCM. Each bar represents the mean and standard deviation of samples from the triplicate flasks.
**A**

**Effect of *T. congolense* and AnTat 1.1 90.13 conditioned media on AnTat CAT PAD reporter line**

- 75% *T. congolense* CM and 25% HMI-9 (TcCM)
- 75% TcBSF3 -ve and 25% HMI-9 (TcN)
- 75% AnTat CM and 25% HMI-9 (TbCM)
- 75% HMI-9 -ve and 25% HMI-9 (TbN)

Parasites/ml vs Day

**B**

**Effect of *T. congolense* and AnTat 1.1 90.13 conditioned media on 427 CAT PAD reporter line**

- 75% *T. congolense* CM and 25% HMI-9 (TcCM)
- 75% TcBSF3 -ve and 25% HMI-9 (TcN)
- 75% AnTat CM and 25% HMI-9 (TbCM)
- 75% HMI-9 -ve and 25% HMI-9 (TbN)

Parasites/ml vs Day
CAT concentration/cell is elevated following treatment of pleomorphic reporter line with *T. congolense* and *T. b. brucei* conditioned media

<table>
<thead>
<tr>
<th>Condition</th>
<th>Day 0</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% T. congolense CM and 25% HMI-9 (TcCM)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% TcBSF3 -ve and 25% HMI-9 (TcN)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% AnTat CM and 25% HMI-9 (TbCM)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% HMI-9 -ve and 25% HMI-9 (TbN)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
</tbody>
</table>

Note: 427 TcCM and TcN day 3 counts done on haemocytometer as many dead cells. No count for most cells in 427 flasks were dead (overgrown).

**Effect of** following treatment of monomorphic and CAT concentration/cell is elevated

<table>
<thead>
<tr>
<th>Condition</th>
<th>Day 0</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% T. congolense CM and 25% HMI-9 (TcCM)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% TcBSF3 -ve and 25% HMI-9 (TcN)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% AnTat CM and 25% HMI-9 (TbCM)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
<tr>
<td>75% HMI-9 -ve and 25% HMI-9 (TbN)</td>
<td><img src="https://example.com/image1" alt="" /></td>
<td><img src="https://example.com/image2" alt="" /></td>
<td><img src="https://example.com/image3" alt="" /></td>
<td>ns</td>
</tr>
</tbody>
</table>
CAT concentration/cell is elevated following treatment of pleomorphic reporter line with T. congolense and T. b. brucei conditioned media

F) The effect of AnTat CM and TcCM on pleomorphic and monomorphic T. b. brucei CAT-PAD reporter lines – Experiment 3. A) The effect of conditioned medium on growth of the pleomorphic cell line. B) The effect of conditioned medium on growth of the monomorphic cell line. Each data point represents the mean and SEM of triplicate flasks. By day 3 of the experiment most cells in the TbCM and TbN conditions were dead so no count is given, many dead cells were found in the TcCM and TcN conditions, so a haemocytometer was used for counts. C) The effect of AnTat CM and TcCM on CAT concentration/cell in the pleomorphic reporter line. D) The effect of AnTat CM and TcCM on CAT concentration/cell in the monomorphic reporter line. No CAT sample was available on day 3 due to cell death. All comparisons between categories were made using unpaired t tests (* P<0.05, *** P<0.001). Analyses were carried out in GraphPad Prism. E) The fold change in CAT concentration/cell when the pleomorphic cell line was treated with TcCM or AnTat CM. F) The fold change in CAT concentration/cell when the monomorphic cell line was treated with TcCM or AnTat CM. Each bar represents the mean and standard deviation of samples from the triplicate flasks.
In the final analysis of the effect of TcCM on pleomorphic and monomorphic CAT-PAD reporter lines, AnTat CM was included as a positive control. In this experiment the increase in CAT concentration/cell with TcCM relative to negative control treatment was of a greater magnitude than in the previous experiment, and was comparable to the increase observed with AnTat CM treatment on days 1 and 2 (Figure 5.12). Additionally, the increase in CAT expression on day 2 following treatment of the monomorphic reporter line with TcCM, but not AnTat CM, was once again observed (Figure 5.12D).

To summarise, *T. congolense* conditioned medium induced inhibition of the growth of the pleomorphic *T. b. brucei* CAT-PAD reporter line. Moreover, this was associated with an increased expression of the CAT reporter, here used to report on ‘stumpiness’. This effect was reproducible in three separate experiments. Whereas only pleomorphs upregulated expression of the stumpy reporter in response to AnTat CM, both pleomorphs and monomorphs demonstrated increased CAT expression in response to TcCM. If *T. congolense* is indeed producing a SIF-like factor, it is conceivable that though monomorphs have lost the response to their own factor through culture passages, they could have retained sensitivity to a *T. congolense* factor.

**5.8 Does a cell line depleted of a SIF-response gene produce less of a stumpy reporter when treated with conditioned medium?**

Previous efforts to link conditioned medium to the characterised stumpy induction pathway had been unsuccessful, possible due to the inability to disentangle effects of nutritional stress from differentiation effects when looking at growth alone. The CAT-PAD reporter provides the more specific output required, and therefore a pleomorphic Hyp2 RNAi line was transfected with the CAT-PAD 3’UTR reporter construct. The generated cell line could then be tested with *T. b. brucei* and *T. congolense* conditioned medium to determine whether RNAi induction inhibited the normal increase in CAT expression when the reporter line was treated with conditioned medium (Figure 5.13).
New strategy to investigate the mechanism of the observed conditioned medium effect. When a pleomorphic reporter line with CAT expression under the control of the PAD1 3’UTR is treated with conditioned medium there is a detectable increase in CAT reporter expression. If a factor(s) in the conditioned medium were acting through the SIF-response pathway, then knock down of a component of this pathway, Hyp2, would be expected to prevent this increase in reporter expression.

5.8.1 Validation of the new reporter line in vitro

Alongside the CAT-PAD 3’UTR reporter construct, the Hyp2 RNAi line was also transfected with a control CAT 449 construct. In the CAT-PAD 3’UTR reporter line, CAT expression will be held repressed in slender parasites unless they receive a trigger that initiates stumpy formation, whereas in the control cell line CAT expression should be constitutive. Both cell lines were treated with doxycycline to induce Hyp2 RNAi, prior to treatment with AnTat CM. The growth responses of both the test and control cell line to the T. b. brucei conditioned medium were comparable, and in both cases Hyp2 RNAi resulted in resistance to 8-CPT-cAMP and AnTat CM (Figure 5.14A). Despite the growth inhibition of the control CAT449 cell line following 8-CPT-cAMP or conditioned medium treatment, CAT reporter expression did not increase whether or not Hyp2 RNAi was induced (Figure 5.14B). This is in contrast to the marked increase in CAT concentration with AnTat CM and 8-CPT-cAMP treatments, observed in the uninduced CAT-PAD reporter line. However, when Hyp2 RNAi was induced in the CAT-PAD line, the increase in the CAT reporter following AnTat CM treatment was entirely absent. A response was observed to 8-CPT-cAMP, but it was delayed in the induced relative to uninduced conditions (Figure 5.14B).
Figure 5.14 The effect of *T. b. brucei* conditioned medium and 8-CPT-cAMP on a pleomorphic reporter line with inducible Hyp2 RNAi. A) The growth effect of treatment with AnTat CM or 8-CPT-cAMP when Hyp2 RNAi is either induced (+Dox) or uninduced (-Dox) in a control cell line (left) or in a cell line with CAT expression under the control of the PAD1 3’UTR (right). For each condition tested n=1. B) The effect of AnTat CM or 8-CPT-cAMP treatment on CAT reporter concentration/ cell (blue bars = control cell line, green bars = CAT-PAD reporter line).
The relatively consistent CAT expression observed under each condition for the CAT449 constitutive control confirmed that the variation in CAT concentration in the CAT-PAD reporter line is a result of regulation through the PAD1 3’UTR. Further, the prevention of the increase in CAT reporter expression in response to AnTat CM caused by induction of Hyp2 RNAi supports a scenario where SIF in AnTat CM is acting through a pathway involving Hyp2. Having validated this approach this preliminary experiment was repeated in triplicate with T. congolense conditioned medium, to see whether the action of this conditioned medium could also be linked to the SIF-response pathway.

### 5.8.2 Validation of the new reporter line in vivo

The newly generated AnTat CAT-PAD 3’UTR reporter line was also tested for pleomorphism in vivo (with Hyp2 RNAi uninduced), by examination of whether differentiation to the stumpy form was accompanied by expression of the CAT reporter. During this test infection, high parasitaemia was accompanied by increased PAD1 expression, stumpy morphology, and a reduction in proliferating cells (%2K1N, 2K2N), all indicating maintenance of pleomorphism (Figure 5.15).

![Figure 5.15](image)

Figure 5.15 The TbHyp2 RNAi CAT-PAD 3’UTR reporter line is capable of differentiating to stumpy forms at high parasitaemia as evidenced by morphology, increased PAD1 expression and reduction in proliferating cells. Measurements presented are from a single infection. In the image on the left stumpy morphology cells were shown to be PAD1 positive (green).
A stumpy sample collected on the final day of infection was also found to demonstrate high CAT expression, higher than that of a dense *in vitro* culture (Figure 5.16). This supported use of the CAT reporter as a good surrogate marker of PAD1 expression and therefore ‘stumpiness’.

![Figure 5.16](image)

**Figure 5.16** The effect of high density *in vitro* or *in vivo* on CAT reporter expression in the TbHyp2 RNAi CAT-PAD reporter line, when Hyp2 RNAi remained uninduced. Each column represents the CAT concentration/cell from a single infection or flask.

### 5.8.3 Effect of *T. congolense* conditioned medium on the TbHyp2 RNAi CAT-PAD reporter line

AnTat CM and TcCM were tested in parallel on the CAT-PAD reporter line with or without induction of Hyp2 RNAi. Whereas AnTat CM and TcCM both impacted cell growth relative to the negative control medium, the effect was reduced when Hyp2 RNAi was induced (Figure 5.17A). Consequently, changes in CAT concentration/cell were examined and, as observed previously, uninduced cells responded to both conditioned media by increasing CAT reporter expression on days 1 and 2 of the experiment relative to the negative controls. As with the trial experiment, when Hyp2 RNAi was induced, the increase in CAT concentration with AnTat CM treatment was lost (Figure 5.17B). This link between the effects of *T. b.*
*brucei* conditioned medium and the SIF-response pathway had not previously been demonstrated and supports the presence of SIF in this conditioned medium and the role of Hyp2 in responding to it. Similarly, the increase in CAT expression on day 1 of the experiment following treatment with TcCM observed for uninduced cells did not occur when Hyp2 RNAi was induced (Figure 5.17B). However, unlike the response to AnTat CM, CAT expression did increase on days 2 and 3 of the experiment after exposure to TcCM treatment even when Hyp2 RNAi was induced, albeit with delayed kinetics relative to the uninduced control.
Effect of AnTat 1.1 90.13 conditioned media on AnTat TbHyp2RNAi CAT PAD reporter line

Effect of *T. congolense* conditioned media on AnTat TbHyp2RNAi CAT PAD reporter line

Fold change in CAT concentration/cell when AnTat Hyp2 RNAi CAT PAD 3'UTR treated with AnTat conditioned medium

Fold change in CAT concentration/cell when AnTat Hyp2 RNAi CAT PAD 3'UTR treated with *T. congolense* conditioned medium
Effect of AnTat 1.1 90.13 conditioned media on AnTat TbHyp2RNAi CAT PAD reporter line

<table>
<thead>
<tr>
<th>Parasites/ml</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% AnTat CM +Dox</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>75% AnTat CM -Dox</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>75% HMI-9 +Dox</td>
<td>****</td>
<td>****</td>
<td>****</td>
</tr>
<tr>
<td>75% HMI-9 -Dox</td>
<td>****</td>
<td>****</td>
<td>****</td>
</tr>
</tbody>
</table>

Effect of T. congolense conditioned media on AnTat TbHyp2RNAi CAT PAD reporter line

<table>
<thead>
<tr>
<th>Parasites/ml</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% T.con CM +Dox</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>75% T.con CM -Dox</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>75% TcBSF3 +Dox</td>
<td>***</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>75% TcBSF3 -Dox</td>
<td>****</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

CAT conc/cell when AnTat Hyp2 RNAi CAT PAD 3'UTR treated with AnTat conditioned medium

<table>
<thead>
<tr>
<th>Day 0</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% AnTat CM +Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% HMI-9 +Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% AnTat CM -Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% HMI-9 -Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
</tbody>
</table>

CAT conc/cell when AnTat Hyp2 RNAi CAT PAD 3'UTR treated with T. congolense conditioned medium

<table>
<thead>
<tr>
<th>Day 0</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>75% T.con CM +Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% TcBSF3 +Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% T.con CM -Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
<tr>
<td>75% TcBSF3 -Dox</td>
<td>6×10^-7</td>
<td>4×10^-7</td>
<td>2×10^-7</td>
</tr>
</tbody>
</table>
Figure 5.17 The effect of TcCM and AnTat CM on the TbHyp2 RNAi CAT-PAD reporter line with or without induction of RNAi. A) The growth effect of conditioned medium with or without induction of Hyp2 RNAi by doxycycline (upper = AnTat CM, lower = TcCM). Each data point represents the mean and SEM of triplicate flasks. B) CAT concentration/cell following treatment with AnTat CM (upper) or TcCM (lower). Comparisons between categories were made using two-way ANOVA followed by Tukey’s multiple comparisons test (* P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001). In the case of both the AnTat CM and TcCM the interaction between time and treatment was found to be significant, so data from individual days were considered separately. Analyses were carried out in GraphPad Prism. C) The fold change in CAT concentration/cell following treatment with AnTat CM (upper) or TcCM (lower). Each bar represents the mean and SEM of samples from the triplicate flasks. It was assumed that any variation within the cell line was minimal in comparison to the effect of treatment.
The different CAT reporter expressions observed between conditions were supported by cell cycle analyses, such that high CAT expression was accompanied by an accumulation in 1K1N (Figure 5.18). For instance while the proportion of proliferating cells (%2K1N, 2K2N) was greatly reduced by AnTat CM and TcCM when Hyp2 RNAi was uninduced, when Hyp2 RNAi was induced conditioned medium treated cultures showed comparable cell cycle profiles to negative controls. This difference between induced and uninduced cultures was most pronounced on days 1 and 2 of the experiment. By day 3 of the experiment, even the induced cultures showed a reduction in proliferating cells in response to TcCM, matching the levels in the uninduced cultures. This was consistent with the previous observation that Hyp2 RNAi delayed CAT expression in response to TcCM.
The effect of AnTat CM and TcCM on the proportion of proliferating cells (2K1N, 2K2N) in the CAT-PAD reporter line with or without the induction of Hyp2 RNAi. Comparisons between categories made using a two-way ANOVA followed by Tukey’s multiple comparisons test (* P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001). The interaction between time and treatment was found to be significant, so data from individual days were considered separately. Analyses were carried out in GraphPad Prism. Each bar represents the mean and standard deviation of samples from the triplicate flasks. 250 cells were counted for each replicate. It was assumed that any variation within the cell line was minimal in comparison to the effect of treatment.

The results of these analyses indicate that SIF in AnTat CM is acting through a pathway involving Hyp2. Additionally, stumpy reporter expression induced by TcCM may also result from a SIF-like molecule, this being signalled through Hyp2. However, Hyp2 RNAi only delays reporter expression in response to TcCM, similar to the effect in the presence of 8-CPT-cAMP. Another common feature of TcCM and 8-CPT-cAMP was that both inhibited monomorph growth, unlike AnTat CM which cannot. Therefore, TcCM and 8-CPT-cAMP may be able to act through an additional, slower-acting pathway bypassing Hyp2, which is still present in monomorphs.
5.9 Co-infection

While conditioned medium experiments provide a convenient and simplified framework for analysing signalling between species, they do not account for the complexity of interactions occurring in the host. Therefore, an analysis of the impact of coinfections on the participating species was performed, supplementing findings from the in vitro experiments.

5.9.1 Simultaneous infection with *T. b. brucei* and *T. congolense*

To monitor *T. congolense* and *T. b. brucei* in a co-infection, a method was devised to easily distinguish the two species. In order to definitively identify parasites as *T. congolense* or *T. b. brucei* in mixed infections, AnTat 1.1 90.13 parasites were transfected with a construct to tag the *T. b. brucei* paraflagellar rod (PFR) structure. This tag could then be detected by antibody binding and the fluorescence used to identify *T. b. brucei*. Primers were designed to amplify the 5’end of the ORF and 3’end of the 5’UTR for endogenous N-terminal tagging of TbPFR2 (Tb927.8.4970) (Appendix 2.4) using the vector pEnT6BY (Kelly et al., 2007) (Figure 5.19).

