Elsevier Editorial System(tm) for Parasitology International
Manuscript Draft

Manuscript Number: PARINT-D-14-00094R1

Title: Global distribution of polymorphisms associated with delayed Plasmodium falciparum parasite clearance following artemisinin treatment: genotyping of archive blood samples

Article Type: SI: Tanabe's Gedenkschrift

Keywords: Plasmodium falciparum, drug resistance, MAL10-688956, MAL13-1718319, Artemisinin combination therapy

Corresponding Author: Dr. Toshihiro Mita, MD, PhD

Corresponding Author's Institution: Juntendo University School of Medicine

First Author: Kenji Murai, MD

Order of Authors: Kenji Murai, MD; Richard Culleton, PhD; Teruhiko Hisaoka, MD, PhD; Hiroyoshi Endo, MD, PhD; Toshihiro Mita, MD, PhD

Abstract: The recent emergence and spread of artemisinin-resistant Plasmodium falciparum isolates is a growing concern for global malaria-control efforts. A recent genome-wide analysis study identified two SNPs at genomic positions MAL10-688956 and MAL13-1718319 which are linked to delayed clearance of parasites following artemisinin combination therapy (ACT). It is expected that continuous artemisinin pressure will affect the distribution of these SNPs. Here, we investigate the worldwide distribution of these SNPs using a large number of archived samples in order to generate baseline data from the period before the emergence of ACT resistance. The presence of SNPs in MAL10-688956 and MAL13-1718319 was assessed by nested PCR RFLP and direct DNA sequencing using 653 global P. falciparum samples obtained before the reported emergence of ACT resistance. SNPs at MAL10-688956 and MAL13-1718319 associated with delayed parasite clearance following ACT administration were observed in 8% and 3% of parasites, respectively, mostly in Cambodia and Thailand. Parasites harbouring both SNPs were found in only eight (1%) isolates, all of which were from Cambodia and Thailand. Linkage disequilibrium was detected between MAL10-688956 and MAL13-1718319, suggesting that this SNP combination may have been selected by ACT drug pressure. Neither of the SNPs associated with delayed parasite clearance were observed in samples from Africa or South America. Baseline information of the geographical difference of MAL10-688956 and MAL13-1718319 SNPs provides a solid basis for assessing whether these SNPs are selected by artemisinin-based combination therapies.

Juntendo University School of Medicine

Department of Molecular and Cellular Parasitology
Hongo 2-1-1, Bunkyo-ku, Tokyo, 113-8421, Japan
Tel 81-3-5802-1042/Fax 81-3-5800-0476

Date: 30 Sep, 2014
To: Editor
Subject: PARINT-D-14-00094
Title: "Global distribution of polymorphisms associated with delayed *Plasmodium falciparum* parasite clearance following artemisinin treatment: genotyping of archive blood samples"

Thank you very much for your mail on 28 September 2014 regarding our manuscript entitled "Global distribution of polymorphisms associated with delayed Plasmodium falciparum parasite clearance following artemisinin treatment: genotyping of archive blood samples". Following your encouragement and support, we have revised the manuscript.

We are very grateful to the reviewers for their comments on this manuscript and for their sincerest acknowledgement for the work that went into the paper. We have carefully considered the points raised by the reviewers and had adopted the suggestions resulting in a further improved paper. All changes in the manuscript are highlighted in red in the revised text. Our detailed point-by-point responses are attached to this cover letter.

Herein I state that all the authors concur with this revision and that this manuscript has not been submitted or accepted for publication elsewhere. All authors fulfill the criteria and no writing assistance other than copy editing was provided in the preparation of the manuscript. Two coauthors who are native English users carefully checked English usage of the text.

Thank you for your time and consideration. I look forward to hearing from you.

Sincerely yours,

Toshihiro MITA M.D., Ph.D.
E-mail: tmita@juntendo.ac.jp

Reviewer #1:

(1) In discussing the evolution of the artemisinin resistance phenotype, the authors suggest that the two SNPs examined might be prerequisites for resistant mutations elsewhere in the genome. The authors also aptly draw a parallel with resistance to pyrimethamine and sulfadoxine, where the sequential accumulation of point mutations confers high-level resistance. I can't help but wonder whether the authors, in the context of this manuscript, can address which of the two SNPs associated with delayed parasite clearance might be selected first by ACT drug pressure. I find the Cambodian sample set to be extremely interesting, in that the two mutant SNPs show statistically significant non-random association not seen in the rest of the parasite genome. Given that the Cambodian samples were collected in three successive years from 2004 to 2006, is it possible to see a year-to-year increase in the frequency of either or both of these SNPs?

<Reply>

Thank for your suggestion. The analysis suggested by the reviewer is interesting. However, we could not find any trend of a year-to-year increase in the frequency of either or both of these SNPs. Please see the results bellow.

MAL10-688956

	T	A
2004	11	19
2005	4	2
2006	0	2

MAL13-1718319

	A	T
2004	23	8
2005	6	1
2006	1	1

MAL10-688956 + MAL13-1718319

	T/A	A/A	T/T	A/T
2004	17	5	2	6
2005	2	2	0	1
2006	1	0	1	0

(2) The authors note the unexpectedly high prevalence of the MAL 10-688956 (A) allele in the Pacific countries, even before ACT was officially adopted as first-line treatment for malaria. Do the authors see evidence of gene flow (e.g. shared microsatellite alleles) between parasites from the Pacific and Southeast Asia, especially among parasites harboring this particular SNP? What do the authors think about the possibility that artemisinin resistance can evolve in the Pacific, independently of the Southeast Asian focus, given the prior example of chloroquine resistance?

<Reply>
We previously determined the microsatellites flanking drug-resistance genes pfcrt, dhps, and dhfr to investigate the possible gene flow between PNG and Southeast Asian countries (Antimicrobial Agent Chemother, 2007, J Infect Dis 2011, Malaria J 2012) and found that gene flow was evidenced in the dhfr and dhps mutants. However, human movement between Southeast Asia and Melanesia seems not to be frequent because of geographical obstacles. Therefore, the observed considerable frequencies of Southeast-origin resistant parasites in PNG are very likely to the infrequent migration from Southeast Asia to Pacific countries and subsequent selection due to the usage of SP in PNG.
In this analysis, however, we found high prevalences of MAL 10-688956 (A) allele in the Pacific countries (14% in Papua New Guinea, 18% in Vanuatu and 2% in the Solomon Islands). If we assume that the observed MAL 10-688956 (A) allele in the Pacific countries is because of the migration of this allele from Southeast Asian countries, it is difficult to explain the observed considerable prevalence of the allele, since artemisinin was not introduced as the first-line antimalarial.
So, we have changed the paragraph that discussed the unexpectedly high prevalence of the MAL 10-688956 (A) allele in the Pacific countries as follow.

(Original)

In the MAL10-688956, the delayed-clearance associated SNP may exist naturally in the parasite populations without any artemisinin selection. A considerable number of delayed-clearance associated SNPs were observed at MAL10-688956 in the Pacific region, despite the fact that artemisinin derivatives were not implemented at the time of sampling.

(Revised)

In the Pacific region, prevalences of the MAL10-688956 delayed-clearance associated SNP were unexpectedly high, although nearly all parasites harbored the non-delayed clearance associated SNP at MAL13-1718319. Artemisinin combination

therapies were not implemented at the time of sampling in any of the countries considered here. Hence, it seems likely that the MAL10-688956 delayed-clearance associated SNP is a parasite polymorphism that exists naturally in the parasite populations of this region, and was not, initially, selected by ACT pressure. The other possibility is the migration of the MAL10-688956 delayed-clearance associated SNP from Southeast Asia. However, since the prevalences of this SNP were high, this may be unlikely, especially in the absence of ACT selection pressure. Further analysis using microsatellite alleles flanking the SNP will clarify the possible migration of from the MAL10-688956 SNP to Southeast Asia.

