var gene diversity and their serological recognition by naturally exposed individuals

Inauguraldissertation

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

von
Nicole Ingrid Falk
aus Brombach, Deutschland

Basel, 2008

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel auf Antrag von

Prof. Dr. Hans-Peter Beck, Prof. Dr. Norbert Müller, Prof. Dr. Till Voss

Basel, den 9. Dezember 2008

Prof. Dr. Eberhard Parlow
Dekan

Table of contents

Acknowledgements

First and foremost, I would like to thank my supervisor Hans-Peter Beck for giving me the opportunity to conduct my PhD thesis at the molecular parasitology group at the STI. It was anything else but a straight forward thesis with many negative results, setbacks and changes of the project, but most of the times you somehow managed to make me see things a bit more positive than I thought they were.....even though this feeling didn't last for a long time ☺. I am also very grateful to Ingrid Felger, who always was sort of a second supervisor since my diploma thesis. Ingrid, thank you very much for your all your input and remarks and your innumerable attemtps to erouse lively scientific discussions. And of course a big thanks to both of you – and Leo ☺- for the annual "Chefilaus" events at your place.

Furthermore, I would like to thank Norbert Müller and Till Voss for their rather spontaneous decision of being part of my thesis committee. I really appreciate your efforts!!!

I am specially grateful to my predecessor Mirjam Kästli for all her work and help with the sequencing story, despite working on a completely different subject for the last 3 years. Miri, I really couldn't have done without you and I am really impressed by your numerous inputs and your valuable advice. I really owe you one (probably more ☺)!!!

I would like to say thank you to our former Master students Florian Haas and Christof Grüring who did all the cloning and first expression experiments of the recombinant 3D7-*var4* and FCR3S1.2-*var1* fragments which I inherited for the second part of my thesis. Many thanks also go to Weihong Qi for her bioinformatic input and to Amanda Ross and Nicolas Maire for their help with statistical issues. Thanks for being so patient!

I would like to specially thank Sebastian Rusch who introduced me into the lab work when I started and from whom I learned so much throughout my thesis. Sebi, I am deeply impressed by your knowledge and I really think you are an enrichment for this group. Thank you so much for your advice and your infinite patience.

I would like to thank my fellow-PhD students Sonja Schöpflin and Bianca Plüss for their support during our studies and our PhD theses. We've come a long way together and I honestly hope this won't be the end. Thank you so much for being there for me.

Acknowledgements

Thanks to Caroline Kulangara for digging a bit deeper into Immunology with me, to Claudia List for elp on peptide issues, to Eva-Maria for her advice as a pharmacist, to Annette Gaida, thank God we have a postdoc again ☺, to Dania Müller for trying to help me rescue some *var4* fragments, to Sylvia Steiger for cloning and her excellent documentation style, to Kathrin Witmer for her input for the GeneMapper story, to Cristian Köpfli for never getting tired of being called “Bub” (the original !), and Lukas von Tobel for the all-morning Trämli rides. Thanks to Esther Pachlatko, Melanie Hug, Simone Edelmann and Marie Ballif. Thanks to all immunologists for hosting me and the great Badminton sessions. Special thanks to Marco Tamborrini for helping me so many times with PRISM software and always being the devil’s advocate ☺ and to Therese Ruf for the many personal discussions.

Thanks to the GR group members, Christian Flück, Igor Niederwieser, Sophie Oehring, Oliver Dietz and Nicolas Brancucci for their input in the various labmeetings. Thanks to all the ones who already left the institute but with whom I had the pleasure to work with: Selina Bopp, Shinji Okitsu, Cornelia Spycher, Daniela Schütte, Christian Nsanzabana, Jutta Marfurt, Serej Ley, Francesca Valsangiacomo, Martin Maire and Samuel Züricher. Thanks to Christian Scheurer, who will probably never leave the insitute ☺. Thanks for being like the big brother I never had.

Thanks to the IT people, Simon Schlumpf, Lukas Camenzind, Dominique Forster and Marco Clementi for their support when my computer was (again) doing things I didn’t want it to do... ☺.

Thanks to Yvette Endriss for the many meals and the opportunity to have an “add-on” on my PhD salary.

Mein grösster Dank geht an meine Eltern, Hermann und Erna Falk, sowie an meinen Freund Benjamin Leisinger. Ihr wisst am besten, was die letzten 3 3/4 Jahre für mich bedeutet haben und deshalb danke ich Euch von ganzem Herzen für Euren Beistand, Eure Unterstützung, Eure Ermutigungen und Eure Kraft. Ich liebe Euch.

Zusammenfassung

Plasmodium falciparum verursacht die schwerste Form von menschlicher Malaria und ist jährlich für 1-2 Millionen Todesfälle verantwortlich, wovon die meisten auf Kinder unter 5 Jahren entfallen, die südlich der Sahara leben. Der Schweregrad der Krankheit variiert von asymptomatischen Infektionen, d.h. ohne jegliche Symptome, bis hin zur schweren Malaria mit Organversagen, schwerer Blutarmut und Koma. Die meisten dieser Symptome werden auf das Anheften von infizierten roten Blutkörperchen an Endothelzellen der Kapillargefässe zurückgeführt. Durch Anheften an Wirtsrezeptoren wie CD36, ICAM1 oder CSA verhindert der Parasit, dass die Wirtszelle in der Milz beseitigt wird und verlängert dadurch seine Lebensspanne.

Eine Schlüsselrolle in diesem Prozess kommt dem variablen Oberflächenprotein *Plasmodium falciparum* Erythrozyten Membran Protein 1 (PfEMP1) zu, einem vom Parasiten produzierten Protein, welches zur Oberfläche der roten Blutkörperchen transportiert wird um Zytoadhärenz zu vermitteln. Durch diesen Prozess setzt sich der Parasit aber dem Immunsystem des Wirtes aus, was zur Produktion von spezifischen Antikörpern führt. Um dieser Immunantwort zu entgehen variiert der Parasit dieses Antigen (antigenic variation), in dem er ein anderes Protein der gleichen Familie an der Oberfläche präsentiert. PfEMP1 wird von circa 60 Genen pro haploidem Parasitengenom codiert und wird auf eine sich gegenseitig ausschliessende Art und Weise exprimiert, d.h. nur eines von 60 *var* Genen ist aktiv, während die Aktivität der verbleibende 59 unterdrückt wird. Es gibt Hinweise darauf, dass sich der Schutz vor schwerer Malaria grösstenteils durch die allmähliche Ansammlung von anti-PfEMP1 Antikörpern während der frühen Kindheit entwickelt, da Erwachsene zwar noch immer infiziert werden aber nur sehr selten Symptome einer klinischen oder schweren Malaria aufweisen.

Neuste Beobachtungen deuten darauf hin, dass nicht alle PfEMP1 Proteine gleichermassen virulent sind, sondern, dass nur eine bestimmte Anzahl von *var* Genen dazu führt, dass manche Parasiten eine deutlich schwerere Krankheit verursachen als andere. Zur Entwicklung von Methoden um eine schwere Malaria zu verhindern, müssten nun diese bestimmten *var* Gene identifiziert werden. Bis heute gibt es nur 6 Studien, die das Repertoire von exprimierten *var* Genen in Menschen untersucht haben. Wir haben Proben aus Papua Neu Guinea aus einer Fall-Kontroll-Studie benutzt und das *var* Gen Repertoire mittels reverser Transkription und

anschliessender Klonierung und Sequenzierung analysiert. Wir haben zusätzlich die 3 verschiedenen *var* Gen-Gruppen (upsA, B, und C) ermittelt und die Daten in Bezug auf den Schweregrad der Krankheit der Kinder analysiert.

In der Anzahl gefundener upsB und upsC Transkripte gab es keine signifikanten Unterschiede zwischen Kindern mit asymptomatischer, milder oder schwerer Malaria, wohingegen eine grössere Anzahl von upsA Genen in Kindern mit Symptomen gefunden wurden als in Kindern ohne Symptome. Eine Substitution einer Aminosäure konnte identifiziert werden, die vor allem in Kindern ohne Symptome aber mit sehr hohen Parasitendichten zu finden war. Möglicherweise beeinflusst die Expression dieser Variante die Bindungsaffinität der infizierten roten Blutzelle. Mit phylogenetischen Analysen war es uns jedoch nicht möglich, bestimmte *var* Gene oder *var* Gen Gruppen zu identifizieren, die mit schwerer Malaria assoziiert waren. Um Informationen über die Gruppenzugehörigkeit der jeweiligen DBL1α-Sequenz (upsA, B, oder C) zu erhalten, wurde diese mit dem 3D7 Genom verglichen. Hierbei zeigte sich, dass diese Methode nur für upsA Gene geeignet ist, während 28% der upsB und 62% der upsC *var* Gene der falschen Gruppe zugeordnet wurden. Obwohl 7% der identifizierten Sequenzen in mehr als einem Kind vorkamen, zeigten bioinformatische Analysen, dass das *var* Gen Repertoire in dieser Region von PNG nicht begrenzt ist.

Es wurde bereits gezeigt, dass Parasiten, die eine schwere Malaria verursachen, häufiger durch Antikörper erkannt werden, als solche die nur eine milde Form der Krankheit verursachen. Im zweiten Teil dieser Arbeit galt es Informationen über die Bedeutung/Wichtigkeit bestimmter PfEMP1 Domänen bei der immunologischen Erkennung durch den Wirt zu erlangen. Hierfür wurden 2 repräsentative *var* Gene, die mit schwerer Malaria assoziiert werden, rekombinant in *E. coli* exprimiert und untersucht ob Seren von natürlich exponierten Individuen aus unterschiedlichen geographischen Regionen dieses Antigen erkennen. Synthetische Peptide komplementierten diese ELISA Experimente mit rekombinanten Proteinen, wenn die Expression bestimmter Domänen nicht möglich war. ELISA und Western Blot Analysen konnten 3 rekombinante Fragmente und 2 synthetische Peptide identifizieren, die möglicherweise bei der Produktion von schützenden Antikörpern eine Rolle spielen. Die Anzahl untersuchter Proben war jedoch sehr klein und weitere Untersuchungen sind nötig, um diese Ergebnisse zu bestätigen.

Im dritten Teil dieser Arbeit sollte die Anwendbarkeit des GeneMapper® Analyse-Programms bei der Genotypisierung von *var* Genen untersucht werde. Diese Methode wurde bereits erfolgreich für die Genotypisierung des polymorphen Markergens *msp2* etabliert und da *var* Gene auch einen gewissen Grad an Längenpolymorphismus aufweisen, wurde untersucht, ob diese Technik die bisherige Analyse von *var* Gen-Diversität mittels mühsamen Klonierens und anschliessender Sequenzierung ersetzen kann. Dazu wurden aufgereinigte PCR Produkte der UTR-DBL1α Domänen, die während des ersten Teils dieser Arbeit (Sequenzanalyse) generiert wurden, mit fluoreszenz-markierten, DBL-spezifischen Primern re-amplifiziert und mit GeneMapper® analysiert. Die Ergebnisse wurden dann mit den Daten aus der Sequenzanalyse verglichen. Die Grössenbestimmung mit GeneMapper® war mit einer durchschnittlichen Abweichung von 1 Basenpaar sehr genau und zeigte grosse Übereinstimmung mit den Sequenzierdaten. Des Weiteren wurden mit GeneMapper® 141 Sequenzen detektiert, die durch Sequenzierung nicht identifiziert wurden. Im umgekehrten Fall gab es nur 16 Sequenzen, die mit GeneMapper® nicht detektiert wurden. Es gab allerdings auch einen grossen Anteil an Sequenzen, die mit GeneMapper® nicht unterschieden werden konnten, da deren DBL1α Domänen die gleiche Länge aufwiesen. Trotz dieses Nachteils sind wir der Meinung, dass GeneMapper® die Analyse von exprimierten *var* Genen und deren Dynamik bedeutend vereinfachen könnte.

Summary

Plasmodium falciparum causes the worst form of human malaria and leads to 1-2 million deaths annually, most of them children below the age of 5 living in sub-saharan Africa. Morbidity varies from asymptomatic infections with no symptoms to severe malaria accompanied by organ failure, severe anemia and coma. Most of these clinical presentations are associated with sequestration of infected red blood cells (iRBC) on host endothelium. By attaching the parasitized erythrocyte to host receptors such as CD36, ICAM or CSA the parasite prevents the cell from being cleared by the spleen and therefore prolongs its own survival.

A key protein involved in this process is the variant surface antigen *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) which is a parasite derived protein transported to the RBC surface to mediate cytoadherence. With this process exposes the parasite itself to the host immune system leading to the production of specific antibodies. In order to evade this host immune response the parasite undergoes antigenic variation by switching to another member of the same protein family. PfEMP1 is encoded by approximately 60 *var* genes per haploid genome and is expressed at the surface in a mutually exclusive manner, i.e. only 1 of the 60 proteins is expressed and exposed at any one time whilst the others remain silenced.

Protection against severe malaria is thought to be mediated to a large degree by the piecemeal acquisition of anti-PfEMP1 antibodies during early childhood, since adults still get infected but rarely develop severe malaria symptoms.

Recent observations suggest that not all PfEMP1 proteins expressed by a parasite are equally virulent, but only a subset of distinct *var* genes might render a parasite more pathogenic than parasites expressing different *var* gene variants. To generate potential anti-severe disease interventions members of this particular subset need to be identified. To date, only 6 studies have been published investigating the repertoire of expressed *var* genes *in vivo*. We have further used samples collected in Papua New Guinea from a case control study and analyzed *var* transcripts by RT-PCR followed by cloning and sequencing. We determined the 3 main groups of 5'UTR and analysed the data with respect to the clinical presentation of the children they were collected from.

The detected number of different *var* group B and C transcipts was not significantly different between asymptomatic, mild or severe malaria cases, whereas an increase

of group A *var* genes was observed in symptomatic cases when compared to children without any malaria symptoms. We identified an amino acid substitution mainly occurring in asymptomatic children with high parasitemia that might influence the binding affinity of parasites expressing these variants. However, using phylogenetic analyses we were not able to identify other distinct *var* genes or subsets associated with severe malaria. Blasting DBL1α domains against the 3D7 genome to obtain information on the upstream region was found to be suitable for group A *var* genes only, whereas 28% of group B and 62% of group C sequences were assigned to the wrong subgroup using this method. Even though we observed a 7% sequence overlap, bioinformatic analyses estimated the *var* gene repertoire in this region of PNG to be unlimited.

It has previously been shown, that isolates causing severe disease are recognized more frequently than those causing mild malaria. In the second part of this thesis, we wanted to obtain information on the importance of distinct PfEMP1 domains in the recognition by the host immune system. For that purpose, fragments of 2 representative *var* genes shown to be associated with severe malaria were recombinantly expressed in *E.coli* and analyzed for their recognition by naturally exposed sera of different origin. Analysis of synthetic peptides using the same sera served to complement the results of ELISAs using recombinant proteins if expression of distinct domains was not possible. ELISA and Western blot analysis determined that 3 recombinant fragments and 2 synthetic peptides harbor epitopes that might play a role in the generation of protective antibodies. However, since sample size was small further investigations are required to confirm these findings.

In the third part of this thesis, we tested the usefulness of the GeneMapper® analysis software to genotype *var* genes. It has been successfully established for genotyping the polymorphic marker gene *msp2* and since *var* genes also show some length polymorphism it was investigated whether this technique could replace tedious cloning and sequencing approaches, used so far to dissect *var* gene diversity. Therefore, purified PCR products of UTR-DBL domains generated during the sequence analysis were reamplified with fluorescently labeled DBL-specific primers and analyzed by GeneMapper®. The results were then compared to the sequencing data. GeneMapper® sizing was highly accurate with a mean deviation of 1bp and showed a high consistency with sequencing data. Furthermore, GeneMapper® detected 141 sequences which were not identified with the sequencing approach,

whereas vice verca, this was only the case for 16 sequences. However, a significant proportion of *var* genes could not be distinguished because the analyzed DBL domains were identical in size. Despite this shortcoming, we belive that GeneMapper® would greatly facilitate the analysis of expressed *var* genes and their dynamics.

Abbreviations

AP	Alkaline phosphatase
ATS	Acidic terminal segment
AVG	Average value
BCIP	5-bromo-4-chloro-3-indolyl phosphate
CIDR	Cysteine-rich interdomain region
CSA	Chondroitin sulphate A
DBL	Duffy-binding like domain
DHFR	Dihydrofolate reductase
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacidic acid
ELISA	Enzyme-linked immunosorbent assay
EPT	End point titer
EtOH	Ethanol
g	$9.81 m/s^2$
gDNA	genomic DNA
HA	Hyaluronic acid
HS	Heparan sulfate
ICAM-1	Inter-cellular adhesion molecule 1
IDV	Integrated density value
IgG	Immunoglobulin G
IL	Interleukin
IPTG	Isopropyl-beta-D-thiogalacto-pyranoside
iRBC	Infected RBC
KAHRP	Knob-associated histidine-rich protein
kDa	Kilo Dalton
MC	Maurer's clefts
MHC	Major histocompatibility complex
NBT	Nitroblue Tetrazolium
Ni-NTA	Nickel-nitrilotriacidic acid
NTS	N-terminal segment

OD	Optical density
PAM	Pregnancy-associated malaria
PBS	Phosphate buffered saline
PBS/T	PBS/Tween
PCR	Polymerase chain reaction
PEG	Polyethylene Glycol
PfEMP1	*Plasmodium falciparum* erythrocyte membrane protein 1
PNG	Papua New Guinea
PNP	Para-nitro-phenol
RBC	Red Blood Cell
Rif	Repetitive interspersed family
RT	Reverse Transcription
SDS-PAGE	Sodiumdodecylsulfate-poly-acrylamide gel electrophoresis
Stevor	Subtelomeric variable open reading frame
TB	Teriffic broth
TBE	Tris-Borate-EDTA
TE	Tris-EDTA
TNFα	Tumor necrosis factor α
TNT	Tris-NaCl-Tween
TZA	Tanzania
TSP	Thrombospondin
Ups	Upstream
UTR	Untranslated region
VCAM-1	Vascular cell adhesion molecule-1

Chapter 1
General Introduction

INTRODUCTION:

Malaria is endemic in about 90 countries of the world, half of which are in Africa. Every year, 300-500 million people get infected, with 90% of all cases occurring in Africa. *Plasmodium falciparum* is only one of four *Plasmodium ssp*. infecting humans, but with an estimated 1 to 3 million deaths annually, the most devastating of its kind. Despite extensive research over the last decades, drug resistance is on the advance and an effective vaccine is still lacking.

Malaria and its clinical presentations

Malaria presents with a large spectrum of disease outcome ranging from asymptomatic infections with no clinical symptoms but parasites detectable in the blood to uncomplicated and severe malaria. Uncomplicated malaria symptoms are unspecifc and resemble those of a minor viral infection comprising headache, fatigue, abdominal pains, fever, chills and vomiting.[1] Without treatment the parasite burden continues to increase and severe malaria causing severe anemia, prostration, convulsions and respiratory distress might develop within a few hours. Involvement of the brain might lead to cerebral malaria with impaired consciousness and coma. If untreated, severe malaria is fatal and even if proper treatment is given, the mortality rate lies between 5 and 15%.[2,3] The exact mechanism underlying the development of a severe malaria episode is incompletely understood but host factors such as age, immune status and genetical background, as well as transmission intensity and composition of the circulating parasite population were proposed to be involved.[4] Additionally, in recent years, the hypothesis of excessive release of pro-inflammatory cytokines like TNFα and IL-1 as the primary driving force of disease and death has been on the advance.[5,6]

The course of a malaria infection

Plasmodium falciparum is transmitted by the female *Anopheles* mosquito. Injected sporozoites migrate through the skin and enter the blood circulation. They actively invade liver cells and undergo a first round of asexual multiplication forming thousands of infective merozoites. This phase of the life cycle does not cause any clinical symptoms. However, infected hepatocytes burst and released merozoites commence the blood stage of an infection during which clinical malaria symptoms

may develop. Having entered the RBC merozoites undergo a second round of multiplication and differentiate into ring – trophozoite - and schizont stage, which finally releases another 16-32 daughter merozoites upon rupture. The majority of merozoites, will again infect red blood cells and a new cylce of infection starts. A small proportion of merozoites will undergo an alternative pathway and differentiate into male and female gametocytes which are eventually taken up by a feeding mosquito where sexual reproduction occurs. A further complicated series of differentiation and growth, leads to the production of infective sporozoites in the salivary glands of the mosquito.

On the one hand, red blood cells perfectly qualify as host cells since they do not express MHC molecules on their surface, which renders them immunologically inert. On the other hand, erythrocytes are devoid of any organelles and do not possess a trafficking machinery used by most eukaryotic cells to transport proteins to their final destination. Therefore, in order to facilitate its survival, the parasite needs to set up a new sytem on its own for trafficking soluble and membrane-associated or membrane-integrated proteins.[7,8] Extensions of the parasite's vacuolar membrane, known as the tubulovesicular network, and structures referred to as Maurer's clefts (MC) occur in the RBC cytosol. Several parasite-derived proteins become associated with the erythrocyte cytoskeleton or are inserted into the host cell membrane.[9] On the outer membrane, this results in electron dense structures called "knobs" as observed by electron microscopy. Knobs are mainly composed of the knob-associated histidine rich protein (KAHRP)[10,11], but also *Plasmodium falciparum* erythrocyte membrane protein 1 and 3 (PfEMP1 and 3)[12,13] are part of the knob structure. PfEMP1 is inserted into the erythrocyte membrane, probably anchored via PfEMP3 and KAHRP and attached to the cytoskeleton at spectrin/actin junctions.These parasite-induced modifications increase the rigidity of the cell which would cause the iRBC to be cleared by the spleen. However, the parasite-derived proteins inserted into the surface of the infected erythrocyte membrane enable the parasite to sequester away from the blood circulation, thus preventing spleenic clearance, by binding to different host cell receptors – a phenomenon unique to *P. falciparum* called "cytoadherence".

Cytoadherence

Cytoadherence describes the adhesion of erythrocytes infected with mature stages of *P. falciparum* to other cell types of the host. Infected RBCs can either adhere to endothelial cells[14,15,16] known as sequestration, to uninfected RBCs - a phenotype called rosetting[17,18] or to other iRBCs (auto-agglutination/clumping).[19,20] For the host, this binding ability has severe consequences. Sequestered parasites occlude the thin blood vessels which causes impaired oxygen supply and disfunction of affected organs may occur.[21,22]

Numerous host molecules have been identified which can act as receptors for iRBCs. CD36, thrombospondin (TSP), ICAM-1, VCAM-1, CD31, P- and E-Selectin serve as receptors on endothelial cells.[23,24,25,26,27,28,29,30,31,32] Chondroitin sulfate A (CSA) and hyaluronic acid (HA) have been shown to be involved in placental malaria.[33,34,35] For rosetting, the following molecules have been implicated: IgM[36], heparan sulfate (HS)[37], CR1[38], blood group antigens A and B[39] and FactorD[40]. Via CD36 iRBC can also adhere to monocytes and dendritic cells with effects on phagocytosis and clearance of iRBCs[41] and immunosuppression.[42]

Despite this huge number of host molecules, field studies have demonstrated pronounced differences in the host receptor specificity and extent of cytoadherence. Whereas most isolates can adhere to CD36 and ICAM-1, minimal or no adhesion to E-Selectin, VCAM-1 or CSA has been reported.[43] Therefore, it was hypothesized that cytoadherence to certain of these receptors might be associated with particular clinical syndromes, such as cerebral malaria. A paradigm for such a relationship is malaria in pregnancy. Women living in endemic areas who already established semi-immunity become susceptible again during their first pregnancy with dileterious effects for both mother and child. Parasites sequestered in the placenta specificially adhere to syncitial trophoblasts via CSA or HA. The placenta presents a new niche for circulating parasites and probably selects for parasites expressing ligands with the ability to occupy this new niche. In subsequent pregnancies women acquire specific antibodies blocking this adherence, even if the original parasite which caused disease came from another continent.[44] Thus, parasites causing malaria during pregnancy seem to express *var* genes of a restricted subset which might qualify as future vaccine candidates. The relationship between other receptors and morbidity seems less obvious: some evidence supports a role for ICAM-1 in the

development of severe malaria, in particular cerebral malaria[45,43,46], however, negative association between ICAM-1 and severe disease could also be found[47]. Similar results were obtained when searching for a link between rosette formation and severe disease outcome. While rosetting in The Gambia[48,49], and Kenya[50] could be associated with cerebral malaria, studies conducted in Thailand[51] and Papua New Guinea[52,53] could not confirm this finding. However, there is good evidence that this might be due a CR1 polymorphism occurring at high frequencies in South East Asia.[54] An association between autoagglutination and severe malaria was demonstrated in children from Kenya.[55]

Even though several receptors have been identified for cytoadherence in the human host, from a parasite-point–of-view only few molecules have been proposed to act as ligands for cellular adhesion. The best described molecule implicated in sequestration so far is the *P. falciparum* erythrocyte membrane protein 1 (*Pf*EMP1) having been shown to mediate adhesion to a range of host receptors.[37,38,56,57,58]

Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) and the *var* gene family

PfEMP1 is encoded by the *var* gene family. These genes vary in size, from 6 to 15kb and are extremely divergent in sequence. In 3D7 59 *var* genes per haploid genome were identified[59], scattered throughout the 14 chromosomes of *P. falciparum*. *var* genes can be classified into 3 major groups (*var* group A, B, C)[60] and 2 intermediate groups (B/A and B/C) according to their chromosomal location and their upstream (Ups) 5' untranslated region (UTR).[61] UpsC *var* genes are located in the center of the chromosome. UpsB *var* genes are either subtelomeric and transcribed towards the centromere or chromosome central in tandem arrays with other upsB and UpsC *var* genes. UpsA *var* genes are subtelomerically located and transcribed into the opposite direction, towards the telomeres. Chromosomal location seems to influence transcriptional regulation. Transcription of central *var* genes lasts 4-8 hours longer than transcription of subtelomeric *var* genes.[62] Furthermore, there is evidence of specific *var* groups being involved in disease morbidity. UpsA *var* genes were found to be upregulated in culture after selecting for *var* genes expressed by parasites causing severe malaria.[63] *In vivo*, UpsA and/or UpsB *var* genes, both from subtelomeric regions were associated with severe malaria in children from

Tanzania[64] and Papua New Guinea (PNG)[65], whereas upsC *var* genes were mainly found in asymptomatic children.[65] In 2005, Bull *et al*.[66] proposed a *var* gene classification model based on specific sequence tags with differering numbers of cysteine residues within a specific domain. The distribution of sequence tags was found to be different among the 3 *var* subgroups: whereas in subgroup B and C mainly 4 cysteine residues (cys4) were found, the DBL1α domain of subgroup A mostly contained only 2 cysteines (cys2). In Brazil, the cys2 sequences could be associated with severe non-cerebral malaria.[67] Further evidence for a specific subset of *var* genes being involved in malaria morbidity comes from Mali[68], where children with cerebral malaria predominantly expressed *var* genes with DBL-1-like domains corresponding to cys2, whereas isolates from children with no symptoms of severe malaria predominantly transcribed *var* genes with DBL-0-like domains which correspond to cys4 type *var* genes.

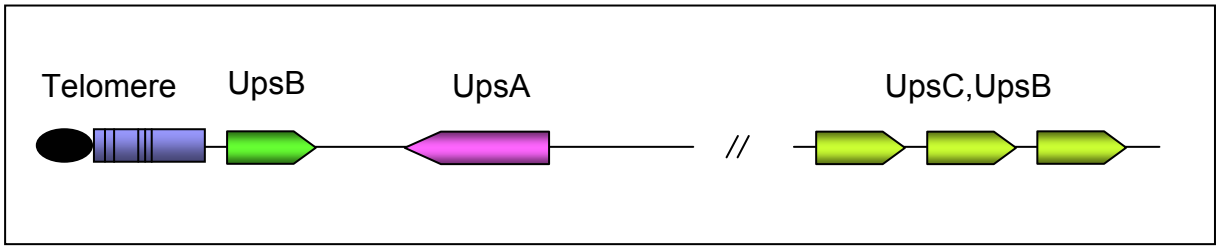


Figure 1. Chromosomal orientation of *var* genes. *var* genes are classified according to their upstream region (UpsABC). Arrows indicate direction of transcription. In blue chromosomal ends (telomeres) are indicated, the black dot represents the telomere associated repreat elements. (Figure adapted from [69])

Despite the huge sequence polymorphism *var* genes share a similar gene organization with 2 exons being separated by an intron. The polymorphic exon1 is building up the extracellular part of the protein whereas the rather conserved exon2 encodes the cytoplasmic acidic terminal segment (ATS), which is thought to anchor the protein to the cytoskeleton of the RBC via binding to proteins of the knob structure.The extracellular domain is highly variable but mainly consists of 4 building blocks: the N-terminal segment (NTS), the Duffy-binding like domain (DBL because of its homology to the Duffy blood group antigen in *P.vivax*), the cysteine-rich interdomain region (CIDR) and C2. On the basis of sequence homology DBL domains can be divided into 6 (α, β, γ, δ, ε and X) and CIDR into 3 (α, β and γ) different sequence types.[70] PfEMP1 proteins have a related protein architecture, but

the sequence, number, location and type of domains differ significantly[60]. Semi-conserved head structures like DBL1α - CIDR1α[13] and DBLδ-CIDRβ make up a protein "prototype" (Figure 2a) found in nearly all small PfEMP1s.[71] By the integration of additional domains the flexibility of large proteins (Figure 2b) may increase and provide a selection advantage, such as the ability to bind to multiple receptors.

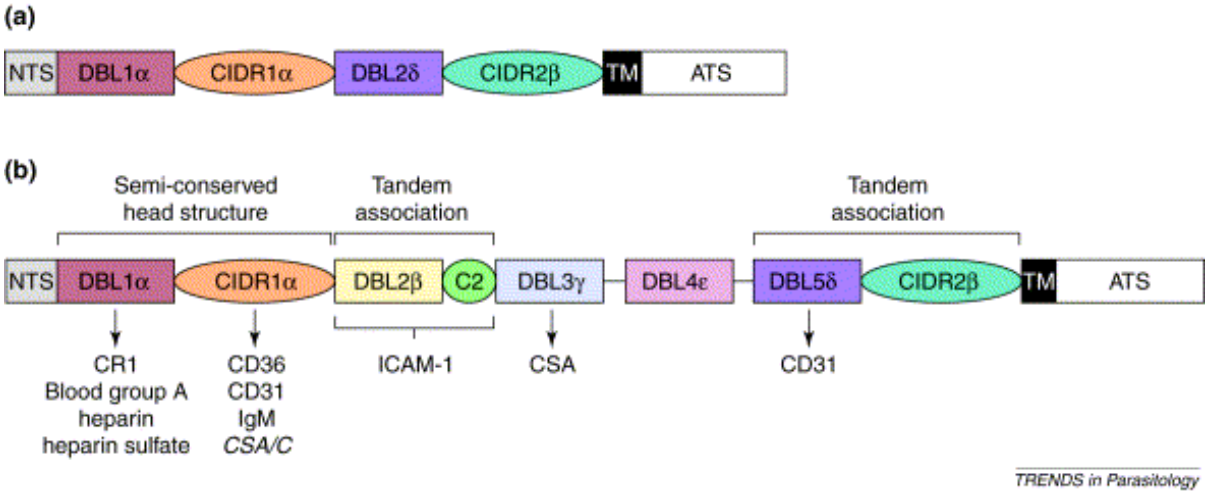


Figure 2. Illustration of PfEMP1 domain architecture.

(a) A small PfEMP1 representative is shown consisting of the NTS domain, the DBL1α-CIDR1α head structure and the ATS domain. In (b) a larger variant is shown harboring additional DBL and CIDR domains, as well as C2 regions. Host cell receptors found to be involved in binding are indicated at their corresponding PfEMP1 domains (Figure from [71]).

Association of host receptors and specific PfEMP1 domains

For a number of host receptors specific binding sites within the various domains have already been mapped (Figure 2b). The most detailed information is available for the interaction of CD36 and the CIDR1α domain. A minimal binding motif of 179aa required for binding could be identified.[21] The rosetting phenotype via CR1 and binding to heparin sulfate is mediated by the DBL1α domain.[72,39] ICAM-1 binding may be achieved by DBL2β domains usually in combination with a downstream C2 element.[73,58,74,75] An adhesion trait to CSA in the placenta has been mapped to CIDR1α[76] and DBL3[57,77], but another study reported about a CSA-binding PfEMP1 lacking this specific DBL domain.[78] Finally, CD31/PECAM binding was shown to be mediated via CIDR1α, DBL2δ and the DBL5δ domain.[31,79]

Antigenic variation and *var* gene switching

As a consequence of placing "alien" proteins or antigens into the membrane of an otherwise immunologically "silent" host cell, the parasite exposes itself to the host immune system which will eventually result in the generation of antibodies against the proteins displayed on the infected RBC surface. These antibodies successfully clear the parasites from the blood circulation resulting in a reduction of parasite density. However, before complete clearance can be achieved, subpopulations of parasites arise that express another variant of PfEMP1 on their surface, a process called antigenic variation. By switching the expressed *var* gene, the parasite changes its antigenic properties and antibodies produced against this protein variant become ineffective. This change in *var* gene expression was correlated with functional changes resulting in altered adhesive phenotypes[80] which in turn might influence the virulence of a parasite isolate during the course of an infection. Therefore, antigenic switching allows the parasite to subvert the host immune response and to eventually establish persistent chronic infections. Information about the sequence of *var* gene switching e.g. whether it occurs randomly or follows a defined order is still missing. However, results by Horrocks *et al.*[81] propose that the switching history might influence the ability to switch to certain *var* genes and the presence of short-lived cross-protective antibodies might lead to the sequential dominance of *var* genes.[82]

The speed of switching *var* genes in *P. falciparum* was investigated *in vitro* in the absence of immune pressure and a switching rate of 2% per generation was calculated.[19] However, in another study much slower rates between 0.25% and 0.025% have been reported[81]. Results from *in vivo* studies are available from laboratory-induced *P. falciparum* infections in naive human volunteers.[83] Switching rates as high as 16% or even higher for the initial switch were proposed, however subsequent switching occurred at much lower rates. Mathematical modeling predicts 0.03%[84] and 18%.[85] Parameters like switching-on and switching-off rates have been proposed[84] and differences in switching rates between isolates distinguish fast- from slow-switching *var* genes[85,86] with different expression patterns during the course of a malaria infection. In a previous publication, these differences were shown to be independent of the 5'UTR of *var* genes[81], however recent reports do find differences that correlate with chromosomal location.[87]

Besides PfEMP1 four other multigene families undergoing antigenic variation have been described namely the *rif* [88], *stevor*[89], *surf* [90] and the *Pfmc-2TM* family.[91] All of these genes are located subtelomerically in close proximity to the *var* gene family. The *rif* (repetitive interspersed family) gene family comprises about 200 genes encoding the variant RIFIN proteins of 30-40kDa. They are expressed at trophozoite stage and can be targeted by the host immune system.[92] Cotransport with PfEMP1 via MCs to the iRBC surface have been reported.[93] They were implicated to act as cofactors in rosetting[88] and binding to CD31[92], however their actual function is yet unclear. Highly similar to *rifs* are the members of the *stevor* (subtelomeric variable open reading frame) family. Due to their high degree of polymorphism they were used as genotyping tools to distinguish *P. falciparum* isolates.[94,95] The *Stevor* multigene family consist of 30-40 genes depending on the parasite isolate. They are expressed in gametocytes and sporozoites[96] as well as in trophozoites and schizont stage[97] and therefore are thought to have multiple functions. They were found to be located in MC and trafficked to the iRBC membrane at schizont stage.[98] Recently, they were shown to be expressed at the apical end of merozoites[98] and they might be associated with erythrocyte invasion.[99] The third multigene family is the recently described *Pfmc-2TM* family comprising 13 gene members.[91] Indirect immunofluorecent studies have shown that Pfmc-2TM localize to Maurer's clefts and are transported to the erythrocyte surface. Complete expression profiles of these proteins are missing but a restricted expression in mid-trophozoite stage has been suggested recently.[100]

In 2005, Winter *et al.*[90] identified another class of polymophic proteins. The surface associated interspersed proteins called SURFINs are encoded by a small family of only 10 *surf* genes, including 3 pseudogenes. They were shown to be contransported to MCs and the iRBC surface with RIFINs and PfEMP1. SURFINs were also found to be part of an amorphous layer attached to the apex of released merozoites.

Information about the function of all these proteins remains elusive.

Var gene regulation

The ability to switch between different *var* genes enables the parasite to prolong its survival in the human host. However, this can only be achieved if not all *var* genes are expressed at the same time. If that was the case, the host immune system would

produce antibodies against all of them at once and eliminate the iRBC immediately. Therefore, the parasite has to be economical with the expression of its *var* genes, in order to not expend the repertoire before transmission is completed. The process used by the parasite to have only one *var* gene active at any one time while the others remain silent is called “mutual exclusion”.

For long time the mechanisms of *var* gene switching and mutual exclusion remained unclear. However, advances in recent years have contributed to shed light on this topic. According to recent publications mutual exclusion seems to be regulated at the level of transcription and it seems to involve multiple layers of control. The interaction between promoters of different regions of *var* genes (UTR and intron)[101,102] and the presence of sterile RNAs[103,104] seem to play important roles in *var* gene silencing. Furthermore, epigenetic mechanisms, such as chromatin - and histone modifications were shown to be involved in the control of *var* gene expression.[105,106,107] Methylation processes of histones were described that “bookmark” certain *var* genes for activation or silencing at the onset of the next cycle, providing “transcriptional memory” for antigenic variation.[108,109] Another level of *var* gene regulation involves the subnuclear localization of *var* genes[105,110,111] with transciptionally active *var* genes in the center of the nucleus and silenced *var* genes at the nuclear periphery in the region of condensed heterochromatin.

