Molecular Epidemiology and Population Genetics of

*Plasmodium vivax* in Papua New Guinea

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Cristian Köpfli

aus

Hochdorf LU

Basel, 2012

Original document stored on the publication server of the University of Basel edoc.unibas.ch

This work is licenced under the agreement „Attribution Non-Commercial No Derivatives – 2.5 Switzerland“. The complete text may be viewed here:

creativecommons.org/licenses/by-nc-nd/2.5/ch/deed.en
Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Dr. PD Ingrid Felger, Dr. Ian Hastings, Prof. Dr. Marcel Tanner


Prof. Dr. Martin Spiess

Dekan

Attribution-Noncommercial-No Derivative Works 2.5 Switzerland

You are free:

to Share — to copy, distribute and transmit the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

• For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
• Any of the above conditions can be waived if you get permission from the copyright holder.
• Nothing in this license impairs or restricts the author’s moral rights.

Your fair dealing and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license) available in German: http://creativecommons.org/licenses/by-nc-nd/2.5/de/legalcode.de

Disclaimer:
The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license. Creative Commons is not a law firm and does not provide legal services. Distributing, displaying of, or linking to this Commons Deed does not create an attorney-client relationship.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/de/deed.de
Datum: 3.4.2009
# Table of Contents

Acknowledgements ........................................................................................................... 5  
Summary ................................................................................................................................. 9  
Zusammenfassung .................................................................................................................. 11  
Abbreviations ....................................................................................................................... 15  
Introduction .......................................................................................................................... 17  
Malaria: Affecting Humanity from Individuals to the Wealth of Nations ......................... 17  
The Burden of *Plasmodium vivax* Malaria ....................................................................... 18  
Comparison of *P. vivax* and *P. falciparum* Morbidity .................................................... 19  
Specific Aspects of *P. vivax* Cell Biology, Life Cycle and Epidemiology ....................... 20  
*P. vivax* Specific Challenges in Malaria Elimination and Eradication ......................... 23  
Malaria in Papua New Guinea ............................................................................................ 24  
Context of this PhD Project ................................................................................................. 25  
Goal and Objectives of the Thesis ...................................................................................... 28  
References ........................................................................................................................... 30  
How Much Remains Undetected? Probability of Molecular Detection of Human *Plasmodia* in the Field ............................................................................................................. 39  
Multiplicity and Diversity of *Plasmodium vivax* Infections in a Highly Endemic Region in Papua New Guinea .................................................................................................................. 47  
*Plasmodium vivax* Populations in Papua New Guinea: Highly Diverse and Unstructured.... 65  
A High Force of Infection Drives the Rapid Natural Acquisition of Immunity to *Plasmodium vivax* in Papua New Guinean Children .......................................................................................... 93  
General Discussion ............................................................................................................... 119  
The Complex Pattern of Malaria Infection and Disease .................................................... 119  
*P. vivax* and Malaria Elimination ...................................................................................... 119  
*P. vivax* Epidemiology in Papua New Guinea ................................................................... 120  
Technical Considerations ..................................................................................................... 124  
Conclusions .......................................................................................................................... 126  
References ........................................................................................................................... 127  
Appendix ............................................................................................................................... 131  
Curriculum Vitae .................................................................................................................. 135
Acknowledgements

First and most of all I would like to thank my supervisor Ingrid Felger for giving me the chance to do my PhD with her at the Swiss Tropical and Public Health Institute. I very much appreciated to have the chance to work independently on my project according to my own scheme and to follow my interests, while always having the opportunity to discuss questions and ideas. I am also grateful to Ingrid that she encouraged me throughout my thesis to present my results at several conferences.

I few other people have worked closely with me throughout the last three years, and I would like thank especially Amanda Ross, Tom Smith and Ivo Mueller. Amanda has helped a lot with data analysis and interpretation and in addition she has corrected many of my manuscript drafts. Her comments on my writing were often sharp, but always for a good reason. A few examples? Here you go: “Two ‘can’s in one sentence sounds un-smooth!” – “Figure 2a and 2b are ugly!!! Please delete the gridlines. We also need to add CI – I can calculate them from the regression…. else I’m afraid it’s a bit embarrassing.” – “OK, enough! It’s OK to use the sentence structure. ‘blah likely does blah’ rather than ‘blah is likely to do blah’ once or twice in a paper, but every time comes across as pompous and repetitive. Sorry!” I am grateful for her help and I am sure I improved my scientific writing skills thanks to her. Much of the credit for the following comment, given by an anonymous reviewer, goes to Amanda: “The manuscript is very well written and the data clearly presented and discussed.”

Tom has discussed my results at many opportunities, given input and helped with manuscript preparation. My favored quote from him: “That’s fine. Glad you got the point in the end. Learning is what a PhD should be about”, after I finally understood a small inconsistency in my analysis he had mentioned several times. Likewise I would like to thank Ivo for discussing my projects. Ivo knows the latest developments in malaria research in PNG as few others do, by sharing this knowledge he helped a lot with the interpretation of my results. Further I thank Katie Benton for the analysis of the force of infection data. And of course a big “Tenkyu tru!” to all the study participants and all the people in PNG who were involved in the collection of samples.

I am looking forward to the upcoming defense as last step of my PhD, and I would like to thank Marcel Tanner for being my faculty representative, and Ian Hastings for joining my exam as external expert.

Special thanks go to my predecessor as PhD student, Sonja Schöpflin. It was a pleasure to find a perfectly organized sample collection in the freezer, as well her good documentation of the P. falciparum clones in the cohort. And of course her advice on capillary electrophoresis at the beginning of my project had helped a lot.
A big “thank you!” for all the help in the lab goes to Dania, especially for all the orderings of lab material he has done for me. As the number 1 PCR reagents user in the group I was heavily depended of our stocks of Taq, dNTPs and buffers, and I am very grateful to Dania that I have always found full fridges!

I am also grateful to Lincoln for his help with the microsatellite typing for the population genetics project. He has done an excellent job. When I first met him in PNG in April 2010 he had never left his country. I am impressed about his curiosity when he came here a few months later for an internship. Not only did he want to learn about science, but also about Switzerland and Basel. He even picked up some German, as proven by the following e-mail, sent to me only a few weeks after his arrival: “Kein Problem! Ich werde beginnen!”

For the population genetics project I would further like to thank Tim Davis and Harin Karunajewa for providing samples from Alexishafen, as well as Jutta Marfurt for the samples from Kunjingini and Sigimaru. A big thanks goes to Alyssa Barry for her advice on that project, and to Tiago Antão. Tiago’s excellent skills in population genetics data analysis helped me a lot to understand my own data, and thanks to his hint I could attend the Congen course in the US.

The open atmosphere at Swiss TPH led to many discussions on diverse scientific topics. I am grateful to those group leaders, postdocs, technicians and students from other groups who have shared their scientific experiences with me. I learned a lot through the discussion of other research projects I was not directly involved in, and the input of other people for my own work was always very valuable. Especially I would like to thank all members of my group, Peter Beck and the molecular parasitology group, and Till Voss and the gene regulation group with whom I have discussed most.

During my PhD I had the opportunity to visit Papua New Guinea, and I am glad that I met many people who made my stay there a very good experience. I thank Peter Siba for welcoming me at the institute, and Lincoln, Livingstone, Dulcie, Elisheba and Mak for being motivated students. I am grateful to Serej, Manuel, Céline and Claire that I could stay in their houses, and Brioni for the many times she picked me up or drove me home. Thanks to Beverlyn for guiding me in Madang, and to Leanne, Lisa, Haydn and all the other people in Madang for celebrating my birthday with me. Big thanks go to Mama Janet and Papa John for their hospitality in Kina Beach.

I am also grateful to Ingrid and Peter for the many opportunities to discuss other things that science, such as exhibitions of the latest developments of Peter’s own art collection. And of course the exceptional Christmas dinners were excellent, no other group will be able to top them! For those who have never attended one: a wonderful journey of excellent food and drinks from 7 pm to somewhat around 4 am – it can’t be described.
I had very interesting discussions with Till where he shared some of his experiences as a postdoc and young professor. I am grateful for that and I hope that some of his hints will help me during my own scientific career.

Further I would like to thank all my colleagues from the PhD student office, Urs (the Knuddelbear), Katrin (Cindy Fairweather), Pax (“Nataka kula pundamilia”), Theresia (Meppen!!), Lucienne (“****wurscht”), Ralf (the trader), Marie (“Un jour sur deux, ça va comme sur des roulettes”), Philipp (the hobby-palindromist), and all the other people on the floor, Matthias, Sergio, Angèle, Christoph, Sonja, Mireia; it was good to have people around for coffee breaks or just for a little chat to relax from scientific thinking.

I also would like to thank to those people at Swiss TPH who have helped in the background to make this thesis a success. I thank the library team for finding copies of old articles or of papers published in slightly peculiar journals, the IT people for solving all problems with computers, and the administration and the technical services for keeping the institute running.

Finally I would like to thank those anonymous reviewers who have given comments on my manuscripts submitted so far. Their input has helped a lot to improve my publications, and of course it is a pleasure to read the following statement: “It was a pleasure reading this revised manuscript by Koepfli et al.”

Last but not least I would like to thank my parents and my girlfriend Anita. Anita has helped me with some computation tasks during my PhD and had a good bottle of wine in the cellar when I needed some relaxation from work. My parents have always supported my decision to study biology and I am grateful for that!

I thank the Swiss National Science Foundation for financial support, as well as the Swiss Society of Tropical Medicine and Parasitology for their financial support to attend a conference and a course.
Summary

In recent years a new focus has been put on malaria, and elimination and eradication of this disease has been brought back to the agenda. While *Plasmodium falciparum* causes most of the disease in Africa, *P. vivax* is predominant in malaria endemic regions in Latin America, Asia and the South Pacific. In the last 10 years increasing reports of severe disease and even mortality caused by *P. vivax* raised new awareness for this parasite. Forty percent of the world population lives in areas of *P. vivax* transmission, and severe disease is caused especially in infants and young children. *P. vivax* has the ability to form semi-dormant liver hypnozoites that can relapse and lead to clinical disease after an extended period of time. As most currently applied drugs are not effective against hypnozoites, they present a major obstacle towards malaria control.

In the lowlands of Papua New Guinea (PNG), malaria prevalence reaches levels similar to those in sub-Saharan Africa. Both *P. falciparum* and *P. vivax* are frequent, and *P. vivax* reaches prevalence higher than anywhere else in the world. Individuals from these regions are often co-infected by both parasite species. Multiple concurrent infections with different strains of *P. falciparum or P. vivax* are common. For this thesis I have genotyped over 3000 *P. vivax* positive blood samples collected in a cohort study of 264 children. These children, aged 1 to 4.5 years, were followed over 16 months and visited every 8 weeks for collection of two blood samples taken 24 hours apart. Additional samples were collected whenever the children presented sick to the local health centre. In parallel morbidity data was collected. *P. vivax* was genotyped using two size polymorphic markers, which both distinguished individual parasite clones. Genotyping data of *P. falciparum* clones was available from a previous study.

Extensive diversity of the genotyping markers and high multiplicity of infection (MOI, the number of co-occurring clones per carrier) was observed. Each *P. vivax* positive child carried a mean of 2.7 concurrent infections, mean MOI was nearly twice as high as *P. falciparum* multiplicity in the cohort. We did not detect a seasonal trend in MOI, and only a moderate increase of MOI with age, most pronounced in very young children up to 3 years. Most likely relapses increase *P. vivax* mean MOI, and they also lead to high MOI during seasons of less transmission.

The detection of parasitemia in field surveys is imperfect owing to parasite densities below or fluctuating around the detection limit of light microscopy or PCR. We have compared >1000 pairs of samples collected 24 hours apart. The collection of a second sample was found to have a limited effect on parasite prevalence: a single PCR missed only 6 to 9% of the combined parasite positivity. The proportion of individual clones missed was more pronounced. Depending on the marker, 19% or 31% of all *P. vivax* clones were missed when sampling on a single day. As a consequence, mean MOI increased from 2.7 to 3.4 when
combining results from paired samples. Detectability does not differ between *P. vivax* and *P. falciparum* or between age groups. Thus comparisons of prevalence and MOI between species and age trends are not biased by detectability.

To investigate structuring of *P. vivax* population in PNG, we have typed a subset of samples with 13 additional markers. Allele frequencies and haplotypes were compared to those from samples collected at other sites in PNG. No differences were detected between parasite populations, most likely because of high gene flow between sites. This is in contrast to the situation in countries of lower endemicity, and differs from the earlier observed moderate structuring in *P. falciparum* populations from PNG.

In this cohort study, the incidence of *P. vivax* malaria decreased with age, whereas the incidence of *P. falciparum* peaked later at the age of 3.5 years. These contrasting trends are in line with results from other countries where both parasites are co-endemic. Differences in the parasite biology, e.g. a more limited reservoir of surface antigens of *P. vivax* compared to *P. falciparum*, might cause faster immunization against *P. vivax* and thus the different age trends in incidence. But also differences in transmission intensity might be responsible. Our typing of clones over an extended period of time allowed estimations of the molecular force of infection (molFOI), i.e. the number of infections per child per year. Each child acquired 14 *P. vivax* clones per year, but only 6 *P. falciparum* clones. molFOI did not change with age, suggesting that acquired immunity builds up gradually with each new clone, leading to a decrease in *P. vivax* incidence over time. This does not rule out that biological differences between *P. vivax* and *P. falciparum* also play a role.

PNG is among the countries with the poorest health infrastructure and the highest prevalence of malaria. *P. vivax* prevalence is in some regions as high as *P. falciparum* prevalence. High MOI and molFOI, as well as high levels of gene flow imply great resilience of *P. vivax* towards antimalarial interventions. Fast acquisition of semi-immunity leads to a large number of asymptomatic *P. vivax* carriers, potentially contributing to transmission, and control is further complicated by relapses, occurring in the blood weeks or months after transmission from the mosquito. The high proportion of *P. vivax* will be a major challenge towards malaria control and elimination in PNG.
**Zusammenfassung**


Die gewählten Marker wiesen eine hohe Diversität auf, und die Multiplizität (die Anzahl verschiedener Stämme pro Träger) war hoch. Jedes infizierte Kind trug im Schnitt 2.7 *P. vivax* Stämme, dieser Wert ist fast doppelt so hoch wie derjenige von *P. falciparum*. Es gab keine saisonalen Unterschiede in der Multiplizität, und nur einen geringen Anstieg der Multiplizität mit dem Alter der Kinder; dieser war am deutlichsten in sehr jungen Kindern unter 3 Jahren. Es ist davon auszugehen, dass wiederausbrechende Leberstadien die Multiplizität erhöhen, dies auch während Jahreszeiten mit geringerer Übertragung durch Mücken.

In Feldstudien ist die Detektion von Parasiten häufig unvollständig, d.h. die Dichte eines Teils der Parasiten liegt unter der Nachweigrenze von Mikroskopie oder molekularen Methoden wie PCR. Wir haben über 1000 Blutproben-Paare verglichen, welche im Abstand von 24 Stunden gesammelt wurden. Die zusätzliche zweite Blutprobe hatte einen eher geringen
Einfluss auf die Prävalenz, mittels PCR wurden nur 6 bis 9% der Proben falsch negativ diagnostiziert. Der Anteil der Stämme, die in einer einzelnen Probe nicht detektiert wurden, war dagegen ausgeprägter. Je nach Marker wurden 19 oder 31% aller *P. vivax* Stämme, die an mindestens einem der 2 Tage nachgewiesen wurden, am Tag 1 nicht entdeckt. Als Folge stieg die Multiplizität von 2.7 auf 3.4 wenn die Resultate von beiden Tagen kombiniert wurden. Die Wahrscheinlichkeit, einen Stamm zu detektieren, unterscheidet sich nicht zwischen *P. vivax* und *P. falciparum* oder in verschiedenen Altersklassen, folglich werden Vergleiche der Prävalenz und Multiplizität nicht beeinflusst.

Um die Populationsstruktur von *P. vivax* in Papua Neuguinea zu untersuchen haben wir einen Teil der Proben aus unserer Kohorte mit 13 weiteren molekularen Markern untersucht. Die Allele-Frequenzen und die Haplotypen wurden verglichen mit denjenigen von drei anderen Orten in Papua Neuguinea. Es wurden keine Unterschiede zwischen diesen Parasitenpopulationen festgestellt, sehr wahrscheinlich aufgrund eines häufigen genetischen Austauschs zwischen den Populationen. Im klaren Gegensatz dazu wurden in Gebieten mit geringer *P. vivax* Prävalenz ausgeprägte Unterschiede zwischen Populationen beobachtet, und in früheren Studien wurde auch eine moderat ausgeprägte Populationsstruktur von *P. falciparum* in Papua Neuguinea festgestellt.


### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ama1</td>
<td>apical membrane protein 1</td>
</tr>
<tr>
<td>CE</td>
<td>capillary electrophoresis</td>
</tr>
<tr>
<td>csp</td>
<td>circumsporozoite protein</td>
</tr>
<tr>
<td>DALY</td>
<td>disability-adjusted life-year</td>
</tr>
<tr>
<td>dbp</td>
<td>Duffy binding protein</td>
</tr>
<tr>
<td>dhfr</td>
<td>dihydrofolate reductase</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>FOI</td>
<td>force of infection</td>
</tr>
<tr>
<td>HE</td>
<td>expected heterozygosity = virtual heterozygosity</td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional Review Board</td>
</tr>
<tr>
<td>ITN</td>
<td>insecticide treated bed net</td>
</tr>
<tr>
<td>LD</td>
<td>linkage disequilibrium</td>
</tr>
<tr>
<td>LDR-FMA</td>
<td>ligase detection reaction-fluorescent microsphere assay</td>
</tr>
<tr>
<td>LM</td>
<td>light microscopy</td>
</tr>
<tr>
<td>MOI</td>
<td>multiplicity of infection</td>
</tr>
<tr>
<td>mFoI</td>
<td>molecular force of infection</td>
</tr>
<tr>
<td>MRAC</td>
<td>Medical Research Advisory Committee</td>
</tr>
<tr>
<td>MS</td>
<td>Microsatellite</td>
</tr>
<tr>
<td>msp</td>
<td>merozoite surface protein</td>
</tr>
<tr>
<td>msp1F3</td>
<td>F3 region of the merozoite surface protein 1</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PIEMP-1</td>
<td>P. falciparum Erythrocyte Membrane Protein 1</td>
</tr>
<tr>
<td>PNG</td>
<td>Papua New Guinea</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
</tr>
<tr>
<td>PNG IMR</td>
<td>Papua New Guinea Institute of Medical Research</td>
</tr>
<tr>
<td>PvAPI</td>
<td><em>P. vivax</em> annual parasite index</td>
</tr>
<tr>
<td>RBC</td>
<td>red blood cell</td>
</tr>
<tr>
<td>RDT</td>
<td>rapid diagnostic test</td>
</tr>
<tr>
<td>RFLP</td>
<td>restriction fragment length polymorphism</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>Swiss TPH</td>
<td>Swiss Tropical and Public Health Institute</td>
</tr>
</tbody>
</table>
Introduction

Malaria: Affecting Humanity from Individuals to the Wealth of Nations

Malaria is one of the most deadly parasitic diseases in the recent history of mankind (Hay et al., 2004). Despite a recent decline in the number of deaths an estimated 781,000 people have died from malaria in 2009 (World Health Organization, 2010b). It affects humans from the individual level to that of health systems or wealth of nations.

Many individuals in the tropics are affected by malaria throughout their life. During pregnancy millions of mothers and their unborn babies are especially vulnerable to malaria (Dellicour et al., 2010), and life starts for many newborns in endemic areas with reduced birth weight because of malaria during pregnancy. Early childhood is characterized by repeated clinical episodes of malaria. Morbidity and mortality peak in children below 5 years. In school children malaria causes repeated absence from school and caretakers are kept away from work. Adults, despite developing partial immunity during teenage years in areas of intense transmission, still suffer occasionally from sporadic malarial episodes, keeping them away from studies or work. Malaria accounts for up to 10% of all disability-adjusted life-years (DALYs) in regions of intense transmission (Breman et al., 2006). The United Nations have declared to combat malaria and other infectious diseases as one of the millennium development goals (www.un.org/millenniumgoals).

Health costs caused by malaria and loss of income due to disease account for a substantial part of household costs in some countries (Russell, 2004), and it has been observed that in countries of high malaria prevalence economic growth lags behind that in comparable non-malarial countries (Sachs and Malaney, 2002). It has been estimated that the growth is reduced by 1.3% per year in countries with intensive malaria as compared to non-endemic countries (Gallup and Sachs, 2001).

Malaria has been called “the strongest known force for evolutionary selection in the recent history of the human genome” (Kwiatkowski, 2005). The sickle-cell disease, a haemoglobin disorder that leads to resistance of severe forms of malaria in heterozygous individuals, but is lethal when homozygous, has become a textbook example of adaptation (Barreiro and Quintana-Murci, 2010).

Malaria is caused by different Plasmodium species, unicellular parasites of the Phylum Apicomplexa. Parasites are transmitted by mosquitoes of the genus Anopheles and have probably infected humans for ten thousands or even million of years (Carter, 2003; Escalante et al., 2005; Pick et al., 2011; Silva et al., 2011).
Introduction

Hundreds of *Plasmodium* species infect mammals, birds and reptiles, and four species – *P. falciparum*, *P. vivax*, *P. malariae* and *P. ovale* – are frequently infecting humans. *P. falciparum* accounts for most of the malaria cases in sub-Saharan Africa, the region suffering most from malaria, while *P. vivax* is the predominant species outside Africa and *P. malariae* and *P. ovale* account for smaller proportions of cases (Guerra et al., 2010; Hay et al., 2009; Mueller et al., 2007; World Health Organization., 2010b). More recently, reports from Malaysia have shown *P. knowlesi*, usually known as an old-world *Macaca* monkey malaria parasite, to infect humans (Cox-Singh et al., 2008; Singh et al., 2004). It is currently unclear whether human-to-human transmission of *P. knowlesi* occurs under natural conditions.

The Burden of *Plasmodium vivax* Malaria

It has been estimated that 2.85 billion people, around 40% of the world population, live in areas of *P. vivax* transmission (Figure 1) (Guerra et al., 2010). Estimations of the total number of cases per year range from 60-70 million (Mendis et al., 2001) to 391 million (Price et al., 2007) or up to 435 million (Baird, 2007). Even if the latter numbers represent overestimates and no recent publications have come to a similar number, there is no doubt that *P. vivax* accounts for a significant proportion of all human malaria cases.

---

**Figure 1:** The global spatial limits of *Plasmodium vivax* malaria transmission in 2009. Red = stable transmission (*PvAPI* (*P. vivax* annual parasite index) $\geq 0.1$ per 1,000 people p.a.), pink=unstable transmission (*PvAPI<0.1* per 1,000 p.a.). Areas where the prevalence of the Duffy negative blood group (conferring nearly complete resistance to *P. vivax*) was estimated to $>90\%$ are hatched. Source: Malaria Atlas Project (Guerra et al., 2010)
Despite the high number of *P. vivax* cases and global distribution in tropical areas, malaria research has focussed for most of the 20th century on *P. falciparum*. *P. falciparum* is the most deadly malaria parasite, and the dominating parasite in sub-Saharan Africa, the region by far most affected by malaria. The number of people living in areas of *P. falciparum* transmission is estimated to 1.38 billion (about half of them living in Africa and the other half in Asia) (Hay et al., 2009). This number of individuals at risk is much lower than 2.85 billion people living in *P. vivax* transmission areas (Guerra et al., 2010). The number of yearly *P. falciparum* cases is estimated to over 450 million (Hay et al., 2010), and the vast majority of the up to one million people that die each year from malaria are infected with *P. falciparum*, (World Health Organization., 2010b). Most of them are children from sub-Saharan Africa below 5 years (Snow et al., 1999). The high *P. falciparum* prevalence and mortality led to neglect of research on other *Plasmodium* spp. Only in recent years intensified research on *P. vivax* was called for, and *P. vivax*-specific research questions were raised (Carlton et al., 2011; de Lacerda et al., 2007; Galinski and Barnwell, 2008; Mendis et al., 2001; Mueller et al., 2009a; Price et al., 2007; Sina, 2002).

**Comparison of *P. vivax* and *P. falciparum* Morbidity**

Numerous reports on severe and even lethal outcome of *P. vivax* malaria led to a change in perception of *P. vivax* malaria. Traditionally called “benign tertian malaria”, *P. vivax* proved to be far more harmful than earlier appreciated (Anstey et al., 2009; Kochar et al., 2005).

A number of studies have compared severity of disease caused by *P. falciparum* and *P. vivax*. Some studies compared patients from areas with sympatric *P. falciparum* and *P. vivax* transmission who were infected with either species and admitted to the same hospital. This allows direct comparison of disease outcome of, with little confounding of external factors such as quality of treatment, general level of immunization in the population, use of bed nets and history of drugs used in the area. Genton and coworkers analyzed over 9’500 malaria cases with confirmed parasitemia in the Wosera area of Papua New Guinea (Genton et al., 2008). Overall, *P. falciparum* was over 3 times as frequent as *P. vivax* based on microscopic diagnosis, however, the ratio of cases that lead to severe malaria (defined as parasitemia plus a recent history of fits, coma, breathing problems, or anemia) was similar for both species: 11.7% for *P. falciparum* and 8.8% for *P. vivax* (Genton et al., 2008). Poespoprodjo and coworkers have studied 1560 infants that were admitted with malaria to a hospital in Papua, Indonesia, with similar numbers of *P. falciparum* and *P. vivax* infections (Poespoprodjo et al., 2009). The rate of deaths attributed to *P. falciparum* was higher compared to *P. vivax* (2.2% vs. 1.0%, not significant), while severe anemia was significantly
higher in *P. vivax* patients (Poespoprodjo et al., 2009). In another study in Papua, Indonesia, *P. vivax*-associated coma occurred 23 times less often than in *P. falciparum* malaria (Lampah et al., 2011). In the same hospital, *P. falciparum* accounted for two out of three hospital admissions due to Malaria, but the risk of severe disease was slightly, but significantly higher (23% vs. 20%) in patients presenting with *P. vivax* compared to *P. falciparum* (Tjitra et al., 2008). In contrast in northeastern Papua, Indonesia, among all patients admitted to the hospital with malaria, the risk of severe disease was about twice as high in patients positive for *P. falciparum* compared to *P. vivax* (7% vs. 3%) (Barcus et al., 2007). However, the rate of fatal outcome in those patients suffering from severe malaria was identical for both species (Barcus et al., 2007). Papua New Guinean children aged 5 to 14 years acquired similar numbers of *P. falciparum* and *P. vivax* infections, but the risk of developing symptomatic malaria was 21 times higher in *P. falciparum* infections (Michon et al., 2007).

In summary, these studies show the general picture of a more severe outcome of falciparum malaria compared to *P. vivax* infections, however, depending on the symptoms measured and on the age group studied, vivax malaria can cause nearly as much morbidity. An age shift between *P. vivax* and *P. falciparum* incidence illustrates best the different patterns of infections by both malaria species. *P. vivax* malaria reaches highest prevalence and causes most disease in the first year of life. *P. falciparum* becomes the dominating parasite at the age of 2 to 5 few years and caused most clinical cases throughout childhood and adolescence (Barcus et al., 2007; Genton et al., 2008; Kochar et al., 2010; Poespoprodjo et al., 2009).