![Figure 5.19 Illustration of the strategy used for endogenous tagging of *T. b. brucei* PFR2.](image)

The plasmid construct was linearised using Xhol before transfection. The construct was integrated by homologous recombination between the 5’UTR and ORF sequences in the construct and those in the endogenous TbPFR2 locus, enabling insertion of a TY-epitope tag at the N-terminus of the TbPFR2 protein. On the right is an image of a successfully transfected parasite, the epitope tagged TbPFR2 is detected using BB2 antibody (red).
The TY-tagged paraflagellar rod of transfected parasites can be identified by BB2 antibody (Bastin et al., 1996), which can be used in combination with PAD1 antibody to identify stumpy parasites, with DAPI staining for cell cycle analysis.

In a trial experiment one mouse was simultaneously infected with T. congolense and TY-tagged PFR expressing T. b. brucei (PFR.TY T. b. brucei). In parallel one mouse was infected with PFR.TY T. b. brucei alone, and one mouse was infected with T. congolense alone for comparison. The T. b. brucei parasitaemia followed a similar profile in the co-infected mouse as in the single species infection (Figure 5.20A). The proportion of the total parasitaemia occupied by T. congolense in the mixed infection was less than that occupied by T. b. brucei on all days monitored (days 3-6), but increased from days 3 to 6 (Figure 5.20B). The proportion of the parasitaemia occupied by each species was determined by counting at least 250 cells on each day of infection, with the BB2 staining of the paraflagellar rod structure distinguishing parasites as T. b. brucei (Figure 5.20C). The maximal parasitaemia reached by T. congolense in the co-infection was lower than in the T. congolense infection alone, possibly indicating suppression by the T. b. brucei parasites (Figure 5.20A).

Following this, a cell cycle analysis was carried out for each parasite species on each day of infection. Images of parasites stained with DAPI, PAD1 antibody and BB2 antibody were captured and overlaid in Image J (Rasband, 1997-2015) for analysis. BB2 staining of the TY-epitope tagged paraflagellar rod of T. b. brucei allowed this parasite to be distinguished from the T. congolense in the co-infection (Figure 5.20C). KN analysis revealed that T. b. brucei had a slightly higher proportion of 2K1N, 2K2N cells in the mixed infection than in the single species infection on days 3 and 4 (Figure 5.21A). T. congolense also had an increased proportion of 2K1N, 2K2N cells in the mixed infection compared to in the single species infection, although any difference was very slight (Figure 5.21B; note low numbers of cells counted on days 3 and 4). Despite the high degree of 1K1N accumulation, indicative of G0/G1 arrest in the T. b. brucei population by day 6 of the mixed infection, T. congolense continued to proliferate to a reasonable extent in the mixed infection. This would indicate that the growth limitation experienced by T. congolense in the
mixed infection was not a result of a *T. b. brucei* produced factor driving growth arrest in *T. congolense*.

The proportion of *T. b. brucei* cells that were PAD1 positive was also followed on each day of the single species and mixed infections (Figure 5.21C). The proportion of *T. b. brucei* PAD1 positive cells in a mixed infection did not appear to differ significantly from that in a single species infection. There was, however, a slightly higher proportion of PAD1 positive cells in the single species infection on day 4, compared to the mixed infection. At the end of each infection, on day 6, parasites were harvested and purified, and were then induced to differentiate to procyclic forms through addition of cis-aconitate and incubation at 27°C in SDM-79. Samples for flow cytometry were taken at 0h, 4h and 24h after addition of cis-aconitate, and were stained with an antibody against EP procyclin. This antibody was only expected to detect *T. b. brucei* procyclin, so that *T. congolense* cells were expected to be negative. Indeed, the majority of *T. congolense* cells were EP procyclin negative. There was no significant difference between the proportion of EP procyclin positive cells in the mixed infection and the *T. b. brucei* only infection (Figure 5.21D).
Figure 5.20 Simultaneous co-infection of *T. congolense* and *T. b. brucei*. A) Parasitaemia in single or mixed infections (n=1), total parasitaemia of the mix is shown on the left graph, a breakdown of parasitaemia by species is shown on the right graph. Total parasitaemia – PFR.TY positive cells = number of *T. congolense* cells. B) The percentage of cells belonging to each species on each day of the experiment. The number of cells counted for each condition is shown below each chart. C) Representative image taken from day 5 of the co-infection. The TY-tagged parflagellar rod is in red, PAD1 staining is in green and DAPI staining is in white. A *T. b. brucei* parasite is marked Tb, and a *T. congolense* parasite is marked Tc. The scale bar represents 10µm.
Overall, when *T. congolesens* and *T. b. brucei* were injected together to establish a co-infection, there was no marked effect on *T. b. brucei* parasitaemia or on the extent of
differentiation to stumpy forms. However, the *T. congoense* parasitaemia was limited in the mixed infection, such that *T. congoense* peaked at a lower parasitaemia than in the single species infection. In this experiment, the *T. b. brucei* parasitaemia increased more rapidly than the *T. congoense* parasitaemia. Therefore, it was thought necessary to initiate the *T. congoense* infection earlier, to allow its establishment before *T. b. brucei* challenge, in order for any potential impact of *T. congoense* on *T. b. brucei* to take effect.

### 5.9.2 *T. congoense* infection followed by *T. b. brucei* challenge

In a subsequent trial infection, *T. congoense* was injected on day 0 and *T. b. brucei* on day 1 of the experiment. Once again, parallel mice were set up with single species infections for comparison. The parasitaemia was monitored from days 4-6 of the experiment, and the parasitaemia attributable to each species in the co-infection was estimated as described in section 5.9.1. *T. congoense* dominated the parasitaemia in the mixed infection on day 4, but was overtaken by *T. b. brucei* on days 5 and 6 of the experiment (Figure 5.22A). Correspondingly, the proportion of the mixed infection attributed to *T. congoense* decreased from day 4 to day 6 (Figure 5.22B).
Figure 5.22 *T. congolense* infection followed by *T. b. brucei* challenge. A) Parasitaemia in single or mixed infections (n=1); total parasitaemia of the mix is shown on the left graph, a breakdown of parasitaemia by species is shown on the right graph. B) The percentage of cells belonging to each species on each day of the experiment. The number of cells counted for each day is shown below each chart.

Cell cycle analysis was performed as described in section 5.9.1. The proportion of 2K1N, 2K2N *T. b. brucei* cells was comparable on each day of infection whether *T. b. brucei* was present alone in a single species infection or with *T. congolense* in a mixed infection. There were slightly fewer 2K1N, 2K2N *T. b. brucei* cells in the mixed infection than in the single species infection on day 4 of the experiment, but there was not strong evidence for a greater degree of cell cycle arrest in the mixed infection (Figure 5.23A). The proportion of 2K1N, 2K2N *T. congolense* cells fluctuated in both the single species infection and the co-infection (Figure 5.23B). There was a slight reduction in the *T. congolense* 2K1N, 2K2N proportion in the mixed infection relative to the single species infection on day 6, but again there was not significant evidence of cell cycle arrest.

The proportion of *T. b. brucei* PAD1 positive cells was not markedly affected by the presence of co-infecting *T. congolense* parasites (Figure 5.23C). There was a lower
proportion of *T. b. brucei* PAD1 positive cells in the mixed infection relative to the single species infection on day 5. However, this is more likely to reflect the lower *T. b. brucei* parasitaemia in the mixed infection than in the single species infection at this time point, rather than a direct effect of *T. congolense* on *T. b. brucei* differentiation. On day 6 of the experiment parasites were harvested and purified, and then induced to differentiate to procyclic forms as described in section 5.9.1. Both *T. b. brucei* from the single species infection and those from the mixed infection demonstrated comparable ability to differentiate to procyclic forms, as judged by the proportion of EP procyclin positive cells at 4h and 24h after induction of differentiation (Figure 5.23D).
Figure 5.23 Effect of co-infection on the cell cycle and proportion of *T. b. brucei* stumpy forms.

A) An analysis of the *T. b. brucei* cell cycle on each day of the infection, with or without the presence of co-infecting *T. congolense*. B) Analysis of the *T. congolense* cell cycle on each day of infection, with or without the presence of co-infecting *T. b. brucei*. The numbers of cells counted in each case are displayed under each chart. C) The effect of co-infecting *T. congolense* parasites on the proportion of PAD1 positive *T. b. brucei* cells, as compared to a single *T. b. brucei* infection. D) The presence of co-infecting *T. congolense* did not effect the ability of the *T. b. brucei* population to differentiate to procyclic forms upon induction compared to *T. b. brucei* from a single species infection, as determined by the proportion of EP Procyclin positive cells at 4h or 24h post-induction. Each measurement represents one single species infection or one co-infection.
Overall the results from this trial experiment indicate that a pre-established *T. congolense* infection can limit *T. b. brucei* population expansion initially. However, the cell cycle, PAD1 and onward differentiation analyses do not suggest that *T. congolense* limitation of *T. b. brucei* parasitaemia is related to increased differentiation to stumpy forms. As observed in the previous trial experiment, the peak of *T. congolense* parasitaemia was lower in the co-infection than in a single species infection.

### 5.9.3 *T. congolense* infection followed by *T. b. brucei* challenge in triplicate

In order to assess the reproducibility of the trial experiment in section 5.9.2, infections were repeated in triplicate. On day 0 of the experiment, 6 mice were infected with *T. congolense* bloodstream forms, on the subsequent day, 3 of these mice were challenged with *T. b. brucei* bloodstream forms, along with 3 other previously uninfected mice. The parasitaemia was monitored from day 3 to 6 post infection.

The total parasitaemia reached in all infections was of comparable magnitude by the end of the experiment. When the parasitaemia of *T. b. brucei* in single species infections was compared to that in co-infections, it was observed that the parasitaemia remained lower in the co-infections for the duration of the experiment, but was still able to reach high parasitaemia (>1x10⁸/ml). On day 4 *T. congolense* infections had reached higher parasitaemia in the mixed infections than the single species infections. However, from day 4 to day 6 the *T. congolense* parasitaemia in the mixed infections plateaued, whereas the parasitaemia of single species infections continued to increase to greater than 1x10⁸/ml by day 6 (Figure 5.24A). The parasitaemia for each individual species in the co-infections was estimated by applying the proportion of the parasitaemia attributed to each species to the total parasitaemia. The proportion of the parasitaemia attributed to each species was calculated by scoring PFR labelled cells by immunofluorescence, as described in section 5.9.1. It was observed that on days 3 and 4 of the co-infections the parasitaemia was dominated by *T. congolense*. However, by days 5 and 6 *T. b. brucei* had become the dominant species in the co-infection (Figure 5.24B).
Both single species and co-configuration parasitaemia were found to be no significant difference on any given day (Figure 5.24). Superseded on days 5 and 6 by a growth-arrested population, indicated by a significant proportion of 2K1N, 2K2N cells, was species infection, and in a co-infection on the left graph, a breakdown of parasitaemia by species is shown on the right graph. Each data point represents the mean and SD of three biological replicates. B) The percentage of cells belonging to each species on each day of the experiment. Each bar represents the mean and standard deviation of three biological replicates.

Cell cycle analysis revealed a comparable trend for T. b. brucei parasites in a single species infection, and in a co-infection (Figure 5.25A). The initially proliferative population, indicated by a significant proportion of 2K1N, 2K2N cells, was superseded on days 5 and 6 by a growth-arrested population, as indicated by accumulation in 1K1N. Indeed, when the percentage of 1K1N cells was compared between T. b. brucei parasites in single species infections or co-infections, there was found to be no significant difference on any given day (Figure 5.25B). Likewise, the KN configurations for T. congolense remained relatively consistent on all days in both single species and co-infections, and no significant difference in the percentage of 1K1N on any given day was observed (Figure 5.25B).
Figure 5.25 Effect of co-infection on the cell cycle and *T. b. brucei* differentiation to the stumpy form. A) An analysis of the *T. b. brucei* and *T. congolense* cell cycle on each day of the infection, with or without the presence of another co-infecting species. Each bar represents the mean and standard deviation of three biological replicates. B) The percentage of cells with a 1K1N configuration was not significantly different between parasites in a single species infection as compared to a co-infection, for either species, when individual days were considered separately. Separate 2-way ANOVAs followed by Bonferroni’s multiple comparisons test were performed for each species using GraphPad Prism. C) The effect of co-infecting *T. congolense* parasites on the proportion of PAD1 positive *T. b. brucei* cells, as compared to a single *T. b. brucei* infection. Unpaired t tests found no significant difference between *T. b. brucei* cells in a single species infection as compared to a co-infection on any day analysed. Statistical analysis was carried out in GraphPad Prism. Each bar represents the mean and standard deviation of 3 biological replicates. D) The presence of co-infecting *T. congolense* did not significantly effect the ability of the *T. b. brucei* population to differentiate to procyclic forms upon induction compared to *T. b. brucei* from a single species infection, as determined by the proportion of EP Procyclin positive cells at 4h or 24h post-induction. Unpaired t tests were carried out using GraphPad Prism. Each bar represents the mean and standard deviation of 3 biological replicates, except for the *T. congolense* samples where n=2 and the procyclic controls where n=1.
The proportion of *T. b. brucei* PAD1 positive cells was similar when comparing parasites from single species infections with those from mixed infections (Figure 5.25C). There was an increase in the proportion of PAD1 positive cells in the single species infections relative to the mixed infections on day 5, but this was found not to be significant. *T. b. brucei* parasites from single species or mixed infections were isolated and induced to differentiate to procyclic forms through addition of cis-aconitate (as described in section 5.9.1). *T. b. brucei* parasites from co-infections displayed comparable ability to differentiate to procyclic forms as those from single species infections, as assessed by EP procyclin expression after 4h or 24h (Figure 5.25D).

In summary, the presence of co-infecting *T. congolense* is able to limit *T. b. brucei* parasitaemia relative to that in a single species infection. However, this limitation is not the result of increased cell cycle arrest, or an increased proportion of stumpy forms. Likewise, the reduced *T. congolense* peak parasitaemia observed in the co-infections, as compared to the single species infections, was not associated with an increase in the proportion of cells in 1K1N.

**5.10 Conclusions and future directions**

Conditioned medium from a dense culture of *T. b. brucei* was known to inhibit growth of a low-density *T. b. brucei* culture, and co-incidentally increase expression of a stumpy reporter (CAT under the control of the PAD1 3’UTR). To test the potential for interaction in a co-infection, *T. congolense* conditioned medium from dense cultures was tested on low density *T. b. brucei*.

*T. congolense* conditioned medium was demonstrated to be most effective at limiting *T. b. brucei* growth when collected after 3 days of culture, and diluted to 75% using HMI-9. Therefore, these conditions were used for the majority of the experiments. *T. congolense* conditioned medium was more effective at inhibiting growth of pleomorphic AnTat 1.1 90.13 cells, than most of the monomorphic cell lines tested, and in most cases did not appear to have any effect on *T. congolense* itself. This could indicate that *T. congolense* was producing a factor that only *T. b. brucei* was capable of responding to, in which case *T. congolense* cells would be behaving like
monomorphic *T. b. brucei* parasites, remaining able to produce a factor but incapable of responding to it themselves. However, the differing growth effects of *T. congoense* conditioned medium on various cell lines were not always reproducible. This was likely to be a result of multiple variable factors in the conditioned medium having an impact on growth, such as nutrient depletion, toxic metabolite build up and possible accumulation of a quorum sensing molecule(s). When looking at the effect of conditioned medium on growth alone, any variations in nutrient depletion or toxic metabolite accumulation, which were difficult to control for, could have overshadowed any effect of possible signalling molecules. Growth alone was not therefore considered to be an adequate measure of the effect of *T. congoense* conditioned medium.

Evidence for an effect of *T. congoense* conditioned medium on differentiation was obtained when treatment with *T. congoense* conditioned medium was found to increase expression of a stumpy reporter. Expression of the CAT reporter, that was under the control of the PAD1 3′UTR, was increased by *T. congoense* conditioned medium treatment in a pleomorphic reporter line. This effect was reproducible and comparable to the effect observed following treatment of the pleomorphic reporter line with a *T. b. brucei* AnTat 1.1 90.13 conditioned medium. However, unlike the effect of the *T. b. brucei* conditioned medium, the *T. congoense* conditioned medium was also able to induce CAT expression in a monomorphic 427 90.13 reporter line. It may be that while monomorphic *T. b. brucei* had lost sensitivity to *T. b. brucei* SIF through extended culture, it had retained sensitivity to a *T. congoense* signalling molecule to which it has not been previously exposed. The effect of the *T. congoense* conditioned medium is comparable to that of 8-CPT-cAMP, which also inhibits monomorph growth with co-incident induction of stumpy reporter expression.

