Transcriptional Regulation of Virulence Gene Families in *Plasmodium falciparum*

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Kathrin Witmer
aus Langendorf (SO)

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von
Dr. Till S. Voss und Dr. Tim-Wolf Gilberger

Basel, den 22. Februar 2011

Prof. Dr. Martin Spiess
Dekan

Table of contents

Acknowledgements

I want to give a big thank-you to Till Voss for making it possible for me to perform my thesis in his group. It was exactly what I wanted to do. It was a very nice time for me, and I think, we all grew a little bit older and learned a lot during the last four years. It was a pleasure to work with you!

I am grateful to Tim Gilberger, for taking all the long way from Canada to be part of my thesis committee in Basel.

A very big thank you goes to Hans-Peter Beck, who infected me with the work on *Plasmodium*, and for being a brilliant travelling mate, I enjoyed it a lot!

I also want to thank all the former and past GR group members, who created a lovely working and fun atmosphere in the lab as well as outside. Chrigu Flück for helping with a lot of lab things and thesis stuff and for being so great; Igor (always hungry) Niederwieser, for being one of the smartest guys I know and being a priceless help for all sorts of experiments; Niggi MB Brancucci for sharing this project with me, I am convinced it would not have come to an end without you and I really enjoyed our daily “discussions” which always turned into a fight; Sophie Oehring for being well-organized and fighting with me against the chaos in the lab =); Johanna Wetzel (many nice coffee breaks =)), Nicole Bertschi (you should sleep more =)), Andi (unändlech) Pulver and Sandra Birrer.

I want to give props to the Beck group members, who are part of the family. Sebi Rusch, this is the second time for you being in my acknowledgements, and you are still very clever =) but not head-to-head anymore, which is a bit of a shame; Dania Müller for sharing qPCR myths, the newest gossip, recipes and for sharing nice years; Esthi Pachlatko for sharing the PhD-time with me, I loved to share it with you, and endless discussions in the culture lab about parasites and the future; Cristian Köpfli (the thinnest moustache ever!); Annette Gaida ; Oli Dietz (the yeast-whisperer); Claudia List ; Anna Perchuc; Caro Kulangara; Marie Ballif; Pax Masimba; Ingrid Felger; Olivia Rudin; Miri Moser; Sandra Brenneisen; Felista Mwingira; Mark Finlayson; Sarah Javati, Christof Grüring, Jessie Hug, Melanie Hug, Simone Edelmann, Patrick Seitz, Sämi Lüdin, Janet Gare, Damien Jacot.

Thanks go also to other members at the SwissTPH for creating a lovely atmosphere: Yvette Endriss (food-supplier); Christian Scheuri (the man who talks to parasites); Urs Duthaler (the wormator); Ralf Brunner (the hard-core-experiment-planner); Philipp Ludin, Kathrin Ingram, Theresia Manneck, Lucienne Tritten and Matthias Fügi (the office-people); Eva Greganova, Sergio Wittlin; Matze Rottmann (you are NOT forgotten this time!)

A special thank you goes to Christoph Schmid who helped me a lot with statistic analysis and for showing me that you can see things from a different angle all the time.

I also want to thank Zbynek Bozdech and Peter Preiser at the Nanyang Technological University for making it possible for me to travel to Singapore and to perform microarrays in their lab. I learned a lot and enjoyed it even more! I want to thank the people in the NTU lab for having a nice time and for teaching me a lot about Chinese/Malay/Indian/Singapori food! I eat a lot! =) Sachel Mok, Hoonie Lua, Kek Yee Liong, Bernardo Foth, Eng How Lim, Ramya Ramadoss, Sabna Cheemadan and all I forgot. I also want to thank specially Conny Hunke for taking care of me and to providing me with a bed and tips and tricks, and Phat Vinh Dip for showing me around Singapore!

I also like to thank the SSTMP and the experimentelle Stiftung für Zoologie, especially Reto Brun, for making it possible for me to travel to several continents and participate at outstanding conferences. Thank you!

I also like to thank again the Kanton Solothurn for supporting me financially during my undergraduate studies.

Last but not least, I like to thank Colette and Urs Witmer over all. This work is dedicated to Suzanne and Hector Schmassmann.

Summary

To date, malaria caused by *Plasmodium falciparum* is still a major health threat. It contributes to illness and severe disease and is responsible for up to one million deaths per year. The intra-erythrocytic asexual life cycle stage is responsible for the pathology associated with malaria. The major virulence factor *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) is exposed at the surface of infected red blood cells (iRBC) and mediates binding to endothelial cells. This leads to sequestration of iRBC in the microvasculature and consequently to evasion of removal in the spleen. PfEMP1 is encoded by the 60-member *var* gene family, which undergoes antigenic variation by *in-situ* switching. Importantly, *var* genes are expressed in a mutually exclusive way, such that only one member is expressed whereas all other copies remain silenced. *var* genes as well as other gene families such as *rif*, *stevor*, *phist* and *pfmc-2tm* are located in subtelomeric heterochromatic regions. The function of these additional families is largely unknown, but they are thought to be implicated in host-parasite interactions and to contribute to antigenic variation.

With this work, I provide deeper insights into the transcriptional regulation of virulence gene families in *P. falciparum* by using transfection-based approaches. We functionally identified autonomous *cis*-acting *var* promoter elements including an upstream activating sequence that is essential for promoter activation. Notably, an element downstream of the transcriptional start site determines mutually exclusive locus recognition. Further, I used comparative transcriptional profiling to show that mutually exclusive expression is restricted to the *var* gene family and is not used in the transcription of other subtelomeric gene families. I show for the first time that knock-down of endogenous *var* gene transcription is also conferred by promoters of a *var* gene subfamily that is implicated in severe malaria. Taken together, this work provides important insight into the mechanisms involved in the regulation of virulence gene families and antigenic variation in *P. falciparum*. Moreover, the findings presented here are consistent with a novel mechanism of mutually exclusive gene choice in eukaryotes.

Introduction

Malaria

Malaria is a disease caused by Apicomplexan parasites of the genus *Plasmodium*, commonly occurring in Sub-Saharan Africa, Asia and South America. Infectious forms of this protozoan parasite are transmitted by female *Anopheles* mosquitoes and cause morbidity in humans, monkeys, rodents and birds. Four species of *Plasmodium* can infect humans: *P. vivax*, *P. ovale*, *P. malariae* and *P. falciparum*, which causes the most severe form of human malaria. However, it has been reported that *P. knowlesi*, a monkey malaria parasite, can also infect humans (Singh et al, 2004). Moreover, human *P. falciparum* malaria is of western gorilla origin and not of ancient human origin (Liu et al, 2010).

About 3.3 billion people - half of the world's population - are at risk of malaria. Every year, this leads to about 250 million malaria cases and in 2008, malaria caused almost 1 million deaths (WHO, 2009). In areas of intense transmission the main burden of disease is carried by young children (0-5 years) in sub-Saharan Africa. It affects also other risk groups including pregnant women and non-immune adults travelling to endemic areas.

People living in endemic areas which are subjected to repeated infections do not gain sterile immunity. They develop a so called semi-immunity, manifesting in asymptomatic infections with low numbers of parasites in the blood.

Clinical manifestations of malaria include severe anaemia, respiratory distress in relation to metabolic acidosis and cerebral malaria. Multi-organ involvement is also frequent (MacPherson et al, 1985; Carlson et al, 1990; Berendt et al, 1994; Rowe et al, 1995; Pongponratn et al, 1991; Montgomery et al, 2007). Resistance to anti-malarial drugs is emerging and so far, effective vaccines against the disease are not available. Several vaccines are in the test pipeline in preclinical studies up to phase 2b trials (WHO, vaccine projects 2010). The most promising vaccine candidate against *Plasmodium falciparum*, RTS,S, is in phase 3 trial and vaccine efficacy has been reported between 11 and 71% (Casares et al, 2010; Olotu et al, 2011).

Plasmodium falciparum has a complex life cycle which includes two hosts, humans and the female *Anopheles* mosquito. The life cycle is complex and characterised by extracellular forms and intracellular stages infecting different cell types. This implies the transformation of the parasite into different morphological forms and developmental stages (Figure 1).

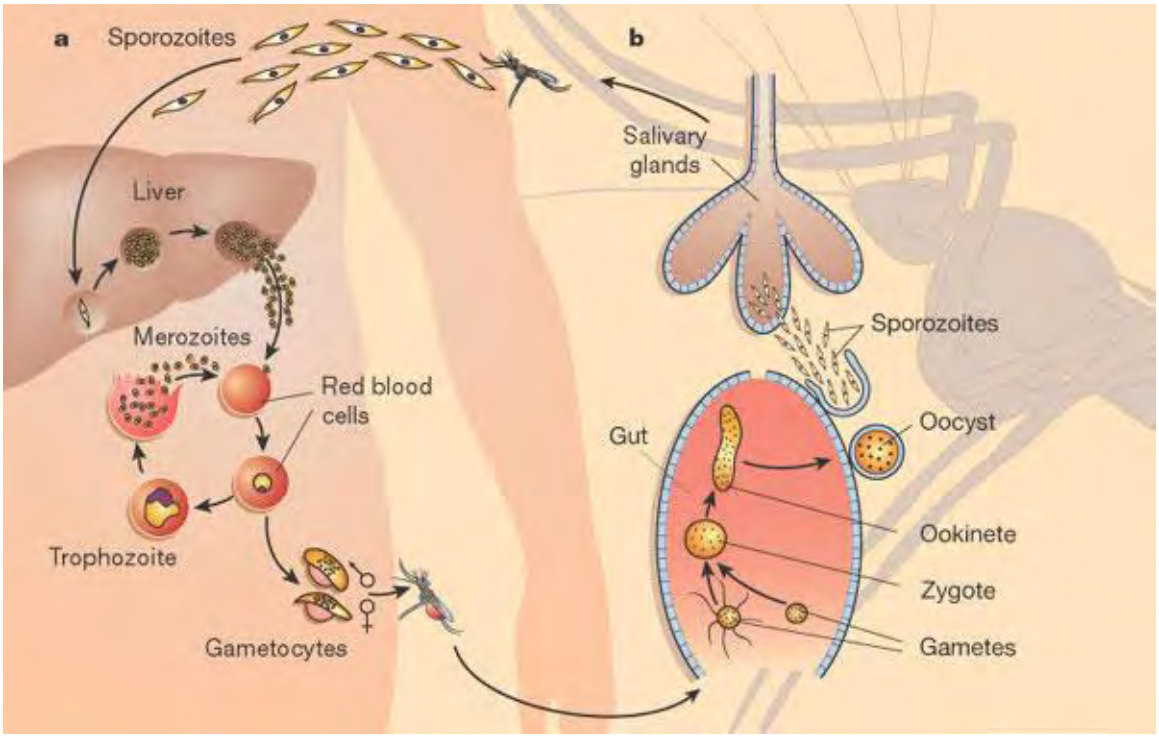


Figure 1. The life cycle of *Plasmodium* (Wirth, 2002)

Shortly, sporozoites are injected into the human skin tissue during the meal of an infected mosquito. These infective stages migrate via the blood stream to the liver and invade hepatocytes where they develop into a liver schizont containing thousands of merozoites. After seven to twelve days the host cell bursts and releases the merozoites into the blood stream where they invade red blood cells (RBC) (Mazier et al, 1985; Roestenberg et al, 2009). Within the RBC, the parasite undergoes another round of asexual multiplication within a 48 hour cycle. The intra-erythrocyte stages are named ring stage, trophozoite and schizont. The bursting of a mature schizont leads to the release of up to 32 merozoites which invade new RBCs to start the cycle again. Importantly, all clinical symptoms of malaria are associated with the intra-erythrocytic cycle. A subset of infected red blood cells develops into

gametocytes, which are infectious to mosquitoes and once taken up, undergo sexual reproduction and meiosis, and further development into ookinetes, oocysts and finally sporozoites.

Host-parasite interactions

The blood stages of *Plasmodium falciparum* are able to evade antibody–mediated host immunity by altering the profile of parasite-encoded antigens exposed at the erythrocyte surface (Biggs et al, 1991; Roberts et al, 1992; Smith et al, 1995; Kyes et al, 2001). The generation of this antigenic variation is the result of the amplification of extensive repertoires of hypervariable subtelomeric multigene families encoding erythrocyte surface proteins, coupled with the ability of parasites to switch the expression of individual genes or subsets of genes within these families. Furthermore, high recombination rates facilitate their evolution and increase diversity (Scherf et al., 2004). The extent of hypervariability within these multigene families suggests that the antigenic diversity on the surface of the infected red blood cell (iRBC) is virtually unlimited (Lavazec et al, 2006; Trimnell et al, 2006). The *var* gene family encodes the most extensively studied family of erythrocyte surface antigens, termed *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). On one hand, this major virulence factor undergoes antigenic variation resulting in immune evasion and chronic infection. On the other hand, members of this large variant family mediate adhesion of iRBCs to a number of host receptors, including ICAM-1 (intracellular adhesion molecule 1), CD36 (platelet glycoprotein IV), complement receptor-1 (CR1) and chondroitin sulphate A (CSA) (Baruch et al, 2002; Kyes et al, 2001; Rogerson & Brown, 1997; Rowe et al, 1997). The adhesion process sequesters iRBCs away from the peripheral circulation and prevents phagocytic clearance in the spleen. Accumulation of infected erythrocytes within the microvasculature in vital organs such as the brain and the placenta leads to complications such as cerebral malaria which is often fatal, and placental malaria which predisposes to fetal and maternal death (Kraemer & Smith, 2006; Rowe & Kyes, 2004). Repeated exposure to multiple infections with *P. falciparum* leads to the development of semi-immunity that is partially based on immune recognition of PfEMP1 and other antigens. As a

consequence, severe disease is seen less commonly in adults; however, during pregnancy women are again susceptible to disease due to adhesion of iRBCs to a novel population of binding sites (CSA) in the placenta (Reeder et al, 1999; Fried & Duffy, 1996, 1998).

PfEMP1 and the *var* gene family

Genomic organization of the *var* gene family

In the sequenced *Plasmodium falciparum* strain 3D7, 56 *var* genes plus three small *var*-like genes are found and most of them are located subtelomerically (Gardner et al, 2002). Each chromosome end typically contains one to three *var* genes, followed by members of other multigene families; chromosome 14 is an exception with only two *var* pseudo-genes located close to each telomere (Gardner et al, 2002). Many subtelomeric regions have two *var* genes arranged in tail-to-tail orientation relative to each other with one or more *rif* genes located in between. Chromosome-central *var* genes can appear singly or in groups that are nearly always tandem arrays (head-to-tail) containing three up to seven *var*.

var genes are preceded by a 5' non-coding sequence which is linked to its chromosomal location and transcriptional orientation (Gardner et al, 2002; Voss et al, 2000; Lavstsen et al, 2003). Based upon sequence similarity, the 5' promoter regions can be defined into three major upstream (ups) sequence groups (upsA, upsB, upsC) (Gardner et al, 2002; Kraemer et al, 2007; Lavstsen et al, 2003). Interestingly, in all three sequenced parasite strains (3D7, HB3, IT4) upsC *var* genes are chromosome-central, upsB *var* genes are predominantly subtelomeric and transcribed away from the telomere and some are also chromosome-central in tandem arrays with other upsB or upsC *var* genes, and upsA *var* genes are subtelomeric but transcribed towards the telomere in opposite direction to upsB *var* genes.

Furthermore, three unusual semi-conserved *var* genes (*var1csa*, *var2csa* and Type3 *var*) have been described which have unique features and are different from all other *var* genes (Kraemer et al, 2007; Lavstsen et al, 2003). The *var2csa* gene is flanked by the unique upsE upstream region (Lavstsen et al, 2003) and is the dominant *var* gene transcribed in CSA-binding parasites and

is also transcribed at high levels in parasites isolated from placentas (Salanti et al, 2003; Kyes et al, 2003). It plays a critical role in the pathogenesis of pregnancy-associated malaria (Salanti et al, 2004). However, no function has yet been ascribed to the proteins encoded by *var1csa* and Type3 *var*.

PfEMP1 protein structure

PfEMP1 is a highly polymorphic protein that varies greatly in size (200-350kDa) and amino acid sequence (Su et al, 1995). However, all *var* genes share a similar gene organization where two exons are separated by an intron of approximately 1 kb (Su et al, 1995; Smith et al, 1995; Baruch et al, 1995). Exon I encodes the highly variable extracellular part of PfEMP1, which is extremely variable in both sequence and length, although it consists of a few fundamental building blocks put together with some minimal rules (Smith et al, 2000): a short region of the N-terminal segment (NTS) contains sequence features sufficient for transport beyond the parasitophorous vacuole that surrounds the intra-erythrocytic parasite (Hiller et al, 2004; Marti et al, 2004). The remainder of the extracellular part consists of two main adhesion domains; Duffy binding-like (DBL; classes α-ε) and cysteine-rich interdomain regions (CIDR; classes α-γ). PfEMP1 is anchored to parasite-induced 'knob' structures underneath the infected RBC membrane through the conserved acidic terminal segment (ATS) that is encoded by the second exon. The extracellular portion of PfEMP1 is responsible for the adherence of iRBCs to endothelial cells or uninfected erythrocytes via interactions with various host surface receptors (Gardner et al, 1996; Reeder et al, 1999; Rowe et al, 1997; Baruch et al, 1996). PfEMP1 was also shown to play central roles in rosetting (the binding of infected to uninfected erythrocytes) (Rowe et al, 1997). Consequently, iRBCs and cell aggregates sequester to microvasculatory capillaries of various organs, resulting in complications that significantly contribute to the severe morbidity and mortality associated with malaria tropica (MacPherson et al, 1985; Carlson et al, 1990; Pongponratn et al, 1991; Berendt et al, 1994; Rowe et al, 1995).

A further role has been assigned to PfEMP1 in modulating host immune cell activity. While dendritic cell modulation by PfEMP1 remains a controversial issue (Elliott et al, 2007; Urban et al, 1999) the down-regulation of IFN-γ

production by γδ-T cells and natural killer cells has recently been demonstrated to be mediated by PfEMP1 (D'Ombrain et al, 2007).

Antigenic variation of PfEMP1

var genes are expressed in a mutually exclusive manner, i.e. only one member is expressed in a single parasite (Scherf et al, 1998; Chen et al, 1998). Switching of *var* gene expression occurs through *in situ* transcriptional activation of a previously silenced copy (Scherf et al, 1998) and this is the basis for antigenic variation of PfEMP1. This mechanism is fundamentally different from those employed by other pathogens such as African trypanosomes and *Borrelia hermsii* where antigenic variation of VSGs and VMPs, respectively, is mainly achieved by DNA recombination events (gene conversion, homologous recombination, reciprocal exchange), translocating silent copies into an active expression site (Deitsch et al, 2009). Understanding the mechanisms of *var* gene switching, and analysing switching rates, has been problematic because of the extreme difficulty of monitoring switching *in vivo*. This made it impossible to determine if the reported *in vitro* rates of antigenic variation in *P. falciparum* (Roberts et al, 1992; Horrocks et al, 2004; Frank et al, 2007) (Witmer, MSc thesis 2006) reflect those in a natural infection. Over the last years many studies investigated the switching rates of *var* genes for both *in vitro* and *in vivo* and resulted in different findings. An *in vitro* study using *P. falciparum* line IT 4/25/5 showed that antigenic variation results in ~2% of the parasite population switching away from the original antigenic type per generation (Roberts et al, 1992). Switching away from a single *var* gene occurs without switching preferences for any subtype and the switching rate is dissimilar for different variants (Horrocks et al, 2004). In contrast, no switching was observed during six months of continuous *in vitro* culture of unselected 3D7 parasites (Witmer, Master Thesis 2006). A longitudinal study by Kaestli et al. (Kaestli et al, 2004) described *var* gene expression in naturally infected semi-immune children from Papua New Guinea during a 4-month period. Many *var* genes were transcribed simultaneously in the infecting population and switched at short intervals, but some identical transcripts recurred in the same child, even after 10 weeks. Lavstsen et al. (Lavstsen et al, 2005) analyzed

samples from volunteers infected with the *P. falciparum* strain NF54 and harvested parasites on day 8 post infection and cultured them for another 27 or 33 days *in vitro*. These cultures showed transcripts of all *var* genes and most were transcribed at roughly similar levels. Interestingly, nine of the 10 lowest transcribed genes belonged to *var* group A or B/A, which have been associated with severe malaria (Jensen et al, 2004).

Frank et al. (Frank et al, 2007) showed that an unselected population of NF54 parasites preferentially expresses central *var* genes and central loci display low off rates ranging from 0 to 0.3% per generation. Conversely, subtelomeric wild–type and transgenic *var* loci exhibited higher off rates of at least 1-2%, indicating that *var* genes in subtelomeric locations may be subject to additional constraints (Frank et al, 2007). Montgomery et al. (Montgomery et al, 2007) analyzed *var* gene transcription by *P. falciparum* parasites from the brain, lung, heart and spleen of samples taken from patients with fatal pediatric malaria. Despite detecting expression of up to 102 different *var* genes in a single host, the authors found that only one or two of these genes were expressed at high levels within the brain and heart tissue of these patients. These results provide the first evidence of organ-specific accumulation of *P. falciparum* variant types and suggest that parasitized erythrocytes can exhibit preferential binding in the body, supporting the hypothesis of cytoadherence-linked pathogenesis.

It has been observed that *var* gene expression and switching behave differentially *in vitro* and *in vivo* (Peters et al, 2007; Bachmann et al, 2009). Short term *in vitro* cultivation of patient samples leads to a reduction of the overall abundance of *var* transcripts (Peters et al, 2007) and *P*. *falciparum* infecting a splenectomized patient showed no *var* gene transcripts (Bachmann et al, 2009), undermining that *var* gene switching and *var* expression is also strongly linked to the infected patient and a severe cross-talk between the immune system of the host and the parasite

Transcription and epigenetic aspects of the *P. falciparum* genome

In *P. falciparum* transcriptional initiation of protein coding genes, as in all eukaryotes, involves the RNA polymerase II complex (RNAPII) assembled within a larger pre-initiation complex (PIC), which contains additional general

transcription factors (GTF). Bioinformatics approaches unveiled key components of the PIC and a range of GTFs by identifying first or secondary structure conservation (Coulson et al, 2004; Callebaut et al, 2005). Comparing 5'UTRs of *P. falciparum* to other eukaryotes revealed that they are unusually long, suggesting a role in mRNA stability (Horrocks et al, 2009). The apparent paucity of *P. falciparum* in specific transcription factors was overcome by the discovery of a lineage-specific family named ApiAP2, related to plant transcription factors (Balaji et al, 2005). The ApiAP2 family is present in all Plasmodia, as well as in all other apicomplexans investigated so far, with 26 members in *P. falciparum* (De Silva et al, 2008). One member, PfSIP2, has been shown to interact specifically with SPE2 elements in subtelomeric regions, and putative binding sites cluster within telomere-associated repeat sequences and subtelomeric *var* upstream sequences (Flueck et al, 2010).

A range of epigenetic molecular marks have been found in *P. falciparum* (Salcedo-Amaya et al, 2009; Miao et al, 2006), with the exception of reversible DNA methylation (Choi et al, 2006). Further, the RNAi machinery is found to be absent (Baum et al, 2009). Moreover, *P. falciparum* lacks histone 1 and contains four different histone variants (H2A.Z, H2Bv, H3.3 and CenH3) (Miao et al, 2006).

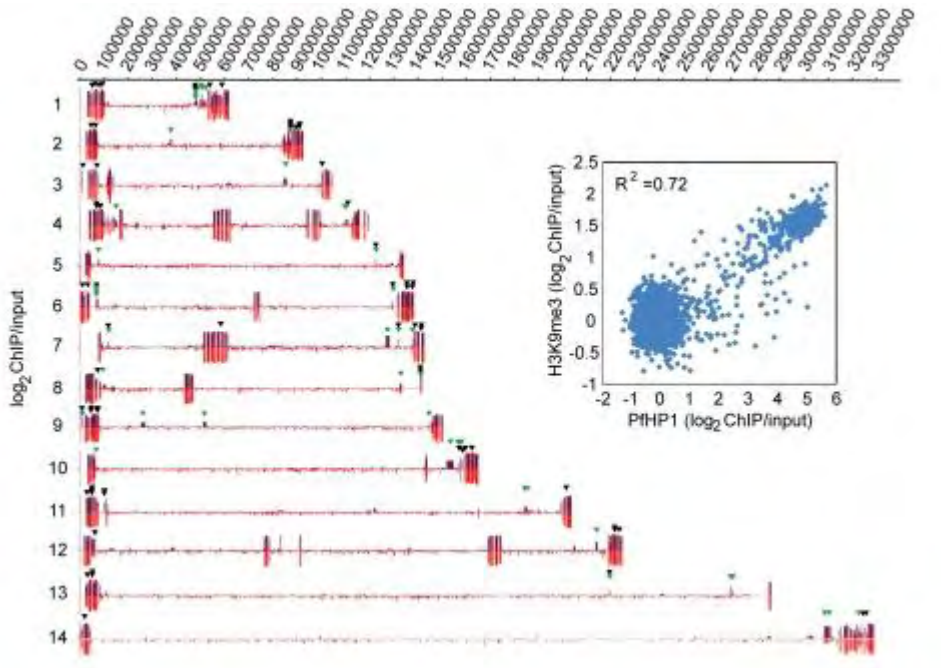


Figure 2. Colocalization of PfHP1 and H3K9me3. Purple bars represent the genome-wide distribution of PfHP1 and red bars indicate H3K9me3, respectively (taken from Flueck et al., 2009).

Genome-wide chromatin immunoprecipitation showed the association of PfHP1 (*P. falciparum* heterochromatin protein 1) with histone 3 lysine 9 trimethylation (H3K9me3) marks (Pérez-Toledo et al, 2009; Flueck et al, 2009). PfHP1/H3K9me3 cover subtelomeric and chromosome-internal virulence gene clusters representing about 10% of the genome (Fig. 2) (Flueck et al, 2009). In analogy to model organisms genes in PfHP1/H3K9me3-enriched domains are transcriptional silenced (Flueck, 2009; Lopez-Rubio, 2009; Salcedo-Amaya, 2009). The rest of the genome, however, is dominated by histone modifications that are associated with transcriptionally active states of genes in yeast and higher eukaryotes (H3K4me3, H3K9ac, H3K14ac and H4ac) (Trelle et al, 2009; Salcedo-Amaya et al, 2009). The histone variant H2A.Z has been shown to localize to euchromatic intergenic regions (Bártfai et al, 2010).

Transcriptional control of the *var* gene family

The regulatory mechanisms of *var* gene activation, silencing and allelic exclusion are highly complex and under intense investigation in several laboratories. *In vitro* culturing of parasites in the absence of immune selection leads to antigenically heterogenous populations; hence *var* gene transcription is difficult to investigate. Northern blot analysis of mRNA isolated from parasite populations selected for binding of iRBCs to specific host endothelial receptors (e.g. CSA, ICAM-1) detect a single *var* transcript only, and these transcripts encode the PfEMP1 variant mediating the binding to their relevant receptor (Taylor et al, 2000a; Kyes et al, 2003; Salanti et al, 2003; Horrocks et al, 2004; Duffy et al, 2005). Recently, Kyes et al. demonstrated that mutually exclusive control of *var* gene expression occurs at the level of transcription initiation during ring stages and is most probably mediated by RNA polymerase II (Kyes et al, 2007). Furthermore, mutually exclusive expression of *var* genes depends solely on non-coding elements at each *var* gene and is

independent of production of a functional PfEMP1 protein (Dzikowski et al, 2006; Voss et al, 2006).

The conserved 5' sequences of *var* genes have promoter activity since they are able to mediate transcription of episomal reporter genes in transient transfection experiments (Voss et al, 2000; Deitsch et al, 1999). *var* gene promoters are regulated by epigenetic mechanisms and upsC and upsB promoters are required to maintain chromosome-internal and subtelomeric *var* genes, respectively, in their silenced default state (Voss et al, 2006, 2007; Dzikowski et al, 2006). Furthermore, regulatory *cis*-acting *var* promoter elements interact with sequence-specific DNA-binding proteins found in the parasite nucleus (Voss et al, 2003). Another regulatory role has been assigned to the *var* gene intron, which was shown to be involved in upsC promoter silencing (Deitsch et al, 2001). This finding has been confirmed (Dzikowski et al, 2006; Voss et al, 2006; Calderwood et al, 2003; Gannoun-Zaki et al, 2005; Frank et al, 2006) but the underlying mechanisms, however, remain obscure. These results established that each *var* gene locus is associated with functional regulatory elements. Interestingly, Voss et al. (Voss et al, 2007) recently demonstrated that two episomal *var* gene promoters *in cis* are activated simultaneously, implying that mutual exclusion is not based on single promoter competition but rather on locus activation. *var1csa* is unusual in that it is transcribed constitutively in all parasites, even as a truncated pseudogene (PFE1640w in 3D7), and thus falls outside the controls of mutually exclusive gene expression (Kyes et al, 2003).

Epigenetic regulation of *var* gene expression

Reversible histone-tail modifications and ATP-dependent nucleosome remodelling are important for epigenetic control in eukaryotic gene regulation. Indeed, activation and silencing of *var* genes correlates with specific histone tail marks, suggesting that epigenetic memory is involved in maintenance and switching of the transcriptional state of the *var* gene repertoire. H3K9ac and H3K4me have been shown to be associated with *var* gene activation (Lopez-Rubio et al, 2007), whereas H3K9me3 is associated with *var* gene silencing (Lopez-Rubio et al, 2007; Chookajorn et al, 2007).

The histone deacetylase PfSIR2, a homologue of the yeast silent information regulator, has been shown to associate with silenced *var* genes and localise to chromosome-end clusters at noncoding telomere-associated repeat elements (TARE 1–6) and telomere repeats (Mancio-Silva et al, 2008; Freitas-Junior et al, 2005). *P. falciparum* possesses two SIR2 paralogues, SIR2A and SIR2B (Tonkin et al, 2009). Knock-out of the PfSIR2A gene leads to de-repression of *var* genes transcribed towards the telomere and those in internal chromosomal clusters (i.e., controlled by UpsA, UpsE, and UpsC, respectively) (Merrick et al, 2010; Duraisingh et al, 2005; Tonkin et al, 2009; Tham et al, 2007). This holds true for endogenous *var* genes as well as episomal reporter constructs whose expression is driven by a *var* promoter (Tonkin et al, 2009). In contrast, the most telomere proximal *var* genes (i.e., controlled by an UpsB promoter) were highly activated in ΔPfSir2B (Tonkin et al, 2009). As in yeast, *P. falciparum* telomeres form physical clusters that are anchored to the nuclear periphery (Freitas-Junior et al, 2000) and form an average of four to seven clusters per nucleus which appear to be cross-linked by protein (Marty et al, 2006). As a consequence, subtelomeric *var* genes inherently assume a perinuclear location, and it was recently shown that central *var* clusters are also positioned at the nuclear periphery, independent of their transcriptional state (Ralph et al, 2005; Voss et al, 2006; Marty et al, 2006). In yeast and other eukaryotes the nuclear periphery is associated with enhanced transcriptional silencing (Gasser, 2001). Similarly, Duraisingh et al. (Duraisingh et al, 2005) demonstrated that a subtelomeric transgene, as well as the subtelomeric *var2csa* gene, were reversibly silenced in *P. falciparum* and subject to position effect variegation, a phenomenon commonly observed for genes in or close to heterochromatic regions. Alterations in the transcriptional states of both genes were found to involve perinuclear locus repositioning. It is currently being discussed if the activation of a *var* gene involves its migration outside of a cluster or not (Duraisingh et al, 2005; Freitas-Junior et al, 2005; Marty et al, 2006; Lopez-Rubio et al, 2009).

Interference of activated episomal *var* gene promoters with endogenous *var* gene transcription

Stable transfection experiments with episomal *var* gene promoters driving expression of a drug selectable marker gene proved to be a powerful strategy to investigate *var* transcriptional control mechanisms. Upon transfection, *var* promoters are silenced by default (Calderwood et al, 2003; Frank et al, 2006; Voss et al, 2006, 2007). Multiple evidence indicate that the silenced state is only efficiently induced if the promoter region is paired in *cis* with the *var* gene intron (Dzikowski et al, 2007; Epp et al, 2008; Voss et al, 2006). However, drug-induced selection for active *var* promoters showed that the upstream sequence alone is sufficient to infiltrate a transgene into the mechanism of mutually exclusive *var* gene transcription (Voss et al, 2006, 2007; Dzikowski et al, 2006). Consequently, an active episomal *var* promoter driving expression of a drug-selectable marker causes a knock-down of the endogenous *var* gene family and this is dependent solely on the regulatory information contained within the promoter sequence.

Telomeres and therefore also subtelomerically positioned *var* genes cluster together at the nuclear periphery (Freitas-Junior et al, 2000). Additionally, chromosome internal *var* genes are also found at the nuclear periphery (Ralph et al, 2005; Voss et al, 2006). Generally heterochromatic, the nuclear periphery is thought to contain a specialized *var* expression site within a euchromatic zone (Ralph et al, 2005). Consistent with this hypothesis, activation of *var* genes occurs in this perinuclear compartment and was shown to be accompanied by locus repositioning (Duraisingh et al, 2005). Interestingly, episomally active *var* gene promoters are associated with this region as well lending further support to the idea that *var* gene promoters are the major target of the regulatory steps involved in mutually exclusive transcription.

Other variant multigene families in *P. falciparum*

Introduction to other families

A number of subtelomeric *P. falciparum* gene families have been identified including *var* (Su et al, 1995), *rif* (repetitive interspersed family) (Weber, 1988), *stevor* (subtelomeric variable open reading frame) (Cheng et al, 1998), *clag* (cytoadherence-linked asexual gene) (Holt et al, 1999), *Pf60* (Grellier et al, 1994), *phist* (*Plasmodium* helical interspersed subtelomeric family) (Sargeant et al, 2006) *pfmc-2tm* (*P. falciparum* Maurer's cleft two transmembrane proteins) (Sam-Yellowe et al, 2004) and *surf* (surface-associated interspersed genes) (Winter et al, 2005).

Unlike PfEMP1, the RIFIN, STEVOR, and PfMC-2TM proteins are predicted to have two membrane-spanning domains flanking a hypervariable loop forming a superfamily in terms of structure and, perhaps, function (Cheng et al, 1998; Sam-Yellowe et al, 2004; Lavazec et al, 2006). The first exon encodes a signal peptide sequence and a longer second exon encodes the PEXEL/VTS trafficking motif (Hiller et al, 2004; Marti et al, 2004) followed by family-specific sequences. The length of the predicted loop differs between RIFIN, STEVOR and PfMC-2TM proteins and the predominant sequence variability that exists within the respective gene families is found in the loop region. In contrast, the N-terminal regions are highly conserved within each family, as well as the short, positively charged C-terminal regions that follow the second TM domain.

The first report of the *rif* family dates back to 1988 (Weber, 1988) and further characterization followed 10 years later (Gardner et al, 1998). *rif* is the largest gene family in *P. falciparum* with 134 members in the laboratory strain 3D7, and can be subdivided into A-type and B-type RIFINs (Joannin et al, 2008) due to a 25aa strech present only in A-type RIFINs (Gardner et al, 2002) and the number of conserved cysteine residues (Joannin et al, 2008). They are small two-exon genes (~1000 base pairs), with a conserved domain architecture (Cheng et al, 1998; Gardner et al, 1998). Two transmembrane regions have been predicted on both sides of the variable region; with this stretch predicted to be exposed to immune pressure (Kyes et al, 1999; Cheng et al, 1998). B-type RIFINs are co-transported with PfEMP1 to the surface of

infected erythrocytes, whereas the localization of A-type RIFINs is restricted to the parasite boundaries (Haeggström et al, 2004; Petter et al, 2007, 2008). Recently, an additional role in merozoite invasion or immune evasion has been proposed for RIFIN based on the expression of a subset of RIFINs in merozoites (Petter et al, 2007). Due to their location at the RBC surface, RIFINs are immunogenic in natural infections and are recognized by human immune sera (Abdel-Latif et al, 2002, 2003; Fernandez et al, 1999). Moreover, they were shown to be expressed in a clonally variant manner and to play a role in rosette formation in some parasite isolates (Fernandez et al, 1999; Kyes et al, 1999)

The proteins most closely related to RIFINs are of the subtelomeric variable open reading frame (STEVOR) family (Cheng et al, 1998; Finn et al, 2006), numbering ~30 copies in 3D7 (Gardner et al, 2002). However, STEVORs and RIFINs differ in that the hypervariable region of STEVOR is up to 300bp smaller than the equivalent region in RIFIN (Gardner et al, 2002). STEVORs are exported to the iRBC membrane and exposed on the surface (Lavazec et al, 2006; Khattab et al, 2008; Blythe et al, 2008; Niang et al, 2009).

PfMC-2TM family members are exported to the Maurer's clefts and also to the erythrocyte surface (Sam-Yellowe et al, 2004; Lavazec et al, 2006). Maurer's clefts are *P. falciparum* specific organelles thought to function as a surrogate golgi and exported proteins transiently locate to these structures (Sam-Yellowe, 2009). The *pfmc-2tm* family contains 12 members (and one pseudogene) in 3D7, and sequences between different isolates are highly conserved with the only divergent sequence comprised within the 20aa long loop region which is proposed to be exposed on the iRBC surface (Lavazec et al, 2006).

The *phist* gene family contains 71 members in P. *falciparum* and shares orthologs with other *Plasmodium* species. All members share a PHIST domain with unknown function which is predicted to be solely composed of alpha-helices (Sargeant et al, 2006). PHIST proteins have a PEXEL-motif and can be subdivided into a, b and c groups, from which only the *phist* a group is P. *falciparum* specific.

A recent *in silico* study identified additional 53 putative exported proteins that can be subgrouped into different families termed *hyp1-17*, containing up to

eight members (Sargeant et al, 2006). Among these, the *hyp4* and *hyp5* family are a curiosity because they are amplified nine times each and exhibit no differences between paralogs and isolates (Lavazec et al, 2006; Mok et al, 2008; Sargeant et al, 2006) Finally, some additional contingency gene families exist within the *Plasmodium* genus that are predominantly found in sub-telomeric regions. *P. vivax,* the most widely distributed human *Plasmodium* species, contains a major subtelomeric multigene superfamily termed *vir* (*P. vivax* variant genes), which corresponds to approximately 10% of coding sequences and is composed of twelve subfamilies (A-L) (del Portillo et al, 2001; Carlton et al, 2008). Subfamilies A and D display structural features shared with the *P. falciparum* SURFIN and PfMC-2TM proteins, respectively (Merino et al, 2006). Overall, some of multigene families are species-specific, others are shared between some, though not all species, and some are conserved within the genus *Plasmodium*. Indeed, it was recently proposed that *vir* genes should be included within a new variant gene superfamily (*Plasmodium* interspersed repeats, *pir*) together with *rif*/*stevor* in *P. falciparum*, *kir* in *P. knowlesi*, and the *cir/yir/bir* family in *P. chabaudi*, *P. yoelii* and *P. berghei,* respectively (Janssen et al, 2004; Hall & Carlton, 2005).

In *P. knowlesi*, the discovery of the *SICAvar* (*Schizont-infected cell agglutination variant antigen*) gene family led to the first evidence of antigenic variation in *Plasmodium* species (Al-Khedery et al, 1999; Brown & Brown, 1965; Howard et al, 1983). It encodes proteins expressed on the surface of infected erythrocytes and implicated in antigenic variation in this species. Despite these parallels there is little homology between the *SICAvar* and the *var* genes. These studies suggest that species-specific evolution of antigen genes, most probably in response to pressure from differing host immune systems, has led to the current diverse repertoire of malaria antigens found in different species.

It has been hypothesized that subtelomeric gene families show high rates of ectopic recombination and that this would be important in the generation of antigenic diversity. This has been demonstrated in the model yeast *S. pombe* (Cooper et al, 1998) as well as in *P. falciparum*, where meiotic recombination among *var* genes is at least eight times more frequent than the estimated genomic average (Freitas-Junior et al, 2000; Taylor et al, 2000b). In addition,

it is also possible that the proximity of *Plasmodium* subtelomeric antigen genes to telomeres modulates the regulation of their expression, just like telomere position effect (TPE) in yeast is responsible for variegated silencing of genes placed near them (Gottschling et al, 1990).

Transcriptional control of multigene families

Despite multigene families in *P*. *falciparum* have been known for many years, there is still a big gap in our understanding of the processes that regulate their transcription. IFA (indirect immunofluorescence assay) experiments showed that RIFIN variants of both subtypes (A and B) can be expressed simultaneously in a single parasite (Petter et al, 2007); suggesting that the *rif* gene family is not transcribed by strict mutual exclusion. Further, *rif* genes do not have a critical role in determining the cytoadhesion specificity of infected erythrocytes for CD36, ICAM1 and e-selectin (Cabral & Wunderlich, 2009). The 5' UTR of *rif* genes is rather short (~245bp) compared to that of *var* genes (about 1000bp) (Tham et al, 2007; Deitsch et al, 1999). The *rif* upstream region possesses functional promoter activity, and two *cis*-acting regions have been identified (Tham et al, 2007)

The analysis of clonal parasite lines showed that the expression of STEVOR and PfMC-2TM is clonally variant and undergoes switching thus providing evidence for a role in *P. falciparum* antigenic variation (Lavazec et al, 2007). However, it remains unknown if these gene families are also transcribed in a mutually exclusive manner.

A recent publication investigated the effect of varying copy numbers of plasmid carrying active multigene family promoters in transfected parasites, namely promoters of *var*, *rif*, *stevor* and *pfmc-2tm* genes (Howitt et al, 2009). Surprisingly, increasing numbers of plasmids carrying the *var, rif* and *stevor*, but not the *pfmc-2tm*, promoter introduced an overall down-regulation of the endogenous *var*, *rif* and *stevor* families. Hence, the authors proposed the existence of a common activation factor for the gene families *var*, *rif* and *stevor*.

Main aim of this thesis

Multigene families in *P. falciparum* undergo antigenic variation and hence, are responsible for persistence of chronic infection and in the case of the *var* gene family, for severe morbidity and death. Despite the efforts in many different laboratories, the mechanisms involved in the global and individual transcriptional control pathways of multigene family members remain largely unknown. In addition, in case of the *var* family the enigma of how mutually exclusive expression and switching are controlled, and hence how antigenic variation of PfEMP1 is achieved, still persists. With this thesis I aimed at contributing to our understanding of the regulation of *var* genes and other *P. falciparum* multigene families. In particular, I was interested in (1) identifying functional *cis*-regulatory elements in *var* gene promoters; (2) investigating if other subtelomeric gene families are regulated by mechanisms similar to those controlling *var* gene expression and in testing if gene families are co-regulated; and (3) testing the existence of a unique DNA element responsible for the mutual exclusion mechanism.

Specific objectives of this thesis

1. Functional identification of *cis*-regulatory elements in *P. falciparum var* gene promoters

Establishment of a functional assay to map regulatory elements mediating *var* gene activation and mutually exclusive recognition followed by ientification of such elements by stable transfection and downstream analyses including quantitative reverse transcriptase PCR (qRT-PCR), Southern and Western blot and immunofluorescence.

2. Shedding light on the proposed regulatory linkage of multigene families in *P. falciparum*

Using the transfection-based approach developed in Objective 1 to test if promoters of *rif*, *stevor*, *pfmc-2tm*, *phist* genes are also silenced by default. Furthermore, the aim of this project was to use whole transcriptome microarray analyses to test if (1) these additional gene families are also transcribed by strict mutual exclusion; (2) regulatory crosstalk between individual gene families exists; and (3) to identify putative transcriptional regulators that may be responsible for differential activation of *var* gene subgroups A, B and/or C.

3. Testing the existence of a unique enhancer element nucleating mutually exclusive *var* gene activation

The perinuclear clusters of *P. falciparum* chromosome ends, and the debatable hypothesis that the active *var* gene moves out of a cluster led to the theory that a unique DNA sequence may exist which controls mutual exclusive *var* gene transcription. The aim of this objective was to engineer and integrate a suitable transfection construct to test this hypothesis and to identify such a DNA element within the *P. falciparum* genome by chromosome confirmation capture (3C).

References

Abdel-Latif MS, Dietz K, Issifou S, Kremsner PG & Klinkert M (2003) Antibodies to Plasmodium falciparum rifin proteins are associated with rapid parasite clearance and asymptomatic infections. *Infect. Immun* **71:** 6229-6233

Abdel-Latif MS, Khattab A, Lindenthal C, Kremsner PG & Klinkert M (2002) Recognition of variant Rifin antigens by human antibodies induced during natural Plasmodium falciparum infections. *Infect. Immun* **70:** 7013-7021

Al-Khedery B, Barnwell JW & Galinski MR (1999) Antigenic Variation in Malaria: a 3' Genomic Alteration Associated with the Expression of a P. knowlesi Variant Antigen. *Molecular Cell* **3:** 131-141

Bachmann A, Esser C, Petter M, Predehl S, von Kalckreuth V, Schmiedel S, Bruchhaus I & Tannich E (2009) Absence of erythrocyte sequestration and lack of multicopy gene family expression in Plasmodium falciparum from a splenectomized malaria patient. *PLoS ONE* **4:** e7459

Balaji S, Babu MM, Iyer LM & Aravind L (2005) Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* **33:** 3994-4006

Bártfai R, Hoeijmakers WAM, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, Kaan A, Treeck M, Gilberger T, Françoijs K & Stunnenberg HG (2010) H2A.Z Demarcates Intergenic Regions of the Plasmodium falciparum Epigenome That Are Dynamically Marked by H3K9ac and H3K4me3. *PLoS Pathog* **6:** e1001223

Baruch DI, Gormely JA, Ma C, Howard RJ & Pasloske BL (1996) Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc. Natl. Acad. Sci. U.S.A* **93:** 3497-3502

Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC, Feldman M, Taraschi TF & Howard RJ (1995) Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell* **82:** 77-87

Baruch DI, Rogerson SJ & Cooke BM (2002) Asexual blood stages of malaria antigens: cytoadherence. *Chem. Immunol* **80:** 144-162

Baum J, Papenfuss AT, Mair GR, Janse CJ, Vlachou D, Waters AP, Cowman AF, Crabb BS & de Koning-Ward TF (2009) Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res* Available at: http://www.ncbi.nlm.nih.gov/pubmed/19380379 [Accessed May 12, 2009]

Berendt AR, Ferguson DJ, Gardner J, Turner G, Rowe A, McCormick C, Roberts D, Craig A, Pinches R & Elford BC (1994) Molecular mechanisms of sequestration in malaria. *Parasitology* **108 Suppl:** S19-28

Biggs BA, Goozé L, Wycherley K, Wollish W, Southwell B, Leech JH & Brown GV (1991) Antigenic variation in Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **88:** 9171-9174

Blythe JE, Yam XY, Kuss C, Bozdech Z, Holder AA, Marsh K, Langhorne J & Preiser PR (2008) Plasmodium falciparum STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. *Infect. Immun* **76:** 3329-3336

Brown KN & Brown IN (1965) Immunity to malaria: antigenic variation in chronic infections of Plasmodium knowlesi. *Nature* **208:** 1286-1288

Cabral FJ & Wunderlich G (2009) Transcriptional memory and switching in the Plasmodium falciparumrif gene family. *Mol. Biochem. Parasitol* **168:** 186-190

Calderwood MS, Gannoun-Zaki L, Wellems TE & Deitsch KW (2003) Plasmodium falciparum var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J. Biol. Chem* **278:** 34125-34132

Callebaut I, Prat K, Meurice E, Mornon J & Tomavo S (2005) Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes. *BMC Genomics* **6:** 100

Carlson J, Helmby H, Hill AV, Brewster D, Greenwood BM & Wahlgren M (1990) Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. *Lancet* **336:** 1457-1460

Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, Crabtree J, Angiuoli SV, Merino EF, Amedeo P, Cheng Q, Coulson RMR, Crabb BS, Del Portillo HA, Essien K, Feldblyum TV, Fernandez-Becerra C, Gilson PR, Gueye AH, Guo X et al (2008) Comparative genomics of the neglected human malaria parasite Plasmodium vivax. *Nature* **455:** 757-763

Casares S, Brumeanu T & Richie TL (2010) The RTS,S malaria vaccine. *Vaccine* **28:** 4880-4894

Chen Q, Fernandez V, Sundström A, Schlichtherle M, Datta S, Hagblom P & Wahlgren M (1998) Developmental selection of var gene expression in Plasmodium falciparum. *Nature* **394:** 392-395

Cheng Q, Cloonan N, Fischer K, Thompson J, Waine G, Lanzer M & Saul A (1998) stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens. *Mol Biochem Parasitol* **97:** 161-76

Choi S, Keyes MK & Horrocks P (2006) LC/ESI-MS demonstrates the absence of 5-methyl-2'-deoxycytosine in Plasmodium falciparum genomic DNA. *Mol. Biochem. Parasitol* **150:** 350-352

Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL & Deitsch KW (2007) Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A* **104:** 899-902

Cooper JP, Watanabe Y & Nurse P (1998) Fission yeast Taz1 protein is required for meiotic telomere clustering and recombination. *Nature* **392:** 828-831

Coulson RMR, Hall N & Ouzounis CA (2004) Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. *Genome Res* **14:** 1548-1554

De Silva EK, Gehrke AR, Olszewski K, León I, Chahal JS, Bulyk ML & Llinás M (2008) Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc. Natl. Acad. Sci. U.S.A* **105:** 8393-8

Deitsch KW, Calderwood MS & Wellems TE (2001) Malaria. Cooperative silencing elements in var genes. *Nature* **412:** 875-876

Deitsch KW, del Pinal A & Wellems TE (1999) Intra-cluster recombination and var transcription switches in the antigenic variation of Plasmodium falciparum. *Mol. Biochem. Parasitol* **101:** 107-116

Deitsch KW, Lukehart SA & Stringer JR (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat. Rev. Microbiol* **7:** 493-503

D'Ombrain MC, Voss TS, Maier AG, Pearce JA, Hansen DS, Cowman AF & Schofield L (2007) Plasmodium falciparum erythrocyte membrane protein-1 specifically suppresses early production of host interferon-gamma. *Cell Host Microbe* **2:** 130-138

Duffy MF, Byrne TJ, Elliott SR, Wilson DW, Rogerson SJ, Beeson JG, Noviyanti R & Brown GV (2005) Broad analysis reveals a consistent pattern of var gene transcription in Plasmodium falciparum repeatedly selected for a defined adhesion phenotype. *Mol. Microbiol* **56:** 774-788

Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, Freitas-Junior LH, Scherf A, Crabb BS & Cowman AF (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. *Cell* **121:** 13-24

Dzikowski R, Frank M & Deitsch K (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog* **2:** e22

Dzikowski R, Li F, Amulic B, Eisberg A, Frank M, Patel S, Wellems TE & Deitsch KW (2007) Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep* **8:** 959-965

Elliott SR, Spurck TP, Dodin JM, Maier AG, Voss TS, Yosaatmadja F, Payne PD, McFadden GI, Cowman AF, Rogerson SJ, Schofield L & Brown GV (2007) Inhibition of dendritic cell maturation by malaria is dose dependent and does not require Plasmodium falciparum erythrocyte membrane protein 1. *Infect. Immun* **75:** 3621-3632

Epp C, Raskolnikov D & Deitsch KW (2008) A regulatable transgene expression system for cultured Plasmodium falciparum parasites. *Malar J* **7:** 86

Fernandez V, Hommel M, Chen Q, Hagblom P & Wahlgren M (1999) Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. *J. Exp. Med* **190:** 1393-1404

Finn RD, Mistry J, Schuster-Böckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, Eddy SR, Sonnhammer ELL & Bateman A (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* **34:** D247-251

Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BTF, Moes S, Bozdech Z, Jenoe P, Stunnenberg HG & Voss TS (2010) A Major Role for the Plasmodium falciparum ApiAP2 Protein PfSIP2 in Chromosome End Biology. *PLoS Pathog* **6:** e1000784

Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BTF, Ehlgen F, Ralph SA, Cowman AF, Bozdech Z, Stunnenberg HG & Voss TS (2009) Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog* **5:** e1000569

Frank M, Dzikowski R, Amulic B & Deitsch K (2007) Variable switching rates of malaria virulence genes are associated with chromosomal position. *Mol. Microbiol* **64:** 1486-1498

Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E & Deitsch K (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. *J. Biol. Chem* **281:** 9942-9952

Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, Nehrbass U, Wellems TE & Scherf A (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407:** 1018-1022

Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, Montiel-Condado D, Ruvalcaba-Salazar OK, Rojas-Meza AP, Mâncio-Silva L, Leal-Silvestre RJ, Gontijo AM, Shorte S & Scherf A (2005) Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121:** 25-36

Fried M & Duffy PE (1996) Adherence of Plasmodium falciparum to chondroitin sulfate A in the human placenta. *Science* **272:** 1502-1504

Fried M & Duffy PE (1998) Maternal malaria and parasite adhesion. *J. Mol. Med* **76:** 162-171

Gannoun-Zaki L, Jost A, Mu J, Deitsch KW & Wellems TE (2005) A silenced Plasmodium falciparum var promoter can be activated in vivo through spontaneous deletion of a silencing element in the intron. *Eukaryotic Cell* **4:** 490-492

Gardner JP, Pinches RA, Roberts DJ & Newbold CI (1996) Variant antigens and endothelial receptor adhesion in Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **93:** 3503-3508