**Specific Aspects of *P. vivax* Cell Biology, Life Cycle and Epidemiology**

*P. vivax* and *P. falciparum* are of different phylogenetic origin. Most probably two independent host switches from monkey to human occurred, with the origin of *P. vivax* in Asia and the origin of *P. falciparum* in Africa (Cornejo and Escalante, 2006; Escalante et al., 2005; Krief et al., 2010; Pick et al., 2011; Silva et al., 2011). As a consequence *P. vivax* differs in many aspects from *P. falciparum*, from molecular and cell biology to epidemiology.

*P. vivax* infects only reticulocytes, young erythrocytes that account for only a few percent of all red blood cells. As a consequence, parasitemia of *P. vivax* is usually an order of magnitude lower than those of *P. falciparum*, and rarely reaches values above 0.5%. Furthermore, *P. vivax* was long believed to depend fully on the Duffy receptor for invasion into the red blood cell (Miller et al., 1976), while *P. falciparum* invades cells independently of this receptor. In large parts sub-Saharan Africa, over 90% of the population carry the Duffy negative trait, and as a consequence, *P. vivax* is rare in this part of the world. Recent reports showed also Duffy-negative individuals to become eventually infected with *P. vivax* (Mendes et al., 2011; Mercereau-Puijalon and Menard, 2010). This suggests that *P. vivax* is more frequent that
previously appreciated in Africa. Serological evidence suggests that 13% of the population in Congo has been in contact with *P. vivax* (Culleton et al., 2009), and *P. vivax* transmission was also reported from Uganda (Dhorda et al., 2011). Future research will show whether *P. vivax* has been more common in Africa but has been frequently misdiagnosed as other *Plasmodium* species.

A consequence of the clear preference of *P. vivax* for reticulocytes, which are difficult to obtain, so far no continuous in-vitro culture could be established (Udomsangpetch et al., 2008). In spite of intensified efforts to develop culture techniques, for example based on human blood enriched for reticulocytes, no continuous in vitro culture system was reported yet.

A further characteristic of *P. vivax* is the early appearance of gametocytes, the sexual forms of the parasite, which are transmitted by mosquitoes. *P. vivax* gametocytes appear a few days after infection, and before the onset of clinical symptoms (reviewed in (Bousema and Drakeley, 2011)).

*P. falciparum* parasites sequester for part of their 48 hours life cycle, i.e. the infected red blood cells cytoadhere in the deep capillaries of inner organs (Garnham, 1966). This prevents passage of parasites through the spleen where infected red blood cells are filtered out. *P. vivax* does not sequester as completely as *P. falciparum*, but partial sequestration has been discussed also from *P. vivax* (del Portillo et al., 2004). The presence of all stages of the red blood cell cycle of *P. vivax* in the blood is one of the characteristics to identify *P. vivax* by light microscopy. Already in 1966 Garnham noted a slight tendency of late *P. vivax* stages to be absent in peripheral circulation (Garnham, 1966), and recently adhesion of *P. vivax* infected erythrocytes to lung endothelial cells was shown, however at a 10-fold lower level than *P. falciparum* adhesion (Carvalho et al., 2010).

One of the most outstanding hallmarks of *P. vivax* is its ability to relapse weeks or months after the initial transmission by mosquitoes. Semi-dormant liver stages, called hypnozoites, are released into the blood stream, leading to blood stage parasitemia and even clinical episodes in absence of transmission. Duration between infection and relapse seems to differ between regions of different transmission intensity. In regions where transmission only occurs during limited time of the year relapses occur after several months or even after a year (Mason, 1975). But in regions of intense transmission relapses generally occur within the first few months after initial infection. The mechanism to trigger relapses is not known. It has been speculated that relapse could be triggered by new *Anopheles* bites or by an infection with a new *P. vivax* strain, both indicating the potential for transmission to the hypnozoites (Hulden, 2011).
Introduction

It has long been unclear whether a relapsing parasite clone is genetically identical to the last infection. Early studies with limited resolution of molecular markers and low numbers of samples have concluded that relapses correspond to the previous blood stage infection (Craig and Kain, 1996; Kirchgatter and del Portillo, 1998). More recent studies with larger numbers of study participants were conducted in settings were re-infection by mosquitoes could be excluded and showed that relapsing parasites differ from those observed at the last infection in most of the cases (Imwong et al., 2007b; Restrepo et al., 2011). As a consequence relapsing parasites cannot be distinguished from new infection. It is not fully clear whether all relapsing parasites had been present as blood stage infection at an earlier time point. It is possible that a number of blood stage infections classified as relapses are the result of extended latency periods of *P. vivax* liver stages.

Genotyping of malaria parasites has become standard in many drug efficacy field studies. Genotypes of parasites at baseline are compared to those at the day of recurrent parasitemia to distinguish between treatment failure and new infection. Interpretation of results from drugs affecting *P. vivax* blood stage stages are complicated as no differentiation between relapses and new infections is possible. Relapsing parasites also add to the complexity of infection. When they coincide with transmitted parasites, they add to multiplicity of infection, and thus to the potential for sexual recombination between parasite lines in the mosquito.

Another hallmark of *P. vivax* is a global distribution (Figure 2). *P. vivax* transmission occurs in the tropics but also in tempered zones, e.g. Middle East and Asia. In the past even large areas of Europe, up to England and Sweden, were *P. vivax* endemic, as well as tempered zones of Asia up to Siberia (Hay et al., 2004; Mendis et al., 2001) (Figure 2).

![Figure 2: global distribution of malaria from 1900 (prior to malaria control) to 2002 (Hay et al., 2004). *P. vivax* had been transmitted in much of the area that became malaria free during the 20th century in Europe, America and Asia.](image)
**P. vivax Specific Challenges in Malaria Elimination and Eradication**

In October 2007 eradication of malaria was brought back to the agenda at a meeting of the Bill and Melinda Gates foundation (Roberts and Enserink, 2007; Tanner and de Savigny, 2008). Since then, numerous reports, articles and even special editions of entire journals have pointed out aspects to consider to achieve elimination (absence of malaria transmission in a defined geographical area) and eradication (absence of malaria from the planet) (Das and Horton, 2010; Singh, 2009). A detailed research agenda ranging from drugs and vaccines over vector control to health system strengthening and monitoring was presented by the international “malERA” initiative in January 2011 (Alonso et al., 2011).

In a first step elimination shall be achieved in countries of comparable low endemicity and unstable transmission. It is reasonable to expect fast success by this approach of fighting malaria from the boarders of its transmission area, while elimination in countries of very high transmission will only be achieved at a later time point (Feachem et al., 2010). Many countries with lower transmission are found in Latin America and Asia, and in most of these countries *P. vivax* is the dominant cause of malaria.

Thus *P. vivax* will likely play a key role on the path to eradication, and intensified research on species-specific intervention strategies is needed. In October 2008 the genome of *P. vivax* was published (Carlton et al., 2008) as well as the transcriptome of the red blood cell cycle (Bozdech et al., 2008). For transcriptome studies parasites were obtained from patients and maintained in short term culture for 48 hours. In 2010 a second genome was published, adding a lot to the knowledge on frequency and distribution of polymorphisms between different *P. vivax* genomes (Dharia et al., 2010). These studies provide the basis for research towards a better understanding of basic *P. vivax* biology.

Biological differences between *P. falciparum* and *P. vivax* have implications for control strategies. One of the biggest hurdles to stop *P. vivax* transmission is the presence of relapses. So far only one approved drug, Primaquine, eliminates hypnozoites from the liver. However, Primaquine can lead to severe side effects in individuals suffering from glucose-6-phosphate dehydrogenase (G6PD) deficiency (Clyde, 1981), the most common enzyme deficiency worldwide (Frank, 2005). Due to high prevalence of G6PD deficiency in some malaria endemic regions, it has been speculated that it confers some resistance to malaria (Ruwende and Hill, 1998). As a consequence of its side effects primaquine is not recommended for use in G6PD deficient patients (World Health Organization., 2010a).

As primaquine is not used as standard treatment against *P. falciparum* malaria even in regions of *P. vivax* co-endemicity, relapsing *P. vivax* parasitemia is frequently observed after *P. falciparum* treatment (Douglas et al., 2011). Presence of *P. falciparum* gametocytes seems to promote *P. vivax* relapse. A recent study from Cambodia has shown *P. vivax* to be present
in 47% of patients at 28 days after initial treatment for falciparum malaria and presence of gametocytes at baseline (Lin et al., 2011). As *P. vivax* was not detected by PCR diagnosis in most of the samples at day 0, most likely relapsing parasites led to the high numbers of *P. vivax* infections during follow-up (Lin et al., 2011).

A further complication in *P. vivax* control is the development of gametocytes before the onset of clinical disease. As a consequence even non-immune individuals might transmit *P. vivax* before they even know that they are infected and before drugs to eliminate blood stage parasites can be applied. On the other hand, as a proportion of gametocytes is transmitted before drug pressure selects for potentially resistant strains, it is expected that *P. vivax* drug resistance develops slower compared to *P. falciparum* (Mendis et al., 2001).

**Malaria in Papua New Guinea**

Papua New Guinea comprises of a number of very diverse ecosystems, from tropical lowlands and islands to the tempered zones of the highlands and mountains reaching above 4500 m. Also biodiversity and cultural diversity is extraordinary high in PNG, with over 800 languages spoken in an area of about 460'000 km² (Attenborough and Alpers, 1992). In concordance with the geographic diversity malaria prevalence ranges from low to very high levels, with temperature as main determinant of transmission intensity (Cattani, 1992; Mueller et al., 2003). Four human malaria species are present in PNG, *P. falciparum, P. vivax, P. malariae* and *P. ovale* (Mehlotra et al., 2002; Mehlotra et al., 2000; Mueller et al., 2002; Peters, 1960), with mixed species infections frequently observed (Burkot et al., 1987; Mueller et al., 2009b). It is believed that *P. vivax* was the dominating parasite before the start of vector control programs in the middle of the 20th century (Cattani, 1992; Hairston et al., 1947; Muller et al., 2003). Pesticide spraying and mass drug administration have meanwhile changed the pattern, today *P. falciparum* is predominant in most part of the country except some regions in the lowlands (reviewed in (Mueller et al., 2003)).

It has even been speculated that population distribution in PNG is a direct consequence of malaria prevalence (Riley, 1983). The PNG population of approximately 6.7 million people resides mainly in the lowlands, which are highly endemic for malaria, and in the highlands at an altitude of around 1500 m, where little malaria transmission prevails. Intermediate altitudes are less populated. The population distribution hypothesis suggested that this pattern is a consequence of malaria transmission. In the lowlands, high malaria prevalence leads on the one hand to high infant mortality, but on the other hand to high human fertility rates and fast immunization of individuals that survive early childhood. Malaria prevalence in highland valleys is too low to substantially influence population growth. In intermediate zones malaria is frequent enough to lead to repeated disease, but immunity does not develop rapidly.
In the tropical lowlands malaria prevalence is usually high with perennial transmission (Cattani et al., 1986; Genton et al., 1995), leading to high morbidity and mortality. In 1960, an infant mortality rate of 571/1000 was reported from Maprik, a region of high malaria endemicity (Peters, 1960), and today malaria is the most frequent outpatient diagnosis and the second most frequent cause of hospital admission in PNG (Mueller et al., 2003). In the highlands at altitudes of 1300 to 1600 m prevalence is lower and transmission usually peaks at the end of the rainy season in April to July (Mueller et al., 2003). During local outbreaks prevalence can reach high levels of above 60%, and very high morbidity and mortality is observed (Mueller et al., 2005).

In 2009 a total of 1.36 million malaria cases were reported from PNG, although most of them were not confirmed parasitologically (World Health Organization., 2010b). Malaria prevalence in PNG reaches levels similar to those in sub-Saharan Africa, and P. vivax in PNG reaches higher levels than anywhere else in the world.

Malariologist Robert Koch has made fundamental discoveries in PNG at the turn of the 20th century (Stanisic et al., 2010). Even today several aspects make malaria research in PNG highly relevant, many of which highlight the specific role of P. vivax: (i) P. vivax prevalence is higher than anywhere else observed in the world, (ii) P. vivax is a mayor public health burden (iii) different levels of transmission and different species composition can be studied in close proximity, and (iv) frequent mixed species infections allow studies on how parasites influence each other (multiple-species multiple-clones interactions).

**Context of this PhD Project**

This thesis is part of a larger study conducted by an international group of collaborators under the leadership of PNG IMR. The aim of this study is to obtain a profound understanding of the relationship between prevalence of different malaria parasites and development of disease. For this purpose a cohort study was conducted over 16 months in PNG. Blood samples from a large number of children were collected in a site of high prevalence of P. falciparum and P. vivax, and of lower prevalence of P. malariae and P. ovale. Samples were collected by active follow-up and by passive case detection at the local health centres. In parallel parameters such as health status and bed net use was recorded every second week from each child.

**Study site, sampling, parasite detection and incidence of disease**

The cohort study was conducted in several villages in the Ilahita area, Maprik district, Papua New Guinea from March 2006 to July 2007. Maprik district is located at the north coast in the
tropical lowlands of PNG, the climate is tropical with high temperatures year round. The study consisted of an active follow up with blood samples taken every two months from every child, and of passive case detection during the whole period. Two hundred sixty-four children aged 0.5 to 3 years at age of enrolment were included in the study, after written parental consent was obtained from the parents or legal guardians. Details of the study site, enrolment and sample collection were published previously (Lin et al., 2010).

In brief, the active follow up consisted of 2-monthly cross sectional surveys repeated 9 times and was carried out as follows: At baseline and in the last round of the study 5 ml venous blood was taken. In between, children were visited every eight weeks and 250 µl finger prick blood was taken regardless of presence of symptoms for malaria. With the exception of the first and the last round, two blood samples were taken 24 hours apart from the same child. In addition children were visited every two weeks and checked for signs of malaria using a questionnaire. In case of presence of symptoms (fever above 37.5°C) or if fever during the 48 hours prior to the visit was reported, a 250 µl finger prick blood sample was taken. From all febrile children two blood slides were examined and rapid diagnostic test (RDT) was made. Children with confirmed parasitemia were treated with Coartem. In few cases other antimalarials (amodiaquine, chloroquine, sulphadoxine, artesunate, quinine and primaquine), received from health posts outside the study area, were given to the children; treatment details were recorded.

In parallel to the periodical active follow-up, passive case detection was maintained at the Ilahita health centre and Sikamo aid post. Regardless of the cause of presentation to the health centre or the clinic, all study participants attending these health facilities were examined for malaria. A blood sample was taken and children were treated in case of positive blood slides or RDT.

Four *Plasmodium* spp. were found in the study population. Species detection was done by light microscopy (LM) (at least two independent reads by expert microscopists), as well as by ligase detection reaction-fluorescent microsphere assay (LDR-FMA), a molecular method to detect human malaria parasites (McNamara et al., 2006). Prevalence at baseline was as follows: *P. vivax* 53.0% by LDR-FMA and 44.3% by LM, *P. falciparum* 49.6% by LDR-FMA and 32.6% by LM, *P. malariae* 9.9% by LDR-FMA and 4.2% by LM, *P. ovale* 2.7% by LDR-FMA and 0.0% by LM (Lin et al., 2010). Incidence of *P. falciparum* and *P. vivax* malaria was similar, however age trends were contrasting: *P. falciparum* incidence peaked at around 3 years of age with approx. 3.5 clinical episodes per child per year, *P. vivax* incidence dropped throughout the age group studied from approx. 3.5 clinical episodes per child per year in the youngest age group of 1 year old children to 1.5 episodes in the oldest children (Lin et al., 2010).
Genotyping

In recent years genotyping based on size polymorphic markers has become the method of choice for many epidemiological studies, where differentiation between individual clones is needed (Collins et al., 2006; Falk et al., 2006; Sutherland, 2008; World Health Organisation, 2008). In addition it is well known that Plasmodium densities are often below the detection limit of light microscopy, and PCR – either to distinguish species or for genotyping – reveals substantially higher prevalence than microscopy alone (reviewed in (Okell et al., 2009)). Also in drug trials, where persisting infections (i.e. drug failure) must be distinguished from new infections or relapses, genotyping polymorphic markers has become a standard method (Gatton and Cheng, 2008; Juliano et al., 2010). This PCR correction of drug trial outcomes often shows that a substantial proportion of patients with recurrent parasitemia at day 28 have acquired new infections, hence the drug applied was more effective than estimated based on microscopy results only (Hwang et al., 2011; Karunajeewa et al., 2008; Ratcliff et al., 2007; Yavo et al., 2011).

The prime criteria for genotyping markers for this study were high diversity and robust PCR amplification. A small number of markers (ideally a single marker) with a large number of alleles is preferred to the combination of several markers with limited diversity. Although guidelines define the procedure for drug efficacy genotyping based on three markers (World Health Organisation, 2008), the laboratory workload to analyze a single markers is lower. In epidemiological studies, genotyping is used to differentiate between clones and to track them over time. High diversity of markers is a special advantage in situations of multiple clone infections, when construction of haplotypes by compiling genotyping data from different loci is not possible or error prone.

The \textit{P. falciparum} parasites in this cohort have been genotyped using the \textit{merozoite surface protein 2} (\textit{msp2}) as marker, results have been published (Schoepflin et al., 2009). A number of size polymorphic molecular markers have been identified in \textit{P. vivax} (Imwong et al., 2007a; Imwong et al., 2005; Imwong et al., 2007b; Karunaweera, 2006). In a preliminary study we have assessed the diversity of nine size polymorphic molecular markers in \textit{P. vivax} samples from PNG (Koepfli et al., 2008). Based in this study, we have selected the most diverse coding sequence, the F3 region of the \textit{merozoite surface protein 1} gene, and the most diverse microsatellite, MS16, for genotyping. There is no need that markers for epidemiological studies are neutral, as long as high diversity is maintained, markers under selection (such as genes coding for surface proteins of the parasite) can be used for genotyping. Care is needed with markers that are linked to drug resistance phenotypes. Instead of the overall diversity of the parasite populations they will more likely represent the history of drug use.
Introduction

Goal and Objectives of the Thesis

The main goal of this thesis was to describe the genetic diversity of the population of *P. vivax* parasites in the cohort study in PNG and to use the genotyping data to investigate the infection dynamics of *P. vivax*. To make further statements on population genetic parameters the parasite population from the cohort study was compared to *P. vivax* populations from other locations in PNG. The specific research objectives and the rationale were:

I. **Determine parameters of *P. vivax* epidemiology in a highly endemic region**

a) Over 5000 blood samples were collected in the course of the cohort study, and over 3000 of them were diagnosed *P. vivax* positive by light microscopy or LDR-FMA. By genotyping these samples we aimed at assessing the diversity (number and distribution of alleles per locus, number of haplotypes) and the number of concurrent infections per patient (multiplicity of infection, MOI). Incidence of falciparum and vivax malaria show clearly contrasting age trends in our cohort (Lin et al., 2010). Hence a special focus was put on the investigation of age dependency of these molecular-epidemiological parameters.

b) For the first time for *P. vivax* a new parameter was to be established: the “force of infection” (molFOI), i.e. the number of new *P. vivax* infections per child per year. FOI is a central parameter in the study of any infectious disease. Disease is a consequence of infection, and many interventions to control malaria, such as administration of preventive drugs or bed nets, aim at reducing the number of new infections. In areas of lower endemicity estimates of FOI can be made based on parasite positivity, without genotyping clones. In situations of high endemicity, as in our cohort, infections usually overlap and genotyping is needed to obtain reliable estimates of FOI.

II. **Estimate the share of parasites that are not detected in standard epidemiological surveys**

Every field survey of parasitemia is confronted with the fact that a number of parasites are not detected, because their density is below the detection limit at the time point of sampling. This applies for microscopical detection of parasites as well as for molecular methods. Due to the combined effect of sequestration and fast asexual replication in the red blood cell, the number of parasites detected in the peripheral blood can change dramatically from one day to the other. The acquisition of a new infection or the loss of an infection (in absence of drug treatment) is neglectable within a 24 h period.
We have compared samples collected 24 hours apart from the same individual, with both respect to species detection of *P. falciparum* and *P. vivax* by light microscopy and PCR, and with respect to composition of individual clones as determined by molecular genotyping. This allows estimations of the detectability of parasites, and as both *P. vivax* and *P. falciparum* are highly prevalent in the study population, we can make direct comparisons between species. The parameter “clone detectability” is of great importance in malaria epidemiology where multiple clone co-infections are the rule.

**III. Compare the genetic composition of *P. vivax* populations in PNG**

Differences between parasite populations can reflect historic changes in epidemiology (such as recent expansions of populations), influence emergence and spread of drug resistance, and inform intervention strategies. In the third part of my thesis I aimed at investigating the population genetic differences between a subset of the samples of the study in Ilahita and three other sites in PNG. Based on multi-locus genotyping, parameters such as the extent of linkage disequilibrium, differences in allelic frequencies, levels of gene flow and population size were assessed.

In the course of this thesis a further practical application of our *P. vivax* genotyping led to a collaboration on PCR correction of results from a drug trial. Results are presented in the appendix (manuscript published in “Antimicrobial Agents and Chemotherapy”). Four different drug regimens were tested in 195 children in PNG, and in a substantial number of cases recurrent parasitemia on day 28 or day 42 was observed (Karunajeewa et al., 2008). Comparison of genotyping results of day 0 samples with samples taken on the day of recurrent parasitemia allows distinction from true treatment failure (same genotype in both samples) and new infection or relapse (different genotype).

Beside these scientific goals technology transfer to PNG and training of local researchers in genotyping has been an additional aspect of this thesis. Training was provided in workshops at the Papua New Guinea Institute of Medical research in Goroka and in Madang, as well as for a visiting student from PNG at the Swiss Tropical and Public Health Institute.
Introduction

References


Introduction


Introduction


How Much Remains Undetected? Probability of Molecular Detection of Human Plasmodia in the Field

Cristian Koepfli1,2,*, Sonja Schoepflin1,2,*, Michael Bretscher1,2, Enmoore Lin3, Benson Kiniboro3, Peter A. Zimmermann4, Peter Siba3, Thomas A. Smith1,2, Ivo Mueller3, Ingrid Felger1,2,*

1 Swiss Tropical and Public Health Institute, Basel, Switzerland, 2 University of Basel, Basel, Switzerland, 3 Papua New Guinea Institute of Medical Research, Goroka, Eastern Highlands Province, Papua New Guinea, 4 Center for Global Health and Diseases, Case Western Reserve University, Cleveland, Ohio, United States of America

Abstract

Background: In malaria endemic areas, most people are simultaneously infected with different parasite clones. Detection of individual clones is hampered when their densities fluctuate around the detection limit and, in case of P. falciparum, by sequestration during part of their life cycle. This has important implications for measures of levels of infection or for the outcome of clinical trials. This study aimed at measuring the detectability of individual P. falciparum and P. vivax parasite clones in consecutive samples of the same patient and at investigating the impact of sampling strategies on basic epidemiological measures such as multiplicity of infection (MOI).

Methods: Samples were obtained in a repeated cross-sectional field survey in 1 to 4.5 years old children from Papua New Guinea, who were followed up in 2-monthly intervals over 16 months. At each follow-up visit, two consecutive blood samples were collected from each child at intervals of 24 hours. Samples were genotyped for the polymorphic markers msp2 for P. falciparum and msp1F3 and MS16 for P. vivax. Observed prevalence and mean MOI estimated from single samples per host were compared to combined data from sampling twice within 24 h.

Findings and Conclusion: Estimated detectability was high in our data set (0.79 [95% CI 0.76–0.82] for P. falciparum and, depending on the marker, 0.61 [0.58–0.63] or 0.73 [0.71–0.75] for P. vivax). When genotyping data from sequential samples, collected 24 hours apart, were combined, the increase in measured prevalence was moderate, 6 to 9% of all infections were missed on a single day. The effect on observed MOI was more pronounced, 18 to 31% of all individual clones were not detected in a single bleed. Repeated sampling revealed little difference between detectability of P. falciparum and P. vivax.

Introduction

Detection of malaria parasites is essential for many malarialogical investigations. For instance, the detailed maps of global malaria risk created by the Malaria Atlas Project (MAP) use parasite prevalence as the main indicator for transmission [1,2,3,4]. Prevalence is also the key malarialogical measure assessed in Malaria Indicator Surveys (MIS) [5,6]. Within populations, the identification of individual parasite clones by PCR based genotyping techniques has substantially increased knowledge of the infection dynamics of malaria by providing estimates of multiplicity of infection, incidence and clearance rates [7,8,9,10]. In addition, it allows classification of drug failures into epidemiological measures such as multiplicity of infection (MOI).

However, all DNA based techniques for detection of malaria parasites in a blood sample are imperfect. Parasites can remain undetected because their densities fall below the detection limit of the diagnostic technique. PCR based methods generally have a better detection limit compared to microscopy [11], but both methods likely miss a proportion of clones.

Several biological factors add to the probability of low numbers of parasite in the blood stream and subsequently of missing clones. In P. Falciparum infections late blood stage forms cytoadhere to the endothelial wall of blood vessels. This sequestration of P. Falciparum in deep organs lasts for 24–28 hrs of the 48 hrs blood stage cycle. During this period parasites are absent from the peripheral circulation and escape detection. A parasite clone defined by a shared genotype and by common ancestry is completely absent only if all individual parasites belonging to this brood are tightly synchronized. By rare chance a clone might be superinfected by another clone sharing the same genotype. The probability of this occurring is determined by the resolution of the typing scheme and the mean multiplicity of infection in the study area [12].

In blood stage schizogyon a single schizont divides into numerous merozoites [13] within a few hours. This has particular implications for P. Vivax, because these parasites usually divide...
synchronously. As a result, the number of DNA copies of a given clone may drastically increase from one day to another. In addition, parasite densities of *P. Vivax* are usually lower than those of *P. Falciparum* [14,15]. Taken these factors together, imperfect detection of *P. Vivax* clones is likely, however, only for *P. Falciparum* this has been suspected because of sequestration.

Besides these issues of parasite biology, also methodological constraints may add to the risk of missing clones. These causes equally apply for both parasite species. In multi-clonal infections minority variants constituting a small proportion of the total parasite load, might be missed by PCR-based detection, because in competition for primers or other constituents of the reaction mix, they are outcompeted by the more abundant clones.”

Although sequestration of *P. Falciparum* has been reported several decades ago [16,17], low parasite densities in *P. Vivax* and the sensitivity threshold of both light microscopy and PCR-based diagnostic methods have been well studied [18,19,20], the consequences of imperfect detectability for estimates of prevalence and other epidemiological parameters have often not been addressed adequately.

The daily dynamics of *P. Falciparum* clones has been investigated previously in a longitudinal study with daily follow up bleeds collected over a period of 14 days from 20 children from Tanzania [21]. A complex dynamics of *P. Falciparum* clones was observed. The composition of infecting clones found in a single host was found to be unstable over time or even changing from one day to another [21]. In a drug efficacy trial in Tanzania children were sampled on two consecutive days and the results were compared to the standard protocol with single bleeds. An increased number of multiple clone infections and additional recrudesence cases were detected [22]. Several studies have described statistical models to estimate infection dynamics of *P. Falciparum* allowing for this imperfect detection [23,24,25,26]. The model by Sama et al. [23] was applied for a longitudinal field study in Ghana, where blood samples were collected once at each of 6 cross sectional surveys conducted in 2-monthly intervals. Statistical models indicated that at any time of sampling on average only 47% of all parasite clones present in a host were detected by genotyping [23].