In order to determine whether *T. b. brucei* and *T. congoense* conditioned media were acting through the characterised stumpy induction pathway, a Hyp2 RNAi cell line was transfected with the CAT PAD 3′UTR reporter construct. Induction of Hyp2 RNAi with doxycycline was found to prevent the usual increase in CAT reporter expression in response to *T. b. brucei* conditioned medium treatment. This provided
evidence that *T. b. brucei* conditioned medium was indeed having an effect through the SIF-signalling pathway. Induction of Hyp2 RNAi also delayed the induction of CAT expression in the reporter line in response to *T. congolense* conditioned medium but did not prevent it entirely. This was similar to the effect observed for Hyp2 RNAi on the response of the reporter line to 8-CPT-cAMP. This indicates that though the initial effect of *T. congolense* conditioned medium and 8-CPT-cAMP may involve signalling through Hyp2, this pathway may be bypassed so that parasites are able respond in a delayed fashion through an alternative pathway. Alternatively, low levels of Hyp2 expression that remain following induction of RNAi provide a delayed response.

Cell cycle analysis revealed that there was an accumulation in 1K1N following treatment with *T. b. brucei* or *T. congolense* conditioned medium when Hyp2 RNAi was uninduced. When Hyp2 RNAi was induced by doxycycline this cell cycle arrest was lost on days 1 and 2. However, in the case of *T. congolense* conditioned medium, a significant reduction in proliferating (2K1N, 2K2N) cells was observed by day 3, even when Hyp2 RNAi was induced. Therefore, the cell cycle analysis supported the CAT expression data, with increased cell cycle arrest correlated with increased stumpy reporter expression.

The ability of *T. congolense* conditioned medium to inhibit growth and induce expression of a stumpy marker in a monomorphic cell line provides the opportunity to investigate the pathway(s) through which *T. congolense* conditioned medium is having its effect. A monomorphic RNAi library was generated in *T. b. brucei* (Alsford et al., 2011), and though not responsive to *T. b. brucei* conditioned medium this cell line was responsive to 8-CPT-cAMP, a sensitivity used to identify SIF-responsive genes (Mony et al., 2014). This method was restricted to identifying genes whose products formed part of the pathway for which the SIF and 8-CPT-cAMP responses overlapped, and therefore missed potential cell surface receptors bypassed by cell permeable 8-CPT-cAMP. If *T. congolense* conditioned medium is able to restrict growth of the *T. b. brucei* library to a sufficient extent to allow outgrowth of a resistant population, it may be possible to identify components of the pathway(s) responsible for the *T. congolense* conditioned medium response. As a
A preliminary test of this, conditioned media from three different densities of *T. congolense* culture were tested on the library either undiluted (100%) or diluted to 75% with HMI-9 (Figure 5.26). All conditioned media were collected after 3 days of culture, but cultures had been established at different densities. Conditioned media collected from high-density (7x10⁶/ml or 8x10⁶/ml) *T. congolense* cultures was able to inhibit growth of the RNAi library severely when diluted to 75% with HMI-9. Conditioned medium collected from a less dense culture (5x10⁶/ml) was only able to limit growth effectively when 100% conditioned medium was used, and only after passaging the culture back to a low density on day 1 after treatment. Therefore, conditioned medium from very dense cultures could be used to screen the RNAi library for genes implicated in sensitivity to *T. congolense* conditioned medium effects.

**Figure 5.26 The effect of *T. congolense* conditioned medium on the monomorphic RNAi library.** Conditioned media from high-density *T. congolense* cultures (7x10⁶/ml – orange lines, 8x10⁶/ml – red lines) were sufficient to inhibit growth of the RNAi library severely when 75% *T. congolense* conditioned medium was used. Each data point represents a measurement from a single well.

While an effect of *T. congolense* on *T. b. brucei* growth and differentiation was detected in conditioned media experiments, no such effect could be detected in co-infections. Although the *T. b. brucei* parasitaemia was reduced when co-infected with *T. congolense* relative to in a single species infection, this was not associated
with increased growth arrest or differentiation to the stumpy form. The *in vitro* and *in vivo* results are not mutually exclusive. In the conditioned medium experiments, the implication is that a dense population of *T. congolense* is able to inhibit growth of low-density *T. b. brucei* resulting in increased expression of stumpy markers. In contrast, in the co-infection studies, on the first day parasitaemia was monitored *T. b. brucei* parasitaemia was already approaching that of *T. congolense*, and both species were in the ascending phase of parasitaemia. This was despite infecting a day earlier and with a much a higher dose of *T. congolense* than of *T. b. brucei*. A possible future experiment could involve infecting with *T. congolense* on day 0, and *T. b. brucei* on day 4, to see whether the *T. b. brucei* population is able to expand in the context of a peak *T. congolense* population. Such an experiment may be more comparable to the conditioned medium experiments performed. These initial co-infection experiments were performed with *T. congolense* IL3000, and revealed that the peak parasitaemia reached by this strain was limited by the presence of co-infecting *T. b. brucei*. Future experiments could combine infection with *T. b. brucei* with various *T. congolense* strains of different virulences, to observe whether these strains behaved differently in the presence of *T. b. brucei*. The inhibition of *T. congolense* IL3000 parasitaemia coincided with the point of *T. b. brucei* parasitaemia where SIF production would be expected to be high, thus it would be interesting to test whether more virulent *T. congolense* strains were less responsive to the *T. b. brucei* effect.

Additionally, it should be considered that in these co-infection studies, *T. congolense* in a single species infection did not appear to accumulate in 1K1N at the peak of parasitaemia, as had been observed previously. Therefore, it is possible that the *T. congolense* parasites in this experiment had lost an aspect of growth control, and this may have had an impact on any potential interaction with *T. b. brucei*.

Finally, as a *T. b. brucei* Hyp2 RNAi cell line was able to show a degree of resistance to *T. congolense* conditioned medium, this cell line could be used in co-infection with *T. congolense* to determine whether Hyp2 RNAi could enable *T. b. brucei* to resist any effects of *T. congolense*. However, this experiment would only
be of relevance if an effect of *T. congolense* on *T. b. brucei* differentiation in the bloodstream were first observed.

Overall, these analyses are supportive of potential interactions between *T. congolense* and *T. b. brucei* populations, which may have consequences for virulence and transmission in co-infections.
Chapter 6: Discussion
No parasite lives in isolation. Firstly, all parasites must be able to detect and respond appropriately to the diverse environments they find themselves in. This includes the ability to adapt their metabolism to differing nutrient availabilities and the ability to counteract or evade host defences in order to ensure establishment and maintenance of an infection. Secondly, single-celled parasites of the same species frequently depend on quorum sensing to coordinate population-wide responses that promote the longevity of parasites of the same genotype. For instance, the density-dependent signalling that drives differentiation of *T. b. brucei* slender bloodstream forms to the growth-arrested and transmissible stumpy forms promotes both chronicity of infection and transmission, favouring survival of the parasite population. Thirdly, the nutrient rich environment of the mammalian bloodstream can represent a complex ecosystem. For example, co-infections comprising multiple trypanosome species, including *T. b. brucei, T. congolense* or *T. vivax*, are not uncommon. Despite belonging to different species, these co-infecting trypanosomes share a common goal; to permit host survival long enough to allow transmission of their kin to the next host. Each species or strain needs to ensure its own survival over that of its competitors.

This thesis aimed to establish whether African trypanosomes of species other than *T. b. brucei* had comparable capabilities for growth control. Furthermore, the capacity of African trypanosomes of different species to detect and respond to each other’s signals was investigated (Figure 6.1). The potential for communication between different species of African trypanosomes was explored in a number of ways. Firstly, the presence of differential forms of *T. congolense* during bloodstream infection was investigated. The *T. b. brucei* quorum sensing mechanism that drives differentiation to the stumpy form has important ramifications for the parasites’ survival in the host and transmission. Therefore, initial investigations examined the likelihood of a comparable parasite-intrinsic mechanism of growth control during *T. congolense* infections. Secondly, it was considered whether *T. congolense* had the capacity to respond to the unidentified signalling molecule(s) of *T. b. brucei*: SIF. *T. congolense* parasites would be readily exposed to SIF when in co-infection with *T. b. brucei*. Thirdly, the ability of *T. congolense* to interfere with *T. b. brucei* growth and differentiation was explored.
Figure 6.1 An overview of the questions addressed in this thesis: Do *T. b. brucei* and *T. congolense* share quorum sensing signal responses? Do *T. congolense* parasites share components of the *T. b. brucei* SIF-responsive pathway, and if so can they respond to SIF? Do *T. congolense* parasites produce a SIF-like molecule of their own, and if so how does this impact on *T. b. brucei*?

In order for *T. congolense* parasites to be responsive to the SIF signal from *T. b. brucei*, they should not only have the SIF-response pathway components necessary to perceive the signal, but they should also be able to enter a quiescent state comparable to that of the *T. b. brucei* stumpy forms whose differentiation is triggered by SIF. A number of convincing orthologues for the *T. b. brucei* SIF-responsive genes (Mony et al., 2014) were identified in *T. congolense* and *T. vivax*, but focus was placed on the so-called hypothetical protein 2 (TbHyp2, Tb927.9.4080), the orthologues of which had good reciprocal BLASTP scores. TbHyp2 was predicted to be a post-transcriptional regulator (Erben et al., 2014). Several of the interaction motifs predicted for TbHyp2 (by the eukaryotic linear motifs database (Dinkel et al., 2012)) were indicative of links with cell cycle regulation and quiescence pathways, as well as with other members of the SIF-response pathway. Many of these interactions were also predicted for the *T. congolense* and *T. vivax* orthologues of TbHyp2 (TcHyp2 and TvHyp2 respectively). As a follow up to this bioinformatics analysis, future work could include pull-down studies with the Hyp2 orthologues to obtain experimental evidence as to whether these proteins share interaction partners and so likely have functional similarity.

Experimental evidence for a conserved link to quiescence pathways was obtained by overexpression of the Hyp2 orthologues in *T. b. brucei* with accompanying analysis
of the effect on cell proliferation. It was demonstrated *in vitro* that overexpression of TbHyp2, TcHyp2 or TvHyp2 resulted in growth arrest in the 1K1N configuration. This result, alongside the increased proportion of parasites in the 1K1N configuration observed at peak *T. congoense* parasitaemia, was consistent with conservation of a cellular quiescence pathway in *T. congoense*.

Technical replicates were used to assess the reproducibility of measurements within the *in vitro* experiments described above. It will be of interest to undertake independent experiments with additional clones to investigate the reproducibility of the key biological effects that have been observed in this, and other *in vitro* experiments reported in this thesis.

The overexpressed TbHyp2 orthologues were N-terminally tagged with a small (10 amino acid) TY motif. Detection of this motif with the BB2 antibody (Bastin et al., 1996) allowed the localisation of the overexpressed proteins to be compared. All were found to have a similarly punctate distribution in *T. b. brucei*, whether a high or lower level of overexpression was induced. This localisation could reflect a protein distributed in the endosomal-sorting pathway. Recently, a study investigating the transfer of materials between parasites via extracellular vesicles identified trafficking of vesicle-transferred materials through the endosomal pathway (Szempruch et al., 2016). Localisation to the endosomal pathway could place components of the SIF-response pathway in an ideal position for inter-parasite signalling. Co-localisation studies with concanavalin A (a marker for the endosomal pathway) could establish whether or not Hyp2 orthologues localise to the endosomal pathway.

The strongest support for a functional SIF-response pathway in *T. congoense* was provided by overexpression of the TcHyp2 orthologue in a TbHyp2 null mutant of *T. b. brucei in vivo*. Growth arrest in the 1K1N configuration, expression of the stumpy-specific marker PAD1, and the capacity for onward differentiation to procyclic forms following exposure to cis-aconitate were all used alongside morphology as indicators of the proportion of *T. b. brucei* cells in an *in vivo* population that were stumpy. Thus, when TcHyp2 overexpression was not induced, the TbHyp2 null mutant reached high parasitaemia without indication of differentiation to stumpy forms. However, when TcHyp2 overexpression was induced, growth limitation was
accompanied by differentiation to stumpy forms. Contrary to expectation it was not possible to observe a complete rescue of stumpy formation in the TbHyp2 null mutant when a TbHyp2 rescue copy was overexpressed. Here, difficulty in controlling the levels of the overexpressed protein in vivo may have contributed to the lack of a synchronous differentiation to stumpy forms. It is also possible that TbHyp2, and perhaps other components of the SIF-response pathway, are under tight regulation, such that when differentiation to stumpy forms has been triggered the protein is down regulated. This could account for the low or undetectable levels of the Hyp2 orthologues several days post-induction of overexpression in vivo.

It appeared that TvHyp2 overexpression was also able to partially restore stumpy formation in the TbHyp2 null mutant in vivo, although this was complicated by leaky expression in the uninduced cell line. This indicates that there may also be conservation of the SIF-response pathway in T. vivax. Further investigations of the potential for T. vivax interactions with other African trypanosomes in the bloodstream would be of interest, considering that this parasite is the most divergent of the African trypanosomes, and undergoes a different pathway of development in the tsetse fly vector. While T. b. brucei and T. congolense parasites must be prepared to quickly adapt to survival in the tsetse fly midgut on transmission from the bloodstream, T. vivax parasites restrict their development in the fly to the mouthparts, so T. vivax bloodstream forms must be prepared to adapt to a different environment. Additionally, T. vivax can be transmitted mechanically by biting insects other than tsetse flies (Osório et al., 2008). Therefore, while an ability to respond to extracellular signals released from competitor parasites in the bloodstream could be considered useful for T. vivax survival in the mammalian host, differences might be expected in how T. vivax would respond to such signals, given possible differences in their transmission prerequisites.

Another similarity between Hyp2 orthologues was the effect of their overexpression in increasing the kinetoplast-posterior distance in T. b. brucei as a result of increased microtubule extension at the posterior end of the cell. This elongation of the posterior end is typical for T. b. brucei cells differentiating to procyclic forms (Matthews and Gull, 1997). However, overexpression of the Hyp2 orthologues appeared to be
driving premature differentiation only as far as stumpy forms and not to procyclic forms, as evidenced by an increase in parasites arrested in 1K1N and expressing PAD1 when TcHyp2 was overexpressed in the context of a TbHyp2 null mutant. These parasites did not express EP Procyclin in the bloodstream, and only expressed this procyclic form marker on induction of differentiation in vitro by cis-aconitate. Evidence from proteomic studies found that TbHyp2 protein abundance did not differ in stumpy forms, procyclic forms or during differentiation to procyclic forms relative to the abundance in slender forms (Dejung et al., 2016; Domingo-Sananes et al., 2015). Nonetheless, we have found that overexpression of TbHyp2, TcHyp2 or TvHyp2 was able to induce differentiation of some T. b. brucei parasites to the stumpy form at lower densities than normally required for differentiation in a pleomorphic infection. Thus, by exposing a cell to high levels of Hyp2, differentiation to the stumpy form could be triggered without the usual requirement for high cell density.

Considering that many transcripts of increased abundance in T. b. brucei stumpy forms relative to slender forms are indicative of preadaptation for transmission to the tsetse fly vector, transcript abundance at peak relative to ascending T. congolense parasitaemia was investigated to see whether transcript abundance supported an accumulation of a transmissible form at peak parasitaemia.

Changes in transcript abundance from T. congolense ascending parasitaemia to peak parasitaemia were compared to changes in abundance observed between T. b. brucei slender and stumpy forms. Consistent with the lesser extent of growth arrest at peak parasitaemia observed for this strain of T. congolense relative to T. b. brucei, the transcriptomic changes indicative of differentiation to a quiescent form were more extensive in the T. b. brucei data set than the T. congolense one. For example, in stumpy forms paraflagellar rod proteins and histones were consistently down regulated, but reduced abundance of these protein categories were less consistently observed at peak T. congolense parasitaemia.

Transcripts for variant surface glycoproteins (VSGs), an important group of proteins in the parasite-host interaction, were found to have increased abundance at T. congolense peak parasitaemia, but not in T. b. brucei stumpy forms. A large number
of VSG transcripts were up regulated at peak *T. congolense* parasitaemia. This would be consistent with the need to generate new antigenic variants before parasites expressing the dominant variant were cleared by the host immune response. This strategy for parasite survival also occurs in *T. b. brucei* infections (Mugnier et al., 2015). Thus an increase in abundance of multiple VSG transcripts would also have been expected at peak *T. b. brucei* parasitaemia. The difference between the *T. congolense* and *T. b. brucei* data is likely to be a result of the immune status of the host. The mice used to generate the *T. b. brucei* transcriptomic data had been pretreated with the immunosuppressant cyclophosphamide, while the mice used to generate the *T. congolense* transcriptomic data had an intact immune system. It is possible that if the transcriptomics analysis for *T. congolense* parasites had been carried out in immunosuppressed mice, the removal of the stress of the host immune response could have allowed greater parasite investment in transmissible forms and more extensive changes in transcripts considered pre-adaptive for transmission at peak parasitaemia.

