ANA FILIPA DE ALMEIDA RIJO FERREIRA

CIRCADIAN RHYTHMS IN A *TRYPANOSOMA BRUCEI* INFECTION: HOST, PARASITE OR BOTH?

Tese de Candidatura ao grau de Doutor em Biologia Básica e Aplicada submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Luísa M. Figueiredo

Categoria – Investigadora Principal

Afiliação – Instituto de Medicina Molecular

Orientador – Joseph S. Takahashi

Categoria – Professor

Afiliação – Howard Hughes Medical Institute, University of Texas Southwestern Medical Center at Dallas

Co-orientador – Ana Tomás

Categoria – Professor Associado

Afiliação – Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto

Para ser grande, sê inteiro: nada
Teu exagera ou exclui.
Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.
Assim em cada lago a lua toda
Brilha, porque alta vive.

~

To be great, be whole: nothing that's you
Should you exaggerate or exclude.
In each thing, be all. Give all you are
In the least you ever do.
The whole moon, because it rides so high,
Is reflected in each pool.

In Odes de Ricardo Reis, 1933
Fernando Pessoa

This work was supported by Fundação para a Ciências e Tecnologia (FCT) by means of a Ph.D. fellowship SFRH/BD/51286/2010 awarded to Ana Filipa de Almeida Rijo Ferreira through the Graduate Program in Areas of Basic and Applied Biology (GABBA), Universidade do Porto, Portugal. The research was also supported by the Howard Hughes Medical Institute (HHMI) and the European Molecular Biology Organization (EMBO).

Acknowledgments

My mom always told me ‘that people will give you what you give them’. I was given so much during my Ph.D. that I hope I have been able to give in return. I want to express my gratitude, my respect and also my deep admiration to all of you that made my Ph.D. journey a memorable experience. Muito obrigada.

First of all, I would like to thank both my mentors Luisa and Joe, who believed in this project and started a collaboration allowing me to pursue it. You were very complementary, obviously in your field of expertise but also with other general skills you taught me. However there were important features I cherish that you both share: a good heart and being very passionate about science and discovering new things, something I hope I absorbed. By having a shared co-mentorship, always traveling between continents, you both made me realize that I would need to be the driver of my project. Although scary, it turned out to be fantastic. You trusted and supported my decisions, guided me when needed and were always available even when apart. I feel I had the best of two worlds, I was very fortunate to have both as mentors. Thank you.

To the GABBA directors, Professor Maria de Sousa, Fátima Carneiro, Manuel Sobrinho Simões, António Amorim and Alexandre do Carmo, I thank you the opportunity and your trust when accepting me as a GABBA student and part of the GABBA family. I thank Prof, Ana Tomás, for being my mentor and always looking out for me, making sure I kept on track. I thank the GABBA program and my fellow colleagues, for organizing the annual GABBA meeting, where I could always find good input and interesting discussions, great moments reuniting with my friends and colleagues, and Port wine. Especially, I would like to thank the GABBA 14th: my dear Milene, Ines, Mariana, Netinho, Catarina, McLindo, Pedrito, Bruno, Daniel, ReNA e Von for all the great moments we shared and for always being an email away to help dealing with frustrations and decisions, even when on the other side of the world.

The daily hurdles couldn't have been overcome without the help of my colleagues and friends from both labs: Sandrini, Penita, Xico, Fabics, Margarida, Daniel, Leonor, Mafaldinha, Helena, Dragão, Idálio & Jen, Marleen, Vicky, Lucia, Nobu, Shu, Junmei, Iza, Yoga, Chryshanthi, Lisa, Noheon, Pin, Guocun, Laura, Sally, Vivek, Shelley, Prachi, Neha, Yan, Yongli, Mariko, Delali, Ivan. In particular Danieeeeel and Leo, who I kept always very very busy, even when I was in the US, with tons of emails, texts and skype calls. And my friends Vicky, Marleen, Jen, Lucia and Crystal, who became my family in the US. The colleagues from Green and Yamazaki labs and the extended Figueiredo lab and Malaria group at IMM. I also thank Tânia, who gave me her very skilled hand and guided me when uncovering the mouse model.

I would also like to thank Meg's lab and the Parasitology community at UTSW, for all the help in parasite related things, giving feedback on my project and for always making me feel as one of them. Especially I must thank Meg, who has been amazing to me, always made me feel very welcome and valued, included me in the parasitology group at UTSW and has been looking out for me- a real scientific godmother.

I also want to thank my dear friends Mafalda, Maria, Sara, Arnaldo, Catarina, mini Di, Navalhetes, Garras, Mica, Patrick, Andrew, priminha Joana, Ferreira and Arturito for always keeping me in their lives, giving me a shoulder to cry and celebrating the good moments.

Agradeço muito a minha Mãe, Pai, Avózinha, tia Cristina e querida familia alargada (Bela e Zé), por me apoiarem sempre, por todo o amor e por poder contar sempre convosco. E claro, a vida não era a mesma sem uma chamada da Avó todos os Domingos. Muito obrigada.

Finally the person I must thank the most is Fernando, for always being by my side and being my biggest critic and number one fan. For joining me in this adventure, for all his love, support and encouragement, and for keeping me fed with things other than avocado bagels. For helping and teaching me tricks for

visual aspects of science communication and this final document represents something we achieved together. I also would like to thank him and Carlos, our little companion, for all the laughs and ventures. For together making this ride so much more pleasant and exciting to ride.

Table of contents

Abstract / Resumo

Abstract

Sleeping sickness is a fatal human infectious disease caused by the unicellular *Trypanosoma brucei* parasite. This disease can be divided into two stages: an early stage in which parasites are restricted to the bloodstream and interstitial spaces of peripheral organs; and a late stage, in which parasites invade the central nervous system. The most characteristic symptom of this disease is the disruption of the sleep/wake cycle, however disruption of temperature and hormone release rhythms are also observed. Since in humans these rhythms are regulated by the circadian clock, we hypothesized that the parasite could disrupt daily rhythms by deregulating the clock genes that drive them.

Initially we carefully characterized the C57BL/6J mouse model of sleeping sickness, to acquire a comprehensive description of the distribution of the parasites among the different organs/tissues of the host. From this study emerged the surprising conclusion that the so-far neglected adipose tissue is in fact a major parasite reservoir. With further RNA sequencing and biochemical analyses we discovered the adaptation of the parasite the lipid-rich environment of adipose tissue by utilizing the beta-oxidation of fatty acids, a pathway not previously shown to be used by this parasite.

After documenting the distribution of the parasite in the mouse, we proceeded to test whether this mouse model mimicked the sleep-wake cycle disruptions observed in sleeping sickness patients. Locomotor activity and temperature analysis confirmed the sleep disruption and provided new insight into how the parasite influences circadian rhythms, by shortening the period the host. Characterization of the clock gene expression in infected mice confirmed this shortening of the clock period occurred in the organs with highest parasite load – the adipose tissue. This period shortening could be reversed upon clearance of the parasite with suramin, an anti-*Trypanosoma* drug commonly used in the field. By contrast, infection with the rodent malaria parasite, *Plasmodium chabaudi*, did not affect the period of the host clock suggesting an effect specific to *T. brucei*. In addition, *T. brucei* co-cultured with fibroblasts *in*

vitro shortened the period of the clock of the fibroblasts suggesting a rather direct effect on host cells.

Finally, we tested whether *T. brucei* parasites have intrinsic circadian rhythms themselves, through analysis of their daily gene expression *in vitro*, by RNA sequencing. We found robust circadian oscillations in gene expression that could be entrained by temperature and to a lesser extent by light, in the two stages of the parasite life cycle. Such daily rhythmic gene expression lead to differences in the parasite population during 24 h, which impacts their susceptibility to treatment with suramin at different times of the day.

In summary we identified a new type of host-parasite interaction that may underlie one of the most characteristic symptoms of sleeping sickness in humans and opens new therapeutic possibilities. Our work also showed the existence of circadian rhythms in *T. brucei*, suggesting that other pathogens may also be able to sense, adapt and anticipate the environment.

Resumo

A doença do sono é uma doença infeciosa fatal causada pelo parasita *Trypanosoma brucei*. Esta doença pode ser dividida em dois estados: um estado inicial, no qual os parasitas se encontram em circulação no sangue e espaços intersticiais de vários órgãos periféricos; e um estado tardio, no qual os parasitas invadem o sistema nervoso central. O sintoma mais característico desta doença é a desregulação do ciclo de sono. No entanto, a desregulação da temperatura corporal e dos níveis rítmicos de hormonas no sangue também são observados nos doentes de doença do sono. Em humanos, estes ritmos de sono, temperatura e hormonas são regulados pelo ritmo circadiano, o que nos levou a formular a hipótese de que os parasitas poderiam desregular os genes de ritmo circadiano que criam estes ritmos.

Inicialmente caracterizámos o modelo de ratinho C57BL/6J de doença do sono para uma melhor compreensão da distribuição dos parasitas nos diferentes órgãos/tecidos do hospedeiro. Desta caracterização emergiu a conclusão surpreendente de que, no estado tardio, o tecido adiposo, até ao momento negligenciado, é o tecido com maior número de parasitas infiltrados. Com análises de transcriptómica e bioquímicas, descobrimos que o parasita se adapta ao ambiente rico em lípidos do tecido adiposo através da utilização da via da beta-oxidação de ácidos gordos, uma via nunca anteriormente observada neste parasita.

Após a documentação da distribuição do parasita no hospedeiro, testámos se este modelo animal reproduzia as manifestações de desregulação de sono observada nos doentes. Através da medição da atividade e temperatura corporal dos ratinhos, confirmámos que este modelo recapitula os sintomas dos doentes e que tal pode ser em parte pelo facto do parasita acelerar o período do ritmo circadiano do hospedeiro. Molecularmente, quando analisámos a expressão dos genes de ritmo circadiano, confirmámos a observação do período acelerado nos órgãos com maior número de parasitas infiltrados. Curiosamente, se eliminarmos os parasitas com o tratamento pelo fármaco suramina, o período retorna ao normal. Pelo contrário, quando infectámos ratinhos com o agente causador da malária, *Plasmodium chabaudi*,

não observámos diferenças no período dos genes de ritmo circadiano, o que sugere que este efeito é específico do parasita *T. brucei*. Cultivámos ainda *T. brucei* juntamente com fibroblastos e observámos que o período do genes de ritmo circadiano dos fibroblastos também acelerou, o que sugere que este efeito possa ser direto.

Por fim, testámos se o parasita *T. brucei* tem, tal como o ratinho, o seu próprio ritmo circadiano, através de análise de transcriptómica. Descobrimos oscilações diárias na expressão de vários genes do parasita, que podem ser sincronizadas principalmente pela temperatura, mas também pela luz exterior. Estas oscilações na expressão génica induzem diferenças na população dos parasitas a diferentes horas do dia, o que tem impacto na susceptibilidade deste ao tratamento com suramina, dependente da hora do dia.

Em conjunto, estes resultados abrem novas vias de investigação no estudo do parasita responsável pela doença do sono, bem como sugerem a existência de ritmos circadianos nos microrganismos patogénicos em si.

List of abbreviations

3'UTR	3' untranslated region
5'UTR	5' untranslated region
a.a.	Amino acid
ATF	Adipose tissue forms
BBB	Blood brain barrier
BCB	Blood-cerebrospinal fluid barrier
BSF	Bloodstream forms
CSF	Cerebrospinal fluid
CT	Circadian time
DD	Constant darkness
FDR	False discovery rate
gDNA	Genomic DNA
GPEET	Procyclin with Gly-Pro-Glu-Glu-Thr repeats
HAT	Human African trypanosomiasis
IC_{50}	Half maximal inhibitory concentration
Kb	Kilobase
LD	Light / dark cycle
LUC	Luciferase
Mb	Megabase
mRNA	messenger RNA
NREM	Non Rapid Eye Movement
PER2	Period 2
PCF	Procyclic form
PCU	Polycistronic unit
PI	Propidium iodide
Pol I	RNA polymerase I
Pol II	RNA polymerase II
Pol III	RNA polymerase III
qPCR	quantitative real-time PCR
RBP	RNA binding protein
REM	Rapid Eye Movement

RNA-Seq	RNA sequencing
RPKM	Reads Per Kilobase of transcript per Million mapped reads
SCN	Suprachiasmatic nuclei
SWS	Slow Wave Sleep
VSG	Variant surface glycoprotein
WHO	World Health Organization
ZT	Zeitgeber time

CHAPTER I

CHAPTER I – Introduction

1. Sleeping sickness

1.1. Clinical features and treatment

Human African African trypanosomiasis (HAT), best known as sleeping sickness, is a fatal human infectious disease caused by the unicellular *Trypanosoma brucei* parasite. This disease can be included in the group of African trypanosomiasis, which comprises a variety of human and animal pathologies caused by different *Trypanosoma* species transmitted through the bite of a blood-feeding tsetse fly (*1*). It is endemic mainly in sub-Saharan Africa, where the tsetse fly vector (*Glossina* spp.) finds optimal environmental conditions to survive (*2*).

Only *Trypanosoma brucei* spp. are pathogenic to humans, causing sleeping sickness: *T. b. gambiense* and *T. b. rhodesiense* - which are widespread in western and eastern Africa, respectively, where they cause approximately 97% and 3% of current cases (*3*). Both species lead to coma and death if patients are left untreated. A third *T. brucei* subspecies, *T. b. brucei* is non-infective to humans, but parasitizes domestic and wild animals, causing a disease called Nagana, which severely hampers cattle production in one-third of the African continent (*2*). The difference between these species is that *T. b. brucei* cannot escape an innate immune response that humans have: trypanosome lytic factor (TLF) serum complexes. TLF-1 and TLF-2 contain haptoglobin related protein (HPR) and apolipoprotein L1 (APOL1), which enter via the haptoglobin–hemoglobin receptor of the trypanosome (TbHpHbR) killing the parasite (*4*). Both human infective trypanosomes escape this TLF lytic mechanism: *T.b. rhodesiense* expresses serum resistance-associated (SRA) protein, which neutralizes the lytic activity of APOL1 (*5*); and *T.b. gambiense* expresses Trypanosoma gambiense-specific glycoprotein (TgsGP), which stiffens endo/lysosomal membranes, presumably blocking the internalization of APOL1 (*6*). Despite this difference *T. b. brucei* is very close to the human

infective subspecies, making it a widely used experimental model of HAT in murine and rat models (*7, 8*).

Clinically, sleeping sickness can be divided into two stages. In the early stage, parasites can be found in the bloodstream and interstitial spaces of several organs, after which they actively invade the central nervous system, starting the late stage. *T. b. gambiense* infection is chronic, with estimated average duration around 3 years, which is evenly divided between the two stages (*9*). *T. b. rhodesiense* disease is usually acute, and death occurs within weeks to months (*10*) likely due to this parasite being less adapted to human beings, affecting mostly animals (livestock and wildlife), which act as dangerous reservoirs of the parasite (*2*).

Patients experience a myriad of symptoms, from neuropsychiatric, motor and weight loss disturbances, among others (*2, 11*). The leading signs and symptoms of the early stage are chronic and intermittent fever, headache, pruritus, lymphadenopathy, and, to a lesser extent, hepatosplenomegaly. In the late stage, sleep disturbances and neuropsychiatric disorders dominate the clinical presentation. This common sleep disruption experienced by patients gives the disease its common name of ‘sleeping sickness’. However, studies in which patients sleep state was measured have shown that, unlike a hypersomnia, the total amount of time spent asleep by these patients is similar to healthy individuals, instead it is the sleep structure and the timing at which sleep occurs that are affected (*12*).

The resurgence of sleeping sickness epidemics in the 1980’s and 1990’s alerted the international community for the risks of neglecting this disease. Since then, an international coordinated effort led by the World Health Organization reinforced control and surveillance programs which led to a drop in sleeping sickness cases since 2009 to less than 10,000 new cases per year (*2*). Nevertheless, it is believed that a substantial number of cases remain unreported since the disease is endemic in rural areas and more recent estimates predict that 70 million people in tropical Africa are at risk of infection (*2*). Although, sleeping sickness is lethal if left untreated; however sporadic

cases of natural progression to asymptomatic carriage or even apparent spontaneous resolution of the infection have been reported for gambiense infection, resembling the trypanotolerance phenomena described for some African cattle species (*13, 14*).

As a whole, African trypanosomiasis represents a serious health and socio-economic burden. Besides the severe life threat posed by sleeping sickness to humans, animal African trypanosomiasis, or Nagana causes around 3 million cattle deaths and total annual losses of US $4.75 billion. Nagana is an incredibly heavy burden particularly in rural areas where livestock production is the main livelihood, perpetrating underdevelopment and poverty in these regions (*15*).

In humans, a major obstacle for eradication of sleeping sickness is the insufficiency of the existing treatments. No prophylactic treatment or vaccine is available and the drugs used to treat early and late stage sleeping sickness are largely not accessible orally, often very toxic, and sometimes ineffective (*16*). The first-line of treatment for early stage *T. b. gambiense* is pentamidine, which has been used since 1940 and is usually administered intramuscularly (sometimes intravenously) (*11*). Pentamidine is usually effective but has the potential complications of hyperglycemia or hypoglycemia, hypotension, and gastrointestinal features (*17*). Early stage of *T. b. rhodesiense* infection is treated with intravenous suramin, used since the early 1920s and, although usually effective, especially when given early in the disease, can result in potential complications such as renal failure, skin lesions, anaphylactic shock, bone marrow toxicity, and neurological complications such as peripheral neuropathy (*17*) (Table 1).

Table 1. Treatment used for sleeping sickness

		First-line treatment	Second-line treatment
Early stage	*T. b. gambiense*	Pentamidine	Suramin
	T. b. rhodesiense	Suramin	Pentamidine
Late stage	*T. b. gambiense*	Eflornithine + nifurtimax (oral)	Melarsoprol
	T. b. rhodesiense	Melarsoprol	Non-available

Treating late stage sleeping sickness is more complicated because the drugs that can cross the blood brain barrier (BBB) are more toxic. The effective drug for treating late stage *T. b. rhodesiense* sleeping sickness is the arsenical melarsoprol, which acts on trypanothione, a parasite molecule that maintains an intracellular reducing environment (*18*). Melarsoprol injections are extremely painful and this drug is extremely toxic. It induces post-treatment reactive encephalopathy in 10% of the patients, half of whom die, leading to an overall mortality from treatment of about 5% (*19*). Although this drug is also effective for late stage *T. b. gambiense* infection, it is no longer used due to its toxicity. Instead, eflornithine, an ornithine decarboxylase inhibitor, is the treatment used for late stage *T. b. gambiense* since 1981 (*17*). Unfortunately, eflornithine is ineffective against *T. b. rhodesiense*, treatment is still intravenous, patients suffer from side effects (bone marrow toxicity, alopecia, seizures, and gastrointestinal deregulation) (*17*), and there is the potential for drug resistance to arise through loss of the putative amino acid transporter TbAAT6 (*20*). The standard procedure in case of successful treatment requires a regular follow-up with repeated lumbar punctures for 1 year, after which patients are considered cured if the cerebrospinal fluid (CSF) remains clear - no parasites nor more than 5 lymphocytes/mm^3.

While some new promising drugs are in the pipeline (e.g. fexinidazole and oxaborole SCYX-7158) (http://www.dndi.org) (*16*), investment in the search for more effective treatments has been set as a priority. That, together with continued surveillance, improved diagnostics and vector control policies will be vital to meet the World Health Organization's target to eliminate sleeping sickness as a public health issue by 2020 (*21*).

1.2. Sleep disruption and other physiological parameters

The sleep/wake cycle in humans is remarkably stable and consolidated. As diurnal animals, humans are mostly active during the day period and sleep ~6-8 hours at night. Sleep starts with non-REM (NREM – non rapid eye movement) sleep or slow wave sleep (SWS) followed by rapid eye movement

(REM) sleep (Fig. 1B – healthy individual). These two types of sleep are defined by electrophysiological signs detected in humans by a combination of electroencephalography (EEG, which records cortical activity), electrooculography (records eye movement) and electromyography (records muscle tone). The measurement of such electrophysiological parameters is collectively termed polysomnography (*22*).

In sleeping sickness patients, sleep abnormalities include deregulation of the sleep/wake cycle, when patients experience periods of nocturnal insomnia and daytime sleepiness; episodes of uncontrollable sleep and fragmentation of the normal sleep structure (*12, 23*). Sleep/wake cycle disruption, meaning the time at which patients sleep have been both documented by EEG but also by actigraphy, with wrist-worn watches allowing to measure rest/activity pattern in humans (*24*) (Fig. 1). The timing at which patients sleep is the most remarkable sign of this disease, the quality of their sleep is altered. Both *T. b. gambiense* and *T. b. rhodesiense* patients show particular poor sleep efficiency with frequent awakenings and arousals at night and recurrent short naps during the day (*25, 26*).

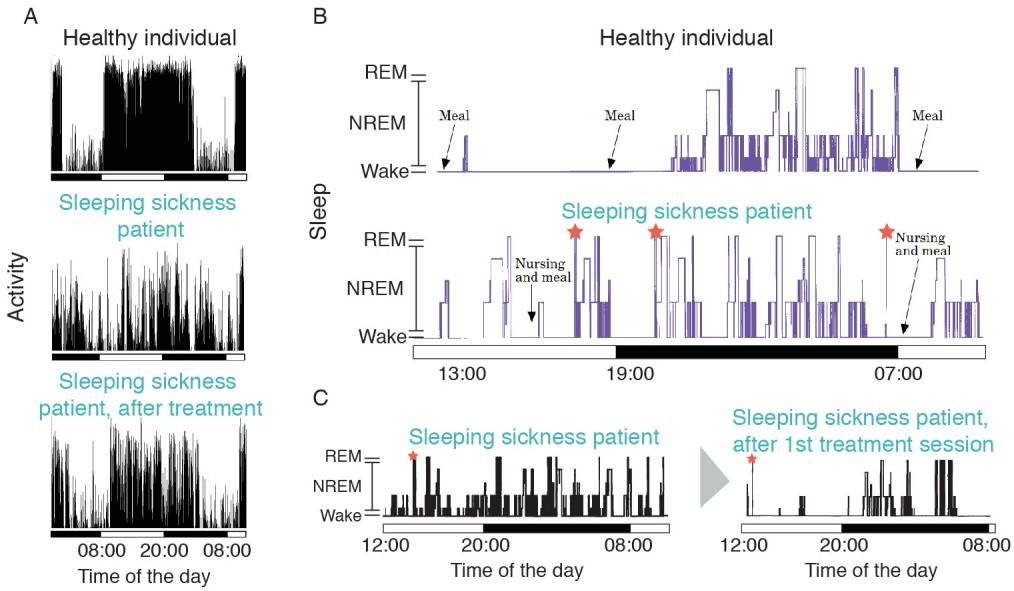


Figure 1. Activity and sleep disruption in sleeping sickness patients. (A) Activity record of both a healthy individual and a sleeping sickness patient, measured by using activity watches. The same patient had his/her activity measured both before and after

pentamidine treatment, in which is clear the reversal of the abnormal activity profile. (adapted from (*24*)). (B) EEG recording of a healthy individual and a sleeping sickness patient. Sleeping sickness patient record shows both increased sleep during the daytime and transitions from wake to REM state, highlighted with orange stars (adapted from (*27*)). (C) EEG recording of a sleeping sickness patient before and after the first treatment session, showing a clear reversal of the sleeping time to the nighttime, upon treatment (adapted from (*28*)).

SOREM (sleep onset rapid eye movement) episodes are characteristic of narcolepsy. SOREM episodes mean a sudden abnormal transition to REM sleep episodes while the person is awake or after a very short NREM sleep period, thus disrupting wakefulness or the NREM–REM cycle during sleep (*29*). In sleeping sickness patients the alterations of the internal sleep structure include SOREM episodes (Fig. 1 – orange star), which may appear earlier than the sleep/wake alterations. At least the Gambian form of HAT shares other features, besides SOREM episodes, with narcolepsy, such as excessive daytime sleepiness, and sleep fragmentation (*30, 31*). This observation was reproduced in a rat model of sleeping sickness (*32-34*).

Narcoleptic patients have low or undetectable CSF levels of orexin (also known as hypocretin) due to the loss of the orexin neurons in the lateral hypothalamus (*35*). Intriguingly, the CSF orexin levels measured in patients affected by *T. b. gambiense* HAT are significantly lower than in control cases. Although not as low as narcoleptic patients, the low CSF orexin concentrations in sleeping sickness patients still indicates a dysfunction of this hypothalamic peptide may be involved in sleeping sickness (*29*). Lower orexin levels than control were also observed in rat models of sleeping sickness (*36*).

Interestingly, prostaglandin D2 (PGD2), which induces NREM sleep when injected into the ventricle system (*37*), is significantly increased in CSF of late stage sleeping sickness patients (*38*). Trypanosomes can themselves produce PGD2 (*39*) that they also use for cell density regulation. Further studies are needed to assess if this contributes to the physiological deregulation observed in sleeping sickness patients, since these patients do not experience overall increase of NREM, but perhaps the timing of PGD2 release is disrupted.

Humans and other animals, in addition to cyclic sleep/wake cycle, also have daily secretion of hormones and daily temperature fluctuations, with a high/low temperature fluctuation during day/night, respectively.

Similar to the sleep/wake cycle disruption observed in sleeping sickness, daily rhythms of hormonal secretion are also affected in *T. b. gambiense* HAT patients. However this is not the case for all hormones: secretion of the pineal hormone melatonin is maintained in HAT patients in the encephalitic stage, though being secreted earlier than expected (phase advanced) (*40*). The cortisol, prolactin plasma hormone levels and the renin activity (an index of release of renin, a key enzyme in the renin–angiotensin system) are severely disrupted in HAT patients at an advanced stage of disease (*41, 42*).

The studies available on daily temperature regulation in sleeping sickness patients seem to have controversial results. In healthy humans, there is a subtle fluctuation in body temperature, being lower during the nighttime., but overall body temperature regulation appears to be preserved (*40*). However, in rat models, core temperature differences of morning and evening are reduced in some of the animals (*33, 43*), indicating disturbances in the daily regulation of body temperature rhythm, as well as sudden hypothermia (*44*).

It is notable that sleep structure and sleep/wake cycle disruption in sleeping sickness patients can both be reversed upon treatment (*45*) (Fig. 1A and 1C). In addition, based upon autopsy studies, despite strong inflammatory response and microglia activation, there does not seem to be massive neuronal damage (*46*). These two observations suggest that the presence of the parasite itself might be responsible for the symptoms observed in sleeping sickness patients.

2. *Trypanosoma brucei*

2.1. Evolutionary origins

Deciphering the evolution of the first eukaryotic lineages is an enigma. Many groups of protists have claimed phylogenetic positioning at the early branches of the eukaryotic tree of life, diverging substantially from the fungi, plant and animal branches. Despite their apparent simplicity, among protists there is tremendous diversification in lifestyles and unique biological features. The supergroup Excavata, phylum Euglenozoa to which the parasite *Trypanosoma brucei* belongs is a good example. Trypanosomes such as *T. brucei* are evolutionary highly distant organisms from the well-known eukaryotic groups of animals, fungi and plants (*47*). Within Euglenozoa, trypanosomes belong to a very particular clade, the Kinetoplastida, distinct by the presence of special mitochondrial DNA, the kinetoplast or kinetoplast DNA (kDNA). This is a large disc-shaped mass consisting of an extensive network of interlocked circular DNA molecules, located adjacent to the basal body at the base of the flagellum (*48*) (Fig. 2).

Trypanosomatida, such as *T. brucei*, offer a rare opportunity to study the origin and evolution of parasitism, because trypanosomatids were the only kinetoplastids acquiring an obligatory parasitic lifestyle (*49*). Trypanosoma and Leishmania genus contain species that spend their life between a mammalian host and insect vector and might either develop intracellularly in host cells at some stage of its life cycle, such as *T. cruzi* and *Leishmania* spp., or, as in the case of *T. brucei*, live exclusively as an extracellular parasite.

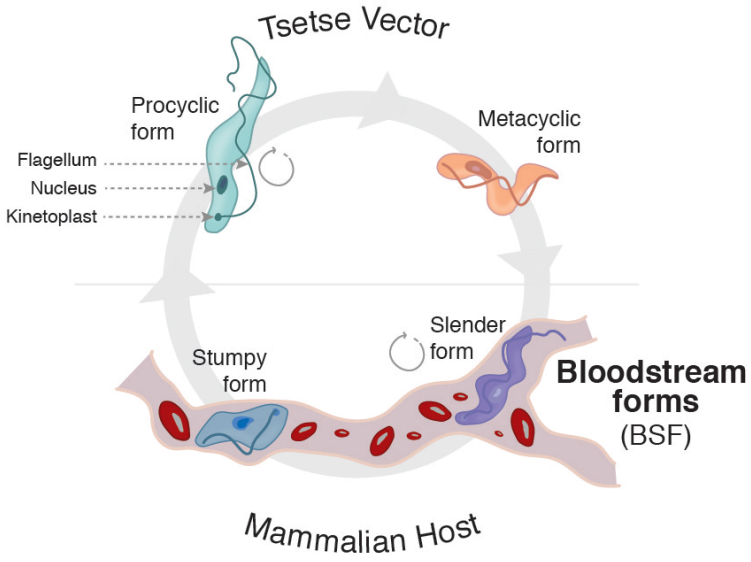


Figure 2. Life cycle of *Trypanosoma brucei*. In the bloodstream of the mammalian host, parasites exist as a polymorphic population of bloodstream forms (BSF) consisting of dividing (circle gray arrows) slender forms and cell cycle arrested stumpy forms. In the tsetse fly vector, after entering the midgut, stumpy forms differentiate to the procyclic forms (PCF). In the salivary glands there is the final differentiation into the infective metacyclic forms, which are injected during the next blood meal of the fly into the mammalian host.

African trypanosomes, such as *T. brucei,* have a mechanism of antigenic variation of their protective coat of variant surface glycoproteins (VSGs), providing an extremely efficient adaptation to extracellular life inside its host. The long history of research on *T. brucei* has made it a great model system to study evolution, disease mechanism and even basic eukaryotic biology: it has now a small fully sequenced 35 Mb genome (*50*), simple genetic manipulation tools available and an RNA interference (RNAi) system (*51*). Indeed, various discoveries were primarily made in cultured *T. brucei* such as RNA editing (*52*), *trans*-splicing (*53*) and glycosyl-phosphatidylinositol (GPI) anchors (*54*).

2.2. A life of adaptations: from vector to host and back

Trypanosoma brucei is a heteroxenous parasite, i.e., it requires more than one obligatory host to complete its life cycle: a mammalian host and the

blood-feeding tsetse fly vector. Mammalian infection starts with a bite of an infected tsetse that inoculates the infective cell cycle arrested metacyclic stage into the mammal bloodstream and draining lymphatics. *T. brucei* then differentiates to a long slender bloodstream form (BSF) that divides asexually by binary fission. In the mammal, parasites live extracellularly in the bloodstream or in the extravascular fluids and interstitial spaces of organs and tissues, such as the brain, heart and lungs (*55*). In the bloodstream, at high parasite density, a yet unidentified parasite-released factor induces differentiation from long slender forms to short stumpy BSF, passing through an intermediate form (*56*). These stumpy forms are cell cycle arrested and the only life cycle stage transmissible to the tsetse upon a blood meal of the fly (Fig. 2).

Once in the fly midgut, the stumpy forms differentiate into procyclic forms (PCF), which re-enter cell cycle, actively multiplying and colonizing the fly gut. PCF undergo a complex differentiation process while migrating to the fly proventriculus. During this process each PCF divides asymmetrically into one long and one short epimastigote; the long epimastigotes arrest their development while the short epimastigotes migrate to the tsetse salivary glands. Here, they attach to epithelial cells and start to multiply. Attached epimastigotes are the only life cycle stage known to perform meiosis (*57*). Finally, in the salivary glands, epimastigotes differentiate to metacyclic forms, ready to be transmitted back to the mammalian host by a tsetse bite (Fig. 2).

Throughout its life cycle, *T. brucei* exhibits a series of specific changes that allow it to adapt to completely different environments it encounters within the mammal and the fly. The process of differentiating from one life cycle stage to the other is complex and implicates dramatic changes in cell morphology, metabolism, cell cycle and cell surface proteins, accompanied by a strong reprogramming in gene expression (*58*).

Recent studies uncovered the genome-wide changes in gene expression between these BSF and PCF, identifying 6–40 % of genes differentially expressed between them (depending on the method used) (*59-62*). This highlights the parasite capacity of rapidly remodeling its gene expression pattern upon sensing the environmental changes. The most striking change the

parasite undergoes is the cyclic full renovation its surface protein composition: i) inside the fly, PCF are coated with highly acidic and repetitive proteins - the procyclins (GPEET and EP procyclins) (*63, 64*); ii) these are then exchanged for a coat of brucei alanine-rich protein (BARP) coat in epimastigotes (*65*); iii) differentiation into the metacyclic stage replaces BARP for a dense coat of metacyclic variant surface glycoproteins (MVSGs); iv) in the mammal, a similar type of VSG is maintained as the major surface protein in the BSF slender and stumpy forms of *T. brucei* (*66*); v) finally, when stumpy forms differentiate to PCF, the VSG coat is shed and replaced by procyclins (*38)*.

Over the years, researchers documented *T. brucei* dramatic changes in adaptation to different environments, from the mammalian host bloodstream to the fly midgut. Although, it is well established that the parasites in the mammalian host are not restricted to the bloodstream, but in fact infiltrate different tissues, it remains unknown whether parasites adapt to these new environments.

2.3. Life inside the mammalian host

Inside the mammalian host parasites enter the bloodstream where they must face a strong immune response. As extracellular parasites, they cannot hide from the highly resourceful immune system of the host; instead they developed ways to deceive it.

The dense VSG coat of *T. brucei* acts as protective barrier against the innate and adaptive immune system of the mammal. It protects parasites from both recognition and lysis by the alternative complement pathway (*67*) and from anti-VSG antibody-mediated phagocytosis (*14*), in part by internalization of anti-VSG antibody complexes from the parasite surface via high rates of endocytosis when recycling the existent VSG (*68-70*) and by their release due to hydrodynamic forces generated by its flagellar movement (*71*). This constant VSG shedding is also actively triggered by the parasite, led by a trypanosome membrane-associated phospholipase C (GPI-PLC), which cleaves the GPI anchor. Puzzlingly, the substantial release of soluble VSG (sVSG) is sufficient

to activate macrophages and Th1 cell response (which may be a distractive tactic), but also downregulate the production of trypanocidal factors such as nitric oxide (NO) and tumor necrosis factor alpha (TNF-α) at later time points of infection (*72*). Finally, the densely packed VSG coat probably acts as a protective cover for other cell-surface molecules from immune recognition (*73, 74*). Although all these phenomena contribute to deceive the immune system, the most sophisticated mechanism this parasite has for immune evasion is called antigenic variation. As anti-VSG antibodies can eventually eliminate the parasite population, a few parasites have switched their VSG coat to a new VSG that have not yet been recognized by the immune system. As a result, these parasites with new VSG will successfully replicate and replace the previous parasite population that is being targeted by the antibodies. The continuous clearing of the recognized VSG followed by the emergence of newly switched variants leads to the characteristic periodic waves of parasitemia described for the first time in humans more than a century ago (*75*) (Fig. 3).

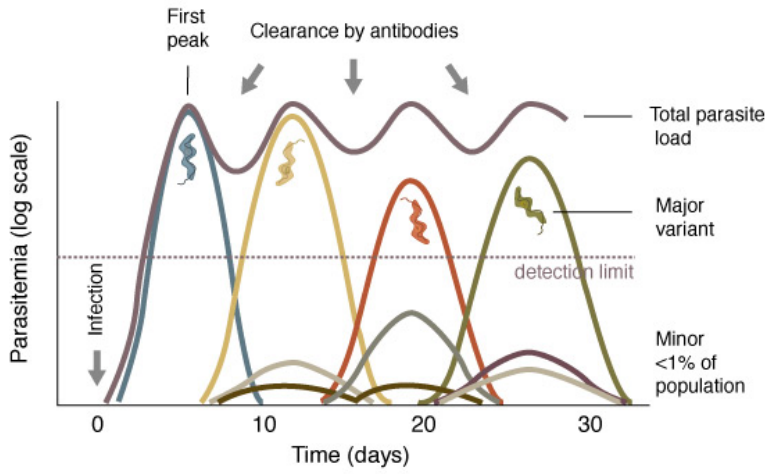


Figure 3. Antigenic variation: a mechanism to evade the immune system. A population of parasites expressing the same variant surface glycoprotein (VSG) dominates (blue curve). This population will be targeted by the antibody response but a second population expressing a new VSG will proliferate (yellow line), and the cycle continues. The outcome is the purple line of total parasite load, representing the waves of parasitemia characteristic of this infection.

In the mammalian host the parasites initially reach the blood, however they are not restricted to the bloodstream. Evidence from patient autopsies and from animal models allowed the tracking of parasites over the course of infection and identified them in the interstitial spaces of multiple tissues/organs.

The parasite entry into the brain has been extensively studied, as it marks the late stage of this disease and consequently the type of treatment to administer. Two possible entry points exist: the blood-brain barrier (BBB) and the blood-csf barrier (BCB). Data have been presented for the use of both pathways (*76, 77*), although it is believed that trypanosomes initially cross the BCB, as their massive accumulation in the choroid plexus area is obvious earlier in infection, whereas presence in the brain parenchyma is only sporadically observed. This preferential crossing at different locations is likely due to a more permeable anatomy of the choroid plexus barrier, BCB (*78*). The BBB is formed by endothelial cells connected by tight junctions, which on the brain side of the vessel are also enfolded by astrocyte feet, forming an additional cellular layer (*79, 80*). In contrast, the BCB does not contain the astrocyte layer and consists of two spatially isolated cellular barriers: i) fenestrated endothelial cells line out the capillaries of the choroid plexus and attach to a basal lamina; ii) epithelial cells form the surface of the choroid (*81*).

Later during infection, however, trypanosomes and white blood cells pass across the BBB into the brain parenchyma. Whether this process happens via diapedesis (as lymphocytes cross, by adhesion to endothelial cells and actively crossing between them) (*82*) or through extravasation due to BBB leakiness (*83, 84*), is still controversial. However, no matter which of these processes are responsible for the penetration in the brain parenchyma, it seems that it happens during later time points of infection when high levels of interferon gamma (IFN-γ) are detected, which seems to facilitate the penetration, since IFN-$\gamma^{-/-}$ and IFN-γ receptor $^{-/-}$ mice have fewer parasites in the brain (*82*).

In human infections the late stage symptoms, which correlate with appearance of parasites in brain, has a latency of at least a few weeks. In rats or mouse models, similarly trypanosomes can be detected in the brain ~20 days post-infection (*77, 85*). Controversially, one recent study showed that with high inoculum, parasites could be detected in the brain within 24 h of infection (*86*).

The heart is another organ with infiltrated trypanosomes. More specifically parasites can be found on the endocardial, epicardial side, valves and in the lymphatic system draining the heart, the latter being particularly evident in late infection (*87*). In *T. b. brucei*-infected dogs parasites were also identified in the heart lymphatic system (*88*). Moreover, cardiac involvement in sleeping sickness has been reported, as recently demonstrated by electrocardiography (ECG) alterations, and appears early in the course of the disease in humans (*89*).

Many other organs have few infiltrated parasites since the early stages of infection. In addition to the heart, parasites also infiltrate the lung, spleen, liver, kidney, testis and epididymis (*90-92*). However not all these organs show viable parasites, in cases such as the liver and the spleen they are mostly found degraded or inside phagocytic cells (*93*). Recent studies with bioluminescent parasites confirmed the general distribution of parasites reported previously with histological studies (*55*). Prior to the work presented in this dissertation, only indirect evidence has described the presence of parasites in the adipose tissue, as massive infiltration of inflammatory cells in this tissue in a model of *T. b. gambiense* (*90*).

Despite these studies characterizing parasite presence in organs/tissues other than the bloodstream, brain and heart, the correlation between anatomical changes, clinical symptoms and pathogenesis remains largely unknown.

2.4. Genome organization

Trypanosoma brucei is a diploid organism with a nuclear genome of approximately 35 Mb/haploid DNA content, which contain ~10,000 genes. Despite its small genome, the karyotype in *T. brucei* is remarkable, consisting of more than 120 chromosomes divided into three different classes according to their size: 11 megabase chromosomes (1-6 Mb), 1-5 intermediate chromosomes (200-900 kb) and ~100 minichromosomes (30-150 kb) (*50, 94*).

The core of minichromosomes consists mostly of repetitive palindromes of 177-bp repeats. Intermediate chromosome structure is not entirely known but it shares with minichromosomes a large core of 177-bp repeats (*95*). Some strains of *T. brucei* also contain nuclear extrachromosomal circular DNAs, termed NlaIII repeat-elements of unknown function (*96*).

About 15% of the total DNA in trypanosomes corresponds to the kDNA which consists on a intertwined network of thousands of double-stranded circular DNA molecules which include large maxicircles (~23 kb) and smaller minicircles (~1 kb) (*97*). Similarly to mitochondrial DNA in other eukaryotes, maxicircles encode ribosomal RNA (rRNA) genes plus genes mostly coding for subunits of the mitochondrial membrane complexes involved in oxidative phosphorylation.

2.5. Gene expression in *T. brucei*

In trypanosomatids large clusters of unrelated genes are co-transcribed by RNA polymerase II (Pol II) from the same coding strand as a single polycistronic unit (PCU) (*50*). Typically, the regions between PCUs are strand-switch regions (SSRs) because they separate PCUs with opposite coding strands. Experimental evidence supports that Pol II transcription initiates bidirectionally at divergent SSRs and terminates at convergent SSRs (*98-100*). Pol II transcription can also initiate internally within transcripton units and at other non-SSRs, for instance where a change of RNA polymerase type likely occurs (*98*). Regions between Pol II transcription units often contain transfer RNAs (tRNAs) and other RNA polymerase III (Pol III)-transcribed small RNAs. Besides, in *T. brucei* the Pol I-transcribed rDNA units and procyclin loci also lie between Pol II transcription units (TriTrypDB 9.0) (Fig. 4).

Remarkably, *T. brucei* Pol I is the only known multifunctional Pol I among eukaryotes. Besides transcribing the rDNA unit (18S, 5.8S and 28S-like genes) it also transcribes three other protein-coding loci: the VSG bloodstream expression sites (BESs), the procyclin loci and the MVSG expression sites (MESs) (*101, 102*). Only ~10 rDNA transcription units are predicted in *T. brucei*, a smaller number compared with the hundreds estimated for other eukaryotes. Pol I transcription of rDNA genes is the most active transcription in a eukaryotic

cell, generally accounting for over 50% of the total transcription activity. This is also achieved with high transcription initiation rates by ~150 Pol I enzymes simultaneously transcribing an rDNA gene (*103*). Evolutionarily, it seems likely that in order to be able to generate ~10^7 VSG surface copies from a single gene, Pol I was the selected RNA polymerase to maintain the cell surface protein production.

T. brucei Pol II transcribes most mRNAs, the spliced leader (SL) RNA gene required for mRNA processing by *trans*-splicing and small nucleolar RNAs (snoRNAs) (*104, 105*). No Pol II promoters or associated transcription factors have been found except for the SL RNA gene promoter (*106*). Chromatin immunoprecipitation (ChIP) sequencing revealed epigenetic marks associate with putative Pol II transcription start sites (TSSs) and transcription termination sites (TTSs): H2A.Z, H2B.V acetylated H4K10, tri-methylated H3K4 and a putative acetyl-binding protein, the bromodomain factor 3 (TbBDF3) are enriched at predicted transcription start sites (TSSs) whereas H3.V and H4.V preferentially locate at potential TTSs (*107, 108*) (Fig. 4).

Pol III transcribes tRNAs, 5S rRNA, small nuclear RNAs (snRNAs), similarly to the other eukaryotes (*109*).

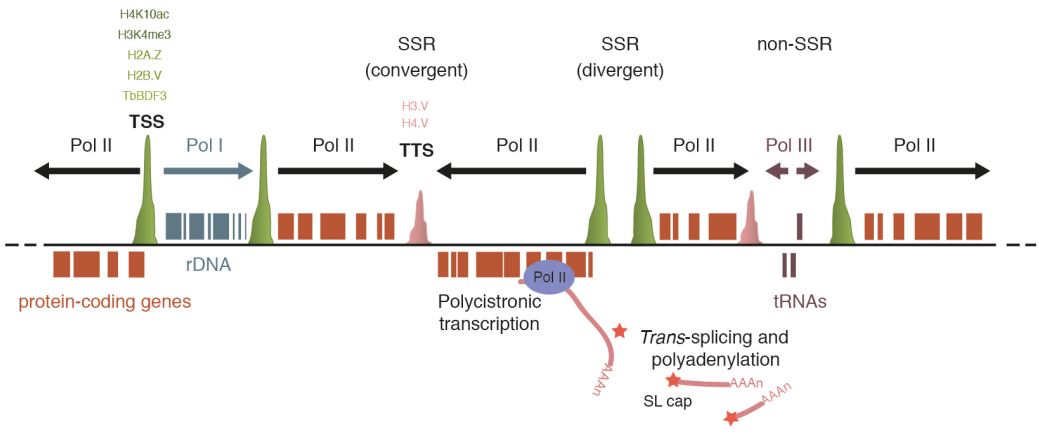


Figure 4. Transcription units and flanking epigenetic signals in *T. brucei*. Most of *T. brucei* genome is organized in polycistronic transcription units (PCUs) transcribed by Pol II. Neighboring PCUs are usually separated by strand-switch regions (SSR). Pol II transcription initiates at divergent SSRs and terminates at convergent SSRs. Transcription might also initiate at non-SSRs. Arrows indicate direction of transcription. Regions in-between Pol II units often contain Pol I-transcribed rDNA units and procyclin loci, and Pol III-transcribed genes such as tRNAs. Predicted transcription

start sites (TSS) are enriched for acetylated H4K10 (H4K10ac), tri-methylated H3K4 (H3K4me3), H2A.Z, H2B.V and TbBDF3. Putative Pol II transcription termination sites (TTSs) are enriched for H3.V and H4.V. *Trans*-splicing and polyadenylation events occur simultaneously to transcription, splicing of the pre-mRNA with the addition of a 5'-Splice leader cap (SL) and a poly(A) tail.

One of the most curious distinct characteristics of kinetoplastids is their lack of transcriptional regulation, a conserved feature among most eukaryotes. Instead genes are transcribed as long polycistrons of 10–100 genes (*110*) (Fig. 4). It is believed that all polycistronic precursor RNAs are transcribed at roughly the same rate; consequently, the regulation of gene expression occurs entirely post-transcriptionally. Within the same polycistronic unit the steady-state mRNA may vary 50-fold (*37*), with the regulation of transcript levels achieved by post-transcriptional regulation. There are few known exception to this: the genes coding for the very abundant *T. brucei* cell surface proteins procyclin and VSG and a few associated genes which are transcribed by Pol I from a well-characterized promoter (*102*). The unusual genome organization may have evolved to allow kinetoplastids to have a very gene-dense genome (even introns are almost absent) and reduce cell cycle time. However, how the post-transcriptional regulation in *T. brucei* occurs remains largely a mystery. One of the few known cases is the upregulation of phosphoglycerate kinase B gene upon differentiation of BSF to PCF, which results from the control of mRNA half-live mediated by U-rich regulatory elements at the 3'UTR that destabilize mRNAs in BSF (*111*). Inroads are being made with genome-wide approaches, providing a comprehensive view on the dominant post-transcriptional mechanisms. RNA-seq studies have gathered evidence for the importance of alternative splicing and polyadenylation (*59, 60, 98*) and control of mRNA half-live and decay rates (*112*) in regulation of mRNA levels. Widespread RNA translation control, assessed by ribosome profiling in BSF and PCF, add to the list of post-transcription regulation mechanisms in *T. brucei* (*113, 114*). Finally, *trans*-acting RNA-binding proteins have been emerging as crucial players in remodeling gene expression patterns during the *T. brucei* life cycle: overexpression of a single *T. brucei* RNA-binding protein (RBP), TbRBP6, is sufficient to trigger the entire differentiation pathway of BSF to metacyclic

parasites in axenic cultures (*115, 116*). In addition, several other studies have identified subsets of target mRNAs potentially regulated by RBPs throughout the life cycle development in *T. brucei (117, 118)*.

In *T. brucei,* the mRNA maturation process is also unusually different from other eukaryotes. Polycistronic precursor transcripts are processed into mature monocistronic mRNAs via *trans*-splicing, which adds a m7G-capped 39-nucleotide Spliced Leader RNA to the 5' untranslated region (5'UTR), and by polyadenylation of the 3' untranslated region (3'UTR) (*53*). These two events – *trans*-splicing and polyadenylation – occur during transcription itself, while polyadenylating the upstream mRNA *trans*-splicing of the downstream occurs (*119*) (Fig. 4). The common mRNA maturation process of eukaryotes consists of *cis*-splicing, which excises non-coding introns from the pre-mRNAs. In *T. brucei*, however, *cis*-spliced introns seem to be restricted to only two genes, the poly(A) polymerase (*120*) and a putative RNA helicase gene (*50, 60*).

3. Circadian rhythms

Circadian rhythms control a variety of biological processes in numerous organisms, ranging from bacteria to humans. These rhythms are an intrinsic way to count the time that provides the capacity to anticipate the rhythmic daily changes in our environment. Perhaps the most obvious function regulated by circadian rhythms is the daily sleep and wake cycle in animals. However, many other physiological processes are regulated by circadian rhythms, including body temperature, feeding behavior, hormone secretion, glucose homeostasis, and cell cycle progression. When circadian cycles are disrupted, either by genetic or environmental insults (e.g. jet lag and alternating shift work), disorders of diverse physiological processes can occur, spanning from obesity to cancer (*121*).

3.1. An evolutionary advantage among kingdoms

The regular 24 h rotation of the earth has led to the evolution of circadian oscillators in virtually all life forms, from prokaryotes to eukaryotes. Geological records suggest that non-heterocystous cyanobacteria represent one of the most ancient cellular life forms on Earth, first appearing perhaps more than 3.8 billion years ago. These cyanobacteria have circadian oscillators which indicates circadian clocks are a truly ancient biological timing system (*122*).

What selective pressures would these ancient organisms experience? For cyanobacteria the pressure is likely to have been the need of having temporally separated the processes of oxygenic photosynthesis and nitrogen fixation. For all organisms, the direct effects of ultraviolet radiation and photo-oxidative damage have given rise to the “escape from UV” hypothesis, in which the restriction of S phase (DNA synthesis) of the cell cycle to nighttime would have strong selective value (*123*). But in general, having a clock provides an organism with a predictive mechanism to tune its internal physiology to the external world, and several studies have shown that a robust internal clock offers a competitive advantage. A plant, instead of starting to transcribe and translate all the proteins required for the photosynthetic process upon sensing the sun in the early morning, would, with a circadian clock, benefit from having them ready at the end of the night anticipating the first sun rays. And this is in fact what having a circadian clock allows, which has obvious beneficial consequences. When wild-type plants (circadian rhythmicity period of ~24 h) are compared with mutant plants whose circadian rhythm period is long (>24 h) or short (<24 h), plants with a clock period matched to the environment grow faster, contain more chlorophyll, fix more carbon and survive better than plants with circadian periods differing from their environment (*119*). Similar observations were made in cyanobacteria, which grow faster than mutant cyanobacteria whose circadian clock period differs from 24 h (*124, 125*).

So, how can organisms count the time and anticipate the rhythmic changes of their environment?