Back in 1966 Garnham reported a slight tendency of late *P. Falciparum* stages to retreat from the peripheral circulation [27], but sequestration is generally thought to be absent in *P. Vivax*. This view was questioned only recently when cytoadherence of *P. Vivax* was shown in vitro [28]. Few studies have addressed the infection dynamics of individual *P. Vivax* clones [7]. Daily fluctuations in detectability of *P. Vivax* clones have not yet been investigated.

To obtain a more precise picture of the parasite population present in a host, blood sampling should be repeated within short time intervals. However, collection of consecutive blood samples from a study participant will translate into considerably increased efforts in the field and laboratory, added costs, and additional discomfort for participants. Limited knowledge has been gathered so far on the effect of such a sampling scheme on epidemiological measures, such as prevalence or multiplicity of infection.

Here we present data from a large set of paired samples that were collected 24 hours apart. Within this time interval re-infection with a new parasite clone is very unlikely and may be ignored. Samples were derived from a cohort study performed in Papua New Guinean children 1 to 4.5 years of age living in an area highly endemic for both *P. Falciparum* and *P. Vivax*. Parasites of both species were genotyped to calculate the detectability of infection and to investigate the benefit of collecting 24 h bleeds on basic and molecular measures of epidemiology such as prevalence and multiplicity of infection. Figure 1 gives a schematic overview on the effect of combining results obtained on two consecutive days.

Methods

Ethics Statement

Scientific approval and ethical clearance for the study was obtained from the Medical Research and Advisory Committee (MRAC) of the Ministry of Health in PNG and from the Ethikkommission beider Basel in Switzerland. Informed written consent was sought from all parents or guardians prior to recruitment of each child.

Field survey and patients

This study was conducted in a rural area near Maprik, East Sepik Province, Papua New Guinea. A detailed description of the study was given previously [15]. Briefly, 263 study participants were enrolled at an age of 1 to 3 years starting in March 2006 and regular follow-up visits were conducted over a period of 16 months until July 2007. At seven time points separated by 8-weekly intervals two consecutive 250 μl finger prick blood samples were

![Schematic overview of possible outcomes of 24 h bleeds. A sample is positive as soon as a parasite is detected on either day. Different colors of PCR results indicate different clones detected. The combined multiplicity of infection includes all clones detected in two corresponding bleeds.](image-url)
collected at intervals of 24 hours (in the following termed: 24 h bleed) from each study participant. Two blood slides were made and a rapid diagnostic test was performed upon presentation of malaria symptoms. Antimalarial treatment with Coartem® (Novartis, Switzerland) was administered upon a positive test plus iron and folate supplementation if haemoglobin levels were <7.5 g/dL. Children that were treated on day 1 of the 24 h bleed were excluded from the analysis (n = 97). Over the 16 months follow-up period, study participants adhered to 93 to 94% of all study visits.

### Laboratory procedures

Diagnosis of the different malaria species was performed in parallel by two methods, light microscopy (LM) and post-PCR Ligase Detection Reaction (LDR) [29]. Only samples positive by light microscopy and/or LDR were genotyped for our markers.

All finger prick blood samples were separated into plasma and cells. DNA was extracted from cell pellets using QiAamp® 96 DNA Blood Kit (Qiagen, Australia) according to the manufacturer’s instructions. All samples were genotyped for the polymorphic marker gene merozoite surface protein 2 (msp2) by use of capillary electrophoresis (PCR-CE) for fragment sizing as previously described by Schoepflin et al. [12] with some minor changes and adaptations of PCR conditions for highly purified DNA as described by Falk et al. [18].

### P. vivax genotyping

Genotyping was performed as described previously [30] with the following modifications: A multiplex primary PCR was done with the primers for the 2 markers mspF3/MS16 followed by individual nested PCRs for mspF3 and MS16. Primary PCR was performed in a volume of 20 μL containing 1 μL template DNA, 0.25 μM of each primer (Eurofins MWG Operon), 0.3 mM dNTPs (Solis BioDyne), 2 mM MgCl2, 2 μL Buffer B (Solis BioDyne) and 5 μL Taq FIREPol (Solis BioDyne). 1 μL of primary PCR product was used as template for nested PCR performed in a volume of 20 μL containing 0.25 μM of each primer (Applied Biosystems), 0.2 mM dNTPs (Solis BioDyne), 2 mM MgCl2, 2 μL Buffer B (Solis BioDyne) and 1.5 μL Taq FIREPol (Solis BioDyne). The forward primers for the nested PCR were labelled with fluorescent dyes: 6-FAM for mspF3, NED for MS16. Cycling conditions were as follows: Initial denaturation 95°C for 1 minute, then 30 cycles (primary PCR) or 25 cycles (nested PCR) with 15 seconds denaturation at 95°C, 30 seconds annealing at 59°C and 30 seconds elongation at 72°C plus a final elongation of 5 minutes at 72°C.

All samples negative after the first round of PCR amplification were repeated once. Repeats and all microscopy negative samples (due to an expected lower parasitaemia) were done under similar conditions with the exception that 2 μL DNA solution was used as template for the primary PCR. Capillary electrophoresis was done as described earlier [30]. As the mspF3 nested PCR in general led to more amplification product compared to the MS16 PCR, twice as much MS16 PCR product (2.5 μl of 1:10 dilution of PCR product) was analysed by capillary electrophoresis. All alleles were grouped into bins of 3 base pairs according to the possible size of insertions and deletions in coding regions and the repeat size of the microsatellite.

### Data analysis

#### Analysis of 24 h interval bleeds

Sample pairs collected 24 hours apart from the same patient were compared. Sample pairs were excluded from the analysis if antimalarial treatment was given on the first day of paired sampling. Individual genotypes were classified by positivity on each of two consecutive days, leading to two categories for each genotype: one day positive (genotype observed on either day of paired sampling, n1) and both days positive (n2). An estimate \( \hat{q} \) of the detectability \( q \) was calculated as suggested by Bretscher et al. [31]:

\[
\hat{q} = \frac{2n_2}{n_1 + 2n_2}
\]

An approximate confidence interval was calculated as follows:

\[
\text{CI} = \frac{2 \sqrt{n_1 n_2 (n_1 + n_2)}}{(n_1 + 2n_2)^2}
\]

Comparison of detectability between day 1 and day 2 was done by McNemar’s exact test for paired data. Detectability was calculated for different age groups of patients, for samples with different MOI (combined over 24 h) and for different parasite densities at day 1. Comparison of different groups was done by nonparametric test for trend across ordered groups. Statistical analysis was performed using STATA® 9.1 statistical analysis software (Stata Corporation, College Station, TX).

### Results

#### Parasite detection by light microscopy and PCR-capillary electrophoresis

For the analysis of paired samples collected in 24 h interval, 1019 pairs, equal to 2038 individual blood samples, were eligible. LM diagnosed *P. falciparum* in 398 samples, in 362 samples (91%) this was confirmed by PCR-CE. Additional 210 microscopy-negative samples were *P. falciparum* positive by PCR-CE, leading to a total of 572 samples for which *P. falciparum* genotyping results were obtained. *P. vivax* was diagnosed in 1001 samples by LM, 987 (98.6%) of them were also positive by PCR-CE. Additional 187 samples were positive by PCR-CE, leading to a total of 1174 blood samples with *P. vivax* genotyping results. Thus, sensitivity of LM was lower than that of PCR based diagnosis. Detection of *P. vivax* by MS16 PCR (1161 positive samples) was more sensitive than by mspF3 PCR (1120 positive samples). This difference was significant (McNemar’s test: \( \chi^2 = 9.48, p = 0.002 \)).

### Effect of repeated sampling on observed parasite prevalence

By light microscopy 242 and 563 out of 1019 pairs were positive for *P. falciparum* and *P. vivax*, respectively, at least on either day (Table 1). Observed prevalence did not differ between individual days but it increased when both days were combined. *P. falciparum* parasites were detected in 19.1% of samples at day 1 and 20.0% at day 2 (McNemar’s test: \( \chi^2 = 2.21, p = 0.14 \)). Overall, parasites were detected in 23.7% of sample pairs. Observed prevalence of *P. vivax* was 48.6% on day 1 and 49.7% on day 2 (McNemar’s test: \( \chi^2 = 9.48, p = 0.002 \)).

The 1019 sample pairs were genotyped using msp2 as marker for *P. falciparum* and both mspF3 and MS16 as markers for *P. vivax*. After PCR the number of pairs positive at least on one of both consecutive days increased to 311 for *P. falciparum* and 616 for *P. vivax*. Table 1 summarizes the detection of parasites by PCR on day 1 and 2 of paired samples. Again the observed prevalence of *P. falciparum* as well as *P. vivax* infection did not differ significantly between both days (*P. falciparum*: 27.8% on day 1 vs. 28.5% on day 2, McNemar’s test: \( \chi^2 = 1.0, p = 0.3 \); *P. vivax*: 58.5% on day 1 vs. 60.5% on day 2, McNemar’s test: \( \chi^2 = 2.21, p = 0.14 \)). When typing results from both days were combined, the observed...
prevalence increased only marginally: from 28% to 30.6% for *P. falciparum* and from 59.3% to 64.0% for *P. vivax*.

Effect of repeated sampling on detection of individual clones

When assessing the persistence of individual clones on the consecutive days of sampling, considerable turn-over in allele composition was observed (Table 2). Examples from 2 patients are given in Figure 2. For *P. falciparum* msp2 64.7% of the infecting clones were observed on both days. For *P. vivax*, 57.0% of the mspF3 alleles and 43.5% of the MS16 alleles were observed on both consecutive days. These two markers differed slightly in their genetic diversity, with MS16 (virtual heterozygosity *H*~v~ = 0.90) showing greater polymorphism than mspF3 (*H*~v~ = 0.88).

Combining the genotyping results from 24 h bleeds made it possible to assess the effect of repeated sampling on other molecular epidemiological parameters, i.e. prevalence and MOI. In *P. falciparum* combining results from both days lead to a small increase in observed mean MOI to 1.68 compared to an observed mean MOI of 1.52 at day 1. In *P. vivax* the observed mean MOI based on mspF3 increased from 2.21 detected on a single day to 2.60 detected on two days. For the more diverse marker MS16, the observed mean MOI increased from 2.43 to 3.10. When both markers were considered to establish mean MOI, i.e. for each pair the highest number of clones observed was counted, observed MOI increased from 2.78 based on a single day bleed to 3.37 based on results of both consecutive days (Table 2).

Detectability, *q*, was in the same range for both *Plasmodium* species: for *P. falciparum* clones detectability was 0.79 and for *P. vivax* detectability was 0.73 based on mspF3 marker and 0.61 based on microsatellite MS16.

The relationship of clone detectability with age, MOI and parasite density was determined. Figure 3 depicts detectability by age group (0–2 years, 2–3 years, >3 years) for light microscopy and PCR-CE detection. No major difference was observed except for *P. vivax* detection by microsatellite marker MS16, which revealed a slightly lower detectability in children above 3 years compared to younger children (no overlap of 95% CI).

The influence of MOI on detectability was investigated (Figure 4). All markers showed a significant decrease in detectability with increasing MOI (nonparametric test for trend across groups, *P. falciparum* msp2: *z* = −3.46, *P* < 0.001; *P. vivax* mspF3: *z* = −3.72, *P* < 0.001; *P. vivax* MS16: *z* = −4.82, *P* < 0.001).

Detectability increased with parasite density for both parasite species (Figure 5). *P. falciparum* msp2 detectability increased with increasing density. *P. vivax* detectability increased with density until 5000 to 10’000 parasites/μl and decreased thereafter. In samples negative by LM detectability was very low, 0.69 for *P. falciparum* msp2, 0.48 for *P. vivax* mspF3 and 0.36 for *P. vivax* MS16.

Discussion

Genotyping of malaria parasites has become an integral part of many malariological field studies. Since more than a decade genotyping has been considered imperative for clinical trials of antimalarials performed in endemic countries. The hallmark of PCR correction of clinical trial outcomes is discrimination of new infections versus recrudescences. Quantification of clone detectability at any time point of blood sampling contributes relevant information on the reliability of PCR corrections. Furthermore, the number of newly acquired clones per time interval might be a suitable outcome measurement of antimalarial interventions; the parameter “clone detectability” might also correct this estimate.

Imperfect detectability for *P. falciparum* has been common knowledge, but to date this effect has been quantified only in few molecular epidemiological studies. The detectability of *P. vivax* clones has been largely ignored due to the reported absence of sequestration. However, the generally lower parasite densities in *P. vivax* compared to *P. falciparum* has potential to contribute to compromised detectability in a major way. The precise estimation of the detection probabilities of both species, undertaken in the

---

**Table 1. Effect of repeated sampling on *P. falciparum* and *P. vivax* prevalence by light microscopy or PCR.**

<table>
<thead>
<tr>
<th></th>
<th><em>P. falciparum</em></th>
<th><em>P. falciparum</em></th>
<th><em>P. vivax</em></th>
<th><em>P. vivax</em></th>
<th><em>P. vivax</em></th>
<th><em>P. vivax</em></th>
<th>PCR 2 markers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Microscopy</td>
<td>msp2 PCR</td>
<td>Microscopy</td>
<td>mspF3 PCR</td>
<td>MS16 PCR</td>
<td>MS16 PCR</td>
<td></td>
</tr>
<tr>
<td>No. of positive samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st day pos., 2nd day neg.</td>
<td>39 (16.1%)</td>
<td>21 (6.8%)</td>
<td>57 (10.1%)</td>
<td>39 (6.3%)</td>
<td>49 (7.8%)</td>
<td>40 (6.1%)</td>
<td></td>
</tr>
<tr>
<td>1st day neg., 2nd day pos.</td>
<td>47 (19.4%)</td>
<td>28 (9.0%)</td>
<td>68 (12.1%)</td>
<td>56 (9.1%)</td>
<td>58 (9.1%)</td>
<td>55 (8.4%)</td>
<td></td>
</tr>
<tr>
<td>Both days positive</td>
<td>156 (64.5%)</td>
<td>262 (84.2%)</td>
<td>438 (77.8%)</td>
<td>521 (84.6%)</td>
<td>527 (83.1%)</td>
<td>557 (85.5%)</td>
<td></td>
</tr>
<tr>
<td>Total no. of pairs positive at least on one day</td>
<td>242</td>
<td>311</td>
<td>563</td>
<td>616</td>
<td>634</td>
<td>652</td>
<td></td>
</tr>
<tr>
<td>Prevalence (parasite positivity)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevalence on 1st day (n = 1019 samples)</td>
<td>0.19 (95/1019)</td>
<td>0.29 (283/1019)</td>
<td>0.49 (495/1019)</td>
<td>0.55 (560/1019)</td>
<td>0.57 (576/1019)</td>
<td>0.59 (597/1019)</td>
<td></td>
</tr>
<tr>
<td>Prevalence on 2nd day (n = 1019 samples)</td>
<td>0.20 (203/1019)</td>
<td>0.28 (290/1019)</td>
<td>0.50 (506/1019)</td>
<td>0.57 (577/1019)</td>
<td>0.57 (585/1019)</td>
<td>0.60 (597/1019)</td>
<td></td>
</tr>
<tr>
<td>Prevalence 2 days combined (n = 1019 pairs)</td>
<td>0.24 (242/1019)</td>
<td>0.31 (311/1019)</td>
<td>0.55 (563/1019)</td>
<td>0.60 (616/1019)</td>
<td>0.62 (634/1019)</td>
<td>0.64 (652/1019)</td>
<td></td>
</tr>
<tr>
<td>Detectability q (95% CI)</td>
<td>0.78 [0.74–0.83]</td>
<td>0.91 [0.89–0.94]</td>
<td>0.88 [0.86–0.89]</td>
<td>0.92 [0.90–0.93]</td>
<td>0.91 [0.89–0.93]</td>
<td>0.92 [0.91–0.94]</td>
<td></td>
</tr>
</tbody>
</table>
same field study and under perfectly matching experimental conditions, allows assessing the combined effects of parasite sequestration, synchronicity and low parasitaemia and differences among *Plasmodium* species. It is clear that the individual factors, contributing jointly to detectability, cannot be determined by our genotyping approach.

Our analysis of samples collected 24 hours apart revealed limited day-to-day fluctuations in the detection of *P. falciparum* and *P. vivax* infection. This indicates that short-term sampling has only a small impact on prevalence estimates – regardless of whether infections are detected by light microscopy or PCR. For both species the observed prevalence by PCR increased, when 2 days were combined, by less than 10%. A more pronounced difference in prevalence based on one versus two days of sampling was only observed for microscopic detection of *P. falciparum*, where prevalence increased by 24%. It is unclear in how far the effect...
of parasite synchronization and sequestration add to this discrepancy between PCR and microscopy in *P. falciparum* detectability. These findings suggest that for *Plasmodium* species conducting repeated sampling within 24 h does not substantially increase the observed prevalence.

Detection of individual clones was very high in *P. falciparum* (~q = 0.79). Accordingly, combining genotypes from both days resulted in a small increase in observed mean MOI rising from 1.52 based on one day to 1.68 for both days. Children <5 years have not yet developed a strong immunity to *P. falciparum* [15] and therefore carry high parasite densities (mean parasite density: 2558 parasites/μl), which leads to a better chance to detect by PCR most of the parasite clones present. Recently in a similar study in Ghana blood samples from individuals up to 20 years were repeatedly collected in intervals of 1, 4, 5 and 7 days [31]. Detectability was calculated using the same approach as in our study. Clone detectability was around 0.6 in sample pairs collected 24 hours apart [31].

In contrast to *P. falciparum*, sequestration of late stage parasites has not been reported from *P. vivax*. Despite this biological difference, the detectability of individual parasite clones was lower in *P. vivax* than in *P. falciparum*. A larger number of *P. vivax* clones was only detected on either day for both *P. vivax* markers analyzed. In *P. vivax* parasite densities are generally much lower than in *P. falciparum*, in our study mean *P. vivax* density was 498 parasites/μl compared to the 5-fold higher *P. falciparum* density. Our results suggest that the overall low parasitaemia combined with synchronized replication, as generally seen in *P. vivax*, has a larger impact on detectability than sequestration plus sporadic low parasitaemia, as observed in *P. falciparum*.

The number of concurrent *P. falciparum* or *P. vivax* infections had a pronounced effect on detectability: Increasing MOI lead to decreasing detectability. It has been suggested that in multiple-clone infections, clones representing a minority of the total parasite population in a host might escape detection by PCR [32]. In experimental mixtures of DNA from two different *P. falciparum* clones up to ratios of 1:100, both genotypes were detected by PCR-CE [33], as well as in mixtures of DNA from two different *P. falciparum* clones in a ratio of 1:5 [34]. However, our observation that detectability decreased with increasing MOI, suggests impaired amplification of minority clones. Hence, the higher...
mean MOI in *P. vivax* (MOI = 3.37) compared to *P. falciparum* (MOI = 1.68) could be seen as a further reason for the slightly lower detectability in *P. vivax*.

Detectability of the two *P. vivax* markers *msp*1F3 (\(\bar{q} = 0.73\)) and MS16 (\(\bar{q} = 0.61\)) differed. Overall, more samples were positive for *P. vivax* and more clones were detected with the MS16 PCR. This suggests a higher detection threshold of the *msp*1F3 PCR in combination with a lower influence of fluctuations in parasite density on detectability. In addition, a difference in the genetic diversity of these markers could add to the discrepant values: two independent clones are expected to share more often the same *msp*1F3 allele.

Our results highlight that a single bleed does not reflect the full complexity of concurrently infecting *P. vivax* clones. The true MOI of *P. vivax* is underestimated to a greater extent than the MOI of *P. falciparum*. The effect of repeated sampling on prevalence, however, was in the same range in *P. falciparum* and *P. vivax* for both species 6 to 9% of all infections were missed on any single day. The known biology suggests that low parasite densities should have different causes in *P. falciparum* and *P. vivax*. *P. falciparum* parasites sequester periodically but a clone is absent from the peripheral circulation only if its erythrocytic cycle is tightly synchronized, which according to our data seems not to be the rule. *P. vivax* generally occurs at lower densities, and the timing of a parasite clone within the erythrocytic cycle seems to be well synchronized. These are major differences between both species, but with respect to detectability of parasite clones in the blood stream, both species differ less than previously thought.

In our study participants parasite densities for *P. vivax* dropped with increasing age, while for *P. falciparum* densities remained at the same level [15]. If parasite densities would directly impact detectability, a similar decrease in detectability with increasing age would thus be expected for *P. vivax*. However, we did not observe a clear effect of parasite density or age on detectability. Despite the fact that we did not detect an age trend in detectability in our study participants aged 1 to 4.5 years, this relationship might be different in older individuals. It remains open whether the often observed decrease of prevalence over the entire age range in moderate to high levels of transmission might be due at least partially to lower detectability associated with decreasing parasite densities.

Our *P. falciparum* results were generated by using a similar experimental and analytical approach as in a previous study, conducted in a highly endemic area in Ghana, and results can thus be compared. When all age groups were included in the Ghana study a strong age-dependency of detectability was noted, with a detectability of over 60% in younger individuals and only 10% in adults [24]. Overall, only 35% of all clones present in the host were detected in a single blood sample [18]. In individuals of the same age group as our participants, detectability ranged from 0.51 to 0.53 [31]. This lower detectability in Ghana could be affected by higher malaria endemicity and higher mean MOI. A lower transmission and therefore slower acquisition of immunity in PNG might lead to lower age dependency in detectability. It remains open whether in older children the often-observed decrease of prevalence by age in moderate to high levels of transmission might be due at least partially to lower detectability associated with decreasing parasite densities. Our current results reflect the situation in children harbouring high parasite densities. The situation in adults might be different.

In conclusion, when both 24 h bleeds were combined, we observed an increase in precision of estimates of epidemiological parameters in our study for both *P. falciparum* and *P. vivax*. While the increase in observed prevalence was limited, the effect on detection of individual alleles was more pronounced. Especially in highly endemic countries where most patients carry multiple clone infections repeated sampling substantially increases the precision of observed epidemiological parameters such as MOI. There was surprisingly little difference between the two parasites; just as studies of *P. falciparum* should recognize that they only detect a proportion of infections, the same is true for *P. vivax*.

**Acknowledgments**

We are grateful to the study participants and their parents or guardians, and to the IMR field team and microscopists. We thank two anonymous reviewers for their comments on the manuscript.

**Author Contributions**

Conceived and designed the experiments: IF IM TAS PAZ PS. Performed the experiments: CK SS EL BK. Analyzed the data: CK SS MB TAS. Wrote the paper: CK SS IF.
References


Multiplicity and Diversity of Plasmodium vivax Infections in a Highly Endemic Region in Papua Guinea

Cristian Koepfli1,2, Amanda Ross1,2, Benson Kiniboro3, Thomas A. Smith1,2, Peter A. Zimmerman4, Peter Siba3, Ivo Mueller3,5,6, Ingrid Felger1,2,*

1 Swiss Tropical and Public Health Institute, Basel, Switzerland, 2 University of Basel, Basel, Switzerland, 3 PNG Institute of Medical Research, Goroka, Papua New Guinea, 4 Centre for Global Health and Disease, Case Western Reserve University, Cleveland, Ohio, United States of America, 5 Walter and Eliza Hall Institute, Parkville, Victoria, Australia, 6 Barcelona Centre for International Health Research, Barcelona, Spain

Abstract

Plasmodium vivax is highly endemic in the lowlands of Papua New Guinea and accounts for a large proportion of the malaria cases in children less than 5 years of age. We collected 2117 blood samples at 2-monthly intervals from a cohort of 268 children aged 1 to 4.5 years and estimated the diversity and multiplicity of P. vivax infection. All P. vivax clones were genotyped using the merozoite surface protein 1 F3 fragment (msp1F3) and the microsatellite MS16 as molecular markers. High diversity was observed with msp1F3 (H = 88.1%) and MS16 (H = 97.8%). Of the 1162 P. vivax positive samples, 74% harbored multi-clone infections with a mean multiplicity of 2.7 (IQR = 1–3). The multiplicity of P. vivax infection increased slightly with age (P = 0.02), with the strongest increase in very young children. Intensified efforts to control malaria can benefit from knowledge of the diversity and MOI both for assessing the endemic situation and monitoring the effects of interventions.

Introduction

Malaria caused by Plasmodium vivax infection is increasingly recognized as a public health burden. The worldwide population at risk is estimated to be 2.85 billion, with high prevalences observed in locations throughout Southeast Asia and the Pacific [1]. Even though P. vivax epidemiology is less well studied and understood compared to that of P. falciparum, it is thought that P. vivax will present a greater challenge on the way to elimination of malaria outside Africa. Papua New Guinea (PNG) presents a variety of different climatic and ecological zones which have differing levels of malaria transmission [2] with a high burden of P. vivax in the tropical lowlands. The occurrence of high prevalence and morbidity marks locations in PNG as suitable field sites for P. vivax drug trials [3,4] and potentially also for future vaccine trials [5].

In Maprik in northern PNG, both P. vivax and P. falciparum are highly prevalent [6]. The incidence of P. vivax clinical episodes has been shown to peak in the second year of life, while that of P. falciparum increases until the fourth year [7,8]. Children between one and five years are considered to be a target age-group for P. vivax vaccine trials [5]. While previous studies have provided baseline data on clinical incidence rates, epidemiological patterns and the age-distribution of disease [8,9], genotyping data on individual P. vivax clones describing their diversity and molecular epidemiological parameters are scarce. For P. falciparum, the mean number of concurrent infections per patient (multiplicity of infection, MOI) has been used as one of several measures of the impact of interventions. MOI is crucial to assess the risk that an individual carries a drug resistant parasite and to evaluate levels of inbreeding [10].

Here we describe the genetic diversity and multiplicity of infection of a P. vivax population in an area of high malaria prevalence in the Maprik District in PNG. Data were obtained from children aged 1 to 4.5 years who were followed-up over 16 months [8]. We make use of two markers to genotype individual P. vivax clones, one a microsatellite and the other a region of the msp1 gene, encoding the Merozoite Surface Protein 1. While the microsatellite MS16 (located on NCBI contig XM_001615468.1) is considered a neutral marker that has been used in a number of population genetics studies [11,12], msp1 (XM_001614792) encodes a potential vaccine candidate (reviewed in [13]) and its diversity has been studied in different settings [14,15,16]. In a previous study, the two markers showed robust PCR amplification and high diversity with a risk of less than 1% that two clones share the same two-loci haplotype [17].

Methods

Ethics statement

The cohort study was approved by institutional review boards of the PNG Medical Research Advisory Committee (approvals 05.19 and 09.24), University Hospitals Case Medical Center (Cleveland, Ohio USA), and the Ethikkommission beider Basel
Study site and design

The cohort study was conducted in the Ilaita area, Maprik District, East Sepik Province, PNG between April 2006 and August 2007. The study area has hyper- to holoendemic perennial transmission with moderate seasonal variation [8,18,19,20]. P. vivax infections are the most prevalent infection in young children and remain frequent into adulthood, while P. falciparum is the predominant infection in children over 4 years of age [6,21].

268 children aged 1 to 3 years at enrolment were followed-up over a period of 16 months. As part of the cohort study, the children were visited every two months with blood samples taken at least for one and, for some surveys, on two consecutive days. In the analysis presented here, only blood samples taken on the first day were included. The prevalence of Plasmodium species by microscopy in the study population was 44.3% for P. vivax malaria and also of relapses, the release of long-lasting, silent liver stages to the blood stream. This is the first dataset from a highly endemic setting that presents data on a large number of individual P. vivax clones genotyped with highly diverse markers.