Minor corrections/comments/suggestions

(1) Line 138: …in rural villages located on 4 islands…
<Reply>
We revised it according to the reviewer's suggestion.

(2) Lines 158-9: …implementation of ACT except in Cambodia and…before the first official report of…
<Reply>
We revised it according to the reviewer's suggestion.

(3) Lines 175-179: The names of restriction enzymes should be italicized. Since the SNPs are only associated with the delayed clearance phenotype, it might be more appropriate to describe the alleles as wild type (and mutant), instead of sensitive (and implicitly resistant).
<Reply>
The reviewer is absolutely correct that we should not say "sensitive" and "resistant", then we were very careful only to refer to them as "delayed clearance associated SNP" and non-delayed clearance associated SNP". However, since the terms, "mutant" and "wild-type", only refer to SNPs linked to a phenotype that has changed between two isogenic parasite lines, we consider it may not be appropriate to refer to "mutant" and "wild-type" SNPs. Thus, we agree that we need to remove reference to "sensitive" and "resistant" parasites, but we don't think we should replace with "mutant" and "wild-type". We use the original nomenclature, "delayed clearance associated SNP" and non-delayed clearance associated SNP".

(4) Lines 179-86: This part is very confusing to me. Was the sequencing done for all undigested amplicons of both loci? Was the sequencing done in only one direction using the reverse primer of the nested PCR? How did the sequencing reaction (lines 185-6)

involve two rounds of PCR reaction, when the previous sentence states that the PCR products were sequenced directly after purification?

<Reply>
We sequenced all undigested amplicons in only one direction using the reverse primer of the nested PCR. For the sequencing, we conducted initial and nested PCR using Gflex DNA Polymerase (Takara), but this looks detail and may cause misunderstanding, so we have deleted this.

(Original)

All undigested samples, in order to confirm the presence of mutant allele, nested PCR amplicons were purified with ExoSAP - IT Kit (Amersham Biosciences, Buckinghamshire, UK) and directly sequenced (50 cycles of 95 °C for 20 sec, 50 °C for 30 sec, and 60 °C for 1 min) in one direction using the reverse primer of the nested PCR (TTATATGTAATGGGTGAAAAGAATGTGG) with a BigDye Terminator 1.1 cycle sequencing kit in the Applied Biosystems 3500xL genetic analyzer (Life Technologies, Carlsbad, California, U.S.). For the sequence reaction, 0.4 unit of Gflex DNA Polymerase (Takara) was used in the both initial and nested PCR reactions.

(Revised)

In all undigested samples, in order to confirm the presence of mutant allele, nested PCR amplicons were purified with ExoSAP - IT Kit (Amersham Biosciences, Buckinghamshire, UK) and were directly sequenced (50 cycles of 95 °C for 20 sec, 50 °C for 30 sec, and 60 °C for 1 min) in one direction using the reverse primer of the nested PCR (TTATATGTAATGGGTGAAAAGAATGTGG) with a BigDye Terminator 1.1 cycle sequencing kit in the Applied Biosystems 3500xL genetic analyzer (Life Technologies, Carlsbad, California, U.S.).

(5) Lines 242-252: The authors use the terms "linkage" and "linkage disequilibrium" interchangeably, but these terms are not equivalent.
<Reply>
The terms "linkage" was rewritten to "linkage disequilibrium" accordingly.

(6) Line 263: In both countries,
<Reply>
We revised it according to the reviewer's suggestion.

(7) Acronyms are used without prior definition, e.g. LD (line 206) = linkage

disequilibrium; GWAS (lines 280 and 283) = genome-wide association study
<Reply>
We revised it according to the reviewer's suggestion.

(8) Figure 1A: The key in the figure is labelled incorrectly. It should say MAL 10-688956 instead of MAL 13-1718319.
<Reply>
We revised it according to the reviewer's suggestion.

(9) Legends for figures 1 and 2: The colors used to represent different genotypes/genotype combinations, as stated in the legends, are different than the actual colors used in the figures.
<Reply>
We revised it according to the reviewer's suggestion.

Reviewer #2:

1. Figure 1 A contains serious mistakes of the SNP ID number. Correction is needed.
<Reply>
This error was also pointed out by the reviewer 1. We revised it.

2. In Itroduction, line 91-92; The description "This suggests that the two proposed SNPs might be widely distributed, and their selection unrelated to ACT pressure" needs reference.
<Reply>
We revised it according to the reviewer's suggestion.

3. If possible, the prevalence of PF3D7_1343700 kelch 324 propeller domain ('K13-propeller') gene in the analyzed samples should be included.
<Reply>
This is an interesting suggestion.
We have already started the analysis of K13-propella gene using same sample-set. We are thinking to assess the baseline polymorphisms of the gene and, more importantly, to investigate whether this gene was under positive selection using several population-genetics methods. Since the data-set will be enormous, we are hoping to submit the results as a separate publication.

***Suggested Reviewers**

Suggested reviewers

- **Osamu Kaneko**

okaneko@nagasaki-u.ac.jp
Department of Protozoology
Institute of Tropical Medicine (NEKKEN)
Nagasaki University
1-12-4 Sakamoto, Nagasaki 852-8523
JAPAN

- Anders Bjorkman

anders.bjorkman@karolinska.se
Division of Infectious Diseases
Karolinska University Hospital
SE-171 76 Stockholm
Sweden

Meera Venkatesan
meera.venkatesan@wwarn.org
Worldwide Antimalarial Resistance Network (WWARN) Molecular Module, Baltimore MD, USA

Highlights

Worldwide distribution for ACT-resistant SNPs was assessed in P. falciparum isolates.

Combination of two ACT-resistant SNPs were localised to Cambodia and Thailand.

Linkage disequilibrium was detected between two ACT-resistant SNPs.

These SNPs were entirely absent in the parasites from Africa or South America.

Global distribution of polymorphisms associated with delayed *Plasmodium falciparum* parasite clearance following artemisinin treatment: genotyping of archive blood samples

Kenji Murai[a,b], Richard Culleton[c], Teruhiko Hisaoka[b], Hiroyoshi Endo[d], Toshihiro Mita[a,d]*

[a] Department of Molecular and Cellular Parasitology, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo, 113-8421, Japan

[b] Department of General Medicine, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo, 113-8421, Japan

[c] Malaria Unit, Department of Pathology, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki, 852-8523, Japan

[d] Department of International Affairs and Tropical Medicine, Tokyo Women's Medical University School of Medicine, 8-1 Kawada-cho, Shinjuku-ku, Tokyo, 162-8666, Japan

***Corresponding author contact information:**

Toshihiro Mita M.D., Ph.D.,

Department of Molecular and Cellular Parasitology, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo, 113-8421, Japan. Tel: +81-3-5802-1043.

Fax: +81-3-5800-0476.