Organisms have cell autonomous circadian oscillators that include positive and negative elements, which together form feedback loops. In most

eukaryote organisms, positive elements of the loop are transcription factors that activate the transcription of genes encoding the negative elements (Fig. 5). As a result, the concentrations of the negative elements rise allowing them to interact with the positive elements, inhibiting their activity. This inhibition decreases transcription rates of negative element genes. Phosphorylation-induced decay of the negative elements decreases their concentrations, which leads to reactivation of the positive elements, allowing the cycle to start again. These positive and negative elements are the so-called ‘clock genes’ that together keep the time keeping mechanism of the core loop. However, it is now known that this is a very complex process with more loops interlocked, important for maintaining the stability and robustness of the oscillator. All of these events impose time delays within the core feedback loop, such that the molecular cycle takes ~24 hours (*126*).

Although a similar timekeeping mechanism, involving positive and negative elements is present across all kingdoms of life, the key regulatory elements that drive this are different and very little conserved. This suggests that the appearance of circadian clocks occurred in more than one event during evolution. Animals such as drosophila and mice, which diverged from a common ancestor some 600–700 million years ago, share many of the genes comprising the circadian clock. Although the general feedback-loop mechanism is similar among these species, there are differences in specific functions between orthologues for several of the components. Furthermore, gene duplication has led to increased complexity among vertebrate clock genes (*126*) (Fig. 5).

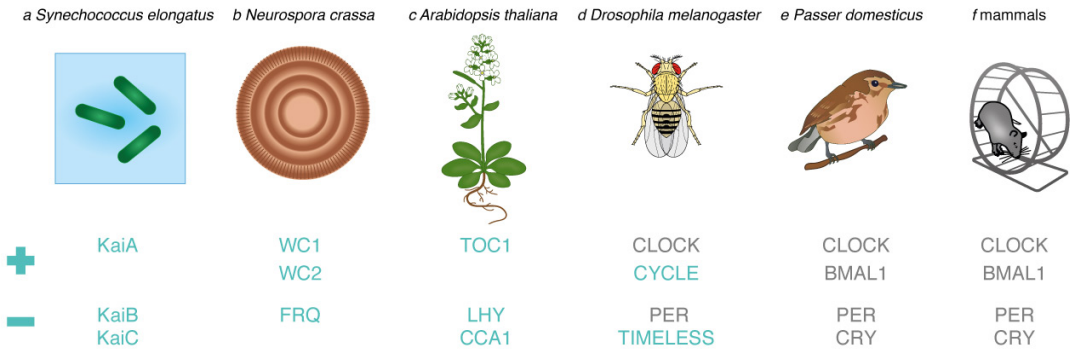


Figure 5. Molecular circadian clock in model organisms. Positive and negative regulators of the core clock of different organism are represented. In blue are the clock proteins that are not conserved in mammals, however sharing similar feedback loop mechanism.

Because the targets of the positive elements of the timekeeping mechanism are not restricted to the negative elements, but include other genes involved in various cellular pathways, daily rhythms are observed in genes involved in multiple pathways resulting rhythmic in physiological functions (Table 2).

Table 2. Examples of processes regulated by circadian clock

Organism	Model system	Master clock tissue / oscillator	Processes regulated by pacemaker
Cyanobacteria	*Synechoccus elongatus*	Kai periodosome	Cell division, photosynthesis, carbohydrate synthesis, gene expression, amino-acid uptake
Yeast	*Neurospora crassa*	FRQ/WC oscillator	Conidiation, gene expression
Plant	*Arabidopsis thaliana*	Vasculature (not fully coupled)	Leaf movement, photosynthesis
Fly	*Drosophila melanogaster*	Ventral lateral neurons Olfactory sensory neurons	Locomotor activity Odor-dependent electrophysiological responses Gene expression
Birds	*Passer domesticus*	Brain (SCN) Retina Pineal gland	Locomotor activity, noradrenaline levels, electrical firing, sympathetic tone Melatonin levels
Mammals	Mouse	Brain (SCN)	Locomotor activity, electrical firing, hormone secretion, cytosolic calcium levels, metabolic activity, immune response, neuropeptide secretion, gene expression

An additional clock exists in cells: the presence of reduction–oxidation (redox) cycles of ROS-scavenging peroxiredoxins. This seems to be present in all domains of life, which has led to the speculation that circadian oscillators co-evolved from this redox homeostasis system following oxygenation on Earth (*127*).

3.2. Mammalian circadian clock

In all circadian systems identified to date, regardless of phylogenetic origin, three major components are present: 1) a light input, 2) the circadian pacemaker itself, and 3) output pathways by which the circadian pacemaker regulates physiology, and behavior throughout the organism (*128*). In humans and other mammals, entrainment of the circadian system by light relies on retinal photoreceptors (rods, cones and retinal ganglion cells, ipRGC). These ipRGCs express melanopsin photopigment, their photoresponse has slow kinetics and a relatively high threshold to light, making them ideally suited to function as circadian photoreceptors. This is because an optimal circadian photoreceptor must integrate light information over relatively long durations and must be insensitive to transient light signals that are not associated with the solar light cycle. The photic information received by the retina is projected to the hypothalamus via the retinohypothalamic tract (RHT). In the hypothalamus there is an area designated suprachiasmatic nuclei (SCN), composed of ~20,000 neurons, each of which is thought to contain a cell autonomous circadian oscillator that together make the master circadian clock. Conclusive evidence that the SCN is the master pacemaker came from lesioning studies and from experiments showing that transplantation of SCN tissue from donor animals harboring circadian clock gene mutations into SCN-lesioned wild-type hosts conferred upon the host the mutant circadian phenotype (*129, 130*).

Light information entering the SCN is transduced into neural and humoral output signals that influence various rhythms in the body including, for example,

temperature, sleep-wake cycle, levels of activity and hormone secretion (*131*) (Fig. 6).

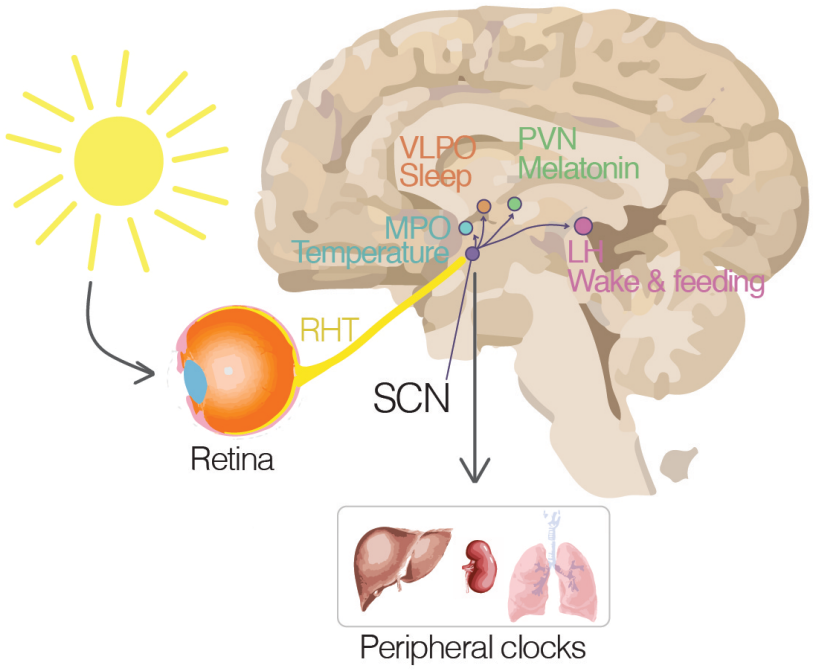


Figure 6 - A schematic diagram of the suprachiasmatic nucleus and its input and output pathways. The mammalian circadian pacemaker in the hypothalamic suprachiasmatic nucleus (SCN) receives light information through the retinohypothalamic tract (RHT). Both neuronal and humoral signals function as output signals from the SCN to other regions of the brain and the periphery. The SCN output pathways are responsible for proper timing of diverse physiological functions, including hormone release, sleep–wake cycle, feeding behavior, and thermoregulation. The medial preoptic region (MPO) controls circadian rhythms of body temperature, paraventricular nucleus (PVN) controls daily hormone secretion and lateral hypothalamus (LH) and ventrolateral preoptic nucleus (VLPO) regulate and sleep–wake cycles.

In circadian biology, the term Zeitgeber (literally, “time giver” in German) is used to describe any daily environmental cue to which the circadian system can synchronize or entrain. Light is the most important cue for most animals, but other ~24 h recurring fluctuations also qualify as Zeitgebers including temperature (*132*) and social cues (*133*). In the laboratory, the daily period of light and dark can be controlled and Zeitgeber time (ZT) is defined relative to the experimental light/dark (LD) cycle. In a LD cycle of 12 hours light and 12 hours darkness (LD 12:12), by definition ZT0 = lights are turned on, and ZT12 =

lights are turned off. When animals are experimentally isolated from all external time cues, as for example either in constant darkness (DD) or constant light (LL), it is possible to measure the endogenous or “free-running” period of the circadian clock. For rodents, the gold standard assay to measure circadian period of the free-running locomotor activity rhythm using running wheels connected to computerized data acquisition systems (*134*). Animals are first entrained to an LD cycle for a few weeks, after which they are exposed to constant conditions (e.g., DD) for several additional weeks during which the endogenous circadian period is measured by analyzing daily locomotor activity records (Fig 7). Each day of the free-running circadian rhythm measured in constant conditions is divided into 24 hours of circadian time (CT). The first half of the cycle (CT0–CT12) is termed subjective day, while the second half of the cycle (CT12–CT24) is termed subjective night. For nocturnal rodents, active period in DD is mainly restricted to the animal’s subjective night. When in DD, the mice period = subjective day + subjective night, is extremely precise (a standard deviation in period that is ~0.2 h or 12 min in mice) (*135*), however it is not exactly 24 h but ~23.7 h, as observed by the slope in activity in the actogram of figure 7.

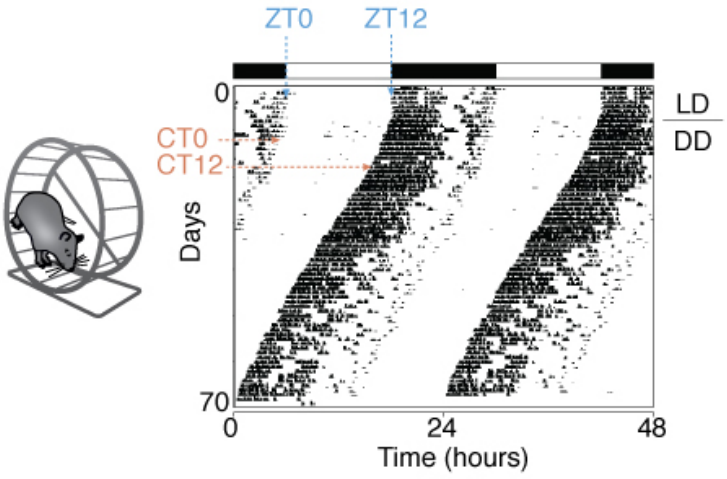


Figure 7. A representative wheel-running locomotor activity record from a wild-type mouse. The record is double-plotted so that 48 h are shown for each horizontal trace. Dark regions represent locomotor activity. For the first seven days the animal was housed in an LD 12:12 cycle, denoted by the bar above the record. The animal was transferred to DD conditions on day 8, as indicated by the horizontal line to the right of the record. Indicated are also the environmental times ZT0, ZT12 and circadian times CT0 and CT12, defined when the animal is in DD.

As I mentioned above, the molecular circadian clock is primarily regulated by a core transcription-translation feedback loop. In mammals, the primary loop involves the activators circadian locomoter output cycles kaput (*Clock*; and its paralog neuronal PAS domain protein 2, *Npas2*), *Bmal1* (also known as aryl hydrocarbon receptor nuclear translocator-like; *Arntl*); and their repressors period homologue 1 (*Per1*), *Per2*, Cryptochrome 1 (*Cry1*) and *Cry2* (Fig. 8). During the day, the basic helix-loop-helix PAS-domain containing transcription factor CLOCK (or NPAS2) interact with BMAL1 to activate transcription of the *Per* and *Cry* genes, resulting in high levels of these transcripts. The resulting PER and CRY proteins heterodimerize, translocate to the nucleus and interact with the CLOCK–BMAL1 complex to inhibit their own transcription (*131*). During the night, the PER–CRY repressor complex is degraded, and CLOCK–BMAL1 can then activate a new cycle of transcription. This feedback cycle takes ~24 h, and the turnover of the PER and CRY proteins is tightly regulated by E3 ubiquitin ligase complexes. There are additional feedback loops interlocked with the core CLOCK-BMAL1/PER-CRY loop. Prominent among these is a loop involving *Rev-erbα* (*Nr1d1*) and *Rora*, which are also direct targets of CLOCK-BMAL1. The feedback effects of this loop affect *Bmal1* transcription (and to a lesser extent *Clock*). Other feedback loops involve the PAR-bZip family members, DBP, HLF, and TEF; the bZip protein, E4BP4 (*Nfil3*); and the bHLH proteins, DEC1 and DEC2 (*Bhlhb2*, *Bhlhb3*), all of which are transcriptional targets of CLOCK-BMAL1 (*131, 136*).

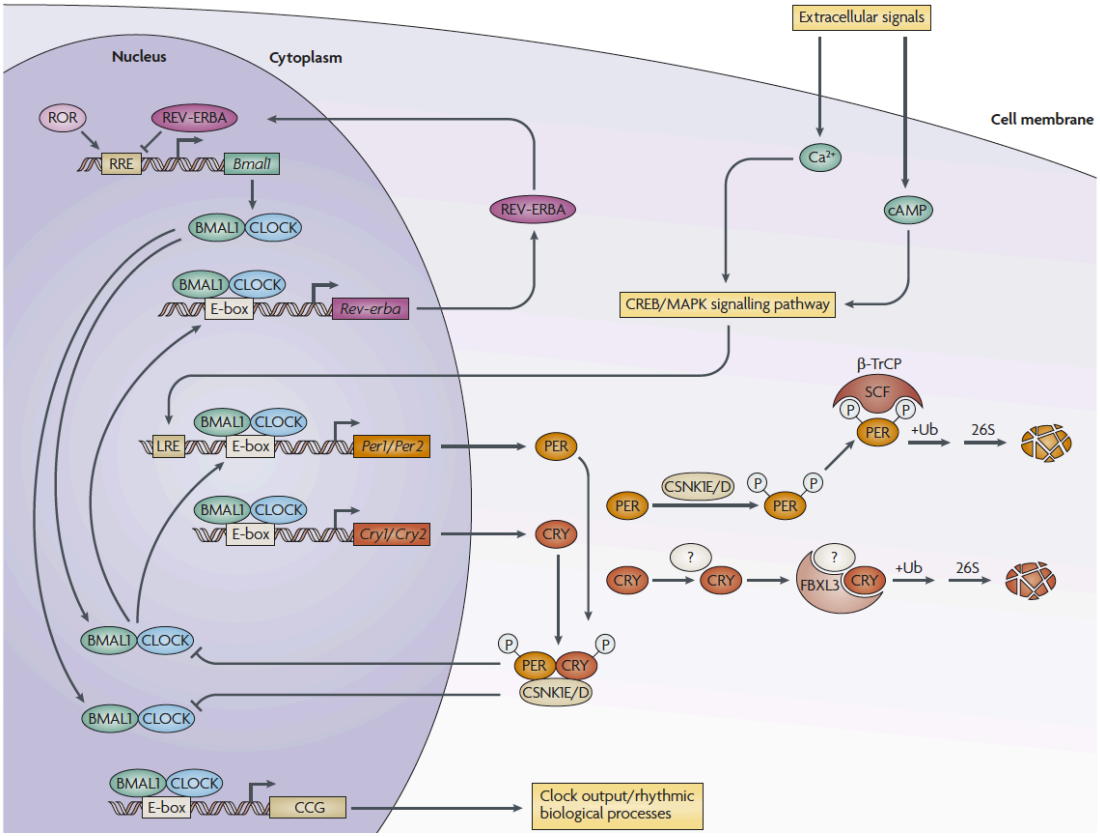


Figure 8. The mammalian circadian clock is composed of a transcriptional–translational feedback network. In mammals, the circadian clock is composed of a primary negative feedback loop involving the genes *Clock*, *Bmal1*, *Per1*, *Per2*, *Cry1* and *Cry2*. CLOCK and BMAL1 are basic helix-loop-helix PAS-domain containing transcription factors that activate transcription of the Per and Cry genes. The resulting PER and CRY proteins heterodimerize, translocate to the nucleus and interact with the CLOCK–BMAL1 complex to inhibit their own transcription. After a period of time, the PER–CRY repressor complex is degraded and CLOCK–BMAL1 can then activate a new cycle of transcription. The secondary autoregulatory feedback loop is composed of *Rev-erba*, which is a direct target of the CLOCK–BMAL1 transcription activator complex. REV-ERBA feeds back to repress Bmal1 transcription and competes with a retinoic acid-related orphan receptor (ROR) to bind ROR response elements (RREs) in the Bmal1 promoter. In addition to the transcriptional activators and repressors, post-translational modification and degradation of circadian clock proteins are crucial steps for determining circadian periodicity. Key kinases for PER (and CRY) phosphorylation are casein kinase 1 delta (CSNK1D) and CSNK1E. One of the roles for phosphorylation of clock proteins is to target them for polyubiquitylation and degradation by the 26S proteosomal pathway. The β-TrCP1 and FBXL3 E3 ubiquitin ligase complexes target the PER and CRY proteins, respectively, for degradation. CCG, clock-controlled genes; CREB, cAMP response element-binding; E-box, CACGTG/T consensus sequence; MAPK, mitogen-activated protein kinase; SCF, SCF E3 ubiquitin ligase; Ub, ubiquitin. Figure from (*131*).

CLOCK–BMAL1 rhythmically bind to E-box *cis*-regulatory elements of multiple genes, the clock controlled genes (ccg) (*137*). This rhythmically binding to their promoter leads to rhythmic transcription and therefore rhythmic levels of numerous mRNAs in the cell. This is the mechanism by which the clock can regulate a variety of cellular and metabolic processes.

3.3. Circadian clock and immunity

43% of all protein coding genes showed circadian rhythms in transcription somewhere in the body, largely in an organ-specific manner (*138*). This clearly shows the importance of the circadian clock regulating multiple physiological functions in the body, among those the immune response. A cell autonomous circadian clock has been found in multiple types of immune cells (*139*), including macrophages (*140*), monocytes (*141*), Th cells (*142*), dendritic cells, and B cells (*143*).

Classic studies in the 1970s showed a remarkable daily variation in survival when mice were infected with pneumococcal bacteria (*144*). There is increased lethality when infection occurs at the end of the rest phase, approximately 2 h before onset of activity. The reason for differences is susceptibility of the mice at different times of the day appears to be due to the daily rhythms in immune response. The immune system is more reactive by the end of the rest phase. The induction of proinflammatory cytokines and chemokine ligands is far greater when mice are challenged with LPS at ZT12 versus ZT0. This includes a greater induction of interleukin-6 (IL6), IL-12 (p40), chemokine (C-C motif) ligand 2 (CCL2), and CCL5 from macrophages challenged at ZT12 versus ZT0 (*145*). These natural rhythms in immune system are probably to prepare the mice for the active phase, when it is more likely to encounter dangers and get wounds. But equally, a reactive immune system corresponds to a greater risk of sepsis resulting from an overactive response.

The time of the day at which mice get infected determines the outcome of the infection. When mice are infected with *Salmonella enterica* Typhimurium at different times, during the day (ZT4) and night (ZT16), the degree of intestinal

inflammation and cytokine gene expression, as well as bacterial colonization, varies. These changes seem to be because the antimicrobial peptide lipocalin-2, to which Salmonella is resistant, is higher during the day than at night. The lipocalin-2 suppresses the resident microbiota, allowing for the massive outgrowth of Salmonella during the day versus the night (*146*).

The SCN and adrenal clock drive rhythms in glucocorticoid secretion from the adrenal glands. The SCN controls the release of adrenocorticotropin (ACTH) from the pituitary gland and ACTH in turn stimulates release of glucocorticoids from the adrenal gland (*147*). Glucocorticoids exhibit broad anti-inflammatory properties and can control cytokine production and leukocyte trafficking (*148*).

Time of day is critical in terms of the type of immune response generated by an organism. The type of immune response can be broadly partitioned into two states, one of high alert and the other of rest and repair.

3.4. Daily rhythms of parasites and microbiota

To this point we have been focusing on the existence of circadian rhythms in free-living organisms that are exposed to dramatic changes, for example light/dark cycles, in their environment. However, organisms living within hosts have also been described to have daily alterations. What remains unknown is whether those daily alterations are due to the circadian clock on those non-free-living organisms, or whether they are simply responding to daily changes within their host.

If the host is a mammal, a non-free-living organism in the bloodstream would be exposed to multiple daily changes: variations in the number and reactivity of patrolling immune cells, temperature, nutrient availability (consequence of the feeding behavior of the host), redox, blood pressure, etc.

In order to be considered an intrinsic circadian rhythm, the daily variations need to have:

i) Self-sustaining periodicity

The rhythms repeat once a day, with a period of 24h, even in the absence of external cues (constant conditions). This criteria allows to distinguish circadian rhythms from simple responses to daily external cues

ii) Entrainment

The rhythms can be adjusted to match the local time (entrainable). When travelling across time zones, human experience jet lag, however, within a few days the sleep/wake cycle and feeding schedule becomes synchronized/entrained to the time of the destination, it resets to match the external stimuli.

iii) Temperature compensation

The rhythms maintain circadian periodicity over a range of physiological temperatures. Some organisms experience a broad range of temperatures in their habitat, and those differences in thermal energy will affect the kinetics of all molecular processes in their cell(s). In order to keep track of time, the organism's circadian clock must maintain a roughly 24-hour periodicity despite the changing kinetics, a property known as temperature compensation.

Daily variations have been observed in parasites, which are eukaryotic organisms living within hosts. Avian *Isospora* intestinal parasites are transmitted when a host ingests sporulated oocysts. In both 12:12 h light/dark of the tropics (*149, 150*) and 24 h light of the artic (*150, 151*), the emergence of the transmissible oocysts in the bird feces peaks in the afternoon, coinciding with evening bird feeding activity. Also observation of rhythmic fluctuations have been made in human parasites. Human species of filariae show daily fluctuation of microfilariae density (the transmissible form) in their hosts' peripheral blood. These parasites can be usually found in the lymphatic system, which leads to the characteristic limb swelling common in filariasis patients. The microfilaria parasites migrate from the lymphatic system to the peripheral blood at night,

which ensures transmission through their vector's blood-meal (*152*). Human malaria parasite species also display synchronous cell cycle durations of 48 or 72 h causing fever every 2 or 3 days, a hallmark of this disease (*153*).

Recent studies showed that the intestinal microbiota, in both mice and humans, exhibits diurnal oscillations that are influenced by feeding rhythms of the animal. This dictates the microbiome composition at different times of the day, leading to daily functional profiles. Ablation of host molecular clock components or induction of jet lag leads to aberrant microbiota diurnal fluctuations, which imposes damaging metabolic consequences for the host (*154, 155*).

Despite such daily fluctuations, it has not been previously established whether non-free-living have intrinsic clocks that anticipate the changes in the environment or whether they are simply responding to host rhythmic physiological cues.

4. Aims

The main goal of this work has been to test the hypothesis that the sleep/wake cycle disruption observed in sleeping sickness patients is due to changes in expression of circadian clock genes.

First, we carefully characterized the C57BL/6J mouse model of sleeping sickness in order determine which organs carry with higher parasite load. A careful parasite distribution analysis would allow correlations with the symptoms observed and potentially with circadian clock disruption. From these studies arose the surprising conclusion that the so-far neglected adipose tissue is in fact the organ with higher parasite load, later in infection. With further analyses we uncovered a parasite adaptation to the adipose tissue lipid-rich environment by activating the beta-oxidation of fatty acids, a pathway not previously documented to be used by this parasite.

With the parasite distribution in the mouse, we proceeded to test whether this model mimicked the sleep-wake cycle disruptions observed in humans, followed by the characterization of the clock gene expression in infected mice.

Finally, we sought to determine whether *T. brucei* parasites have intrinsic circadian rhythms through analysis of their daily gene expression *in vitro*, by RNA sequencing.

These data open new avenues for the research on the pathogen responsible for the fatal sleeping sickness, as well as for the studying the existence of circadian rhythms in other pathogens.

CHAPTER II

CHAPTER II – Results

In the chapter II, I documented the result section of my research projects in the form of three stories:

1. *Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice

2. *Trypanosoma brucei* infection accelerates the mouse circadian rhythm

3. Post-Transcriptional Circadian Regulation of Genes Related to Metabolism and Suramin Sensitivity in *Trypanosoma brucei*

Chapter II – section 1

Trypanosoma brucei parasites occupy and functionally adapt to the adipose tissue in mice

The "Chapter II – Results – 1. *Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice" consists of a co-first authorship publication. Here is clarified the contribution of each of the authors:

Publication:

> Sandra Trindade[#] & **Filipa Rijo-Ferreira**[#], Tânia Carvalho, Daniel Pinto-Neves, Fabien Guegan, Francisco Aresta-Branco, Fabio Bento, Simon A. Young, Andreia Pinto, Jan Van Den Abbeele, Ruy M. Ribeiro, Sérgio Dias, Terry K. Smith, Luisa M. Figueiredo 2016. *Trypanosoma brucei* parasites occupy and functionally adapt to host adipose tissue. *Cell Host&Microbe* (in press)

[1]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1990-375 Lisboa, Portugal
[2]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9111, United States
[3]Graduate Program in Areas of Basic and Applied Biology, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal
[4]Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, United States
[5]Department of Biomedical Sciences, Unit of Veterinary Protozoology, Institute of Tropical Medicine Antwerp, B-2000 Antwerp, Belgium.
[6]Department of Physiology, Laboratory of Zoophysiology, University of Ghent, B-9000 Ghent, Belgium
[7]Biomedical Sciences Research Complex, University of St Andrews, North Haugh, St Andrews, Fife, KY16 9ST, United Kingdom

* Corresponding author
E-mail: lmf@medicina.ulisboa.pt

[#]These authors contributed equally to this work
[$] Guest Professor of Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal

Filipa Rijo-Ferreira, the candidate for the PhD degree at the University of Porto, contributed to the paper by initially establishing single-handed a mouse model of *Trypanosoma brucei* infection in my laboratory. Filipa also contributed by designing original experiments that led to the publication of this paper. As part of this initial work, she observed that parasites preferentially accumulate in

the adipose tissue, characterizing this accumulation by histological analysis, genomic DNA and RNA quantifications. She demonstrated that throughout infection there is a gradual increase in the parasite load in many organs, including the brain, but that is in the adipose tissue where more parasites can be found. Filipa also designed and prepared the RNA samples for RNA sequencing, as she was further involved in the planning of the transplantation experiments presented in the publication.

The contribution of each author to the publication is clarified in the section “Authors contribution” with the following text:

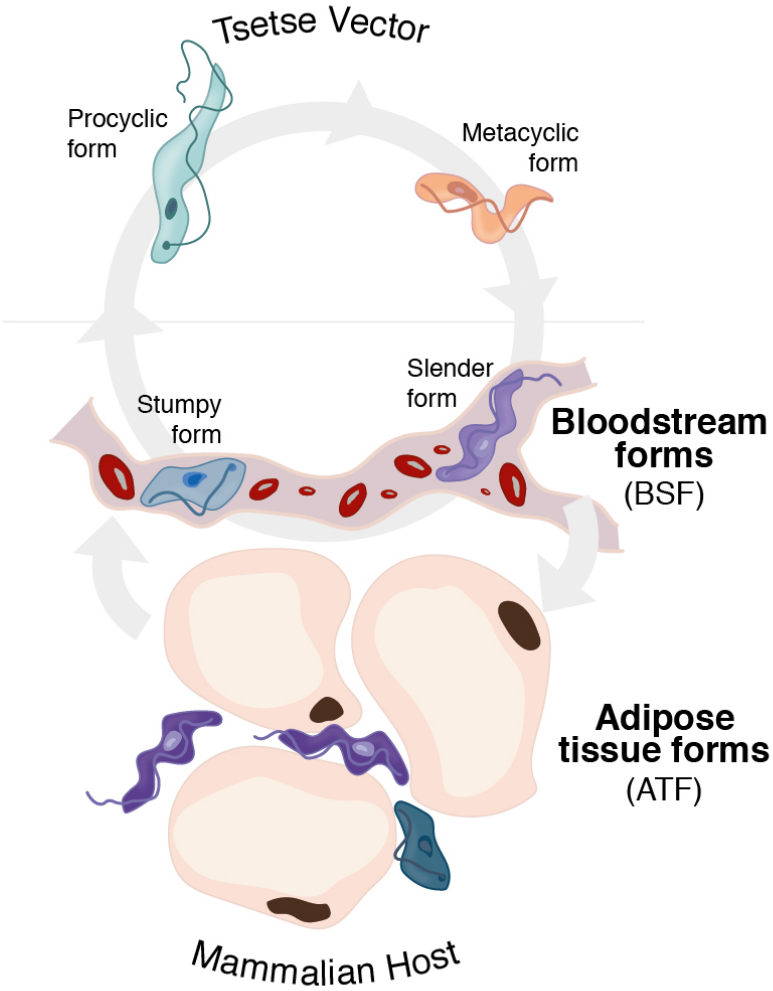
“Authors contribution: S.T., F.R.F., T.K.S. and L.M.F. designed the experiments and wrote the paper; S.T., F.R.F., T.C., F.A.-B., F.B., A.P., T.K.S. conducted the experiments; D.P.N. and S.A.Y. analyzed the RNA-seq data; J.V.D.A. designed and conducted experiments with Tsetse flies; R.M.R. conducted the statistical analysis; S.D. designed experiments.”

1. *Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice

Keywords: African trypanosomes, fat, mouse infection, fatty acid b-oxidation, metabolism, transcriptome

1.1. Summary

Trypanosoma brucei is an extracellular parasite that causes sleeping sickness. In mammalian hosts, trypanosomes are thought to exist in two major niches: early in infection they populate the blood; later, they breach the blood brain barrier, at which point the disease is terminal. Here we report the unexpected finding that adipose tissue is a third major functional reservoir for *T. brucei*. We show that adipose tissue contains replicative parasites capable of infecting a naïve mouse. We also demonstrate they are transcriptionally very different from parasites in the blood, and that these differences are functionally significant, with parasites in fat having adapted their metabolism to the utilization of lipids as a carbon source. Together, these findings identify the adipose tissue as a novel niche, which is populated by a new form of *T. brucei* life cycle (which we termed ATF) and which has potential consequences for pathogenesis.



1.2. Introduction

Human African Trypanosomiasis (HAT), also known as sleeping sickness, is a neglected tropical disease that is almost always fatal if left untreated. This disease is caused by *Trypanosoma brucei,* a unicellular parasite that lives in the blood, lymphatic system and interstitial spaces of organs (reviewed in(*16*)).

Disease pathology often correlates with sites of accumulation of the infectious agent within its host, including the brain, which is associated with characteristic neuropsychiatric symptoms and sleep disorder. Weight loss is another typical clinical feature of sleeping sickness pathology (*16*), but is essentially unstudied.

T. brucei is transmitted through the bite of a tsetse and quickly adapts to the mammalian host to become what is known as a “slender” bloodstream form (BSF). As parasitemia increases, slender forms are capable of sensing population density and this triggers differentiation to the stumpy form, which is pre-adapted to life in the transmitting tsetse vector and, once there, further differentiates into procyclic form (PCF). Several studies have shown 10-30% of genes being differentially expressed between BSF and PCF (reviewed in(*156*)), including genes involved in metabolism, organelle activity, cell cycle regulation and endocytic activity. Recent proteomic studies also revealed around 33% of proteins that are developmentally regulated(*157*).

A major difference between BSF and PCF is their energy production, with the former utilizing glucose via glycolysis within the glycosome and the latter utilizing proline and to a lesser extent other amino acids as their carbon source, via the Krebs cycle in its mitochondrion (reviewed in (*158*)). To date, no fatty acid β-oxidation has been observed as a carbon source in any life cycle stage of this parasite. This has been a puzzling observation, as the genes required for productive β-oxidation, including the carnitine-acyltransferases (for mitochondrial import of fatty acids) are present in the genome.

Here, we describe a new form of *T. brucei* in mammalian hosts: we demonstrate that *T. brucei* accumulates in adipose tissue (consistent with recent studies showing accumulation of parasites in the lower abdomen (*55, 159, 160*) and that this accumulation is functional. Namely, fat resident *T. brucei* have a different metabolic profile from either slender or stumpy forms in the blood, and this profile is consistent with their utilization of fatty acids (Myristic Acid) as a carbon source. These experiments describe a new form of *T. brucei* life cycle, and possibly explain weight loss (wasting), one of the characteristic pathological features of this disease.

1.3. Results

T. brucei parasites are heterogeneously distributed in mice

The well-established mouse model (C57BL/6J mice with a pleomorphic clone AnTat1.1^{E}) was used to confirm weight loss during infection as observed

in humans with sleeping sickness. Parasitemia followed a previously described pattern: the first peak of parasitemia occurred 5-6 days post-infection, at around $2x10^8$ parasites/mL and, after approximately 4 days of undetectable parasitemia, parasites could be detected again with a fluctuating parasitemia of 10^6 - 10^7 parasites/mL (Figure 1A). After the first peak of parasitemia, all infected animals showed reduced food intake and a 10-15% decrease in body weight. Eventually all mice recovered normal food intake, although their body weight remained 5% lower than that of non-infected animals (Figure 1B and S1A). The weight of most organs from mice sacrificed on 6 and 28 post-infection showed minimal changes relative to day 0, except fat depots that decreased on average 43% (±12%). Spleen size and weight increased dramatically as previously reported (Figure 1C). Infected mice died 35 ± 2.5 days post-infection (Figure 1D).

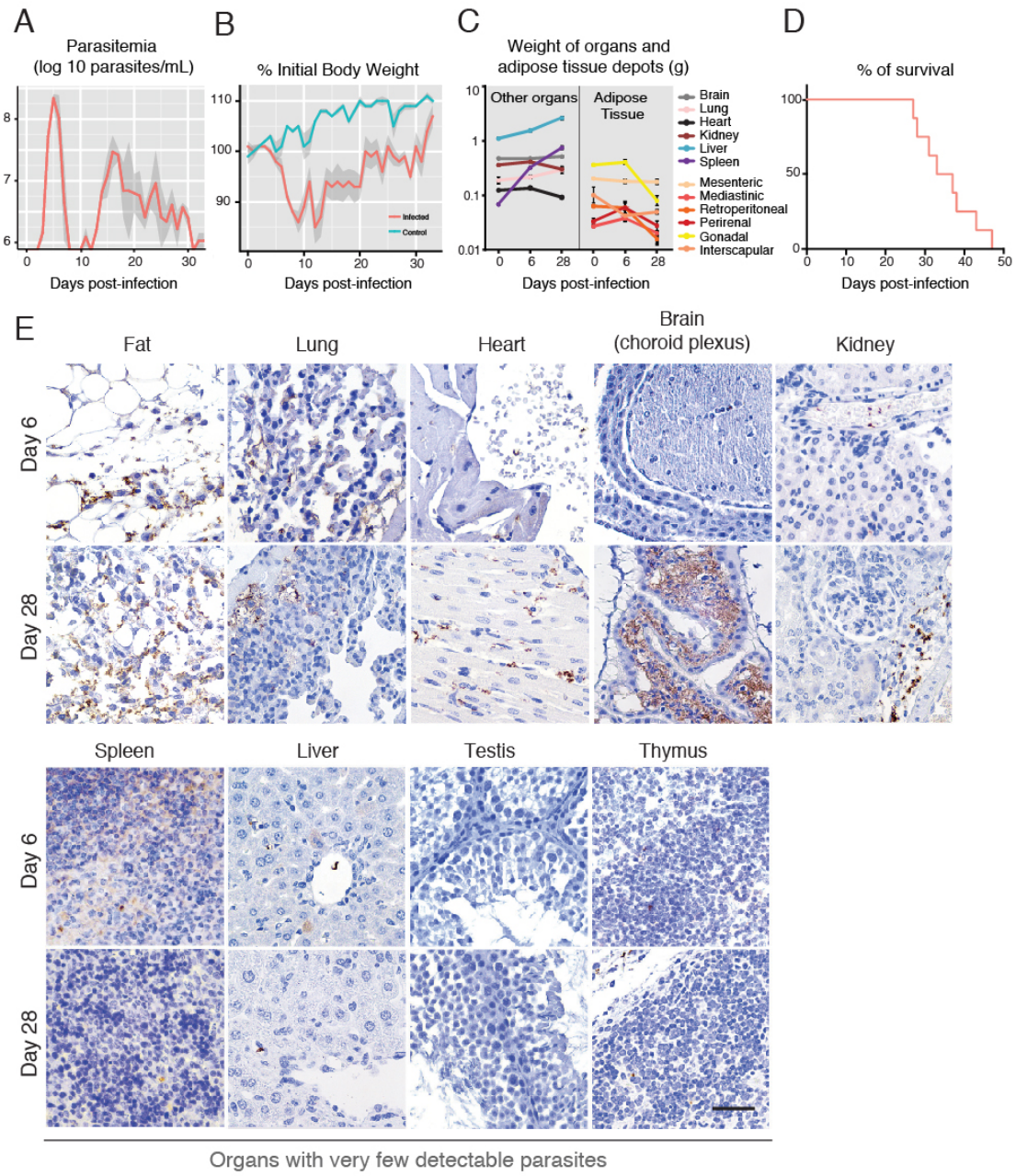


Figure 1. Tissue distribution of *T. brucei* during a mouse infection is heterogeneous. (A) Mean parasitemia profile of 20 mice infected with *T. brucei* AnTat1.1^{E}. Parasitemia was assessed from tail-blood using a hemocytometer (limit of detection is around 4x10^{5} parasites/mL). Light grey shaded area represents SEM. (B) Variation of body weight during infection. Daily body weight measurement of control and infected mice (n = 15 per group). Light grey shaded area represents SEM. (C) Variation of organ weight during infection (n = 4 per group). (D) Survival curve of *T. brucei* infected mice (red curve) (n = 8). (E) Representative light microscopic images of *T. brucei* distribution in several organs/tissues, at days 6 and 28 post-infection, assessed by immunohistochemistry with a non-purified rabbit anti-VSG antibody (parasites appear brown). n = 5 per time point. Original magnification, 40x.

To assess the parasite load in different organs, we used immunohistochemistry at different days of infection (Figure 1E and Figure S1B). Parasites were consistently detected in the fat 6 days post-infection and at later time points, while in other organs they were seen sporadically and at very low densities. As infection progressed, we observed an increase in parasite load in most organs, with fat, heart, brain, lung and kidney being the most visibly infiltrated. Parasites were always found extracellularly within the interstitium of these organs. In the brain, our data corroborate the extensively reported evidence for the localization of parasites being restricted to the choroid plexuses and meninges (*16*) (Figure 1E).

Histologically, thymus, lymph nodes, bone marrow, skin, salivary glands, spleen, gastrointestinal mucosa, testis and liver displayed few or no parasites (Figure 1E and data not shown). Although parasites in the stroma of the testis were absent, the epidymal fat body and stroma of the epididymis (a small paired organ in the posterior end of the testis) contained many parasites many of which appeared as debris, but which could explain the bioluminescence detected by (*55*).

Early in infection, *T. brucei* accumulate in fat

Immunohistochemical staining showed parasites in the stroma of several fat depots: gonadal, mediastinic, mesenteric, retroperitoneal, perirenal, and interscapular (Figure 2A). TEM confirmed that these parasites were indeed extravascular as numerous trypanosomes were observed in the interstitial space, either between adjacent adipocytes or between the adipocytes and the capillaries (Figure 2B).

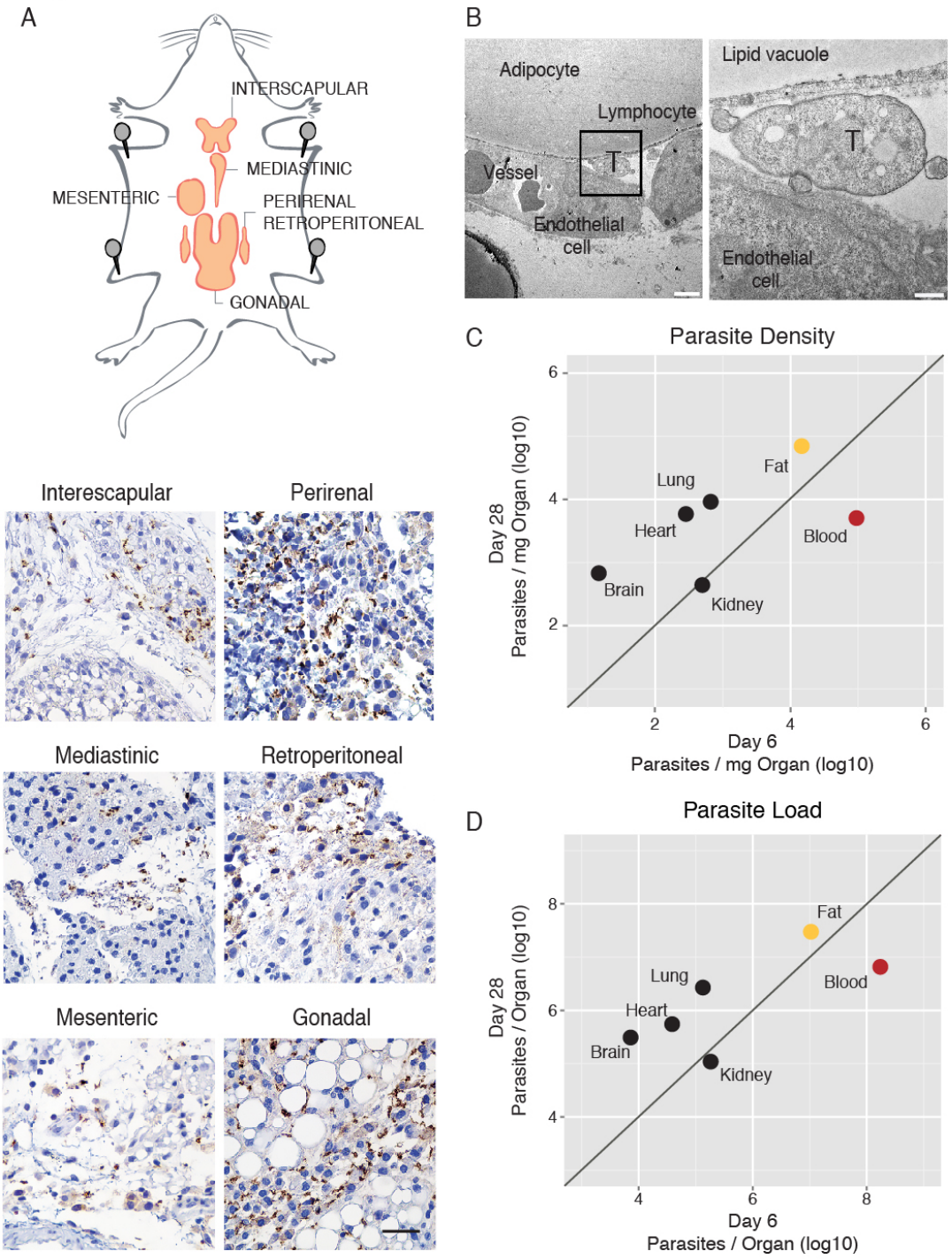


Figure 2. Fat depots are a major parasite reservoir. (A) Schematic representation of mice fat depots and Anti-VSG immunohistochemistry images of six different fat depots, collected 28 days post infection. Original magnification, 40x. (B) Transmission electron micrograph of a gonadal fat depot 6 days post infection. Trypanosome (T) and lymphocyte in the interstitial space, adjacent to an adipocyte and next to a small capillary. Scale bars represent 2 and 0.5 µm in the left and right panels, respectively. (C) Parasite density in multiple organs/tissues (6 and 28 days post-infection) was measured by qPCR of gDNA (*T. brucei* 18s rDNA normalized to the weight of the tissue/organ). Blood density was assumed 1.05 g/mL. Fat value is the average of quantification of the six depots indicated in panel A. Each point represents the

geometric mean of the parasite density on days 6 (n = 3-9) and 28 post-infection (n = 3-6). (D) Parasite load in multiple organs/tissues estimated by multiplying parasite density with organ weight at the corresponding day of infection. Each point represents the geometric mean of the parasite density on days 6 (n =3-9) and 28 post-infection (n = 3-6).

To quantify parasite density, we used as a proxy Trypanosome DNA, which was quantified at 6 and 28 days post-infection in the organs/tissues where parasites had been detected by histology, i.e. fat, lung, heart, kidney, brain and blood (Figure 2C and Figure S2A). The blood had the highest parasite density on day 6. Among solid organs/tissues, for the same day of infection, parasite density is relatively low, except for fat that has on average 60-fold more parasites than lung, heart, kidney and brain and 7-fold less than blood. On day 28 of infection, parasite density remained equally high in fat (10^4-10^5 parasites/mg), while it increased, on average 20-fold in the brain, heart and lung (Figure 2C). The overall high parasite density was detected in all fat depots characterized in this study, with no significant differences between white and brown adipose depots (Figure S2B and Supplemental Experimental Procedures). The blood is the only site where we observed a decrease in parasite density during infection, which is consistent with parasitemia dynamics (Figures 1A and 2C). As a consequence, on day 28, fat is the compartment with the highest parasite density (LME, $p < 0.0001$).

Overall the density of parasites per mg of organ/tissue correlated well with the density calculated as a ratio of parasite DNA versus mouse DNA in each tissue (Figure S2C). We also observed essentially the same pattern of parasite density and the same preferential accumulation in the fat when we quantified parasite RNA (qPCR), instead of DNA (Figure S2D), suggesting that DNA quantification reflects accurately the number of live parasites. Immunohistochemistry also showed accumulation of parasites in fat regardless of parasite strain (EATRO1125 AnTat1.1^E, Lister 427), infection route (intraperitoneal, intravascular), mouse strains (C57BL/6J, BALB/c), animal gender (male, female) or rodent species (mice, rat) (Figure S2E).

Fat represents around 14% of the body weight of a healthy mouse (Jackson Mouse Phenome Database), thus it is potentially a very large reservoir of parasites. The number of parasites in the organs/tissues (parasite load) was

determined by multiplying parasite density by the weight of the organ at the corresponding time of infection. For fat, we used the weight of the six depots characterized in this work, which comprises around 25% of the total body fat. We observed that, while 6 days post-infection, the blood contains the majority of the parasites (around 10^8 parasites), on day 28 the six depots of fat contain overall more parasites than the blood, brain and all other tested organs combined (LME, $p < 0.0001$) (Figure 2D). A similar preference for accumulation in adipose tissue was observed when the mouse infection was initiated by a tsetse bite, which deposits metacyclic forms in the skin of the mouse (Figure S3) (LME, $p < 0.0001$). Overall, these data revealed that fat represents a major reservoir of parasites, regardless of whether the infection was initiated by BSF or metacyclic forms.

Adipose tissue contains replicative and infective parasites

In the blood, parasites can be either replicative slender or G1-arrested stumpy forms (*161*). To investigate whether the parasites from fat (referred hereafter to as adipose tissue forms (ATF)) are replicative or not, we infected mice with a *GFP::PAD1$_{utr}$* reporter cell-line, in which a GFP gene is followed by a *PAD1* 3'UTR that confers maximum expression in stumpy forms (J Sunter, A Schwede and M Carrington, unpublished; (*162*)). Four and six days post-infection, blood and fat were collected and parasites isolated and purified. As described by(*163*), on day 4 we observed that most parasites in the blood were GFP-negative (98% ± 0.3%); while on day 6 most parasites expressed GFP (86% ± 2.6%). Interestingly, we noted that intermediate forms can also be detected in the blood as a transient population of intermediate GFP-intensity around day 5 post-infection (data not shown). As the yield of isolation of ATF parasites was very low on day 4 and 5, we analyzed these parasites only on day 6. The majority of the ATF parasites were GFP-negative (79% ± 4.6%), while 21% ± 4.6% expressed GFP, indicating that fat contains fewer stumpy forms than blood (Figure 3A).

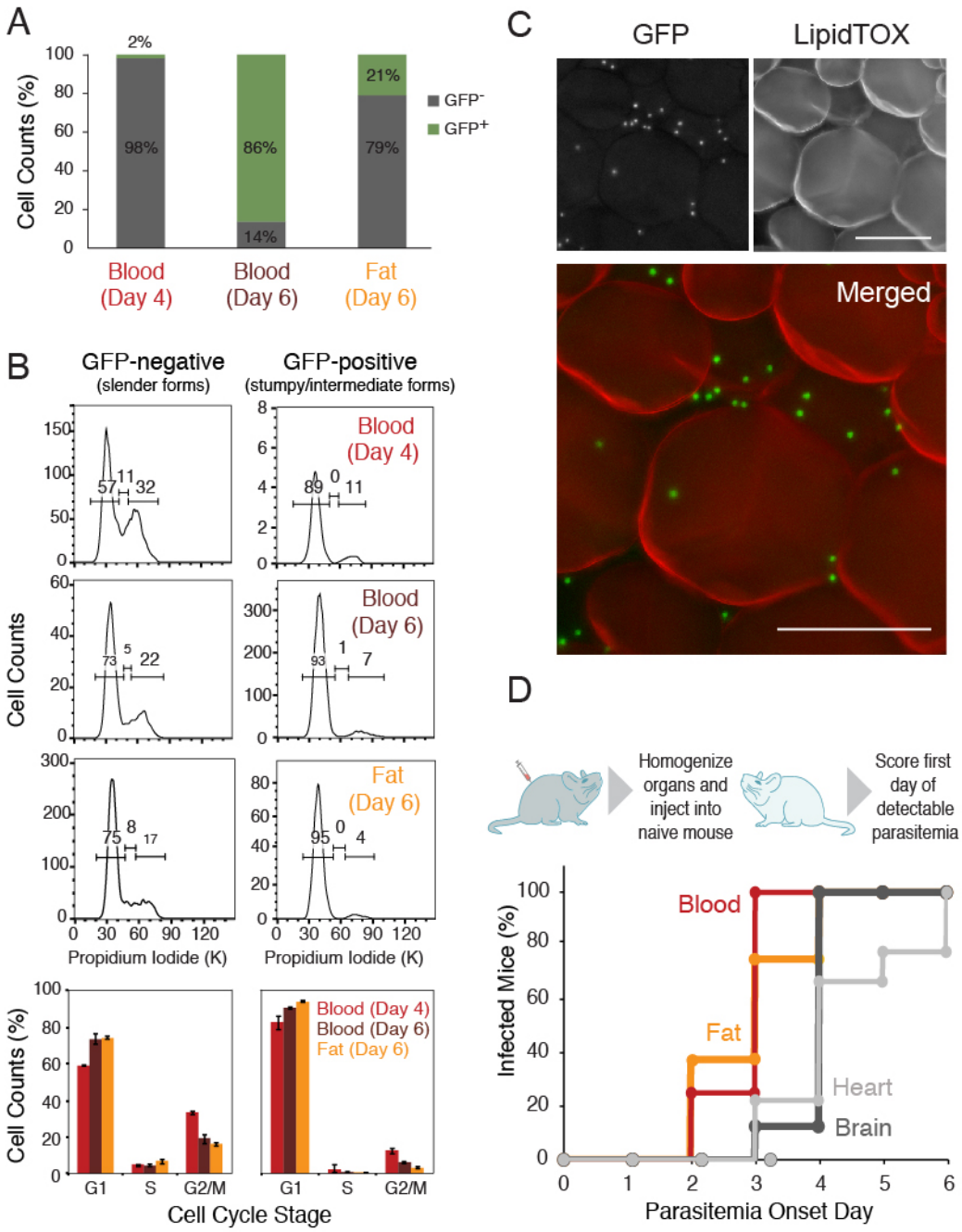


Figure 3. Adipose tissue harbors replicative forms that can establish a new infection. (A) Frequency of GFP expression measured by flow cytometry in parasites isolated from blood and fat, 4 and 6 days post-infection with a *GFP::PAD1*$_{utr}$ *T. brucei* reporter cell-line (n = 2-3). (B) Cell cycle analysis assayed by flow cytometry of propidium iodide stained parasites (n = 2-3). (C) Fluorescence microscopy of gonadal adipose tissue from a mouse infected for 6 days with *GFP::PAD1*$_{utr.}$ Lipid droplets were stained with LipidTOX (red) and nuclei of GFP-expressing parasites (stumpy forms) are green. (D) Onset of parasitemia curves in mice that were injected i.p. with infected organs/tissues lysates from a first mouse. Lysates from blood, heart, brain and gonadal fat depot were prepared from mice sacrificed between 21-28 days post-infection to ensure presence of a larger number of parasites (n = 9).