Var gene diversity

PfEMP1 proteins have a central role in the biology of *P. falciparum* infections[26,112] and its dual character is of major importance. On the one hand it mediates cytoadherence to various host receptors in various organs, which is a critical process for the survival and transmission of the parasite, but also the cause of the high virulence and severe disease outcome. On the other hand PfEMP1 undergoes antigenic variation to evade the host immune system which is the underlying basis for the establishment of chronic infections. With 50-60 *var* genes per haploid genome the *var* gene repertoire is huge. Recombination and gene shuffeling events during meiosis[113,114] and probably mitosis generate additional diversity. The telomeric location of *var* genes further facilitates recombination. Clustering of telomers into “bouquets” has been suggested as a potential mechanism for bringing *var* genes into close proximity for recombination to generate diversity within the family.[113] Global

sequence comparisons have reveiled that the diversity of *var* genes is immense with almost no overlap among repertoires from different geographical regions[115] or between different isolates.[116,66] The overall sequence similarity in the 3D7 genome ranges from 24-98%.[117] The DBL1α which is the most conserved domain of the extracellular part of PfEMP1 shows a sequence similarity of only 25%. Therefore, PfEMP1 as vaccine candidate is considered rather unfeasible, however adhesive functions probably require the restriction of its variability and keep the molecule semi-conserved to a certain degree.

Immunity

Constant exposure to *P. falciparum and* the survival of repeated malaria infections in endemic areas will eventually lead to semi-immunity. While clinical immunity protecting against severe disease is acquired relatively rapidly[118], asymptomatic infection and infections leading to only mild malaria symptoms still occur. The vast diversity of *var* genes is probably the reason why individuals get repeatedly infected and never develop sterile immunity.

Specific antibodies from reconvalescent serum, probably directed against PfEMP1 were shown to agglutinate the infecting parasite strain after acute infection but not other isolates.[119,120] These variant specific antibodies were strongly implicated in the acquisition of protective immunity.[120,121,122] However other studies also show the existence of cross-reactive antibodies suggesting the presence of conserved epitopes.[123,124,125,126,127,128] If the development of cross-reactive antibodies confers protection as suggested by Marsh&Howard[129] PfEMP1 might be considered a potential vaccine candidate. If, however most antibodies are variant specific as suggested by Newbold *et al*. [130], then the situation is more challenging due to the tremendous sequence diversity. The finding of isolates causing severe malaria being more frequently recognized than those that cause mild disease[131] and studies showing that isolates causing severe malaria are antigenically less diverse[132] imply the existence of a restricted subset of *var* genes. By identifying representatives of this subset and including them into a multicomponent vaccine candidate, the creation of an anti-severe disease intervention might be feasible.

Aims and objectives

1. Analysis of *Plasmodium falciparum var* genes expressed in children from Papua New Guinea.

Specific objectives of this topic:

- To investigate the overall diversity of expressed *var* genes in children from Papua New Guinea.
- To identify differences in diversity among children with different clinical presentations.
- To test whether disease outcome was related to *var* genes determined by their upstream sequences
- To identify potential subgroups of *var* genes associated with severe malaria.

2. Identification of immunodominant epitopes of PfEMP1.

Specific objectives of this topic:

- To obtain information about the importance of distinct PfEMP1 domains in the recognition by the host immune system.
- To indentify domains recognized by semi-immune adults but not by children which might be implicated in the generatin of protective antibodies.

3. Application of capillary electrophoresis sizing technique as new *var* gene genotyping tool.

Specific objectives of this topic:

- To investigate whether GeneMapper could replace approaches of expressed *var* gene typing by cloning and sequencing
- To investigate whether GeneMapper could be applied to trace expressed *var* genes in naturally occurring infections over time

Chapter 2.
Analysis of *Plasmodium falciparum var* genes expressed in children from Papua New Guinea

running title: *var* gene expression in children

Analysis of *Plasmodium falciparum var* genes expressed in children from Papua New Guinea

Nicole Falk[1,#], Mirjam Kaestli[1,#,a], Weihong Qi[1,b], Michael Ott[2], Kay Baea[3], Alfred Cortés[3,c], Hans-Peter Beck[1,*]

Affiliation of authors:

[1] Swiss Tropical Institute, Socinstrasse 57, 4002 Basel, Switzerland

[2] F. Hoffmann-La Roche AG, Basel, Switzerland

[3] PNG Institute of Medical Research, PO Box 378, Madang, MP 511, Papua New Guinea

[#] These authors contributed equally

Present address:

[a] Menzies School of Health Research, Emerging and Infectious Diseases Division, Charles Darwin University, Darwin, Australia

[b] Functional Genomics Center Zurich, Uni/ETH Zurich, Zurich, Switzerland

[c] ICREA and Institute for Research in Biomedicine, Cell & Developmental Biology, Barcelona, Spain

[*] corresponding author

Word count manuscript: 3423

Word count abstract:200

Footnotes:

The authors declare not conflict of interest

This project obtained funding from the Swiss National Science Foundation (Grant number 3100A0-104043/1) and the Optimus Foundation.

This material has not been presented at any meeting or conference

Corresponding author: Hans-Peter Beck, Swiss Tropical Institute, Socinstrasse 57, CH-4051 Basel, Switzerland, Tel: +41-61-284 8116, Fax: +41-61-271 8654, E-mail: hans-peter.beck@unibas.ch

Abstract:

Background: The variable antigen PfEMP1 is a major virulence factor in malaria. A large number of *var* genes encode PfEMP1 and we hypothesized that a restricted PfEMP1 repertoire is involved in clinical disease. To test this we conducted a case control study in Papua New Guinea and analysed expressed *var* genes in naturally infected individuals.

Methods: *var* mRNA was isolated from 79 children with asymptomatic, mild, and severe malaria. We prepared cDNA from the upstream region into the DBL1α domain, and picked 20 clones from each reaction for sequencing.

Results: 25% of centrally located *var* genes were shared between children whilst only 5% of subtelomeric genes were shared, indicating a lower diversity in the former group. Linkage between upstream sequences of group B or C *var* genes and DBL1α groups was not observed making determination of the *var* gene group by Blast analysis of DBL1α against 3D7 impossible. Although diversity of *var* genes is vast but varying between sites, we could identify certain amino acid stretches in the DBL1α domain which seem to determine severity of disease.

Conclusion: Despite this vast diversity restricted disease associated *var* genes can be identified and might be used for innovative interventions based on PfEMP1.

Keywords: malaria, PfEMP1, field study, *var* gene expression, phylogenetic trees, case control study, species richness estimation

Introduction

Antigenic variation is an important evasion mechanism associated with sequestration and virulence of *Plasmodium falciparum* malaria. Adherence of parasitized red blood cells to endothelial cells and to non-infected erythrocytes (rosetting) characteristic for *P. falciparum* leads to microvascular obstructions in various organs [1-3]. This is mediated by the variant surface antigen *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1). This protein is encoded by 1 of approximately 60 *var* genes [4] scattered throughout all chromosomes. *var* genes can be classified into 3 major groups (*var* group A, B and C) and 2 intermediate groups (B/A and B/C) according to their chromosomal location and their 5' upstream region [5]. *var* group A and B genes are subtelomerically located in contrast to *var* group C genes which are arranged in the centre of chromosomes. Differences in transcriptional regulation between *var* group B and C have been described [6] and might indicate differences in their function.

PfEMP1 is structured into several distinct domains with an extra-cellular part composed of a variable N-terminal segment (NTS), various Duffy binding-like domains (DBL), and cysteine-rich interdomain regions (CIDR)[5]. These domains have been associated with different binding specificities for host receptors [7-10]. The extracellular part is followed by a transmembrane domain and the conserved intracellular acidic terminal segment (ATS) anchoring the protein to the cytoskeleton [9, 11, 12]. PfEMP1 enables the parasite to avoid splenic clearance by sequestration but consequently is presented to the immune system eliciting an immune response. Therefore, the parasite undergoes antigenic variation by switching expression to alternative PfEMP1s to subvert the immune response. This switch is often accompanied by changes in adhesive properties [13].

Because of antigenic variation PfEMP1 is often not considered a feasible malaria vaccine candidate. However, recent studies have shown that structural similarities exist in the molecules head structure [14], and only few PfEMP1 variants have been associated with certain clinical presentations such as variants binding chondroitin sulfate A (CSA) with pregnancy associated malaria (PAM)[15]. Other PfEMP1 molecules have been associated with more severe disease [16-19], e.g. PfEMP1 molecules with DBL1-like domains lacking 1 or 2 cysteines characteristic for group A and B/A *var* genes [20,21]. Up-regulation of group A and B *var* genes in children with severe malaria was observed in Tanzania but not in a study from Kenya [22, 23]. In Papua New Guinea, upregulation of *var* group B genes in severe malaria was evident, but not of *var* group A [24].

Despite inconsistencies, these results support the notion that a limited number or structures of disease-related *var* genes exist and that their identification would enable the development of anti-severe disease interventions.

To study the association of expressed *var* genes and clinical malaria presentation, we analysed *var* transcripts of parasites from children with severe, mild and asymptomatic malaria of a case-control study in PNG. *var* mRNA was reverse transcribed and PCR amplified, followed by cloning and sequencing. Sequences were classified on amino acid motifs and numbers of cysteine residues in the DBL1α domain as proposed previously [22]. This is one of few studies addressing *var* expression *in vivo* and which adds to understand the clinical relevance of PfEMP1. Using bioinformatic tools we assessed the diversity but also highlight the limits of the current approach to identify specific subsets of expressed *var* genes.

Materials and Methods

Study samples

Samples were collected during a case-control study in Madang in PNG as described [24]. Ethical clearance was obtained from the MRAC of PNG. Blood samples were analysed from 16 children with severe malaria as defined by WHO criteria [25], from 29 mild malaria cases with no further symptoms of another disease, and from 34 parasitemic but asymptomatic, age-matched children.

Isolation of *var* transcripts and cDNA synthesis

var mRNA isolation and reverse transcription (RT) was performed as described [26]. Briefly, parasite RNA was extracted using TRIzol (Invitrogen) following the manufacturer's instructions. RNA was treated twice with RQ1 DNase (Promega). Full-length *var* transcripts were obtained by incubation of RNA with biotinylated oligonucleotides complementary to the ATS domain and captured using streptavidin-linked Dynal beads (Dynal,UK). RT was performed on captured hybrids using Sensiscript reverse transciptase (Qiagen). Sample aliquots without reverse transcriptase served as negative control.

Amplification of DBL1α -domain and var group-specific PCR

Degenerated DBL1α-5' and DBL1α-3' primers were used to amplify DBL1α-domains as described [26] (referred to as DBL1only). Upstream sequences were amplified using *var* group A, B and C-specific forward primers (figure 1) and DBL1α-3' reverse primers resulting in products of approximately 1.4 kb (referred to as 5'UTR-DBL1). PCR conditions were 94°C, 5 min followed by 35 cycles, 30 sec at 95°C, 60 sec at 52°C, and 90 sec at 68 °C. PCR products were purified using the High Pure PCR Purification Kit (Roche).

Cloning and sequencing

PCR products were processed for sequencing as described [26]. Briefly, fragments were ligated into pGEM-T (Promega) or pET vectors for TOPO cloning (Invitrogen) according to the manufacturer's instructions, and transfected into *E. coli* SURE cells (Stratagene). If possible, at least 20 clones were picked per transfection and sequenced on an automated sequencer (Applied Biosystems) using the forward primer T7 (pGEM-T) or M13 (TOPO).

Sequence analysis

Sequences were checked, edited, and analyzed using MT Navigator (ABI, version 1.02b3), DNASTAR (version 4; http://www.dnastar.com), NCBI BLAST (http://www.ncbi.nlm.nih.gov) and CLUSTALW (http://www.ebi.ac.uk).

Mostly, a consensus sequence was represented by several sequences from the same transfection. Sequences were considered identical if they differed by <5 nucleotides. Nucleotide sequences were translated using ExPASY translate tool (http://www.expasy.org) and checked for identity using a 95% identity cut-off.

Sequences were classified based on the number of cysteine residues and specific amino acid motifs (positions of limited variability (PoLV)) within DBL1α domains as proposed by Bull *et al.* [22]. DBL1α sequences were assigned to one of six groups proposed and their distribution was analyzed with regard to *var* groups A, B and C and the clinical status of the child from whom the sequences were isolated.

Multiple sequence alignments were performed in ClustalX (version 1.83; http://www.clustal.org) and unrooted, minimum evolution, phylogenetic trees (based on amino acids, p-distance) were created with 1,000 bootstrap replicates in MEGA (version 3.1; Molecular Evolutionary Genetics Analysis). Multidimensional scaling (MDS) analysis and Analysis of Similarities (ANOSIM) were performed using Primer 6.1.9 (Primer-E Ltd., UK). 1-way ANOSIM, a non-parametric permutation procedure (999 permutations) tests the null hypothesis of no differences between two groups and was based on a Poisson-corrected distance matrix of amino acid sequences, as was MDS, a non-metric multivariate ordination method.

All sequences have been submitted to GenBank with accession numbers EU787517-EU787985.

Statistical analysis

Intercooled Stata (version 8.2) was used for univariate analysis, Fisher's Exact, and Mann-Whitney U tests. All tests were 2-tailed and considered significant if $p<0.05$.

Species richness estimation

For sequence richness determination, sequences from PNG were compared to a sequence set from Tanzania (unpublished), to published sequence data from Mali [20], and from a global sequence collection [27]. DBL1α sequences from all samples were compared against each other with BLASTCLUST [28]. 95% sequence identity were assigned the same sequence type. Number of sequence types depended on how many

samples were taken into account. Perl scripts were developed to simulate the sequential inclusion of all samples in all possible combinations. Accumulation curves were generated by plotting the increase in recovery of new sequences as a function of sampling effort.

Results

Sequence data on 5'UTR-DBL1α was available for 24 asymptomatic, 29 uncomplicated, and 14 severe children. DBL1α-only sequences were available for 28 asymptomatic, 23 mild, and 14 severe malaria cases (Table 1). DBL1α-only sequences were primarily used to analyze diversity of *var* transcripts, while phylogenetic analysis was performed on 5'UTR-DBL1α sequences .

var group distribution in clinical subgroups

On average, we found 6.4 different DBL1α-only sequences per child with no differences between asymptomatic, mild or severe malaria, even when adjusting for numbers of infecting parasite strains per child. *var* group B and C transcripts were found at equal numbers in the three clinical subgroups, but only one *var* group A sequence was observed in asymptomatic malaria cases (Table 1).

27 of 109 (24.8%) group C 5'UTR-DBL1α sequences were detected in several children, but only 11 of 231 (4.8%) group B sequences were shared by several children (Fisher's Exact Test, $p<0.001$). Most *var* group B or C sequences found in several children were shared in asymptomatic and uncomplicated cases, whilst only 19 group A sequences were found once in these children.

Distribution of motifs in DBL1α sequences

The sequences from block D (ARSFADIGDI) to block H (WFEEW)(Figure 1)[9] of all DBL1α sequences were grouped according to the number of cysteines in this region and distinct amino acid motifs called positions of limited variability (PoLV) as previously proposed [22].

Figure 2A shows the distribution of sequence groups 1 to 6 in *var* group A, B or C transcripts. The majority of *var* group B and C transcripts contained sequences of group 4 to 6, with group 4 being the dominant. *var* group A almost exclusively contained group 1 to 3 sequences.

Group 1 to 3 sequences were shown previously to be more likely associated with severe disease [20-22, 29-31]. Although the majority of *var* transcripts in this study belonged to *var* group 4 to 6 (Figure 2B), group 1 to 3 sequences were significantly less common in asymptomatic children than in clinical cases (Fisher's Exact test $p=0.007$)(Figure 2B). A more detailed analysis of transcripts and associations with sequence length and PoLV motifs can be found in the supplement.

Upstream region of DBLonly sequences

The importance of upstream regions for classification of *var* genes has been recognized, but its determination is difficult. To test whether blasting DBL1α sequences against the 3D7 genome might predict the upstream region in field samples we blasted the DBL1α domains of UTR-DBL1α sequences for which we knew the upstream sequence against the 3D7 genome to determine the *var* group in silico.

DBL1α domains in linkage with upsA sequences were correctly determined in all but one case (17/18), whereas 28% (21/74) of group B sequences were assigned wrongly as group C sequences and 62% (30/48) of confirmed group C sequences were wrongly assigned to group B sequences. Overall, this lead to assignments to the wrong *var* group of 37% of sequences (51/140), showing that prediction of *var* groups is impossible using the DBL1α sequence.

Phylogenetic analysis of NTS-DBL1α sequences

A distance-matrix based, radial phylogenetic tree of NTS-DBL1α sequences of obtained *var* transcripts and 3D7 *var* genes as reference shows seven main clusters (Figure 3). Several clusters have an increased proportion of *var* sequences with specific characteristics as described below.

var group A cluster: this was the only cluster clearly separated and transcripts from rosetting parasites were found in this group.

Clusters with commonly expressed *var* genes: common transcripts found in several children mainly grouped into 2 clusters (Figure 3, blue clusters). One cluster comprised group B and C transcripts and three 3D7 *var* genes of group BA or BC, representing *var* genes with complex domain structures including DBL2β-C2. In 2 of 4 children with cerebral malaria the group C sequence s44s36a35C13 was the most abundant after cloning and sequencing. The second cluster contained exclusively 3D7 *var* type B sequences.

var group C cluster: this cluster comprised mainly *var* group C or BC sequences (Figure 3, green cluster). Two closely related sequences (a59C6, a38C5) originated from two children with asymptomatic malaria but with very high parasitemia of 108,000 and 44,000 parasites/μl and low haemoglobin levels. These *var* group C sequences were the dominant *var* transcript in these children and no *var* group A or B transcript was amplified confirming previous quantitative PCR data [24]. These sequences have a distinct DBL1α block A with a cysteine substitution to a tyrosine (Figure 4). Subsequently, this substitution was found in 19 different *var* transcripts of which 14 were *var* group C genes, deriving from 18 children (median age 46 months; 4 severe, 5 mild, 9 asymptomatic cases). These children had a 7

times higher parasitemia (median parasitemia: 64,000 parasites/μl) than children not expressing this *var* variant (median parasitemia: 8,920 parasites/μl) (Mann-Whitney U Test, p=0.067).

Clustering of *var* transcripts of severe and asymptomatic malaria cases

While no clustering of full-length NTS-DBL1α sequences of severe malaria cases was obvious, a multidimensional scaling (MDS) plot showed significant clustering of the DBL1α stretch from block E to F [9] (Figure 5, Figure S4 in supplement) of dominant *var* transcripts of severe malaria cases (1-way Analysis of Similarities (ANOSIM) permutation test, p=0.002). 7 of 12 *var* sequences from severe malaria clearly grouped apart (Figure 5) also evident in a distance tree of this sequence stretch containing all *var* transcripts (supplement Figure S4). This distinct cluster contained *var* transcripts of parasites from 14 children, of which 7 had severe malaria (Fisher's Exact test, p=0.006). *var* transcripts expressed in 3 of 4 children with cerebral malaria grouped in this cluster (Fisher's Exact test, p=0.025).

Species richness determination

To estimate diversity of *var* DBL1α sequences, we simulated species accumulation curves based on four datasets (Figure 6). For each dataset, the number of DBL1α sequence types was plotted against the number of samples studied, and empirical plots were well fitted by a linear function. Different non-parametric estimators of species richness implemented in eco-tool (http://www.eco-tools.net) were applied to all datasets but none was stabilized before reaching the full number of samples.

In all cases, curves did not reach a plateau which indicated that the diversity of DBL1α sequences is vast, and more sampling efforts are needed to capture the complete sequence diversity. The slopes of the curves varied among datasets (Figure 6) reflecting different sequence diversity. Other factors such as the source of sequences (cDNA or genomic DNA) or the number of clones sequenced per sample also contributed to the observed differences.

Discussion

Little information is available on *var* gene expression in naturally infected malaria cases [20-24, 32]. To describe the expression of *var* genes, we conducted a case control study in PNG and analyzed *var* transcripts of parasites isolated from 79 children with asymptomatic, mild, and severe malaria.

The number of different *var* transcripts of group B or C detected by cloning and sequencing was not significantly different between asymptomatic, mild or severe malaria cases which is in concordance with data from Mali [20]. We reported previously a significant quantitative shift by real-time PCR from *var* group B to C transcripts in the same symptomatic and asymptomatic malaria cases [24]. However, quantitative analysis based on cloning and sequencing of PCR products introduces bias through primers, amplification plateaus, and cloning, and cannot be compared directly with quantitative PCR [33]. A combination of quantitative and qualitative information on *var* transcripts provides the most meaningful data.

An increase of *var* group A transcripts in symptomatic malaria cases was observed when compared to asymptomatic malaria. This is in agreement with a study by Bull et al. [22] and subsequent studies [20, 29-31] which showed that DBL1α sequences of cys2 type (groups 1 to 3, mainly group A *var* genes) were mainly found in symptomatic children. However, this shift in distribution of DBL1α groups between clinical presentations was only evident in dominantly expressed sequences.

7% of all DBL1αonly sequences were found in more than one child in a total of 370 different detected sequences. This indicates a larger diversity in DBL1α sequences than observed by Barry and colleagues [27] in the Amele population, a small subpopulation within our study area who identified only 187 different DBL1α sequences. This might indicate a geographic population structure with locally restricted subpopulations of *var* types.

Approximately 25% of all group C sequences were detected more than once in several children whilst only 5% of group B sequences were shared. Because there are approximately three times more group B than group C *var* genes in 3D7, this confirms previous findings of high recombination rates especially between group B genes [30, 34-36]. Phylogenetic analyses also suggested frequent genetic exchange between group B and C genes because they did not cluster separately, and Blast analysis using the DBL1α sequence was unable to determine the respective *var* group.

Only 16 children with severe malaria (including 4 cerebral cases) were admitted to the hospital during the 4 months study period. The low incidence of severe malaria has been attributed to good access to antimalarials, omnipresence of *P. vivax*, and to the genetic background of Papua New Guineans [37]. Therefore, finding the same *var* NTS-DBL1α sequence dominantly transcribed in 2 of 4 children with cerebral malaria was surprising. In a distance tree, this sequence clustered with other frequently detected *var* transcripts and with 3 genes of complex structure of the 3D7 group BA/BC genes. These 3D7 *var* genes contain additional domains such as DBL2β-C2 which previously was found associated with binding to the intercellular adhesion molecule 1 (ICAM-1) and cerebral malaria [38-40]. Another of these genes, PFL0020w, was previously found to be the second most transcribed *var* gene in an artificial *P. falciparum* 3D7 challenge infection in a non-immune human host [41]. PFL0020w has an identical DBL1α sequence (DQ519151) as a *var* transcript of parasites isolated from heart tissue of two fatal paediatric malaria cases in Malawi [32]. Thus, special attention should be paid to *var* genes within this cluster as they might be candidates strongly associated with morbidity of malaria.

While no clustering of NTS-DBL1α sequences of severe malaria cases was obvious, MDS analysis within block E to F of the DBL1α showed significant sequence clustering in a large subset derived from severe malaria cases. This cluster contained *var* transcripts from 3 of 4 children with cerebral malaria, suggesting a restriction associated with severe malaria.
Finding a subset of *var* sequences associated with severe disease and detected in several children agrees with previous studies [16, 18, 42] and suggests the existence of conserved *var* genes that are readily expressed in less immune hosts increasing parasite survival by providing distinct binding abilities to vital organs.

Two older asymptomatic children had unusually high parasitemia (>40,000 parasites/μl) with no accompanying symptoms. Parasites from these children almost exclusively expressed only one *var* group C transcript and previously had shown a strong proportional increase of *var* group C transcripts [24]. These two sequences were closely related to each other and the 3D7 *var* gene PF08_0107 that previously was associated with weak IgG recognition in semi-immune children and with high expression in unselected 3D7 parasites [19]. In these *var* sequences and in several others within a cluster containing mainly *var* group C types a cysteine was replaced by tyrosine. It remains to be confirmed whether this motif is a marker for a distinct *var* type or whether it is involved in immune

recognition or binding. Such replacement of a cysteine providing disulfide bridges with an aromatic tyrosine might lead to changes in the tertiary structure of PfEMP1.

Our finding suggests that *var* group C genes are primarily expressed in older, semi-immune hosts who already developed immunity against those *var* genes that are involved in severe pathogenesis. Unselected cultured parasites showed *var* group C genes highly expressed [own observation, 19, 43] with low switch-off rates [43]. This might explain the exclusive transcription of group C genes in some semi-immune children. It is yet unclear whether high parasite loads in these children are a result of a *var* group C expression. A "hole in the antibody repertoire" [19, 44, 45] due to the rare expression of these antigens or by other strategies to evade protective immune responses could explain this. The fact that these children were asymptomatic despite the high parasitemia suggests that these PfEMP1 molecules could not mediate cytoadherence in vital organs.

In conclusion, we showed that in symptomatic malaria cases a higher proportion of *var* group A or cys2 DBL1α sequences are present, and we provide evidence that a subset of frequently encountered *var* genes with complex DBL1α structure might be associated with more severe forms of malaria. We also described another subset of *var* group C genes which are frequently expressed in older children with asymptomatic high parasitemia. However, we were unable to clearly identify an association of defined *var* gene expression with severe malaria. Apart from technical reasons such as the small number of severe malaria cases and technical challenges inherent to a cloning and sequencing approach from field samples, this lack of an association could also be due to the mixture of parasites found in peripheral blood samples expressing many different *var* genes as previously reported by Montgomery and colleagues [32]. Further research on *var* gene expression in natural infections and in different settings is urgently needed to understand the dynamics of *var* gene expression and the associated pathogenesis. Only then, innovative advanced proteomic studies can be facilitated to identify PfEMP1 molecules that might become targets of new anti-disease interventions.

Acknowledgments

This project has been supported by the Optimus Foundation and the Swiss National Science Foundation (Grant number 3100A0-104043/1). We would like to thank the children of PNG and their parents for participating in the study and we acknowledge support from the staff at the PNG IMR, Madang. We would like to thank Sylvia Steiger, Eric Huber, Mark Gillingwater, and Sonja Keller-Maerki for assistance in cloning and sequencing. We would like to acknowledge the support in sequencing from Dorothee Foernzler of Hoffmann La Roche, Basel. We thank the Basel Computational Biology Center [BC] for computing hardware and software support.

References

1 Marsh K, English M, Crawley J, Peshu N. The pathogenesis of severe malaria in African children. Ann Trop Med Parasitol. **1996;** 90:395-402.

2 Dondorp AM, Pongponratn E, White NJ. Reduced microcirculatory flow in severe *falciparum* malaria: pathophysiology and electron-microscopic pathology. Acta Trop. **2004**; 89:309-17.

3 Dondorp AM, Ince C, Charunwatthana P, et al. Direct in vivo assessment of microcirculatory dysfunction in severe *falciparum* malaria. J Infect Dis. **2008**; 197:79-84.

4 Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature. **2002**; 419:498-511.

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

6 Voss TS, Thompson JK, Waterkeyn J, et al. Genomic distribution and functional characterisation of two distinct and conserved *Plasmodium falciparum var* gene 5' flanking sequences. Mol Biochem Parasitol. **2000**;107:103-15.

7 Baruch DI, Gormely JA, Ma C, Howard RJ, Pasloske BL. *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. **1996**; 93:3497-502.

8 Smith JD, Kyes S, Craig AG, et al. Analysis of adhesive domains from the A4VAR *Plasmodium falciparum* erythrocyte membrane protein-1 identifies a CD36 binding domain. Mol Biochem Parasitol. **1998**; 97:133-48.

9 Smith JD, Subramanian G, Gamain B, Baruch DI, Miller LH. Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. Mol Biochem Parasitol. **2000**; 110:293-310.

10 Bir N, Yazdani SS, Avril M, Layez C, Gysin J, Chitnis CE. Immunogenicity of Duffy binding-like domains that bind chondroitin sulfate A and protection against pregnancy-associated malaria. Infect Immun. **2006**; 74:5955-63.

11 Su XZ, Heatwole VM, Wertheimer SP, et al. The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. Cell. **1995**; 82(1):89-100.

12 Robinson BA, Welch TL, Smith JD. Widespread functional specialization of *Plasmodium falciparum* erythrocyte membrane protein 1 family members to bind CD36 analysed across a parasite genome. Mol Microbiol. **2003**; 47:1265-78.

13 Roberts DJ, Craig AG, Berendt AR, et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature. **1992**; 357:689-92.

14 Rowe JA, Kyes SA. The role of *Plasmodium falciparum var* genes in malaria in pregnancy. Mol Microbiol. **2004**; 53:1011-9.

15 Bull PC, Kortok M, Kai O, et al. *Plasmodium falciparum*-infected erythrocytes: agglutination by diverse Kenyan plasma is associated with severe disease and young host age. J Infect Dis. **2000**; 182:252-9.

16 Klein MM, Gittis AG, Su HP, et al. The cysteine-rich interdomain region from the highly variable *Plasmodium falciparum* erythrocyte membrane protein-1 exhibits a conserved structure. PLoS Pathog. **2008**; 4:e1000147.

17 Nielsen MA, Staalsoe T, Kurtzhals JA, et al. *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity. J Immunol. **2002**; 168:3444-50.

18 Nielsen MA, Vestergaard LS, Lusingu J, et al. Geographical and temporal conservation of antibody recognition of *Plasmodium falciparum* variant surface antigens. Infect Immun. **2004**; 72:3531-5.

19 Jensen AT, Magistrado P, Sharp S, et al. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. J Exp Med. **2004**;199:1179-90.

20 Kyriacou HM, Stone GN, Challis RJ, et al. Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia. Mol Biochem Parasitol. **2006**;150:211-8.

21 Kirchgatter K, Portillo H del A. Association of severe noncerebral *Plasmodium falciparum* malaria in Brazil with expressed PfEMP1 DBL1 alpha sequences lacking cysteine residues. Mol Med. **2002**; 8:16-23.

22 Bull PC, Berriman M, Kyes S, et al. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. PLoS Pathog. **2005**; 1:e26.

23 Rottmann M, Lavstsen T, Mugasa JP, et al. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. Infect Immun. **2006**; 74:3904-11.

24 Kaestli M, Cockburn IA, Cortés A, Baea K, Rowe JA, Beck HP. Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. J Infect Dis. **2006**; 193:1567-74.

25 World Health Organization, Communicable Diseases Cluster. Severe *falciparum* malaria. Trans R Soc Trop Med Hyg. **2000**; 94 Suppl 1:S1-90.

26 Kaestli M, Cortes A, Lagog M, Ott M, Beck HP. Longitudinal assessment of *Plasmodium falciparum var* gene transcription in naturally infected asymptomatic children in Papua New Guinea. J Infect Dis. **2004**; 189:1942-51.

27 Barry AE, Leliwa-Sytek A, Tavul L, et al. Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*. PLoS Pathog. **2007**; 3:e34.

28 Altschul SF., Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol. **1990**; 215:403-10.

29 Trimnell AR, Kraemer SM, Mukherjee S, et al. Global genetic diversity and evolution of *var* genes associated with placental and severe childhood malaria. Mol Biochem Parasitol. **2006**;148:169-80.

30 Kraemer SM, Kyes SA, Aggarwal G, et al. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. BMC Genomics. **2007**; 8:45.

31 Bull PC, Kyes S, Buckee CO, et al. An approach to classifying sequence tags sampled from *Plasmodium falciparum var* genes. Mol Biochem Parasitol. **2007**; 154:98-102.

32 Montgomery J, Mphande FA, Berriman M, et al. Differential *var* gene expression in the organs of patients dying of *falciparum* malaria. Mol Microbiol. **2007**; 65:959-67.

33 Gölnitz U, Albrecht L, Wunderlich G. *Var* transcription profiling of *Plasmodium falciparum* 3D7: assignment of cytoadherent phenotypes to dominant transcripts. Malar J. **2008**; 7:14.

34 Freitas-Junior LH, Bottius E, Pirrit LA, et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. Nature. **2000**; 407:1018-22.

35 Taylor HM, Kyes SA, Harris D, Kriek N, Newbold CI. A study of *var* gene transcription in vitro using universal *var* gene primers. Mol Biochem Parasitol. **2000**; 105:13-23.

36 Frank M, Kirkman L, Costantini D, et al. Frequent recombination events generate diversity within the multi-copy variant antigen gene families of *Plasmodium falciparum*. Int J Parasitol. **2008**; 38:1099-109.

37 Müller I, Bockarie M, Alpers M, Smith T. The epidemiology of malaria in Papua New Guinea. Trends Parasitol. **2003**; 19:253-9.

38 Springer AL, Smith LM, Mackay DQ, Nelson SO, Smith JD. Functional interdependence of the DBLbeta domain and c2 region for binding of the *Plasmodium falciparum* variant antigen to ICAM-1. Mol Biochem Parasitol. **2004**; 137:55-64.

39 Smith JD, Craig AG, Kriek N, et al. Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria. Proc Natl Acad Sci U S A. **2000**; 97:1766-71.

40 Howell DP, Levin EA, Springer AL, et al. Mapping a common interaction site used by *Plasmodium falciparum* Duffy binding-like domains to bind diverse host receptors. Mol Microbiol. **2008**; 67:78-87.

41 Peters J, Fowler E, Gatton M, Chen N, Saul A, Cheng Q. High diversity and rapid changeover of expressed *var* genes during the acute phase of *Plasmodium falciparum* infections in human volunteers. Proc Natl Acad Sci U S A. **2002**; 99:10689-94.

42 Lindenthal C, Kremsner PG, Klinkert MQ. Commonly recognised *Plasmodium falciparum* parasites cause cerebral malaria. Parasitol Res. **2003**; 91:363-8.

43 Frank M, Dzikowski R, Amulic B, Deitsch K. Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol Microbiol. **2007**; 64:1486-98.

44 Bull PC, Lowe BS, Kortok M, Molyneux CS, Newbold CI, Marsh K. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. Nat Med. **1998**; 4:358-60.

45 Ofori MF, Dodoo D, Staalsoe T, et al. Malaria-induced acquisition of antibodies to *Plasmodium falciparum* variant surface antigens. Infect Immun. **2002**; 70:2982-8.

Table 1: Number of detected *var* transcripts grouped according to clinical status (n = number of children) **A**) Number of 5'UTR - DBL1α sequences with number of total sequences, number of different sequences (with full-length DBL1a sequence) and average number of different sequences per child (range) **B)** Identical to A) but only the DBL1α domain was amplified

A)

	Number of total *var* sequences	**Number of different *var* sequences**	**Average number / child (range)**
upsA	133	19	0.4 (0-3)
Asymptomatic	**4**	**1**	**0 (0-1)**
Uncomplicated	**107**	**13**	**0.6 (0-3)**
Severe	**22**	**5**	**0.4 (0-2)**
upsB	607	78	4.3 (0-13)
Asymptomatic	**227**	**28**	**3.2 (0-11)**
Uncomplicated	**252**	**28**	**4.8 (0-13)**
Severe	**128**	**22**	**5.1 (0-9)**
upsC	603	51	2.4 (0-6)
Asymptomatic	**256**	**21**	**2.3 (0-6)**
Uncomplicated	**237**	**22**	**2.3 (0-6)**
Severe	**110**	**8**	**2.8 (0-5)**
Total	1343	148	**7.1 (0-14)**

B)

	Number of total DBL1α sequences	**Number of different DBL1α sequences**	**Average number / child (range)**
Asymptomatic	**482**	**148**	**6.3 (1-14)**
Uncomplicated	**280**	**141**	**6.7 (2-14)**
Severe	**163**	**75**	**6.1 (1-9)**
Total	925	364	**6.4 (1-14)**

Figure legends

Figure 1: *var* gene structure and PCR products generated with *var* group specific 5' UTR forward or DBL1α forward and DBL1α reverse primer.

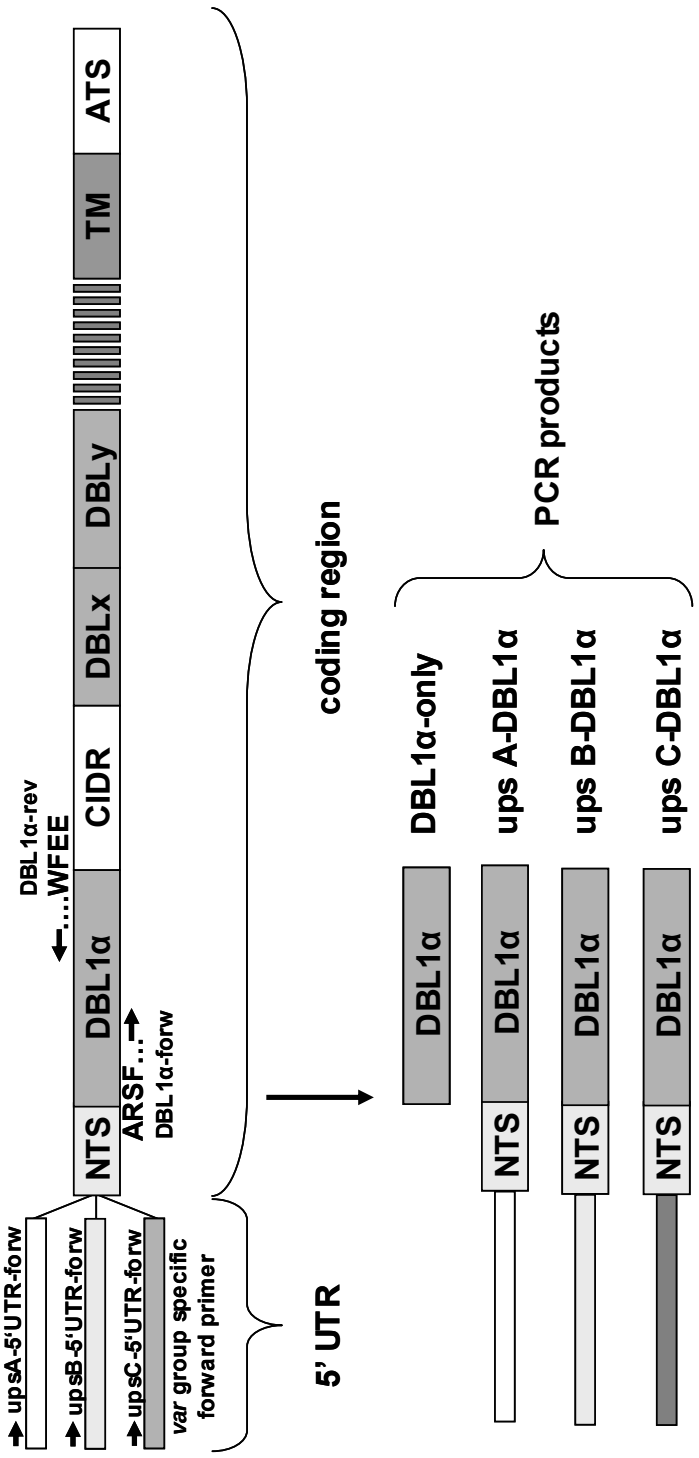
Figure 2: Grouping of different DBL1α sequences into 6 sequence groups according to the classification model of Bull *et al.* [2005]. A) DBL1α sequence groups in *var* group A, B and C. B) DBL1 α sequence groups and clinical status

Figure 3: Unrooted minimum evolution, consensus radial tree (1,000 bootstrap replicates, bootstraps above 50 indicated) of the *var* NTS-DBL1α amino acids stretch of study *var* transcripts with 3D7 *var* genes as reference. "■" *var* sequences detected in several children, "●"sequences with the Y motif in block B of DBL1α, "▲" an identical NTS-DBL1α sequence found in 2 children with cerebral malaria and "♦" *var* sequences of rosetting isolates. Labeling of study sequences: small letters a, u or s mark "asymptomatic", "uncomplicated" or "severe" malaria; the first number refers to age of children in months, big letters A, B or C refer to *var* groups with 2nd number indicating the sequence number. Successive small letters and numbers refer to the same sequence found in several children.