Species detection, genotyping and data analysis

Finger prick and venous blood samples were collected and DNA was extracted as previously described [8]. The presence of P. falciparum, P. vivax, P. ovale and P. malariae was detected by light microscopy as well as by post-PCR Ligase Detection Reaction (LDR) [8], a molecular method for the detection and species identification of malaria parasites [23]. All samples which were P. vivax positive by light microscopy or LDR plus 38 negative baseline samples were selected for genotyping.

The selection of highly diverse molecular markers is crucial for assessing MOI in molecular epidemiology studies. Based on our previous results [17], we selected the polymorphic marker gene mspF3 and the microsatellite MS16 for genotyping. In contrast to population genetic studies, where large numbers of neutral markers of moderate to high diversity (e.g. microsatellites) are generally analyzed, a small number of highly polymorphic markers are ideal for tracking clones in epidemiological studies for two reasons. Where concurrent infections with several clones are common construction of haplotypes combining data from several PCR amplified markers is difficult. In addition, Selecting MOI is defined as the maximum number of clones by any of the markers typed, the risk of overestimating MOI due to PCR artefacts increases with the number of markers analyzed, in particular for microsatellite amplification where PCR artefacts caused by polymerase slippage are of concern [24] (Table A and B in Text S1). Any highly polymorphic marker is suitable for studying the epidemiology of multiple infections. As long as high diversity is maintained, there is no need for selective neutrality.

In this molecular epidemiological study, the coding sequence mspF3 was chosen because it harbours a more complex repeat structure than microsatellites. This has the advantage that PCR artefacts due to slippage are rare for mspF3. Our second marker was MS16, a highly polymorphic microsatellite that lacks dominant alleles.

PCR and capillary electrophoresis were performed with slight modifications of the published protocol [17] to save costs and labour time: a multiplex primary PCR was done with the primers for the 2 markers mspF3 and MS16 followed by individual nested PCRs for mspF3 and MS16. The primary PCR was done in a volume of 20 μl containing 1 μl template DNA, 0.25 μM of each primer (Eurofin MWG Operon), 0.3 mM dNTPs (Solis BioDyne), 2 mM MgCl₂, 2 μl Buffer B (Solis BioDyne) and 5 U TaqFIREPol (Solis BioDyne). As we expected low parasitaemia in samples negative by microscopy, 2 μl DNA instead of 1 μl were used for the primary PCR. 1 μl primary PCR product was used as the template for the nested PCR, which was performed in a volume of 20 μl containing 0.25 μM of each primer (Applied Biosystems), 0.2 mM dNTPs (Solis BioDyne), 2 mM MgCl₂, 2 μl Buffer B (Solis BioDyne) and 1.5 U TaqFIREPol (Solis BioDyne). The forward primers for the nested PCR were labelled with fluorescent dyes: 6-FAM for mspF3, NED for MS16. Cycling conditions were as follows: initial denaturation 95 °C for 1 minute, then 30 cycles (primary PCR) or 25 cycles (nested PCR) with 15 seconds denaturation at 95 °C, 30 seconds annealing at 59°C and 30 seconds elongation at 72 °C plus a final elongation of 5 minutes at 72 °C. Subsequently, capillary electrophoresis was performed as described [17].

The PCR data was analysed using the GeneMarker® programme version 1.85 (SoftGenetics). Based on experience from preliminary studies, peaks above a cut-off of 1000 units relative fluorescent intensity (RFU) were considered true amplification products, all peaks below this cut-off were considered background noise as well as lesser peaks in the vicinity of strong peaks reaching 40% (mspF3) and 70% (MS16) of their height. Occasionally the fluorescence intensity differed between plates or samples as indicated by varying signal intensities of the commercial size standard. To compensate for this technical shortfall, the standard cut off value was lowered from 1000 to 300 RFU if signal intensities of both sample peaks and size standard peaks were low (generally below 1000 RFU). This practice was justified by a greater agreement in positivity at both loci. As a consequence, the proportion of samples positive only for a single marker dropped from 12% to 10%. All samples were checked visually (after blinding of samples) for stutter peaks thereby excluding one mspF3 and 25 MS16 samples from further analyses.
The genotyping method was validated by typing a subset of 28 samples for both markers in duplicate. 80% of msp1/F3 clones and 88% of MS16 clones were detected in both replicates (Tables C and D in Text S1). An important reason for the imperfect detection of clones is the low concentration of template DNA in samples with scanty parasitemia, where partial amplification of all templates seems to be governed by chance. In serial dilutions of DNA in field samples, we have demonstrated this effect by performing PCR amplification in triplicate for each dilution (out of several clones detected in undiluted DNA).

Data analysis

Alleles were grouped into bins of 3 base pairs, defined by the expected size differences in the two markers: 3 base pairs (bp) for the coding region of P. vivax as well as for microsatellite MS16 harbouring a 3 bp repeat unit. When a single genotype was observed with both markers, a blood sample was defined as single clone infection. In multiple clone infections, the highest number of clones observed for either marker defined the combined MOI of a blood sample. We used the kappa statistic to describe agreement between the molecular markers after correcting for chance agreement.

The determination of MOI for single and double clone infections was validated by genotyping a subset of samples with 12 additional markers. MOI = 1 was confirmed in 67/92 (72.8%) samples and MOI = 2 in 31/32 (96.9%) samples (Tables A and B in Text S1).

Although the distributions of MOI are skewed, we present the mean MOI, a common measure, to allow comparisons with other studies. We estimated the effect of age (at the time of the survey in 6 months age groups) and season on prevalence and MOI using regression models. To account for multiple visits per child, we included a random effect for child. The models were implemented in STATA version 10 [25] and WinBUGS version 1.4 [26].

The genetic diversity of a given locus in a population is expressed by the virtual heterozygosity $H_E$, i.e. the probability that two clones taken at random from the population carry different alleles. $H_E$ was calculated using the formula

$$H_E = \frac{n}{n-1} \left(1 - \sum p_i^2\right)$$

where $n$ is the number of clones analysed and $p_i$ is the frequency of allele $i$. $H_E$ of msp1/F3 and MS16 were determined by using only the first $P$. vivax positive sample of each study participant; $H_E$ of msp1/F3/MS16 haplotypes by using the first single clone infection per patient. This procedure prevents potential sampling bias due to repetition of persisting clones from the same individual. Linkage between markers was assessed using LIAN 3.5 [27]. Linkage disequilibrium measured from only two markers cannot provide information on inbreeding and was used only to provide evidence that the markers occur independently of each other.

Results

Of the 2117 blood samples collected in cross sectional surveys, 1340 were genotyped since they were positive by microscopy or LDR. Of the 88 microscopy and LDR negative samples that were selected for genotyping, only 2 (2.3%) were positive. This low proportion did not justify genotyping all negative samples.

Nested PCR for the two $P$. vivax genotyping markers provided an amplification product from at least one marker from 1162 samples. msp1/F3 PCR products were obtained from 1094 samples (93.3%), positive results were obtained for both markers. The two markers agreed well on $P$. vivax positivity (kappa = 0.71), although the difference in PCR efficiency (McNemar’s test: $P = 0.026$) suggests a slightly higher sensitivity of the MS16 PCR.

Prevalence

The overall prevalence of $P$. vivax based on positivity by PCR was 55%. The prevalence was lowest in children under 1.5 years at 44% and reached 62% in children aged 3 to 3.5 years. The increase of prevalence by age at the time of the survey was significant ($P = 0.005$) largely driven by the lower prevalence in children less than 1.5 years. Without this youngest age group no evidence of a trend was observed. No major seasonal trend in $P$. vivax prevalence was observed with the exception of a slight peak in September ($P = 0.17$).

Allelic diversity

In 1162 samples positive for $P$. vivax, 57 different msp1/F3 and 103 different MS16 alleles were detected (Figures 1A and 1B). Virtual heterozygosity $H_E$ was 97.8% for MS16, 88.1% for msp1/F3 and 99.1% for msp1/F3-MS16 haplotypes determined in single-clone infections (Table 1). In 219 single clone infections from 148 patients, a total of 134 different haplotypes were observed with the most common haplotype detected in only six individuals (Figure 1C). We tested this data set for independence of the two molecular markers. No linkage disequilibrium was observed ($L^2_s = -0.0001, P = 0.53$).

Multiplicity of infection

The MOI was determined for each marker separately, as well as for both markers combined. In $P$. vivax positive samples, the mean

| Table 1. Diversity and multiplicity of infection of $P$. vivax in Papua New Guinea. |
|---------------------------------|-----------------|----------------|-----------------|-----------------|
| **No. of positive samples** | **No. of clones** | **Multiplicity of infection** | **Allelic richness** | **Virtual Heterozygosity ($H_E$)** |
| msp1/F3 | 1094 (79)* | 2480 (173)* | 2.27 | 57 (31)* | 0.881 (0.874)* |
| MS16 | 1118 (80)* | 2542 (175)* | 2.27 | 103 (65)* | 0.976 (0.977)* |
| combined | 1162 | NA** | 2.69 | NA** | NA** |
| Single clone infections | 219 (25)* | 219 (25)* | 1 | 154 (25)* | 0.991 (0.96)* |

*In brackets numbers for baseline only. **NA = Not applicable. doi:10.1371/journal.pntd.0001424.t001

**NA = Not applicable.**
MOI was 2.27 for each marker individually and 2.69 when calculated from the maximal number of clones per sample by any marker. Among samples for which positive results were obtained from both markers, MOI was concordant in 38% (397/1050) (kappa = 0.17). A difference of one clone was observed in 38.5% (404/1050) of samples. The frequency distribution of MOI plotted separately for msp1F3 and MS16 was compared to the combined MOI (Figure 2). A single marker slightly underestimated MOI. Multiple clone infections were observed in 63% of all positive samples by msp1F3 and in 61% by MS16. When results of both markers were combined, the proportion of multiple clone infections increased to 74% (Table S1). Likewise, the proportion of samples with a MOI of 3 or higher was underestimated based on a single marker. Of the 531 samples with a combined MOI between 3 and 9, the combined MOI result was reproduced only in 279 samples (52%) by msp1F3 and in 61% by MS16. When results of both markers were combined, the proportion of multiple clone infections increased to 74% (Table S1). Likewise, the proportion of samples with a MOI of 3 or higher was underestimated based on a single marker. Of the 531 samples with a combined MOI between 3 and 9, the combined MOI result was reproduced only in 279 samples (52%) by msp1F3 and in 61% by MS16 alone. The mean MOI of *P. vivax* was associated with age. In children up to 1.5 years the mean MOI was 2.4 increasing slightly up to 2.8 in children 3.5 to 4.5 years of age (Figure 3, P=0.02). If the youngest children below 1.5 years were excluded, no significant trend was observed (P=0.23). The increase of the proportion of children bearing more than two clones was more pronounced (Figure 4). In the 18 youngest children aged 300 to 400 days, we observed a low MOI of 1.67 and only two children (11%) carried more than 2 clones. There was no significant seasonal variation in MOI (P=0.50).

**Discussion**

We have genotyped *P. vivax* parasites in a cohort of 268 children from an area of Papua New Guinea with sympatric *P. falciparum* and *P. vivax* with prevalences of 49.6% and 53.0% respectively in this cohort at enrolment [8]. The ecology and epidemiology of *P. vivax* differs from that of *P. falciparum* in several aspects. Parasite densities are generally lower, which is likely to affect prevalence data generated by both microscopically or molecular diagnosis [28,29]. Gametocytes appear shortly after an infection is established in a host [30], with implications for transmission and the frequency of sexual recombination, and the occurrence of relapses leads to appearance of new genotypes in the blood stream independent of mosquito transmission.

Our genotyping enables the distinction of individual parasites within the human host and thus the assessment of MOI. The mean MOI for *P. vivax* in our cohort was 2.7 and 73.6% of samples carried multiple clones (combined results from two independent markers). This compares to a substantially lower *P. falciparum* MOI of 1.5 in the same cohort determined by the marker *msp2* and only 35.2% of samples carrying multiple clone infections [31].

The multiplicity of *Plasmodium* infections depends on a number of factors including transmission intensity or the duration of infection as the result of loss of infection and antimalarial treatment. An additional factor, unique to *P. vivax*, contributes to the number of blood stage infections circulating in the blood: relapses of semi dormant liver stages. Previous studies have shown that relapses often genetically differ from already present blood stage parasites [32,33] and thus lead to increased MOI. As Coartem does not clear liver stage parasites and the level of treatment in our cohort was high, the combined effect of treatment and relapses is likely to enhance differences between *P. vivax* and *P. falciparum* MOI. In addition, as mosquitoes biting people harbouring multi-clone infections are more likely to transmit several clones concurrently [34,35], the higher MOI among *P. vivax* blood-stage parasites will increase the likelihood that multi-clone *P. vivax* infections are transmitted in a single mosquito bite. Under intense transmission such as found in lowland PNG, the MOI of *P. vivax* species increases with age, with the increase most pronounced in children below 1.5 years. In an earlier study, no evidence of differences between children aged 4 to 14 were found [36]. This increase in early childhood may at least in part be related to the increased exposed body surface with child growth thus leading to higher rates of mosquito bites and consequently risk of infection [37]. In addition, with rapidly increasing immunity fewer *P. vivax* infections may reach high densities which are associated with febrile illness and antimalarial treatment [8] and the average duration of a *P. vivax* infection may increase with age.
Both the mean MOI and prevalence using genotyping data showed no pronounced seasonality, concurring with previous findings using light microscopy and LDR detection [8]. In contrast, the incidence of clinical disease increased in the wet season [8]. The lack of annual fluctuations in *P. vivax* prevalence and MOI observed in PNG is likely to be caused by relapses during periods where there is less mosquito transmission.

The proportion of multiple-clone infections clearly differs from observations from countries of lower *P. vivax* endemicity. We observed polyclonal infections in three out of four samples. Even if numbers cannot be compared directly to other studies using other genotyping protocols (higher numbers of markers increase the chance of observing several clones at least in one marker), this proportion only reaches 11 to 49% in the Amazon [12,38], 55% in Sri Lanka [11] but 73% in Myanmar [11].

The molecular markers *msp1*F3 and MS16 showed a high degree of genetic diversity. While three *msp1*F3 alleles reached frequencies above 10%, the distribution of MS16 alleles is more homogenous with the highest frequency of 5%. The selected markers are suitable for studies where high resolution discrimination between *P. vivax* clones is required, such as longitudinal tracking of clones or discrimination between existing and incoming infections. The probability of two individual clones sharing the same 2-loci haplotype was below 1%. Analysing additional polymorphic markers would lead only to a minimal improvement of discrimination. As the mean MOI increases, the chance that two clones within a host share the same haplotype increases. Simulations indicate that mean MOI would be unlikely to be substantially underestimated with either marker, unless the mean MOI was greater than 6 (Amanda Ross, manuscript in

**Figure 3. Boxplot of MOI by age group.** The median is represented by the central line. The box represents the interquartile range from the 25th to 75th centiles. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box, points beyond this are plotted individually.

doi:10.1371/journal.pntd.0001424.g003
Due to its antigenicity and surface exposed location, MSP1 is considered a candidate for a malaria vaccine (reviewed in [13]). We identified 57 MSP1F3 alleles with the 3 most abundant alleles adding up to a frequency of 53%. The predominant alleles maintained stable allelic frequencies throughout the study. In a similar number of Papua New Guinean children, 27 haplotypes were observed for another potential vaccine candidate, the Duffy Binding Protein II (DBPII, XM_001615397). The 3 most frequent DBPII alleles were present in 57% of infections [40]. These results indicate high diversity of P. vivax antigens in PNG. With respect to vaccine development based on PfMSP1, the allelic frequencies generated in our genotyping study provide useful information on genetic diversity of this antigen.

In summary, this study provides one of the first large data sets of P. vivax genotypes from a highly endemic area. Our high resolution typing technique accurately determined allelic frequencies and clone multiplicity. This adds to the knowledge about P. vivax epidemiology and may serve as reference data for high endemicity P. vivax populations. The molecular parameters established could be utilized as one of several measures for effective monitoring of intervention and control of P. vivax, for surveillance and to parameterize mathematical models of transmission dynamics [40].

Supporting Information

Table S1  Genotyping results from 1162 P. vivax positive field samples using the markers msp1F3 and MS16. This table contains the number of samples for MOI = 1 to MOI = 9 for each marker separately and for the combination of both markers. (PDF)

Text S1  Validation of Plasmodium vivax genotyping based on msp1F3 and MS16 as molecular markers. This text contains results from confirmation of multiplicity of infection in field samples by genotyping additional molecular markers, congruence of samples typed in duplicate, and detection of clones in serial dilutions of DNA. This file contains Tables A–F and Figures A and B. (PDF)

Checklist S1  STROBE checklist. (DOC)

Acknowledgments

We thank the study participants and their parents and guardians, and the IMR field team.

Author Contributions

Conceived and designed the experiments: CK BK. Performed the experiments: CK AR IF. Analyzed the data: CK AR. Wrote the paper: CK AR IF.

References


Table S1: Genotyping results from 1162 *P. vivax* positive field samples using the markers *msp1F3* and MS16

<table>
<thead>
<tr>
<th>MOI</th>
<th>MOI determined by <em>msp1F3</em></th>
<th>MOI determined by MS16</th>
<th>MOI combined from both markers</th>
<th>MOI identical by both markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>404</td>
<td>439</td>
<td>305</td>
<td>230</td>
</tr>
<tr>
<td>2</td>
<td>291</td>
<td>295</td>
<td>305</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>208</td>
<td>181</td>
<td>235</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>118</td>
<td>107</td>
<td>166</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>49</td>
<td>89</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>32</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>15</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Sample negative for one marker and positive for the other marker: 67, 19, 0
Sample excluded because of PCR artifacts: 1, 25, 0

Total number of samples: 1162, 1162, 1162, 397
Total number of multiclonal infections: 690 (63.1%)**, 679 (60.7%)**, 857 (73.6%)**
Total number of single clone infections: 404 (36.9%)**, 439 (39.3%)**, 305 (26.2%)**

* If values from both markers were discrepant, the higher value was accepted

** Not taking into account samples negative for this marker and samples excluded because of PCR artifacts
Text S1: Validation of *Plasmodium vivax* genotyping based on *msp1*F3 and MS16 as molecular markers

1. Confirmation of multiplicity of infection in field samples from Papua New Guinea by genotyping additional molecular markers

92 single clone infections and 32 double clone infections (based on combined results from *msp1*F3 and MS16) were typed with 12 additional markers: MS1, MS2, MS4, MS5, MS6, MS7, MS8, MS9, MS10, MS12, MS15, MS20 [1] (manuscript in preparation).

Assessing multiplicity of infection (MOI) based on a large number of markers increases the risk that MOI is overestimated because of potential within-host clonal variation [2] or stutter artefacts occurring in one marker. Intra-host clonal variation is thought to be the result of strand-slippage in a repetitive sequence, such as a microsatellite, during mitotic DNA replication. Stutter peaks are PCR artefacts observed especially when microsatellites with a simple repeat structure are amplified [3]. Despite visual checking of electropherograms for stutter peaks and excluding them, some PCR artifacts may be overlooked. Both processes lead to higher MOI at one locus, while the other loci remain unchanged.

To compensate for within-host clonal variation and PCR stutter peaks, we set a cut-off as follows: a sample was accepted as multi-clone infection if at least 3 loci indicated a MOI>1. Applying this cut-off, MOI=1 (based on *msp1*F3/MS16 typing) was confirmed in 67/92 samples (72.8%) (Table A).

**Table A: Confirmation of single clone infections (MOI=1) by genotyping 12 additional markers**

<table>
<thead>
<tr>
<th>Total no. of samples</th>
<th>all markers</th>
<th>1 marker with MOI&gt;1</th>
<th>2 markers with MOI&gt;1</th>
<th>3 markers with MOI&gt;1</th>
<th>&gt;3 markers with MOI&gt;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>92</td>
<td>31 (33.7%)</td>
<td>23 (25.0%)</td>
<td>13 (14.1%)</td>
<td>18 (19.6%)</td>
<td>7 (7.6%)</td>
</tr>
</tbody>
</table>

In sample sets typed with many markers it has been frequently observed that the highest MOI detected by any marker is not supported by other markers [4,5], suggesting occurrence of intra-host clonal variation and PCR stutter peaks. Their frequency is not known. However we can estimate it from our data in table A that includes results from 1288 individual PCRs (92 samples typed with *msp1*F3, MS16 and 12 microsatellites). In a total of 49 PCRs additional alleles were detected (23 samples with 1 PCR indicating MOI >1 and 13 samples with 2 PCRs indicating MOI > 1). Based on these values, we estimate the frequency of PCR results
that overestimate MOI to be 49/1288 = 3.8%. The following calculation illustrates the increased risk for overestimating MOI when many markers are analyzed: assuming a 4% chance of a false-positive peak for a single marker, 96% of all samples will reveal the true MOI for this marker. However, with 14 independent markers, the chance of observing the true MOI drops to 56%.

In addition to single clone infections, we have genotyped 32 samples with MOI=2 revealed by msp1F3/MS16 typing for 12 additional markers. Results are given in table B. Applying the same cut off as above (maximal two markers with higher MOI), MOI=2 was confirmed in 31/32 (96.9%) of samples.

For 4/32 (12.5%) samples, MOI=2 was not confirmed by any additional marker. We cannot rule out that these samples are in fact single clone infections with an incorrect result for either msp1F3 or MS16.

**Table B: Confirmation of MOI=2 by genotyping 12 additional markers**

<table>
<thead>
<tr>
<th>Total no. of samples</th>
<th>all markers MOI=1 or 2</th>
<th>1 marker with MOI&gt;2</th>
<th>2 markers with MOI&gt;2</th>
<th>3 markers with MOI&gt;2</th>
<th>&gt;3 markers with MOI&gt;2</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>20 (62.5%)</td>
<td>9 (28.1%)</td>
<td>2 (6.2%)</td>
<td>1 (3.1%)</td>
<td>0</td>
</tr>
</tbody>
</table>

**Conclusion:** When choosing the number of markers to be genotyped, there is a trade-off between number of artefacts, which increases with each additional marker, and limited resolution of multiple infections when there are high frequencies of some alleles. Using the 12 additional markers, we reproduced the msp1F3/MS16 genotyping results for a MOI of 1 or 2 in a reasonable proportion of the samples.
2. Congruence of samples typed in duplicate

We genotyped 28 samples in duplicate for both markers *msp1F3* and MS16. Congruence in MOI results and allelic composition in each sample was assessed.

MOI was identical in 50% (*msp1F3*) and 71% (MS16) of duplicates (Table C). The agreement, estimated using kappa (Table C), was fair to high.

**Table C: Congruence in MOI after genotyping 28 samples in duplicate**

<table>
<thead>
<tr>
<th>Difference in MOI</th>
<th><em>msp1F3</em></th>
<th>MS16</th>
<th>Combined results</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>14 (50%)</td>
<td>20 (71.4%)</td>
<td>15 (53.6%)</td>
</tr>
<tr>
<td>1</td>
<td>10 (35.7%)</td>
<td>4 (14.3%)</td>
<td>9 (32.1%)</td>
</tr>
<tr>
<td>2</td>
<td>4 (14.3%)</td>
<td>2 (7.1%)</td>
<td>2 (7.1%)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>2 (7.1%)</td>
<td>2 (7.1%)</td>
</tr>
</tbody>
</table>

**kappa** 0.33 0.62 0.4

Detection of individual genotypes was compared between duplicates. The *msp1F3* allelic composition was identical in 12/28 (42.9%) sample pairs, that of MS16 in 19/28 (67.9%).

The probability of detecting a clone in the duplicate samples (n=28) was estimated using the formula presented by Bretscher et al [6] (Table D).

**Table D: Detection of clones in 28 samples genotyped in duplicate**

<table>
<thead>
<tr>
<th></th>
<th><em>msp1F3</em></th>
<th>MS16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clones detected twice</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td>Clones detected only once</td>
<td>26</td>
<td>16</td>
</tr>
<tr>
<td>Detectability [0.95 CI]</td>
<td>0.80 [0.73-0.88]</td>
<td>0.88 [0.83-0.94]</td>
</tr>
</tbody>
</table>

The congruence obtained from these 28 samples genotyped in duplicate was compared to detectability of clones. Detectability was estimated from sample pairs collected 24 hours apart from the same child [7]. Fluctuating parasite densities from one day to another lead to differences in bleeds collected within short time intervals, but no gain or loss of infection is expected within such a short time period. When day 1 and day 2 bleeds from the same child
were analyzed, detectability was 0.73 for *msp1F3* clones and 0.61 for MS16 clones [7]. This comparison revealed that the congruence of results obtained from duplicated typing of the same blood sample was much higher, 0.80, for *msp1F3* and 0.88 for MS16.

Imperfect detectability of parasite clones in field samples is caused by low densities of parasites. Depending on the presence or absence of a parasite clone in the DNA template added to the PCR reaction mix, PCR detection of low-density parasites leads to a chance result [8]. Results obtained at the detection limit of PCR can be exemplified by performing serial dilutions of a DNA solution followed by PCR (see paragraph 3 below).

**Conclusion:** The agreement of MOI and clone detection in the duplicate samples was reasonable and so genotyping the same sample in duplicate is not necessary.
3. Detection of clones in serial dilutions of DNA

Any method to detect parasites at densities close to the detection limit is imperfect. To assess the impact of low concentrations of template DNA on the outcome of our *P. vivax* genotyping PCR, DNA from field samples was diluted 1:10^2, 1:10^3, 1:10^4 and 1:10^5. Subsequently, *P. vivax* genotyping was performed in triplicate using the standard protocol [7]. Six field samples were analyzed. MOI and allelic composition in each sample were assessed. Results for marker *msp1F3* are listed in Table E and for MS16 in Table F.

**Table E: Number of *msp1F3* clones detected in triplicates of serial DNA dilutions**

<table>
<thead>
<tr>
<th>DNA dilution</th>
<th>1:1</th>
<th>1:10^2</th>
<th>1:10^3</th>
<th>1:10^4</th>
<th>1:10^5</th>
<th>Total no. of clones observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>triplicate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>Sample 1*</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sample 2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Sample 3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Sample 4*</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Sample 5</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sample 6</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Total no. of pos. samples (n=18)</td>
<td>18/18</td>
<td>16/18</td>
<td>12/18</td>
<td>4/18</td>
<td>0/18</td>
<td></td>
</tr>
</tbody>
</table>

* Electropherogramms are displayed in Figures A and B.
Table F: Number of MS16 clones detected in triplicates of serial DNA dilutions.

<table>
<thead>
<tr>
<th>DNA dilution</th>
<th>1:1</th>
<th>1:10²</th>
<th>1:10³</th>
<th>1:10⁴</th>
<th>1:10⁵</th>
<th>Total no. of clones observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>triplicate</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>Sample 1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Sample 2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Sample 3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sample 4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Sample 5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sample 6</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total no. of pos. samples (n=18)</td>
<td>18/18</td>
<td>14/18</td>
<td>10/18</td>
<td>4/18</td>
<td>1/18</td>
<td></td>
</tr>
</tbody>
</table>

Genotyping results from undiluted and serial dilutions of DNA from field samples 1 and 4, done in triplicate, are shown in Figure A and B, respectively. At a DNA dilution of 1:10³, some clones are no longer amplified, as compared to undiluted DNA. At such low DNA concentrations, the distribution of detected genotypes seems to be random, as seen in sample 4 (Figure B): one or two genotypes per PCR reaction were amplified at a DNA dilution of 1:10³, but the allelic composition of the 3 replicates differed. Triplicate A showed two fragments of 236 bp and 262 bp, respectively. Triplicate B contained a single 262 bp fragment only, and triplicate C a single 274 bp fragment. All 3 clones were amplified when DNA was not diluted.
Figure A: Electropherograms of *msp1F3* genotyping PCR performed in triplicate (relicates A, B, and C) on serial dilutions of field sample 1. The standard cut-off of 1000 RFU was applied for all samples. Alleles above this cut-off are indicated with a grey line.