To confirm if GFP-negative ATF parasites are replicative and GFP-positive parasites are cell cycle arrested, we stained the parasite nuclear DNA with propidium iodide and quantified it by flow cytometry. In all samples (blood day 4, blood day 6 and fat day 6) we observed that GFP-negative parasites displayed a cell cycle profile characteristic of replicative cells (around 60-70% of cells in G1, 5% in S-phase and 20-30% in G2/M), while GFP-positive cells were cell cycle arrested in G1/G0 (90-95%) (Figure 3B). Similar data were obtained by performing cell cycle analysis with DyeCycle Violet (Figures S4A and S4B), further confirming the presence of slender and stumpy forms in adipose tissue.

To validate the presence of stumpy forms in adipose tissue, we used fluorescence microscopy on an intact gonadal depot infected with *GFP::PAD1*$_{utr}$ expressing parasites (Figure 3C). LipidTOX stains the lipids in the large lipid droplet of adipocytes. Among the adipocytes, we could clearly observe many green foci, which represent the parasite nuclei where GFP accumulates, thus confirming the presence of GFP-positive parasites (stumpy forms) in close proximity to adipocytes. The presence of replicating parasites in intact tissue was confirmed by immunohistochemistry with an anti-H2A antibody. Dividing nuclei were clearly visible in close proximity to adipocytes (Figure S4C), further confirming the FACS cell cycle data.

To test whether ATF parasites are capable of establishing a new infection, an infected donor mouse was sacrificed, perfused and several organs collected, homogenized and injected intraperitoneally into recipient naïve mice. Parasitemia was assessed daily thereafter, and scored on the first day it became detectable (Figure 3D). Mice that received blood or a fat homogenate showed parasitemia earlier (around 3 days post-transplantation) than animals injected with heart and brain homogenates (around 4 days post-transplantation), consistent with the observed parasite load in these organs (Figure 2D and Figure 3D). Transplant of intact gonadal fat depot also led to successful infection of the recipient naïve mice (data not shown); suggesting that parasites can exit from an intact tissue. These results showed that parasites from fat, heart and brain are capable of reinvading the bloodstream and establishing a new infection.

Morphology of Adipose Tissue Forms

Although T. brucei is always extracellular, its morphology changes during the life cycle, which may reflect a specific adaptation to the host niche (*164*). ATF, like BSF, have an ondulating appearance, with a flagellum attached to the cell body. Kinetoplast is positioned between nucleus and flagellar pocket.

To characterize the morphology of ATF, we compared the length and width of GFP::PAD1utr parasites isolated from adipose tissue (day 6 post-infection) and blood (day 4 and 6 post-infection) (Figure 4A and 4B). Automatic measurements of phase contrast microcopy images were generated via HTIAoT (*165*) and confirmed with manual measurements (Figure S5A). We observed that slender forms (GFP-negative) from blood and adipose tissue were very similar both in length (blood: 24.39 ± 2.50 µm; fat: 24.57 ± 2.99 µm) and width (blood: 2.15 ± 0.26 µm; fat: 2.12 ± 0.26 µm). This average length is consistent with previous reports (*166, 167*). In blood day 6 post-infection, GFP-positive parasites were, as expected, shorter (18.43 ± 1.81 µm) and wider (3.11 ± 0.38 µm) than slender counterparts of day 4, corresponding to the morphology of stumpy forms (*168*). Interestingly, in adipose tissue, we found not only stumpy forms, but also GFP-positive parasites that morphologically were in-between slender and stumpy forms (length: 21.32 ± 2.73 µm; width: 2.29 ± 0.31 µm). These probably correspond to the previously described blood intermediate forms, which as the name suggests, are not fully differentiated into stumpy forms, but that already express GFP::PAD1 (*162*). These results indicate that adipose tissue is populated by parasites whose morphology has been previously found in the blood. The only significant difference is their relative distribution: on day 6 of infection, while blood is mostly populated by stumpy forms, adipose tissue appears to be “delayed” as we detected both intermediate and stumpy forms.

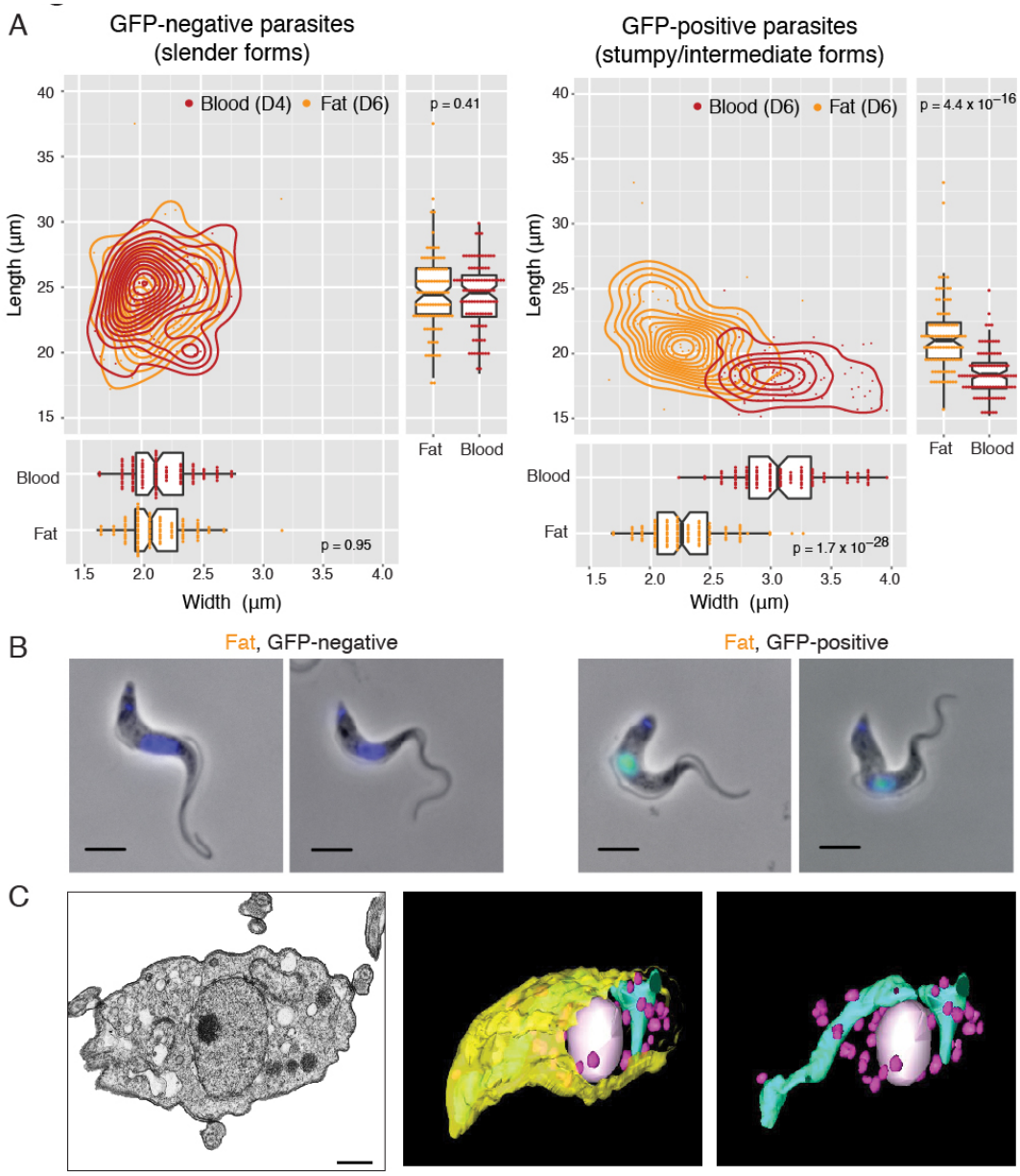


Figure 4. ATF parasites are morphologically similar to parasites from the blood. (A) Morphological features (length and width) of fixed parasites isolated from fat and blood of mice infected with GFP::PAD1utr reporter. Fat gonadal tissue was collected on day 6 post-infection. The blood "controls" were obtained as follows: GFP-negative parasites were collected on day 4 post-infection (mostly slender forms); GFP-positive parasites were collected on day 6 post-infection (mostly stumpy forms). Morphometric measurements were scored from phase contrast microscopy images and analyzed via HTIAoT and confirmed by manual measurement. GFP-negative: slender form; GFPpositive: stumpy and intermediate forms. n = 100 per group. Statistical significance was assessed using a Wilcoxon rank sum test. (B) Representative images of parasites isolated from fat. Replicating parasites (such as the second from the left) were excluded from morphometric analysis. (C) Transmission electron micrograph and 3D Tomography images of a parasite isolated from gonadal adipose tissue. Mitochondrion is represented in blue/green, glycosomes in pink, nucleus in white and plasma membrane in yellow. Scale bar represents 500 nm.

At the ultrastructure level, ATF contain all major structures described in other stages of life cycle (*169*), including an electron dense coat, nucleus, mitochondrion, endoplasmic reticulum, Golgi apparatus, glycosomes, dense granules and numerous vesicles compatible with endosomes, an internal subpellicular corset of microtubules underneath plasma membrane and a flagellum attached to the cell body (Figure 4C and Figure S5B). Using serial 3D-Tomography, we observed that the single mitochondrion of ATF occupies a small volume of parasite body and it is not highly branched (Figure 4C and Supplemental Movie). This organization was confirmed by Mitotracker Green staining, which showed no major differences between the mitochondrion of parasites in blood and adipose tissue (Figure S5C and Supplemental Movie).

Transcriptome of ATF parasites reveals differences in several key regulatory processes

During its life cycle, *T. brucei* adapts to its environment by changing gene expression (reviewed in(*170*). To test whether parasites within fat also adapt to the lipid-rich environment, total RNA was extracted from infected gonadal fat depot (n = 3) on day 6, along with parasites from blood (n = 2) on day 4 (maximizing slender and minimizing stumpy forms), and was subjected to RNA-seq analysis. As expected, sequence reads from blood samples corresponded mainly to parasite transcripts, while sequence reads from fat corresponded mainly to host transcripts. Nevertheless, the 1% to 9% of the sequence reads from *T. brucei* provided enough statistical power to detect changes in the transcriptomes of ATF parasites (Table S1). Two previously published RNA-seq datasets of BSF parasites grown in culture were also included in this analysis (*171*).

Unbiased clustering of gene expression profiles revealed that ATF parasites cluster separately from parasites isolated from blood or culture (Figure 5A), suggesting significant changes in their overall transcriptome. Changes were identified using three methods of differential expression analysis, and only those genes identified by at least two methods with an adjusted p value < 0.01

were considered. These analyses showed that 2,328 genes (around 20% of transcriptome) were differentially expressed between BSF and ATF parasites: 1,160 were up-regulated in ATF parasites and 1,178 were up-regulated in BSF (Figure 5B-5C and Table S2).

Significant changes were found in genes involved in gene expression regulation, cell cycle and cell signaling (Table 1). RNA binding proteins play an important role in gene expression and differentiation throughout the *T. brucei* life cycle. RNA binding protein 42 (RBP42, Tb927.6.4440) binds many mRNAs involved in cellular energy metabolism (*172*). Up-regulation of RBP42 in ATF parasites could be involved in the metabolic rewiring when parasites enter the fat. Overexpression of the RNA helicase DHH1 leads to up-regulation of RBP5 in BSF (*173*). Interestingly, ATF parasites showed a simultaneous up-regulation of both genes (11 and 7-fold, respectively) (Table 1).

ATF parasites transcriptome also showed dramatic changes in gene expression of various post- and co-translational modifying enzymes that might have considerable influence on diverse cellular processes (Table 1). A small number of genes potentially acting in the cell cycle and cytokinesis was identified with significant differential expression, including the cytoskeleton associated AIR9 protein and spastin, which were up-regulated (Table 1), suggesting differences in cell cycle regulation in these parasites. Consistent with a minor stumpy form population in fat, we did not find enrichment of stumpy specific genes in the transcriptome of ATF parasites. Extracellular signaling mechanisms also seem to be affected in ATF parasites, including up-regulation of TOR3, which can relate the supply of external nutrients to internal energy levels to regulate cellular growth (*174*).

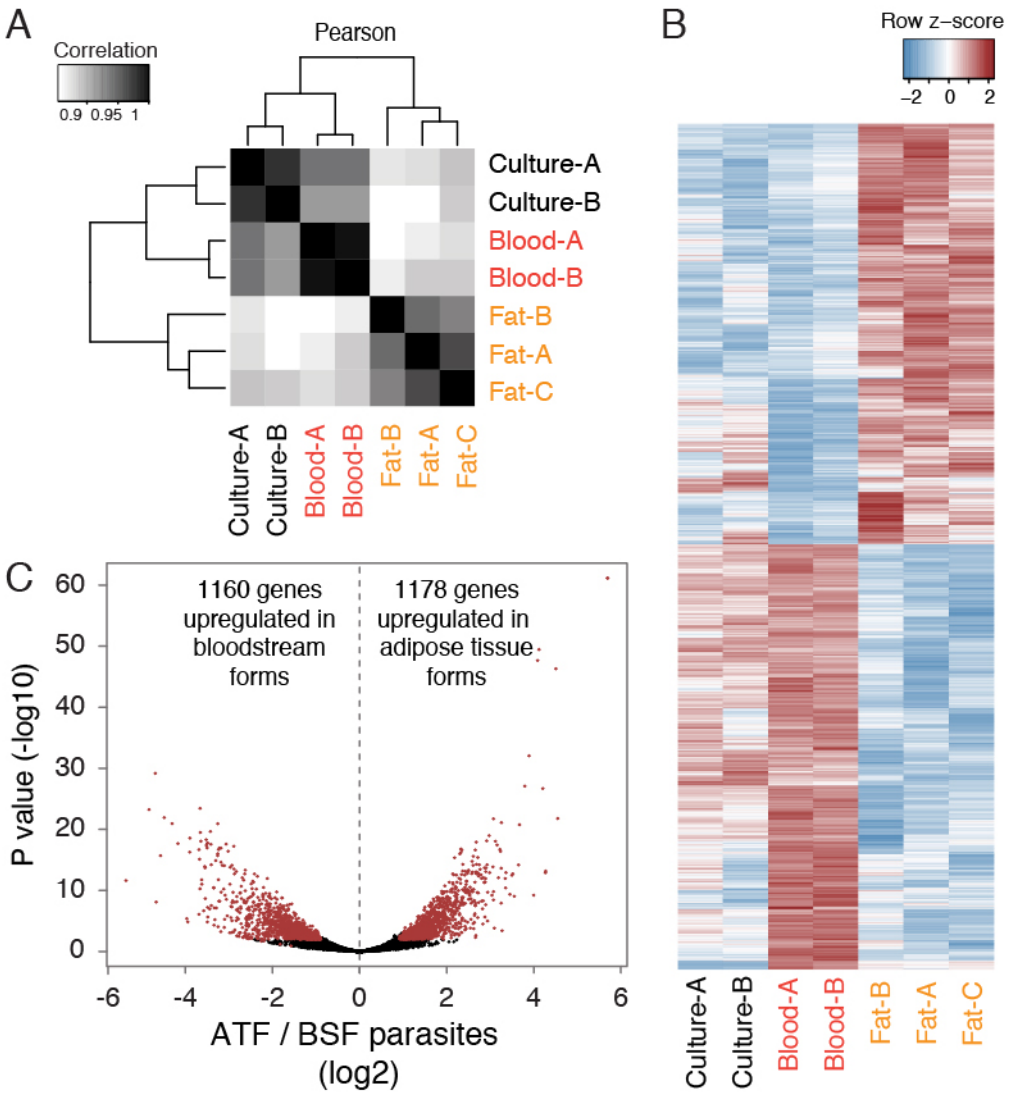


Figure 5. ATF parasites are transcriptionally different from bloodstream forms. (A) Hierarchically clustered heat map of Pearson correlations of transcript levels (log_2 transformed RPKM)) from independent RNA-seq datasets: Lister427 parasites grown in culture (n = 2); parasites isolated from blood of AnTat1.1-infected mice on day 4 post-infection (n = 2); parasites isolated from gonadal fat on day 6 post-infection (n = 3). (B) Heat map view of relative transcript levels for differentially expressed genes from culture and *in vivo* in parasites isolated from the two tissues (adjusted p < 0.01 in at least two of three methods). (C) Volcano plot displaying in red the differentially expressed genes represented in (B). Displayed p values and fold-changes are from DESeq2.

Interestingly, although by TEM, an electron dense coat can be observed at the parasite surface (Figure 2B), we found that the transcripts levels of the active VSG (VSG AnTat1.1, CAA25971.1) are 3-fold down-regulated in adipose tissue. As the VSGnome of AnTat1.1E clone is currently unknown, we could not test whether silent VSGs were up-regulated as a compensatory mechanism.

Genes encoding for other surface molecules, such as the haptoglobin-receptor and most procylins, were not differentially expressed.

Transcriptome of ATF parasites reveals metabolic adaptations

One of the most evident changes in ATF transcriptome is the up-regulation of many metabolic pathways including glycolysis, pentose phosphate, purine salvage, sterol and lipid metabolism and, surprisingly, β-oxidation. 13 of the 14 enzymatic steps of glycolysis are up-regulated relative to BSF (Table 1). This may either be a response to the lower glucose concentration in fat interstitial fluid relative to bloodstream or alternatively, ATF parasites may have up-regulated gluconeogenesis, which relies mostly on the same enzymes.

In ATF parasites, genes involved in 3 out of the 5 biosynthetic steps in the pentose-phosphate pathway were up-regulated, including the rate-limiting glucose-6-phosphate dehydrogenase (Tb927.10.2490). This observation, taken together with the fact that numerous enzymes (16 in total) involved in purine salvage are also up-regulated suggests that ATF parasites may increase purine production. Interestingly, the purine phosphatases (Tb927.8.3800 and Tb927.7.1930) and cAMP phosphodiesterase PDEA (Tb927.10.13000) are up and down-regulated, respectively, suggesting that the increased purine production may be directed towards cAMP signaling (Table 1).

Table 1. Classes of transcripts differentially expressed between parasites in adipose tissue and blood.

	Gene ID	Annotation	FAT / Blood (Log2 FC)*
Glycolysis	Tb927.10.2010	Hexokinase (HK1)	2.5
	Tb927.10.2020	Hexokinase (HK2)	2.8
	Tb927.3.3270	ATP-dependent phosphofructokinase	1.6[a]
	Tb927.10.5620	Fructose-bisphosphate aldolase, glycosomal	1.9[a]
	Tb927.11.5520	Triosephosphate isomerase	1,5
	Tb927.6.4280	GAPDH, glycosomal	1.3[a]
	Tb927.6.4300	GAPDH, glycosomal	1.3[a]
	Tb927.10.6880	GAPDH, cytosolic	1.4
	Tb927.9.9820	Glyceraldehyde-3-phosphate dehydrogenase, flagellar	-1.3
	Tb927.1.710	Phosphoglycerate kinase (PGKB)	1.4
	Tb927.1.700	Phosphoglycerate kinase (PGKC)	1.8
	Tb927.10.2890	Enolase	1.7[a]
	Tb927.11.6280	Pyruvate phosphate dikinase	2,6
	Tb927.10.12700	Pyruvate dehydrogenase E1 alpha subunit, putative	1.6[a]
	Tb927.3.1790	Pyruvate dehydrogenase E1 beta subunit	2.9
	Tb927.10.7570	Dihydrolipoamide acetyltransferase E2 subunit, putative	1.3[a]
	Tb927.11.16730	Dihydrolipoamide dehydrogenase (GCVL-2)	2.5
	Tb927.3.4390	Dihydrolipoamide dehydrogenase (GCVL-1)	-1.2
Pentose Phosphate Pathway	Tb927.10.2490	Glucose-6-phosphate 1-dehydrogenase	1.6
	Tb927.9.12110	6-phosphogluconate dehydrogenase, decarboxylating	1.2[a]
	Tb927.7.5680	Deoxyribose-phosphate aldolase, putative	2.4
	Tb927.11.3030	Phosphoribosylpyrophosphate synthetase, putative	1.9
Purine Salvage	Tb927.7.7040	Methylthioadenosine phosphorylase, putative	2.7
	Tb927.7.1790	Adenine phosphoribosyltransferase, putative	3.5
	Tb927.10.10810	Adenosine deaminase-like protein	1.3[a]
	Tb927.7.4570	Inosine-guanine nucleoside hydrolase (IG-NH)	2
	Tb927.11.3650	Adenylosuccinate synthetase, putative	1.7[a]
	Tb927.10.16120	Inosine-5'-monophosphate dehydrogenase	2
	Tb927.5.4560	Guanine deaminase, putative	1.8[a]
	Tb927.8.3800	Guanosine diphosphatase, putative	1.1[a]
	Tb927.5.2080	Guanosine monophosphate reductase, putative	1.3[a]
	Tb927.7.1930	Nucleoside diphosphatase, putative	1.7[a]
	Tb927.5.1360	Nucleoside 2-deoxyribosyltransferase	1.5
	Tb927.10.830	Adenylate kinase (ADKB), flagellar	-2.3
	Tb927.9.8450	Adenylate kinase (ADKG), cytosolic	-1
	Tb927.2.6150	Adenosine transporter P1 (TbNT2)	1.7[a]
	Tb927.2.6200	Nucleoside transporter, putative (TbNT3)	1.6[a]
	Tb927.2.6220	Nucleoside transporter, putative (TbNT4)	2
	Tb927.2.6320	Adenosine/inosine transporter, putative (TbNT6)	1.8[a]
	Tb927.2.6280	Adenosine/inosine transporter, putative (TbNT7)	1.4[a]
	Tb927.8.6060	2-amino-3-ketobutyrate coenzyme A ligase, putative	2.9
	Tb927.5.930	NADH-dependent fumarate reductase (FRDg)	2,7
	Tb927.11.5050	Fumarate hydratase, class I (FHm)	1.6
	Tb927.5.940	NADH-dependent fumarate reductase, (FRDm2)	2.7
	Tb927.11.8910	NADH dehydrogenase subunit NB6M (complex I)	1.9
	Tb927.10.9440	Alternative NADH dehydrogenase (NDH2)	1.5[a]
	Tb927.8.6580	Succinate dehydrogenase SDH1 (complex II)	2.4[a]
	Tb927.8.5640	Succinate dehydrogenase subunit SDH6 (complex II)	1.8
	Tb927.7.3590	Succinate dehydrogenase subunit (complex II)	2.0
	Tb927.1.1580	Cytochrome c oxidase assembly factor (complex IV)	3
	Tb927.7.7430	ATP synthase F1 alpha chain, mitochondrial precursor	1.6
	Tb927.3.1380	ATP synthase F1 beta chain, mitochondrial precursor	1,7
	Tb927.11.5280	ATP synthase F0 subunit 9, putative	2.4
	Tb927.7.1470	ATP synthase F0 subunit 9, putative	1.8
	Tb927.10.1570	ATP synthase F0 subunit 9, putative	2.3[a]
Amino Acid catabism, TCA and Electron Transport Chain	Tb927.1.3950	Alanine aminotransferase (ALAT)	4.6
	Tb927.10.3660	Aspartate aminotransferase	1.3[a]
	Tb927.11.1450	2-oxoglutarate dehydrogenase E1 component, putative	1.5[a]
	Tb927.11.11680	2-oxoglutarate dehydrogenase E2 component, putative	2.0[a]
	Tb927.3.2230	Succinyl-CoA synthetase alpha subunit, putative	1.6[a]
	Tb927.10.7410	Succinyl-CoA ligase [GDP-forming] beta-chain, putative	2.6
	Tb927.8.6060	2-amino-3-ketobutyrate coenzyme A ligase, putative	2.9
	Tb927.5.930	NADH-dependent fumarate reductase (FRDg)	2,7
	Tb927.11.5050	Fumarate hydratase, class I (FHm)	1.6
	Tb927.5.940	NADH-dependent fumarate reductase, (FRDm2)	2.7
	Tb927.11.8910	NADH dehydrogenase subunit NB6M (complex I)	1.9
	Tb927.10.9440	Alternative NADH dehydrogenase (NDH2)	1.5[a]
	Tb927.8.6580	Succinate dehydrogenase SDH1 (complex II)	2.4[a]
	Tb927.8.5640	Succinate dehydrogenase subunit SDH6 (complex II)	1.8
	Tb927.7.3590	Succinate dehydrogenase subunit (complex II)	2.0
	Tb927.10.3100	Glycerol-3-phosphate acyltransferase, putative	1.3
	Tb927.3.4820	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative	2.4
	Tb927.11.15150	1-acyl-sn-glycerol-3-phosphate acyltransferase, putative	1.2

	Gene ID	Annotation	FAT / Blood (Log2 FC)*
Lipid Metabolism	Tb927.3.1700	Diacylglycerol acyltransferase, putative	2.4
	Tb927.8.6390	Lysophospholipase, putative	1.8
	Tb927.7.220	CDP-diacylglycerol synthetase, putative	-1.3
	Tb927.1.4830	Phospholipase A1	-1.5
	Tb927.11.20730	Glycerophosphoryl diester phosphodiesterase, putative	-1.5
	Tb927.5.1140	Ethanolamine kinase	1.8
	Tb927.10.12810	Cholinephosphate cytidylyltransferase	1.0[a]
	Tb927.9.10080	Phosphatidylserine decarboxylase, putative	-1.3
	Tb927.7.5450	Lipin phosphatidic acid phosphatase, putative	-2.5
	Tb927.10.13930	Phosphatidic acid phosphatase, putative	-1.2[a]
	Tb927.8.7440	Triglyceride lipase, putative	1.2[a]
	Tb927.3.3870	Triglyceride lipase, putative	1.2[a]
	Tb927.8.7730	Dihydroceramide synthase, putative	-2.7
	Tb927.4.4740	Dihydroceramide synthase, putative	-2.2
	Tb927.11.4700	Prostaglandin f synthase	2.2
	Tb927.7.4170	Fatty acid elongase 2	-2.2
	Tb927.5.4530	Fatty acid elongase 4	-1.4
Sterol Metabolism	Tb927.11.6210	Sterol 14-alpha-demethylase	-1.9[a]
	Tb927.11.15530	C-14 sterol reductase, putative	1.5[a]
	Tb927.8.3240	Lathosterol oxidase, putative	-2
	Tb927.4.4080	C-5 sterol desaturase, putative	-1.4
	Tb927.8.2480	3-oxo-5-alpha-steroid 4-dehydrogenase-like, putative	-2.1
Fatty Acid β-Oxidation	Tb927.9.4210	Fatty acyl CoA synthetase 3 (ACS3)	1.5
	Tb927.9.4230	Fatty acyl CoA synthetase 4 (ACS4)	1.5
	Tb927.3.3900	Carnitine O-palmitoyltransferase II, putative	1.4[a]
	Tb927.11.2230	Carnitine O-acetyltransferase, putative	-2,6
	Tb927.8.1440	Enoyl-CoA hydratase, putative	1.7
	Tb927.3.4850	Enoyl-CoA hydratase, mitochondrial precursor, putative	-1.8
	Tb927.8.1820	3-ketoacyl-CoA thiolase, putative	1.7
	Tb927.8.2540	3-ketoacyl-CoA thiolase, putative	1.3[a]
RNA metabolism and chromatin	Tb927.8.2780	RNA-binding protein RBP10	-2.6[a]
	Tb927.9.4560	RNA-binding protein RBP32	-3.2
	Tb927.6.4440	RNA-binding protein RBP42	2.0
	Tb927.11.12100	RNA-binding protein RBP5	3.5
	Tb927.10.3990	RNA helicase DHH1	2,9
	Tb927.11.14950	Zinc finger protein 2 (ZFP2)	2.0
	Tb927.5.810	Zinc finger protein (ZC3H11)	2
	Tb927.1.2600	Pumilio RNA binding protein 9 (PUF9)	-3,3
	Tb927.2.2670	Histone H4 variant	-2.3
Protein Post-Transcriptional modifications	Tb927.10.13480	Protein kinase, putative	2.7
	Tb927.11.4470	Serine/threonine protein kinase, putative	-2.4
	Tb927.3.1240	Protein phosphatase 2A	3.4
	Tb927.11.4990	Kinetoplastid-specific phospho-protein	-1.9
	Tb927.11.12650	Palmitoyl acyltransferase 5, putative	1.8
	Tb927.7.3350	Palmitoyl acyltransferase 11, putative	1.1
	Tb927.9.7700	Palmitoyl acyltransferase 1, putative	-1.4
	Tb927.11.7230	Palmitoyl acyltransferase 8, putative	-1.1[a]
	Tb927.7.5490	Arginine N-methyltransferase, type III (PRMT7)	3.2
	Tb927.10.3560	Arginine N-methyltransferase, type I (PRMT3)	1.5
Cell Cycle Regulation	Tb927.7.5770	Nuclear Dbf2-related kinase (PK53)	-2.3
	Tb927.7.5570	Cell cycle associated protein MOB1-A	1.4
	Tb927.7.5580	Cell cycle associated protein MOB1-B	1.3
	Tb927.10.2440	Metacaspase MCA4, cysteine peptidase	3.8
	Tb927.11.3220	Metacaspase MCA1, cysteine peptidase	3.1
	Tb927.9.14220	Metacaspase MCA5, cysteine peptidase	-1.2
	Tb927.11.17000	AIR9	3.9
	Tb927.3.1440	Spastin	2.5
	Tb927.11.9530	14-3-3 (I) protein	2.4
Cell Signaling	Tb927.10.13000	cAMP phosphodiesterase A (PDEA)	-1.3
	Tb927.11.12860.1	cAMP Response Protein 2 (CARP2)	-2.7[a]
	Tb927.3.1040	cAMP Response Protein 4A (CARP4A)	-1.8
	Tb927.3.1060	cAMP Response Protein 4B (CARP4B)	-2.1
	Tb927.4.800	TOR3 (TOR-like 1)	2.2
	Tb927.8.2770	Inositol 1,4,5-trisphosphate receptor	1.8
	Tb927.11.6270	Inositol-1,4,5-trisphosphate (IP3) 5-phosphatase, put.	1.5[a]
	Tb927.11.15330	Phosphatidylinositol 3-kinase catalytic subunit, putative	-1.6

ATF parasites showed significant up-regulation of the alanine and aspartate aminotransferases and the glutamate shunt, which feed products into the tricarboxylic acid (TCA) cycle. Additionally, this cycle also appeared to be more active, given the up-regulation in three key steps, allowing it to process succinate, fumarate and 2-oxoglutarate, resembling the TCA cycle of the *T.*

brucei insect form (reviewed in (*158*)). These changes suggest that the F0/F1 ATP synthase complex is functional and that the associated electron transport chain is operating in ATF parasites in a manner similar to that in PCF parasites.

Significant changes in gene expression of lipid and sterol metabolic pathways were also observed in ATF parasites. But one of the most striking observations in the RNA-seq data was the potential presence of an active fatty acid β-oxidation, which produces energy from fatty acid catabolism. This was unexpected, as β-oxidation activity has never been detected in any *T. brucei* life cycle stage to date. ATF parasites showed up-regulation of the putative genes responsible for the 2nd and 4th steps of the β-oxidation cycle (enoyl-CoA hydratase and 3-ketoacyl-CoA thiolase, respectively) (Figure 6A). Moreover, fatty acid transport across the mitochondrial membrane (facilitated by acyl-CoA synthases and carnitine-acyltransferases) were up-regulated, while fatty acid elongases 2 and 4 were down-regulated, suggesting that in ATF parasites, endocytosed fatty acids are not being elongated and anabolized into more complex molecules. Instead they may enter the glycosomes and/or mitochondrion, where they are catabolized via a β-oxidation pathway to form acetyl-CoA (experimentally validated, see below and Figure 6), which feeds into the now active TCA cycle.

ATF parasites have active fatty acid b-oxidation

To investigate whether ATF parasites are capable of β-oxidation, labeled myristate was used in a pulse-chase experiment with living trypanosomes and potential labeled β-oxidation intermediates were identified by gas chromatography-mass spectrometry (GC-MS). Myristate (C14:0) was chosen as it is efficiently uptaken up and incorporated into lipids and glycosylphosphatidylinositol (GPI) anchors in the slender BSF parasites (reviewed in(*175*)). Isolated ATF parasites were labeled with deuterated-myristate (D_{27}-C14:0) for 1 hour and then chased with serum, following which labeled myristate metabolites were identified. As expected, BSF showed accumulation of D_{27}-C14:0 during the pulse and chase periods (Figure 6B-6C). ATF parasites also showed D_{27}-C14:0 accumulation during the pulse (Figure 6B

upper panel and Figure 6C), but the amount of D_{27}-C14:0 decreased significantly during the chase (Figure 6B lower panel and Figure 6C). Importantly, the decrease of D_{27}-C14:0 in ATF parasites coincided with the detection of β-oxidation metabolites derived from the labeled myristate, including myristoleic acid (D_{25}-C14:1), 3-hydroxy-myristate (D_{25}-3-OH-C14:0) and 3-oxo-myristic acid (D_{24}-3-keto-C14:0) (Figure 6B, right panels, Figure 6C and Figure S6 and S7). The latter two metabolites were also observed to some minor extent during the pulse (Figure 6B, upper panel and Figure 6C), while D_{25}-C14:1 was present in higher amount during the chase period. Minor amounts of unlabeled 3-hydroxy-myristrate (H_{26}-3-OH-C14:0) were also observed in ATF parasites, but not BSF (Figure 6C).

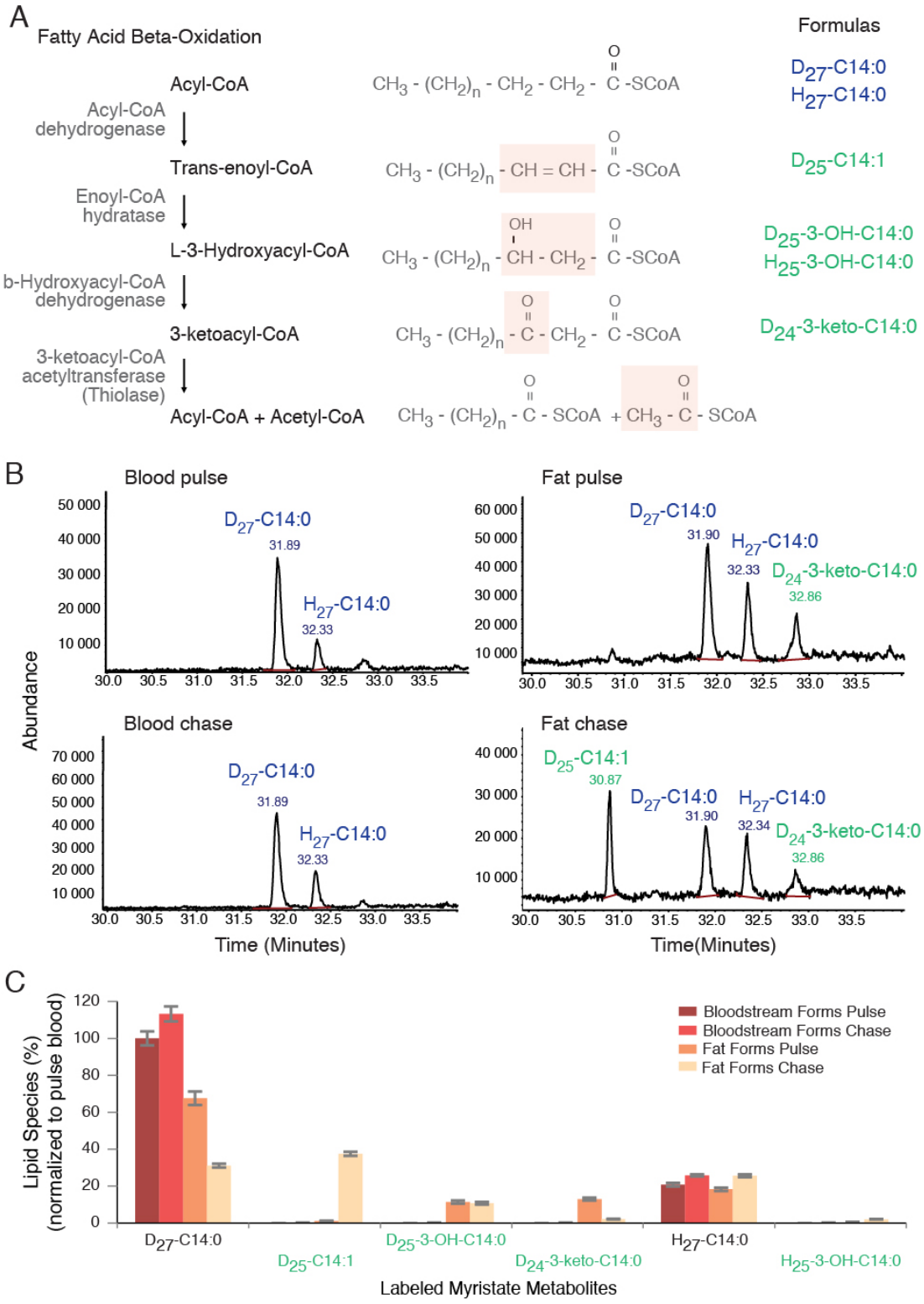


Figure 6: Fatty acid b-oxidation is active in ATF parasites. (A) Schematic of fatty acid b-oxidation pathway. Four enzymatic modifications are indicated by shaded box on the fatty acid structures where biotransformation takes place. Formulas in blue and green indicate the myristate and b-oxidation metabolites, thereof from the non- and labeled myristate, respectively, identified in this work. (B) FAME analysis by GC-MS of D_{27}-C14:0 labeled bloodstream (left) or ATF (right) parasites for 1 hour (upper) and chased for a further 1hour (lower). GC-MS trace shows 30-34 minutes (n = 3). (C)

Uptake of D_{27}-C14:0 and b-oxidation metabolites after normalization to the added internal standard C17:0. 100% equates to the amount of D_{27}-C14:0 taken up by bloodstream form in the 1 hour labeling (pulse) (n = 3). Represented are the means and the respective average absolute deviations of the means.

Collectively these data show that ATF parasites are able to actively uptake exogenous Myristic acid and form β-oxidation intermediates, demonstrating for the first time the existence of this pathway in trypanosomes, and suggesting that ATF parasites could in part use fatty acid β-oxidation to satisfy their energy requirements.

1.4. Discussion

A well-established feature of the unicellular, extracellular *T. brucei* parasite is their ability to invade the central nervous system (CNS). Here we show that, while blood is the major site of parasite accumulation on the first peak of parasitemia, fat contains the highest density and total number of parasites later in infection (around 100- to 800-fold more than the brain). Although the reason(s) why parasites accumulate in adipose tissue remain unknown, we clearly show it has dramatic consequences for the parasites. They functionally adapt to the tissue environment by rewiring gene expression, including the possibility of using lipids as a carbon source.

Possible advantages to parasite accumulation in the adipose tissue

The differential parasite distribution among blood and organs/tissues could be due to several non mutually-exclusive reasons: parasites may be more efficiently eliminated by some tissue-specific immune responses parasites may grow at different rates in each organ/tissue and/or parasites may not equally penetrate and exit organs/tissues. Depending on the dynamics of entry and exit of parasites to/from adipose tissue, it is possible that fat acts as a source of parasites that can repopulate the blood. This reversible movement between blood and fat could have important implications for (i) the transmission dynamics, since stumpy formation is triggered by a quorum-sensing mechanism

(*163*); and (ii) antigenic variation, if fat, for example, would favor the appearance of new VSG variants that could later go to the blood (*176*).

An intriguing question is whether stumpy forms could be directly ingested by a tsetse fly from the subcutaneous fat, which would be a novel root of transmission. In our histology analysis, we did not find a significant number parasites in this fat depot, but we have obviously not screened the entire skin of the mouse. Moreover, we performed this analysis in mice infected by intraperitoneal injection, which bypasses the skin as the first entry point of metacyclic forms. So, it remains to be determined whether subcutaneous fat is important for accumulation of metacyclics and/or ATFs and how it impacts transmission.

Not all *Trypanosoma* species occupy the same niche in the host. *T. brucei* and *T. evansi* are mainly tissue-invading parasites, while *T. congolense* stays in smaller capillaries and venules of tissues and *T. vivax* remains mainly in circulation (*177*). These differences have been associated to the different swimming properties of each *Trypanosoma* species (*167*). Previous reports had indirectly suggested that *T. brucei* parasites could be present in adipose tissue (*7, 178*). Our study demonstrates that *T. brucei* parasites accumulate in high numbers in the fat of rodent animals. Although mouse is an accepted model to study *T. brucei* infection (*7*), we cannot exclude the possibility that accumulation in adipose tissue is a result of the selection process that happens when *T. brucei* infects a non-natural host. In the future, it would be interesting to confirm whether fat preference is a common feature of this and other *Trypanosoma* species in their natural hosts and to compare their swimming properties in different tissues/organs.

It is intriguing to note that several pathogens infect adipose tissue. *T. cruzi*, the causative agent of Chagas disease in Latin America, invades adipocytes during acute infections in mice and humans (*179*). Also *Plasmodium berghei*, a causative agent of rodent malaria, sequesters in lungs and fat in a CD36-dependent fashion. *Mycobacterium tuberculosis* infects adipocytes where it accumulates in intracytoplasmic lipid inclusions and survives in a “dormant” non-replicating state that is insensitive to anti-mycobacterial drugs(*182*). Human immunodeficiency viruses (HIV) take advantage of the fat as a viral reservoir

during the chronic stage of infection and persistence on this reservoir is actually an obstacle for treatment(*183*). It is possible that persistence of *T. brucei* in the fat may account for some of the treatment failures in humans(*184*).

Functional adaptation to host adipose tissue

A major observation in this work is that 20% of the genes are differentially expressed between ATF and BSF, which is comparable with the differences between BSF and PCF (around 30%) and between slender and stumpy BSF forms (around 12%) (reviewed in(*170*). Parasites adapt to the fat environment by changing transcript levels of genes involved in metabolism, signaling, cell cycle control and RNA binding. Using biochemical assays, we confirmed that ATF can uptake free fatty acids (Myristic Acid) and catabolize them by fatty acid β-oxidation, which could lead to the production of ATP via the TCA cycle and oxidative phosphorylation. Therefore, it seems that parasites can sense and adapt to the fat, by rewiring their gene expression, including the ability of using lipids as a carbon source.

The major carbon source of BSF and PCF parasites is glucose and proline, respectively, both of which are readily available nutrients in the host and tsetse (reviewed in(*158*)). Fat is in its essence a lipid-rich environment. Therefore, it is possible that free fatty acids or some other form of lipid are readily available for the parasites to endocytose or actively transport via a receptor-mediated process (*185*). So far only one receptor has been identified in *T. brucei* as necessary for the import of LDL particles(*186*). Its transcript levels are not altered in ATF, suggesting either that this protein can be up-regulated (post-) translationally or that lipid/fatty acid import is mediated by yet uncharacterized transporters.

An abnormal consumption of the host lipids during a *T. brucei* infection could contribute to the weight loss observed in patients with sleeping sickness, cattle with Nagana and mice infected with *T. brucei* (*16, 187*). Interestingly, obese mice (*db*-/- knock-out mice) infected with *T. brucei* live 3-fold longer than their littermates, suggesting that having more adipose tissue partially protected mice from a *T. brucei* infection(*188*). Because obesity is associated with

persistent low-grade chronic inflammation in adipose tissue (*189*), it is possible that in obese mice parasites get more efficiently eliminated (or controlled), thus prolonging the survival of the host.

Most of what is known today about the mechanisms of virulence, persistence and transmission of *T. brucei* results from studies performed in bloodstream form parasites. The fact that adipose tissue is major reservoir of functionally differentiated parasites inevitably brings a new perspective to the current state of the art. First of all, what is the relative contribution of BSF and ATF for pathogenicity and host metabolic alterations? Could fat act as a source of parasites expressing novel VSGs? What are the implications of such a large reservoir of ATF in terms of transmission? What is the dynamics of parasite entry and exit from fat? Given the brain is a lipid-rich organ that is also invaded by *T. brucei*, it is obvious to ask whether parasites in the brain also adapt their gene expression and how these are related to brain-associated pathology? Do ATF induce changes in the host metabolism providing an advantage to the parasitic infection? Is the immune response of the adipose tissue more permissive to *T. brucei* parasites? Are anti-trypanosome drugs equally efficient at eliminating ATF and BSF? In sum, our findings have important consequences for the understanding of parasite biology, disease and drug treatment efficacy.

1.5. Experimental Procedures

Ethical Statements

Animal experiments were performed according to EU regulations and approved by the Animal Ethics Committee of Instituto de Medicina Molecular (iMM) (AEC_2011_006_LF_TBrucei_IMM). The animal facility of iMM complies with the Portuguese law for the use of laboratory animals (Decreto-Lei 113/2013); and follows the European Directive 2010/63/EU and the FELASA (Federation of European Laboratory Animal Science Associations) guidelines and recommendations concerning laboratory animal welfare. The tsetse fly mediated *T. brucei* infection work was performed in compliance with the

regulations for biosafety and animal ethics (VPU2014_1) and under approval from the Environmental administration of the Flemish government.

Animal Experiments

Unless otherwise indicated, all infections were performed in wild-type male C57BL/6J mice, 6-10 week old (Charles River, France), by intraperitoneal injection of 2000 *T. brucei* AnTat 1.1^{E} 90-13 parasites. For parasite counts, blood samples were taken daily from the tail vein. Organs/tissues of infected mice were collected at days 6 and 28 post-infection unless otherwise stated. Animals were sacrificed by CO_2 narcosis, blood collected by heart puncture and mice immediately perfused. Collected organs were snap frozen in liquid nitrogen or fixed in 10% neutral-buffered formalin. In transplants, homogenates as well as 600 µL of blood were transplanted into age and sex-matched naïve mice.

Cell Lines

The majority of the infections described in this manuscript were performed using *T. brucei* AnTat 1.1E, a pleomorphic clone derived from an EATRO1125 clone. AnTat 1.1E 90-13 is a transgenic cell-line encoding the tetracyclin repressor and T7 RNA polymerase(*190*). Tsetse infections were performed with *T. brucei* AnTAR1 strain. We also used *T. brucei* Lister 427, a monomorphic strain derived from antigenic type MiTat 1.2, clone 221a(*191*). The stumpy reporter cell line GFP::PAD1utr derives from AnTat 1.1E 90-13 in which the green fluorescent protein, GFP, is coupled to PAD1 3'UTR. A nuclear localization signal (NLS) targets GFP protein into the nucleus. Prior to infection, *T. brucei* cryostabilates were thawed and parasite mobility was checked under an optic microscope.

Mice Infections

Mice were group-housed in filter-top cages and maintained in a Specific-Pathogen-Free barrier facility. The facility has standard laboratory conditions: 21 to 22°C ambient temperature and a 12 h light/12 h dark cycle. Chow and water were available *ad libitum*. Unless otherwise stated, all infections were performed in wild-type male C57BL/6J mice, 6-10 week old (Charles River,

France), by intraperitoneally (i.p.) infection of 2000 *T. brucei* AnTat 1.1E 90-13 parasites.

To test if the presence of parasites in fat was dependent on the model, we performed the following variations in the infection protocol: a) male C57BL/6J mice were infected with a more virulent strain, the *T. b. brucei* strain Lister 427; b) male C57BL/6J mice were infected intravenously (i.v.) in the tail vein; c) a different strain of mice, BALB/c, and 7 week old Wistar rats (Charles River, France) were also infected (the latter with 4000 parasites); d) female C57BL/6J mice and finally e) male C57BL/6J mice naturally infected by tsetse bite. For this latter protocol, freshly emerged *Glossina morsitans morsitans* flies were fed their first blood meal on a *T. brucei* AnTAR1 infected mice at the peak of parasitemia. Subsequently, flies were maintained on commercially available defibrinated horse blood through *in vitro* membrane feeding. Thirty days after the infective blood meal, individual flies were evaluated for the presence of metacyclic trypanosomes in their salivary glands by salivation on pre-warmed (37°C) glass slides. To initiate a natural infection, one individual tsetse fly with a mature salivary gland infection was allowed to probe and feed per mouse.

Clinical Parameters and Organ Collection

All measurements in mice were made between 17:00 and 18:00. For parasite counts, blood samples were taken daily from the tail vein, and parasitemia was determined by manual counting using a Neubauer chamber. Organs/tissues of infected C57BL/6J mice (male and female) were collected at days 6, 13, 20 and 28 post-infection; for the infection with *T. b. brucei* strain Lister 427, organs were collected once parasitemia reached $1x10^8$ parasites/mL; for BALB/c mice, at day 6 post-infection; and for Wistar rats and C57BL/6J mice i.v. infections, at days 6, 13 and 20 post-infection. Animals were sacrificed by CO_2 narcosis, blood was collected by cardiac puncture and perfusion was performed to eliminate blood and parasites from circulation. Briefly, mice were perfused transcardially with pre-warmed heparinized saline (50 mL 1X phosphate buffered saline (1X PBS) with 250 ml heparine per animal) using a peristaltic pump, ranging its speed from 2 mL/min to 8 mL/min. Organs were then collected and either used immediately for parasites isolation, snap frozen in

liquid nitrogen for molecular analysis, or fixed in 10% neutral-buffered formalin for histopathology.

Transplantations

Donor mice were sacrificed with CO_2 narcosis (at days 21 and 28 post-infection) and the gonadal fat depot, brain, heart, and 600 µL of blood were harvested. Organs were manually homogenized through a 70 µm mesh into 1X PBS. Cell suspension was centrifuged at 1000 g for 10 minutes and resuspended in 800 µl of HMI11. Tissue lysates were injected intraperitoneally in naïve mice. Blood was diluted in 1X PBS and centrifuged for 5 minutes at 2800 g. Cell pellet was diluted in 800 µl of 1X PBS and injected intraperitoneally in naïve mice.

Histology, electron microscopy and fluorescence analysis

Formalin-fixed organs were embedded in paraffin and 3µm sections were stained with hematoxylin and eosin (H&E). For immunohistochemistry, 3µm sections were stained for VSG using a non-purified rabbit serum anti-*T. brucei* VSG13 antigen (cross-reactive with most VSGs via the C-terminal domain) or non-purified rabbit serum anti-*T. brucei* H2A (generated against a recombinant protein) (kind gift of Christian Janzen), diluted 1:5000 and 1:3000, respectively. Briefly, antigen heat-retrieval was performed in a microwave oven (800 w) for 15 minutes with pH 9 Sodium Citrate buffer (Leica Biosystems, MO, USA). Incubation with ENVISION kit (Peroxidase/DAB detection system, Dako Corp, Santa Barbara, CA) was followed by Mayer's hemalumen counterstaining. No staining was observed in the negative control (without primary antibody). Tissue sections were examined by a pathologist (TC), blinded to experimental groups, in a Leica DM2500 microscope coupled to a Leica MC170 HD microscope camera.