Gardner MJ, Tettelin H, Carucci DJ, Cummings LM, Aravind L, Koonin EV, Shallom S, Mason T, Yu K, Fujii C, Pederson J, Shen K, Jing J, Aston C, Lai Z, Schwartz DC, Pertea M, Salzberg S, Zhou L, Sutton GG et al (1998) Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum. *Science* **282:** 1126-1132

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M, Nene V, Shallom SJ et al (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419:** 498-511

Gasser SM (2001) Positions of potential: nuclear organization and gene expression. *Cell* **104:** 639-642

Gottschling DE, Aparicio OM, Billington BL & Zakian VA (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell* **63:** 751-762

Grellier P, Précigout E, Valentin A, Carcy B & Schrével J (1994) Characterization of a new 60 kDa apical protein of Plasmodium falciparum merozoite expressed in late schizogony. *Biol. Cell* **82:** 129-138

Haeggström M, Kironde F, Berzins K, Chen Q, Wahlgren M & Fernandez V (2004) Common trafficking pathway for variant antigens destined for the surface of the Plasmodium falciparum-infected erythrocyte. *Mol. Biochem. Parasitol* **133:** 1-14

Hall N & Carlton J (2005) Comparative genomics of malaria parasites. *Curr. Opin. Genet. Dev* **15:** 609-613

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estraño C & Haldar K (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306:** 1934-1937

Holt DC, Gardiner DL, Thomas EA, Mayo M, Bourke PF, Sutherland CJ, Carter R, Myers G, Kemp DJ & Trenholme KR (1999) The cytoadherence linked asexual gene family of Plasmodium falciparum: are there roles other than cytoadherence? *Int. J. Parasitol* **29:** 939-944

Horrocks P, Pinches R, Christodoulou Z, Kyes SA & Newbold CI (2004) Variable var transition rates underlie antigenic variation in malaria. *Proc. Natl. Acad. Sci. U.S.A* **101:** 11129-11134

Horrocks P, Wong E, Russell K & Emes RD (2009) Control of gene expression in Plasmodium falciparum - ten years on. *Mol. Biochem. Parasitol* **164:** 9-25

Howard RJ, Barnwell JW & Kao V (1983) Antigenic variation of Plasmodium knowlesi malaria: identification of the variant antigen on infected erythrocytes. *Proc. Natl. Acad. Sci. U.S.A* **80:** 4129-4133

Howitt CA, Wilinski D, Llinás M, Templeton TJ, Dzikowski R & Deitsch KW (2009) Clonally variant gene families in Plasmodium falciparum share a common activation factor. *Mol. Microbiol* **73:** 1171-1185

Janssen CS, Phillips RS, Turner CMR & Barrett MP (2004) Plasmodium interspersed repeats: the major multigene superfamily of malaria parasites. *Nucleic Acids Res* **32:** 5712-5720

Jensen ATR, Magistrado P, Sharp S, Joergensen L, Lavstsen T, Chiucchiuini A, Salanti A, Vestergaard LS, Lusingu JP, Hermsen R, Sauerwein R, Christensen J, Nielsen MA, Hviid L, Sutherland C, Staalsoe T & Theander TG (2004) Plasmodium falciparum associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes. *J. Exp. Med* **199:** 1179-1190

Joannin N, Abhiman S, Sonnhammer EL & Wahlgren M (2008) Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. *BMC Genomics* **9:** 19

Kaestli M, Cortes A, Lagog M, Ott M & Beck H (2004) Longitudinal assessment of Plasmodium falciparum var gene transcription in naturally infected asymptomatic children in Papua New Guinea. *J. Infect. Dis* **189:** 1942-1951

Khattab A, Bonow I, Schreiber N, Petter M, Schmetz C & Klinkert M (2008) Plasmodium falciparum variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. *Malar. J* **7:** 137

Kraemer SM, Kyes SA, Aggarwal G, Springer AL, Nelson SO, Christodoulou Z, Smith LM, Wang W, Levin E, Newbold CI, Myler PJ & Smith JD (2007) Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics* **8:** 45

Kraemer SM & Smith JD (2006) A family affair: var genes, PfEMP1 binding, and malaria disease. *Curr. Opin. Microbiol* **9:** 374-380

Kyes S, Horrocks P & Newbold C (2001) Antigenic variation at the infected red cell surface in malaria. *Annu. Rev. Microbiol* **55:** 673-707

Kyes SA, Rowe JA, Kriek N & Newbold CI (1999) Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **96:** 9333-9338

Kyes S, Christodoulou Z, Pinches R, Kriek N, Horrocks P & Newbold C (2007) Plasmodium falciparum var gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. *Mol. Microbiol* **63:** 1237-1247

Kyes SA, Christodoulou Z, Raza A, Horrocks P, Pinches R, Rowe JA & Newbold CI (2003) A well-conserved Plasmodium falciparum var gene shows an unusual stage-specific transcript pattern. *Mol. Microbiol* **48:** 1339-1348

Lavazec C, Sanyal S & Templeton TJ (2006) Hypervariability within the Rifin, Stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. *Nucleic Acids Res* **34:** 6696-707

Lavazec C, Sanyal S & Templeton TJ (2007) Expression switching in the stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. *Mol Microbiol* **64:** 1621-34

Lavstsen T, Magistrado P, Hermsen CC, Salanti A, Jensen ATR, Sauerwein R, Hviid L, Theander TG & Staalsoe T (2005) Expression of Plasmodium falciparum erythrocyte membrane protein 1 in experimentally infected humans. *Malar. J* **4:** 21

Lavstsen T, Salanti A, Jensen ATR, Arnot DE & Theander TG (2003) Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar. J* **2:** 27

Liu W, Li Y, Learn GH, Rudicell RS, Robertson JD, Keele BF, Ndjango JN, Sanz CM, Morgan DB, Locatelli S, Gonder MK, Kranzusch PJ, Walsh PD, Delaporte E, Mpoudi-Ngole E, Georgiev AV, Muller MN, Shaw GM, Peeters M, Sharp PM et al (2010) Origin of the human malaria parasite Plasmodium falciparum in gorillas. *Nature* **467:** 420-425

Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R & Scherf A (2007) 5' flanking region of var genes nucleate histone modification patterns linked to

phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol* **66:** 1296-1305

Lopez-Rubio J, Mancio-Silva L & Scherf A (2009) Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5:** 179-190

MacPherson GG, Warrell MJ, White NJ, Looareesuwan S & Warrell DA (1985) Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *Am. J. Pathol* **119:** 385-401

Mancio-Silva L, Rojas-Meza AP, Vargas M, Scherf A & Hernandez-Rivas R (2008) Differential association of Orc1 and Sir2 proteins to telomeric domains in Plasmodium falciparum. *J. Cell. Sci* **121:** 2046-2053

Marti M, Good RT, Rug M, Knuepfer E & Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306:** 1930-1933

Marty AJ, Thompson JK, Duffy MF, Voss TS, Cowman AF & Crabb BS (2006) Evidence that Plasmodium falciparum chromosome end clusters are cross-linked by protein and are the sites of both virulence gene silencing and activation. *Mol. Microbiol* **62:** 72-83

Mazier D, Beaudoin RL, Mellouk S, Druilhe P, Texier B, Trosper J, Miltgen F, Landau I, Paul C & Brandicourt O (1985) Complete development of hepatic stages of Plasmodium falciparum in vitro. *Science* **227:** 440-442

Merino EF, Fernandez-Becerra C, Durham AM, Ferreira JE, Tumilasci VF, d'Arc-Neves J, da Silva-Nunes M, Ferreira MU, Wickramarachchi T, Udagama-Randeniya P, Handunnetti SM & Del Portillo HA (2006) Multi-character population study of the vir subtelomeric multigene superfamily of Plasmodium vivax, a major human malaria parasite. *Mol. Biochem. Parasitol* **149:** 10-16

Merrick CJ, Dzikowski R, Imamura H, Chuang J, Deitsch K & Duraisingh MT (2010) The effect of Plasmodium falciparum Sir2a histone deacetylase on clonal and longitudinal variation in expression of the var family of virulence genes. *Int. J. Parasitol* **40:** 35-43

Miao J, Fan Q, Cui L, Li J, Li J & Cui L (2006) The malaria parasite Plasmodium falciparum histones: organization, expression, and acetylation. *Gene* **369:** 53-65

Mok BW, Ribacke U, Sherwood E & Wahlgren M (2008) A highly conserved segmental duplication in the subtelomeres of Plasmodium falciparum chromosomes varies in copy number. *Malar. J* **7:** 46

Montgomery J, Mphande FA, Berriman M, Pain A, Rogerson SJ, Taylor TE, Molyneux ME & Craig A (2007) Differential var gene expression in the organs of patients dying of falciparum malaria. *Mol. Microbiol* **65:** 959-967

Niang M, Yan Yam X & Preiser PR (2009) The Plasmodium falciparum STEVOR multigene family mediates antigenic variation of the infected erythrocyte. *PLoS Pathog* **5:** e1000307

Olotu A, Lusingu J, Leach A, Lievens M, Vekemans J, Msham S, Lang T, Gould J, Dubois M, Jongert E, Vansadia P, Carter T, Njuguna P, Awuondo KO, Malabeja A, Abdul O, Gesase S, Mturi N, Drakeley CJ, Savarese B et al (2011) Efficacy of RTS,S/AS01E malaria vaccine and exploratory analysis on anti-circumsporozoite antibody titres and protection in children aged 5-17 months in Kenya and Tanzania: a randomised controlled trial. *The Lancet Infectious Diseases* **11:** 102-109

Pérez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernández-Cuevas NA, Delgadillo DM, Vargas M, Martínez-Calvillo S, Scherf A & Hernandez-Rivas R (2009) Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. *Nucleic Acids Res* **37:** 2596-2606

Peters JM, Fowler EV, Krause DR, Cheng Q & Gatton ML (2007) Differential changes in Plasmodium falciparum var transcription during adaptation to culture. *J. Infect. Dis* **195:** 748-755

Petter M, Bonow I & Klinkert M (2008) Diverse expression patterns of subgroups of the rif multigene family during Plasmodium falciparum gametocytogenesis. *PLoS ONE* **3:** e3779

Petter M, Haeggström M, Khattab A, Fernandez V, Klinkert M & Wahlgren M (2007) Variant proteins of the Plasmodium falciparum RIFIN family show distinct subcellular localization and developmental expression patterns. *Mol Biochem Parasitol* **156:** 51-61

Pongponratn E, Riganti M, Punpoowong B & Aikawa M (1991) Microvascular sequestration of

parasitized erythrocytes in human falciparum malaria: a pathological study. *Am. J. Trop. Med. Hyg* **44:** 168-175

del Portillo HA, Fernandez-Becerra C, Bowman S, Oliver K, Preuss M, Sanchez CP, Schneider NK, Villalobos JM, Rajandream MA, Harris D, Pereira da Silva LH, Barrell B & Lanzer M (2001) A superfamily of variant genes encoded in the subtelomeric region of Plasmodium vivax. *Nature* **410:** 839-842

Ralph SA, Scheidig-Benatar C & Scherf A (2005) Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. *Proc. Natl. Acad. Sci. U.S.A* **102:** 5414-5419

Reeder JC, Cowman AF, Davern KM, Beeson JG, Thompson JK, Rogerson SJ & Brown GV (1999) The adhesion of Plasmodium falciparum-infected erythrocytes to chondroitin sulfate A is mediated by P. falciparum erythrocyte membrane protein 1. *Proc. Natl. Acad. Sci. U.S.A* **96:** 5198-5202

Roberts DJ, Craig AG, Berendt AR, Pinches R, Nash G, Marsh K & Newbold CI (1992) Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357:** 689-692

Roestenberg M, McCall M, Hopman J, Wiersma J, Luty AJF, van Gemert GJ, van de Vegte-Bolmer M, van Schaijk B, Teelen K, Arens T, Spaarman L, de Mast Q, Roeffen W, Snounou G, Rénia L, van der Ven A, Hermsen CC & Sauerwein R (2009) Protection against a malaria challenge by sporozoite inoculation. *N. Engl. J. Med* **361:** 468-477

Rogerson SJ & Brown GV (1997) Chondroitin sulphate A as an adherence receptor for Plasmodium falciparum-infected erythrocytes. *Parasitol. Today (Regul. Ed.)* **13:** 70-75

Rowe A, Obeiro J, Newbold CI & Marsh K (1995) Plasmodium falciparum rosetting is associated with malaria severity in Kenya. *Infect. Immun* **63:** 2323-2326

Rowe JA & Kyes SA (2004) The role of Plasmodium falciparum var genes in malaria in pregnancy. *Mol. Microbiol* **53:** 1011-1019

Rowe JA, Moulds JM, Newbold CI & Miller LH (1997) P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature* **388:** 292-295

Salanti A, Dahlbäck M, Turner L, Nielsen MA, Barfod L, Magistrado P, Jensen ATR, Lavstsen T, Ofori MF, Marsh K, Hviid L & Theander TG (2004) Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *J. Exp. Med* **200:** 1197-1203

Salanti A, Staalsoe T, Lavstsen T, Jensen ATR, Sowa MPK, Arnot DE, Hviid L & Theander TG (2003) Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol. Microbiol* **49:** 179-191

Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AMG, Cohen AM, Janssen-Megens EM, van de Vegte-Bolmer M, Selzer RR, Iniguez AL, Green RD, Sauerwein RW, Jensen ON & Stunnenberg HG (2009) Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **106:** 9655-9660

Sam-Yellowe TY (2009) The role of the Maurer's clefts in protein transport in Plasmodium falciparum. *Trends Parasitol* **25:** 277-284

Sam-Yellowe TY, Florens L, Johnson JR, Wang T, Drazba JA, Le Roch KG, Zhou Y, Batalov S, Carucci DJ, Winzeler EA & Yates JR (2004) A Plasmodium gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. *Genome Res* **14:** 1052-1059

Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP & Cowman AF (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* **7:** R12-R12

Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, Pouvelle B, Gysin J & Lanzer M (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. *EMBO J* **17:** 5418-5426

Singh B, Sung LK, Matusop A, Radhakrishnan A, Shamsul SS, Cox-Singh J, Thomas A & Conway DJ (2004) A large focus of naturally acquired Plasmodium knowlesi infections in human beings. *The Lancet* **363:** 1017-1024

Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, Peterson DS, Pinches R, Newbold CI & Miller LH (1995) Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected

erythrocytes. *Cell* **82:** 101-110
Smith JD, Subramanian G, Gamain B, Baruch DI & Miller LH (2000) Classification of adhesive domains in the Plasmodium falciparum erythrocyte membrane protein 1 family. *Mol. Biochem. Parasitol* **110:** 293-310
Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, Peterson DS, Ravetch JA & Wellems TE (1995) The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. *Cell* **82:** 89-100
Taylor HM, Kyes SA, Harris D, Kriek N & Newbold CI (2000a) A study of var gene transcription in vitro using universal var gene primers. *Mol. Biochem. Parasitol* **105:** 13-23
Taylor HM, Kyes SA & Newbold CI (2000b) Var gene diversity in Plasmodium falciparum is generated by frequent recombination events. *Mol. Biochem. Parasitol* **110:** 391-397
Tham W, Payne PD, Brown GV & Rogerson SJ (2007) Identification of basic transcriptional elements required for rif gene transcription. *International Journal for Parasitology* **37:** 605-615
Tonkin CJ, Carret CK, Duraisingh MT, Voss TS, Ralph SA, Hommel M, Duffy MF, Silva LMD, Scherf A, Ivens A, Speed TP, Beeson JG & Cowman AF (2009) Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. *PLoS Biol* **7:** e84
Trelle MB, Salcedo-Amaya AM, Cohen AM, Stunnenberg HG & Jensen ON (2009) Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. *J. Proteome Res* **8:** 3439-3450
Trimnell AR, Kraemer SM, Mukherjee S, Phippard DJ, Janes JH, Flamoe E, Su X, Awadalla P & Smith JD (2006) Global genetic diversity and evolution of var genes associated with placental and severe childhood malaria. *Mol. Biochem. Parasitol* **148:** 169-180
Urban BC, Ferguson DJ, Pain A, Willcox N, Plebanski M, Austyn JM & Roberts DJ (1999) Plasmodium falciparum-infected erythrocytes modulate the maturation of dendritic cells. *Nature* **400:** 73-77
Voss TS, Thompson JK, Waterkeyn J, Felger I, Weiss N, Cowman AF & Beck HP (2000) Genomic distribution and functional characterisation of two distinct and conserved Plasmodium falciparum var gene 5' flanking sequences. *Mol. Biochem. Parasitol* **107:** 103-115
Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, Beeson JG, Reeder JC, Crabb BS & Cowman AF (2006) A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. *Nature* **439:** 1004-1008
Voss TS, Kaestli M, Vogel D, Bopp S & Beck H (2003) Identification of nuclear proteins that interact differentially with Plasmodium falciparum var gene promoters. *Mol. Microbiol* **48:** 1593-1607
Voss TS, Tonkin CJ, Marty AJ, Thompson JK, Healer J, Crabb BS & Cowman AF (2007) Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. *Mol. Microbiol* **66:** 139-150
Weber JL (1988) Interspersed repetitive DNA from Plasmodium falciparum. *Mol. Biochem. Parasitol* **29:** 117-124
Winter G, Kawai S, Haeggström M, Kaneko O, von Euler A, Kawazu S, Palm D, Fernandez V & Wahlgren M (2005) SURFIN is a polymorphic antigen expressed on Plasmodium falciparum merozoites and infected erythrocytes. *J Exp Med* **201:** 1853-63
Wirth DF (2002) The parasite genome: Biological revelations. *Nature* **419:** 495-496

Mutually Exclusive Activation of *Plasmodium falciparum var* Genes is Mediated by a Novel Mechanism of Singular Gene Choice

Nicolas M. B. Brancucci[1,2,3], Kathrin Witmer[1,2,3], Igor Niederwieser[1,2], Christoph D. Schmid[1,2], Christian Flueck[1,2], and Till S. Voss[1,2,*]

[1]Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, 4002 Basel, Switzerland; [2]University of Basel, 4003 Basel, Switzerland

[3]These authors contributed equally to this work.

Abstract

Plasmodium falciparum is responsible for the most severe form of malaria in humans. Antigenic variation of *P. falciparum* erythrocyte membrane protein 1 leads to immune evasion and occurs through switches in mutually exclusive *var* gene transcription. The recent progress in *Plasmodium* epigenetics notwithstanding, the mechanisms by which singularity of *var* activation is achieved are unknown. We employed a functional approach to dissect the role of *var* promoters in mutually exclusive activation. Besides identifying elements involved in promoter activation, transcriptional initiation and translational repression we describe a sequence-specific DNA-protein interaction required to maintain singular gene choice. Activation of promoters lacking this motif occurs no longer in competition with endogenous *var* genes. Together, our findings identified a novel mechanism for the control of mutually exclusive gene transcription in eukaryotes. In addition to its importance in *P. falciparum* antigenic variation, our results may also help to explain singular gene choice in other systems.

Introduction

Many unicellular pathogens use antigenic variation to escape adaptive immune responses elicited in the host. The widespread occurrence of this strategy in evolutionary distant species underscores its key role in pathogen survival and spreading. While the underlying control pathways are highly diverse in different systems, both mechanistically and in terms of complexity, antigenic variation is defined by two basic concepts. First, the antigens are encoded by gene families, the members of which are expressed in a mutually exclusive manner. Second, switches in the expression of individual members lead to antigenic variation of surface-exposed antigens. In several medically important pathogens such as *Borrelia* spp., *Neisseria* spp., *Giardia lamblia*, *Plasmodium falciparum* and *Trypanosoma brucei*, this paradigm of clonal phenotypic variation reaches a remarkable yet poorly understood level of sophistication (Morrison et al., 2009; Deitsch et al., 2009; Prucca and Lujan, 2009; Dzikowski and Deitsch, 2009).

The apicomplexan parasite *P. falciparum* causes several hundred million malaria cases and close to one million deaths annually (World Health

Organisation, 2010). Malaria-associated morbidity and mortality is a result of the intra-erythrocytic developmental cycle (IDC) where repeated rounds of parasite invasion into red blood cells (RBCs) are followed by intra-cellular maturation and replication. Intra-erythrocytic parasites expose the major virulence factor *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) on the host cell surface (Leech et al., 1984). This highly polymorphic antigen, encoded by the 60-member *var* gene family, undergoes antigenic variation to facilitate chronic infection and transmission (Su et al., 1995; Smith et al., 1995; Gardner et al., 2002; Roberts et al., 1992; Biggs et al., 1991). Furthermore, PfEMP1 mediates sequestration of infected RBC aggregates in the microvasculature of various organs and is thus directly responsible for severe outcomes, including cerebral and placental malaria (Pongponratn et al., 1991; MacPherson et al., 1985; Beeson and Duffy, 2005; Baruch et al., 1996; Gardner et al., 1996; Reeder et al., 1999).

Mutually exclusive expression of PfEMP1 is controlled at the level of *var* gene transcription and is carried out by RNA polymerase II (RNA polII) in ring stage parasites during the first half of the IDC (Kyes et al., 2007; Scherf et al., 1998; Dzikowski et al., 2006). Only one *var* gene is transcribed at any time while all other members are silenced (Scherf et al., 1998). Switches in *var* gene transcription, and consequently antigenic variation of PfEMP1, occur by *in situ* activation and are independent of detectable recombination events (Scherf et al., 1998). Silencing is explained by the fact that all *var* genes are positioned in subtelomeric and some chromosome-internal heterochromatic regions (Gardner et al., 2002; Salcedo-Amaya et al., 2009; Flueck et al., 2009; Lopez-Rubio et al., 2009). These chromosomal domains cluster at the nuclear periphery and are uniformly enriched in histone 3 lysine 9 tri-methylation (H3K9me3) and *P. falciparum* heterochromatin protein 1 (PfHP1) (Freitas-Junior et al., 2000; Salcedo-Amaya et al., 2009; Flueck et al., 2009; Lopez-Rubio et al., 2009; Perez-Toledo et al., 2009; Freitas-Junior et al., 2005; Duraisingh et al., 2005). The presence of these epigenetic marks is directly linked to *var* gene silencing (Chookajorn et al., 2007; Lopez-Rubio et al., 2007). In contrast, the active *var* locus is associated with H3K9ac and H3K4me2/me3 instead (Lopez-Rubio et al., 2007). Additionally, the process of *var* gene activation is linked to perinuclear locus re-positioning (Voss et al.,

2006; Duraisingh et al., 2005; Marty et al., 2006; Ralph et al., 2005; Dzikowski et al., 2007).

In recent years, *var* gene promoters emerged as the key component in all layers of *var* gene regulation. They play essential roles not only in *var* gene silencing and activation, but also in recognition of the active family member by the enigmatic mutual exclusion mechanism (Voss et al., 2007; Voss et al., 2006; Dzikowski et al., 2006). Functional analysis has been most informative in experiments using *var* promoters to drive transcription of drug-selectable reporter genes. In absence of drug selection *var* promoters are predominantly silenced, whereas drug challenge selects for parasites carrying active promoters (Voss et al., 2006; Voss et al., 2007). This forced activation is sufficient to infiltrate a drug-selectable reporter into the mutual exclusion program (Dzikowski et al., 2007; Voss et al., 2007; Voss et al., 2006). In addition to *var* promoters, the conserved intron of *var* genes acts as a cooperative partner in silencing and mutual exclusion when placed downstream of a *var* promoter-driven expression cassette (Frank et al., 2006; Gannoun-Zaki et al., 2005; Calderwood et al., 2003; Deitsch et al., 2001; Dzikowski et al., 2007).

Based on the current state of knowledge we postulated that transcriptional control of *var* genes is mediated by unknown sequence information contained within the promoter region. To date, however, functional entities responsible for silencing, activation or mutual exclusion have not been discovered. In this study, we developed a comprehensive functional promoter mapping approach tailored to identify and characterise *var* gene-specific regulatory information. We identified an upstream activating sequence (UAS) that is essential for promoter activation. This region acts autonomously in the context of a minimal heterologous promoter reminiscent of classical RNA polII-dependent transactivation. Moreover, we identified a DNA-protein interaction downstream of the transcriptional start site (TSS), which is required for inclusion of the locus into the mutual exclusion pathway. In absence of this element *var* promoters are fully active but, unlike wild-type promoters, do not compete with endogenous *var* gene transcription. We also show that the *var* 5' untranslated region (5' UTR) is a target of post-transcriptional regulation. Our findings show for the first time that the complex regulation of mutually exclusive *var* gene

transcription is mediated by functional *cis*-acting modules with intrinsic and position-dependent activities. They are furthermore consistent with a novel mechanism in sustaining singular gene choice in eukaryotes.

Results

Functional *var* promoter mapping by bidirectional deletion analysis

We made use of a system suitable to analyse promoter activity in stably transfected parasites. All reporter constructs are based on the parental plasmid pBC (Figure 1A) where the blasticidin deaminase (*bsd*) resistance cassette selects for stable episomes. A 2.5kb *var* upsC upstream sequence controls transcription of the dual reporter encoding human dihydrofolate reductase fused to green fluorescent protein (h*dhfr-gfp)*. A *var* gene intron element is located downstream of the h*dhfr-gfp* cassette to account for its role in *var* gene regulation. A telomere-associated repeat element 6 (TARE6 or rep20) sequence is included for improved plasmid segregation (O'Donnell et al., 2002). In such a context, homogenous populations carrying active upsC promoters are obtained via selection with the antifolate drug WR99210 (WR) (Voss et al., 2007; Voss et al., 2006).

To identify elements involved in promoter activation and mutual exclusion we sequentially truncated the upsC promoter from either the 5' or 3' end (Figure 1B). We chose this bi-directional approach to identify possible functional regions both up- and downstream of the putative TSS (Deitsch et al., 1999). Transfected parasites were challenged with WR and resistant populations were obtained for all but one cell line, 3D7/pBC3 (Figure 1B). Several attempts to select for WR-resistant 3D7/pBC3 parasites failed showing that the region between -1656 to -1217 comprises an important UAS and/or the core promoter.

To test if any of the deletions affected promoter strength we determined relative h*dhfr-gfp* transcript levels in ring stage parasites by quantitative reverse transcriptase PCR (qRT-PCR). As shown in Figure 1B, transcript levels in 3D7/pBC1 and 3D7/pBC2 were similar to those in 3D7/pBC indicating that the sequence upstream of -1656 does not contribute to *var* promoter activity. The promoter in pBC5, lacking 491bp of the 5' UTR, was

also fully active. In contrast, the truncation encompassing bps -1057 to -1 in pBC4 caused a significant reduction in transcript levels. Hence, two deletions, located upstream and downstream of the TSS, respectively, caused a defect in promoter activity.

Functional identification of an autonomous upsC upstream activating sequence

To learn more about the nature of the putative UAS we set out to analyse its function in the context of a minimal heterologous promoter. We decided to use the knob-associated histidine rich protein (*kahrp*) gene promoter for three reasons. First, the TSS of this gene has been mapped to 849bp upstream of the ATG (Lanzer et al., 1992). Second, similar to *var* genes timing of *kahrp* transcription peaks in ring stage parasites. Third, the *kahrp* locus is not enriched in H3K9me3/PfHP1 (Salcedo-Amaya et al., 2009; Flueck et al., 2009; Lopez-Rubio et al., 2009), which is important to prevent heterochromatin-mediated masking of autonomous *cis*-acting activities. Hence, we generated plasmid pBK_{min} where bps -1115 to -1 of the *kahrp* upstream sequence control transcription of the h*dhfr-gfp* reporter. The rep20 and intron elements were removed for initial experiments to avoid potential interference of these modules with *kahrp* promoter activity (pBK_{min}-RI) (Figure 2A). Parasites carrying pBK_{min}-RI episomes were readily obtained. Notably, the disposition of pBK_{min}-RI concatamers to integrate into the endogenous *kahrp* locus allowed us to measure K_{min} activity also in a chromosomal environment. This integration event essentially causes a promoter swap where K_{min} drives expression of the endogenous *kahrp* gene and the endogenous *kahrp* promoter controls transcription of the h*dhfr-gfp* reporter (Figures 2B and S1). As shown in Figure 2C, the episomal K_{min} promoter exhibited three-fold higher activity compared to the chromosomal K_{min} promoter. Importantly, however, compared to the endogenous full-length *kahrp* promoter, the episomal and chromosomal minimal promoters displayed a 300-fold and 1000-fold reduced activity, respectively (Figure 2C). Hence, K_{min} clearly fulfilled the requirements for a minimal promoter.

Next, we asked if the upsC region implicated in transcriptional activation was able to activate the minimal *kahrp* promoter. We kept the rep20 and intron

elements in the following set of constructs for consistency with the setup used in the upsC deletion constructs. We cloned two overlapping fragments containing the putative upsC UAS upstream of the minimal *kahrp* promoter in pBK_{min} to create upsC-K_{min} hybrid promoters ($pBC1K_{min}$ and $pBC2K_{min}$) (Figure 2D). The region downstream of the upsC TSS encompassing bps -463 to -20, which has no effect on upsC promoter activity (Figure 1B), was used as control ($pBC3K_{min}$). qRT-PCR analysis revealed that upsC fragments C1 (-1679 to -1200) and C2 (-1401 to -727) consistently activated K_{min} to a similar extent in both the episomal and chromosomal context whereas fragment C3 had no effect (Figures 2D and S1). Furthermore, neither the *var* intron nor the rep20 element altered K_{min} activity. These findings corroborate the results obtained with the upsC deletion constructs and are consistent with the presence of a *var* UAS located between bps -1401 and -1217. Remarkably, this element activates transcription from a heterologous minimal promoter demonstrating an autonomous, context-independent function in activating RNA polII-mediated transcription.

The upsC TSS and downstream sequence modulate levels of expression

To date, the TSS has only been mapped for a single upsC *var* gene (-1056 upstream of the ATG) (Deitsch et al., 1999). Based on a multiple sequence alignment we expected the TSS of PFL1960w at position -1167 (Voss et al., 2000). We therefore assumed the deletion of bps -1057 to -1 on plasmid pBC4, which caused a substantial reduction in promoter activity (Figure 1B), spanned most of the 5' UTR but did not remove the TSS. Northern blot analysis confirmed the reduced abundance of steady state transcripts in 3D7/pBC4, but revealed that the size difference between pBC- and pBC4-derived transcripts was much smaller than expected (Figure 3A). In spite of the 1057bp deletion in the 5' UTR, pBC4-derived transcripts were similar in size to the transcripts originating from pBC6, where only 460bp of 5' UTR were deleted. This suggested that transcription from the truncated promoter in pBC4 initiated from an alternative upstream TSS. Consequently, the reduced promoter activity observed for pBC4 was likely due to the lack of proper transcriptional initiation from the natural upsC TSS.

To investigate this possibility in more detail, we created another plasmid termed pBCK (Figure 3B). The CK hybrid promoter consists of the upsC fragment spanning bps -1679 to -914, which includes the identified UAS and the suspected core promoter and TSS. This region was fused upstream to 764bp of the *kahrp* 5' UTR to maintain the approximate distance of the core promoter in relation to the ATG. This hybrid promoter produced h*dhfr-gfp* transcript levels comparable to those achieved by the full-length promoter in 3D7/pBC. Therefore, the upsC sequence spanning bps -1679 to -914 specifies a sub-region of the upsC promoter capable of proper transcriptional activation and initiation.

We were still intrigued by the fact that 3D7/pBC4 parasites were able to survive WR challenge at largely reduced h*dhfr-gfp* transcript levels. Interestingly, parallel qRT-PCR and semi-quantitative Western blot analysis demonstrated that despite 20-fold higher h*dhfr-gfp* transcript levels in WR-selected 3D7/pBC compared to 3D7/pBC4 both lines expressed similar amounts of hDHFR-GFP (Figure 3C). An independent time-course experiment confirmed these results and excluded the possibility of altered transcriptional timing and/or transcript accumulation in 3D7/pBC4 parasites (Figure S2). Comparable levels of hDHFR-GFP expression in both lines were also evident from live fluorescence microscopy analysis (Figure S3). These findings indicate a role of the upsC 5' UTR in translational repression. This hypothesis is backed up by the fact that pBCK-derived transcripts, which carry the *kahrp* instead of the upsC 5' UTR, also give rise to substantially higher protein levels (Figure 3C). Our results further demonstrate that (1) the upsC TSS of this particular promoter is located in between bps -1057 to -914; (2) a 750bp upsC core region fully recapitulates the activation potential of the 2.5kb upsC upstream regulatory region; and (3) the region downstream of the TSS is dispensable for upsC promoter activation.

Mutually exclusive *var* locus recognition is mediated by a regulatory element downstream of the TSS

Transgenic parasites carrying activated full-length *var* promoters do not transcribe endogenous *var* genes and fail to express PfEMP1 (Voss et al., 2007; Voss et al., 2006; Dzikowski et al., 2007; Howitt et al., 2009;

Chookajorn et al., 2007; Dzikowski et al., 2006). This implies that mutually exclusive locus recognition may be mediated by *cis*-linked regulatory sequence information. Therefore, we tested if any of the activated truncated promoters escaped mutually exclusive activation (Figure 4A). As expected, the control line 3D7/pBM, in which the unrelated ring stage-specific *mahrp* promoter controls h*dhfr-gfp* transcription, expressed PfEMP1 at the expected levels, whereas 3D7/pBC parasites exhibited a PfEMP1 knock-down phenotype. PfEMP1 expression was also abolished in 3D7/pBC2 parasites showing that the most upstream 832 bps of the upsC promoter are not involved in mutually exclusive recognition. In contrast, 3D7/pBC5, 3D7/pBC4 and 3D7/pBCK parasites expressed PfEMP1 at wild-type levels similar to the control line 3D7/pBM. The promoters on these three constructs all lack the same 419bp element downstream of the TSS.

To map this functional region more precisely we cloned three additional truncated promoters in pBC6, pBC7 and pBC8. WR-selected 3D7/pBC6 and 3D7/pBC7 parasites expressed PfEMP1 showing that these truncated promoters were also not recognised by the mutual exclusion mechanisms (Figure 4B). In contrast, 3D7/pBC8 parasites failed to express PfEMP1 demonstrating that the pBC8 promoter was activated in a mutually exclusive manner similar to the full-length promoter. This series of experiments pinpoint a 101bp mutual exclusion element (MEE) (bps -316 to -215) that drives the upsC promoter into mutually exclusive activation; in absence of this element promoters escape this restriction and can be fully activated in parallel to endogenous *var* transcription.

Since such a proposed function of the MEE may be directly linked to the interaction with an unknown nuclear factor we tested three overlapping fragments (MEE1 - MEE3) in electromobility shift assays (EMSA) (Figure 4C). Whereas MEE1 and MEE3 showed no sign of specific binding (data not shown), the 47bp MEE2 fragment formed a specific DNA-protein complex when incubated with parasite nuclear extracts. This interaction was specifically competed in a dose-dependent manner by an excess of homologous competitor only. Interestingly, the core 20 bp sequence of this motif occurs in a strictly conserved position upstream of 32 *var* genes of the upsC, upsB, upsB/C and upsB/A types (Figure S4). In summary, these

experiments identified a *cis*-acting element downstream of the TSS that is required for mutually exclusive *var* gene activation. Our functional data combined with the demonstrated binding of a nuclear factor assign a potential role to the MEE-protein interaction in active prevention of transcription. Further, the conservation of this motif in over half of all *var* loci provides circumstantial evidence for a conserved role of this element in mutually exclusive *var* expression

The upsC downstream element inhibits expression from a heterologous promoter in a position-dependant manner

Next, we asked if the upsC downstream sequence including the MEE is able to modulate expression from a heterologous promoter when placed in a conserved position downstream of the TSS. Noteworthy, insertion of the same element upstream of the *kahrp* TSS in pBC3K$_{min}$ had no measurable effect on transcription (Figure 2D). We replaced a stretch of the *kahrp* 5' UTR (-445 to -1) in pBK$_{min}$ with the upsC sequence ranging from bps -519 to -1 to obtain pBK$_{min}$C4 (Figure 5A). Transfected parasites were selected on BSD and the plasmid was integrated into the *kahrp* locus. This recombination event created the *kahrp*-upsC hybrid promoter *kahrp*C4, which controls transcription of the first h*dhfr-gfp* gene on the integrated concatamer (Figure 5B). Strikingly, 3D7/pBK$_{min}$C4 parasites were completely refractory to WR selection in twelve independent challenge experiments. Attempt number 13 was eventually successful and we obtained a WR-resistant 3D7/pBK$_{min}$C4 population (Figure 5B). Interestingly, comparison of unselected and WR-selected 3D7/pBK$_{min}$C4 to the control parasite 3D7/pBK$_{min}$ revealed a dual role for the upsC downstream element in (1) inhibiting translation; and (2) repressing transcription. Firstly, unselected and WR-selected 3D7/pBK$_{min}$C4 populations transcribed h*dhfr-gfp* to similar levels, yet unselected parasites showed no evidence for hDHFR-GFP expression. Importantly, this discrepancy was not observed in 3D7/pBK$_{min}$ control parasites where the wild-type *kahrp* promoter controls h*dhfr-gfp* transcription (Figure 5C). This result provides direct second evidence for an important function for the upsC 5' UTR in translational repression. Secondly, qRT-PCR analysis indicates a role for the upsC downstream element in repression of RNA PolII-dependent transcription. We

observed two-fold lower relative h*dhfr-gfp* transcript levels in 3D7/pBK$_{min}$C4 compared to the control line 3D7/pBK$_{min}$ (52.04% +/- 17.3 SEM) (Figure 5D). This was not observed for an endogenous control gene that was transcribed at virtually identical levels in both lines (95.51% +/- 16.33 SEM).

In light of the exceptional difficulties in generating WR-resistant 3D7/pBK$_{min}$C4 parasites, we considered a genomic rearrangement the most probable cause for this altered phenotype. Indeed, we noted an additional restriction fragment hybridising with the h*dhfr* probe in WR-resistant 3D7/pBK$_{min}$C4 parasites that was absent in the unselected population (Figure 6A). A subsequent elaborate mapping strategy based on Southern blotting and ligation-mediated PCR uncovered a remarkable recombination event (Figures S5 and S6). Our results are consistent with a gene conversion process that resulted in the loss of the end of chromosome four and replacement with a duplicated copy of the end of chromosome two (Figure 6B). This was mediated by a single crossover between the most telomere-proximal h*dhfr-gfp* gene on chromosome two, and the 3' end of exon 1 of *var* gene PFD0005w (Figure 6B and Figures S5 and S6). Consequently, transcription of a single functional h*dhfr-gfp* gene was now under control of a *var* gene intron, which was previously shown to have bi-directional promoter (Epp et al., 2009). Indeed, qRT-PCR using primers specific to this recombined locus identified active h*dhfr-gfp* transcription (Figure 6C). We therefore believe expression of hDHFR-GFP in the resistant population was conferred by *var* intron-derived h*dhfr-gfp* transcripts rather than those produced by the *kahrp*C4 promoter.

In conclusion, our reciprocal approach corroborates the findings obtained with the upsC deletion constructs. We conclude that the upsC downstream element acts in a position-dependent manner in inhibiting both RNA PolII-dependent transcription and protein translation. Bypassing of this restriction was only possible through an extremely rare recombination event, which underscores the efficiency at which the upsC 5' UTR inhibits translation.

Discussion

Mutually exclusive transcription of gene families is widespread in nature and facilitates phenotypic plasticity within an isogenic background. The importance

of this type of regulation is exemplified by antigenic variation in unicellular pathogens as a prime strategy to secure survival and transmission. In *Trypanosoma brucei*, the causing agent of African sleeping sickness, mutually exclusive transcription of variant surface glycoprotein genes is carried out by an extra-nucleolar RNA polI-containing body (Navarro and Gull, 2001). Another paradigm of mutual exclusion is that of singular odorant receptor (OR) gene choice in individual olfactory neurons in mammals (McClintock, 2010). Regulatory DNA elements are implicated in exclusive transcription of one out of over a thousand OR genes including the DNA enhancer H (Lomvardas et al., 2006; Fuss et al., 2007), promoter sequences (Vassalli et al., 2002; Qasba and Reed, 1998) and OR coding regions (Nguyen et al., 2007). Negative feedback on the protein level is also implicated in the one receptor-one neuron rule (Shykind et al., 2004; Lewcock and Reed, 2004; Serizawa et al., 2003). In remarkable analogy to mutually exclusive *var* regulation, Lomvardas and colleagues recently described a functional association of H3K9me3 and H3K4me3 with silenced and active OR loci, respectively (Magklara et al., 2011). However, due to the complexity of the control mechanisms involved, we still lack detailed knowledge as to how mutually exclusive transcription is achieved in any system. In this study, we developed and successfully applied a complementary functional approach to study mutual exclusion in *P. falciparum var* gene transcription. For the first time, we identified *cis*-acting entities as essential mediators of *var* gene activation and singular gene choice.

var gene transcription is mediated by RNA polII and occurs stage-specifically by activation in ring stage parasites and subsequent repression or poising during the rest of the IDC (Lopez-Rubio et al., 2007; Kyes et al., 2007). We identified an UAS element essential for upsC promoter activation. This element activates a heterologous minimal promoter both on episomal plasmids and at the endogenous locus. The position of this element upstream of the natural TSS, and the competence to activate transcription from a heterologous promoter are attributes inherently associated with the role of UAS elements in transcriptional activation (Levine and Tjian, 2003). Our results are therefore consistent with the sequence-specific recruitment of a transcriptional activator(s) by the UAS to orchestrate the assembly of the pre-

initiation complex (PIC) and/or to activate PolII-dependent transcription. The fact that this element functions autonomously in a euchromatic context implies a rather ubiquitous presence of this *trans*-acting factor and precludes a restricted role in mutually exclusive activation.

The alternative upstream TSS identified in this study may also play a role in *var* regulation. Active transcription through the promoter region was shown to be required for chromatin remodelling and transcription of the *S. pombe fbp1+* locus (Hirota et al., 2008) and it will be interesting to test if similar processes act on the level of individual *var* gene activation. Analysis of the transcripts emanating from this upstream TSS uncovered a role for the upsC 5' UTR in translational repression. These alternative transcripts gave rise to significantly higher protein levels compared to transcripts carrying the natural 5' UTR. Conversely, insertion of bps -519 to -1 of the upsC 5' UTR into the context of the heterologous *kahrp* promoter rendered these transcripts incompetent for translation. In this particular example the level of translational inhibition was compelling and irreversible; 3D7/pBK$_{min}$C4 parasites were only able to overcome this repression through a spontaneous gene conversion event placing the drug resistance marker under control of a different promoter. Interestingly, two recent studies reported inefficient translation of the PfEMP1 variant VAR2CSA despite the presence of high levels of *var2csa* transcripts (Amulic et al., 2009; Mok et al., 2008). Deitsch and colleagues showed that this translational repression was controlled by an upstream open reading frame (uORF) in the *var2csa* 5' UTR (Amulic et al., 2009). Similar to uORFs, upstream translation initiation codons (uAUGs) can lead to a substantial decrease in translation efficiency and have important roles in translational control during development and conditions of cell stress (van der Velden and Thomas, 1999; Spriggs et al., 2010). The upsC 5' UTR investigated here contains one sense uORF and the remarkable number of 17 uAUGs within 150 bps upstream of the ATG. Notably, we find that uAUGs are generally enriched in *var* 5' UTRs compared to other ring stage-specific genes (Figure S7). Together, these findings indicate that *P. falciparum* may in fact use this type of control to modulate expression of PfEMP1 under different environmental conditions.

The current model of mutually exclusive *var* transcription is based on a number of recent studies (Voss et al., 2007; Voss et al., 2006; Duraisingh et al., 2005; Dzikowski et al., 2007; Lopez-Rubio et al., 2009; Ralph et al., 2005). It postulates the existence of a physically restricted perinuclear zone dedicated to the expression of a single *var* gene. Activation requires entry into this zone with concomitant substitution of the formally active locus. The process of activation is linked to the removal of H3K9me3/PfHP1 and deposition of H3K9ac and H3K4me2/3 marks predominantly along the region downstream of the TSS (Lopez-Rubio et al., 2007). Here, we identified a deletion downstream of the TSS as the common denominator of all six promoter variants that escaped mutually exclusive activation. Unlike full-length promoters, activation of promoters lacking this element does not occur at the expense of, but in parallel to, the transcription of an endogenous *var* gene. Notably, this deletion does not alter the relative activity of the upsC promoter indicating that the MEE represents a functional entity with intrinsic activity that is uncoupled from transcription initiation. The specific binding of an unknown nuclear protein or protein complex to the MEE corroborates this hypothesis. It is important to appreciate that this binding activity was soluble in high salt nuclear extracts, ruling out a restricted distribution of this factor to a sequestered perinuclear expression site.

To the best of our knowledge, this is the first description of a DNA-protein interaction implicated in mutually exclusive locus recognition. Although the exact function of this interaction remains to be discovered, we favour a role in transcriptional repression of *var* genes. For instance, the mutual exclusion element may orchestrate the assembly of a repressive complex that prevents PIC recruitment or RNA polII progression. Our finding that the upsC downstream element including the MEE caused a two-fold reduction in *kahrp* promoter activity is consistent with such a scenario, although the extent of repression in this heterologous context was only moderate. However, given the specific role of mutually exclusive transcription in *var* regulation, it is entirely feasible that cooperative interactions with additional *var*-specific promoter elements would mediate solid repression in the natural context. At the same time, the binding of a specific factor to the mutual exclusion element may earmark *var* loci for mutually exclusive activation. Clearly, additional

experiments tailored towards dissecting the exact role of this interaction are now required to test these hypotheses.

Using promoter deletion analyses combined with ectopic insertion of regulatory *var* elements into a euchromatic locus we were able to systematically reconstruct some of the control steps of *var* gene activation and mutual exclusion. Based on these novel findings, and by integrating current knowledge, we propose a speculative mechanistic model for mutually exclusive *var* gene activation (Figure 7). The position of *var* loci in heterochromatic perinuclear clusters prevents accessibility to specific and general transcription factors and is probably the most important determinant of transcriptional inactivity (Freitas-Junior et al., 2005; Duraisingh et al., 2005; Freitas-Junior et al., 2000; Flueck et al., 2009; Voss et al., 2006; Lopez-Rubio et al., 2009; Ralph et al., 2005; Perez-Toledo et al., 2009). A MEE-interacting factor or complex (MIF) may bind downstream of the TSS to reinforce repression and to prevent leaky transcription from silenced loci. Importantly, based on the escape of mutually exclusive activation in absence of the MEE, and the repression conferred by this element in the *kahrp* locus, this complex may also support transcriptional repression in absence of heterochromatic marks. Such a function may be crucial in keeping *var* genes repressed that are positioned within the previously identified euchromatic zone at the nuclear periphery (Ralph et al., 2005). Singular *var* gene choice may occur through the recognition of MEE/MIF, or an alternative *var* locus-specific sequence tag, by the unique *var* gene expression site (VES) (Voss et al., 2006; Duraisingh et al., 2005; Dzikowski et al., 2007; Lopez-Rubio et al., 2009). Once locked in, the VES may trigger the exchange of H3K9me3/PfHP1 with H3K4me2/3 and H3K9ac marks and the dissociation of the repressive MIF complex. Physical association of the active *var* locus with the VES may also play a crucial role in epigenetic memory, i.e. in keeping the *var* gene in place for re-activation in daughter cells (Lopez-Rubio et al., 2007).

This model describes a novel logic in mutually exclusive gene expression and provides us with an informed working hypothesis for further functional dissection of the mechanisms orchestrating singular *var* gene choice. In particular, targeted identification of the proteins or protein complexes interacting with the regulatory elements characterised in this study will be a

promising and exciting avenue to pursue. Detailed insight into this complex regulatory system is important for our understanding of immune evasion and virulence of *P. falciparum* and other pathogens. Furthermore, our results will also help to understand conceptually similar systems in other organisms.

Experimental Procedures

Parasite culture and transfection

P. falciparum 3D7 parasites were cultured as described previously (Trager and Jenson, 1978). Growth synchronisation was achieved by repeated sorbitol lysis (Lambros and Vanderberg, 1979). Transfections were performed as described (Voss et al., 2006). Parasites were selected on 2.5ug/ml blasticidin-S-HCl and 4nM WR99210. Transfection constructs are described in Supplemental Experimental Procedures.

Quantitative reverse transcription PCR

qPCR was performed on reverse transcribed total RNA and gDNA isolated from synchronous parasite cultures. A detailed protocol, relative transcript calculation and primer sequences are provided in Supplemental Experimental Procedures.

Southern and Northern blot analysis

gDNA was digested with appropriate restriction enzymes overnight and separated on 0.7% agarose gels. Total RNA was isolated from saponin-released parasites using Tri Reagent (Ambion). RNA was glyoxylated for 1h at 60°C in five volumes glyoxal reaction mixture and electrophoresis was performed using 1xBPTE-buffered agarose gels (Sambrook and Russell, 2001). Blots were probed with ^{32}P-dATP-labeled h*dhfr*, *kahrp*, *hsp86* and PFD0005w PCR fragments. Membranes were stripped by boiling in 0.1% SDS for 15min in between hybridisations.

Western blot analysis

Detection of hDHFR-GFP and GAPDH (loading control) was performed on whole cell lysates. Primary antibody dilutions were: mouse anti-GFP (Roche

Diagnostics, 11814460001), 1:1000; monoclonal mouse anti-GAPDH 1-10B (kind gift of Claudia Daubenberger), 1:20'000. PfEMP1 was extracted from trophozoite-infected RBC pellets (Triton X-100-insoluble/SDS soluble fraction) as described (van Schravendijk et al., 1993). Extracts were separated by SDS-PAGE using 5% polyacrylamide gels using Tris-glycine or Tris-acetate buffers. PfEMP1 was detected using the monoclonal mouse anti-PfEMP1 antibody 1B/6H-1 (Duffy et al., 2002) diluted 1:500.

Electromobility shift assay

High salt nuclear extracts and EMSAs were prepared and carried out as described (Voss et al., 2002) with the following modifications. Proteins were extracted with 500mM KCl and incubated with 20fmol of radiolabeled probe in 1xEMSA buffer containing 100ng of poly(dA-dT) as nonspecific competitor. Complementary oligonucleotide sequences used to generate double stranded probes and competitors are listed in Table S1.

Acknowledgments

We are grateful to Dania Mueller for help in qRT-PCR experiments and Sandra Birrer for help in cloning and transfection. We thank Tim-Wolf Gilberger for critical reading of the manuscript. N.M.B.B. received a Boehringer Ingelheim PhD scholarship (www.bifonds.de). This work is supported by the Swiss National Science Foundation (PP00A-110835; PP00P3_130203; www.snf.ch). The authors declare that no conflict of interest exists.

References

Amulic,B., Salanti,A., Lavstsen,T., Nielsen,M.A., and Deitsch,K.W. (2009). An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. PLoS. Pathog. *5*, e1000256.

Baruch,D.I., Gormely,J.A., Ma,C., Howard,R.J., and Pasloske,B.L. (1996). Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. Proc. Natl. Acad. Sci. U. S. A *93*, 3497-3502.

Beeson,J.G. and Duffy,P.E. (2005). The immunology and pathogenesis of malaria during pregnancy. Curr. Top. Microbiol. Immunol. *297*, 187-227.

Biggs,B.A., Gooze,L., Wycherley,K., Wollish,W., Southwell,B., Leech,J.H., and Brown,G.V. (1991). Antigenic variation in Plasmodium falciparum. Proc. Natl. Acad. Sci. U. S. A *88*, 9171-9174.

Calderwood,M.S., Gannoun-Zaki,L., Wellems,T.E., and Deitsch,K.W. (2003). Plasmodium falciparum var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. J. Biol. Chem. *278*, 34125-34132.

Chookajorn,T., Dzikowski,R., Frank,M., Li,F., Jiwani,A.Z., Hartl,D.L., and Deitsch,K.W. (2007). Epigenetic memory at malaria virulence genes. Proc. Natl. Acad. Sci. U. S. A *104*, 899-902.

Deitsch,K.W., Calderwood,M.S., and Wellems,T.E. (2001). Malaria. Cooperative silencing elements in var genes. Nature *412*, 875-876.

Deitsch,K.W., del,P.A., and Wellems,T.E. (1999). Intra-cluster recombination and var transcription switches in the antigenic variation of Plasmodium falciparum. Mol. Biochem. Parasitol. *101*, 107-116.

Deitsch,K.W., Lukehart,S.A., and Stringer,J.R. (2009). Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat. Rev. Microbiol. *7*, 493-503.

Duffy,M.F., Brown,G.V., Basuki,W., Krejany,E.O., Noviyanti,R., Cowman,A.F., and Reeder,J.C. (2002). Transcription of multiple var genes by individual, trophozoite-stage Plasmodium falciparum cells expressing a chondroitin sulphate A binding phenotype. Mol. Microbiol. *43*, 1285-1293.

Duraisingh,M.T., Voss,T.S., Marty,A.J., Duffy,M.F., Good,R.T., Thompson,J.K., Freitas-Junior,L.H., Scherf,A., Crabb,B.S., and Cowman,A.F. (2005). Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell *121*, 13-24.

Dzikowski,R. and Deitsch,K.W. (2009). Genetics of antigenic variation in Plasmodium falciparum. Curr. Genet. *55*, 103-110.

Dzikowski,R., Frank,M., and Deitsch,K. (2006). Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. PLoS. Pathog. *2*, e22.

Dzikowski,R., Li,F., Amulic,B., Eisberg,A., Frank,M., Patel,S., Wellems,T.E., and Deitsch,K.W. (2007). Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO Rep. *8*, 959-965.