E-mail: tmita@juntendo.ac.jp

Abstract

The recent emergence and spread of artemisinin-resistant *Plasmodium falciparum* isolates is a growing concern for global malaria-control efforts. A recent genome-wide analysis study identified two SNPs at genomic positions MAL10-688956 and MAL13-1718319 which are linked to delayed clearance of parasites following artemisinin combination therapy (ACT). It is expected that continuous artemisinin pressure will affect the distribution of these SNPs. Here, we investigate the worldwide distribution of these SNPs using a large number of archived samples in order to generate baseline data from the period before the emergence of ACT resistance. The presence of SNPs in MAL10-688956 and MAL13-1718319 was assessed by nested PCR RFLP and direct DNA sequencing using 653 global *P. falciparum* samples obtained before the reported emergence of ACT resistance. SNPs at MAL10-688956 and MAL13-1718319 associated with delayed parasite clearance following ACT administration were observed in 8% and 3% of parasites, respectively, mostly in Cambodia and Thailand. Parasites harbouring both SNPs were found in only eight (1%) isolates, all of which were from Cambodia and Thailand. Linkage disequilibrium was detected between MAL10-688956 and MAL13-1718319, suggesting that this SNP combination may have been selected by ACT drug pressure. Neither of the SNPs associated with delayed parasite clearance were observed in samples from Africa or South America. Baseline information of the geographical difference of MAL10-688956 and MAL13-1718319 SNPs provides a solid basis for assessing whether these SNPs are selected by artemisinin-based combination therapies.

Keywords: *Plasmodium falciparum*, drug resistance, MAL10-688956, MAL13-1718319, Artemisinin combination therapy

1. Introduction

There were 207 million cases of malaria and 627 000 deaths due to the disease world-wide in 2012 [1]. One of the most serious threats to the successful control of malaria is the emergence of parasites that are resistant to antimalarial drugs. The World Health Organization currently recommends artemisinin-based combination therapies (ACTs) as the first-line treatment for uncomplicated malaria. However, there is great concern that artemisinin-resistant *Plasmodium falciparum* parasites have emerged in the Cambodia/Thailand border region, where parasites resistant to other antimalarial drugs also originated [2-6]. These ACT “resistant” strains are characterized by a delay in the time it takes parasites to clear from the body following treatment [7].

For many anti-malarial drugs such as chloroquine and pyrimethamine/sulfadoxine, the genetic mutations that underlie resistance have been largely elucidated. These mutations can be used as molecular markers to monitor the appearance and geographical spread of resistant parasites. For artemisinin and its derivatives, a genetic region associated with a delay in parasite clearance following ACT treatment was identified in 2012 [8]. Soon after, four single nucleotide polymorphisms (SNPs) linked to the delayed clearance phenotype were identified on chromosomes 10, 13, and 14 [9]. Two of these, MAL10-688956 (A) and MAL13-1718319 (T), were proposed to be suitable molecular markers for the resistance phenotype.

MAL10-688956 is located on chromosome 10 in the 3’ untranslated region of the DNA polymerase delta catalytic subunit gene and MAL13-1718319 is in a RAD5 homolog. These two proteins are thought to be involved in post-replication repair [9,

10]. In particular, RAD5 is a DNA clamp that is involved in the DNA damage tolerance pathway which promotes the repair of discontinuities [11-13]. In yeast, mutations in this gene have been implicated in cell cycle arrest [13, 14] and thus, a similar role may be expected in *P. falciparum* [15], which might lead to delayed clearance following artemisinin treatment.

We have previously analysed 53 travellers' malaria samples collected from patients that had returned to Scotland from 11 African and nine South-eastern Asia/Oceania countries. We found that two samples harboured the delayed-clearance associated SNPs in both MAL10-688956 and MAL13-1718319 [16]. These two isolates were obtained from Thailand and Cambodia, the epicentre of the apparent emergence of resistance to artemisinin as well as to other antimalarial drugs [4, 17, 18]. This finding supports the notion that SNPs MAL10-688956 and MAL13-1718319 could be applicable as molecular markers for the surveillance of artemisinin resistance.

However, the delayed-clearance associated SNPs are also found in several *P. falciparum* laboratory maintained clones, e.g., V1/S, IT, 106/1, and FCR3 (MAL10-688956-A) and V1/S and IT (MAL13-1718319-T), which were isolated from patients from diverse geographic regions and well before the first reports of artemisinin resistance. This suggests that the two proposed SNPs might be widely distributed, and their selection unrelated to ACT pressure [9]. Thus, we consider that robust information on the global prevalence of these SNPs before the widespread implementation of ACTs would provide the baseline data necessary to infer whether they are, indeed, reliable markers for the spread of ACT resistance.

Here, we determine the distribution of delayed-clearance associated SNPs MAL10-688956 and MAL13-1718319 using a large number of *P. falciparum* isolates

from East/West Africa, Asia, Pacific Oceania and South America. All samples were obtained before the first report of the emergence of artemisinin resistance [2]. We report the complete absence of delayed-clearance associated SNPs in parasites collected from South America and Africa. Delayed-clearance associated SNPs were found singularly in a number of samples from areas outside Africa and South America, but only parasites from Cambodia and Thailand harboured both SNPs simultaneously.

2. Material and methods

2.1 Study Sites

Blood samples were obtained from *P. falciparum*–infected patients in all age groups, unless otherwise stated, living in 13 malaria-endemic countries as follows (Table 1):

1. Bangladesh: Samples were collected from patients infected with *P. falciparum* in the Bandarban district hospital in 2007. This study was approved by the Bangladesh Medical Research Council and the local health regulatory body in Bandarban, Bangladesh [19].

2. Cambodia: Samples were collected from *P. falciparum* infected individuals during a cross-sectional survey of rural villages in Chumkiri, Kampot province in 2004, 2005 and 2006. The study was approved by the National Center for Parasitology, Entomology, and Malaria Control of Cambodia [20].

3. Thailand: Samples were collected from patients infected with *P. falciparum* at town clinics located in the western border of Tak, Kanchanaburi, and Ratchaburi provinces from 2001 to 2002. The study was approved by the Ethics Committee of the Faculty of Tropical Medicine, Mahidol University.

4. Lao People's Democratic Republic (Lao PDR): Samples were collected from *P. falciparum* infected individuals during cross-sectional surveys of rural villages in Khammouanne province in 1999. The study was approved by the Laos Ministry of Health [21].

5. Philippines: Samples were collected from patients infected with *P. falcip*arum in hospitals on Palawan Island in 1997. This study was approved by the Palawan Provincial Health Office [22].

6. Papua New Guinea: Samples were collected from *P. falciparum* infected individuals at villages in Dagua district, East Sepik in 2002 and 2003. The study was approved by the National Department of Health Medical Research Advisory Committee of Papua New Guinea.

7. Solomon Islands: Samples were collected from *P. falciparum* infected individuals during cross-sectional surveys in northeastern Guadalcanal Island from 1995 to 1996. The study was approved by the Ethics Committee of the Solomon Islands for Medical Research.

8. Vanuatu: Samples were collected from *P. falciparum* infected individuals during cross-sectional surveys in rural villages located on 4 islands; Gaua, Santo, Pentecost and Malakula, in 1996 and 1998. The study was approved by the Vanuatu Department of Health [23].

9. Kenya: Samples were collected from *P. falciparum* infected individuals during cross-sectional surveys at 4 villages in Kisii District in 1998. The study was approved by the Kenyan Ministry of Health and Education [24].

10. Tanzania: Samples were collected from *P. falciparum* infected individuals during cross-sectional surveys in the Rufiji River Delta in eastern coastal Tanzania in

1998 and 2003. The study was approved by the Ethics Committee of the National Institute for Medical Research of Tanzania [25].

11. Republic of the Congo: Samples were collected from patients with *P. falciparum* in Pointe-Noire, Brazzaville, and Gamboma in 2006. The study was approved by the Ministry of Research and Ministry of Health of the Republic of the Congo [26].