X-axis: size of DNA fragment in base pairs

Y-axis: Relative fluorescent units (RFU)

Red: commercial size standard

Green: *msp1F3* alleles
**Figure B:** Electropherograms of *msp1F3* genotyping PCR performed in triplicate (replicates A, B, and C) on serial dilutions of field sample 4.

**Conclusion:** When low parasitemia is simulated by serial DNA dilutions, detection of clones becomes stochastic and differs between replicates.
References


Plasmodium vivax Populations in Papua New Guinea: Highly Diverse and Unstructured

Cristian Koepfli¹,², Lincoln Timano¹,³, Peter Siba³, Ivo Mueller⁴, Alyssa Barry¹, Tiago Antão⁵, Ingrid Felger¹,²

¹ Swiss Tropical and Public Health Institute, Basel, Switzerland
² University of Basel, Basel, Switzerland
³ PNG Institute of Medical Research, Goroka, Papua New Guinea
⁴ Infection & Immunity Division, Walter & Eliza Hall Institute, Melbourne, Australia
⁵ Liverpool School of Tropical Medicine, Liverpool, UK

This chapter will be submitted for publication after further revisions.
Abstract

The clinical burden of *Plasmodium vivax* malaria is increasingly recognized, and the importance of this parasite is highlighted by its predominance in focus areas of intensified anti-malaria interventions towards elimination. Yet little is known about its population genetic structure, in particular in areas of high *P. vivax* transmission. Knowledge on genetic exchange within and between populations and gene flow has implications for drug and vaccine based control strategies. Papua New Guinea (PNG) comprises a variety of *P. vivax* transmission settings, with the worldwide highest transmission ever documented.

We have genotyped 254 samples from 3 lowland and 1 highland sites in PNG for 14 microsatellite loci. Diversity is high with expected heterozygosity values ranging from 0.66 to 0.97 for the different markers. 252 individual multi-locus haplotypes were observed in single clone infections or in dominant clones of multiple clone infections. Separation between sites was very low ($F_{ST}$=0.0015-0.0251), with the slightly more pronounced difference occurring between lowland and highland locations. These findings imply high effective population size and levels of high gene flow. The population genetic parameters for *P. vivax* in PNG are in sharp contrast to a strong sub-structuring reported from other countries with low *P. vivax* endemicity. The high level of gene flow in *P. vivax*, and thus extensive parasite or vector migration, strongly argues for country-wide rather than patchy control efforts in order to be effectively control vivax malaria in PNG.
Introduction

Papua New Guinea (PNG) is outstanding with respect to biodiversity as well as diversity of human cultures and languages (Attenborough and Alpers, 1992). Ecological zones range from tropical lowlands to mild and tempered climate zones in the highlands and high mountain ranges. PNG harbours four *Plasmodium* species that infect humans, *P. falciparum*, *P. vivax*, *P. malariae* and *P. ovale*, with probably the highest *P. vivax* prevalence anywhere in the world (Mueller et al., 2003). *P. vivax* is the predominant cause of malaria outside Africa (Guerra et al., 2010), and there are increasing reports of severe outcome of disease, especially from PNG and West Papua (Genton et al., 2008; Poespoprodjo et al., 2009; Tjitra et al., 2008).

Population structure can inform interventions towards malaria elimination, as marked differences between parasite populations indicate little gene flow, suggesting restricted parasite or vector migration. In such settings with restricted gene flow, antimalarial interventions in small areas seem possible. At a later stage, when transmission is reduced to a low level, genotyping could play an important part in tracking outbreaks.

Diversity, population structure, gene flow and linkage disequilibrium (LD) between loci influence the emergence and spread of drug resistance and affects efficiency of potential future vaccines (Ariey et al., 2003; Buckee and Gupta, 2010; Yuan et al., 2011). LD between loci is high in nearly clonal populations with little recombination between strains, but associations between markers are quickly reverted in panmictic populations. The extent of LD influences the marker density needed for genome wide association studies (Hayton and Su, 2008; Mu et al., 2005). Likewise LD influences spread of drug resistant parasites. A number of current antimalarial drugs combine compounds acting on different targets of the parasite metabolism, and as a consequence accumulation of mutations in different genes is needed to confer drug resistance (Cammack, 2011). Once emerged by chance, such drug resistance haplotypes can persist and spread quickly in conditions of low genetic exchange. But if parasites undergo frequent recombination with non-resistant parasites, resistance haplotypes may be quickly interrupted (Talisuna et al., 2003).

Multi-locus genotyping of neutral markers of *P. falciparum* and *P. vivax* provide data for the analysis of diversity and population structure. Population genetic studies from village- to intercontinental level were undertaken with samples from Latin America and several South-East Asian countries, where *P. vivax* prevalence is generally lower than in PNG. *P. vivax* populations from these countries have revealed considerable genetic differentiation between populations, suggesting limited gene flow (Ferreira et al., 2007; Gunawardena et al., 2010; Imwong et al., 2007; Karunaweera et al., 2008; Van den Eede et al., 2010b).

Because of much higher *P. vivax* transmission intensity in PNG, a different parasite population structure is likely. High diversity has been reported from studies of *P. vivax* surface
antigens in PNG (Bruce et al., 2000; Henry-Halldin et al., 2011; Kolakovich et al., 1996), and moderate differences between populations based on the *Duffy binding protein gene* (Cole-Tobian and King, 2003), but not on subtypes of the *circumsporozoite protein* (Henry-Halldin et al., 2011). Population genetic analyses based on neutral markers were conducted for *P. falciparum*, which in PNG is equally prevalent as *P. vivax*. High haplotype diversity and moderately structured population were observed at 4 lowland sites (Schultz et al., 2010). Results from a local malaria outbreak in the highlands could indicate that species-specific differences occur: clonality was confirmed for all *P. falciparum* infections, however, *P. vivax* samples from this study were highly diverse and included multiple infections (Mueller et al., 2002).

Our previous finding of a *P. vivax* MOI twice as high as that of *P. falciparum* (Koepfli et al., 2011) also suggested that sympatric populations of these parasites differ substantially despite their similar prevalence. This lead us to investigate the population parameters in 254 *P. vivax* samples collected from patients at three sites in the lowlands and one site in the highlands of PNG, using 14 size polymorphic neutral microsatellite markers. Diversity of parasites and population structure was assessed village- and country-wide, and compared to results obtained from the *msp1* gene, a polymorphic surface antigen likely under selection.
Methods

Ethics statement

This study was performed on archived samples from previous studies with approval from the Institutional Review Board of PNG Institute of Medical Research (Amendment to IRB 0919 and to MRAC 09.24 from the PNG Medical Research Advisory Committee).

Study sites and patients

*P. vivax* positive blood samples from four regions in PNG were used. Three study sites (Ilaita, Kunjingini and Alexishafen) were located in the tropical lowlands; one site (Sigimaru) was in the central highlands at an altitude of 1100 metres (table 1, figure 1).

At the Ilaita study site (Maprik District, East Sepik Province) cross-sectional sampling was done in several villages from April 2006 to August 2007 from 264 children aged 0.5 to 4.5 years (Lin et al., 2010). Samples were collected in the course of a drug efficacy studies in the Kunjingini health center (Maprik District, East Sepik Province) from April 2004 to February 2005 from children aged 0.5 to 7 years (Marfurt et al., 2007), at the Alexishafen health centre (Madang Province) from April 2005 to July 2007 from children aged 0.5 to 5 years (Karunajeewa et al., 2008), and in the Sigimaru health center (Karimui area, Simbu Province) between 2002 and 2005 from children aged 0.5 to 7 years (Marfurt et al., 2007).

Large parts of the coastal lowlands of PNG are characterized by high prevalence of *P. vivax* and *P. falciparum* malaria with perennial transmission and mild seasonal variation. In study participants from the Ilaita site, East Sepik Province, prevalence of *Plasmodium* species by microscopy was 44.3% for *P. vivax*, 32.6% for *P. falciparum* and 4.2% of *P. malariae*. Incidence rate of malaria was 1.92 episodes per child per year for *P. vivax* and 1.59 for *P. falciparum* (Lin et al., 2010). In contrast, in Simbu province in the highlands, malaria prevalence varies from below 5% at higher altitudes in the northern parts of the province to 35% in the southern and lower altitude areas, where our study site is located (Mueller et al., 2004). In a survey in 2001 and 2002 *P. vivax* prevalence in South Simbu was 8% (Mueller et al., 2004).

From all 4 study sites *P. vivax* positive samples had been genotyped in the course of previous epidemiological and drug efficacy studies, using the two markers *msp1F3* and MS16 (Barnadas et al., 2011; Koepfli, 2011). To facilitate construction of multi-locus haplotypes (see below), single clone infections according to these earlier typing results were selected for genotyping additional loci. Because the number of single clone infections was not sufficient, we also included samples with a multiplicity of 2 or 3.
Genetic markers, PCR and genotyping by capillary electrophoresis

We have selected a panel of 14 well-described and frequently used size polymorphic *P. vivax* markers for genotyping: MS1, MS2, MS4, MS5, MS6, MS7, MS8, MS9, MS10, MS12, MS15, MS16, MS20 (Karunaweera, 2006), Pv3.27 (Imwong et al., 2007; Koepfli et al., 2008). In addition the polymorphic F3 region of the *merozoite surface protein 1* (*msp1F3* (Koepfli et al., 2008)) was typed. MS3, that had also been described as polymorphic *P. vivax* marker (Karunaweera, 2006), was excluded after preliminary testing because no amplification product was obtained from over 50% of samples. These markers were selected for comparability with previous studies in other countries.

Genotyping results for MS16 and *msp1F3* were retrieved from earlier studies (Barnadas et al., 2011; Koepfli, 2011; Koepfli et al., 2008). For 13 additional microsatellite markers (MS1, MS2, MS4, MS5, MS6, MS7, MS8, MS9, MS10, MS12, MS15, MS16, MS20, Pv3.27) a semi-nested PCR protocol with 13-plex primary followed by individual reactions for nested PCR was applied. Forward primers for the 13 primary PCR assays were designed (supplementary table S1) and used in combination with previously published nested primers. Nested PCR primers were adopted from earlier publications (Imwong et al., 2007; Karunaweera, 2006). Multiplex primary PCR was conducted under conditions optimized for multiplex reactions as follows:

- Each 50 µl reaction contained 1 µl template DNA, 5 µl Buffer B (Solis Biodyne), 0.3 mM dNTPs, 3 mM MgCl$_2$, 0.25 µM of each primer (MWG Operon, 26 primers in total) and 10 U Taq Firepol (Solis Biodyne). Cycling conditions were 1 minute initial denaturation at 95°C followed by 25 cycles of 15 seconds at 95°C, 30 seconds at 60°C, 30 seconds at 72°C and a final elongation step for 5 minutes at 72°C.

5 µl of each nested amplification product were run on a 1.5% agarose gel. According to the intensity of bands on this gel, PCR products were diluted in H$_2$O either in a ratio of 1:10, 1:20 or 1:40. If a very weak or no band was observed on the gel (e.g. for microsatellite MS4), the PCR product was diluted 1:4. This usually resulted in a detectable peak in capillary electrophoresis (CE). 2.5 µl of such diluted amplification product was analysed by CE as described (Koepfli et al., 2008).
Plasmodium vivax Populations in Papua New Guinea

Data analysis

Data retrieved from CE was analysed using the GeneMarker® programme version 1.85 (SoftGenetics) as described (Koepfli, 2011). We used a total of 15 markers, 14 of which were presumably neutral microsatellites, whereas msp1F3 represents a size polymorphic domain of intragenic repeats of the surface antigen MSP1. It is not clear whether this domain is under balancing selection, therefore msp1F3 typing results were only included in a separate round of our analyses, and results obtained from microsatellite data alone were compared to those with msp1F3 data.

All available single clone infections were used in this study. However, as up to 75% of all children in our study sites carried multiple clone infections, a sufficient sample size was not reached, and thus samples with multiplicity of 2 and 3 were also included. Samples yielding amplification products for <9/13 markers were excluded from further analysis. The full dataset included all clones detected per sample by any marker, i.e. the predominant clone plus all minority clones from multiple clone infections. For specific analyses the building of multi-locus haplotypes was required. In case of multiple clone infections we have used peak heights to determine the dominant clone and all minority clones were excluded. This strategy has been applied in previous studies (Anderson et al., 2000; Hunt et al., 2005). Due to the likely preferred amplification of smaller alleles or minor variation in peak height ratios, this approach could be error prone if peaks of 2 clones are of similar height. To rule out ambiguities, we have excluded from haplotype building all markers, for which the minor allele reached a peak height of >70% of the predominant peak.

Alleles were binned into 2, 3 or 4 bp bins according to their repeat unit size using TANDEM software Version 1.08 (Matschner and Salzburger, 2009). As TANDEM processes only genotyping data from diploid organisms, results from single clone infections were duplicated for constructing allele bins.

The software Dropout was used to detect genotyping errors due to PCR artefacts leading to overestimation of the number of haplotypes (McKelvey and Schwartz, 2005). Expected numbers of pairwise differences between haplotypes were compared with observed differences. The software Lositan was used to detect loci that are under positive or balancing selection and thus influence the analysis of population differentiation. To detect non-neutral loci 100'000 simulations were run in Lositan under a stepwise mutation model (Antao et al., 2008). For msp1F3 Lositan suggested populations differentiation in the range of that of microsatellite markers. Dropout and Lostian software gave no indication of genotyping errors or non-neutral loci, and all samples and markers were included into further analysis.

The genetic diversity of a locus in a population is expressed by the expected heterozygosity $H_E$, i.e. the probability that two clones taken at random from the population carry different
alleles. \( H_E \) was calculated as described (Gatton and Cheng, 2008). Linkage disequilibrium (LD) was assessed using the program LIAN 3.5 applying a Monte Carlo test with 100'000 random resamplings (Haubold and Hudson, 2000; http://adenine.biz.fh-weihenstephan.de/cgi-bin/lian/lian.cgi.pl). LIAN cannot handle missing data, thus 7 markers and subsequently all samples still containing missing data were removed from the data. The remaining dataset included the markers MS1, MS2, MS8, MS9, MS10, MS12, MS15 and MS20. The final sample size for each study site is shown in table 1. Slightly different datasets of samples without missing data, based on different combinations of markers, led to similar estimates of LD, indicating that results are not influenced by selection of markers. The standardized index of association \( I_{SA} \) was calculated as an estimation of LD using the equation

\[
I_{SA} = \frac{1}{I-1} \left( \frac{V_D}{V_e} - 1 \right)
\]

with \( I \) being the number of loci analyzed, \( V_D \) the variance of the number of loci at which each pair of haplotypes differs and \( V_e \) the expected variance in linkage equilibrium. \( I_{SA} \) is zero for linkage equilibrium (Haubold and Hudson, 2000).

To compare allelic frequencies between populations in Ilaita, Kunjingini, Alexishafen and Sigimaru, Wright's \( F_{ST} \) coefficient was calculated using FSTAT (Goudet, 1995). This program only allows analysis of diploid data. “Virtual diploids” were created from the full dataset by grouping alleles per study site randomly into pairs. This approach can be justified in haploid organisms because \( F_{ST} \) values rely entirely on allele frequencies within populations without requiring that an allele pair was derived from the homologous chromosomes of a diploid individual (Balding et al., 2007). \( P \)-values were adjusted using Bonferroni’s correction for multiple comparisons. \( F_{ST} \) takes values between 0 (indicating that allelic frequencies in all populations are identical) and 1 (indicating maximum differentiation between populations without any allele shared) (Balding et al., 2007; Hartl and Clark, 2007). Correlation between genetic difference, geographic distance and altitude was assessed by Mantel testing in FSTAT with 20'000 randomisations. Distances and altitudes were retrieved from Google Earth.

The samples from Ilaita were collected in cross-sectional surveys in different villages, approximately 2 to 5 km from each other. This allowed comparison of allelic differences between subpopulations. Five sites were compared according to the detailed map given in Figure 2.

As a complementary approach to assess population structure we have analysed haplotypes using the program STRUCTURE version 2.3.2 (Pritchard et al., 2000). This algorithm attempts to form groups of haplotypes without prior information on the origin of a sample. STRUCTURE is able to detect genetic differences among subgroups that do not correlate with
geography. The number of populations (K) was set from 1 to 20 with 3 replications per K, each with 100,000 Markov Chain Monte Carlo runs after a burn-in period of 50,000 steps, using the admixture model. Principal component analysis (PCA) was done in R using the prcomp function. Haplotypes containing missing data were excluded to calculate the covariance matrix, but projected into the final plot if not more than 3 data points were missing.
Results

We have genotyped 15 size polymorphic molecular markers in 254 *P. vivax* positive blood samples from different locations in Papua New Guinea. Out of 3810 expected genotyping data points (254 samples times 15 markers), 3566 (93.6%) data points were generated. For the construction of haplotypes in multiple-clone infections all minority clones were excluded, but in 155 electropherograms concurrent genotypes yielded peaks of almost equal height and thus no dominant genotype could be determined. This led to the exclusion of these data points from haplotype building. The final data set comprised 3411 out of a maximum of 3810 (89.5%) data points. Because these gaps were equally distributed over all samples and markers, excluding certain samples or markers would not improve data completeness. All markers were highly polymorphic with 8-61 alleles per locus and $H_E$ of 0.664-0.997 (Table 2).

We have assessed population structure using different approaches. $F_{ST}$ values were calculated in the full dataset, including also minor clones, as use of haplotypes is not required for this calculation. Very little differences were observed between lowland populations, but comparisons between lowland sites and Sigimaru gave slightly higher $F_{ST}$ values (Table 3A). The only statistically significant difference was found between Ilaita in East Sepik Province and Sigimaru in the highlands ($F_{ST}$=0.0247, $P$=0.0013), most likely due to the larger sample size from Ilaita. Correlation between genetic differentiation and geographical distance or altitude was assessed Mantel testing. Distance alone explained only 40% of genetic differences ($P$=0.449), but altitude explained 89.6% ($P$=0.071), and the combination of both explained 96.1%.

We have then searched for clustering of genotypes in our dataset using the widely used software STRUCTURE. STRUCTURE makes it possible to assess population genetic structure in groups of samples without prior information on the origin of these samples. No structuring was found in our data set, the highland - lowland difference between Ilaita and Sigimaru populations was not reproduced. This suggests that there is either no clustering in our data or only very little, as STRUCTURE is not adequate for detection of very low, yet significant $F_{ST}$ values (Pritchard et al., 2007). Similarly, Principal component analysis using R did not reveal any grouping of haplotypes (Figure 3).

Next we have assessed linkage disequilibrium (LD) between loci. As this analysis cannot handle missing data, we have assessed LD in a reduced sample set of 8 markers, only including samples without missing data (n=142). No LD was observed when assessed separately for all four study sites as well as for all sites in combination (Table 4).

In addition to the 14 presumably neutral microsatellite markers we have genotyped the marker *msp1F3*, which codes for the intragenic repeat region of the MSP1 surface antigen. This marker could be subject to balancing selection. We have repeated our analysis for *msp1F3*...
alone and for a combined dataset of the 14 microsatellites and \textit{msp1}F3. (Table 4B). Population structure analysis of \textit{msp1}F3 only confirmed the difference between Ilaita and Sigimaru ($F_{ST}=0.0332$, $P=0.00017$). Moreover, a difference between Alexishafen and Ilaita was observed ($F_{ST}=0.0299$, $P=0.00017$). When all 15 markers were analyzed in a single dataset, the difference between Alexishafen and Ilaita remained significant ($F_{ST}=0.0251$, $P=0.00167$, Table 3C). The number of individual haplotypes ($n=252$) remained the same after inclusion of \textit{msp1}F3, as well as the result of the analysis using the STRUCTURE software.

In a study from Peru, using the same markers, significant population differentiation was found between villages as close as 2 km (Van den Eede et al., 2010b). Samples from Ilaita were collected in individual villages, and we could thus assess population genetic substructuring between five villages in the study area (Figure 2), 1 to 5 km from each other. All $F_{ST}$ values between villages were close to 0 ($F_{ST} < 0.005$) and non-significant ($P > 0.4$). Inclusion of \textit{msp1}F3 did not change results, nor were differences between villages observed when only \textit{msp1}F3 was analyzed.
Discussion

All 14 microsatellites showed high numbers of alleles and a genetic diversity similar or higher to published results from Asia and Latin America (Imwong et al., 2007; Karunaweera, 2006; Van den Eede et al., 2010a; Van den Eede et al., 2010b). The genetic composition of the 3 lowland *P. vivax* populations differed only marginally, whereas $F_{ST}$ values for comparisons between lowland and highland sites were higher. Most likely this was caused by the fact that Sigimaru is separated from the 3 lowland sites by a sparsely populated region with little malaria transmission. *P*-values were influenced by different sample sizes, leading only for the Ilaita-Sigimaru comparison to a significant result.

This low level of population structuring in *P. vivax* observed in PNG is in sharp contrast to results obtained in Latin America, where *P. vivax* is the predominant malaria parasite, but endemicity is usually low. $F_{ST}$ values were 0.4-0.7 between five sites in Colombia (Imwong et al., 2007) and 0.03-0.25 between sites in Brazil (Rezende et al., 2009). Large geographic distances combined with mountain ranges without malaria transmission might add to the genetic differentiation seen in Colombia. But also on a scale of 2 to 50 km distance between sites, without geographical barriers to transmission, high population differentiation was found (Van den Eede et al., 2010b). *P. vivax* is the predominant Plasmodium species also in many Asian countries. Inter-country $F_{ST}$ values were established between populations from India, Laos and Thailand and ranged from 0.13-0.45 (Imwong et al., 2007).

*P. falciparum* populations from different lowland sites in PNG were investigated previously. *P. falciparum* diversity was equally high at each site and all 318 haplotypes, based on 10 loci, were different. Population differentiation in *P. falciparum* was more pronounced compared to the situation of *P. vivax*, and showed significant differences between all sites with $F_{ST}$ values of 0.05-0.14 (Schultz et al., 2010).

The little genetic structuring of *P. vivax* populations are difficult to explain in view of the geographic distance. Transport within PNG is restricted to few roads, often unpaved and difficult to travel, and expensive air transport. Only in the past decade inhabitants from the East Sepik province migrated to Madang province which could explain the similarity of parasites found in Alexishafen. In contrast, human migration between Sigimaru and the lowlands is likely more restricted. Yet our data suggest, that mobility along the highland lowland trading routes is sufficiently strong to allow genetic exchange among *P. vivax* populations.

No LD was detected between our markers. LD is a result of inbreeding of parasites and high LD in *P. vivax* has been reported previously from low transmission areas, such as the Amazon (Ferreira et al., 2007; Imwong et al., 2007; Rezende et al., 2009; Rezende et al., 2010; Van den Eede et al., 2010b), Sri Lanka or Ethiopia (Gunawardena et al., 2010;
Karunaweera et al., 2008). A report from South Korea, a country with currently emerging P. vivax prevalence, showed a decrease of LD paralleled with an increase in prevalence over a period of 10 years (Honma et al., 2011). Absence of LD was reported from India, Laos and Thailand (Imwong et al., 2007), however LD was observed in neighboring Myanmar (Gunawardena et al., 2010). This pattern of strong LD in low transmission, but no LD in high transmission follows expectations and is equally true for P. falciparum (Anderson et al., 2000; Neafsey et al., 2008).

Recombination among Plasmodium parasites can occur in the mosquito midgut if genetically different gametocytes are taken up at a blood meal. High levels of transmission increase the proportion of humans carrying multiple clone infections and thus augment the chance of simultaneous transmission of different gametocyte clones in a single mosquito bite. In blood samples from our study participants the proportion of children carrying multiple clone infections ranged from 63 to 73% (table 1). In areas of low P. vivax transmission this proportion ranged from 13% in Sri Lanka (Gunasekera et al., 2007) to 10-34% in South America (Imwong et al., 2007; Van den Eede et al., 2010b) and 30-35% in Thailand and Laos (Imwong et al., 2007).

To date it is still unknown whether all P. vivax clones concurrently present in a human host produce gametocytes simultaneously, and in how far multiplicity of infection in the host correlates with recombination frequency in mosquitoes. All P. falciparum and P. vivax clones were genotyped in the Ilaita cohort study (Lin et al., 2010), from which also our P. vivax samples derived (Koepfli et al., 2011). Only 33% of P. falciparum positive children carried multiple clone infections (Sonja Schoepflin, unpublished), in contrast to 73% multi-clone infections for P. vivax (Koepfli, 2011). It is plausible that in P. falciparum the predominance of single clone infections, and as a consequence less recombination in mosquitoes, caused the more substantial P. falciparum population structuring.

While our 14 microsatellite markers are presumable neutral, msp1F3 encodes a highly polymorphic region of the MSP1 antigen. Coding regions, in particular those of surface antigens, are under immune selection. For P. falciparum MSP1 it was shown that high levels of anti-msp1 antibodies protected children from malaria, and that this protection was allele specific (Conway et al., 2000). Parasites infecting a semi-immune host will be favored if they express an allele that is not recognized by antibodies, and as a consequence large number of alleles will be maintained in parasite populations. (Hastings, 1996). Indeed, in P. falciparum population genetic differences for a polymorphic msp1 fragment were less pronounced than for neutral loci (Conway et al., 2000). Our P. vivax msp1F3 locus is located approximately 1 kb upstream of the highly antigenic C-terminal MSP119 fragment, that elicits a strong antigenic response (de Oliveira et al., 1999; Soares et al., 1997) and is a widely studied vaccine candidate (reviewed in (Herrera et al., 2007)). Consequently, balancing selection could be...
expected to act on *msp1*F3, but given the low level of population differences we observed with neutral loci, we did not detect a different pattern for *msp1*. We cannot rule out that in situations were marked population genetic differences are detected by neutral microsatellite markers, balancing selection would be identified to act on *msp1*F3.

**Conclusions**

We present the first comparison of *P. vivax* populations on small to medium geographic scale in a country of generally high endemicity, but also including a site of presumably lower transmission intensity. The nearly complete lack of population structuring is in sharp contrast to results of low endemic areas. Our results highlight a difference in population structure between the sympatric species *P. vivax* and *P. falciparum*, with the latter species being more structured, possibly owing to its much lower mean MOI. Large population size, high gene flow and absence of LD have direct implications for interventions. Single mutations conferring drug resistance are expected to spread fast in the country, however multi-locus genotypes associated with resistance are expected to break down fast in absence of drug pressure (Dye and Williams, 1997). Our results imply frequent migration of parasites between different sites, which likely compromises the outcome of interventions in single locations.
Acknowledgements

We would like to thank the study participants and their parents and guardians, the IMR field team that collected the Ilaita samples, and the health centre staff in Alexishafen, Sigimaru and Kunjingini. We thank Enmoore Lin, Benson Kiniboro, Jutta Marfurt, Sonja Schoepflin, Tim Davis and Harin Karunajeewa for providing field samples for this study, and Anita Lerch for help with analysis. This work was supported by the Swiss National Science Foundation (grants no: 320030-125316/1 and 310030-134889/1).