Figure 4: Multiple sequence alignment of DBL1α block A from cysteine 1 to cysteine 3. Selection of study and 3D7 *var* sequences with amino acid differences between groups in bold.

Figure 5: Multidimensional scaling analysis of a distance matrix of *var* sequences between semi-conserved DBL1α blocks E and F from children with severe and asymptomatic malaria. Letters A, B and C indicate *var* groups of these sequences. △ indicates children with asymptomatic malaria, ▼ indicates children with severe malaria. Letters without triangles are sequences derived of children with mild malaria.

Figure 6: Comparison of PfEMP1 DBL1α accumulation curves simulated for different datasets. ★ ★ Mali, ... Tanzania, **.....** PNG, ____ global. The value of x axis represents the number of patient samples or cloned isolates studied. The data set are described in the figure table.



primers	primer sequence	length of product (bp)
upsA - 5'UTR - forw	5'-AACTTACCATAAATTATCATCAAA-3'	1172-1528
upsB - 5'UTR - forw	5'-ATGTAATTGTTGTTTTTTTTTTTGTTAGAATATTTAAA-3'	
upsC - 5'UTR - forw	5'-CACATATA(A/G)TACGACTAAGAAACA-3'	
DBL1α - forw	5'-GCACGAAGTTTTGCAGATAT(A/T)GG-3'	365 - 485
DBL1α - rev	5'-AA(A/G)TCTTC(T/G)GCCCATTCCTCGAACCA-3'	

Figure 1

Figure 2a

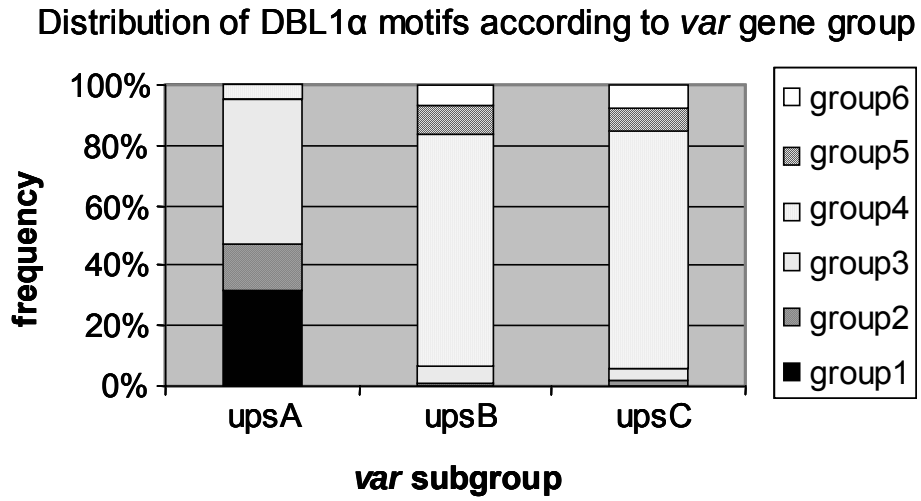
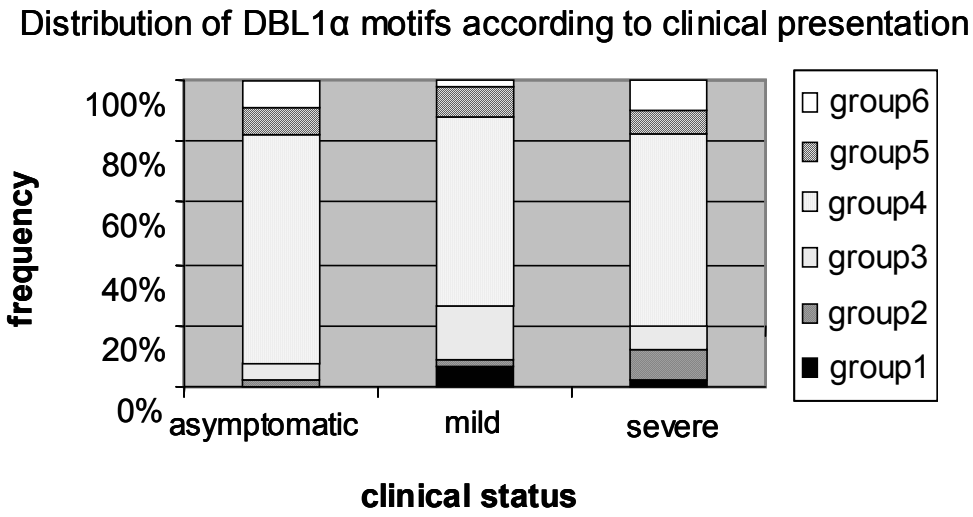


Figure 2b



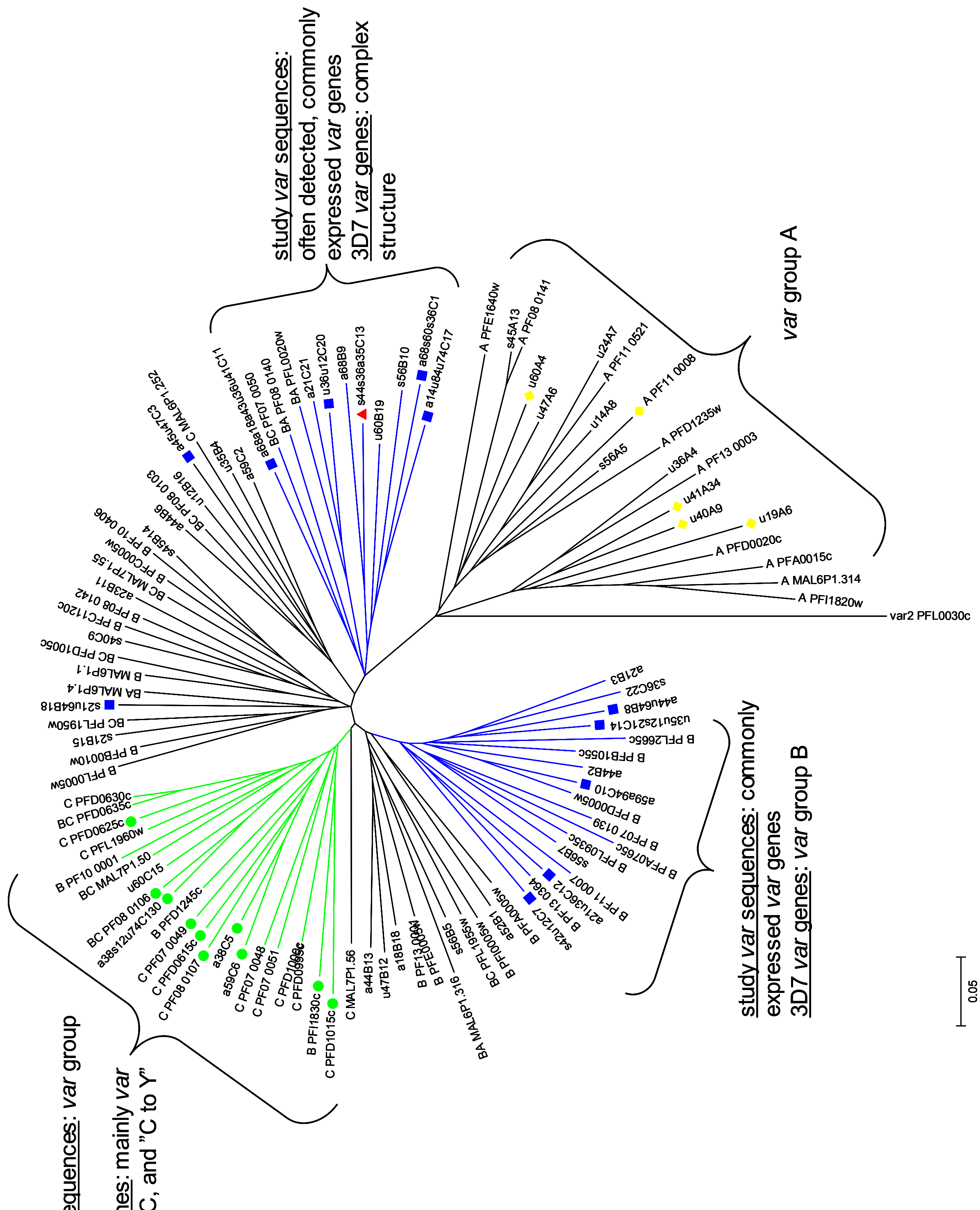

study var sequences: often detected, commonly expressed var genes
3D7 var genes: complex structure
var group A
study var sequences: commonly expressed var genes
3D7 var genes: var group B
study var sequences: var group C, "C to Y"
3D7 var genes: mainly var group C / BC, and "C to Y"
var2 PFL0030c
0.05

Figure 3

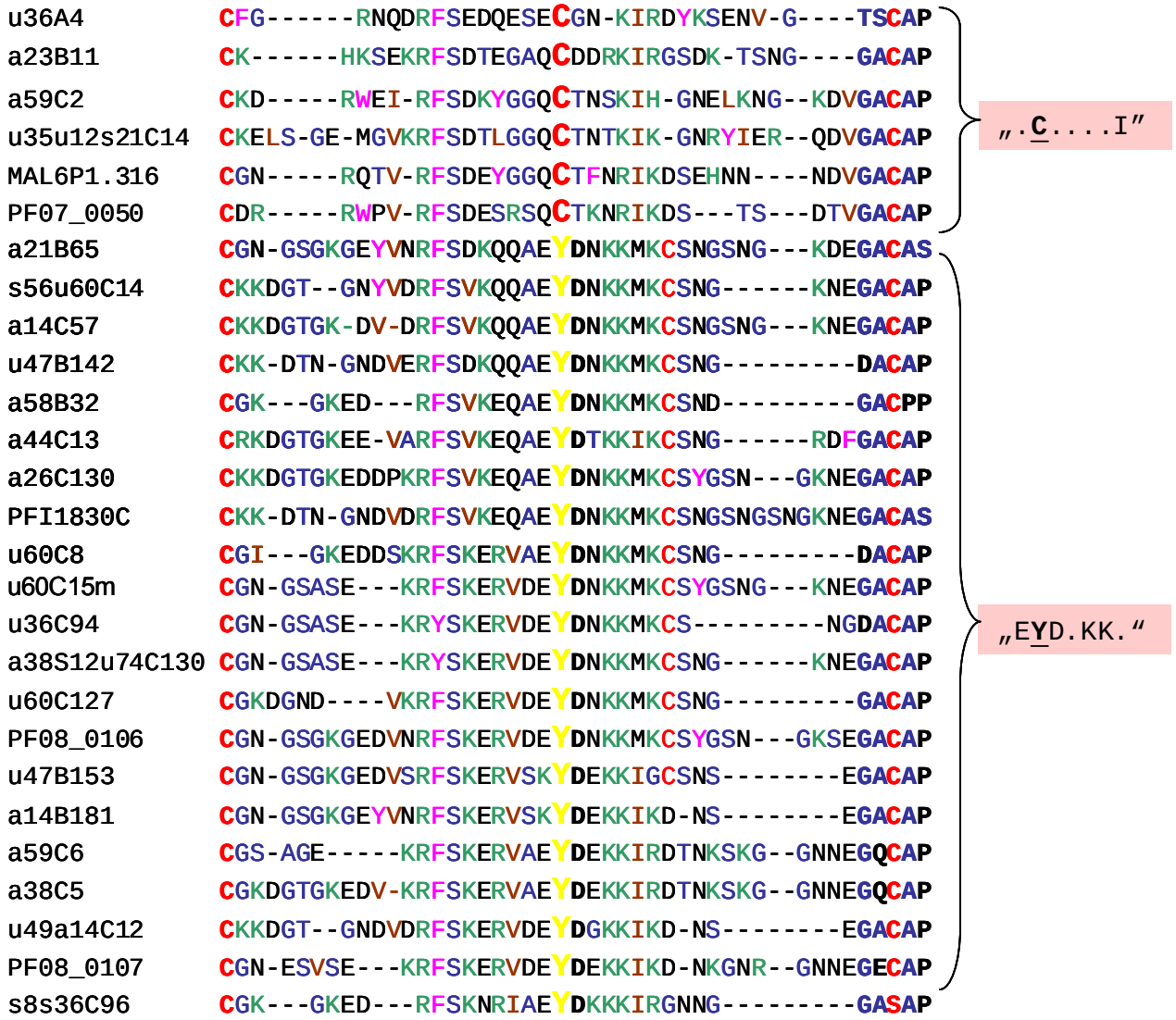

u36A4 CFG------RNQDRFSEDQESECGN-KIRDYKSENV-G----TSCAP
a23B11 CK------HKSEKRFSDTEGAQCDDRKIRGSDK-TSNG----GACAP
a59C2 CKD-----RWEI-RFSDKYGGQCTNSKIH-GNELKNG--KDVGACAP
u35u12s21C14 CKELS-GE-MGVKRFSDTLGGQCTNTKIK-GNRYIER--QDVGACAP
MAL6P1.316 CGN-----RQTV-RFSDEYGGQCTFNRIKDSEHNN----NDVGACAP
PF07_0050 CDR-----RWPV-RFSDESRSQCTKNRIKDS---TS---DTVGACAP
a21B65 CGN-GSGKGEYVNRFSDKQQAEYDNKKMKCSNGSNG---KDEGACAS
s56u60C14 CKKDGT--GNYVDRFSVKQQAEYDNKKMKCSNG------KNEGACAP
a14C57 CKKDGTGK-DV-DRFSVKQQAEYDNKKMKCSNGSNG---KNEGACAP
u47B142 CKK-DTN-GNDVERFSDKQQAEYDNKKMKCSNG---------DACAP
a58B32 CGK---GKED---RFSVKEQAEYDNKKMKCSND---------GACPP
a44C13 CRKDGTGKEE-VARFSVKEQAEYDTKKIKCSNG------RDFGACAP
a26C130 CKKDGTGKEDDPKRFSVKEQAEYDNKKMKCSYGSN---GKNEGACAP
PFI1830C CKK-DTN-GNDVDRFSVKEQAEYDNKKMKCSNGSNGSNGKNEGACAS
u60C8 CGI---GKEDDSKRFSKERVAEYDNKKMKCSNG---------DACAP
u60C15m CGN-GSASE---KRFSKERVDEYDNKKMKCSYGSNG---KNEGACAP
u36C94 CGN-GSASE---KRYSKERVDEYDNKKMKCS---------NGDACAP
a38S12u74C130 CGN-GSASE---KRYSKERVDEYDNKKMKCSNG------KNEGACAP
u60C127 CGKDGND----VKRFSKERVDEYDNKKMKCSNG---------GACAP
PF08_0106 CGN-GSGKGEDVNRFSKERVDEYDNKKMKCSYGSN---GKSEGACAP
u47B153 CGN-GSGKGEDVSRFSKERVSKYDEKKIGCSNS--------EGACAP
a14B181 CGN-GSGKGEYVNRFSKERVSKYDEKKIKD-NS--------EGACAP
a59C6 CGS-AGE-----KRFSKERVAEYDEKKIRDTNKSKG--GNNEGQCAP
a38C5 CGKDGTGKEDV-KRFSKERVAEYDEKKIRDTNKSKG--GNNEGQCAP
u49a14C12 CKKDGT--GNDVDRFSKERVDEYDGKKIKD-NS--------EGACAP
PF08_0107 CGN-ESVSE---KRFSKERVDEYDEKKIKD-NKGNR--GNNEGECAP
s8s36C96 CGK---GKED---RFSKNRIAEYDKKKIRGNNG---------GASAP
„·C....I"
„EYD.KK."

Figure 4

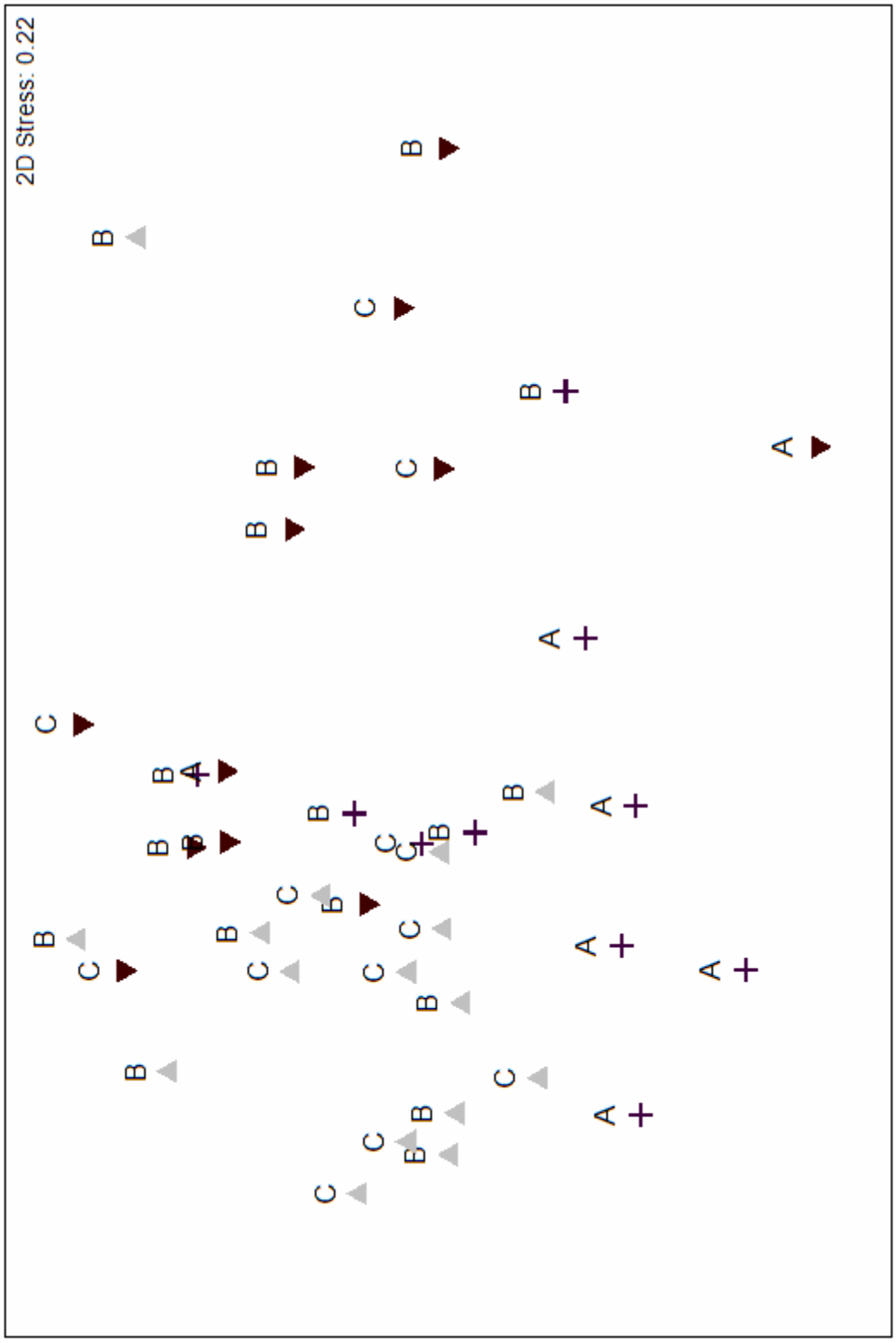


Figure 5

Sample regions	PNG	Tanzania	Mali	Global
Reference	this study	unpublished	Kyriacou HM, 2006	Barry AE, 2007
Sample type	patient samples	patient samples	patient samples	cloned isolates
Multiplicity of infection	1.3	>1	2.4	-
No. samples	61	15	26	25
Mean/median sequence tags per sample	12 (median)	45 (mean)	16 (median)	20 (median)
Type of sequence tags	cDNA	cDNA	cDNA	genomic DNA
Total no. sequence tags	963	668	277	608
Total no. sequence types	327	193	127	538

Figure 6

Supplement:

The distribution of *var* sequences from this study into the 6 sequence groups according to a model proposed by Bull *et al.* [1] was similar to the distribution of *var* sequences from *Plasmodium falciparum* 3D7 strain with four exceptions: one *var* group A sequence was classified as a group 4 sequence, one *var* group C sequence contained a group 2 motif and two *var* group C sequences contained a group 3 motif.

Distinct length distribution of sequences associated with specific PoLV motifs

Bull *et al.* [2] showed an association between sequence length and PoLV motifs, and that the length distribution of the six different groups was similar in sequences originating from three different geographic locations suggesting similar structural features. DBL1α sequences derived from this study were assigned to group 1 to 6 and analyzed for their length distribution. Similar to the results obtained of Bull and colleagues [2] sequences containing the MFK* or the *REY motifs were associated with shorter DBL domains (Figure S1). MFK* and *REY motifs were mutually exclusive, and MFK* was predominantly found in DBL1α domains containing group 1 motifs, but it was also found in three group 6 sequences that lack the second cysteine residue between PoLV3 and PoLV4. When the PoLV distribution of DBL1-only sequences of this study was compared with those occurring in 3D7, there was a high degree of concurrency with most PoLV motifs also found in 3D7 (Figure S2). The distribution of these motifs was also highly similar between the two sets of sequences, which is consistent with data from Kenya [2]. In contrast, PoLV motifs occurring with high frequencies in clinical cases were also found frequently in asymptomatic children (Figure S3).

The length distribution of sequences associated with distinct PoLV motifs was similar to the distribution found in Kenya [2] which might indicate that structural features determine size. Similar to Bull and colleagues [2] we also identified PoLV motifs found in 3D7 in the samples from PNG. However, PoLV motifs frequently occurred both in samples from symptomatic and asymptomatic children which is in contrast to data by Normark et al. [3] who suggested the presence of specific amino acid motifs in certain PfEMP1 DBL sequences predisposing the induction of severe malaria. Motifs associated with rosetting parasites or severe malaria cases were only found at low frequencies and no associations were detected.

Clustering of *var* transcripts of severe versus asymptomatic malaria cases

In a minimum evolution distance tree on a DBL1α stretch from block E to F [4], a group of sequences clustered apart which contained *var* transcripts of parasites of 14 children, of which 7 had severe malaria (Fisher's Exact test, P=0.006)(Figure S4). 3 of 4 children with cerebral malaria expressed *var* transcripts which grouped in this cluster (Fisher's Exact test, P=0.025).

Supplemetary references

1 Bull PC, Berriman M, Kyes S, et al. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. PLoS Pathog. **2005**;1:e26.

2 Bull PC, Kyes S, Buckee CO, et al. An approach to classifying sequence tags sampled from *Plasmodium falciparum var* genes. Mol Biochem Parasitol. **2007**;154:98-102.

3 Normark J, Nilsson D, Ribacke U, et al. PfEMP1-DBL1alpha amino acid motifs in severe disease states of *Plasmodium falciparum* malaria. Proc Natl Acad Sci U S A. **2007**; 104:15835-40.

4 Smith JD, Subramanian G, Gamain B, Baruch DI, Miller LH. Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. Mol Biochem Parasitol. **2000**; 110:293-310.

Figures in supplement:

Figure S1: Length distribution of sequences grouped into group 1 to 6 according to the number of cysteines and PoLV.

Figure S2: Distribution of PoLV motifs in 3D7 and in PNG samples. Sequences not shared between the two data sets are indicated with “OTHER”.

Figure S3: Distribution of PoLV motifs in children with asymptomatic and clinical malaria.

Figure S4: A) Multiple alignment of a sequence stretch between block E and F of DBL1α with dominantly transcribed *var* sequences of parasites of asymptomatic and severe malaria cases. Sequences highlighted in red grouped apart in the MDS analysis which showed a significant clustering of dominant *var* transcripts of severe versus asymptomatic malaria cases. The sequence stretch underlined was used further for phylogenetic analysis (see B). B) Minimum evolution unrooted distance tree with 1,000 bootstrap replicates (bootstraps above 40 are indicated) based on the above underlined sequence stretch using all study *var* sequences. Red circles indicate the 7 *var* sequences of severe malaria cases which grouped apart in the MDS analysis. Yellow triangles indicate further *var* transcripts of severe malaria cases.

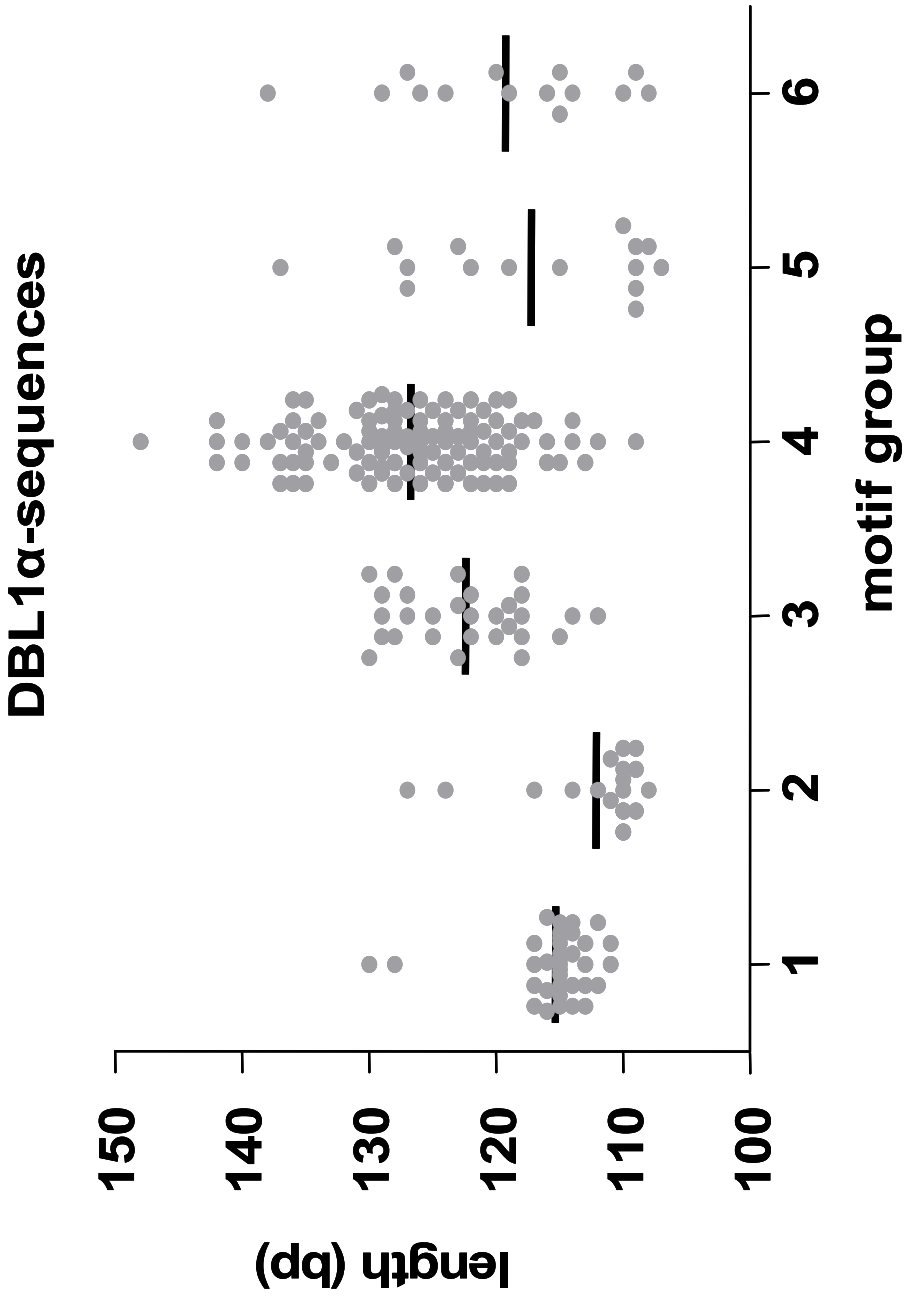

DBL1α-sequences
length (bp)
150
140
130
120
110
100
motif group
1
2
3
4
5
6

Figure S1

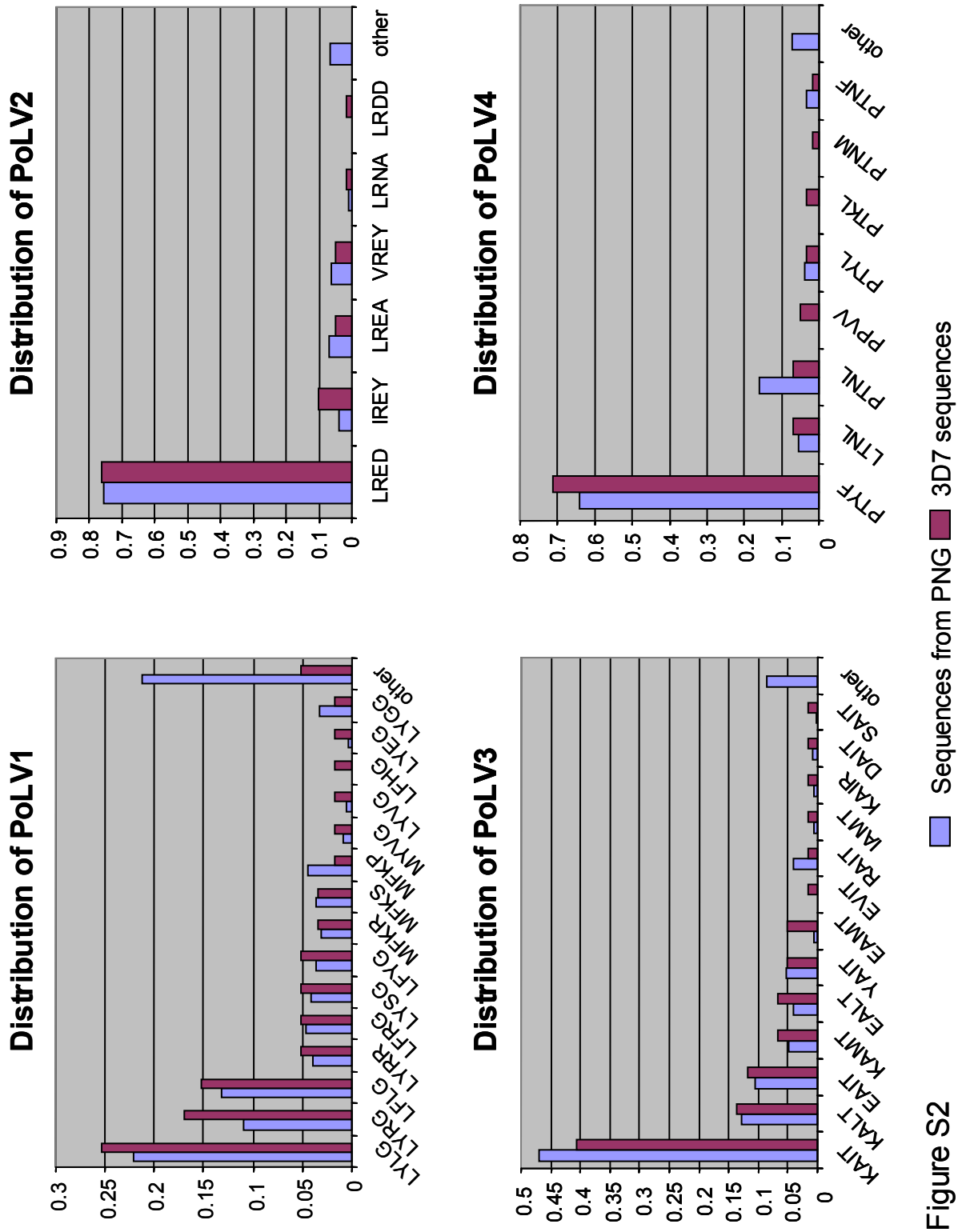

Distribution of PoLV2
LRED IREY LREA VREY LRNA LRDD other
Distribution of PoLV4
PTYF LTNL PTNL PPVV PTYL PTKL PTNM PTNF other
Distribution of PoLV1
LYLG LYRG LFLG LYRR LFRG LYSG LFYG MFKR MFKS MFKP MYVG LYVG LFHG LYEG LYGG other
Distribution of PoLV3
KALT KALT EALT KAMT EALT YALT EAMT EVLT RALT IAMT KAIR DALT SALT other
Sequences from PNG
3D7 sequences

Figure S2

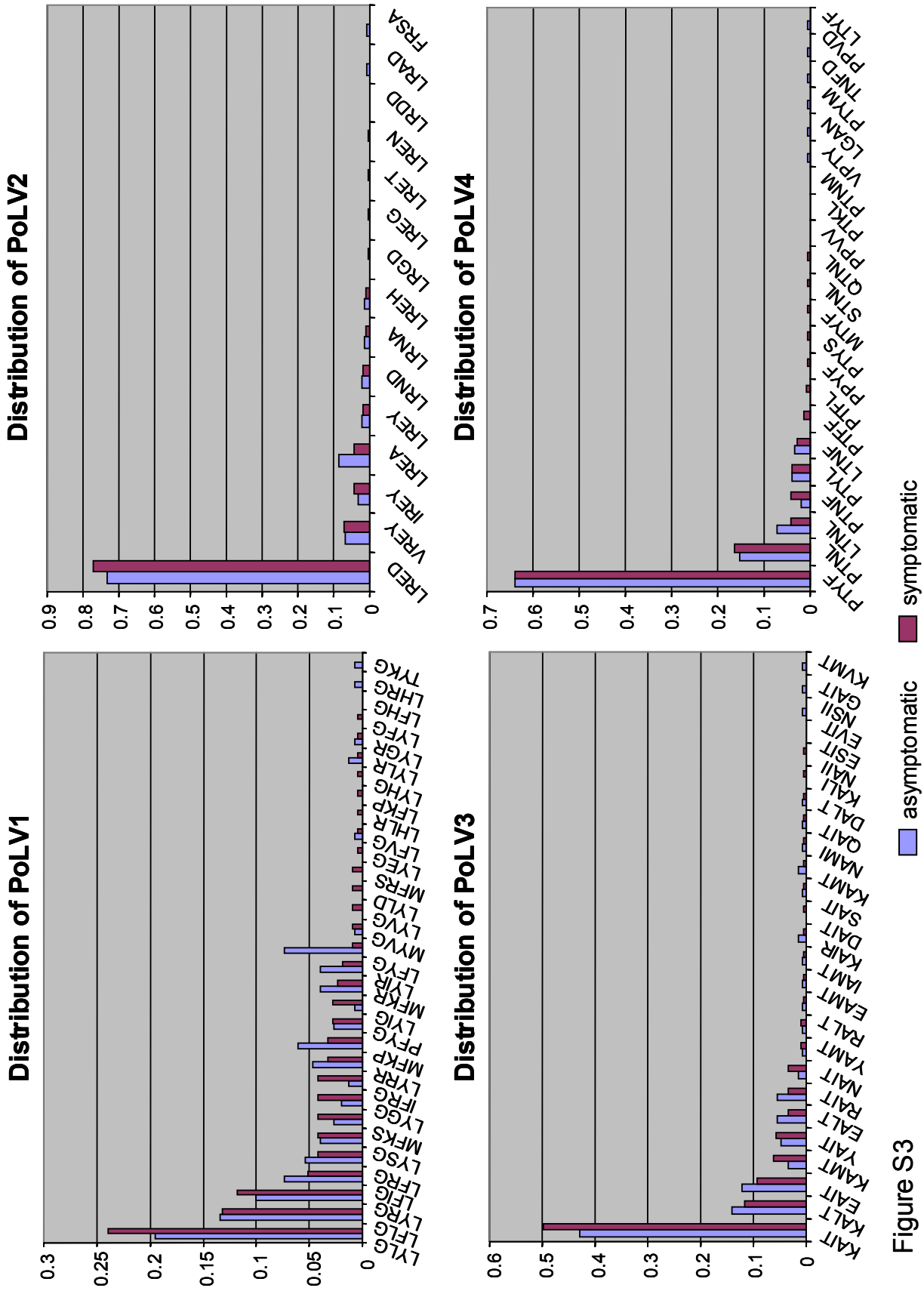

Distribution of PoLV2
Distribution of PoLV4
Distribution of PoLV1
Distribution of PoLV3
asymptomatic
symptomatic