12. Ghana: Samples were collected from *P. falciparum* infected children during cross-sectional surveys in 3 villages near Winneba, a western coastal region, in 2004. This study was approved by the Ministry of Health/Ghana Health Service.

13. Brazil: Samples were collected from *P. falciparum* infected individuals in the eastern part of Acre state in 1985–1986, 1999, and 2004–2005. The study protocol was approved by the ethics review board of the Institute of Biomedical Sciences, University of São Paulo.

All studies were conducted before the official implementation of ACT except in Cambodia and Thailand, and before the first official report of artemisinin resistance [2]. Finger-prick blood samples were collected and transferred on filter paper (ET31CHR; Whatman) in the all studied regions except Thailand in which venous blood samples were used. Parasite DNA was purified using a QIAamp DNA blood mini kit (QIAGEN) or the EZ1 BioRobot™ (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. In all study sites, informed consent was obtained from individual patients or their guardians and antimalarial treatment was provided if necessary.

2.2 Determination of polymorphisms in MAL10-688956 and MAL13-1718319

Nested polymerase chain reaction (PCR) for Plasmodium species typing was conducted to confirm the presence of *P. falciparum* parasites [27, 28]. All *P.*

falciparum positive samples were assayed for the presence of SNPs in MAL10-688956 and MAL13-1718319 [9]. A PCR and restriction fragment length polymorphism (RFLP) protocol was used as described on the worldwide antimalarial resistance network (WWARN) website: http://www.wwarn.org/toolkit/procedures with one modification (1.5 mM $MgCl_2$ in the nested PCR for MAL13-1718319). Briefly, genotyping was conducted by nested PCR followed by RFLP analysis. *NsiI* (New England Biolabs), which digests the amplified product when the allele (T, non-delayed clearance associated SNP) is present at the polymorphic site, was used for MAL10-688956. *Msl I* (New England Biolabs), which produces two digested products when the allele (A, non-delayed clearance associated SNP) is present, was used for MAL13-1718319. In all undigested samples, in order to confirm the presence of delayed-clearance associated SNP, nested PCR amplicons were purified with ExoSAP - IT Kit (Amersham Biosciences, Buckinghamshire, UK) and were directly sequenced (50 cycles of 95 °C for 20 sec, 50 °C for 30 sec, and 60 °C for 1 min) in one direction using the reverse primer of the nested PCR (TTATATGTAATGGGTGAAAAGAATGTGG) with a BigDye Terminator 1.1 cycle sequencing kit in the Applied Biosystems 3500xL genetic analyzer (Life Technologies, Carlsbad, California, U.S.).

2.3 Microsatellite analysis

Ten neutral microsatellite markers with no evidence of genetic hitchhiking were genotyped to examine the possibility that linkage disequilibrium is found between evolutionary neutral markers in the Cambodian parasite population. The loci used were as follows: TA42 and TA81 (chromosome 5), TA1, TA87, and TA109

194 (chromosome 6), TA60 and 2490 (chromosome 10), and ARA2 (chromosome 11),
195 and Pfg377 and PfPK2 (chromosome 12). These markers were amplified by
196 semi-nested PCR using fluorescent end-labelled primers as previously described [29].
197 Size variations of the amplified products were determined by electrophoresis on a
198 DNA sequencer and analysed with GeneScan software (Applied Biosystems).
199 Samples with minor peaks at least 50% in peak height compared to the major peak
200 were considered mixed genotype, and were excluded from the analysis.
201
202 *2.4 Statistical analysis*
203 D' and r^2 were measured to assess potential linkage disequilibrium between
204 delayed-clearance associated SNPs [30, 31]. Linkage disequilibrium for all pairs of ten
205 microsatellite loci was also examined using Genepop version 4.1 under the following
206 Markov chain parameters: dememorization number = 20000), number of batches =
207 500, and number of iterations per batch = 10000. The significance of linkage
208 disequilibrium was assessed using the two-tailed Chi-squared test. $P < 0.05$ was
209 considered statistically significant.
210
211 **3. Results**
212 *3.1 MAL10-688956*
213 Among a total of 653 *P. falciparum* isolates, we successfully determined allele types
214 for 637 isolates (98%) at MAL10-688956 (Figure 1.A). The overall prevalence of the
215 delayed-clearance associated SNP was 8% (52/637). Ten isolates (2%) were found to
216 be dimorphic with both T and A nucleotides present at this position, indicating a mixed
217 infection. The delayed-clearance associated SNP was not detected in any isolates

from Africa or South America, but was prevalent in Cambodia (38%) and Thailand (26%). In both countries, ACTs were already in use as first-line therapy for uncomplicated malaria when the samples were obtained. However, there were very few cases of the presence of this delayed-clearance associated SNP in neighbouring countries, Laos PDR (0%) and Bangladesh (2%). Notably, unexpectedly high prevalences of the delayed-clearance associated SNP was observed in Pacific countries, 14% in Papua New Guinea, 18% in Vanuatu and 2% in the Solomon Islands. In these regions, artemisinin or its derivatives were not implemented at the time of sampling.

3.2 MAL13-1718319

We successfully genotyped MAL13-1718319 in 637 isolates (98%) (Figure 1.B). The delayed-clearance associated SNP (T) was observed in only 18 isolates (3%), none of which were from Africa or South America. This SNP was almost exclusively confined to samples from Cambodia and Thailand, with the exception of one sample from Papua New Guinea, which harboured both alleles.

3.3 Combinations of delayed-clearance associated SNPs at MAL10-688956 and MAL13-1718319

Those samples harbouring a delayed-clearance associated SNP at both MAL10-688956 and MAL13-1718319 are shown in Figure 2. Isolates with mixed alleles at either locus (n = 11) were excluded. Among 615 isolates, eight isolates (1%) harboured delayed-clearance associated SNPs at both loci. Nearly all of these isolates (7/8) were localised in Cambodia, and one was from Thailand.

242

243 *3.4 Analysis of linkage disequilibrium*

244 Linkage disequilibrium (non-random association) between MAL10-688956 and
245 MAL13-1718319 was observed only in Cambodia (D' = 0.5174 and r^2 = 0.1628)
246 (Table 2) with statistical significance (p = 0.0146, Chi-square test). We then assessed
247 the existence of linkage disequilibrium between ten putatively neutral microsatellite
248 loci to clarify whether the observed linkage disequilibrium was an inherent feature of
249 the Cambodian parasite population. Among a total of 36 Cambodian isolates, 25
250 showed multiple alleles at at least one microsatellite locus and were excluded from
251 this analysis (Table S1). Analysis of linkage disequilibrium between each
252 microsatellite locus produced 45 comparisons, but two results were not obtained
253 because only one allele combination was observed (Figure S1). No linkage
254 disequilibrium was observed in the remaining 43 comparisons with the smallest p
255 value 0.18 (comparison between TA60 and TA42), which indicates that linkage
256 disequilibrium is not an inherent feature of the Cambodian isolate population. Rather,
257 the observed linkage disequilibrium may be produced by the selection of parasites
258 that harbour both delayed-clearance associated SNPs.