For transmission electron microscopy, gonadal fat depot from infected mice (days 6 and 28 post-infection) was collected and fixed for three hours at 4ºC in 0.1 M cacodylate buffer, pH 7.4, containing 2.5% (v/v) glutaraldehyde. After staining for 1 hour with 1% (w/v) osmium tetroxide and 30 minutes with 1% (w/v) uranyl acetate, samples were dehydrated in an ethanol gradient (70-95-100%), transferred to propylene oxide and embedded in EPON™ resin. Semi-thin sections (300-400 nm) were stained with toluidine blue for light microscopy

evaluation. Ultra-thin sections (70 nm) were collected in copper slot grids and stained with 2% uranyl acetate and lead citrate (Reynolds recipe). Grids were screened in a Hitachi H-7650 transmission electron microscope at 100 kV acceleration.

For 3D reconstruction of isolated trypanosomes, parasites isolated from gonadal fat depot were centrifuged at 5000rpm and processed as described above for whole tissue. After embedding, approximately 26 serial ultra-thin sections (70 nm) were collected for each individual parasite. Grids of seven parasites were screened in a Hitachi H-7650 transmission electron microscope at 100 kV acceleration, serial section alignment was achieved using the IMOD software package version 4.7.3 for alignment and modeling (*192*). Videos were projected using ImageJ 4.47v.

For fluorescence analysis, the gonadal depot was stained with LipidTox, fixed in Paraformaldehyde and embedded in Fluoromount-G. Fluorescence images were acquired using a 40x objective in a Zeiss Cell Observer wide-field microscope. For morphometric analysis, isolated parasites were fixed in 2% paraformaldehyde, DAPI stained and embedded in vectashield. Images were acquired using a 63x oil objective with optional optovar magnification (1.6x) in the same wide-field microscope. Parasite morphometric analysis was done essentially as described in(*165*).

The mitochondrion of 1-5x10^6 isolated parasites was labeled using MitoTracker Green (Invitrogen/Molecular Probes, M-7514) according to the manufacturer's instructions. Fluorescence and DIC images were acquired using a confocal Laser Point-Scanning Microscope (Zeiss LSM 710).

Parasite Quantification

T. brucei 18S rDNA genes were amplified from genomic DNA of a known mass of tissue and converted into parasite number using a standard curve. For RNA quantification, the DDCt method was used by amplifying *TbZFP3* and mouse *Gapdh* genes from tissue total RNA.

Parasite Isolation from Tissues

Bloodstream parasites were purified over a DEAE column(*193*), while ATF parasites were isolated from gonadal fat depot by incubating the depot in 10mL

of MEM or HMI11 at 37ºC and 150 rpm agitation for up to 40 minutes. After which parasites in media are centrifuged for 10 min at 1800 rpm at RT, and numbers assess with a neubauer chamber. Expected yield is 5 x 10^6 parasites.

Flow Cytometry

Cell cycle analysis was performed using propidium-iodide (PI) or Vybrant DyeCycle violet (DCV) in fixed or live cells, respectively. PI staining was done according to (*194*). For DCV staining, cell suspensions were washed and 0.5 mL of DCV was added per each million of isolated parasites and incubated for 10 minutes at 37ºC. PI, DCV and GFP intensities were measured with BD LSRFortessa™ cell analyzer.

RNA-sequencing

RNA and cDNA library of both blood and gonadal fat depot from days 4 and 6 of infection, respectively, were prepared as described (*171*) and samples sequenced in an Illumina HiSeq2000 platform. Reads were processed and mapped to the *T. brucei* TREU927 genome. Differential gene expression was analyzed and genes were considered differentially expressed if they were detected by at least two of the three considered algorithms (p adjusted < 0.01).

Myristate Metabolic Labeling

To evaluate myristate incorporation and metabolism, the fat isolation protocol was performed in glucose free Minimum Essential Medium (MEM). Parasites were placed in an vented tube with 1 mL of MEM and starved for 30 minutes at 37ºC. Starved parasites were then labeled, with 0.4 mg of radiolabeled D_{27}-C14:0 pre-coupled with defatted BSA, for one hour. 450 µL of the cell suspension were washed, snap frozen in liquid nitrogen and lyophilized in glass vials (Pulse sample). The remaining parasites were re-suspended in 500 µL of MEM and 100 µL of HMI11 for 1 hour at 37ºC and at the end processed as for pulse sample (Chase sample). Metabolite extraction, identification and quantification were conducted as described in (*195*) with the exception that fatty acids were released by acid hydrolysis (200 µL of 6M HCl at 110ºC for 16 hours).

Statistical Analysis

Statistical analyses were performed by fitting linear mixed effect models with mice as random effects unless otherwise indicated. At least three independent experiments were considered in each case and statistical significance was set to a = 0.05 level. Data were analyzed after logarithm transformation.

1.6. Supplementary figures

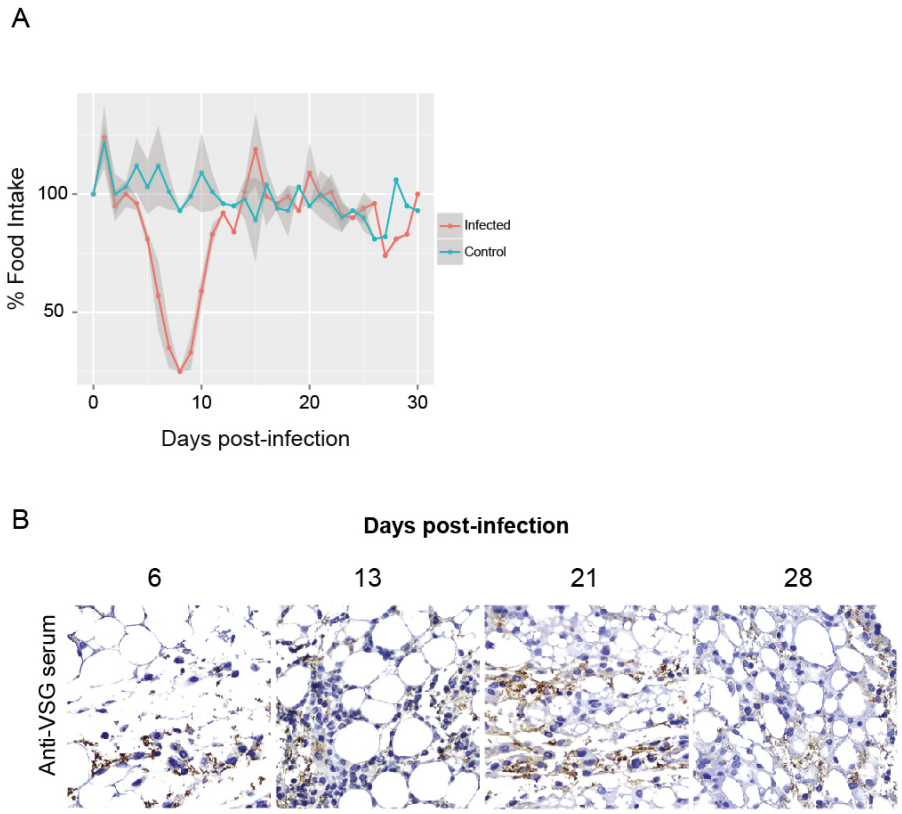


Figure S1. Related to Figure 1 and Figure 2. Clinical and histology details during *T. brucei* infection. C57BL/6J mice were injected i.p. with 2000 AnTat1.1^{E} parasites. (A) Variation of food intake during infection. Animals (n = 15 per group) were group housed and food intake was measured daily by weighting the food and dividing by the number of mice per cage (n = 4 per condition). Light grey shaded area represents SEM. (B) Representative light microscopic images of gonadal fat depots at different days post-infection, assessed by immunohistochemistry with anti-VSG antibody (parasites appear in brown). Original magnification, 400x.

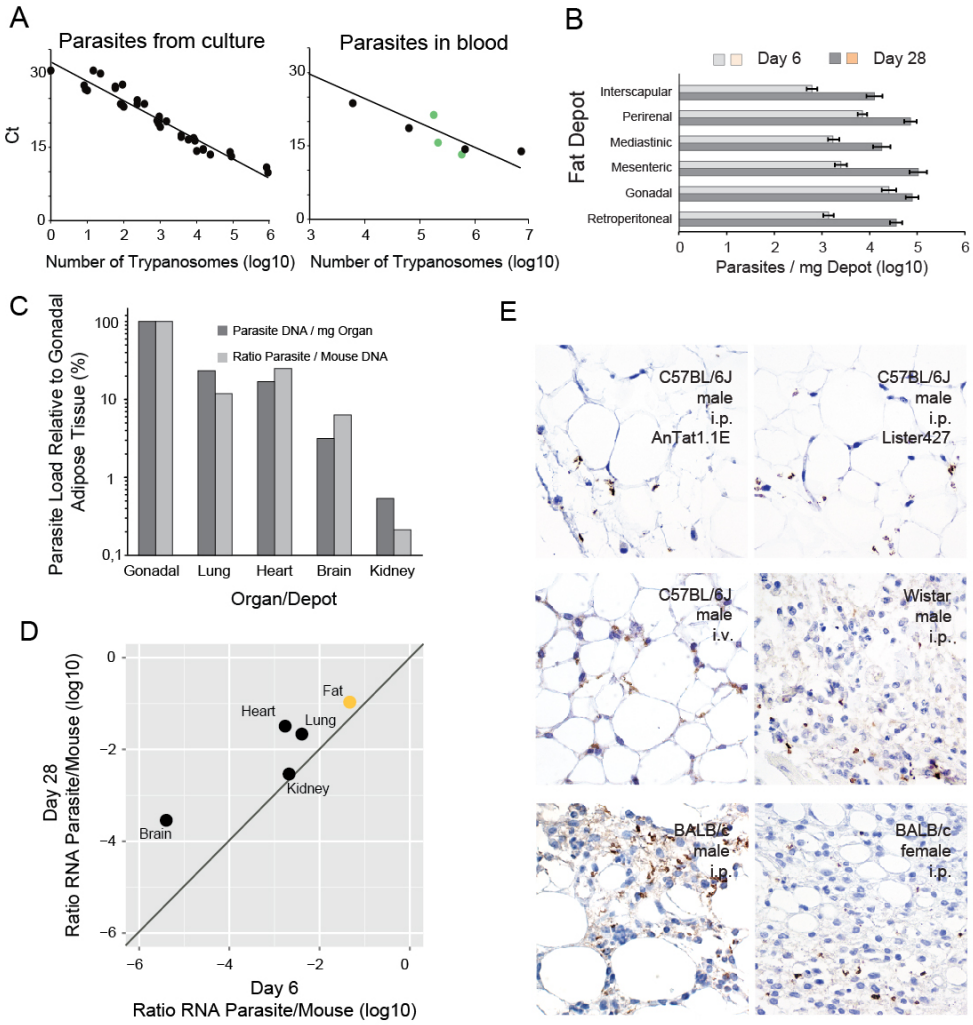


Figure S2. Related to Figure 1 and 2. Further validation of preferential accumulation of parasites in adipose tissue. (A) Calibration curve obtained from four independent in vitro cultures of known cell density. gDNA was extracted, serially diluted and amplified by quantitative PCR using *T. brucei* 18S rDNA primers. The goodness of fit of the linear regression is R2 = 0.933. Calibration curve obtained from blood from infected mice (n = 3) (green dots) and from a culture of parasites diluted in blood from naive mice (n = 1) (black dots). The goodness of fit of the linear regression is R2 = 0.925. (B) Parasite density in six fat depots on day 6 (n = 6-12) and 28 (n = 3-6) post-infection determined by qPCR of gDNA. For each depot, significant differences were found between days 6 and 28 post infection (Student's unpaired t test, $P < 0.05$). The average of all the studied depots shows a clear difference between parasite densities at days 6 and 28 post-infection (Student's unpaired t test, $P < 0.0001$). The geometric means and correspoding standard errors are indicated. (C) Two different gDNA qPCR methods were compared on day 28 post-infection: i. number of parasites per mg of organ (n = 3 - 9) and ii. ratio between *T. brucei* and mouse 18S gDNA (n = 3 - 9). Both methods show a similar parasite density in different tissues (LME, $P = 0.72$). (D) Parasite density on day 6 and 28 post-infection determined by qPCR of RNA. Transcripts of the parasite TbZFP3 gene were normalized to the mouse Gapdh. Each point represents the geometric mean of the parasite density on day 6 (n = 3 - 4) and on

day 28 post-infection (n = 3 - 5). RNA quantification validates the conclusions taken from gDNA qPCR: the adipose tissue is the major parasite reservoir (LME, P = 0.0006). The relative contribution of other organs is similar to what was measured by gDNA, except for brain, in which parasite density was lower than expected. Perhaps TbZFP3 is downregulated in this organ. (E) Representative light microscopic images of *T. brucei* in gonadal adipose tissue in different models of infection, assessed by immunohistochemistry with anti-VSG antibody (parasites appear in brown). Original magnification, 400x.

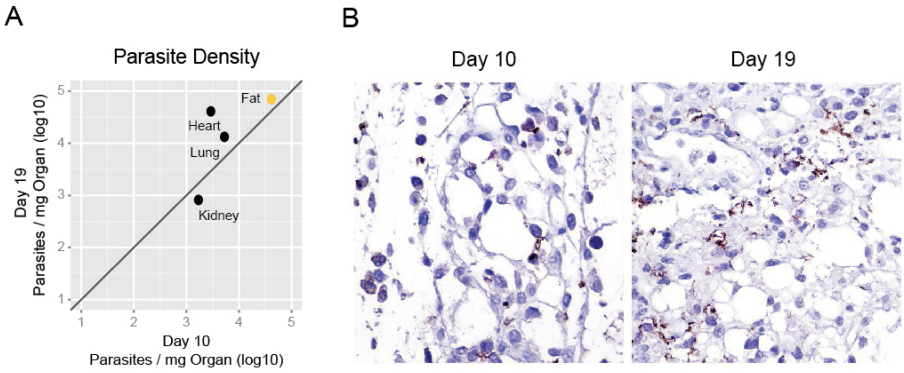


Figure S3. Related to Figure 2. Parasites accumulate in fat when infection is initiated by tsetse bite. (A) Parasite density in multiple organs in mice naturally infected by the bite of a tsetse fly. Mice were sacrificed at the first and second peak of parasitemia (10 and 19 days post-infection respectively) and parasite density determined by gDNA and qPCR as described in Figure 2. Each point represents the geometric mean of the parasite density at day 10 (n = 4) and at day 19 (n = 8). (B) Anti-VSG immunohistochemistry images of gonadal fat depot, 10 and 19 days post-infection by tsetse bite. Original magnification, 40x.

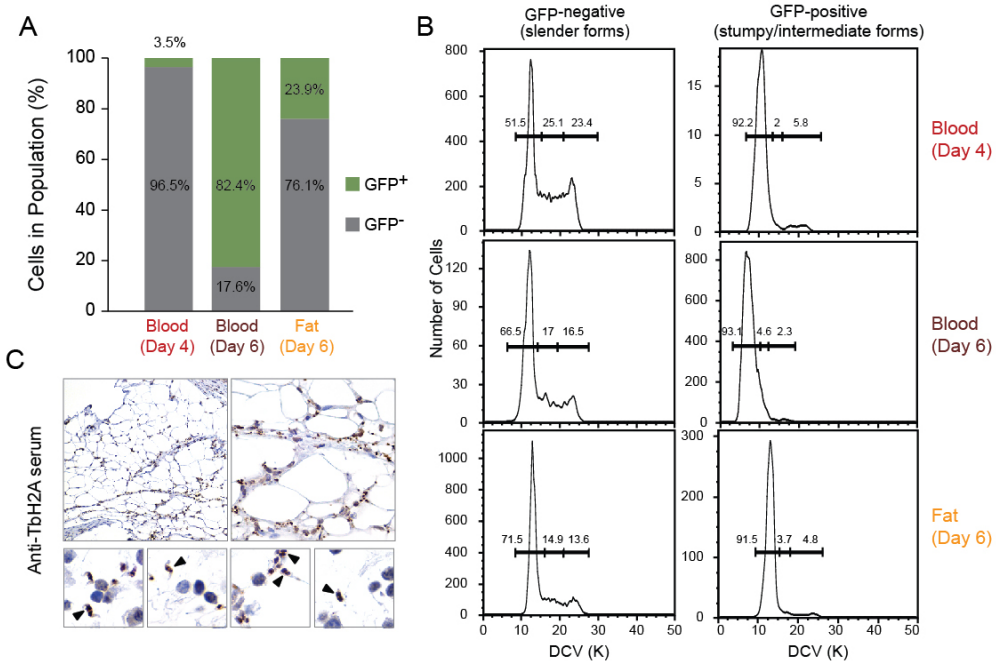


Figure S4. Related to Figure 3. Complementary methods to confirm presence of both replicative and cell cycle arrested parasites in fat on day 6 post-infection.

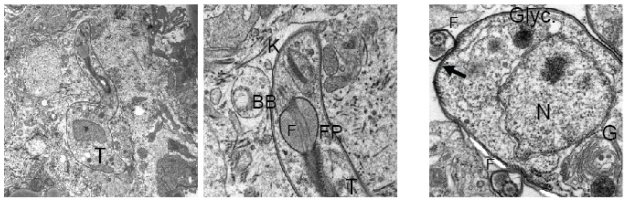
(A) GFP Expression and (B) Cell cycle analysis in parasites isolated from blood and adipose tissue, 4 or 6 days post-infection, assayed by FACS. A GFP::PAD1$_{UTR}$ T. brucei reporter cell-line was stained with DyeCycle Violet. The histograms represent the distribution and the percentage of parasites in each cell cycle stage. DyCycle Violet staining validates the conclusions taken from propidium iodide staining analysis. (C) Representative light microscopy images of gonadal fat depots on day 6 post-infection, assessed by immunohistochemistry with anti-*T. brucei* H2A rabbit serum (parasites appear in brown). Original magnification, 20x, 40x and 100x. Arrowheads indicate parasites undergoing nuclear division.

A

GFP expression	Tissue	Length (µm)	Width (µm)
GFP-negative (slenders)	Fat	24.57 ± 2.99	2.12± 0.26
	Blood	24.39 ± 2.50	2.15 ± 0.26
GFP-positive (stumpy/interm.)	Fat	21.32 ± 2.73	2.29 ± 0.31
	Blood	18.43 ± 1.81	3.11 ± 0.38

n= 100, Numbers are mean values with ± standard deviations .

B



C

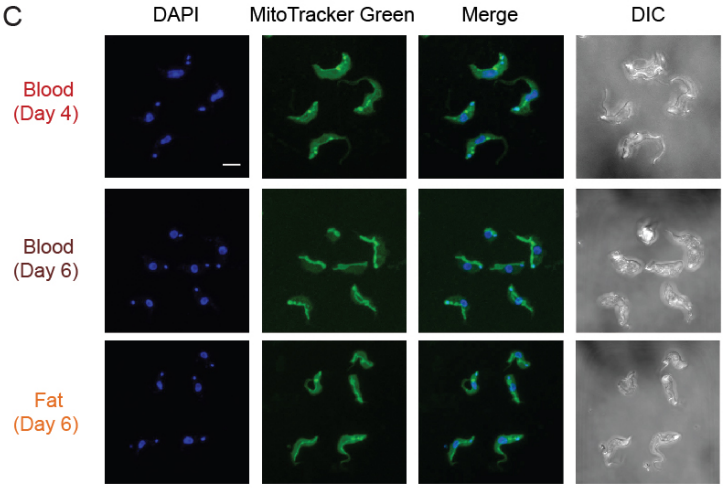


Figure S5. Related to Figure 4. Subcellular organization of parasites from adipose tissue. (A) Length and width mean values of GFP-negative (slender) and GFP-positive stumpy/intermediate) parasites isolated from blood or fat. (B) Transmission electron micrograph images of parasites in gonadal adipose tissue (day 28 post-infection). T, trypanosome; K, kinetoplast; BB, basal body; F, flagellum; FP, flagellar pocket; N, nucleus; Glyc, Glycosomes; Arrow: subpellicular microtubules. Scale bars represent 2 and 0,5 µm in the left and right panels, respectively. (C) MitoTracker Green, which stains in live cells the mitochondrion membrane, regardless of its membrane potential, was used to assess mitochondrion morphology. DNA was

stained by DAPI and images were captured under a confocal microscope. Bloodstream form parasites from day 4 of infection showed a punctate pattern typical of mitochondrion in slender parasites, while on day 6, were wider and shorter, mitochondrion displayed a tubular structure with a few branches in what appeared to be stumpy forms. The mitochondrion of adipose tissue forms present a tubular structure, but with fewer branches and thinner. All panels are shown with the same magnification. Scale bar represents 5 µm.

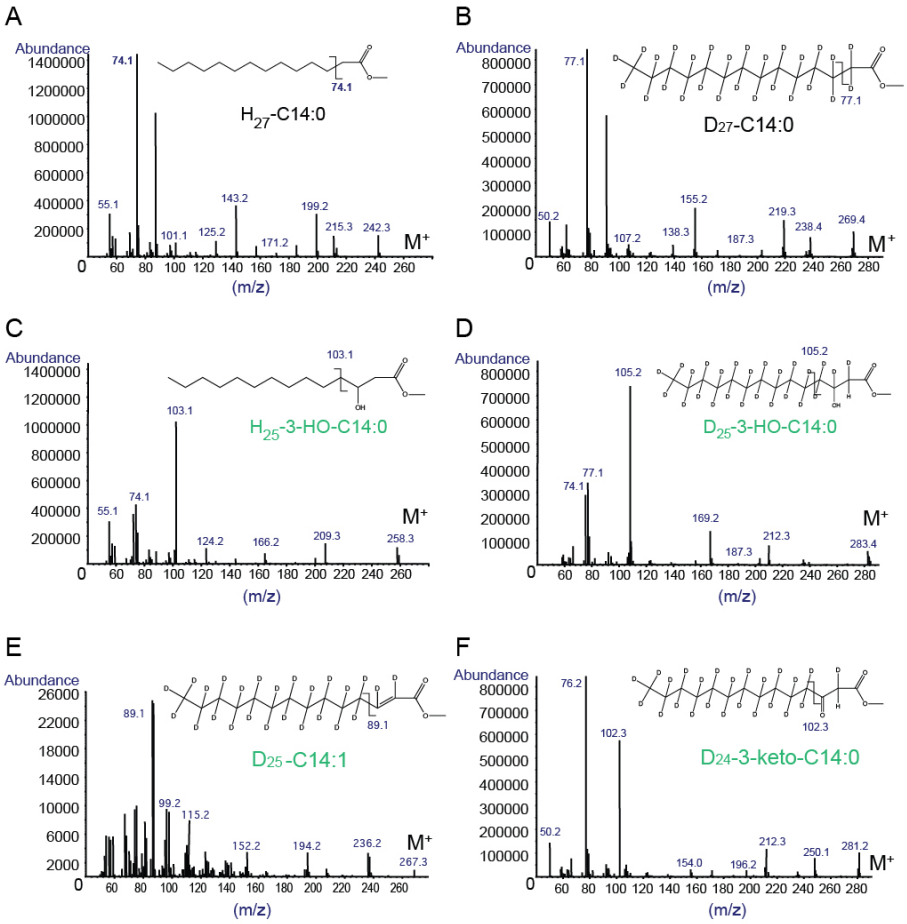


Figure S6. Related to Figure 6. Lipid metabolites identified by GC-MS.
All panels show the structure and fragmentation pattern of methyl ester derivatives of:
(A) myristic acid (C14:0)
(B) fully deuterated-myristic acid (D27-C14:0)
(C) 3-hydroxy- myristic acid (3-HO-C14:0)
(D) 3-hydroxy-deuterated- myristic acid (D24-3-HO-C14:0)
(E) deuterated-myristoleic acid (D25-C14:1)
(F) 3-keto-deuterated- myristic acid (D24-3-keto-C14:0)

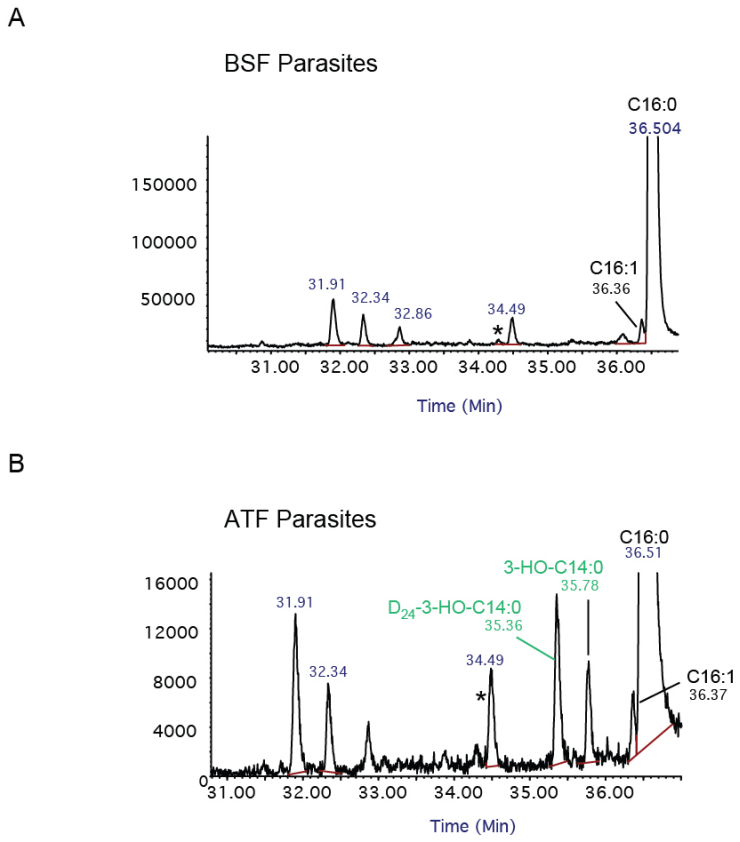


Figure S7. Related to Figure 6. Only ATF parasites produce hydroxyl-fatty acids as part of beta-oxidation catabolism of fatty acids. FAME analysis by GC-MS of D27-Mys labeled and subsequently chased bloodstream (A) and adipose tissue (B) forms. Trace 31-37 minutes showing positions of (3-HO-C14:0) and (D24-3-HO-C14:0) in adipose tissue forms only.

1.7. Accession Numbers

The ArrayExpress accession number for Lister427 culture parasites is E-MTAB-1715. Sequence data generated as part of this study have been submitted to the ArrayExpress database (EMBL-EBI) under accession number E-MTAB-4061.

1.8. Acknowledgments

The authors thank Keith Matthews (University of Edinburgh) for providing AnTat1.1^{E} clone, Christian Janzen (University of Wurzburg) for the *GFP::PAD1*$_{utr}$ cell-line, Dave Barry (University of Glasgow) for critical reading of

the manuscript, Catarina Gadelha (University of Nottingham) for guidance in electron microscopy studies, Rita Ventura (ITQB, Portugal) for support with lyophilization of samples, João Rodrigues (iMM) and Eugénia Carvalho (University of Coimbra) for useful discussions of the data, Leonor Pinho and Margarida Vaz for help with mouse work and and Fernando Augusto for drawing the graphical abstract. The authors would like to acknowledge the Rodent Facility, the Histology and Comparative Pathology and Bioimaging Laboratories of the Instituto de Medicina Molecular and all members of Luisa Figueiredo's lab. This work was supported by 55007419 (HHMI) and 2151 (EMBO) to L.M.F, D.P.N., F.B., F.G.; FCT fellowships to S.T., F.R.F., F.A.-B. (SFRH/BPD/89833/2012, SFRH/BD/51286/2010, SFRH/BD/80718/2011, respectively); Wellcome Trust grant (093228), MRC MR/M020118/1 and European Community Seventh Framework Programme under grant agreement No. 602773 (Project KINDRED) to S.A.Y. and T.K.S.; PAI 7/41 (Belspo) and ERC-NANOSYM to J.V.D.A.

Chapter II – section 2

Trypanosoma brucei infection accelerates the mouse circadian rhythm

The "Chapter II – Results – 2. *Trypanosoma brucei* infection accelerates the mouse circadian rhythm" consists of a second manuscript that will be soon submitted for publication. Here is clarified the contribution of each of the authors:

Manuscript:

Filipa Rijo-Ferreira[1,2,3], Margarida Sanches-Vaz[2], Luisa M. Figueiredo[2*#], Joseph S Takahashi[1,4*#] *Manuscript in preparation.*

[1]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, United States
[2]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[3]Graduate Program in Areas of Basic and Applied Biology, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal
[4]Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA
*To whom correspondence should be addressed. E-mail: Joseph.Takahashi@UTSouthwestern.edu or lmf@medicina.ulisboa.pt
[#]These authors contributed equally to this work

Filipa Rijo-Ferreira, the candidate for the PhD degree at the University of Porto, contributed to original idea of this project, and was responsible for the design and performance of all experiments as well as for the analysis and presentation of the results, following intensive discussions with her mentors Dr Figueiredo and Professor Takahashi. In addition, she wrote the manuscript as an integral part of her future paper (in preparation for submission) and of the thesis. Dr Figueiredo and Professor Takahashi designed and supervised the project and revised the manuscript. She was helped by Margarida Sanches-Vaz in one of the experiments with *Plasmodium chabaudi*.

2. *Trypanosoma brucei* infection accelerates the mouse circadian rhythm

2.1. Abstract

Sleeping sickness is a fatal disease caused by *Trypanosoma brucei* - a unicellular parasite that lives in the bloodstream and interstitial spaces of several organs, including the brain. Sleeping sickness patients show alterations in the sleep/wake cycle, which have led to the hypothesis that this may be a circadian rhythm disorder. To test this, we first recorded the circadian running-wheel activity of *T. brucei*-infected mice. We observed that infected mice run 7-fold more during the rest phase, with some animals running ~50% of their daily activity during the rest phase. These observations indicate that as infection progresses the circadian rhythm of the host is disrupted. In constant dark, we detected a gradually shorter period of running-wheel activity a few days after infection. This is surprising because parasites take >15 days to accumulate in the brain, suggesting that the early effects in mice behavior may be a consequence of a peripheral signal that feedback on the master clock. When we assessed PERIOD2::LUC expression *ex vivo,* we observed that, although all organs have robust circadian rhythms, those with higher parasite load have ~2 h shorter period than non-infected organs, indicating that the infection has a major but specific impact on peripheral circadian rhythm. Finally, when we co-cultured *T. brucei* parasites with PER2::LUC fibroblasts, we also observed a shortening of the circadian period, suggesting that *T. brucei* parasites could directly modulate the circadian clock of their host. Overall, we conclude that *T. brucei* causes a circadian rhythm disorder, which by changing the host molecular clock may be responsible for the changes in sleep/wake cycle observed in mice and humans and that give the name to the disease, as well as other physiological characteristics controlled by the clock as metabolism.

2.2. Introduction

Sleeping sickness is an infectious disease caused by *Trypanosoma brucei* - a unicellular and extracellular parasite. *T. brucei* can be initially found in the bloodstream and interstitial spaces of several organs and eventually crosses the blood brain barrier and reaches the brain, leading to coma and death if untreated. Sleeping sickness patients experience a myriad of different symptoms: from hypothermia, disturbed appetite, weight loss, motor weakness, tremor to even aggressiveness (*23*). But, the hallmark of sleeping sickness is the disruption of the sleep pattern, giving name to the disease. Patients experience somnolence during the day and insomnia at night, but with similar total time spent sleeping as healthy individuals (*28, 41*). This curious type of sleep/wake cycle disruption, together with changes in body temperature regulation and wrong time of endocrine secretion (*27, 41, 78, 196*), have led to the hypothesis that sleeping sickness may be a circadian rhythm disorder (*78*).

By living in a 24-hour world, organisms are subjected to daily environmental changes. Many organisms, from bacteria to humans have evolved molecular mechanisms to anticipate such changes. In humans, the master internal clock is located in the hypothalamic suprachiasmatic nuclei (SCN) of the brain, which with both neuronal and humoral signals can regulate output pathways that control diverse physiological functions: including sleep/wake cycle, core body temperature and metabolism (*131*). Within each cell of our body we can find an accurate transcription–translation feedback loop comprised of a core set of genes: the activators *Clock* and *Bmal1* and their repressors *Per* (Period 1 and 2) and *Cry* (Cryptochromes 1 and 2) (*126*). This loop takes 24 h to be completed, leading to the rhythmic expression of multiple downstream genes that will then impact a variety of cellular pathways.

It has been known since 1895, when David Bruce first discovered *T. brucei* as the causative agent of sleeping sickness (*197*) that this infection was in the source of the sleep disruption. However, even though the parasite's life cycle and evasion from host immune system has been studied in detailed (*161, 198, 199*), very little is known about how it affects the sleep/wake cycle. We do know that, after treatment, the sleep/wake cycle disruption of patients reverts to normal (*28*). Consistently, autopsies of patients who died of sleeping sickness

revealed very little neuro-degeneration (*46, 78*). Together these observations suggest that the presence of the parasite, rather than immune response or neuronal death, is the direct cause of this disease.

In this work, we used a mouse model to test if T. brucei causes a circadian rhythm disorder. We show that (i) *T. brucei* mouse infection reproduces the circadian behavior and body temperature changes observed in humans; (ii) *T. brucei* infection accelerates the mouse circadian rhythm behavior and PER2 expression and (iii) this effect may be partly caused by a direct interaction with the parasite.

2.3. Results

To test whether *T. brucei* infection disrupts a mouse circadian behavior, we measured the activity of infected mice using running wheels. Mice were individually caged and running wheel revolutions scored with an automated system. To confirm that the circadian activity of mice prior to infection was normal, mice were first entrained in light/dark (LD) for 7 days and then constant darkness (DD) for 10 days. As expected, we observed that every mice, prior to infection, had a normal circadian rhythm, perceived by the robust running activity in DD. On day 0, mice were injected i.p. either with 2000 *T. brucei* parasites of a pleomorphic strain or with vehichle only. We observed that globally infected mice were 4-fold less active than control mice (Fig. 1A-B), in particular in the periods following high parasitemia, when anemia is ~50-70% (day 6-10 days post-infection, fig. S1A-B). This sickness-like activity reduction is likely due to high parasitemia that leads to exacerbated inflammation, as it has been shown that injection of pro-inflammatory cytokines reduces global running activity of a mouse (*200*).

In humans, *T. brucei* causes relatively low and often undetectable parasitemia. In order to make the infection in mice more similar to humans, mice are typically treated around day 20 post-infection with suramin, a drug that cannot cross the blood brain barrier and thus only kills parasites in circulation (*201*). As a result the sickness-like phenotype disappears, mouse survival is prolonged with undetectable parasitemia.

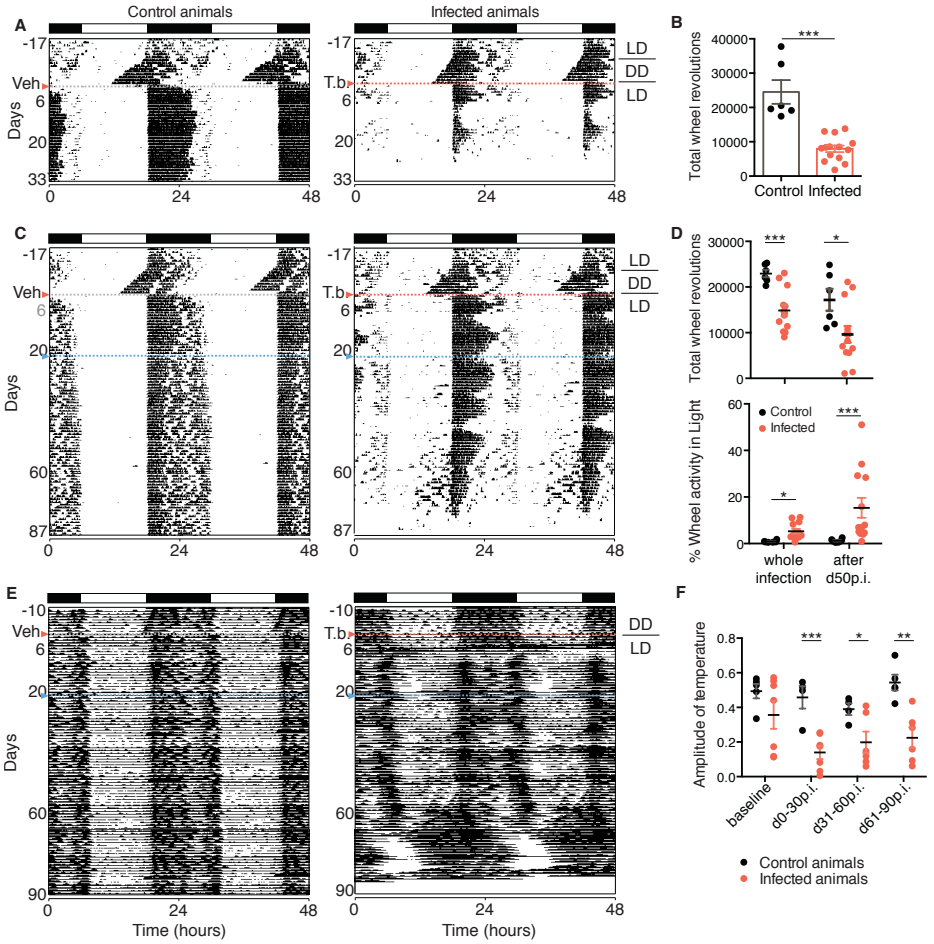


Figure 1. Circadian activity and body temperature disruption in *Trypanosoma*-infected mice. (A) Representative actograms of daily wheel-running activity of control and infected mice in light-dark cycles. All running-wheel experiments involve 7 days acclimatization period to wheels followed by 10 days dark period to confirm all animals have a normal circadian rhythm. Animals are either infected or injected with vehicle. Activity records were double plotted, with each day being represented beneath and also to the right of the preceding day. Horizontal black and white bars at the top of each actogram represent lights off and on, respectively. (B) Activity levels of control (n = 6) and infected (n = 14) mice. Error bars show mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (C) Representative actograms of daily wheel-running activity of control and infected mice in light-dark cycles. Animals were treated with suramin i.p. on day 21 post-infection. (D) Daily and relative rest-period activity levels of control (n = 6) and infected (n = 14) mice. Error bars show mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (E) Representative actograms of core body temperature of control and infected mice in light-dark cycles. Animals were treated i.p. with suramin on day 21 post-infection. (F) Amplitude of circadian temperature oscillation control (n = 5) and infected (n = 7) mice. Bar graphs show mean ± SEM.

When we treated mice with suramin at day 21 post-infection, we observed that infected mice still run 1.5-fold less than control mice (fig. S1D). However, when considering the percentage of daily activity spent running during the rest period (day) was only 1% in control mice and ~15% in infected mice 50 days post-infection (Fig. 1C-D and fig. S1C). Altogether (day and night) infected mice run less than control mice, but even when considering the total number of wheel revolutions during the light period (day), infected mice run 4.5-fold more than controls (fig. S1D, unpaired t test $p < 0.05$). This is extremely surprising because the inhibitory effects of light in running-wheel behavior have been extensively studied (*202, 203*). Remarkably, after day 50 four infected mice displayed dramatic changes in their behavior, running more than 25% of their daily activity during the rest period (Fig. 1C-D and fig. S1C).

To further confirm that *T. brucei* infection disrupts circadian rhythm, we measured the core body temperature by telemetry (*204*). Mice were implanted under deep anesthesia during the light phase with an i.p. temperature sensor. We observed an initial fever-like peak that lasted ~2 days matching the first peak of parasitemia (day 5-6 post-infection), followed by a normal oscillation of high temperature during the night and low during the day (Fig. 1E and fig. S1E). However, around 60 days post-infection the highest core temperature shifted from the night to the day period, resembling the activity increase we observed in the wheel activity experiment (Fig.1B and E). This indicates that circadian temperature regulation is also affected, which is not necessarily always synchronized with the regulation of running-wheel activity (*205*).

Interestingly, since the beginning of infection the temperature amplitude, which assesses the robustness of the circadian oscillation, was reduced in infected mice (Fig. 1F). The telemetry system also allowed us to measure general cage activity. We observed, once again that later in the infection the general cage activity is increase during the rest (light) period (fig. S1F).

Using an additional independent assay for determination of sleep-like behavior by automated video analysis (CleverSys), we confirmed that later in infection (day 95) infected mice spend more time resting during the active phase than control animals, even though they do not rest more over the 24 h

(fig. S2). Together these results show that *T. brucei*-infected mice show a similar disruption of circadian activity along with daily temperature deregulation as described in humans.

Because mice started running during light period, which is incredibly rare with wild-type mice, we wondered whether they vision was compromised affecting the light perception and therefore inhibition of running-wheel activity during the day. The elevated plus maze, in which animals are preferentially restricted to the closed arms, avoiding the open/exposed arms, showed no differences between trypanosome-infected and control mice. This suggested that even later during infection, mice have visual depth perception. Next we asked if this infection disrupts circadian rhythms in constant darkness, when the internal circadian clock is the only temporal cue. In the absence of light, circadian clock imposes a stable wheel running activity with a period of ~23.7 h, with more activity during the night phase (Fig 2A-C) (*134*). To address if *T. brucei* infection altered the internal clock period, we infected and let the infection progress in the darkness only. We observed that running-wheel activity period was shorter in infected mice than in control mice (Fig. 2A-D, 30 min on average, unpaired t test). Unlike control mice, whose period is stable over time, infected mice showed a progressively shorter period (Fig. 2B-D): from ~23.5 h in the first 10 days to ~23.1 h in d70-80. Interestingly, this shorter period was immediately detected in the first ten days post-infection, a period in which very few or no parasites can be detected in the brain (Fig. 2E) (*82*). This short period is also detected in later time points (after day 40), when parasites are exclusively in the brain due to suramin treatment. Our results suggest that changes in mice behavior may be not be exclusively due to presence of parasites in the brain, but rather a consequence of a peripheral signal, such as a metabolite or hormone that feedback on the master clock.

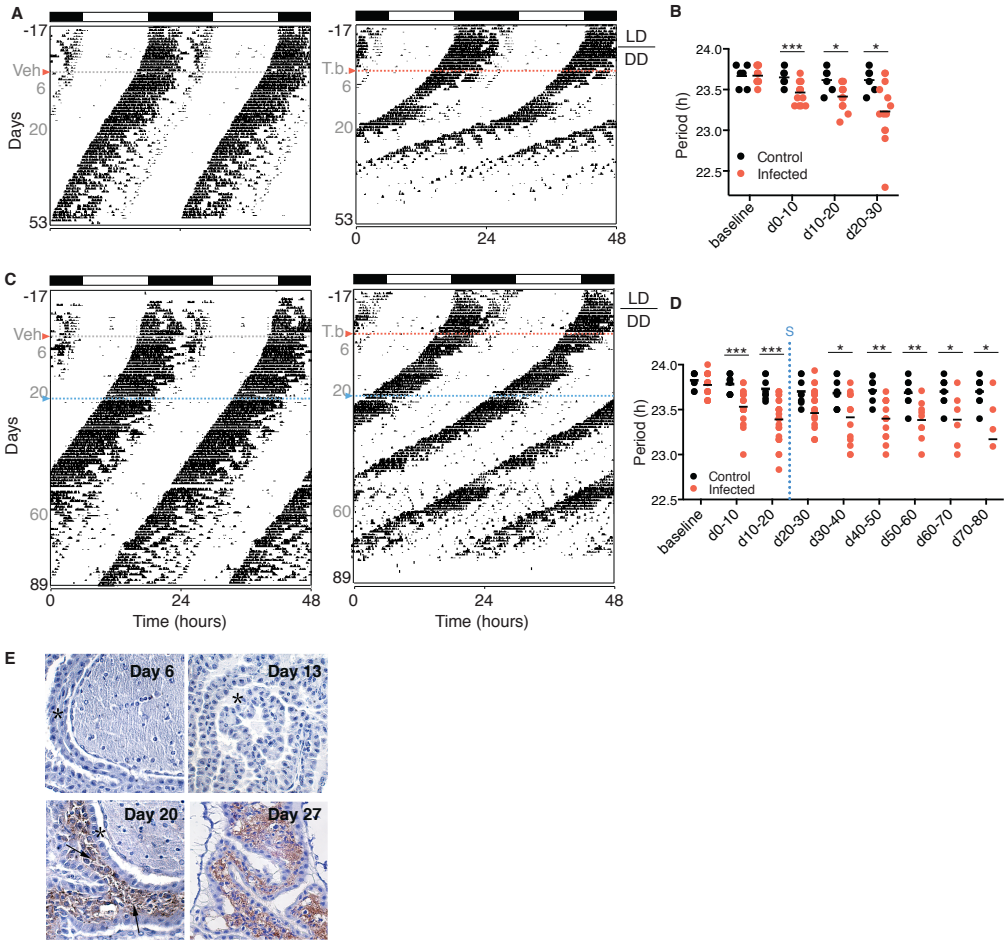


Figure 2. Shorter period of running-wheel activity of *Trypanosoma*-infected mice. (A) Representative actograms of daily wheel-running activity of control and infected mice in constant darkness (DD). All running-wheel experiments involve 7 days acclimatization period to wheels followed by 10 days dark period to confirm all animals have a normal circadian rhythm, after which animals are either infected or injected with vehicle in the dark. Activity records were double plotted, with each day being represented beneath and also to the right of the preceding day. Horizontal black and white bars at the top of each actogram represent lights off and on, respectively. (B) Period of running-wheel activity of control (n = 6) and infected (n = 14) mice. Bar graphs show mean ± SEM. (C) Representative actograms of daily wheel-running activity of control and infected mice in constant dark. Animals were treated with suramin i.p. on day 21 post-infection. (D) Period of control (n = 8) and infected (n = 27) mice. Error bars show mean ± SEM. (E) Temporal histological analysis of infected brain. Representative images of the choroid plexus area (n = 5/group).

Since both behavior and core temperature of the mouse were disrupted upon infection, next we interrogated if this could be explained by changes in the

expression of core clock genes. For this, we infected PERIOD2::LUCIFERASE mice (PER2::LUC) (*206*), we sacrificed animals and we collected multiple organs to measure *ex vivo* the daily PER2 expression. When mice were sacrificed 6 days post-infection no changes in period were detected, but interestingly the phase of the adipose tissue depots was as advanced. To test whether this change in PER2 expression was stronger once more parasites have infiltrated the organs, we tested the expression of PER2::LUC in the same organs on day 20. 20 days post-infection, we observed that most infected organs kept the normal circadian parameters of phase, damping and period of the controls (Fig. 3A-C and fig. S3A). Adipose tissue, however, which has the highest parasite load (chapter I of this thesis), had a ~2 h shorter period (Fig. 3A-B and S3B, unpaired t test, $p < 0.001$). This tissue also showed a phase advance as observed on day 6 and importantly remained healthy, with damping remaining unaffected. This was observed in both gonadal and perirenal adipose tissue depots (Fig. 3B-C). The period shortening observed on day 20 and phase advance for both day 6 and 20 in adipose tissue suggests that even though on day 6 the presence of parasites was not enough to lead to a shorter period, these depots are already having their maximum PER2:LUC expression earlier in the day than controls, and as the parasite load increases a shorter period is detected, even *ex vivo* (Fig. 3C).

To test whether the shorter period of infected adipose tissue could be reverted by eliminating the parasites, we treated animals with suramin and assessed PER2::LUC expression profile at day 60. When parasites were eliminated from the peripheral organs, the period of the adipose tissue was the same as in control organs (Fig. 3D-E), suggesting that the shorter period was a consequence of high parasite load. Remarkably, on day 60 the period of the SCN of infected mice was 30 min shorter than control SCNs (Fig. 3D-F and S3B, unpaired t test, $p < 0.01$). This is particularly interesting because, in light/dark conditions only around day 60 we observed disruption of the circadian activity and body temperature of infected mice (Fig.1C-F). Together these data suggest that the cumulative parasite load in the peripheral organs leads to changes in the period of the tissue circadian clock. The SCN seems more

resistant to circadian changes than peripheral organs and only at day 60 could we detect a statistically significant reduction of the period.

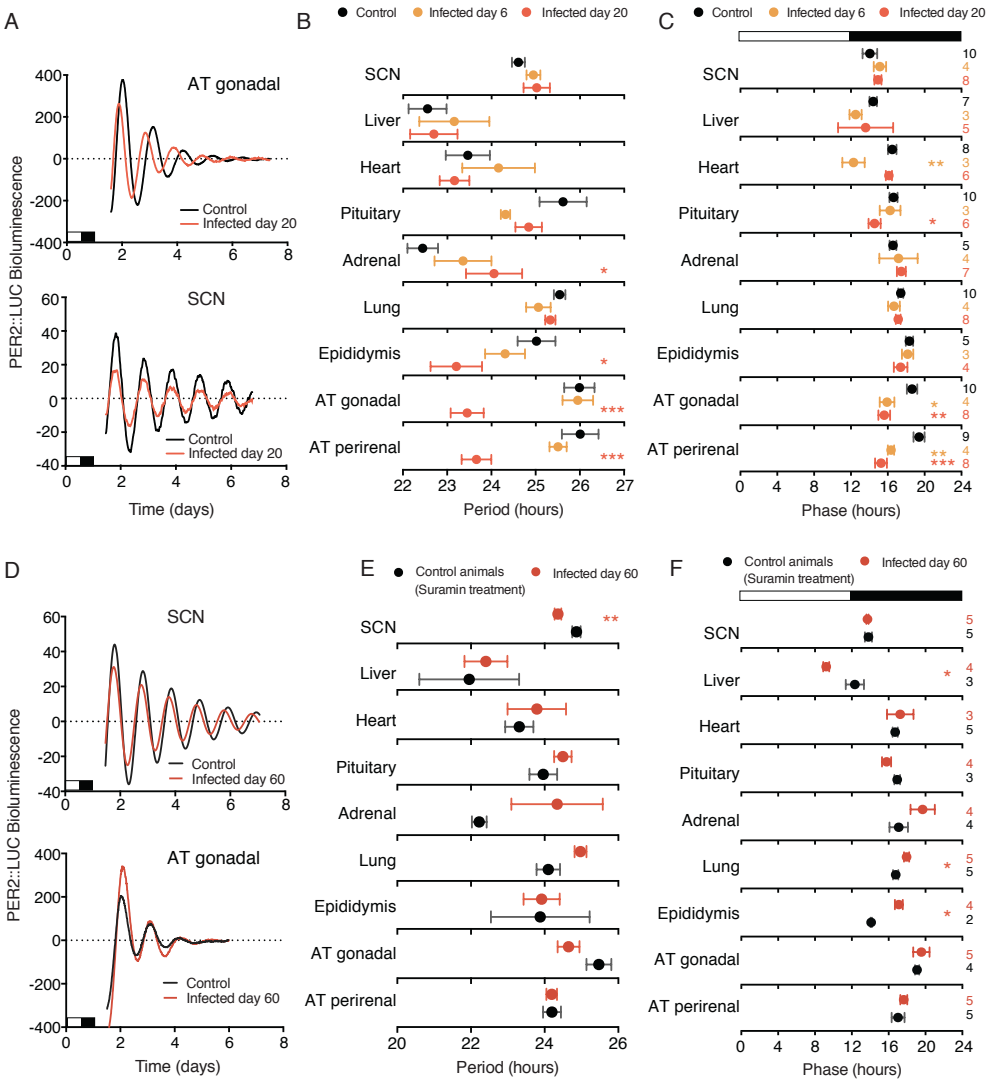


Figure 3. Shorter period of PER2::LUC explants when high number of parasites is present. (A) Representative normalized records of bioluminescence reporting of circadian expression from gonadal adipose tissue and SCN on day 20 post-infection/vehicle. Tissues were prepared from mice in LD. Shown are 7 days of continuous recording after explant preparation. (B, C) Period and phase plots of various tissues harvested from control (black) and infected mice (day 6 yellow, day 20 orange) mice. The sample size is indicated on right. Shown are mean period ± SD. *p < 0.05, **p < 0.01, ***p < 0.0001 by unpaired t test. (D) Representative records of bioluminescence reporting of circadian expression from gonadal adipose tissue and SCN on day 60 post-infection/vehicle. Animals were treated with suramin. Tissues were prepared from mice in LD. Shown are 7 days of continuous recording after explant preparation.