One key similarity in the type of transcripts that changed abundance from ascending to peak parasitaemia in *T. b. brucei* and *T. congolense* was the predominance of transcripts predicted to encode cell surface proteins. In the case of *T. b. brucei*, transcripts encoding EP Procyclins were increased in abundance in stumpy forms, while transcripts encoding proteins belonging to a *T. congolense*–specific cell surface family differed in abundance between peak and ascending parasitaemia. Consequently, like *T. b. brucei*, *T. congolense* may be preparing to rapidly change its surface architecture in response to transmission to a new host environment by holding transcripts required for the next life cycle stage in readiness. In a proteomic study that compared *T. congolense* IL3000 bloodstream forms with procyclic, epimastigote and metacyclic forms derived in vitro, a number of cell surface proteins were found to be differentially regulated through the *T. congolense* life cycle (Eyford et al., 2011). Likewise, surface proteins have been shown to change abundance early during the differentiation from *T. b. brucei* bloodstream to procyclic forms (Dejung et al., 2016; Domingo-Sananes et al., 2015; Shimogawa et al., 2015).
Despite some similarities in the types of transcripts increased in abundance at peak parasitaemia, when *T. congolense* and *T. b. brucei* encoded proteins from the transcriptome analysis were placed into orthologous groups, the orthologous groups increased in abundance at peak *T. congolense* parasitaemia were not the same as those increased in abundance in *T. b. brucei* stumpy forms. These findings are indicative of divergent strategies evolved by *T. congolense* and *T. b. brucei* to ensure transmission. In the future, the generation of a transcriptomic data set for *T. vivax* ascending and peak parasitaemia, could make a useful comparison to the *T. b. brucei* and *T. congolense* data, due to the different challenges faced by each parasite on uptake by the vector.

The results of these investigations highlight the potential for a parasite-intrinsic mechanism of growth control in *T. congolense* that appears similar to that of *T. b. brucei* bloodstream forms in a number of ways. Nevertheless, different species of parasites may be expected to have developed different strategies for bloodstream survival given their different life styles in the host. For example, *T. congolense* bloodstream forms have a tendency to adhere to host blood vessels and red blood cells (Banks, 1978, 1979), a strategy not employed by *T. b. brucei*. *T. b. brucei* instead has a tendency to invade tissues (including the brain parenchyma and adipose tissue (Trindade et al., 2016)) whereas *T. congolense* is restricted to blood vessels (Masocha et al., 2007). In addition, the absence of a morphologically stumpy form in *T. congolense* already suggested that differences between *T. congolense* and *T. b. brucei* development in the bloodstream might be expected. Indeed, when the results presented in this thesis are considered in detail, clear differences in the manner in which *T. b. brucei* and *T. congolense* parasites responded to high density are revealed.

Vickerman (Vickerman, 1965) had found indications of an active mitochondrion in *T. b. brucei* stumpy forms, but not slender forms, in contrast he had found indications of an active mitochondrion in all bloodstream form cells of *T. congolense*. It is possible that the transcript abundance changes between *T. congolense* ascending and peak parasitaemia were not extensive because transmissible forms were present in both the ascending and peak populations. After the first peak of parasitaemia, it has
been demonstrated that *T. b. brucei* maintains a high level of transmissible forms during chronic infection (MacGregor et al., 2011). Therefore, both *T. congolense* and *T. b. brucei* may share a strategy that maintains transmissible forms throughout infection, which may be a requirement due to the poor vectoral capacity of the tsetse fly. Furthermore, *T. congolense* is more infective to tsetse than *T. b. brucei* (ILRAD, 1981), and this may mean that the considerable investment in transmissible forms observed at peak *T. b. brucei* parasitaemia is not necessary to ensure the transmission of *T. congolense*.

An alternative explanation for the less extensive changes observed between *T. congolense* ascending and peak parasitaemia could be that the *T. congolense* IL3000 strain used in this study has developed some monomorphic traits as a result of serial passage in rodents without transmission through the tsetse fly. This is a phenomenon commonly observed with *T. b. brucei*. Monomorphism in the *T. congolense* strain used could also explain why these parasites were not inhibited by their own conditioned medium, despite *T. congolense* conditioned medium having an inhibitory effect on pleomorphic *T. b. brucei*. This is the case for *T. b. brucei* monomorphs, which produce conditioned medium containing SIF, but no longer respond to the inhibitory effects of conditioned medium themselves (Vassella et al., 1997a).

Various lines of evidence generated *in vitro* point to production of a molecule(s) by *T. congolense* that is able to inhibit growth of *T. b. brucei* in a manner resembling the action of SIF. Treatment of pleomorphic *T. b. brucei* (AnTat 1.1 90.13) with either *T. b. brucei* or *T. congolense* conditioned medium resulted in an increased proportion of cells in the 1K1N configuration consistent with the action of a SIF-like molecule. Furthermore, both *T. congolense* and *T. b. brucei* conditioned media had a greater effect on the growth of pleomorphic *T. b. brucei* AnTat 1.1 90.13 than on various monomorphic strains derived from Lister 427. Additionally, neither conditioned medium had a considerable effect on the *T. congolense* IL3000 strain used. Despite having the ability to induce comparable responses in *T. b. brucei*, a higher density of *T. congolense* culture (~7x10⁶/ml) was required to generate an effective conditioned medium than the density required to generate an effective *T. b. brucei* conditioned medium (2-3x10⁶/ml). One possible interpretation of this discrepancy would be that
*T. congolense* parasites produce less of a density-sensing molecule per cell, or that the threshold at which such a molecule is produced is at a higher parasite density.

*In vitro*, *T. b. brucei* parasites do not appear morphologically stumpy at high density in liquid culture. However, the accumulation of cells in the 1K1N configuration and expression of a stumpy reporter (under the control of the PAD1 3’UTR) were used as surrogate markers to report on the degree of stumpy formation in culture. The growth inhibitory effects elicited by both *T. congolense* and *T. b. brucei* conditioned medium were linked to an increased differentiation to the stumpy form *in vitro*, as indicated by increased expression of a CAT reporter under the control of the PAD1 3’UTR by conditioned medium-treated *T. b. brucei*. The response to conditioned medium was also linked to the SIF-response pathway, as knock down of TbHyp2 reduced the growth effect of *T. b. brucei* conditioned medium on AnTat 1.1 90.13 cells and prevented the expression of the stumpy reporter in response to conditioned medium treatment. In addition, the accumulation of cells in the 1K1N configuration, which is usually observed following conditioned medium treatment, was not observed. In contrast, TbHyp2 RNAi only delayed the induction of stumpy formation in response to 8-CPT-cAMP or *T. congolense* conditioned medium treatment. These results implicate a role for the SIF-response protein TbHyp2 in the response to *T. congolense* conditioned medium, but suggest that alternative pathways for the action of 8-CPT-cAMP and *T. congolense* conditioned medium may exist that bypass action through Hyp2 and permit a response with delayed kinetics.

Further evidence for differences in pathways to perceive the *T. b. brucei* and *T. congolense* conditioned medium signals is the effect of these two treatments on a monomorphic reporter line of *T. b. brucei*. While a Lister 427 monomorphic reporter line with a CAT reporter under the control of the PAD1 3’UTR does not respond to *T. b. brucei* conditioned medium with increased CAT expression, treatment with *T. congolense* conditioned medium was able to trigger a significant increase in CAT reporter expression. This effect is similar to the effect of the cell permeable cAMP analogue, 8-CPT-cAMP, indicating that the pathway of action of this molecule may overlap with that of *T. congolense* conditioned medium. Monomorphic cell lines have lost the ability to respond to *T. b. brucei* SIF during multiple passages in the
laboratory without cyclical development in the fly. These monomorphic parasites would not have been exposed to a *T. congoense* density sensing molecule during laboratory passage, and the observation that they have retained sensitivity to a *T. congoense* produced molecule, but not to a *T. b. brucei* one, suggests that the *T. congoense* signal is not the same as SIF. The response of monomorphic *T. b. brucei* to *T. congoense* conditioned medium could be applied in future to identify the signalling components necessary for the response by treatment of a monomorphic RNAi library with *T. congoense* conditioned medium and identification of the RNAi targets in the resistant population.

Co-infection experiments where *T. congoense* infections were initiated one day before superinfection with *T. b. brucei* resulted in a reduction in *T. b. brucei* parasitaemia relative to in a single species infection. Nevertheless, these initial studies were unable to reproduce the effect of *T. congoense* on *T. b. brucei* differentiation that had been observed in the *in vitro* experiments. However, the infection schedule used in these *in vivo* studies was not designed to optimally reflect the *in vitro* experiments. In the co-infection studies, *T. congoense* parasitaemia is in an ascending phase when superinfection with *T. b. brucei* is carried out. In contrast, the *T. congoense* culture from which conditioned medium is derived has reached a peak of high density, before use of the conditioned medium in treatment of low-density *T. b. brucei* cultures. Therefore, it may be that the *T. congoense* parasitaemia peak must occur when the *T. b. brucei* parasitaemia is sufficiently low, in order to observe an effect on *T. b. brucei* differentiation to stumpy forms relative to a single species infection. Thus, future investigations could allow establishment of *T. congoense* infection, followed 4 days later with *T. b. brucei* superinfection, such that peak *T. congoense* parasitaemia coincides with a low density *T. b. brucei* population. This experiment would more closely parallel the conditioned medium experiments but would only be possible for *T. congoense* strains of low virulence. Another difference between the *in vitro* and *in vivo* experiments was the source of the IL3000 strain used. Thus, testing the *T. congoense* cells used *in vitro* in a co-infection could provide a better comparison. Further co-infection studies will be essential in determining the physiological relevance of the *T. congoense* effect on *T. b. brucei* observed *in vitro.*
The competitive suppression of *T. b. brucei* parasitaemia observed during these co-infection studies with *T. congolense* was compatible with previous reports of interference in establishment of a superinfecting parasite of a different species or strain to the already established parasite population (Dwinger et al., 1989; Morrison et al., 1982). Nevertheless, without evidence of an increased proportion of cells in the 1K1N configuration or a greater proportion of stumpy cells in the *T. b. brucei* population present in the co-infection, the interference observed cannot be linked to a *T. congolense*-derived differentiation trigger. Additionally, during the co-infection studies, superinfection with *T. b. brucei* was shown to cause *T. congolense* parasitaemia to peak at a lower level, but once again this was not associated with an increased proportion of parasites in the 1K1N configuration. In these co-infection experiments, the *T. congolense* cells used had begun to show less sign of 1K1N arrest at peak parasitaemia in single-species control infections than had been previously observed. If these *T. congolense* parasites were beginning to show monomorphic characteristics, this could have had an impact on the response to co-infection. Therefore, returning to an earlier pass of this *T. congolense* strain may yield different results. Additionally, co-infection of pleomorphic *T. b. brucei* with *T. congolense* parasites of recent field origins, and also different virulence, may help to better address the potential for *T. congolense* to interfere with *T. b. brucei* differentiation in the bloodstream. Finally, if an effect of *T. congolense* co-infection on *T. b. brucei* differentiation were observed, the *T. b. brucei* Hyp2 RNAi cell line could be used to test whether this action was mediated through the SIF-response pathway as indicated by the *in vitro* data.

In summary, a number of the results presented in this thesis support potential interactions between African trypanosome species (Figure 6.2).

*T. congolense* parasites from ascending and peak parasitaemia were characterized by examination of their KN configuration, as well as RNA-seq analysis. The identification of a *T. congolense* form at peak parasitaemia that shared characteristics with the stumpy form of *T. b. brucei* would have been consistent with a shared model of growth control. While differences between *T. congolense* parasites from ascending and peak parasitaemia were identified in this thesis, the changes were not as
extensive as those accompanying the T. b. brucei slender to stumpy transition. This highlights the potentially different strategies employed by the two species in the bloodstream to ensure survival and transmission. Nevertheless, the increased proportion of T. congolense parasites in the 1K1N configuration that was observed at peak parasitaemia supports an element of growth control, such that an ability to respond to a parasite-derived density signal remains a possibility.

While differentiation to the stumpy form of T. b. brucei can be monitored by use of various markers, such as PAD1, our previous analyses of T. congolense infection had indicated that identification of stumpy-equivalent forms in this parasite would not be straightforward. Therefore, a different strategy was applied to investigate whether T. congolense had the tools necessary to receive the T. b. brucei differentiation signal. It was possible to identify orthologues for many of the components of the T. b. brucei SIF-response pathway (Mony et al., 2014) in T. congolense and T. vivax by BLASTP. Overexpression of the T. congolense orthologue for the SIF-response pathway component Tb927.9.4080 (TbHyp2) was able to restore the capacity for differentiation to stumpy forms in an otherwise SIF-resistant T. b. brucei TbHyp2 null mutant in vivo. The conservation of a functional component of the SIF–response pathway in the T. congolense genome supports a model by which T. congolense parasites may detect and respond to the T. b. brucei density signal in a co-infection.

Finally, it was demonstrated that treatment of T. b. brucei with conditioned medium from a dense culture of T. congolense was able to cause increased growth arrest and expression of a stumpy form marker in vitro. This effect was found to be delayed by RNAi targeting TbHyp2, indicating an intersection with the SIF-signalling pathway in the response to T. congolense conditioned medium. These results suggest that T. congolense may be able to modify the growth dynamics of T. b. brucei in a co-infection. Further investigation of co-infections, using recently isolated field strains, will help to unravel the physiological relevance of such interactions in vivo.

The response of Plasmodium parasites to co-infection with genetically diverse strains has been investigated in multiple studies. Some studies have indicated that co-infection would favour the selection of more virulent parasites (Bell et al., 2006; Wargo et al., 2007). However, it has also been suggested that at high levels of stress,
if survival in the host is at risk, then parasites may switch to a strategy of terminal investment in transmissible stages to ensure their long-term survival. The threshold at which different parasite species switch to such a terminal investment strategy may be linked to the life cycle characteristics of the parasite; for instance *P. falciparum*, which has a longer cell cycle and whose gametocytes require longer for maturation than mouse-infective species of *Plasmodium* may switch to investment in transmission at lower levels of stress (Carter et al., 2013). Thus, parasite replication strategies in the context of a co-infection may vary depending on the level of stress exerted by the competition, and the respective parasite species’ tolerance to stress. One study has demonstrated that *P. malariae* co-infection could increase *P. falciparum* gametocyte production (Bousema et al., 2008). Therefore, evidence from *Plasmodium* studies points to impacts of co-infection on both parasite virulence and transmission. Likewise, there have been descriptions of reduced virulence in trypanosome co-infections, for instance co-infection with *T. vivax* was shown to cause less of an effect on packed cell volume in horses or donkeys than infection with *T. congolense* alone (Pinchbeck et al., 2008). The results presented in this thesis have indicated that *T. b. brucei* development in the host and investment in the transmissible, stumpy form could also be affected by interactions with competitors. As in the case of *Plasmodium* responses to co-infection (Reece et al., 2009), the responses of African trypanosomes may be characterised by plasticity depending on the nature of the host and the competitors.

The potential for similarities in transmission strategies between malaria parasites and African trypanosomes had been proposed previously (Pollitt et al., 2011). Therefore, as in the case of *Plasmodium* infection (Reece et al., 2009), it is important to understand how trypanosomes respond to environmental cues, for example in co-infection or drug treatment, in order to develop therapeutics that force parasites into a suboptimal development strategy, or to predict how parasites may respond to treatments.

Interactions between *T. b. brucei*, *T. congolense* and *T. vivax* during mixed infections could result in the component species placing more investment in replication than growth control and transmission, in order to better compete with co-infecting species.
for the host’s resources. This could result in selection of more virulent parasites, which when removed from the competitive environment may inflict greater pathology on a subsequent host. Alternatively, if African trypanosomes of different species are able to mutually respond to each other’s growth limiting signals, this could increase transmission but may play a role in limiting the pathogenicity of mixed infections. If this were to be the case, removal of one of the component species with a species-specific drug treatment, could result in resurgence of the remaining species. Development of a therapeutic SIF-mimic to drive *T. b. brucei* bloodstream forms prematurely stumpy could allow clearance by the host immune system; the applicability of such a therapeutic to *T. congolense* and *T. vivax* infections would depend on the conservation of the signal perception pathway.

The existence of quorum sensing systems in bacteria has led to the evolution of cheats (Diggle et al., 2007). There are cells that are signal-blind and will not respond to a signal or cells that do not produce an energetically costly signal molecule. In trypanosome infections, it could be envisaged that one trypanosome species (e.g. a *T. congolense* strain) could become more successful in competing for host resources in a co-infection by reducing its response to a density-sensing molecule produced by another co-habiting species (e.g. pleomorphic *T. b. brucei*). Thus, one might expect to find strains or species that have become desensitised to inter-species density control in nature.