Figure S3

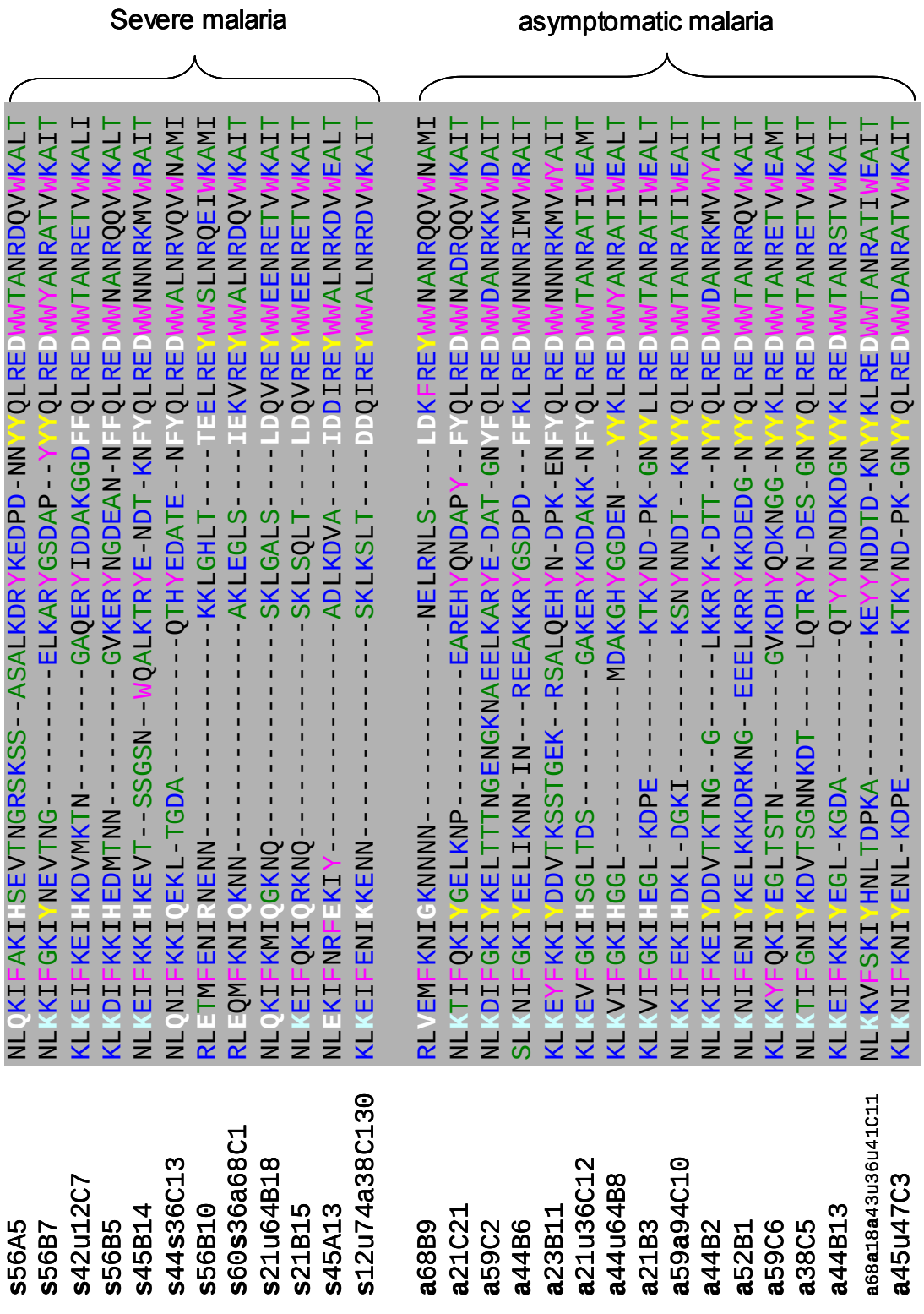

Severe malaria
asymptomatic malaria
s56A5 NLQKIFAKIHSEVTNGRSKSS--ASALKDRYKEDPD-NNYYQLREDWWTANRDQVWKALT
s56B7 NLKKIFGKIYNEVTNG----------ELKARYGSDAP--YYYQLREDWWYANRATVWKAIT
s42u12C7 KLKEIFKEIHKDVMKTN----------GAQERYIDDAKGGDFFQLREDWWTANRETVWKALI
s56B5 KLKDIFKKIHEDMTNN----------GVKERYNGDEAN-NFFQLREDWWNANRQQVWKALT
s45B14 NLKEIFKKIHKEVT--SSGSN-WQALKTRYE-NDT-KNFYQLREDWWNNNRKMVWRAIT
s44s36C13 NLQNIFKKIQEKL-TGDA----------QTHYEDATE--NFYQLREDWWALNRVQVWNAMI
s56B10 RLETMFENIRNENN----------KKLGHLT----TEELREYWWSLNRQEIWKAMI
s60s36a68C1 RLEQMFKNIQKNN----------AKLEGLS----IEKVREYWWALNRDQVWKAIT
s21u64B18 NLQKIFKMIQGKNQ----------SKLGALS----LDQVREYWWEENRETVWKAIT
s21B15 NLKEIFQKIQRKNQ----------SKLSQLT----LDQVREYWWEENRETVWKAIT
s45A13 NLEKIFNRFEKIY----------ADLKDVA----IDDIREYWWALNRKDVWEALT
s12u74a38C130 KLKEIFENIKKENN----------SKLKSLT----DDQIREYWWALNRRDVWKAIT
a68B9 RLVEMFKNIGKNNNN----------NELRNLS----LDKFREYWWNANRQQVWNAMI
a21C21 NLKTIFQKIYGELKNP----------EAREHYQNDAPY--FYQLREDWWNADRQQVWKAIT
a59C2 NLKDIFGKIYKELTTTNGENGKNAEELKARYE-DAT-GNYFQLREDWWDANRKKVWDAIT
a44B6 SLKNIFGKIYEELIKNN-IN--REEAKKRYGSDPD--FFKLREDWWNNNRIMVWRAIT
a23B11 KLKEYFKKIYDDVTKSSTGEK--RSALQEHYN-DPK-ENFYQLREDWWNNNRKMVWYAIT
a21u36C12 KLKEVFGKIHSGLTDS----------GAKERYKDDAKK-NFYQLREDWWTANRATIWEAMT
a44u64B8 KLKVIFGKIHGGL----------MDAKGHYGGDEN--YYKLREDWWYANRATIWEALT
a21B3 KLKVIFGKIHEGL-KDPE----------KTKYND-PK-GNYYLLREDWWTANRATIWEALT
a59a94C10 NLKKIFEKIHDKL-DGKI----------KSNYNNDT--KNYYQLREDWWTANRATIWEAIT
a44B2 NLKKIFKEIYDDVTKTNG--G----LKKRYK-DTT--NYYQLREDWWDANRKMVWYAIT
a52B1 NLKNIFENIYKELKKKDRKNG-EEELKRRYKKDEDG-NYYQLREDWWTANRRQVWKAIT
a59C6 KLKKYFQKIYEGLTSTN----------GVKDHYQDKNGG-NYYKLREDWWTANRETVWEAMT
a38C5 NLKTIFGNIYKDVTSGNNKDT----LQTRYN-DES-GNYYQLREDWWTANRETVWKAIT
a44B13 KLKEIFKKIYEGL-KGDA----------QTYYNDNDKDGNYYKLREDWWTANRSTVWKAIT
a68a18a43u36u41C11 NLKKVFSKIYHNLTDPKA----------KEYYNDDTD-KNYYKLREDWWTANRATIWEAIT
a45u47C3 KLKNIFKNIYENL-KDPE----------KTKYND-PK-GNYYQLREDWWDANRATVWKAIT

Figure S4a

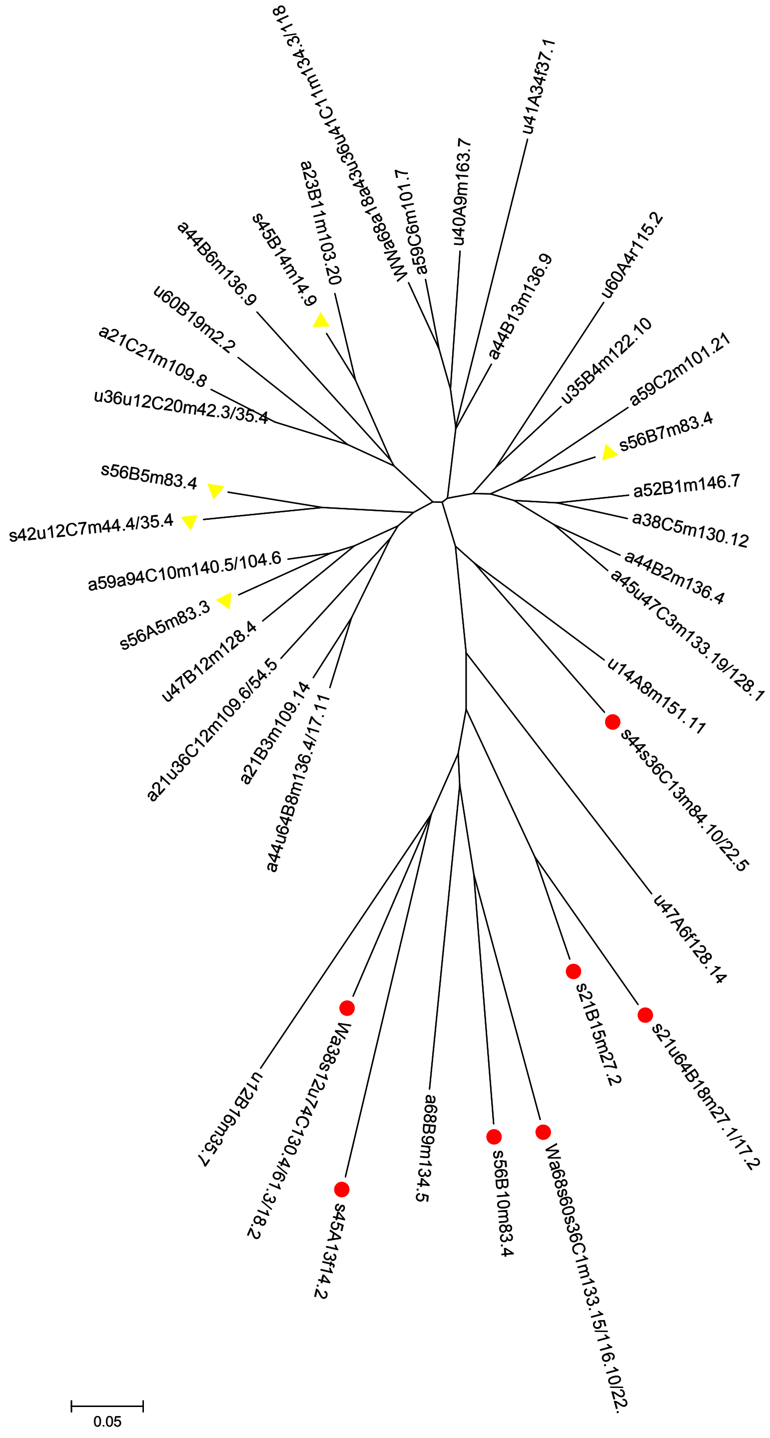

Wa68a184a3c36u41C11m134.3/118
u41A34f37.1
u40A9m163.7
a59C6m101.7
a23B11m103.20
s45B14m14.9
a44B6m136.9
u60B19m2.2
a21C21m109.8
u36u12C20m42.3/35.4
a44B13m136.9
u60A4r115.2
u35B4m122.10
a59C2m101.21
s56B7m83.4
a52B1m146.7
a38C5m130.12
a44B2m136.4
a45u47C3m133.19/128.1
s56B5m83.4
s42u12C7m44.4/35.4
a59a94C10m140.5/104.6
s56A5m83.3
u47B12m128.4
a21u36C12m109.6/54.5
a21B3m109.14
a44u64B8m136.4/17.11
u14A8m151.11
s44s36C13m84.10/22.5
u47A6f128.14
s21B15m27.2
s21u64B18m27.1/17.2
u12B16m35.7
Wa38s12u74C130.4/61.3/18.2
s45A13f14.2
a68B9m134.5
s56B10m83.4
Wa68s60s36C1m133.15/116.10/22.
0.05

Figure 4b

Chapter 3
Identification of immunodominant epitopes of PfEMP1

Introduction

In malaria endemic areas, semi-immunity is acquired after repeated exposure to the malaria parasite. Passive transfer of antibodies from immune to non-immune individuals suggested that immunity is, at least in part, mediated by antibodies.[1,2] Parasite-derived antigens inserted into the RBC membrane are potential targets for this protective immunity since they are exposed for a long period on the erythrocyte surface and serve critical biological functions.[3] Currently, *Plasmodium falciparum* erythrocyte membrane portein 1 (PfEMP1) is the best characterized of these parasite-induced proteins proposed as targets for naturally acquired immunity[4] and the presence of anti–PfEMP1 antibodies has been associated with clinical immunity.[5,6,7,8]

Despite the large *var* gene repertoire coding for PfEMP1, the architecture of this protein is rather conserved. Most PfEMP1 variants are build up by four 4 building blocks: the N-terminal segment (NTS), the Duffy-binding like domain (DBL), the cysteine-rich interdomain region (CIDR) and the more conserved domain C2. Sequence comparisons identified 6 DBL (α, β, γ, δ, ε and X) and 3 CIDR (α, β and γ) subgroups.[9] The number and the location of these domains varies considerably among different PfEMP1 proteins[10], however certain headstructures or tandem repeats, always made up of the same domains, are also oberserved.[11,12] Recently, structural conservation within these head-structures has been reported.[13]

Several studies in different epidemiological settings were performed, investigating a possible link between severe malaria and the expression of particular PfEMP1 protein sequences. Upregulation of upsB *var* genes was found in severe and mild malaria cases in Africa and PNG, whereas upsC *var* genes were upregulated in asymptomatic children.[14,15] UpsA *var* genes have also been shown to be upregulated in the 3D7 laboratory strain after selection for severe malaria.[16] Other studies reported on particular DBL domains predominantly transcribed in patients with severe malaria.[17] The best understood correlation between disease and expressed *var* gene variant is that of pregnancy associated malaria (PAM) and *var2csa* expression.[18,19] After one or two pregnancies transcendent antibodies develop that recognize placental iRBC from different geographical regions and correlate with protection from malaria.[20,21] The binding region of *var2csa* is probably located in the DBL3X domain, since it showed cross-reaction to antibodies raised against the DBL3γ domain of *var1csa*, which block adhesion to CSA.[22] A more recent report showed that the DBL6ε domain might also play a role in protection against PAM.[23]

In 1992, Newbold *et al.*[24] postulated that most anti-PfEMP1 antibodies are variant-specific. However, more recent data also report about the existence of cross-reactive antibodies raised against different DBL and CIDR domains.[25,26,27,28] Therefore, understanding the naturally occurring antibody response to the various DBL and CIDR domains of PfEMP1 is an important part of evaluating the usefulness of PfEMP1 as a vaccine candidate.

Previous serological studies using recombinant protein fragments of different parasite isolates implicated the CIDR1α[29], CIDR2β[30] and the DBL1α[31] in protection against malaria episodes.

In the third part of this thesis, we tried to add information to these existing data by recombinant expression of DBL and CIDR domains of 3D7 PFD1235w/*var4* and FCR3S1.2-*var1*. The former was shown to be up-regulated in 3D7 culture after selecting for *var* genes expressed by parasites causing severe disease, whereas the latter has been identified as the rosetting ligand which is in turn associated with severe malaria. Screening with naturally exposed sera of different origin should provide insight into the importance of these PfEMP1 domains in the recognition by the host immune system.

METHODS

Generation of recombinant *var* gene fragments

Saponin lysis of parasites

Saponin lysates of 3D7 and FCR3S1.2 parasite cultures were spun down for 5 min at 500g in a 15ml Falcon tube. The supernatant was discarded and the pellet was resuspended in 10 pellet- volumes of freshly prepared lysis buffer containing 0.03% of saponin in PBS. After 5min of incubation on ice and centrifugation at 400g for 10min, the supernatant was carefully removed and the pellet was washed twice with PBS.

DNA extraction

The pellet was dissolved in 465µl of TE and 20µl of 20% SDS, 10µl 0.5M EDTA and 5µl ProteinaseK (20mg/ml) were added. The mixture was incubated in a 60°C heat-block overnight. To extract DNA 400µl of H_2O-saturated phenol and 400µl chloroform were added and the mixture was shaken for 5min and centrifuged for 10min at 12000g at 4°C. The aqueous upper phase was transferred to a new tube and the extraction was repeated twice albeit in the last cycle 500µl of chloroform were used instead of phenol/chloroform mixture.

DNA precipitation

DNA was precipitated by adding 1/10 volume of 3M Na-acetate (pH 5.2) and 3 volumes of absolute EtOH to 1 volume of dissolved DNA. The mixture was incubated at -80°C for at least 1 hour and centrifuged for 30min at 12000g at 4°C. The supernatant was removed and the pellet was washed with 75% EtOH and incubated for 3min at 60°C until residual EtOH had evaporated. The precipitated DNA was then dissolved in 40µl of TE buffer and stored at 4°C.

Amplification of *var* gene fragments

Amplification of PFD1235w/*var4* fragments

The nucleotide and amino acid sequences of PFD1235w/*var4* were available at www.plasmodb.org. The extracellular part of the *var* gene was divided into 10 fragments (Figure 1a) and PCR reactions with specific primers (Appendix 1) were carried out on a Mastercycler gradient machine (Eppendorf).

The PCR contained the following reagents: 2U FIREPol® DNA polymerasel and 10x BD buffer (both Solis BioDyne), 200µM dNTPs, 400nM forward and reverse primer each, 1.5mM $MgCl_2$ and 1µl of 3D7 gDNA. The reaction was carried out in a total volume of 50µl. PCR conditions were as follows: After an initial denaturation step of 7min at 94°C, 30 cycles of denaturation for 45s at 95°C, primer annealing for 60s at 50°C and extension for 60s at 72°C followed. A final elongation step for 5min at 72°C was added.

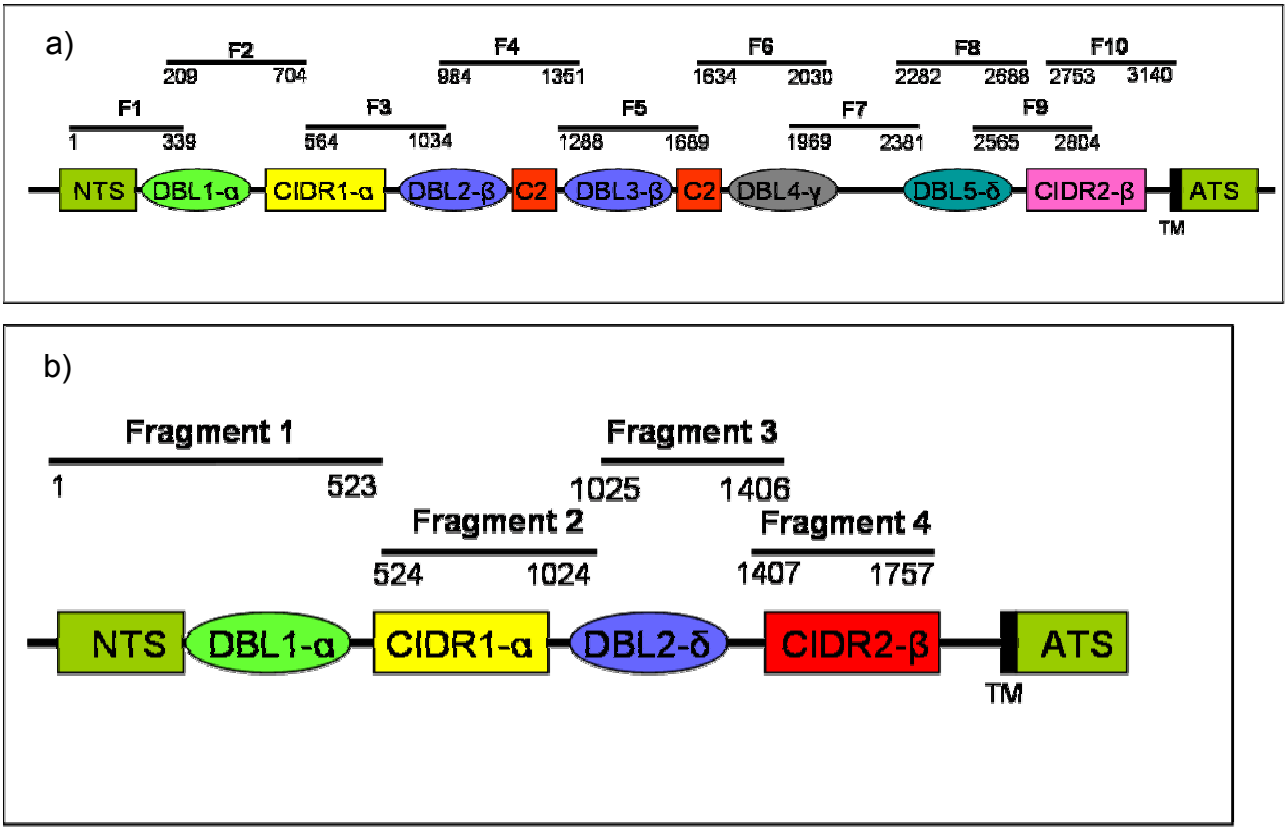


Figure 1. Domain structure of PFD1235w/*var4* (a) and FCR3S1.2-*var1* (b) and relative location of PCR fragments. PFD1235w/*var4* is a rather huge *var* gene with 7 domains compared to *var1* of FCR3S1.2 with only 4 domains. Fragments schematically represent the 10 and 4 regions that were intended to be recombinantly expressed in *E.coli*.

Amplification of FCR3S1.2-*var1* fragments

The nucleotide and amino acid sequences of FCR3S1.2-*var1* were retrieved from www.pubmed.org (accession number AF003473). The extracellular part of this *var* gene was divided into 4 fragments (Figure 1b) and PCR reactions with specific primers (Appendix 1) were carried out on a Mastercycler gradient machine (Eppendorf). The PCR contained the following reagents: 2U FIREPol® DNA polymerasel and 10x BD buffer (both Solis BioDyne), 200µM dNTPs, 400nM forward and reverse primer each, 1.5mM $MgCl_2$ and 2µl of FCR3S1.2 gDNA. The reaction was carried out in a total volume of 50µl. PCR conditions were as follows: After an initial denaturation step of 7min at 94°C, 30 cycles of denaturation for 60s at 95°C, primer annealing for 30s at primer specific temperatures (Appendix 1) and extension for 30s at 72°C followed. A final elongation step for 5min at 72°C was added.

Gel Electrophoresis

1.5% agarose was dissolved in 0.5x TBE buffer and poured into a gel chamber. PCR products were loaded with Blue Juice (30% glycerol, a spatula tip of bromphenol blue and xylene cyanol, 70% TE), run at 120V constant current and stained in ethidium bromide (1µg/ml in 0.5x TBE) for visualization under a White/UV Transilluminator.

Purification of PCR products

Positive PCR products were purified using the Nucleospin® PCR purification columns (Macherey-Nagel) according to the manufacturer's instructions. The concentration of the purified PCR products was estimated on a 0.7% agarose gel or using a Nanodrop spectrometer (Witec AG).

Cloning of PCR products into cloning vectors

For PFD1235w/*var4*, PCR fragments were initially cloned into pGEM-T and TOPO cloning vectors in order to avoid a frameshift causing a stop codon when ligating certain fragments into the expression vector. Subsequently, the PCR products were subcloned into expression vector pQE30 (carrying a N-terminal tag of 6 histidine residues) using restriction enzymes *Sac*I/*Xho*I and *Pst*I/*Sph*I for fragments cloned into TOPO and pGEM-T, respectively.

PCR products of FCR3S1.2-*var1* were cloned into expression vector pQE60 (carrying a 6xhis-tag at the C-terminus of the multiple cloning site) using restriction enzymes *Nco*I and *Bgl*II.

For both *var* genes ligation using T4 DNA ligase was done overnight in a water bath at 16°C. The ligation mix was purified by phenol/chloroform extraction before precipitation with NaAc (3M pH 4.5), absolute ethanol and t-RNA at -80°C for 1 hour. After 30min of centrifugation at 12000g at 4°C, the supernatant was removed and the pellet was washed with 75% ethanol and airdried. The precipitated DNA was dissolved in 10µl ddH_2O and stored at -20°C until further use.

Preparation of electrocompetent *E. coli* cells

M15 *E.coli* cells were grown overnight at 37°C in 5 ml LB containing kanamycin (25mg/µl). *E.coli* of this strain contain a pREP4 plasmid which confers resistance to kanamycin. 1ml of the overnight culture was transferred to 800ml LB containing kanamycin. Bacteria were grown for 4 hours until an optical density at 600nm (OD_{600}) of 0.5-0.6 was reached. Cells were kept on ice for 30min at 4°C. After centrifugation at 2500g for 15min, the supernatant was discarded and the pellet was resuspended in 800ml cold and sterile ddH_2O. Cells were spun again for 10min, the supernatant was discarded and the pellet was resuspended in 400ml ddH_2O as described before. This step was repeated twice before resuspending the pellet in ddH_2O containing 10% glycerol. The solution was centrifuged for 10min, the supernatant was discarded and the pellet was resuspended in 1ml of ddH_2O again containing 10% glycerol. Aliquots of 45µl were transferred into 500µl Eppendorf tubes, immediately frozen in liquid nitrogen and stored at -80°C until further use.

Electroporation of *E.coli*

40 µl of electrocompentent M15 cells were transformed by electroporation (voltage:2500V, capacity: 25µF, shunt: 201Ω, pulse: 5ms) and immediately transferred to an Eppendorf tube containing 1ml of pre-warmed LB medium. Bacteria were allowed to recover in a water bath at 37°C for 30min, plated on agar plates containing 100µg/ml ampicillin and kanamycin (25µg/ml) and incubated at 37°C overnight.

Screening of electroporated cells and isolation of plasmids

To check whether bacteria had taken up the plasmid with the correct insert, a PCR with the corresponding vector primers was performed and a masterplate of the colonies to be checked was prepared and incubated at 37°C overnight. The PCR products were analyzed on an agarose gel as described before. Clones with a correct insert were picked from the masterplate and grown in LB containing ampicillin at 37°C overnight. Plasmids from the overnight culture were isolated by miniprep and inserts were cut out with the appropriate restriction enzymes to test for correct sizes.

Small scale expression

For all recombinant fragments expression was tried first in a small scale experiment. 20ml of TB medium containing ampicillin and kanamycin were inoculated with 1ml of overnight bacterial culture grown in 4ml of LB medium containing the same antibiotics. The freshly inoculated culture was shaken at 37°C and OD_{600} was measured in constant intervals. When the culture had reached an OD_{600} of approximately 0.6, 1ml of culture was removed and kept at -20°C as uninduced control. Expression in the remaining culture was induced by adding IPTG to a final concentration of 1mM and incubated for 4 hours. 1ml of the induced culture was transferred to an Eppendorf tube. Remaining cells were harvested, centrifuged at 2000g for 5 min and stored at -20°C until further use ("determination of solubility" see below). Both samples, induced and uninduced, were centrifuged for 5min at 10000g. The pellet was resuspended in 100µl of Laemmli buffer and heated up to 96°C for 5min to denature proteins and break disulfide bonds. The protein samples were separated by SDS-PAGE (see below).

SDS-PAGE and Western Blot

Protein samples were separated on a 12.5% SDS polyacrylamide gel for 1.5 hours at 30mA. Separated proteins were transferred to a nitrocellulose membrane (Hybond™-C Extra, Amersham Biosciences) using a semi-dry blotter (BioRad). Blotting was carried out for 3 hours at 42mA in Tris-glycine buffer containing 20% methanol. Membranes were blocked for 1 hour with TNT (0.1M Tris-HCl, 0.15M NaCl, 0.05% Tween 20, pH 8) containing 5% non-fat milk powder. The primary mouse anti-6xHis antibody (dilution 1:2000 in 1% TNT) was incubated for 4 hours at room temperature. Afterwards, the membrane was washed 6 times for 5min in TNT. As secondary antibody, a goat anti-mouse antibody labeled with alkaline phosphatase (Sigma; dilution 1:20000 in 1% TNT) was used and the membrane was incubated for 1 hour. After another round of extensive washing in TNT, the membrane was incubated with Tris-buffer containing 300µg/ml 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and 150µg/ml nitro blue tetrazolium (NBT). The color reaction was stopped with ddH_2O.

Determination of solubility and large scale expression

Cell pellet of remaining culture from small scale expression was used to determine the solubility of the recombinant protein fragments following QIAexpressionist™ protocol 5 "Determination of target protein solubility". In brief, the cell pellet was resuspended in 5ml of lysis buffer A for native purification and frozen in dry ice/methanol. After thawing in cold

water, lysate was treated with 1mg/ml of lysozyme and 20µg/ml RNase and DNase for 30min. Lysate was sonicated 6x 10 seconds using a sonicator (Branson Sonifier). After centrifugation for 30min at 10000g the supernatant containing the soluble fraction (extract A) was transferred into a new tube. The pellet containing the insoluble fraction (extract B) was suspended in 5ml of lysis buffer. An aliquot of soluble and insoluble fractions were combined with 2x SDS sample buffer and analyzed on a 12.5% SDS gel. Once the solubility of the proteins was determined protein lysates of 500ml large scale expression cultures were prepared according to protocols 8 and 9 of the QIAexpressionist™.

Purification of recombinant *var* fragments

Affinity purification was carried out according to QIAexpressionist™ protocols 11 and 15 for soluble and insoluble protein fragments, respectively. In short, the clear lysate obtained under native conditions was mixed with 1ml of Ni-NTA agarose and shaken for 1 hour at 4°C to allow the recombinant protein fragments to bind to the Ni-NTA via the His-tag. The slurry mixture was then transferred into a plastic column containing a fibreglas filter previously saturated with lysis buffer A .The Ni-NTA protein mixture was run through the column and the flow-through was collected. Subsequently, the column was washed twice with 4ml wash buffer and protein fragments were eluted in 4 fractions (0.5ml) of elution buffer pH 8.0.

For purification under denaturing conditions the incubation with Ni-NTA agarose was performed at room temperature. The lysate, stored in buffer B containing 8M urea, was loaded onto a column previously saturated with buffer B and the flow-through was collected. After washing the column twice with 4ml of wash buffer (pH 6.3) under denaturing conditions the recombinant fragments were eluted 4 times in 0.5ml elution buffer pH 5.9 and 4 times in 0.5ml elution buffer pH 4.5.

Fractions were analyzed by Western blot and protein concentration was roughly estimated by Nanodrop at 280nm. Fractions containing the highest concentration of the recombinant protein were pooled and stored at -20°C and 4°C for soluble and insoluble fragments, respectively.

ELISA on recombinant protein fragments of PFD1235w/*var4*

Recombinantly expressed protein fragments of PFD1235w/*var4* were used as antigens in an Enzyme-linked Immunosorbent Assay (ELISA) and screened for recognition by sera from 48 adults from PNG, 37 pairs baseline/follow-up samples from Tanzanian children as well as from 7 children with asymptomatic and 8 children with severe malaria from Tanzania and PNG. 2.5µg/ml of recombinant protein in fresh PBS were coated on

Maxisorp 96-well plates (Nunc) and incubated over night at 4°C (Figure 2). Plates were washed once in ddH_2O containing 0.05% Tween20 using an ELISA washer and blocked for 1 hour at room temperature with 5% non-fat milk powder in PBS/0.01% Tween20 (PBS/T). Plates were washed again and incubated with serial dilutions of human sera starting from dilution 1:200 for adults and 1:100 for children in 1% non-fat milk powder in PBS/T. After incubation for 2 hours at room temperature, the plates were washed 3 times on the ELISA washer and incubated with the secondary goat anti-human IgG AP-labeled antibody (1:5000) in 1% non-fat milk powder in PBS/T for 1 hour at room temperature. After washing, the plate was incubated with PNP (1mg/ml) in alkaline substrate buffer (160mM $NaHCO_3$, 130mM Na_2CO_3, 1mM $MgCl_2$, pH 8.6) and the signal was detected by an ELISA reader at 405nm.

Sera	Dilution											
	1: 200	1: 400	1: 800	1: 1600	1: 3200	1: 6400	1: 12800	1: 25600	1: 51200	1: 102400	1: 204800	1: 409600
posSP												
negSP												
Serum1												
Serum2												
Serum3												
Serum4												
Serum5												
Serum6												

Figure 2. ELISA plate layout. 2.5µg/ml of each of the 4 recombinant PFD1235w/*var4* fragments and the unrelated control fragment were coated in each well. 6 different sera as well as a positive and a negative serum pool were serially diluted starting from 1:200 to a final dilution of 1:409600.

ELISA analysis and data management

Adults' and children's sera described above were tested for recognition on recombinant protein fragments of PFD1235w/*var4*. Recombinantly expressed DHFR was used as a control to assess background levels. A serum pool consisting of 20 individual sera from PNG served as positive control and a serum pool of malaria-negative blood donors was used to calculate the threshold of recognition. All OD_{405} values were recorded and corrected for DHFR recognition by subtraction: OD_{serum}-OD_{DHFR}. The corrected values were plotted against the logarithmic dilution. The mean value of the negative serum pool plus two standard deviations was used to calculate the threshold for each plate

individually. The endpoint titer corresponded to the dilution where the curve of the plotted sera met the threshold line.

Synthetic peptides of PFD1235w/*var4*

Because only some fragments could be obtained as recombinant proteins, synthetic peptides were used to bridge the gaps. Coverage was not complete but representative peptides were chosen (Figure 3).

Selection of synthetic peptides

Amino acid sequences of parts of PFD1235w/*var4* which could not be recombinantly expressed were loaded into the PeptideSelect™ DesignTool avaiable at http://peptideselect.invitrogen.com and checked for antigenicity, hydrophobicity and surface accessibility. According to the STABLECOIL analysis tool (http://www.bionmr.ualberta.ca/bds/software/stablecoil; website currently under reconstruction), PFD1235w/*var4* does not contain any stable coiled-coil structures which are thought to adapt their natural conformation in aqueous solution and which would have rendered recognition in an ELISA more reliable. Therefore peptides were also chosen from regions with predicted intrinsically unstructured characteristics as well as from regions with predicted globular structure (for sequence information see Appendix 2). In order to obtain strong adherence of the peptides to the plastic surface of the ELISA plates, peptides were ordered with a biotin-label at the N-terminus and ELISAs were carried out on streptavidin-coated 96-well plates (Nunc). A PEG-linker between the biotin and the peptide sequence should facilitate the accessibility to the peptide when bound to the plate.

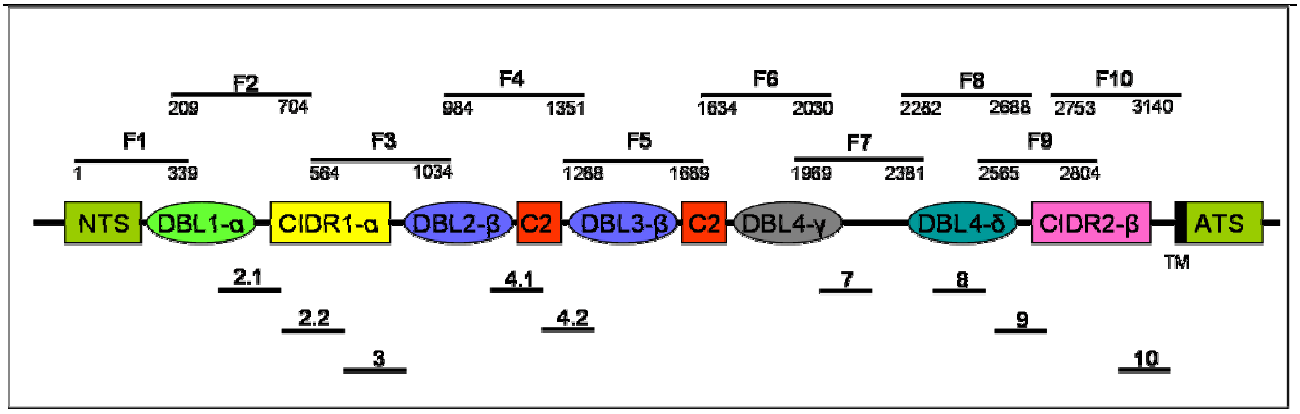


Figure 3. Location of synthetic peptides. Detailed schematic representation of the 9 synthetic peptides 2.1, 2.2, 3, 4.1, 4.2, 7, 8, 9 and 10 and their relative location to the recombinant fragments of PFD1235w/*var4* are indicated. Numbers below black bars indicate amino acid boundaries of recombinant fragments.

Peptide ELISA

Streptavidin-coated 96-well plates (Nunc) were pre-washed 4 times on the ELISA washer. 2.5µg/ml of the 9 synthetic peptides 2.1, 2.2, 3, 4.1, 4.2, 7, 8, 9 and 10 were coated in fresh PBS (pH 7.2) for 1.5 hours at room temperature. Plates were washed 3 times on the ELISA washer before blocking the plates for 1 hour in PBS containing 5% non-fat milk powder. Plates were washed twice and human sera (same as used above) were applied in a high salt dilution buffer (2.5 % non-fat milk powder, 0.05% Tween20 in PBS and 0.363M of NaCl) in a dilution of 1:200 for adults and 1:100 for children (Figure 4). Plates were incubated at room temperature for 60min and washed again twice. Secondary goat anti-human IgG labeled with alkaline phosphatase (Sigma) was added in a dilution of 1:5000 in PBS/5%milk/0.05% Tween20 and incubated for 60min at room temperature on a rotary shaker. Afterwards, plates were washed and PNP substrate was added in alkaline substrate buffer (see above) in a 1:100 dilution. OD_{405} was measured by an ELISA reader.

	1	2	3	4	5	6	7	8	9	10	11	12
A	posSP	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
B	posSP	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22
C	posSP	S23	S24	S25	S26	S27	S28	S29	S30	S31	S32	S33
D	posSP	S34	S35	S36	S37	S38	S39	S40	S41	S42	S43	S44
E	negSP	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
F	negSP	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22
G	negSP	S23	S24	S25	S26	S27	S28	S29	S30	S31	S32	S33
H	negSP	S34	S35	S36	S37	S38	S39	S40	S41	S42	S43	S44

Figure 4. Layout of peptide ELISA. 44 individual sera per plate were analyzed in duplicates. Positive and negative serum pools were pipetted in quatruplicates.

Analysis of peptide ELISA

Streptavidin-coated plates are very expensive and therefore we decided not to go for serial dilutions, but only to look for “responder” and “non-responder” (serum-wise) or “recognition “and “no recognition” (peptide-wise). A serum pool of 20 semi-immune sera from an endemic area served as positive control. Background recognition was accounted for by subtracting OD_{405} values of an uncoated plate without the corresponding peptide. A serum pool of European blood donors without any history of malaria served as negative control. The threshold of recognition for each individual peptide was set as the mean OD_{405} of the negative serum pool on each plate plus two standard deviations.