(E, F) Period and phase plots of various tissues harvested from control (black) and infected mice (day 60 dark orange) mice. The sample size is indicated on right. Shown are mean period ± SD. *p < 0.05, **p < 0.01, ***p < 0.0001 by unpaired t test.

To test if the expression of other clock genes was also affected upon infection, we infected 18 mice kept in the darkness and on day 20 post-infection we collected and extracted RNA from organs. 3 mice were sacrificed every 4 h, with the time defined for each time point estimated based on their running-wheel activity circadian time (*203*). We measured the gene expression of genes involved in the molecular clock: transcriptional activator *Bmal1*, transcriptional repressors *Per1* and *Rev-erba,* and an immediate downstream target of the clock D site of albumin promoter (albumin D-box) binding protein (*Dbp*). As expected, even after 20 days in the dark, control mice's clock gene expression cycled in liver, adipose tissue and hypothalamic area (Fig. 4A and S4A). It is clear the expected daily oscillations of *Bmal1*, with its maximum expression at circadian time 24 h (CT24). BMAL1 protein heterodimerizes with CLOCK and activates transcription of *Per1*, *Rev-erba* and *Dbp*, whose maximum expression is detectable at CT8-12. On the contrary, the circadian transcript profile of clock genes is highly disrupted when mice are infected by *T. brucei* (Fig. 4A and S4A, two-way ANOVA).

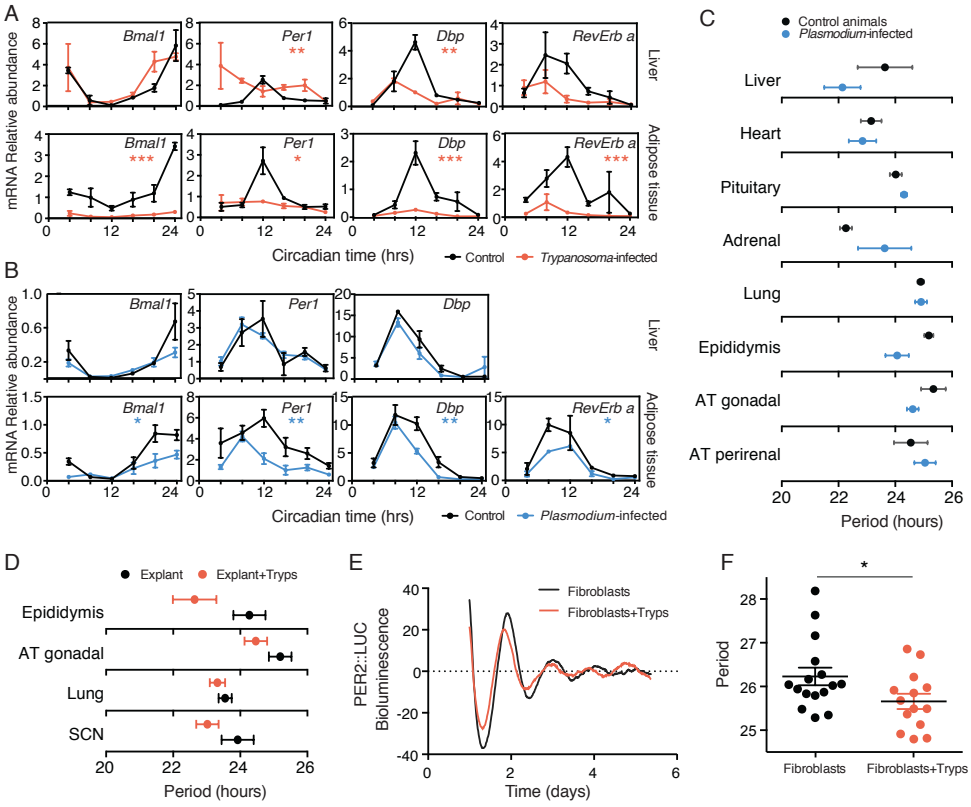


Figure 4. Shorter period is not common in every infection. (A) Real-time RT-PCR analysis of clock gene expression in control (black) and *Trypanosoma*-infected (orange) mice. Two organs are shown, liver and gonadal adipose tissue. Error bars represent SEM for each time point from three independent replicates. Two-way ANOVA shows significant statistical differences between control and *Trypanosoma*-infected mice, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (B) Real-time qPCR analysis of clock gene expression in control (black) and *Plasmodium*-infected (blue) mice. Two organs are shown, liver and gonadal adipose tissue. Error bars represent SEM for each time point from three independent replicates. Two-way ANOVA shows significant statistical differences between control and *Trypanosoma*-infected mice, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (C) Period plot of various tissues harvested from control (black) and *Plasmodium*-infected mice (blue) mice. The sample size is indicated on right. No significance was found, unpaired t test. (D) Period plot of various tissues harvested from PER2::LUC mice and co-cultured with medium only (black) or 10^5 *Trypanosoma brucei* (orange). The sample size is indicated on right. No significance was found, unpaired t test. (E) Representative records of bioluminescence reporting of circadian expression from PER2::LUC ear fibroblasts co-cultured with 10^5 *Trypanosoma brucei* (orange). (F) Period comparison from PER2::LUC ear fibroblast with medium (black) and co-cultured 10^5 *Trypanosoma brucei* (orange). Unpaired t test, * $p < 0.05$.

To test if the disruption of clock gene circadian expression is a specific effect of *T. brucei* infection, we infected mice with *Plasmodium chabaudi*

parasite. Both these infections are chronic in the mouse model, they have comparable immune response and similar tissue sequestration (fig. S4B-D). We infected 18 mice with *P. chabaudi* and right before the second peak of parasitemia mice were housed in the dark for 48 h. Organs were collected during the second peak of parasitemia, every 4 h and the same transcript levels of the same clock genes were measured by qPCR. We observed that the expression profile of clock genes in the liver was similar in both control and *Plasmodium*-infected mice. However, in the adipose tissue expression of clock genes was in general lower in *Plasmodium*-infected mice (Fig. 4B, two-way ANOVA). Even though the effects observed in clock gene expression of *T. brucei*-infected mice, the fact that to some extent there was downregulation of clock gene expression in the adipose tissue of *Plasmodium*-infected mice is likely to be due to inflammation, since it has been previously demonstrated that TNF-α impairs expression of clock genes in synchronized NIH 3T3 fibroblasts (*200*). The period of oscillations however, remained the same in the control fibroblasts in the same study (*200*). To assess wether the period of clock gene oscillations was affected, we infected PER2::LUC mice with *P. chabaudi*. None of the organs tested showed significantly different period than control mice (Fig.4C, unpaired t test, $p > 0.05$), indicating that unlike *T. brucei*, a *P. chabaudi* infection does not cause any detectable circadian alterations.

Together our data seem to dissect two different effects of the *T. brucei* infection: i) down-regulation of clock gene expression (common to other infections, inflammation-driven); ii) period shortening (probably not inflammation-driven).

To further exclude the contribution of inflammation, we co-cultured healthy explants of SCN, lung, gonadal adipose tissue and epididymis of PER2::LUC mice with 10^5 *T. brucei* parasites. Luminescence was recorded for seven days and circadian parameters estimated as above. Whereas no significant changes were detected in terms of phase, amplitude and period, in the presence of parasites period tended to be shorter (Fig. 4D). Perhaps the number of parasites co-cultured was not enough to cause changes in such a high number of host cells. Unfortunately due to technical limitations we could not test with higher number of parasites (*T. brucei* in culture differentiate into a

non-replicative form and die if their density is higher than 1-3x10^6/mL), instead we tested with lower number of mouse PER2::LUC cells, primarily cultured ear fibroblasts. We observed that, in the presence of *T. brucei*, fibroblasts period was 30 min shorter (Fig.4E-F, unpaired t test, $p < 0.05$). These data suggest that the presence of parasites, and not the immune system is responsible for the changes in period of host circadian clock.

2.4. Discussion

In this study, we showed that *T. brucei* affects both behavior and temperature regulation of its host, which seem to happen since the beginning of the infection, prior to the detection of parasites in the brain. The period of PER2::LUC infected adipose tissue showed 2h shorter period than control, which could be reverted when parasites were cleared by suramin treatment. The period of PER2::LUC SCN of infected mice was shorter around day 60, when we also detected increase activity and body temperature during the rest phase of these mice. Interestingly, *P. chabaudi*-infected organs showed a slight downregulation of clock gene expression but their period remained similar to controls. Finally, in co-culture experiments with *T. brucei*, fibroblasts showed shorter period, which may mean that parasites, and not immune response, directly accelerate the host's molecular clock.

Throughout the years evidence has been gathered that connects inflammation and in particular cytokines with fever, fatigue, and sleep disturbances, which are collectively referred to as sickness behavior syndrome. By injecting TNF-α, a pro-inflammatory cytokine, the locomotor activity of mice is dramatically decreased (*200*). TNF-α and IL-1β are also both somnogenic, increasing nonrapid eye-movement (NREM) sleep (*207, 208*). TNF-α leads to the suppression of the expression of clock downstream genes *Dbp*, *Tef*, and *Hlf* and of the *Period* genes *Per1*, *Per2*, and *Per3* (*200*). So, in general cytokines (*209-211*) and even bacterial infection (*146*) have been shown to affect the circadian clock, which may be a common factor to all infections and responsible for the sickness behavior.

Although this may explain why *Trypanosoma*-infected mice run less than healthy control mice, and why both *Trypanosoma*- and *Plasmodium*-infected mice show lower transcript levels of clock genes, it cannot explain the differences in timing of expression. Although TNF-α injection decreases activity, it does not alter mice circadian rest-activity cycles (*200*). Similarly, cytokines damped the oscillations of clock gene expression but always kept the same period (*200*). The fact that the period of PER2::LUC *Trypanosoma*-infected mice did not change on day 6, when there is already very high inflammation, at least in the bloodstream (*198, 201*), and that despite inflammation (*212*), *Plasmodium*-infected mice organs also maintained their period, suggests that period shortening induced during *T. brucei* infection is not due to inflammation. Furthermore, *in vitro* PER2::LUC fibroblasts co-cultured with *T. brucei* showed shorter period. Together these results suggest that *Trypanosoma brucei* appears be the first infection capable to disrupt the period of the host circadian clock.

Ex vivo analysis of PER2::LUC expression in neurons of the SCN of *T. brucei*-infected mice showed that this remained similar to controls on day 6 and day 20 post-infection. On the contrary, on day 60 post-infection the period of PER2::LUC expression of SCN was shorter than controls. This could be due to either higher parasite infiltration in the brain parenchyma on day 60 than on day 20, or to the fact that the master clock has a more robust molecular clock (*213*), with higher intrinsic resistance that requires more time to be affected. One or both of these possibilities are likely the cause for why, even when significantly shorter than controls, SCN period is only 30 min shorter and not 2 h as observed for the high infiltrated adipose tissue.

Interestingly, when behavior of infected mice was recorded in the drak, we detected an immediate shorter running period, even before the parasites accumulate in the brain. Since running behavior is controlled by the master clock in the brain, SCN, these observations mean that the presence of high numbers of *T. brucei* in the brain is not necessary for the master clock to be disrupted. Probably a peripheral molecule(s), metabolite or hormone, from the host or directly produced by the parasite, is released from the periphery and it

feeds back to the master clock. This observation is consistent with a study in Uganda, in which 58% and 54% of *T. brucei rhodesiense* infected patients experienced somnolence in both early- and late-stage, respectively (*214*). Similar observations were made in *T. brucei gambiense (23, 215).* This suggests that, even though this disease is divided into two stages: an early-stage with parasites only in the periphery and a late-stage when parasites penetrate the blood brain barrier, patients may experience sleep disruption in both stages.

In this study we described the first infection that interferes with the host circadian rhythm at the behavior, cellular and molecular levels. These changes are likely the cause of sleep alterations that are typical of sleeping sickness. Further studies are required to identify what systemic signal (secreted by the parasite, or a molecule produced by the host in response to this infection) that could be responsible for such changes in the period of the circadian clock.

2.5. Materials and methods

Ethics Statement

All animal care and experimental procedures were performed in accordance with University of Texas Southwestern Medical Center (UTSW) IACUC guidelines, approved by the Ethical Review Committee at the University of Southwestern Medical Center and performed under the IACUC-2012-0012 protocol and according to European Union (EU) Directive 2010/63/EU and approved by the Animal Ethics Committee of Instituto de Medicina Molecular (IMM) (AEC-2011-006-LF-TBrucei-IMM), following FELASA guidelines concerning laboratory animal welfare.

Parasites and Culture Conditions

T. brucei AnTat 1.1^{E}, a pleomorphic clone, derived from an EATRO1125 clone was originally isolated from blood of *Tragelaphus scriptus* in Uganda. For all the experiments, we used AnTat 1.1^{E} 90-13, a transgenic cell-line encoding the tetracyclin repressor and T7 RNA polymerase (23). For all mice infections, the parasite cryostabilates used were obtained by a previous infection, on day 5

post-infection. For co-culture experiments, bloodstream forms were grown routinely in HMI-11 at 37°C in 5% CO2 (24). Parasite numbers were calculated using a Hemocytometer.

Trypanosoma brucei mice-infection

The infections of wild-type male C57BL/6J mice, 6-10 week old (UT Southwestern Medical Center Mouse Breeding Core Facility or Charles Rivers, France) described in this manuscript were performed by intraperitoneally (i.p.) injection of 2,000 *T. brucei* AnTat 1.1^{E} parasites. Prior to infection, *T. brucei* cryostabilates were thawed and parasite viability and numbers were assessed by mobility under an optic microscope. For circadian luminescence experiments mice were group-housed and maintained in a Specific-Pathogen-Free barrier facility. The facility has standard laboratory conditions: 21 to 22°C ambient temperature and a 12 h light/12 h dark cycle. For behavioral and telemetry experiments mice were individually housed. Chow and water were available *ad libitum*.

Circadian behavioral experiments

80 mice were placed in individual running wheel cages and their activity recorded continuously using the ClockLab data collection system (Actimetrics, Evanston, IL) (*203*). After the first week on an LD12:12 cycle, the mice were released into DD for 10 days and then infected with *T. brucei* either back on LD cycle (28 mice) or remaining in DD (28 mice) for the rest of the experiment. 14 infected mice in LD cycle and 14 mice in DD were treated with suramin on day 21 post-infection, at ZT10-11. Free running periods were measured by line fitting of activity onsets from data collected during the DD period of the assay using ClockLab data analysis software.

Core body temperature measurement

12 mice were implanted with minimitters inside the peritoneal cavity. Animals were monitored for 3 weeks to recover from the surgery prior to recordings. Baseline recordings of 4 days in LD12:12 cycles and 10 days in DD. Seven mice were infected whereas 5 were injected with HMI-11 only. For the

remaining period of the experiment the light schedule was LD12:12 cycles. On day 21 post-infection all animals were injected i.p. with suramin.

CleverSys behavioral assay

6 mice were injected i.p. with either HMI-11 or *T. brucei*. 21 days post-infection animals were injected with suramin. On days 50, 65 and 95 post-infection animals were individually housed and their movement recorded for 48 h using CleverSys software. Data were auto scored by CleverSys.

RNA Isolation and Real-Time PCR of Trypanosoma brucei-infected Mice

Animals were housed individually and activity monitored using a running wheel. 36 animals were either infected or injected with medium and released into constant darkness. Locomotor activity was recorded and analyzed using ClockLab software (Actimetrics, Wilmette, IL) to determine the circadian phase, as previously described (*203*), for each animal on day 20 post-infection. Three animals per phase cluster (6 time points) were euthanized by cervical dislocation, organs collected and snap frozen.

RNA isolation and real-time PCR were performed as described previously (*216*). Transcript levels were normalized to *Gapdh*. Primer sequences are listed in table S1.

RNA Isolation and Real-Time PCR of *Plasmodium chabaudi*-infected Mice

36 wild-type male C57BL/6J mice, 6-10 week old (Charles Rivers, France) were either infected by i.p. injection of $1x10^5$ *P. chabaudi* infected red blood cells (iRBC) in 200 μL PBS or injected with PBS as control animals. Parasitemia was assessed daily. On the second peak of parasitemia animals were released into constant darkness for 48 h. Three animals per each of the 6 time points (every 4 h) were euthanized by cervical dislocation, organs collected and snap frozen. RNA extracted and gene expression quantified by real-time PCR (*216*).

Circadian Bioluminescence Experiments

Trypanosoma-Infected and uninfected *mPer2::Luc* mice (*206*) were euthanized by cervical dislocation between ZT10 and ZT12 at days 6, 20 and 60 post-

infection. SCN tissues were isolated from 300 μm coronal sections and pituitary, heart, lung, liver, adrenal, adipose tissue and epididymis were dissected and kept in chilled Hanks' buffered salt solution (Invitrogen) until culture. All dissected tissues were cultured on Millicell culture membranes (PICMORG50, Millipore) and were placed in 35 mm tissue culture dishes containing 2 mL TbM50 medium, composed of 50% (v/v) DMEM media (Mediatech) supplemented with 2 mM L-Glutamine, 25 units/mL penicillin, 25 μg/mL streptomycin (Invitrogen) and 50% (v/v) HMI-11 no phenol 5% FBS and 0.1 mM luciferin potassium salt (L-8240, Biosynth AG). Sealed dishes were placed in LumiCycle luminometer machines (Actimetrics, Wilmette, IL) and bioluminescence was recorded continuously.

Plasmodium-infected *mPer2::Luc* mice were euthanized on day 16 post-infection and organs were processed similarly.

Primary ear fibroblasts were isolated from the ear of PER2::LUC mice using 5 mg/mL collagenase I and 0.05% trypsin incubating 30 min at 37ºC.

2.6. Supplementary figures and tables

Table S1 – primers

Gene	Fw/Rev	Seq (5' - 3')
Reverb a	Fw	AGA CTT CCC GCT TCA CCA AG
Reverb a	Rev	AGC TTC TCG GAA TGC ATG TT
Period 1	Fw	CCC AGC TTT ACC TGC AGA AG
Period 1	Rev	ATG GTC GAA AGG AAG CCT CT
Dbp	Fw	CGAAGAACGTCATGATGCAG
Dbp	Rev	GGTTCCCCAACATGCTAAGA
Gapdh	Fw	caaggagtaagaaaccctggacc
Gapdh	Rev	CGAGTTGGGATAGGGCCTCT
Bmal1	Fw	CCACCTCAGAGCCATTGATACA
Bmal1	Rev	GAGCAGGTTTAGTTCCACTTTGTCT
IL-10	Fw	TGCTATGCTGCCTGCTCTTA
IL-10	Rev	TCATTTCCGATAAGGCTTGG
TNF-a	Fw	AATGGCCTCCCTCTCATCAGTT
TNF-a	Rev	CCACTTGGTGGTTTGCTACGA
IL-6	Fw	TTCCATCCAGTTGCCTTCTT
IL-6	Rev	CAGAATTGCCATTGCACAAC
IL-1	Fw	GCCCATCCTCTGTGACTCAT
IL-1	Rev	AGGCCACAGGTATTTTGTCG

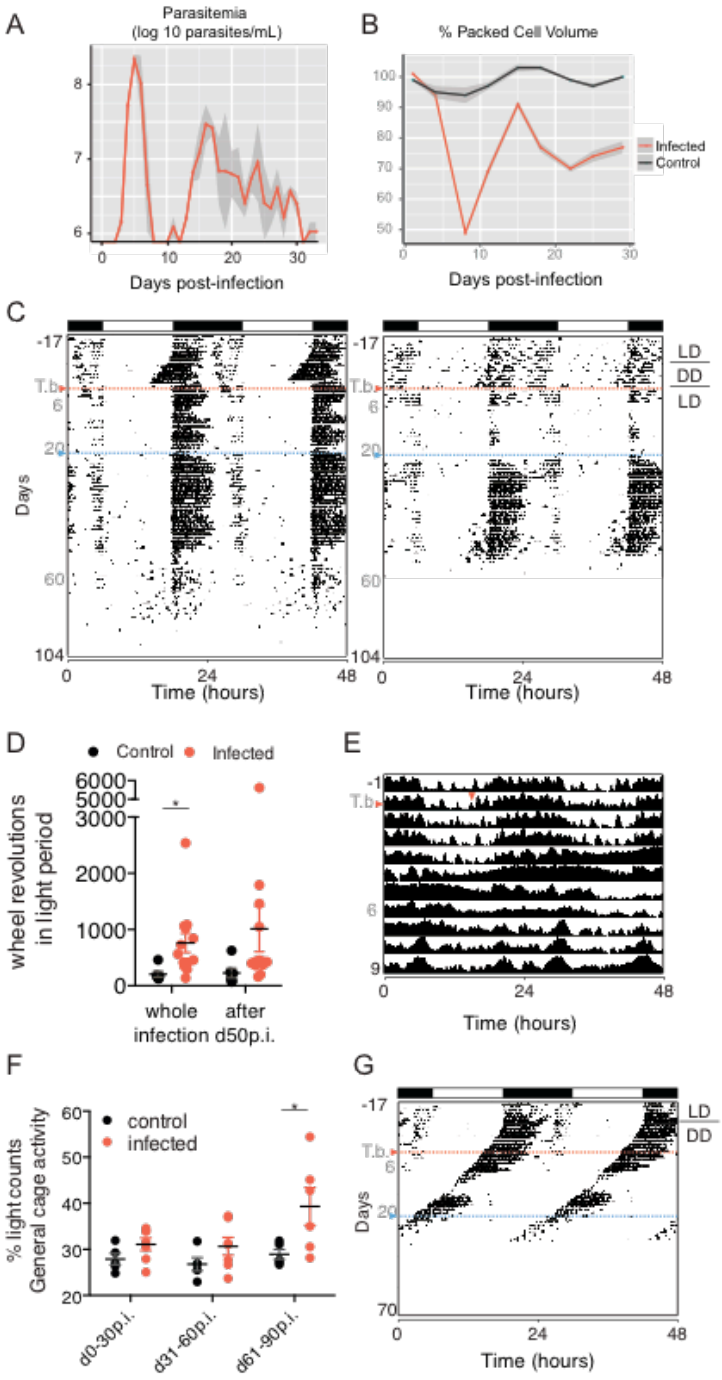


Figure S1. Circadian activity and body temperature disruption in *Trypanosoma*-infected mice. (A) Parasitemia profile of *T. brucei* infected mice. (B) Hematocrit of both control (black) or *Trypanosoma*-infected (orange) mice. (C) Two additional actograms of daily wheel-running activity of infected mice in light-dark cycles. Animals were treated with suramin i.p. on day 21 post-infection. (D) Total activity levels during rest period of mice infected and treated with suramin on day 21. Error bars show mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (E) Representative actogram of telemetry measured body temperature in the initial nine days of infection revealing a ~2 day fever period in infected mice. (F) Telemetry measured general cage

activity of control (n = 5) and infected (n = 7). Error bars show mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (G) Additional actogram of daily wheel-running activity of infected mice in DD.

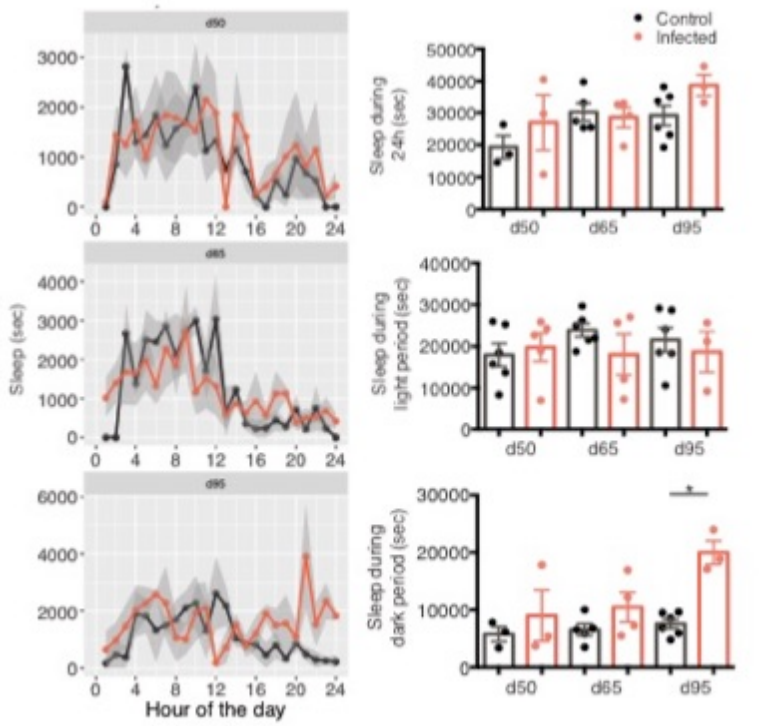


Figure S2. Automated video analysis revealed sleep-like behavior disruption in *Trypanosoma brucei*-infected mice. (A) Daily amount of time spent resting in both control (black) and *Trypanosoma*-infected mice (orange) on day 50, 65 and 95 post-infection. Shaded area represent SEM. (B) Quantification of amount spent resting on day 50, 65 and 95 post-infection in the entire 24 h, day- or night-period. error bars represent SEM.

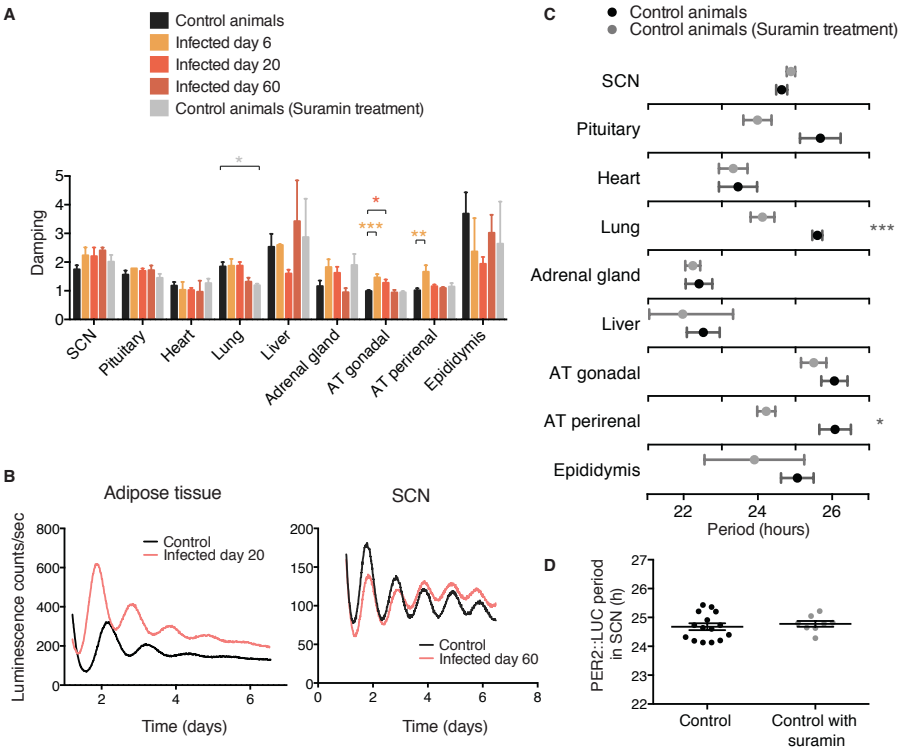


Figure S3. Shorter period of PER2::LUC explants when high number of parasites is present. (A) Damping analysis of various tissues harvested from control or infected mice in DD. Control (black, pooled from day 6 and day 20), infected day 6 (yellow), infected day 20 (orange), infected day 60 (dark orange, treated with suramin on day 21) and control on day 60 treated with suramin (gray). Shown are mean period ± SD. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (B) Representative records of bioluminescence reporting of circadian expression from gonadal adipose tissue, day 20 and SCN on day 60 post-infection/vehicle. (C) Period plots of various tissues harvested from control (black) and control day 60 treated with suramin (gray) mice. Shown are mean period ± SD. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test. (D) Period plots of SCN harvested from control (black) and control day 60 treated with suramin (gray) mice show no effect of suramin in this tissue.

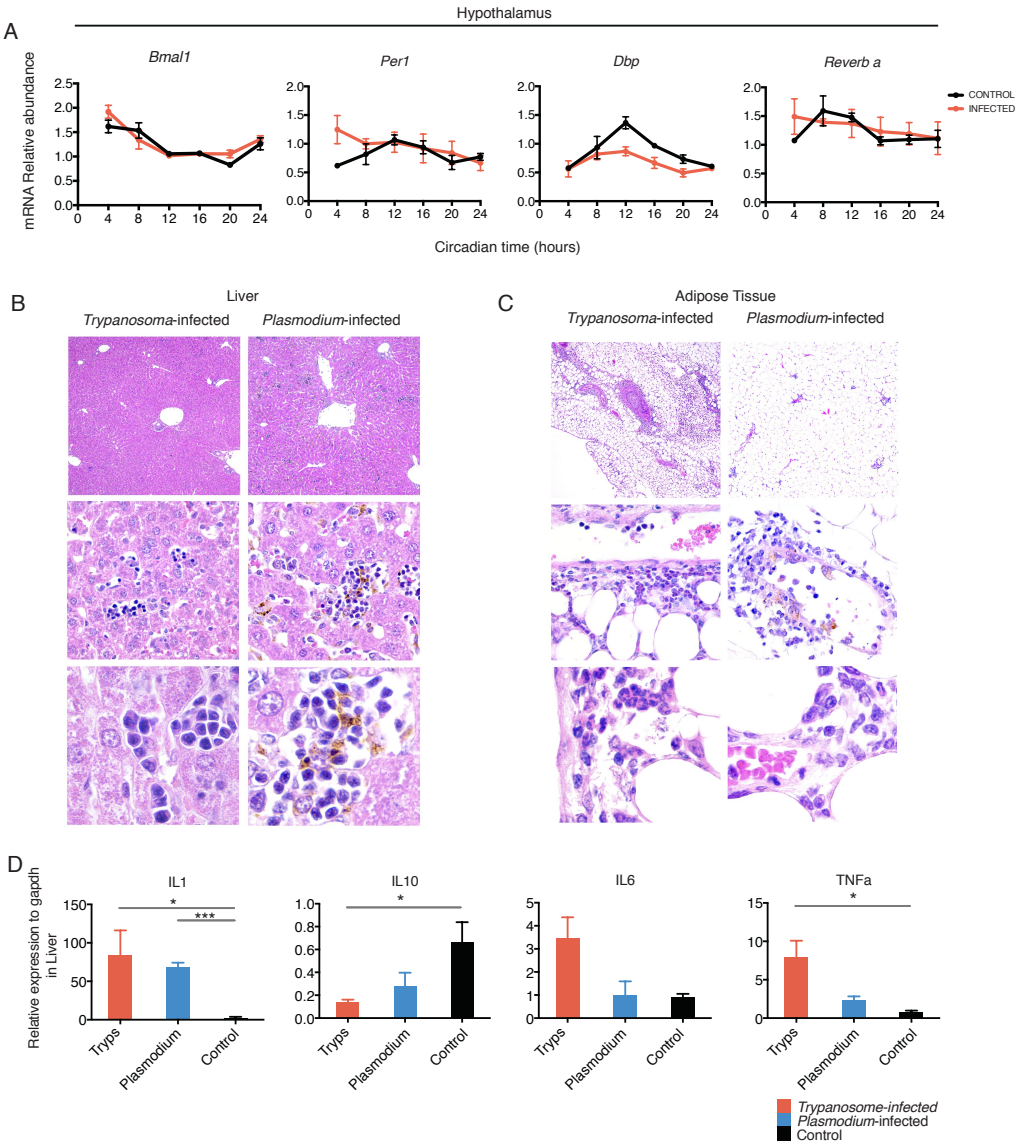


Figure S4. Shorter period is not common in every infection. (A) Real-time RT-PCR analysis of clock gene expression in control (black) and *Trypanosoma*-infected (orange) mice hypothalamus. Error bars represent SEM for each time point from three independent replicates. (B) Histological analysis of Liver from mice infected with *T. brucei* (left column) and *P. chabaudi* (right column), 20 and 16 days post-infection, respectively. For *T. brucei*, there is infiltration by mononuclear cells, periportal and parenchymal, multifocal and of minimal to mild severity. For *P. chabaudi*, inflammatory cell infiltrates are also mononuclear-cell rich, frequently laden with hemozoin (brown pigment), and distribution is mainly parenchymal, multifocal and of moderate severity. (C) Histological analysis of adipose tissue from mice infected with *T. brucei* (left column) and *P. chabaudi* (right column), 20 and 16 days post-infection, respectively. For both infections, adipose tissue displays inflammatory cell infiltrates, but the cell type, topography, distribution and severity are distinct. For *T. brucei*, mixed inflammatory cell infiltrates are rich in polymorphonuclear eosinophils and mononuclear cells, macrophages, lymphocytes and some plasma cells; cells are located in the adipose tissue stroma, with diffuse distribution and the lesion is severe. Trypanosomes are frequently seen, scattered among the inflammatory infiltrates. For *P. chabaudi*

inflammatory cell infiltrates are mononuclear-cell rich, composed of macrophages and lymphocytes; are mainly ascribed to the perivascular compartment, with multifocal distribution and of mild to moderate severity. Infected red blood cells are frequently seen, sequestered in the adipose tissue vessels/capillaries. (D) Cytokine expression levels in liver of control and infected mice (n = 5/group). Error bars show mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 by unpaired t test.

2.7. Acknowledgments

We thank Chryshanthi Joseph and Victoria Acosta Rodriguez for help during the implant surgeries of the telemetry system, Lisa Thomas for breeding the PER2::LUC mice, Tânia Carvalho for help taking the histological images and discussing results, Cristina Afonso for help in the CleverSys recording, and Vanessa Luis and Jeremy Stubblefield for help during the 24h tissue collection. The work presented has been supported by Howard Hughes Medical Institute to J.S.T, by HHMI International Early Career Scientist (55007419) to L.M.F., by Fundação para a Ciência e Tecnologia SFRH/BD/51286/2010 to F.R-F

Chapter II – section 3

Post-Transcriptional Circadian Regulation of Genes Related to Metabolism and Suramin Sensitivity in *Trypanosoma brucei*

The "Chapter II – Results – 3. Post-Transcriptional Circadian Regulation of Genes Related to Metabolism and Suramin Sensitivity in *Trypanosoma brucei*" consists of a third manuscript has been submitted for publication. Here is clarified the contribution of each of the authors:

Manuscript:

Filipa Rijo-Ferreira[1,2,3], Daniel Pinto-Neves[2], Nuno L Barbosa-Morais[2], Joseph S Takahashi[1,4*#], Luisa M. Figueiredo[2*#] *Manuscript submitted.*

[1]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA
[2]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[3]Graduate Program in Areas of Basic and Applied Biology, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal
[4]Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA
*To whom correspondence should be addressed. E-mail: joseph.takahashi@utsouthwestern.edu and lmf@medicina.ulisboa.pt
[#]These authors contributed equally to this work

[#] these authors contributed equally to the present work

Filipa Rijo-Ferreira, the candidate for the PhD degree at the University of Porto, contributed to original idea of this project, and was responsible for the design and performance of all experiments. Filipa also participated in the bioinformatics analysis, which was mainly done by Daniel Pinto-Neves. In addition, she wrote the manuscript as an integral part of her future paper (submitted) and of the thesis. Dr Figueiredo and Professor Takahashi designed and supervised the project and revised the manuscript. Nuno L. Barbosa-Morais helped and supervised the bioinformatics analysis performed in this manuscript.

3. Post-Transcriptional Circadian Regulation of Genes Related to Metabolism and Suramin Sensitivity in *Trypanosoma brucei*

3.1. Abstract

Free-living organisms have evolved circadian molecular clocks to anticipate daily changes in their environments; however, it is unclear whether parasites have intrinsic clocks. *Trypanosoma brucei,* the causative agent of sleeping sickness, is a unicellular parasite that switches between two hosts, a vertebrate and a tsetse fly vector. Here we show that approximately 10% of genes in *T. brucei* are expressed with a circadian rhythm in both entrained and constant conditions. The maximum expression of these genes occurs at two different phases of the day and depend on a post-transcriptional mechanism. Circadian genes are enriched in cellular metabolic pathways, and coincide with two peaks of intracellular ATP concentration. Moreover, daily changes in the parasite population lead to differences in suramin sensitivity, a drug commonly used to treat this infection.

One Sentence Summary: *Trypanosoma brucei* parasites have a post-transcriptional intrinsic circadian clock that regulates metabolism and sensitivity to suramin.

3.2. Main text

The Earth's rotation forced life to evolve under cyclic day and night environmental changes. In order to anticipate such daily cycles, prokaryote and eukaryote free-living organisms evolved intrinsic clocks that regulate physiological and behavioral processes. These molecular clocks are robust and run with a period of approximately 24 h even in the absence of external environmental stimuli, such as light or temperature cycles (*217*). Daily rhythms have been observed in organisms living within hosts, such as parasites. Human species of filariae show daily fluctuation of microfilariae density (the

transmissible form) in their hosts' peripheral blood. These parasites migrate from the lymphatic system to the peripheral blood at night, which ensures transmission through their vector's blood-meal (*152*). Human malaria parasite species display synchronous cell cycle durations of 48 or 72 h causing fever every 2 or 3 days, a hallmark of this disease (*153*). Despite such clear daily rhythms, it has not been previously established whether parasites have intrinsic clocks that anticipate the changes in the environment or whether they are simply responding to host rhythmic physiological cues.

Sleeping sickness is a fatal disease caused by *Trypanosoma brucei*. It is transmitted by the tsetse fly and is endemic in 36 sub-saharan African countries, where 70 million people are at risk (*218*). The fact that this unicellular parasite is extracellular, easily cultured, genetically amenable and its genome has been sequenced, makes *T. brucei* an ideal system to test whether an obligatory pathogen has an intrinsic circadian clock. In eukaryotes, the current model for circadian timekeeping is based on a transcription/translation feedback loop (TTFL), in which transcriptional activators are negatively regulated by target gene repressors (*219*). However, clock genes are not phylogenetically conserved among fungi, plants and vertebrates, indicating that each phylum evolved an intrinsic clock with different components (*220, 221*).

Here we asked whether *T. brucei* has an intrinsic circadian clock. We probed the parasite transcriptome by RNA-seq, searching for transcripts oscillating with a 24 h period. For this we synchronized parasites *in vitro* for three days to alternating temperature cycles of 32ºC and 37ºC at 12 h intervals. On the 4^{th} and 5^{th} days, parasites were kept in either these alternating conditions or constantly at 37ºC. During these two days, parasite RNA was collected every four hours and subjected to RNA-seq analysis (Fig. 1A). We measured circadian gene expression after entrainment with temperature cycles in both *T. brucei* bloodstream (32ºC/37ºC cycles) and insect procyclic forms (23ºC/28ºC cycles), as well as entrainment with light in the bloodstream form (fig. S1).

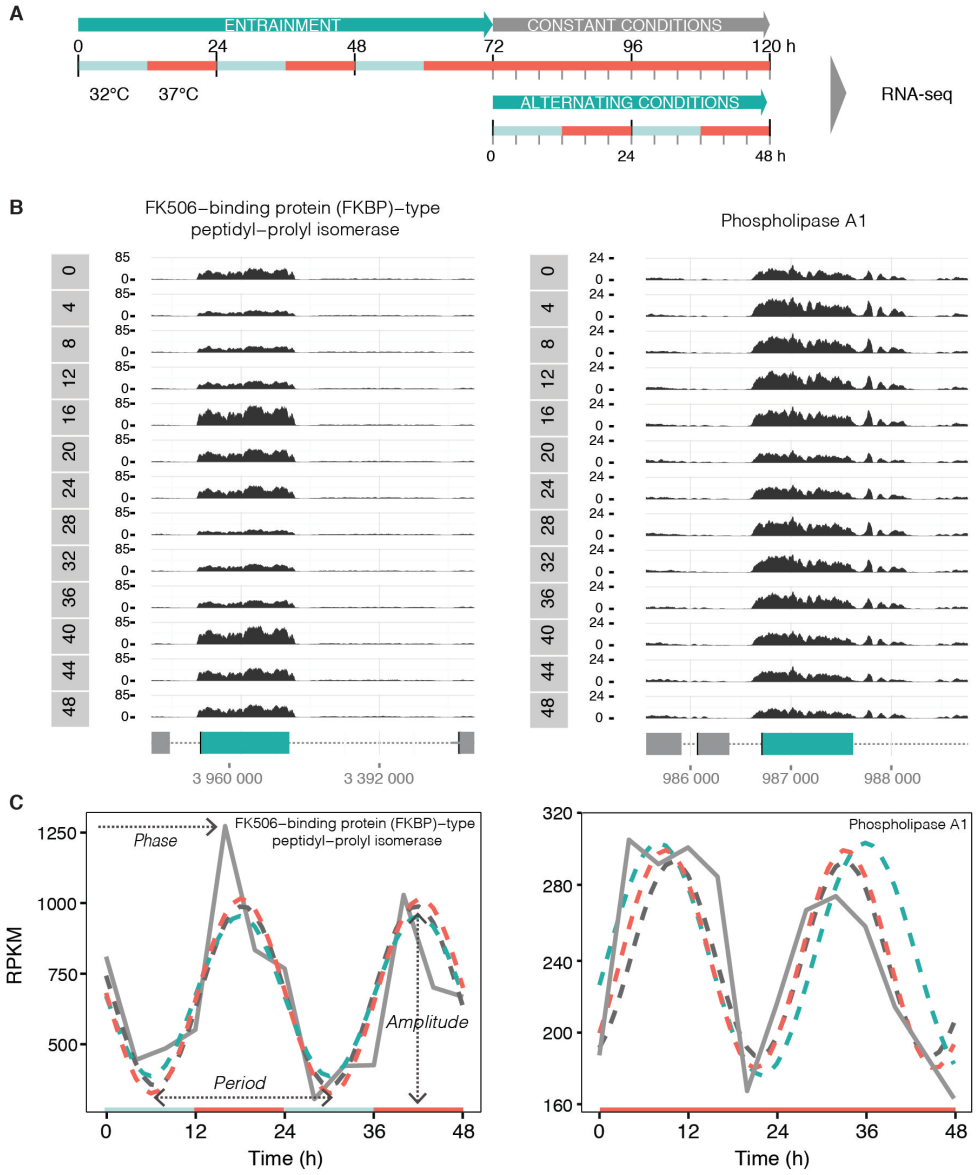


Fig. 1. Experimental design and RNA-seq analysis of circadian gene expression of *Trypanosoma brucei*. (**A**) Populations of parasites were entrained to 12 h : 12 h temperature intervals for three days, after which they were released into constant conditions or kept in alternating conditions for two days. During these two days, RNA was collected every 4 h for RNA-seq (see Materials and Methods and fig. S1). (**B**) Genome browser views of RNA-seq data from bloodstream form parasites in temperature-entrained constant conditions for two genes: Tb927.10.16100, FK506-binding protein (FKBP)-type peptidyl-prolyl isomerase, putative and Tb927.1.4830, phospholipase A1 (genes represented in teal). CDS (coding sequence) is represented as a rectangle and intergenic regions as dotted line. Reads are shown in black as reads per million total reads (RPM). (**C**) RNA-seq read coverage in RPKM and circadian algorithm fits. ARSER fit is represented in a dark gray dashed line, JTK_CYCLE in teal and Fisher's G-Test in orange.

Within each dataset, an unbiased comparison of the 13 samples using hierarchical clustering and principal component analysis showed that a cyclic pattern component accounted for 16-27% of the total variance, surprisingly being the first component in five of the six datasets. This shows that the time of sample collection is a key factor in most samples (fig. S2). To identify which genes showed a circadian expression pattern, we used three well-established algorithms to test if transcript levels cycles with a 24 h period (Fig. 1B-C). We considered a gene as cycling when its mean expression was above 10 RPKM (reads per kilobase of transcript per million mapped reads) and its circadian oscillation was detected by at least two of the three algorithms ($p \leq 0.05$, fig. S3). Two examples of cycling transcripts are shown in Fig. 1B-C.

Global analysis (Fig. 2 and data tables S1-3) revealed that hundreds of cycling transcripts were expressed in both stages of the life cycle. In bloodstream forms entrained with temperature, we found 1,490 genes (~15% of genome) oscillating in alternating conditions and 1,092 genes (~11% of genome) oscillating in constant conditions (Fig. 2A and S5). To determine the false discovery rate (FDR) for detection of cycling transcripts empirically, we performed a permutation test in which we randomized the original time of sample collection and calculated the number of cycling genes 10,000 times. We observed that the number of cycling genes identified in the correct sampling order was significantly higher than when sampling order was permutated (FDR < 0.05, fig. S4).

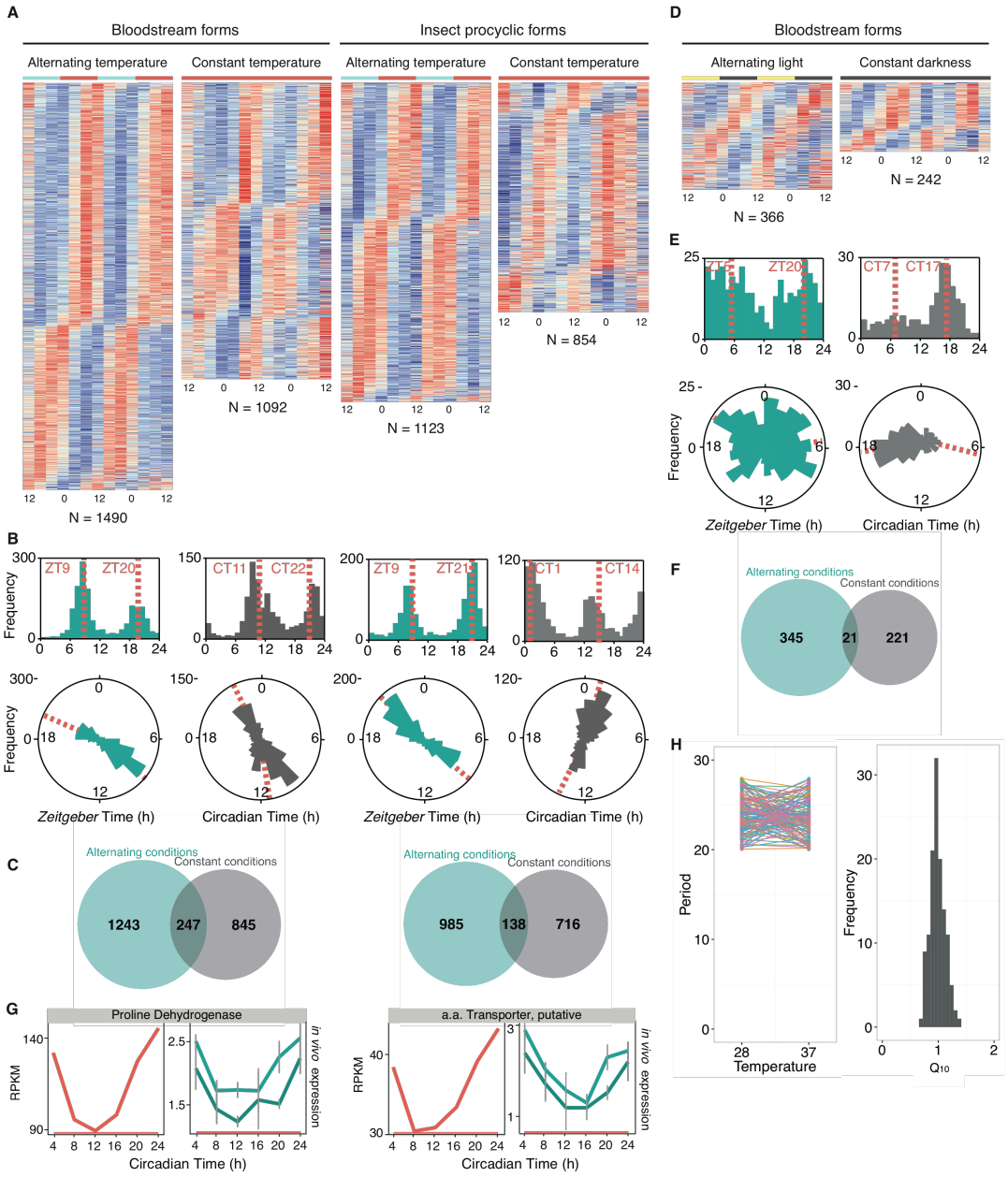


Fig. 2. *T. brucei* circadian transcriptome in bloodstream and insect procyclic forms. (**A**) Heat-map views of temperature-entrained cycling genes of bloodstream and procyclic forms. Each row represents a gene, ordered vertically by phase, determined by ARSER. (**B**) Phase distribution of cycling genes entrained by temperature. The phase of each gene rhythm across the day is represented in a histogram plot (top) and rose plot (bottom). The mean circular phase of the different phase clusters is indicated in orange. (**C**) Venn diagram of number of cycling genes identified in temperature-entrained cycling and constant conditions for bloodstream (left) and insect procyclic forms (right). (**D**) Heat-map views of light-entrained cycling genes of bloodstream for parasites. (**E**) Phase distribution of cycling genes entrained by light. (**F**) Venn diagram of number of cycling genes identified in light-entrained cycling and constant conditions for bloodstream forms. (**G**) Cycling gene expression *in vivo*. RNA was extracted from blood of infected mice. Transcript levels of proline dehydrogenase (Tb927.7.210) and putative amino acid transporter (Tb927.8.7650)

were normalized to non-cycling transcripts of zinc finger protein 3 (ZFP3, Tb927.3.720, teal) and acidic phosphatase (Tb927.5.610, dark teal). N = 18 (3 mice/time point). (**H**) Period of oscillation of 127 common cycling genes at constant temperatures of 28ºC and 37ºC. Distribution of the estimated Q_{10} for the period of the 127 common cycling genes.

As observed in other systems, the distribution of the phases of maximal expression of cycling genes is bimodal, with the majority of genes peaking at environmental Zeitgeber Time (ZT) ZT8 and ZT20 (Fig. 2B). In constant temperature, phases are shifted ~2-3 h corresponding to Circadian Time (CT) CT11 and CT22 (Fig. 2B). As previously described (*222, 223*), we found that the oscillatory transcriptome is divergent between alternating and constant conditions and three categories were identified: i) genes oscillating in alternating conditions only (stimulus-driven and clock independent) (1,243 genes, in the case of temperature-entrained bloodstream forms); ii) genes oscillating in both alternating and constant conditions (clock-driven genes) (247 genes); iii) genes oscillating in constant conditions only (genes whose cycling is suppressed or masked during entrainment conditions) (845 genes). The vast majority of the 247 genes retained their mean, amplitude, phase and period of oscillation in both conditions (Fig. 2C and S8).

When the host is bitten by a tsetse fly, *T. brucei* parasites differentiate into procyclic forms that are adapted to survive in the mid-gut of the fly where the temperature is much lower (~28ºC). Even though the transcriptome of bloodstream and insect procyclic forms is ~30% distinct (*59*), we tested whether insect procyclic trypanosomes share a circadian transcriptome. Similarly to the protocol used for bloodstream forms, we temperature-entrained cultures of procyclic forms for three days using 12 h intervals of 23ºC and 28ºC, followed by two days of constant (28ºC) or alternating temperature (fig. S1). In these trypanosomes, we identified 1,123 genes cycling in alternating conditions and 854 genes cycling endogenously (Fig. 2A-C and S6). Of these, 127 genes (~1-2% of transcriptome) oscillate in both life cycle stages, while 965 are specific to bloodstream and 727 are specific to procyclic forms (fig. S9). These data show that in two different stages of the life cycle of *T. brucei,* ~10% of the transcriptome undergoes circadian oscillations, suggesting that having a circadian rhythm conferred an evolutionary advantage throughout the parasite

life cycle. The fact that most cycling genes differ between the two stages indicates that the circadian clock can sense and adapt to the different host environments, another hallmark of circadian clocks.