259

260 **4. Discussion**

261 We assessed the distribution of SNPs at MAL10-688956 and MAL13-1718319
262 using a large number of archived worldwide *P. falciparum* isolates prior to the reported
263 emergence of ACT resistance [2]. Parasites harbouring a combination of two
264 delayed-clearance associated SNPs (A allele in MAL10-688956 and T allele in
265 MAL13-1718319) were found exclusively in Thailand and Cambodia. In both countries,

ACT was already implemented when our sampling was carried out. This may be due to the fact that there were no clear criteria for assessing ACT 'resistance' until the mid-2000s [32] and thus, the potential emergence of ACT-resistant parasites might have been missed at the time our sampling was carried out. Additionally/alternatively, delayed-clearance associated SNPs at MAL10-688956 and MAL13-1718319 alone may not be sufficient for the acquirement of the 'resistance' phenotype; rather, these polymorphisms might be necessary prior to the attainment of 'resistance' through mutation elsewhere in the genome. A similar mechanism is thought to be involved in the attainment of resistance to pyrimethamine/sulfadoxine. Amino acid substitutions at position 108 in the dihydrofolate reductase gene and at position 437 in dihydropteroate synthase gene are not, in themselves, sufficient to confer a high degree of in vivo resistance, but they are required as an initial step for the further acquisition of other mutations in these genes that eventually result in pyrimethamine and sulfadoxine resistance, respectively [33, 34]. It is possible that resistance to artemisinin would require mutations in more than one gene in the parasite's genome. Therefore, one or more undetermined gene(s) that were not identified in the previous genome-wide association study by Takala-Harrison et al [9] might be required for the 'resistance' phenotype. A recent deep whole-genome sequencing study has indicated that linkage disequilibrium decayed within 1 kb [35]. Hence, as noted by the authors, the SNP markers used for the previous genome-wide association study, 7 kb apart on average, may not be enough to detect all loci associated with parasite clearance [9].

The delayed-clearance associated SNP combination was mainly distributed in Cambodia and these SNPs were significant linked. Since the two SNPs are located on different chromosomes (10 and 13), physical linkage does not explain the observed

290 linkage disequilibrium. Thus, the following two mechanisms are proposed for the
291 observed linkage disequilibrium; (1) low diversity within the Cambodian population,
292 leading to the signature of linkage disequilibrium between markers separated by large
293 physical distances on the genome (2) natural selection of a particular allele
294 combination that has a sufficient selective advantage over others [36]. To test the
295 former possibility, we examined linkage disequilibrium using 10 microsatellite markers
296 that are thought to be selectively neutral. We did not find linkage disequilibrium
297 between any of the possible pairs of microsatellite loci, suggesting that the observed
298 linkage disequilibrium between two delayed-clearance associated SNPs is not the
299 result of low diversity within the Cambodian parasite population. Hence, it is probable
300 that the observed linkage disequilibrium is a result of selective pressure favouring
301 parasites with both delayed-clearance associated SNPs. The continuous use of
302 artemisinin in this area is one candidate for possible selecting factors.
303 We found no evidence for the existence of delayed-clearance associated SNPs at
304 either MAL10-688956 or MAL13-1718319 in Africa or South America. All the samples
305 from African regions were collected between 1998 and 2006, at the time period before
306 the wide-scale implementation of artemisinin derivatives. Practically, the development
307 of a credible molecular marker of ACT resistance is urgently required, particularly in
308 sub-Saharan Africa. Since many individuals living in this region develop protective
309 immunity to malaria after repeated infections, this immunity would enhance the
310 effectiveness of antimalarial drugs [37]. As a result, these patients may respond to
311 antimalarial drugs even if they are infected with drug-resistant parasites [38-40] and,
312 in such cases, it would be impossible to monitor the emergence of
313 artemisinin-resistant parasites based on clinical information alone [16]. In this regard,

the absence of either delayed-clearance associated SNPs before wide-scale implementation of artemisinin derivatives in Africa suggests that assessment of MAL10-688956 or MAL13-1718319 may allow the detection of the emergence of artemisinin resistance before the appearance of clinical failure cases.

The following points, however, should be further considered for the application of MAL10-688956 and MAL13-1718319 as molecular markers for the surveillance of the emergence of ACT resistance. In the Pacific region, prevalences of the MAL10-688956 delayed-clearance associated SNP were unexpectedly high, although nearly all parasites harbored the non-delayed clearance associated SNP at MAL13-1718319. Artemisinin combination therapies were not implemented at the time of sampling in any of the countries considered here. Hence, it seems likely that the MAL10-688956 delayed-clearance associated SNP is a parasite polymorphism that exists naturally in the parasite populations of this region, and was not, initially, selected by ACT pressure. The other possibility is the migration of the MAL10-688956 delayed-clearance associated SNP from Southeast Asia. However, since the prevalences of this SNP were high, this may be unlikely, especially in the absence of ACT selection pressure. Further analysis using microsatellite alleles flanking the SNP will clarify the possible migration of from the MAL10-688956 SNP to Southeast Asia.

Recently, another artemisinin-resistance related marker, PF3D7_1343700 kelch propeller domain ('K13-propeller') gene, has been identified using whole-genome sequencing of an artemisinin-resistant parasite line [41]. A number of SNPs (at least 17) were described in the propeller domains of K13, some of which might be predictive SNPs of resistance in different geographical settings. This marker may

prove more suitable as a moleciular marker for ACT resistant parasites than the two described here.

5. Conclusions

Parasites bearing delayed-clearance associated SNPs at both MAL10-688956 and MAL13-1718319 are localised to Cambodia and Thailand. Both the delayed-clearance associated SNPs were entirely absent from Africa or South America prior to the introduction of ACTs as first-line treatments for malaria in these countries. This study provides baseline information for geographical differences in the distribution of malaria parasites carrying delayed-clearance associated SNPs at both MAL10-688956 and MAL13-1718319, and so lays the groundwork for developing molecular markers for monitoring the emergence and spread of ACT resistance.

Acknowledgements

We thank Francis Hombhanje, Hikota Osawa, Takahiro Tsukahara, Hideaki Eto, Akira Kaneko, Hiroshi Ohmae, Masatoshi Nakamura, Lek Dysoley, Jun Kobayashi, Aung S. Marma, Willis S. Akhwale, Anders Björkman, Mathieu Ndounga, Mawuli Dzodzomenyo, and Marcelo U. Ferreira for assistance with sample collection, Nobuyuki Takahahashi, Maiko Okochi for technical assistance, and Takatoshi Kobayakawa for the organization of the study. This study was supported by a Cooperative Research Grant(s) of NEKKEN, 2013, Grants-in-aid for scientific research [23659211, 23590498] and Foundation of Strategic Research Projects in Private Universities [S0991013] from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Potential conflicts of interest for each author: none reported.

Authors' contributions

KM conducted data collection and prepared the manuscript. RC conducted a survey and made substantial corrections to the manuscript, and helped with the interpretation of the data.TH and HE organized the study. TM was involved in the study design, sampling collection, data analysis, and manuscript preparation. All authors read and approved the final manuscript.

Transparency declarations

None to declare.

Figure legends

Figure 1. Geographical distribution of *Plasmodium falciparum* SNPs associated with delayed clearance following ACT treatment at (A) MAL10-688956 (n=637) and (B) MAL13-1718319 (n=637) in 13 countries is shown. At MAL10-688956 genotypes are classified delayed-clearance associated SNP (dark green), non-delayed clearance associated SNP (white), and mixture of two SNPs (yellow-green). At MAL13-1718319, genotypes consist of, delayed-clearance associated SNP (dark blue), non-delayed clearance associated SNP (white), and mixture of two SNPs (light blue).