One study reported that co-infection between a less virulent and extremely virulent *T. congolense* strain reduced the pathogenicity associated with infection as compared to an infection with the virulent strain alone (Masumu et al., 2009). In another study that investigated co-infection between a virulent strain of *T. b. brucei* with an avirulent strain (Balmer et al., 2009) the total parasitaemia in the mixed infection was found to be reduced relative to a single infection with the virulent species alone, with an associated reduction in the deleterious effects on the host. It was thus suggested that the inhibition experienced by both strains was the result of active interference rather than a consequence of reaching maximum population size capacity. It was postulated that the mutual inhibition observed could be the result of parasite-released factors having a direct effect on competitors (allelopathy) or that the interference
observed might be immune mediated. However, in this case no effect on differentiation was observed as the majority of parasites were described as slender during the infections. The lack of differentiation in any of the parasite strains tested in this study even at high parasitaemia indicates that the virulent strain may have become monomorphic. Nevertheless, the co-infection results presented in this thesis also found evidence of mutual suppression that was not linked to differentiation, despite the use of a pleomorphic strain of *T. b. brucei*. This indicates that other factors such as allelopathic interference or immune mediated effects may be playing a role in mediating interspecies interactions between *T. b. brucei* and *T. congolense*.

Anaemia is a common feature of the pathology resulting from infection with different species of African trypanosomes. Nevertheless, co-infection with two species of trypanosomes does not necessarily result in a cumulative effect on pathological indicators such as PCV (Pinchbeck et al., 2008). This could relate to the interference phenomenon observed by which an established species of trypanosome may effect the growth of a superinfecting species (Dwinger et al., 1989). Pathogenic factors released by African trypanosomes into the bloodstream (Authie, 1994; Coustou et al., 2012; Guegan et al., 2013) may become more abundant as parasitaemia increases, and thus these factors could be reduced by the interference in parasite growth occasionally observed in co-infections. Differentiation from the replicative slender form to the growth-arrested stumpy form is an important component of growth control and thus virulence limitation in *T. b. brucei* infections (MacGregor et al., 2012). Therefore, the impact of a co-infecting species on differentiation would be expected to have an impact on host pathology. *T. b. brucei* parasites have a capacity to invade the adipose tissue and it has been suggested that this could contribute to wasting pathology (Trindade et al., 2016). The ability of the *T. b. brucei* differentiation signal to accumulate at this site has not yet been determined. Likewise, the ability of *T. congolense* parasites to invade the adipose tissue has not been assessed. The ability of different parasites to exchange signals with possible effects on parasite differentiation status and host pathology could be impacted by their ability to occupy shared compartments.
Interference between two species in a shared environment is found in many diverse ecosystems in nature. Eukaryotic species may develop alternative strategies to deal with competitors, other than simply increasing their own rate of growth and depletion of available resources. For example, certain allelopathic strains of rice have been shown to inhibit growth of the weed barnyardgrass (He et al., 2012). This interference was demonstrated to be the result of products released from rice in co-culture with barnyardgrass as the residual solution from this mixed culture was found to inhibit the growth of barnyardgrass seedlings. The nutrients in the residual solution had been returned to normal levels before the test, so that it was unlikely that the effect was caused by nutrient depletion. There are parallels between this plant interaction and the effect of *T. congolense* conditioned medium on *T. b. brucei*. Thus, studies on interactions between species in one system can inform studies on possible inter-species interactions in evolutionarily divergent systems. Another example is found in prokaryotes; *Escherichia coli* that produce an antibacterial toxin (colicin) are able to kill neighbouring sensitive *E. coli* in order to get a better share of available nutrients. This strategy is more effective in the structured environment of a soft agar matrix than in liquid culture as the colicin-producing bacteria are then able to increase the nutrients available to them locally (Chao and Levin, 1981). If *T. congolense* parasites were releasing a molecule to interfere with *T. b. brucei* growth, such a molecule could be particularly effective locally as *T. congolense* parasites are known to adhere to host blood vessels (Banks, 1978). It has additionally been proposed that bacterial released extracellular death factors that cause interspecies death could be used to develop novel antimicrobials (Kumar and Engelberg-Kulka, 2014). Thus understanding mechanisms of intra- and interspecies signalling in diverse organisms can have useful implications for agriculture and veterinary and human health.

In conclusion, the hitherto unexpected capacity for cross-talk between different African trypanosome species in a co-infection indicated by this thesis, has implications for the virulence, transmission and evolution of these devastating parasites of animals and humans. Moreover, understanding the mechanisms of inter-species interactions has a broad relevance across many diverse species, many of which are also of medical or veterinary importance.
Figure 6.2 Evidence in support of cross-talk between two different species of African trypanosome. The possible intra- and inter-species interactions investigated in this thesis are represented diagrammatically, and potential interactions were numbered 1-3. The evidence for and against these proposed interactions are summarised in the numbered tables below the diagram.
Bibliography


Bengaly, Z., Sidibe, I., Ganaba, R., Desquesnes, M., Boly, H., and Sawadogo, L. (2002b). Comparative pathogenicity of three genetically distinct types of


Matthews, K.R., and Gull, K. (1997). Commitment to differentiation and cell cycle re-entry are coincident but separable events in the transformation of African trypanosomes from their bloodstream to their insect form. Journal of Cell Science 110 (Pt 20), 2609-2618.


procyclic stage of African trypanosomes. Molecular and Biochemical Parasitology 31, 203-216.


Appendix 1: Supplementary information on *T. congoense* and *T. b. brucei* RNA-seq analysis

Appendix 1.1 – Orthologues of PAD1 and PAD2 in *T. congoense* and *T. vivax*

Appendix 1.2 – Additional information for *T. congoense* transcripts that were at least 4-fold more abundant at peak relative to ascending parasitaemia

Appendix 1.3 – Additional information for the 8 *T. congoense* transcripts with the largest reduction in abundance at peak relative to ascending parasitaemia

Appendix 1.4 – *T. congoense* transcripts that were of reduced abundance at peak relative to ascending parasitaemia for which the protein product was classified as other

Appendix 1.5 - *T. congoense* transcripts that were of increased abundance at peak relative to ascending parasitaemia for which the protein product was classified as other

Appendix 1.6 – Additional information for *T. b. brucei* transcripts that were at least 4-fold less abundant in stumpy relative to slender parasites for which the protein product was described as hypothetical protein

Appendix 1.7 - Trypanosome orthologue clusters containing members for which transcripts demonstrated pronounced reduction in abundance in *T. b. brucei* stumpy relative to slender forms
Appendix 1.1 *T. b. brucei* PAD1 and PAD2 amino acid sequences were used in a BLASTP search against both *T. vivax* and *T. congolense* proteins. There was not a close PAD1 orthologue found in *T. vivax*, but there was in *T. congolense*. This was consistent with the different life cycle stages of these African trypanosome species. *T. congolense*, like *T. b. brucei*, differentiates to procyclic forms in the midgut of the tsetse fly, whereas *T. vivax* development is limited to the mouthparts of the fly.

<table>
<thead>
<tr>
<th>Gene function</th>
<th><em>T. brucei</em> ID</th>
<th><em>T. vivax</em> ID</th>
<th>T. vivax hit 1</th>
<th>Result from input of <em>T. vivax</em> hit 1 into BLAST (E value, hit number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAD1</td>
<td>Tb927.7.5930</td>
<td>TVY486_1117600 (these values are for the largest (aa 184-589 of query sequence) of 3 sections of sequence similarity)</td>
<td>2.30E-89</td>
<td>39% 56% Tb927.7.5930 (1.9e-89, 5)</td>
</tr>
<tr>
<td>PAD2</td>
<td>Tb927.7.5940</td>
<td>TVY486_1117600 (these values are for the largest (aa 184-589 of query sequence) of 3 sections of sequence similarity)</td>
<td>4.20E-90</td>
<td>39% 56% Tb927.7.5940 (3.5e-90, 4)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene function</th>
<th><em>T. brucei</em> ID</th>
<th><em>T. congolense</em> ID</th>
<th>T. congolense hit 1</th>
<th>Result from input of <em>T. congolense</em> hit 1 into BLAST (E value, hit number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAD1</td>
<td>Tb927.7.5930</td>
<td>TcIL3000.0.18180 (section of sequence similarity covers aa 6-579 of query sequence)</td>
<td>1.30E-239</td>
<td>75% 86% Tb927.7.5930 (9.8e-240, 1)</td>
</tr>
<tr>
<td>PAD2</td>
<td>Tb927.7.5940</td>
<td>TcIL3000.0.18180 (section of sequence similarity covers aa 6-571 of query sequence)</td>
<td>4.10E-238</td>
<td>75% 87% Tb927.7.5940 (3.0e-238, 2)</td>
</tr>
<tr>
<td>SysID</td>
<td>Updated Description</td>
<td>Top_Tb927_BLASTP (e value)</td>
<td>Top_Tb927_BLASTP_description</td>
<td>LogFC</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------</td>
<td>----------------------------</td>
<td>-------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TcIL3000_0 23590</td>
<td>hypothetical protein, conserved</td>
<td>Tb527.10 15600 (2.0)</td>
<td>PKD2-type peptidyl-prolyl cis-trans isomerase, conserved expression site-associated gene (ESAG) protein, putative</td>
<td>4.3351</td>
</tr>
<tr>
<td>TcIL3000_0 34550</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.7 3260 (7e-13)</td>
<td>variant surface glycoprotein (VSG, pseudogene), putative</td>
<td>3.6334</td>
</tr>
<tr>
<td>TcIL3000_0 41620</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.5 4870 (0.28)</td>
<td>receptor-type adenylate cyclase GRESAG 4, putative procyclin-associated gene 2 (PAG2) protein</td>
<td>3.5275</td>
</tr>
<tr>
<td>TcIL3000_0 43500</td>
<td>hypothetical protein, conserved</td>
<td>Tb11.5 0821 (0.24)</td>
<td>65 kDa invariant surface glycoprotein</td>
<td>3.4135</td>
</tr>
<tr>
<td>TcIL3000_0 13130</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>Tb927.10 10220 (2e-14)</td>
<td>procyclin-associated gene 5 (PAG5) protein</td>
<td>3.3888</td>
</tr>
<tr>
<td>TcIL3000_0 35150</td>
<td>Invariant surface glycoprotein, putative</td>
<td>Tb927.2 3280 (4e-19)</td>
<td>65 kDa invariant surface glycoprotein</td>
<td>3.3481</td>
</tr>
<tr>
<td>TcIL3000_0 01540</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.10 10230 (3e-05)</td>
<td>procyclin-associated gene 5 (PAG5) protein</td>
<td>3.3063</td>
</tr>
<tr>
<td>TcIL3000_0 05070</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.7 7200 (0.98)</td>
<td>hypothetical protein, conserved</td>
<td>3.0385</td>
</tr>
<tr>
<td>TcIL3000_0 32040</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.10 10230 (2e-05)</td>
<td>procyclin-associated gene 5 (PAG5) protein</td>
<td>2.9728</td>
</tr>
<tr>
<td>TcIL3000_0 14750</td>
<td>Trypanosomai VSG domain containing protein, putative</td>
<td>Tb927.1 4890 (7e-05)</td>
<td>expression site-associated gene 2 (ESAG2) protein, putative procyclin-associated gene 2 (PAG2) protein</td>
<td>2.9302</td>
</tr>
<tr>
<td>TcIL3000_0 01330</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>Tb927.10 10220 (3e-17)</td>
<td>procyclin-associated gene 2 (PAG2) protein</td>
<td>2.9183</td>
</tr>
</tbody>
</table>

LogFC: Logarithmic Fold Change
adj.P.Val: Adjusted P Value
Tb orthologue on geneDB: Gene DB orthologue
<table>
<thead>
<tr>
<th>TcIL3000_0_19220</th>
<th>Trypanosomal VSG domain containing protein, putative</th>
<th>Tb927.5.4610 (0.005)</th>
<th>variant surface glycoprotein (VSG, pseudogene)-related, putative</th>
<th>2.8953</th>
<th>0.0161</th>
<th>NA</th>
<th>T.con orthologues VSG</th>
</tr>
</thead>
<tbody>
<tr>
<td>TcIL3000_0_58370</td>
<td>Transferin receptor-like, ESAG6-like</td>
<td>Tb927.10.5690 (1e-12)</td>
<td>intracytoplasmic-associated gene 2 (PAG2), putative</td>
<td>2.8744</td>
<td>0.0228</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
</tr>
<tr>
<td>TcIL3000_0_37210</td>
<td>Invariant surface glycoprotein, putative (fragment)</td>
<td>Tb927.5.1390 (0.022)</td>
<td>64 kDa invariant surface glycoprotein</td>
<td>2.7233</td>
<td>0.0119</td>
<td>NA</td>
<td>Y (1390)</td>
</tr>
<tr>
<td>TcIL3000_0_60600</td>
<td>Hypothetical protein, conserved</td>
<td>Tb11.v5.0214 (2e-05)</td>
<td>retrotransposon hot spot (RHS), putative</td>
<td>2.6998</td>
<td>0.0288</td>
<td>NA</td>
<td>T.con orthologues hypothetical</td>
</tr>
<tr>
<td>TcIL3000_0_26040</td>
<td>Hypothetical protein, conserved</td>
<td>Tb927.2.830 (0.002)</td>
<td>retrotransposon hot spot protein (RHS, pseudogene), putative</td>
<td>2.6591</td>
<td>0.0133</td>
<td>NA</td>
<td>T.con orthologues hypothetical</td>
</tr>
<tr>
<td>TcIL3000_0_50230</td>
<td>Expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.10.5690 (1e-16)</td>
<td>procyclin-associated gene 2 (PAG2), putative</td>
<td>2.6259</td>
<td>0.0013</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
</tr>
<tr>
<td>TcIL3000_0_23900</td>
<td>Trypanosomal VSG domain containing protein, putative (fragment)</td>
<td>Tb927.1.05 (5e-05)</td>
<td>variant surface glycoprotein (VSG), putative</td>
<td>2.6222</td>
<td>0.0182</td>
<td>NA</td>
<td>Y (Tb05)</td>
</tr>
<tr>
<td>TcIL3000_0_24580</td>
<td>Hypothetical protein, unlikely</td>
<td>Tb11.v5.0595 (9e-05)</td>
<td>variant surface glycoprotein (VSG), putative</td>
<td>2.6138</td>
<td>0.0469</td>
<td>NA</td>
<td>T.con orthologues VSG</td>
</tr>
<tr>
<td>TcIL3000_0_44800</td>
<td>Trypanosomal VSG domain containing protein, putative (fragment)</td>
<td>Tb927.9.820 (1e-07)</td>
<td>variant surface glycoprotein (VSG, pseudogene), putative</td>
<td>2.5974</td>
<td>0.0264</td>
<td>NA</td>
<td>Y (820)</td>
</tr>
<tr>
<td>TcIL3000_0_50940</td>
<td>Transferin receptor-like, ESAG6-like (fragment)</td>
<td>Tb27.7.3250 (4e-28)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.5873</td>
<td>0.0044</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
</tr>
<tr>
<td>TcIL3000_0_43840</td>
<td>Expression site-associated gene (ESAG) protein, putative</td>
<td>Tb27.7.3260 (4e-19)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.5827</td>
<td>0.0250</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
</tr>
</tbody>
</table>