In addition to temperature entrainment, a canonical property of circadian clocks is temperature compensation, which is the ability of the period of a rhythm to remain relatively constant at various physiologically permissive temperatures (*204*). 127 genes cycled in constant conditions at 28ºC and 37ºC, which allowed us to estimate the Q_{10} for these genes. Period values do not significantly change between these two temperatures and the average Q_{10} is 0.99 (Fig. 2H) showing that the circadian clock in *T. brucei* is temperature compensated.

Light is a very strong cue to which most organisms entrain. Hence we interrogated whether light could entrain bloodstream forms. We identified 366 genes cycling in the presence of the light stimulus and 242 genes cycling in constant darkness (Fig. 2D-F and S1 and S7). Although a principal component analysis identified a cyclic pattern in gene expression through the day (fig. S2C), a permutation test was not significant for these datasets (fig. S4B), which clearly shows that light is a weaker environmental cue than temperature for *T. brucei*, not surprising for a parasite that is never free-living. Nevertheless, 42 genes were entrained by both temperature and light (fig. S9).

Because the identification of the circadian transcriptome was performed in parasites in culture, next we tested whether transcript oscillations could also be found *in vivo*, i.e. in parasites from an infected mouse. For this, we collected blood every 4 h from mice kept in constant darkness. RNA was extracted and subjected to qPCR. We confirmed that transcript levels of several genes also cycled *in vivo* (Fig. 2G and S10), including proline dehydrogenase and a putative amino acid transporter. Taken together, these results show that host physiological rhythms (*in vivo*) and temperature (in culture) are capable of synchronizing *T. brucei* parasites, and that the transcriptome circadian oscillation is driven by an endogenous clock.

Even though in eukaryotes the circadian timekeeping mechanism is based on a transcription/translation feedback loop model, recent studies have shown that post-transcriptional and post-translational steps impose further levels of

circadian regulation, in addition to non-transcriptional mechanisms (*224*). *T. brucei* and other Kinetoplastida are peculiar eukaryotes as most of the genome is organized in polycistronic units (PCUs) that are constitutively transcribed and, as a result, gene expression is mainly regulated post-transcriptionally (*199, 225*). To determine whether *T. brucei* circadian gene expression was also post-transcriptionally regulated, we tested whether cycling genes clustered in specific PCUs. We found that cycling genes show a uniform distribution among most PCUs (similar proportion of cycling and non-cycling genes, Kolmogorov-Smirnov $p > 0.1$), with no bias for a specific position within a PCU (Kolmogorov-Smirnov $p > 0.5$) nor enrichment for genes peaking at a specific phase (Kolmogorov-Smirnov $p > 0.1$, Fig. 3A-D and S11-13). The fact that co-transcribed genes can either not cycle or cycle with a maximum expression at opposing phases suggests that the timekeeping mechanism used by *T. brucei* is primarily based on post-transcriptional regulation, which represents a novel mechanism of timekeeping in eukaryotes.

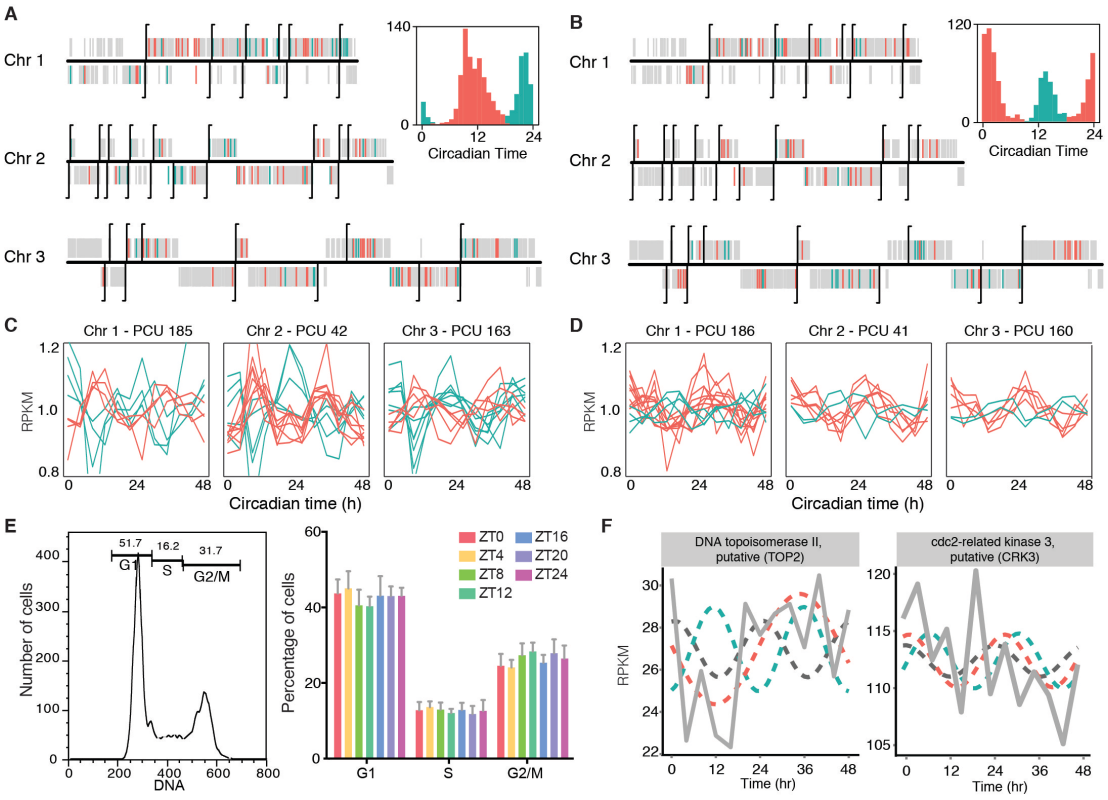


Fig. 3. ***T. brucei*** **cycling gene expression is post-transcriptionally regulated.** (**A-B**) Distribution of cycling genes genes across chromosomes 1, 2 and 3. The transcription

start site (TSS) at the beginning of each polycistronic unit (PCUs) is indicated by a vertical black line. Genes are either: gray when non-cycling; orange when cycling with maximal expression between CT3-CT18 for (A) bloodstream or CT18-CT9 for (B) insect procyclic forms; and teal when cycling with maximal expression between CT19-CT2 for (A) bloodstream or CT10-CT19 for (B) insect procyclic forms. (**C-D**) Examples of cycling genes with different phases of expression encoded in the same PCU in (C) bloodstream forms and in (D) insect procyclic forms. A representative PCU from each of the first three chromosomes is depicted. (**E**) Cell cycle profile analysis of bloodstream parasites throughout the 4th day of entrainment to temperature. Parasites were fixed and stained with propidium iodide. (**F**) Expression profile of two cell cycle associated genes (Tb927.11.11540 and Tb927.10.4990) and the algorithms fitted curves, showing no circadian oscillation. Gray line represents the raw RPKM in constant conditions.

When we compared the mean expression of cycling and non-cycling gene sets, we observed that cycling genes have higher expression (fig. S14A) (*226*). Given that the 5' and 3' untranslated regions (UTRs) are determinants of the fate of an mRNA, we interrogated the length and GC content of both 5' and 3'UTR sequences of the two major clusters of cycling genes. We found no major differences in length and GC content of cycling and non-cycling UTRs (fig S14B-C). Similar to other species, sequence elements in the 3'UTR have been identified in *T. brucei* as regulatory motifs of the mRNA levels of EP Procyclin and cytosolic phosphoglycerate kinase (PGKC) (*227, 228*). As a potential post-transcriptional regulation mechanism, we searched for common motifs in 5' or 3'UTRs of cycling genes of the two major gene clusters using six motif finding algorithms. Although this analysis identified hundreds of motifs, when we compared with the number of motifs identified in randomly selected sequences, the statistical significance enrichment dropped to 29 enriched motifs in the 5'UTR and 32 in 3'UTRs of bloodstream forms' cycling genes (4 and 28 respectively in procyclic forms) (FDR < 0.05, fig. S15A-D). Since relevant regulatory sequences are often conserved among species (*229*), we tested whether motifs identified in cycling genes were more conserved among other Kinetoplastida species than those identified from a random set of genes. Interestingly, motifs identified on bloodstream 5' and 3'UTRs, and procyclic 3'UTRs of cycling genes were more conserved in other pathologically relevant related species (*T. b. gambiense*, *T. cruzi*, *T. congolense*, *T. vivax* and *Leishmania major*) than those found in random sequences (FDR < 0.001, fig.

S15E-F). These results suggest that conserved motifs of the 5' and 3' UTR regulatory sequences may play a role in the regulation of the cyclic gene expression pattern of *T. brucei.*

To rule out that the cell cycle was responsible for the pattern of cycling transcripts, we temperature-entrained *T. brucei* bloodstream cultures and collected cells throughout the day to measure DNA content. Not surprisingly, we observed that parasite cultures in alternating conditions grew slower than in constant 37°C (7:33 h versus 6:53 h, doubling time, fig. S16B). However, the frequency of dividing cells in the population was constant throughout the day (~25%) suggesting that parasite cell division was not synchronized to occur at a certain time of the day (Fig. 3E and S16A). Furthermore, among cycling transcripts, we detected no enrichment of cell cycle associated genes (table S1 and fig. S16C), as illustrated by the expression profile of DNA topoisomerase II and cdc2-related kinase 3 (Fig. 3F). Together, these data indicate that the cyclic pattern of *T. brucei* bloodstream transcriptome is independent of the cell cycle.

To explore the biological relevance of a circadian clock in *T. brucei,* we performed a temporal Gene Ontology (GO) analysis. We assigned the clock-regulated genes into 12 groups based on the phase of maximal expression and evaluated the enrichment of GO terms (Fig. 4A and databases S5 and S6). We found that 95% of cycling GO terms are enriched in only one phase cluster in bloodstream form (93% in insect procyclic forms) ($p < 0.05$), suggesting that specific cellular processes are upregulated at different times of the day. One such process is carbohydrate metabolism, in which 13 out of 31 genes annotated to this GO term peaked expression at CT22-24 (Fig. 4A). Among those genes, ten belong to the glycolysis pathway (fig. S17). The insect-stage parasite circadian gene expression also seems to upregulate different cellular functions throughout the day. For example, vesicle-mediated transport GO term is composed of 13 genes, nine of which peak at CT0-2 (Fig. 4A).

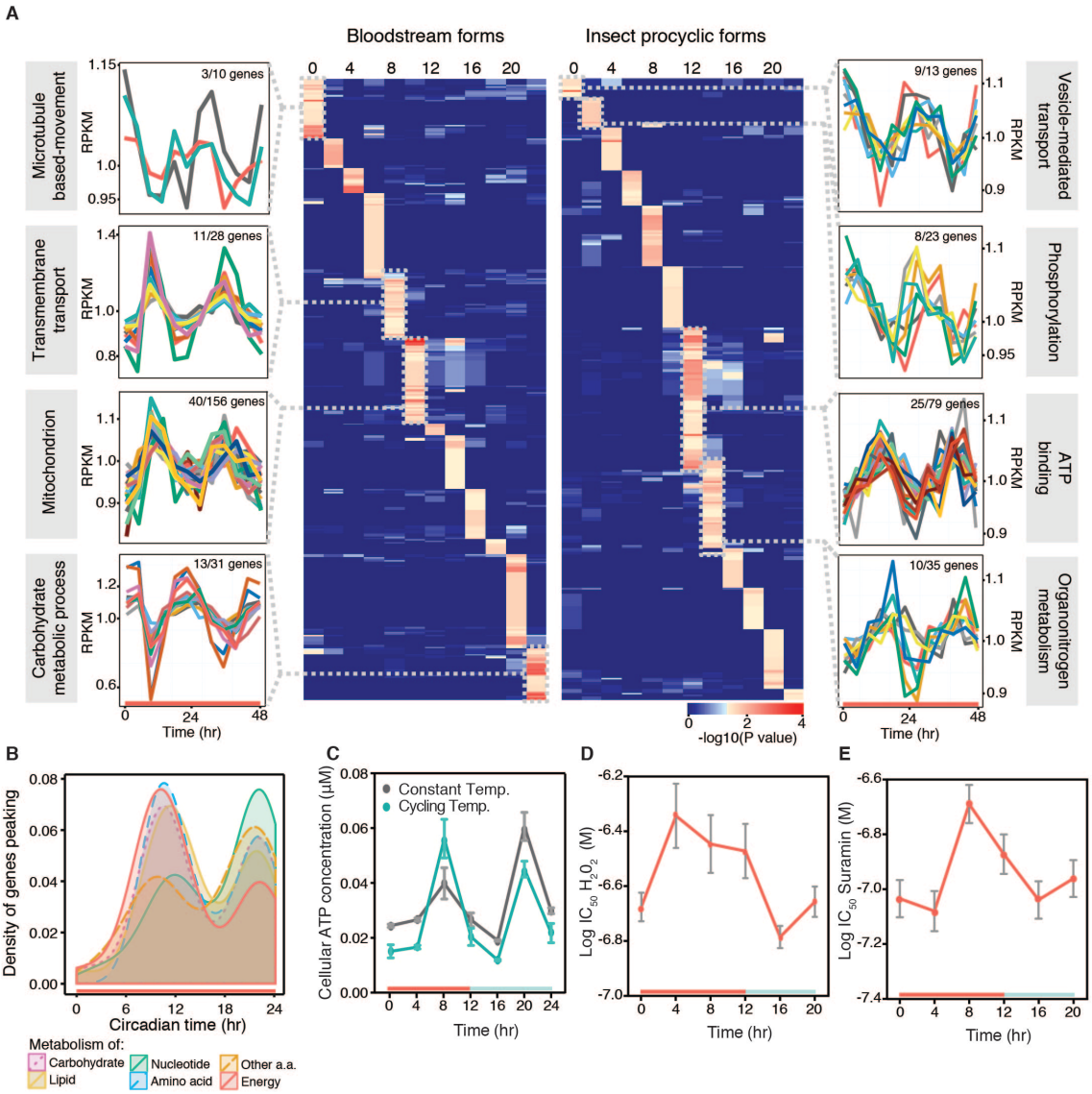


Fig. 4. The *T. brucei* circadian transcriptome comprises many metabolism related genes. (**A**) Heat-map view of GO term enrichment of both bloodstream and insect-stage parasite circadian gene expression throughout the day (2 h phase cluster, $p < 0.05$). Side plots show individual gene expression profile of the manually curated most significantly enriched GO term in the selected clusters. (**B**) Metabolic pathways are enriched in intrinsic cycling genes expressed at different times of the day. (**C**) Parasite intracellular ATP concentrations were measured on day four from cultures in alternating or constant temperature. (**D**) IC_{50} calculated from oxidative stress sensitivity experiment of parasites at different times of the day. N = 6 biological replicates. Error bars represent standard error. (**E**) IC_{50} calculated from suramin treatment of parasites at different times of the day. N = 9 biological replicates. Error bars represent standard error.

A pathway analysis (KEGG) confirmed that many cycling genes are involved in metabolism or metabolism-associated functions. Even though a

cycling gene peaks only once a day, other genes from the same metabolic pathway may not oscillate, peak at the same phase or an opposite phase (Fig. 4B). This overall transcript oscillation of metabolism-associated genes suggests that during the 24 h day the parasite population undergoes qualitative and quantitative metabolic adaptations. Because expression of most cycling genes peaks at two opposing phases of the day (Fig. 4B), it is likely that, as observed in mammals, parasites experience metabolic 'rush hours' twice a day (*138*). To test this hypothesis, we temperature-entrained *T. brucei* cultures and collected cells throughout the day to measure intracellular ATP concentrations. We found that ATP content was higher at ZT/CT8 and ZT/CT20, which coincided with the time at which metabolic genes present the highest transcript levels (Fig. 4C).

From the pathway analysis we also identified some cycling genes involved in redox metabolism (metabolism of other amino acids – glutathione pathway, Fig. 4B and S18). To test if such gene expression oscillations lead to different levels of sensitivity to oxidative stress throughout the day, we temperature-entrained bloodstream forms and, starting every 4 h through the day we incubated parasites with H_2O_2 for 1 h and measured their viability. We observed a time-dependent sensitivity to H_2O_2 treatment, which was 2.5-fold higher at ZT4 than at ZT16 ($p < 0.001$, Fig. 4D and fig. S18). These results confirm that circadian metabolic gene expression patterns have functional consequences to the overall metabolic and redox state of the parasite.

Because parasites display a different transcriptome at different times of the day, we wondered whether this would affect the sensitivity of the parasite to suramin, a drug commonly used in the field. Similar to the oxidative stress experiment, we temperature-entrained parasites and beginning every 4 h through the day we tested cell viability upon a 24 h *in vitro* treatment with suramin. We observed that the parasites are more resistant to suramin treatment beginning at ZT8, which is reflected in a higher IC_{50} ($p < 0.001$), as 2.5-fold higher drug concentration is needed to kill these parasites (Fig. 4E and S19). Thus, we conclude that during the day parasites are not equally sensitive to suramin treatment.

Various studies in pathogens (*149-153*) and microbiota (*154, 155*) have described daily rhythms in microorganisms living inside hosts. However in these

experiments it was not established whether such behavior was endogenously controlled by the pathogen, or whether rhythms were imposed by the host. Indeed in recent microbiome experiments, the 24-hour variations in the microbiome appear to be driven by the host and by its feeding regime (*154, 155*). In contrast, our results show that T. brucei has an intrinsic mechanism to count time, which results in a circadian oscillating transcriptome. This clock imposes cyclic changes in the parasite population with a 24 h period, primarily at the metabolic level. As a result, upon environment entrainment in vitro, parasites in the morning are different from those in the evening, which is likely an important adaptation since in vivo their host environment (mammal or insect) also undergoes circadian changes (*230, 231*). This study demonstrates the potential of high-throughput approaches for identifying circadian patterns in the transcriptome of non-model organisms and it provides a foundation for the search of the master regulators of this process in T. brucei and for the search for endogenous clocks in other important infectious agents such as the malaria parasite.

3.3. Experimental Procedures

Ethics Statement

All animal care and experimental procedures were performed in accordance with University of Texas Southwestern Medical Center (UTSW) IACUC guidelines, approved by the Ethical Review Committee at the University of Southwestern Medical Center and performed under the IACUC-2012-0012 protocol.

Parasites and Culture Conditions

T. brucei AnTat 1.1^{E}, a pleomorphic clone, derived from an EATRO1125 clone was originally isolated from blood of *Tragelaphus scriptus* in Uganda. For all the experiments, we used AnTat 1.1^{E} 90-13, a transgenic cell-line encoding the tetracyclin repressor and T7 RNA polymerase (*190*).

Bloodstream forms were grown routinely in HMI-11 at 37°C in 5% CO_2 (*232*). For all RNA-seq experiments individual cultures of parasites were prepared, adjusting the initial parasite density so that parasite cultures would be at 10^6 parasites/mL at each collection time point. Parasite numbers were calculated using a Hemocytometer. Samples were collected every 4 h throughout two consecutive days. For the bloodstream forms - Temperature RNA-seq experiment, culture flasks were moved every 12 h between incubators either at 32°C or 37°C or remained at constant 37°C. In the bloodstream forms - Light/Dark entrainment RNA-seq experiment, a warm white LED 3W lamp was used to illuminate the cultures inside the incubator. Temperature was kept at 37°C and fluctuations were monitored and shown to be less than 0.1°C.

Differentiation of bloodstream forms to procyclic forms was induced by adding 6 mM cis-acconitate to DTM medium and by reducing temperature to 28°C. The newly differentiated procyclic cultures were maintained as described previously (*233*). Differentiation was assessed by EP Procyclin expression, using anti-*Trypanosoma brucei* Procyclin, FITC mouse IgG1 (Cedarlane Labs). For the procyclic form RNA-seq experiment, culture flasks were moved every 12 h between incubators either at 23°C or 28°C or remained at constant 28°C.

Transcriptome Sequencing (RNA-seq)

RNA was isolated from 10^7 *Trypanosoma brucei* cells with TRIzol reagent according to the manufacturer's instructions (Life Technologies). 1 µg of total RNA was enriched for mRNA using Poly-A beads for RNA-seq according to the manufacturer's instructions (Invitrogen). The removal of ribosomal RNAs was confirmed on a Bioanalyzer Nano Chip (Agilent Technologies). Sequencing libraries were constructed using TruSeq RNA Sample preparation protocol (Illumina). RNA-sequencing of libraries was performed in HiSeq2000 platform (Illumina) with 50-bp reads according to manufacturer's instructions by the UTSW McDermott Next Generation Sequencing Core and Beijing Genomics Institute (BGI). Read quality was assessed using the FASTQC quality control tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The SolexaQA suite of programs (*234*) was used to trim raw reads to their longest contiguous segment above a PHRED quality threshold of 28, and reads smaller than 25

nucleotides long were discarded. Reads were mapped to the *T. brucei* TREU927 reference genome using bowtie (v1.0.0) (*235*) allowing for 2 mismatches and only non-ambiguous alignments (options –v 2 –m 1). The number of reads mapping to each gene was determined, and then normalized to RPKM (excluding a highly expressed VSG gene from the calculation) using the R software environment and the packages *GenomicAlignments* (*236*), *Biostrings* (*237*) and *rtracklayer* (*238*) from Bioconductor.

Time Series Analysis for Circadian Cycling

In figure S2, hierarchical clustering analysis was performed and heatmaps of Spearman correlations from centered log2 transformed RPKM values were done in the R software environment using the function heatmap.2 from the *gplots* package (*239*). Principal component analysis (PCA) was done on centered and log2 transformed RPKM values using the function *princomp*.

RNA cycling was assessed by three programs: GeneCycle (*240*), that implements Fisher's G-Test, JTK_CYCLE (*241*) and ARSER (*242*). For Fisher's G-Test and JTK_CYCLE analyses, RPKM data was detrended by linear regression. A gene was considered cycling if two out of three programs detected periodic expression with threshold of $p \leq 0.05$ and mean expression higher than 10 RPKM. This cutoff was defined by assessing the coefficient of variation in relation to the mean expression across all time points. The amplitude, period and phase reported by ARSER were used for further analysis. The heatmaps of phases in figure 2A and 2D plot the z-score transformed RPKM values, ordered by the phase determined by ARSER. The peaks of expression phase distributions were determined by fitting a mixed von Mises-Fisher model to the bimodal phase distributions using the R software environment and the *movMF* package (*243*), extracting the means of the two von Mises-Fisher distributions.

To determine the false discovery rate (FDR) of identification of cycling genes the time point of collection were randomized and number of cycling genes assessed. These permutation tests were run 10,000 times.

Chromosome Distribution of Cycling Genes

Uniform distribution of cycling genes among PCUs was tested by comparing the distribution of proportions of cycling/non-cycling genes per PCU with a distribution obtained by randomly sampling 10,000 times the same number of genes from the genome. To test if cycling genes displayed a bias in their positioning within PCUs, the distance of each gene to its nearest upstream transcription start site (TSS) was calculated and then the distributions of these distances for cycling genes only and for all genes were compared. In order to test if cycling genes within a PCU tend to peak at the same phase, the distribution of proportions of each phase cluster within PCUs with a random distribution obtained by randomly permuting the cluster assignment of cycling genes 10,000 times were compared. For these three analyses significance was assessed by a Kolmogorov-Smirnov test.

5' and 3'UTR length, GC content and Motif identification

Cycling genes were clustered according to their phase of oscillation. 4 hour clusters were centered at the peaks of the bimodal phase distribution and expanded if the consecutive cluster had at least 100 genes. 5' and 3'UTR sequences were obtained based on the *T. brucei* TREU 927 annotation obtained from TriTrypDB (v8.0). Those sequences were used to compare the 5' and 3'UTR sequence length and GC content of genes assigned to each cluster with the set of non–cycling genes. Significance was assessed with Wilcoxon rank sum test.

For motif identification, annotated UTR sequences less than 25 nucleotides (nts) long were discarded, and sequences longer than 1000 nts were trimmed to the first 1000 nts. Six motif discovery programs (Weeder2 (*244*), Trawler (*245*), DREME (*246*), MEME (*247*), FIRE (*248*) and XXmotif (*249*)) were used to identify motifs enriched in the 5' and 3'UTR region of cycling genes from each phase cluster. Because some programs elicit many motifs that are highly similar to each other, the motifs elicited by all programs were compared to each other (by computing the Euclidean distance between each pair of motifs) and clustered using hierarchical clustering with a distance threshold of 0.25. For further analysis, a representative motif was chosen from each cluster by

selecting the motif with the highest total information content. These representative motifs were mapped back to the 3'UTR sequences of all genes using FIMO (*250*) with a threshold p value of 1×10^{-4}. Motif enrichment in each cluster group was assessed using Fisher's Exact Test. The number of motifs identified in the cycling genes was compared with the number of motifs identified from 25 experiments in which we randomly selected the genes from the entire gene set.

To determine the preferred location of occurrence of enriched motifs, FIMO was used to map the motifs to the flanking 1000 nts upstream and downstream of all annotated coding sequences (CDSs), and for each motif, the number of matches inside and outside the UTR region was quantified. A motif was considered specific to UTR sequences if it occurred more frequently inside the UTR region than expected based on the annotated UTR lengths (binomial test; $p < 0.01$).

To evaluate the conservation of enriched motifs, the genome references of 5 species closely related to *T. brucei* (*Leishmania major*, *T. b. gambiense*, *T. cruzi*, *T. congolense*, *T. vivax*) were obtained from TriTrypDB, and representative UTR sequences of orthologous genes were extracted (100 nts and 400 nts for 5' and 3' UTRs respectively). FIMO was used to match the motifs to these UTR sequences. For each *T. brucei* UTR containing a given motif, the presence or absence of that motif in the UTRs of the annotated orthologs of that gene was determined, and the motif was considered "conserved" in that UTR if it also appears in at least one ortholog UTR sequence. The hypothesis that motifs identified from circadian genes are more conserved than motifs identified from random groups of genes was tested with Wilcoxon Rank Sum test.

Mice Infection and Real-Time Quantitative PCR Analysis

The infections of wild-type male C57BL/6J mice, 6-10 week old (UT Southwestern Medical Center Mouse Breeding Core Facility) described in this manuscript were performed by intraperitoneal (i.p.) injection of 2,000 *T. brucei* AnTat 1.1^{E} parasites. Prior to infection, *T. brucei* cryostabilates were thawed and parasite viability and numbers were assessed by mobility under a

microscope. Mice were individually housed in activity wheel-equipped cages under LD 12:12 for 7 days after which animals were kept in dark conditions. Chow and water were available *ad libitum*. Locomotor activity was recorded and analyzed using ClockLab software (Actimetrics, Wilmette, IL) to determine the circadian phase, as previously described (*203*), for each animal on day 20 post-infection. Blood was sampled by cardiac puncture and RNA extracted with TRIzol LS according to the manufacturer's instructions (Life Technologies). Reverse transcription and real-time PCR were performed as described previously (*216*). Primer efficiencies were determined using standard curves with 3-$\log_{10}$ coverage. Transcript levels were normalized to genes whose expression remained constant in both temperature and light alternating conditions. Primer sequences are listed in table S2.

Assessment of Metabolic Activity and Cell Cycle Stage

Bloodstream form parasites were cultured and entrained by temperature as described above for the RNA-seq experiment. Samples were collected every 4 h throughout the day.

For metabolic activity assessment, parasitemia was assessed and metabolic activity was measured according to the manufacturer's instructions of CellTiter-Glo® Luminescent Cell Viability Assay (Promega) from 1 x 10^5 parasites centrifuged and resuspended in 25µl trypanosome dilution buffer (TDB, 5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose, pH 7.7).

For cell cycle analysis, 2 x 10^6 parasites were fixed by slowly adding ethanol to a final concentration of 70%. Fixed trypanosomes were pelleted and stained with 0.5 mL in PBS / 2 mM EDTA containing 10 µg RNAse A and 1 mg propidium iodide for 30 min at 37°C. Intensity of red fluorescence was measured using a FACSCalibur flow cytometer (BD Biosciences) and data were analyzed using FlowJo.

Functional analysis of cycling genes

Cycling genes identified by RNA-seq were clustered in 12 groups (CT0-CT2, CT2-CT4, etc.) based on their expression peak. *T. brucei* GO term annotations were obtained from TriTrypDB. GO term enrichment was assessed in each group by Hypergeometric test using GOstats (*251*) and by Fisher's Exact Test and plotted as heatmap in figure 4A. Manually curated GO term was defined when both statistical tests show enrichment ($p < 0.05$) and more than 3 genes annotated to a GO term were cycling with the determined phase (see table S5 and S6).

Suramin and H2O2 sensitivity assay

Bloodstream form parasites were cultured and entrained by temperature as described above for the RNA-seq experiment. Every 4 h, parasites were harvested from exponential phase cultures, counted and plated in 96 well flat-bottom microtiter plates at a cell density of 10,000–20,000 cells/well. Serial dilution concentrations (1:3) of suramin (Sigma) were added. The compound was applied in triplicate at eight concentrations and incubated for 24 h at 37°C. For the oxidative stress experiment, H_2O_2 (Sigma) at six different dilutions were tested (1:10) and parasites were incubated for one hour at 37°C. Alamar Blue (Sigma) was used to determine cell viability by adding at ten percent of the well volume followed by 4h incubation at 37°C. Fluorescence was measured with 530ex/590em nm and percentage of live cells calculated by normalizing to non treated parasites. Calculation of IC_{50} values was done by 4-parameter nonlinear curve fit (GraphPad Prism). H_2O_2 sensitivity was tested in two independent experiments from a total of six biological replicates. Suramin resistance was tested in three independent experiments and IC_{50} values are shown as the mean of those experiments (N = 9).

3.4. Supplementary figures

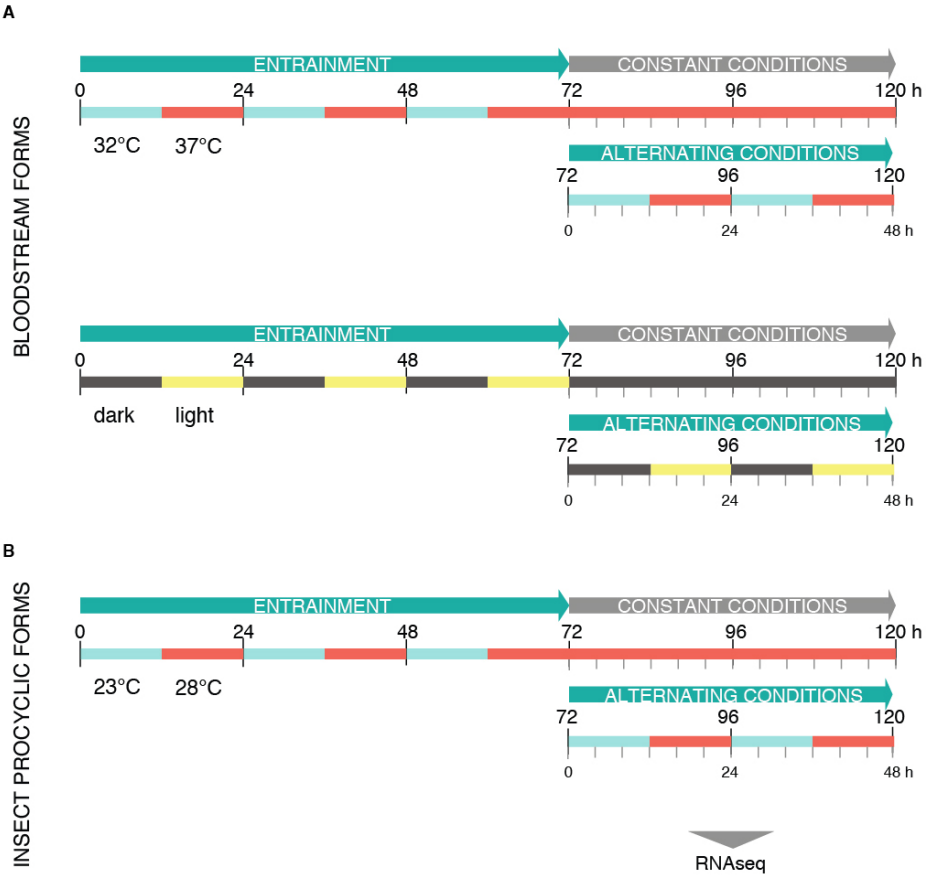


Fig. S1. Experimental design for RNA-seq experiments. Parasites were cultured in individual flasks and their density was calculated to reach 1 x 10^6 at each collection time point. (A) Bloodstream parasites were synchronized either with temperature or light. (B) Insect procyclic parasites were synchronized with temperature only. Synchronizations were done for three days in alternating temperature or light conditions. At the end of day three (72 h), culture flasks were split into alternating or constant conditions groups. RNA samples were collected for two days (a total of 13 samples per condition, with the second cycle acting as biological replicate).

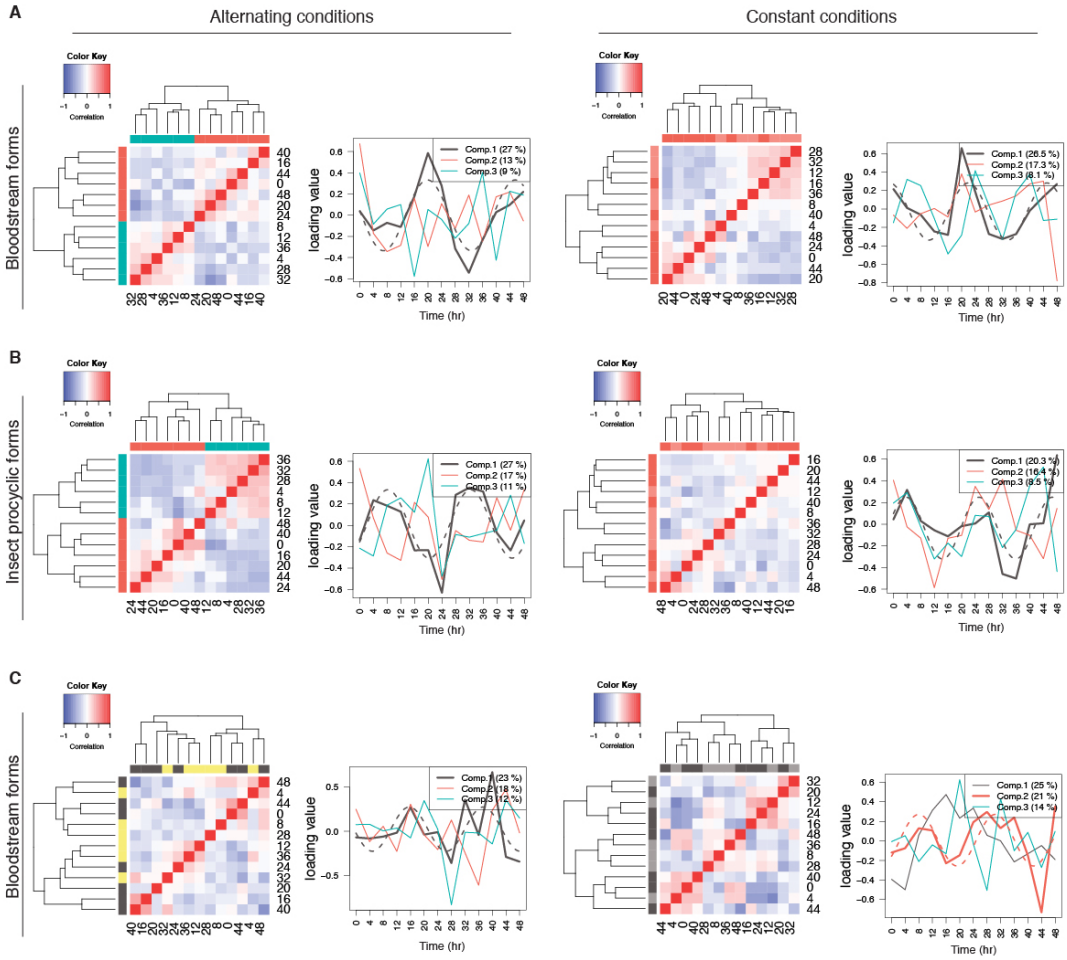


Fig. S2.Unbiased assessment of cycling gene expression patterns. Clustered heatmap of sample correlations and PCA from (A) bloodstream temperature-entrained, (B) insect procyclic temperature-entrained and (C) bloodstream light-entrained datasets. Cyclic component of the PCA is identified as a solid line, in the same color as the dashed line of its fitted cosine curve.

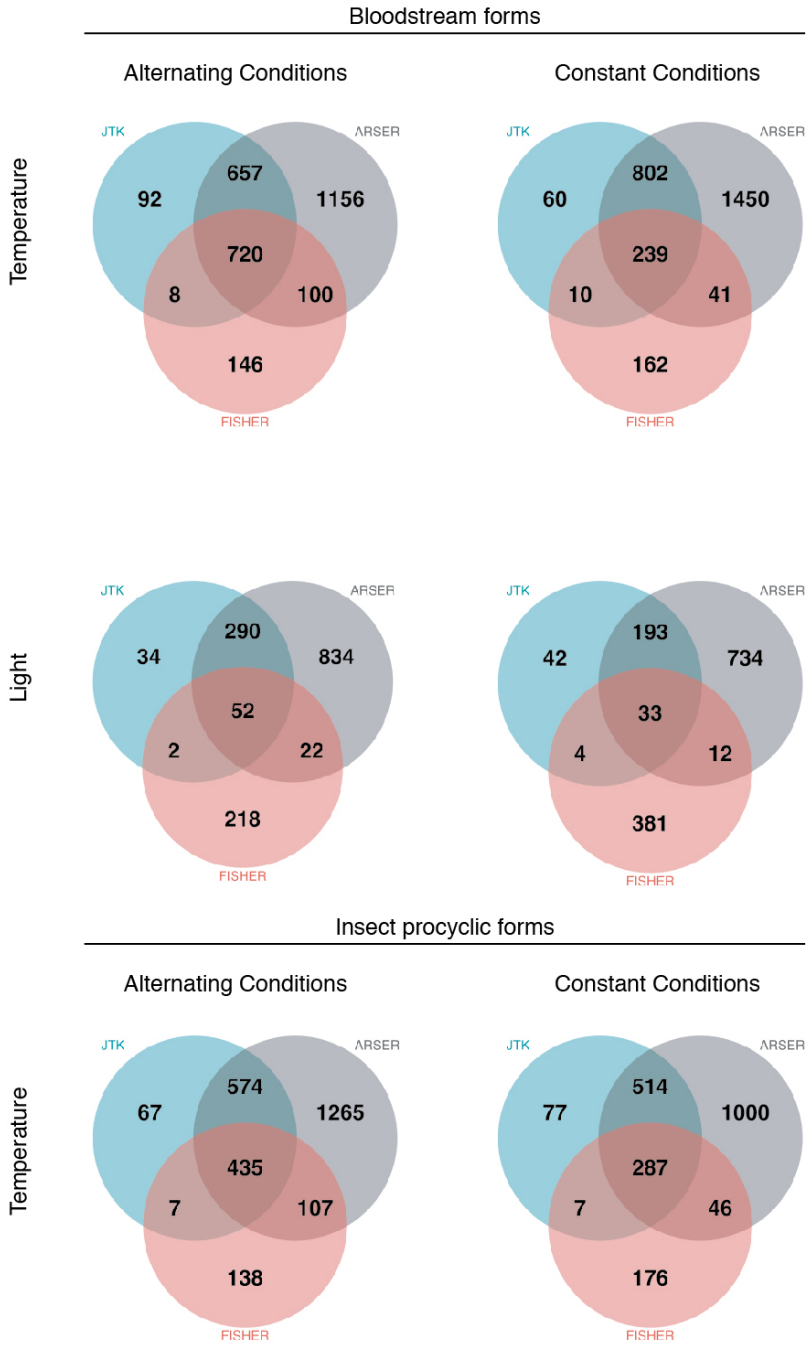


Fig. S3. Venn diagram of the number of cycling genes identified by each algorithm (ARSER, JTK_CYCLE and Fisher's G-Test, $p < 0.05$).

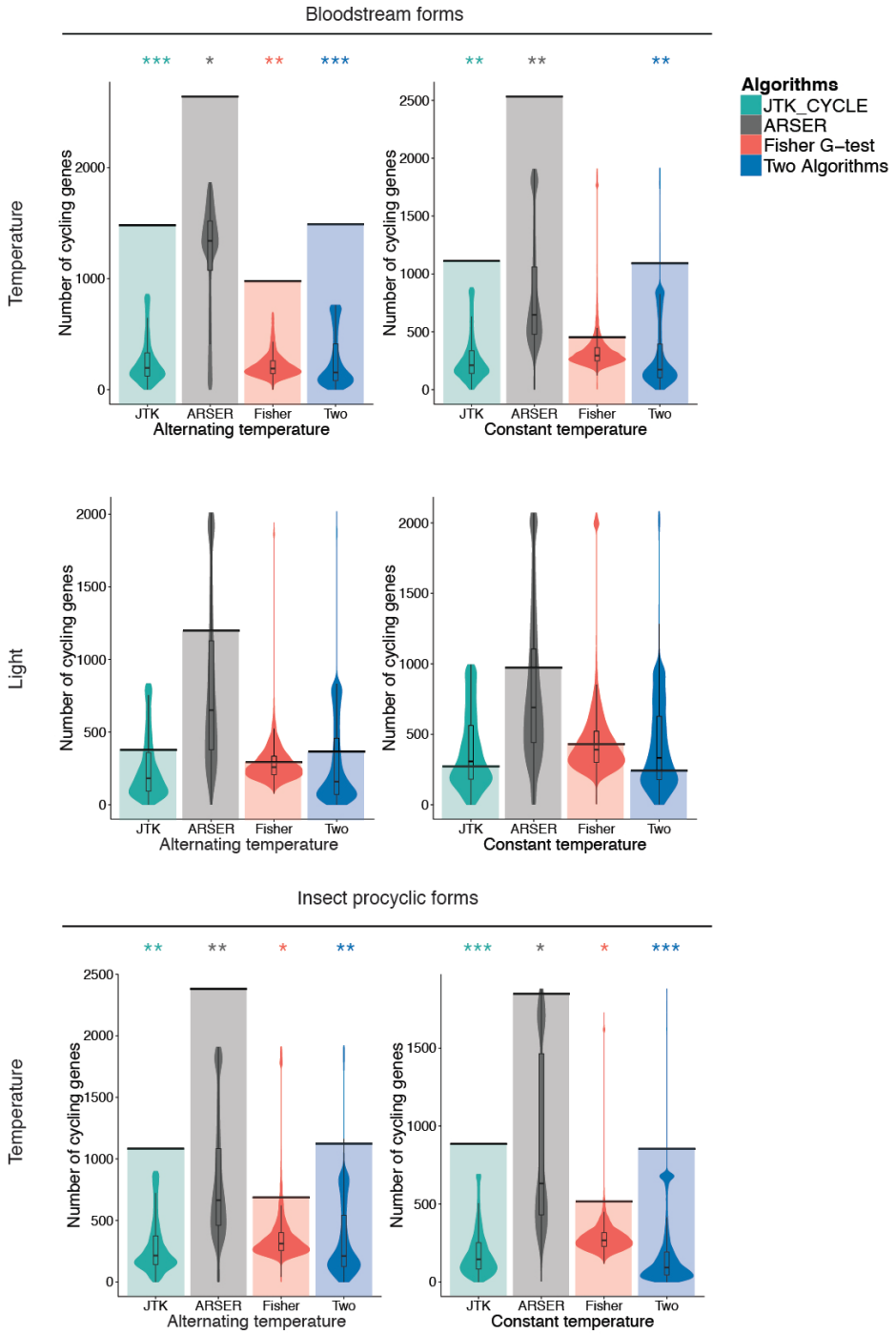


Fig. S4. Number of cycling genes identified by each algorithm or in two algorithms in the correct sampling order (colored bars) and the distribution of the number of cycling genes identified by permutating 10,000 times the time point order of sample collection (violin plots). Median and quartiles are represented inside the violin plots in black and * represents statically significance between the number of cycling genes in the real order and permutated order at FDR < 0.05, ** represents FDR < 0.01 and *** represents FDR < 0.001.

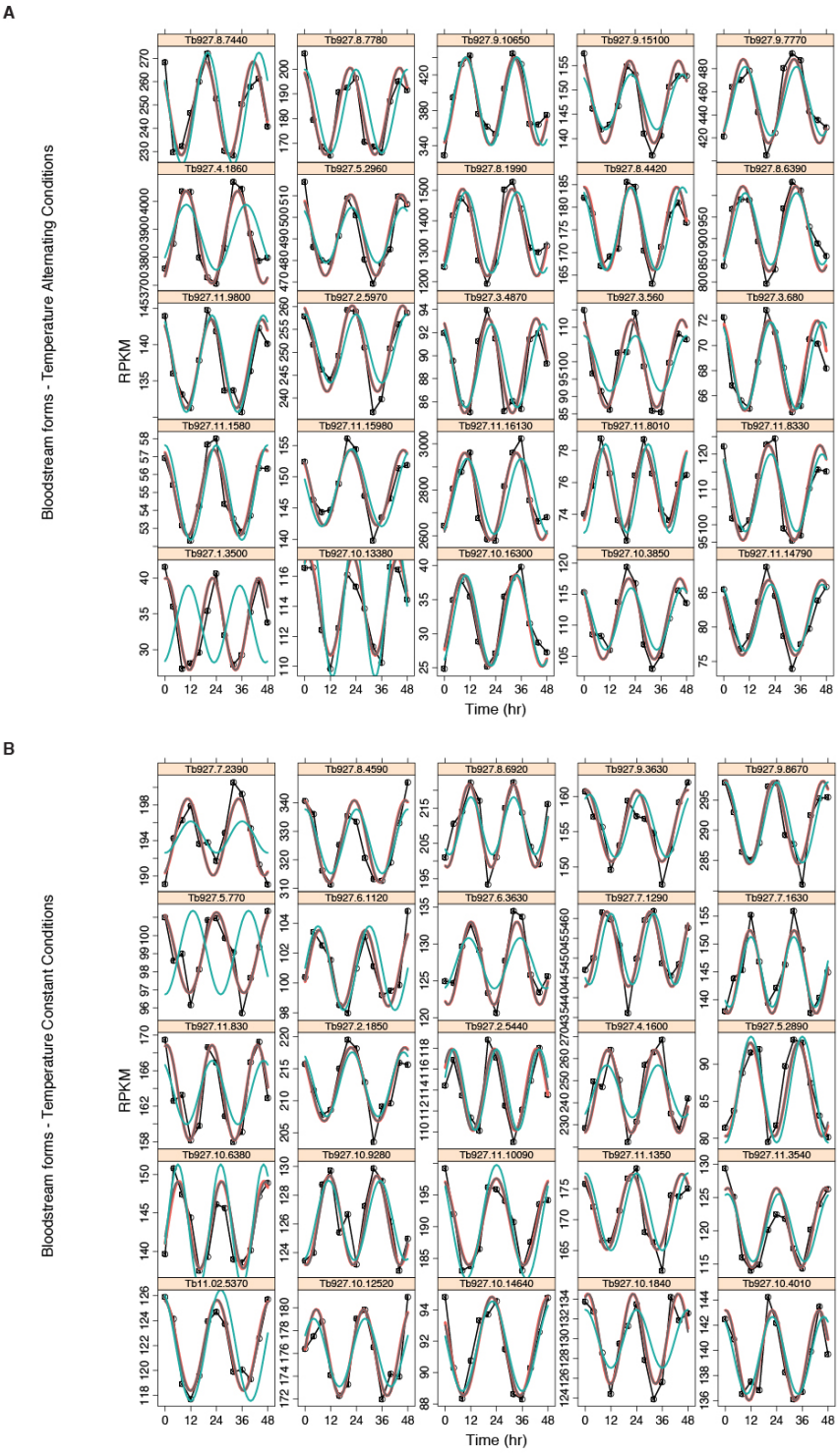


Fig. S5. Top 25 most significant cycling genes of temperature-entrained bloodstream forms in alternating and constant conditions. Order defined by ARSER p value. Black line represents the raw RPKM, teal represents JTK_CYCLE fit, gray represents ARSER fit and orange represents Fisher's G-Test fit.

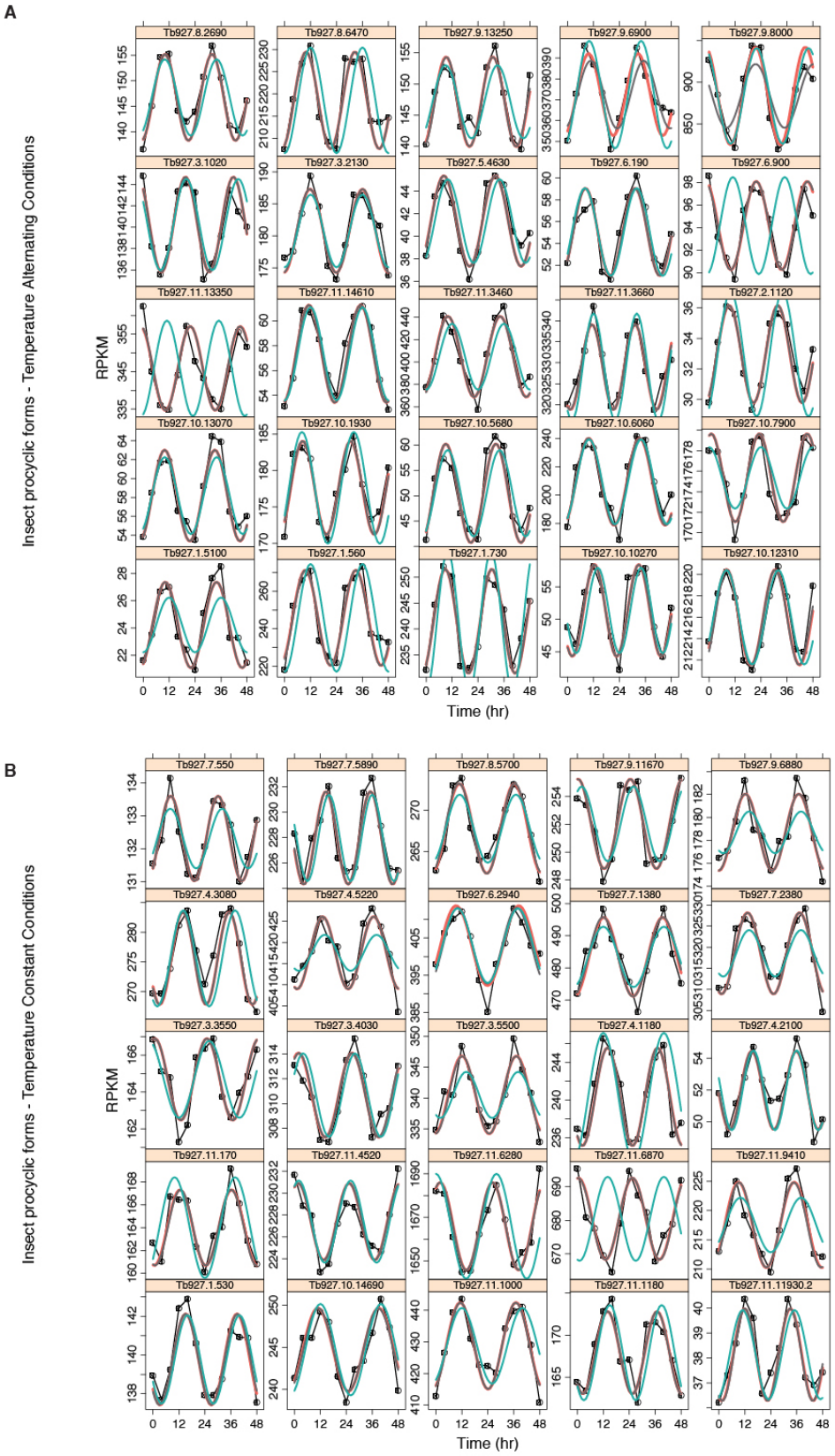


Fig. S6. Top 25 most significant cycling genes of temperature-entrained insect procyclic forms in alternating and constant conditions. Order defined by ARSER p value. Black line represents the raw RPKM, teal represents JTK_CYCLE fit, gray represents ARSER fit and orange represents Fisher's G-Test fit.