Figure 2. Geographical distribution of allele combinations of MAL10-688956 and MAL13-1718319 in 13 countries is shown (n=615). MAL10-688956 delayed-clearance associated SNP + MAL13-1718319 delayed-clearance associated SNP (red), MAL10-688956 delayed-clearance associated SNP + MAL13-1718319 non-delayed clearance associated SNP (green), MAL10-688956 non-delayed clearance associated SNP + MAL13-1718319 delayed-clearance associated SNP (black) and MAL10-688956 non-delayed clearance associated SNP + MAL13-1718319 non-delayed clearance associated SNP (white). Isolates harbouring mixed genotypes at either MAL10-688956 or MAL13-1718319 were excluded.

Supplementary data

Figure S1. Linkage disequilibrium between 10 microsatellite loci was assessed to clarify whether linkage disequilibrium was observed in Cambodia. *P* values of each comparison are shown.

References

[1] WHO. World Malaria Report 20132013.
[2] Noedl H, Se Y, Schaecher K, Smith BL, Socheat D, Fukuda MM, et al. Evidence of artemisinin-resistant malaria in western Cambodia. N Engl J Med. 2008;359:2619-20.
[3] Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, et al. Artemisinin resistance in Plasmodium falciparum malaria. N Engl J Med. 2009;361:455-67.
[4] Mita T, Venkatesan M, Ohashi J, Culleton R, Takahashi N, Tsukahara T, et al. Limited geographical origin and global spread of sulfadoxine-resistant dhps alleles in Plasmodium falciparum populations. J Infect Dis. 2011;204:1980-8.
[5] Mita T, Tanabe K. Evolution of Plasmodium falciparum drug resistance: implications for the development and containment of artemisinin resistance. Jpn J Infect Dis. 2012;65:465-75.
[6] Phyo AP, Nkhoma S, Stepniewska K, Ashley EA, Nair S, McGready R, et al. Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. Lancet. 2012;379:1960-6.
[7] WHO. Global plan for artemisinin resistance containment 20112011.
[8] Cheeseman IH, Miller BA, Nair S, Nkhoma S, Tan A, Tan JC, et al. A major genome region underlying artemisinin resistance in malaria. Science. 2012;336:79-82.
[9] Takala-Harrison S, Clark TG, Jacob CG, Cummings MP, Miotto O, Dondorp AM, et al. Genetic loci associated with delayed clearance of Plasmodium falciparum following artemisinin treatment in Southeast Asia. Proc Natl Acad Sci U S A. 2013;110:240-5.
[10] Ginsburg H. Progress in in silico functional genomics: the malaria Metabolic Pathways database. Trends Parasitol. 2006;22:238-40.
[11] Ginsburg H. Should chloroquine be laid to rest? Acta Trop. 2005;96:16-23.
[12] Blastyak A, Pinter L, Unk I, Prakash L, Prakash S, Haracska L. Yeast Rad5 protein required for postreplication repair has a DNA helicase activity specific for replication fork regression. Mol Cell. 2007;28:167-75.
[13] Boiteux S, Jinks-Robertson S. DNA repair mechanisms and the bypass of DNA damage in Saccharomyces cerevisiae. Genetics. 2013;193:1025-64.
[14] Hishida T, Kubota Y, Carr AM, Iwasaki H. RAD6-RAD18-RAD5-pathway-dependent tolerance to chronic low-dose ultraviolet light. Nature. 2009;457:612-5.
[15] Fidock DA. Microbiology. Eliminating malaria. Science. 2013;340:1531-3.
[16] Hunja CW, Unger H, Ferreira PE, Lumsden R, Morris S, Aman R, et al. Travellers as sentinels: Assaying the worldwide distribution of polymorphisms associated with artemisinin combination therapy resistance in Plasmodium falciparum using malaria cases imported into Scotland. Int J Parasitol. 2013;43:885-9.
[17] Wootton JC, Feng X, Ferdig MT, Cooper RA, Mu J, Baruch DI, et al. Genetic diversity and chloroquine selective sweeps in Plasmodium falciparum. Nature. 2002;418:320-3.
[18] Roper C, Pearce R, Nair S, Sharp B, Nosten F, Anderson T. Intercontinental spread of pyrimethamine-resistant malaria. Science. 2004;305:1124.

443 [19] Marma AS, Mita T, Eto H, Tsukahara T, Sarker S, Endo H. High prevalence of
444 sulfadoxine/pyrimethamine resistance alleles in Plasmodium falciparum parasites
445 from Bangladesh. Parasitol Int. 2010;59:178-82.
446 [20] Dysoley L, Kaneko A, Eto H, Mita T, Socheat D, Borkman A, et al. Changing
447 patterns of forest malaria among the mobile adult male population in Chumkiri District,
448 Cambodia. Acta Trop. 2008;106:207-12.
449 [21] Kobayashi J, Phompida S, Toma T, Looareensuwan S, Toma H, Miyagi I. The
450 effectiveness of impregnated bed net in malaria control in Laos. Acta Trop.
451 2004;89:299-308.
452 [22] Sakihama N, Nakamura M, Palanca AA, Jr., Argubano RA, Realon EP, Larracas
453 AL, et al. Allelic diversity in the merozoite surface protein 1 gene of Plasmodium
454 falciparum on Palawan Island, the Philippines. Parasitol Int. 2007;56:185-94.
455 [23] Sakihama N, Mitamura T, Kaneko A, Horii T, Tanabe K. Long PCR amplification
456 of Plasmodium falciparum DNA extracted from filter paper blots. Exp Parasitol.
457 2001;97:50-4.
458 [24] Akhwale WS, Lum JK, Kaneko A, Eto H, Obonyo C, Bjorkman A, et al. Anemia
459 and malaria at different altitudes in the western highlands of Kenya. Acta Trop.
460 2004;91:167-75.
461 [25] Tanabe K, Sakihama N, Rooth I, Bjorkman A, Farnert A. High frequency of
462 recombination-driven allelic diversity and temporal variation of Plasmodium
463 falciparum msp1 in Tanzania. Am J Trop Med Hyg. 2007;76:1037-45.
464 [26] Tsumori Y, Ndounga M, Sunahara T, Hayashida N, Inoue M, Nakazawa S, et al.
465 Plasmodium falciparum: differential selection of drug resistance alleles in contiguous
466 urban and peri-urban areas of Brazzaville, Republic of Congo. PLoS One.
467 2011;6:e23430.
468 [27] Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, do Rosario VE, et al.
469 High sensitivity of detection of human malaria parasites by the use of nested
470 polymerase chain reaction. Mol Biochem Parasitol. 1993;61:315-20.
471 [28] Culleton RL, Mita T, Ndounga M, Unger H, Cravo PV, Paganotti GM, et al. Failure
472 to detect Plasmodium vivax in West and Central Africa by PCR species typing. Malar
473 J. 2008;7:174.
474 [29] Anderson TJ, Haubold B, Williams JT, Estrada-Franco JG, Richardson L,
475 Mollinedo R, et al. Microsatellite markers reveal a spectrum of population structures in
476 the malaria parasite Plasmodium falciparum. Mol Biol Evol. 2000;17:1467-82.
477 [30] Lewontin R. The interaction of selection and linkage. I. General considerations;
478 heterotic models. Genetics. 1964;49:49.
479 [31] Hill W, Robertson A. Linkage disequilibrium in finite populations. Theoretical and
480 Applied Genetics. 1968;38:226-31.
481 [32] Noedl H. Artemisinin resistance: how can we find it? Trends Parasitol.
482 2005;21:404-5.
483 [33] Peterson DS, Walliker D, Wellems TE. Evidence that a point mutation in
484 dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in
485 falciparum malaria. Proc Natl Acad Sci U S A. 1988;85:9114-8.
486 [34] Triglia T, Cowman AF. Primary structure and expression of the dihydropteroate
487 synthetase gene of Plasmodium falciparum. Proc Natl Acad Sci U S A.
488 1994;91:7149-53.