Epp,C., Li,F., Howitt,C.A., Chookajorn,T., and Deitsch,K.W. (2009). Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite Plasmodium falciparum. RNA. *15*, 116-127.

Flueck,C., Bartfai,R., Volz,J., Niederwieser,I., Salcedo-Amaya,A.M., Alako,B.T., Ehlgen,F., Ralph,S.A., Cowman,A.F., Bozdech,Z., et al. (2009). Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS. Pathog. *5*, e1000569.

Frank,M., Dzikowski,R., Costantini,D., Amulic,B., Berdougo,E., and Deitsch,K. (2006). Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. J. Biol. Chem. *281*, 9942-9952.

Freitas-Junior,L.H., Bottius,E., Pirrit,L.A., Deitsch,K.W., Scheidig,C., Guinet,F., Nehrbass,U., Wellems,T.E., and Scherf,A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature *407*, 1018-1022.

Freitas-Junior,L.H., Hernandez-Rivas,R., Ralph,S.A., Montiel-Condado,D., Ruvalcaba-Salazar,O.K., Rojas-Meza,A.P., Mancio-Silva,L., Leal-Silvestre,R.J., Gontijo,A.M., Shorte,S., et al. (2005). Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell *121*, 25-36.

Fuss,S.H., Omura,M., and Mombaerts,P. (2007). Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell *130*, 373-384.

Gannoun-Zaki,L., Jost,A., Mu,J., Deitsch,K.W., and Wellems,T.E. (2005). A silenced Plasmodium falciparum var promoter can be activated in vivo through spontaneous deletion of a silencing element in the intron. Eukaryot. Cell *4*, 490-492.

Gardner,J.P., Pinches,R.A., Roberts,D.J., and Newbold,C.I. (1996). Variant antigens and endothelial receptor adhesion in Plasmodium falciparum. Proc. Natl. Acad. Sci. U. S. A *93*, 3503-3508.

Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., et al. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. Nature *419*, 498-511.

Hirota,K., Miyoshi,T., Kugou,K., Hoffman,C.S., Shibata,T., and Ohta,K. (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature *456*, 130-134.

Howitt,C.A., Wilinski,D., Llinas,M., Templeton,T.J., Dzikowski,R., and Deitsch,K.W. (2009). Clonally variant gene families in Plasmodium falciparum share a common activation factor. Mol. Microbiol. *73*, 1171-1185.

Kyes,S., Christodoulou,Z., Pinches,R., Kriek,N., Horrocks,P., and Newbold,C. (2007). Plasmodium falciparum var gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. Mol. Microbiol. *63*, 1237-1247.

Lambros,C. and Vanderberg,J.P. (1979). Synchronization of Plasmodium falciparum erythrocytic stages in culture. J. Parasitol. *65*, 418-420.

Lanzer,M., de Bruin,D., and Ravetch,J.V. (1992). A sequence element associated with the Plasmodium falciparum KAHRP gene is the site of developmentally regulated protein-DNA interactions. Nucleic Acids Res. *20*, 3051-3056.

Leech,J.H., Barnwell,J.W., Miller,L.H., and Howard,R.J. (1984). Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. J. Exp. Med. *159*, 1567-1575.

Levine,M. and Tjian,R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.

Lewcock,J.W. and Reed,R.R. (2004). A feedback mechanism regulates monoallelic odorant receptor expression. Proc. Natl. Acad. Sci. U. S. A *101*, 1069-1074.

Lomvardas,S., Barnea,G., Pisapia,D.J., Mendelsohn,M., Kirkland,J., and Axel,R. (2006). Interchromosomal interactions and olfactory receptor choice. Cell *126*, 403-413.

Lopez-Rubio,J.J., Gontijo,A.M., Nunes,M.C., Issar,N., Hernandez,R.R., and Scherf,A. (2007). 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol. Microbiol. *66*, 1296-1305.

Lopez-Rubio,J.J., Mancio-Silva,L., and Scherf,A. (2009). Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell Host. Microbe *5*, 179-190.

MacPherson,G.G., Warrell,M.J., White,N.J., Looareesuwan,S., and Warrell,D.A. (1985). Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. Am. J. Pathol. *119*, 385-401.

Magklara,A., Yen,A., Colquitt,B.M., Clowney,E.J., Allen,W., Markenscoff-Papadimitriou,E., Evans,Z.A., Kheradpour,P., Mountoufaris,G., et al. (2011). An epigenetic signature for monoallelic olfactory receptor expression. Cell *145*, 555-570.

Marty,A.J., Thompson,J.K., Duffy,M.F., Voss,T.S., Cowman,A.F., and Crabb,B.S. (2006). Evidence that Plasmodium falciparum chromosome end clusters are cross-linked by protein and are the sites of both virulence gene silencing and activation. Mol. Microbiol. *62*, 72-83.

McClintock,T.S. (2010). Achieving singularity in mammalian odorant receptor gene choice. Chem. Senses *35*, 447-457.

Mok,B.W., Ribacke,U., Rasti,N., Kironde,F., Chen,Q., Nilsson,P., and Wahlgren,M. (2008). Default Pathway of var2csa switching and translational repression in Plasmodium falciparum. PLoS. ONE. *3*, e1982.

Morrison,L.J., Marcello,L., and McCulloch,R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cell Microbiol. *11*, 1724-1734.

Navarro,M. and Gull,K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. Nature *414*, 759-763.

Nguyen,M.Q., Zhou,Z., Marks,C.A., Ryba,N.J., and Belluscio,L. (2007). Prominent roles for odorant receptor coding sequences in allelic exclusion. Cell *131*, 1009-1017.

O'Donnell,R.A., Freitas-Junior,L.H., Preiser,P.R., Williamson,D.H., Duraisingh,M., McElwain,T.F., Scherf,A., Cowman,A.F., and Crabb,B.S. (2002). A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. EMBO J. *21*, 1231-1239.

Perez-Toledo,K., Rojas-Meza,A.P., Mancio-Silva,L., Hernandez-Cuevas,N.A., Delgadillo,D.M., Vargas,M., Martinez-Calvillo,S., Scherf,A., and Hernandez-Rivas,R. (2009). Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. *37*, 2596-2606.

Pongponratn,E., Riganti,M., Punpoowong,B., and Aikawa,M. (1991). Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. Am. J. Trop. Med. Hyg. *44*, 168-175.

Prucca,C.G. and Lujan,H.D. (2009). Antigenic variation in Giardia lamblia. Cell Microbiol. *11*, 1706-1715.

Qasba,P. and Reed,R.R. (1998). Tissue and zonal-specific expression of an olfactory receptor transgene. J. Neurosci. *18*, 227-236.

Ralph,S.A., Scheidig-Benatar,C., and Scherf,A. (2005). Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proc. Natl. Acad. Sci. U. S. A *102*, 5414-5419.

Reeder,J.C., Cowman,A.F., Davern,K.M., Beeson,J.G., Thompson,J.K., Rogerson,S.J., and Brown,G.V. (1999). The adhesion of Plasmodium falciparum-infected erythrocytes to chondroitin sulfate A is mediated by P. falciparum erythrocyte membrane protein 1. Proc. Natl. Acad. Sci. U. S. A *96*, 5198-5202.

Roberts,D.J., Craig,A.G., Berendt,A.R., Pinches,R., Nash,G., Marsh,K., and Newbold,C.I. (1992). Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature *357*, 689-692.

Salcedo-Amaya,A.M., van Driel,M.A., Alako,B.T., Trelle,M.B., van den Elzen,A.M., Cohen,A.M., Janssen-Megens,E.M., van de Vegte-Bolmer,M., Selzer,R.R., Iniguez,A.L., et al. (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc. Natl. Acad. Sci. U. S. A.

Sambrook,J. and Russell,D. (2001). Molecular Cloning: A Laboratory Manual (CSHL Press).

Scherf,A., Hernandez-Rivas,R., Buffet,P., Bottius,E., Benatar,C., Pouvelle,B., Gysin,J., and Lanzer,M. (1998). Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. EMBO J. *17*, 5418-5426.

Serizawa,S., Miyamichi,K., Nakatani,H., Suzuki,M., Saito,M., Yoshihara,Y., and Sakano,H. (2003). Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. Science *302*, 2088-2094.

Shykind,B.M., Rohani,S.C., O'Donnell,S., Nemes,A., Mendelsohn,M., Sun,Y., Axel,R., and Barnea,G. (2004). Gene switching and the stability of odorant receptor gene choice. Cell *117*, 801-815.

Smith,J.D., Chitnis,C.E., Craig,A.G., Roberts,D.J., Hudson-Taylor,D.E., Peterson,D.S., Pinches,R., Newbold,C.I., and Miller,L.H. (1995). Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell *82*, 101-110.

Spriggs,K.A., Bushell,M., and Willis,A.E. (2010). Translational regulation of gene expression during conditions of cell stress. Mol. Cell *40*, 228-237.

Su,X.Z., Heatwole,V.M., Wertheimer,S.P., Guinet,F., Herrfeldt,J.A., Peterson,D.S., Ravetch,J.A., and Wellems,T.E. (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. Cell *82*, 89-100.

Trager,W. and Jenson,J.B. (1978). Cultivation of malarial parasites. Nature *273*, 621-622.

van der Velden,A.W. and Thomas,A.A. (1999). The role of the 5' untranslated region of an mRNA in translation regulation during development. Int. J. Biochem. Cell Biol. *31*, 87-106.

van Schravendijk,M.R., Pasloske,B.L., Baruch,D.I., Handunnetti,S.M., and Howard,R.J. (1993). Immunochemical characterization and differentiation of two approximately 300-kD erythrocyte membrane-associated proteins of Plasmodium falciparum, PfEMP1 and PfEMP3. Am. J. Trop. Med. Hyg. *49*, 552-565.

Vassalli,A., Rothman,A., Feinstein,P., Zapotocky,M., and Mombaerts,P. (2002). Minigenes impart odorant receptor-specific axon guidance in the olfactory bulb. Neuron *35*, 681-696.

Voss,T.S., Healer,J., Marty,A.J., Duffy,M.F., Thompson,J.K., Beeson,J.G., Reeder,J.C., Crabb,B.S., and Cowman,A.F. (2006). A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. Nature *439*, 1004-1008.

Voss,T.S., Mini,T., Jenoe,P., and Beck,H.P. (2002). Plasmodium falciparum possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript. J. Biol. Chem. *277*, 17493-17501.

Voss,T.S., Thompson,J.K., Waterkeyn,J., Felger,I., Weiss,N., Cowman,A.F., and Beck,H.P. (2000). Genomic distribution and functional characterisation of two distinct and conserved Plasmodium falciparum var gene 5' flanking sequences 109. Mol. Biochem. Parasitol. *107*, 103-115.

Voss,T.S., Tonkin,C.J., Marty,A.J., Thompson,J.K., Healer,J., Crabb,B.S., and Cowman,A.F. (2007). Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum 23. Mol. Microbiol. *66*, 139-150.

World Health Organisation. (2010). World Malaria Report 2010 (Geneva: WHO Press).

Figure legends

Figure 1. Functional *var* promoter activity mapping by serial deletion analysis

(A) Schematic map of pBC. The PFL1960w upsC upstream sequence controls transcription of h*dhfr-gfp*. The approximate position of the TSS is indicated (Deitsch et al., 1999). The *bsd* resistance cassette selects for stably transfected parasites. The *var* intron is indicated by a thick dashed line. pBC descendants were obtained by replacing the upsC promoter with truncated sequences using *Bgl*II and *Not*I. hsp86 5', *hsp86* promoter; Pb DT 3', *P. berghei dhfr*-thymidilate synthase terminator; rep20, 0.5kb TARE6 repeat element; hrp2 3'; histidine-rich protein 2 terminator.

(B) Activities of full-length and truncated promoters in WR-selected parasites. Deletions are represented by dashed lines. Numbers represent nucleotide positions in relation to the ATG. Successful WR-selection is indicated by check marks. Values represent relative h*dhfr-gfp* transcripts normalised against plasmid copy number and transcription of PF13_0170. Values represent the average of three independent experiments (two replicates for 3D7/pBC1 and 3D7/pBC2) (mean +/- SEM). n.a., not applicable.

Figure 2. An upsC UAS element activates the minimal promoter K_{min}

(A) Schematic map of pBK_{min}-RI. The position of the *kahrp* TSS is indicated (Lanzer et al., 1992). See also Figure S1.

(B) Schematic map of pBK_{min}-RI concatamers integrated into the endogenous *kahrp* locus (PFB0100c). See also Figure S1.

(C) Comparison of relative transcript levels produced by the episomal (h*dhfr-gfp* transcripts) and chromosomal (*kahrp* transcripts) minimal promoters, and the *kahrp* wild-type promoter (h*dhfr-gfp* transcripts) in 3D7/pBK_{min}-RI parasites. Values are derived from three independent experiments and represent *msp8*-normalised transcripts (mean +/- SEM). Values for the episomal K_{min} promoter were adjusted for plasmid copy numbers.

(D) Analysis of upsC-K_{min} hybrid promoters. upsC insertions are depicted by bold grey lines. The rep20 element is indicated by a vertical array and the *var* intron by a dashed line. The graph compares relative transcript levels (*msp8*-

normalised) produced by the episomal (h*dhfr-gfp* transcripts, grey bars) and chromosomal (*kahrp* transcripts, black bars) K_{min} and upsC-K_{min} hybrid promoters. Values for episomal promoters are adjusted for plasmid copy numbers and derived from three independent experiments (mean +/- SEM). Values for 3D7/pBK_{min}-RI are identical to those in Figure 2C.

Figure 3. The upsC 5' UTR and TSS affect expression levels

(A) Identification of an alternative upstream TSS (dashed arrow). Full-length and truncated promoters are schematically depicted on top. Transcript size and abundance were estimated by Northern analysis. Ethidium bromide-stained 18S and 28S rRNAs serve as loading control.

(B) Schematic map of the expression cassette in pBCK. The hybrid sequence consists of a functional upsC promoter (-1679 to -914) (grey line) including the TSS fused to the *kahrp* 5' UTR (bps -764 to -1) (black line). The graph shows relative h*dhfr-gfp* transcript levels in WR-selected 3D7/pBC (control) and 3D7/pBCK parasites (normalised against PF13_0170 and plasmid copy number). Values for 3D7/pBCK are derived from three independent experiments (mean +/- SEM).

(C) Semi-quantitative analysis of transcript and protein abundance by parallel qRT-PCR and Western blot. Values represent PF13_0170-normalised h*dhfr-gfp* transcripts produced by a single promoter (first graph) or by a single parasite (second graph) in WR-selected ring stage parasites. *msp*8 transcripts serve as stage-specific control (third graph). Expression of hDHFR-GFP and GAPDH (loading control) in the same parasite samples is shown below. See also Figures S2 and S3.

Figure 4. Functional identification of a mutual exclusion element downstream of the TSS

(A) Promoters are schematically depicted on top. upsC sequences are shown in grey. Deletions are represented by dashed lines. The orange box highlights the region required for mutually exclusive activation. PfEMP1 expression was monitored by Western blot using a mAb against the conserved ATS region. The antibody cross-reacts with human spectrin in uninfected RBCs. PfEMP1

is detected at various sizes above 250kDa. The signal at 160kDa probably represents smaller PfEMP1 species (asterisk). RBC, uninfected RBCs.

(B) The mutual exclusion element maps to a 101bp region downstream of the TSS. Promoters are schematically depicted on top. The orange box identifies the mutual exclusion element (MEE). WR-unselected 3D7/pBC carrying a silenced episomal upsC promoter served as a second negative control. M, size standard; -WR, unselected; +WR, WR-selected.

(C) The MEE is bound by a nuclear factor. EMSA probes and competitors are schematically depicted on top. The autoradiograph shows a specific DNA-protein complex between radio-labeled MEE2 and an unknown nuclear factor(s). Competition was carried out in presence of 5-, 25- and 100-fold molar excess of unlabeled DNA fragments. See also Figure S4.

Figure 5. The upsC downstream element inhibits expression from the heterologous *kahrp* promoter

(A) Schematic map of the $pBK_{min}C4$ plasmid.

(B) Schematic map of the $pBK_{min}C4$ concatamer integrated into the endogenous *kahrp* locus. Restriction sites used in Southern analysis and fragment lengths are indicated and colour-coded. S, *Stu*I; B, *Bgl*II. The Southern blot on *Bgl*II/*Stu*I-digested gDNA shows integration of $pBK_{min}C4$ into the endogenous *kahrp* locus. The membrane was hybridised with h*dhfr* (top) and *kahrp* (bottom). Fragments are colour-coded according to the integration map. wt, size of the *kahrp* fragment in 3D7 wild-type parasites. i, integration event; p, plasmid fragment.

(C) Semi-quantitative analysis of transcript and protein abundance in 3D7/pBK_{min} (control) and 3D7/$pBK_{min}C4$ parasites. h*dhfr-gfp* and *hsp86* (loading control) transcripts are shown in the top panels. Ethidium bromide-stained 18S and 28S rRNAs serve as second loading control. Expression of hDHFR-GFP and GAPGH (loading control) in the same parasite samples is shown below.

(D) The upsC downstream element represses *kahrp* promoter activity. The bars represent the ratio of relative h*dhfr-gfp* (grey bar) and PF13_0170 (control; black bar) transcript levels in 3D7/$pBK_{min}C4$ parasites compared to the 3D7/pBK_{min} control. Results are the mean +/- SEM of six independent

experiments using unselected and WR-selected parasites (three biological replicates each) ($p<0.002$; unpaired t-test).

Figure 6. The upsC 5' UTR mediates irreversible translational repression.

(A) Southern analysis on digested gDNA from unselected and WR-selected 3D7/pBK$_{min}$C4 parasites. Additional h*dhfr*-containing fragments specifically detected in WR-selected parasites are highlighted by pink arrows. Fragments are colour-coded according to the map in Figure 5A. S, *Stu*I; B, *Bgl*II; K, *Kpn*I; i, integration event; p, plasmid fragment.

(B) Schematic map of the gene conversion event in WR-selected 3D7/pBK$_{min}$C4 parasites. The ends of chromosome 2 and 4 (4/2) in unselected (WR-selected) parasites are schematically depicted. Gene IDs (www.plasmoDB.org) are indicated for a subset of genes as reference. The dashed arrow highlights the site of gene conversion. The blue box represents the duplicated region of chromosome 2. The green box represents the region of chromosome 4 that was deleted. The beige box displays a zoom-in view of the gene conversion event and the resulting recombined locus. Detailed mapping and identification of the recombination site is presented in Figures S5 and S6.

(C) h*dhfr-gfp* transcripts are produced from the *var* gene intron on chromosome 4. Values represent relative *var* intron-derived h*dhfr-gfp* (black bars) and *msp8* (open bars, control) transcript levels at three consecutive time points in WR-selected 3D7/pBK$_{min}$C4 parasites (PF13_0170-normalised).

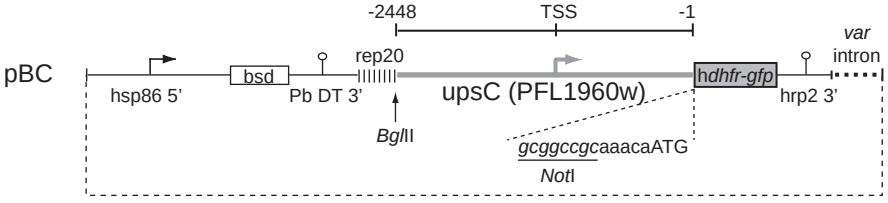
Figure 7. A novel model for singular *var* gene choice

A chromosome end cluster located in a transcriptionally permissive perinuclear region is schematically depicted on top. The unique *var* gene expression site (VES) recognises a single *var* gene through specific unknown DNA motifs at the locus (white hexagon) and/or the MEE element itself (red oval). Interaction of the locus with the VES leads to establishment of a permissive chromatin conformation (green circles) and dissociation of the MIF complex (blue), and allows for RNA polII-dependent transcriptional initiation/elongation. Transcriptional activation involves interaction between

unknown transcription factors (orange) and the UAS (green oval). Additional *var* genes within this subnuclear domain are excluded from the VES and protected from illegitimate transcription by the repressive MIF complex. The function of MIF may be to either block transcriptional elongation or to prevent transcriptional initiation or PIC assembly on the core promoter. *var* genes in heterochromatic perinculear zones that are silenced primarily through their association with H3K9me3/PfHP1 are shown below. The MIF complex may be assembled on each *var* locus to support transcriptional repression.

Figure 1

A


-2448
TSS
-1
var
intron
pBC
rep20
bsd
hdhfr-gfp
hsp86 5'
Pb DT 3'
upsC (PFL1960w)
hrp2 3'
BglII
gcggccgcaaacaATG
NotI

B

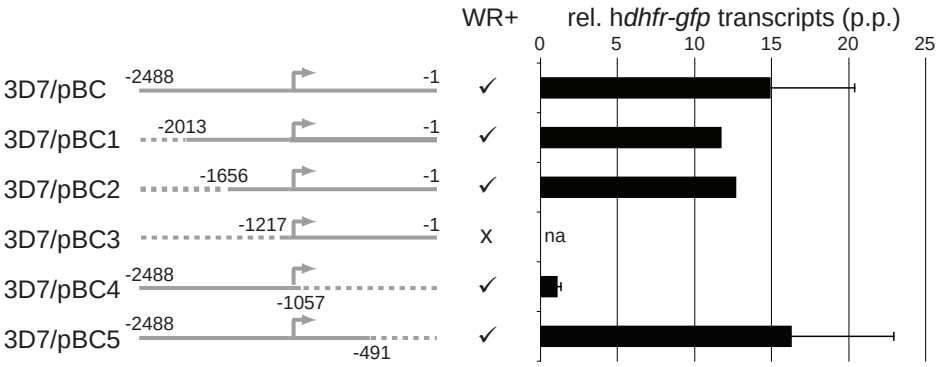

WR+
rel. hdhfr-gfp transcripts (p.p.)
0
5
10
15
20
25
3D7/pBC
-2488
-1
✓
3D7/pBC1
-2013
-1
✓
3D7/pBC2
-1656
-1
✓
3D7/pBC3
-1217
-1
X
na
3D7/pBC4
-2488
-1057
✓
3D7/pBC5
-2488
-491
✓

Figure 2

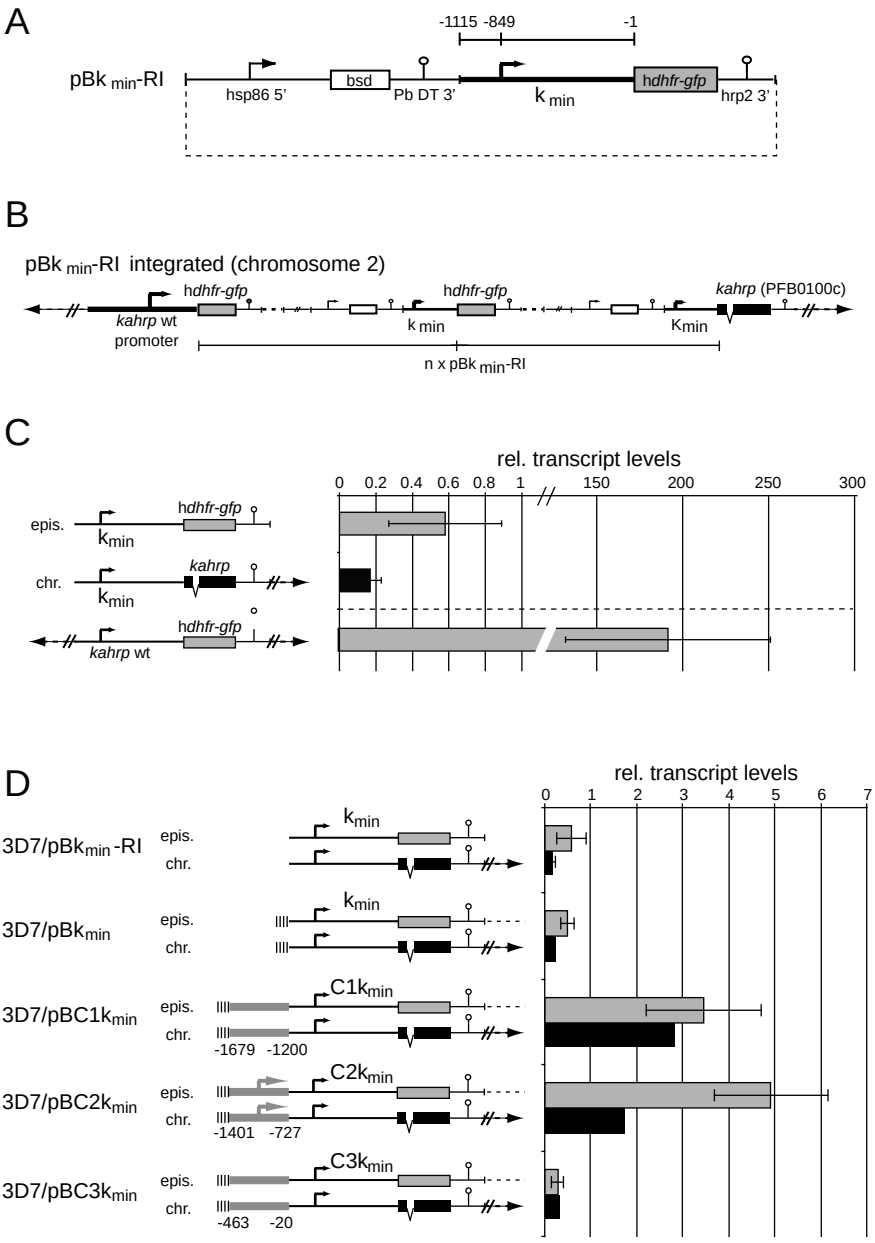

A
-1115 -849 -1
pBk min-RI
hsp86 5'
bsd
Pb DT 3'
k min
hdhfr-gfp
hrp2 3'
B
pBk min-RI integrated (chromosome 2)
hdhfr-gfp
hdhfr-gfp
kahrp (PFB0100c)
kahrp wt
promoter
k min
Kmin
n x pBk min-RI
C
rel. transcript levels
0 0.2 0.4 0.6 0.8 1 150 200 250 300
epis.
hdhfr-gfp
kmin
chr.
kahrp
kmin
hdhfr-gfp
kahrp wt
D
rel. transcript levels
0 1 2 3 4 5 6 7
3D7/pBkmin-RI
epis.
chr.
kmin
3D7/pBkmin
epis.
chr.
kmin
3D7/pBC1kmin
epis.
chr.
C1kmin
-1679 -1200
3D7/pBC2kmin
epis.
chr.
C2kmin
-1401 -727
3D7/pBC3kmin
epis.
chr.
C3kmin
-463 -20

Figure 3

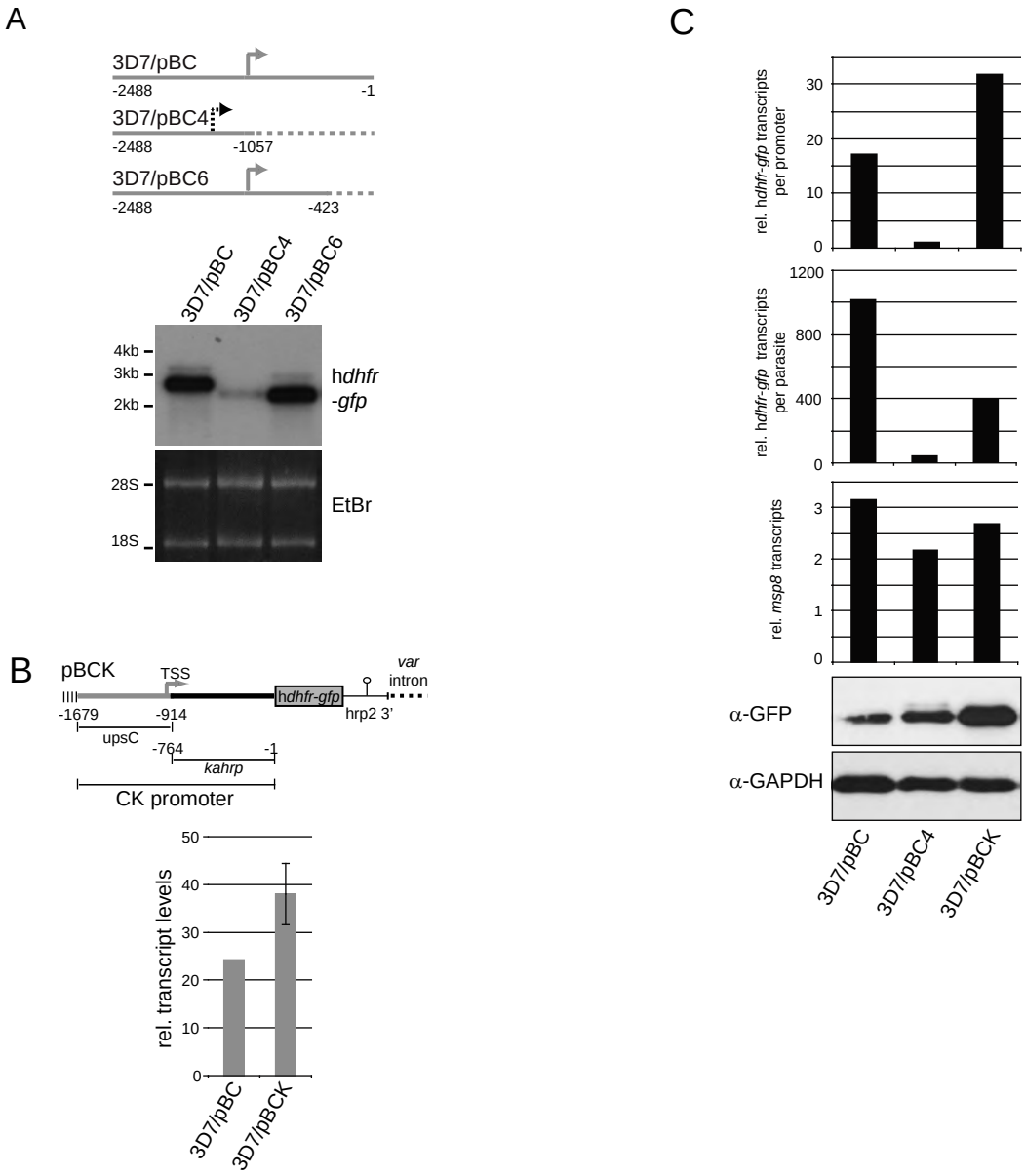

A
3D7/pBC
-2488
-1
3D7/pBC4
-2488
-1057
3D7/pBC6
-2488
-423
3D7/pBC
3D7/pBC4
3D7/pBC6
4kb
3kb
2kb
hdhfr
-gfp
28S
18S
EtBr
B
pBCK
TSS
var
intron
hdhfr-gfp
-1679
-914
hrp2 3'
upsC
-764
-1
kahrp
CK promoter
rel. transcript levels
50
40
30
20
10
0
3D7/pBC
3D7/pBCK
C
rel. hdhfr-gfp transcripts per promoter
30
20
10
0
rel. hdhfr-gfp transcripts per parasite
1200
800
400
0
rel. msp8 transcripts
3
2
1
0
α-GFP
α-GAPDH
3D7/pBC
3D7/pBC4
3D7/pBCK

Figure 4

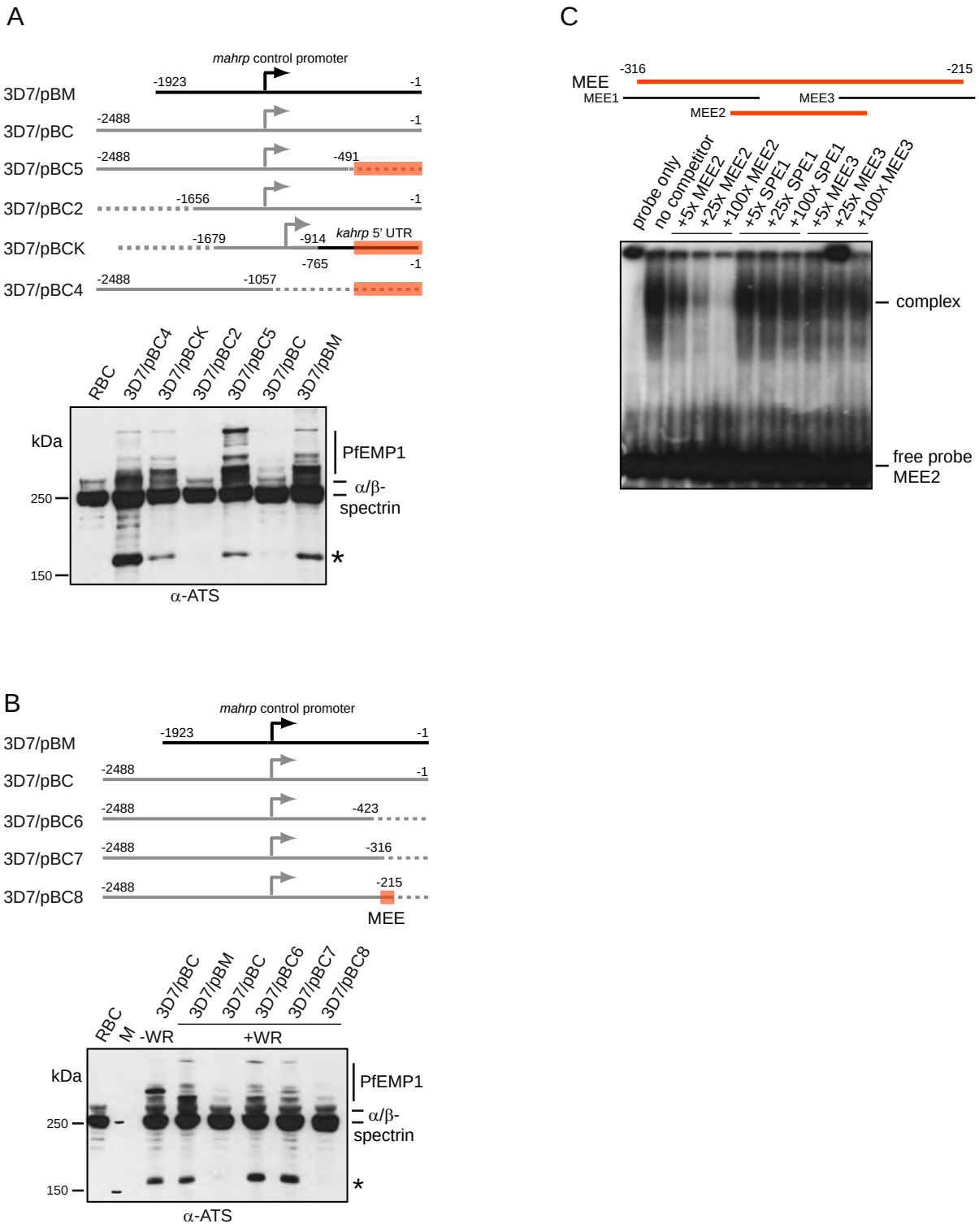

A
mahrp control promoter
3D7/pBM
-1923
-1
3D7/pBC
-2488
-1
3D7/pBC5
-2488
-491
3D7/pBC2
-1656
-1
3D7/pBCK
-1679
-914
kahrp 5' UTR
-765
-1
3D7/pBC4
-2488
-1057
RBC
3D7/pBC4
3D7/pBCK
3D7/pBC2
3D7/pBC5
3D7/pBC
3D7/pBM
kDa
250
150
PfEMP1
α/β-spectrin
*
α-ATS
B
mahrp control promoter
3D7/pBM
-1923
-1
3D7/pBC
-2488
-1
3D7/pBC6
-2488
-423
3D7/pBC7
-2488
-316
3D7/pBC8
-2488
-215
MEE
RBC
M
-WR
+WR
3D7/pBC
3D7/pBM
3D7/pBC
3D7/pBC6
3D7/pBC7
3D7/pBC8
kDa
250
150
PfEMP1
α/β-spectrin
*
α-ATS
C
MEE
-316
-215
MEE1
MEE2
MEE3
probe only
no competitor
+5x MEE2
+25x MEE2
+100x MEE2
+5x SPE1
+25x SPE1
+100x SPE1
+5x MEE3
+25x MEE3
+100x MEE3
complex
free probe
MEE2

Figure 5

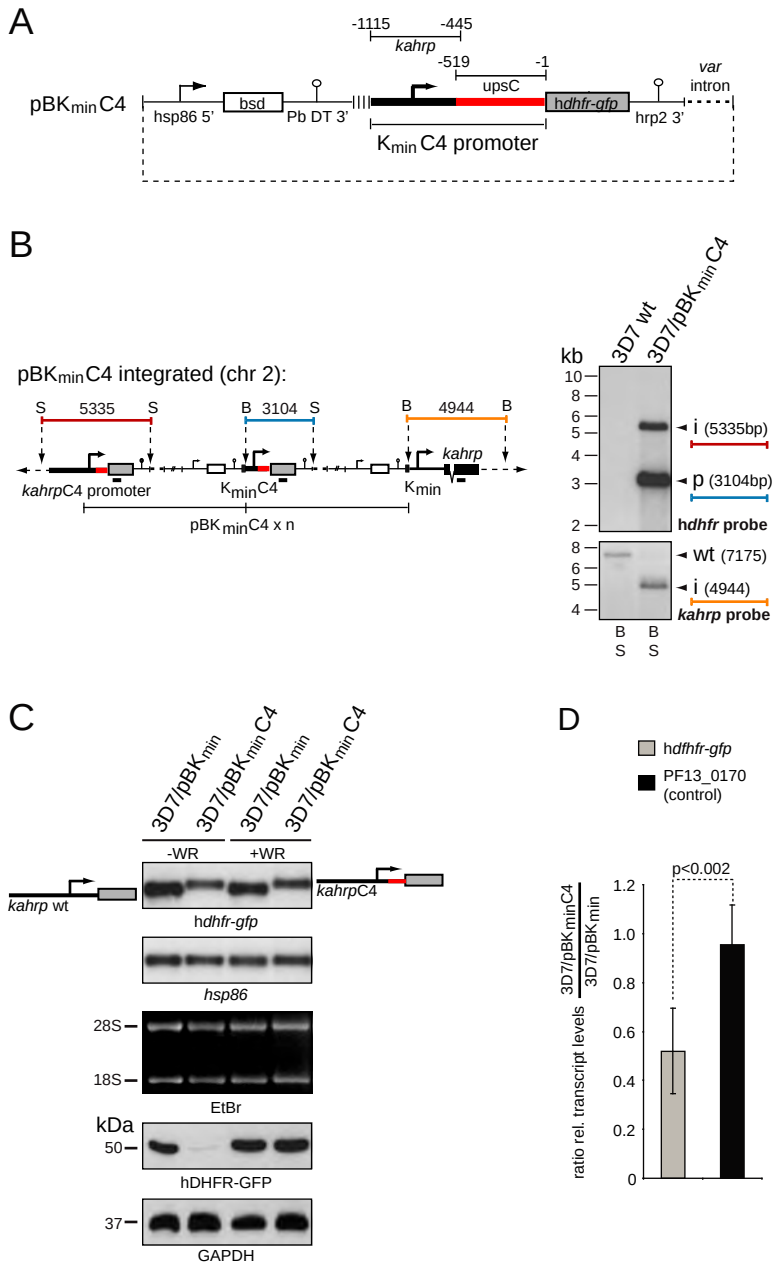

A
-1115
-445
kahrp
-519
-1
upsC
var
intron
pBKminC4
bsd
hsp86 5'
Pb DT 3'
hdhfr-gfp
hrp2 3'
KminC4 promoter
B
pBKminC4 integrated (chr 2):
S
5335
S
B
3104
S
B
4944
B
kahrp
kahrpC4 promoter
KminC4
Kmin
pBKminC4 x n
3D7 wt
3D7/pBKminC4
kb
10
8
6
5
4
3
2
i (5335bp)
p (3104bp)
hdhfr probe
wt (7175)
i (4944)
kahrp probe
B S
B S
C
3D7/pBKmin
3D7/pBKminC4
3D7/pBKmin
3D7/pBKminC4
-WR
+WR
kahrp wt
kahrpC4
hdhfr-gfp
hsp86
28S
18S
EtBr
kDa
50
hDHFR-GFP
37
GAPDH
D
hdhfr-gfp
PF13_0170
(control)
p<0.002
ratio rel. transcript levels
3D7/pBKminC4
3D7/pBKmin
1.2
1.0
0.8
0.6
0.4
0.2
0

Figure 6

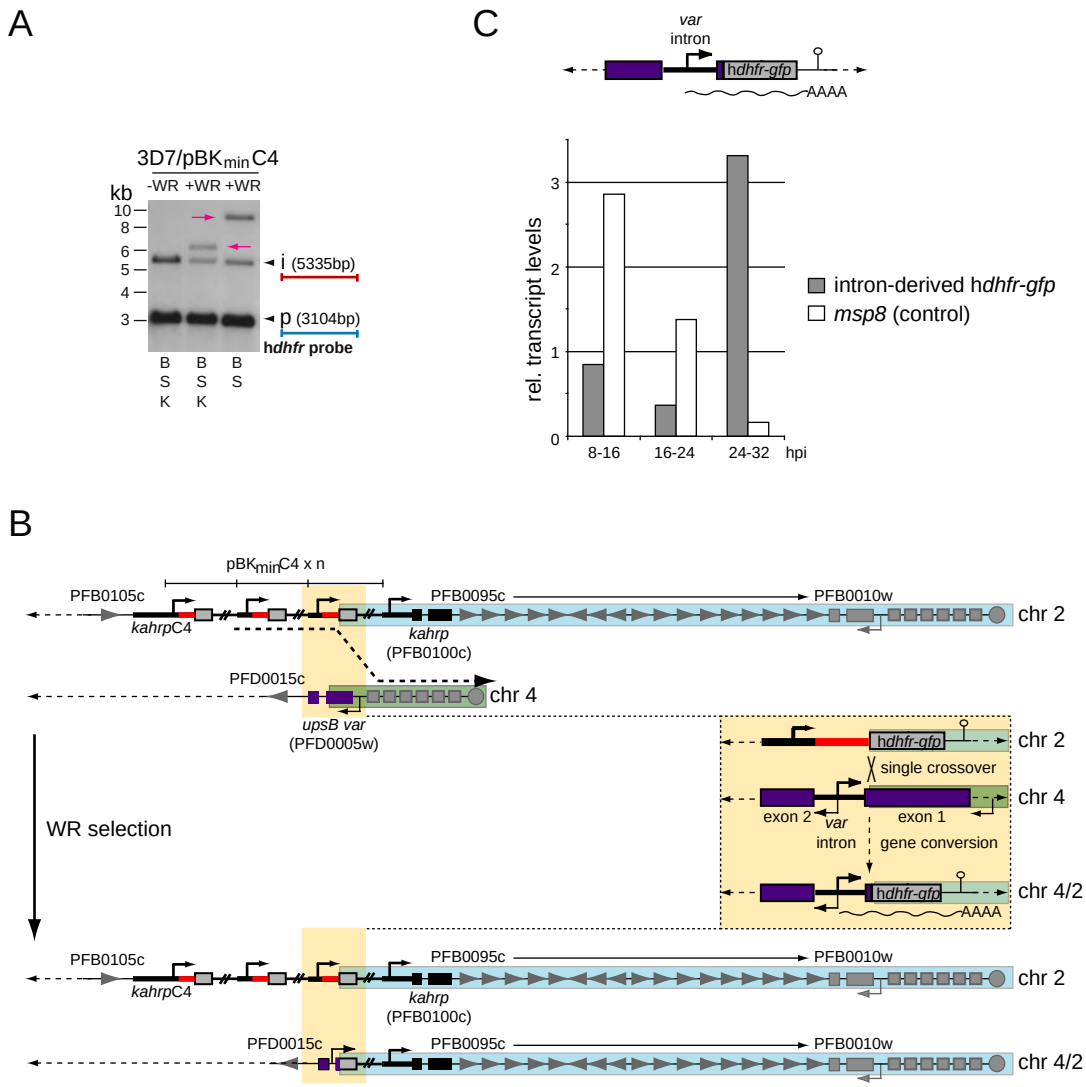
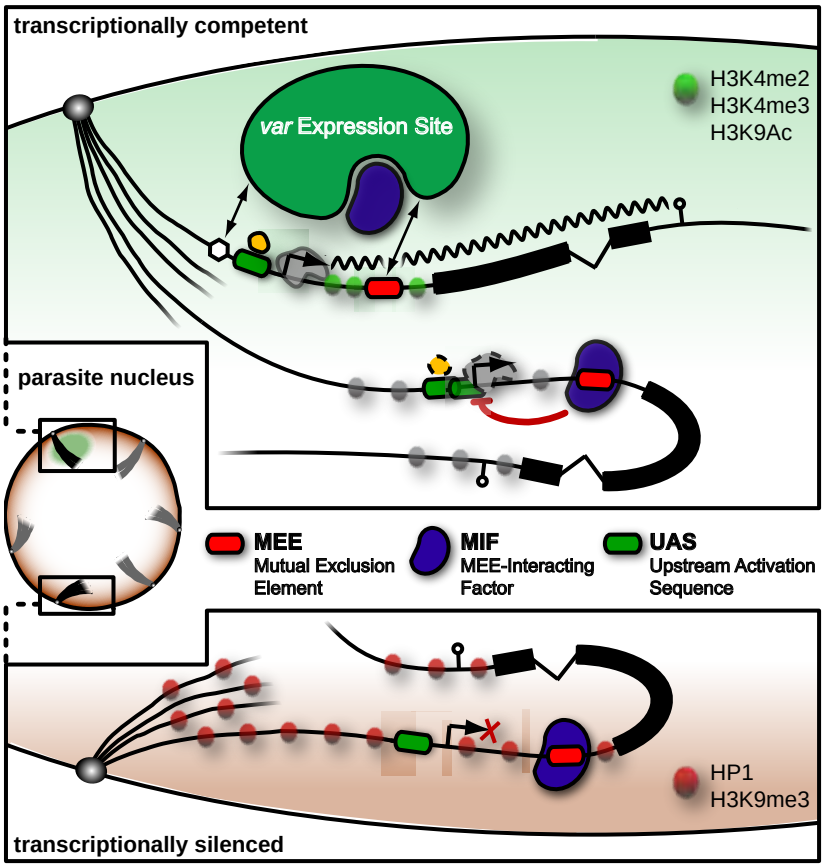

A
3D7/pBK$_{min}$C4
-WR +WR +WR
kb
10
8
6
5
4
3
i (5335bp)
p (3104bp)
hdhfr probe
B S K
B S K
B S
C
var intron
hdhfr-gfp
AAAA
rel. transcript levels
intron-derived hdhfr-gfp
msp8 (control)
8-16
16-24
24-32
hpi
B
pBK$_{min}$C4 x n
PFB0105c
kahrpC4
PFB0095c
kahrp
(PFB0100c)
PFB0010w
chr 2
PFD0015c
upsB var
(PFD0005w)
chr 4
WR selection
hdhfr-gfp
single crossover
exon 2
var intron
exon 1
gene conversion
chr 4/2

Figure 7


transcriptionally competent
var Expression Site
H3K4me2
H3K4me3
H3K9Ac
parasite nucleus
MEE
Mutual Exclusion Element
MIF
MEE-Interacting Factor
UAS
Upstream Activation Sequence
HP1
H3K9me3
transcriptionally silenced

Supplemental Data

Figure S1. Southern analysis of gDNA isolated from parasites presented in Figure 2. (A) Autoradiographs of Southern blots showing episomal maintenance or plasmid integration into the endogenous *kahrp* locus in 3D7/pBK$_{min}$ and 3D7/pBK$_{min}$-RI. gDNA was digested with *Bgl*II and *Hind*III. Blots were probed with a radiolabeled *kahrp* fragment. E, episomal; I, integrated. (B) Autoradiographs of Southern blots showing episomal maintenance or plasmid integration into the endogenous *kahrp* locus in 3D7/pBC1K$_{min}$, 3D7/pBC2K$_{min}$ and 3D7/pBC3K$_{min}$. gDNA was digested with *Bgl*II and *Hind*III. Blots were probed with a radiolabeled *kahrp* fragment. (C) Schematic map of the endogenous *kahrp* locus. (D-F) Schematic maps of the integration events in 3D7/pBK$_{min}$ (D), 3D7/pBC1K$_{min}$ and 3D7/pBC2K$_{min}$ (E), and 3D7/pBC3K$_{min}$ (F). *Bgl*II and *Hind*III restriction sites and length of the corresponding fragments are indicated.

Figure S2. The upsC 5' UTR mediates translational repression. The promoters in pBC and pBC4 are schematically depicted on top. Semi-quantitative analysis of transcript and protein abundance by parallel Northern and Western blot in a time-course experiment. Total protein and RNA were harvested simultaneously from synchronised 3D7/pBC and 3D7/pBC4 parasites at three consecutive time points during the IDC (ring stages, 8-18hpi; late ring stages/early trophozoites, 16-26hpi; late trophozoites/early schizonts, 24-34hpi). Expression of hDHFR-GFP and GAPDH (loading control) was detected with anti-GFP and anti-GAPDH antibodies, respectively (upper panels). Steady-state h*dhfr-gfp* and *hsp86* (loading control) transcripts were detected using radiolabeled h*dhfr* and *hsp86* probes, respectively.

Figure S3. Assessment of hDHFR-GFP expression by live cell microscopy. 3D7/pBC, 3D7/pBC4 and 3D7/pBM (control) cultures were analysed by fluorescence microscopy (see Supplemental Experimental Procedures) to determine the proportion of hDHFR-GFP-expressing parasites in unselected (-WR) and WR-selected (+WR) populations. Images were acquired at 40x magnification. iRBCs were detected using DAPI. The

percentage of hDHFR-GFP-expressing parasites (positive for DAPI and GFP) is indicated and was determined by counting over 200 individual iRBCs (mean 277, +/- 61.2 s.d.). Note that the absence of hDHFR-GFP signal in unselected 3D7/pBC as compared to the control line 3D7/pBM indicates silencing of the upsC promoter.

Figure S4. The core sequence of the 47bp MEE2 motif occurs in a conserved position upstream of 32 *var* genes. The schematic shows the presence and relative position of a 17-22bp consensus motif upstream of 32 *var* genes. This motif forms the core of the 47bp MEE2 element that is bound in a sequence-specific manner by a nuclear protein (see Fig. 4). Gene accession numbers are indicated on the left (www.plasmoDB.org). The colour code clusters *var* genes into the different *var* gene subgroups upsA, upsB, upsC, upsE, upsB/C, and upsB/A. The consensus sequence is shown on top. Red boxes indicate the position of this motif in each upstream region. Numbers on the right represent the position of the first nucleotide of the motif relative to the translation initiation ATG.

Figure S5. Identification of a gene conversion event by Southern blotting and ligation-mediated PCR. (A) The upper map schematically depicts the end of chromosome 2 including the integrated plasmid concatamer (blue box) in 3D7/pBK$_{min}$C4 parasites. *kahrp* promoter sequences are depicted by thick black lines. The upsC 5' UTR sequence is depicted in red. The grey circles and squares represent the telomeric tract and telomere-associated repeat elements (TAREs) 1-6, respectively. Triangles indicate ORFs. The gene accession number refers to the upsB *var* gene PFB0010w. The lower map shows a zoom-in view of the integrated concatamer (blue box). Restriction sites used in Southern analysis are shown by vertical dashed arrows, and expected fragment lengths are indicated and colour-coded. The h*dhfr* probe used for hybridisation is shown below the h*dhfr-gfp* coding sequence (grey box). (B) The autoradiograph shows the hybridisation results obtained after digesting 3D7/pBK$_{min}$C4 gDNA from unselected (-WR) and selected (+WR) populations with *EcoR*V/*Nco*I (red), *EcoR*V/*Spe*I (blue) or *EcoR*V/*Stu*I (green). Note the presence of an additional h*dhfr*-containing fragment after each

double-digest specifically in WR-selected, but not unselected, parasites (highlighted by purple arrows). In each case, the size of the additional fragment (schematically depicted to the right, bottom) is approximately 2 kb smaller than the size of the *EcoR*V/*Nco*I, *EcoR*V/*Spe*I or *EcoR*V/*Stu*I plasmid fragments (depicted to the right, top). This result suggested the presence of a novel *EcoR*V site upstream of a single copy of h*dhfr-gfp* (highlighted in purple). i, integration event; p, plasmid fragment. (C) Ligation-mediated PCR. gDNA from WR-selected 3D7/pBK$_{min}$C4 parasites was digested with *EcoR*V and *Nco*I and ligated into *EcoR*V/*Nco*I-digested pET-41 (EMD Biosciences). To amplify *EcoR*V/*Nco*I restriction fragments containing the h*dhfr* coding sequence a primary PCR reaction was performed using T7 and h*dhfr*_R1 (R1) as forward and reverse primers, respectively. The primary PCR product was diluted 1:200 and used as template for a semi-nested PCR reaction using T7 and h*dhfr*_R2 (R2) as forward and reverse primers, respectively. This reaction yielded three specific amplicons each of which was sequenced using primer h*dhfr*_R3 (R3). Two fragments represented plasmid backbone sequences as a result of semi-specific amplification (data not shown). The nucleotide sequence of the third fragment is shown at the bottom (reversed sequence). It begins with an *EcoR*V site within the intron of *var* gene PFD0005w on chromosome 4 (orange letters) and continues into the 3' end of exon 1 (purple letters). The green letters highlight the 10 bp sequence involved in the recombination event between the *var* and h*dhfr-gfp* loci. The grey box represents the start of the h*dhfr-gfp* coding sequence. Above the nucleotide sequence a detailed view of the recombination event is depicted. A single cross-over occurred between the 10 bp sequence (green letters) directly upstream of the h*dhfr-gfp* reporter (grey box) on chromosome 2, and an identical sequence (green letters) at the very 3' end of exon 1 of *var* gene PFD0005w (purple box) on chromosome 4. As a result, the h*dhfr-gfp* reporter (grey box) was placed under control of the *var* gene intron promoter (orange line) on the reverse strand via gene conversion.