266
<table>
<thead>
<tr>
<th>GenBank ID</th>
<th>Protein Type</th>
<th>Description</th>
<th>Accession</th>
<th>E-value</th>
<th>Q-value</th>
<th>Fold Change</th>
<th>Description</th>
<th>Accession</th>
<th>E-value</th>
<th>Q-value</th>
<th>Fold Change</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.8.6880</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.8.6880 (4.6)</td>
<td>2.5708</td>
<td>0.0370</td>
<td>NA</td>
<td>N (IF-2 and normal proliferation)</td>
<td>Tb927.8.6880 (4.6)</td>
<td>2.5708</td>
<td>0.0370</td>
<td>NA</td>
<td>N (IF-2 and normal proliferation)</td>
<td></td>
</tr>
<tr>
<td>Tb11.v5.0624 &amp; Tb927.7.3250 (5e-08)</td>
<td>hypothetical protein, conserved</td>
<td>retrotransposon hot spot protein (PHS), putative</td>
<td>2.5665</td>
<td>0.0090</td>
<td>NA</td>
<td>T.con orthologues hypothetical.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb11.v5.0624 &amp; Tb927.7.3250 (2e-06)</td>
<td>hypothetical protein, conserved</td>
<td>retrotransposon hot spot protein (PHS), putative</td>
<td>2.5648</td>
<td>0.0238</td>
<td>NA</td>
<td>T.con orthologues hypothetical.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.7.3260 (5e-22)</td>
<td>ESAG6-like protein, putative</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.5150</td>
<td>0.0295</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.con)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.7.3260 (4e-08)</td>
<td>ESAG6-like protein, putative</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.4264</td>
<td>0.0332</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (3260)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.5680 (1e-15)</td>
<td>ESAG6 protein, putative</td>
<td>procyclin-associated gene 1 (PAG1) protein, putative</td>
<td>2.4232</td>
<td>0.0428</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N (3260 and PAG1)</td>
<td>Y (T.con)</td>
<td>Pb536</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.5680 (1e-07)</td>
<td>hypothetical protein, conserved</td>
<td>retrotransposon hot spot protein (PHS), putative</td>
<td>2.4150</td>
<td>0.0118</td>
<td>NA</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.7.3260 (7e-17)</td>
<td>ESAG6-like protein, putative</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.4102</td>
<td>0.0051</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>Y (3260)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.7.3260 (8e-23)</td>
<td>ESAG6-like protein, putative</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.3917</td>
<td>0.0062</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>Y (T.con)</td>
<td>Pb536</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.11.13360 (D.69)</td>
<td>hypothetical protein, conserved</td>
<td>AAA ATPase, putative</td>
<td>2.3877</td>
<td>0.0035</td>
<td>NA</td>
<td>N (13360)</td>
<td>Tb927.11.13360 (D.69)</td>
<td>2.3877</td>
<td>0.0035</td>
<td>NA</td>
<td>N (13360)</td>
<td></td>
</tr>
<tr>
<td>Tb927.11.12710 (7e-11)</td>
<td>Trypanosomal VSG domain containing protein, putative</td>
<td>variantsurfaceglycoprotein (VSG)-related, putative</td>
<td>2.3801</td>
<td>0.0235</td>
<td>NA</td>
<td>Y (12710)</td>
<td>Tb927.11.12710 (7e-11)</td>
<td>2.3801</td>
<td>0.0235</td>
<td>NA</td>
<td>Y (12710)</td>
<td></td>
</tr>
<tr>
<td>Tb927.9.16260 (7e-06)</td>
<td>Trypanosomal VSG domain containing protein, putative</td>
<td>expression site-associated gene 3 (ESAG3, pseudogene), putative</td>
<td>2.3507</td>
<td>0.0383</td>
<td>NA</td>
<td>Y (16260)</td>
<td>Tb927.9.16260 (7e-06)</td>
<td>2.3507</td>
<td>0.0383</td>
<td>NA</td>
<td>Y (16260)</td>
<td></td>
</tr>
<tr>
<td>Tb11.v5.0214 (4e-04)</td>
<td>hypothetical protein, conserved</td>
<td>retrotransposon hot spot (RHS) protein, putative</td>
<td>2.3346</td>
<td>0.0312</td>
<td>NA</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.11.17930 (2.7)</td>
<td>hypothetical protein, conserved</td>
<td>variant surface glycoprotein (VSG), degenerate and frameshifted</td>
<td>2.3200</td>
<td>0.0119</td>
<td>NA</td>
<td>T.con orthologues hypothetical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The table includes entries with hypothetical, putative, and conserved protein types, along with expression site-associated genes. The descriptions include various biological contexts such as immune evasion, metabolism, and translation. E-values and Q-values are also provided for each entry.*
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>T.b br ortho</th>
<th>T.b con orthologues</th>
<th>T. con orthologues</th>
<th>T.b con ortho</th>
<th>T.b ortho</th>
<th>T.b ortho</th>
<th>T.b ortho</th>
</tr>
</thead>
<tbody>
<tr>
<td>TcIL3000_001570</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.7.3260 (4e-14)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.3198</td>
<td>0.0447</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (3260)</td>
</tr>
<tr>
<td>TcIL3000_045120</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>Tb927.10.10220 (1e-14)</td>
<td>procyclin-associated gene 2 (PAG2) protein</td>
<td>2.3106</td>
<td>0.0133</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (PAG2)</td>
</tr>
<tr>
<td>TcIL3000_10120</td>
<td>hypothetical protein, unlikely</td>
<td>Tb927.11.5940 (1.5)</td>
<td>receptor-type adenylyl cyclase GRESAG 4, putative</td>
<td>2.3198</td>
<td>0.0063</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (5940)</td>
</tr>
<tr>
<td>TcIL3000_29730</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.4.5690 (0.12)</td>
<td>variant surface glycoprotein (VSG, pseudogene), putative</td>
<td>2.2295</td>
<td>NA</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (VSG)</td>
</tr>
<tr>
<td>TcIL3000_56150</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>Tb927.7.3260 (7e-22)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.2199</td>
<td>0.0485</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (3260)</td>
</tr>
<tr>
<td>TcIL3000_5260</td>
<td>hypothetical protein, unlikely</td>
<td>Tb927.3.4460 (0.21)</td>
<td>hypothetical protein, conserved</td>
<td>2.1863</td>
<td>0.0062</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>N (4460)</td>
</tr>
<tr>
<td>TcIL3000_22950</td>
<td>hypothetical protein, conserved (fragment)</td>
<td>Tb927.2.6380 (0.096)</td>
<td>variant surface glycoprotein (VSG, pseudogene), putative</td>
<td>2.1814</td>
<td>0.0033</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (VSG)</td>
</tr>
<tr>
<td>TcIL3000_40140</td>
<td>Trypanosomal VSG domain containing protein, putative</td>
<td>Tb927.9.7330 (0.018)</td>
<td>expression site-associated gene 2 (ESAG2) protein, putative</td>
<td>2.1788</td>
<td>0.0428</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>N (7330)</td>
</tr>
<tr>
<td>TcIL3000_60190</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.8.2370 (4.6)</td>
<td>dicer-like protein</td>
<td>2.1568</td>
<td>0.0331</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.br ortho)</td>
</tr>
<tr>
<td>TcIL3000_29290</td>
<td>invariant surface glycoprotein, putative</td>
<td>Tb927.2.3280 (3e-17)</td>
<td>65 kDa invariant surface glycoprotein</td>
<td>2.1527</td>
<td>0.0186</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.br ortho)</td>
</tr>
<tr>
<td>TcIL3000_29130</td>
<td>hypothetical protein, conserved</td>
<td>Tb11.155.0880 (3e-75)</td>
<td>UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase, putative</td>
<td>2.1329</td>
<td>0.0265</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.br ortho)</td>
</tr>
<tr>
<td>TcIL3000_43120</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>Tb927.7.3260 (7e-17)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.1297</td>
<td>0.0495</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.br ortho)</td>
</tr>
<tr>
<td>TcIL3000_18170</td>
<td>Transferrin receptor-like, ESAG6-like</td>
<td>Tb927.7.3260 (7e-13)</td>
<td>expression site-associated gene (ESAG) protein, putative</td>
<td>2.1168</td>
<td>0.0235</td>
<td>Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6)</td>
<td>N</td>
<td>Y (T.br ortho)</td>
</tr>
<tr>
<td>TcIL3000_0_13560</td>
<td>VSG-associated, congolesen-specific ORF Tb11.v5, 6655 (0.015) hypothetical protein, conserved 2.1160 0.0205 NA</td>
<td>Y (T.con) T.con orthologues hypothetical or VSG-associated, congolesen-specific ORF, 6655-trypanosome RH5 family putative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_13560</td>
<td>hypothetical protein, conserved Tb927.8, 7140 (1e-70) UDP-Gal or UDP-GlcNac-dependent glycosyltransferase, putative 2.1090 0.0312 Tb927.8.7140</td>
<td>Y (T.con) T.con orthologues hypothetical or VSG-associated, congolesen-specific ORF, 7140 - GPI-anchor biosynthetic processes, protein glycosylation, UDP-glycosyltransferases, galactosyltransferase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_52560</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.10, 5690 (1e-14) procyclin-associated gene 2 (PAG2) protein, putative 2.1019 0.0236 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) Y (PAG2) - abnormal proliferation in pcf and diff.) Y (T.con) Y (PAG2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_51720</td>
<td>hypothetical protein Tb927.8, 7170 (6.0) inositol polyphosphate 1-phosphatase, putative 2.0913 0.0400 NA N (7170) 7170 - phosphotidylinositol phosphorylation, mitochondrion, in pcf and bsf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_01560</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.10, 10220 (2e-20) procyclin-associated gene 2 (PAG2) protein 2.0785 0.0235 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG2) Y (PAG2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26370</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3260 (6e-21) procyclin-associated gene 1 (PAG1) protein, putative 2.0164 0.0312 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG1) Y (PAG1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_25680</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3260 (9e-22) procyclin-associated gene 1 (PAG1) protein, putative 2.0138 0.0209 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG1) Y (PAG1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_25680</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.7, 3260 (9e-22) procyclin-associated gene 1 (PAG1) protein, putative 2.0155 0.0415 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N immune evasion/tolerance, transferrin receptor, transport, 3250 in pcf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26370</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3260 (8e-14) procyclin-associated gene 1 (PAG1) protein, putative 2.0138 0.0209 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG1) Y (PAG1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26370</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3250 (8e-14) procyclin-associated gene 1 (PAG1) protein, putative 2.0155 0.0415 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N immune evasion/tolerance, transferrin receptor, transport, 3250 in pcf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26370</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3260 (9e-22) procyclin-associated gene 1 (PAG1) protein, putative 2.0138 0.0209 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG1) Y (PAG1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_01560</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.10, 10220 (2e-20) procyclin-associated gene 2 (PAG2) protein 2.0785 0.0235 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG2) Y (PAG2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_16180</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.10, 10220 (8e-22) procyclin-associated gene 2 (PAG2) protein 2.0951 0.0268 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N (PAG2) Y (PAG2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_51720</td>
<td>hypothetical protein Tb927.8, 7170 (6.0) mositol polyphosphate 1-phosphatase, putative 2.0913 0.0400 NA N (7170) 7170 - phosphotidylinositol phosphorylation, mitochondrion, in pcf and bsf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_00950</td>
<td>Trypanosomal VSG domain containing protein, putative Tb927.11, 14620 (8e-06) expression site-associated gene 2 (ESAG2) protein, putative 2.0725 0.0101 NA T.con orthologues trypanosomal VSG, 14620 - integral membrane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_12380</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.7, 3260 (6e-21) expression site-associated gene (ESAG) protein, putative 2.0707 0.0119 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N Y (T.con) Y (3260) immune evasion/ tolerance, transferrin receptor, transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_01560</td>
<td>Transferrin receptor-like, ESAG6-like Tb927.7, 3260 (6e-21) expression site-associated gene (ESAG) protein, putative 2.0707 0.0119 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N Y (T.con) Y (3260) immune evasion/ tolerance, transferrin receptor, transport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_29460</td>
<td>Trypanosomal VSG domain containing protein, putative Tb927.11, 14620 (8e-06) expression site-associated gene 2 (ESAG2) protein, putative 2.0725 0.0101 NA T.con orthologues trypanosomal VSG, 14620 - integral membrane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_05950</td>
<td>Trypanosomal VSG domain containing protein, putative Tb927.11, 20270 (0.004) variant surface glycoprotein (VSG, pseudogene), putative 2.0628 0.0235 NA Y (20270) T.con orthologues trypanosomal VSG, 20270 - degenerate and frameshifted.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TcIL3000_0_26370</td>
<td>expression site-associated gene (ESAG) protein, putative Tb927.7, 3250 (8e-14) expression site-associated gene (ESAG) protein, putative 2.0155 0.0415 Tb927.7.3260 (ESAG7), Tb927.7.3250 (ESAG6) N immune evasion/ tolerance, transferrin receptor, transport, 3250 in pcf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 1.2 T. congolense transcripts that were at least 4-fold more abundant at peak relative to ascending parasitaemia (excluding genes described as VSG or with no protein product assigned). The T. b. brucei orthologue for each T. congolense gene identified by BLASTP is listed alongside its product description. Additionally, any T. b. brucei orthologue information listed on GeneDB is included. If RNAi targeting a T. b. brucei orthologue resulted in reduced fitness during differentiation from bloodstream to procyclic forms (implicating a role in growth in either life cycle stage or in differentiation) in the study by Alsford et al (Alsford et al., 2011), this was highlighted in green. If either the T. congolense protein or T. b. brucei orthologue were described as cell surface expressed on GeneDB (Jackson et al., 2013) then this was highlighted in yellow.
<table>
<thead>
<tr>
<th>SysID</th>
<th>Updated Description</th>
<th>Top_Tb927.BLASTP (e value)</th>
<th>Top_Tb927.BLASTP description</th>
<th>logFC</th>
<th>adj.P. Val</th>
<th>Tb orthologue on geneDB</th>
<th>Abnormal cell proliferation in differentiated bloodstream (Tbr)</th>
<th>Cell surface phylum</th>
<th>Comments from Gene DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>TcIL3000_6_6310</td>
<td>hypothetical protein, unlikely</td>
<td>Tb927.9.16260 (0.15)</td>
<td>expression site-associated gene 3 (ESAG3, pseudogene), putative</td>
<td>-0.9713</td>
<td>0.0258</td>
<td>NA</td>
<td>Y (16260)</td>
<td>NA</td>
<td>16260 - degenerate and frameshifted</td>
</tr>
<tr>
<td>TcIL3000_3_960</td>
<td>hypothetical protein</td>
<td>Tb927.10.8920 (0.22)</td>
<td>ras-like small GTPase, putative</td>
<td>-0.911</td>
<td>0.0368</td>
<td>NA</td>
<td>Y (8920) abnormal proliferation - 6d bsf, pcf and diff.</td>
<td>NA</td>
<td>Possible signalling protein based on tertiary structure prediction. GTP binding, cytoplasm, nucleus</td>
</tr>
<tr>
<td>TcIL3000.11_16690</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.7.6310 (1.4)</td>
<td>polo-like protein kinase</td>
<td>-0.9104</td>
<td>0.0489</td>
<td>NA</td>
<td>Y(6310)</td>
<td>NA</td>
<td>T.congolense orthologues hypothetical. 6310: RNAi phenotype - no basal body duplication in pcf, abnormal mitochondrial organization in bsf and pcf, abnormal cytokinesis in bsf, no growth defect in bsf, lethal in bsf, expressed in pcf. GO: mitotic cell cycle, polo kinase, tyrosine kinase, ser/thr kinase activity, ATP binding, nucleus</td>
</tr>
<tr>
<td>TcIL3000_0_27530</td>
<td>hypothetical protein, conserved (fragment)</td>
<td>Tb927.7.3330 (5e-14)</td>
<td>hypothetical protein, conserved</td>
<td>-0.8813</td>
<td>0.0309</td>
<td>NA</td>
<td>Y(3330) abnormal proliferation bsf and diff</td>
<td>NA</td>
<td>T.congolense orthologues hypothetical. 3330 - Enriched in acidocalcisomes in T.brucei pcf, detected in bsf PM and cytoskeletal fraction. GO: transcription regulator, vesicle-mediated transport, acidocalcisome, nucleus, membrane</td>
</tr>
<tr>
<td>TcIL3000_0_35940</td>
<td>hypothetical protein, conserved (fragment)</td>
<td>Tb927.7.3330 (2e-04)</td>
<td>hypothetical protein, conserved</td>
<td>-0.8784</td>
<td>0.0295</td>
<td>NA</td>
<td>Y(3330) abnormal proliferation bsf and diff</td>
<td>NA</td>
<td>T.congolense orthologues hypothetical. 3330 - Enriched in acidocalcisomes in T.brucei pcf, detected in bsf PM and cytoskeletal fraction. GO: transcription regulator, vesicle-mediated transport, acidocalcisome, nucleus, membrane</td>
</tr>
<tr>
<td>TcIL3000_3_970</td>
<td>3-oxo-5-alpha-steroid 4-dehydrogenase, putative</td>
<td>Tb927.3.1840 (0.0)</td>
<td>3-oxo-5-alpha-steroid 4-dehydrogenase, putative</td>
<td>-0.8615</td>
<td>0.0295</td>
<td>Tb927.3.1840</td>
<td>N(but abnormal proliferation pcf)</td>
<td>NA</td>
<td>1840- enriched in T.brucei pcf acidocalcisomes, in bsf PM and cytoskeletal fraction, in pcf. GO: dolichol biosynthetic process, glycosylation, oxidoreductase activity on CH-CH group of donors, mitochondrial inner/outer membrane, cytoplasm, integral component of membrane</td>
</tr>
<tr>
<td>TcIL3000_0_35810</td>
<td>hypothetical protein, conserved</td>
<td>Tb927.10.2880 (2.9)</td>
<td>calcium channel protein, putative</td>
<td>-0.8488</td>
<td>0.0115</td>
<td>NA</td>
<td>Y(2880)</td>
<td>NA</td>
<td>T.congolense orthologues hypothetical. In bsf PM fraction, RNAi compromised flagellum attachment. Located to FAZ; expressed in pcf. GO: ion channel activity, cation, contractile vacuole 1620 - degenerate and frameshifted</td>
</tr>
<tr>
<td>TcIL3000_0_59550</td>
<td>hypothetical protein</td>
<td>Tb927.9.16260 (1.5)</td>
<td>expression site-associated gene 3 (ESAG3, pseudogene), putative</td>
<td>-0.7635</td>
<td>0.0434</td>
<td>NA</td>
<td>Y (16260)</td>
<td>NA</td>
<td>16260 - degenerate and frameshifted</td>
</tr>
</tbody>
</table>
Appendix 1.3 *T. congolense* transcripts that were less abundant at peak relative to ascending parasitaemia (excluding genes described as VSG or with no protein product assigned). The eight transcripts with the largest reduction in abundance are shown. The *T. b. brucei* orthologue for each *T. congolense* gene identified by BLASTP is listed alongside its product description. Additionally, any *T. b. brucei* orthologue information listed on GeneDB is included. If RNAi targeting a *T. b. brucei* orthologue resulted in reduced fitness during differentiation from bloodstream to procyclic forms (implicating a role in growth in either life cycle stage or in differentiation) in the study by Alsford *et al* (Alsford *et al*., 2011), this was highlighted in green. If either the *T. congolense* protein or *T. b. brucei* orthologue were described as cell surface expressed on GeneDB (Jackson *et al*., 2013) then this was highlighted in yellow.
Other: Number of transcripts