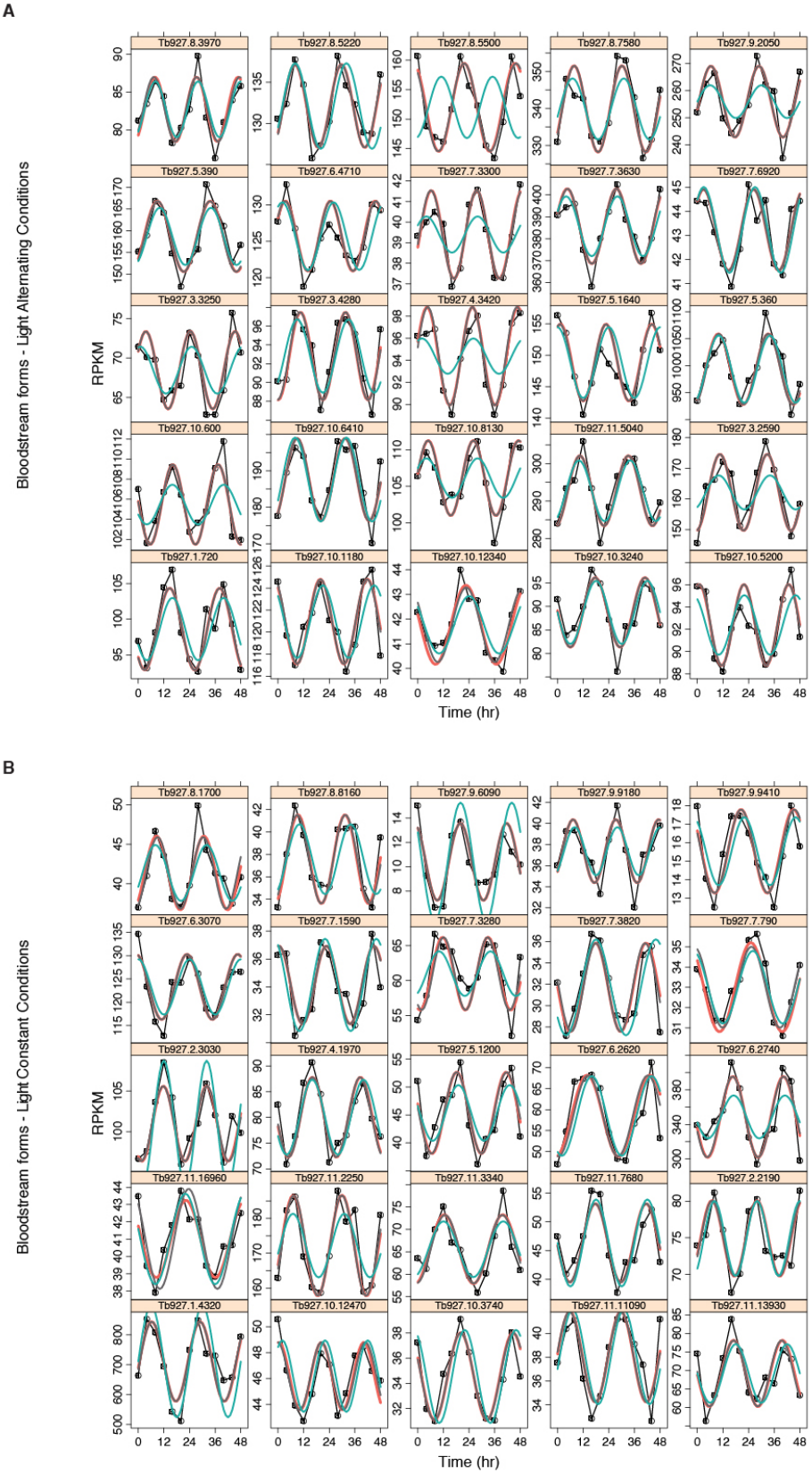


Fig. S7. Top 25 most significant cycling genes of light-entrained bloodstream forms in alternating and constant conditions. Order defined by ARSER p value. Black line represents the raw RPKM, teal represents JTK_CYCLE fit, gray represents ARSER fit and orange represents Fisher's G-Test fit.

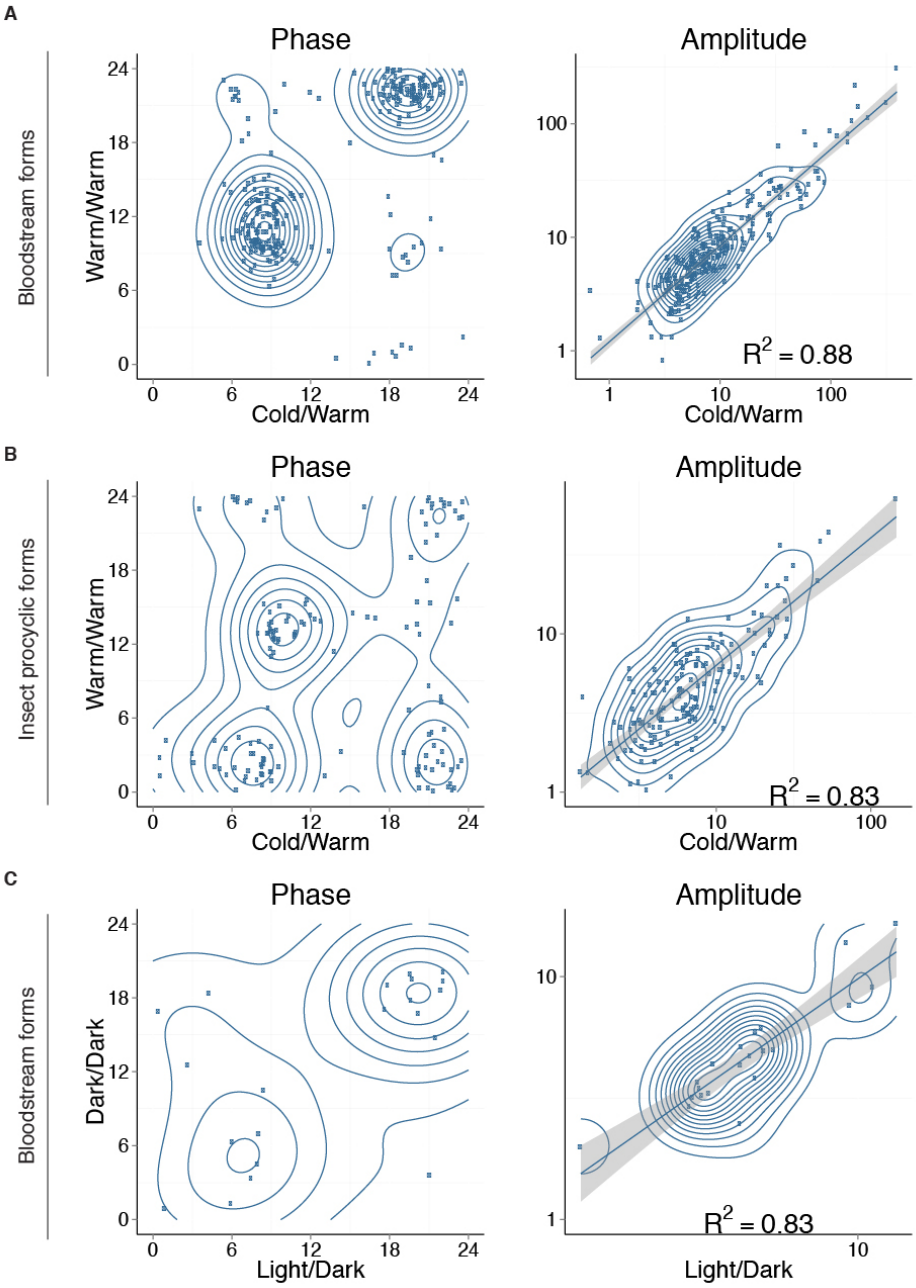


Fig. S8. The correlation plots of circadian parameters phase and amplitude between constant and alternating conditions show that most cycling genes maintain their circadian characteristics in both conditions. (A) bloodstream temperature-entrained, (B) insect-stage temperature-entrained and (C) bloodstream light-entrained datasets.

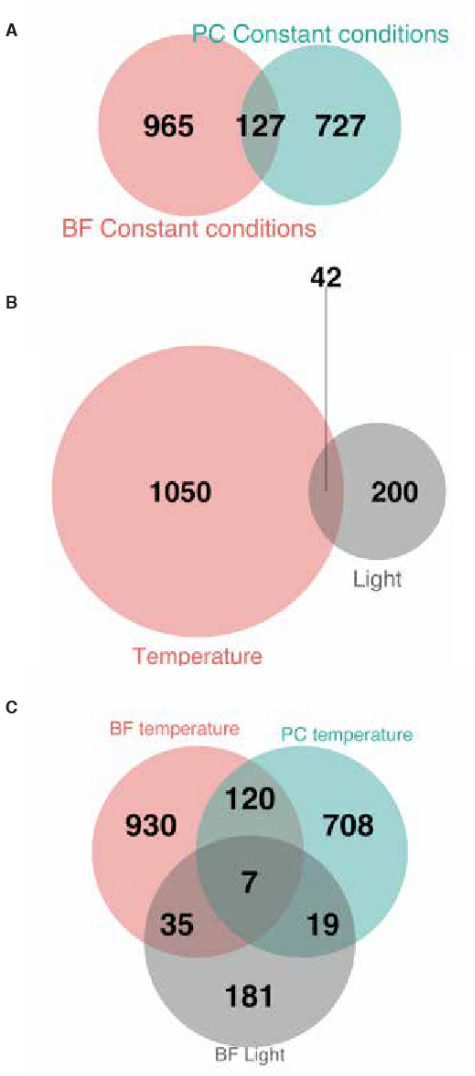


Fig. S9. Venn diagram of the intersection of endogenous cycling genes from different RNA-seq datasets. (A) Samples collected in constant conditions after entrainment to temperature in bloodstream (rose) or insect procyclic-forms (teal). (B) Samples collected in constant conditions after entrainment to temperature or light (gray) in bloodstream forms. (C) Samples collected in constant conditions after entrainment to temperature (bloodstream and procyclic forms) or light (bloodstream forms).

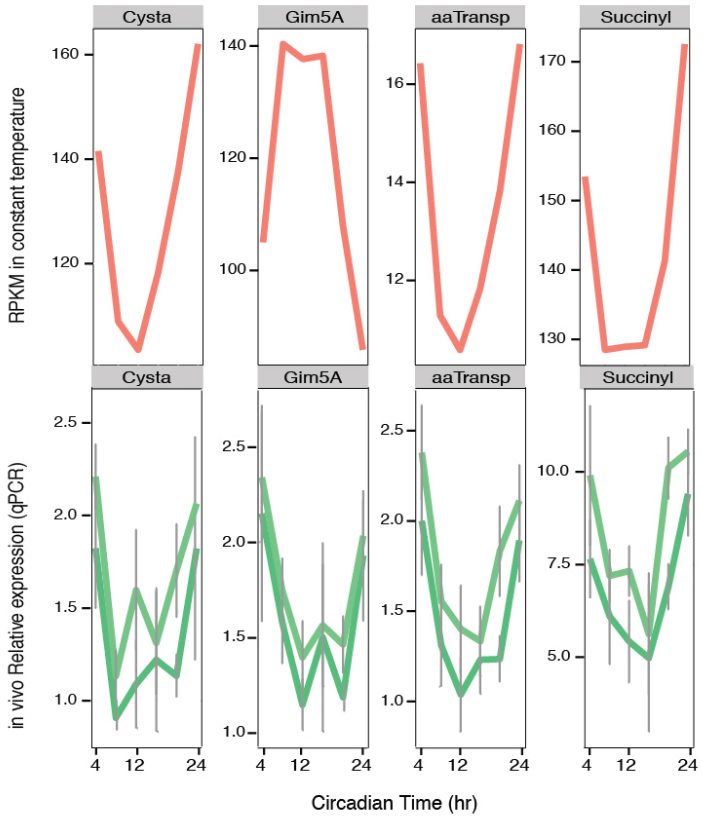


Fig. S10. Transcript levels oscillate in parasites isolated from blood of infected mice. Upper panel represents endogenous cycling gene expression from culture of temperature-entrained bloodstream form parasites, detected by RNA-seq. Bottom panel represents the *in vivo* expression profile of the same genes. Represented genes are Tb927.9.12320 (cystathionine gamma lyase, putative), Tb927.9.11580 (glycosomal membrane protein, gim5a), Tb927.5.900 (oligosaccharyl transferase subunit, putative), Tb927.1.4830 (phospholipase A1), Tb927.8.7740 (aminoacid transporter, putative) and Tb927.11.2690 (succinyl-coA:3-ketoacid-coenzyme A transferase, putative). Cycling genes expression levels were normalized to non-cycling genes zinc finger protein 3 (ZFP3, Tb927.3.720, green) or to a putative acidic phosphatase (Tb927.5.610, dark green). N = 18 (3 mice/time point). Gim5A is the example of a transcript for which *in vitro* RNA-Seq and *in vivo* qPCR have opposite phases.

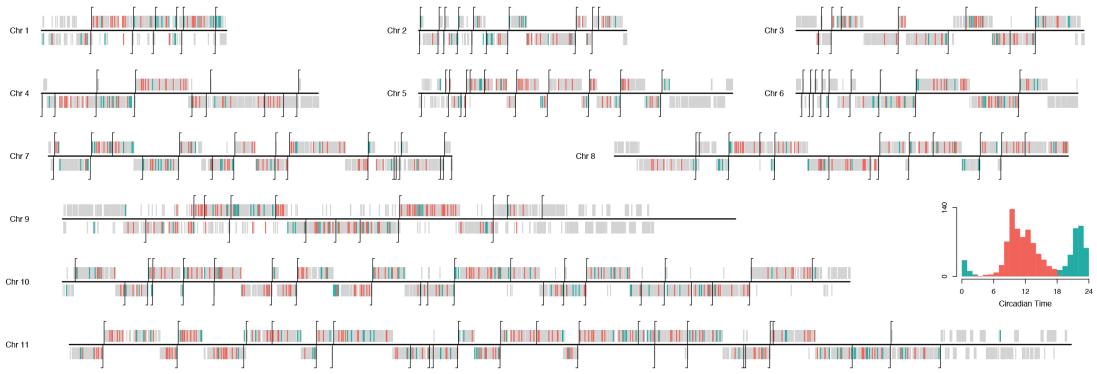


Fig. S11. Bloodstream form clock driven genes are distributed randomly across the genome (continuation of figure 3A)**.** Individual chromosomes are represented with polycistronic units (PCUs) marked by its Transcription start site (TSS) in black line. Genes are either gray when non-cycling, orange when cycling with maximal expression between CT3-CT18 and teal when cycling with maximal expression between CT19-CT2.

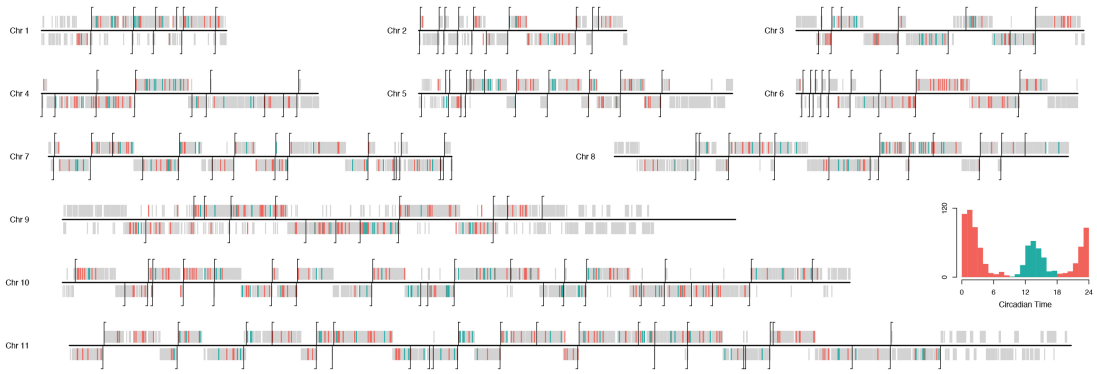


Fig. S12. Insect procyclic clock driven genes are distributed randomly across the genome (continuation of figure 3B)**.** Individual chromosomes are represented with Polycistronic units (PCUs) marked by its Transcription start site (TSS) in black line. Genes are either gray when non-cycling, teal when cycling with maximal expression between CT10-CT19 and orange when cycling with maximal expression between CT18-CT9.

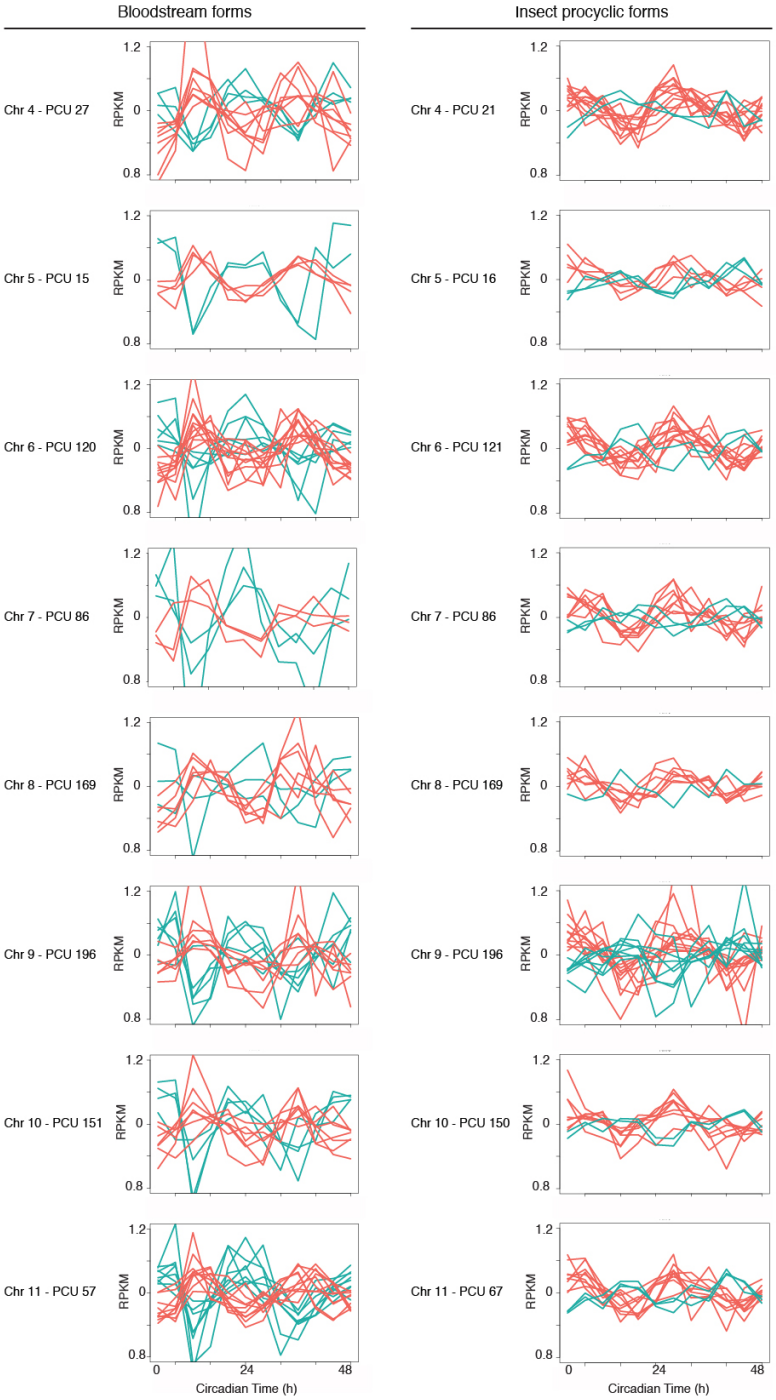


Fig. S13. Gene expression profile of genes belonging to the same polycistronic unit. Representative polycistronic units (PCUs) from each chromosome are shown (continuation of figure 3C and 3D). For the bloodstream forms, teal represents RPKM levels of cycling genes with maximal expression between CT19-CT2 and orange when cycling with maximal expression between CT3-CT18. For the insect procyclic forms, teal represents genes with maximal expression levels between CT10-CT19 and orange for genes with maximal expression between CT18-CT9.

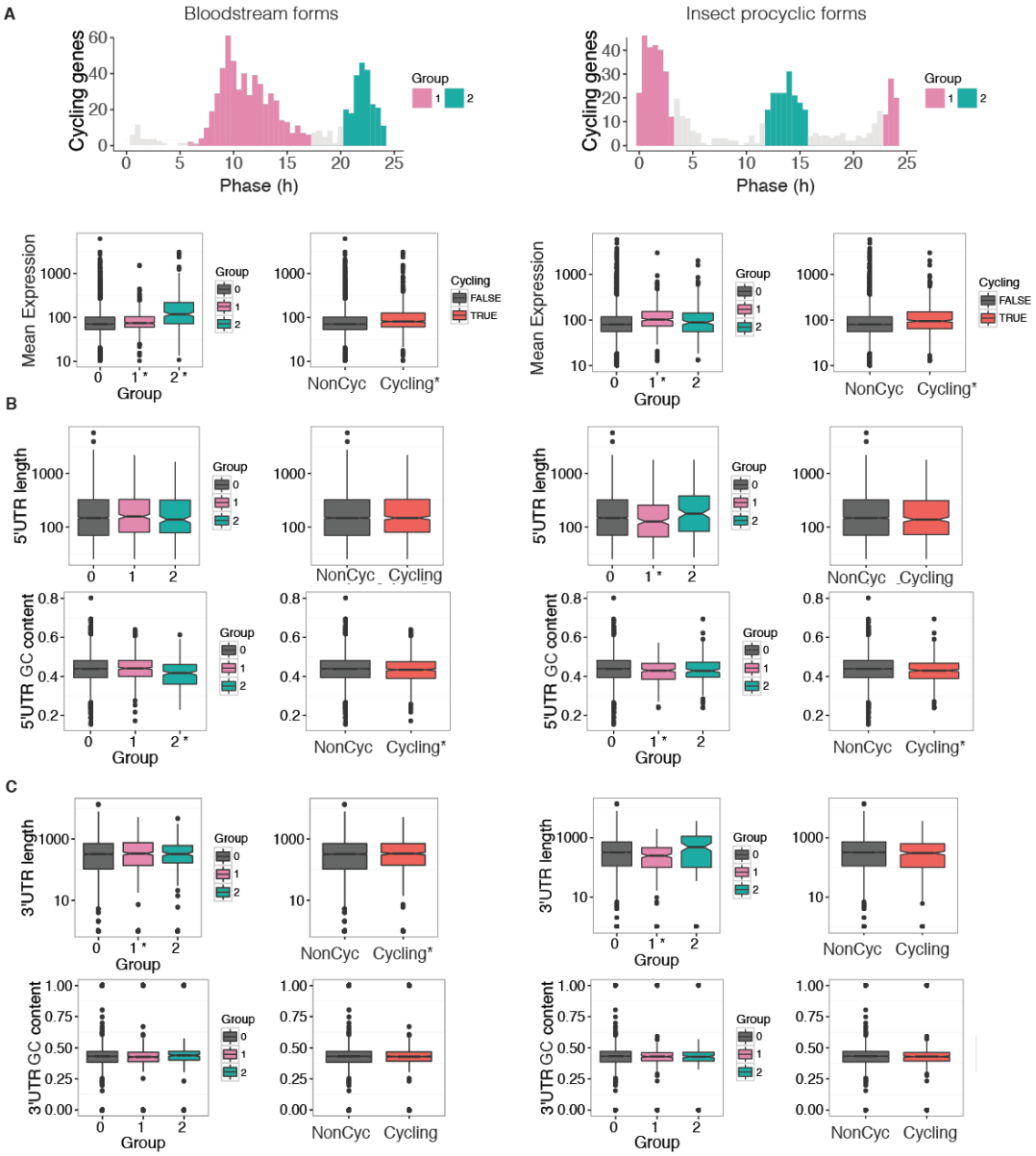


Fig. S14. Untranslated regions (UTRs) of cycling genes in *T. brucei* (A) bloodstream and insect procyclic forms. Cycling genes from two phase clusters were defined by centering at the phase mean of the cluster ± 2h, which was expanded if the consecutive cluster had more than 100 genes. (B) 5'UTR and (C) 3'UTR sequences length, GC content and mean expression were compared with non-cycling genes (Group 0). Both clusters of cycling genes groups were pooled (red) to compare with non-cycling genes (NonCyc, gray). * in the x axis label following the group number indicates mean of cycling genes is significantly different from mean of non-cycling genes, $p < 0.05$, Wilcoxon test.

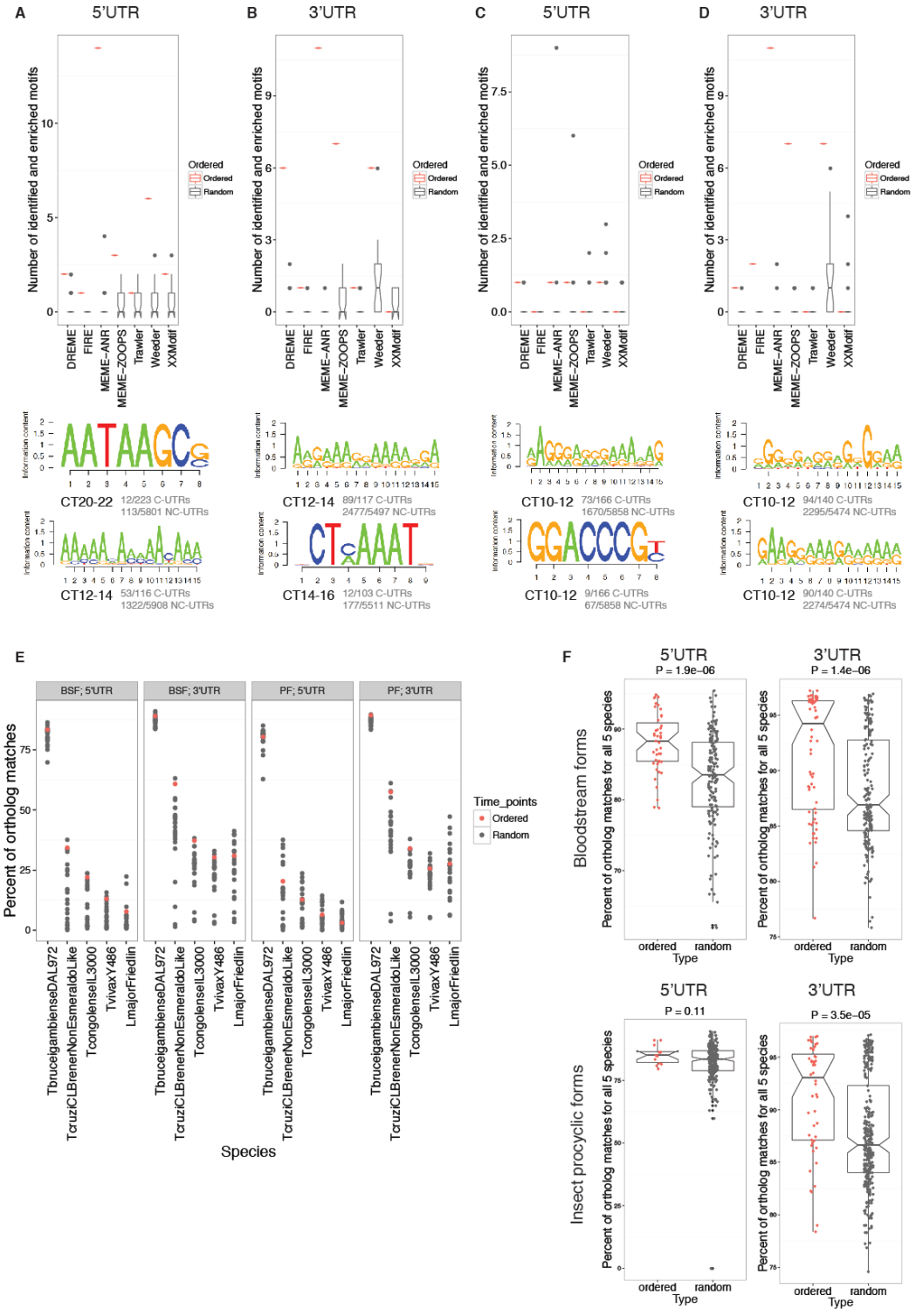


Fig. S15. Enriched motifs identified in (A-B) bloodstream and (C-D) procyclic UTR sequences using six different motif finder algorithms. Results in red represent the number of enriched motifs identified in the UTRs of cycling genes and the gray represents the distribution of the motifs identified when the genes in each cluster assignment were randomly selected 25 times (FDR < 0.05). In the bottom panel are representative sequence logos of two top putative *cis*-regulatory elements (pCREs)

enriched in the UTRs of cycling genes. (E) Percentage of orthologous genes that contain the conserved motif in the UTR. Red represents the percentage of all motifs identified in the UTR of cycling genes, and gray represents the conservation of motifs found in 25 random groups of genes (F) Global conservation of *T. brucei* motifs relative to five other kinetoplastida species. Motifs identified in cycling genes (red) are more conserved than motifs identified in randomly selected genes (gray). P values are annotated on the plot title.

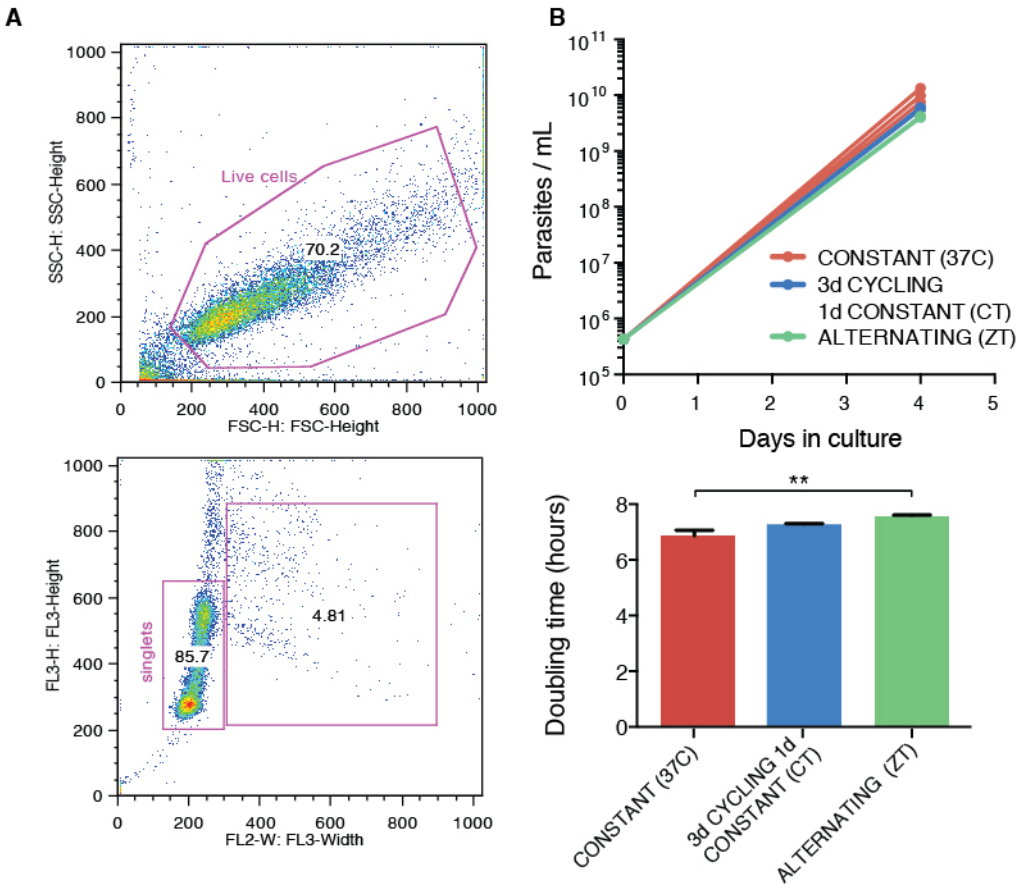


Fig. S16. Circadian gene expression is not correlated with cell cycle stage. (A) Gating strategy to identify frequency of live single parasites in each cell cycle stage (G1, S and G2/M phase) based on DNA content profile. (B) Parasite growth curve in different culture conditions and calculated doubling times. N = 6, Šídák multiple comparison of the means.

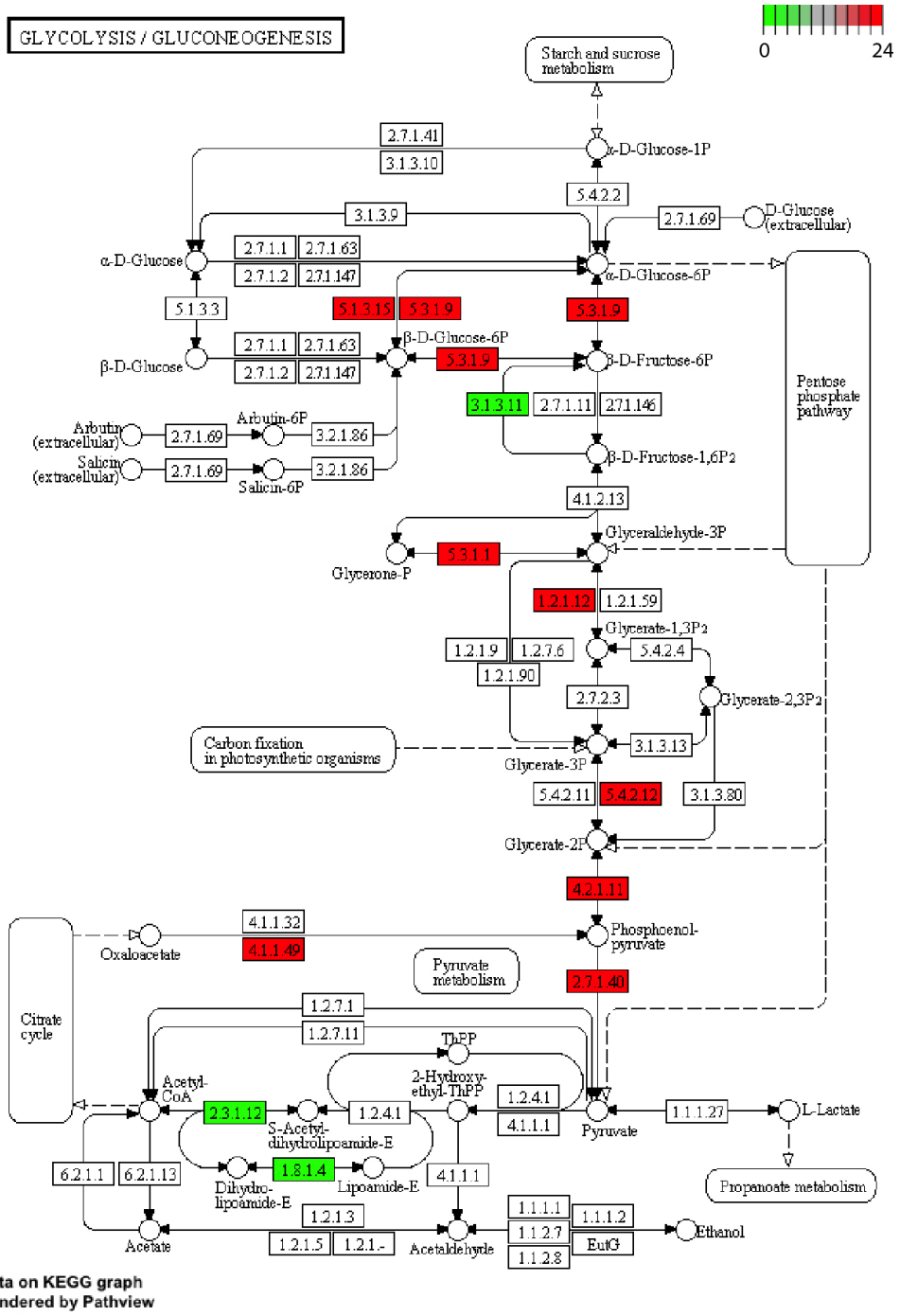


Fig. S17. Eleven genes of the glycolytic pathway cycle through the day, three have maximum expression close to phase CT0 (green) and eight have their maximum expression closer to CT24 (red).

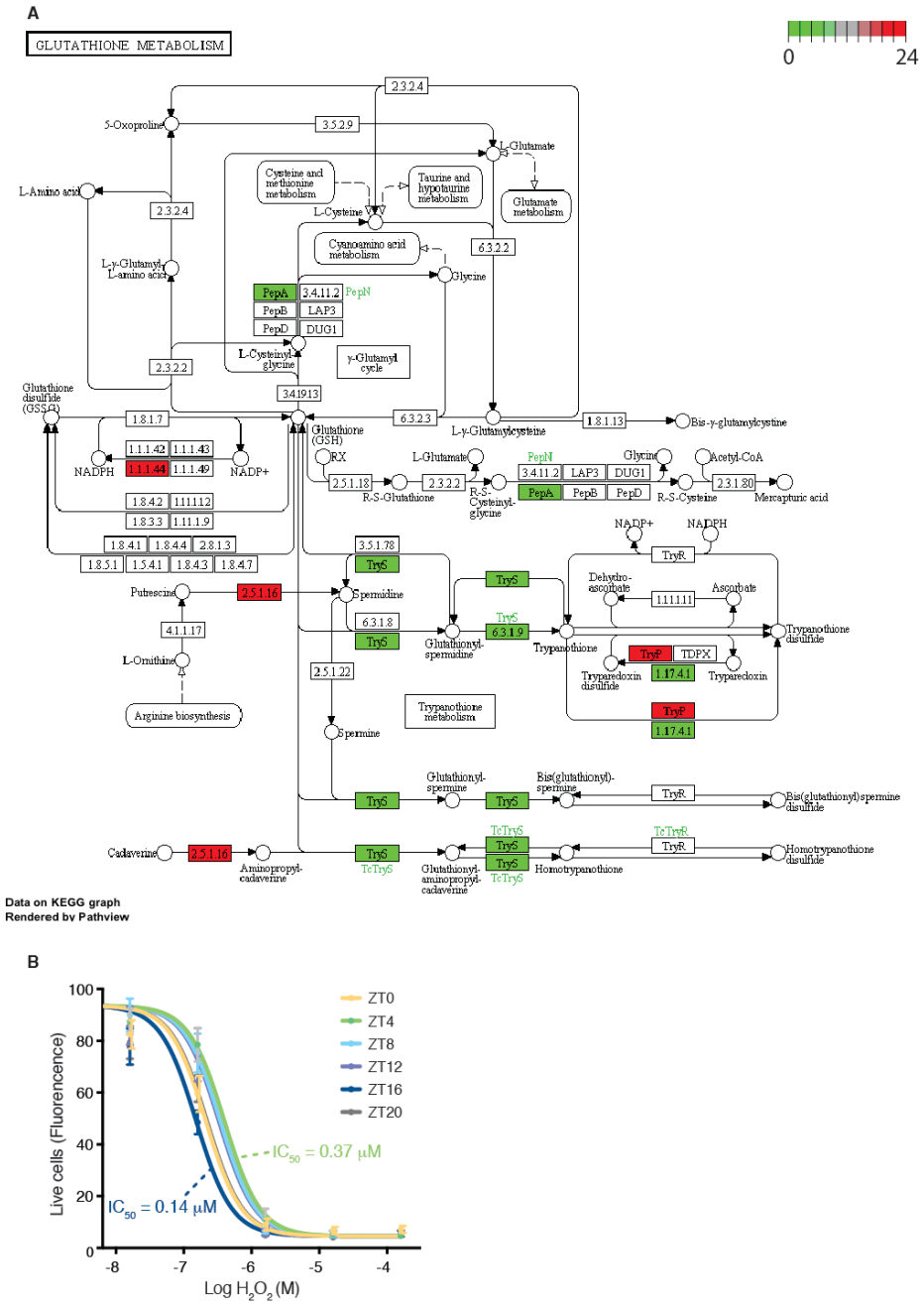


Fig. S18. Parasite sensitivity to oxidative stress changes throughout the day. (A) Glutathione pathway has multiple genes whose mRNA is cycling (colored red or green). Color gradient represents phase of maximum gene expression from CT0 (green) to CT24 (red). (B) Percentage of live cells collected around the clock and treated with serial dilution of H_2O_2 concentrations. $p < 0.0001$ non linear regression (variable slope, four parameters) comparison shows LogIC50 is different between data sets.

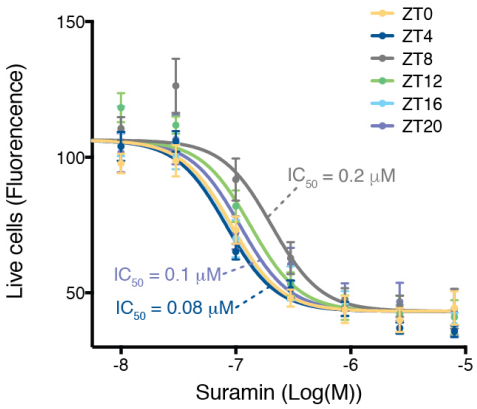


Fig. S19. Parasite sensitivity to suramin drug treatment changes throughout the day. Percentage of live cells collected around the clock and treated with serial dilution of suramin concentrations. p < 0.0001 non linear regression (variable slope, four parameters) comparison shows LogIC50 is different between data sets. Error bars represent standard error. N = 9 biological replicates tested in three independent experiments.

Table S1.
Cell cycle associated genes are not enriched among the clock-regulated gene pool.

Life cycle stage	Synchronization	Condition during collection	Number genes	Number of cell cycle genes	Proportion	p value
Bloodstream	Temperature	32/37°C	1490	117	0.0599	0.2702
		37°C	1092	44	0.0504	0.7815
	Light	LD	366	16	0.0734	0.0957
		DD	242	19	0.0547	1.0000
Insect-stage	Temperature	32/37°C	1123	82	0.0595	0.3723
		37°C	854	43	0.0646	0.1519
Bloodstream	-	-	9499	530 #	0.0528	-

cell cycle genes identified in Archer *et al.*

Table S2.
Primer sequences used for real-time PCR

Gene	gene ID	Fw/Rev	Seq (5' - 3')
Gim5A protein,glycosomal membrane protein (gim5A)	Tb927.9.11580	Fw	TGCGAATCTCTTGCACTGTC
Gim5A protein,glycosomal membrane protein (gim5A)	Tb927.9.11580	Rev	GTACATCCAACACAGCACGG
oligosaccharyl transferase subunit, putative	Tb927.5.900	Fw	GTCGGTGGACGACGAGTTAT
oligosaccharyl transferase subunit, putative	Tb927.5.900	Rev	CATGTGTTACCATCCGCAAG
amino acid transporter, putative	Tb927.8.7650	Fw	CTCGCAGTGGTTGCTGATAA
amino acid transporter, putative	Tb927.8.7650	Rev	GGTTGCCTGACTTCTGCTTC
proline dehydrogenase	Tb927.7.210	Fw	TGTTCTTGACTATGCCGCTG
proline dehydrogenase	Tb927.7.210	Rev	TTTTCGTCAAAGGCTTGCTT
cystathione gamma lyase, putative	Tb927.9.12320	Fw	AGGAGCTGTGCTTACTCCCA
cystathione gamma lyase, putative	Tb927.9.12320	Rev	TCACCAGCATTCATCAGAGC
phospholipase A1 (PLA1)	Tb927.1.4830	Fw	ACATCTTCCTGATGGTTGGC
phospholipase A1 (PLA1)	Tb927.1.4830	Rev	ATTTGCCGCTGCTAACTTGT
acidic phosphatase	Tb927.5.610	Fw	AGCGAGGGCTGAAGTTATCA
acidic phosphatase	Tb927.5.610	Rev	GAGCTTCATCTTCGTCCCAG
amino acid transporter, putative	Tb927.8.7740	Fw	GAAGAAGCGTTCCGTCACTC
amino acid transporter, putative	Tb927.8.7740	Rev	GATTTGCGTAGGGGTCGTAA
succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor, putative	Tb927.11.2690	Fw	ATGGACCTTACGATGCTTGG
succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor, putative	Tb927.11.2690	Rev	ACGCAGTGTTTTCCCGTAAC
constitutively expressed posttranscriptional regulator TbZFP3	Tb927.3.720	Fw	CAGGGGAAACGCAAAACTAA
constitutively expressed posttranscriptional regulator TbZFP3	Tb927.3.720	Rev	TGTCACCCCAACTGCATTCT

3.5. Acknowledgments

We thank Dr. Meg Phillips for help in starting parasite cultures at UT Southwestern and for advice on the suramin sensitivity experiments, Prachi Nakashe for the library preparations of the light/dark datasets, Michael Broderick for help during the oxidative stress experiment, Leonor Pinho for technical support, Jeremy Stubblefield for help during the blood collection, Gokhul Kilaru for technical support when running the permutation tests and Drs. Dave Barry, Meg Phillips, Carla Green and Margarida Vaz for reading the manuscript. The work presented has been supported by Howard Hughes Medical Institute to J.S.T, by HHMI International Early Career Scientist (55007419) to L.M.F., by Fundação para a Ciência e Tecnologia SFRH/BD/51286/2010 to F.R-F and IF/00595/2014 to N.L.B-M. The data reported in the paper are available in the Supplementary Materials Databases S1-S6.

CHAPTER III

CHAPTER III – General discussion and future directions

1. Main findings

The main findings of this thesis are summarized below:

T. brucei parasites adapt to adipose tissue

- Later in infection the adipose tissue is the host tissue with highest parasite load.
- *T. brucei* adapts to the lipid-rich environment of adipose tissue by remodeling its gene expression.
- In the adipose tissue *T. brucei* activates the beta-oxidation of fatty acids as carbon source.

T. brucei accelerates the circadian clock of its host

- *T. brucei* deregulates the daily locomotor activity and body temperature of its mouse host.
- Since early in infection the daily locomotor activity period is shortened by the infection.
- Organs with higher parasite load – adipose tissue – show a short period of PER2 expression, which can be rescued with suramin treatment.
- Later in infection, PER2 expression period in SCN is shorter than in controls, coinciding with the period when disruption of daily locomotor activity and temperature are observed.
- *Plasmodium chabaudi* infection does not interfere with the period of clock genes.
- *In vitro*, in the absence of immune cells, *T. brucei* shortens the period of fibroblasts.

T. brucei parasites have an endogenous circadian clock

- 10% of *T. brucei* gene expression oscillates daily.

- The daily gene expression oscillation is temperature compensated and cell cycle independent.
- In *T. brucei* the circadian regulation of gene expression is post-transcriptional.
- Most of the genes whose expression cycles are related to metabolic pathways.
- During the day *T. brucei* population experiences two 'rush hours' of gene expression, which correlates with higher ATP production.
- Parasites are more resistant to oxidative stress during the morning.
- During the day, there is a window when parasites are more susceptible to suramin treatment.

2. Discussion and potential future approaches

Extensive research and efforts to understand *Trypanosoma brucei* infection have been made, which have led to significant understanding of the process of antigenic variation, the parasite life cycle, as well as the parasite distribution in the mammalian host. However, with this thesis I hope I have emphasized that many aspects of this infection remain a mystery, even among the themes in which more work has been done.

What causes the symptoms in sleeping sickness patients? Why are there so many different manifestations: from sleep/wake cycle disruption and mood alterations, to weight loss? These are the aspects of this disease that have been lacking more attention due to their complexity. For my Ph.D. project I was particularly intrigued by the circadian features of this infection. I hope these projects have, even if humbly, contributed to a better understanding of it, and that they will serve as a foundation for future and more detailed studies.

Why do parasites occupy the adipose tissue?

The surprising observation of the adipose tissue as a major parasite reservoir raises many questions: Why do the parasites infiltrate preferentially this tissue? Is it to access the lipids it needs? Is it to hide from the massive antibody response in the blood circulation? If so, is immune response less effective against this parasite in this particular tissue versus others? Does the

parasite always need to infiltrate the host adipose tissue in order to progress in its life cycle? Perhaps it is going to turn out to be a combination of multiple factors, but with the high numbers of infiltrated parasites observed in this tissue and the metabolic changes it undergoes, it seems unlikely that the parasite is there just by chance.

In order to test if the upregulated pathway of beta-oxidation of fatty acids is indeed crucial for the parasite presence in the adipose tissue, or for the general fate of the infection, it would be interesting to knock out genes in this pathway and compare the parasitemia, the parasite distribution and host survival upon infection with beta-oxidation deficient parasites.

T. brucei invades the brain parenchyma through a mechanism of diapedesis, which seems to be in part dependent on IFN-γ, since IFN-$\gamma^{-/-}$ and IFN-γ receptor $^{-/-}$ mice have fewer parasites in the brain. This process of diapedesis is how inflammatory cells cross from the bloodstream to the tissues. Therefore it would be interesting to test whether IFN-$\gamma^{-/-}$ and IFN-γ receptor $^{-/-}$ mice have also fewer parasites in adipose tissue. It would be an approach to assess if crossing the vessels to the adipose tissue involve, at least in part, a similar process to crossing the blood vessels to the brain.

Why would parasites disrupt the host circadian clock?

T. brucei infection disrupts the circadian rhythm of its host. Why would such consequence be evolutionarily advantageous for the parasite? It seems plausible to assume that sleepiness during daytime increases the chances of trypanosomes to be transmitted to the insect. Tsetse flies are pool feeders, i.e. they cut the skin and suck up blood from the respective lesion. This bite is rather painful and sufficient feeding takes time. Humans can reach any part of the body with their hands therefore feeding would be more effective if the person does not react. Plus, tsetse flies are diurnal feeder and attracted to movement, so a disease in which the person is relatively active, with little sudden naps during the day, seems an optimal scenario for the fly blood meal. On the contrary, we could assume that the disruption of the sleep/wake cycle is not the most advantageous circadian disruption the parasite induces. The circadian clock regulates multiple physiological and metabolic functions in the

body, including immune response. The deregulation of the host clock might mean a deregulation of the very sophisticated and well-orchestrated immune response to pathogens, which would lead to a rather easier evasion to this response by the parasite. Finally we could assume the deregulation of the host metabolism is advantageous for the parasite. This parasite divides every 7 hours and therefore the need for nutrients is demanding. The deregulation of the circadian metabolism of the host, including the potential deregulation of its feeding behavior would provide more frequent abundance of nutrients, rather than only once a day.

It would be interesting to test these last two possibilities, if immune response and metabolism are dramatically affected in *T. brucei*-infected mice. I would start by collecting both blood and adipose tissue of control and infected mice during 48 h and extract RNA for RNAseq. With this we could analyze the circadian parameters of immune and metabolic functions among the two mice groups.

Interestingly, with our studies we provided insights of the potential specific changes *T. brucei* has on circadian clock of the host: a period shortening. This detailed characterization has not been studied in humans, however is curious to correlated with the indirect observations made. Pineal hormone melatonin, which is the hormone that anticipates the daily onset of darkness, is maintained in sleeping sickness patients in the encephalitic stage, although it is still cycling, it is being secreted rather earlier than expected (phase advanced) (*40*). Phase advance means the peak of the expression of melatonin is happening earlier, which is also what we observe in the expression of PER2 protein. Since our approach allows measuring continuously PER2 expression and without the interference of external stimuli, we uncovered that this phase advance in PER2 expression coincided with a shorter period in mice. Do melatonin levels also peak earlier in the mouse model? Perhaps this melatonin phase adavance indicates that in humans there is also a shortening of the host period upon infection. This would be complicated but interesting to address in patients, by isolating them from external cues.