Figure S6. Verification of the gene conversion event between chromosomes 2 and 4 in WR-selected 3D7/pBK$_{min}$C4 parasites. (A) The map schematically depicts the end of chromosome 2 including the integrated

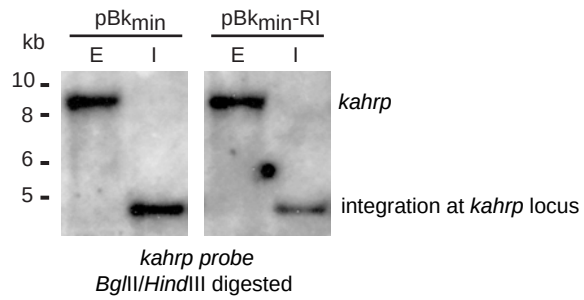
plasmid concatamer (blue box) in 3D7/pBK$_{min}$C4 parasites. *kahrp* promoter sequences are depicted by thick black lines. The upsC 5' UTR sequence is depicted in red. The grey circles and squares represent the telomeric tract and telomere-associated repeat elements (TAREs) 1-6, respectively. Triangles indicate ORFs. The gene accession number refers to the upsB *var* gene PFB0010w. The lower map shows a zoom-in view of the integrated concatamer (blue box). Restriction sites used in Southern analysis are shown by vertical dashed arrows, and expected fragment lengths are indicated and colour-coded. The h*dhfr* probe used for hybridisation is shown below the h*dhfr-gfp* coding sequence (grey box). *EcoR*I sites are absent from the plasmid sequence. Hence, the *EcoR*I sites up- and downstream of the integrated concatamer release a restriction fragment in the size of 6228 bp (chromosomal DNA) plus n times 9475 bps (entire plasmid length) according to the number of copies in the concatamer. (B) The map schematically depicts the end of wild-type chromosome 4 including *var* gene PFD0005w (purple box) in unselected 3D7/pBK$_{min}$C4 parasites. The PFD0005w exon 1 probe used for hybridisation is shown below the coding sequence. The position of the *EcoR*I restriction site downstream of the *var* locus and expected fragment length are indicated. (C) The map schematically depicts the end of chromosome 4 after the gene conversion event between chromosomes 2 and 4 in WR-selected 3D7/pBK$_{min}$C4 parasites ("chromosome 4/2 end"). The border between the green and blue boxes identifies the site of single cross-over recombination. The green and blue boxes represent sequences of the acceptor (chromosome 4) and donor (chromosome 2), respectively, of the gene conversion event. Restriction sites used in Southern analysis are shown by vertical dashed arrows, and expected fragment lengths are indicated and colour-coded. (D) The autoradiograph shows the hybridisation results obtained after digesting 3D7/pBK$_{min}$C4 gDNA from unselected (-WR) and selected (+WR) populations with *EcoR*I, *EcoR*I/*Nco*I or *EcoR*I/*Stu*I. The membrane was hybridised with h*dhfr* (top) and PFD0005w (bottom). Arrows are colour-coded according to the integration maps in A-C and identify the expected restriction fragments. The red, orange and yellow arrows highlight the restriction fragments that contain the single h*dhfr-gfp* cassette driven by the *var* intron promoter on chromosome 4/2 specifically in WR-selected

parasites. Hybridisation with the PFD0005w exon 1 probe highlights the terminal chromosome 4 *EcoR*I fragment in unselected 3D7/pBKminC4 parasites. In WR-selected parasites this signal is absent demonstrating the gene conversion-induced loss of the end of chromosome 4. i, integration event; p, plasmid fragment.

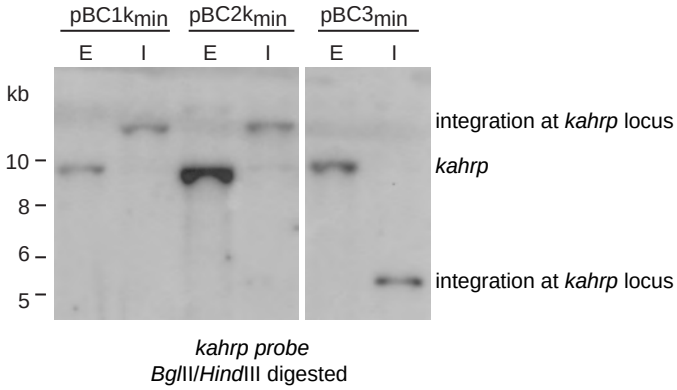
Figure S7. uAUGs are significantly enriched in *var* 5' UTRs. The box reflects the lower and upper quartiles, respectively. The median is indicated by a thick vertical line. Outliers are shown as dots. For each gene, sequences were downloaded from PlasmoDB version 7.2 (www.plasmoDB.org) and the counts of the trinucleotide sequence 'ATG' were assessed in sets of 150bp sequences upstream of the ATG annotated as translation initiation codon. The counts of 'ATG' within each sequence were determined using custom-made perl scripts, and the distribution of counts within each sequence set was visualized as boxplots using R. The *var* gene set includes UTR sequences of 60 genes. The control set consists of UTR sequences of 404 genes with peak transcription in ring stages. Selection of these sequences was based on RNA Seq data (Otto et al., 2010) according to the following criteria: timing of maximal expression: 8 hpi and 16 hpi; maximal expression ratio: 8-fold induction; maximum expression percentile: 30th percentile. uAUGs are significantly enriched in *var* 5' UTRs compared to the control set of ring stage-specific genes ($p=7.56x10^{-11}$; Welch t-test).

Figure S1

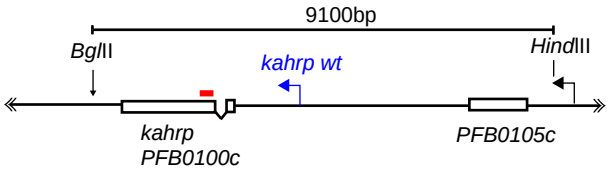
A



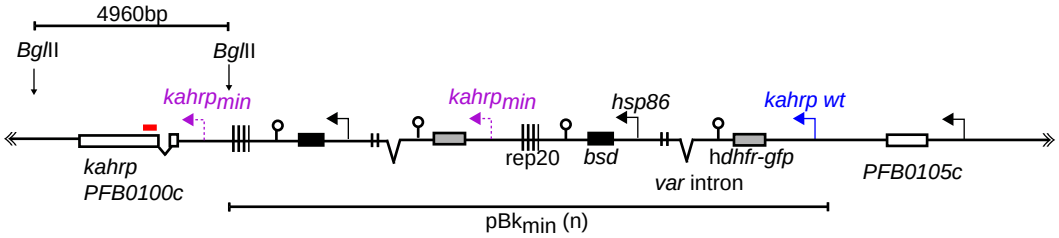
B



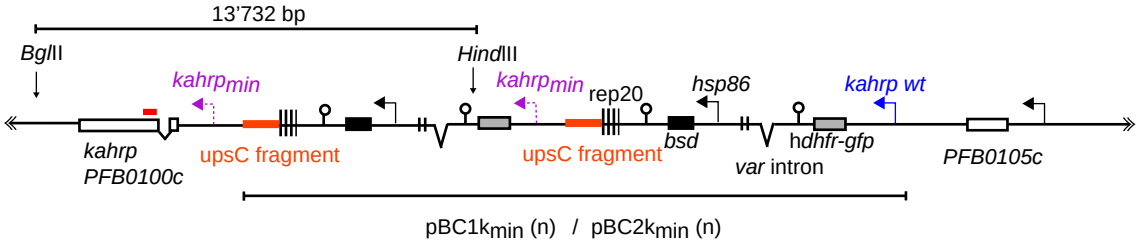
C *kahrp* locus



D Integration of pBk$_{min}$ at the *kahrp* locus



E Integration of pBC1k$_{min}$ and pBC2k$_{min}$ at the *kahrp* locus



F Integration of pBC3k$_{min}$ at the *kahrp* locus

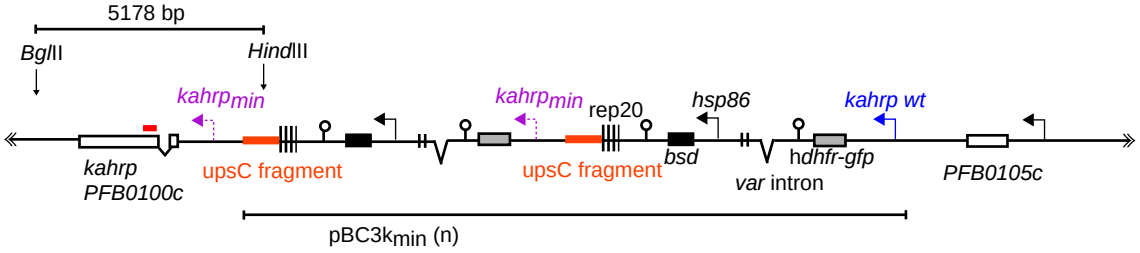


Figure S2

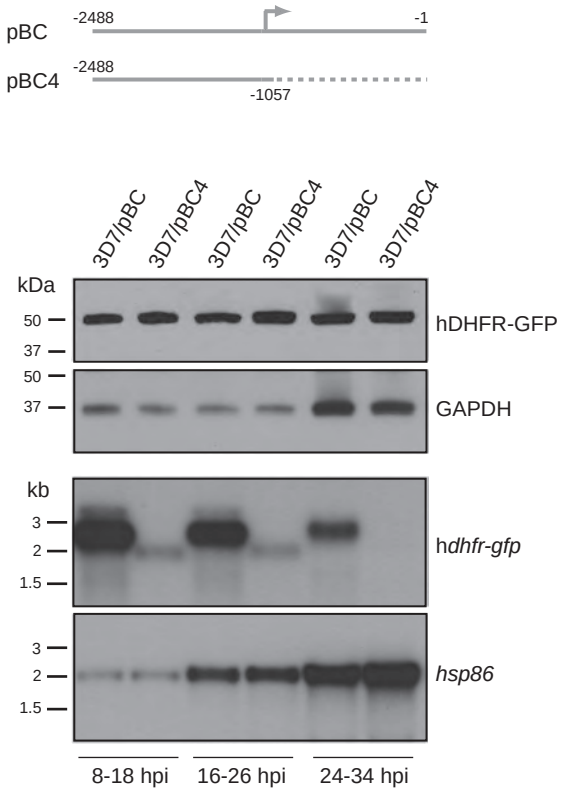

pBC
-2488
-1
pBC4
-2488
-1057
3D7/pBC
3D7/pBC4
3D7/pBC
3D7/pBC4
3D7/pBC
3D7/pBC4
kDa
50
37
50
37
hDHFR-GFP
GAPDH
kb
3
2
1.5
3
2
1.5
h*dhfr-gfp*
hsp86
8-18 hpi
16-26 hpi
24-34 hpi

Figure S3

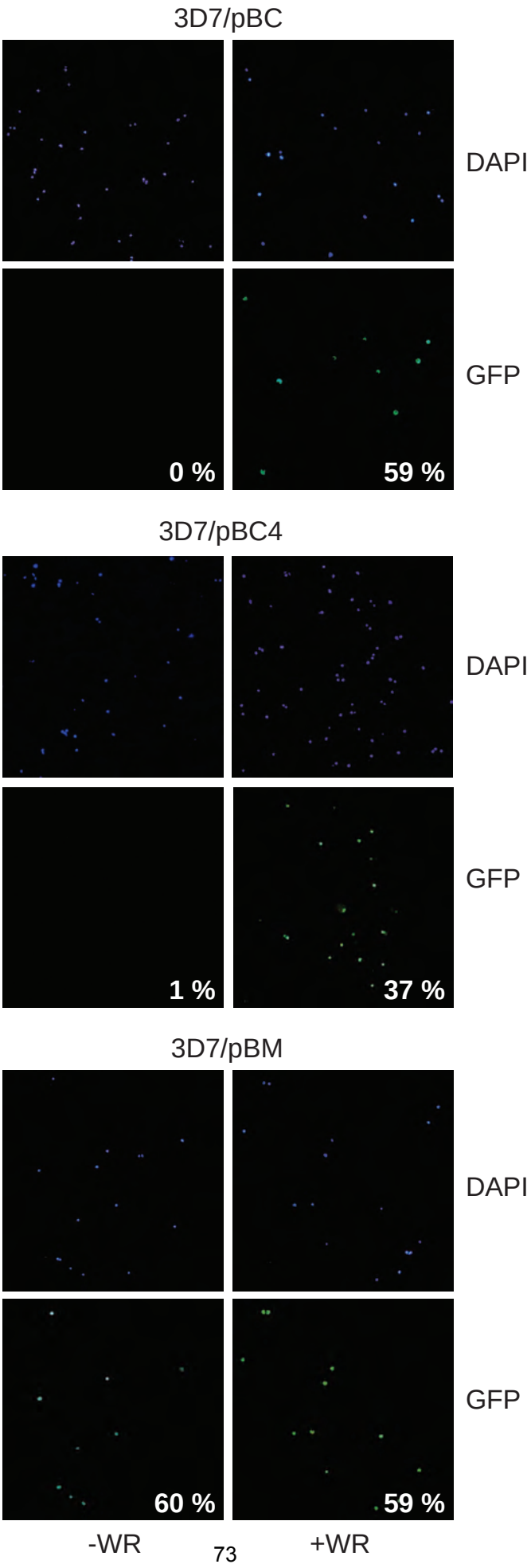

3D7/pBC
DAPI
GFP
0 %
59 %
3D7/pBC4
DAPI
GFP
1 %
37 %
3D7/pBM
DAPI
GFP
60 %
59 %
-WR
+WR

Figure S4

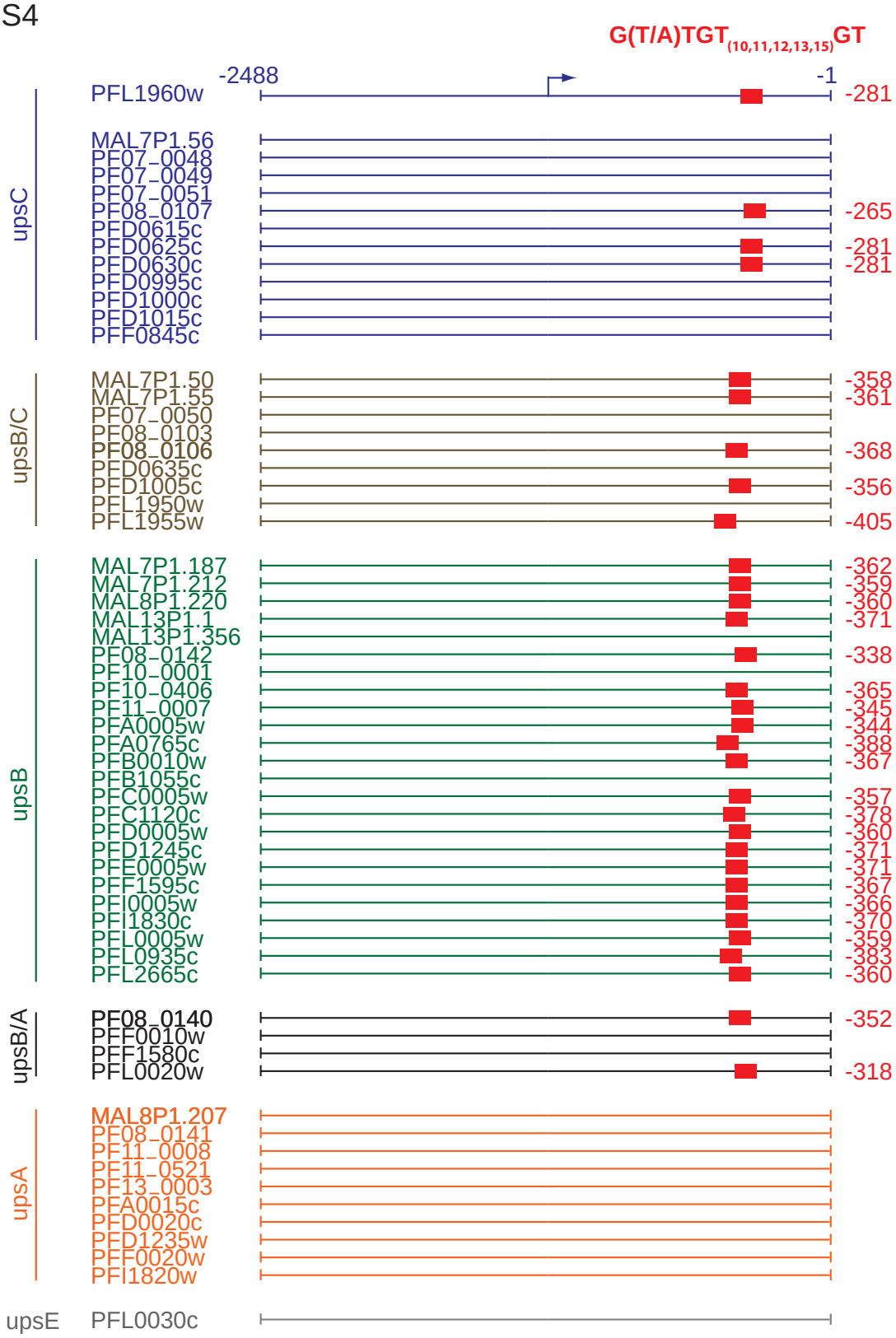

G(T/A)TGT(10,11,12,13,15)GT
-2488
-1
upsC
PFL1960w
-281
MAL7P1.56
PF07_0048
PF07_0049
PF07_0051
PF08_0107
-265
PFD0615c
PFD0625c
-281
PFD0630c
-281
PFD0995c
PFD1000c
PFD1015c
PFF0845c
upsB/C
MAL7P1.50
-358
MAL7P1.55
-361
PF07_0050
PF08_0103
PF08_0106
-368
PFD0635c
PFD1005c
-356
PFL1950w
PFL1955w
-405
upsB
MAL7P1.187
-362
MAL7P1.212
-359
MAL8P1.220
-360
MAL13P1.1
-371
MAL13P1.356
PF08_0142
-338
PF10_0001
PF10_0406
-365
PF11_0007
-345
PFA0005w
-344
PFA0765c
-388
PFB0010w
-367
PFB1055c
PFC0005w
-357
PFC1120c
-378
PFD0005w
-360
PFD1245c
-371
PFE0005w
-371
PFF1595c
-367
PFI0005w
-366
PFI1830c
-370
PFL0005w
-359
PFL0935c
-383
PFL2665c
-360
upsB/A
PF08_0140
-352
PFF0010w
PFF1580c
PFL0020w
-318
upsA
MAL8P1.207
PF08_0141
PF11_0008
PF11_0521
PF13_0003
PFA0015c
PFD0020c
PFD1235w
PFF0020w
PFI1820w
upsE
PFL0030c

Figure S5

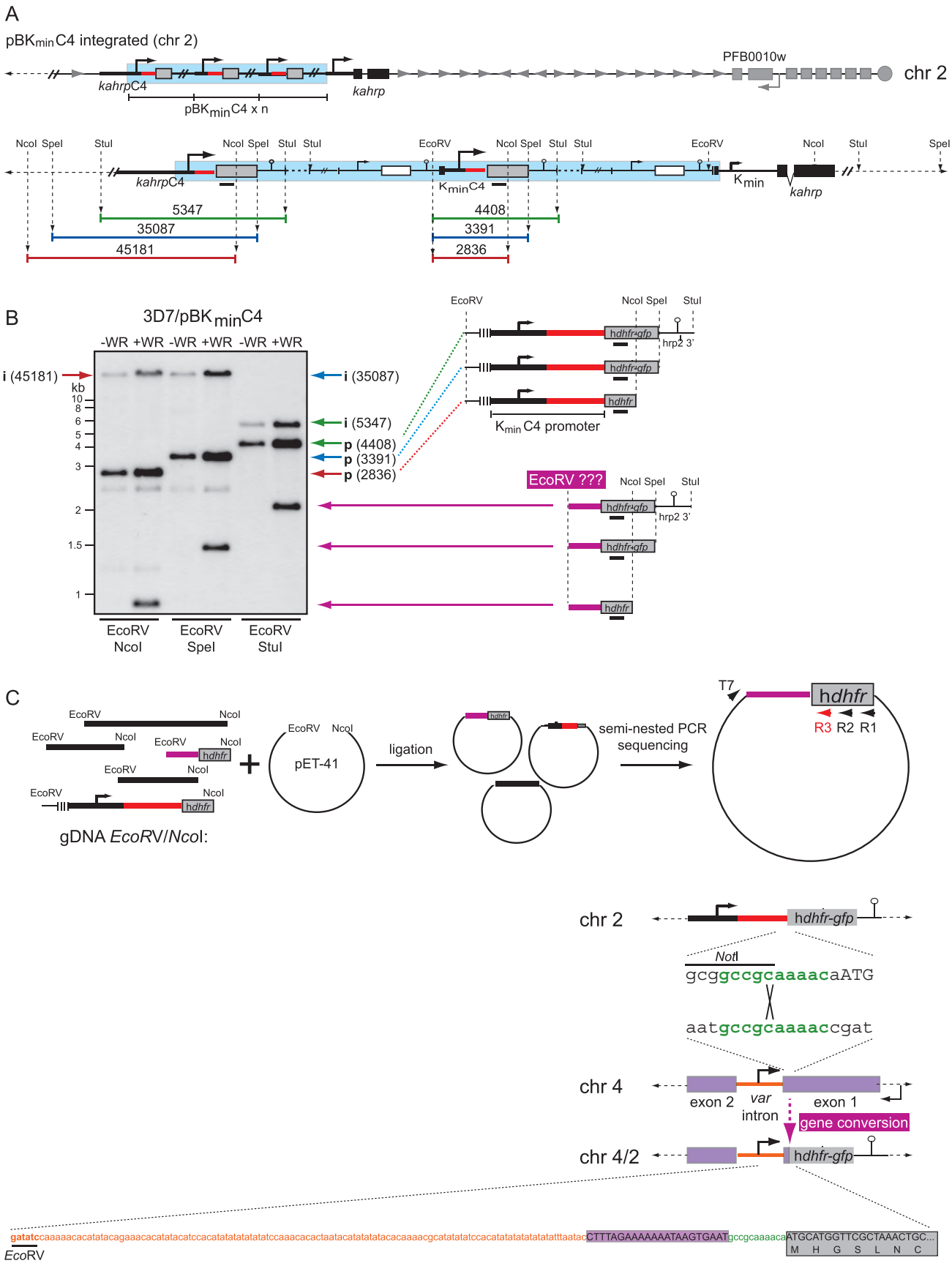
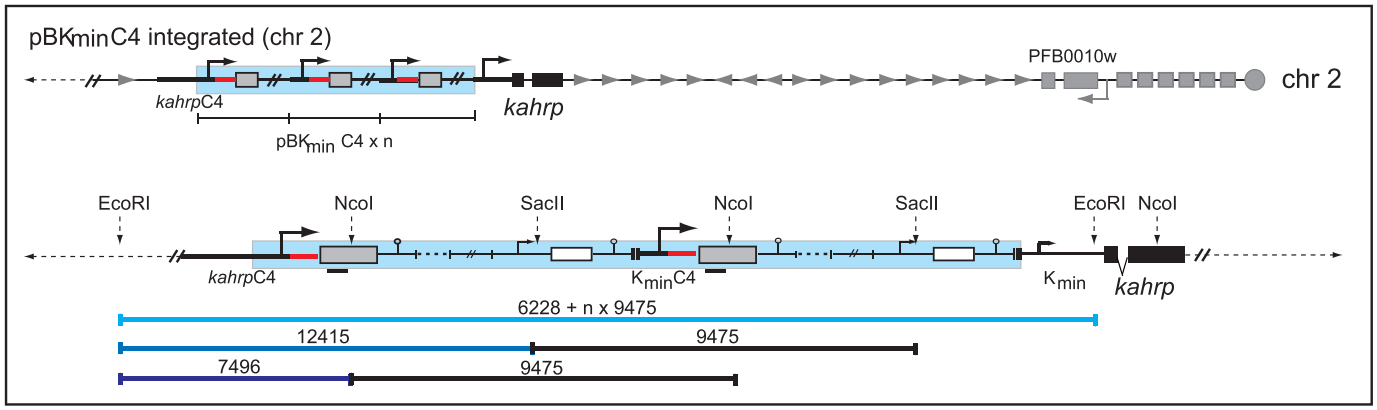
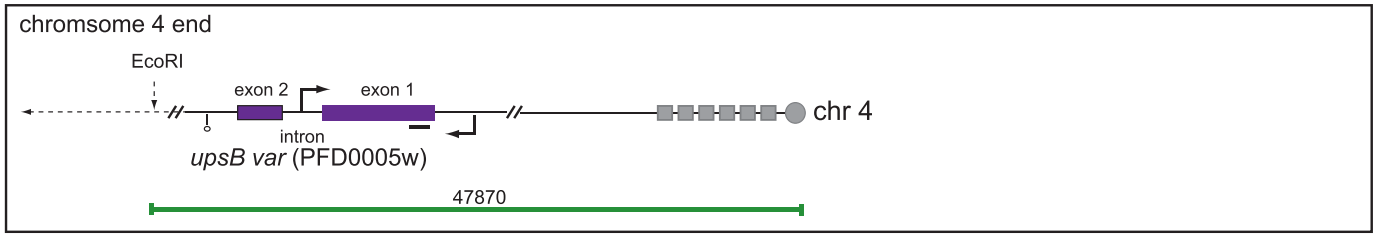

A
pBKminC4 integrated (chr 2)
PFB0010w
chr 2
kahrpC4
kahrp
pBKminC4 x n
NcoI SpeI StuI
EcoRV
KminC4
Kmin
5347
35087
45181
4408
3391
2836
B
3D7/pBKminC4
-WR +WR -WR +WR -WR +WR
i (45181)
i (35087)
i (5347)
p (4408)
p (3391)
p (2836)
kb
10
8
6
5
4
3
2
1.5
1
hdhfr-gfp
hrp2 3'
hdhfr
KminC4 promoter
EcoRV ???
EcoRV NcoI
EcoRV SpeI
EcoRV StuI
C
gDNA EcoRV/NcoI:
pET-41
ligation
semi-nested PCR sequencing
T7
R3 R2 R1
chr 2
NotI
gcggccgcaaaacaATG
aatgccgcaaaaccgat
chr 4
exon 2
var intron
exon 1
gene conversion
chr 4/2
gatatccaaaaacacatatacagaaacacatatacatccacatatatatatatccaaacacactaatacatatatatacacaaaacgcatatatatccacatatatatatatatttaatacCTTTAGAAAAAAATAAGTGAATgccgcaaaacaATGCATGGTTCGCTAAACTGC...
EcoRV
M H G S L N C

Figure S6

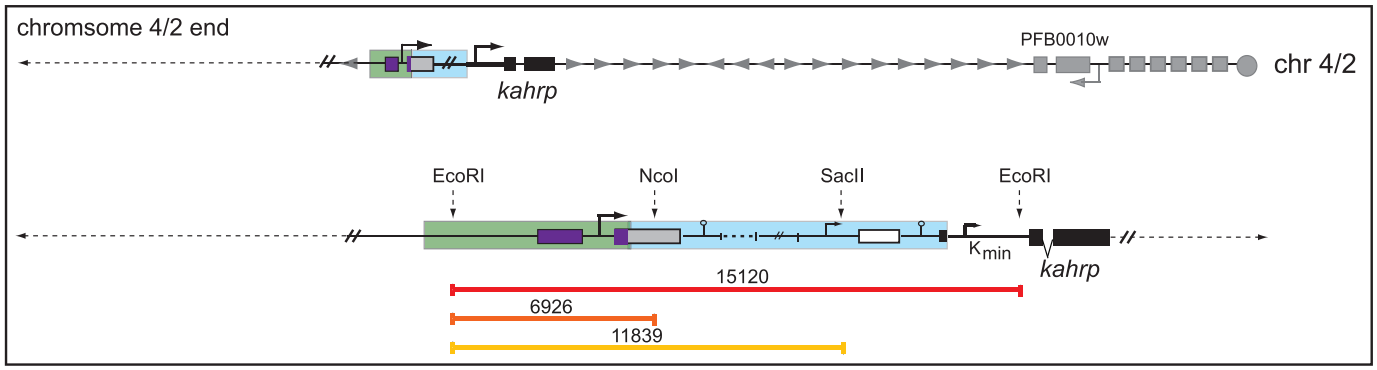
A



B



C



D

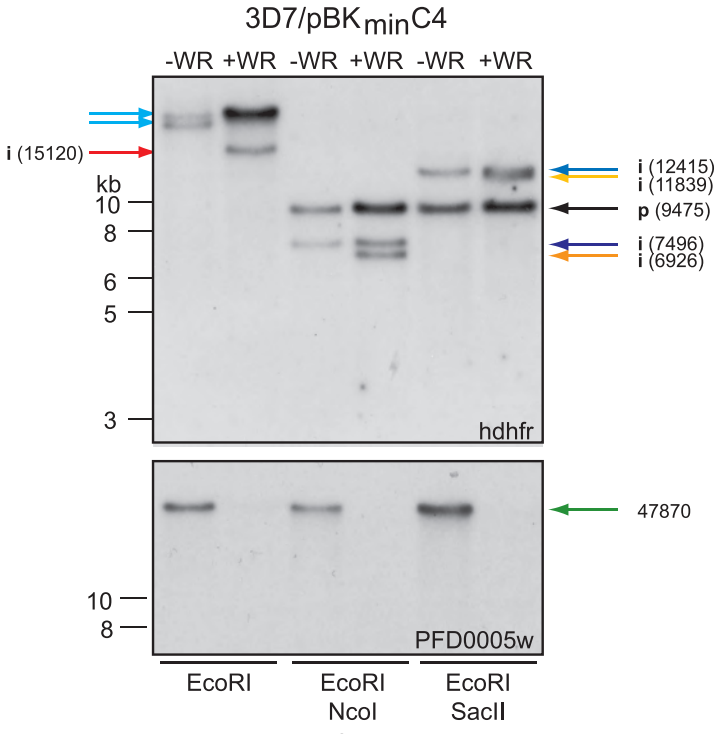


Figure S7

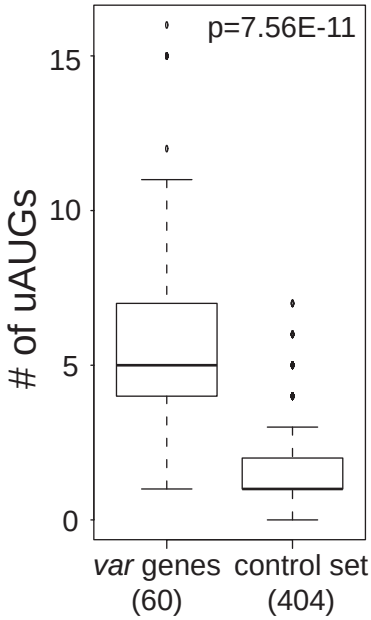


Table S1. All primers used in this study are listed. Restriction sites are indicated in bold.

	primer sequence (5`→3`)	RE sites
pBKmin		
kahrp-Rev-N-N	gatc**gcggccgc**ata**gctagc**gattctctaataattatgtacg	NotI, NheI
kahrp-Fwd-P-B-C	gatc**ctgcag**atgact**agatct**acgtgta**atcgat**cctaaaactgcatgtagtg	PstI, BglII, ClaI
upsC-Kmin hybrid promoters		
upsC-1_forw	cagt**ggatcc**atttcatcattattaaagtagag	BamHI
upsC-1_rev	cagt**ggatcc**ttctatctatattatctaccac	BamHI
upsC-2_forw	cagt**ggatcc**ttttttttctttttgatgttgtac	BamHI
upsC-2_rev	cagt**ggatcc**atatttcatactaatattattcac	BamHI
upsC-3_forw	cagt**ggatcc**ccactacatggtattaccac	BamHI
upsC-3_rev	cagt**ggatcc**gttcgtgactacatgatgtc	BamHI
upsC deletion constructs		
upsC-F	cagt**agatct**ctttatgttggtacattatacatg	BglII
upsC-R	cagt**gcggccgc**ttttgtttttttgtttatcgttcg	NotI
upsC1-F	cagt**agatct**atagaaatattactgtttggag	BglII
upsC2-F	cagt**agatct**atttcatcattattaaagtagag	BglII
upsC3-F	cagt**agatct**atttattttcatagaaatgtgg	BglII
upsC4-R	cagt**gcggccgc**aaaaagaattataatcgaagaac	NotI
upsC5-R	cagt**gcggccgc**tgtttcttagtcgtactatatgtg	NotI
upsC6-R	cagt**gcggccgc**tatttaatactttatattatgtgg	NotI
upsC7-R	cagt**gcggccgc**atttatattaccatgatgccg	NotI
upsC8-R	cagt**gcggccgc**cacattattactctaatatgcg	NotI
pBCK		
kahrp-F-B	cagt**agatct**ctatgttagttatgataggacc	BglII
kahrp-R-N-S-B	cagt**gcggccgc**agt**cctgcagg**agt**ggatcc**atttttcaatctttttttttcagc	NotI, SbfI, BamHI
upsC-F-A	cagt**ggcgcgcc**atttcatcattattaaagtagag	AscI
upsC-R-M	cagt**acgcgt**tatatttttaaaagattaaatatagc	MluI
pBKminC4		
kahrp-F-B	cagt**agatct**ctatgttagttatgataggacc	BglII
kahrp-R-N-S-B	cagt**gcggccgc**agt**cctgcagg**agt**ggatcc**atttttcaatctttttttttcagc	NotI, SbfI, BamHI
upsC-F-B	cagt**ggatcc**cccatcacatatagtaggac	BamHI
upsC-R	cagt**gcggccgc**ttttgtttttttgtttatcgttcg	NotI
qPCR		
PF13_0170F	tggctaggatatgattggaaagaaca	
PF13_0170R	tacggttctatttctatatggtgaatca	
gfpF	acacttgtcactactttcgcgtatggtcttc	
gfpR	accttcaaacttgacttcagcacgtgtcttgtagt	
kahrpF	acggatccggtgactccttcgat	
kahrpR	tggtgaacctgtggtgcttggtgat	
msp8F	tgacgcaaaagcaagggacaacaataataatgatga	
msp8R	tcatcgtcatcatcattatcatcatcatcacc	
hd-I-F	cgacgatgcagtttagcgaac	
hf-I-R	atacagaaacacatatacatccacat	
hybridisation probes		
kahrpF	acggatccggtgactccttcgat	
kahrpR	tggtgaacctgtggtgcttggtgat	
hdhfrF	agctggatccgcggccgcaaaacatgcatggttcgctaaactg	
hdhfrR	agctgtcgacagcagcatcattcttctcatatacttcaa	

hsp86F	gaattgattagtaatgctagtg
hsp86R	gtttcatccttagtaactgtg
PFD0005wF	ccaaaaacacatatacagaaacaca
PFD0005wR	caacgaatgcagtttagcgga
EMSA oligonucleotides	
MEE2-F	cagttgataatgatgtttttttttgttatagattatgataacaagcttta
MEE2-R	gacttaaagcttgttatcataatctataacaaaaaaaaaacatcattatca
MEE3-F	cagtacaagctttatgaatcgcatattagagtaataatgtgcatgcatgac
MEE3-R	gactgtcatgcatgcacattattactctaatatgcgattcataaagcttgt
SPE1M-F	cacggacaaaaaaagtaaccgagaattattatatataaatat
SPE1M-R	atatttatatataataattctcggttacttttttgtccgtg
Ligation-mediated PCR	
T7 terminator	gctagttattgctcagcgg
hdhfr_R1	agctgtgcagagcagcatcattcttctcatatacttcaa
hdhfr_R2	ccttgtggaggttccttgag
hdhfr_R3	ggtagtccccgttcttgcc

Supplemental Experimental Procedures

Transfection constructs

All transfection constructs generated in this study are derivatives of pBK_{min}, which was itself obtained by replacing the *cam* promoter in pBcam (unpublished) with the 1115bp minimal *kahrp* promoter (K_{min}) amplified from 3D7 gDNA and *Bgl*II/*Not*I restriction. Constructs pBC and pBC1 to pBC8 were generated by replacing K_{min} with *Bgl*II/*Not*I-digested upsC upstream sequences (PFL1960w) amplified from pCAT5B1 (Voss et al., 2000). upsC-*kahrp* hybrid promoters were obtained by cloning *BamH*I-digested upsC fragments into *Bgl*II-digested pBK_{min}. To obtain $pBK_{min}C4$, the K_{min} promoter was replaced by a *Bgl*II/*Not*I-digested *kahrp* promoter fragment (-1115 to -445 bps) containing a penultimate 3` *BamH*I restriction site. The resulting plasmid was digested with *BamH*I/*Not*I to insert the upsC 5' UTR element (-519 to -1). To generate pBCK the $K_{min}$ promoter was replaced by a *BamH*I/*Not*I-digested *kahrp* 5' UTR fragment (-764 to -1) containing *Asc*I and *Mlu*I sites at the 5` end to allow subsequent fusion to the *Asc*I/*Mlu*I-digested upsC promoter sequence (-1679 to -914). The *mahrp1* promoter was amplified from gDNA and cloned into *Bgl*II/*Not*I-digested pBK_{min} to obtain pBM. All primer sequences are listed in Table S1.

Quantitative reverse transcription PCR

Pre-synchronised parasites cultures were synchronised twice 16 hours apart to obtain an eight hour growth window. Total RNA was isolated using Tri Reagent (Ambion) and further purified using the RNeasy Plus Mini Kit (Qiagen) for removal of gDNA. Residual gDNA was digested with TURBO DNA-*free*™ (Ambion). All samples were tested negative for contaminating gDNA by qPCR. RNA was reverse transcribed using the RETROscript Kit (Ambion). qPCR reactions for absolute transcript quantification of h*dhfr-gfp*, *kahrp,* PF13_0170 (glutaminyl-tRNA synthetase), *msp8* and *var* intron-derived h*dhfr-gfp* were performed at final primer concentrations of 0.4uM using SYBR Green Master Mix (Applied Biosystems) on a StepOnePlus™ Real-Time PCR System (Applied Biosystems) in a reaction volume of 12ul. Plasmid copy numbers were determined by qPCR on gDNA isolated from the same parasite samples and calculated by dividing the absolute h*dhfr-gfp* copy numbers by the average value obtained for *msp8* and PF13_0170. All reactions were run in duplicate yielding virtually identical Ct values. Serial dilutions of gDNA and plasmid DNA were used as standards for absolute quantification. Relative transcript values were calculated by normalisation against the house-keeping gene PF13_0170 or *msp8*. Primer sequences are listed in Table S1.

Live cell fluorescence microscopy

500ul culture (5% hematocrit, 3-8% trophozoites) were pelleted for 1min at 200g and washed twice in 1ml RPMI-HEPES. The pellet was resuspended in 400ul RPMI-HEPES containing 2.5ug/ml DAPI and incubated for 10min at 37°C. Cells were washed once in 1ml PBS. Three ul of packed RBCs were mixed with 2ul Vectashield (Vector Laboratories Inc.) on a glass slide and parasites were immediately analysed on a Leica DM 5000B microscope. Images were taken at 40x magnification with a Leica DFC 300 FX camera (exposure time for GFP and DAPI was 1.3s and 130ms, respectively) and acquired via the Leica IM 1000 software. All images were processed using Adobe Photoshop CS2 and identical settings.

Supplemental Reference

Otto,T.D., Wilinski,D., Assefa,S., Keane,T.M., Sarry,L.R., Bohme,U., Lemieux,J., Barrell,B., Pain,A., Berriman,M., Newbold,C., and Llinas,M. (2010). New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. Mol. Microbiol. *76*, 12-24.

Mutually Exclusive Transcription of Subtelomeric Gene Families in *Plasmodium falciparum* is Restricted to *var* Genes

Kathrin Witmer[1,2], Christoph D. Schmid[1,2], Yen-Hoon Luah[3], Peter R. Preiser[3], Zbynek Bozdech[3], Till S. Voss[1,2]

[1] Swiss Tropical and Public Health Institute, Basel, Switzerland

[2] University of Basel, Switzerland

[3] School of Biological Sciences, Nanyang Technological University, Singapore

Abstract

The *P. falciparum* genome is equipped with several subtelomeric gene families that are implicated in parasite virulence and immune evasion. The members of these gene families are uniformly positioned within heterochromatic domains of the genome and are thus subject to variegated expression. The best-studied example is that of the *var* gene family encoding the major parasite virulence factor *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). PfEMP1 undergoes antigenic variation through switches in mutually exclusive *var* gene transcription. *var* gene promoters function as crucial regulatory elements in the underlying epigenetic control strategy. Transcriptional regulation of other subtelomeric gene families and their role in parasite biology is much less understood. Here, we investigated the mode of transcriptional control of *var*, *rif*, *stevor*, *phist* and *pfmc-2tm* families by comparative genome-wide transcriptional profiling of transgenic parasite lines. Our results establish a clear functional distinction between *var* and non-*var* transcriptional control mechanisms. Unlike *var* promoters, we find that promoters of non-*var* families are not silenced by default. Moreover, we show that mutually exclusive transcription is unique to the *var* gene family. The three major *var* promoter types are functionally equal and play an essential role in singular gene choice. Further, we provide evidence suggesting that mutually exclusive activation of all *var* types is regulated by a common control mechanism. Our findings identified a differential logic in the regulation of *var* and other subtelomeric virulence gene families, which will have important implications for our understanding and future analyses of phenotypic variation in malaria parasites.

Author Summary

Plasmodium falciparum causes the most severe form of malaria in humans. During blood stage infection intra-erythrocytic parasites express different surface antigens that are encoded by gene families. These are located in a repressive genome environment, which facilitates expression of individual antigens in a clonally variant manner. Consequently, *P. falciparum* continuously escapes immune recognition and clearance. We used a

comparative approach to study the function of gene promoters in transcriptional regulation of clonally variant expression. Transcription of the *var* gene family, encoding the major virulence factor PfEMP1, is controlled by a different logic compared to other gene families. *var* gene transcription is mutually exclusive; only a single family member is activated at any time. The three major *var* promoter types act as equally crucial elements in this specialised mode of regulation. In contrast, promoters of other gene families lack this competence and their cognate gene families are not transcribed by mutual exclusion. The inherent necessity to restrict *var* expression to a single gene underscores the vulnerability of *P. falciparum* to immune responses targeting PfEMP1. Our results are important to understand the significance of antigenic variation in parasite survival and to elucidate the mechanism underlying mutually exclusive *var* gene expression.

Introduction

The blood stages of *Plasmodium falciparum* evade antibody–mediated host immunity by altering parasite-encoded antigens exposed on the surface of infected red blood cells (iRBCs) through antigenic variation [1–4]. Antigenic variation results from the amplification of hypervariable gene families, coupled to the ability to switch expression of individual genes. Notably, these gene families are predominantly located near the telomeres [5]. The positional clustering of antigen gene families in subtelomeric regions is not unique to *P. falciparum* but is common to a wide range of pathogenic protozoa and fungi [6–10]. Subtelomeric gene families show high rates of recombination to facilitate diversity and this has been demonstrated in the model yeast *S. pombe* [11] and in *P. falciparum*, where meiotic recombination among *var* genes is at least eight times more frequent than the estimated genomic average [12,13]. Proximity to telomeres also modulates gene expression as exemplified by variegated silencing of genes nearby telomeres [14]. Indeed, *P. falciparum* subtelomeric regions are heterochromatic and predominantly occupied by the epigenetic silencing marks trimethylated histone 3 lysine 9 (H3K9me3) and heterochromatin protein 1 (PfHP1) [15–17]. This repressive

environment facilitates silencing and variegated gene expression thus promoting antigenic variation in *P. falciparum*.

Subtelomeric *P. falciparum* gene families include *var* [3,18,19], *rif* (repetitive interspersed family) [20], *stevor* (subtelomeric variable open reading frame) [21], *phist* (*Plasmodium* helical interspersed subtelomeric) [22], *pfmc-2tm* (*P. falciparum* Maurer's clefts two transmembrane) [23] and the *hyp* families 1 to 17 [22]. The best studied example is that of the ~60 member *var* gene family. *var* gene tanscription is controlled by conserved promoter sequences that can be sub-grouped into three main classes according to sequence and chromosomal location [5,24–26]. upsB *var* genes are most closely located to telomeres and transcribed towards the centromere. upsA *var* genes are also telomerically located but transcribed towards the telomeres, and upsC *var* genes appear exclusively in chromosome-internal clusters. PfEMP1, encoded by *var* genes, is exposed on the iRBC surface and mediates adherence to various endothelial receptors in the human host [4,18,27]. The resulting sequestration of RBC aggregates in the microvasculature is a major cause of severe clinical outcomes such as cerebral and pregnancy-associated malaria [28–31]. Moreover, antigenic variation of PfEMP1 contributes substantially to chronic disease and transmission.

Unlike PfEMP1, the function of other exported protein families in parasite biology and disease is unclear, but there is indication for roles in host-parasite interactions [32–37]. With over 130 members, the *rif* gene family is the largest in *P. falciparum* [36,38]. A-type RIFINs are associated with parasite-induced membranous structures in the iRBC cytosol termed Maurer’s clefts (MC) and are exported to the surface of the infected RBC. B-type RIFINs have been shown to reside within the parasite [39,40]. Protein products of the ~35-member *stevor* family localize to the surface of infected erythrocytes and are also associated with MCs [41–44]. The *pfmc-2tm* family (twelve members) are associated with MCs and RBC surface [23,41]. Members of the *phist* and *hyp* families are predicted to be exported to the iRBC [22], but this has not yet been experimentally confirmed.

var genes are transcribed in a mutually exclusive manner such that in single parasites only one *var* gene is active and all other members are silenced [45]. *var* promoter regions are essential in these alternative states of activity [46–

51]. Transfection experiments using episomal promoters to drive expression of drug selectable markers proved a great strategy to investigate transcriptional regulation of *var* genes [46,49,51]. Upon transfection, *var* promoters are silenced by default [46,49,52,53] and this silenced state is only induced properly in presence of the *var* gene intron [46,50,53,54]. Interestingly, drug-induced selection for active promoters is sufficient to infiltrate drug-selectable markers into the pathway of mutually exclusive transcription [46,49,51]. In such parasites, the active episomal *var* promoter substitutes the active *var* gene, which leads to a shutdown of endogenous *var* transcription and, consequently, absence of PfEMP1 on the iRBC surface. These studies have been instrumental not only in understanding the regulatory mechanisms underlying mutually exclusive transcription, but also in reconstructing earlier observations with regard to the variable phenotype of iRBCs [46,55–57].

Despite most other multigene families in *P. falciparum* have been known for many years, our understanding of the processes that regulate their transcription is limited. Furthermore, it is also unclear if and to what extent transcriptional co-regulation and possible crosstalk between subtelomeric virulence gene families exists. Expression of *rif*, *stevor* and *pfmc-2tm* genes is also clonally variant *in vitro* and *in vivo* [32,35,39,40,43,58]. However, in contrast to *var* genes transcription of the *rif* family is not mutually exclusive as *rif* variants of both subtypes can be expressed simultaneously in a single parasite [39]. Analysis of parasites selected for monogenic *var* expression provided no evidence for co-regulation between the active *var* and neighbouring *rif* or *stevor* genes [59,60]. These studies indicate that juxtaposition to active *var* genes does not predispose the activity of other gene family members. Interestingly, however, interference of several episomal gene family promoters with transcription of various endogenous gene families has been reported recently indicating that they may share a common activation factor [61].

Here, we used genome-wide comparative transcriptional profiling to investigate the extent to which activation of extra-chromosomal virulence gene promoters affects transcription of endogenous gene families. Our results establish a clear functional distinction between *var* and non-*var* promoters. In

contrast to *var* genes, promoters of other gene families are active by default. Activation of all three major *var* gene promoter types leads to the specific down-regulation of endogenous *var* genes. In contrast, the *rif*, *stevor*, *pfmc-2tm* and *phistb* promoters do not interfere with transcription of their respective gene families. Moreover, we find only limited evidence for potential co-regulation between subtelomeric gene families. Together, our findings demonstrate that mutually exclusive transcription is unique to the *var* gene family and that *var* promoters of all types play an essential part in singular gene choice.

Results

Generation of stable reporter parasite lines

Parasites of the strain 3D7 were transfected with constructs carrying eight different promoters: upsA-type *var* (PF13_0003), upsB-type *var* (PFL0005w), upsC-type *var* (PFL1960w), *rif* (PF13_0004), *stevor* (PFL2610w), *phistb* (PFL2540w), *pfmc-2tm* (final gene ID unknown due to high sequence conservation between *pfmc-2tm* promoters) and *cam* (PF14_0323) (control). Note that the *rif* and *upsA* promoters are present on the sense and antisense strand, respectively, of the 2.8 kb intergenic sequence separating these head-to-head oriented genes. All sequences were cloned into the context of the parental reporter vector pBcam by replacement of the *cam* promoter (Fig. 1A). The blasticidin-S deaminase (*bsd*) resistance cassette was used to obtain transfectants carrying stable episomes. Importantly, this initial selection is independent of the activity of the test promoters driving transcription of the downstream drug-selectable reporter h*dhfr-gfp* (human dihydrofolate reductase fused to green fluorescent protein). A *var* gene intron element is located downstream of the h*dhfr-gfp* cassette on all constructs for consistency. Transfection and subsequent selection on BSD-S-HCl (BSD) was successful and generated stable reporter cell lines 3D7/upsA†, 3D7/upsB, 3D7/upsC, 3D7/rif, 3D7/stevor, 3D7/phist, 3D7/pfmc-2tm and the control 3D7/cam. All lines exhibited similar plasmid copy numbers ranging from two to eight copies per parasite (Fig. 1B). After WR99210 (WR) selection, average plasmid numbers increased in each transfectant with the

highest number of 33 observed in 3D7/*rif*. Unexpectedly, 3D7/upsA† was completely refractory to WR selection in seven challenge experiments on three independently generated BSD-resistant populations. This failure to select for an active upsA promoter may have been related to efficient upsA silencing. To exploit the possibility that *rif* promoter activity may predispose the head-to-head upsA promoter for activation, we replaced the entire *bsd* cassette in pBrif with an inverted *bsd* gene and terminator element to create pBupsA (Fig. 1C). In this context, the reverse complement sequence of the *rif* upstream region is expected to act as an upsA promoter driving *bsd* transcription. Indeed, parasites transfected with this construct were successfully selected on BSD. Hence, this BSD-resistant population (3D7/upsA) carried an activated upsA promoter and was used in all subsequent experiments. Southern blot analysis revealed that plasmids were maintained episomally in all lines except 3D7/stevor where integration into the endogenous *stevor* locus (PFL2610w) occured (Fig. S1).

Initial assessment of overall transcript profiles at four consecutive intervals during intra-erythrocytic development.

We harvested total RNA at four consecutive time points during the intra-erythrocytic developmental cycle (IDC) from 3D7 wild-type parasites and each of the WR-selected transfected lines (except for 3D7/upsA where a BSD-selected population was used). This sampling strategy was designed to include the temporal peaks of transcription for each endogenous gene family [23,36,45,62,63]. Transcript levels for each gene were determined relative to a 3D7 reference cDNA pool by hybridisation to a genome-wide long oligonucleotide microarray [64] (Dataset S1). Hierarchical average linkage clustering revealed four distinct clusters corresponding to the four sampling time points (Fig. 2A). Within each time point, the samples showed a high degree of correlation but the individual datasets clustered differently. The overall similarity between individual transcriptomes was lowest in early ring stages and increased gradually with progression through the IDC (Fig. 2B). This is explained by the variegated expression of a large number of genes expressed specifically in ring stage parasites [65–67]. Compared to both control parasites (3D7 wild-type and 3D7/cam), the ring stage transcriptomes

of 3D7/upsA, 3D7/upsB, 3D7/upsC and 3D7/rif consistently displayed a higher number of deregulated genes than the 3D7/stevor, 3D7/phist and 3D7/pfmc-2tm transfectants. Notably, the transcriptomes of all reporter lines consistently showed a higher degree of similarity to the control transfectant 3D7/cam than to untransfected parasites (Fig. 2B). We explain this important observation by the likelihood that perturbations associated with the generation of transgenic parasites *per se* cause alterations in mRNA profiles. Indeed, we found 127 genes either up- or down-regulated by more than two-fold in at least one time point in all transfected lines compared to the untransfected control (Fig. S2). Interestingly, this set was characterised by a large number of genes coding for proteins involved in DNA/RNA metabolism, protein translation and export, and stress response. These changes are compatible with possible responses to plasmid maintenance and replication as well as to the mode of BSD action, which inhibits protein synthesis [68].

var and non-*var* subtelomeric gene family promoters display distinct functional characteristics

As an important consideration for comparative analyses, we tested whether the stage-specific activities of the episomal promoters matched the temporal transcription profile of the cognate endogenous gene families. To do this, we compared reporter gene mRNA abundance profiles to the average relative expression of the corresponding gene family members (Fig. S3). This analysis revealed high correlations for the *var* upsA, upsB, upsC, *stevor, pfmc-2tm* and *cam* promoters. The episomal *rif* promoter profile correlated with endogenous *var* rather than *rif* transcription implying that this bi-directional *rif* promoter displays a *var*-specific activation pattern. For the *phistb* promoter we observed divergence in early schizonts where episomal promoter activity increased and endogenous *phist* transcription decreased. Hence, except for *phistb* the episomal promoters analysed here contain sufficient information to recapitulate the temporal expression of endogenous gene families.