[35] Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, et al. Analysis of Plasmodium falciparum diversity in natural infections by deep sequencing. Nature. 2012;487:375-9.
[36] Hartl DL, Clark AG. Principles of population genetics. 4 ed. Sunderland: Sirauer Associates, INC; 2007.
[37] Rogerson SJ, Wijesinghe RS, Meshnick SR. Host immunity as a determinant of treatment outcome in Plasmodium falciparum malaria. Lancet Infect Dis. 2010;10:51-9.
[38] Cravo P, Culleton R, Hunt P, Walliker D, Mackinnon MJ. Antimalarial drugs clear resistant parasites from partially immune hosts. Antimicrobial agents and chemotherapy. 2001;45:2897-901.
[39] Djimde A, Doumbo OK, Cortese JF, Kayentao K, Doumbo S, Diourte Y, et al. A molecular marker for chloroquine-resistant falciparum malaria. N Engl J Med. 2001;344:257-63.
[40] Dorsey G, Kamya MR, Singh A, Rosenthal PJ. Polymorphisms in the Plasmodium falciparum pfcrt and pfmdr-1 genes and clinical response to chloroquine in Kampala, Uganda. J Infect Dis. 2001;183:1417-20.
[41] Ariey F, Witkowski B, Amaratunga C, Beghain J, Langlois AC, Khim N, et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. Nature. 2014;505:50-5.