References


http://adenine.biz.fh-weihenstephan.de/cgi-bin/lian/lian.cgi.pl.


Plasmodium vivax Populations in Papua New Guinea


Plasmodium vivax Populations in Papua New Guinea


Plasmodium vivax Populations in Papua New Guinea

Figure 1: Location of sites of sample collection in Papua New Guinea and altitude above sea level

Fig. 4: Sub-structure of the Ilaita study site. Allelic frequencies of the villages in the red circles were compared against each other. Adapted from (Lin et al., 2010).
Figure 3: Principal component analysis of \textit{P. vivax} haplotypes genotyped with 14 microsatellite markers.
Table 1: *P. vivax* samples included into this study

<table>
<thead>
<tr>
<th>Site</th>
<th>year of collection</th>
<th>No. of <em>P. vivax</em> pos. samples collected</th>
<th>Mean MOI of all <em>P. vivax</em> pos. samples*</th>
<th>Proportion of multiple clone infections in all <em>P. vivax</em> pos. samples*</th>
<th>No. of samples genotyped</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilaita</td>
<td>2006-2007</td>
<td>2096</td>
<td>2.67</td>
<td>73.1%</td>
<td>132</td>
<td>(Koepfli, 2011; Lin et al., 2010)</td>
</tr>
<tr>
<td>Kunjingini</td>
<td>2003-2005</td>
<td>94</td>
<td>2.07</td>
<td>63.8%</td>
<td>38</td>
<td>(Marfurt et al., 2007; Marfurt et al., 2008)</td>
</tr>
<tr>
<td>Alexishafen</td>
<td>2005-2007</td>
<td>150</td>
<td>2.27</td>
<td>72.7</td>
<td>45</td>
<td>(Barnadas et al., 2011; Karunajeewa et al., 2008)</td>
</tr>
<tr>
<td>Sigimaru</td>
<td>2004-2005</td>
<td>48</td>
<td>2.6</td>
<td>70.8%</td>
<td>39</td>
<td>(Marfurt et al., 2007)</td>
</tr>
</tbody>
</table>

* based on the markers msp1F3 and MS16
Table 2: Diversity of 15 *P. vivax* markers in PNG

<table>
<thead>
<tr>
<th></th>
<th>no. of pos. samples</th>
<th>no. of alleles</th>
<th>observed size range (bp)</th>
<th>expected heterozygosity $H_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS1</td>
<td>237</td>
<td>9</td>
<td>221 - 251</td>
<td>0.692</td>
</tr>
<tr>
<td>MS2</td>
<td>228</td>
<td>24</td>
<td>171 - 379</td>
<td>0.913</td>
</tr>
<tr>
<td>MS4</td>
<td>206</td>
<td>13</td>
<td>155 - 227</td>
<td>0.664</td>
</tr>
<tr>
<td>MS5</td>
<td>213</td>
<td>14</td>
<td>154 - 196</td>
<td>0.866</td>
</tr>
<tr>
<td>MS6</td>
<td>221</td>
<td>10</td>
<td>211 - 253</td>
<td>0.840</td>
</tr>
<tr>
<td>MS7</td>
<td>214</td>
<td>15</td>
<td>141 - 243</td>
<td>0.774</td>
</tr>
<tr>
<td>MS8</td>
<td>225</td>
<td>39</td>
<td>181 - 334</td>
<td>0.965</td>
</tr>
<tr>
<td>MS9</td>
<td>231</td>
<td>8</td>
<td>152 - 173</td>
<td>0.807</td>
</tr>
<tr>
<td>MS10</td>
<td>231</td>
<td>20</td>
<td>156 - 213</td>
<td>0.898</td>
</tr>
<tr>
<td>MS12</td>
<td>237</td>
<td>9</td>
<td>167 - 233</td>
<td>0.666</td>
</tr>
<tr>
<td>MS15</td>
<td>231</td>
<td>20</td>
<td>231 - 291</td>
<td>0.884</td>
</tr>
<tr>
<td>MS16</td>
<td>241</td>
<td>61</td>
<td>194 - 452</td>
<td>0.977</td>
</tr>
<tr>
<td>MS20</td>
<td>240</td>
<td>28</td>
<td>158 - 251</td>
<td>0.927</td>
</tr>
<tr>
<td>Pv3.27</td>
<td>220</td>
<td>33</td>
<td>184 - 460</td>
<td>0.934</td>
</tr>
<tr>
<td>msp1F3</td>
<td>236</td>
<td>28</td>
<td>237 - 372</td>
<td>0.849</td>
</tr>
</tbody>
</table>
Table 3: Genetic differentiation of *P. vivax* populations in Papua New Guinea

<table>
<thead>
<tr>
<th></th>
<th>Kunjingini</th>
<th>Alexishafen</th>
<th>Sigimaru</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ilaita</td>
<td>0.0012</td>
<td>0.0091</td>
<td>0.0247*</td>
</tr>
<tr>
<td>Kunjingini</td>
<td>0.0052</td>
<td>0.0207</td>
<td></td>
</tr>
<tr>
<td>Alexishafen</td>
<td>0.0252</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilaita</td>
<td>0.0053</td>
<td>0.0332*</td>
<td>0.0299*</td>
</tr>
<tr>
<td>Kunjingini</td>
<td>0.0489</td>
<td>0.0267</td>
<td></td>
</tr>
<tr>
<td>Alexishafen</td>
<td>0.0034</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilaita</td>
<td>0.0015</td>
<td>0.0107*</td>
<td>0.0251*</td>
</tr>
<tr>
<td>Kunjingini</td>
<td>0.0082</td>
<td>0.0211</td>
<td></td>
</tr>
<tr>
<td>Alexishafen</td>
<td>0.0238</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*F*<sub>ST</sub> values for inter-site comparisons are given.

(A) results obtained from 14 microsatellite markers

(B) results obtained from marker *msp1F3* only

(C) results obtained from 14 microsatellite markers and *msp1F3*

* significant at the Bonferroni-corrected 5% level (*P*<0.00833)
Table 4: Linkage disequilibrium between 8 markers (MS1, MS2, MS8, MS9, MS10, MS12, MS15, MS20), determined by LIAN software

<table>
<thead>
<tr>
<th>Location</th>
<th>No. of samples</th>
<th>$I^2_A$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ilaita</td>
<td>67</td>
<td>0.0075</td>
<td>0.0613</td>
</tr>
<tr>
<td>Kunjingini</td>
<td>22</td>
<td>0.0271</td>
<td>0.0255</td>
</tr>
<tr>
<td>Alexishafen</td>
<td>31</td>
<td>0.0121</td>
<td>0.13</td>
</tr>
<tr>
<td>Sigimaru</td>
<td>22</td>
<td>-0.0086</td>
<td>0.744</td>
</tr>
</tbody>
</table>

All       | 142            | 0.0052 | 0.023 |

Supplementary table S1: Primer sequences

<table>
<thead>
<tr>
<th>Marker</th>
<th>Primer name</th>
<th>Sequence 5' - 3'</th>
<th>Fluorescent dye</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS1</td>
<td>primary forward</td>
<td>CATCTCGACATGTCGACGTAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nested forward</td>
<td>TCAACTGTTGGAAGGGCAAT</td>
<td>6FAM</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttTGCTGCGTTTTTTGGTTCTG</td>
<td></td>
</tr>
<tr>
<td>MS2</td>
<td>primary forward</td>
<td>AGCACGACCAACAAGAGAGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nested forward</td>
<td>GAGCTAGCCAAGGGTTCAACA</td>
<td>VIC</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttTGGGGAGAGACTCCCTTTTC</td>
<td></td>
</tr>
<tr>
<td>MS4</td>
<td>primary forward</td>
<td>TAAACCAATGGGTAACGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nested forward</td>
<td>CGATTTACTGTTGACGCTGAA</td>
<td>6FAM</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttCAAAGGAACATGCTCGATGA</td>
<td></td>
</tr>
<tr>
<td>MS5</td>
<td>primary forward</td>
<td>TTCGGCTGGTTTCCAATTAGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nested forward</td>
<td>CGTCCTCTATCGCGTACACA</td>
<td>NED</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttAAAGGGAGAGGAGCGAAAAC</td>
<td></td>
</tr>
<tr>
<td>Primer</td>
<td>Forward Sequence</td>
<td>Nested Forward Sequence</td>
<td>Reverse Sequence</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------</td>
<td>------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>MS6</td>
<td>primary forward</td>
<td>GAGCTGCTGCTTCTATTTTGGG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttGGAGGACATCAACGGGATT</td>
<td></td>
</tr>
<tr>
<td>MS7</td>
<td>primary forward</td>
<td>ACATCAAAGCAAGAAGAGAGGG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttAGGGTCTTCAGCGTGTTGTTT</td>
<td></td>
</tr>
<tr>
<td>MS8</td>
<td>primary forward</td>
<td>AAACGTAAAACCTTTGGCGG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttAGGCCCTTTGCGTTT</td>
<td></td>
</tr>
<tr>
<td>MS9</td>
<td>primary forward</td>
<td>TGAATTTCATCCCCATTTGCCG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttGAAGCTGCCCCTATGTAAT</td>
<td></td>
</tr>
<tr>
<td>MS10</td>
<td>primary forward</td>
<td>AGGACCAAACGGAGGACATG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttTCCTTCAGGGAGACTT</td>
<td></td>
</tr>
<tr>
<td>MS12</td>
<td>primary forward</td>
<td>AACGTTCCTTGCCCACCTTG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttCTGCTTGGGGCTGCT</td>
<td></td>
</tr>
<tr>
<td>MS15</td>
<td>primary forward</td>
<td>CGCACTCTTCATCTCTCAGC</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>ctgtcttCGGCCAGATGAAAGGATAA</td>
<td></td>
</tr>
<tr>
<td>MS16</td>
<td>primary forward</td>
<td>TTCCTGATGACAAATTTGACCG</td>
<td>nested forward</td>
</tr>
<tr>
<td></td>
<td>primary reverse</td>
<td>TCTTTCATTTTGAGCATCGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nested forward</td>
<td>CTTGTTGTTGGTTGATGGTG</td>
<td></td>
</tr>
</tbody>
</table>

*Plasmodium vivax* Populations in Papua New Guinea
nested reverse  ctgcttAGTACGTCAACCATGTGGGTAG
MS20 primary forward  CAAGGTGCAGATGGAAGATTGG
nested forward  GCACAACAAATGCAAGATCC  VIC
reverse  ctgcttGTGGCAGTGGCTCATCTTCT
Pv3.27 fwd primary  TTTTTCAACTTGCTGCCCCCTG
fwd nested3)  GGACATTCCAAATGTATGTGCAGTCG  6-FAM
reverse  ctgcttCGTCATCGTCATTGCTCTGGAG
msp1F3 primary forward  GGAGAACATAAGCTACCTGTCC
primary reverse  GTTGTTACTTTGCTCTCTCCTCC
nested forward  CAAGCCTACCAAGAATTGATCCCCAA  VIC
nested reverse  ctgcttATTACTTTTGTAGTCTCCTCGGCGTAGTCC

In case of listing only one reverse primer sequence, a semi-nested protocol was used with identical reverse primers for both primary and nested PCR. A 7-pb tail promoting addition of a terminal adenine to amplification product is given in lower-case letters.

a primer originally described in (Karunaweera, 2006).
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

# A High Force of Infection Drives the Rapid Natural Acquisition of Immunity to *Plasmodium vivax* in Papua New Guinean Children

Cristian Koepfli\(^1\), Katie Benton\(^2,3\), Enmoore Lin\(^1\), Benson Kiniboro\(^4\), Peter A Zimmerman\(^5\), Terry Speed\(^6\), Peter Slba\(^4\), Thomas A. Smith\(^1\), Ingrid Felger\(^1\), Ivo Mueller\(^2,4,6\)

\(^\circ\) corresponding authors: ivomueller@fastmail.fm, ingrid.felger@unibas.ch

\(^1\) Swiss Tropical and Public Health Institute, Basel, Switzerland

\(^2\) Walter & Eliza Hall Institute, Parkville, Australia

\(^3\) University of California, Berkeley, CA, USA

\(^4\) Papua New Guinea Institute of Medical Research, Papua New Guinea

\(^5\) Center for Global Health and Diseases, Case Western Reserve University, Cleveland, OH, USA

\(^6\) Barcelona Centre for International Health Research, Barcelona, Spain.

This chapter will be submitted for publication after further revisions.
Abstract

The burden of disease caused by *Plasmodium vivax* is increasingly recognized. Morbidity and mortality are high especially in very young children. In areas where *P. vivax* and *P. falciparum* are co-endemic, age trends in incidence of clinical disease and prevalence differ significantly, with *P. vivax* peaking in younger children compared to *P. falciparum*. This suggests differences in the rate of acquisition of immunity. As natural immunity to malaria is acquired over an extended period of time and requires repeated infection, it has been suggested that the number of infections over time drives the speed of immunization.

We have estimated the molecular force of infection (molFOI, i.e. the number of times an individual is infected per year) of *P. vivax* in a cohort of 264 children in Papua New Guinea. Children were sampled bi-monthly over a period of 16 months in an area with high prevalence of *P. falciparum* and *P. vivax*. In addition passive case detection was maintained. *P. vivax* incidence decreased throughout the age group, but *P. falciparum* incidence peaked at 3.5 years. All *P. vivax* clones were genotyped using two highly polymorphic markers. The molFOI was estimated to 14.6 *P. vivax* clones per year and did not change with age. This contrasts earlier findings of a much lower molFOI in *P. falciparum*. We conclude that the different numbers of *P. vivax* and *P. falciparum* clones that infect children in early childhood lead to different speed in acquisition of immunity, and thus causes different age trends in incidence.
Introduction

People who live in malaria endemic areas achieve immunity to disease after some years of exposure (Koch, 1900). The immunization is gradual, and the speed of acquisition of immunity depends on transmission intensity. While incidence and parasitemia decline fast under intense transmission, partial immunity will build up over several decades (if at all) in areas of low transmission. The rate to immune acquisition also seems to differ between parasite species, with immunity to P. vivax acquired faster than immunity to P. falciparum. In numerous field studies conducted in areas co-endemic for these species the burden of P. vivax infections and disease was found to peak at a younger age than that due to P. falciparum (Balfour, 1935; Earle, 1939; Gunewardena et al., 1994; Maitland et al., 1996; Mendis et al., 2001; Michon et al., 2007; Mueller et al., 2009; Phimpraphi et al., 2008). Similarly, during the course of malaria therapy, were neurosyphilis patients were artificially infected with malaria, P. vivax densities remained below the pyrogenic threshold after only a few doses, while each of several infections with P. falciparum induced fever (Ciucă et al., 1934; Collins and Jeffery, 1999; Collins et al., 2004).

Malaria infections consist of different clones (or strains) that infect individuals successively or in parallel. Strains can be characterized by their surface antigens (Anders et al., 1993). Their diversity is well documented for a number of P. falciparum antigens as well as for some P. vivax antigens. Examples include the circumsporozoite protein (csp) (Henry-Halldin et al., 2011; Imwong et al., 2005), different merozoite surface proteins (msp) (Fenton et al., 1991; Koepflı et al., 2008; Raj et al., 2004), the apical membrane protein 1 (ama1) (Cortes et al., 2003), or the P. vivax Duffy binding protein (dbp) (Cole-Tobian and King, 2003). The large number of antigen variants allows parasites to escape the immune system, as newly infecting strains likely express different alleles as previous ones, and thus are not recognized by existing antibodies (Bull and Marsh, 2002; Osier et al., 2008). As a consequence, immunity is thought to be largely clone specific, providing less protection against disease caused by heterologous clones (Ciucă et al., 1934; Cole-Tobian et al., 2009; Collins et al., 2004; Genton et al., 2002; Jeffery, 1966; Lyon et al., 2008; Smith et al., 1999).

The differences in acquisition to immunity against P. falciparum and P. vivax infection observed in population-based studies could be caused by several factors. Biological differences between parasites might account for differences in immunization, such as a more limited antigenic repertoire of P. vivax compared to the one of P. falciparum. P. falciparum expresses sequentially up to 60 variants of the Erythrocyte Membrane Protein 1 (PFEMP-1) in a mutually exclusive manner (Voss et al., 2006) and acquisition of antibodies against broad repertoire of PFEMP-1 variants is thought to be key to developing anti-falciparum immunity (Bull et al., 1998; Dodoo et al., 2001; Giha et al., 2000). This is not the case for the P. vivax multi-gene vir genes, the products of which are also expressed on the surface of the host.
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

...reticulocyte (Fernandez-Becerra et al., 2005), but where multiple genes are concurrently expressed and whose function is yet to be fully understood (Bernabeu et al., 2011). Similarly, *P. vivax* relies almost fully on the Duffy receptor for invasion into the red blood cell (RBC) (Menard et al., 2010; Miller et al., 1976), while *P. falciparum* is able to invade through a variety of partially redundant pathways (Cowman and Crabb, 2006). Antibodies targeting these essential ligands can block invasion in-vitro (Chen et al., 2011; Grimberg et al., 2007; Jiang et al., 2011; Tham et al., 2009; Triglia et al., 2011) and are associated with significant protection in-vivo (King et al., 2008; Reiling et al., 2010; Richards et al., 2010). As a broad inhibitory response against multiple ligands may be required for effective *P. falciparum* immunity (Persson et al., 2008), different rates of immune acquisition might reflect the broader spectrum of *P. falciparum* invasion pathways.

Immunity to malaria is poorly understood (Doolan et al., 2009; Langhorne et al., 2008). Both the total number of infections as well as the number of different clones seem to influence speed of immune acquisition (Jeffery, 1966). According to this hypothesis individuals acquire complete resistance against parasites once they had been infected with all parasite strains circulating in a population (Aguas et al., 2012; Weisman et al., 2001). As alternative to differences between parasite species in antigenic diversity or invasion, different numbers of transmitted clones could explain contrasting trends in acquisition of immunity. High numbers of infections per time can be the results of frequent bites of infected mosquitoes, or human hosts can acquire different parasite clones through a single bite of a mosquito carrying genetically diverse sporozoites (Annan et al., 2007; Mzilahowa et al., 2007).

The number of infections acquired over time largely depends on transmission intensity. In the case of *P. vivax* relapsing hypnozoites add to the frequency of blood stage parasitemia. Relapsing parasites are mostly genetically different from parasites detected during last acute blood stage infection (Chen et al., 2007; Imwong et al., 2007; Restrepo et al., 2011), either because the relapsing clones might correspond to earlier (acute) blood stage infections, or because the primary and relapsing infections are genetically different clones transmitted in a single mosquito bite. Relapses could therefore boost existing immune responses, and contribute to broadening the immune repertoire, thereby contributing to the natural acquisition of immunity to *P. vivax*.

To determine the molecular force of infection (molFOI, the number of infections acquired per time) under natural conditions genotyping of all parasite clones of both *P. falciparum* and *P. vivax* in an individual over an extended period of time is required. We have followed 264 children aged 1 to 3 years at enrolment over 16 months in an area of comparable high endemicity of *P. falciparum* and *P. vivax* (Lin et al., 2010). As in earlier studies significant difference in age specific burden of *P. vivax* and *P. falciparum* malaria were observed during...
follow-up: While *P. vivax* incidence decreased throughout the age group, *P. falciparum* incidence increased from 1 to 3.5 years with little change thereafter (Lin et al., 2010).

In order to assess the relationship between (individual) exposure and risk of malaria, we genotyped all *P. falciparum* and *P. vivax* infections detected over 16 months (Koepfli et al., 2011a; Mueller et al., submitted). By tracking *msp2* PCR-fragments in consecutive infections we found that on average children acquired 5.9 new *P. falciparum* infections/child/year (Mueller et al., submitted). The molFOI was found to increase significant with age, was strongly reduced in children using insecticide treated bed nets ITNs, and showed significant spatial and temporal variation. Adjusting for individual differences in molFOI completely explained the observed age trends and effect of ITN as well as ~50% of the seasonal and almost all of the spatial variation in the incidence of *P. falciparum* malaria.

In an earlier analysis, we have presented diversity and multiplicity of the *P. vivax* clones in the cohort (Koepfli et al., 2011a). We now present a more in-depth analysis of *P. vivax* genotyping data with the aim to estimate molFOI for *P. vivax*, to determine its association with age, ITN use and season and investigate its relationship to the age-specific burden of *P. vivax* malaria in this cohort.
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

**Methods**

**Ethics statement**

Informed written consent was sought from all parents or guardians prior to recruitment of each child. Scientific approval and ethical clearance for the study was obtained from the Medical Research and Advisory Committee (MRAC 05.19) of the Ministry of Health in PNG and from the Ethikkommission beider Basel in Switzerland (no 03/06).

**Field survey and patients**

This study was conducted in Ilaita, a rural area near Maprik, East Sepik Province, Papua New Guinea. A detailed description of the study was given elsewhere (Lin et al., 2010). Briefly, 264 study participants were enrolled at an age of 10 to 38 months between March and September 2006. Following enrolment children were followed actively every 2 weeks for malaria morbidity for a period of up to 16 months (until July 2007). In addition, children were actively checked every 8 to 9 weeks for the presence of malarial infections. Except for the first and last round of active case detection, two consecutive blood samples were collected by finger prick at intervals of 24 h from each study participant at each follow-up visit. An individual thus contributed up to 16 samples, 14 of which were paired samples collected 24 h apart. A passive case detection system was maintained at the local health centre and aid post throughout the entire study period. At each episode of febrile illness a blood sample collected from all participants, a rapid diagnostic test (RDT) was performed and haemoglobin measured using Hemacue (Angholm, Sweden). Antimalarial treatment with Coartem® (Novartis, Switzerland) was administered upon a positive RDT or if haemoglobin levels were <7.5 g/dl. In children with negative RDT, blood slides were read within 24 h and microscopy positive children treated with Coartem®.

**Laboratory procedures**

All finger prick blood samples were separated into plasma and cells. DNA was extracted from cell pellets using QIAamp® 96 DNA Blood Kit (Qiagen, Australia) according to the manufacturer’s instructions.

For genotyping individual *P. vivax* clones, the molecular markers msp1F3 and MS16 were typed using capillary electrophoresis for highly precise fragment sizing, as required for longitudinal follow up of parasite clones. Both markers proved to be highly polymorphic in the
cohort with a virtual heterozygosity of 97.8% for MS16 and 88.1% for msp1F3. Details of the genotyping technique have been described elsewhere (Koepfli et al., 2011a).

**Data analysis**

In a previous analysis of the samples collected 24 hours apart we have shown that in a single blood sample often not all alleles of a marker are detected. Twenty-one percent of all msp1F3 alleles and 28% of all MS16 clones were missed on a single day (Koepfli et al., 2011b). For all analyses we therefore used the combined genotyping data from both day 1 and day 2, except for enrolment and final time point where only 1 sample was available. Consequently, multiplicity of infection (MOI) in each host was defined as the number of different msp1F3 and MS16 clones, detected in the day 1 sample combined (where available) with the day 2 sample. For all discrepancies in typing results obtained by these 2 markers, the higher MOI was used, because very low abundant templates cannot be consistently amplified.

The force of new *P. vivax* infections (molFOI) was calculated as follows: For each interval of 8 to 9 weeks the number of newly acquired infections was determined. New clones were defined as msp1F3 and MS16 alleles not present in any preceding interval. An interval started on the first day after a regular cross-sectional visit; it thus included all samples collected during passive case detection over two months plus the samples collected at the end of the interval. For each interval, children were considered at risk of acquiring new infection from the 1st day after the last blood sample from the preceding interval was taken. In line with the pharmacokinetic properties of the drugs (Hombhanje et al., 2005; Salman et al., 2011a; Salman et al., 2011b), children were not considered at risk for 2 weeks after treatment with Coartem® and 4 weeks after treatment with amodiaquine (AQ) plus sulphadoxine-pyramethamine (SP). The force of infection for each child and interval was subsequently converted into the number of new clones acquired per year-at-risk.

As in the earlier analyses of *P. falciparum* molFOI (Mueller et al., submitted), longitudinal random-effects models, or generalized linear mixed models (GLMMs) were used for analyses of prevalence, multiplicity and force of infection well as for incidence of *P. vivax* episodes. These models were chosen because they allowed the fixed effects, or subject-specific covariates, to be specified separately from the random effects, i.e. repeated measurements from the same child over time and unmeasured village factors. Furthermore, unlike other models, the random-effects model allowed for decomposition of the random effects into between-village and within-village variation.

We fitted a Poisson-Gaussian random-effects model with a log link function to relate the fixed and random effects to the number of clinical episodes experienced during a two month
interval (defined as febrile illness plus any \( P.\) \( vivax\) density or \( P.\) \( vivax\) >500 parasites/\( \mu l\)). Similarly, for the prevalence of \( P.\) \( vivax\) infection by PCR (\( msp1F3\) and/or MS16), we fitted a Binomial-Gaussian random effects model with a logit link function. Covariates were selected based on earlier analyses of the same data (Lin et al., 2010). Exploratory data analysis guided the specification of the model when the relationship between the covariates and the outcome did not appear to be linear, such as with age, FOI rate and season. Seasonality was characterized by two readily interpretable parameters: the amplitude, which is traditionally half the range between the peak(s) and trough(s), and the phase, which is the location of the first zero crossing in a cycle relative to the origin in time. When included directly, these two parameters define a non-linear regression term. For computational convenience, they are usually replaced by \( \sin\) and \( \cos\) terms with fixed phase, which enter linearly, and whose coefficients (amplitudes) are thus more easily estimated together with those of other terms in the linear part of the model. Their coefficients are then combined via a standard trigonometric identity to give the original amplitude and phase parameters (Diggle, 2002). For all outcomes except prevalence, an offset was fit to adjust for years at risk. Estimation of these models was done using the LME4 package in R version 2.12.

All point estimates except those for seasonal effects were obtained by smoothing raw raw data using generalized additive models. Point estimates for seasonal peaks and trough were obtained from the GLMM-models.
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

**Results**

A total of 264 children aged between 0.9 – 3.2 yrs of age (median 1.7, Inter-quartile range (IQR): [1.3, 2.4]) were enrolled and followed for up to 69 weeks. 248 of 264 (93.9%) children were retained until the end of the study with between 96.0% - 100.0% of children see at each scheduled 2-monthly survey. Over the entire follow-up the age of children thus ranged from 0.9 to 4.5 years of age. A more detailed description of the cohort was given in (Lin et al., 2010).

**Prevalence by PCR and multiplicity of *P. vivax* infections**

All children except 5 had at least one *msp1*F3 or MS16 PCR positive sample during the 16 months follow-up. 51.6% and 52.7% of all samples collected were positive for *msp1*F3 or MS16, respectively. A prevalence of 54.8% was obtained for the two markers combined.