Several studies reported that upsB and upsC *var* promoters are regulated by epigenetic mechanisms and adopt either a silenced (default) or an active (activated) state [46,49,51,53]. By quantifying reporter transcript levels in unselected (default) and drug-selected (activated) populations we asked if the

promoters of upsA *var* and subtelomeric non-*var* gene families show the same behaviour. The non-epigenetically regulated house-keeping gene promoter *cam* and the upsB promoter served as important negative and positive controls for silencing, respectively. Quantitative reverse transcription PCR (qRT-PCR) analysis revealed increased activities in drug-selected compared to unselected populations for all promoters tested (Fig. 3 and S4). However, only the upsB (18.2-fold induction +/- 1.0 s.d.) and upsA promoter (3.8-fold +/- 0.6 s.d.) showed a significantly elevated level of induction relative to the negative control promoter *cam* (2.1-fold +/- 0.4 s.d.). In contrast, we found no evidence for alternative states of activity for the *stevor*, *phistb* and *pfmc-2tm* promoters. The head-to-head *rif* promoter displayed a 7.1-fold mean induction (+/- 3.4 s.d.; not significant) in WR-selected versus unselected parasites, which may be indicative for silencing in a considerable fraction of the unselected population. In summary, these results demonstrate for the first time that promoters of subtelomeric non-*var* gene families are unable to recruit transcriptional silencing autonomously and are thus active by default. This is in clear distinction to *var* promoters, which are characterised by reversible responsiveness to epigenetic control mechanisms.

The capability to interfere with mutually exclusive *var* gene transcription is restricted to *var* gene promoters

Activation of upsB and upsC promoters leads to down-regulation of endogenous *var* transcription [46,49–51]. Here, we studied the specificity of this crosstalk by comparative analysis of relative *var* transcription levels in all cell lines. As shown in Figure 4A, mean relative *var* expression was strikingly reduced in early and late ring stages in 3D7/upsA, 3D7/upsB and 3D7/upsC compared to all other lines. Interestingly, this was also evident in 3D7/rif to a level comparable to that induced by active *var* promoters. As expected, *var* transcription was repressed/absent in trophozoites and early schizonts (TP3 and TP4) in all lines investigated, underscoring the specificity of the changes observed for TP1 and TP2.

The box plot in Fig. 4A compares the distribution of relative *var* expression levels at TP1. With the exception of 3D7/stevor, *var* transcription was reduced in all transfected lines compared to 3D7. Importantly, however, *var* down-

regulation was much less pronounced in 3D7/phist, 3D7/pfmc-2tm and 3D7/cam compared to the lines carrying activated *var* promoters. When compared to 3D7/cam, significant down-regulation occurred specifically only in 3D7/upsA, 3D7/upsB, 3D7/upsC and 3D7/rif parasites. This is also illustrated by comparing the relative expression levels for each individual *var* gene (Fig. 4B). In 3D7/upsA, 3D7/upsB, 3D7/upsC and 3D7/rif *var* gene transcription was largely reduced or absent. In contrast, transcripts of all *var* genes were detected in the 3D7/stevor, 3D7/phist, 3D7/pfmc-2tm and 3D7/cam populations. Compared to 3D7, the relative abundance of most transcripts was slightly lower and some were transcribed at higher levels in these lines. It is worth mentioning that 3D7/upsC parasites exhibited similar or even elevated relative expression of a subset of *var* genes when compared to the 3D7 control. Interestingly, all of these genes belong to the upsA and upsB/A *var* subgroups. These genes are hardly expressed in 3D7 parasites [69,70] and are thus only present in minute amounts in the 3D7 cDNA reference pool. Consequently, these hybridisation signals have to be interpreted with caution as they merely represent transcription slightly above background. At this stage, however, we can not exclude the possibility of an incomplete *var* knock-down effect in a small subset of 3D7/upsC parasites.

In addition to *var* genes, other heterochromatic genes also showed a trend towards down-regulation in transfected versus untransfected parasites, irrespective of the promoter used to drive h*dhfr-gfp*. However, when comparing global transcription profiles to the control transfectant 3D7/cam this effect disappears except for *var* genes, which remain specifically down-regulated (Fig. S5). We therefore conclude that transcription of heterochromatic genes is generally lower in episomally transfected parasites. This hypothesis is corroborated by the finding that 3D7/stevor, the only cell line where integration of the plasmid concatamer into the endogenous locus had occurred, was the only cell line displaying wild-type *var* transcript levels.

In summary, this analysis shows that activation of the three main *var* promoter types, and a head-to-head *rif* promoter, results in silencing of endogenous *var* genes. Promoters of other gene families do not have this competence underscoring the specificity of *var* promoter-mediated interference with mutually exclusive *var* transcription.

Variegated expression of other gene families occurs independently of promoter activation and is not mutually exclusive

In striking contrast to *var* genes, neither of the *rif*, *stevor*, *phist* and *pfmc-2tm* promoters caused a significant knock-down in transcription of their cognate endogenous gene families (Fig. 5). These data clearly show that transcriptional control of the *rif*, *stevor*, *phist* and *pfmc-2tm* families does not obey the same strategy employed for *var* genes, and that these families are not transcribed by strict mutual exclusion.

However, variegated expression of these families was still obvious. Except for the *phist* families (Fig. 5C), we observed altered average expression in some transgenic lines relative to 3D7/cam. For instance, *rif* transcription was slightly but significantly altered in 3D7/upsA, 3D7/upsB, 3D7/stevor, 3D7/phist and 3D7/pfmc-2tm (Fig. 5A). Transcription of individual *stevor* and *pfmc-2tm* members was characterised in general by a large dispersion of relative expression about the mean (Figs. 5B and 5D). In trophozoites, average *pfmc-2tm* transcription was significantly reduced in 3D7/stevor, 3D7/phist and 3D7/pfmc-2tm. Interestingly, in late ring stages average expression of the *rif*, *stevor* and *pfmc-2tm* families was significantly lower in 3D7/upsA and higher in 3D7/upsB, which was most pronounced for the *pfmc-2tm* family. If these altered profiles are indeed linked to the activation of particular *var* promoters, or if they are simply due to stochastic variation remains to be investigated using more targeted approaches.

We were also interested in asking if any of the recently described *hyp* families [22] were differentially expressed in our transfected cell lines. While the temporal transcription profile of most *hyp* genes was similar in all lines, we observed a striking up-regulation in relative *hyp4* transcription in all time points in 3D7/pfmc-2tm parasites compared to all other lines (Fig. S6). This observation is particularly interesting given that all nine *hyp4* paralogs are located in head-to-tail orientation directly upstream of a *pfmc-2tm* gene locus.

Preliminary evidence suggests that activation of different *var* promoter types is carried out by a common regulatory pathway

To date, it is unknown if *var* gene activation involves a single regulatory pathway, or if distinct transcriptional regulators mediate subtype-specific

activation of *var* genes. To gain insight into this interesting question we interrogated our datasets to identify genes with differential expression in the 3D7/upsA, 3D7/upsB and 3D7/upsC transfectants. This analysis pinpointed a total of only 20 transcripts with greater than three-fold changes in relative expression in TP1 and TP2 in any one compared to both other cell lines. Nine genes were differentially expressed in 3D7/upsC, all of which are heterochromatic [16] and encode known or predicted exported proteins [22] including one PHISTa paralog (Fig. 6A). Interestingly, a previous study reported up-regulation of a *phista* gene in parasites selected to express an upsC-type PfEMP1 [71]. Further, two pairs of directly adjacent genes, PFB0975c/PFB0980w and PF07_0004/MAL7P1.6, were similarly up- and down-regulated, respectively, suggesting that local alterations in the heterochromatic environment affected their transcription equally. Two conserved *Plasmodium* proteins with unknown function and *ctrp* [72] were up-regulated in 3D7/upsB (Fig. 6B). Transcription of these three genes is higher in gametocytes compared to intra-erythrocytic stages [73]. Hence, it is possible that these changes were due to a higher proportion of committed sexual stages in the 3D7/upsB population. Eight genes were differentially expressed in 3D7/upsA, seven of which are associated with PfHP1 [16] (Fig. 6C). Two of the three down-regulated genes, PFB0972w and PFB0970c, fall into the same heterochromatic cluster that was up-regulated in 3D7/upsC (see above). Two up-regulated genes encode PHISTb paralogs and two encode members of the ApiAP2 family of putative transcription factors. One of the *apiap2* genes codes for PfSIP2 (PFF0200c), which has been implicated in upsB silencing [74]. The second *apiap2* gene (PFL1085w) is the only member of this family that is positioned in facultative heterochromatin [16] and thus predisposed to variegated expression.

In summary, this comparative approach identified only minor transcriptional differences between parasites selected for activation of upsA, upsB or upsC promoters. Nearly all changes affected heterochromatic genes that are subject to variegated expression. Despite the differential regulation of two *apiap2* genes, this preliminary analysis failed to detect a clear link between distinct signatures of potential regulatory factors and activation of different *var* promoter subtypes.

Discussion

var upstream regions are the dominant functional elements involved in the three main pillars of *var* gene regulation – silencing, activation and mutually exclusive locus recognition [46,47,49–53,75,76]. Here, we were able to expand our knowledge on this complex control strategy onto a genome-wide level, and to incorporate the analysis of upsA and several additional non-*var* promoters. Our comparative functional analysis separates the *var* and non-*var* subtelomeric gene families into two groups based on a different logic of transcriptional regulation.

The three main *var* promoter types share several functional attributes. First, they display similar relative activities and similar temporal activation profiles restricted to ring stages. Second, as shown here and elsewhere, upsA, upsB and upsC promoters exist in two different states of activity; silenced or active [46,49,51,77]. Third, activation of episomal copies of each promoter type leads to a specific down-regulation of endogeneous *var* transcription. Together, these findings suggests that despite considerable sequence variation between upsA, upsB and upsC upstream regions mutually exclusive activation of all *var* genes is governed by a common regulatory mechanism. This assumption is supported by our failure to identify distinct signatures of transcripts coding for potential transcriptional regulators in 3D7/upsA, 3D7/upsB and 3D7/upsC parasites.

Unlike *var* promoters, the *stevor*, *phistb* and *pfmc-2tm* promoters showed no evidence for reversible states of activity. Similar to the house-keeping gene promoter *cam*, they were active by default already in unselected parasites. Moreover, we observed no specific down-regulation in transcription of the endogenous *rif*, *stevor*, *phistb* and *pfmc-2tm* families in parasites carrying extra copies of the corresponding promoters. Analysis of the *rif* promoter delivered several lines of evidence suggesting co-regulation of head-to-head *rif* and upsA *var* genes. First, the temporal activity profile of this *rif* promoter matches that of endogenous *var* rather than *rif* genes. Second, both promoters display reduced activity levels in the default compared to the activated states. Third, and most compelling, selection for *rif* promoter activation induced an endogenous *var* knock-down similar to that achieved by

var promoters. These results imply that activation of this subset of *rif* genes may be carried out by the same mechanisms that are also required for mutually exclusive activation of upsA genes. Interestingly, head-to-head upsA and *rif* genes were also co-regulated in parasites selected with IgG antibodies from highly immune adults [78], and are similarly up-regulated in absence of silent information regulator 2A (PfSIR2A) [79,80]. In light of this apparent co-regulation *in vitro* and *in vivo* it will be interesting to test if upsA PfEMP1 and RIFINs also co-operate on the protein level.

Our initial attempts to select for an active upsA promoter by WR-selection in 3D7/upsA† parasites were unsuccessful. Likewise, a previous study failed to obtain activated upsA promoters based on drug selection [77]. Intriguingly, the relative h*dhfr-gfp* mRNA levels in unselected 3D7/upsA† parasites were higher than those in most other WR-selected lines (Fig. S4). Hence, the refractoriness of 3D7/upsA† to WR challenge was not due to irreversible upsA silencing. Rather, our results point towards a role for the upsA 5' untranslated region (5' UTR) in translational repression, and we recently made similar observations for the upsC 5' UTR (unpublished). Alternatively, it is also plausible that the phenotypic background in 3D7 parasites is rather incompatible with expression of upsA genes. These *var* subtypes are frequently expressed in clinical isolates and are associated with severe disease [70,81–84]. In cell cultures, however, upsA genes are preferably silenced and have faster switch off rates than upsB and upsC genes [69,70]. Hence, environmental cues are important in determining the predisposition to express certain classes of *var* genes. The differential role of PfSIR2A and PfSIR2B in silencing of *var* gene subgroups [79,80,85] may provide an example of how such signals could be translated into changes in *var* expression. Similarly, the observed increase in transcription of two *apiap2* members in the 3D7/upsA population may indicate a role in modulating *var* transcription. For instance, increased nuclear levels of PfSIP2 may be required for enhanced silencing of upsB genes in upsA-expressing parasites. Expression of the second *apiap2* gene (PFL1085w) is also elevated in parasite field isolates from pregnant women and children [86]. It is therefore tempting to speculate that variegated expression of this regulator may be implicated in expression of PfEMP1 variants linked to severe disease.

Tailored functional genomics approaches are now required to test if these ApiAP2 proteins are indeed involved in regulating differential *var* gene expression.

A currently accepted model proposes the existence of a unique and physically restricted perinuclear zone dedicated to the expression of a single *var* locus only [46,50,79,87]. According to this concept, transcriptional switching would occur through competition between a silenced and the active *var* gene. Our results are fully concordant with this hypothesis. Amongst the gene families investigated here, the mechanism by which singular *var* gene choice occurs is unique and specific to the *var* gene family. Promoters of other gene families do not interfere with this system. We envisage that specific *cis*-acting elements in *var* upstream regions could target *var* loci to such an expression site. Default silencing may be a prerequisite to protect *var* loci from illegitimate expression elsewhere in the nucleus. In fact, the reversible responsiveness of *var* promoters to silencing and activation may represent an inherent contributor to the overall mechanism controlling mutually exclusive transcription. On the contrary, activation of *stevor*, *phistb* and *pfmc-2tm* promoters is not mutually exclusive and appears to be mediated by an existing pool of readily available specific transcription factors.

In a previous study, Howitt and colleagues observed that with increasing plasmid copy numbers, where *var*, *rif*, *stevor* or *pfmc-2tm* promoters controlled transcription of the *bsd* resistance gene, transcription of endogenous gene families decreased in a family-transcending manner [61]. To explain these observations, the authors proposed the existence of a common titratable factor in activating subtelomeric gene families in *P. falciparum* [61]. In spite of a similar experimental setup, and similar plasmid copy numbers per parasite, our analysis is inconsistent with such a concept. Similar to Howitt *et al.* we observed down-regulation of *var* and to some extent also *rif*, *stevor*, *pfmc-2tm* and other heterochromatic families in several transfected cell lines. However, this reduced transcription was not caused by absorbance of a limiting factor specifically by subtelomeric virulence gene promoters. We draw this conclusion since this effect was not specific to plasmids carrying gene family promoters but was also observed in the control transfectant 3D7/cam. Additionally, the episomal promoters maintained their

characteristic temporal activation profiles suggesting that distinct sets of transcriptional activators control transcription of different gene families.

Recent genome-wide ChIP experiments demonstrated the ubiquitous enrichment of H3K9me3 and PfHP1 over *P. falciparum*-specific gene family islands throughout the genome [15–17]. The reversibility of these epigenetic marks provides the basis for clonally variant expression of subtelomeric genes. For instance, *var* gene activation is linked to the local exchange of H3K9me3/PfHP1 with H3K9ac and methylation of H3K4 predominantly in the upstream region [48]. Such a functional epigenetic footprint is probably also involved in transcriptional activation of silenced non-*var* family members including *stevor*, *pfmc-2tm*, *rif* and others. However, the apparent lack of autonomous *cis*-acting elements capable of recruiting the silencing machinery to these promoters suggests that variant expression of these families is passively controlled by *cis* and/or *trans* effects of neighbouring loci, which actively participate in structuring the local chromatin environment. Interestingly, parasites selected for the expression of a unique *var* gene show no evidence for co-activation of neighbouring *rif* and *stevor* members [59,60]. Similarly, knock-down in endogenous *var* transcription in the four cell lines analysed here was unlinked to concomitant down-regulation of juxtaposed genes. Overall, this indicates that variant expression of non-*var* family members is not directly linked to the activity of adjacent *var* loci, and that the chromatin alterations required to permit *var* gene activation are locally confined.

Overall, our study demonstrates a lack of systematic regulatory cross-talk between different subtelomeric gene families in *P. falciparum*. The most interesting circumstantial evidence for potential co-regulation is related to the *pfmc-2tm* family. First, we observed a significant down- and up-regulation of *pfmc-2tm* genes in 3D7/upsA and 3D7/upsB, respectively. It is worth mentioning that similar to our *in vitro* data a previous study reported inordinate *pfmc-2tm* expression patterns *in vivo* [58]. However, if *pfmc-2tm* expression is truly co-regulated with subtype-specific *var* gene transcription, or if these variations result from stochastic fluctuations in temporal *pfmc-2tm* activity, needs to be tested in future studies. Second, *hyp4* genes displayed a striking up-regulation in 3D7/pfmc-2tm compared to all other cell lines. Due to 99%

sequence identity between eight of the ten *hyp4* paralogs [88], our hybridisation data fail to distinguish between transcription of individual members. It thus remains unclear if a single or several *hyp4* genes were up-regulated. Intriguingly, however, all *hyp4* genes are located in head-to-tail orientation directly upstream of *pfmc-2tm* genes. The specific up-regulation of *hyp4* transcription in 3D7/pfmc-2tm together with the strict positional link between both gene types suggests that these two gene families are co-expressed and that their exported gene products may be functionally inter-connected.

To the best of our knowledge, this is the first study comparing multiple transfected *P. falciparum* lines by global transcriptional profiling. Our data suggest that drug selection of transfected parasites *per se* causes changes in overall gene transcription. Such alterations may be linked to the initial selection of subpopulations after electroporation, or may be elicited by continuous drug challenge and/or forced replication of episomes. As an important notion for future studies, we advise to use mock transfectants rather than untransfected parasites as controls for comparative transcriptome analyses involving transgenic lines. Another important technical aspect of this study relates to the use of transgenic *P. falciparum* to study the function of variant surface antigens. While the ability to create PfEMP1 knock-down parasites using such a transfection-based approach continues to be extremely informative for the analysis of immune responses to and binding properties of PfEMP1 [46,55–57], our results show that this approach may not be feasible for the analysis of other suspected surface antigens and virulence factors.

In this work we shed important light on the transcriptional control of virulence gene families in *P. falciparum* and possibly that of other pathogens. We show that activation of the three major types of *var* gene promoters is equally important in the process of singular *var* gene choice in *P. falciparum*. The transition of a *var* gene promoter from the silenced to the activated state most likely represents an integral part of the overall mutual exclusion mechanism. We also demonstrate that this mode of regulation is unique to *var* genes; non-*var* families such as *rif*, *stevor*, *phist* and *pfmc-2tm* do not obey the same principle of mutually exclusive transcription. This striking difference in the

control of *var* versus non-*var* gene families underscores the high significance of antigenic variation of PfEMP1 in parasite survival and transmission.

Materials and methods

Plasmid construction

All transfection constructs generated in this study are derivatives of pBcam which was itself obtained by the following procedure. The h*dhfr* gene in pBcam_MCS [16] was replaced with a h*dhfr-gfp* fusion gene. The h*dhfr-gfp* fusion was generated by ligation of a h*dhfr* PCR product into *BamH*I/*Sal*I-digested pGEM3-Zf(+) (Promega). The *Not*I site at the 5' end of the forward primer is followed by a 5bp AAAACA sequence, which naturally occurs directly upstream of the PFL1960w *var* ATG. The *gfp* gene was amplified from pARLmTGFP [89] and ligated in frame with the h*dhfr* gene into *Sal*I/*Avr*II-digested pGEM-h*dhfr*. The fusion gene was excised with *BamH*I/*Avr*II and cloned into *BamH*I/*Avr*II-digested pBcam_MCS to obtain pBcamHG. The *var* gene intron was amplified from pHBupsCRI [46], digested with *Mfe*I and cloned into *EcoR*I-digested pBcamHG to obtain pBcam. 5' upstream regions of interest were amplified by PCR from 3D7 gDNA, digested with *Bgl*II/*Not*I and cloned upstream of the h*dhfr-gfp* reporter gene into *Bgl*II/*Not*I-digested pBcam. The resulting constructs were named pBupsA†, pBupsB, pBupsC, pBrif, pBstevor, pBphistb and pBpfmc-2tm. To create plasmid pBupsA, the *bsd* gene and PbDT3' terminator sequence was amplified from pBcam, digested with *Bgl*II/*Kpn*I and cloned in reverse orientation into *Bgl*II/*Kpn*I-digested pBrif. Primer sequences, accession numbers and lengths of the respective promoter sequences are shown in Table S1.

Parasite culture and transfection

P. falciparum 3D7 parasites were cultured as described previously [90]. Transfections were performed as described [46] and parasites were selected on 2.5□g/ml BSD and 4nM WR99210. Growth synchronisation was achieved by repeated sorbitol lysis [91]. To obtain an eight-hour growth window pre-synchronized parasites were synchronized twice 16 hours apart.

Transcriptional profiling by microarray analysis

For microarray analysis, total RNA was harvested at four consecutive timepoints: 6-14 hours post-invasion (hpi), 14-22 hpi, 22-30 hpi and 30-38 hpi. A pool of total RNA isolated from sychronized 3D7 parasites at five consecutive timepoints (10 hour growth window each) covering the complete intra-erythrocytic developmental cycle was used as reference sample. Total RNA was prepared by disrupting 1.5 ml of RBC pellet (>3% parasitaemia) with 10ml of TriReagent (Sigma-Aldrich) following manufacturer's instructions. cDNA was generated from total RNA and labelled as previously described [65] and hybridised to a long oligonucleotide genome-wide microarray containing 10166 probes [64]. The raw microarray data representing mRNA abundance ratios between each sample and the reference pool were subjected to lowess slide normalization and background filtering as implemented by the Acuity 4.0 program (Molecular Devices). For genes represented by more than one probe ratios were averaged (Dataset S1). tRNA and rRNA signals and genes for which signals were detected in less than 75% of all samples were excluded from the analysis.

Data analysis

Genes and arrays were clustered by average linkage clustering using Gene Cluster 3.0 [92]. The array tree and heatmaps were generated using the Java Treeview program [93]. Boxplots and scatterplots were generated using the statistical package R (version 2.10.0). Paired and unpaired t-tests were used to assess statistical significance between means of log2 ratios obtained from microarray hybridisations and qRT-PCR triplicate values, respectively. Differentially expressed genes presented in Fig. S2 were analysed by functional pathway enrichment using the KEGG and MPMP databases [94,95] and manual curation.

Genomic DNA extraction, RNA extraction and cDNA synthesis

Parasites were released from iRBCs in 0.15% saponin. Total RNA was isolated using TriReagent (Sigma-Aldrich) and further purified using the RNeasy Plus Mini Kit (Qiagen) for removal of gDNA. Residual gDNA was digested with TURBO DNA-*free*™ (Ambion). All samples were tested negative

for contaminating gDNA by quantitative PCR (qPCR). RNA was reverse transcribed using the RETROscript Kit® (Ambion). gDNA was extracted from the same parasite samples by addition of 3% SDS/0.27 M Na-acetate to the parasite pellet followed by phenol/chloroform extraction and precipitation. gDNA was resuspended in ddH_2O and diluted 1:100 for qPCR analysis.

Quantitative real-time PCR

For determination of the temporal promoter activity profiles RNA and gDNA were harvested from the same cultures that were also used for microarray analysis. For analysis of promoter activities in unselected and drug-selected populations parasites were synchronized as described above and gDNA and total RNA was harvested at 6-14hpi (TP1) for 3D7/upsA, 3D7/upsB, 3D7/upsC, 3D7/rif, 3D7/phist and 3D7/cam, and at 22-30hpi (TP3) for 3D7/stevor and 3D7/pfmc-2tm, respectively. qRT-PCR experiments were performed on three independent biological samples. Plasmid copy numbers were determined by qPCR on gDNA and calculated by dividing the absolute copy numbers obtained for h*dhfr-gfp* by the value obtained for the PFL0900c (arginyl-tRNA synthetase) locus [53]. All reactions were run in triplicate yielding virtually identical Ct values. Five serial 1:10 dilutions of 3D7 genomic or plasmid DNA were used as standard for absolute quantification. Relative transcript profiles were calculated by normalisation against the house-keeping gene PFL0900c. Cycling conditions were: 50°C, 2min; 95°C 10min, followed by 40 cylces of 95°C, 15sec; 58°C 1min. Product-specific amplification was ensured by performing melting curves for each reaction. Reactions were performed at final primer concentrations of 0.4uM using Power SYBR® Green Master Mix (Applied Biosystems) on a StepOnePlus™ Real-Time PCR System (Applied Biosystems) in a reaction volume of 12ul. Primers are listed in Table S1.

Southern blotting

gDNA was subjected to Southern analysis using standard procedures. gDNA was digested with *EcoR*I/*Pvu*II and separated on 0.7% agarose gels. Southern blots were probed with random-primed ^{32}P-dATP-labelled h*dhfr*.

Accession numbers

PlasmoDB (www.plasmoDB.org) accession numbers for genes and proteins discussed in this publication are: upsA-type *var* (PF13_0003), upsB-type *var* (PFL0005w), upsC-type *var* (PFL1960w), *rif* (PF13_0004), *stevor* (PFL2610w), *phistb* (PFL2540w), *cam* (PF14_0323), PfSIP2 (PFF0200c), ApiAP2 factor (PFL1085c), PfHP1 (PFL1005c).

Acknowledgements

We thank Sachel Mok for help in microarray analysis and Christian Flueck for critically reading the manuscript. This work was supported by a Swiss National Science Foundation award (PP00A-110835; PP00P3_130203; www.snf.ch), a Singapore Biomedical Research Council grant (09/1/22/19/614; www.a-star.edu.sg), and a National Medical Research Council grant (M48080019; www.nmrc.gov.sg).

References

1. Biggs BA, Gooze L, Wycherley K, Wollish W, Southwell B, et al. (1991) Antigenic variation in Plasmodium falciparum. Proc Natl Acad Sci U S A 88: 9171-9174.
2. Roberts DJ, Craig AG, Berendt AR, Pinches R, Nash G, et al. (1992) Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature 357: 689-692.
3. Smith JD, Chitnis CE, Craig AG, Roberts DJ, Hudson-Taylor DE, et al. (1995) Switches in expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell 82: 101-110.
4. Kyes S, Horrocks P, Newbold C (2001) Antigenic variation at the infected red cell surface in malaria. Annu Rev Microbiol 55: 673-707.
5. Gardner MJ, Hall N, Fung E, White O, Berriman M, et al. (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419: 498-511.
6. Barry JD, Ginger ML, Burton P, McCulloch R (2003) Why are parasite contingency genes often associated with telomeres? Int J Parasitol 33: 29-45.
7. Deitsch KW, Lukehart SA, Stringer JR (2009) Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat Rev Microbiol 7: 493-503.
8. Myler PJ, Allison J, Agabian N, Stuart K (1984) Antigenic variation in African trypanosomes by gene replacement or activation of alternate telomeres. Cell 39: 203-211.
9. De Las PA, Pan SJ, Castano I, Alder J, Cregg R, et al. (2003) Virulence-related surface glycoproteins in the yeast pathogen Candida glabrata are encoded in subtelomeric clusters and subject to. Genes Dev 17: 2245-2258.
10. Keely SP, Renauld H, Wakefield AE, Cushion MT, Smulian AG, et al. (2005) Gene arrays at Pneumocystis carinii telomeres. Genetics 170: 1589-1600.
11. Cooper JP, Watanabe Y, Nurse P (1998) Fission yeast Taz1 protein is required for meiotic telomere clustering and recombination. Nature 392: 828-831.
12. Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, et al. (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature 407: 1018-1022.
13. Taylor HM, Kyes SA, Newbold CI (2000) Var gene diversity in Plasmodium falciparum is generated by frequent recombination events. Mol Biochem Parasitol 110: 391-397.

14. Gottschling DE, Aparicio OM, Billington BL, Zakian VA (1990) Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. Cell 63: 751-762.
15. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, et al. (2009) Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc Natl Acad Sci U S A
16. Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, et al. (2009) Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog 5: e1000569.
17. Lopez-Rubio JJ, Mancio-Silva L, Scherf A (2009) Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell Host Microbe 5: 179-190.
18. Su XZ, Heatwole VM, Wertheimer SP, Guinet F, Herrfeldt JA, et al. (1995) The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. Cell 82: 89-100.
19. Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC, et al. (1995) Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. Cell 82: 77-87.
20. Weber JL (1988) Interspersed repetitive DNA from Plasmodium falciparum. Mol Biochem Parasitol 29: 117-124.
21. Cheng Q, Cloonan N, Fischer K, Thompson J, Waine G, et al. (1998) stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens. Mol Biochem Parasitol 97: 161-176.
22. Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, et al. (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. Genome Biol 7: R12.
23. Sam-Yellowe TY, Florens L, Johnson JR, Wang T, Drazba JA, et al. (2004) A Plasmodium gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. Genome Res 14: 1052-1059.
24. Voss TS, Thompson JK, Waterkeyn J, Felger I, Weiss N, et al. (2000) Genomic distribution and functional characterisation of two distinct and conserved Plasmodium falciparum var gene 5' flanking sequences. Mol Biochem Parasitol 107: 103-115.
25. Kraemer SM, Smith JD (2003) Evidence for the importance of genetic structuring to the structural and functional specialization of the Plasmodium falciparum var gene family. Mol Microbiol 50: 1527-1538.
26. Lavstsen T, Salanti A, Jensen AT, Arnot DE, Theander TG (2003) Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. Malar J 2: 27.
27. Leech JH, Barnwell JW, Miller LH, Howard RJ (1984) Identification of a strain-specific malarial antigen exposed on the surface of Plasmodium falciparum-infected erythrocytes. J Exp Med 159: 1567-1575.
28. MacPherson GG, Warrell MJ, White NJ, Looareesuwan S, Warrell DA (1985) Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. Am J Pathol 119: 385-401.
29. Carlson J, Helmby H, Hill AV, Brewster D, Greenwood BM, et al. (1990) Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. Lancet 336: 1457-1460.
30. Rowe A, Obeiro J, Newbold CI, Marsh K (1995) Plasmodium falciparum rosetting is associated with malaria severity in Kenya. Infect Immun 63: 2323-2326.
31. Pongponratn E, Riganti M, Punpoowong B, Aikawa M (1991) Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. Am J Trop Med Hyg 44: 168-175.
32. Lavazec C, Sanyal S, Templeton TJ (2007) Expression switching in the stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. Mol Microbiol 64: 1621-1634.
33. Abdel-Latif MS, Khattab A, Lindenthal C, Kremsner PG, Klinkert MQ (2002) Recognition of variant Rifin antigens by human antibodies induced during natural Plasmodium falciparum infections. Infect Immun 70: 7013-7021.
34. Abdel-Latif MS, Dietz K, Issifou S, Kremsner PG, Klinkert MQ (2003) Antibodies to Plasmodium falciparum rifin proteins are associated with rapid parasite clearance and asymptomatic infections. Infect Immun 71: 6229-6233.
35. Fernandez V, Hommel M, Chen Q, Hagblom P, Wahlgren M (1999) Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are

encoded by the rif gene family and are the target of human immune responses. J Exp Med 190: 1393-1404.
36. Kyes SA, Rowe JA, Kriek N, Newbold CI (1999) Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. Proc Natl Acad Sci U S A 96: 9333-9338.
37. Schreiber N, Khattab A, Petter M, Marks F, Adjei S, , et al. (2008) Expression of Plasmodium falciparum 3D7 STEVOR proteins for evaluation of antibody responses following malaria infections in naive infants. Parasitology 135: 155-167.
38. Joannin N, Abhiman S, Sonnhammer EL, Wahlgren M (2008) Sub-grouping and sub-functionalization of the RIFIN multi-copy protein family. BMC Genomics 9: 19.
39. Petter M, Haeggstrom M, Khattab A, Fernandez V, Klinkert MQ, et al. (2007) Variant proteins of the Plasmodium falciparum RIFIN family show distinct subcellular localization and developmental expression patterns. Mol Biochem Parasitol 156: 51-61.
40. Petter M, Bonow I, Klinkert MQ (2008) Diverse expression patterns of subgroups of the rif multigene family during Plasmodium falciparum gametocytogenesis. PLoS ONE 3: e3779.
41. Lavazec C, Sanyal S, Templeton TJ (2006) Hypervariability within the Rifin, Stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. Nucleic Acids Res 34: 6696-6707.
42. Khattab A, Bonow I, Schreiber N, Petter M, Schmetz C, et al. (2008) Plasmodium falciparum variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. Malar J 7: 137.
43. Blythe JE, Yam XY, Kuss C, Bozdech Z, Holder AA, et al. (2008) Plasmodium falciparum STEVOR proteins are highly expressed in patient isolates and located in the surface membranes of infected red blood cells and the apical tips of merozoites. Infect Immun 76: 3329-3336.
44. Niang M, Yan Y, X, Preiser PR (2009) The Plasmodium falciparum STEVOR multigene family mediates antigenic variation of the infected erythrocyte. PLoS Pathog 5: e1000307.
45. Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, et al. (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. EMBO J 17: 5418-5426.
46. Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, et al. (2006) A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. Nature 439: 1004-1008.
47. Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, et al. (2007) Epigenetic memory at malaria virulence genes. Proc Natl Acad Sci U S A 104: 899-902.
48. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez RR, et al. (2007) 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol 66: 1296-1305.
49. Voss TS, Tonkin CJ, Marty AJ, Thompson JK, Healer J, et al. (2007) Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. Mol Microbiol 66: 139-150.
50. Dzikowski R, Li F, Amulic B, Eisberg A, Frank M, et al. (2007) Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. EMBO Rep 8: 959-965.
51. Dzikowski R, Frank M, Deitsch K (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. PLoS Pathog 2: e22.
52. Calderwood MS, Gannoun-Zaki L, Wellems TE, Deitsch KW (2003) Plasmodium falciparum var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. J Biol Chem 278: 34125-34132.
53. Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E, et al. (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. J Biol Chem 281: 9942-9952.
54. Deitsch KW, Calderwood MS, Wellems TE (2001) Malaria. Cooperative silencing elements in var genes. Nature 412: 875-876.
55. D'Ombrain MC, Voss TS, Maier AG, Pearce JA, Hansen DS, et al. (2007) Plasmodium falciparum erythrocyte membrane protein-1 specifically suppresses early production of host interferon-gamma. Cell Host Microbe 2: 130-138.

56. Elliott SR, Spurck TP, Dodin JM, Maier AG, Voss TS, et al. (2007) Inhibition of dendritic cell maturation by malaria is dose dependent and does not require Plasmodium falciparum erythrocyte membrane protein 1. Infect Immun 75: 3621-3632.
57. Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, et al. (2008) Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell 134: 48-61.
58. Mackinnon MJ, Li J, Mok S, Kortok MM, Marsh K, et al. (2009) Comparative transcriptional and genomic analysis of Plasmodium falciparum field isolates. PLoS Pathog 5: e1000644.
59. Cabral FJ, Wunderlich G (2009) Transcriptional memory and switching in the Plasmodium falciparumrif gene family. Mol Biochem Parasitol 168: 186-190.
60. Sharp S, Lavstsen T, Fivelman QL, Saeed M, McRobert L, et al. (2006) Programmed transcription of the var gene family, but not of stevor, in Plasmodium falciparum gametocytes. Eukaryot Cell 5: 1206-1214.
61. Howitt CA, Wilinski D, Llinas M, Templeton TJ, Dzikowski R, et al. (2009) Clonally variant gene families in Plasmodium falciparum share a common activation factor. Mol Microbiol 73: 1171-1185.
62. Kaviratne M, Khan SM, Jarra W, Preiser PR (2002) Small variant STEVOR antigen is uniquely located within Maurer's clefts in Plasmodium falciparum-infected red blood cells. Eukaryot Cell 1: 926-935.
63. Kyes S, Pinches R, Newbold C (2000) A simple RNA analysis method shows var and rif multigene family expression patterns in Plasmodium falciparum. Mol Biochem Parasitol 105: 311-315.
64. Hu G, Llinas M, Li J, Preiser PR, Bozdech Z (2007) Selection of long oligonucleotides for gene expression microarrays using weighted rank-sum strategy. BMC Bioinformatics 8: 350.
65. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, et al. (2003) The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol 1: E5.
66. Marti M, Baum J, Rug M, Tilley L, Cowman AF (2005) Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. J Cell Biol 171: 587-592.
67. Llinas M, Bozdech Z, Wong ED, Adai AT, DeRisi JL (2006) Comparative whole genome transcriptome analysis of three Plasmodium falciparum strains. Nucleic Acids Res 34: 1166-1173.
68. Yamaguchi H, Yamamoto C, Tanaka N (1965) Inhibition of protein synthesis by blasticidin S. I. Studies with cell-free systems from bacterial and mammalian cells. J Biochem 57: 667-677.
69. Frank M, Dzikowski R, Amulic B, Deitsch K (2007) Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol Microbiol 64: 1486-1498.
70. Peters JM, Fowler EV, Krause DR, Cheng Q, Gatton ML (2007) Differential changes in Plasmodium falciparum var transcription during adaptation to culture. J Infect Dis 195: 748-755.
71. Mok BW, Ribacke U, Winter G, Yip BH, Tan CS, et al. (2007) Comparative transcriptomal analysis of isogenic Plasmodium falciparum clones of distinct antigenic and adhesive phenotypes. Mol Biochem Parasitol 151: 184-192.
72. Trottein F, Triglia T, Cowman AF (1995) Molecular cloning of a gene from Plasmodium falciparum that codes for a protein sharing motifs found in adhesive molecules from mammals and plasmodia. Mol Biochem Parasitol 74: 129-141.
73. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, et al. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. Science 301: 1503-1508.
74. Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BT, et al. (2010) A major role for the Plasmodium falciparum ApiAP2 protein PfSIP2 in chromosome end biology. PLoS Pathog 6: e1000784.
75. Voss TS, Kaestli M, Vogel D, Bopp S, Beck HP (2003) Identification of nuclear proteins that interact differentially with Plasmodium falciparum var gene promoters. Mol Microbiol 48: 1593-1607.
76. Gannoun-Zaki L, Jost A, Mu J, Deitsch KW, Wellems TE (2005) A silenced Plasmodium falciparum var promoter can be activated in vivo through spontaneous deletion of a silencing element in the intron. Eukaryot Cell 4: 490-492.

77. Muhle RA, Adjalley S, Falkard B, Nkrumah LJ, Muhle ME, et al. (2009) A var gene promoter implicated in severe malaria nucleates silencing and is regulated by 3' untranslated region and intronic cis-elements. Int J Parasitol 39: 1425-1439.
78. Wang CW, Magistrado PA, Nielsen MA, Theander TG, Lavstsen T (2009) Preferential transcription of conserved rif genes in two phenotypically distinct Plasmodium falciparum parasite lines. Int J Parasitol 39: 655-664.
79. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, et al. (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. Cell 121: 13-24.
80. Tonkin CJ, Carret CK, Duraisingh MT, Voss TS, Ralph SA, et al. (2009) Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS Biol 7: e84.
81. Bull PC, Pain A, Ndungu FM, Kinyanjui SM, Roberts DJ, et al. (2005) Plasmodium falciparum antigenic variation: relationships between in vivo selection, acquired antibody response, and disease severity. J Infect Dis 192: 1119-1126.
82. Kyriacou HM, Stone GN, Challis RJ, Raza A, Lyke KE, et al. (2006) Differential var gene transcription in Plasmodium falciparum isolates from patients with cerebral malaria compared to hyperparasitaemia. Mol Biochem Parasitol 150: 211-218.
83. Rottmann M, Lavstsen T, Mugasa JP, Kaestli M, Jensen AT, et al. (2006) Differential expression of var gene groups is associated with morbidity caused by Plasmodium falciparum infection in Tanzanian children. Infect Immun 74: 3904-3911.
84. Kirchgatter K, Portillo HA (2002) Association of severe noncerebral Plasmodium falciparum malaria in Brazil with expressed PfEMP1 DBL1 alpha sequences lacking cysteine residues. Mol Med 8: 16-23.
85. Merrick CJ, Dzikowski R, Imamura H, Chuang J, Deitsch K, et al. (2010) The effect of Plasmodium falciparum Sir2a histone deacetylase on clonal and longitudinal variation in expression of the var family of virulence genes. Int J Parasitol 40: 35-43.
86. Vignali M, Armour CD, Chen J, Morrison R, Castle JC, et al. (2011) NSR-seq transcriptional profiling enables identification of a gene signature of Plasmodium falciparum parasites infecting children. J Clin Invest 121: 1119-1129.
87. Ralph SA, Scheidig-Benatar C, Scherf A (2005) Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. Proc Natl Acad Sci U S A 102: 5414-5419.
88. Templeton TJ (2009) The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, Plasmodium falciparum. Mol Biochem Parasitol 166: 109-116.
89. Struck NS, de Souza DS, Langer C, Marti M, Pearce JA, et al. (2005) Re-defining the Golgi complex in Plasmodium falciparum using the novel Golgi marker PfGRASP. J Cell Sci 118: 5603-5613.
90. Trager W, Jenson JB (1978) Cultivation of malarial parasites. Nature 273: 621-622.
91. Lambros C, Vanderberg JP (1979) Synchronization of Plasmodium falciparum erythrocytic stages in culture. J Parasitol 65: 418-420.
92. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863-14868.
93. Saldanha AJ (2004) Java Treeview--extensible visualization of microarray data. Bioinformatics 20: 3246-3248.
94. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. Nucleic Acids Res 32: D277-D280.
95. Ginsburg H (2006) Progress in in silico functional genomics: the malaria Metabolic Pathways database. Trends Parasitol 22: 238-240.

Figure Legends

Figure 1. Transgenic reporter cell lines. (A) Schematic map of the parental plasmid pBcam. The *bsd* resistance cassette (white box) selects for stable plasmid maintenance. The h*dhfr-gfp* cassette (grey box) allows selecting for active promoters of interest (red arrow). The *var* gene intron is indicated by a v-shape. A 0.5kb TARE6 element is shown by vertical lines. hsp86 5', *hsp86* promoter; *Pb* DT 3', *P.* berghei *dhfr*-thymidilate synthase terminator; *hrp2* 3'; histidine-rich protein 2 terminator. **(B)** Descendants of pBcam were obtained by replacing the *cam* promoter with promoters of interest (red arrows) using *Bgl*II and *Not*I. Numbers represent nucleotide positions in relation to the ATG start codon. Grey dashed arrows represent the bi-directional upsA/*rif* promoter. Plasmid names and cell lines are indicated. White and grey bars represent average plasmid copy numbers in unselected and WR-selected cell lines, respectively. n.a., not available. **(C)** Top: schematic map of pBupsA where the *bsd* and h*dhfr-gfp* selectable markers are controlled by the bi-directional upsA/*rif* promoter. Bottom: average plasmid copy number in BSD-selected 3D7/upsA parasites.

Figure 2. Hierarchical clustering of 36 transcriptomes. (A) Hierarchical average linkage clustering of 3D7 and eight transgenic cell lines at four time points during the IDC. Numbers indicate similarity scores of two joined elements by uncentered Pearson correlation. TP1, 6-14 hpi; TP2, 14-22hpi; TP3, 22-30hpi; TP4, 30-38hpi. **(B)** Bars indicate the number of genes displaying >4-fold change in relative expression (open bars, down-regulated; grey bars, up-regulated) in each cell line and time point relative to 3D7 wild-type (left panel) or the 3D7/cam control transfectant (right panel). Parasites transfected with *var* promoter constructs are shaded grey. hpi, hours post-invasion

Figure 3. Gene family promoters other than *var* are not silenced by default. Relative reporter transcript levels were determined by qRT-PCR (normalised against transcription of arginyl-tRNA synthetase (PFL0900c) and adjusted for plasmid copy numbers). The bars represent the average increase

in relative promoter activity in drug-selected compared to unselected parasites at the time of peak transcription (fold increase). Values derive from three independent biological replicates each (mean +/- s.d.) (see also Fig. S4). Note that the value for the upsA promoter was calculated by dividing the relative *bsd* transcript levels in 3D7/upsA by the relative h*dhfr-gfp* levels in unselected 3D7/upsA†. Significant levels of induction compared to the negative control line 3D7/cam are indicated. hpi, hours post-invasion.

Figure 4. Mutually exclusive transcription is unique and specific to *var* gene family. (A) Knock-down of endogenous *var* gene transcription. The heat-map reflects the mean relative expression of all *var* genes at four time points in each cell line. The box plot visualizes the distribution of relative *var* expression levels in all cell lines at the time of peak transcription (TP1). Parasites transfected with *var* promoter constructs are shaded. The box reflects the lower and upper quartiles, respectively, and the median is indicated by a horizontal line. Outliers are shown as dots. The solid and dotted grey lines are drawn at the median for 3D7/cam and 3D7, respectively. Significant changes in relative *var* expression compared to the control transfectant 3D7/cam are indicated by asterisks ($p<0.01$). **(B)** The heat-map reflects relative transcript abundance for each *var* gene at all four time points in all cell lines.

Figure 5. Multigene families other than *var* are not subject to promoter-induced mutual exclusion. Heat-maps reflect the mean relative expression of all members of the endogenous gene families at four time points in each cell line. Box plots visualize the distribution of relative expression levels for each gene family in all cell lines at the time of peak transcription. The box reflects the lower and upper quartiles, respectively, and the median is indicated by a horizontal line. Outliers are shown as dots. The solid and dotted grey lines are drawn at the median for 3D7/cam and 3D7, respectively. Significant changes in relative expression compared to the control transfectant 3D7/cam are indicated by asterisks ($p<0.01$). ns, not significant. **(A)** Relative *rif* expression. **(B)** Relative *stevor* expression. **(C)** Relative *phist*

expression. **(D)** Relative *pfmc-2tm* expression. Parasite lines transfected with the cognate promoter construct are shaded.

Figure 6. Differential expression in parasite lines selected for activation of different *var* promoter subtypes. (A) Genes with >3-fold changes in relative expression at TP1 and TP2 in 3D7/upsC compared to both 3D7/upsA and 3D7/upsB. (B) Genes with >3-fold changes in relative expression at TP1 and TP2 in 3D7/upsB compared to both 3D7/upsA and 3D7/upsC. (C) Genes with >3-fold changes in relative expression at TP1 and TP2 in 3D7/upsA compared to both 3D7/upsB and 3D7/upsC. The colour code refers to fold changes in relative expression ($2^{\square log2\ ratios}$).