Figure1A

Figure 1A

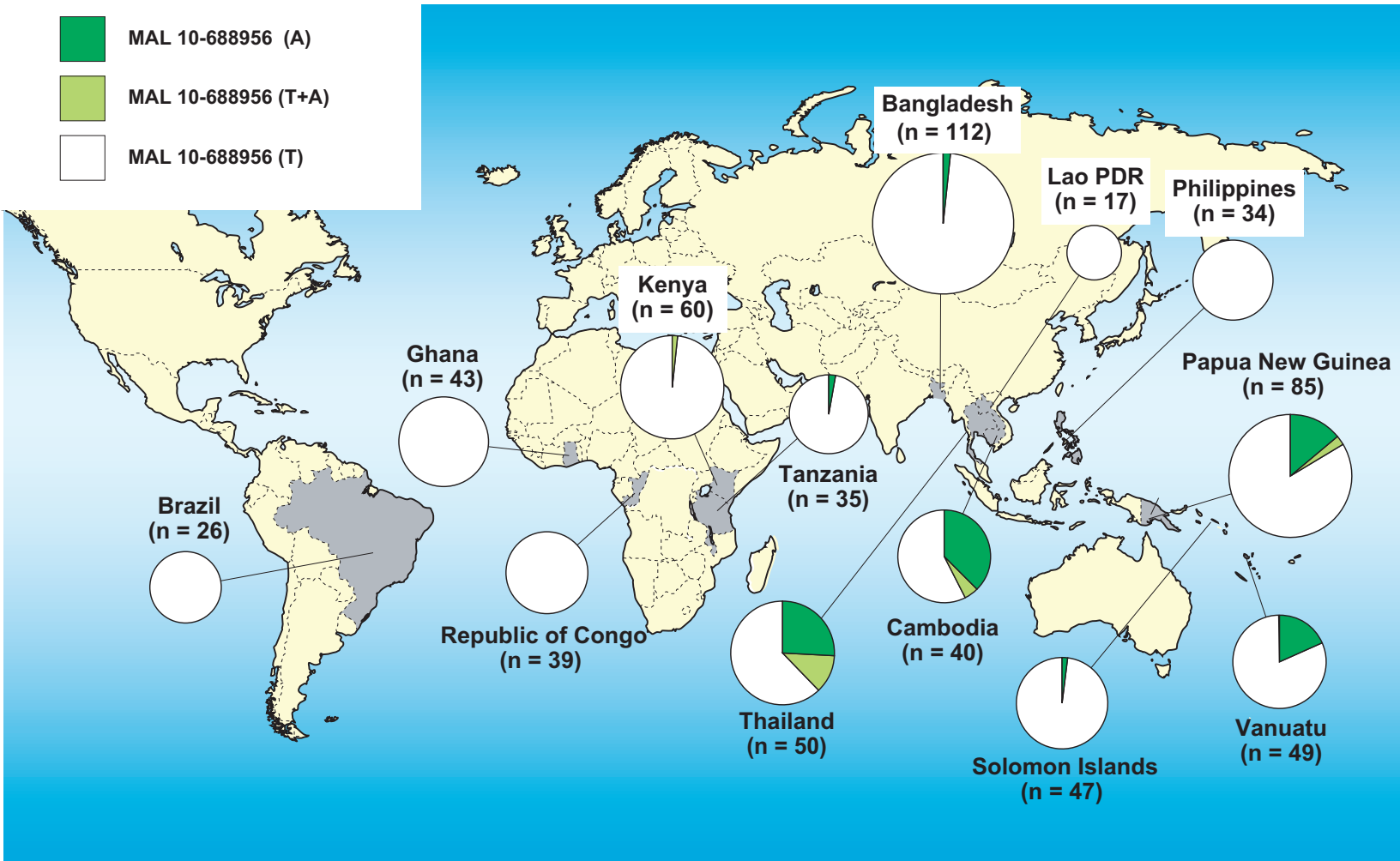

Figure1B

Figure 1B

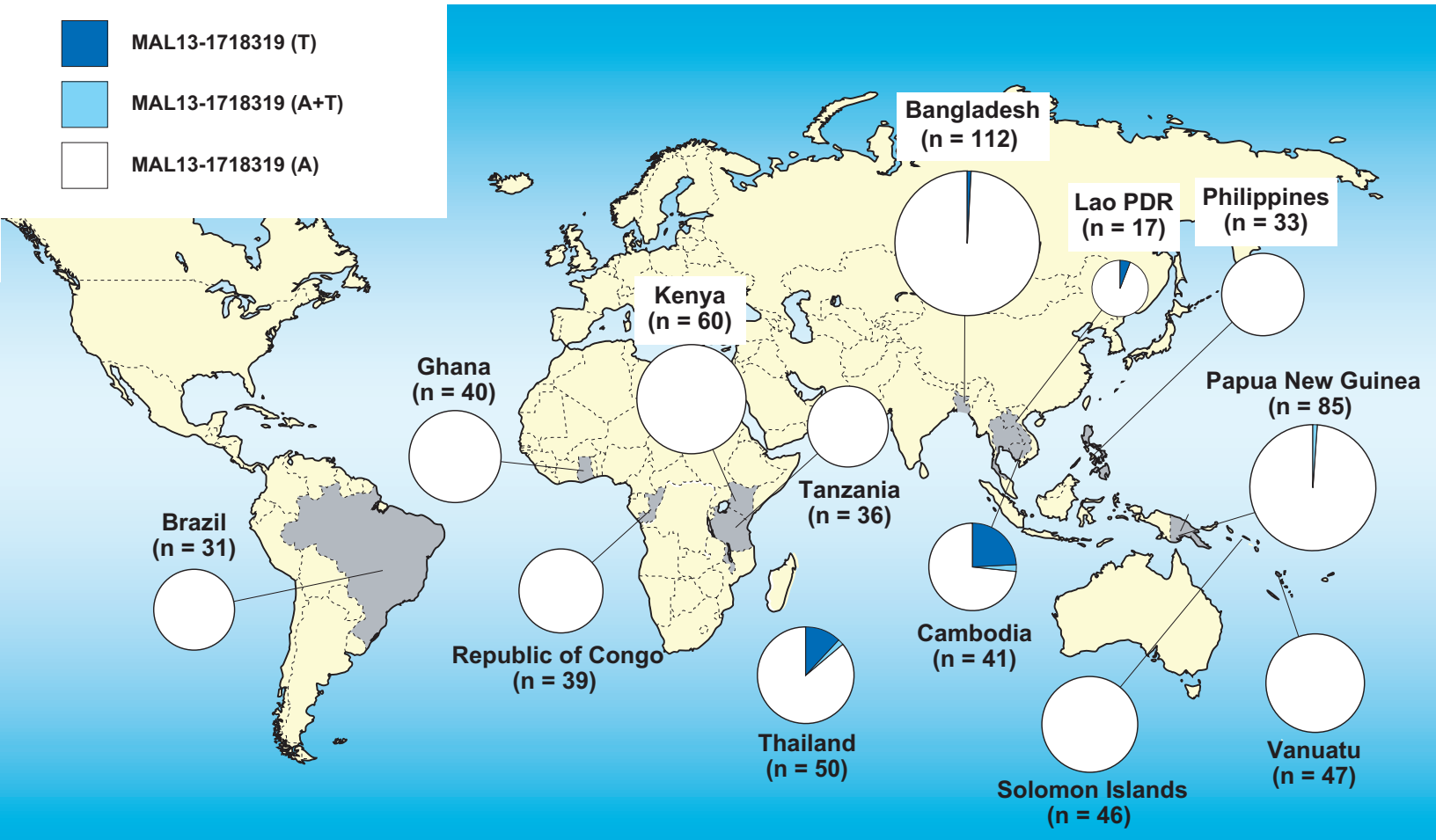

Figure2

Figure 2

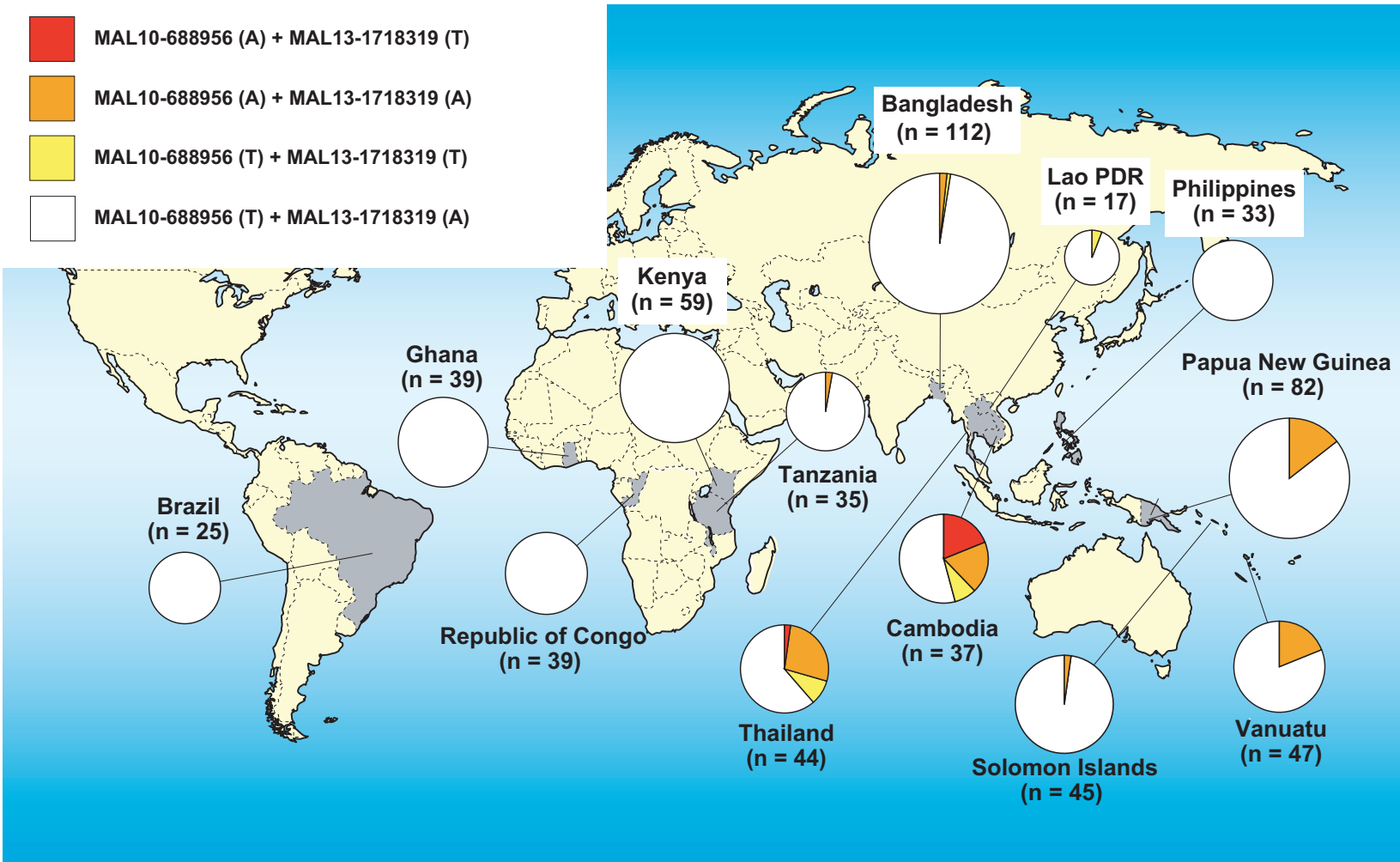

Table 1. *Plasmodium falciparum* isolates from 13 countries used in this study

Country	Area	No. of isolates	Year of sampling
Southeast Asia			
Bangladesh	Bandarban	112	2007
Cambodia	Chumkiri, Kampot	43	2004, 2005, 2006
Thailand	Tak, Kanchanaburi, Ratchaburi	50	2001-2
LaoPDR	Khammouanne	17	1999
Philippines	Palawan Island	34	1997
Pacific Oceania			
Papua New Guinea	East Sepik	85	2002, 2003
Solomon Islands	Guadalcanal Island	48	1995-6
Vanuatu	Gaua, Santo, Pentecost, Malakula	50	1996, 1998
Africa			
Kenya	Kisii	62	1998
Tanzania	Rufiji River Delta	36	1998, 2003
Republic of Congo	Pointe-Noire, Brazzaville, Gamboma	40	2006
Ghana	Winneba	44	2004
South America			
Brazil	Acre	32	1985-6, 1999, 2004-5
Total		**653**	

Table 2. Linkage disequilibrium in the *Plasmodium falciparum* between MAL10-688956 and MAL13-1718319 in Bangradesh, Cambodia and Thailand.

Country	n	D'	r^2	*p* value[a]
Bangladesh	112	ND	0.00016	0.8491
Cambodia	37	0.5174	**0.1629**	**0.0146**
Thailand	44	0.3231	0.00561	0.6074

ND, not determined.

[a] $\chi 2$ test.

Isolates in which allele types were not determined were excluded (n=4), and then isolates harbouring mixed SADCAT/non-SADCAT in either MAL10-688956 or MAL13-1718319 were excluded (n= 8).

*Graphical abstract (for review)