Re-expression of recombinant fragments of FCR3S1.2-*var1*

To secure sufficient protein amounts, additional expression of recombinant fragments had to be done. Overnight cultures from glycerol stocks and large scale expression experiments as described above were made to re-express protein fragments in appropriate amounts. However, even after several rounds of re-expression, protein yields remained low and were not enough to be tested in ELISA. Transformation into several different *E.coli* cell types did not help to solve the problem. Therefore, instead of ELISA, Western blot analysis requiring much less material was performed.

Western blot on recombinant fragments of FCR3S1.2-*var1*

100µl of recombinant protein were combined with 100µl of 2xSDS sample buffer and incubated for 5min at 96°C. The mixture was loaded on a 12.5% SDS acrylamide gel and SDS-PAGE was performed as described above. In order to test human sera of different origin (same as used in ELISA on PFD1235w/*var4*) the membrane was cut into 24 strips. Strips were blocked for 1 hour in TNT containing 5% non-fat milk powder. After washing the membrane once in TNT the primary antibody was applied in a 1:1000 dilution in TNT

containing 1% non-fat milk powder for 1 hour at room temperature. Positive and negative serum pools were the same as described previously. Since all recombinant proteins contained a 6xhis-tag, a mouse anti-6xhis antibody was included as additional control. After washing 3 times for 5min in TNT membrane strips were incubated for 1 hour with secondary goat anti-human (1:5000) and goat anti-mouse antibodies (1:10000). Washing 3 times for 5min with TNT and once 5min in Tris for Western (0.1M Tris, 0.5mM $MgCl_2$, 1.5M NaCl, pH 9.5), BCIP and NBT (1:100 in Tris for Western) were added and the color reaction was stopped with ddH_2O.

Analysis of Western Blot signals

Signal intensity was determined using the Spot Density tool of AlphaEase FC software® (Alpha Innotech Corporation). The positive serum pool served as internal control and was used to normalize the signals of the individual sera. The signal of the anti-6xhis antibody was used to distinguish between specific recognition of the recombinant protein fragments and additional unspecific bands in the positive serum pool probably due to recognition of residual *E. coli* fragments that failed to be removed during purification. To account for differences in the “window size” used to measure the singals, not the absolute values but the average value (AVG) was measured which is the integrated density value (IDV) divided by the area in which the corresponding signal was detected. The negative serum pool was subtracted as unspecific background levels.

RESULTS

Recognition of PFD1235w/*var4* in naturally exposed individuals

With the predicted size of 405kDa PFD1235w/*var4* is a rather large *var* gene with multiple domains conferring different receptor bindings. Several *var* gene domains like DBL1α and the CIDR1α were implicated in binding to different host receptors. Since PFD1235w/*var4* was associated with severe malaria previously[16], we were interested in finding possible immunodominant epitopes of this specific *var* gene that might be associated with the generation of protective anti-PfEMP1 antibodies. Due to the large size, the protein could not be expressed as a whole, but should have been expressed in 10 recombinant overlapping fragments.

Amplification of PFD1235w/*var4* fragments was previously carried out in our lab. Unfortunately, only 5 of 10 fragments (fragments 1,5,6,8 and 9) could be expressed (Figure 5), 4 of which were available as glycerol stocks (fragments 1,5,6 and 8) and were re-expressed in order to obtain reasonable amounts of protein to perform ELISA.

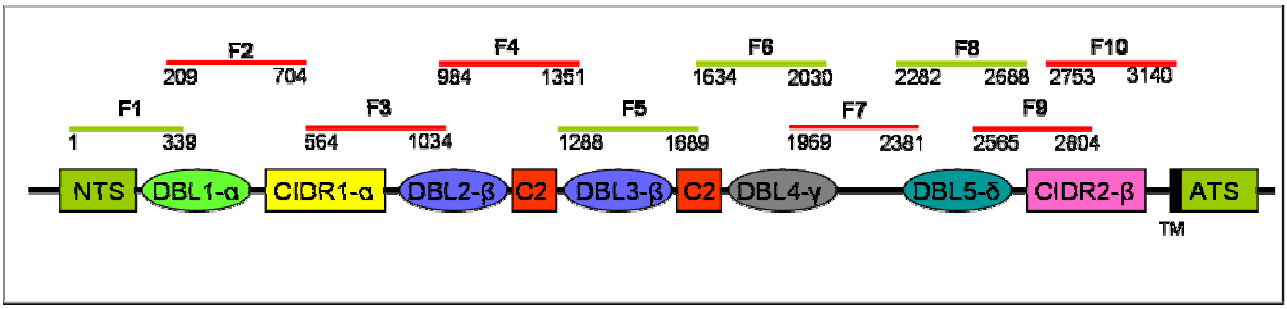


Figure 5. Schematic representation of PFD1235w/*var4* and the recombinant protein fragments at their relative position. Protein fragments which could be expressed are indicated in green, fragments which failed to express are indicated in red.

It is preferrable to use native protein in ELISA experiments, but the failure to express the protein fragments under native conditions made it necessary to extract them under strong denaturing conditions. In several large scale expression experiments sufficient amounts of protein could be produced (Figure 6). Predicted sizes of the 4 fragments were 43.4kDa, 53.3kDa, 49.1kDa and 45.6kDa for fragment 1, 5, 6, and 8, respectively. Figure 6 shows the 4 recombinant protein fragments after purification. Most of the additional bands are probably due to degradation, however, some of them might also be attributed to histidine-containing *E. coli* proteins.

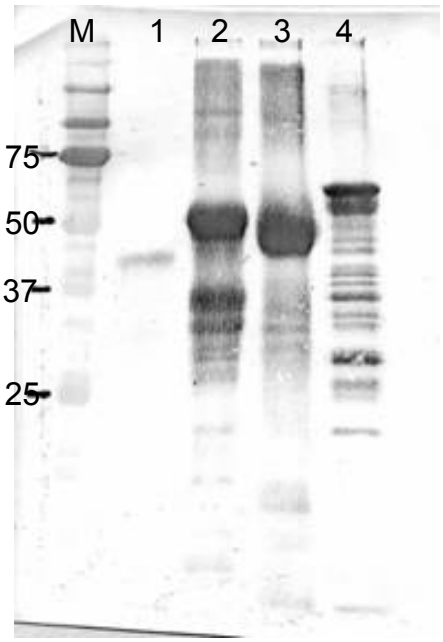


Figure 6. Western Blot of recombinantly expressed PFD1235w/*var4* domains.
Proteins were purified over a Nickel column, size-separated on a 12.5% SDS-PAGE and detected with a mouse anti-6xhis and an alkaline phosphatase (AP)-labeled anti-mouse IgG antibody. In the first lane Precision Plus Protein all blue standard marker is indicated with its 75kDa, 50kDa, 37kDa and 25kDa bands highlighted. Lanes 1-4 correspond to recombinant fragments 1, 5, 6 and 8, respectively.

ELISA on recombinant PFD1235w/*var4* fragments

ELISA on recombinant PFD1235w/*var4* fragments using semi-immune adults' sera

The 4 recombinantly expressed fragments 1, 5, 6, and 8 were tested in ELISA for their differential recognition by sera of different origin. In brief, proteins were immobilized on 96-well plates and endpoint titers were measured for 48 sera from asymptomatic, semi-immune adults collected during a case-control study in PNG. The data were assessed as shown in Figure 2.

In general, there were striking differences in recognition (Table 1+2). Fragments 5 and 6 corresponding to the DBL3β and C2-DBL4γ region were well recognized by nearly all sera (42/48 and 45/48) with a mean endpoint titer (EPT) of approximately 1:14000. Fragment 8 corresponding to DBL5δ was recognized equally well (41/48) but with a much lower mean EPT of 1:5500. Fragment 1 corresponding to the NTS-DBL1α domain was the least recognized protein (11/48) with a mean EPT of only 1:3600. Most sera (28/48) recognized 3 of the 4 recombinant fragments. Similar numbers of sera (8/48 and 9/48) recognized 2

and 4 fragments and 3 sera were infrequent responders with only one recognized fragment. Interestingly, individual sera did not only differ in their ability to recognize different protein fragments but also in the strength these fragments were recognized with. As an example, SUK 80 recognized fragment 6 in a dilution of 1:102400, fragment 8 was only recognized up to 1:1600, but fragments 3 and 5 were not recognized at all This indicates, that differences in recognition frequencies are really due to the recognition of different domains and not simply because the donor was a good responder with high titers against all recognized antigens.

	Fragment 1	Fragment 5	Fragment 6	Fragment 8
FoR_{total}	11/48	42/48	45/48	41/48
FoR_{males}	4/27	26/27	25/27	23/27
$FoR_{females}$	7/21	16/21	20/21	18/21
Mean EPT Total	3618	14319	13844	5507
Mean EPT Male	2000	16515	16672	3748
Mean EPT Female	4543	10750	10310	7756

Table 1. Frequency of recognition (FoR) and mean endpoint titers (EPT) of adult sera from PNG. FoR and mean EPT are indicated for all sera tested (total) as well as for men and women separately.

ID	Fragment 1	Fragment 5	Fragment 6	Fragment 8
45	0	25600	12800	0
46	0	1600	12800	200
47	0	12800	25600	6400
49	0	800	800	1600
52	0	6400	3200	3200
53	0	0	6400	102400
43	0	25600	12800	12800
44	0	6400	6400	400
50	0	1600	3200	1600
51	0	3200	6400	1600
53	0	800	1600	800
59	0	6400	200	800
61	0	0	3200	0
62	0	6400	25600	12800
64	0	3200	0	0
65	400	1600	200	400
66	0	0	0	800
67	200	12800	3200	1600
68	0	0	6400	3200
72	0	12800	12800	6400
73	0	12800	51200	6400
76	0	6400	51200	6400
80	0	0	102400	1600
81	0	3200	6400	200
83	0	3200	12800	0
86	800	12800	6400	1600
88	800	12800	12800	3200
89	0	400	6400	1600
90	0	6400	6400	800
93	0	6400	12800	1600
94	0	102400	25600	3200
95	3200	6400	3200	200
98	0	12800	3200	1600
103	800	200	800	800
105	0	3200	12800	0
108	0	800	200	0
109	3200	51200	6400	400
110	0	25600	6400	6400
113	800	25600	12800	3200
115	0	102400	3200	12800
122	3200	0	1600	6400
127	25600	6400	0	1600
128	0	6400	12800	3200
131	0	1600	3200	1600
134	0	6400	51200	1600
136	800	3200	3200	800
137	0	51200	12800	0
139	0	3200	51200	1600

Table 2. Endpoint titers of semi-immune adults from PNG tested on the 4 recombinant fragments of PFD1235w/*var4*.

Differences in recognition are indicated by a color code ranging from white (no recogntion) to dark brown (strong recognition) and their corresponding reciprocal endpoint titers in the range between 0 and 102400.

Differences in recognition frequencies regarding sex

Of the 48 donors, 27 were male and 21 were female. In order to test, whether recognition was dependent on sex, frequencies of recognition were analyzed for male and female, seperately. Fragments 6 and 8 seem to be recognized equally well by both sexes. However, whereas females tended to better recognize fragment 1 the frequency of recognition for fragment 5 was slightly higher in serum samples of male donors (Table 1 and Figure 7) However differences in recognition were not significant (p= 0.17 for fragment 1 and p=0.07 for fragment 5 using Fisher's exact test).

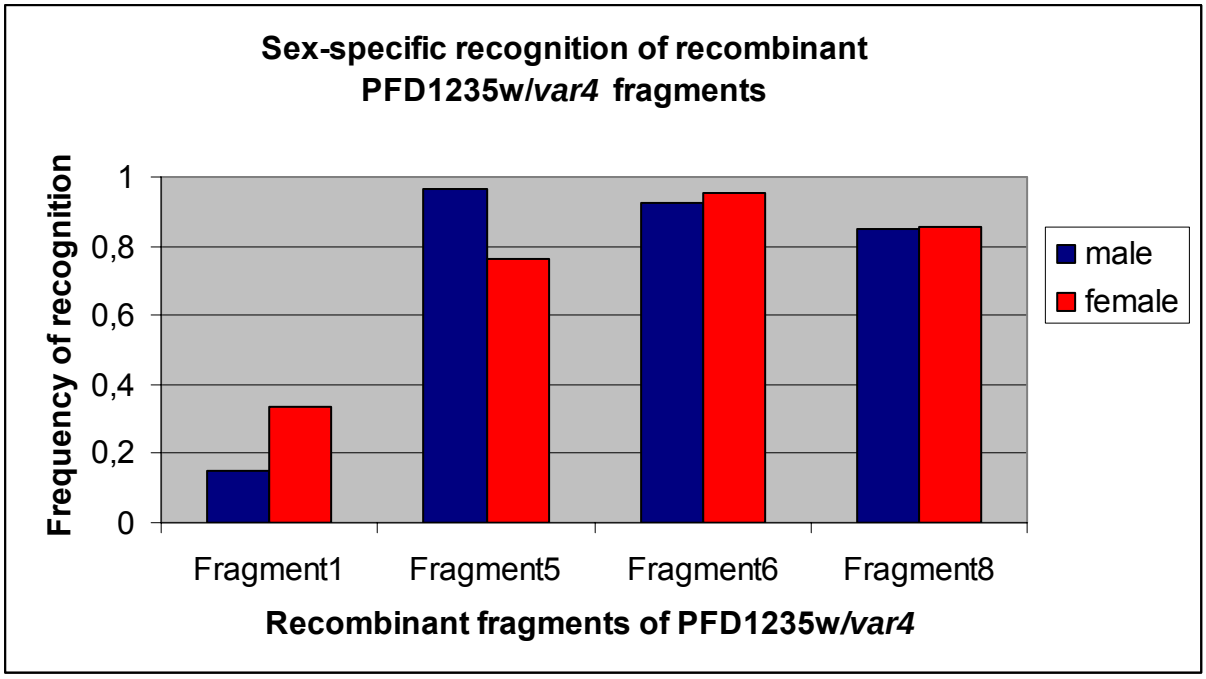


Figure 7. Gender-specific frequencies of recognition

Frequency of recognition of the 4 recombinant PFD1235w/*var4* fragments were analyzed for male (blue) and female (red) donors separately.

ELISA on recombinant PFD1235w/*var4* fragments using children's sera

In order to investigate whether recognition frequencies in adults differed from that in children, which might indicate the presence of protective antibodies against certain epitopes, we analyzed serum samples from 36 African children collected in a longitudinal study. In general, both, frequency of recognition and antibody titers were much lower in children's sera when compared to adults' sera (Table 3 and Table 4).

	Fragment 1		Fragment 5		Fragment 6		Fragment 8	
	Baseline	6 months later	Baseline	6 months later	Baseline	6 months later	Baseline	6 months later
BC602	0	0	0	0	0	400	0	0
BC622	0	0	200	0	3200	1600	0	0
BC612	0	400	800	800	800	800	800	200
BC617	0	0	0	0	0	200	0	0
BC619	0	0	0	200	200	1600	0	200
BC600	0	0	0	0	1600	3200	400	400
BC586	0	0	0	0	200	3200	0	800
BC583	0	0	0	0	100	800	0	0
BC601	0	0	0	0	100	1600	400	800
BC598	0	0	0	0	800	400	400	400
BC555	0	0	0	0	0	0	0	0
BC552	0	0	100	100	1600	0	800	1600
BC548	0	0	0	0	0	400	800	3200
BC640	0	0	0	0	0	200	800	400
BC645	100	200	0	0	200	200	200	0
BC646	0	0	0	0	200	400	200	400
BC648	100	0	100	0	0	400	800	200
BF309	0	800	0	0	0	0	0	800
BC572	200	0	0	0	3200	800	400	200
BC574	0	0	0	0	800	1600	200	0
BC578	0	0	0	0	0	800	0	0
BC580	0	0	0	0	400	3200	0	1600
BC575	0	0	0	200	1600	1600	800	1600
BC623	0	0	800	0	6400	3200		
BC624	0	0	0	0	6400	800	0	0
BC626	0	200	0	0	100	800	100	400
BC635	100	100	0	200	0	1600	0	0
BC643	800	800	0	0	0	0	0	100
BC561	200	100	0	0	0	0	0	0
BC559	800	400	0	0	0	0	1600	100
BF321	400	1600	0	0	400	1600	100	200
BF323	400	0	200	0	200	800	100	0
BF328	0	400	0	0	0	200	0	200
BF329	0	400	0	100	200	200	0	0
BF332	0	0	0	100	0	0	100	0
BF334	0	100	0	0	0	0	0	0

Table 3. Recognition frequencies of and dynamics of antibody titers to recombinant fragments in children from Tanzania.

Differences in recognition are indicated by a color code ranging from white (no recognition) to orange (strongest recognition) and their corresponding reciprocal endpoint titers in the range between 0 and 6400.

In children, as in adults, fragment 6 was the one most frequently recognized (table 3 and 4; with a mean EPT of 1:1367 at baseline), fragment 8 was moderately (mean EPT= 1:524) and fragment 1 was only poorly (mean EPT 1:344) recognized. Interestingly,

fragment 5 which was well recognized in adults was hardly recognized in children. Children sera were collected in Tanzania and adults sera were coming from PNG and therefore this difference in recognition might be due to the different geographical origin of children and adults sera. However, it should be mentioned that 3D7 is a parasite strain of unknown origin.

	Fragment 1	Fragment 5	Fragment 6	Fragment 8
$FoR_{baseline}$	9/36	6/36	21/36	17/36
$FoR_{follow-up}$	12/36	7/36	28/36	20/36
Mean $_{EPT\ baseline}$	344	367	1367	524
Mean $_{EPT\ follow-up}$	458	243	1164	690

Table 4. Frequency of recognition (FoR) and mean endpoint titers (EPT) of children sera from Tanzania. FoR of the 4 recombinant PFD1235w/*var4* fragments and mean EPT of baseline and follow-up samples are indicated separately.

Recognition patterns of baseline and follow-up samples were included to gain insight into dynamics of antibodies against the 4 recombinant fragments (Figure 8).

The frequency of recognition slightly increased for all recombinant fragments when baseline and follow-up samples were compared. Comparison of mean EPT of the these two sample groups showed a slight increase for fragments 1 and 8 and a slight decrease for fragments 5 and 6. However, neither of the two changes was statistically significant and therefore no real increase or decrease in antibody titers to any of the 4 fragments could be observed if baseline and follow-up sample were compared.

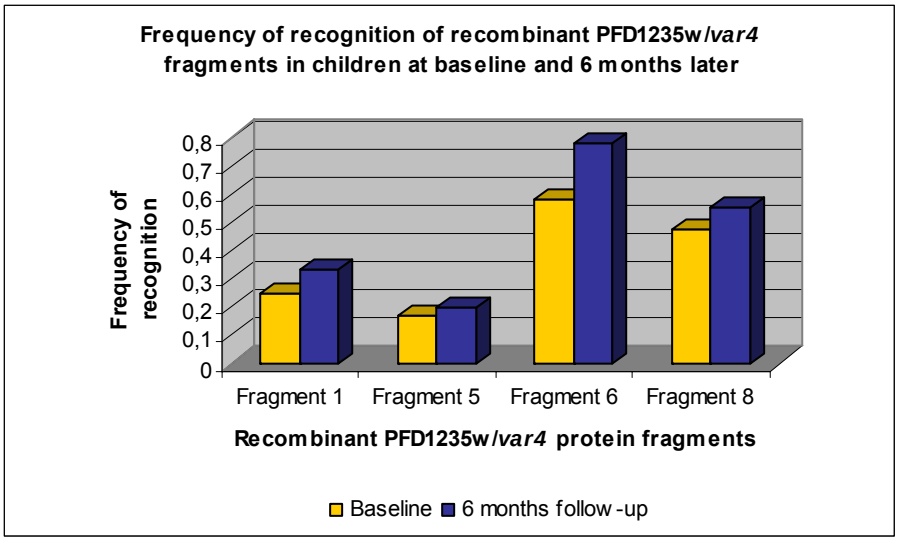


Figure 8. Antibody dynamics of longitudinal samples. FoR of the 4 recombinant PFD1235w/*var4* fragments for baseline and follow-up samples are indicated.

In order to test whether recognition of any of the 4 recombinant fragments might be due to protective antibodies, a small set of children's sera with 7 asymptomatic and 8 well defined

severe malaria cases from Tanzania was analyzed. Due to the small sample size, age-matched children with asymptomatic and severe malaria from PNG were included as well. ELISA revealed that none of the children with severe malaria was able to recognize fragment 5, whereas 57% of children with asymptomatic malaria did recognize it (Figure 9) with no difference between children coming from PNG or Tanzania. However, due to the small sample size, this difference was only borderline significant (p=0.08). A difference in recognition of the remaining fragments between children with asymptomatic and severe malaria could not be observed.

In conclusion, recognition patterns of fragments 1, 6 and 8 were similar between adults and children. In contrast, fragment 5 was well recognized by adults' sera whereas children could only hardly recognize this specific protein. Additional experiments comparing recognition between children with asymptomatic and severe malaria revealed a borderline significant difference between these two groups with no recognition in severe children and 57% of recognition in asymptomatic children. Therefore, fragment 5, corresponding to the DBL3β domain, might contain epitopes implicated in the generation of protective antibodies.

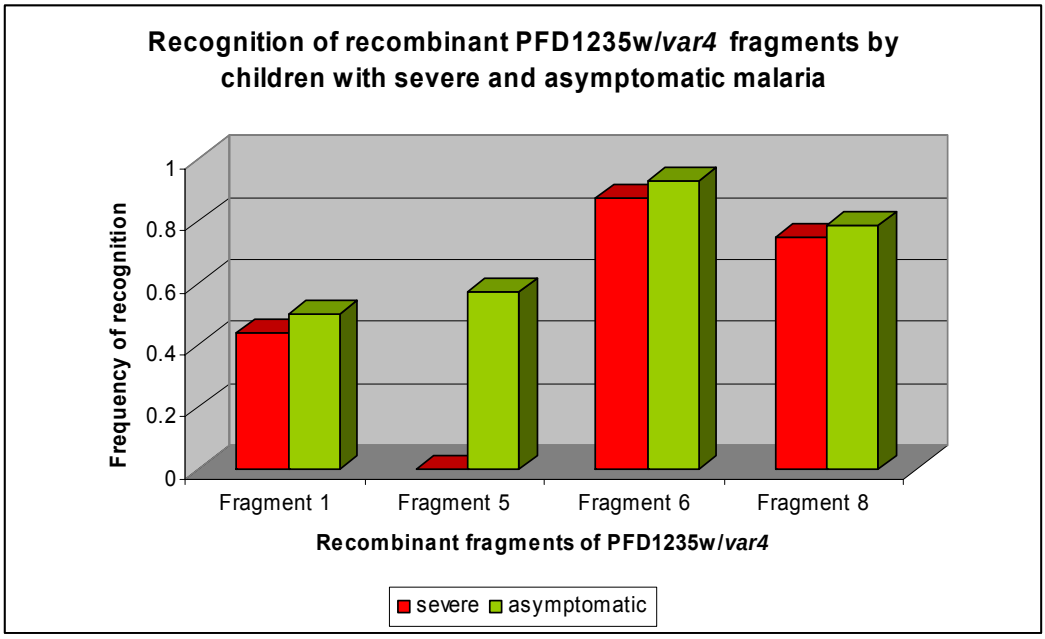


Figure 9. Recognition of recombinant PFD1235w/*var4* fragments by children with severe and asymptomatic malaria from PNG and Tanzania.

ELISA on PFD1235w/*var4* synthetic peptides using semi-immune adults' sera

9 synthetic peptides fused to biotin were incubated on streptavidin-coated 96-well plates and tested for recognition by children's and adults' sera from endemic areas used before. In semi-immune adults from PNG peptides 2.2 and 7 corresponding to parts of the CIDR1α

and the DBL4γ domain of PFD1235w/*var4* (Figure 3) were best recognized reaching median OD_{405} values of 0.46 and 0.33, respectively (Figure 10). Peptides 2.1, 4.1, 4.2, 8 and 9 were moderately recognized (median OD_{405} = 0.18, 0.21, 0.11, 0.26 and 0.15, respectively) whereas peptides 3 (median OD_{405}=0) and 10 (median OD_{405}=0.02) were hardly recognized. Differences in the recognition of synthetic peptides were found to be statistically significant (p< 0.01; Kruskal-Wallis test).

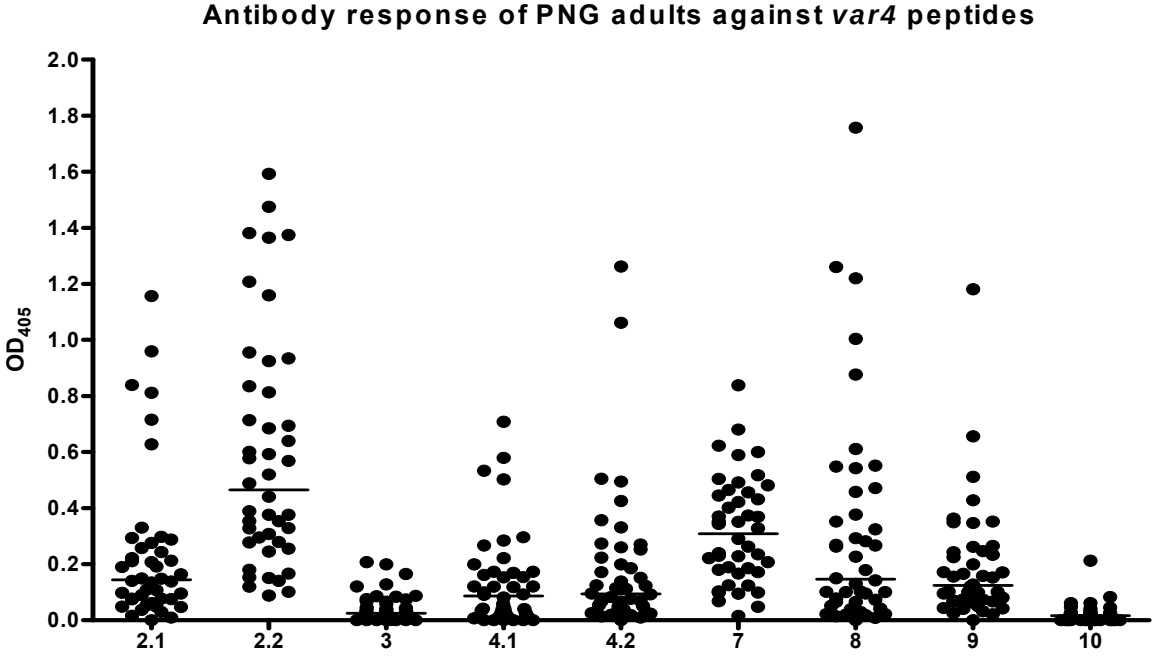


Figure 10. Signal intensities of semi-immune adult sera tested on the 9 synthetic peptides of PFD1235w/*var4*.

The x-axis shows the 9 different peptides. On the y-axis the signal intensity (OD_{405}) is indicated. The horizontal bar marks the median of recognition for each of the 9 peptides separately.

When recognition patterns of adult and children sera (same as used before) were compared (Figure 11), adults had significantly higher antibody titers for peptides 2.1 (p= 0.0007), 2.2 (p= 0.0001), 4.1 (p=0.002), 4.2 (p=0.0001), 7 (p=0.004) 8 (p=0.0001) and 10 (p=0.0001; Wilcoxon-Mann-Whitney-U test).

When children with asymptomatic and severe malaria from PNG and Tanzania were compared (Figure 12), differences in recognition were only significant for peptide 2.1 (p=0.03, Wilcoxon-Mann-Whitney-U test) corresponding to the interface between the DBL1α and the CIDR1α domain, and peptide 4.1 (p=0.049) covering part of the DBL2β-C2 region, with the latter one being only borderline significant. Recognition of peptides was found to be independent from the geographical origin of the tested sera (data not shown).

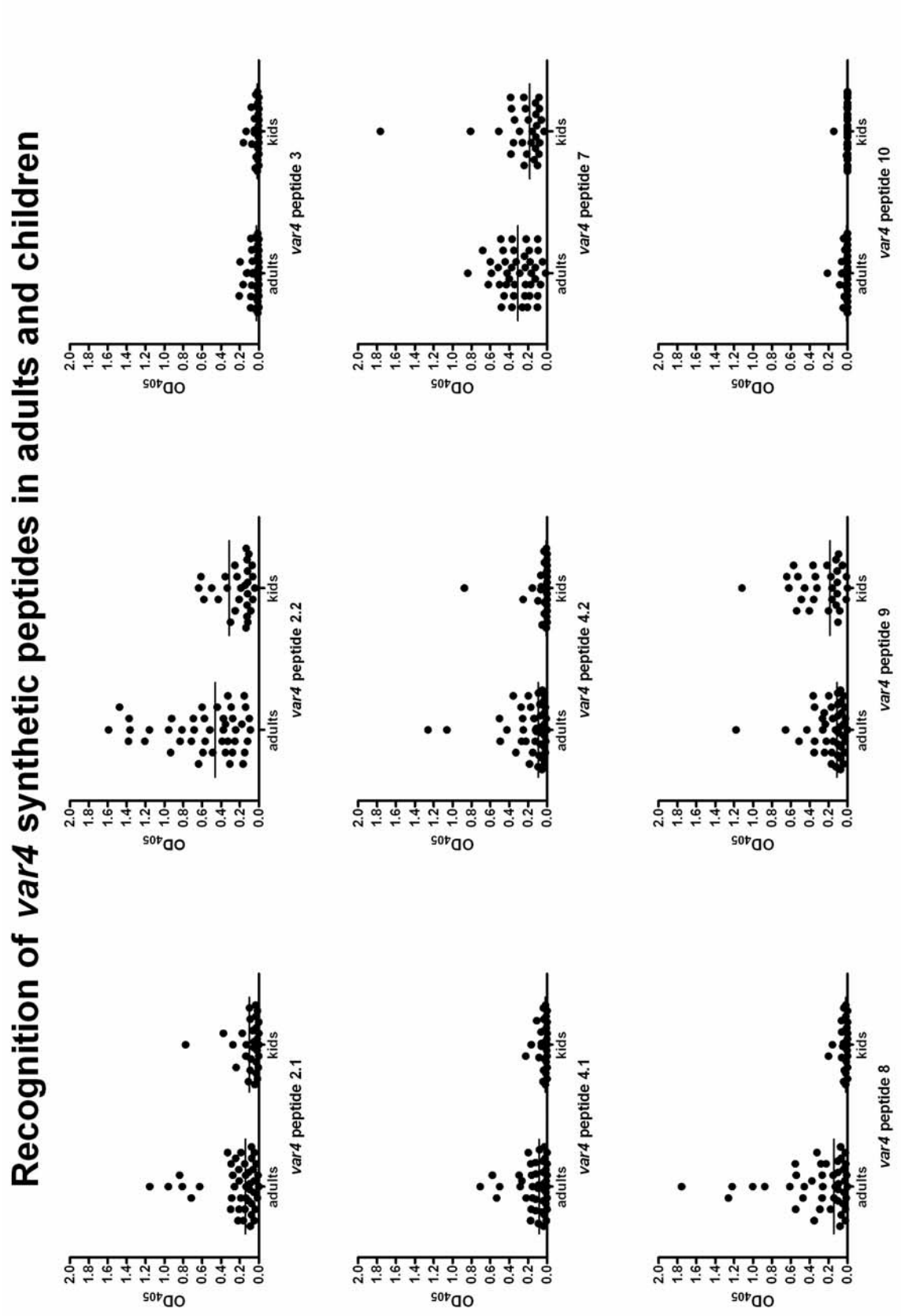


Figure 11. Recognition of synthetic *var4* peptides in adults and children. Comparison of antibody titers against peptides 2.1, 2.2, 3, 4.1, 4,2. 7, 8,9 and 10 in adults and children.

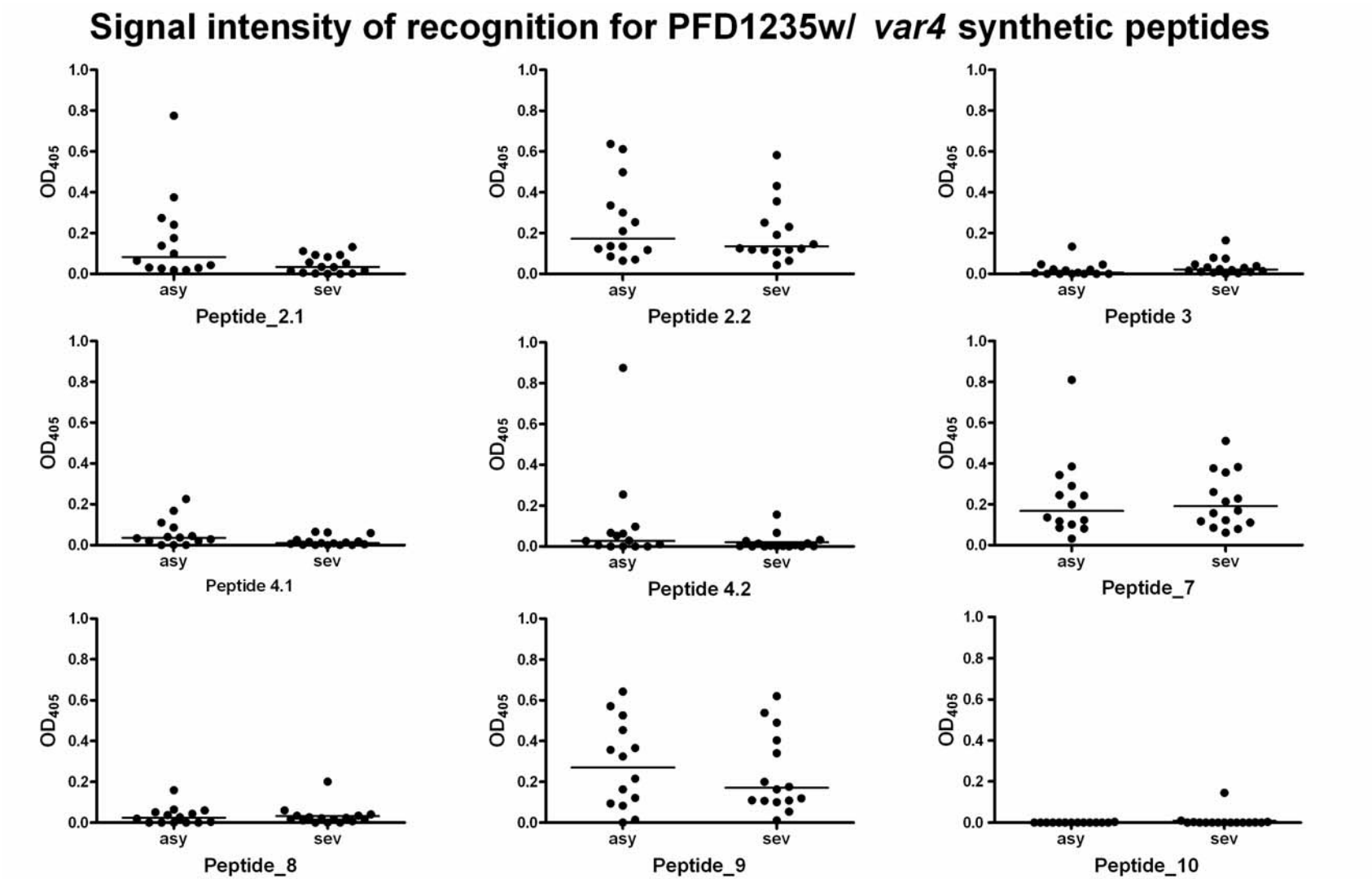


Figure 12. Signal intensity of synthetic peptide recognition in children with asymptomatic and severe malaria.

Children sera from 7 asymptomatic and 8 severe cases from PNG and TZA were tested for their ability to recognize the 9 synthetic peptides. Recognition of each peptide is shown for asymptomatic and severe cases separately. The horizontal bar indicated the median signal intensity.

Western blot analysis for recombinant protein fragments of FCR3S1.2-*var1*

The *var1* of FCR3S1.2 is a rather small protein and thus, it was intended to be expressed as 4 recombinant fragments. Previously, 3 fragments could be cloned and expressed (Fragment 2, 3 and 4). Fragments 2 and 3 were isolated from the insoluble pellet, whereas fragment 4 was found in the soluble protein fraction. However, the very N-terminal part containing the DBL1α domain could not be expressed in reasonable amounts, even after several rounds of trying and dividing fragment 1 into smaller pieces. In general, protein yields were very low. Attempts to transform three alternative *E.coli* strains were not successful. Thus, it was not possible to perform ELISA with such little protein amounts and therefore we decided to analyze their intensity of recognition by Western Blot.

Analysis of signal intensity showed that in adults antibody titers were significantly higher for fragment 2 ($p<0.001$) and 4 ($p<0.001$) compared to children whereas recognition of fragment 3 was very similar in these two groups (Figure 13a). Difference in recognition was also significant when antibody responses against fragment 2 ($p=0.011$) and 4 ($p=0.02$) where compared in children with asymptomatic and severe malaria (Figure 13b). However, whereas for fragment 2 intensity of recognition was higher in asymptomatic children, fragment 4 was better recognized in children with severe malaria. For children with severe malaria, the geographical origin of the sera did not seemed to play a role. However, this was not the case for children with asymptomatic malaria (Figure 13c): on the one hand children from PNG showed significantly higher signal intensities than asymptomatic children from Tanzania for fragment 2 ($p=0.013$) and fragment 3 ($p=0.009$). On the other hand, asymptomatic children from Tanzania had significantly higher immune responses for fragment 4 ($p=0.0017$) than their counterparts from PNG.