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-hydroxy-3-methylglutaryl-CoA synthase, putative</td>
<td>1</td>
</tr>
<tr>
<td>3-oxo-5-alpha-steroid 4-dehydrogenase, putative</td>
<td>1</td>
</tr>
<tr>
<td>5-methyltetrahydropteroylglutamate-homocysteine S-methyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>6-phosphofructo-2-kinase/fructose-2,6-biphosphatase, putative</td>
<td>1</td>
</tr>
<tr>
<td>65 kDa invariant surface glycoprotein-like protein</td>
<td>1</td>
</tr>
<tr>
<td>acid phosphatase, putative</td>
<td>1</td>
</tr>
<tr>
<td>amino acid transporter, putative</td>
<td>3</td>
</tr>
<tr>
<td>ankyrin-repeat protein, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>ATP-dependent DEAD/H DNA helicase recQ, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Calpain family cysteine protease, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>calpain, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>calpain, putative, cysteine peptidase, Clan CA, family C2, putative</td>
<td>1</td>
</tr>
<tr>
<td>cAMP-specific phosphodiesterase (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Cell division protein kinase, putative</td>
<td>1</td>
</tr>
<tr>
<td>chaperone protein DNAj, putative</td>
<td>2</td>
</tr>
<tr>
<td>Concanavalin A-like lectin/glucanases superfamily, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Concanavalin A-like lectin/glucanases superfamily/SPRY domain/HECT-domain (ubiquitin-transferase), putative</td>
<td>1</td>
</tr>
<tr>
<td>conserved protein</td>
<td>1</td>
</tr>
<tr>
<td>cyclin 3, putative, mitotic cyclin, putative</td>
<td>1</td>
</tr>
<tr>
<td>cytoskeleton-associated protein, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>DNA polymerase epsilon catalytic subunit, putative</td>
<td>1</td>
</tr>
<tr>
<td>DNA polymerase zeta catalytic subunit, putative</td>
<td>1</td>
</tr>
<tr>
<td>dolichyl-P-Man:Man7GlcNAc2-PP-dolichylalpha6-mannosyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>Domain of unknown function (DUF4486), putative</td>
<td>1</td>
</tr>
<tr>
<td>dynein heavy chain, putative</td>
<td>3</td>
</tr>
<tr>
<td>EB1-like C-terminal motif containing protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>endonuclease/exonuclease/phosphatase, putative</td>
<td>1</td>
</tr>
<tr>
<td>fatty acid desaturase, putative</td>
<td>1</td>
</tr>
<tr>
<td>fatty acyl CoA synthetase 2, putative</td>
<td>1</td>
</tr>
<tr>
<td>Galactose oxidase, central domain containing protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>glucosamine-fructose-6-phosphate aminotransferase, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Got1/Sft2-like family, putative</td>
<td>1</td>
</tr>
<tr>
<td>Hydin (pseudogene)</td>
<td>1</td>
</tr>
<tr>
<td>intraflagellar transport protein IFT140, putative</td>
<td>1</td>
</tr>
<tr>
<td>Ion channel, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>leucine-rich repeat protein (LRRP), putative</td>
<td>1</td>
</tr>
<tr>
<td>lipase domain protein, putative</td>
<td>3</td>
</tr>
<tr>
<td>mannosyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>microtubule-associated protein 2</td>
<td>1</td>
</tr>
<tr>
<td>mitochondrial malate dehydrogenase</td>
<td>1</td>
</tr>
<tr>
<td>mitogen-activated protein kinase 2, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>multicopper oxidase, putative</td>
<td>1</td>
</tr>
<tr>
<td>multidrug resistance-associated protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>NAD or NADP dependent oxidoreductase, putative, short chain dehydrogenase, putative</td>
<td>1</td>
</tr>
<tr>
<td>NADH-cytochrome b5 reductase, putative</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix 1.4 *T. congolense* transcripts that were of reduced abundance at peak relative to ascending parasitaemia and for which the protein product was classified as other.

<table>
<thead>
<tr>
<th>Protein/Domain/Fragment</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>neurobeachin/beige protein, putative</td>
<td>2</td>
</tr>
<tr>
<td>Octanoyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>Organic solute transport protein 1, putative</td>
<td>1</td>
</tr>
<tr>
<td>pantothenate kinase subunit, putative</td>
<td>1</td>
</tr>
<tr>
<td>paraflagellar rod component, putative (fragment)</td>
<td>2</td>
</tr>
<tr>
<td>paraflagellar rod protein</td>
<td>2</td>
</tr>
<tr>
<td>Pep3/Vps18/deep orange family/Region in Clathrin and VPS, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>PGAP1-like protein/Alpha/beta hydrolase family, putative</td>
<td>1</td>
</tr>
<tr>
<td>PhD-like phosphatase, putative</td>
<td>1</td>
</tr>
<tr>
<td>phosphatidylinositol kinase related protein, putative (fragment)</td>
<td>2</td>
</tr>
<tr>
<td>phosphatidylinositolN-acetylgucosaminytransferase subunit c, putative</td>
<td>1</td>
</tr>
<tr>
<td>phospholipid-translocating P-type ATPase (flippase), putative</td>
<td>1</td>
</tr>
<tr>
<td>phospholipid-transporting ATPase, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Predicted membrane protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>Protein Associated with Differentiation, putative</td>
<td>1</td>
</tr>
<tr>
<td>protein phosphatase 2C, putative</td>
<td>1</td>
</tr>
<tr>
<td>pteridine transporter, putative</td>
<td>1</td>
</tr>
<tr>
<td>Putative cell surface-expressed gene family (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Putative integral membrane protein conserved region (DUF2404), putative</td>
<td>1</td>
</tr>
<tr>
<td>RAVE protein 1 C terminal, putative</td>
<td>1</td>
</tr>
<tr>
<td>ribosomal protein L21, putative</td>
<td>1</td>
</tr>
<tr>
<td>Right handed beta helix region containing protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>RNA-binding protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>Serine/threonine-protein kinase Nek1-related, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>small GTP-binding rab protein, putative</td>
<td>2</td>
</tr>
<tr>
<td>SNF2 family protein</td>
<td>1</td>
</tr>
<tr>
<td>SPRY domain containing protein, putative</td>
<td>3</td>
</tr>
<tr>
<td>TLC domain containing protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>TOR-like 1 (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Triose-phosphate Transporter family, putative</td>
<td>1</td>
</tr>
<tr>
<td>tryparedoxin peroxidase, putative</td>
<td>1</td>
</tr>
<tr>
<td>ubiquitin carboxyl-terminal hydrolase, putative</td>
<td>1</td>
</tr>
<tr>
<td>UDP-Gal or UDP-GlcnAc-dependent glycosyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>UDP-glucose:glycoprotein glucosyltransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>UDP-N-acetylgucosamine-dolichyl-phosphate N-acetylgucosamineprophosphatetransferase, putative</td>
<td>2</td>
</tr>
<tr>
<td>Unc104-like kinesin, putative</td>
<td>1</td>
</tr>
<tr>
<td>vacuolar assembly protein vps41, putative</td>
<td>1</td>
</tr>
<tr>
<td>vacuolar sorting protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>WD40 repeats, putative</td>
<td>1</td>
</tr>
<tr>
<td>zinc finger protein family member, putative</td>
<td>1</td>
</tr>
</tbody>
</table>
## Other:

<table>
<thead>
<tr>
<th>Description</th>
<th>Number of transcripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>60S ribosomal protein L44, putative</td>
<td>1</td>
</tr>
<tr>
<td>65 kDa invariant surface glycoprotein, putative</td>
<td>1</td>
</tr>
<tr>
<td>69 kDa paraflagellar rod protein, PFR2, putative</td>
<td>1</td>
</tr>
<tr>
<td>aspartate aminotransferase, putative</td>
<td>1</td>
</tr>
<tr>
<td>C2H2-type zinc finger/Zinc finger, C2H2 type, putative</td>
<td>1</td>
</tr>
<tr>
<td>CCCH zinc finger in TRM13 protein/U11-48K-like CHHC zinc finger containing protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>chromatin binding protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>cyclophilin type peptidyl-prolyl cis-trans isomerase, cyclophilin a, putative</td>
<td>1</td>
</tr>
<tr>
<td>cysteine peptidase C (CPC)</td>
<td>2</td>
</tr>
<tr>
<td>DNA-directed RNA polymerase I subunit RPB5z, putative</td>
<td>1</td>
</tr>
<tr>
<td>Endonuclease-reverse transcriptase/Reverse transcriptase (RNA-dependent DNA polymerase), putative</td>
<td>1</td>
</tr>
<tr>
<td>enoyl-CoA hydratase/isomerase family protein, putative</td>
<td>1</td>
</tr>
<tr>
<td>eukaryotic translation initiation factor 4e, putative</td>
<td>1</td>
</tr>
<tr>
<td>glutamyl-tRNA synthetase, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>Guanyl transferase family 8, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>guanine nucleotide-binding beta subunit-like protein, WD40/YVTN repeat-like-containing protein</td>
<td>1</td>
</tr>
<tr>
<td>histone H2B, putative</td>
<td>1</td>
</tr>
<tr>
<td>histone H2B, putative</td>
<td>1</td>
</tr>
<tr>
<td>histone H4, putative</td>
<td>1</td>
</tr>
<tr>
<td>Leucine Rich Repeat, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>kinetoplastid kinetochore protein 8, putative</td>
<td>4</td>
</tr>
<tr>
<td>Nodulin-like, putative</td>
<td>3</td>
</tr>
<tr>
<td>peptidyl-prolyl cis-trans isomerase/rotamase, putative, PPlase, putative</td>
<td>1</td>
</tr>
<tr>
<td>proteasome regulatory non-ATP-ase subunit 10, putative</td>
<td>2</td>
</tr>
<tr>
<td>Protein translation factor SUI1 homolog, putative</td>
<td>2</td>
</tr>
<tr>
<td>reverse transcriptase (RNA-dependent DNA polymerase), putative</td>
<td>1</td>
</tr>
<tr>
<td>ribonucleoside-diphosphate reductase small chain</td>
<td>1</td>
</tr>
<tr>
<td>ribose 5-phosphate isomerase, putative</td>
<td>1</td>
</tr>
<tr>
<td>S-adenosyl-L-homocysteine hydrolase, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>S-adenosyl-L-homocysteine hydrolase/S-adenosyl-L-homocysteine hydrolase, NAD binding domain containing protein, putative (fragment)</td>
<td>1</td>
</tr>
<tr>
<td>S-adenosylhomocysteine hydrolase, putative</td>
<td>2</td>
</tr>
<tr>
<td>succinyl-CoA synthetase alpha subunit, putative</td>
<td>1</td>
</tr>
<tr>
<td>T-complex protein 1, theta subunit, putative</td>
<td>1</td>
</tr>
<tr>
<td>Transferrin receptor-like protein</td>
<td>1</td>
</tr>
<tr>
<td>ZFP family member, putative</td>
<td>1</td>
</tr>
</tbody>
</table>

Appendix 1.5 *T. congolense* transcripts that were of increased abundance at peak relative to ascending parasitaemia and for which the protein product was classified as other.
<table>
<thead>
<tr>
<th>SysID</th>
<th>Description2</th>
<th>logFC</th>
<th>adj.P.Val</th>
<th>Phobius</th>
<th>TMHMM</th>
<th>SignalP</th>
<th>Coils</th>
<th>PFAM, Gene3D, Superfamily, Panther</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tb927.1.500</td>
<td>hypothetical protein, unlikely</td>
<td>-5.145</td>
<td>1.1E-03</td>
<td>1-18 SIGNAL PEPTIDE, 19-85 NON-CYT</td>
<td>1-20 SignalP no TM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.9.13170</td>
<td>hypothetical protein, unlikely</td>
<td>-4.117</td>
<td>2.3E-02</td>
<td>1-20 SIGNAL PEPTIDE, 21-141 NON-CYT</td>
<td>1-20 SignalP no TM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.13060</td>
<td>hypothetical protein, unlikely</td>
<td>-3.589</td>
<td>7.0E-03</td>
<td>1-20 CYT, 21-38 TM, 39-52 NON-CYT, 53-77 TM, 78-164 CYT</td>
<td>21-38, 53-76 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.1010</td>
<td>hypothetical protein</td>
<td>-3.535</td>
<td>2.0E-03</td>
<td>1-30 SIGNAL PEPTIDE, 31-133 NON-CYT</td>
<td>15-37 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.7150</td>
<td>hypothetical protein</td>
<td>2.902</td>
<td>8.8E-03</td>
<td>1-20 CYT, 21-38 TM, 39-52 NON-CYT, 53-77 TM, 78-164 CYT</td>
<td>21-38, 53-76 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.9.16070</td>
<td>hypothetical protein</td>
<td>-2.743</td>
<td>7.5E-03</td>
<td>1-20 CYT, 21-38 TM, 39-52 NON-CYT, 53-77 TM, 78-164 CYT</td>
<td>21-38, 53-76 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.4.2740</td>
<td>hypothetical protein, conserved</td>
<td>-2.574</td>
<td>4.0E-05</td>
<td>1-47 CYT, 48-68 TM, 67-85 NON-CYT, 86-104 TM, 105-195 CYT</td>
<td>50-172, 87-104 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.10.1020</td>
<td>hypothetical protein</td>
<td>-2.455</td>
<td>8.2E-03</td>
<td>1-47 CYT, 48-68 TM, 67-85 NON-CYT, 86-104 TM, 105-195 CYT</td>
<td>50-172, 87-104 TMhelix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.8.5340</td>
<td>hypothetical protein</td>
<td>-2.233</td>
<td>3.7E-03</td>
<td>1-24 SIGNAL PEPTIDE, 25-1222 NON-CYT</td>
<td>1-19 SignalP no TM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tb927.11.890</td>
<td>hypothetical protein, conserved</td>
<td>-2.161</td>
<td>8.9E-05</td>
<td>1-24 SIGNAL PEPTIDE, 25-1222 NON-CYT</td>
<td>1-19 SignalP no TM</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
T. b. brucei transcripts that were at least 4-fold less abundant in stumpy relative to slender parasites and were described as hypothetical proteins. The InterPro search tool was used to investigate predicted domains for these hypothetical proteins.