Figure 1

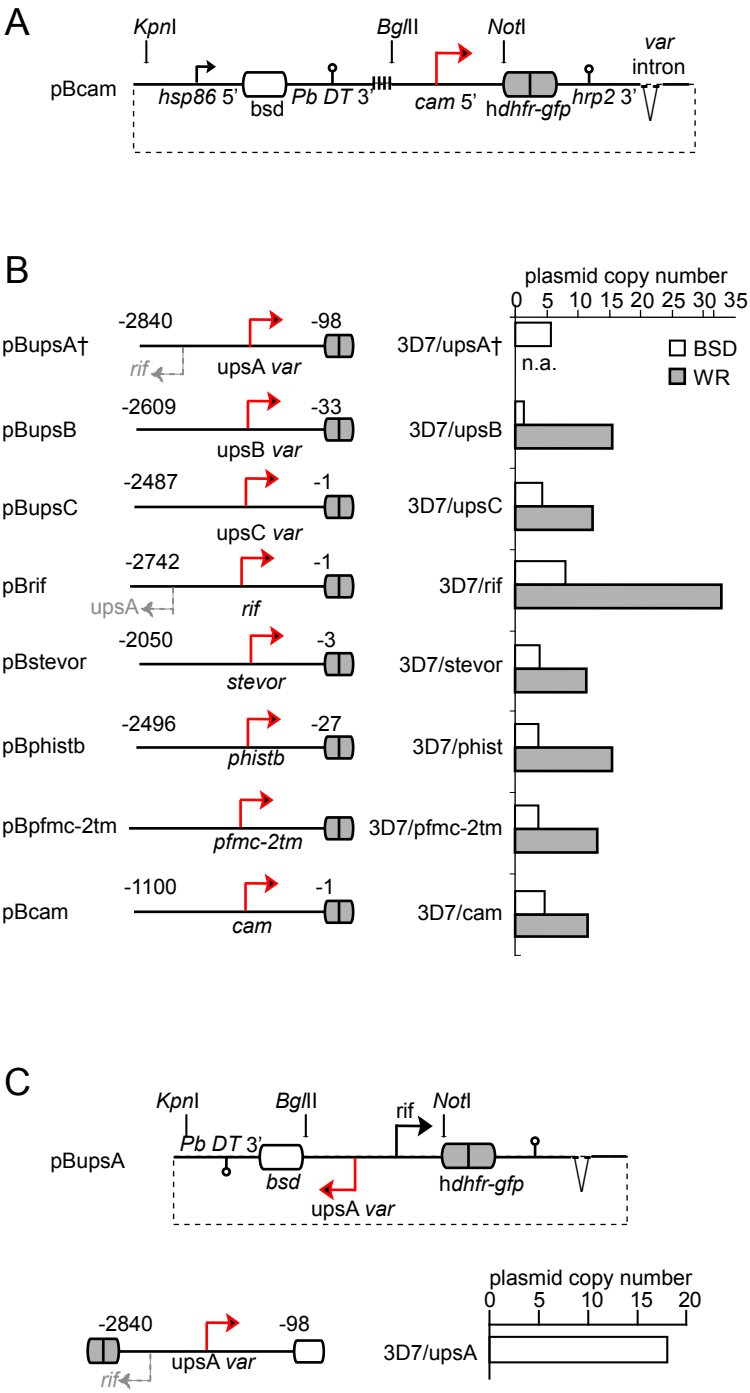

A
KpnI
BglII
NotI
var
intron
pBcam
hsp86 5'
bsd
Pb DT 3'
cam 5'
hdhfr-gfp
hrp2 3'
B
plasmid copy number
0 5 10 15 20 25 30 35
BSD
WR
pBupsA†
-2840
-98
rif
upsA var
3D7/upsA†
n.a.
pBupsB
-2609
-33
upsB var
3D7/upsB
pBupsC
-2487
-1
upsC var
3D7/upsC
pBrif
-2742
-1
upsA
rif
3D7/rif
pBstevor
-2050
-3
stevor
3D7/stevor
pBphistb
-2496
-27
phistb
3D7/phist
pBpfmc-2tm
pfmc-2tm
3D7/pfmc-2tm
pBcam
-1100
-1
cam
3D7/cam
C
KpnI
BglII
rif
NotI
pBupsA
Pb DT 3'
bsd
hdhfr-gfp
upsA var
plasmid copy number
0 5 10 15 20
-2840
-98
upsA var
rif
3D7/upsA

Figure 2

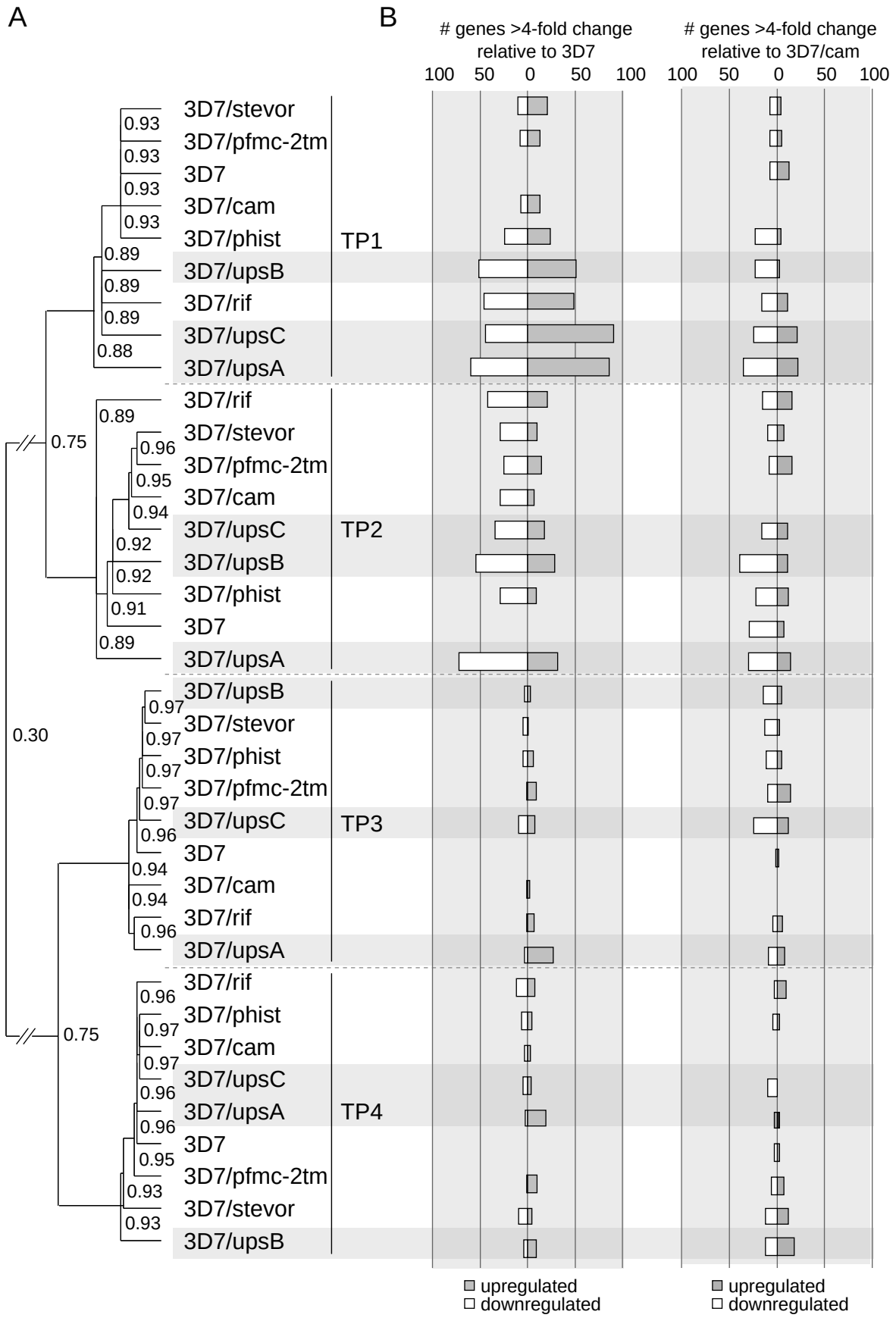

A
B
genes >4-fold change relative to 3D7
genes >4-fold change relative to 3D7/cam
100 50 0 50 100
100 50 0 50 100
3D7/stevor
3D7/pfmc-2tm
3D7
3D7/cam
3D7/phist
3D7/upsB
3D7/rif
3D7/upsC
3D7/upsA
TP1
3D7/rif
3D7/stevor
3D7/pfmc-2tm
3D7/cam
3D7/upsC
3D7/upsB
3D7/phist
3D7
3D7/upsA
TP2
3D7/upsB
3D7/stevor
3D7/phist
3D7/pfmc-2tm
3D7/upsC
3D7
3D7/cam
3D7/rif
3D7/upsA
TP3
3D7/rif
3D7/phist
3D7/cam
3D7/upsC
3D7/upsA
3D7
3D7/pfmc-2tm
3D7/stevor
3D7/upsB
TP4
0.93
0.93
0.93
0.93
0.89
0.89
0.89
0.88
0.89
0.75
0.96
0.95
0.94
0.92
0.92
0.91
0.89
0.30
0.97
0.97
0.97
0.97
0.96
0.94
0.94
0.96
0.75
0.96
0.97
0.97
0.96
0.96
0.95
0.93
0.93
upregulated
downregulated
upregulated
downregulated

Figure 3

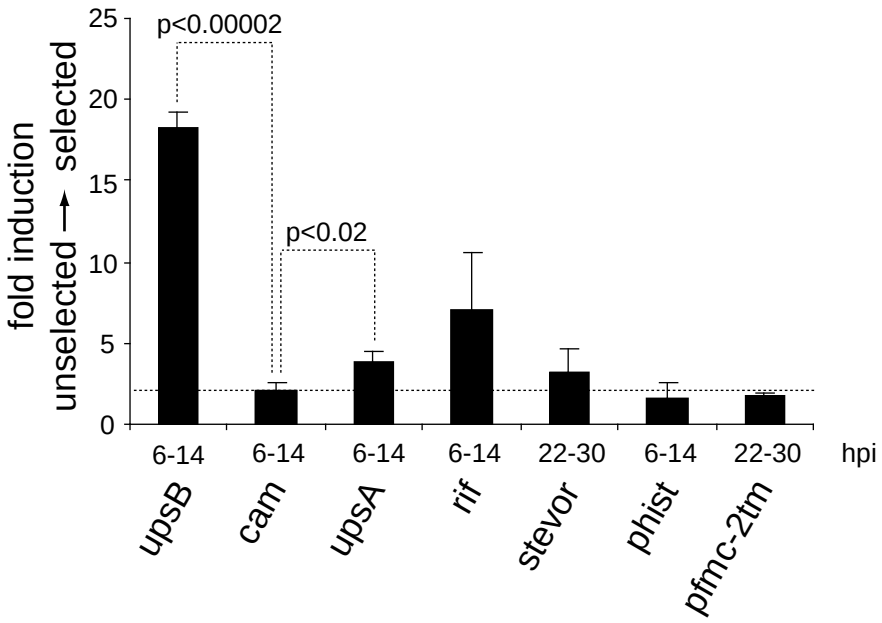

fold induction
unselected → selected
25
20
15
10
5
0
p<0.00002
p<0.02
6-14
6-14
6-14
6-14
22-30
6-14
22-30
hpi
upsB
cam
upsA
rif
stevor
phist
pfmc-2tm

Figure 4

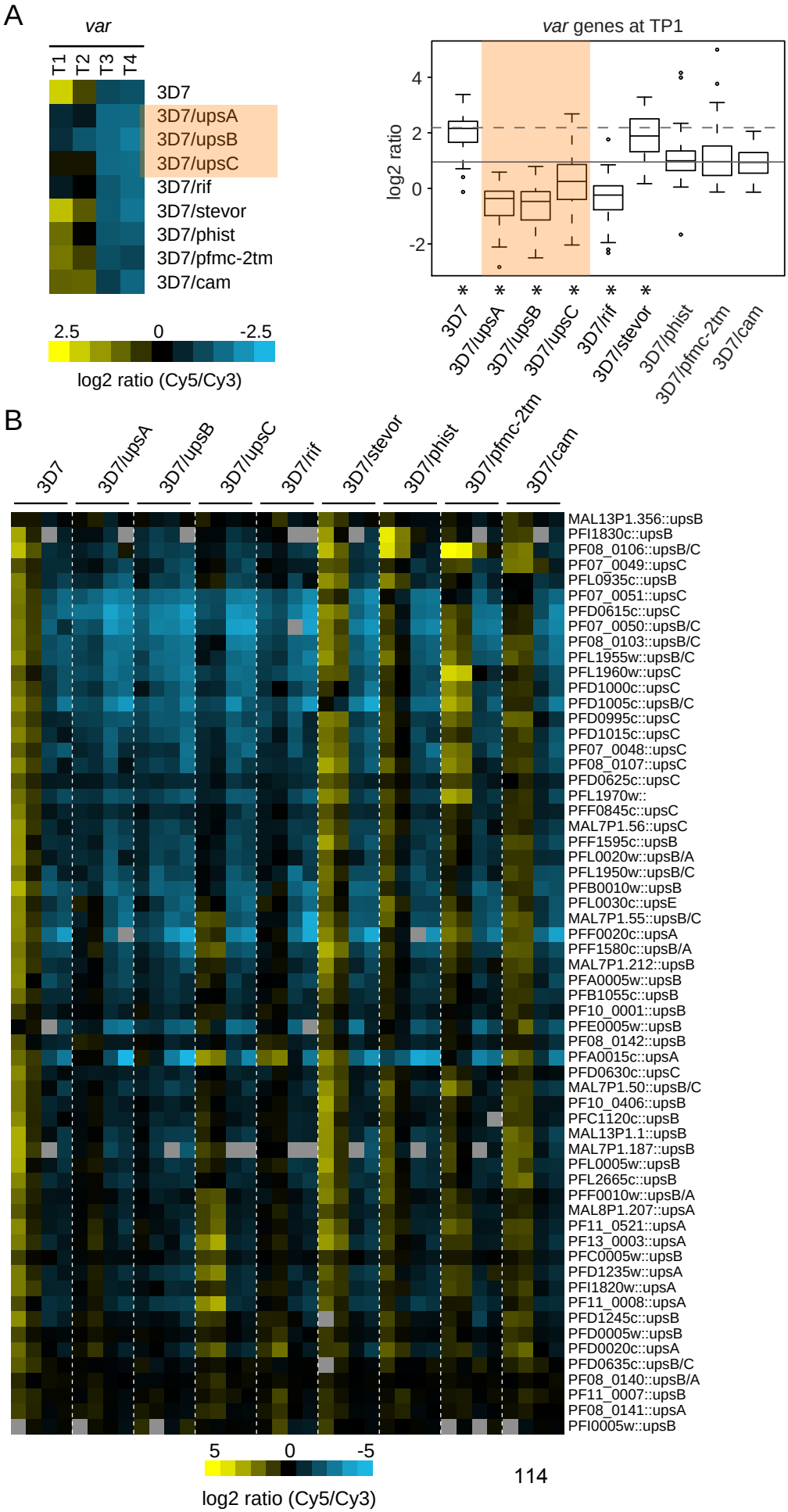

A
var
T1
T2
T3
T4
3D7
3D7/upsA
3D7/upsB
3D7/upsC
3D7/rif
3D7/stevor
3D7/phist
3D7/pfmc-2tm
3D7/cam
2.5
0
-2.5
log2 ratio (Cy5/Cy3)
var genes at TP1
log2 ratio
4
2
0
-2
B
MAL13P1.356::upsB
PFI1830c::upsB
PF08_0106::upsB/C
PF07_0049::upsC
PFL0935c::upsB
PF07_0051::upsC
PFD0615c::upsC
PF07_0050::upsB/C
PF08_0103::upsB/C
PFL1955w::upsB/C
PFL1960w::upsC
PFD1000c::upsC
PFD1005c::upsB/C
PFD0995c::upsC
PFD1015c::upsC
PF07_0048::upsC
PF08_0107::upsC
PFD0625c::upsC
PFL1970w::
PFF0845c::upsC
MAL7P1.56::upsC
PFF1595c::upsB
PFL0020w::upsB/A
PFL1950w::upsB/C
PFB0010w::upsB
PFL0030c::upsE
MAL7P1.55::upsB/C
PFF0020c::upsA
PFF1580c::upsB/A
MAL7P1.212::upsB
PFA0005w::upsB
PFB1055c::upsB
PF10_0001::upsB
PFE0005w::upsB
PF08_0142::upsB
PFA0015c::upsA
PFD0630c::upsC
MAL7P1.50::upsB/C
PF10_0406::upsB
PFC1120c::upsB
MAL13P1.1::upsB
MAL7P1.187::upsB
PFL0005w::upsB
PFL2665c::upsB
PFF0010w::upsB/A
MAL8P1.207::upsA
PF11_0521::upsA
PF13_0003::upsA
PFC0005w::upsB
PFD1235w::upsA
PFI1820w::upsA
PF11_0008::upsA
PFD1245c::upsB
PFD0005w::upsB
PFD0020c::upsA
PFD0635c::upsB/C
PF08_0140::upsB/A
PF11_0007::upsB
PF08_0141::upsA
PFI0005w::upsB
5
0
-5
log2 ratio (Cy5/Cy3)

Figure 5

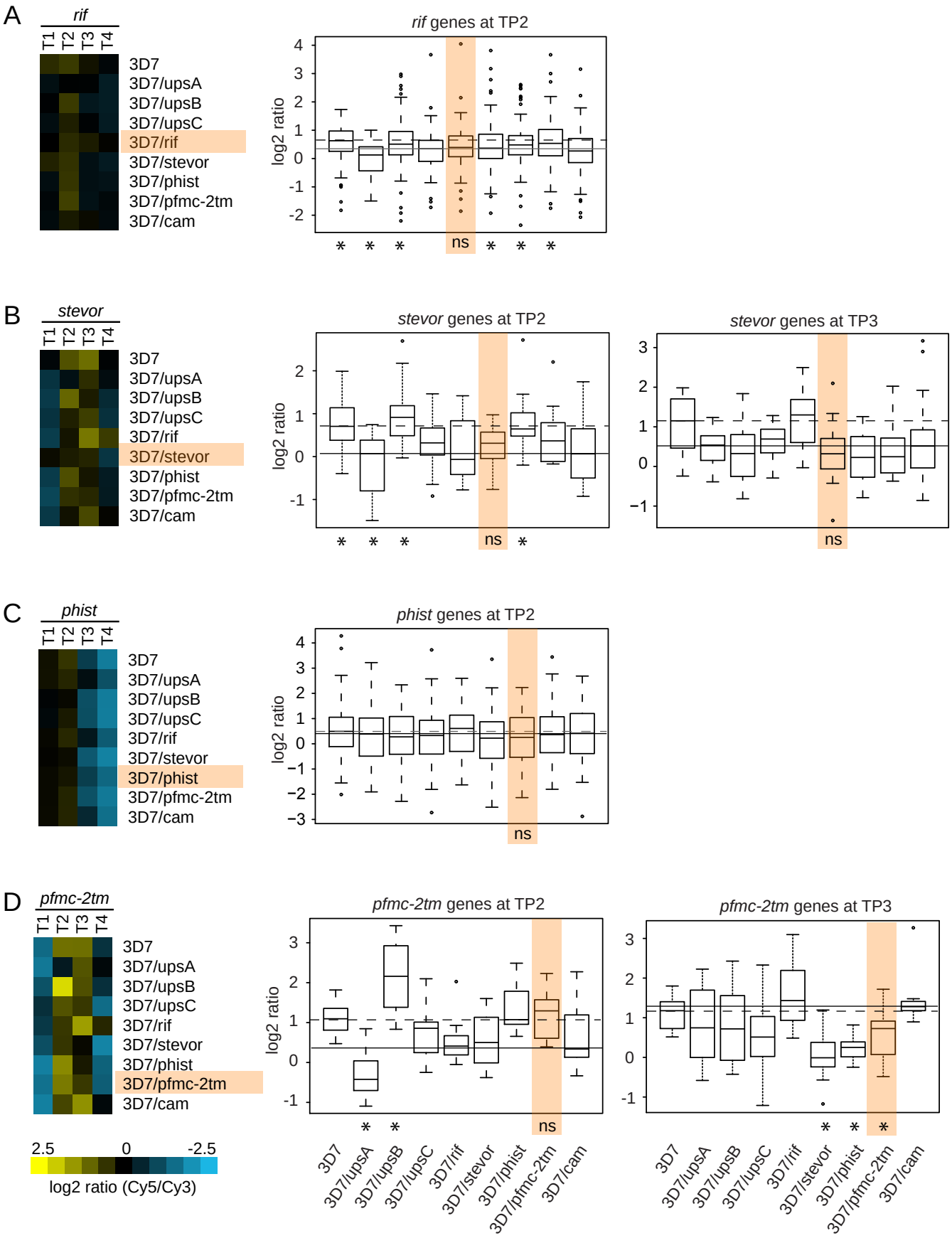
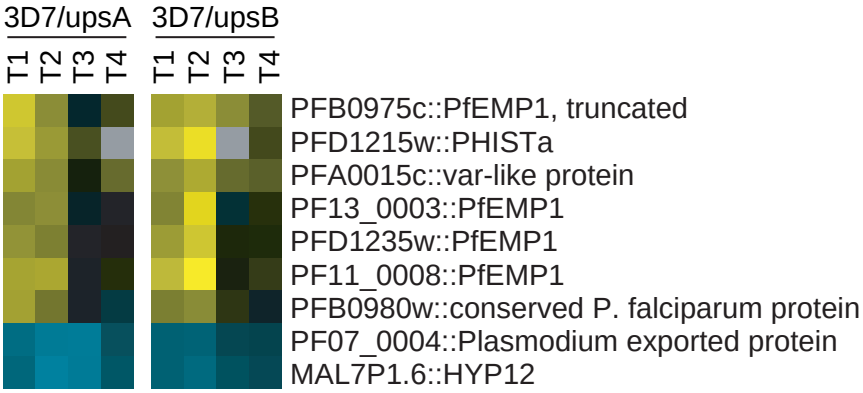

A
rif
T1 T2 T3 T4
3D7
3D7/upsA
3D7/upsB
3D7/upsC
3D7/rif
3D7/stevor
3D7/phist
3D7/pfmc-2tm
3D7/cam
rif genes at TP2
log2 ratio
ns
B
stevor
stevor genes at TP2
stevor genes at TP3
C
phist
phist genes at TP2
D
pfmc-2tm
pfmc-2tm genes at TP2
pfmc-2tm genes at TP3
2.5 0 -2.5
log2 ratio (Cy5/Cy3)

Figure 6

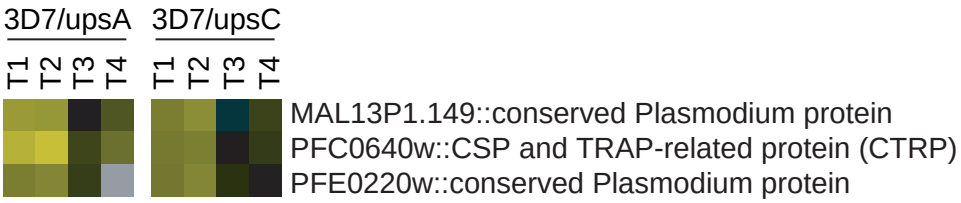
A

differentially expressed in **3D7/upsC** vs 3D7/upsA and 3D7/upsB



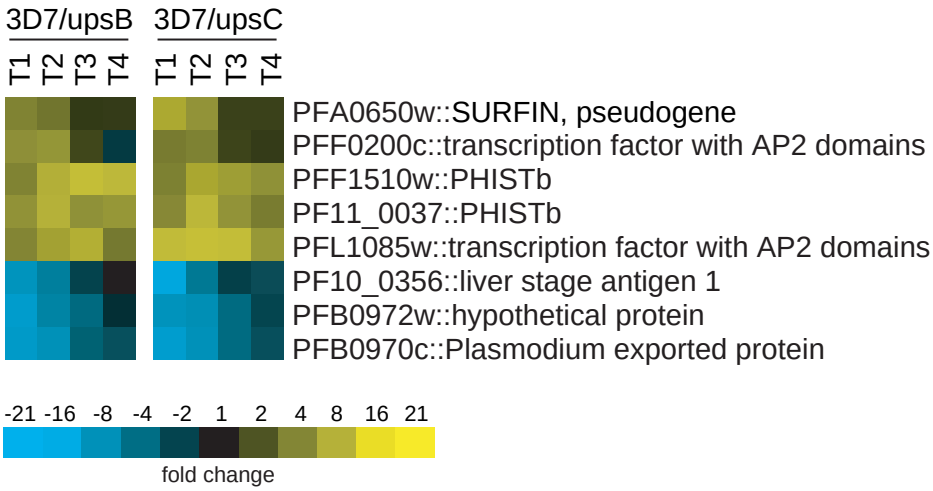
B

differentially expressed in **3D7/upsB** vs 3D7/upsA and 3D7/upsC



C

differentially expressed in **3D7/upsA** vs 3D7/upsB and 3D7/upsC



Supplementary Information

Dataset S1. Averaged Cy5/Cy3 log2 ratios for all genes in all nine cell lines at four timepoints each.

Table S1. List of primers used in this study.

Figure S1. (A) Southern blot of gDNA isolated from drug-selected transgenic cell lines. gDNA was digested with *PvuII*/*Eco*RI. h*dhfr* was used as a probe to visualize episomal fragments or plasmid integration into endogenous loci. The size of the resulting fragments is indicated for each cell line. The second fragment in 3D7/stevor reflects integration of pBstevor into the PFL2610w locus (see D). **(B)** Plasmid map of pBcam and derivatives. **(C)** Schematic map of the PFL2610w *stevor* locus. **(D)** Integration of pBstevor at the stevor locus PFL2610w. The positions of *EcoR*V and *Pvu*II restriction sites and the size of the expected fragments hybridising with the h*dhfr* probe are highlighted.

Figure S2. Heat map showing all genes commonly de-regulated >2-fold in at least one time point in all transgenic cell lines compared to 3D7/wild type parasites. The colour code refers to fold changes in relative expression. The colour scale indicates fold changes in relative transcript abundance ($2^{\log 2 \text{ ratios}}$) between each of the transfected lines compared to 3D7 wild-type parasites. GeneIDs and annotations are according to PlasmoDB version6.3. Coloured boxes group genes into four generic functional clusters.

Figure S3. Correlation between episomal promoter activity and endogenous gene transcription. Grey lines refer to the left axis and represent relative h*dhfr-gfp* mRNA levels produced per single promoter as determined by qRT-PCR. Values are normalised against transcription of arginyl-tRNA Synthetase (PFL0900c) and adjusted for plasmid copy numbers. Black lines refer to the right axis and represent the mean relative expression of the cognate gene family members in 3D7/cam. Time points one to four are plotted on the x-axis. Pearson correlation coefficients are indicated. **(A**) Correlation between

episomal upsA, upsB and upsC *var* promoter activities and endogenous *var* gene transcription. Note that upsA activity is reflected in relative *bsd* transcripts (black diamond). **(B)** Episomal *rif* promoter activity in correlation with endogenous *rif* (top) or *var* transcription (bottom). **(C)** Episomal *stevor, phistb* and *pfmc-2tm* activity in correlation with transcription of the corresponding endogenous gene families. (D) Correlation between relative activities of the episomal and endogenous *cam* promoters.

Figure S4. Relative promoter activities in unselected (default) and drug-selected (activated) populations. Relative reporter transcript levels reflect episomal promoter activities in each transfectant before (black bars) and after (open bars) selection. Note that in 3D7/upsA the activated upsA promoter drives transcription of *bsd* rather than h*dhfr-gfp* (diamond). Transcription of the house-keeping gene arginyl-tRNA synthetase (PFL0900c) was used for normalisation. Values have been adjusted for plasmid copy numbers. Total RNA was harvested three times independently at time points of peak activity. Values represent the mean +/- s.d. hpi, hours post-invasion.

Figure S5. (A) The scatter plots show the correlation of relative expression levels of all genes at TP1 for 3D7/upsB (top), 3D7/pfmc-2tm (center) and 3D7/cam (bottom) in comparison to 3D7 wild-type parasites. **(B)** The scatter plots show the correlation of relative expression levels of all genes at TP1 for 3D7/upsB (top) and 3D7/pfmc-2tm (bottom) in comparison to the 3D7/cam control transfectant. Pearson correlation coefficients are indicated. *var* genes are highlighted in blue, heterochromatic genes [1] are highlighted in red. All other genes are shown in grey. Values on the y- and x-axis represent log2 ratios for each gene in the two cell lines.

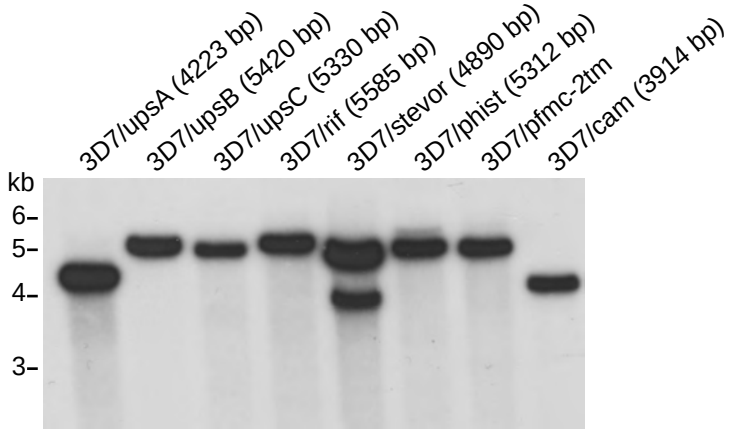
Figure S6. Heat map showing relative expression levels (log2 ratios) of all *hyp* family genes in all cell lines and time points. *hyp4* family members up-regulated in 3D7/pfmc-2tm are highlighted in grey.

References

1. Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, et al. (2009) *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog 5: e1000569.

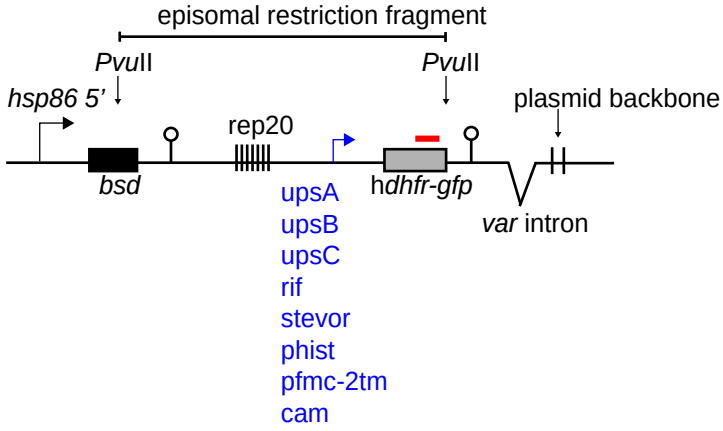
Figure S1

A

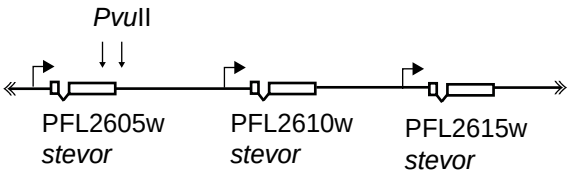


*Pvu*II/*EcoR*I-digested gDNA; probed with *gfp*

B plasmid map of pB_(promoter of interest)



C PFL2610w *stevor* locus



D Integration of pBstevor in the *PFL2610w* locus

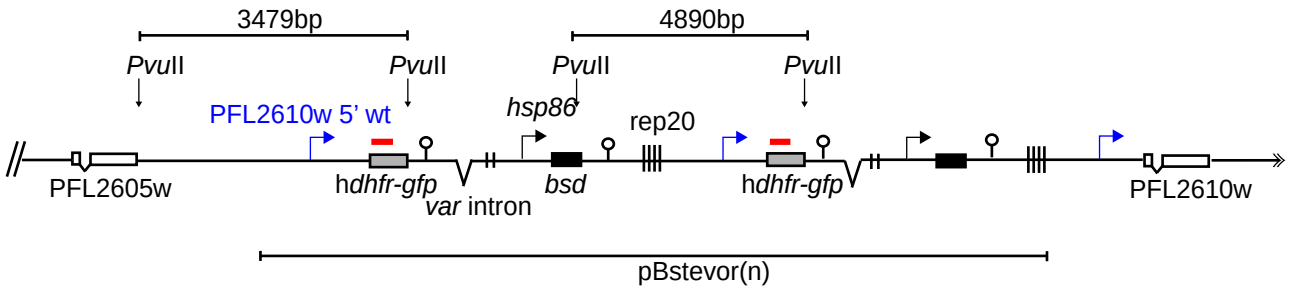


Figure S2

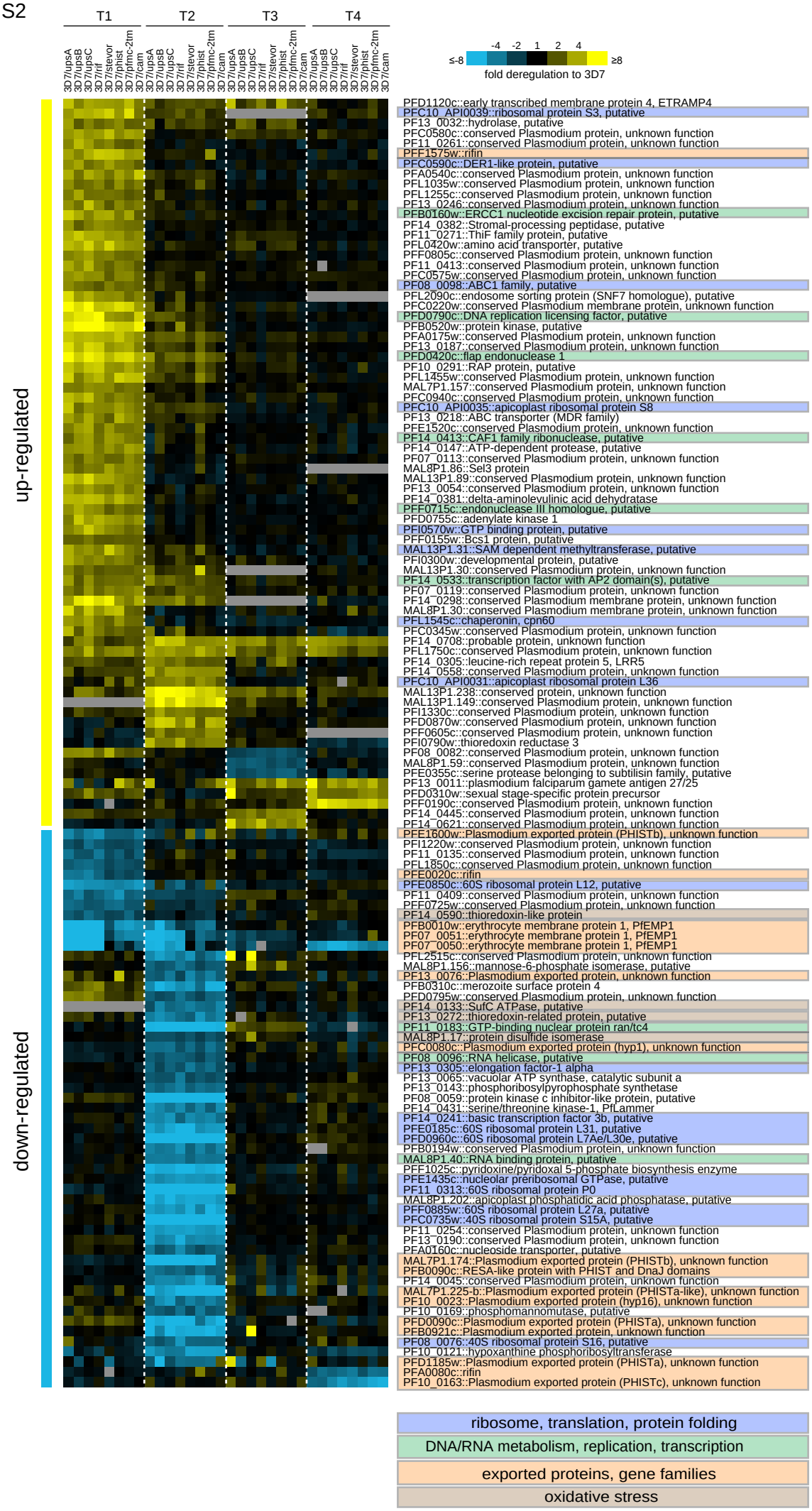

T1
T2
T3
T4
3D7/upsA
3D7/upsB
3D7/upsC
3D7/rif
3D7/stevor
3D7/phist
3D7/pfmc-2tm
3D7/cam
-4 -2 1 2 4
≤-8
≥8
fold deregulation to 3D7
up-regulated
down-regulated
PFD1120c::early transcribed membrane protein 4, ETRAMP4
PFC10_API0039::ribosomal protein S3, putative
PF13_0032::hydrolase, putative
PFC0580c::conserved Plasmodium protein, unknown function
PF11_0261::conserved Plasmodium protein, unknown function
PFF1575w::rifin
PFC0590c::DER1-like protein, putative
PFA0540c::conserved Plasmodium protein, unknown function
PFL1035w::conserved Plasmodium protein, unknown function
PFL1255c::conserved Plasmodium protein, unknown function
PF13_0246::conserved Plasmodium protein, unknown function
PFB0160w::ERCC1 nucleotide excision repair protein, putative
PF14_0382::Stromal-processing peptidase, putative
PF11_0271::ThiF family protein, putative
PFL0420w::amino acid transporter, putative
PFF0805c::conserved Plasmodium protein, unknown function
PF11_0413::conserved Plasmodium protein, unknown function
PFC0575w::conserved Plasmodium protein, unknown function
PF08_0098::ABC1 family, putative
PFL2090c::endosome sorting protein (SNF7 homologue), putative
PFC0220w::conserved Plasmodium membrane protein, unknown function
PFD0790c::DNA replication licensing factor, putative
PFB0520w::protein kinase, putative
PFA0175w::conserved Plasmodium protein, unknown function
PF13_0187::conserved Plasmodium protein, unknown function
PFD0420c::flap endonuclease 1
PF10_0291::RAP protein, putative
PFL1455w::conserved Plasmodium protein, unknown function
MAL7P1.157::conserved Plasmodium protein, unknown function
PFC0940c::conserved Plasmodium protein, unknown function
PFC10_API0035::apicoplast ribosomal protein S8
PF13_0218::ABC transporter (MDR family)
PFE1520c::conserved Plasmodium protein, unknown function
PF14_0413::CAF1 family ribonuclease, putative
PF14_0147::ATP-dependent protease, putative
PF07_0113::conserved Plasmodium protein, unknown function
MAL8P1.86::Sel3 protein
MAL13P1.89::conserved Plasmodium protein, unknown function
PF13_0054::conserved Plasmodium protein, unknown function
PF14_0381::delta-aminolevulinic acid dehydratase
PFF0715c::endonuclease III homologue, putative
PFD0755c::adenylate kinase 1
PFI0570w::GTP binding protein, putative
PFF0155w::Bcs1 protein, putative
MAL13P1.31::SAM dependent methyltransferase, putative
PFI0300w::developmental protein, putative
MAL13P1.30::conserved Plasmodium protein, unknown function
PF14_0533::transcription factor with AP2 domain(s), putative
PF07_0119::conserved Plasmodium protein, unknown function
PF14_0298::conserved Plasmodium membrane protein, unknown function
MAL8P1.30::conserved Plasmodium membrane protein, unknown function
PFL1545c::chaperonin, cpn60
PFC0345w::conserved Plasmodium protein, unknown function
PF14_0708::probable protein, unknown function
PFL1750c::conserved Plasmodium protein, unknown function
PF14_0305::leucine-rich repeat protein 5, LRR5
PF14_0558::conserved Plasmodium protein, unknown function
PFC10_API0031::apicoplast ribosomal protein L36
MAL13P1.238::conserved protein, unknown function
MAL13P1.149::conserved Plasmodium protein, unknown function
PFI1330c::conserved Plasmodium protein, unknown function
PFD0870w::conserved Plasmodium protein, unknown function
PFF0605c::conserved Plasmodium protein, unknown function
PFI0790w::thioredoxin reductase 3
PF08_0082::conserved Plasmodium protein, unknown function
MAL8P1.59::conserved Plasmodium protein, unknown function
PFE0355c::serine protease belonging to subtilisin family, putative
PF13_0011::plasmodium falciparum gamete antigen 27/25
PFD0310w::sexual stage-specific protein precursor
PFF0190c::conserved Plasmodium protein, unknown function
PF14_0445::conserved Plasmodium protein, unknown function
PF14_0621::conserved Plasmodium protein, unknown function
PFE1600w::Plasmodium exported protein (PHISTb), unknown function
PFI1220w::conserved Plasmodium protein, unknown function
PF11_0135::conserved Plasmodium protein, unknown function
PFL1850c::conserved Plasmodium protein, unknown function
PFE0020c::rifin
PFE0850c::60S ribosomal protein L12, putative
PF11_0409::conserved Plasmodium protein, unknown function
PFF0725w::conserved Plasmodium protein, unknown function
PF14_0590::thioredoxin-like protein
PFB0010w::erythrocyte membrane protein 1, PfEMP1
PF07_0051::erythrocyte membrane protein 1, PfEMP1
PF07_0050::erythrocyte membrane protein 1, PfEMP1
PFL2515c::conserved Plasmodium protein, unknown function
MAL8P1.156::mannose-6-phosphate isomerase, putative
PF13_0076::Plasmodium exported protein, unknown function
PFB0310c::merozoite surface protein 4
PFD0795w::conserved Plasmodium protein, unknown function
PF14_0133::SufC ATPase, putative
PF13_0272::thioredoxin-related protein, putative
PF11_0183::GTP-binding nuclear protein ran/tc4
MAL8P1.17::protein disulfide isomerase
PFC0080c::Plasmodium exported protein (hyp1), unknown function
PF08_0096::RNA helicase, putative
PF13_0305::elongation factor-1 alpha
PF13_0065::vacuolar ATP synthase, catalytic subunit a
PF13_0143::phosphoribosylpyrophosphate synthetase
PF08_0059::protein kinase c inhibitor-like protein, putative
PF14_0431::serine/threonine kinase-1, PfLammer
PF14_0241::basic transcription factor 3b, putative
PFE0185c::60S ribosomal protein L31, putative
PFD0960c::60S ribosomal protein L7Ae/L30e, putative
PFB0194w::conserved Plasmodium protein, unknown function
MAL8P1.40::RNA binding protein, putative
PFF1025c::pyridoxine/pyridoxal 5-phosphate biosynthesis enzyme
PFE1435c::nucleolar preribosomal GTPase, putative
PF11_0313::60S ribosomal protein P0
MAL8P1.202::apicoplast phosphatidic acid phosphatase, putative
PFF0885w::60S ribosomal protein L27a, putative
PFC0735w::40S ribosomal protein S15A, putative
PF11_0254::conserved Plasmodium protein, unknown function
PF13_0190::conserved Plasmodium protein, unknown function
PFA0160c::nucleoside transporter, putative
MAL7P1.174::Plasmodium exported protein (PHISTb), unknown function
PFB0090c::RESA-like protein with PHIST and DnaJ domains
PF14_0045::conserved Plasmodium protein, unknown function
MAL7P1.225-b::Plasmodium exported protein (PHISTa-like), unknown function
PF10_0023::Plasmodium exported protein (hyp16), unknown function
PF10_0169::phosphomannomutase, putative
PFD0090c::Plasmodium exported protein (PHISTa), unknown function
PFB0921c::Plasmodium exported protein, unknown function
PF08_0076::40S ribosomal protein S16, putative
PF10_0121::hypoxanthine phosphoribosyltransferase
PFD1185w::Plasmodium exported protein (PHISTa), unknown function
PFA0080c::rifin
PF10_0163::Plasmodium exported protein (PHISTc), unknown function
ribosome, translation, protein folding
DNA/RNA metabolism, replication, transcription
exported proteins, gene families
oxidative stress

Figure S3

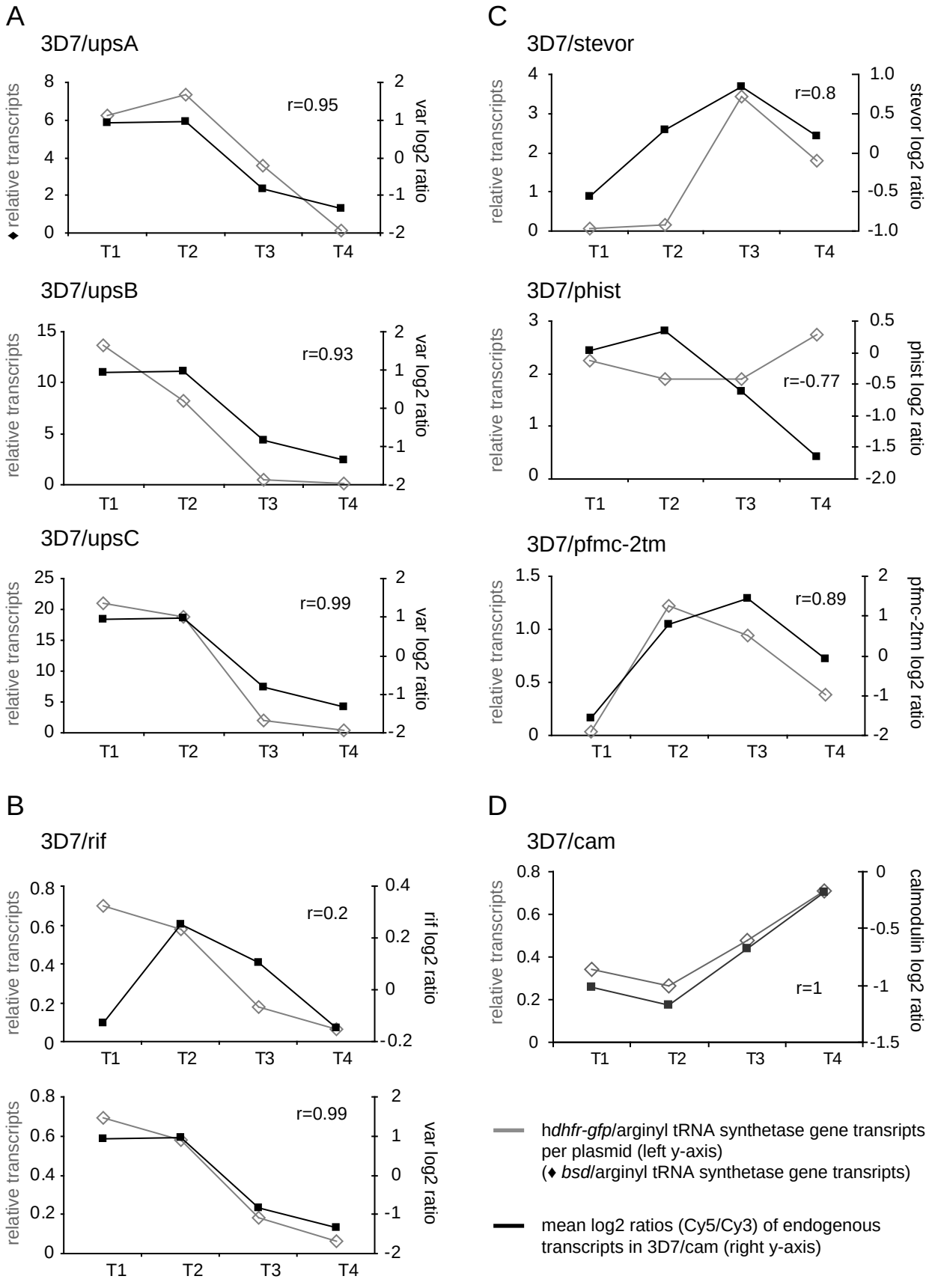

A
3D7/upsA
relative transcripts
var log2 ratio
r=0.95
T1
T2
T3
T4
3D7/upsB
r=0.93
3D7/upsC
r=0.99
C
3D7/stevor
r=0.8
stevor log2 ratio
3D7/phist
r=-0.77
phist log2 ratio
3D7/pfmc-2tm
r=0.89
pfmc-2tm log2 ratio
B
3D7/rif
r=0.2
rif log2 ratio
r=0.99
D
3D7/cam
r=1
calmodulin log2 ratio
h*dhfr-gfp*/arginyl tRNA synthetase gene transripts per plasmid (left y-axis)
(♦ *bsd*/arginyl tRNA synthetase gene transripts)
mean log2 ratios (Cy5/Cy3) of endogenous transcripts in 3D7/cam (right y-axis)

Figure S4

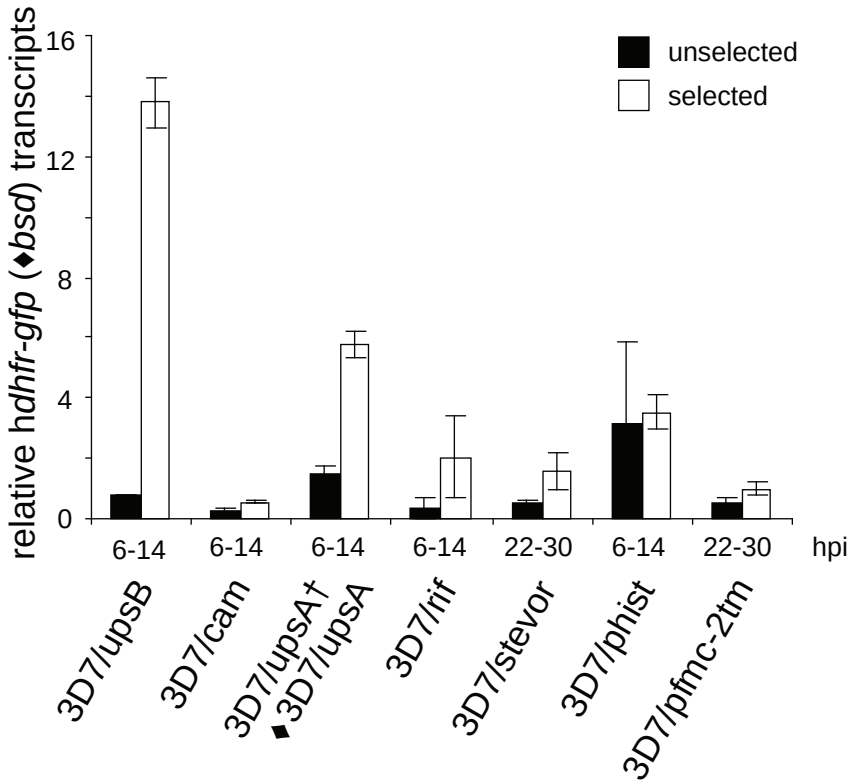

relative hdhfr-gfp (♦bsd) transcripts
16
12
8
4
0
unselected
selected
6-14
6-14
6-14
6-14
22-30
6-14
22-30
hpi
3D7/upsB
3D7/cam
3D7/upsA† ▮3D7/upsA
3D7/rif
3D7/stevor
3D7/phist
3D7/pfmc-2tm

Figure S5

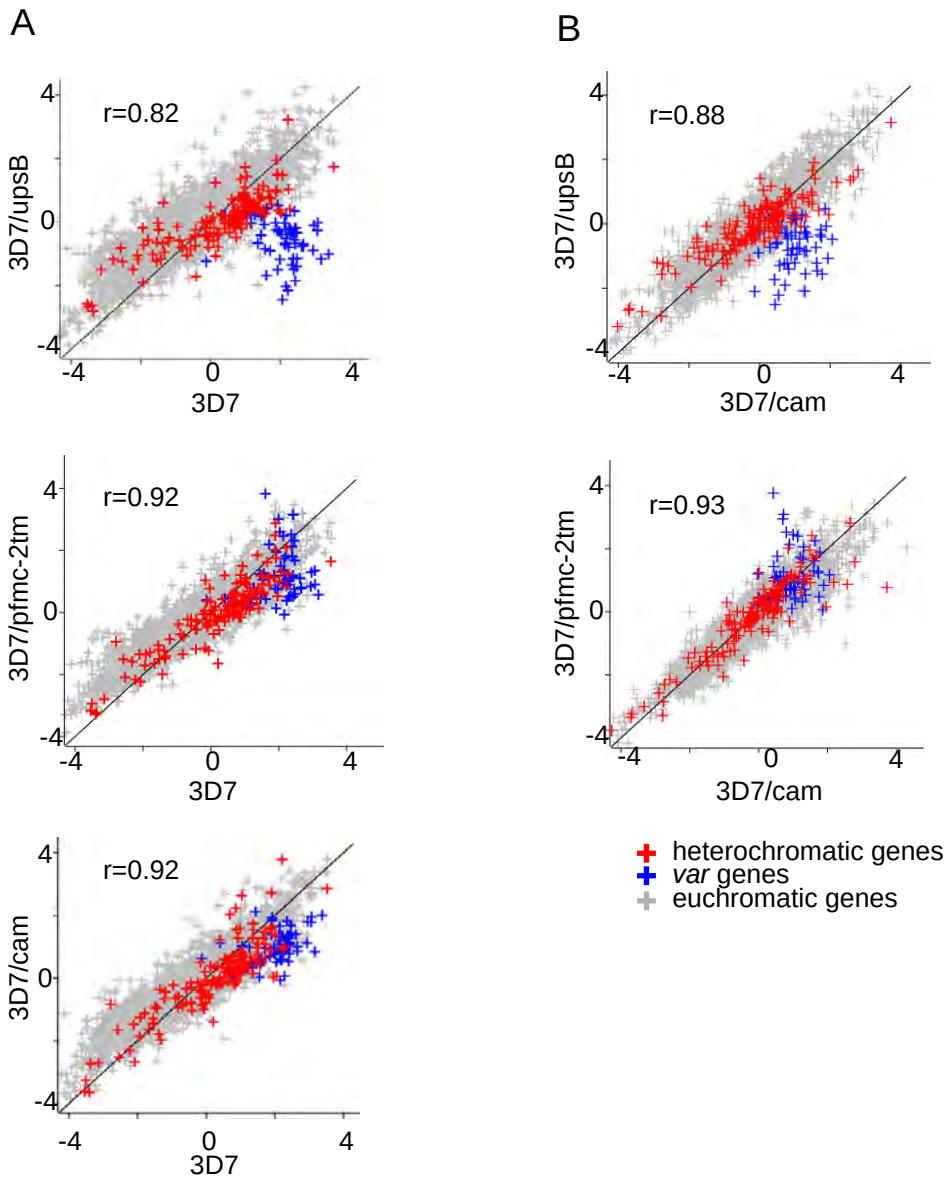
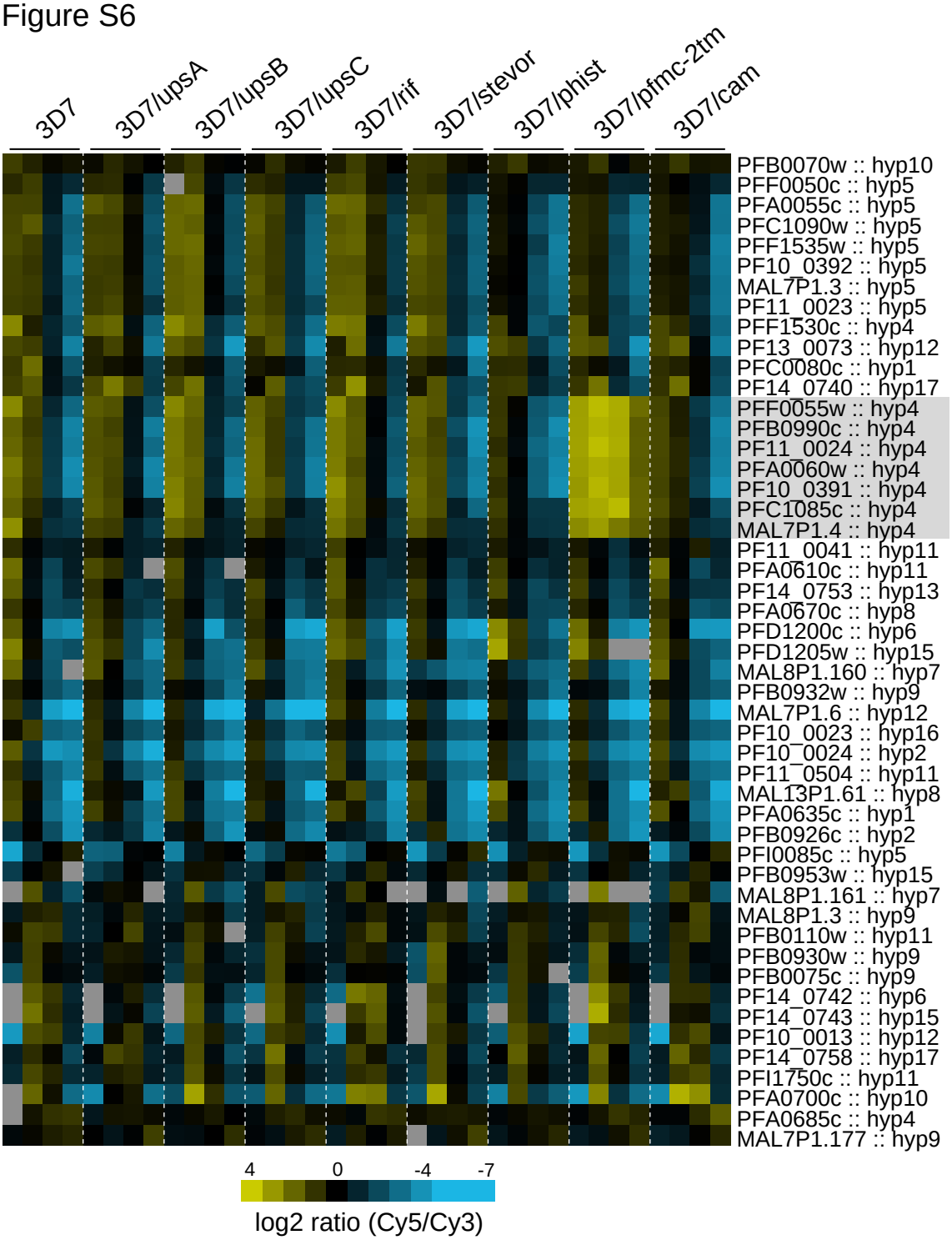

A
B
r=0.82
r=0.88
r=0.92
r=0.93
r=0.92
3D7/upsB
3D7/pfmc-2tm
3D7/cam
3D7
3D7/cam
heterochromatic genes
var genes
euchromatic genes

Figure S6


3D7
3D7/upsA
3D7/upsB
3D7/upsC
3D7/rif
3D7/stevor
3D7/phist
3D7/pfmc-2tm
3D7/cam
PFB0070w :: hyp10
PFF0050c :: hyp5
PFA0055c :: hyp5
PFC1090w :: hyp5
PFF1535w :: hyp5
PF10_0392 :: hyp5
MAL7P1.3 :: hyp5
PF11_0023 :: hyp5
PFF1530c :: hyp4
PF13_0073 :: hyp12
PFC0080c :: hyp1
PF14_0740 :: hyp17
PFF0055w :: hyp4
PFB0990c :: hyp4
PF11_0024 :: hyp4
PFA0060w :: hyp4
PF10_0391 :: hyp4
PFC1085c :: hyp4
MAL7P1.4 :: hyp4
PF11_0041 :: hyp11
PFA0610c :: hyp11
PF14_0753 :: hyp13
PFA0670c :: hyp8
PFD1200c :: hyp6
PFD1205w :: hyp15
MAL8P1.160 :: hyp7
PFB0932w :: hyp9
MAL7P1.6 :: hyp12
PF10_0023 :: hyp16
PF10_0024 :: hyp2
PF11_0504 :: hyp11
MAL13P1.61 :: hyp8
PFA0635c :: hyp1
PFB0926c :: hyp2
PFI0085c :: hyp5
PFB0953w :: hyp15
MAL8P1.161 :: hyp7
MAL8P1.3 :: hyp9
PFB0110w :: hyp11
PFB0930w :: hyp9
PFB0075c :: hyp9
PF14_0742 :: hyp6
PF14_0743 :: hyp15
PF10_0013 :: hyp12
PF14_0758 :: hyp17
PFI1750c :: hyp11
PFA0700c :: hyp10
PFA0685c :: hyp4
MAL7P1.177 :: hyp9
4
0
-4
-7
log2 ratio (Cy5/Cy3)

Supplementary table 1

Primers used for cloning	Gene ID	primer sequence forward	restriction sites	length of fragment
dhfr-Fwd-B-N		agct**ggatccgcggccgc**aaaacaatgcatggttcgctaaactg	BamHI, NotI	
dhfr-Rev-S		agcta**gtcgacgcagc**atcattcttctcatatacttcaa	SalI	
gfp-Fwd-S		agct**gtcgacgctgct**atgagtaaaggagaagaactT	SalI	
gfp-Rev-A-H		agcta**agcttcctagg**ttaactagtaccggtttgta	AvrII, HindIII	
intron-F-M-S		catg**caattg**aaggccttatgtatttatattgaaaaag	MfeI, StuI	
intron-R-M-S		catg**caattg**aaggccttgtgtatatgtatgtgtatg	MfeI, StuI	
upsA_Fwd	PF13_0003	cagt**agatct**tttttctttttttattgtgatacg	BglII	2742 bp
upsA_Rev		cagt**gcggccgctt**taaaaacaatattatttcttcatgg	NotI	
upsB_Fwd	PFL0005w	cagt**agatct**cctataggtacgaatatagg	BglII	2576 bp
upsB_Rev		cagt**gcggccgc**gtcatacatatatttggagc	NotI	
upsC_Fwd	PFL1960w	cagt**agatct**ttatgttggtacattatacatg	BglII	2487 bp
upsC_Rev		cagt**gcggccgc**ttttgttttttgtttatcgttcg	NotI	
rif_Fwd	PF13_0004	cagt**agatct**ttttaaaaacaatattatttcttcatgg	BglII	2742 bp
rif_Rev		cagt**gcggccgc**tttttcttttttattgtgatacg	NotI	
stevor_Fwd	PFL2610w	tacg**agatct**attaaatcaat**c**aatatatggttcc	BglII	2047 bp
stevor_Rev		cagt**gcggccgc**taattcaaaaaaaaataatgttatatttc	NotI	
phist_Fwd	PFL2540w	cagt**ggatcc**acttatatttttatgaattgtagg	BamHI	2469 bp
phist_Rev		cagt**gcggccgc**gatttatacatcgtatagatacg	NotI	
pfmc-2tm_Fwd		tcga**agatct**gaataaaagaatactatacattaacag	BglII	~2500bp
pfmc-2tm_Rev		cagt**gcggccgc**gtgtattctaatatatagatgtg	NotI	
Bsd_F_BglII		cagt**agatct**atggcaccttgtctcaag	BglII	
PbDT_R_KpnI		cagt**ggtacc**ggccgcctaccctgaag	KpnI	
Primers used for qPCR				
arginyl-tRNA synthetase_Fwd	PFL0900c	aagagatgcatgttggtcattt		117
arginyl-tRNA synthetase_Rev		gagtacccaatcacctaca		
bsd_Fwd		acagcgtcgccagcgcagctctcta		151
bsd_Rev		atcgcgacgatacaagtcaggttgccagct		
gfp_Fwd		acacttgtcactacttcgctatggtcttc		174
gfp_Rev		accttcaaacttgacttcagcacgtgtcttgtagt		

Testing for the existence of a singular *var* gene enhancer element using 3C (chromosome conformation capture)

Introduction

Investigations into the structure of the *P. falciparum* nucleus suggest the existence of a region of uncondensed euchromatin in the largely heterochromatic nuclear periphery (Ralph et al., 2005). This sub-nuclear region has been proposed to function as a specific expression site for *var* genes. The *var* gene family encompasses ~60 members per haploid genome and can be subdivided into three major groups (upsA, upsB and upsC) according to their upstream region (Gardner et al, 2002; Lavstsen et al, 2003). *var* genes are transcribed in a mutually exclusive manner where only one member is expressed whereas all others remain silenced (Scherf et al, 1998). Switching in *var* gene expression leads to antigenic variation of PfEMP1 (*P. falciparum* erythrocyte membrane protein 1), the main virulence factor of *P. falciparum*-infected red blood cells (Scherf et al, 1998). A model for mutually exclusive expression involves that the proposed expression site can activate only one *var* locus at a time, similar to the transcriptional body mediating mutually exclusive expression of *vsg* genes in African trypanosomes (Navarro & Gull, 2001). Furthermore, recent publications reported that *var* genes move to a different position in the nucleus upon transcriptional activation (Duraisingh et al, 2005; Ralph et al, 2005; Marty et al, 2006). The *var* upstream region plays a pivotal role in *var* gene silencing, activation and mutual exclusion (Voss et al, 2006, 2007; Dzikowski et al, 2006). An active episomal *var* gene promoter expressing a reporter gene is recruited to chromosome end clusters where it interferes with endogenous *var* gene transcription, causing a knock-down of the entire *var* gene family (Voss et al, 2006). However, to date the factor(s) controlling this sophisticated mode of expression have not been identified.