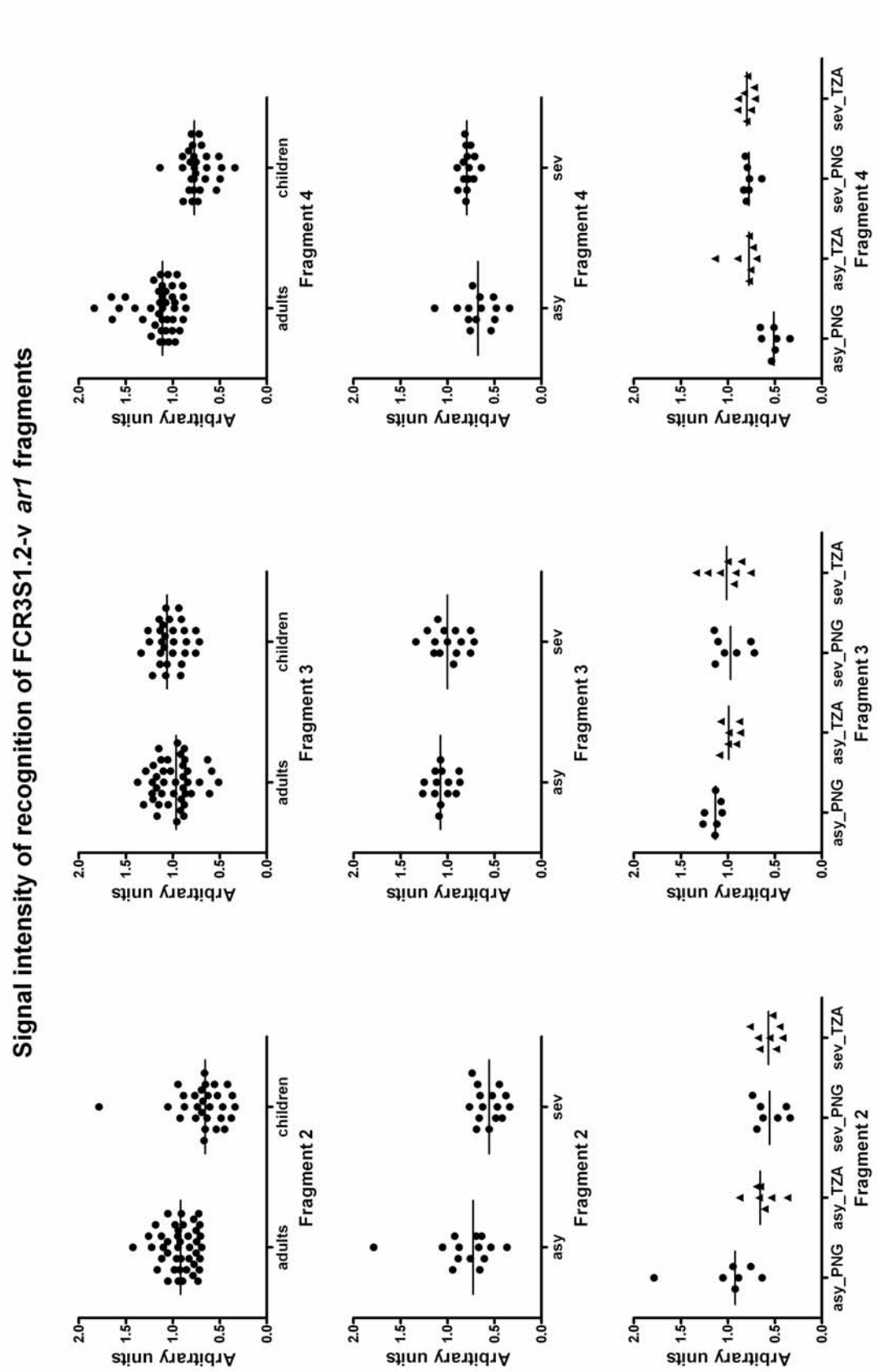


Figure 13. Signal intensity of recognition for recombinant FCR3S1.2-var1 fragments. (a) shows the comparison between adults and children. In (b) differences in recognition by children with asymtpomatic and severe malaria are shown and in (c) differences in between asymptomatic and severe children are shown if children from PNG● and Tanzania▲ were anaylzed

Discussion

The ability of *P. falciparum* to attach to endothelial receptors in various tissues and to undergo antigenic variation poses a tremendous pressure on the host immune system. Since PfEMP1 mediates both of these evasion strategies it represents a key virulence factor. Despite the fact that the *var* gene repertoire is huge, the overall architecture of PfEMP1 proteins seems to be conserved among different parasite isolates. Adhesion traits of several domains have already been mapped[12] and a few specific domains were even implicated in severe malaria.[32,3,33,34,23] Previously, a serological study using recombinant protein fragments of the 3D7 group A *var* gene PF11_0008 revealed that antibodies against the CIDR2β domain might be involved in protection against malaria episodes in Tanzania[30]. Another study also using sera from Tanzania found out that antibodies against the CIDR1α of *var4*, predict protection against malarial anemia and febrile episodes[29]. The latest publication reports a positive correlation between the presence of anti-DBL1α antibodies and the protection from subsequent clinical malaria in individuals being parasite negative at the time of bleed in Kenya.[31]

In this part of the thesis, we tried to identify immunodominant epitopes of two representative *var* genes which have been associated with severe malaria previously.[16,35] For that purpose, fragments of about 700-1500bp in size were expressed in *E.coli*. ELISA screening with sera of different origin should provide information about the frequency of recognition of the recombinant fragments and potentially reveal epitopes involved in the generation of protective antibodies.

Expression of recombinant fragments in *E.coli*

Cloning of all 10 PFD1235w/*var4* fragments was successful, however expression was only possible for 4 of them and protein yield was generally very low. Similarly, only 3 of 4 overlapping fragments could be expressed of the rather short *var1* of the rosetting strain FCR3S1.2. Using *E.coli* as heterologous expression system has both advantages and disadvantages. In the first place, the transfection and cultivation of *E.coli* is simple and fast and a cheap method to produce large amounts of antigens for biological examinations. However, malaria antigens are among the most difficult proteins to express with *in vitro* methods because of their extreme genetic codon usage.[36] Codons like arginine, leucine, isoleucine and proline are frequently found in *P. falciparum* but are rarely present in *E.coli*, which will inhibit the translation process and result in truncated products of the desired protein. This might be a possible explanation for the multiple bands of different molecular

weights obtained after purification. Additional transcriptional start and termination sites in *E.coli* and the fact that *E.coli* also contains some histidine rich proteins might also account for these additional bands. As expression was done in the absence of protease inhibitors protein products might in principle also be due to degradation by bacterial proteases. However, all but one of the recombinant fragments were isolated from inclusion bodies which are confined structures used by bacteria to avoid toxic effects of accumulating aggregates of incorrectly folded proteins. These insoluble protein contents needed to be isolated under very harsh conditions with 8M urea and therefore all potentially active proteases would have been denaturated. In order to increase the solubility, expression was only initiated at post-log growth phase as suggested by Flick *et al.*[36]. However, the desired effect of increased soluble protein fractions was not achieved. Expression at room temperature to allow for proper folding did not help to increase solubility either. Further attempts to improve the protein quality and yield by transformation of *E.coli* strain BL21 Condon plus RIL which contains an additional plasmid that codes for the rare t-RNAs, were not successful (data not shown). Another option to generate correctly folded proteins would be to change the expression system. Several organisms like baculovirus[37,38], Dictyostelium discoideum[39,40,41,42], Xenopus[41], yeast[43,44,45,46], transgenic tobacco plants[47], goats[48,49] and mice[50] were used to produce recombinant malaria antigens. However, toxicity, low protein yields or time-consuming codon optimization and removal of glycosylation sites renders these expression systems cost-ineffective in large-scale production.

In conclusion, there is no standard protocol for the production of recombinant antigens which are soluble, intact and functional. Temperature, pH, density, amino acid composition and cell line, might influence protein expression and therefore optimization has to be done for each recombinant protein individually.

Recognition of recombinant PFD1235w/*var4* fragments by naturally exposed individuals

Investigating the recognition frequency using sera of semi-immune adults

In a first experiment, 4 recombinant fragments expressed in amounts sufficient to perform ELISA, were screened for recognition by 48 semi-immune adults from PNG. Overall, fragments 5,6 and 8 corresponding to the DBL3β, C2-DBL4γ and the DBL4δ domain of PFD1235w/*var4* were well recognized by adults. In contrast, fragment 1 representing the DBL1α domain was only poorly recognized which might be due to low antigenicity or that this particular DBL domains is not presented in the parasite population of PNG.

Furthermore, DBL1α domains might contain several structural epitopes which would have been destroyed during purification under denaturing conditions. The remaining linear epitopes might not be as immunogenic as the ones present in other recombinant fragments. However, these hypotheses are highly speculative.

Recognition was found to be selective. The variation in recognition, e.g. sera recognizing some recombinant proteins strongly and others not at all, indicates that these differences are indeed due to specific recognition and not to overall low or high responsiveness. Similar findings were also reported by Mackintosh *et al.*[31] for *var1* of the A4 paraiste strain. Since approximately half of our study samples came from female and half from male donors, we tested whether recognition of the recombinant proteins was sex-dependent. Fragments 6 and 8 were equally well recognized by both sexes but there was a slight difference in antibody titers between men and women for fragment 1 and 5. Whereas females tended to recognize fragment 1 more frequently, the frequency of recognition for fragment 5 was slightly higher in serum samples of male donors. However, these differences were not significant. Gender-specific recognition of malaria antigens has only been observed in PAM. Multigravid women generate antibodies blocking the adhesion of iRBC to placental CSA whereas sera collected from men are not able to recognize these parasites[51,52] or the level and prevalence of antibodies are substantially lower.[53] Since PFD1235w/*var4* neither shows much homology to the *var2csa* nor does it contain a DBL3γ or any DBLx or DBLε domains implicated in PAM previously[34,22,23], gender specific differences in recognition were not expected.

Investigating the recognition frequency using sera obtained from children

In order to identify potential epitopes playing a role in the production of protective antibodies, recognition of sera was assessed and compared to data found previously in adults. In sera collected longitudinally in Tanzania, antibodies titers in children were considerably lower than in adults which was expected since exposure time of children has been much shorter. Fragments 1, 6 and 8 were similarly recognized in children as in adults. However, fragment 5 which was well recognized in adults was only poorly recognized in children. Since children and adult sera originated from different continents this difference in recognition was analyzed again with a small sample set of children sera from PNG. Results of this second study led to the same finding and further investigations with sera from children with asymptomatic and severe malaria showed that only children with asymptomatic malaria were able to recognize fragment 5, whereas children with severe malaria did not. Fragment 5 corresponds to a DBLβ-C2 headstructure, which is

only present in 24% of the 3D7 *var* genes and almost exclusively found in *var* group A previously associated with severe malaria. However, difference in recognition was only borderline significant and further analysis would be needed to test whether fragment 5 really elicits protective antibodies.

In a previous study by Joergensen *et al*.[54] the very same domain was the least recognized part of the three PFD1235w/*var4* domains tested with Tanzanian samples, with similar frequency of recognition in children aged 2-4 years and adolescent between 15-19 years of age. Whereas our proteins were all purified under denaturing conditions, Joergensen *et al.*[54] managed to express and purify their fragments with glutathione transferase tags which are highly soluble. Therefore, these differences in recognition might simply be due to differences in protein preparation.

Samples from baseline and 6 months follow-up samples were compared in order to analyze potential dynamics in the antibody repertoire of children from Tanzania. We observed a slight increase in the frequency of recognition for all recombinant fragments. However, when EPT of baseline and follow-up sample were compared, an increase in antibody titer could not be observed anymore for all fragments. Furthermore, when analyzing paired samples individually, antibody titers against fragments 1, 6 and 8 increased in some children during the six months period whereas those of others decreased. This might be further evidence that antibodies against these domains do not contribute to clinical immunity.

Peptide ELISA on 9 synthetic peptides of PFD1235/*var4* using semi-immune adult and children sera

Since only 4 of 10 overlapping fragments spanning the whole PFD1235w/*var4* could be expressed as recombinant proteins, 9 biotinylated synthetic peptides were analyzed to bridge protein gaps at least in part.

Recognition of synthetic fragments was tested in ELISA with the same set of adult and children sera used before. Signal intensities were much lower when compared to recombinant proteins which was expected since recombinant proteins were much bigger in size and therefore contained more epitopes accessible for antibodies present in the different sera. The overall recognition pattern was similar in adults and children with peptides 2.2 and 7 being well recognized, peptides 2.1 4.1, 4.2 and 8 were morderately recognized and peptides 3 and 10 were the least recognized proteins in both adults and children. Fragment 9 was the best recognized peptide in children whereas in adults this

peptide was only moderately recognized. Peptides 2.1, 2.2, 4.1, 4.2, 8 and 10 were found to be significantly better recognzied by adults than by children which might be explained by prolonged exposure time of adults compared to children. However, only the recognition of peptides 2.1 and 4.1 remained significantly different when children with asymptomatic and severe malaria were compared, rendering these two peptides potential candidates which might play a role in the generation of protective antibodies. This difference in recognition was found to be independent of the geographical origin of the sera since recognition patterns of children from Tanzania and PNG were similar.

In summary, we detected a significant difference in recognition for two of the ten PFD1235w/*var4*-derived peptides. Synthetic peptides 2.1 and 4.1 were significantly better recognized by asymptomatic children when compared to children with severe malaria which might indicate their potential role in the generation of protective antibodies. However, the number of tested sera was very small and therefore further studies are required to confirm these findings.
We did not observe significant differences in recognition (or any recognition at all) for the remaining peptides which might have several reasons:

Firstly, the peptides were chosen by screening the *var* gene for regions of antigenicity, hydrophobicity and surface accessibility. These calculations are based on amino acid composition only and since to date no 3D structure of complete PfEMP1 proteins exists, we cannot be sure that the chosen peptides are really accessible to antibodies. Secondly, completely different results might have been obtained if peptides of other regions would have been chosen or if peptides would have been longer or shorter than the chosen 30 amino acids. Peptides of this size probably contain several epitopes improving their antigenicity but they might also form stable secondary structures, which might not necessarily mimic the ones in the native protein. Thirdly, the peptides chosen - even if exposed - might be located in regions of the protein which are simply not important for the generation of protective antibodies. Fourthly, antibody titers were generally very low and therefore, recognition of peptides might not have been sufficiently high to make a potential difference visible.

3D7 has been in culture for more than 20 years and probably none of the donors tested has ever been infected with this strain. Nevertheless, some recombinant fragments and synthetic peptides were recognized by numerous adults and children sera which clearly

indicates, that there is serological cross-reactivity between PfEMP1 proteins expressed by different parasites. The selective recognition of only some but not all fragments or peptides further indicates the exchange of specific domains between *var* genes rather than the preservation of full-length proteins within parasite populations[54].

Western blot with 3 recombinant fragments of FCR3S1.2-*var1* using semi-immune adults' and children's sera

Recombinant fragments of FCR3S1.2-*var1* were tested for their recognition by semi-immune adults' sera and sera from children with asymptomatic and severe malaria and signal intensities were compared. Whereas signal intensities for fragment 3, corresponding to the DBL2δ domain, were similar for adults and children, fragments 2 and 4, corresponding to domains CIDR1α and CIDR2β, seemed to be recognized with higher antibody titers than the ones found in children. This pattern of signal intensity hold true when children with asymptomatic and severe malaria were compared and was found to be independent of the geographical origin of children sera for children with severe malaria. However, when asymptomatic children from PNG and Tanzania were analyzed separately, fragment 2 and 3 were significantly better recognized by children from PNG whereas for fragment 4 it was the other way round which might indicate that in PNG epitopes contained in fragment 2 and 3 are implicated in the generation of protective antibodies whereas in Tanzania epitopes of fragment 4 are more important for the acquisition of semi-immunity. However, due to the many comparative analyses, the finding of significant differences might simply be due to chance.

In the present study, the very N-terminal part of the FCR3S1.2-*var1*, corresponding to the DBL1α domain could not be expressed which would have been important as FCR3S1.2 is a rosetting strain and rosetting was shown to be mediated by this particular DBL domain. It would have been interesting to see whether this region was significantly better recognized in adults compared to children and especially in children with asymptomatic malaria when compared to children with severe malaria.

It should be worth mentioning that such a study would rather focus on samples from Tanzania where rosetting was associated with severe malaria. Contrarily, in PNG, such an association could not be found due to a frequent mutation in the CR1 receptor, which was found to be the binding partner of the DBL1α domain. Therefore, one would not expect a singificant difference in recognition between asymptomatic and severe children from PNG if binding was mdiated by CR1. However, the DBL1α domain for FCR3S1.2-*var1* was

shown to bind to negatively charged sulfated glycosaminoglycans (GAGs)[35], e.g. heparan sulfate on RBCs and therefore samples from PNG might still be included.

I would like to conclude by pointing out that using recombinant fragments as well as synthetic peptides can only give us a rough idea about the presence or absence of specific antibodies and their corresponding titers. Probably most of the detected antibodies are directed against linear epitopes since proper folding of eukaryotic proteins by prokaryotic organisms seems very unlikely. Therefore, we might miss part of the protective antibody repertoire as soon as the antibodies are directed against structural epitopes.

References

1. COHEN, S., McGREGOR, I.A. & CARRINGTON, S. Gamma-globulin and acquired immunity to human malaria. *Nature* **192**, 733-7(1961).
2. Edozien, J.C., Gilles, H.M. & Udeozo, I.O.K. ADULT AND CORD-BLOOD GAMMA-GLOBULIN AND IMMUNITY TO MALARIA IN NIGERIANS. *The Lancet* **280**, 951-955(1962).
3. Newbold, C.I. et al. PfEMP1, polymorphism and pathogenesis. *Ann Trop Med Parasitol* **91**, 551-7(1997).
4. Bull, P.C. & Marsh, K. The role of antibodies to Plasmodium falciparum-infected-erythrocyte surface antigens in naturally acquired immunity to malaria. *Trends Microbiol* **10**, 55-8(2002).
5. David, P.H. et al. Parasite sequestration in Plasmodium falciparum malaria: spleen and antibody modulation of cytoadherence of infected erythrocytes. *Proc Natl Acad Sci U S A* **80**, 5075-9(1983).
6. Reeder, J.C. & Brown, G.V. Antigenic variation and immune evasion in Plasmodium falciparum malaria. *Immunol Cell Biol* **74**, 546-54(1996).
7. Bull, P.C. et al. Antibody recognition of Plasmodium falciparum erythrocyte surface antigens in Kenya: evidence for rare and prevalent variants. *Infect. Immun* **67**, 733-9(1999).
8. Giha, H.A. et al. Antibodies to variable Plasmodium falciparum-infected erythrocyte surface antigens are associated with protection from novel malaria infections. *Immunol Lett* **71**, 117-26(2000).
9. Smith, J.D. et al. Classification of adhesive domains in the Plasmodium falciparum erythrocyte membrane protein 1 family. *Mol Biochem Parasitol* **110**, 293-310(2000).
10. Kraemer, S.M. et al. Patterns of gene recombination shape *var* gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics* **8**, 45(2007).
11. Su, X.Z. et al. The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes. *Cell* **82**, 89-100(1995).
12. Smith, J.D. et al. Decoding the language of *var* genes and Plasmodium falciparum sequestration. *Trends Parasitol* **17**, 538-45(2001).

13. Klein, M.M. et al. The cysteine-rich interdomain region from the highly variable plasmodium falciparum erythrocyte membrane protein-1 exhibits a conserved structure. *PLoS Pathog* **4**, e1000147(2008).
14. Rottmann, M. et al. Differential expression of *var* gene groups is associated with morbidity caused by Plasmodium falciparum infection in Tanzanian children. *Infect. Immun* **74**, 3904-11(2006).
15. Kaestli, M. et al. Virulence of malaria is associated with differential expression of Plasmodium falciparum *var* gene subgroups in a case-control study. *J. Infect. Dis* **193**, 1567-74(2006).
16. Jensen, A.T.R. et al. Plasmodium falciparum associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. *J. Exp. Med* **199**, 1179-90(2004).
17. Kirchgatter, K. & Portillo, H.D.A. Association of severe noncerebral Plasmodium falciparum malaria in Brazil with expressed PfEMP1 DBL1 alpha sequences lacking cysteine residues. *Mol. Med* **8**, 16-23(2002).
18. Gamain, B. et al. Pregnancy-associated malaria: parasite binding, natural immunity and vaccine development. *Int J Parasitol* **37**, 273-83(2007).
19. Rogerson, S.J. & Boeuf, P. New approaches to pathogenesis of malaria in pregnancy. *Parasitology* **134**, 1883-93(2007).
20. Duffy, P.E. & Fried, M. Plasmodium falciparum adhesion in the placenta. *Curr Opin Microbiol* **6**, 371-6(2003).
21. Staalsoe, T. et al. Variant surface antigen-specific IgG and protection against clinical consequences of pregnancy-associated Plasmodium falciparum malaria. *The Lancet* **363**, 283-289(2004).
22. Bir, N. et al. Immunogenicity of Duffy binding-like domains that bind chondroitin sulfate A and protection against pregnancy-associated malaria. *Infect Immun* **74**, 5955-63(2006).
23. Fernandez, P. et al. Var2CSA DBL6-epsilon domain expressed in HEK293 induces limited cross-reactive and blocking antibodies to CSA binding parasites. *Malar J* **7**, 170(2008).
24. Newbold, C.I. et al. Plasmodium falciparum: the human agglutinating antibody response to the infected red cell surface is predominantly variant specific. *Exp Parasitol* **75**, 281-92(1992).

25. Gamain, B., Miller, L.H. & Baruch, D.I. The surface variant antigens of *Plasmodium falciparum* contain cross-reactive epitopes. *Proc Natl Acad Sci U S A* **98**, 2664-9(2001).
26. Gratepanche, S. et al. Induction of crossreactive antibodies against the *Plasmodium falciparum* variant protein. *Proc Natl Acad Sci U S A* **100**, 13007-12(2003).
27. Elliott, S.R. et al. Antibody recognition of heterologous variant surface antigens after a single *Plasmodium falciparum* infection in previously naive adults. *Am J Trop Med Hyg* **76**, 860-4(2007).
28. Mo, M. et al. The C-terminal segment of the cysteine-rich interdomain of *Plasmodium falciparum* erythrocyte membrane protein 1 determines CD36 binding and elicits antibodies that inhibit adhesion of parasite-infected erythrocytes. *Infect Immun* **76**, 1837-47(2008).
29. Lusingu, J.P.A. et al. Levels of plasma immunoglobulin G with specificity against the cysteine-rich interdomain regions of a semiconserved *Plasmodium falciparum* erythrocyte membrane protein 1, VAR4, predict protection against malarial anemia and febrile episodes. *Infect Immun* **74**, 2867-75(2006).
30. Magistrado, P.A. et al. Immunoglobulin G antibody reactivity to a group A *Plasmodium falciparum* erythrocyte membrane protein 1 and protection from P. falciparum malaria. *Infect Immun* **75**, 2415-20(2007).
31. Mackintosh, C.L. et al. Failure to respond to the surface of *Plasmodium falciparum* infected erythrocytes predicts susceptibility to clinical malaria amongst African children. *Int. J. Parasitol* (2008).doi:S0020-7519(08)00142-2
32. Turner, G.D. et al. An immunohistochemical study of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. *Am J Pathol* **145**, 1057-69(1994).
33. Fernandez-Reyes, D. et al. A high frequency African coding polymorphism in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet* **6**, 1357-60(1997).
34. Buffet, P.A. et al. *Plasmodium falciparum* domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection. *Proc Natl Acad Sci U S A* **96**, 12743-8(1999).
35. Chen, Q. et al. Identification of *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) as the rosetting ligand of the malaria parasite P. falciparum. *J. Exp. Med* **187**, 15-23(1998).

36. Flick, K. et al. Optimized expression of *Plasmodium falciparum* erythrocyte membrane protein 1 domains in *Escherichia coli*. *Malar. J* **3**, 50(2004).
37. Pang, A.L.Y. et al. In vivo expression and immunological studies of the 42-kilodalton carboxyl-terminal processing fragment of *Plasmodium falciparum* merozoite surface protein 1 in the baculovirus-silkworm system. *Infect Immun* **70**, 2772-9(2002).
38. Pizarro, J.C. et al. Crystal structure of a Fab complex formed with PfMSP1-19, the C-terminal fragment of merozoite surface protein 1 from *Plasmodium falciparum*: a malaria vaccine candidate. *J Mol Biol* **328**, 1091-103(2003).
39. Fasel, N. et al. Dictyostelium discoideum as an expression host for the circumsporozoite protein of *Plasmodium falciparum*. *Gene* **111**, 157-63(1992).
40. van Bemmelen, M.X. et al.Expression and one-step purification of *Plasmodium proteins* in dictyostelium. *Mol Biochem Parasitol* **111**, 377-90(2000).
41. Muhia, D.K. et al. Multiple splice variants encode a novel adenylyl cyclase of possible plastid origin expressed in the sexual stage of the malaria parasite *Plasmodium falciparum*. *J Biol Chem* **278**, 22014-22(2003).
42. Naudé, B. et al. Dictyostelium discoideum expresses a malaria chloroquine resistance mechanism upon transfection with mutant, but not wild-type, *Plasmodium falciparum* transporter PfCRT. *J Biol Chem* **280**, 25596-603(2005).
43. Kocken, C.H.M. et al.High-level expression of the malaria blood-stage vaccine candidate *Plasmodium falciparum* apical membrane antigen 1 and induction of antibodies that inhibit erythrocyte invasion. *Infect Immun* **70**, 4471-6(2002).
44. Pan, W. et al. Fusion of two malaria vaccine candidate antigens enhances product yield, immunogenicity, and antibody-mediated inhibition of parasite growth *in vitro*. *J Immunol* **172**, 6167-74(2004).
45. Baruch, D.I. et al. Immunization of Aotus monkeys with a functional domain of the *Plasmodium falciparum* variant antigen induces protection against a lethal parasite line. *Proc. Natl. Acad. Sci. U.S.A* **99**, 3860-5(2002).
46. Miles, A.P. et al.Large-scale purification and characterization of malaria vaccine candidate antigen Pvs25H for use in clinical trials. *Protein Expr Purif* **25**, 87-96(2002).
47. Turpen, T.H. et al. Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. *Biotechnology (N Y)* **13**, 53-7(1995).
48. Behboodi, E. et al. Health and reproductive profiles of malaria antigen-producing transgenic goats derived by somatic cell nuclear transfer. *Cloning Stem Cells* **7**, 107-18(2005).
49. Nicholls, H. Milking goats for malaria vaccine. *Drug Discov Today* **9**, 248(2004).

50. Stowers, A.W. et al. A recombinant vaccine expressed in the milk of transgenic mice protects Aotus monkeys from a lethal challenge with *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **99**, 339-44(2002).
51. Fried, M. et al. Maternal antibodies block malaria. *Nature* **395**, 851-852(1998).
52. Khattab, A. u. a. Analysis of IgG with specificity for variant surface antigens expressed by placental *Plasmodium falciparum* isolates. *Malar J* **3**, 21(2004).
53. Beeson, J.G. et al. Antibodies among men and children to placental-binding *Plasmodium falciparum*-infected erythrocytes that express var2csa. *Am. J. Trop. Med. Hyg* **77**, 22-8(2007).
54. Joergensen, L. et al. 3D7-Derived *Plasmodium falciparum* erythrocyte membrane protein 1 is a frequent target of naturally acquired antibodies recognizing protein domains in a particular pattern independent of malaria transmission intensity. *J. Immunol* **178**, 428-35(2007).

..

Chapter 4
Application of capillary electrophoresis sizing technique as new *var* gene genotyping tool

Application of capillary electrophoresis sizing technique as new *var* gene genotyping tool

Nicole Falk, Kathrin Witmer, Hans-Peter Beck*

Affiliation of authors:
Swiss Tropical Institute, Socinstrasse 57, 4002 Basel, Switzerland

* Corresponding author

Email addresses of authors: nicole.falk@stud.unibas.ch
kathrin.witmer@unibas.ch
hans-peter.beck@unibas.ch

Word count manuscript: 3191
Word count abstract: 443

Abstract

Background: *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) is encoded by approximately 60 *var* genes per haploid genome and the total *var* gene repertoire in the population is highly diverse. Switching from one PfEMP1 to another has been shown to not only alter the protein being expressed at the erythrocyte surface but also to cause functional changes resulting in altered adhesive phenotypes which potentially influence disease outcome. *var* gene dynamics and switching of *var* genes at the time of transition from an asymptomatic to a mild or severe malaria status is of major interest in patients naturally infected with *P. falciparum*. To identify switches and to record the dynamics of *var* gene expression in naturally occurring infections specific *var* cDNA is currently synthesized, cloned, and subsequently sequenced. This is tedious, laborious, expensive, and most probably biased. Because the *var* gene repertoire is vast the outcome of this strategy depends highly on the number of clones analyzed, which can sum up to thousands of clones for sequencing even in small studies involving few patients. Technical limitations further decrease the chance of analyzing a representative population of expressed *var* genes. Therefore, a new and efficient technology is urgently needed.

Methods: Here we present an innovative genotyping tool to potentially replace the tedious and error prone cloning and sequencing technique. The approach is based on capillary electrophoresis and fragment sizing using the GeneMapper® program which initially has been tested to study *var* gene transcription in a 3D7 *in vitro* culture. Subsequently, we used the technique on field samples to validate its applicability. The target sequence was the DBL1α domain, which is highly diverse in sequence and size, and GeneMapper® was evaluated by comparing sizing data to data obtained by cloning and sequencing during a previous study.

Results: GeneMapper® sizing is highly accurate with a mean deviation of ~1bp from the size determined by sequencing. In samples of 42 infected children GeneMapper® showed a high consistency with sequencing data, and 83.2% of clones identified by sequencing were also detected by GeneMapper®. Within the overall population of *var* genes, a significant proportion cannot be distinguished because the analyzed DBL1α domains were identical in size. However, in only four children two *var* gene sequences were present which could not be identified as different by this technique because the respective domains were of identical size. Furthermore, GeneMapper® detected many additional fragments which were not observed by cloning and sequencing.

Conclusions: Despite some limitations, GeneMapper® greatly facilitates studies of *var* gene transcription and dynamics because switches in expression can rapidly be detected. The technique is straight forward, and costs and labour are a fraction of the approach using cloning and sequencing.

Introduction

Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1) is a major virulence factor. It is encoded by approximately 60 *var* genes albeit only one type of PfEMP1 is dominantly expressed at the erythrocyte membrane at any one time [1]. According to their chromosomal location and their 5' untranslated region (UTR), *var* genes can be divided into 3 major groups (*var* groups A, B and C) [2]. The extracellular part of PfEMP1 molecules is mainly composed of several Duffy binding like (DBL) and cysteine-rich interdomain region (CIDR) domains. The number, location and sequence of DBL and CIDR domains varies significantly among different PfEMP1 proteins. PfEMP1 is responsible for sequestration by attaching iRBCs to various host cell receptors resulting in the occlusion of tissue capillaries and contributes significantly to the pathogenicity of *P. falciparum*. PfEMP1 is a surface exposed protein and is targeted by the host immune system. To escape the adaptive immune system, PfEMP1 undergoes antigenic variation by switching the expressed *var* gene(s). Several attempts have been made to examine the composition and the dynamics of *var* gene transcripts in longitudinal studies *in vitro* and *in vivo*[3,4,5,6,7,8] and studies investigating the speed of antigenic switching have shown that switching rates might vary significantly from 0.025% to 18%. Other studies have been conducted to identify *var* genes or groups of *var* genes involved in severe malaria.[9,10,11,12,13,14]

Previously, analyzing *var* gene transcription required the isolation of RNA and subsequent reverse transcription into cDNA. This was then cloned and in most studies between 50 and 100 clones were sequenced for each sample. This approach has been seriously hampered by the fact that within a patient several *var* transcripts were found and in order to capture the true diversity of expression such large number of clones had to be sequenced. The isolation of RNA can be a major undertaking due to low parasite densities, and subsequent cloning and sequencing increases the difficulties of this appraoch because of differences in the performance of primers, differences in PCR efficiency, and the limitations in cloning and ligation. To overcome these shortcomings, we have used an automatic sizing technique based on capillary electrophoresis to distinguish different *var* gene domains by their specific sizes. GeneMapper® analysis software was tested as a new genotyping tool to potentially replace the tedious and error prone cloning and sequencing approach.

GeneMapper® is a semi-automatic capillary-electrophoresis-based genotyping tool exploiting sequence length polymorphism. Fluorescently labeled primers render PCR products detectable for a laser and according to an internal size standard, individual PCR fragments are assigned a specific length. This technique has been extensively used for

polymorphic marker genes such as *msp2* [15,16] but also for polymorphic marker genes of *P. vivax*.[17]Since *var* genes also exhibit length polymorphism we established GeneMapper® analysis on the amplified DBL1α domain with fluorescently-labeled primers. Initially, we tested this approach on a 3D7 culture monitored over a long period of time and finally analyzed samples from a case-control study in Papua New Guinea (PNG). With these samples we were able to compare the GeneMapper® approach with data obtained from cloning and sequencing (Falk 2008, submitted).

METHODS

GeneMapper® Software (Applied Biosystems, version 3.7) is an analytical tool that exploits sequence length polymorphism. Target sequences are PCR amplified using fluorescently-labeled primers and products are detected by a laser after capillary electrophoresis. DNA fragments of defined length serve as internal size standard for the creation of a size calibration curve. According to the manufacturer this allows for a size differentiation of up to 1bp. The length of the labeled PCR products is then determined using this calibration curve and each peak is assigned a defined size. *var* gene domains also exhibit size polymorphism to a certain degree, and a fluorescently 6-FAM-labeled DBL1α reverse primer was used in combination with 'tailed' DBL 1α forward primer. A specific 7-basepair tail of the forward primer increases the likelihood of the addition of a non-template A (Figure 1) and reduces ambiguity in sizing of fragments.[18]

P. falciparum in vitro culture

To evaluate the Genemapper® method, subsequent samples of a 3D7 *in vitro* culture were analyzed. The culture was monitored for *var* gene expression for a period of 168 days and in total 18 samples of 3-21 days intervals were taken. For each time point 10ml of 3D7 parasite culture at 5% hematocrit were harvested at ring stages, washed once with PBS, pelleted, dissolved in 3ml TRIzol (Invitrogen), incubated for 5 min at 37°C and stored at -80°C.

Isolation of RNA and reverse transcription

The isolation of (full-length *var*) mRNA and reverse transcription was performed as described elsewhere[7]. Briefly, total RNA of ring-stage parasites was extracted using Trizol in accordance with the manufacturer's protocol. RNA was treated twice with 3U of RQ1 DNase (Promega). To obtain full length *var* transcripts, RNA was incubated with biotinylated oligonucleotides complementary to the conserved exon 2. 200µg Dynabeads® M-280 Streptavidin were added to the RNA. After washing, reverse transcription (RT) was performed on the captured full-length *var* transcripts using random primers and Sensiscript™ reverse transcriptase (Qiagen). A sample without reverse transcriptase was included as a contamination control (RT(-)). After RT, cDNA was treated with RNase A. All cDNA samples were checked by PCR with degenerated DBL1α primers (DBLα-5' and DBLα-3') as described previously.[7]

Polymerase chain reaction (PCR) with samples obtained from culture

To amplify the target sequence, a PCR was set up consisting of 2U FIREPol® DNA polymeraseI, 10x PCR buffer BD (Solis BioDyne), containing 80mM Tris pH 9.4, 20mM $(NH_4)_2SO_4$, 1.5mM $MgCl_2$, 500nM of forward (5'-GC ACGAAGTTTTGCAGATAT(A/T)GG-3') and reverse primer (5'-AA(A/G)TCTTC(T/G)GCCCATTCCTCGAACCA-3') each and 200µM dNTPs in a volume of 30µl. 1µl of cDNA was added to the reaction mix. PCR conditions were initially 5 min at 94°C, followed by 30 cycles 95°C, 30 sec, 54°C, 1 min, and 68°C, 40 sec. Amplification was stopped after a final elongation step at 68°C for 7 min. PCR products were run on a 1% agarose gel and positive RT(-) samples were discarded whilst remaining samples were amplified with GeneMapper®- specific primers. PCR conditions were identical except for primers which were a tailed forward (5'-7bp-tail- GC ACGAAGTTTTGCAGATAT(A/T)GG-3') and a 6-FAM-labeled DBL1α reverse primer (5'-6-FAM- AA(A/G)TCTTC(T/G)GCCCATTCCTCGAACCA-3').

Sample preparation for GeneMapper® analysis of samples obtained from culture

Since higher salt concentrations disturb analyses on automated sequencers, labeled PCR samples were precipitated over night at -20°C with 2.5 volumes of absolute ethanol and 1/10 volume of 3M NaAc pH 5.2. After washing with 100µl 75% ethanol pellets were resolved in 25µl 5mM Tris-HCl pH7. DNA concentration was estimated on a 1% agarose gel and samples were diluted to ~2ng/µl with 5 mM Tris-HCl. 10µl of diluted DNA were added to 10µl of a 1:40 dilution of GeneScan™ 500 ROX™ size standard (Applied Biosystems) and dried over night. For electrophoresis on a 96-capillary sequencer, the samples were sent to the Genomics Core Facility of the MRC Clinical Science Centre in London. After electrophoresis results were analyzed with GeneMapper® software version 3.7 (Applied Biosystems).

Polymerase chain reaction and GeneMapper® analysis for field samples

1µl of purified primary PCR product previously generated during a case control study in Papua New Guinea (PNG) (Falk et al. 2008, manuscript submitted) served as template. In brief, primary PCR products were generated using subgroup-specific forward primers (upsA-5'-AACTTACCATAAATTATCATCAAA-3',upsB-5'ATGTAATTGTTGTTTTTTTTTTT GTTAGAATATTTAAA-3' or upsC-5'- CACATATA(A/G)TACGACTAAGAAACA-3') and the DBL1α reverse primer (see above) with the following conditions: Initial denaturation 94°C for 5 min followed by 25 cycles of 30 s at 95°C, 1 min at 52°C, and 1 min at 64°C. A final elongation of 7 min at 64°C was performed at the end of the reaction.

PCR products were diluted 10- to 40-fold in water. 2.5µl of this mixture were added to 10µl of a 1:40 dilution of GeneScan™ 500 ROX™ size standard and treated as described above.