### Trypanosome clusters for *T. brucei* transcripts reduced in stumpy forms

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Gene ID</th>
<th>Description</th>
<th>Log2FoldChange</th>
<th>P-Value</th>
<th>Function</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>VSG 286</td>
<td>glyceraldehyde 3-phosphate dehydrogenase, glycosomal</td>
<td>-2.120</td>
<td>4.7E-05</td>
<td>VSG</td>
<td>hypothetical, <em>T. b. specific</em></td>
</tr>
<tr>
<td>26</td>
<td>ESAG3 356</td>
<td>pyruvate kinase</td>
<td>-2.033</td>
<td>4.2E-05</td>
<td>ESAG3</td>
<td>hypothetical, <em>T. b. specific</em></td>
</tr>
<tr>
<td>42</td>
<td>histone H4 519</td>
<td>VSG</td>
<td>4.2E-05</td>
<td>4.7E-05</td>
<td>VSG</td>
<td>hypothetical, <em>T. b. specific</em></td>
</tr>
<tr>
<td>65</td>
<td>paraflagellar rod 723</td>
<td>hypothetical/myosin-like protein</td>
<td>187-207</td>
<td>Coil</td>
<td>Cilia/flagella-associated protein 20/WDR90/C3orf67 PFAM 1-184 (PF05018 ), PANTHER 2-219 (PTHR12458)</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>histone 2B 768</td>
<td>phosphatidylinositol phosphatase</td>
<td>14 VSG 286</td>
<td></td>
<td>VSG</td>
<td>flagellum targeting protein kharon 1</td>
</tr>
<tr>
<td>81</td>
<td>beta tubulin 870</td>
<td>VSG</td>
<td>7648</td>
<td>VSG</td>
<td>11386</td>
<td>hypothetical</td>
</tr>
<tr>
<td>83</td>
<td>ESAG2 1579</td>
<td>PUF nine target 1</td>
<td>9273</td>
<td>cytoskeleton corset-associated protein</td>
<td>11419</td>
<td>hypothetical</td>
</tr>
<tr>
<td>85</td>
<td>VSG 1741</td>
<td>arginine kinase</td>
<td>9846</td>
<td>VSG</td>
<td>11455</td>
<td>VSG, pseudogene, <em>T. b. specific</em></td>
</tr>
<tr>
<td>93</td>
<td>glucose transporter 1752</td>
<td>VSG</td>
<td>9859</td>
<td>Leucine Rich Repeat Protein (LRRP), pseudogene</td>
<td>11467</td>
<td>ESAG3, pseudogene</td>
</tr>
<tr>
<td>95</td>
<td>Leucine Rich Repeat Protein (LRRP) 2437</td>
<td>chromosomal passenger protein</td>
<td>10921</td>
<td>hypothetical</td>
<td>11484</td>
<td>hypothetical, <em>T. b. specific</em></td>
</tr>
<tr>
<td>124</td>
<td>calmodulin 3891</td>
<td>p25o, hypothetical</td>
<td>11037</td>
<td>hypothetical, <em>T. b. specific</em></td>
<td>11486</td>
<td>VSG</td>
</tr>
<tr>
<td>174</td>
<td>VSG 4108</td>
<td>par1 paraflagellar rod component</td>
<td>11057</td>
<td>VSG</td>
<td>11492</td>
<td>hypothetical, <em>T. b. specific</em></td>
</tr>
<tr>
<td>205</td>
<td>VSG 4455</td>
<td>VSG</td>
<td>11058</td>
<td>hypothetical, <em>T. b. specific</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>206</td>
<td>VSG 4917</td>
<td>ATP-dependent 6-phosphofructokinase</td>
<td>11075</td>
<td>hypothetical, <em>T. b. specific</em></td>
<td>*</td>
<td>Just 2 members in cluster, <em>T. b. brucei</em> and <em>T. b. gambiense</em></td>
</tr>
<tr>
<td>207</td>
<td>VSG 5057</td>
<td>hypothetical, tetratricopeptide</td>
<td>11165</td>
<td>VSG-related, <em>T. b. brucei only</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>243</td>
<td>VSG 5102</td>
<td>mitochondrial DNA ligase</td>
<td>11174</td>
<td></td>
<td></td>
<td>retrotransposon hot spot protein, pseudogene</td>
</tr>
</tbody>
</table>

Appendix 1.7 Trypanosome orthologue clusters containing members for which transcripts demonstrated pronounced reduction in abundance in *T. b. brucei* stumpy relative to slender forms.
Appendix 2 : Supplementary Information on Hypothetical Protein 2 and orthologues

Appendix 2.1 – Protein sequence alignment of hypothetical protein 2 orthologues

Appendix 2.1 – Location of short linear motifs found in Hyp2 orthologues

Appendix 2.3 – pDex577-Y vector map

Appendix 2.4 – Primer sequences

Appendix 2.5 – Localisation of TY-tagged Hyp2 orthologues in relation to stress granule marker SCD6

Appendix 2.6 – KN configurations following induction of Hyp2 orthologue overexpression in vitro for 24 hours

Appendix 2.7 – KN configurations following induction of Hyp2 orthologue overexpression in vitro for 48 hours

Appendix 2.8 – The growth effect of TcHyp2 overexpression in vitro with different concentrations of doxycycline used for induction

Appendix 2.9 – pEnT6B-Y and pEnT6P-Y vector maps

Appendix 2.10 – Titration of doxycycline to test the effect of Hyp2 orthologue overexpression in the context of TbHyp2 dKO
Appendix 2.1: Protein sequence alignment of Hypothetical Protein 2 orthologues found in African trypanosomes. Alignment was performed with CLC genomics workbench.
Appendix 2.2: Short linear motifs found in Hyp2 orthologues. Motifs were identified in the Eukaryotic linear motif database. Figure was prepared using Protter.
Appendix 2.3 pDex577-Y vector – vector map created with SnapGene® software (from GSL Biotech; available at snapgene.com).

Appendix 2.4 Primers
Appendix 2.5 Localisation of TY-tagged Hyp2 orthologues (methanol fixation) (A). Cells were incubated with high (1\(\mu\)g/ml) or low (10\(^{-4}\)\(\mu\)g/ml) concentrations of doxycycline for 1 or 2 days to induce overexpression. Cells were methanol fixed prior to detection of TY-tagged Hyp2 with BB2 antibody (red). Scale bar represents 10\(\mu\)m.

Overexpressed Hyp2 does not co-localise with starvation stress granule marker SCD6 (B). Punctate pattern of TY-tagged Hyp2 expression did not overlap with SCD6 staining in methanol fixed bloodstream forms.
Appendix 2.6 Induction of Hyp2 overexpression in vitro (24h). A reduction in dividing cells following Hyp2 overexpression is complicated by a considerable proportion of abnormal KN morphologies with 1µg/ml doxycycline induction. There is a comparable reduction in dividing cells with 10^{-4} µg/ml induction, with fewer abnormal cells. Each bar represents the measurement of a single well.
Appendix 2.7 Hyp2 overexpression in vitro reduces the proportion of proliferating cells and causes accumulation in 1K1N. The number of abnormal cells is minimal when $10^{-4}\mu g/ml$ doxycycline is used for induction. Each bar represents the mean and standard deviation of triplicate wells.
Appendix 2.8 Induction of TcHyp2 overexpression with $2 \times 10^{-5}$ µg/ml doxycycline had a minimal effect on growth (A) (each data point represents a measurement from a single well), but TY-tagged protein was still detectable by western blot with BB2 antibody (B). G6PDH was used as a loading control.

Appendix 2.9 pEnT6B-Y and pEnT6P-Y vector maps showing regions from vector replaced with Hyp2 5’UTR and 3’UTR – vector maps created with SnapGene® software (from GSL Biotech; available at snapgene.com).
Appendix 2.10 Titration of doxycycline to test the effect of Hyp2 orthologue overexpression in the context of TbHyp2 dKO. Each data point represents a measurement from a single well.
Appendix 3: Supplementary information on *T. congoense* and *T. b. brucei* interactions *in vitro* and *in vivo*

Appendix 3.1 – Example morphologies following treatment of monomorphic or pleomorphic *T. b. brucei* with *T. congoense* conditioned medium

Appendices 3.2-3.5 – The effect of *T. congoense* and *T. b. brucei* conditioned medium on the growth of various cell lines
Appendix 3.1 Example morphologies following treatment of monomorphs (427) or pleomorphs (AnTat 1.1 90.13) with *T. congolense* conditioned medium or negative control medium. The upper two panels demonstrate comparable morphology between monomorphic (427) cells treated with *T. congolense* conditioned medium and those treated with negative control medium. The lower two panels demonstrate that pleomorphic cells treated with *T. congolense* conditioned medium may have a ‘stumpier’ morphology than those treated with negative control medium. The scale bar represents 10µm.
Appendix 3.2 The effect of *T. congolense* and AnTat 1.1 90.13 conditioned medium on various cell lines. *T. congolense* conditioned medium was collected on day 2 at $7.8 \times 10^6$ cells/ml (pH 7.5). AnTat conditioned medium was collected on day 2 at $2.45 \times 10^6$ cells/ml (pH 8). Conditioned medium was tested on cell lines described previously. Each data point represents the mean and SEM of triplicate wells.
Appendix 3.3 The effect of *T. congolense* and AnTat 1.1 90.13 conditioned medium on *T. congolense*. In this experiment all *T. b. brucei* cultures died. *T. congolense* conditioned medium was collected on day 2 at $7.1 \times 10^6$ cells/ml (pH7). AnTat conditioned medium was collected on day 2 at $2.3 \times 10^6$ cells/ml (pH7.5). Each data point represents the mean and SEM of triplicate wells.
Appendix 3.4 The effect of *T. congolense* and AnTat 1.1 90.13 conditioned medium on various cell lines. *T. congolense* conditioned medium was collected on day 2 at 6x10^6 cells/ml (pH7.5). AnTat conditioned medium was collected on day 2 at 2.4x10^6 cells/ml (pH8). Each data point represents the mean and SEM of triplicate wells.
Appendix 3.5 The effect of *T. congolense* and AnTat 1.1 90.13 conditioned medium on various cell lines. *T. congolense* conditioned medium was collected on day 3 at 6.6x10^6/ml, and AnTat 1.1 90.13 conditioned medium was collected on day 2 at 3.4x10^6/ml. Each data point represents the mean and SEM of triplicate wells.
Appendix 4: Buffers and reagents
4.1 *T. congoense* TcBSF3 media

The recipe for TcBSF3 media was taken from (Coustou et al., 2010).

To prepare the basal media one bottle of MEM powder (Sigma M0643) was mixed with:

- 5.96 g acid HEPES (Sigma H3375)
- 2.2 g NaHCO$_3$ (Sigma 55761)
- 1 g D-glucose (VWR)
- 110 mg sodium pyruvate (Sigma P3662)
- 10.68 mg adenosine (Sigma A4036)
- 14 mg hypoxanthine
- 4 mg thymidine
- 14.1 mg bathocuproinedisulfonic acid disodium salt (Sigma B1125)

Aguettant Versol water was added to a final volume of 1 litre. The pH was adjusted to 7.3, and the basal media was filter sterilised with a 0.22µm filter unit. The basal media can be stored at 4°C for a number of months.

To complete the media combine:

75% TcBSF3 basal media
20% Goat serum (Sigma, G6767)
5% Serum plus (Sigma, 14008C)
0.0014% β-mercaptoethanol
0.8% 200mM glutamine
1% Penicillin/ streptomycin solution (GIBCO by Life technologies, 15140-122)

The completed media was filter sterilised by passing through a 0.22µm filter unit and was stored at 4°C for up to 2 weeks.
4.2 HMI-9 Media

One bottle of HMI-9 powder (Life technologies, 90915N) was supplemented with:

45mM NaHCO₃

256µM β-mercaptoethanol

The medium was made up to 4 litres with autoclaved dH₂O, the pH was adjusted to 7.5, and the medium was passed through a 0.22µm filter.

HMI-9 media was completed by addition of 10% heat-inactivated FBS (GIBCO by Life technologies, 10500-064), 10% sterile dH₂O and 1% Penicillin/ Streptomycin (GIBCO by Life technologies, 15140-122).

4.3 TV3 media

This medium for culturing *T. vivax* epimastigotes was based on (D'Archivio et al., 2011).

One bottle of Iscove's Modified Dulbecco's Medium (IMDM) containing 25mM D-glucose, 25mM HEPES and 4mM glutamine (GIBCO, 21980065) was supplemented with:

10% heat-inactivated FBS (GIBCO by Life technologies, 10500-064)

10% heat-inactivated Goat serum (Sigma, G6767)

0.03mM bathocuproinedisulfonic acid disodium salt (Sigma B1125)

0.14mM β-mercaptoethanol

0.05mM Thymidine

0.2mM Hypoxanthine

0.4mM Cysteine HCl

10mM Proline
4.4 **TE solution**

10mM Tris HCl pH8
1mM EDTA pH8
Sterile filtered through 0.22µm filter.

4.5 **TELT (lysis buffer)**

50mM Tris HCl pH8
62.5mM EDTA pH 8
2.5M LiCl
4% v/v Triton
Sterile filtered through 0.22µm filter.

4.6 **6x DNA loading buffer**

15% (v/v) ficoll 400
0.5% (w/v) orange G

4.7 **TAE buffer**

0.4mM Tris-HCl
0.1mM sodium acetate
0.1mM EDTA

4.8 **Miniprep solution I**

50mM glucose
25mM Tris-HCl pH8
10mM EDTA pH8
4.9 Miniprep solution II

0.2M NaOH

1% sodium dodecyl sulphate (SDS)

4.10 Miniprep solution III

3M potassium acetate

11.5% (v/v) Glacial acetic acid

4.11 Trypanosome Dilution Buffer (TDB)

5mM KCl

80mM NaCl

1mM MgSO\(_4\).7H\(_2\)O

20mM Na\(_2\)HPO\(_4\)

2mM NaH\(_2\)PO\(_4\)

20mM glucose

Adjusted pH to 7.4

4.12 Phosphate buffered saline (PBS)

137mM NaCl

3mM KCl

16mM Na\(_2\)HPO\(_4\)

3mM KH\(_2\)PO\(_4\)

Adjusted to pH 7.4
4.13 SDM-79 Media

1 bottle of SDM-79 powder (Life technologies) was added to 4.5 litres of autoclaved dH$_2$O, and this was supplemented with 9g NaHCO$_3$. This was adjusted to pH 7.3 and passed through a 0.22µm filter.

SDM-79 media was completed by addition of 10% heat-inactivated FBS (GIBCO by Life technologies, 10500-064), 0.1% Hemin and 1% Penicillin/ Streptomycin (GIBCO by Life technologies, 15140-122).

4.14 FACS fix

2% formaldehyde

0.05% gluteraldehyde

1x PBS

4.15 Laemmli buffer (2x)

125mM Tris HCl pH6.8

20% glycerol

4% SDS

0.005% bromophenol blue

To 9 volumes of Laemmli buffer (2x) 1 volume of fresh β-mercaptoethanol was added shortly before use.
4.16 Solutions for protein gels

4.16.1 Separating solution (10ml)

<table>
<thead>
<tr>
<th></th>
<th>8%</th>
<th>10%</th>
<th>12%</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>4.6ml</td>
<td>4.0ml</td>
<td>3.3ml</td>
</tr>
<tr>
<td>30% Acrylamide mix</td>
<td>2.7ml</td>
<td>3.3ml</td>
<td>4.0ml</td>
</tr>
<tr>
<td>1.5M Tris (pH8.8)</td>
<td>2.5ml</td>
<td>2.5ml</td>
<td>2.5ml</td>
</tr>
<tr>
<td>10% SDS</td>
<td>0.1ml</td>
<td>0.1ml</td>
<td>0.1ml</td>
</tr>
<tr>
<td>10% APS</td>
<td>0.1ml</td>
<td>0.1ml</td>
<td>0.1ml</td>
</tr>
<tr>
<td>TEMED</td>
<td>0.006ml</td>
<td>0.004ml</td>
<td>0.004ml</td>
</tr>
</tbody>
</table>

4.16.2 Stacking solution (5ml)

H₂O – 3.4ml
30% Acrylamide mix – 0.83ml
1.0M Tris (pH 6.8) – 0.63ml
10% SDS – 0.05ml
10% APS – 0.05ml
TEMED – 0.005ml

4.17 Tris-glycine electrophoresis buffer

25mM Tris
192mM Glycine
0.1% SDS
4.18 Western Blot transfer buffer
25mM Tris
192mM Glycine
20% Methanol

4.19 Western Blot stripping buffer
50mM glycine
2% SDS
Adjusted to pH2.2

4.20 10X MOPS
221mM MOPS
50mM Sodium acetate, pH7
10mM EDTA
This was autoclaved and stored in the dark at 4°C.

4.21 RNA gel loading buffer
30% formamide
6% formaldehyde
1x MOPS
0.01% bromophenol blue
10% glycerol

4.22 20xSSC
3M NaCl
0.3M tri-sodium citrate
4.23 Hybridisation buffer

5x SSC
50% Formamide
0.02% SDS
2% DIG block

4.24 Maleic Acid Buffer

100mM Maleic acid
150mM NaCl

Adjusted to pH7.5 with NaOH then autoclaved

4.25 10% DIG block

10g blocking reagent (Roche 000000011585762001) was dissolved in 100ml Maleic acid buffer, autoclaved and stored in 10ml aliquots at -20°C.

4.26 Wash buffer (for Northern blot)

0.3% Tween 20 in 1x Maleic acid buffer

4.27 Detection buffer

100mM Tris HCl
100mM NaCl pH 9.5