*P. vivax* prevalence increased significantly (OR 1.44, IC[95][1.17, 1.76], *P* < 0.001, Table 1) with age (at start of interval) raising from 49% in children <1 yrs to 65 - 70% in children >3yrs. (Figure 1). Personal ITN use (OR 0.49, IC[95][0.33, 0.73], *P* < 0.001) and antimalarial treatment < 4 weeks prior (OR 0.46, IC[95][0.37, 0.57], *P* < 0.001) were both associated with a strong reduction in *P. vivax* prevalence. When smoothing raw prevalence using non-parametric regression splines, *P. vivax* prevalence was found to increase during the first 30 weeks of follow-up with little change thereafter. Nevertheless, in the multivariate GLMM model (Table 1, Figure 2) significant seasonal (*P* = 0.002) and between year variation (OR 1.40, IC[95][1.02, 1.92], *P* = 0.03) in prevalence was detected. Based on sin/cos modelling, *P. vivax* prevalence was found to peak in week 38 (i.e. mid September / late dry season) and was lowest in week 12 (late March / late rainy season, Figure 2). Despite adjusting for these factors, prevalence of infection still varied significantly between villages (*P* = 0.008) and among children with villages (*P* < 0.001).

The patterns observed for multiplicity of infection paralleled those observed for prevalence. MOI rose significantly throughout the age range (Table 1, *P* < 0.001), with mean MOI increasing from about 2.5 in 1 year old children to 3.6 at in 4 years old children (Figure 1). Average ITN use was associated with a significant reduction in MOI from 3.3 in non-users to 3.0 in children always sleeping under an ITN (Figure 1, Table 1, *P* = 0.03). Despite small seasonal amplitude, average MOI also showed significant variation over time (Figure 1, Table 2, *P* = 0.002) with MOI estimates to peak in week 40 (early Oct / late dry season). There was significant variation in MOI among children living in the same villages (Table 1, *P* < 0.001) but not between villages (*P* = 0.1).
Force of infection in 8-9 week intervals

Excluding any period with residual drug levels from the time at risk, the average \( mo_{\text{FOI}} \) was 14.6 new \( P. vivax \) infections per child per year.

\( mo_{\text{FOI}} \) showed a very pronounced seasonality (Figure 2, Table 1, \( P < 0.0001 \)). Based on GLMM modeling \( mo_{\text{FOI}} \) was estimated to peak in early January (week 53, 17 clones/child/year) and was lowest in early July (week 27, 11 clones/child/year, Figure 2). \( mo_{\text{FOI}} \) was also found to be significant lower in 2007 compared to 2006 (IRR 0.84, IC\(_{95}\)[0.77, 0.92], \( P = 0.0002 \), Table 1, Figure 1). Regular ITN use was associated with a significant reduction in acquisition of new clones (IRR 0.66, IC\(_{95}\)[0.56, 0.77], \( P < 0.0001 \), Table 1), whereas children with antimalarial treatment in 4 weeks prior to an interval had a higher \( mo_{\text{FOI}} \) (IRR 1.24, IC\(_{95}\)[1.15, 1.33], \( P < 0.0001 \), Table 1). \( mo_{\text{FOI}} \) did however not vary significant with age (Figure 1, \( P \)-value (GLMM) = 0.6). As for MOI, significant variation in \( mo_{\text{FOI}} \) was detected among children in a village (\( P < 0.001 \)) but not between villages (\( P = 0.3 \), Table 1). A summary of the GLMM estimates for \( mo_{\text{FOI}} \) (Fit stats) can be found in Table 1.

Predictors of clinical \( P. vivax \) illness

Over the 69 weeks of follow-up, a total of 1134 febrile episodes (incidence rate (IR) 4.60 / child / yr) with parasitaemia (by light microscopy) were observed (Lin et al., 2010). Overall, \( P. vivax \) was the 2\(^{\text{nd}} \) most common cause of malarial illness (any density: 605 (IR 2.46), \( Pv > 500/\mu l \)) after \( P. falciparum \) (any density: 630 (IR 2.56), \( Pf > 2,500/\mu l \): 472 (IR 1.92)). \( P. malariae \) and \( P. ovale \) episodes were rare. All further analyses are done using only the more specific definition of \( P. vivax \) malaria (i.e. febrile illness plus parasitaemia > 500/\( \mu l \)) (Mueller et al 2009, Malaria J).

As in earlier analyses (Lin et al., 2010), age and season were significant predictors of clinical episodes of \( P. vivax \) malaria (Table 2). The incidence of \( P. vivax \) malaria decreased log-linearly with age (Figure 1,Table 2, \( P < 0.0001 \)) from 2.9 episodes / child / year in children 1 year old children to a minimum of 0.8 episodes at 3.5 years of age. The incidence of \( P. vivax \) malaria peaked at the beginning of rainy season (week 49, early December), and was lowest in the early dry season (week 23, early June, Figure 2). ITN use on the other hand showed no significant association with incidence of \( P. vivax \) malaria. The incidence of \( P. vivax \) malaria varied significantly among villages (\( P < 0.0001 \)) and among children living in the same village (\( P < 0.0001 \)).

When \( mo_{\text{FOI}} \) was added to the model (fitted as the rate of new clones acquired per year at risk), \( mo_{\text{FOI}} \) was highly significantly associated with an increase in the incidence of \( P. vivax \) malaria (Figure 3, \( P < 0.0001 \)). Adjusting for \( mo_{\text{FOI}} \) accounted for a substantial amount of
seasonal variation and resulted in a 45% decrease in season amplitude (Table 2, Figure 4) but contrary to what was observed for *P. falciparum* (Mueller et al., submitted) it had no effect at all on other factor's association with incidence of *P. vivax* malaria.
Discussion

As in other regions where *P. vivax* and *P. falciparum* are co-endemic (Gunewardena et al., 1994; Maitland et al., 1996; Phimpraphi et al., 2008), on the island of New Guinea the burden of *P. vivax* infections and illness peaks in younger children compared to *P. falciparum* (Kasehagen et al., 2006; Lin et al., 2010; Michon et al., 2007; Mueller et al., 2009). The increasingly appreciated burden of *P. vivax* associated severe morbidity and even mortality (Genton et al., 2008; Manning et al., in press; Poespoprodjo et al., 2009; Tjitra et al., 2008) is concentrated in the very youngest children (<3 yrs). This indicates a fast acquisition of immunity to *P. vivax* in individuals with life-long exposure to both *P. falciparum* and *P. vivax* infections. The reason for the faster acquisition of immunity to *P. vivax* compared to that against *P. vivax* has been unclear, but as immunity to malaria builds up gradually and is thought be strain specific (Ciucu et al., 1934; Jeffery, 1966), it is likely that the number of distinct infections acquired over an individuals lifetime is a major driving force for acquired immunity.

Genotyping all PCR positive infections that occurred in individual children over 16 months of follow-up, has revealed that children aged 0.9 - 4.5 years children indeed acquire more than twice as many genetically distinct *P. vivax* (14.6 clones/child/year) compared to *P. falciparum* blood-stage infections (5.9 clones/child/year) (Mueller et al., submitted). In parallel with this higher molecular force of infection (molFOI), we also observed an increased prevalence and multiplicity of *P. vivax* infections compared to *P. falciparum*.

Several reasons could account for the markedly higher *P. vivax* molFOI compared to that of *P. falciparum*. Genotyping techniques used for both species were comparable and experiments were performed in the same laboratory. Typing was based on length polymorphism in highly diverse marker genes as evidenced by their respective expected heterozygosity, $H_e$. *P. falciparum* was typed using msp2 ($H_e=0.933$) (Schoepflin et al., 2009), and MS16 ($H_e=0.978$) and msp1F3 ($H_e=0.881$) were used for *P. vivax* typing (Koepfli et al., 2011a). The detectability (the proportion of clones detected in both of two bleeds 24 h apart) of clones was 79% for Pfmsp2 and 61% and 73% for PfMS16 and Pfmsp1F3 respectively (Koepfli et al., 2011b). Although *P. vivax* genotyping was based on two loci (and MOI and FOI were determined from maximal number of alleles per sample observed by any marker), the overall sensitivity seems comparable between single locus Pfmsp2 genotyping and two loci *P. vivax* typing, when considering the similar values for $H_e$ and detectability.

For this study we have counted all clones detected in regular 2 monthly double bleeds, as well as all samples collected during febrile illness episodes detected during active and passive case detection. Very low-density clones are likely missed, since genotyping studies based on next generation sequencing technology have suggested that multiplicity of infection might be
several times higher than when detected by classical PCR (Campino et al., 2011; Juliano et al., 2010). From this and from our own previous studies on clone detectability, we assume that our estimates of molFOI and MOI of both species are likely underestimates of the true burden and complexity of Plasmodium spp. infection in our cohort. Nevertheless, sensitivity any genotyping likely is imperfect. As long as comparable methodologies are used, the validity of the principal conclusions of epidemiological analyses presented here will not be affected.

In PNG P. vivax and P. falciparum are transmitted by the same mosquito vectors. While no entomological studies were conducted concurrently with the cohort, earlier studies in different PNG lowlands population reported comparable sporozoite rates for P. falciparum and for P. vivax in the local vector populations (Benet et al., 2004; Hii et al., 2001; Michon et al., 2007). Although the relative prevalence of P. vivax sporozoites may rise after ITN introduction (Bockarie and Dagoro, 2006), it is nevertheless unlikely that the differences in molFOI could be explained solely by different EIRs (entomological inoculation rates) for the two species. However, 75% of all P. vivax positive individuals in our cohort carried multiple clone infections, but only 33% of all P. falciparum positive children (Koepfli et al., 2011a) and Sonja Schoepflin, unpublished results). This increases the chance that mosquitoes will concurrently ingest gametocytes from more than one P. vivax clones, which in turn increases the likelihood that individuals become infected with multiple P. vivax clones by a single mosquito bite.

Finally, P. vivax parasites detected in the blood stream do not always derive from a recent mosquito bite, but can also be the result of relapsing hypnozoites from an earlier bite. New Guinea P. vivax strains were found to relapse very rapidly and 63% of 1 - 5 yr old children had a recurrent parasitemia within 6 weeks after treatment with Coartem (Barnadas et al., 2011). Two thirds of these recurrent infections (all post 28 days) were of a different genotype. Although genotyping cannot differentiate between true new infections and relapses of a different genotype as seen in the last infection, relapsing P. vivax infections likely contribute significantly to the higher molFOI. In summary, molFOI of P.falciparum and P. vivax measure different epidemiological parameters: while molFOI of P. falciparum is directly linked to transmission by mosquitoes, molFOI of P. vivax measures both, mosquito-borne transmission intensity and frequency and genetic complexity of relapsing parasites.

Whereas P. falciparum molFOI was found to be strongly age dependent and to increase from 3 clones per year in children <1 year to 8 clones per year in those >4 (Mueller et al., submitted), P. vivax molFOI does not change significantly with age. We thus estimate that during the first years of life, the children in our cohort were infected with three to four times more genetically distinct P. vivax than P. falciparum clones. This large difference in molFOI, and thus the overall genetic diversity to which children are exposed in early childhood, has dramatic impacts on the age specific burden with P. vivax and P. falciparum malaria. Although a higher molFOI was associated with a significant increase in incidence of clinical disease in both species (Mueller
et al., submitted), the incidence of *P. vivax* decreased significantly across the entire age range, while the incidence of *P. falciparum* malaria increased with the age. This suggests that between the age of 1 to 4 years children have been exposed to a sufficient number of *P. vivax* infections to acquire high level of clinical immunity to this parasite.

Such an exposure may even be sufficient for children to acquire a certain degree of sterile immunity against *P. vivax* infection. Mosquito biting rates are known to increase proportional to host body size (Port et al., 1980). Therefore, as children grow in size, their exposure to malarial infection also increases, as confirmed by the strong age-dependence of *P. falciparum* molFOI. *P. vivax* molFOI did not change with age. This suggests that in older children some of the transmitted *P. vivax* sporozoites do not succeed to establish detectable blood stage infections.

In conclusion, we propose that the number of infections acquired over time is a mayor determinant of immunity against clinical malaria. Although we can not rule out that generic differences exist between *P. falciparum* and *P. vivax* with respect to the nature of immune responses elicited and immune acquisition by their hosts, the observed differences in molFOI can largely explain the striking differences in incidence trends between the two species. The faster acquisition of immunity against *P. vivax* disease results in a marked decrease in incidence with increasing age, despite a constant number of blood stage infections.
References


The Force of *Plasmodium vivax* Infections in Papua New Guinean Children


The Force of \textit{Plasmodium vivax} Infections in Papua New Guinean Children


The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

Figure 1: Age and seasonal patterns of prevalence, multiplicity (MOI), force-of-infection (MOI FOI) and incidence of *P. vivax* malaria. Spline smoothed raw data and 95% confidence intervals.
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

**Figure 2**: Effect of ITN use on prevalence, multiplicity (MOI) and force (FOI) of *P. vivax* infection and incidence of clinical episodes.
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

Figure 3: Association between force-of-infection (molFOI) and incidence of *P. vivax* malaria. A) Spline smoothed raw data. B) Incidence rate ratio adjusted for seasonal difference from GLMM model. Predicted means and 95% confidence intervals.

Figure 4: Predicted associations of seasonality with *P. vivax* molecular epidemiology endpoints: A) seasonal patterns in prevalence, MOI and molFOI, B) seasonal effects on incidence of *P. vivax* (>500/µl) before (blue) and after adjusting for difference in molFOI (red).
Table 1: Parameter estimates GLMM models for prevalence, multiplicity of infection (MOI) and force of \textit{P. vivax} infections (\textit{molFOI}).

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Prevalence OR$^a$</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
<th>MOI RR$^b$</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
<th>FOI IRR$^c$</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.44</td>
<td>1.17</td>
<td>1.76</td>
<td>0.0006</td>
<td>1.15</td>
<td>1.09</td>
<td>1.22</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sin(week)</td>
<td>0.72</td>
<td>0.6</td>
<td>0.86</td>
<td>0.002</td>
<td>0.92</td>
<td>0.88</td>
<td>0.97</td>
<td>0.002</td>
<td>0.99</td>
<td>0.93</td>
<td>1.05</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Cos(week)</td>
<td>0.95</td>
<td>0.8</td>
<td>1.12</td>
<td></td>
<td>1.01</td>
<td>0.96</td>
<td>1.06</td>
<td>1.23</td>
<td>1.18</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average ITN use$^d$</td>
<td>0.49</td>
<td>0.33</td>
<td>0.73</td>
<td>0.0004</td>
<td>0.89</td>
<td>0.8</td>
<td>0.98</td>
<td>0.03</td>
<td>0.66</td>
<td>0.56</td>
<td>0.77</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Treated$^e$</td>
<td>0.46</td>
<td>0.37</td>
<td>0.57</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td>1.24</td>
<td>1.16</td>
<td>1.33</td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Year 2007</td>
<td>1.4</td>
<td>1.02</td>
<td>1.92</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td>0.84</td>
<td>0.77</td>
<td>0.92</td>
<td>0.0002</td>
<td></td>
</tr>
</tbody>
</table>

Random effects

| village       | 0.12              |     |     | 0.008 |             |     |     | 0.1  | 0.007     |     |     | 0.3  |
| child         | 0.88              |     |     | <.0001 |             |     |     | <.0001 | 0.2       |     |     | <.0001 |

Seasonal stats

| betaSin       | -                 |     |     | -     |             |     |     | -    | -          |     |     |     |
| betaCos       | 0.328504067       |     |     | 0.083381609 |             |     |     | 0.010050336 |
| Amplitude     | 0.33              |     |     | 0.08  |             |     |     | 0.21 |     |
| offset        | 11.71811035       |     |     | -     |             | 12.01702659 | - | 0.401479259 |
| peak week     | 38                |     |     | 40    |             |     |     | 1    |     |
| trough week   | 12                |     |     | 14    |             |     |     | 27   |     |
| Month of Peak | late September    |     |     | early October |             |     |     | early January |
| Month of Trough| late March       |     |     | early April |             |     |     | early July |

$^a$ OR: Odds ratio, LCL: lower confidence level, UCL: upper confidence level

$^b$ RR: Rate ratio

$^c$ IRR: incidence rate ratio, CI$_{95}$: 95% confidence interval.

$^d$ insecticide treated net use: 0% vs 100% use

$^e$ Treated with antimalarials within 28 days prior to start of interval.

$^f$ joint p-value for sin and cos
Table 2: Incidence of *P. vivax* malaria

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>PvAll IRR</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
<th>PvAll+ FOI</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
<th>Pv 500 IRR</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
<th>Pv 500+ FOI</th>
<th>LCL</th>
<th>UCL</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.72</td>
<td>0.63</td>
<td>0.83</td>
<td>&lt;.0001</td>
<td>0.68</td>
<td>0.6</td>
<td>0.78</td>
<td>&lt;.0001</td>
<td>0.55</td>
<td>0.46</td>
<td>0.67</td>
<td>&lt;.0001</td>
<td>0.52</td>
<td>0.44</td>
<td>0.62</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Sin (week)</td>
<td>0.88</td>
<td>0.78</td>
<td>0.99</td>
<td>&lt;.0001</td>
<td>0.92</td>
<td>0.82</td>
<td>1.03</td>
<td>&lt;.0001</td>
<td>0.89</td>
<td>0.77</td>
<td>1.03</td>
<td>&lt;.0001</td>
<td>0.94</td>
<td>0.81</td>
<td>1.09</td>
<td>0.01</td>
</tr>
<tr>
<td>Cos (week)</td>
<td>1.48</td>
<td>1.31</td>
<td>1.66</td>
<td>1.29</td>
<td>1.14</td>
<td>1.45</td>
<td>1.45</td>
<td>1.25</td>
<td>1.67</td>
<td>1.24</td>
<td>1.07</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOI^1/3</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.88</td>
<td>1.74</td>
<td>2.03</td>
<td>&lt;.0001</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.99</td>
<td>1.8</td>
<td>2.19</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Random effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>village</td>
<td>0.08</td>
<td>0.002</td>
<td>0.17</td>
<td>&lt;.0001</td>
<td>0.2</td>
<td>0.0001</td>
<td>0.35</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>child</td>
<td>0.33</td>
<td>&lt;.0001</td>
<td>0.11</td>
<td></td>
<td>0.003</td>
<td>0.56</td>
<td>&lt;.0001</td>
<td>0.28</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seasonal stats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>betaSin</td>
<td>0.128</td>
<td></td>
<td></td>
<td>-0.083</td>
<td>-0.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>betaCos</td>
<td>0.392</td>
<td></td>
<td></td>
<td>0.255</td>
<td>0.372</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude</td>
<td>0.41</td>
<td></td>
<td></td>
<td>0.27</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>offset</td>
<td>-</td>
<td>2.609</td>
<td>-2.619</td>
<td>-2.515</td>
<td>2.318</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>peak week</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trough week</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Month of Trough</td>
<td>early June</td>
<td>early June</td>
<td>early June</td>
<td>early June</td>
<td>early June</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fit stats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log likelihood</td>
<td>-843</td>
<td>-684</td>
<td>-669</td>
<td>-551</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC^b</td>
<td>1699</td>
<td>1382</td>
<td>1349</td>
<td>1115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a^ joint p-value for sin and cos  ^b^ Akaike Information Criterion
The Force of *Plasmodium vivax* Infections in Papua New Guinean Children

**Supplementary Table 1:** Parameter estimates for multiple random effects models predicting the number of incident clinical episodes of *P. vivax* malaria with any parasites density with and without adjustment for molFOI.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base model</th>
<th>molFOI adjusted</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR</td>
<td>CI 95</td>
<td>p - value</td>
<td>IRR</td>
<td>CI 95</td>
<td>p - value</td>
</tr>
<tr>
<td>Age (at interval)</td>
<td>0.72</td>
<td>[0.62, 0.83]</td>
<td>&lt;0.0001</td>
<td>0.69</td>
<td>[0.60, 0.78]</td>
<td>0.3</td>
</tr>
<tr>
<td>Sin (week)</td>
<td>0.88</td>
<td>[0.78, 0.99]</td>
<td>&lt;0.0001*</td>
<td>0.92</td>
<td>[0.82, 1.03]</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Cos (week)</td>
<td>1.48</td>
<td>[1.31, 1.66]</td>
<td></td>
<td>1.28</td>
<td>[1.14, 1.45]</td>
<td></td>
</tr>
<tr>
<td>FOI</td>
<td>-</td>
<td>-</td>
<td></td>
<td>1.89</td>
<td>[1.74, 2.04]</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Random effects:

- **Village**
  - Base model: 0.08, p < 0.01
  - molFOI adjusted: 0.14, <0.001

- **Child within village**
  - Base model: 0.33, p < 0.001
  - molFOI adjusted: 0.11, <0.01

Log likelihood: -843 vs -683

AIC: 1701 vs 1382

Seasonal effects

- **Amplitude**
  - Base model: 0.41, [0.37, 0.51]
  - molFOI adjusted: 0.26, [0.24, 0.37]

- **Peak (week)**
  - 49

- **Trough (week)**
  - 23

---

*a* IRR: incidence rate ratio, CI 95: 95% confidence interval. 
*b* Insecticide treated net use: 0% vs 100% use

*c* Force of infection (# new clone per year-at-risk), cub-root transformed.

*d* Akaike Information Criterion
General Discussion

The Complex Pattern of Malaria Infection and Disease

The epidemiology of malaria is the result of a complex interplay of the parasites, their human hosts and the mosquito vector. During its life cycle the parasite passes from the mosquito salivary glands through the human skin to the liver, and then invades red blood cells in continuous cycles. A small proportion of parasites transform into sexual forms, gametocytes, and once taken up during a mosquito blood meal they undergo recombination and migrate to the salivary glands, where the sporozoites start the cycle again (reviewed in Bousema and Drakeley, 2011; Prudencio et al., 2006).

All clinical symptoms are caused by the red blood cell cycle, while sporozoite invasion and the liver stage do not cause disease. *P. vivax* parasites can appear in the blood stream after relapses of hypnozoites, i.e. liver stages that stay dormant for an extended period of time after mosquito transmission. These relapsing parasites can induce clinical disease, however the onset and frequency of malaria depends not only on transmission intensity and relapses, but is also strongly influenced by acquired immunity, which develops in areas of intense transmission after a few years (Koch, 1900).

*P. vivax* and Malaria Elimination

Understanding the factors influencing transmission and immunization will become increasingly important in the upcoming era of malaria eradication. In this context, the specific biological characteristics of *P. vivax* are of particular relevance, as antimalarial interventions might not affect all Plasmodium species similarly. Interventions already led to a decrease of malaria in many places of the world. In several countries, however, this was accompanied by a shift in predominance from *P. falciparum* to *P. vivax*. Such results were obtained from the highlands of Papua New Guinea (Mueller et al., 2005), Colombia (Rodriguez et al., 2011), Cambodia, Laos or eastern Thailand (Cui et al., 2011).

The elimination of *P. vivax* is complicated by several specific characteristics of this parasite’s biology. Firstly, hypnozoites that relapse after an unpredictable period of time require long-term surveillance of patients after clearance of acute blood stage parasitemia. This effort is required even in regions of very low transmission and thus a low risk of re-infection. Secondly, *P. vivax* parasitemia is generally lower than that of *P. falciparum*. This might lead to underestimation of *P. vivax* infections by standard light microscopy. Thirdly, fast acquisition of
semi-immunity leads to a large number of asymptomatic *P. vivax* carriers, all potentially contributing to transmission. As a consequence of these discrepancies to *P. falciparum*, the need for *P. vivax* specific research for malaria elimination has been addressed (Feachem et al., 2010; Mueller et al., 2009), as well as the need for new drugs targeting hypnozoites (Wells et al., 2010).

Feasibility of elimination of malaria in a country depends on the epidemiological situation, including the level of transmission and the risk of reintroduction of malaria due to migration, as well as on the health infrastructure, the political situation and the commitment of the government towards malaria elimination. In a recent report, Tatem and coworkers have identified PNG among the countries with the poorest health infrastructure, when comparing all malaria endemic countries (Tatem et al., 2010). Given the high prevalence of *P. vivax*, the report ranked PNG as one of the countries facing the biggest difficulties to eliminate *P. vivax* malaria (Tatem et al., 2010). In the same report the north coast of PNG was identified as one of the few regions outside sub-Saharan Africa, where transmission needs to be reduced by more than 90% in order to eliminate malaria (Tatem et al., 2010).

### *P. vivax* Epidemiology in Papua New Guinea

This thesis addresses a number of questions related to *P. vivax* epidemiology and population genetics in Papua New Guinea (PNG). Samples were collected in a previously described cohort study in PNG in an area with high prevalence of around 50% for *P. vivax* as well as *P. falciparum*. Children aged 1 to 4.5 years were followed up in two-monthly cross-sectional surveys and during all febrile episodes (Lin et al., 2010). To gain a more comprehensive view of *P. vivax* populations in PNG, *P. vivax* clones from three other sites in PNG were typed and population structure was assessed.

In preceding molecular-epidemiological studies, Sonja Schoepflin had genotyped all *P. falciparum* clones using the marker *msp2* (Schoepflin et al., 2009). In the total sample set, over 3000 samples were positive for *P. vivax*. In the course of this thesis all *P. vivax* positive samples were genotyped for the markers *msp1F3* and MS16. Both markers showed high numbers of alleles. The virtual heterozygosity (the chance that two individual clones carry different alleles of a marker) was 88.1% for *msp1F3* and 97.1% for MS16 (details given in chapter 3). By using both highly polymorphic markers in combination only a small number of parasite clones will be missed due to the fact that two concurrent infections within one host share the same marker allele.

Parasite diversity as well as prevalence may be underestimated owing to imperfect detectability of all parasite clones present in a host. One reason for imperfect detection
probably is a parasite density below the detection limit of PCR. While the limits of microscopic
detection are well known, the detection limit of PCR based techniques is less acknowledged.
Repeated analysis of the same individual can be applied to investigate sensitivity in field
samples of very low parasite density. Different factors may cause low densities, or densities
changing from one day to another. Host immunity, or in the case of *P. vivax*, the host cell
preference for reticulocytes, results in overall low parasite density. The effects of
synchronized schizogony, and thus a marked increase in parasite numbers in 48 hour
intervals, and in the case of *P. falciparum* sequestration, resulting in absence of parasites
from the blood stream for part of the red blood cell cycle, can cause dramatical fluctuation in
parasite density in the blood stream within a short period of time.

Molecular epidemiological studies rely on a reliable detection of parasites. As we were
especially interested in age trends and differences between parasite species, we estimated
the proportion of infections missed in a single blood sample for both species, *P. falciparum*
and *P. vivax*, using data collected over 12 months. To this aim, we have compared blood
samples taken from the same individual in a 24-hour interval. Within one day no gain or loss
of an infection is expected. Therefore we concluded that parasites, observed only on one day,
must have been below the detection limit on the other day.

The microscopy and genotyping results from over 1'000 paired blood samples showed that
the impact of repeated sampling on prevalence is moderate: only around 10% of all parasite
positive samples were missed by sampling once only. The only exception to a generally low
impact of single sampling was observed for the microscopical detection of *P. falciparum*,
where 22% of infections were missed when sampling on a single day only. The effect is more
pronounced when individual clones are studied, with 18% of all *P. falciparum* and, depending
on the marker, 22% to 28% of all *P. vivax* clones being missed by sampling only once. We did
not observe an effect of age on detectability, however, with increasing multiplicity of infection
(MOI) the detectability decreased. There were no pronounced differences between
Plasmodium species, suggesting that the generally lower *P. vivax* densities balance the
reduction of detectability in *P. falciparum* caused by sequestration and synchronized
schizogony.

Chapter 3 of this thesis describes in detail the diversity of *P. vivax* clones and MOI in the
cohort study in PNG. Diversity was very high with 154 different two-loci haplotypes detected
in 219 single clone infections. No haplotypes were predominant. Mean MOI was 2.7 with a
slight increase with age, while *P. falciparum* MOI was only 1.5. No seasonal trend in MOI was
observed; likely relapses added to MOI in times of less mosquito transmission.