To evaluate GeneMapper®, obtained fragment sizes of DBL1α PCR products ranging from 373 to 496 bp were either compared to sequence data available at PlasmoDB for 3D7 or to sequencing data obtained from these samples. Peaks were considered when above an artificial cut-off of 100 or 500 fluorescent units for culture or field isolates, respectively. Peaks of low fluorescent intensity in close proximity (3bp) to a high density peak were considered as “shoulder peaks” and excluded from further analysis. The expected length of the DBL1α domains was calculated by counting basepairs between and including primers used for amplification plus 8 basepairs due to the 7-basepair tail and the added nucleotide.

RESULTS

Virtual discrimination power of GeneMapper® in 3D7

Using the published sequence of the 3D7 genome we virtually determined the number of differently sized DBL1α fragments which would be amplified with our primer set. Of 60 published *var* sequences we were able to unequivocally distinguish 39 (65%) when using upstream-specific forward primers for ups A, B and C (Table1). In 4 cases the reverse primer did not match perfectly and it is likely that these *var* genes may not amplify. The *var* gene PFL0030c (*var2*) does not contain a DBL1α domain and cannot be detected.

In the ups A group DBL1α domains fell into 8 different size groups of which 3 contained more than 1 sequence (3x2). For ups B and ups C there were 21 and 10 different DBL1α size groups of which 8 (1x6, 1x4, 6x2) and 3 (3x2) contained more than one sequence, respectively. Sequences of group B/C or B/A were amplified with ups B forward primers.

Monitoring of 3D7 *in vitro* culture

A 3D7 *in vitro* culture was monitored over a time period of 168 days. Samples of 18 time points were taken in 3 to 20 days intervals and analyzed by GeneMapper®. Figure 2 shows an electropherogram, the typical output format of GeneMapper®.

A dominant peak of 422bp was observed throughout the complete monitoring period (Figure 3). The peak at 422bp potentially could have been derived from 3 *var* genes: PFD0995c and PFD1000c, both ups C, and PF07_0050 belonging to the intermediate *var* group upsB/C (Table 1), since the DBL1α domains of these *var* genes are identical in size. In this case, these fragments could not be differentiated by subgroup because DBL1α-specific primers were used only. Besides the 422bp fragment, fragments with the lengths of 413bp, 416bp and 419bp were observed frequently. Overall, at all time points, more than one peak was observed and the number of fragments varied between 2 -18.

GeneMapper® analysis of field samples

GeneMapper® analysis was applied to field samples of 42 patients collected during a case control study conducted in PNG which has been described previously (Kaestli 2006). In order to increase discrimination power, the primary PCR amplification was done using primers binding to sequences in the 5' untranslated region of ups A, B and C *var* genes. A nested PCR was subsequently performed with GeneMapper® primers and analyzed as described above. Because all samples had been amplified, cloned and sequenced previously (Falk et al. 2008, manuscript submitted), it was possible to directly compare the

composition of samples collected. Fragments were identified manually using an artificial cut-off of 500 fluorescent units and the 8bp-tail was subtracted for comparison with obtained sequence data.

From 42 children 132 cloned domains were identified by sequencing (Table 2). Of those, 119 were unique and 13 sequences were found in more than one child.

In the same children GeneMapper® detected 253 genotypes of which 77 were different by size and upstream region. 55 sequences were found repeatedly (between 2-11 times) in more than one child. When sequencing data and GeneMapper® results were compared, 99 genotypes were detected by both methods. 9 cloned fragments identified by sequencing were initially detected by GeneMapper® but fell below the cut-off and 7 were not detected by GeneMapper® at all. Additional 4 fragments could not be detected by GeneMapper® since DBL1α domains of identical size but different sequence occurred in the same child. Therefore, if sequencing is used as a Gold Standard the sensitivity of GeneMapper® in this sample set was 83.2%. However, GeneMapper® detected 141 genotypes which were not detected by sequencing (Figure 4), whereas vice versa this was only the case for 20 sequences. Sizing was extremely precise with a mean deviation of 1bp from the calculated fragment length.

Discussion

PfEMP1 is a major virulence factor. With its dual character of mediating sequestration and preventing effective immune responses by antigenic variation PfEMP1 presents a tremendous challenge for the immune system of the human host. Switching from one PfEMP1 to another not only alters the protein being expressed at the erythrocyte surface but is usually accompanied by functional changes resulting in altered adhesive phenotypes.[19] However, whether *var* gene switching occurs randomly or follows a defined order is yet unknown and remains to be investigated. Changes of this molecule at the time of transition from asymptomatic to mild or severe malaria status would also be of major interest and therefore *var* gene dynamics in patients of different clinical presentations have been studied.[3,7,8] This has been done by cloning and sequencing of isolated and reverse transcribed RNA. In order to capture the true diversity 20 to 100 clones from each sample have been sequenced in various studies on *var* gene expression in naturally occurring infections. This is not only extremely expensive and cumbersome, also differences in PCR and cloning efficiencies hamper this approach. Therefore, new tools allowing a rapid assessment of the dynamics of *var* gene expression are needed. In this study we investigated whether fragment sizing by capillary electrophoresis together with GeneMapper® analysis software could be applied to the analysis of *var* gene expression. This technology has been previously successfully implemented for genotyping of *P. falciparum msp2* [15,16] and for other marker genes of *P. vivax* [17]. In a first approach we tested this techique with samples collected from a longterm 3D7 *in vitro* culture, which was monitored for *var* gene expression and potential switching for 168 days. For this pilot study only DBL1α-specific PCR primers were used on cDNA of full-length *var* genes that did not allow to distinguish between the major *var* gene groups. Nevertheless, a virtual analysis of the 3D7 genome showed that 27 of 60 *var* genes can be distinguished by size. Although the discrimination power in this case is not particularly high, a potential switch would be detected with a probability of 45%.

During 168 days we identified 25 of 27 distinguishable *var* genes in the long term culture of 3D7. This reflects a very large diversity of *var* genes with 2-18 different variants expressed per time point (Figure 3). However, throughout the observation period only one *var* gene was dominantly expressed suggesting that switching from one *var* gene to another did not occur in a concerted manner but that switches occurred randomly at low frequency underlying a stable expression of one gene.

Antigenic switching has been investigated *in vitro*[6,20,21] and *in vivo*[3,5,7,8,22] and switching rates were calculated from as low as 0.025% to up to 18% per generation. This implies that switching rates might differ among *var* genes and recently it was indeed shown that different *var* genes have intrinsically different switching rates dependent on their chromosomal location.[23] We were not able to detect any significant *var* gene switching during 168 days in our 3D7 culture but there were fluctuations in the expression of minor variants. The culture was not selected for a particular phenotype and it is also possible that the repertoire of expressed *var* genes was non-homogenous because we used non-synchronized parasites probably expressing many *var* genes[1,24]. Despite the fact that the dominant peak could have arisen from 3 different *var* genes (PFD1000c, PFD0995c and PF07_0050) we are very confident that we would have detected a switch with our approach.

When fragment sizing and GeneMapper® were applied to field samples, cDNA was amplified from the 5' UTR to the DBL1α domain. In order to improve the discrimination power, a primary PCR was done for each *var* gene subgroup (group A, B, and C) separately and was followed by a DBL1α-specific PCR with labeled primers. Table1 shows the increase of discrimination power by approximately 20% resulting in 39 of 60 potentially distinguishable *var* genes. This approach adds also the information on the expressed *var* gene subgroup which has been show to be an important determinant for disease severity.[10,12,13,21]

When sizing by capillary electrophoresis and GeneMapper® was compared to sequencing data from the same samples (Falk et al. 2008, manuscript submitted), GeneMapper® was able to detect 83.2% of sequences. 9 sequences (7.5%) would have been detected but were below the artificial cut-off and 7 DBL1α sequences (5.8%) were not detected at all. Thus, improving determination of cut-off would increase the sensitivity of the approach significantly. Whereas with sequencing only 13 fragments were found in more than one child this number was much larger for GeneMapper® as in these 42 children many DBL1α domains were present with identical size but different sequence. Hence, discrimination power might be further increased by using sequence domains of higher diversity.

Most importantly, the case that 2 DBL1α domains in the same individual were different in sequence but not in size and thus could not be detected by GeneMapper® was only observed in 4 children (9.5%). This reflects the power by which GeneMapper® can identify switches only by sizing of *var* gene domains.

With the two parameters of DBL1α size and upstream region (upsABC) GeneMapper® was able to distinguish 253 sequences in 42 children. GeneMapper® detected 141 *var* fragments which were not detected with cloning and sequencing whereas vice versa this was only the case for 20 sequences. This is partly due to the higher sensitivity of GeneMapper® as previously shown[16] but also to the much simpler processing of samples comprising a nested PCR only compared to the numerous steps involved in cloning, ligation and sequencing. The representation of diversity is in particular highly dependent on the number of sequences generated by cloning from each sample whilst all fragments are detected in 3 PCRs (upsABC) by GeneMapper®.

It is noteworthy that the analyzed samples derived from a case control study and subsequent *var* gene switches could not be determined in the same child. However, this technology enables the analysis of longitudinal samples and thus will provide information about *var* gene dynamics within the same individual. For that purpose, the frequency of *var* genes with DBL1α domains identical in size but different in sequence occurring in subsequent samples of the same patient remains to be investigated. It is promising to see that in only 4 children two *var* gene domains were observed which had identical sizes but were different by sequence.

In conclusion, sizing by capillary electrophoresis and GeneMapper® as a genotyping tool to study *var* gene dynamics is a great improvement over previously used techniques employing cloning and sequencing of PCR fragments. This cloning strategy is expensive and cumbersome and limitations lie in the 'collector's problem', i.e. how many clones need to be sequenced to obtain a representative sample. A major problem faced by all methods to study expression pattern in naturally occurring infections is the stability of RNA and the isolation of sufficient amouts. Despite the fact that these problems will remain, studies on *var* gene transcription and dynamics as previously performed by others[3,4,7,8] would have greatly benefited from the new GeneMapper® approach.

References

1. Scherf A, Hernandez-Rivas R, Buffet P, et al. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*. EMBO J. 1998;17:5418-26.
2. Gardner MJ, Hall N, Fung E, et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature. 2002;419:498-511.
3. Peters J, Fowler E, Gatton M, et al. High diversity and rapid changeover of expressed *var* genes during the acute phase of *Plasmodium falciparum* infections in human volunteers. Proc. Natl. Acad. Sci. U.S.A. 2002;99:10689-94.
4. Peters JM, Fowler EV, Krause DR, Cheng Q, Gatton ML. Differential changes in *Plasmodium falciparum var* transcription during adaptation to culture. J. Infect. Dis. 2007;195:748-55.
5. Gatton ML, Peters JM, Fowler EV, Cheng Q. Switching rates of *Plasmodium falciparum var* genes: faster than we thought? Trends Parasitol. 2003;19:202-8.
6. Horrocks P, Pinches R, Christodoulou Z, Kyes SA, Newbold CI. Variable *var* transition rates underlie antigenic variation in malaria. Proc. Natl. Acad. Sci. U.S.A. 2004;101:11129-34.
7. Kaestli M, Cortes A, Lagog M, Ott M, Beck H. Longitudinal assessment of *Plasmodium falciparum var* gene transcription in naturally infected asymptomatic children in Papua New Guinea. J. Infect. Dis. 2004;189:1942-51.
8. Wunderlich G, Alves FP, Gölnitz U, et al. Rapid turnover of *Plasmodium falciparum var* gene transcripts and genotypes during natural non-symptomatic infections. Rev. Inst. Med. Trop. Sao Paulo. 47:195-201.
9. Kirchgatter K, Portillo HDA. Association of severe noncerebral *Plasmodium falciparum* malaria in brazil with expressed pfemp1 DBL1 alpha sequences lacking cysteine residues. Mol. Med. 2002;8:16-23.
10. Jensen ATR, Magistrado P, Sharp S, et al. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group a *var* genes. J. Exp. Med. 2004;199:1179-90.
11. Bull PC, Berriman M, Kyes S, et al. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. PLoS Pathog. 2005;1:e26.
12. Kaestli M, Cockburn IA, Cortés A, et al. Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. J. Infect. Dis. 2006;193:1567-74.

13. Rottmann M, Lavstsen T, Mugasa JP, et al. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in tanzanian children. Infect. Immun. 2006;74:3904-11.
14. Montgomery J, Mphande FA, Berriman M, et al. Differential *var* gene expression in the organs of patients dying of falciparum malaria. Mol. Microbiol. 2007;65:959-67.
15. Jafari S, Le Bras J, Bouchaud O, Durand R. *Plasmodium falciparum* clonal population dynamics during malaria treatment. J Infect Dis. 2004;189:195-203.
16. Falk N, Maire N, Sama W, et. al. Comparison of pcr-rflp and genescan-based genotyping for analyzing infection dynamics of *Plasmodium falciparum.* Am J Trop Med Hyg. 2006;74:944-50.
17. Koepfli C, Mueller I, Marfurt J, et al. Evaluation of *Plasmodium vivax* genotyping markers for molecular monitoring in clinical trials. JID, in press. .
18. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques. 1996;20:1004-6, 1008-10.
19. Smith JD, Chitnis CE, Craig AG, et al. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. Cell. 1995;82:101-10.
20. Roberts DJ, Craig AG, Berendt AR, et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. Nature. 1992;357:689-92.
21. Lavstsen T, Magistrado P, Hermsen CC, et. al Expression of *Plasmodium falciparum* erythrocyte membrane protein 1 in experimentally infected humans. Malar. J. 2005;4:21.
22. Brannan LR, Turner CM, Phillips RS. Malaria parasites undergo antigenic variation at high rates *in vivo*. Proc Biol Sci. 1994;256:71-5.
23. Frank M, Dzikowski R, Amulic B, Deitsch K. Variable switching rates of malaria virulence genes are associated with chromosomal position. Mol. Microbiol. 2007;64:1486-98.
24. Chen Q, Fernandez V, Sundström A, et al. Developmental selection of *var* gene expression in *Plasmodium falciparum*. Nature. 1998;394:392-5.

Table1.*var* genes in 3D7. First column indicates predicted length of the DBL1α PCR amplicon in base pairs. In the second and third column the corresponding 3D7 *var* genes and their upstream regions are listed. The colored boxes indicate *var* genes with the DBL1α domains of the same size. The fourth column shows the number of mismatches in the reverse primer. *var2* does not contain a DBL1α domain. In the last column the probability of not detecting a switch because of identical size is indicated.

Predicted DBL1α size	Corresponding *var gene*	Upstream region	Reverse primer matching	Probability of undetected switch
368	PFF0010w	ups B/A	5 bases mismatch	0
371	PF08_0141	ups A		0.034
	PFE1640w	ups A		0.034
380	PF13_0003	ups A		0
383	PFD0020c	ups A		0
389	PFF0020c	ups A	8 bases mismatch	0
	MAL7P1.55	upsB/C		0
392	PFI1820w	ups A		0
	PF08_0140	ups B/A		0
395	PFA0015c	ups A		0.034
	PF11_0521	ups A		0.034
	PFL0020w	ups B/A		0
398	PF07_0139	upsB		0.1
	PFL0005w	upsB		0.1
	PFC0005w	upsB		0.1
	PFF1595c	upsB		0.1
	PFD1005c	upsB/C		0.1
	PFL1950w	upsB/C		0.1
401	PFF0845c	upsC	5 bases mismatch	0
404	PFD1235w	ups A		0.03
	MAL7P1.1	ups A		0.03
	PF10_0406	upsB		0
	PFD1015c	upsC		0
407	PFL2665c	upsB		0.03
	PFA0765c	upsB		0.03
410	PF10_0001	upsB		0
413	PFA0005w	upsB		0
416	PFF1580c	ups B/A		0.07
	PFL1955w	upsB/C		0.07
	PF08_0103	upsB/C		0.07
	PF08_0106	upsB/C	2 bases mismatch	0.07
	MAL7P1.56	upsC		0
419	PF08_0142	upsB		0
	PFL1960w	upsC		0.03
	PF07_0048	upsC		0.03
422	PFD1000c	upsC		0.03
	PFD0995c	upsC		0.03
	PF07_0050	upsB/C		0
425	PFD1245c	upsB		0.03
	PF11_0007	upsB		0.03
	PF07_0049	upsC		0
428	PF11_0008	ups A		0
	PFB0010w	upsB		0
431	PFL0935c	upsB		0.03
	PFC1120c	upsB		0.03
	PF08_0107	upsC		0
437	PFI1830c	upsB		0
440	PFI0005w	upsB		0.03
	PFB1055c	upsB		0.03
443	PFD0005w	upsB		0.03
	MAL7P1.50	upsB/C		0.03
446	PF07_0051	upsC		0
449	PFD0615c	upsC		0
452	PFD0625c	upsC		0.03
	PFD0630c	upsC		0.03
458	PFD0635c	upsB/C		0
461	PF13_0364	upsB		0
470	PFE_0005w	upsB		0.03
	PF13_0001	upsB		0.03
	PFL0030c		no DBL1α	

Table 2. *var* genes observed by sequencing and GeneMapper® analysis. The total number of sequences, the number of different sequences and the number of sequences occurring in several children are indicated for both techniques.

	Fragments detected by cloning and sequencing	Fragments detected by GeneMapper®
Total number of sequences	132	253
Number of different sequences	119	77
Number of sequences occurring more than once	13	55

Figure legends

Figure 1. Schematic drawing of primer modification (Figure adapted from "User Bulletin ABI PRISM™ Linkage Mapping Set Version 2.5").

Figure 2. Example of a GeneMapper electropherogram. The x-axis corresponds to the size of the DBL1α fragments in bp. The y-axis indicates the fluorescent intensities of detected peaks. 6FAM–labeled DBL1α fragments are shown in blue. ROX-labeled fragments of the size standard are indicated in red. Example depicted here derived from a field sample.

Figure 3. *var* gene transcription pattern of the 3D7 *in vitro* culture. Dark purple fields indicate the dominant (highest) peak in every sample. The white number in the dark purple field shows the peak height in fluorescence-units. Other colours indicate the percentage of the particular peak height in relation to the dominant peak. The day of harvesting and the parasitemia of the culture is indicated. The numbers on the left side indicate the predicted *var* DBL1α fragment lengths according to the PlasmoDB database.

Figure 4. Number of *var* genes detected by sequencing and GeneMapper. The total number of clones detected are listed patient-wise. The number of sequences identified by both techniques are indicated in orange. 9 fragments were only detected by sequencing because they fell below the artificial cut-off (blue) and an additional 7 fragments were not detected by GeneMapper at all (yellow). Sequences only detected by GeneMapper but not with sequencing are indicated in green. Asteristics indicate fragments representing two sequences of the same size but different sequence in the same child.

Figures

Figure 1

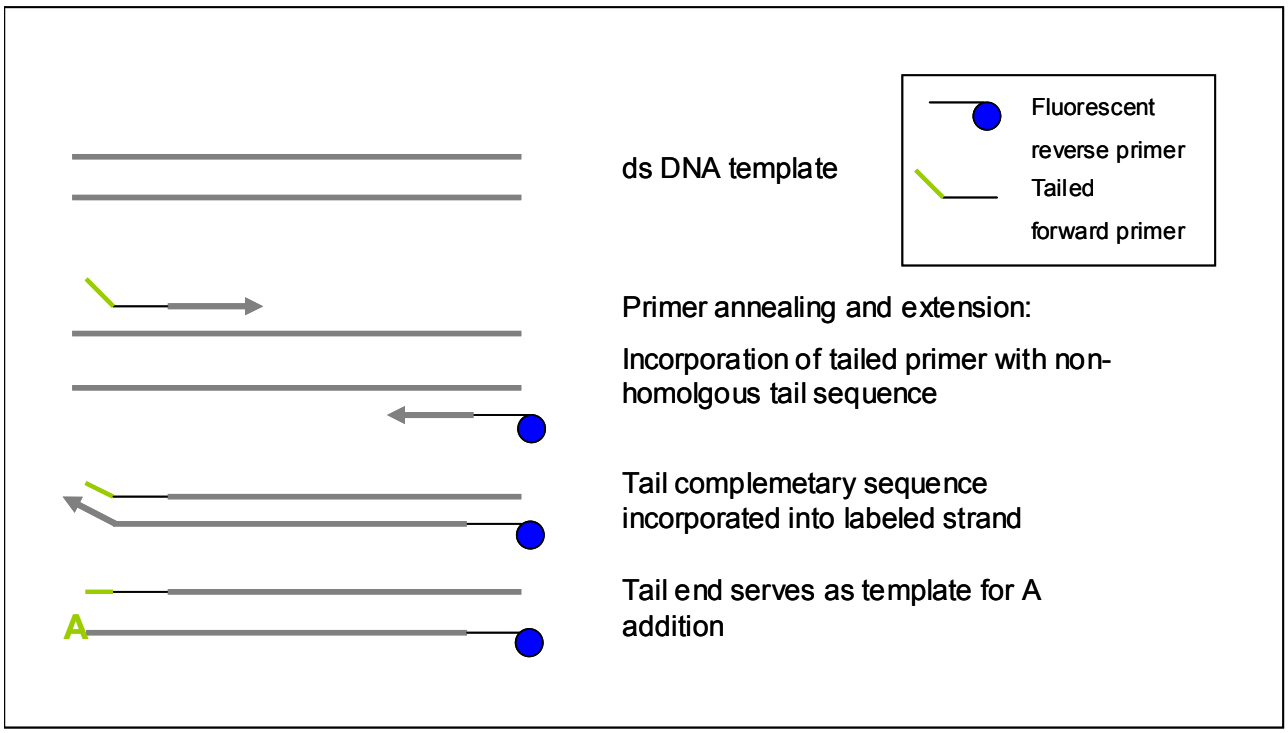
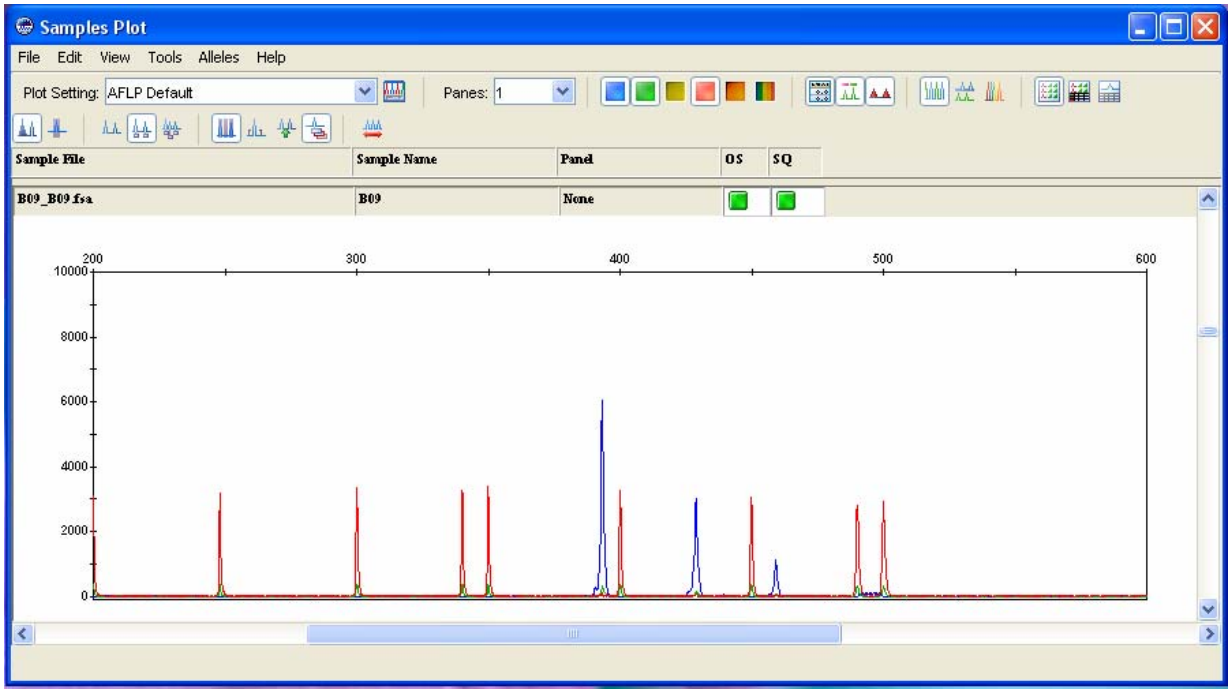
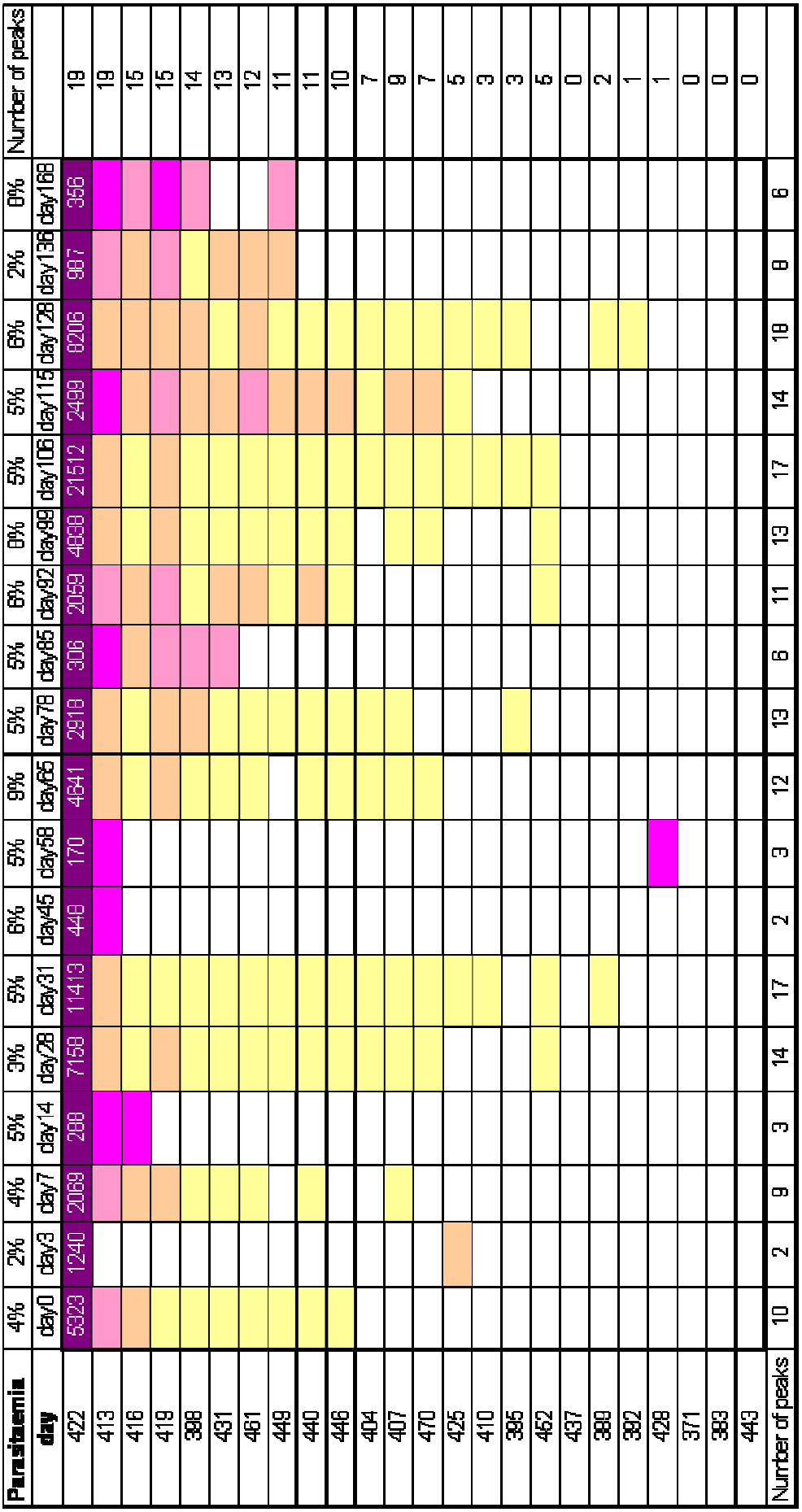


Figure 2





Parasitaemia	4%	2%	4%	5%	3%	5%	6%	5%	9%	5%	5%	5%	6%	0%	5%	6%	2%	0%	Number of peaks
day	day0	day3	day7	day14	day28	day31	day45	day58	day65	day78	day85	day92	day99	day106	day115	day128	day136	day168	
422	5323	1240	2069	288	7158	11413	448	170	4641	2918	306	2059	4838	21512	2499	8206	987	356	19
413																			19
416																			15
419																			15
398																			14
431																			13
461																			12
449																			11
440																			11
446																			10
404																			7
407																			9
470																			7
425																			5
410																			3
395																			3
452																			5
437																			0
388																			2
382																			1
428																			1
371																			0
383																			0
443																			0
Number of peaks	10	2	9	3	14	17	2	3	12	13	6	11	13	17	14	18	8	6	

Figure 3

	Sequences															
ID	**Present in GM+Seq**				**Present only in seq**		**Present only in GM**									
XB2 upsB	407*						392	419	449							
XB3 upsC	476						398	407								
XB6 upsC	407	422					410									
XB9 upsB	437				398	461	365	368	449							
XB9 upsC	458						407	428								
XB17 upsB					404	416	371	389	425	434	437					
XB13 upsB	407	431	443		468		386	401								
XB13 upsC	410						398	419	425	434	440					
XB14 upsA	371						422									
XB14 upsB	380	401					449									
XB18 upsC	395	428					422									
XB25 upsA	383	395					416									
XB25 upsB	467						368	371	380	386	389	395	401	416	434	440
XB25 upsC	404	440					416	431								
XB27 upsB	374	431					362	368	401	413	425	431				
XB27 upsC	398						413	425	437	461						
Xb33 upsA					416		422	440	449							
XB33 upsB	419	422	449*		389		446									
XB35 upsA	380	389														
Xb35 upsB	368	386	422		416		398	434	437	446	461					
XB35 upsC	398	413														
XB36 upsC	386	437			422	398	371	380	395	449						
XB37 upsA	383						389									
XB37 upsB	419						368	392	395	407	425	428	434			
XB37 upsC	404						419	428								
XB39 upsC	395						383	413								
XB42 upsB	449						398	401	416							
XB42 upsC	398	404					458									
XB44 upsB	416				446		404	422	428	449						
XB44 upsC	413						398	416								
XB53 upsB	398				419		368	371	395	401	422	428	443	446		
XB54 upsA	374						383	398								
XB54 upsC	428						410	437	449							
XB55 upsB	437															
XB56 upsA	395															
XB56 upsC	395	425	434				368	389	449							

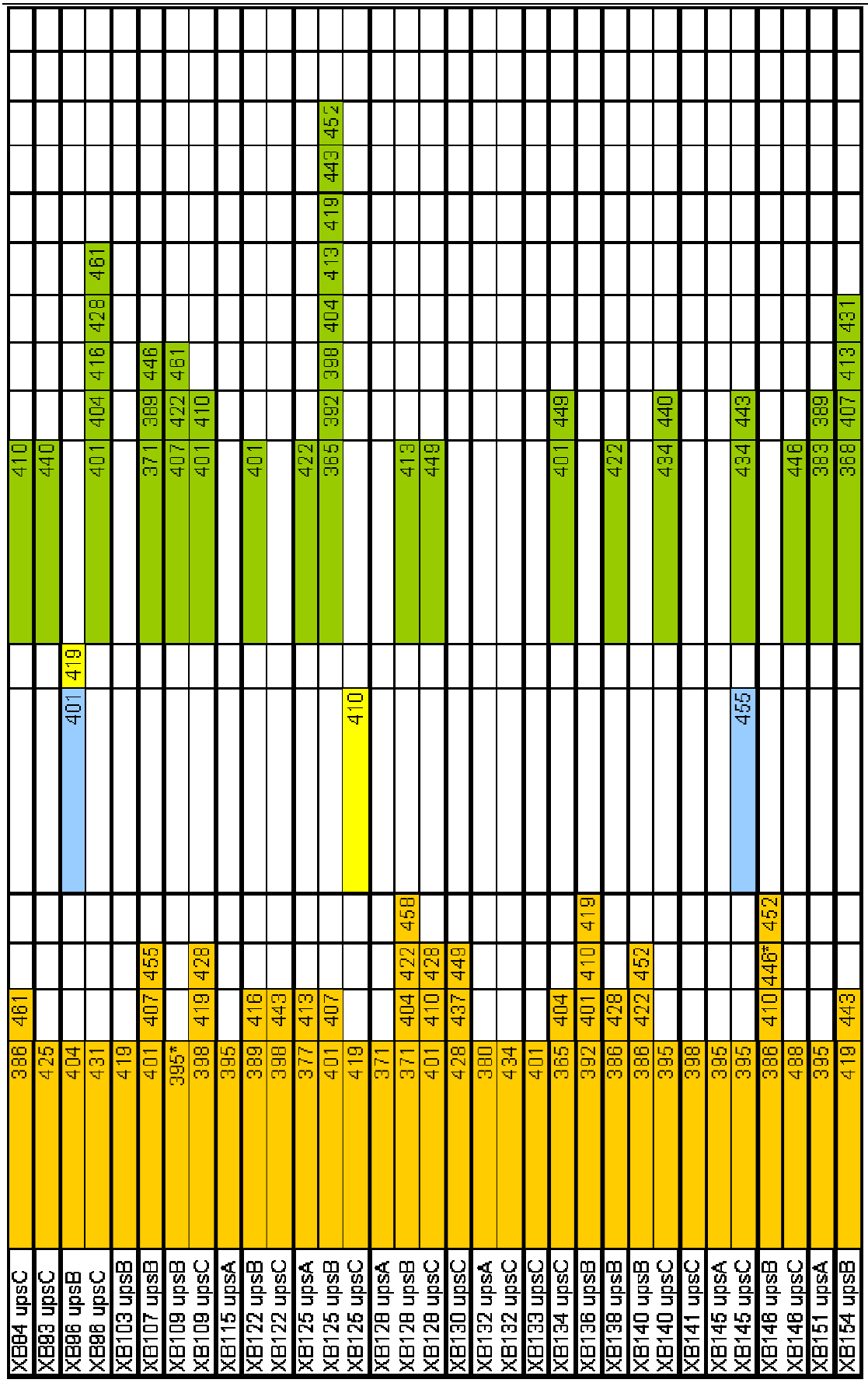

XB04 upsC 386 461 410
XB03 upsC 425 440
XB06 upsB 404 401 419
XB06 upsC 431 401 404 416 428 461
XB03 upsB 419
XB07 upsB 401 407 455 371 389 446
XB09 upsB 395* 407 422 461
XB09 upsC 398 419 428 401 410
XB15 upsA 395
XB22 upsB 389 416 401
XB22 upsC 398 443
XB25 upsA 377 413 422
XB25 upsB 401 407 365 392 398 404 413 419 443 452
XB25 upsC 419 410
XB28 upsA 371
XB28 upsB 371 404 422 458 413
XB28 upsC 401 410 428 449
XB30 upsC 428 437 449
XB32 upsA 380
XB32 upsC 434
XB33 upsC 401
XB34 upsC 365 404 401 449
XB36 upsB 392 401 410 419
XB38 upsB 386 428 422
XB40 upsB 386 422 452 434 440
XB40 upsC 395
XB41 upsC 398
XB45 upsA 395
XB45 upsC 395 455 434 443
XB48 upsB 386 410 446* 452
XB48 upsC 488 446
XB51 upsA 395 383 389
XB54 upsB 419 443 368 407 413 431

Figure 4

Chapter 5
General Discussion

General Discussion

In the first part of this thesis a cloning and sequencing approach was used to identify a previously suggested subset of *var* genes responsible for the development of severe malaria. Despite the fact that DBL1α domains with reduced number of cysteines were mainly found in symptomatic children which support previous findings[66,133,68] we were not able to identify particular *var* gene sequences associated with malaria morbidity. With our sequencing approach we could not confirm an upregulation of upsB in severe cases or upsC in children with asymptomatic infections reported previously[134,65] since the number of detected sequences were similar for upsB and upsC regardless of the children's clinical presentation. Neither did we find specific motifs or homologous sequence stretches that were shared among *var* genes isolated from children with severe disease which - once more - clearly shows the extensive polymorphism of this multi-gene family. In contrast, in some asymptomatic children we identified a cysteine to tyrosine amino acid substitution and *var* genes carrying this substitution were mainly of *var* group C type which were interestingly found in children who showed far higher parasitemias than children infected with parasites not harboring such substitution it could be speculated that children might tolerate such high parasitemias without showing any clinical symptoms because these parasites show a decreased binding affinity, as a substitution of cysteine residues forming disulfide bonds might result in significant structural modifications. Thus, this substitution might lead to less adherence and consequently to higher parasite populations. However, why such an increased number of circulating parasites is not cleared by the spleen more rapidly remains to be elucidated. Furthermore, linking malaria morbidity to specific *var* genes or groups of *var* genes might be hampered by the fact that only the parasite population cirulating in the peripheral blood can be investigated. It has been shown that circulating genotypes only form a subset of those sequestered in the tissues[135] implicating that with the common approaches the most important variants might be missed. A recent study on fatal pediatric malaria patients in Malawi supports this finding as over 100 different *var* variants were expressed in a single patient with up to 49 different variants in a single organ. Additional data from the postmortem study in Malawi point towards organ-specific sequestration of expressed *var* genes which implies that the type of PfEMP1 determines the site of cytoadherence.[136] DBL1α sequence tags identified by Bull *et al*.[66] were not associated with the site of sequestration. The majority of sequences contained 4 cysteine residues compared to only 13% with 2 cysteine residues, a motif previously

associated with severe malaria[67] and many sequences were found to be highly similar to 3D7 which both resembles our findings.

var gene diversity hampers the design of universal primers and therefore most sequencing projects - just like ours - focus on the analysis of distinct sequence stretches rather than whole genes. Analyzing only the DBL1α domain is intrinsingly an analytical restriction in a large gene and other parts might play equally or more important roles in pathology but are not analyzed.