Similar to *var* genes, the odorant receptor gene family in mammals (comprising several hundred genes) is expressed in a mutually exclusive manner (Rodriguez, 2007; Niimura & Nei, 2003). Every olfactory sensory neuron expresses a single olfactory receptor (OR) gene resulting in hundreds of functionally different neuronal populations in the nasal cavity. Recent observations in the mouse argue for the existence of multiple levels of regulation. Using the chromosome conformation capture (3C) technique (Dekker et al, 2002), Lomvardas et al. show that a single enhancer (H) in mice is able to associate with OR gene promoters on different chromosomes but with only one OR gene (the active one) in any given olfactory neuron (Lomvardas et al, 2006). However, in homozygous H knock-out mice, the expression of only three OR family members that are located *in cis* downstream of the H enhancer is abolished. Thus, the authors could not corroborate the hypothesis of H as an essential *trans*-acting enhancer (Fuss et al, 2007). Nevertheless, based on the parallels between the regulation of odorant receptor genes in mammals and *var* genes in *P. falciparum* it is tempting to speculate that a unique DNA enhancer sequence may be employed by the parasite to control mutually exclusive *var* gene expression by interaction with only one *var* promoter at a time.

I therefore proposed to test for the existence of a singular enhancer element mediating the sequestration of the active *var* promoter to a *var*-specific transcriptional factory at the nuclear periphery using a modified 3C approach. Shortly, intact parasite nuclei carrying an active transgenic *var* promoter are isolated and subjected to formaldehyde cross-linking. Cross-linked DNA is digested to completion with a restriction enzyme frequently cutting in the *P. falciparum* genome and releasing the entire transgenic *var* gene promoter. DNA fragments are subsequently ligated at decreasing DNA concentrations to favor intra-molecular end-joining. Cross-linking is then reversed, DNA purified and individual ligation products are detected and quantified by PCR using locus-specific primers. The PCR products are subsequently cloned into a plasmid vector and sequenced. Figure 1 shows the principle of the 3C technique employed to identify the putative *var* gene enhancer element.

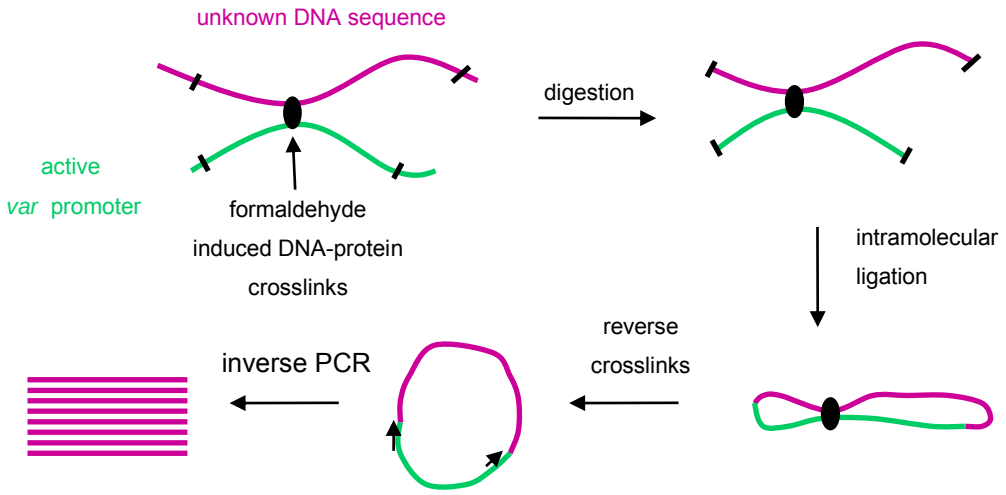


Figure 1. Schematic representation of the modified 3C strategy proposed here (picture taken and modified from Dekker et al., 2002)

Material and Methods

Generation of a transfection construct tailored for 3C

To achieve plasmid integration into the *rh3* locus (*reticulocyte-binding like protein homologue 3*; PFL2520w) via double crossover, pCC4RH3 was used as the parental plasmid. pCC4RH3 is a derivative of pCC4 (Maier et al, 2006) and consists of two 900bp targeting fragments of the *rh3* coding sequence, the negative selection cassette FCU (a gene fusion of the *S. cerevisiae cytosine deaminase* and *uracil phosphoribosyl transferase*) and a blasticidin-S-deaminase *bsd* resistance cassette (T. Voss, unpublished) (Fig. 2A). pCC4RH3 was digested with *Spe*I and *Eco*RI, releasing the entire *bsd* resistance cassette including the *calmodulin* (*cam*) promoter and the *hrp2* (*histidine rich protein 2*) terminator. A DNA linker (linker1) with restriction sites for *Nhe*I-*Mlu*I-*Eco*RI was inserted, eliminating the previous *Spe*I and *Eco*RI sites, to produce pCC4RH3-linker1 (Fig. 2B).

Plasmid pBcamHG (Brancucci and Witmer et al., submitted) was digested with *Pst*I and *Not*I, removing the TARE6 (rep20) sequence and the *cam* promoter. To construct the plasmid pB-linker2 (Fig. 2C), a DNA linker containing ten additional restriction sites (*Nhe*I-*Not*I-*Asc*I-*Ngo*MIV-*Bgl*II-*Mbo*I-

*Pst*I-*Bam*HI-*Mbo*I-*Not*I) was inserted into *Pst*I/*Not*I-digested pBcamHG, eliminating the previous *Pst*I site.

2482 bp of an upsC *var* gene upstream sequence (PFL1960w) was amplified with primers pstI_T7_upsC_forw and bglII_SP6_upsC_rev resulting in a *var* promoter flanked by T7 and SP6 promoter sequences (Table 1). This PCR product was digested with *Pst*I and *Bgl*II and cloned into *Pst*I/*Bam*HI digested pB-linker2, generating plasmid pBupsC(T7_SP6)HG (Fig. 2D). From this plasmid, the *var* upsC sequence driving expression of the h*dhfr-gfp* fusion gene and the *hrp2* 3' region was released by digestion with *Nhe*I and *Eco*RI and then ligated into the *Nhe*I/*Eco*RI digested plasmid pCC4RH3-linker1, generating the final plasmid pCC4RH3-43C, henceforth named pkrass (Fig 2E).

Parasite culture and transfection

P. falciparum 3D7 parasites were cultured as described previously (Trager & Jenson, 1978). Growth synchronisation was achieved by repeated sorbitol lysis (Lambros & Vanderberg, 1979). Transfections were performed as described (Voss et al, 2006). Transfected parasites were selected on 2nM WR99210 (WR) (a kind gift of Jacobus Pharmaceutical Co., Inc.). To select for integration of the plasmid into the *rh3* locus, transfected parasites were subjected to repeated cycles off and on WR as previously reported (Maier et al, 2006). Drug cycling was followed by negative selection against the FCU cassette with 5nM 5-FC (Ancotil® , Valeant Pharmaceuticals, Switzerland) (Maier et al, 2006).

Southern blotting

In order to test for successful integration of pkrass into the *rh3* locus by double crossover recombination, genomic DNA (gDNA) was isolated from transgenic 3D7/pkrass parasites and subjected to Southern blotting using standard procedures. gDNA was digested with *Pvu*II, *Spe*I/*Sph*I or *Cla*I, and separated on 0.7% agarose gels. Southern blots were probed with random-primed ^{32}P-dATP-labelled h*dhfr*, *rh3* 3' and/or *rh3* 5' coding sequences.

Results and Discussion

Identification of an unknown sequence requires a modified 3C approach

The 3C (chromosome conformation capture) technology was originally used to study the spatial organization of the yeast chromosome III (Dekker et al, 2002). Since then, it has been applied to demonstrate long-range looping interactions between genes and regulatory sequences in several mammalian loci (Tolhuis et al, 2002; Palstra et al, 2003). 3C was also successfully used to demonstrate *trans*-chromosomal interactions in mice (Spilianakis et al, 2005). The main difference of the approach chosen here compared to previous studies employing the 3C technique is that we aimed at identifying an unknown unique sequence, rather than demonstrating the suspected interaction of two known loci. The working hypothesis is that the interaction of an active *var* promoter with a unique sequence promotes mutual exclusive *var* gene expression. Hence, both primer recognition sequences used for amplification of an unknown sequence must be positioned within the known sequence, in this case the upsC *var* promoter. Complications arise from the fact that *var* gene promoters occur multiple times in the genome and are conserved by sequence which makes the designing of specific primer pairs unfeasible. This required adjustment or tailoring of the *var* upsC promoter sequence by generating a transgenic cell line carrying specific sequences attached up- and downstream of the active *var* promoter.

With the planned 3C experiment, we rely on a system where an active *var* gene promoter does not allow for switching to another *var* gene. Unselected *P. falciparum* cultures do not fulfil this requirement due to continuous switching such that almost all out of 60 *var* genes are found to be expressed (Salanti et al, 2003) However, parasite populations can be selected to homogenously express a single *var* gene (*var2csa*). This is achieved by repeated rounds of panning parasites on chondroitin sulphate A (CSA), a host receptor bound specifically by one of the 60 PfEMP1 variants, VAR2CSA (Salanti et al, 2003). These parasites switch away from the active *var* gene to other previously silenced variants (Scherf et al, 1998). Therefore, repeated selection rounds are needed to maintain transcription of var2csa. To circumvent this time-consuming and tedious selection procedure and the

abovementioned difficulties to distinguish between individual promoters due to sequence conservation I chose to take a transgenic approach.

Transfection with a plasmid containing an upsC *var* promoter driving expression of a drug-selectable reporter is well established (Voss et al, 2006; Dzikowski et al, 2006; Frank et al, 2006). In these systems the episomal *var* promoter can be kept in its active state by constant drug selection. In addition, the promoter on the reporter construct is competent to infiltrate the mutually exclusive *var* gene expression mechanism (Voss et al, 2006). In other words, active episomal var promoters displace the active endogenous *var* locus resulting in a knock-down endogenous *var* gene transcription. This shows that an episomal *var* gene promoter harbours the same information as a chromosomal *var* gene promoter and hence is a suitable bait to perform 3C. However, to ensure the episomal *var* promoter used is distinguishable from endogenous upsC *var* promoters, bacteriophage T7 and SP6 primer sequences on the 5' and 3' end of the upsC promoter were included (Fig 2).

Upon transfection, parasites create so-called concatamers, consisting of multiple plasmid copies arranged tail-to-head (Williamson et al, 2002; O'Donnell et al, 2002). In such a concatamer, multiple *var* promoters can be active in *cis* (Voss et al, 2007). However, it remains uncertain whether this phenomenon is related to the absence of boundary elements on the plasmids that naturally block activation of *cis*-linked *var* genes or if it is due to a non-natural chromatin environment. In this background, the hypothesis of a unique DNA sequence mediating mutually exclusive *var* expression is difficult to test. The simplest explanation is that one *var* promoter is linked with the unique DNA sequence and other promoters on the concatamer are active due to the absence of boundary elements. Therefore, the *var* promoters on the concatamer that are unlinked to the unique DNA sequence simply outnumber the single *var* promoter linked to this unique sequence. Hence, 3C experiments with parasites harboring concatamers appear to provide suboptimal conditions. As a consequence, we chose the strategy of integrating a single copy of a tailored upsC promoter cassette into a chromosomal locus.

Design of a tailored plasmid construct suitable for 3C

The plasmid pkrass contains a 2472 bp *var* upsC 5' sequence (PFL1960w) driving expression of the drug-selectable reporter h*dhfr-gfp*, that confers resistance to the antifolate drug WR. The upsC promoter in pkrass is flanked at its 5' and 3' end with a T7 and SP6 primer sequence, respectively. T7 and SP6 sequences are of bacteriophage origin and are thus uniquely occurring sequences in parasites transfected with pkrass. Primers to these sequences can be used to specifically amplify any sequences ligated to the reporter promoter in a 3C experiment.

The 3C approach chosen here relies on a restriction enzyme digest to create DNA fragments which can then be ligated in order to identify any sequence bridged to the *var* promoter, such as the proposed putative unique DNA sequence that mediates mutually exclusive *var* gene expression. Furthermore, the restriction fragments need to be in a size range suitable for ligation and subsequent amplification. The four-cutter enzyme *Mbo*I was chosen because (1) it cuts the *P. falciparum* genome frequently into fragments of about 500-1000 base pairs on average; and (2) it does not act in the upsC promoter sequence chosen for this experiment. Hence, the *var* upstream sequence was engineered to be flanked by *Mbo*I sites (GATC) directly upstream and downstream of the Sp6 and T7 sites, respectively. Upon digestion, the 2472 bp *var* 5' sequence including the flanking T7 and SP6 promoter sequences is released, and intra-molecular ligation to DNA sequences that are attached to the upsC promoter via DNA-protein interactions is possible (see Figure 1).

The *rh3* (reticulocyte binding like protein homologue 3, PFL2520w) locus was chosen for integration of the upsC reporter construct. The *rh3* gene, located on chromosome 12, is a pseudogene and has been established as non-essential by knock-out experiments using both double and single crossover recombination strategies (Duraisingh et al, 2002; Taylor et al, 2001). We flanked the upsC promoter reporter cassette with *rh3* 5' and 3' coding sequences to facilitate targeted homologous recombination into the *rh3* locus by double cross-over. Further, the *S. cerevisiae* fusion gene *fcu* is included on the plasmid and used as a negative selection marker against the plasmid

backbone (Erbs et al, 2000; Maier et al, 2006). An overview of the *rh3* locus, the plasmid pkrass and possible crossover events is given in Figure 3.

Attempts to integrate pkrass into the endogenous *rh3* locus by double crossover recombination

The plasmid pkrass was successfully transfected into 3D7 parasites, resulting in the transgenic cell line 3D7/pkrass. Previous reports revealed that the selection for a double crossover event in *P. falciparum* can be achieved upon one to three cycles of off-on drug selection, followed by negative selection against the plasmid backbone (Maier et al, 2006, 2008). Episomal plasmids are lost rapidly in the absence of drug selection as they are distributed unevenly between daughter cells whereas integrated plasmids are inherently segregated with the chromosomes (O'Donnell et al, 2002). Passive selection over time by on-off drug cycles entail that eventually all transfected parasites obtained carry plasmid that are integrated into the genome by homologous recombination. This useful method for gene disruption is limited due to the persistence of stable episomal forms. To overcome these limitations, negative selection of transfected parasites was developed using either the *thymidine kinase* gene of the Herpes simplex virus or the *cytosine deaminase* (*cd*) gene of *E. coli* (Duraisingh et al, 2002; Mullen et al, 1992; Moolten, 1986). The pro-drug 5-fluorocytosine (5-FC) is converted by cytosine deaminase and eventually inhibits thymidylate synthase and hence DNA synthesis. A fusion gene of *S. cerevisiae cd* and *uracil phosphoribosyl transferase* (*uprt*) has been shown to be 1000-fold more potent than *cd* itself (Erbs et al, 2000). This chimeric enzyme (termed FCU) has been utilized for successful negative selection in *P. falciparum* (Maier et al, 2006, 2008).

In our study, populations of stable 3D7/pkrass transfectants were cultured for one month off drug allowing the parasites to shed episomes, followed by drug selection (WR) until resistant parasites grew normally. These steps were repeated twice. Different selection protocols were applied: 3D7/pkrass parasites were subjected to negative selection with 5-FC after the first, second or third drug cycle, respectively. Either way, the selection procedure was terminated by the addition of 5-FC. gDNA from parasites that had undergone different drug cycling schemes was subjected to Southern blot

analysis to monitor the integration of pkrass at the *rh3* locus. gDNA harvested after the first WR cycle followed by 5-FC challenge was digested with *Pvu*II and the Southern blot was probed with either the *rh3* 3' flanking site or the plasmid marker h*dhfr* (Fig. 4A). The *rh3* 3' band at 7087bp reflects either the non-integrated plasmid or a single crossover event via the 5' flank of the *rh3* gene (Fig. 4A, C and E). The band detected with the h*dhfr* probe at 7078bp does not distinguish between an episomal plasmid or a 5' single crossover integration (4C and E). However, the expected band of 1385bp for the endogenous *rh3* gene is not included on the blot.

3D7/pkrass parasites subjected to two WR cycles followed by 5-FC challenge were again monitored by Southern blot (Fig. 4B). As above, no double crossover recombination was observed. These results clearly show that a cycling strategy of one or two drug cycles does not result in a double crossover recombination of pkrass into the *rh3* locus. Rather, plasmids were still maintained episomally or integrated into the locus by single crossover recombination via the 5' flanking sequence (Fig. 4B). The obtained *rh3* 3' bands at 7087bp and 1385bp do not allow to distinguish between these two alternatives (Fig. 4B, C and E).

This drug cycle strategy was repeated with a biological replicate. To further promote the loss of all episomes, the off-drug period in the second cycle was prolonged to three months. In addition, 3D7/pkrass was cycled three times following standard protocols (Alex Maier, personal communication) before 5-FC challenge (Fig. 5A). To distinguish more clearly between episomal plasmid, double crossover or single crossover events, gDNA from 3D7/pkrass was digested with *Spe*I/*Sph*I and probed with the *rh3* 5' flank (Fig. 5). In summary, the different cycling strategies resulted in the same outcome, namely a wild type and a plasmid band (10'566 bp and 3599bp, respectively), proving that neither single nor double crossover events had occurred in these 3D7/pkrass parasites. However, an additional band was observed at around 8000 bp, which cannot be explained by alternative integration via the upsC promoter into the PFL1960w locus (Fig. 5F). This band may be rather due an incomplete digest of the concatamer, since it is also observed in the lane loaded with pkrass only.

Taken together, attempts to integrate an upsC *var* promoter driving expression of a drug-selectable marker with abovementioned cycling strategies were unsuccessful. 3D7/pkrass transfectants were found to carry pkrass episomally still after three repeated off and on drug cycles as well as after a prolonged second off-drug cycle. This is unexpected, since the *rh3* locus has repeatedly been shown to be readily targetable for gene knock-out (Duraisingh et al, 2002; Taylor et al, 2001). We therefore concluded that our failing in integrating pkrass into *rh3* may be directly linked to the sequence attributes of the upsC promoter. Indeed, episomal *var* promoters were shown to remain in the active state up to 10 weeks off drug selection (Chookajorn et al, 2007). This entails that plasmids carrying an active *var* promoter are lost at a much slower rate. We therefore decided to change the cycling strategy and tested the consequences of a four-month period off drug without previous drug cycling. This should minimize the formation of stably replicating episomes (O'Donnell et al, 2002). Newly thawed 3D7/krass parasites from initial stocks were used. WR challenge after four months was followed by negative selection with the addition of the 5-FC prodrug. We observed that after cultivation without drug pressure for four months, the majority of the 3D7/pkrass parasite population had indeed lost their episomes (Fig. 6A). Addition of WR, however, resulted in the selection of transfectants still carrying concatamers (Fig. 6A).

In summary, with the different cycling strategies applied to 3D7/pkrass we were unable to target the *rh3* locus, let alone to yield the desired double crossover integration of the upsC promoter into the *rh3* locus. In all cases, the negative selection with 5-FC did not select against maintenance of episomes arguing for the accumulation of mutations in the *fcu* negative selectable marker resulting in resistance to 5-FC.

Conclusion

I have successfully cloned the plasmid pkrass designed for 3C experiments and effectively transfected it into *P. falciparum* parasites, generating the transgenic cell line 3D7/pkrass. Unfortunately, attempts to integrate the *var* upsC 5' sequence driving expression to the h*dhfr* drug-selectable cassette

into the *rh3* locus by double crossover recombination failed. Therefore, the planned 3C experiment to test for the existence of a unique DNA sequence in *P. falciparum* mediating mutually exclusive *var* gene expression was not performed. The existence of a unique DNA sequence controlling single *var* gene expression thus remains obscure. I set out to explain the obtained negative results and alternative approaches are discussed.

Chromosome ends in *P. falciparum* are clustered in aggregates at the nuclear periphery and clusters are linked by proteins (Freitas-Junior et al, 2000; Marty et al, 2006). Endogenous *var* genes co-localize with telomeric clusters irrespective of their transcriptional status (Marty et al, 2006). Similarly, episomal active *var* gene promoters localize to the nuclear periphery and are associated with a chromosome-end cluster (Voss et al, 2006). This suggests that the perinuclear subcompartment permissive for transcription (Duraisingh et al, 2005) lies within or in close proximity to a telomeric cluster. This strongly suggests that the active episomal *var* gene promoter in pkrass localizes to the nuclear periphery and is associated with a telomeric cluster. The locus for integration of pkrass is *rh3* (*PFL2520w*). This locus is located 140kb away from the right telomere and is thus uncertain to be contained into a telomeric cluster or not. Either way, it is tempting to speculate that the active episomal *var* promoter on pkrass and *rh3* are delocalized in the nucleus. Consequently, their spatial separation prevents the close proximity required for a recombination event between an active *var* promoter and *rh3*.

Episomal plasmids are lost rapidly in the absence of drug selection as they are partitioned unevenly amongst the daughter progeny (O'Donnell et al, 2002). However, it has been reported that an active episomal *var* gene promoter driving expression to a drug-selectable marker persists active for more than 10 weeks, even after the drug selection has been dismissed (Chookajorn et al, 2007). This type of molecular memory for gene expression is mediated through epigenetic factors. Trimethylated histone 3 at lysine-9 (H3K9me3) is significantly enriched at silent *var* genes (Chookajorn et al, 2007; Lopez-Rubio et al, 2007). In contrast, an active *var* gene locus is marked with tri- and dimethylated lysine-4 at histone 3 and in ring stage parasites (Lopez-Rubio et al, 2007). During schizogony, the active *var* gene becomes poised for transcription with modified H3K4 marks that are thought

to be inherited to the next generation and lead to reiterated *var* gene expression. Such inheritance may also rely on spatial memory conferred by epigenetic marks. I therefore conclude that a cycling strategy of one month off-drug may not result in the release of episomes carrying an active *var* promoter due to epigenetic memory. This is consistent with the fact that the pkrass episomes persisted for several months in culture and that WR selection after one month off drug did not directly influence parasite growth (data not shown), indicating that the transfectants do rarely switch away of the active episomal *var* promoter. Culturing of 3D7/pkrass for longer periods off drug may help to select for a rare double crossover event with a single var promoter cassette integrated. Indeed, continuous culturing of 3D7/pkrass for up to four months off drug resulted in a parasite population carrying hardly any episomes. The following drug selection, however, selected for a parasite population still carrying pkrass episomally. This suggests that pkrass can only be shed in a parasite where switching to an endogenous *var* gene occurred. Consequently, the episomal *var* gene promoter would become silenced and these plasmids would be lost upon subsequent WR selection. Whether a double crossover recombination event is facilitated with a silenced episomal *var* promoter is questionable. Episomal *var* promoters in their silenced state still localize to the nuclear periphery (Voss et al, 2006). Locus repositioning between an active and a silenced *var* gene appears to occur in the context of telomeric clusters (Marty et al, 2006). However, it is up to debate whether a silenced *var* promoter may come into closer nuclear vicinity to the *rh3* gene compared to an active *var* promoter.

An alternative approach is to include a second drug-selectable marker on pkrass to select for stable episomes. This strategy would allow the episomal *var* promoter to reside in its default silenced state. To keep the size of the resulting plasmid in a workable range the generation of a *bsd-fcu* fusion gene might be a promising option, combining both positive and negative selection in one gene. Repeated drug cycles can then be performed with BSD-S-HCl and only upon double crossover recombination, WR challenge selects for parasites with an active *var* promoter driving expression of the h*dhfr-gfp* fusion gene.

However, one does not entirely rely upon the use of a transgenic cell line to perfom 3C experiments: one possible approach would be the selection of parasites to an endothelial receptor, e.g. binding to CSA (Pouvelle et al, 1998). Wild-type parasites selected to express var2csa could theoretically be used for 3C experiments. Disadvantages of using an endogenous *var* gene promoter are (i) the difficulty to design unique primer sequences due to sequence homologies within the *var* gene family; and (ii) re-examination of restriction sites suitable for this approach.

References

Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL & Deitsch KW (2007) Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A* **104:** 899-902

Dekker J, Rippe K, Dekker M & Kleckner N (2002) Capturing chromosome conformation. *Science* **295:** 1306-1311

Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, Freitas-Junior LH, Scherf A, Crabb BS & Cowman AF (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. *Cell* **121:** 13-24

Duraisingh MT, Triglia T & Cowman AF (2002) Negative selection of Plasmodium falciparum reveals targeted gene deletion by double crossover recombination. *International Journal for Parasitology* **32:** 81-89

Dzikowski R, Frank M & Deitsch K (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog* **2:** e22

Erbs P, Regulier E, Kintz J, Leroy P, Poitevin Y, Exinger F, Jund R & Mehtali M (2000) In vivo cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene. *Cancer Res* **60:** 3813-3822

Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E & Deitsch K (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. *J. Biol. Chem* **281:** 9942-9952

Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, Nehrbass U, Wellems TE & Scherf A (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407:** 1018-1022

Fuss SH, Omura M & Mombaerts P (2007) Local and cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse. *Cell* **130:** 373-384

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M, Nene V, Shallom SJ et al (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419:** 498-511

Lambros C & Vanderberg JP (1979) Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J. Parasitol* **65:** 418-420

Lavstsen T, Salanti A, Jensen ATR, Arnot DE & Theander TG (2003) Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar. J* **2:** 27

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J & Axel R (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* **126:** 403-413

Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R & Scherf A (2007) 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol* **66:** 1296-1305

Maier AG, Braks JAM, Waters AP & Cowman AF (2006) Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination. *Mol. Biochem. Parasitol* **150:** 118-121

Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, Szestak T, Chesson J, Wu Y, Hughes K, Coppel RL, Newbold C, Beeson JG, Craig A, Crabb BS & Cowman AF (2008) Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. *Cell* **134:** 48-61

Marty AJ, Thompson JK, Duffy MF, Voss TS, Cowman AF & Crabb BS (2006) Evidence that Plasmodium falciparum chromosome end clusters are cross-linked by protein and are the sites of both virulence gene silencing and activation. *Mol. Microbiol* **62:** 72-83

Moolten FL (1986) Tumor chemosensitivity conferred by inserted herpes thymidine kinase genes: paradigm for a prospective cancer control strategy. *Cancer Res* **46:** 5276-5281

Mullen CA, Kilstrup M & Blaese RM (1992) Transfer of the bacterial gene for cytosine deaminase to mammalian cells confers lethal sensitivity to 5-fluorocytosine: a negative selection system. *Proc. Natl. Acad. Sci. U.S.A* **89:** 33-37

Navarro M & Gull K (2001) A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. *Nature* **414:** 759-763

Niimura Y & Nei M (2003) Evolution of olfactory receptor genes in the human genome. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 12235 -12240

O'Donnell RA, Freitas-Junior LH, Preiser PR, Williamson DH, Duraisingh M, McElwain TF, Scherf A, Cowman AF & Crabb BS (2002) A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. *EMBO J* **21:** 1231-1239

Palstra R, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F & de Laat W (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet* **35:** 190-194

Pouvelle B, Fusaï T, Lépolard C & Gysin J (1998) Biological and biochemical characteristics of cytoadhesion of Plasmodium falciparum-infected erythrocytes to chondroitin-4-sulfate. *Infect. Immun* **66:** 4950-4956

Ralph SA, Scheidig-Benatar C & Scherf A (2005) Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. *Proc. Natl. Acad. Sci. U.S.A* **102:** 5414-5419

Rodriguez I (2007) Odorant and pheromone receptor gene regulation in vertebrates. *Curr. Opin. Genet. Dev* **17:** 465-470

Salanti A, Staalsoe T, Lavstsen T, Jensen ATR, Sowa MPK, Arnot DE, Hviid L & Theander TG (2003) Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol. Microbiol* **49:** 179-191

Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, Pouvelle B, Gysin J & Lanzer M (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. *EMBO J* **17:** 5418-5426

Spilianakis CG, Lalioti MD, Town T, Lee GR & Flavell RA (2005) Interchromosomal associations between alternatively expressed loci. *Nature* **435:** 637-645

Taylor HM, Triglia T, Thompson J, Sajid M, Fowler R, Wickham ME, Cowman AF & Holder AA (2001) Plasmodium falciparum homologue of the genes for Plasmodium vivax and Plasmodium yoelii adhesive proteins, which is transcribed but not translated. *Infect. Immun* **69:** 3635-3645

Tolhuis B, Palstra RJ, Splinter E, Grosveld F & de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* **10:** 1453-1465

Trager W & Jenson JB (1978) Cultivation of malarial parasites. *Nature* **273:** 621-622

Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, Beeson JG, Reeder JC, Crabb BS & Cowman AF (2006) A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. *Nature* **439:** 1004-1008

Voss TS, Tonkin CJ, Marty AJ, Thompson JK, Healer J, Crabb BS & Cowman AF (2007) Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. *Mol. Microbiol* **66:** 139-150

Williamson DH, Janse CJ, Moore PW, Waters AP & Preiser PR (2002) Topology and replication of a nuclear episomal plasmid in the rodent malaria Plasmodium berghei. *Nucleic Acids Research* **30:** 726 -731

Figure legends

Figure 2. Cloning strategy for pkrass. The different plasmids and intermediates are schematically depicted. (A) pCC4RH3 (Voss, unpublished). (B) pCC4RH3-linker1. (C) pB-linker2 (modified pBcamHG (Brancucci and Witmer et al., submitted)). (D) pBupsC(T7_SP6)HG with the inserted *var* upsC promoter flanked by T7 and SP6 sequences. (E) Final plasmid pkrass used for transfection. *Pb* DT3', *P. berghei* dihydrofolate reductase-thymidilate synthase terminator; *hsp86* 5', heat shock protein 86 upstream region; *rh3* flank, reticulocyte binding-like protein homologue 3 5' and 3' flanking sites (red and orange box, respectively); *cam* promoter, calmodulin upstream sequence; *bsd*, blasticidin deaminase drug resistance cassette (open box); FCU, *S. cerevisiae* cytosine deaminase and uracil phosphoribosyl transferase fusion gene (black box); *hrp2* 3', histidine rich protein 2 terminator sequence; h*dhfr-gfp*, human dihydrofolate reductase and green fluorescent protein fusion gene (grey box). Plasmid backbones are not shown.

Figure 3. Expected pkrass double and single crossover integrations at the *rh3* locus. (A) schematic depiction of the *rh3* locus and pkrass. (B) pkrass double crossover into the *rh3* locus via coding 5' and 3' flanking sites. (C) pkrass single crossover integration via the *rh3* 5' flanking sequence. (D) pkrass single crossover integration via the *rh3* 3' flanking sequence.

Figure 4. One and two drug cyclings did not result in a double crossover integration of pkrass. (A) Southern blot of *Pvu*II-digested of 3D7/pkrass gDNA after a completed off and on WR cycle and negative selection with 5-FC. pkrass, plasmid control. *rh3* 3' and h*dhfr* probes were radiolabeled with ^{32}P-ATP. The detected band at 7087bp does not distinguish between a 5' single crossover integration or episomal plasmids. (B) Southern blot of *Pvu*II-digested gDNA of 3D7/pkrass after two completed off and on WR cycle and negative selection with 5-FC. 3D7 wild-type gDNA and pkrass plasmid DNA are controls. *rh3* 3' and h*dhfr* probes were radiolabeled with ^{32}P-ATP. The detected band at 7087bp in the h*dhfr*-probed blot indicates either episomal plasmids or a 5' single crossover integration. The 1345bp band in the rh3 3'

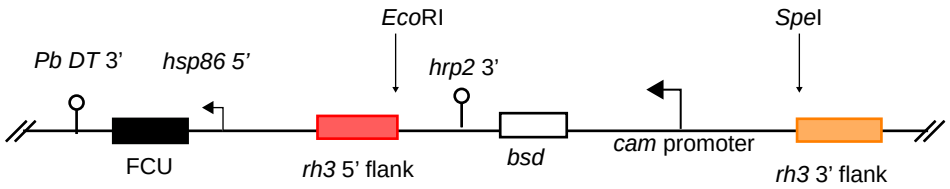
probed blot confers to an intact *rh3* locus or a 5' single crossover integration, the 7087bp band confers to episomal pkrass or a 5' single crossover integration. (C-F) *Pvu*II restriction sites are shown for pkrass, the *rh3* locus and the three integration maps. The resulting bands in Southern blot probed with *rh3* 3' and h*dhfr* are indicated.

Figure 5. Two and three drug cyclings did not result in a double crossover integration of pkrass. (A) Southern blot of *Spe*I/*Sph*I-digested gDNA of 3D7/pkrass after two or three completed off and on WR cycles and negative selection with 5-FC. 3D7 wild-type gDNA and pkrass plasmid DNA are controls. *rh3* 5' probe was radiolabeled with ^{32}P-ATP. The band at about 10000bp confers to the endogenous *rh3* gene, the band at 3699bp corresponds to the fragment carrying the *rh3* 5' flank on the plasmid. Unspecific bands are most probably due to incomplete digestion. (B-F) *Spe*I/*Sph*I restriction sites are shown for pkrass, the *rh3* locus and the three integration maps as well as for the PFL1960w locus. The resulting bands in Southern blot probed with *rh3* 5' are indicated.

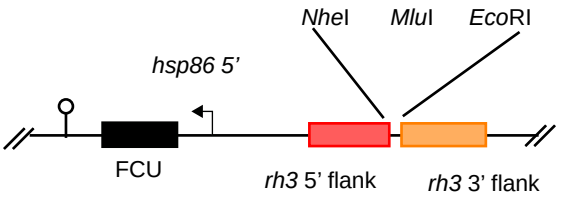
Figure 6. pkrass remains episomal in 3D7/pkrass parasites after 4 months off drug selection. (A) Southern blot of *Cla*I-digested 3D7/pkrass gDNA after four months off WR selection, followed by WR drug selection and negative selection with 5-FC. 3D7 wild-type gDNA and pkrass plasmid DNA are controls. *rh3* 3' probe was radiolabeled with ^{32}P-ATP. The band at 9336bp confers to the endogenous *rh3* gene, the band at 4837bp corresponds to the fragment carrying the *rh3* 3' flank on the plasmid. (B-E) *Cla*I restriction sites are shown for pkrass, the *rh3* locus and the three integration maps. The resulting bands in Southern blot probed with *rh3* 3' are indicated.

Figure 2

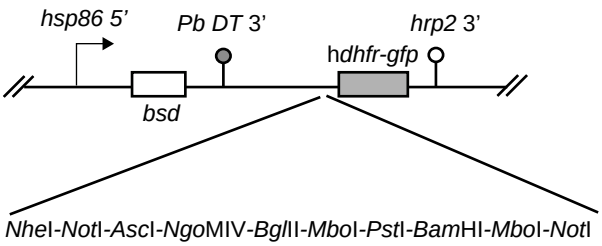
A plasmid map of pCC4RH3



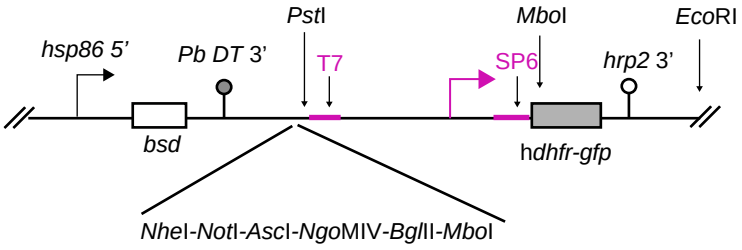
B plasmid map of pCC4RH3-linker1



C plasmid map of pB-linker2



D plasmid map of pBupsC(T7_SP6)HG



E plasmid map of pkrass

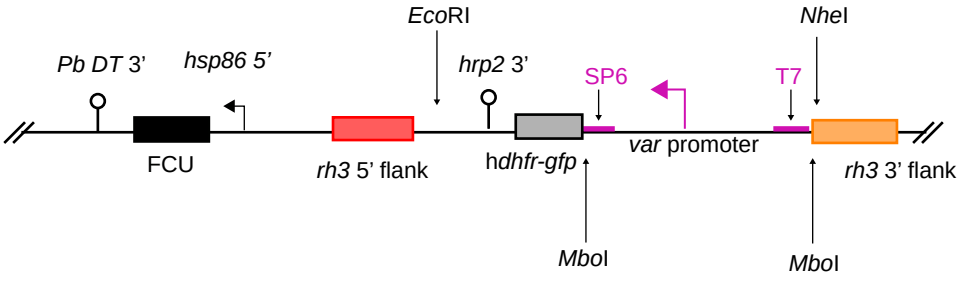
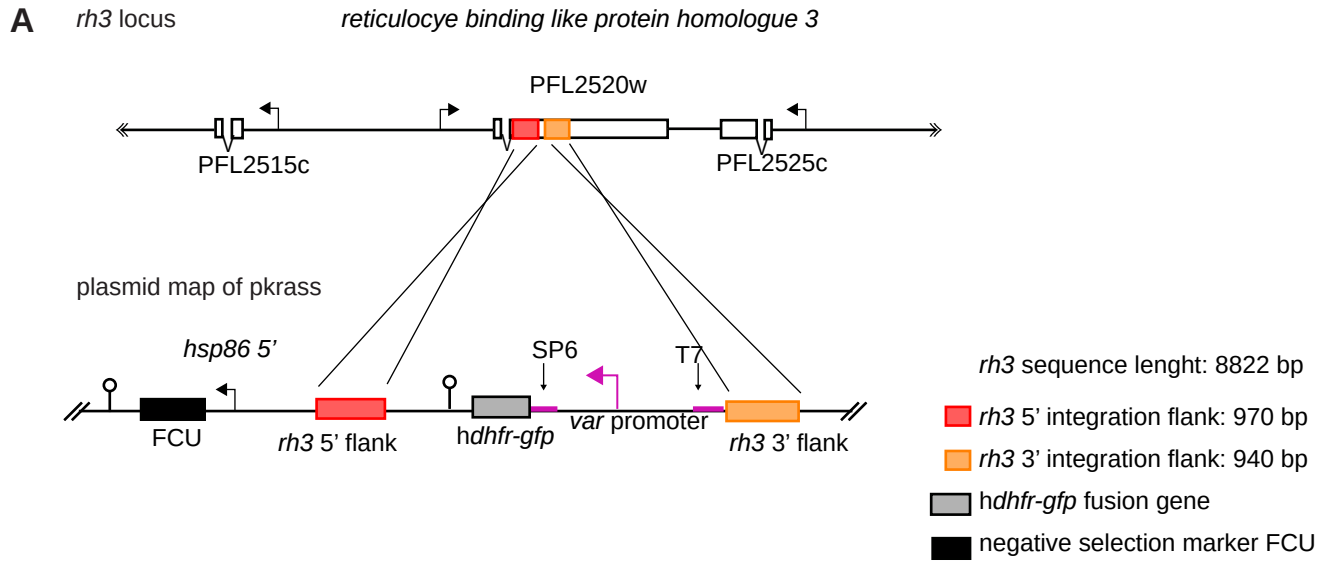
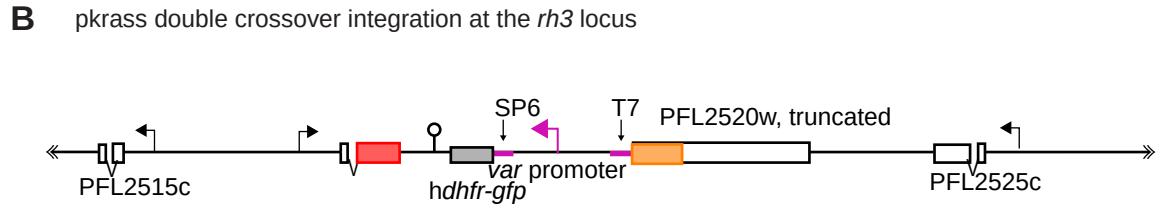
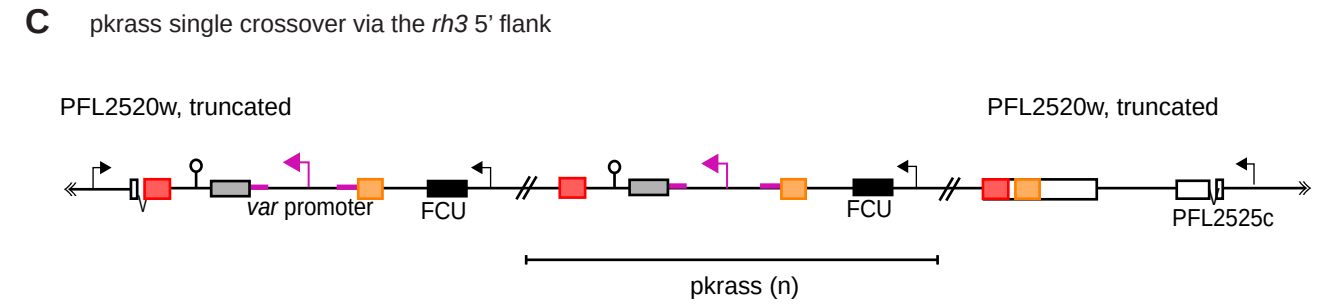


Figure 3


A
rh3 locus
reticulocye binding like protein homologue 3
PFL2520w
PFL2515c
PFL2525c
plasmid map of pkrass
hsp86 5'
SP6
T7
FCU
rh3 5' flank
hdhfr-gfp
var promoter
rh3 3' flank
rh3 sequence lenght: 8822 bp
rh3 5' integration flank: 970 bp
rh3 3' integration flank: 940 bp
hdhfr-gfp fusion gene
negative selection marker FCU


B
pkrass double crossover integration at the rh3 locus
SP6
T7
PFL2520w, truncated
PFL2515c
var promoter
hdhfr-gfp
PFL2525c


C
pkrass single crossover via the rh3 5' flank
PFL2520w, truncated
PFL2520w, truncated
var promoter
FCU
FCU
PFL2525c
pkrass (n)

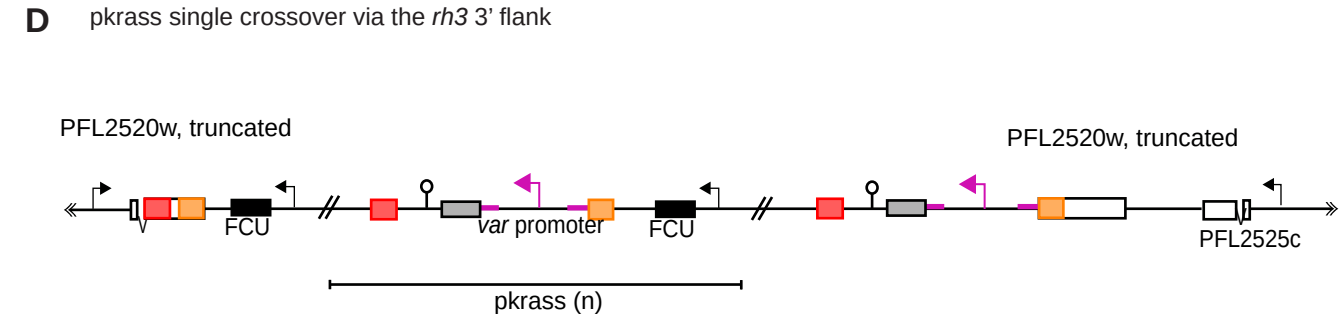
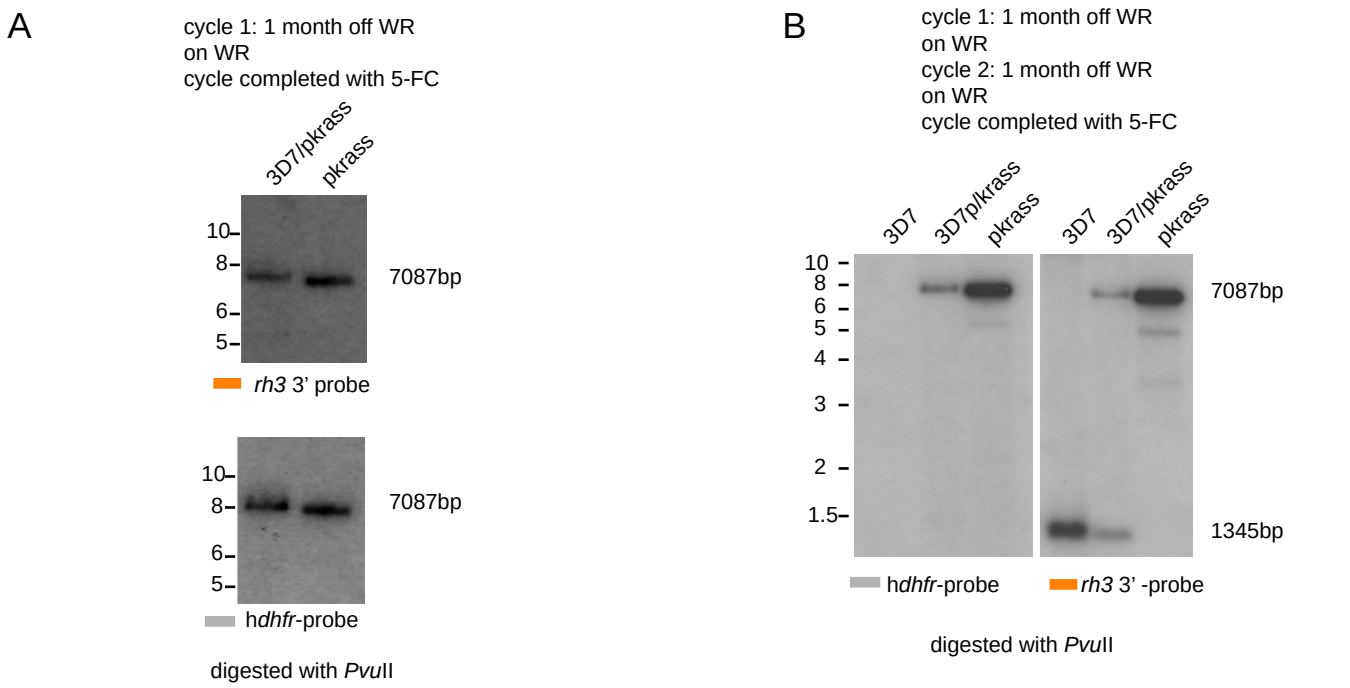
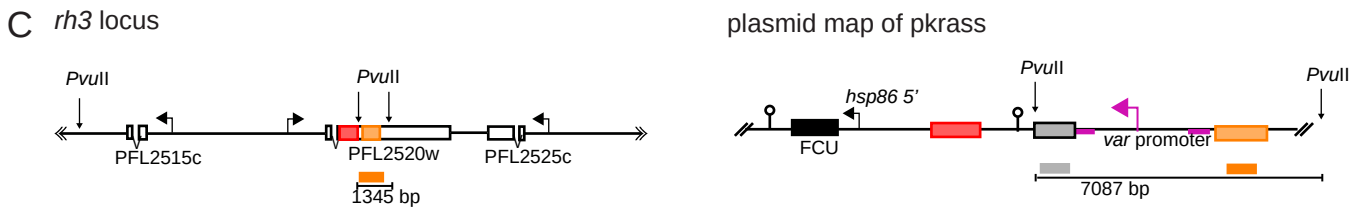
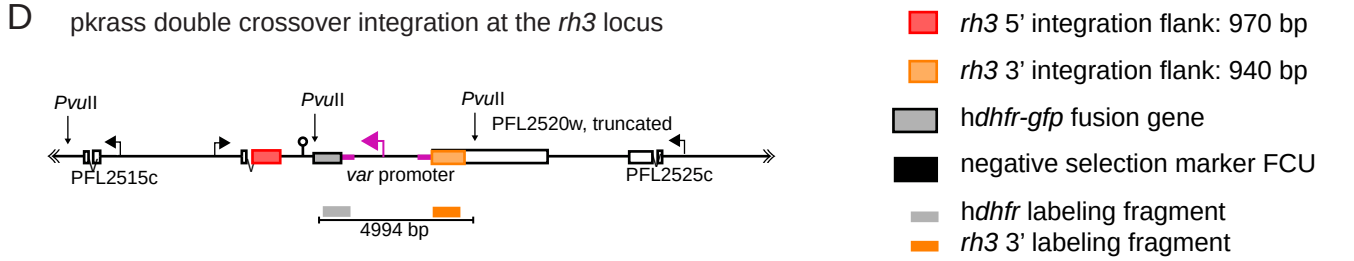
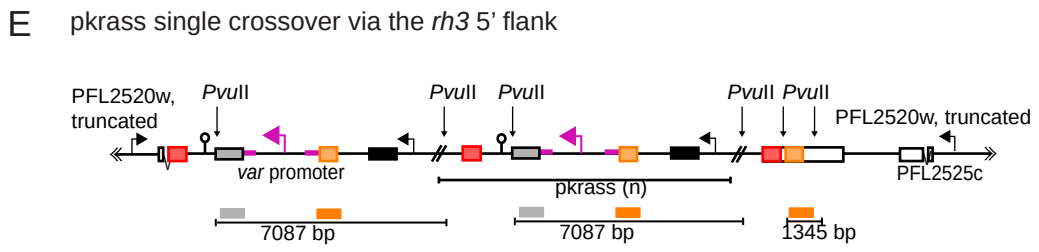