The hypothesis of sleeping sickness being a circadian disorder has been previously considered. In fact Lundkvist *et al*. tested this hypothesis in the rat model of sleeping sickness. They also reported a shorter *Per1* expression period of the pituitary gland of infected rats, although they were not able to detect any period changes in the SCN, nor disruption of the circadian locomotor activity and temperature (*252*). The difference in findings between our studies is likely due to the rodent model used. The rat model of sleeping sickness has a high variability, probably associated with the fact that Wistar rats are outbred, i.e. bred from parents not closely related. This leads to high variability among rats, which is clear when assessing their survival to *T. brucei* infection, which happens on average 15.3 ± 12.4 days post-infection, with death occurring between the tenth and forty-sixth days of infection (*253*). In addition their study of locomotor activity and temperature was assessed in light/dark conditions. When we assessed those parameters in LD, we also did not detect changes between infected and control mice early after infection, other than reduced activity. Only when we used the infection protocol that better mimicks the low parasitemia and chronic infection observed in humans (by treating mice with suramin) we were able to uncover the disruption of the timing at which both locomotor activity and higher body temperature occur. Possibly these circadian disruptions in LD require a more chronic infection in order to emerge.

T. brucei parasites also count the time

Trypanosoma brucei turned out to be an ideal system to test whether parasites have intrinsic circadian rhythms: it is extracellular, more than one life cycle stage can be cultured, the cell cycle is different from 24 h and its genome fully sequenced. The first two features, being extracellular and cultured allowed us to exclude any possible effect from the host on this timekeeping time. All the circadian oscillations have to be driven by the parasite itself.

43% of all protein coding genes showed circadian rhythms in transcription somewhere in the body, largely in an organ-specific manner (*138*). Even, when considering individual tissues the number of cycling genes is

substantial, with approximately ~20% of genes cycling in the liver and the SCN (Table 1).

In our study we found that ~10% of the transcriptome of *T. brucei* was oscillating daily, but the amplitude of such oscillations are not as high as in the liver: ~2.5-fold in liver versus ~1.3-fold in *T. brucei*. This could possibly be due, to some extent, to asynchrony among the parasite population (which would dilute the maximum and minimum expression levels). Perhaps the temperature entrainment we performed *in vitro* is not sufficient to fully synchronize all parasites in culture.

This might be also the reason why the number of cycling genes in immortalized cell lines, of both humans and mice, are extremely low (0.1-1.2%), and those that cycle, do so with very low amplitude.

Table 1. Summary of the percentage of cycling genes identified by RNA sequencing or microarray in different species and organs.

Species	Tissue	Method	% cycling transcripts	Reference
Mouse	Liver	RNAseq	22%	(*137*)
		RNAseq	15%	(*138*)
		microarray	16%	(*254*)
	Hypothalamus	RNAseq	2.9%	(*138*)
	SCN	RNAseq	24%	(*255*)
	Macrophages	RNAseq	8.1%	(*140*)
Mouse cell line	NIH3T3	microarray	0.1%	(*254*)
Human	Blood	RNAseq	6.4%	(*256*)
Human cell line	U2OS cells	RNAseq	1.2%	(*257*)
		microarray	0.1%	(*254*)
Chlamydomonas reinhardtii	Whole photosynthetic algae	RNAseq	50%	(*258*)
T. brucei	Whole parasite	RNAseq	10%	Present thesis

Based on these previous observations, 10% of cycling gene expression *in vitro* is high, and this could anticipate that both number of cycling genes and their amplitude would increase if we measure the *T. brucei* transcriptome *in vivo*. It would be interesting to do a circadian analysis of the in vivo transcriptome in both mammalian tissues (particularly adipose tissue and brain) and in the tsetse fly.

One approach we are attempting to follow up with this project is the development of a circadian transcriptional reporter cell line for the parasite. This means cloning the 3'UTR of a gene whose expression is cycling fused to Firefly Luciferase, which would allow us to measure in the lumicycle (real-time luminescence measurement) the expression of this gene. This tool will allow multiple further characterizations, such as what are the stimuli *T. brucei* entrains to: serum, glucose, lipids etc. Furthermore, this tool would allow us to easily test whether some genes are essential, or at least impact the circadian clock once mutated / knocked out. Since the parasite circadian transcriptome appears to be regulated at the post-transcriptional level, good candidates for *T. brucei* clock genes would be RNA binding proteins and/or potentially non-coding RNAs.

Once circadian mutants have been identified the next exciting experiment would be to assess if these mutants are viable and able to establish infection in mouse. In case they are viable, further analyses would be to test assess parasitemia profile, progression in the parasite's life cycle, mouse survival and ultimately whether these parasites still lead to the circadian manifestations in the mouse model. Since the circadian clock genes across species are so little conserved, it seems plausible that once *T. brucei* clock genes are identified, and if their mutation leads to impaired infection, they could be potential new drug targets to treat this disease.

This study also begs for the study of potential circadian rhythms in gene expression and other outputs in other pathogens, which are possibly not only restricted to parasites.

As little Annie would say:

'The sun will come out tomorrow
So you gotta hang on
'till tomorrow, come what may!
Tomorrow, tomorrow, I love ya, tomorrow
You're only a day away'

Graffiti at Deep Ellum, Dallas, TX by @Joeskilz52

CHAPTER IV

CHAPTER IV – Publications

The following scientific articles and manuscripts in preparation are the result of the Ph.D. research work portrayed in this dissertation:

1. Sandra Trindade[#] & **Filipa Rijo-Ferreira**[#], Tânia Carvalho, Daniel Pinto-Neves, Fabien Guegan, Francisco Aresta-Branco, Fabio Bento, Simon A. Young, Andreia Pinto, Jan Van Den Abbeele, Ruy M. Ribeiro, Sérgio Dias, Terry K. Smith, Luisa M. Figueiredo 2016. *Trypanosoma brucei* parasites occupy and functionally adapt to adipose tissue in mice. *Cell Host&Microbe* (in press)

2. **Filipa Rijo-Ferreira**, Margarida Sanches-Vaz, Luisa M. Figueiredo[#], Joseph S Takahashi[#] 2016. *Trypanosoma brucei* infection accelerates the mouse circadian rhythm (in preparation)

3. **Filipa Rijo-Ferreira**, Daniel Pinto-Neves, Nuno L Barbosa-Morais, Joseph S Takahashi[#], Luisa M. Figueiredo[#] 2016. Post-Transcriptional Circadian Regulation of Genes Related to Metabolism and Suramin Sensitivity in *Trypanosoma brucei. Science* (submitted)

[#] these authors contributed equally to the present work

CHAPTER V

CHAPTER V – References

1. W. Gibson, Resolution of the species problem in African trypanosomes. *Int J Parasitol* **37**, 829-838 (2007).
2. WHO, Control and surveillance of human African trypanosomiasis: report of a WHO expert committee. *WHO Tecnical Report Series. Geneva, Switzerland: World Health Organization*, (2013).
3. P. P. Simarro *et al.*, The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int J Health Geogr* **9**, 57 (2010).
4. B. Vanhollebeke *et al.*, A haptoglobin-hemoglobin receptor conveys innate immunity to Trypanosoma brucei in humans. *Science* **320**, 677-681 (2008).
5. L. Vanhamme *et al.*, Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* **422**, 83-87 (2003).
6. P. Uzureau *et al.*, Mechanism of Trypanosoma brucei gambiense resistance to human serum. *Nature* **501**, 430-434 (2013).
7. C. Giroud *et al.*, Murine Models for Trypanosoma brucei gambiense disease progression--from silent to chronic infections and early brain tropism. *PLoS Negl Trop Dis* **3**, e509 (2009).
8. S. Mogk *et al.*, Cyclical appearance of african trypanosomes in the cerebrospinal fluid: new insights in how trypanosomes enter the CNS. *PLoS One* **9**, e91372 (2014).
9. F. Checchi, J. A. Filipe, D. T. Haydon, D. Chandramohan, F. Chappuis, Estimates of the duration of the early and late stage of gambiense sleeping sickness. *BMC Infect Dis* **8**, 16 (2008).
10. M. Odiit, F. Kansiime, J. C. Enyaru, Duration of symptoms and case fatality of sleeping sickness caused by Trypanosoma brucei rhodesiense in Tororo, Uganda. *East Afr Med J* **74**, 792-795 (1997).
11. R. Brun, J. Blum, F. Chappuis, C. Burri, Human African trypanosomiasis. *Lancet* **375**, 148-159 (2010).
12. A. Buguet, S. Bisser, T. Josenando, F. Chapotot, R. Cespuglio, Sleep structure: a new diagnostic tool for stage determination in sleeping sickness. *Acta Trop* **93**, 107-117 (2005).
13. V. Jamonneau *et al.*, Untreated Human Infections by Trypanosoma brucei gambiense Are Not 100% Fatal. *PLoS Negl Trop Dis* **6**, e1691 (2012).
14. P. Guirnalda, N. B. Murphy, D. Nolan, S. J. Black, Anti-Trypanosoma brucei activity in Cape buffalo serum during the cryptic phase of parasitemia is mediated by antibodies. *Int J Parasitol* **37**, 1391-1399 (2007).
15. FAO, Programme Against African Trypanosomosis (PAAT). *Food and Agriculture Organization of the United Nations* http://www.fao.org/ag/AGAINFO/programmes/en/paat/disease.html**.** , (2014).
16. P. G. Kennedy, Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). *Lancet neurology* **12**, 186-194 (2013).

17. P. G. Kennedy, Human African trypanosomiasis of the CNS: current issues and challenges. *J Clin Invest* **113**, 496-504 (2004).
18. A. H. Fairlamb, G. B. Henderson, A. Cerami, Trypanothione is the primary target for arsenical drugs against African trypanosomes. *Proc Natl Acad Sci U S A* **86**, 2607-2611 (1989).
19. P. G. Kennedy, The continuing problem of human African trypanosomiasis (sleeping sickness). *Ann Neurol* **64**, 116-126 (2008).
20. I. M. Vincent *et al.*, A molecular mechanism for eflornithine resistance in African trypanosomes. *PLoS Pathog* **6**, e1001204 (2010).
21. WHO, Investing to overcome the global impact of neglected tropical diseases: third WHO report on neglected tropical diseases. (2015).
22. C. Cirelli, The genetic and molecular regulation of sleep: from fruit flies to humans. *Nat Rev Neurosci* **10**, 549-560 (2009).
23. J. Blum, C. Schmid, C. Burri, Clinical aspects of 2541 patients with second stage human African trypanosomiasis. *Acta Trop* **97**, 55-64 (2006).
24. A. K. Njamnshi *et al.*, Actigraphy in human african trypanosomiasis as a tool for objective clinical evaluation and monitoring: a pilot study. *PLoS Negl Trop Dis* **6**, e1525 (2012).
25. B. M. Sanner, N. Buchner, S. Kotterba, W. Zidek, Polysomnography in acute African trypanosomiasis. *J Neurol* **247**, 878-879 (2000).
26. A. Buguet *et al.*, 24 hour polysomnographic evaluation in a patient with sleeping sickness. *Electroencephalogr Clin Neurophysiol* **72**, 471-478 (1989).
27. A. Buguet *et al.*, The duality of sleeping sickness: focusing on sleep. *Sleep Med Rev* **5**, 139-153 (2001).
28. A. Buguet, P. Tapie, J. Bert, Reversal of the sleep/wake cycle disorder of sleeping sickness after trypanosomicide treatment. *J Sleep Res* **8**, 225-235 (1999).
29. L. Lin *et al.*, The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98**, 365-376 (1999).
30. M. Bentivoglio, K. Kristensson, Neural-immune interactions in disorders of sleep-wakefulness organization. *Trends Neurosci* **30**, 645-652 (2007).
31. Y. Dauvilliers *et al.*, Hypocretin and human African trypanosomiasis. *Sleep* **31**, 348-354 (2008).
32. A. Darsaud *et al.*, Twenty-four-hour disruption of the sleep-wake cycle and sleep-onset REM-like episodes in a rat model of African trypanosomiasis. *Sleep* **27**, 42-46 (2004).
33. G. Grassi-Zucconi *et al.*, Sleep fragmentation, and changes in locomotor activity and body temperature in trypanosome-infected rats. *Brain Res Bull* **37**, 123-129 (1995).
34. P. F. Seke Etet *et al.*, Sleep and rhythm changes at the time of Trypanosoma brucei invasion of the brain parenchyma in the rat. *Chronobiol Int* **29**, 469-481 (2012).
35. S. Nishino, B. Ripley, S. Overeem, G. J. Lammers, E. Mignot, Hypocretin (orexin) deficiency in human narcolepsy. *Lancet* **355**, 39-40 (2000).
36. M. Palomba *et al.*, Alterations of orexinergic and melanin-concentrating hormone neurons in experimental sleeping sickness. *Neuroscience*, (2015).

37. O. Hayaishi, Y. Urade, N. Eguchi, Z. L. Huang, Genes for prostaglandin d synthase and receptor as well as adenosine A2A receptor are involved in the homeostatic regulation of nrem sleep. *Arch Ital Biol* **142**, 533-539 (2004).
38. V. W. Pentreath, K. Rees, O. A. Owolabi, K. A. Philip, F. Doua, The somnogenic T lymphocyte suppressor prostaglandin D2 is selectively elevated in cerebrospinal fluid of advanced sleeping sickness patients. *Trans R Soc Trop Med Hyg* **84**, 795-799 (1990).
39. B. K. Kubata, M. Duszenko, K. S. Martin, Y. Urade, Molecular basis for prostaglandin production in hosts and parasites. *Trends Parasitol* **23**, 325-331 (2007).
40. B. Claustrat *et al.*, Plasma melatonin rhythm is maintained in human African trypanosomiasis. *Neuroendocrinology* **68**, 64-70 (1998).
41. G. Brandenberger *et al.*, Disruption of endocrine rhythms in sleeping sickness with preserved relationship between hormonal pulsatility and the REM-NREM sleep cycles. *J Biol Rhythms* **11**, 258-267 (1996).
42. M. W. Radomski *et al.*, Disruptions in the secretion of cortisol, prolactin, and certain cytokines in human African trypanosomiasis patients. *Bull Soc Pathol Exot* **87**, 376-379 (1994).
43. B. Berge *et al.*, Disruptions of ultradian and circadian organization of core temperature in a rat model of African trypanosomiasis using periodogram techniques on detrended data. *Chronobiol Int* **22**, 237-251 (2005).
44. C. Chevrier *et al.*, Clinical assessment of the entry into neurological state in rat experimental African trypanosomiasis. *Acta Trop* **95**, 33-39 (2005).
45. A. Buguet *et al.*, Sleep-wake cycle in human African trypanosomiasis. *J Clin Neurophysiol* **10**, 190-196 (1993).
46. G. B. Lundkvist, K. Kristensson, M. Bentivoglio, Why trypanosomes cause sleeping sickness. *Physiology (Bethesda)* **19**, 198-206 (2004).
47. T. Cavalier-Smith, Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree. *Biol Lett* **6**, 342-345 (2010).
48. S. M. Adl *et al.*, The revised classification of eukaryotes. *J Eukaryot Microbiol* **59**, 429-493 (2012).
49. J. Lukes, T. Skalicky, J. Tyc, J. Votypka, V. Yurchenko, Evolution of parasitism in kinetoplastid flagellates. *Mol Biochem Parasitol* **195**, 115-122 (2014).
50. M. Berriman *et al.*, The genome of the African trypanosome Trypanosoma brucei. *Science* **309**, 416-422 (2005).
51. F. Hawking, The resistance of Trypanosoma congolense, T. vivax and T. evansi to human plasma. *Trans R Soc Trop Med Hyg* **72**, 405-407 (1978).
52. R. Benne *et al.*, Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**, 819-826 (1986).
53. J. C. Boothroyd, G. A. Cross, Transcripts coding for variant surface glycoproteins of Trypanosoma brucei have a short, identical exon at their 5' end. *Gene* **20**, 281-289 (1982).
54. M. A. Ferguson, S. W. Homans, R. A. Dwek, T. W. Rademacher, Glycosyl-phosphatidylinositol moiety that anchors Trypanosoma brucei variant surface glycoprotein to the membrane. *Science* **239**, 753-759 (1988).

55. F. Claes *et al.*, Bioluminescent imaging of Trypanosoma brucei shows preferential testis dissemination which may hamper drug efficacy in sleeping sickness. *PLoS Negl Trop Dis* **3**, e486 (2009).
56. E. Vassella, B. Reuner, B. Yutzy, M. Boshart, Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. *Journal of cell science* **110 (Pt 21)**, 2661-2671 (1997).
57. L. Peacock *et al.*, Identification of the meiotic life cycle stage of Trypanosoma brucei in the tsetse fly. *Proc Natl Acad Sci U S A* **108**, 3671-3676 (2011).
58. S. Kramer, Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. *Mol Biochem Parasitol* **181**, 61-72 (2012).
59. D. Nilsson *et al.*, Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei. *PLoS Pathog* **6**, e1001037 (2010).
60. T. N. Siegel, D. R. Hekstra, X. Wang, S. Dewell, G. A. Cross, Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites. *Nucleic Acids Res* **38**, 4946-4957 (2010).
61. S. Kabani *et al.*, Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in Trypanosoma brucei. *BMC genomics* **10**, 427 (2009).
62. B. C. Jensen, D. Sivam, C. T. Kifer, P. J. Myler, M. Parsons, Widespread variation in transcript abundance within and across developmental stages of Trypanosoma brucei. *BMC genomics* **10**, 482 (2009).
63. I. Roditi *et al.*, Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of Trypanosoma brucei. *J Cell Biol* **108**, 737-746 (1989).
64. C. E. Clayton, M. R. Mowatt, The procyclic acidic repetitive proteins of Trypanosoma brucei. Purification and post-translational modification. *J Biol Chem* **264**, 15088-15093 (1989).
65. S. Urwyler, E. Studer, C. K. Renggli, I. Roditi, A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of Trypanosoma brucei. *Mol Microbiol* **63**, 218-228 (2007).
66. G. A. Cross, Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. *Parasitology* **71**, 393-417 (1975).
67. A. Ferrante, A. C. Allison, Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat. *Parasite Immunol* **5**, 491-498 (1983).
68. J. D. Barry, Capping of variable antigen on Trypanosoma brucei, and its immunological and biological significance. *Journal of cell science* **37**, 287-302 (1979).
69. S. K. Natesan, A. Black, K. R. Matthews, J. C. Mottram, M. C. Field, Trypanosoma brucei brucei: endocytic recycling is important for mouse infectivity. *Exp Parasitol* **127**, 777-783 (2011).

70. S. K. Natesan, L. Peacock, K. Matthews, W. Gibson, M. C. Field, Activation of endocytosis as an adaptation to the mammalian host by trypanosomes. *Eukaryot Cell* **6**, 2029-2037 (2007).
71. M. Engstler *et al.*, Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. *Cell* **131**, 505-515 (2007).
72. J. M. Mansfield, D. M. Paulnock, Regulation of innate and acquired immunity in African trypanosomiasis. *Parasite Immunol* **27**, 361-371 (2005).
73. A. Schwede, N. Jones, M. Engstler, M. Carrington, The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. *Mol Biochem Parasitol* **175**, 201-204 (2011).
74. S. J. Black, P. Guirnalda, D. Frenkel, C. Haynes, V. Bockstal, Induction and regulation of Trypanosoma brucei VSG-specific antibody responses. *Parasitology* **137**, 2041-2049 (2010).
75. R. Ross, D. Thomson, A Case of Sleeping Sickness showing Regular Periodical Increase of the Parasites Disclosed. *British medical journal* **1**, 1544-1545 (1910).
76. C. Mulenga, J. D. Mhlanga, K. Kristensson, B. Robertson, Trypanosoma brucei brucei crosses the blood-brain barrier while tight junction proteins are preserved in a rat chronic disease model. *Neuropathol Appl Neurobiol* **27**, 77-85 (2001).
77. H. Schmidt, The pathogenesis of trypanosomiasis of the CNS. Studies on parasitological and neurohistological findings in trypanosoma rhodesiense infected vervet monkeys. *Virchows Arch A Pathol Anat Histopathol* **399**, 333-343 (1983).
78. K. Kristensson, M. Nygard, G. Bertini, M. Bentivoglio, African trypanosome infections of the nervous system: parasite entry and effects on sleep and synaptic functions. *Prog Neurobiol* **91**, 152-171 (2010).
79. L. L. Rubin, J. M. Staddon, The cell biology of the blood-brain barrier. *Annu Rev Neurosci* **22**, 11-28 (1999).
80. N. J. Abbott, L. Ronnback, E. Hansson, Astrocyte-endothelial interactions at the blood-brain barrier. *Nat Rev Neurosci* **7**, 41-53 (2006).
81. K. Kristensson, Microbes' roadmap to neurons. *Nat Rev Neurosci* **12**, 345-357 (2011).
82. W. Masocha *et al.*, Cerebral vessel laminins and IFN-gamma define Trypanosoma brucei brucei penetration of the blood-brain barrier. *J Clin Invest* **114**, 689-694 (2004).
83. J. A. Coles *et al.*, Intravital imaging of a massive lymphocyte response in the cortical dura of mice after peripheral infection by trypanosomes. *PLoS Negl Trop Dis* **9**, e0003714 (2015).
84. J. Rodgers *et al.*, Magnetic resonance imaging to assess blood-brain barrier damage in murine trypanosomiasis. *Am J Trop Med Hyg* **84**, 344-350 (2011).
85. A. A. Poltera, A. Hochmann, W. Rudin, P. H. Lambert, Trypanosoma brucei brucei: a model for cerebral trypanosomiasis in mice--an immunological, histological and electronmicroscopic study. *Clin Exp Immunol* **40**, 496-507 (1980).
86. U. Frevert *et al.*, Early invasion of brain parenchyma by african trypanosomes. *PLoS One* **7**, e43913 (2012).

87. A. A. Poltera, A. Hochmann, P. H. Lambert, A model for cardiopathy induced by Trypanosoma brucei brucei in mice. A histologic and immunopathologic study. *Am J Pathol* **99**, 325-352 (1980).
88. W. I. Morrison, M. Murray, P. D. Sayer, J. M. Preston, The pathogenesis of experimentally induced Trypanosoma brucei infection in the dog. I. Tissue and organ damage. *Am J Pathol* **102**, 168-181 (1981).
89. J. A. Blum *et al.*, Cardiac alterations in human African trypanosomiasis (T.b. gambiense) with respect to the disease stage and antiparasitic treatment. *PLoS Negl Trop Dis* **3**, e383 (2009).
90. T. O. Frommel, Y. Fujikura, J. R. Seed, Tissue alterations in Microtus montanus chronically infected with Trypanosoma brucei gambiense. *J Parasitol* **77**, 164-167 (1991).
91. W. I. Morrison, M. Murray, P. D. Sayer, J. M. Preston, The pathogenesis of experimentally induced Trypanosoma brucei infection in the dog. II. Change in the lymphoid organs. *Am J Pathol* **102**, 182-194 (1981).
92. G. J. Losos, B. O. Ikede, Pathology of experimental trypanosomiasis in the albino rat, rabbit, goat and sheep--A preliminary report. *Can J Comp Med* **34**, 209-212 (1970).
93. J. M. Bafort, H. Schmidt, Experimental chronic Trypanosoma brucei rhodesiense infection in Microtus montanus. *Am J Trop Med Hyg* **32**, 968-975 (1983).
94. G. A. Cross, H. S. Kim, B. Wickstead, Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427. *Mol Biochem Parasitol* **195**, 59-73 (2014).
95. B. Wickstead, K. Ersfeld, K. Gull, The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes. *Genome research* **14**, 1014-1024 (2004).
96. N. S. Alsford *et al.*, The identification of circular extrachromosomal DNA in the nuclear genome of Trypanosoma brucei. *Mol Microbiol* **47**, 277-289 (2003).
97. K. K. Chen, J. E. Donelson, Sequences of two kinetoplast DNA minicircles of Tryptanosoma brucei. *Proc Natl Acad Sci U S A* **77**, 2445-2449 (1980).
98. N. G. Kolev *et al.*, The transcriptome of the human pathogen Trypanosoma brucei at single-nucleotide resolution. *PLoS Pathog* **6**, e1001090 (2010).
99. N. M. El-Sayed *et al.*, The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science* **309**, 409-415 (2005).
100. N. M. El-Sayed *et al.*, The sequence and analysis of Trypanosoma brucei chromosome II. *Nucleic Acids Res* **31**, 4856-4863 (2003).
101. J. M. Kooter, P. Borst, Alpha-amanitin-insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. *Nucleic Acids Res* **12**, 9457-9472 (1984).
102. A. Gunzl *et al.*, RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. *Eukaryot Cell* **2**, 542-551 (2003).
103. B. Albert, J. Perez-Fernandez, I. Leger-Silvestre, O. Gadal, Regulation of ribosomal RNA production by RNA polymerase I: does elongation come first? *Genet Res Int* **2012**, 276948 (2012).

104. D. A. Dunbar, A. A. Chen, S. Wormsley, S. J. Baserga, The genes for small nucleolar RNAs in Trypanosoma brucei are organized in clusters and are transcribed as a polycistronic RNA. *Nucleic Acids Res* **28**, 2855-2861 (2000).
105. G. Gilinger, V. Bellofatto, Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. *Nucleic Acids Res* **29**, 1556-1564 (2001).
106. A. Gunzl *et al.*, Transcription of the Trypanosoma brucei spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. *Mol Biochem Parasitol* **85**, 67-76 (1997).
107. J. R. Wright, T. N. Siegel, G. A. Cross, Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in Trypanosoma brucei. *Mol Biochem Parasitol* **172**, 141-144 (2010).
108. T. N. Siegel *et al.*, Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes Dev* **23**, 1063-1076 (2009).
109. V. Nakaar, A. O. Dare, D. Hong, E. Ullu, C. Tschudi, Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes. *Mol Cell Biol* **14**, 6736-6742 (1994).
110. P. J. Johnson, J. M. Kooter, P. Borst, Inactivation of transcription by UV irradiation of T. brucei provides evidence for a multicistronic transcription unit including a VSG gene. *Cell* **51**, 273-281 (1987).
111. J. Blattner, C. E. Clayton, The 3'-untranslated regions from the Trypanosoma brucei phosphoglycerate kinase-encoding genes mediate developmental regulation. *Gene* **162**, 153-156 (1995).
112. T. Manful, A. Fadda, C. Clayton, The role of the 5'-3' exoribonuclease XRNA in transcriptome-wide mRNA degradation. *RNA* **17**, 2039-2047 (2011).
113. B. C. Jensen *et al.*, Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. *BMC genomics* **15**, 911 (2014).
114. J. J. Vasquez, C. C. Hon, J. T. Vanselow, A. Schlosser, T. N. Siegel, Comparative ribosome profiling reveals extensive translational complexity in different Trypanosoma brucei life cycle stages. *Nucleic Acids Res* **42**, 3623-3637 (2014).
115. N. G. Kolev, E. Ullu, C. Tschudi, The emerging role of RNA-binding proteins in the life cycle of Trypanosoma brucei. *Cell Microbiol*, (2014).
116. N. G. Kolev, K. Ramey-Butler, G. A. Cross, E. Ullu, C. Tschudi, Developmental progression to infectivity in Trypanosoma brucei triggered by an RNA-binding protein. *Science* **338**, 1352-1353 (2012).
117. H. S. Najafabadi *et al.*, Global identification of conserved post-transcriptional regulatory programs in trypanosomatids. *Nucleic Acids Res* **41**, 8591-8600 (2013).
118. M. Wurst *et al.*, Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in Trypanosoma brucei. *Mol Microbiol* **83**, 1048-1063 (2012).
119. E. Ullu, K. R. Matthews, C. Tschudi, Temporal order of RNA-processing reactions in trypanosomes: rapid trans splicing precedes polyadenylation of newly synthesized tubulin transcripts. *Mol Cell Biol* **13**, 720-725 (1993).

120. G. Mair *et al.*, A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA* **6**, 163-169 (2000).
121. C. B. Green, J. S. Takahashi, J. Bass, The meter of metabolism. *Cell* **134**, 728-742 (2008).
122. J. L. Ditty, S. B. Williams, S. S. Golden, A cyanobacterial circadian timing mechanism. *Annu Rev Genet* **37**, 513-543 (2003).
123. C. S. Pittendrigh, Temporal organization: reflections of a Darwinian clock-watcher. *Annual review of physiology* **55**, 16-54 (1993).
124. M. A. Woelfle, Y. Ouyang, K. Phanvijhitsiri, C. H. Johnson, The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. *Curr Biol* **14**, 1481-1486 (2004).
125. A. N. Dodd *et al.*, Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**, 630-633 (2005).
126. P. L. Lowrey, J. S. Takahashi, Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annu Rev Genomics Hum Genet* **5**, 407-441 (2004).
127. R. S. Edgar *et al.*, Peroxiredoxins are conserved markers of circadian rhythms. *Nature* **485**, 459-464 (2012).
128. M. H. Hastings, A. B. Reddy, E. S. Maywood, A clockwork web: circadian timing in brain and periphery, in health and disease. *Nat Rev Neurosci* **4**, 649-661 (2003).
129. M. R. Ralph, R. G. Foster, F. C. Davis, M. Menaker, Transplanted suprachiasmatic nucleus determines circadian period. *Science* **247**, 975-978 (1990).
130. J. S. Takahashi, M. Menaker, Role of the suprachiasmatic nuclei in the circadian system of the house sparrow, Passer domesticus. *J Neurosci* **2**, 815-828 (1982).
131. J. S. Takahashi, H. K. Hong, C. H. Ko, E. L. McDearmon, The genetics of mammalian circadian order and disorder: implications for physiology and disease. *Nat Rev Genet* **9**, 764-775 (2008).
132. R. K. Barrett, J. S. Takahashi, Temperature compensation and temperature entrainment of the chick pineal cell circadian clock. *J Neurosci* **15**, 5681-5692 (1995).
133. J. Aschoff *et al.*, Human circadian rhythms in continuous darkness: entrainment by social cues. *Science* **171**, 213-215 (1971).
134. C. S. P. a. S. Daan, A Functional Analysis of Circadian Pacemakers in Nocturnal Rodents I. The Stability and Lability of Spontaneous Frequency *J. comp. Physiol.* **106**, 223-252 (1976).
135. E. D. Herzog, S. J. Aton, R. Numano, Y. Sakaki, H. Tei, Temporal precision in the mammalian circadian system: a reliable clock from less reliable neurons. *J Biol Rhythms* **19**, 35-46 (2004).
136. J. A. Mohawk, C. B. Green, J. S. Takahashi, Central and peripheral circadian clocks in mammals. *Annu Rev Neurosci* **35**, 445-462 (2012).
137. N. Koike *et al.*, Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. *Science*, (2012).
138. R. Zhang, N. F. Lahens, H. I. Ballance, M. E. Hughes, J. B. Hogenesch, A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111**, 16219-16224 (2014).
139. D. B. Boivin *et al.*, Circadian clock genes oscillate in human peripheral blood mononuclear cells. *Blood* **102**, 4143-4145 (2003).

140. M. Keller *et al.*, A circadian clock in macrophages controls inflammatory immune responses. *Proc Natl Acad Sci U S A* **106**, 21407-21412 (2009).
141. K. D. Nguyen *et al.*, Circadian Gene Bmal1 Regulates Diurnal Oscillations of Ly6Chi Inflammatory Monocytes. *Science*, (2013).
142. T. Bollinger *et al.*, Circadian clocks in mouse and human CD4+ T cells. *PLoS One* **6**, e29801 (2011).
143. A. C. Silver, A. Arjona, M. E. Hughes, M. N. Nitabach, E. Fikrig, Circadian expression of clock genes in mouse macrophages, dendritic cells, and B cells. *Brain Behav Immun* **26**, 407-413 (2012).
144. P. G. Shackelford, R. D. Feigin, Periodicity of susceptibility to pneumococcal infection: influence of light and adrenocortical secretions. *Science* **182**, 285-287 (1973).
145. J. B. Julie E. Gibbs, Stephen Beesley, Laura Matthews, Karen D. Simpson, Susan H. Boyce, Stuart N. Farrow, Kathryn J. Else, Dave Singh, David W. Ray and Andrew S. I. Loudon, The nuclear receptor REV-ERBα mediates circadian regulation of innate immunity through selective regulation of inflammatory cytokines. *PNAS* **2**, (2012).
146. M. M. Bellet *et al.*, Circadian clock regulates the host response to Salmonella. *Proc Natl Acad Sci U S A* **110**, 9897-9902 (2013).
147. T. Dickmeis, N. S. Foulkes, Glucocorticoids and circadian clock control of cell proliferation: at the interface between three dynamic systems. *Mol Cell Endocrinol* **331**, 11-22 (2011).
148. A. E. Coutinho, K. E. Chapman, The anti-inflammatory and immunosuppressive effects of glucocorticoids, recent developments and mechanistic insights. *Mol Cell Endocrinol* **335**, 2-13 (2011).
149. K. M. Lindstrom *et al.*, Feather mites and internal parasites in small ground finches (Geospiza fuliginosa, Emberizidae) from the Galapagos Islands (Equador). *J Parasitol* **95**, 39-45 (2009).
150. A. Gryczynska, O. Dolnik, T. D. Mazgajski, Parasites of Chaffinch (Fringilla coelebs) population. Part I. Coccidia (Protozoa, Apicomplexa). *Wiadomosci parazytologiczne* **45**, 495-500 (1999).
151. O. V. Dolnik, B. J. Metzger, M. J. Loonen, Keeping the clock set under the midnight sun: diurnal periodicity and synchrony of avian Isospora parasites cycle in the High Arctic. *Parasitology* **138**, 1077-1081 (2011).
152. S. M. Asio, P. E. Simonsen, A. W. Onapa, Analysis of the 24-h microfilarial periodicity of Mansonella perstans. *Parasitol Res* **104**, 945-948 (2009).
153. F. Hawking, The clock of the malaria parasite. *Scientific American* **222**, 123-131 (1970).
154. C. A. Thaiss *et al.*, Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell* **159**, 514-529 (2014).
155. A. Zarrinpar, A. Chaix, S. Yooseph, S. Panda, Diet and feeding pattern affect the diurnal dynamics of the gut microbiome. *Cell Metab* **20**, 1006-1017 (2014).
156. T. N. Siegel, K. Gunasekera, G. A. Cross, T. Ochsenreiter, Gene expression in Trypanosoma brucei: lessons from high-throughput RNA sequencing. *Trends Parasitol* **27**, 434-441 (2011).
157. F. Butter *et al.*, Comparative proteomics of two life cycle stages of stable isotope-labeled Trypanosoma brucei reveals novel components of the

parasite's host adaptation machinery. *Molecular & cellular proteomics : MCP* **12**, 172-179 (2013).
158. B. Szoor, J. R. Haanstra, M. Gualdron-Lopez, P. A. Michels, Evolution, dynamics and specialized functions of glycosomes in metabolism and development of trypanosomatids. *Curr Opin Microbiol* **22**, 79-87 (2014).
159. A. P. McLatchie *et al.*, Highly sensitive in vivo imaging of Trypanosoma brucei expressing "red-shifted" luciferase. *PLoS Negl Trop Dis* **7**, e2571 (2013).
160. N. Van Reet, H. Van de Vyver, P. P. Pyana, A. M. Van der Linden, P. Buscher, A panel of Trypanosoma brucei strains tagged with blue and red-shifted luciferases for bioluminescent imaging in murine infection models. *PLoS Negl Trop Dis* **8**, e3054 (2014).
161. P. MacGregor, B. Szoor, N. J. Savill, K. R. Matthews, Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. *Nat Rev Microbiol* **10**, 431-438 (2012).
162. P. MacGregor, K. R. Matthews, Identification of the regulatory elements controlling the transmission stage-specific gene expression of PAD1 in Trypanosoma brucei. *Nucleic Acids Res* **40**, 7705-7717 (2012).
163. P. MacGregor, N. J. Savill, D. Hall, K. R. Matthews, Transmission stages dominate trypanosome within-host dynamics during chronic infections. *Cell Host Microbe* **9**, 310-318 (2011).
164. R. J. Wheeler, E. Gluenz, K. Gull, The limits on trypanosomatid morphological diversity. *PLoS One* **8**, e79581 (2013).
165. R. J. Wheeler, K. Gull, E. Gluenz, Detailed interrogation of trypanosome cell biology via differential organelle staining and automated image analysis. *BMC Biol* **10**, 1 (2012).
166. K. M. Tyler, K. R. Matthews, K. Gull, Anisomorphic cell division by African trypanosomes. *Protist* **152**, 367-378 (2001).
167. J. L. Bargul *et al.*, Species-Specific Adaptations of Trypanosome Morphology and Motility to the Mammalian Host. *PLoS Pathog* **12**, e1005448 (2016).
168. K. Tyler, Differentiation and division of Trypanosoma brucei in the mammalian bloodstream. Manchester: University of Manchester. (1998).
169. K. Gull, The cytoskeleton of trypanosomatid parasites. *Annu Rev Microbiol* **53**, 629-655 (1999).
170. K. G. T. Nicolai Siegel, George A.M. Cross, Torsten Ochsenreiter, Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. *Trends in Parasitology* **27**, 434-441 (2011).
171. A. C. Pena *et al.*, Trypanosoma brucei histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo. *Mol Microbiol* **93**, 645-663 (2014).
172. A. Das *et al.*, The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in Trypanosoma brucei energy metabolism. *RNA* **18**, 1968-1983 (2012).
173. A. Kramer *et al.*, Regulation of daily locomotor activity and sleep by hypothalamic EGF receptor signaling. *Science* **294**, 2511-2515 (2001).
174. T. C. de Jesus *et al.*, Target of rapamycin (TOR)-like 1 kinase is involved in the control of polyphosphate levels and acidocalcisome maintenance in Trypanosoma brucei. *J Biol Chem* **285**, 24131-24140 (2010).

175. J. J. van Hellemond, A. G. Tielens, Adaptations in the lipid metabolism of the protozoan parasite Trypanosoma brucei. *FEBS Lett* **580**, 5552-5558 (2006).
176. M. R. Mugnier, G. A. Cross, F. N. Papavasiliou, The in vivo dynamics of antigenic variation in Trypanosoma brucei. *Science* **347**, 1470-1473 (2015).
177. G. J. Losos, and Ikede, B. O., Review of patholgy of diseases in domestic and laboratory animals caused by *Trypanosoma congolense*, *T. vivax*, *T. brucei*, *T. rhodesiense* and *T. gambiense. Veterinary Pathology*, pp. 1-71 (1972).
178. J. H. Fernandes, J. M. Atouguia, M. C. Peleteiro, F. W. Jennings, V. E. Rosario, Post-treatment hind-leg paralysis in mice infected with Trypanosoma brucei brucei: a light microscopic study. *Acta Trop* **63**, 179-184 (1997).
179. S. M. Ferreira AV, Menezes Z, Macedo AM, Gelape C et al., Evidence for Trypanosoma cruzi in adipose tissue in human chronic Chagas disease. *Microbes and Infection* **13**, 1002-1005 (2011).
180. O. D. Lenzi HL, Lima MT, Gattass CR, Trypanosoma cruzi: paninfectivity of CL strain during murine acute infection. *Experimental Parasitology* **84**, 16-27 (1996).
181. F. F. A. Lewis MD, Taylor MC, Burrell-Saward H, McLatchie AP et al., Bioluminescence imaging of chronic Trypanosoma cruzi infections reveals tissue-specific parasite dynamics and heart disease in the absence of locally persistent infection. *Cellular Microbiology* **16**, 1285-1300 (2014).
182. O. Neyrolles *et al.*, Is adipose tissue a place for Mycobacterium tuberculosis persistence? *PLoS One* **1**, e43 (2006).
183. T. W. Chun, S. Moir, A. S. Fauci, HIV reservoirs as obstacles and opportunities for an HIV cure. *Nat Immunol* **16**, 584-589 (2015).
184. J. B. Richardson *et al.*, Whole genome sequencing shows sleeping sickness relapse is due to parasite regrowth and not reinfection. *Evol Appl* **9**, 381-393 (2016).
185. E. Vassella *et al.*, A major surface glycoprotein of trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. *Genes Dev* **14**, 615-626 (2000).
186. I. Coppens, F. R. Opperdoes, P. J. Courtoy, P. Baudhuin, Receptor-mediated endocytosis in the bloodstream form of Trypanosoma brucei. *J Protozool* **34**, 465-473 (1987).
187. M. Ranjithkumar *et al.*, Hyperlipidaemia in trypanosomiasis of naturally infected horses: possible cachexia-anorexia syndrome? *Tropical animal health and production* **45**, 417-421 (2013).
188. B. O. Amole, M. Wittner, D. Hewlett, H. B. Tanowitz, Trypanosoma brucei: infection in murine diabetes. *Exp Parasitol* **60**, 342-347 (1985).
189. N. Ouchi, J. L. Parker, J. J. Lugus, K. Walsh, Adipokines in inflammation and metabolic disease. *Nature reviews. Immunology* **11**, 85-97 (2011).
190. M. Engstler, M. Boshart, Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei. *Genes Dev* **18**, 2798-2811 (2004).

191. J. G. Johnson, G. A. Cross, Selective cleavage of variant surface glycoproteins from Trypanosoma brucei. *The Biochemical journal* **178**, 689-697 (1979).
192. J. R. Kremer, D. N. Mastronarde, J. R. McIntosh, Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76 (1996).
193. M. C. Taylor, J. M. Kelly, Optimizing bioluminescence imaging to study protozoan parasite infections. *Trends Parasitol*, (2014).
194. F. Aresta-Branco, S. Pimenta, L. M. Figueiredo, A transcription-independent epigenetic mechanism is associated with antigenic switching in Trypanosoma brucei. *Nucleic Acids Res*, (2015).
195. S. O. Oyola *et al.*, Functional analysis of Leishmania cyclopropane fatty acid synthetase. *PLoS One* **7**, e51300 (2012).
196. A. Buguet *et al.*, [Sleeping sickness: major disorders of circadian rhythm]. *Med Trop (Mars)* **61**, 328-339 (2001).
197. D. Steverding, The history of African trypanosomiasis. *Parasit Vectors* **1**, 3 (2008).
198. A. Beschin, J. Van Den Abbeele, P. De Baetselier, E. Pays, African trypanosome control in the insect vector and mammalian host. *Trends Parasitol* **30**, 538-547 (2014).
199. L. M. Figueiredo, G. A. Cross, C. J. Janzen, Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* **7**, 504-513 (2009).
200. G. Cavadini *et al.*, TNF-alpha suppresses the expression of clock genes by interfering with E-box-mediated transcription. *Proc Natl Acad Sci U S A* **104**, 12843-12848 (2007).
201. J. M. Sternberg *et al.*, Meningoencephalitic African trypanosomiasis: Brain IL-10 and IL-6 are associated with protection from neuro-inflammatory pathology. *J Neuroimmunol* **167**, 81-89 (2005).
202. C. S. P. a. S. Daan, A Functional Analysis of Circadian Pacemakers in Nocturnal Rodents IV. Entrainment: Pacemaker as Clock. *J. comp. Physiol.*, 291-331 (1976).
203. S. M. Siepka, J. S. Takahashi, Methods to record circadian rhythm wheel running activity in mice. *Methods Enzymol* **393**, 230-239 (2005).
204. E. D. Buhr, S. H. Yoo, J. S. Takahashi, Temperature as a universal resetting cue for mammalian circadian oscillators. *Science* **330**, 379-385 (2010).
205. T. Cambras *et al.*, Circadian desynchronization of core body temperature and sleep stages in the rat. *Proc Natl Acad Sci U S A* **104**, 7634-7639 (2007).
206. S. H. Yoo *et al.*, PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc Natl Acad Sci U S A* **101**, 5339-5346 (2004).
207. P. A. Bryant, J. Trinder, N. Curtis, Sick and tired: Does sleep have a vital role in the immune system? *Nature reviews. Immunology* **4**, 457-467 (2004).
208. J. M. Krueger, F. J. Obal, J. Fang, T. Kubota, P. Taishi, The role of cytokines in physiological sleep regulation. *Ann N Y Acad Sci* **933**, 211-221 (2001).

209. G. B. Lundkvist, R. H. Hill, K. Kristensson, Disruption of circadian rhythms in synaptic activity of the suprachiasmatic nuclei by African trypanosomes and cytokines. *Neurobiol Dis* **11**, 20-27 (2002).
210. L. Marpegan, T. A. Bekinschtein, M. A. Costas, D. A. Golombek, Circadian responses to endotoxin treatment in mice. *J Neuroimmunol* **160**, 102-109 (2005).
211. Y. Kwak *et al.*, Interferon-gamma alters electrical activity and clock gene expression in suprachiasmatic nucleus neurons. *J Biol Rhythms* **23**, 150-159 (2008).
212. I. Angulo, M. Fresno, Cytokines in the pathogenesis of and protection against malaria. *Clin Diagn Lab Immunol* **9**, 1145-1152 (2002).
213. A. C. Liu *et al.*, Intercellular coupling confers robustness against mutations in the SCN circadian clock network. *Cell* **129**, 605-616 (2007).
214. L. Maclean, H. Reiber, P. G. Kennedy, J. M. Sternberg, Stage Progression and Neurological Symptoms in Trypanosoma brucei rhodesiense Sleeping Sickness: Role of the CNS Inflammatory Response. *PLoS Negl Trop Dis* **6**, e1857 (2012).
215. K. Urech, A. Neumayr, J. Blum, Sleeping sickness in travelers - do they really sleep? *PLoS Negl Trop Dis* **5**, e1358 (2011).
216. S. M. Siepka *et al.*, Circadian mutant Overtime reveals F-box protein FBXL3 regulation of cryptochrome and period gene expression. *Cell* **129**, 1011-1023 (2007).
217. P. L. Lowrey, J. S. Takahashi, Genetics of circadian rhythms in Mammalian model organisms. *Adv Genet* **74**, 175-230 (2011).
218. O. World Health, Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. *World Health Organization technical report series*, v-xii, 1-100 (2012).
219. M. W. Young, S. A. Kay, Time zones: a comparative genetics of circadian clocks. *Nat Rev Genet* **2**, 702-715 (2001).
220. M. Rosbash, The implications of multiple circadian clock origins. *PLoS Biol* **7**, e62 (2009).
221. D. Bell-Pedersen *et al.*, Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet* **6**, 544-556 (2005).
222. A. M. van der Linden *et al.*, Genome-wide analysis of light- and temperature-entrained circadian transcripts in Caenorhabditis elegans. *PLoS Biol* **8**, e1000503 (2010).
223. M. E. Hughes, G. R. Grant, C. Paquin, J. Qian, M. N. Nitabach, Deep sequencing the circadian and diurnal transcriptome of Drosophila brain. *Genome research* **22**, 1266-1281 (2012).
224. A. B. Reddy, G. Rey, Metabolic and nontranscriptional circadian clocks: eukaryotes. *Annu Rev Biochem* **83**, 165-189 (2014).
225. C. E. Clayton, Life without transcriptional control? From fly to man and back again. *EMBO J* **21**, 1881-1888 (2002).
226. G. Z. Wang *et al.*, Cycling Transcriptional Networks Optimize Energy Utilization on a Genome Scale. *Cell reports* **13**, 1868-1880 (2015).
227. F. Mignone, C. Gissi, S. Liuni, G. Pesole, Untranslated regions of mRNAs. *Genome Biol* **3**, REVIEWS0004 (2002).
228. C. Clayton, M. Shapira, Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* **156**, 93-101 (2007).

229. A. C. Nelson, F. C. Wardle, Conserved non-coding elements and cis regulation: actions speak louder than words. *Development* **140**, 1385-1395 (2013).
230. J. Bass, J. S. Takahashi, Circadian integration of metabolism and energetics. *Science* **330**, 1349-1354 (2010).
231. A. M. Curtis, M. M. Bellet, P. Sassone-Corsi, L. A. O'Neill, Circadian Clock Proteins and Immunity. *Immunity* **40**, 178-186 (2014).
232. H. Hirumi, K. Hirumi, Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J Parasitol* **75**, 985-989 (1989).
233. S. Knusel, I. Roditi, Insights into the regulation of GPEET procyclin during differentiation from early to late procyclic forms of Trypanosoma brucei. *Mol Biochem Parasitol* **191**, 66-74 (2013).
234. M. P. Cox, D. A. Peterson, P. J. Biggs, SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC bioinformatics* **11**, 485 (2010).
235. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
236. M. Lawrence *et al.*, Software for computing and annotating genomic ranges. *PLoS computational biology* **9**, e1003118 (2013).
237. P. A. H. Pages, R. Gentleman and S. DebRoy, in *R package version 2.38.4.*
238. M. Lawrence, R. Gentleman, V. Carey, rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* **25**, 1841-1842 (2009).
239. B. B. Gregory R. Warnes, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber, T. L. Andy Liaw, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc, S. a. B. Venables, in *R package version 2.17.0.* . (2015).
240. S. Wichert, K. Fokianos, K. Strimmer, Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics* **20**, 5-20 (2004).
241. M. E. Hughes, J. B. Hogenesch, K. Kornacker, JTK_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *J Biol Rhythms* **25**, 372-380 (2010).
242. R. Yang, Z. Su, Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation. *Bioinformatics* **26**, i168-174 (2010).
243. B. G. Kurt Hornik, movMF: An R Package for Fitting Mixtures of von Mises-Fisher Distributions. *Journal of Statistical Software*, 1-31 (2014).
244. G. Pavesi *et al.*, MoD Tools: regulatory motif discovery in nucleotide sequences from co-regulated or homologous genes. *Nucleic Acids Res* **34**, W566-570 (2006).
245. L. Ettwiller, B. Paten, M. Ramialison, E. Birney, J. Wittbrodt, Trawler: de novo regulatory motif discovery pipeline for chromatin immunoprecipitation. *Nat Methods* **4**, 563-565 (2007).
246. T. L. Bailey, DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653-1659 (2011).
247. T. L. B. a. C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second*

International Conference on Intelligent Systems for Molecular Biology, 28-36 (1994).
248. O. Elemento, N. Slonim, S. Tavazoie, A universal framework for regulatory element discovery across all genomes and data types. *Mol Cell* **28**, 337-350 (2007).
249. S. Luehr, H. Hartmann, J. Soding, The XXmotif web server for eXhaustive, weight matriX-based motif discovery in nucleotide sequences. *Nucleic Acids Res* **40**, W104-109 (2012).
250. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017-1018 (2011).
251. S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257-258 (2007).
252. G. B. Lundkvist *et al.*, Clock gene expression during chronic inflammation induced by infection with Trypanosoma brucei brucei in rats. *J Biol Rhythms* **25**, 92-102 (2010).
253. A. Darsaud *et al.*, Clinical follow-up in the rat experimental model of African trypanosomiasis. *Exp Biol Med (Maywood)* **228**, 1355-1362 (2003).
254. M. E. Hughes *et al.*, Harmonics of circadian gene transcription in mammals. *PLoS Genet* **5**, e1000442 (2009).
255. W. G. Pembroke, A. Babbs, K. E. Davies, C. P. Ponting, P. L. Oliver, Temporal transcriptomics suggest that twin-peaking genes reset the clock. *eLife* **4**, (2015).
256. S. N. Archer *et al.*, Mistimed sleep disrupts circadian regulation of the human transcriptome. *Proc Natl Acad Sci U S A* **111**, E682-691 (2014).
257. J. Hoffmann *et al.*, Non-circadian expression masking clock-driven weak transcription rhythms in U2OS cells. *PLoS One* **9**, e102238 (2014).
258. N. Panchy *et al.*, Prevalence, evolution, and cis-regulation of diel transcription in Chlamydomonas reinhardtii. *G3* **4**, 2461-2471 (2014).