Transformation-associated recombination (TAR) cloning[137] might be a possible approach to sequence full-length *var* genes therefore gaining access to the sequences further downstream of the DBL1α domain. Multiple alignments and bioinformatic analyses might help to identify sequence homologies or distinct sequence motifs in patients with different clinical presentations. Including patients from different geographical region would further broaden the knowledge about the diversity of the *var* gene repertoire existing worldwide. Sequence analyses of full length *var* genes will certainly provide more meaningful data than comparing discrete regions of *var* genes only.

However, it should be kept in mind that despite the extreme polymorphism, PfEMP1s of different sequence exhibit similar protein architectures[60] which might indicate that structural analyses are much more powerful than primary sequence comparisons.

var genes possess up to 7 different domains[61] and up to now, antigenic epitopes within most of them remain elusive, as does their involvement in the generation of protective antibodies. Therefore, we tried to address this issue in the second part of the thesis by recombinant expression of *var* gene regions including domains further downstream of the frequently analyzed DBL1α domain. Screening of recombinant domains with sera from naturally exposed individuals should give information about the importance of these regions for the production of anti-PfEMP1 antibodies and their potential involvement in immunological protection. Of the 2 representative *var* genes reported to be associated with severe malaria[63,37], 3 recombinant fragments and 2 synthetic peptides were found to be significantly differently recognized by adults compared to children as well as asymptomatic children compared to children with severe malaria. However, sample sizes were small and therefore more data are required to confirm a potential involvement of these fragments and peptides in the generation of protective antibodies. Furthermore, since not all fragments could be expressed we might have missed important epitopes and generally, data obtained in ELISA with heterologously expressed protein fragments should be considered with care. Thus, investing into the identification of novel and more suitable expression systems might be crucial.

With *var* genes, there is an added issue, which is the fact that due to the vast diversity, many domains may never be seen by certain individuals. The domains tested here all derived from cultured isolates of different origin and tested patients might not have seen these sequences. Nevertheless, many sera did show reactivity against several domains which implies similarities- to a certain degree- among different parasite isolates, but we did not find any evidence for these strain-transcendent immune responses to play a role in protection.

Identifying such potentially important domains in ELISA might serve as a pre-selection for peptides on a protein microarray which would greatly facilitate this search, since many more sera could be analyzed simultaneously. A protein microarray would greatly facilitate the identification of distinct epitopes conferring protection and thus would allow directly to study the differential recognition in healthy and diseased children. Since such protein arrays would require only little amounts of sera, many more children samples, which are usually only available in very small volumes, could be tested, with paired acute/reconvalescent follow-up samples being the most valuable ones. Alternative systems would be very useful and a recent application of the BioPlex^{100} system by Cham *et al.*[138] showed promising results. In the bead-based BioPlex^{100} approach, recombinant proteins are covalently coupled onto sets of beads which are impregnated with different dyes emitting their specific detection signal upon excitation. Recombinant proteins can be detected by biotinylated secondary antibodies with phycoerythrin-conjugated streptavidin used as a reporter. Theoretically, one hundred different antigens could be analyzed simultaneously in 1µl of sample by this technique. Cham *et al.*[138] evaluated 28 unique bead populations coated with recombinant *Plasmodium falciparum* 3D7 DBL and CIDR domains and found the assay to be sensitive, accurate, reproducible and high throughput.

In the last part of this thesis we tried to establish a new genotyping tool in order to facilitate the analysis of *var* gene dynamics and diversity in the field. GeneMapper® technology was used for *var* gene tracking and differentiation to finally replace the tedious cloning and sequencing approach which is a common strategy to study *var* gene diversity and switching and which was also the method of choice for sequencing analysis in the first part of this thesis. Comparisons with sequencing data revealed that GeneMapper® sensitivity is equally good or even better. Using DBL1α-specific primers on subgroup-specific primary PCR products identified a large *var* gene repertoire present in 42 analyzed children. Sequencing only identified 13 fragments which were found in more than one child, GeneMapper® identified 55 DBL1α types occurring more than once in several children.

This was due to a large number of equally sized fragments with different sequences. *var* gene domains with an upsB promoter region and a size of 401bp were observed in 11 children resulting in a frequency of 0.04. Thus, the probability of detecting two *var* genes of identical size in different children or in sequential samples of the same child is 0.0016. Therefore, despite some shortcomings GeneMapper® holds a large potential to replace cloning and sequencing. The latter technology, not only is extremely cumbersome and expensive but is probably not representative and suffers from bias and other limiting factors. However, further evaluation and optimization of GeneMapper® is required. Experiments using the 5' UTR of *var* genes as target sequence might improve the resolution, as well as the regions downstream of the DBL1α domain. However, since the DBL1α domain, apart from the acidic terminal sequence, is thought to be the most conserved *var* gene domain, it might be challenging to design universal primers for these downstream elements. Further improvement of resolution could be obtained by enzymatic digestions, but would increase hands on time and costs.

The number of publications implicating PfEMP1 in the development of semi-immunity keeps growing. The latest one was presented by Beeson et al. 2008 at the MAM conference, claiming 80% of the antibody responses against the iRBC surface to be directed against PfEMP1. However, it is worth mentioning that more than 50% of the 3D7[139] proteins are still of unknown function and that PfEMP1 is not the only protein on the surface of the iRBC. Other proteins like the RIFINs[88], SURFINs[90] or STEVORs[96] -about which only little is known so far-may also be implicated in malaria morbidity and therefore should be included in further analyses.

Finally, although anti-disease immunity is mainly thought to be mediated by antibodies, the power of the innate immune system should not be neglected. Recent publications by Clark *et al.*[5] even claim that sequestration is only a secondary effect of a systemic inflammation, characterized by the release of pro- and anti-inflammatory cytokines and that an imbalance between those is responsible for disease outcome.

References of General Introduction and Discussion

1.9241546948_text_eng.pdf (application/pdf-Objekt). on <http://whqlibdoc.who.int/publications/2006/9241546948_text_eng.pdf>

2. Waller, D. et al. Clinical features and outcome of severe malaria in Gambian children. *Clin Infect Dis* 21, 577-87(1995).

3. Marsh, K. et al. Indicators of life-threatening malaria in African children. *N Engl J Med* 332, 1399-404(1995).

4. Greenwood, B., Marsh, K. & Snow, R. Why do some African children develop severe malaria? *Parasitol. Today (Regul. Ed.)* 7, 277-81(1991).

5. Clark, I.A. The advent of the cytokine storm. *Immunol Cell Biol* 85, 271-3(2007).

6. Clark, I.A. et al. Human malarial disease: a consequence of inflammatory cytokine release. *Malar J* 5, 85(2006).

7. van Ooij, C. & Haldar, K. Protein export from *Plasmodium parasites*. *Cell Microbiol* 9, 573-82(2007).

8. Charpian, S. & Przyborski, J.M. Protein transport across the parasitophorous vacuole of *Plasmodium falciparum*: into the great wide open. *Traffic* 9, 157-65(2008).

9. Wickham, M.E. et al. Trafficking and assembly of the cytoadherence complex in *Plasmodium falciparum*-infected human erythrocytes. *EMBO J* 20, 5636-49(2001).

10. Pologe, L.G. et al. Primary structure and subcellular localization of the knob-associated histidine-rich protein of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 84, 7139-43(1987).

11. Triglia, T. et al. The complete sequence of the gene for the knob-associated histidine-rich protein from *Plasmodium falciparum*. *EMBO J* 6, 1413-9(1987).

12. Pasloske, B.L. et al. Cloning and characterization of a *Plasmodium falciparum* gene encoding a novel high-molecular weight host membrane-associated protein, PfEMP3. *Mol Biochem Parasitol* 59, 59-72(1993).

13. Su, X.Z. et al. The large diverse gene family *var* encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* 82, 89-100(1995).

14. Udeinya, I.J. et al. Falciparum malaria-infected erythrocytes specifically bind to cultured human endothelial cells. *Science* 213, 555-7(1981).

15. Berendt, A.R., Ferguson, D.J. & Newbold, C.I. Sequestration in *Plasmodium falciparum* malaria: sticky cells and sticky problems. *Parasitol Today* 6, 247-54(1990).

16. Pongponratn, E. et al. Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. *Am J Trop Med Hyg* 44, 168-75(1991).
17. Handunnetti, S.M. et al. Uninfected erythrocytes form "rosettes" around *Plasmodium falciparum* infected erythrocytes. *Am J Trop Med Hyg* 40, 115-8(1989).
18. Treutiger, C.J. et al. Rouleaux-forming serum proteins are involved in the rosetting of *Plasmodium falciparum*-infected erythrocytes. *Exp Parasitol* 93, 215-24(1999).
19. Roberts, D.J. et al. Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* 357, 689-92(1992).
20. Pain, A. et al. Platelet-mediated clumping of *Plasmodium falciparum*-infected erythrocytes is a common adhesive phenotype and is associated with severe malaria. *Proc Natl Acad Sci U S A* 98, 1805-10(2001).
21. Miller, L.H. et al. Definition of the minimal domain of CIDR1alpha of *Plasmodium falciparum* PfEMP1 for binding CD36. *Mol. Biochem. Parasitol* 120, 321-3(2002).
22. Miller, L.H., Good, M.F. & Milon, G. Malaria pathogenesis. *Science* 264, 1878-83(1994).
23. Ockenhouse, C.F. et al. Molecular basis of sequestration in severe and uncomplicated *Plasmodium falciparum* malaria: differential adhesion of infected erythrocytes to CD36 and ICAM-1. *J Infect Dis* 164, 163-9(1991).
24. Barnwell, J.W. et al. A human 88-kD membrane glycoprotein (CD36) functions in vitro as a receptor for a cytoadherence ligand on *Plasmodium falciparum*-infected erythrocytes. *J Clin Invest* 84, 765-72(1989).
25. Ho, M. & White, N.J. Molecular mechanisms of cytoadherence in malaria. *Am J Physiol* 276, C1231-42(1999).
26. Baruch, D.I. Adhesive receptors on malaria-parasitized red cells. *Baillieres Best Pract Res Clin Haematol* 12, 747-61(1999).
27. Roberts, D.D. et al. Thrombospondin binds falciparum malaria parasitized erythrocytes and may mediate cytoadherence. *Nature* 318, 64-6
28. Berendt, A.R. et al. Intercellular adhesion molecule-1 is an endothelial cell adhesion receptor for *Plasmodium falciparum*. *Nature* 341, 57-9(1989).
29. Ho, M. et al. Visualization of *Plasmodium falciparum*-endothelium interactions in human microvasculature: mimicry of leukocyte recruitment. *J Exp Med* 192, 1205-11(2000).

30. Ockenhouse, C.F. et al. Human vascular endothelial cell adhesion receptors for *Plasmodium falciparum*-infected erythrocytes: roles for endothelial leukocyte adhesion molecule 1 and vascular cell adhesion molecule 1. *J Exp Med* 176, 1183-9(1992).
31. Treutiger, C.J. et al. PECAM-1/CD31, an endothelial receptor for binding *Plasmodium falciparum*-infected erythrocytes. *Nat Med* 3, 1405-8(1997).
32. Udomsangpetch, R. et al. Promiscuity of clinical *Plasmodium falciparum* isolates for multiple adhesion molecules under flow conditions. *J Immunol* 158, 4358-64(1997).
33. Beeson, J.G. et al. *Plasmodium falciparum* isolates from infected pregnant women and children are associated with distinct adhesive and antigenic properties. *J Infect Dis* 180, 464-72(1999).
34. Fried, M. & Duffy, P.E. Adherence of *Plasmodium falciparum* to chondroitin sulfate A in the human placenta. *Science* 272, 1502-4(1996).
35. Gysin, J. et al. *Ex vivo* desequestration of *Plasmodium falciparum*-infected erythrocytes from human placenta by chondroitin sulfate A. *Infect Immun* 67, 6596-602(1999).
36. Scholander, C. et al. Novel fibrillar structure confers adhesive property to malaria-infected erythrocytes. *Nat Med* 2, 204-8(1996).
37. Chen, Q. et al. Identification of *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) as the rosetting ligand of the malaria parasite P. falciparum. *J. Exp. Med* 187, 15-23(1998).
38. Rowe, J.A. et al. *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature* 388, 292-5(1997).
39. Barragan, A. et al. Blood group A antigen is a coreceptor in *Plasmodium falciparum* rosetting. *Infect Immun* 68, 2971-5(2000).
40. Luginbühl, A. et al. Complement factor D, albumin, and immunoglobulin G anti-band 3 protein antibodies mimic serum in promoting rosetting of malaria-infected red blood cells. *Infect Immun* 75, 1771-7(2007).
41. McGilvray, I.D. et al. Nonopsonic monocyte/macrophage phagocytosis of *Plasmodium falciparum*-parasitized erythrocytes: a role for CD36 in malarial clearance. *Blood* 96, 3231-40(2000).
42. Urban, B.C. et al. *Plasmodium falciparum*-infected erythrocytes modulate the maturation of dendritic cells. *Nature* 400, 73-7(1999).

43. Newbold, C. u. a. Cytoadherence, pathogenesis and the infected red cell surface in *Plasmodium falciparum*. *Int J Parasitol* 29, 927-37(1999).

44. Fried, M. u. a. Maternal antibodies block malaria. *Nature* 395, 851-852(1998).

45. Turner, G.D. et al. An immunohistochemical study of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role for intercellular adhesion molecule-1 in cerebral sequestration. *Am J Pathol* 145, 1057-69(1994).

46. Fernandez-Reyes, D. et al. A high frequency African coding polymorphism in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet* 6, 1357-60(1997).

47. Rogerson, S.J. et al. Cytoadherence characteristics of *Plasmodium falciparum*-infected erythrocytes from Malawian children with severe and uncomplicated malaria. *Am J Trop Med Hyg* 61, 467-72(1999).

48. Carlson, J. et al.Human cerebral malaria: association with erythrocyte rosetting and lack of anti-rosetting antibodies. *Lancet* 336, 1457-60(1990).

49. Treutiger, C.J. et al. Rosette formation in *Plasmodium falciparum* isolates and anti-rosette activity of sera from Gambians with cerebral or uncomplicated malaria. *Am J Trop Med Hyg* 46, 503-10(1992).

50. Rowe, A. et al. *Plasmodium falciparum* rosetting is associated with malaria severity in Kenya. *Infect Immun* 63, 2323-6(1995).

51. Ho, M. et al. Rosette formation of *Plasmodium falciparum*-infected erythrocytes from patients with acute malaria. *Infect Immun* 59, 2135-9(1991).

52. al-Yaman, F. et al. Human cerebral malaria: lack of significant association between erythrocyte rosetting and disease severity. *Trans R Soc Trop Med Hyg* 89, 55-8

53.Rogerson, S.J. et al. Disruption of erythrocyte rosettes and agglutination of erythrocytes infected with *Plasmodium falciparum* by the sera of Papua New Guineans. *Trans R Soc Trop Med Hyg* 90, 80-4

54. Cockburn, I.A. et al. A human complement receptor 1 polymorphism that reduces *Plasmodium falciparum* rosetting confers protection against severe malaria. *Proc. Natl. Acad. Sci. U.S.A* 101, 272-7(2004).

55. Roberts, D.J. et al. Autoagglutination of malaria-infected red blood cells and malaria severity. *Lancet* 355, 1427-8(2000).

56. Baruch, D.I. et al. Identification of a region of PfEMP1 that mediates adherence of *Plasmodium falciparum* infected erythrocytes to CD36: conserved function with variant sequence. *Blood* 90, 3766-75(1997).

57. Reeder, J.C. et al. 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, 5198-202(1999).
58. Smith, J.D. et al. Identification of a *Plasmodium falciparum* intercellular adhesion molecule-1 binding domain: a parasite adhesion trait implicated in cerebral malaria. *Proc Natl Acad Sci U S A* 97, 1766-71(2000).
59. Gardner, M.J. et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498-511(2002).
60. Kraemer, S.M. et al. Patterns of gene recombination shape *var* gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. *BMC Genomics* 8, 45(2007).
61. Lavstsen, T. et al. Sub-grouping of *Plasmodium falciparum* 3D7 *var* genes based on sequence analysis of coding and non-coding regions. *Malar. J* 2, 27(2003).
62. Voss, T.S. et al. Identification of nuclear proteins that interact differentially with *Plasmodium falciparum var* gene promoters. *Mol. Microbiol* 48, 1593-607(2003).
63. Jensen, A.T.R. et al. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A *var* genes. *J. Exp. Med* 199, 1179-90(2004).
64. Rottmann, M. et al. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infect. Immun* 74, 3904-11(2006).
65. Kaestli, M. et al. Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study. *J. Infect. Dis* 193, 1567-74(2006).
66. Bull, P.C. et al. *Plasmodium falciparum* variant surface antigen expression patterns during malaria. *PLoS Pathog* 1, e26(2005).
67. Kirchgatter, K. & Portillo, H.D.A. Association of severe noncerebral *Plasmodium falciparum* malaria in Brazil with expressed PfEMP1 DBL1 alpha sequences lacking cysteine residues. *Mol. Med* 8, 16-23(2002).
68. Kyriacou, H.M. et al. Differential *var* gene transcription in *Plasmodium falciparum* isolates from patients with cerebral malaria compared to hyperparasitaemia. *Mol. Biochem. Parasitol* 150, 211-8(2006).

69. Kyes, S.A., Kraemer, S.M. & Smith, J.D. Antigenic variation in *Plasmodium falciparum*: gene organization and regulation of the *var* multigene family. *Eukaryotic Cell* 6, 1511-20(2007).

70. Smith, J.D. et al. Classification of adhesive domains in the *Plasmodium falciparum* erythrocyte membrane protein 1 family. *Mol Biochem Parasitol* 110, 293-310(2000).

71. Smith, J.D. et al. Decoding the language of *var* genes and *Plasmodium falciparum* sequestration. *Trends Parasitol* 17, 538-45(2001).

72. Vogt, A.M. et al. Heparan sulfate on endothelial cells mediates the binding of *Plasmodium falciparum*-infected erythrocytes via the DBL1alpha domain of PfEMP1. *Blood* 101, 2405-11(2003).

73. Baruch, D.I. et al. *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, 3497-502(1996).

74. Springer, A.L. et al. Functional interdependence of the DBLbeta domain and c2 region for binding of the *Plasmodium falciparum* variant antigen to ICAM-1. *Mol Biochem Parasitol* 137, 55-64(2004).

75. Chattopadhyay, R. et al. Molecular analysis of the cytoadherence phenotype of a *Plasmodium falciparum* field isolate that binds intercellular adhesion molecule-1. *Mol Biochem Parasitol* 133, 255-65(2004).

76. Degen, R., Weiss, N. & Beck, H.P. *Plasmodium falciparum:* cloned and expressed CIDR domains of PfEMP1 bind to chondroitin sulfate A. *Exp Parasitol* 95, 113-21(2000).

77. Buffet, P.A. et al. *Plasmodium falciparum* domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection. *Proc. Natl. Acad. Sci. U.S.A* 96, 12743-8(1999).

78. Salanti, A. et al. Selective upregulation of a single distinctly structured *var* gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol. Microbiol* 49, 179-91(2003).

79. Chen, Q. et al. The semiconserved head structure of *Plasmodium falciparum* erythrocyte membrane protein 1 mediates binding to multiple independent host receptors. *J. Exp. Med* 192, 1-10(2000).

80. Smith, J.D. et al. Switches in expression of *Plasmodium falciparum var* genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* 82, 101-10(1995).

81. Horrocks, P. et al. Variable *var* transition rates underlie antigenic variation in malaria. *Proc. Natl. Acad. Sci. U.S.A* 101, 11129-34(2004).

82. Recker, M. et al. Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature* 429, 555-8(2004).

83. Peters, J. et al. High diversity and rapid changeover of expressed *var* genes during the acute phase of *Plasmodium falciparum* infections in human volunteers. *Proc. Natl. Acad. Sci. U.S.A* 99, 10689-94(2002).

84. Paget-McNicol, S. et al. The *Plasmodium falciparum var* gene switching rate, switching mechanism and patterns of parasite recrudescence described by mathematical modelling. *Parasitology* 124, 225-35(2002).

85. Gatton, M.L. et al. Switching rates of *Plasmodium falciparum var* genes: faster than we thought? *Trends Parasitol* 19, 202-8(2003).

86. Gatton, M.L. & Cheng, Q. Investigating antigenic variation and other parasite-host interactions in *Plasmodium falciparum* infections in naïve hosts. *Parasitology* 128, 367-76(2004).

87. Frank, M. et al. Variable switching rates of malaria virulence genes are associated with chromosomal position. *Mol. Microbiol* 64, 1486-98(2007).

88. Kyes, S.A. et al. 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, 9333-8(1999).

89. Limpaiboon, T. et al. Characterization of a *Plasmodium falciparum* epitope recognized by a monoclonal antibody with broad isolate and species specificity. *Southeast Asian J Trop Med Public Health* 21, 388-96(1990).

90. Winter, G. et al. SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes. *J Exp Med* 201, 1853-63(2005).

91. Sam-Yellowe, T.Y. et al. A *Plasmodium* gene family encoding Maurer's cleft membrane proteins: structural properties and expression profiling. *Genome Res* 14, 1052-9(2004).

92. Fernandez, V. et al. 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, 1393-404(1999).

93. Haeggström, M. et al. Common trafficking pathway for variant antigens destined for the surface of the *Plasmodium falciparum*-infected erythrocyte. *Mol Biochem Parasitol* 133, 1-14(2004).

94. Limpaiboon, T. et al. 7H8/6, a multicopy DNA probe for distinguishing isolates of *Plasmodium falciparum*. *Mol Biochem Parasitol* 47, 197-206(1991).

95. Prescott, N. et al. *Plasmodium falciparum* genetic diversity can be characterised using the polymorphic merozoite surface antigen 2 (MSA-2) gene as a single locus marker. *Mol Biochem Parasitol* 63, 203-12(1994).

96. McRobert, L. et al. Distinct trafficking and localization of STEVOR proteins in three stages of the *Plasmodium falciparum* life cycle. *Infect Immun* 72, 6597-602(2004).

97. Kaviratne, M. et al. Small variant STEVOR antigen is uniquely located within Maurer's clefts in *Plasmodium falciparum*-infected red blood cells. *Eukaryot Cell* 1, 926-35(2002).

98. Blythe, J.E. et al. *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, 3329-36(2008).

99. Khattab, A. et al. *Plasmodium falciparum* variant STEVOR antigens are expressed in merozoites and possibly associated with erythrocyte invasion. *Malar J* 7, 137(2008).

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

101. Deitsch, K.W., Calderwood, M.S. & Wellems, T.E. Malaria. Cooperative silencing elements in *var* genes. *Nature* 412, 875-6(2001).

102.Frank, M. et al. Strict pairing of *var* promoters and introns is required for *var* gene silencing in the malaria parasite *Plasmodium falciparum*. *J Biol Chem* 281, 9942-52(2006).

103. Calderwood, M.S. et al. *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, 34125-32(2003).

104. Kyes, S.A. et al. A well-conserved *Plasmodium falciparum var* gene shows an unusual stage-specific transcript pattern. *Mol Microbiol* 48, 1339-48(2003).

105. Duraisingh, M.T. et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* 121, 13-24(2005).

106. Freitas-Junior, L.H. et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 121, 25-36(2005).

107. Comeaux, C.A. & Duraisingh, M.T. Unravelling a histone code for malaria virulence. *Mol Microbiol* 66, 1291-5(2007).

108. Lopez-Rubio, J.J. et al. 5' flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. *Mol Microbiol* 66, 1296-305(2007).
109. Chookajorn, T. et al. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A* 104, 899-902(2007).
110. Ralph, S.A., Scheidig-Benatar, C. & Scherf, A. Antigenic variation in *Plasmodium falciparum* is associated with movement of *var* loci between subnuclear locations. *Proc. Natl. Acad. Sci. U.S.A* 102, 5414-9(2005).
111. Voss, T.S. et al. A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature* 439, 1004-8(2006).
112. Cooke, B., Coppel, R. & Wahlgren, M. Falciparum malaria: sticking up, standing out and out-standing. *Parasitol. Today (Regul. Ed.)* 16, 416-20(2000).
113. Freitas-Junior, L.H. et al. Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407, 1018-22(2000).
114. Taylor, H.M., Kyes, S.A. & Newbold, C.I. *Var* gene diversity in *Plasmodium falciparum* is generated by frequent recombination events. *Mol Biochem Parasitol* 110, 391-7(2000).
115. Barry, A.E. et al. Population genomics of the immune evasion (*var*) genes of *Plasmodium falciparum*. *PLoS Pathog* 3, e34(2007).
116. Kaestli, M. et al. Longitudinal assessment of *Plasmodium falciparum var* gene transcription in naturally infected asymptomatic children in Papua New Guinea. *J. Infect. Dis* 189, 1942-51(2004).
117. Flick, K. & Chen, Q. *var* genes, PfEMP1 and the human host. *Mol. Biochem. Parasitol* 134, 3-9(2004).
118. Gupta, S. et al. Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nat Med* 5, 340-3(1999).
119. Reeder, J.C. et al. Diversity of agglutinating phenotype, cytoadherence, and rosette-forming characteristics of *Plasmodium falciparum* isolates from Papua New Guinean children. *Am J Trop Med Hyg* 51, 45-55(1994).
120. Bull, P.C. et al. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nat. Med* 4, 358-60(1998).
121. Giha, H.A. et al. Antibodies to variable *Plasmodium falciparum*-infected erythrocyte surface antigens are associated with protection from novel malaria infections. *Immunol Lett* 71, 117-26(2000).

122. Dodoo, D. et al. Antibodies to variant antigens on the surfaces of infected erythrocytes are associated with protection from malaria in Ghanaian children. *Infect Immun* 69, 3713-8(2001).
123. Giha, H.A. et al. Overlapping antigenic repertoires of variant antigens expressed on the surface of erythrocytes infected by *Plasmodium falciparum*. *Parasitology* 119 (Pt 1), 7-17(1999).
124. Gamain, B., Miller, L.H. & Baruch, D.I. The surface variant antigens of *Plasmodium falciparum* contain cross-reactive epitopes. *Proc Natl Acad Sci U S A* 98, 2664-9(2001).
125. Ofori, M.F. et al. Malaria-induced acquisition of antibodies to *Plasmodium falciparum* variant surface antigens. *Infect Immun* 70, 2982-8(2002).
126. Kinyanjui, S.M. et al. Protection against clinical malaria by heterologous immunoglobulin G antibodies against malaria-infected erythrocyte variant surface antigens requires interaction with asymptomatic infections. *J Infect Dis* 190, 1527-33(2004).
127. Gratepanche, S. et al. Induction of crossreactive antibodies against the *Plasmodium falciparum* variant protein. *Proc Natl Acad Sci U S A* 100, 13007-12(2003).
128. Moll, K. et al. Generation of cross-protective antibodies against *Plasmodium falciparum* sequestration by immunization with an erythrocyte membrane protein 1-duffy binding-like 1 alpha domain. *Infect Immun* 75, 211-9(2007).
129. Marsh, K. & Howard, R.J. Antigens induced on erythrocytes by *P. falciparum*: expression of diverse and conserved determinants. *Science* 231, 150-3(1986).
130. Newbold, C.I. et al. *Plasmodium falciparum*: the human agglutinating antibody response to the infected red cell surface is predominantly variant specific. *Exp Parasitol* 75, 281-92(1992).
131. Bull, P.C. et al. Antibody recognition of *Plasmodium falciparum* erythrocyte surface antigens in Kenya: evidence for rare and prevalent variants. *Infect. Immun* 67, 733-9(1999).
132. Nielsen, M.A. et al.Geographical and temporal conservation of antibody recognition of *Plasmodium falciparum* variant surface antigens. *Infect. Immun* 72, 3531-5(2004).
133. Kirchgatter, K., Mosbach, R. & del Portillo, H.A. *Plasmodium falciparum*: DBL-1 *var* sequence analysis in field isolates from central Brazil. *Exp. Parasitol* 95, 154-7(2000).

134. Rottmann, M. et al. Differential expression of *var* gene groups is associated with morbidity caused by *Plasmodium falciparum* infection in Tanzanian children. *Infect. Immun* 74, 3904-11(2006).
135. Montgomery, J. et al. Genetic analysis of circulating and sequestered populations of *Plasmodium falciparum* in fatal pediatric malaria. *J. Infect. Dis* 194, 115-22(2006).
136. Montgomery, J. et al. Differential *var* gene expression in the organs of patients dying of falciparum malaria. *Mol. Microbiol* 65, 959-67(2007).
137. Kouprina, N. & Larionov, V. TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nat Rev Genet* 7, 805-12(2006).
138. Cham, G.K.K. et al. A semi-automated multiplex high-throughput assay for measuring IgG antibodies against *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1) domains in small volumes of plasma. *Malar J* 7, 108(2008).
139. Rasti, N., Wahlgren, M. & Chen, Q. Molecular aspects of malaria pathogenesis. *FEMS Immunol. Med. Microbiol* 41, 9-26(2004).

Appendix 1

Primer sequences (including restriction sites) used to amplify PFD1235w/*var4* fragments.

Primer name	Primer sequence
FWD1	5'-GGATCCATGGGGAATGCATCATCATCAG-3'
RWD1	5'-GGATCCATCACGACATGCTTCCTTAACC-3'
FWD2	5'-GGATCCAATGATTATTGTGGCCATGGTG-3'
RWD2	5'-GGATCCTGCGCCTTCTGAATCTTTCGTAC-3'
FWD3	5'-GGATCCACACCATCATCACACAAAGTTCC-3'
RWD3	5'-GGATCCTTCGGCCCATTCCGTCATCC-3'
FWD4	5'-GGATCCTTACGTTCAGACTGGTGGGAAG-3'
RWD4	5'-GGATCCAAGTAACACATCGCCCAATAAGG-3'
FWD5	5'-AGATCTGATTATAGTCGTGGAGGTACG-3'
RWD5	5'-AGATCTGCAATCACACGCTGTAGCATAC-3'
FWD6	5'-AGATCTGGTGTGCACACCGTGTATTC-3'
RWD6	5'-AGATCTACTGGCGACGTCATCGATAC-3'
FWD7	5'-GGATCCGGTACTAGTGATGCTACGGG-3'
RWD7	5'-GGATCCAGCGGACTTCACAAAGGCGTG-3'
FWD8	5'-GGATCCGACACAAGTGAGAATGGTGCC-3'
RWD8	5'-GGATCCATTACCTCCATCACCATTACTTC-3'
FWD9	5'-GGATCCGGTGCTAGTGGTACCGGCG-3'
RWD9	5'-GGATCCACCTGCACCCTGACAAATGCC-3'
FWD10	5'-AGATCTGGTAGTTGTGGGAGTGCTAAG-3'
RWD10	5'-AGATCTCGCCGCAAAACCGATACCTAC-3'

Primer sequences (including restriction sites) used to amplify FCR3S1.2-*var1* fragments

Primer name	Primer sequence	Annealing temperature
Frag_fwd1a	5'-AACCATGGCGACTTCAGGAGG-3'	55.2°C
Frag_rev1a	5'-TTTAGATCTTACTGTTTCTCGATTCG-3'	55.2°C
Frag_fwd1b	5'-AACCATGGCAGTATGGGGAGCC-3'	54.0°C
Frag_rev1b	5'-TTTAGATCTGCAATATTCCGATCG-3'	54.0°C
Frag_fwd2	5'-AACCATGGAACCCTGTCCCGACTGTGG-3'	56.5°C
Frag_rev2	5'-TTTAGATCTCGCTCCTTCTTGTGCTACTGCC-3'	56.5°C
Frag_fwd3	5'-AACCATGGGACATGGATTACCACGTGTCG-3'	56.5°C
Frag_rev3	5'-TTTAGATCTTAGACCACTGACCCCACAATTACC-3'	56.5°C
Frag_fwd4	5'-AACCATGGCAAATGGGAACTGCGATGGTAAAG-3'	50.2°C
Frag_rev4	5'-TTTAGATCTGCGTAAAGGTGGAGGGGTATCAG-3'	50.2°C

Appendix 2

Amino acid sequences of PFD1235w/*var4* synthetic peptides

Peptide Name	Peptide Sequence
var4-**2.1**	Biotin-PEG-RNQRNEFEKQKKKYYKEIQTYTSKDAKTDS
var4-**2.2**	Biotin-PEG-WAKTKENEWKKVKTIYKNENGNTNNYYKKL
var4-**3**	Biotin-PEG-SKDSCPPSVDTKTNPCAKPPGSKPTKSVKQ
var4-**4.1**	Biotin-PEG-GGKKGPPPATHPYKSVNTRDKRDATDDTTP
var4-**4.2**	Biotin-PEG-EANETMLKNSSNGNDKDESKLKGKAEEGDY
var4-**7**	Biotin-PEG-TWLKNWKTQYKTQSKKYFDDKRKELYKSID
var4-**8**	Biotin-PEG-YNTDSNGKDKKIQQVKATDNTDLFQKLKKD
var4-**9**	Biotin-PEG-RSYRKWIERKKTEYEKQESAYSKQKSNYVN
var4-**10**	Biotin-PEG-EEEEETDSHIYEDYSDSDAEEDDEDEAVTE

CURRICULUM VITAE

PERSONAL DATA

Name	Nicole Ingrid Falk
Date of Birth	November 22, 1979
Place of Birth	Bühl, Germany
Nationality	German
Affiliation	Swiss Tropical Institute Socinstrasse 57 4051 Basel Switzerland
	Phone: +41 61 284 8211 Fax: +41 62 284 8101
	Email: nicole.falk@stud.unibas.ch

EDUCATION

1990-1999	High School, Hans-Thoma-Gymnasium Lörrach, Germany Grade: 2.1
1999-2004	Diploma in "Integrative Biology", Biozentrum, University of Basel, Switzerland. Grade: 5.3 Diploma thesis in Molecular Epidemiology, Swiss Tropical Institute, University of Basel, Switzerland Subject:

"Automated genotyping of Plasmodium falciparum using the merozoite surface protein 2 as polymorphic marker gene."

2005 - 2008 PhD at the Swiss Tropical Institute, University of Basel, Switzerland

Subject:

"*Var* gene diversity and their serological recognition by naturally exposed individuals"

MEETINGS AND PRESENTATIONS

2005 PhD student meeting of the Swiss Society for Tropical Medicine and Parasitology (SSTMP), Ascona, Switzerland.

Talk: „Expression dynamics of *var* genes in *Plasmodium falciparum"*

Attandance at the "COST action meeting 857", Beatenberg, Switzerland.

2006 Attandance at the Joint Meeting of the Royal Society of Tropical Medicine and Hygiene and the Swiss Society for Tropical Medicine and Parasitology, Basel, Switzerland

2007 Joint Meeting of the French Society for Parasitology, the German Society for Parasitology and the Swiss Society for Tropical Medicine and Parasitology, Strasbourg, France

Poster: „Analysis of Plasmodium falciparum *var* gene repertoire expressed in children from Papua New Guinea"

Union of the Swiss Society for experimental Biology (USGEB), Basel, Switzerland

Poster: „Analysis of Plasmodium falciparum *var* gene repertoire expressed in children from Papua New Guinea"

PhD student meeting of the Swiss Society for Tropical Medicine and Parasitology (SSTMP), Münchenwiler, Switzerland
Talk: "Identification of immunodominant epitopes of *Plasmodium falciparum* erythrocyte membrane protein 1"

TECHNIQUES

Molecular Biology:

General molecular biology technologies such as cloning and sequencing, quantitative real-time PCR, magnetic bead-based isolation of gene-specific mRNA, RT-PCR ELISA and SDS PAGE/Western blot.

Bioinformatics:

Application of various programs for sequence analysis and database search, e.g. Seqman of DNAstar, CLUSTALW, CLUSTALX, Phylip, NCBI BLAST, PlasmoDB and GeneMapper.

Additional Courses:

2006- 2007 Participation in the 6.th Round of the Women into Industry Mentoring Program, a Collaboration between the University of Basel and Novartis

2008 Good Clinical Practice Course training modules 1&2 at the Center for Clinical Research in Zurich.Courses included the following topics:
Study methodology, GCP principles, study documents, study procedures, adverse event reporting, quality assurance, SOPs, archiving and statistical principles.

LANGUAGES German, mother tongue
English, fluent
French, basic knowledge

PUBLICATIONS

Falk N, Maire N, Sama W, Owusu-Agyei S,Smith T, Beck HP,Felger I. Comparison of PCR-RFLP and Genescan-based genotyping for analyzing infection dynamics of *Plasmodium falciparum*. Am J Trop Med Hyg. 2006 Jun;74(6):944-50.

Mugittu K, Abdulla S, **Falk N**, Masanja H, Felger I, Mshinda H, Beck HP, Genton B. Efficacy of sulfadoxine-pyrimethamine in Tanzania after two years as first-line drug for uncomplicated malaria: assessment protocol and implication for treatment policy strategies.Malar J. 2005 Nov 18;4(1):55.

Falk N, Kaestli M, Qi W, Ott M, Baea K, Cortés A, Beck HP. Analysis of Plasmodium falciparum *var* genes expressed in children from Papua New Guinea.
Manuscipt submitted to the Journal of infectious diseases.

Falk N, Wittmer K, Beck HP. Application of capillary electrophoresis sizing technique as new *var* gene genotyping tool
Manuscipt in preparation.

REFERENCES

Prof. Dr. Hans-Peter Beck	Supervisor of my PhD, Swiss Tropical Institute, Basel, Switzerland. E-mail: hans-peter.beck@unibas.ch Phone: 061 284 81 16
.	
PD Dr. Ingrid Felger	Supervisor of my Diploma thesis, Swiss Tropical Institute, Basel, Switzerland. E-mail: ingrid.felger@unibas.ch Phone: 061 284 81 17

During my studies I attended lectures and courses of the following lecturers:

H.C. Imhof, H.J. Güntherodt, C. Schönenberger, H.Sigel, U.Séquin, A.Wiemken, H.P. Hauri, Senn, M. Spiess, M. Affolter, B. Baur, W. Gehring, Strazewski, T. Boller, W.Keller, U. Jenal, U. Aebi, S. Arber, H. Reichert, T. Schwede, M. Rüegg, G. Cornelis, C. Dehio, R. Brun, A.Seelig, I. Felger, H.P. Beck, G. Pluschke, C. Daubenberger, N. Weiss, M. Tanner, P. Vounatsou, T. Smith, M. Lister.