Plasmodium parasites live in a highly structured environment in their hosts and sexual
recombination in the mosquito midgut is only possible if different gametocytes are taken up
(Mzilahowa et al., 2007). This is the case if human hosts are infected simultaneously with different clones and if more than one clone produce gametocytes. High MOI is thus expected to be associated with high recombination between different parasites, and as a consequence, with low levels of linkage disequilibrium (LD). It has been shown in *P. falciparum* that LD decreases with increasing transmission intensity and with a high proportion of human hosts carrying multiple clone infections (Mu et al., 2005).

The extent of LD, as well as the parasite population structure, influence the spread of drug resistant parasites. High gene flow, resulting in little differences between geographically separated parasite populations, leads to high admixture and thus low LD. In such situations linkage between mutated loci that confer resistance to antimalarial drugs, is expected to be low (Talisuna et al., 2003). If a single gene confers drug resistance, the resistance trait likely spreads faster if gene flow is high (Talisuna et al., 2003). This is for example the case for resistance to sulphadoxine-pyrimethamine caused by mutations in the *dihydrofolate reductase* (*dhfr*) gene (Peterson et al., 1990).

Marked population differences and high levels of LD were shown in *P. vivax* populations from regions of low to moderate transmission in Latin America and Asia (Gunawardena et al., 2010; Imwong et al., 2007a; Van den Eede et al., 2010). Also *P. falciparum* in PNG shows a moderate level of population structure (Schultz et al., 2010). To compare these previous results from *P. falciparum* from PNG or *P. vivax* from other endemic settings, with *P. vivax* populations from in PNG, we have genotyped a subset of the samples from our cohort study and samples from 3 other locations in PNG for 13 additional neutral markers (chapter four). We did not detect population genetic differences between sites or LD, not even in the highlands, where substantially lower *P. vivax* transmission was expected. Together with the finding of a high MOI, our population genetic analyses indicated a panmictic population structure for *P. vivax* in PNG. This all implies that interventions to reduce transmission are not feasible in a restricted area due to high numbers of immigrating parasites.

The outcome of disease is strongly influenced by the acquisition of immunity to malaria. The 1 to 4.5 year old children in our cohort showed contrasting trends in episodes caused by *P. falciparum* and *P. vivax*. Incidence caused by *P. vivax* decreased over the whole age group from approximately 3.5 to 1.5 clinical episodes per child per year (Lin et al., 2010). In contrast, *P. falciparum* incidence rose from 1.2 clinical episodes per year in children aged 1 year to 3.5 episodes per year in 3-year-old children. Afterwards incidence decreases. It is generally observed that *P. vivax* causes disease mainly in very young children, while immunity to *P. falciparum* is acquired more slowly and thus *P. falciparum* causes episodes also in older children (Gunewardena et al., 1994; Mendis et al., 2001; Michon et al., 2007).
Laboratory experiments have shown that acquired immunity to malaria is largely clone-specific, i.e. immunity is less effective against infections with heterologous strains (Jeffery, 1966; Lyon et al., 2008). It has been speculated that the total number of infections acquired over time is a driving force behind the speed of acquisition of immunity (Weisman et al., 2001). In chapter five we have estimated the molecular force of infection (molFOI, the number of infections acquired per time) in our cohort. On average, a child was infected with more than 14 P. vivax clones per year, more than twice the number of P. falciparum clones acquired per year (Mueller et al., manuscript in preparation). P. vivax FOI did not change with age, indicating the decrease in P. vivax incidence indeed is a result of acquired immunity and is not caused by a decrease in numbers of new infections. We therefore suggest that the higher number of P. vivax infections acquired over time, as compared to the number of P. falciparum infections, leads to immunity in a shorter time than for P. falciparum. The same argument is true for the higher P. vivax MOI.

While P. falciparum molFOI was found to be a direct measure for transmission (Mueller et al., manuscript in preparation), any P. vivax parasite detected in the blood stream could also be the result of relapsing hypnozoites. Longitudinal sampling in P. vivax patients has demonstrated that the genotype of relapsing parasites differs to the one in the last blood stage infection in most of the cases (Imwong et al., 2007b; Restrepo et al., 2011). Thus relapses further increase molFOI. It is however not known whether the relapsing clone had been present as blood stage infection at an earlier time point. The answer to this question would require studies with infection of P. vivax naïve individuals in a setting where re-infection can be excluded, and subsequent genotyping of all blood stage parasites over several months.

molFOI is a central epidemiological parameter, and is of great relevance to monitor interventions, such as bed nets, which aim at lowering the number of infections acquired. Indeed, use of bed nets has led to a decrease of P. falciparum molFOI and incidence (Mueller et al., manuscript in preparation). This effect of bed nets is less clear for P. vivax incidence, most likely because of relapses that occur independent of bed net use. The higher P. vivax MOI and molFOI in comparison to the corresponding P. falciparum parameters certainly can be attributed to relapses. But also a higher transmission of P. vivax, e.g. due to different vector preference of P. vivax and P. falciparum, might play a role. No entomological studies were undertaken in parallel to our cohort, thus we cannot provide information on entomological inoculations rates for P. vivax and P. falciparum.

Due to the high diversity of P. vivax clones and the high level of gene flow between villages, children are infected within their first years of life with a large number of different genotypes. These infections cause a peak P. vivax morbidity in early childhood, and in parallel, a fast acquisition of immunity against disease.
Technical Considerations

Genotyping *P. vivax* using size polymorphic markers

For all studies presented in this thesis parasites were genotyped by PCR amplification of length polymorphic DNA markers followed by sizing of PCR products by capillary electrophoresis (CE). Length polymorphic markers have been used to genotype *Plasmodium* species since early molecular-epidemiological studies in the 1990s. Early genotyping protocols for *P. vivax* used PCR followed by restriction fragment length polymorphism (RFLP) to study the diversity of parasites, including the number of strains in a population or the multiplicity of infection (Bruce et al., 1999; Imwong et al., 2005). PCR-RFLP also proved useful to classify the outcome of drug treatment by comparing parasites at day 0 with parasites at day of recurrent parasitemia, and thus differentiating between new infection and treatment failure (Ohrt et al., 1997).

Given the relatively low level of technical equipment needed for PCR-RFLP experiments, the method has proven useful in endemic countries until now (Mbacham et al., 2010; Zhong et al., 2011). Nevertheless, the method has some disadvantages: Analysis of PCR or restriction fragments on agarose gels can be difficult, especially if alleles of a given marker differ by only a few base pairs (as it is usually observed in microsatellite typing). The number of size variant alleles that can be accurately discriminated by gel electrophoresis clearly is limited, in particular if for larger fragment lengths. Further problems arise in case of multiple clone infections, when the analysis of large numbers of bands after restriction digest of different clones can be extremely difficult. Finally, minority clones representing only a small fraction of the entire PCR product are likely overseen on a gel.

CE overcomes several of these problems. PCR products differing as little as 2 or 3 base pairs in size can be reliably differentiated, allowing discrimination of a large number of alleles for a single marker. Multiclonal infections can easily be identified. However, rules and cut offs need to be defined to distinguish true, but low intensity signals from background noise. Other authors have suggested to define a threshold of one quarter or one third of the intensity of the mayor peak in a sample, and to count minor peaks only if their intensity reaches at least this intensity (Havryliuk and Ferreira, 2009). However, in natural infections densities of concurrent parasite clones within a host might differ by several orders of magnitude, and the intrinsic ratios of peak intensities of two amplification products are still reflected even after nested PCR (Hunt et al., 2005). Applying a threshold as suggested would eliminate a considerable number of minority clones in a blood sample. We have thus defined a fixed threshold, independent of the intensity of the predominant clone.

In this context it should be noted that clones sharing markers of identical lengths will not be resolved by capillary sizing of fragments, but could be differentiated by sequencing. Thus, it is
important to consider that estimates of multiplicity obtained by PCR-CE are useful epidemiological parameter, but do not necessarily measure absolute MOI.

**Outlook: next generation sequencing**

Next generation sequencing technologies offer new opportunities and might overcome some of the limitations mentioned above. They require, however, sophisticated laboratory facilities as well as sound bioinformatics skills to analyse huge amounts of data produced. Both these factors are in high demand in particular in endemic countries. Analysis of large sample sets obtained in field studies or in the course of drug efficacy trials by next generation sequencing approaches is not yet done, but might become the method of choice in the future.

These technologies have already given interesting insights and helped to interprète field data. Recently Juliano and coworkers did massively parallel pyrosequencing of the *P. falciparum* *msp1* and *msp2* genes in field patient isolates. By analysing on average over 1300 reads per patient, they were able to detect low density variants, and indeed they found up to six times more clones per patient as determined by PCR combined with gel electrophoresis (Juliano et al., 2010). Likewise, whole genome SNP-typing proved useful for population genetic analysis. Campino et al. have genotyped over 300 single nucleotide polymorphis (SNPs) in several *P. falciparum* strains applying an Illumina genotyping assay. They were able to distinguish strains of different geographical origin and confirmed population structure obtained in previous studies (Campino et al., 2011). Typing SNPs rather than length polymorphism in microsatellites makes it possible to analyze further aspects of parasite biology and evolution. By assessing the rate of dN/dS, i.e. the rate of non-synonymous (dN) mutations per synonymous (dS) mutation, loci under selection (characterized by high values of dN/dS) can be identified. In case of field samples carrying multiple-clone infections, however, reconstruction of multi-locus haplotypes is currently impossible.

An increasing number of new technologies for genotyping and genome analysis are made available, and the initially high costs decreased dramatically over the past few years. They overcome some of the limitations of more traditional approaches, but might also lead to disproportional high complexity of laboratory and data analysis methods. For future studies the methods to apply need to be evaluated carefully and selected in accordance with the prime research questions.
Conclusions

We have for the first time genotyped a large number of *P. vivax* clones from a cohort of 264 children in an area highly endemic for *P. vivax* and *P. falciparum* (Lin et al., 2010). Our choice of two highly polymorphic markers allows discrimination of clones with a high resolution, and thus we could assess multiplicity (MOI) and molecular force of infection ($mol\text{FOI}$). All *P. falciparum* clones detected in this cohort have been genotyped previously, allowing comparisons of Plasmodium species. Most likely the values measured are underestimates because a proportion of clones was missed as a result of imperfect detectability of low density clones. However, as detectability does not differ between parasite species or age groups, comparisons and age trends of MOI and FOI were possible. We detected substantially higher *P. vivax* MOI and FOI compared to *P. falciparum*, most likely caused by relapses. High MOI increases recombination and thus seemed to eliminate population substructuring, in contrast to a moderate population structuring of *P. falciparum* in PNG (Schultz et al., 2010) *P. vivax* allelic composition does not differ between sites in PNG. High MOI and $mol\text{FOI}$ as well as high levels of gene flow imply great resilience of *P. vivax* towards antimalarial interventions (Mendis et al., 2001).

*P. vivax* $mol\text{FOI}$ was twice as high as that of *P. falciparum* (Schoepflin et al., manuscript in preparation), which likely leads to faster immunization and could therefore explain the differences in the age distribution of *P. vivax* versus *P. falciparum* incidence (Lin et al., 2010).

This thesis can serve as a reference for molecular epidemiology of *P. vivax* in high transmission areas and might be useful for future comparisons with *P. falciparum*. Control of *P. vivax* malaria is complicated by the large reservoir of dormant liver stages that relapse independently of mosquito transmission. The absence of population structuring implies intensive migration of parasites and affects the spread of drug resistance. *P. vivax* will thus present a major challenge on the way towards control and future elimination of malaria in PNG.
References


Appendix

During the course of this thesis contributions were made to the following manuscript published in “Antimicrobial Agents and Chemotherapy”. It shows a practical application of *P. vivax* genotyping for PCR correction of results from a drug trial and is thus not linked directly to the topic of this thesis.
Characterization of Treatment Failure in Efficacy Trials of Drugs against *Plasmodium vivax* by Genotyping Neutral and Drug Resistance-Associated Markers

Celine Barnadas, Cristian Koepfli, Harin A. Karunajeewa, Peter M. Siba, Timothy M. E. Davis, and Ivo Mueller

Vector Borne Diseases Unit, PNG Institute of Medical Research, Goroka, Papua New Guinea; Center for Global Health & Diseases, Case Western Reserve University, Cleveland, Ohio; Infection & Immunity Division, Walter & Eliza Hall Institute, Melbourne, Australia; Swiss Tropical and Public Health Institute, Basel, Switzerland; School of Medicine and Pharmacology, University of Western Australia, Perth, Australia; and Barcelona Centre for International Health Research, Barcelona, Spain

Received 10 November 2010/Returned for modification 18 January 2011/Accepted 14 June 2011

*Plasmodium vivax* intervention trials customarily report uncorrected treatment failure rates. Application of recrudescence-reinfection genotyping and drug resistance single-nucleotide polymorphism typing to a 4-arm comparative efficacy trial illustrated that molecular approaches can assist in understanding the relative contributions of true drug resistance (recurrant with same genotype) and new infections to treatment failure. The PCR-corrected adequate clinical and parasitologic response may constitute an informative secondary endpoint in future *P. vivax* drug trials.

Approximately 40% of the world’s population is at risk of vivax malaria (7). Recent interest in this infection has been heightened by the emergence of chloroquine (CQ) resistance in 1989 (21) and subsequent reports of severe disease (6, 25). There has been a resultant increase in monitoring of *Plasmodium vivax* drug sensitivity through efficacy trials (2, 12, 19, 20, 22) and identification of molecular markers of resistance (1, 8, 15, 24).

Recurrent *P. vivax* parasitemia in intervention trials may indicate not only treatment failure but also activation of liver-stage hypnozoites (relapses) or a new infection (17). Two studies of patients not at risk of reinfection found that most relapses were genetically distinct from the primary infections (4, 10). Standardized genotyping protocols characterizing treatment failure have not yet been developed for antimalarial trials for vivax malaria. However, candidate markers on genes coding for surface proteins (2, 9, 14) or neutral markers such as microsatellites (10, 13, 14) could, if sufficiently polymorphic, allow discrimination between strains in assessing posttreatment recurrence in a way analogous to that established for falciparum malaria (26). In a recent study of small numbers of children in Papua New Guinea (PNG) treated with amodiaquine or CQ plus sulfadoxine-pyrimethamine (SP), the authors recommended use of two highly polymorphic markers associated with a very low probability of independent infections carrying the same alleles (14). We have utilized this approach in a retrospective analysis of samples taken from a larger number of PNG children participating in an efficacy trial comparing CQ-SP and three artemisinin combination therapies (ACTs) (12), and we performed a complementary analysis of 4-aminoquinoline and SP drug resistance markers (3, 5, 15, 24).

The study was conducted in Madang and East Sepik Provinces between 2005 and 2007 (12) and involved 195 children aged 0.5 to 5 years with >250 *P. vivax* asexual forms/μl and no features of severe malaria who were randomly assigned to CQ-SP, artemunate-SP (ARTS-SP), dihydroartemisinin-piperazine (DHA-PIP), or artemether-lumefantrine (AL) arms of the trial. The non-PCR-corrected clinical and parasitologic failure rates were 49.0%, 48.7%, 15.8%, and 51.5%, respectively, after 28 days of follow-up and 87.0%, 66.7%, 30.6%, and 69.7% after 42 days. There was no difference between the rate of recurrent *P. vivax* parasitemia between the CQ-SP, ARTS-SP, and AL arms (P = 0.28, log rank test) (Fig. 1A).

Genotyping based on length polymorphism of a region of *msp1* (*msp1F3*) and a microsatellite, MS16, was performed; this combination has a probability of <0.25% for two isolates to carry the same alleles (14). Recurrent infections occurring during 42 days of follow-up that contained at least one genotype present at baseline were classified as recurrent infections with the same genotype, and recurrences with a different genotype were classified as new infections. Since *P. vivax* genotyping was not prespecified and in view of limited sample volumes, usable blood samples on the day of recurrent parasitemia were available for 70.1% and 70.3% of the samples to genotyping was not prespecified and in view of limited sample volumes to genotyping was not prespecified and in view of limited sample volumes, usable blood samples on the day of recurrent parasitemia were available for 70.1% and 70.3% of the samples to days 28 and 42, respectively. The present substudy was approved by the PNG IMR Institutional Review Board (approval 1029).

During 28 days of follow-up, there were no significant differences between CQ-SP, ARTS-SP, or AL in the rates of either recurrent parasitemia with the same genotype (Fig. 1B) (P = 0.74) or new infections (Fig. 1C) (P = 0.59). Up to day 42, there were significantly more cases of recurrent parasitemia...
with the same genotype in the CQ-SP arm but not in the number of new infections (Table 1). At days 28 and 42, fewer infections with either the same or different genotypes were observed after DHA-PIP (Table 1). Day 42 treatment failure rates were 87.0%, 66.7%, 69.7%, and 30.6% for CQ-SP, ARTS-SP, AL, and DHA-PIP, respectively, and 51.4%, 28.1%, 22.2%, and 9.7% after PCR correction.

To better understand these differences, we screened mutations in two P. vivax genes related to SP or 4-aminoquinoline resistance, namely, dhfr (5) and mdr1 (3, 23). Of patients allocated to the CQ-SP or ARTS-SP groups, 32 (69.6%) and 38 (71.8%), respectively, were infected with at least one triple or quadruple dhfr mutant parasite at enrolment (57L, 58R, and 61 M, 117T). While the small sample size did not allow firm conclusions regarding selection of mutant parasites, an increase in the frequency of triple/quadruple dhfr mutants was observed in the recurrent parasitemias with the same genotype (based on msp1/MS16 genotyping) in the CQ-SP (15/18; 83.3%) and ARTS-SP (9/9; 100%) arms. No such increase was observed in the AL arm (20/32 [62.5%] versus 4/6 [66.7%]). In the CQ-SP arm, the presence of parasites with triple/quadruple dhfr (57L, 58R, 61M/117T) plus mdr1 976F was associated with treatment failure (recurrent parasitemia with same genotype) with an odds ratio of 3.9 (95% confidence interval, 0.6–28.2; \( P = 0.08 \)). Given the overall high levels of triple/quadruple dhfr mutants, this adds to emerging evidence that mdr1 (976F) mutations may be involved in reduced P. vivax CQ sensitivity (16, 24).

Despite comparable non-PCR-corrected adequate clinical and parasitologic responses (ACPRs), the genotyping/molecular marker data reveal between-treatment differences in recurrent parasitemia that reflect the pharmacodynamic and pharmacokinetic profiles of the antimalarial drugs (11, 18, 27). Given the high prevalence of quadruple dhfr mutant parasites and the relatively short elimination half-lives of the components, SP is likely to have contributed little to either initial parasite clearance or to prevention of new (or relapsing) infections during follow-up in the CQ-SP and ARTS-SP arms.

**FIG. 1.** Kaplan-Meier curves for time to first recurrent P. vivax infection during 6 weeks of follow-up. (A) Time to any recurrent P. vivax parasitemia; (B) time to recurrent infection with the same genotype; (C) time to recurrent infection with a new genotype.

**TABLE 1.** Genotyping results of recurrent parasitemia during 6 weeks of follow-up

<table>
<thead>
<tr>
<th>Follow-up period, genotyping method, and result</th>
<th>No. of patients in assessment group and no. (%) with indicated result</th>
<th>( P ) value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CQ-SP</td>
<td>ART-SP</td>
<td>AL</td>
</tr>
<tr>
<td>28-day assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncorrected ACPR</td>
<td>51 (51.0)</td>
<td>20 (51.3)</td>
</tr>
<tr>
<td>Evaluable by PCR</td>
<td>26 (59.1)</td>
<td>20 (58.8)</td>
</tr>
<tr>
<td>No recurrent parasitemia</td>
<td>26 (59.1)</td>
<td>20 (58.8)</td>
</tr>
<tr>
<td>Same genotype</td>
<td>10 (22.7)</td>
<td>4 (11.8)</td>
</tr>
<tr>
<td>New infection</td>
<td>8 (18.2)</td>
<td>10 (29.4)</td>
</tr>
<tr>
<td>PCR-corrected ACPR†</td>
<td>34 (77.3)</td>
<td>30 (88.2)</td>
</tr>
<tr>
<td>42-day assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noncorrected ACPR</td>
<td>6 (13.0)</td>
<td>13 (33.3)</td>
</tr>
<tr>
<td>Evaluable by PCR</td>
<td>35</td>
<td>32</td>
</tr>
<tr>
<td>No recurrent parasitemia</td>
<td>6 (17.1)</td>
<td>13 (40.6)</td>
</tr>
<tr>
<td>Same genotype</td>
<td>18 (51.4)</td>
<td>9 (28.1)</td>
</tr>
<tr>
<td>New infection</td>
<td>11 (31.4)</td>
<td>10 (31.3)</td>
</tr>
<tr>
<td>PCR-corrected ACPR†</td>
<td>17 (48.6)</td>
<td>23 (71.9)</td>
</tr>
</tbody>
</table>

* Based on a chi-squared test, with \( P \) values for the CQ-SP, ART-SP, and AL treatments arms (single arms) or for DHA-PIP and the other three arms combined ("combined").

† PCR-corrected ACPR, the sum of no recurrent parasitemia and new infections.
Thus, ARTS would have been primarily responsible for successful initial clearance in the latter arm. The predominantly late occurrence of recurrent parasitemia irrespective of origin in the CO-SP arm indicates that CQ remained partially effective despite the positive selection of mdrl mutant (976F) parasites. The difference in efficacy between DHA-PIP and AL may largely reflect the terminal elimination half-lives of PIP (3 to 4 weeks) and lumefantrine (4 to 6 days), with long-lasting PIP suppression of reinfections and relapses regardless of genotype and plasma lumefantrine concentrations beyond 2 weeks posttreatment that were insufficient to prevent recurrence (20).

The present preliminary data highlight the important potential contribution that genotyping/molecular marker typing can make to improved characterization of recurrent parasitemia in *P. vivax* intervention trials. Since genotyping cannot differentiate between true failures and relapses with the same genotype, the primary endpoint should remain ACPR without genotyping. However, ACPR after genotyping based on epidemiologically appropriate markers could be added as a secondary endpoint. Further research may promote harmonization of *P. vivax* genotyping protocols and the adoption of consensus recommendations.

The intervention trial was sponsored by WHO Western Pacific Region, Rotary Against Malaria (PNG), and the National Health and Medical Research Council (NHMRC) of Australia (grant 353663). The present substudy was funded by NHMRC project grant 1010203. T.M.E.D. was supported by an NHMRC Practitioner fellowship.

REFERENCES

Curriculum Vitae

Contact Information

Name
Cristian Koepfli

Affiliation
Swiss Tropical and Public Health Institute
Socinstrasse 57, 4051 Basel, Switzerland

Private address
Kandererstrasse 23, 4057 Basel, Switzerland

Phone
+41 61 284 81 20 (office)
+41 76 418 70 01 (mobile)

E-Mail
cristian.koepfli@unibas.ch

Nationality
Swiss

Education

Oct 2008 – Dec 2011
PhD in molecular Epidemiology at the Swiss Tropical and Public Health Institute, Basel. Project: Epidemiology and population genetics of *Plasmodium vivax* in Papua New Guinea. Supervisor: Dr. Ingrid Felger

Master studies in Developmental Biology, University of Zurich, Switzerland.

Master thesis project: Proteomic characterization of 3 proteins promoting growth in *Drosophila melanogaster*. Supervisors: Dr. Erich Brunner, Prof. Dr. Ernst Hafen

Sept 2002 - July 2005
Bachelor studies in Biology, University of Zurich

March 2001
Swiss federal high school certificate

Language Skills

Mother tongue
Swiss German

English
Excellent skills

Spanish
Excellent skills (several longer stays in South America)

French
Good skills (September – December 1997: 3 month course in France)
Computer Skills

Office  Good skills in all usual office programs (Word, Excel, PowerPoint) and EndNote

Bioinformatics  Experience with a number of statistical and bioinformatics programs such as Stata, R, GeneMarker, Arlequin, Fstat, Structure etc.

Programming  Basic skills in programming in Perl and Java

Work Experience

June 2007 – Sept 2008  Co-project manager of a feasibility study for a new Tibetan medical academy in India in collaboration with the Tibetan Medical and Astro Institute of H. H. the Dalai Lama in Dharamsala, India, and Swiss experts

June – August 2008  Traineeship in molecular biology in the lab of Andrew P. Waters, University of Glasgow. Project: Characterization of the *Plasmodium berghei* Scramblase protein

April 2007 - June 2008  Scientific collaborator in molecular epidemiology at the Swiss Tropical Institute, Basel. Project: Development of a genotyping method for *Plasmodium vivax* based on size polymorphic genetic markers detected by PCR followed by capillary electrophoresis (alternative service in lieu of military service)

Oct 2005 – April 2006  Organization of the Annual Meeting of the “World Student Community for Sustainable Development”, a one-week workshop on “Mind, Knowledge and Sustainability” in the Swiss mountains with 50 participants from 15 countries

April - July 2005  Assistant in mathematics at the University of Zurich (8 hours/week)

Oct 2003 - July 2004  Assistant in chemistry at the University of Zurich (4 hours/week)

May - Sept 2002  Protection of natural reserves with the “Foundation for Economy and Ecology” (alternative service in lieu of military service)
Publications


Evaluation of Plasmodium vivax genotyping markers for molecular monitoring in clinical trials.
J Infect Dis. 2009 Apr 1;199(7):1074-80.

Characterization of Treatment Failure in Efficacy Trials of Drugs against Plasmodium vivax by Genotyping Neutral and Drug Resistance-Associated Markers.
Barnadas C, Koepfli C, Karunajeewa HA, Siba PM, Davis TM, Mueller I.

Presentations

PEG 2011 Oral presentation: *Plasmodium vivax* in Papua New Guinea consists of an unstructured, highly diverse population 9th Malaria meeting of the “Paul Ehrlich Gesellschaft for Chemotherapie”


PEG 2010 Oral presentation: How much is hiding? Probability of detection of human *Plasmodia* in the field 8th Malaria meeting of the “Paul Ehrlich Gesellschaft for Chemotherapie”

MEEGID 2010 Oral presentation: Population characteristics of *Plasmodium vivax* in Papua New Guinea 10th International Conference on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases
Curriculum Vitae

SSTMP 2009
Oral presentation: **Comparison of molecular methods and microscopy for detection of Plasmodium falciparum and P. vivax infections**
Annual Meeting of the Swiss Society of Tropical Medicine and Parasitology

BioMalPar 2009
Poster presentation: **Plasmodium vivax infection dynamics**
5th Annual BioMalPar Conference on the Biology and Pathology of the Malaria Parasite

**Course attendances**

ConGen

Java programming
Introduction into programming with Java. University of Basel, 1.5 days / week during 14 weeks

During my PhD studies I have attended classes from Prof. Dr. H.P. Beck, Prof. Dr. M. Tanner, Prof. Dr. T. Vetter and Dr. PD I. Felger

**Miscellaneous**

Student representative
2008 – 2011 representative of the PhD students at the Swiss Tropical and Public Health Institute

[project 21]
2002 – 2010 active member of [project 21] – students for sustainability of the ETH and the University of Zurich. Vice-president 2004 and president 2005
Responsible for a wide variety of events at the ETH and University of Zurich
Contribution of several articles to “Studio!Sus” (magazine on sustainability)

Photography
Sound experience in documentary photography. See www.crik.ch for pictures.