D
pkrass single crossover via the rh3 3' flank
PFL2520w, truncated
PFL2520w, truncated
FCU
var promoter
FCU
PFL2525c
pkrass (n)

Figure 4


A
cycle 1: 1 month off WR
on WR
cycle completed with 5-FC
3D7/pkrass
pkrass
10
8
6
5
7087bp
rh3 3' probe
10
8
6
5
7087bp
hdhfr-probe
digested with PvuII
B
cycle 1: 1 month off WR
on WR
cycle 2: 1 month off WR
on WR
cycle completed with 5-FC
3D7
3D7p/krass
pkrass
3D7
3D7/pkrass
pkrass
10
8
6
5
4
3
2
1.5
7087bp
1345bp
hdhfr-probe
rh3 3' -probe
digested with PvuII


C
rh3 locus
PvuII
PvuII
PFL2515c
PFL2520w
PFL2525c
1345 bp
plasmid map of pkrass
PvuII
PvuII
hsp86 5'
FCU
var promoter
7087 bp


D
pkrass double crossover integration at the rh3 locus
PvuII
PvuII
PvuII
PFL2520w, truncated
PFL2515c
var promoter
PFL2525c
4994 bp
rh3 5' integration flank: 970 bp
rh3 3' integration flank: 940 bp
hdhfr-gfp fusion gene
negative selection marker FCU
hdhfr labeling fragment
rh3 3' labeling fragment


E
pkrass single crossover via the rh3 5' flank
PFL2520w,
truncated
PvuII
PvuII
PvuII
PvuII
PvuII
PFL2520w, truncated
var promoter
pkrass (n)
PFL2525c
7087 bp
7087 bp
1345 bp

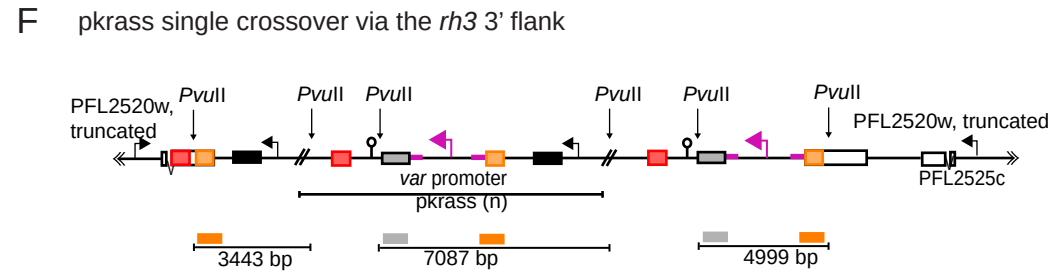

F
pkrass single crossover via the rh3 3' flank
PFL2520w,
truncated
PvuII
PvuII
PvuII
PvuII
PvuII
PvuII
PFL2520w, truncated
var promoter
pkrass (n)
PFL2525c
3443 bp
7087 bp
4999 bp

Figure 5

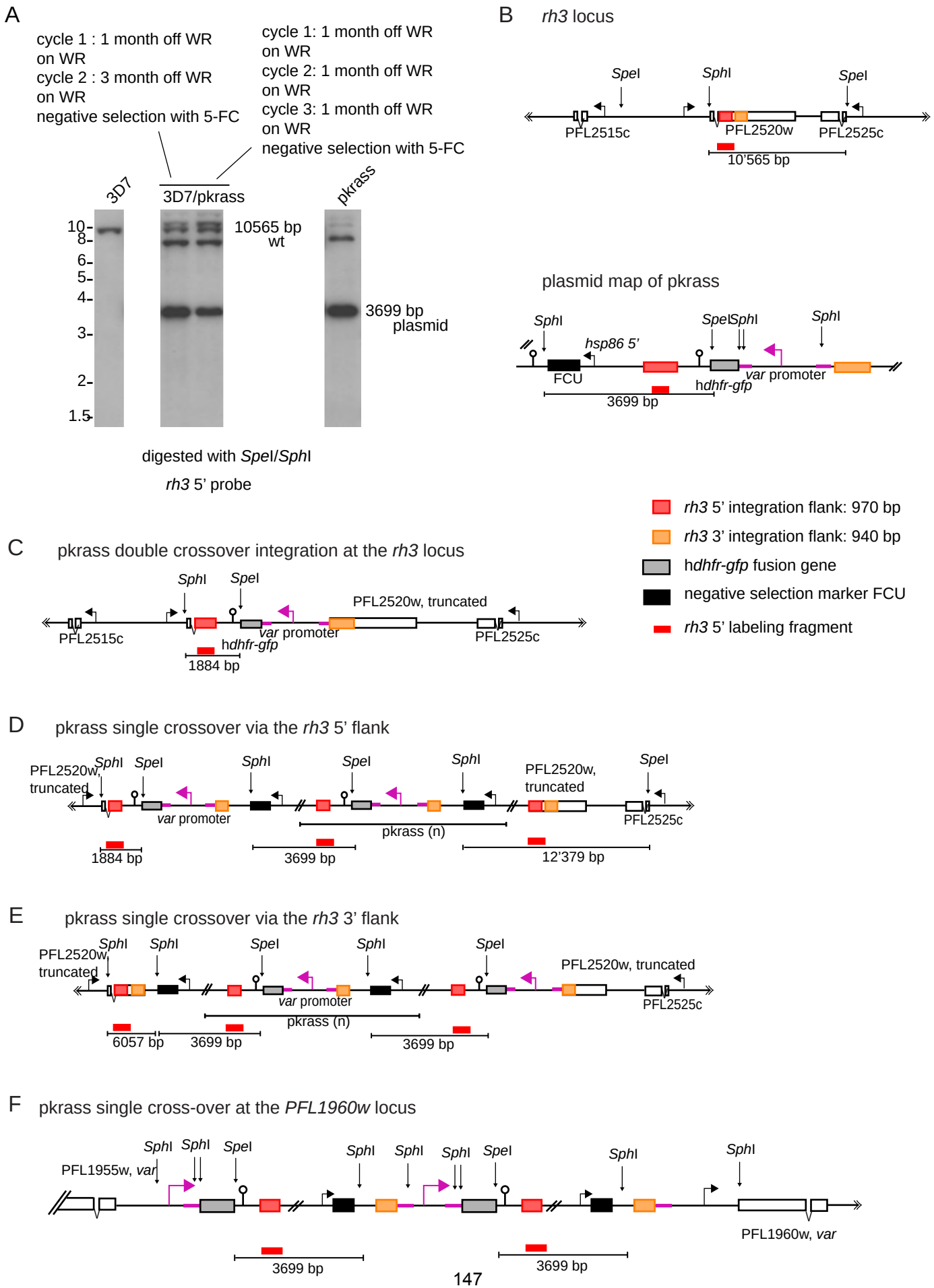
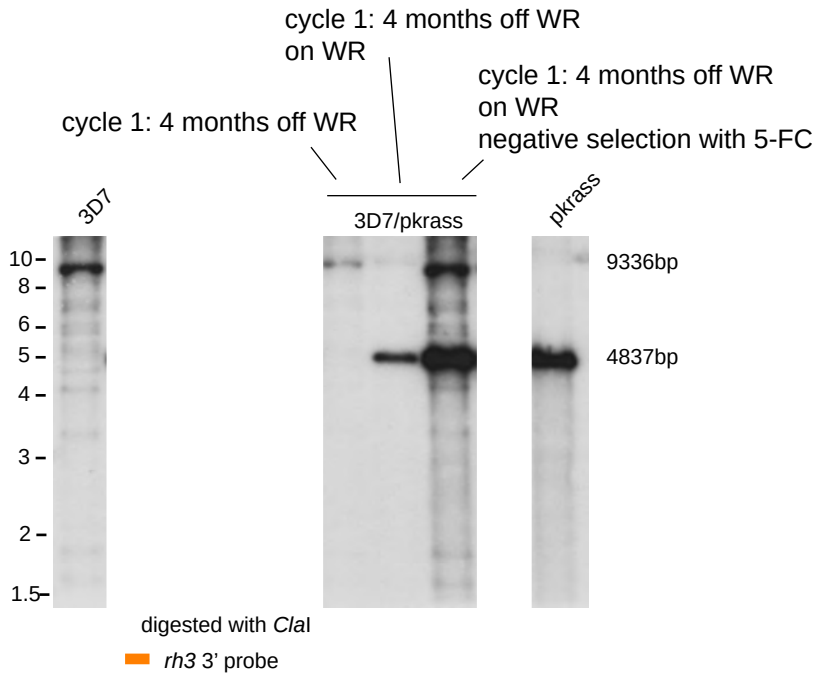
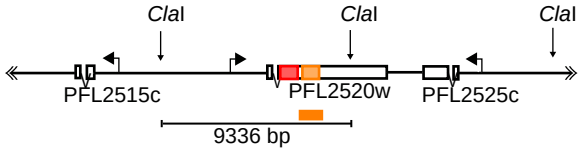

A
cycle 1 : 1 month off WR
on WR
cycle 2 : 3 month off WR
on WR
negative selection with 5-FC
cycle 1: 1 month off WR
on WR
cycle 2: 1 month off WR
on WR
cycle 3: 1 month off WR
on WR
negative selection with 5-FC
3D7
3D7/pkrass
pkrass
10
8
6
5
4
3
2
1.5
10565 bp
wt
3699 bp
plasmid
digested with SpeI/SphI
rh3 5' probe
B
rh3 locus
SpeI
SphI
SpeI
PFL2515c
PFL2520w
PFL2525c
10'565 bp
plasmid map of pkrass
SphI
SpeI
SphI
SphI
hsp86 5'
FCU
var promoter
hdhfr-gfp
3699 bp
rh3 5' integration flank: 970 bp
rh3 3' integration flank: 940 bp
hdhfr-gfp fusion gene
negative selection marker FCU
rh3 5' labeling fragment
C
pkrass double crossover integration at the rh3 locus
SphI
SpeI
PFL2520w, truncated
PFL2515c
var promoter
hdhfr-gfp
PFL2525c
1884 bp
D
pkrass single crossover via the rh3 5' flank
PFL2520w,
truncated
SphI
SpeI
SphI
SpeI
SphI
PFL2520w,
truncated
SpeI
var promoter
pkrass (n)
PFL2525c
1884 bp
3699 bp
12'379 bp
E
pkrass single crossover via the rh3 3' flank
PFL2520w,
truncated
SphI
SphI
SpeI
SphI
SpeI
PFL2520w, truncated
var promoter
pkrass (n)
PFL2525c
6057 bp
3699 bp
3699 bp
F
pkrass single cross-over at the PFL1960w locus
PFL1955w, var
SphI
SphI
SpeI
SphI
SphI
SphI
SpeI
SphI
SphI
PFL1960w, var
3699 bp
3699 bp

Figure 6

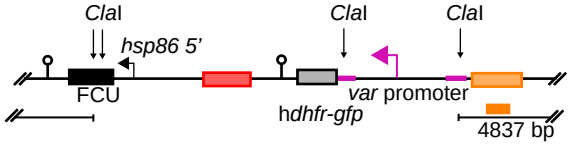
A



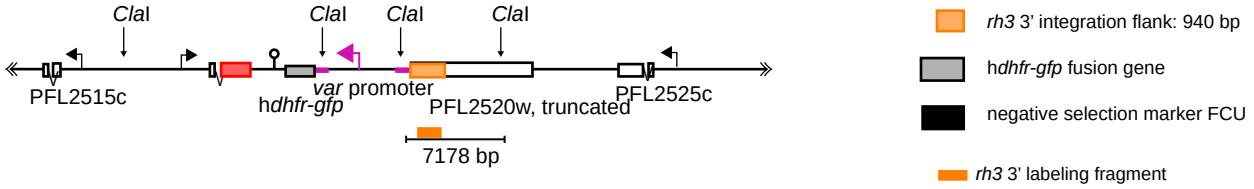
B *rh3* locus



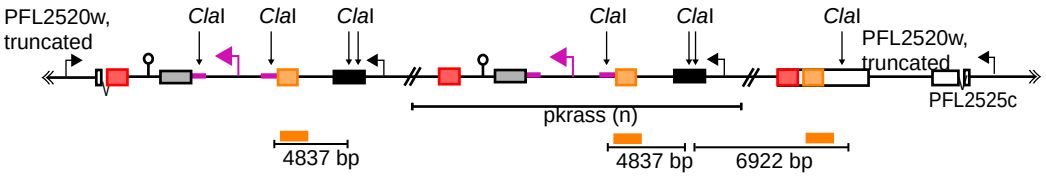
plasmid map of pkrass



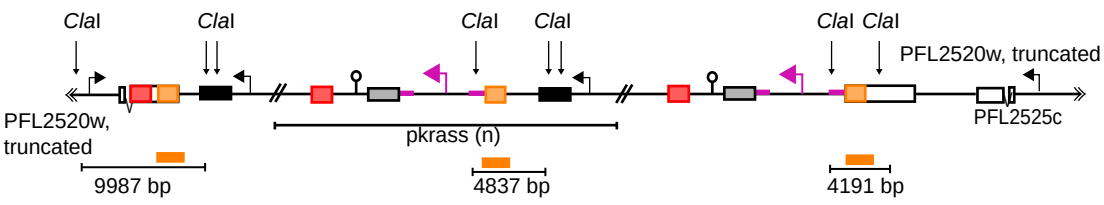
C pkrass double crossover integration at the *rh3* locus



D pkrass single crossover via the *rh3* 5' flank



E pkrass single crossover via the *rh3* 3' flank



Supplementary table 1. Oligo sequences used in this study.

Primers used for cloning

primer name	primer sequence
linker1_CC4_F	aattgctagcacgcgtgaatt
linker1_CC4_R	ctagaattcacgcgtgctagc
pstI_T7_upsC_forw	catgctgcagccctatagtgagtcgtattattggtacattatacatgaatttcatg
bglII_SP6_upsC_rev	catgagatctctatagtgtcacctaaatcttttgttttttgtttatcgttcgtg
linker2_pBHG_F	agctagcgcggccgcggcgcgccgccggcagatctgatcctgcagacgtggatccgatcgc
linker2_pBHG_R	ggccgcgatcggatccacgtctgcaggatcagatctgccggcggcgcgccgcggccgcgctagcttgca

Primers used for Southern probes

primer name	primer sequence
RH3 3'_F	aaaataaatccgctgatagtac
RH3 3'_R	catcaactaaggtttcatca
dhfr_F	agctggatccgcggccgcaaaacatgcatggttcgctaaactg
dhfr_R	agctgtcgacagcagcatcattcttctcatatacttcaa
RH3 5'_F	cagtccgcggtattcaaaggaacaatgaagg
RH3 5'_R	cagtactagtcccaatattctcttgtcc

General Discussion and Outlook

With this work I provide deeper insights into the transcriptional regulation of the *var* gene family and other multigene families in *P. falciparum* by using transfection-based approaches. We show that among parasite virulence gene families, the mutually exclusive mode of transcriptional control is restricted to the *var* gene family. Based on functional data we find that mutually exclusive transcription and *var* promoter activation are uncoupled, and that activation appears to be achieved by a standard upstream activating sequence (UAS) located upstream of the transcriptional start site (TSS). With regard to the mechanism of mutual exclusion we demonstrate that the absence of a 101bp sequence (MEE; bps -315 to -215 relative to the ATG) in a *var* gene promoter disqualifies it from being recognized by the mutual exclusion mechanism although it does not impair promoter activity. Hence, for the first time we were able to identify *cis*-acting elements in a *var* gene promoter that are essential in mediating *var* gene activation and dictating the singularity of gene expression. These findings are clearly in support of the previously proposed existence of a specialized perinuclear compartment where mutually exclusive transcription of *var* genes is accomplished.

The phenomenon of mutually exclusive transcription

The expression of a single gene from a large gene family is a widespread phenomenon used by many different organisms. In pathogens, single gene choice of exposed antigens contributes to antigenic variation. Antigenic variation serves to avoid eradication by the host's immune system and helps to maintain persistent infections. In African trypanosomes antigenic variation is reflected by stochastic switches of the variant surface glycoprotein (VSG) coat. VSG switching is driven by recombination events that bring a previously silenced *vsg* gene into an active expression site (ES), or by transcriptional switches in which the active ES is silenced and a previously silenced ES gets activated (Morrison et al, 2009). In mammals, single gene choice occurs in in B and T cell antigen receptors, as well as in olfactory sensory neurons, where only one of ~1300 members of the odorant receptor genes is expressed

(Chess et al, 1994). Intriguingly, DNA recombination is not involved in single odorant receptor gene choice (Eggan et al, 2004; Li et al, 2004). A current model of single gene choice implies a single *trans*-acting enhancer element H that allows stochastic activation of only one odorant receptor allele (Lomvardas et al, 2006). A similar mechanism may be employed by *P. falciparum*.

The structure and organisation of *P. falciparum* chromosome ends is remarkably conserved. They are characterised by telomere repeats that are followed by telomere-associated repeat elements TARE1 to 6 that are conserved by both sequence and position. This non-coding part of the chromosomes is followed by a semi-conserved arrangement of members of many gene families which encode proteins exported or predicted to be exported into the red blood cell (RBC) and/or onto its surface. Subtelomeric regions in *P. falciparum* are uniformly associated with H3K9me3/PfHP1-enriched heterochromatin, an environment facilitating silencing and variegated gene expression and setting the stage for antigenic variation (Flueck et al, 2009; Pérez-Toledo et al, 2009; Salcedo-Amaya et al, 2009). The best studied gene family is the *var* gene family, encompassing ~60 members in the *P. falciparum* 3D7 genome (Gardner et al, 2002). Based on their conserved upstream regions and chromosomal location, *var* genes can be divided into three main subgroups termed upsA, upsB and upsC, respectively (Lavstsen et al, 2003). *var* genes are transcribed in a mutually exclusive way and *in situ* switching of the expressed member leads to antigenic variation of the encoded protein PfEMP1 (*Plasmodium falciparum* erythrocyte protein 1) (Scherf et al, 1998). PfEMP1 is the major *P. falciparum* virulence factor, mediating sequestration of infected red blood cells to the microvascular system in the human host (Baruch et al, 1996; Gardner et al, 1996). Importantly, antibodies of infected individuals directed against parasite proteins on the surface of iRBCs mainly recognize PfEMP1 antigens (Voss et al, 2006; Maier et al, 2008). Furthermore, PfEMP1 suppresses the IFN-γ production by host immunomodulatory cells (D'Ombrain et al, 2007). Other subtelomeric gene families that are intermingled with *var* are termed *rif*, *stevor*, *phist* and *pfmc-2tm* (Weber, 1988; Cheng et al, 1998; Sargeant et al, 2006; Sam-Yellowe et al, 2004). Unlike PfEMP1, the role of these proteins in

parasite biology and disease is unclear, but there is indication that they might play a role in host-parasite interactions and antigenic variation as well (Lavazec et al, 2007; Abdel-Latif et al, 2002, 2003; Fernandez et al, 1999; Kyes et al, 1999; Schreiber et al, 2008). Nonetheless, the *var*, *rif*, *stevor*, *phist* and *pfmc-2tm* families are highly similar regarding their subtelomeric location and the resulting heterochromatic environment, and they encode the N-terminal export signal termed PEXEL or VTS (Marti et al, 2004; Hiller et al, 2004). However, despite all these resemblances, we find that the transcriptional regulation employed by the var and other gene families is remarkably different.

We show that mutually exclusive transcription is uniquely adopted by the *var* gene family and not by other gene families. We used a genome-wide transcriptomic approach with transgenic cell lines carrying an active multigene family promoter upstream of a selectable marker gene. While the activation of the episomal *var* promoter induces a family-specific knock-down of the endogenous *var* gene family, we did not see such an effect when activating other gene family promoters. These results clearly show that the above-mentioned non-*var* multigene families are not mutually exclusively transcribed. Neither did we observe a downregulation of endogenous multigene families independent of the source of the episomal promoter, a phenomenon that has been previously reported (Howitt et al, 2009). In their study, Howitt and colleagues propose a common activation factor for *P. falciparum* multigene family transcription. A previous study of the same group showed that increased copy numbers of an episomal *var* promoter gradually downregulated the expression of the endogenous *var* gene family, and proposed the existence of a limited nuclear factor that is required for *var* gene activation (Dzikowski & Deitsch, 2008). The study of Howitt and colleagues is based on these results but was expanded to other multigene families (Howitt et al, 2009). The authors compared the transcription of endogenous gene families in transfectants carrying active episomal *var*, *rif*, *stevor* and *pfmc-2tm* promoters. Increasing copy numbers of these plasmids resulted in a downregulation of not only the selfsame, but also of other gene family members. However, in this study, this effect was not constant or consistent since for example increased copy numbers of the *pfmc-2tm* promoter

transfectant resulted in increased endogenous *var* expression. Our findings, in contrast, do not show a downregulation of any gene family upon activation of any promoter, apart from the well-established *var*-specific knock-down induced by an active *var* promoter. Furthermore, in the study of Howitt et al. only one developmental timepoint was investigated, coinciding with the peak expression of the respective family. This approach is more susceptible to be confounded by stage-specific differences in transcription of genes during the IDC (intraerythrocytic developmental cycle). In our study, we monitored the expression of multigene families at four consecutive time points, an experimental design less prone to growth-induced variations. We further included two controls, 3D7 wild-type parasites and mock transfectant where the resistance gene is controlled by a the single-copy house-gene promoter *cam* to build a solid basis for the comparison of transcriptional changes in the seven test transfectants. Surprisingly, we observed a downregulation of *var* and other heterochromatic genes even in the mock transfectant, I therefore propose that transcription of heterochromatic gene families is *per se* downregulated in transfected cell lines carrying episomes, and that increasing plasmid copy numbers intensifies this effect. Parasites replicating high copy numbers may be exposed to increased stress and may use more resources to survive, and might balance this with decreased expression of non-essential virulence genes. Alternatively, parasites carrying episomes are known to have delayed growth compared to wild-type parasites. Thus, comparing wild-type parasites to transgenic cell lines may to detect transcriptional changes that are in fact due to slight differences in life cycle stage. It is therefore difficult to differentiate between an observed effect being caused by a real phenomenon or developmental timepoint.

The induced *var* knock-down in our transgenic cell lines, however, is due to single *var* gene choice and is a specific transcriptional control of this family. In our study, we confirm previous findings for upsB and upsC promoters (Voss et al, 2006, 2007), but we also show for the first time that a knock-down of endogenous *var* gene transcription is also mediated by an active episomal upsA promoter. This clearly demonstrates that mutually exclusive *var* expression is independent of the *var* promoter subtype. Taken together, we demonstrate that the *var* gene family in *P. falciparum* responds to particular

ordinances of transcriptional control. It may be speculated that in the *in vivo* situation simultaneous expression of several PfEMP1 variants may be detrimental to parasite survival whereas for other antigen families this may not be the case. This can be explained by the fact that PfEMP1 is highly immunogenic and is the main inducer of *P. falciparum*-specific antibody responses against the surface of iRBCs. Antigenic variation of this gene family helps to establish chronic infection and consequently, prolongs parasite survival and facilitates transmission. Why such a tight control of expression doesn't appear to be required for other surface antigen families is unclear. It is possible that antigens such as RIFINs and STEVORs are simply not immunodominant, or their exposure on the surface is masked by the much larger PfEMP1 antigens.

Are *cis*-acting elements the main players in singular gene choice in *P. falciparum*?

To identify sequence elements contributing to *var* gene activation and mutually exclusive transcription, we generated transgenic cell lines with truncated upsC promoters and with upsC promoter fragments cloned into the context of the heterologous minimal *kahrp* (knob-associated histidine-rich protein) promoter. We identified an upstream activation sequence (UAS) located upstream of the upsC TSS that acts independently of the promoter context. Moreover, we identified a 101bp sequence (-315 to -215bp) acting independently of transcriptional activity and orchestrating mutually exclusive expression of *var* genes. This element was termed MEE for mutual exclusion element. The absence of the MEE did not impair upsC promoter activity, but excluded the upsC promoter from the mutual exclusion programme. The function of the MEE is location-dependent in a heterologous context since it did not impair *kahrp* promoter activity when positioned upstream of the TSS, but led to a complete repression of the *kahrp* promoter when cloned downstream of the TSS.

The specific binding of a protein or protein complex to a 47bp motif within the MEE suggests its involvement in halting transcription, acting as a physical block to prevent transcriptional elongation. We therefore propose the existence of a mutual exclusion element interacting factor (MIF), consisting of

one or several proteins that bind to the MEE. How MIF acts as a transcriptional repressor remains elusive, but there are several ways how this could be achieved. One possibility is that the MIF acts as a crude physical block of transcription. Alternatively, MIF might act as a classical repressor and/or MIF might recruit factors that establish silencing marks. Intriguingly, the blocking effect of a sequence containing the MEE downstream of the heterologous *kahrp* promoter can only be overcome by a spontaneous deletion of the *var* intron in one copy of the integrated concatamer. These results implicate that the intron is somehow involved in maintaining the function of the MIF complex. Indeed, previous studies have demonstrated an involvement of the var intron in episomal *var* promoter silencing (Deitsch et al, 2001; Voss et al, 2006; Frank et al, 2006). However, the exact role of the intron (harbouring bidirectional promoter activity) (Epp et al, 2008) gets complicated by the fact that its replacement with an unrelated promoter similarly silences an episomal *var* promoter (Dzikowski et al, 2007). This strongly suggests that the silencing effect of the *var* intron is not sequence-specific, but may be rather conferred by its own promoter activity. I therefore propose that the function of the MIF may require an interaction with general transcription factors that are recruited to the intron.

Our proposed model of single *var* gene choice is currently based on the identification of two regulatory elements of one out of 60 *var* promoters: the UAS and the MEE. The assignment to this model to have a pan-*var* validity implies that the UAS and MEE regulatory sequences are present in all *var* genes. However, a conducted MEME search on 60 *var* upstream sequences to define short motifs (an approach that is commonly used to identify binding sites for transcription factors) did not result in the desired identification of similar motifs (UAS and MEE) in other *var* promoters (Christoph Schmid, unpublished). Putative conserved *cis*-acting elements may be masked by the overall sequence conservation between the promoters. Thus, alternative approaches to test for the existence of functional UAS and MEE in all *var* promoters have to be undertaken. A very time-consuming and nerve-wracking approach would be the experimental validation for another *var* promoter with a similar transfection-based approach. Therefore, a modified *in-silico* approach would be preferable. Positionally conserved nucleotide base composition or

nucleosome positioning effects may help to elucidate if motifs similar to UAS and MEE are present in other *var* promoters.

Endogenous *var* genes and active episomal *var* promoters co-localize with telomeric clusters at the nuclear periphery irrespective of their transcriptional state (Freitas-Junior et al, 2000; Marty et al, 2006; Voss et al, 2006). This suggests that the perinuclear subcompartment permissive for *var* transcription (Duraisingh et al, 2005) lies within or in close proximity to a telomeric cluster (Ralph et al, 2005; Lopez-Rubio et al, 2009; Marty et al, 2006). Our repeated attempts to integrate an active *var* promoter into a locus 140kb away from a telomere failed. It is likely that integration of an active *var* promoter into this locus is impossible due to spatial constraints. Taken together, this strongly suggests that the VES (*var* expression site) is a restricted compartment only accommodating the active *var* gene. In this context, we can alternatively explain the strong repression of the construct $pK_{min}C4$: integration of this plasmid in the *kahrp* locus created a promoter swap where the last 500bp of the 5'UTR of the *kahrp* upstream region are substituted with a *var* sequence harbouring the MEE. Assuming that the VES is indeed a restricted compartment, it may be that the *kahrp* locus is spatially remote, similar to the *rh3* locus, such that an interaction of the VES and the MEE is impossible. As a consequence, the removal of the intron may have abolished the blocking effect of the MEE without the requirement for interaction with the VES allowing for VES-independent expression of the reporter gene.

Previous studies revealed a dynamic epigenetic footprint on *var* promoters, suggesting that epigenetic memory is involved in maintenance and switching of the transcriptional state of the *var* gene repertoire. Activation and silencing of *var* genes correlates with specific histone tail marks. H3K9ac and H3K4me are associated with *var* gene activation (Lopez-Rubio et al, 2007), whereas H3K9me3 marks silenced *var* genes (Lopez-Rubio et al, 2007; Chookajorn et al, 2007). Heterochromatin formation can be nucleated via sequence-specific DNA-protein interactions or RNAi (Bernstein & Allis, 2005). However, the RNAi machinery is absent in *P. falciparum* (Baum et al, 2009). A protein has been identified that binds specifically to sequence motifs upstream of subtelomeric upsB *var* genes and within TAREs. This interaction may serve in nucleating formation of subtelomeric heterochromatin (Flueck et al, 2010). To

date, however, it remains completely unknown how chromatin marks are established at *var* promoters.

Previous findings together with this work provide evidence that mutually exclusive transcription of the *var* gene family is accomplished by a novel mechanism. I propose a multilayered regulation of mutually excusive *var* gene transcription incorporating epigenetic marks, *cis*-acting elements, and a specialized subnuclear compartment. The nuclear periphery represents a heterochromatic environment where associated loci including silenced *var* genes are marked with H3K9me3 and HP1 (Fig. 1B). Activation of a previously silenced *var* locus involves a repositioning into a peripheral euchromatic compartment. Consequently, loci residing within this compartment are marked with H3K4me2, H3K4me3 and H3K9ac. Expression of a singular *var* gene involves interaction with the VES and removal of the MIF. The VES interacts with a *var*-specific sequence located in the upstream sequence and/or directly with the MEE (Fig. 1A).

However, the question remains how the parasites ensure the singularity of the *var* expression site VES in a nucleus. One possibility would be the targeted assembly of a single VES on a specific DNA element that is unique in the genome (depicted in pink), similar to the H enhancer sequence in odorant receptors (Lomvardas et al, 2006). However, my attempts to identify such a putative DNA element were unsuccessful due to an unsuited experimental design. But I am still convinced this idea is worth pursuing using a slightly modified approach, i.e. integration of a silenced *var* promoter into the previously chosen *rh3* locus.

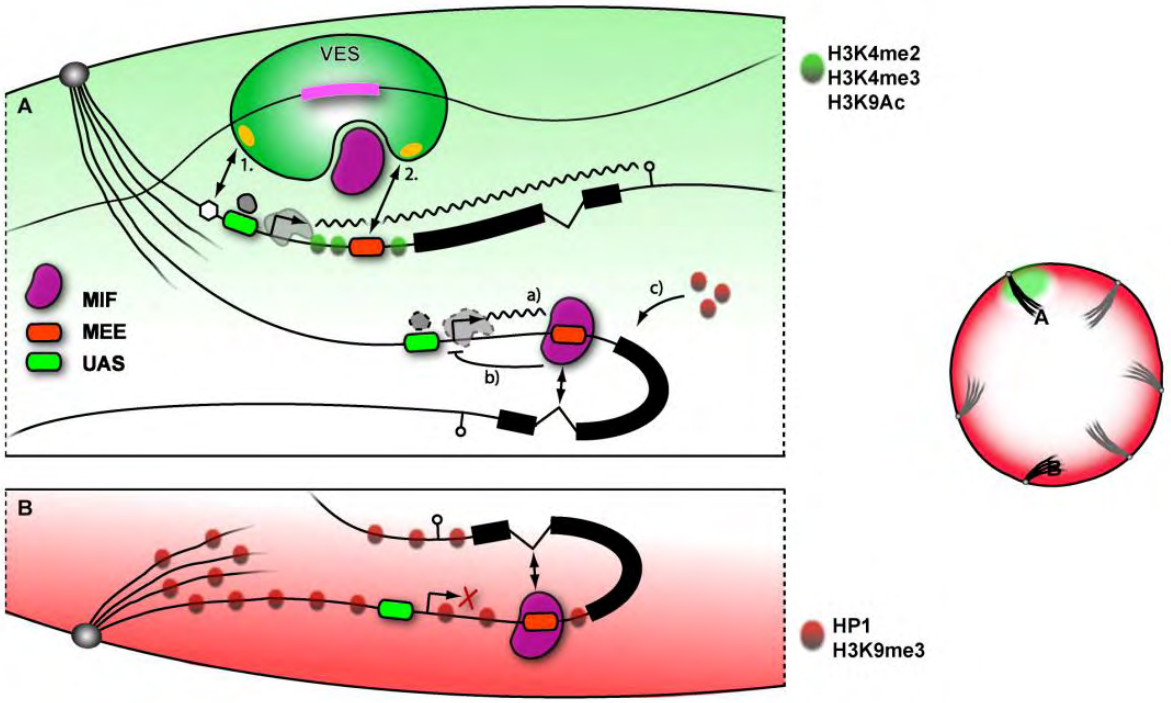


Figure 1. Proposed model of mutually exclusive *var* gene transcription in the apicomplexan parasite *Plasmodium falciparum*. (A) transcription of a single active *var* gene. The active *var* gene locus interacts with the VES that ensures its singularity by interaction with a unique DNA sequence (pink). The VES might unbind the MIF from the MEE and may interact with a specific *var* upstream element (1.), or the MEE itself (2.) resulting in transcriptional initiation of the *var* gene promoter. MIF interacts with the intron sequence. Transcription of *var* genes excluded from the VES but close to/in the nuclear peripheral compartment permissive for transcription is avoided by the interaction of MIF by acting as a road block for transcription (a), by acting as a repressor element on transcriptional initiation (b) and/or by recruiting epigenetic factors establishing silencing marks (c). **(B)** *var* genes localized in perinuclear heterochromatin are silenced by epigenetic marks.
MIF, mutual exclusion element interacting factor; MEE, mutual exclusion element; VES, *var* expression site; UAS, upstream activation sequence; HP1, heterochromatin protein 1.

This proposed model of mutually exclusive var transcription is based on integration of our new findings with current knowledge. How can we explain the transcriptional behaviour of non-*var* multigene families in light of a var-specific VES? *rif* and *stevor* genes have been shown to be transcribed irrespectively of their relative location to the active *var* gene (Cabral & Wunderlich, 2009; Sharp et al, 2006). This is in line with our study, where we find *rif* and *stevor* genes to be similarly expressed independent of a knock-down of endogenous *var* gene transcription. Hence, one could speculate that variegated expression of gene families in *P. falciparum* might be dictated primarily by their proximity to the permissive perinuclear zone harbouring the VES combined with a stochastical activation of these members.

Outlook

Disruption of mutually exclusive *var* gene transcription may be a powerful tool to open new avenues for intervention strategies against *P. falciparum* malaria. Although we propose the existence for MIF and VES and a unique DNA sequence, next steps veering towards identification of these putative regulators of mutually exclusive *var* expression are needed to obtain a more detailed picture of the processes involved. There are two possible drug-based strategies that differently destabilize mutually exclusive *var* expression: First, interference with single gene choice would result in parasites displaying their complete PfEMP1 repertoire at once. Secondly, blocking access to the putative VES could cause a *var* knock-down resulting in parasites that do not expose PfEMP1 on their surface.

For example, experiments with the human intestinal parasite *Giardia lamblia* showed that antigenic variation is essential for its survival (Rivero et al, 2010). Similarl to *P. falciparum*, *G. lamblia* undergoes antigenic variation by expressing only one member of the surface protein family VSP (variant-specific surface proteins) (Nash & Keister, 1985). By interfering with single VSP choice Rivero and colleagues generated parasites that expressed more than one member of the VSP family (Rivero et al, 2010). Primary infection with these parasites protected gerbils from subsequent *Giardia* infections. Similarly, forcing *P. falciparum* to reveal its antigenic repertoire of PfEMP1 may decrease its virulence and in particular its ability to cause chronic

infections. However, PfEMP1 variants vary in different isolates (Kyes et al, 1997; Barry et al, 2007; Fowler et al, 2002). Further, the adherence of PfEMP1 variants to host-endothelial cells causes severe disease (Baruch et al, 2002; Kyes et al, 2001; Rogerson & Brown, 1997; Rowe et al, 1997). Thus, an intervention interfering with single *var* gene choice could be used only in combination with anti-malarial drugs. Blocking the putative VES and therefore inducing a PfEMP1 knock-down can reduce sequestration-mediated morbidity and mortality. PfEMP1 knock-down parasites would circulate freely in the blood stream rendering them vulnerable to clearance by the spleen (Bachmann et al, 2009). However, the future will show if we are able to interfere with single *var* gene choice *in vivo* and support anti-malarial therapy. Our findings lead one step forward in prizing the secret of mutually exlclusive *var* gene expression out of *P. falciparum.*

References

Abdel-Latif MS, Dietz K, Issifou S, Kremsner PG & Klinkert M (2003) Antibodies to Plasmodium falciparum rifin proteins are associated with rapid parasite clearance and asymptomatic infections. *Infect. Immun* **71:** 6229-6233

Abdel-Latif MS, Khattab A, Lindenthal C, Kremsner PG & Klinkert M (2002) Recognition of variant Rifin antigens by human antibodies induced during natural Plasmodium falciparum infections. *Infect. Immun* **70:** 7013-7021

Bachmann A, Esser C, Petter M, Predehl S, von Kalckreuth V, Schmiedel S, Bruchhaus I & Tannich E (2009) Absence of erythrocyte sequestration and lack of multicopy gene family expression in Plasmodium falciparum from a splenectomized malaria patient. *PLoS ONE* **4:** e7459

Barry AE, Leliwa-Sytek A, Tavul L, Imrie H, Migot-Nabias F, Brown SM, McVean GAV & Day KP (2007) Population genomics of the immune evasion (var) genes of Plasmodium falciparum. *PLoS Pathog* **3:** e34

Baruch DI, Gormely JA, Ma C, Howard RJ & Pasloske BL (1996) Plasmodium falciparum erythrocyte membrane protein 1 is a parasitized erythrocyte receptor for adherence to CD36, thrombospondin, and intercellular adhesion molecule 1. *Proc. Natl. Acad. Sci. U.S.A* **93:** 3497-3502

Baruch DI, Rogerson SJ & Cooke BM (2002) Asexual blood stages of malaria antigens: cytoadherence. *Chem. Immunol* **80:** 144-162

Baum J, Papenfuss AT, Mair GR, Janse CJ, Vlachou D, Waters AP, Cowman AF, Crabb BS & de Koning-Ward TF (2009) Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res* Available at: http://www.ncbi.nlm.nih.gov/pubmed/19380379 [Accessed May 12, 2009]

Bernstein E & Allis CD (2005) RNA meets chromatin. *Genes Dev* **19:** 1635-1655

Cabral FJ & Wunderlich G (2009) Transcriptional memory and switching in the Plasmodium falciparumrif gene family. *Mol. Biochem. Parasitol* **168:** 186-190

Cheng Q, Cloonan N, Fischer K, Thompson J, Waine G, Lanzer M & Saul A (1998) stevor and rif are Plasmodium falciparum multicopy gene families which potentially encode variant antigens. *Mol Biochem Parasitol* **97:** 161-76

Chess A, Simon I, Cedar H & Axel R (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78:** 823-834

Chookajorn T, Dzikowski R, Frank M, Li F, Jiwani AZ, Hartl DL & Deitsch KW (2007) Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A* **104:** 899-902

Deitsch KW, Calderwood MS & Wellems TE (2001) Malaria. Cooperative silencing elements in var genes. *Nature* **412:** 875-876

D'Ombrain MC, Voss TS, Maier AG, Pearce JA, Hansen DS, Cowman AF & Schofield L (2007) Plasmodium falciparum erythrocyte membrane protein-1 specifically suppresses early production of host interferon-gamma. *Cell Host Microbe* **2:** 130-138

Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, Freitas-Junior LH, Scherf A, Crabb BS & Cowman AF (2005) Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in Plasmodium falciparum. *Cell* **121:** 13-24

Dzikowski R & Deitsch KW (2008) Active transcription is required for maintenance of epigenetic memory in the malaria parasite Plasmodium falciparum. *J. Mol. Biol* **382:** 288-297

Dzikowski R, Li F, Amulic B, Eisberg A, Frank M, Patel S, Wellems TE & Deitsch KW (2007) Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep* **8:** 959-965

Eggan K, Baldwin K, Tackett M, Osborne J, Gogos J, Chess A, Axel R & Jaenisch R (2004) Mice cloned from olfactory sensory neurons. *Nature* **428:** 44-49

Epp C, Raskolnikov D & Deitsch KW (2008) A regulatable transgene expression system for cultured Plasmodium falciparum parasites. *Malar J* **7:** 86

Fernandez V, Hommel M, Chen Q, Hagblom P & Wahlgren M (1999) Small, clonally variant antigens expressed on the surface of the Plasmodium falciparum-infected erythrocyte are encoded by the rif gene family and are the target of human immune responses. *J. Exp. Med* **190:** 1393-1404

Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BTF, Moes S, Bozdech Z, Jenoe P,

Stunnenberg HG & Voss TS (2010) A Major Role for the Plasmodium falciparum ApiAP2 Protein PfSIP2 in Chromosome End Biology. *PLoS Pathog* **6**

Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BTF, Ehlgen F, Ralph SA, Cowman AF, Bozdech Z, Stunnenberg HG & Voss TS (2009) Plasmodium falciparum Heterochromatin Protein 1 Marks Genomic Loci Linked to Phenotypic Variation of Exported Virulence Factors. *PLoS Pathog* **5**

Fowler EV, Peters JM, Gatton ML, Chen N & Cheng Q (2002) Genetic diversity of the DBLalpha region in Plasmodium falciparum var genes among Asia-Pacific isolates. *Mol. Biochem. Parasitol* **120:** 117-126

Frank M, Dzikowski R, Costantini D, Amulic B, Berdougo E & Deitsch K (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite Plasmodium falciparum. *J. Biol. Chem* **281:** 9942-9952

Freitas-Junior LH, Bottius E, Pirrit LA, Deitsch KW, Scheidig C, Guinet F, Nehrbass U, Wellems TE & Scherf A (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. *Nature* **407:** 1018-1022

Gardner JP, Pinches RA, Roberts DJ & Newbold CI (1996) Variant antigens and endothelial receptor adhesion in Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **93:** 3503-3508

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M, Nene V, Shallom SJ et al (2002) Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419:** 498-511

Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estraño C & Haldar K (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306:** 1934-1937

Howitt CA, Wilinski D, Llinás M, Templeton TJ, Dzikowski R & Deitsch KW (2009) Clonally variant gene families in Plasmodium falciparum share a common activation factor. *Mol. Microbiol* **73:** 1171-1185

Kyes S, Horrocks P & Newbold C (2001) Antigenic variation at the infected red cell surface in malaria. *Annu. Rev. Microbiol* **55:** 673-707

Kyes S, Taylor H, Craig A, Marsh K & Newbold C (1997) Genomic representation of var gene sequences in Plasmodium falciparum field isolates from different geographic regions. *Mol. Biochem. Parasitol* **87:** 235-238

Kyes SA, Rowe JA, Kriek N & Newbold CI (1999) Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **96:** 9333-9338

Lavazec C, Sanyal S & Templeton TJ (2007) Expression switching in the stevor and Pfmc-2TM superfamilies in Plasmodium falciparum. *Mol Microbiol* **64:** 1621-34

Lavstsen T, Salanti A, Jensen ATR, Arnot DE & Theander TG (2003) Sub-grouping of Plasmodium falciparum 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar. J* **2:** 27

Levine M & Tjian R (2003) Transcription regulation and animal diversity. *Nature* **424:** 147-151

Li J, Ishii T, Feinstein P & Mombaerts P (2004) Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature* **428:** 393-399

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J & Axel R (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* **126:** 403-413

Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R & Scherf A (2007) 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol. Microbiol* **66:** 1296-1305

Lopez-Rubio J, Mancio-Silva L & Scherf A (2009) Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* **5:** 179-190

Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, Szestak T, Chesson J, Wu Y, Hughes K, Coppel RL, Newbold C, Beeson JG, Craig A, Crabb BS & Cowman AF (2008) Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes. *Cell* **134:** 48-61

Marti M, Good RT, Rug M, Knuepfer E & Cowman AF (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306:** 1930-1933

Marty AJ, Thompson JK, Duffy MF, Voss TS, Cowman AF & Crabb BS (2006) Evidence that

Plasmodium falciparum chromosome end clusters are cross-linked by protein and are the sites of both virulence gene silencing and activation. *Mol. Microbiol* **62:** 72-83

Morrison LJ, Marcello L & McCulloch R (2009) Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cellular Microbiology* **11:** 1724–1734

Nash TE & Keister DB (1985) Differences in excretory-secretory products and surface antigens among 19 isolates of Giardia. *J. Infect. Dis* **152:** 1166-1171

Pérez-Toledo K, Rojas-Meza AP, Mancio-Silva L, Hernández-Cuevas NA, Delgadillo DM, Vargas M, Martínez-Calvillo S, Scherf A & Hernandez-Rivas R (2009) Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. *Nucleic Acids Res* **37:** 2596-2606

Ralph SA, Scheidig-Benatar C & Scherf A (2005) Antigenic variation in Plasmodium falciparum is associated with movement of var loci between subnuclear locations. *Proc. Natl. Acad. Sci. U.S.A* **102:** 5414-5419

Rivero FD, Saura A, Prucca CG, Carranza PG, Torri A & Lujan HD (2010) Disruption of antigenic variation is crucial for effective parasite vaccine. *Nat Med* **16:** 551-557

Rogerson SJ & Brown GV (1997) Chondroitin sulphate A as an adherence receptor for Plasmodium falciparum-infected erythrocytes. *Parasitol. Today (Regul. Ed.)* **13:** 70-75

Rowe JA, Moulds JM, Newbold CI & Miller LH (1997) P. falciparum rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature* **388:** 292-295

Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AMG, Cohen AM, Janssen-Megens EM, van de Vegte-Bolmer M, Selzer RR, Iniguez AL, Green RD, Sauerwein RW, Jensen ON & Stunnenberg HG (2009) Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. *Proc. Natl. Acad. Sci. U.S.A* **106:** 9655-9660

Sam-Yellowe TY, Florens L, Johnson JR, Wang T, Drazba JA, Le Roch KG, Zhou Y, Batalov S, Carucci DJ, Winzeler EA & Yates JR (2004) A Plasmodium gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. *Genome Res* **14:** 1052-1059

Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP & Cowman AF (2006) Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* **7:** R12-R12

Scherf A, Hernandez-Rivas R, Buffet P, Bottius E, Benatar C, Pouvelle B, Gysin J & Lanzer M (1998) Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in Plasmodium falciparum. *EMBO J* **17:** 5418-5426

Schreiber N, Khattab A, Petter M, Marks F, Adjei S, Kobbe R, May J & Klinkert M (2008) Expression of Plasmodium falciparum 3D7 STEVOR proteins for evaluation of antibody responses following malaria infections in naïve infants. *Parasitology* **135:** 155-167

Sharp S, Lavstsen T, Fivelman QL, Saeed M, McRobert L, Templeton TJ, Jensen ATR, Baker DA, Theander TG & Sutherland CJ (2006) Programmed transcription of the var gene family, but not of stevor, in Plasmodium falciparum gametocytes. *Eukaryotic Cell* **5:** 1206-1214

Voss TS, Healer J, Marty AJ, Duffy MF, Thompson JK, Beeson JG, Reeder JC, Crabb BS & Cowman AF (2006) A var gene promoter controls allelic exclusion of virulence genes in Plasmodium falciparum malaria. *Nature* **439:** 1004-1008

Voss TS, Tonkin CJ, Marty AJ, Thompson JK, Healer J, Crabb BS & Cowman AF (2007) Alterations in local chromatin environment are involved in silencing and activation of subtelomeric var genes in Plasmodium falciparum. *Mol. Microbiol* **66:** 139-150

Weber JL (1988) Interspersed repetitive DNA from Plasmodium falciparum. *Mol. Biochem. Parasitol* **29:** 117-124

Curriculum Vitae

Personal data

Name	Kathrin Witmer
Citizenship:	Swiss
Date of Birth:	07. 01. 1980
Place of Origin:	Solothurn, Switzerland
Affiliation:	Swiss Tropical and Public Health Institute
	Socinstrasse 57
	4051 Basel
	Switzerland
	Phone: +41 61 284 82 27
	Fax: +41 61 271 86 54
email:	kathrin.witmer@gmail.com

Education

Feb 2007 – March 2011	**PhD in cell biology**
	Swiss Tropical and Public Health Institute, Basel
	"Dissecting the regulatory information in *Plasmodium falciparum* multigene family promoters"
	Supervisor: Dr. Till Voss
	Research advisors: Prof. Dr. Hans-Peter Beck, Prof. Dr. Tim Wolf Gilberger
Oct 2004 - May 2006	**Master thesis in infection biology and epidemiology**
	Swiss Tropical Institute, Basel
	"*var* gene transcription dynamics in *Plasmodium falciparum* 3D7 *in vitro*"
	Supervisor: Prof. Dr. Hans-Peter Beck
Oct 2001 – Oct 2004	Undergraduate studies in Organismic Biology at the University of Basel
	Certificate: Bachelor of Science in Animal and Plant Biology
	Minor: Medicine Biology
Jan 2000	Matura Type B Latin in Solothurn, Switzerland

Oral Presentations

Nov 2010	8. Jahrestreffen der AG Malaria der Paul-Ehrlich-Gesellschaft, Basel, Switzerland "Regulatory linkage of Plasmodium *falciparum multigene* families"
Oct 2008	SSTMP (Swiss Society for Tropical Medicine and Parasitology) PhD Student Meeting, Jongny, Switzerland "Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"
Dec 2007	SSTMP PhD students meeting, Schloss Münchenwiler, Switzerland "Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"

Poster Presentations

Sep 2010	21st Molecular Parasitology Meeting, Woods Hole, MA, USA "Regulatory linkage of *Plasmodium falciparum* multigene families" and "Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"
Sep 2009	20th Molecular Parasitology Meeting, Woods Hole, MA, USA "Using the *Plasmodium falciparum kahrp* upstream region to find *var* gene activation elements"
Nov 2008	6. Jahrestreffen der AG Malaria der Paul-Ehrlich-Gesellschaft, Tübingen, Germany "Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"
Feb 2008	Molecular Approaches to Malaria (MAM) conference, Lorne, Australia "Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"

Attended Conferences

May 2009	5th BioMalPar Conference on the Biology and Pathology of the Malaria Parasite, Heidelberg, Germany
Apr 2007	3rd BioMalPar Conference, Heidelberg, Germany

Awards

Posterprize

Sep 2010, 21st Molecular Parasitology Meeting, Woods Hole, MA, USA

Witmer K and Brancucci NMB, Niederwieser I, Birrer S, Flueck C, Voss TS

"Dissecting the regulatory information in *Plasmodium falciparum var* gene promoters"

Publications

Mutually exclusive transcription of subtelomeric gene families in *P. falciparum* is restricted to *var* genes

Witmer K, Schmid CD, Luah YH, Preiser PR, Bozdech Z,. Voss TS

Manuscript in preparation

Mutually exclusive activation of *Plasmodium falciparum var* genes is mediated by a novel mechanism of singular gene choice

Witmer K and Brancucci NMB, Niederwieser I, Birrer S, Schmid CD, Flueck C, Voss TS

Manuscript in preparation

Organellar proteomics reveals hundreds of novel nuclear proteins in the malaria parasite *Plasmodium falciparum*

Oehring S and Woodcroft B, Moes S, Dietz O, Pulfer A, Wetzel J, Dekiwadia C, Maeser P, Flueck C, Witmer K, Brancucci N, Niederwieser I, Jenoe P, Ralph SA, Voss TS

Manuscript in preparation

A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology

Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BT, Moes S, Bozdech Z, Jenoe P, Stunnenberg HG, Voss TS.

PLoS Pathogens, 2010 Feb 26;6(2)

References

Till Voss
Swiss Tropical and Public Health Institute
Socinstrasse 57
4051 Basel
till.voss@unibas.ch

Hans-Peter Beck
Swiss Tropical and Public Health Institute
Socinstrasse 57
4051 Basel
hans-peter.beck@unibas.ch

During my studies, I attended lectures of the following lecturers:

H. Walser, C. Schönenberger, H. Sigel, C.Körner, D. Senn, J. Stöcklin, P. Oehlhafen, M. Nyffeler, T. Kiefhaber, M. Spiess, H.P.. Hauri, M. Hall, T. Bickle, J. Seelig, A. Seelig, U. Jenal, W. Gehring, V. Schmid, J. Meier, W. Keller, M. Affolter, G. Cornelis, C. Thompson, C. Dehio, U. Aebi, S. Arber, A. Engel, S. Grzesiek, H. Reichert, M. Rüegg, T. Schirmer, T. Schwede, A. Engel, T. Boller, B. Baur, R. Brun, N. Weiss, M. Tanner, H.P. Beck, I. Felger, G. Pluschke, W. Salzburger, B. Fahrenkrog