Genetic profiling of *Plasmodium ovale wallikeri* relapses with microsatellite markers and whole-genome sequencing

Authors: Valentin Joste[1,2#], Emma Colard-Itté[2], Émilie Guillochon[1], Frédéric Ariey[3], Romain Coppée[4], Jérôme Clain[1,2] and Sandrine Houzé[1,2] on behalf of the French National Reference Center for Imported Malaria Study Group*

1. Université Paris Cité, IRD, MERIT, F-75006 Paris, France

2. Centre National de Référence du Paludisme, AP-HP, Hôpital Bichat – Claude-Bernard, F-75018 Paris, France

3. INSERM U1016, Institut Cochin, Laboratoire de parasitologie-mycologie, Hôpital Cochin, AP-HP, Université de Paris, Paris, France

4. Université Paris Cité & Sorbonne Paris Nord, Inserm, IAME, F-75018 Paris, France

* The Investigation study group members are listed in the Acknowledgments.

[#]Corresponding author: Valentin Joste (valentinjoste@gmail.com), phone: +33617435543

Alternate corresponding author: Sandrine Houzé (sandrine.houze@aphp.fr)

Running title: Profile of *P. ovale wallikeri* relapses

Abstract count: 131 words

Main text count: 3,486 words

Abstract

Like *Plasmodium vivax*, both *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* have the ability to cause relapse in humans, defined as recurring asexual parasitaemia originating from liver dormant forms subsequent to a primary infection. Here, we investigated relapse patterns in *P. ovale wallikeri* infections from a cohort of travelers who were exposed to the parasite in Sub-Saharan Africa and then experienced relapses after their return to France. Using a novel set of eight highly polymorphic microsatellite markers, we genotyped 15 *P. ovale wallikeri* relapses. For most relapses, the paired primary and relapse infections were highly genetically related (with 12 being homologous), an observation that was confirmed by whole-genome sequencing for the four relapses we further studied. This is, to our knowledge, the first genetic evidence of relapses in *P. ovale* spp.

Keywords: *Plasmodium ovale wallikeri*; microsatellites; relapses; WGS; sWGA

Main article

Introduction

Malaria is one of the deadliest infectious diseases, especially in Sub-Saharan Africa with approximatively 619,000 deaths in 2021 [1]. While *Plasmodium falciparum* is the best characterized species in Africa with a well-described epidemiology [1], *P. ovale* spp. are also endemic in Africa [2] but their exact prevalence are still imperfectly established and have been probably largely underestimated [3]. The increased use of qPCR (quantitative Polymerase Chain Reaction) makes the diagnosis of non-*falciparum* or *Plasmodium* co-infections easier [4] and will help in the future to better define the *P. ovale* spp. prevalence in endemic area.

P. ovale wallikeri and *P. ovale curtisi* were recently classified as two different species based on distinct genetic patterns [5]. These species are sympatric and display different clinical characteristics that are well-described in imported malaria patients, such as deeper thrombocytopenia and shorter latency period for *P. ovale wallikeri* compared to *P. ovale curtisi* [6,7]. Like *P. vivax*, both *P. ovale curtisi* and *P. ovale wallikeri* have the ability to cause relapse [6,8], defined as recurring asexual parasitaemia originating from liver dormant forms subsequent to a primary infection [9]. Despite *P. ovale* spp. relapses are rarer than in *P. vivax* infections, primaquine has demonstrated its ability to reduce their prevalence in endemic settings [10], but not in non-endemic regions [11]. *P. ovale* spp. relapses are still understudied and the development of genomic tools are necessary to study them. We recently developed a method to selectively amplify *P. ovale* spp. genomes over the human genome [12], permitting whole-genome sequencing (WGS) of these parasites by next-generation sequencing (NGS) from clinical blood samples. This is a promising approach but WGS remains expensive and requires bioinformatics skills not available everywhere. Therefore, a set of easy to type and highly polymorphic markers would complement the genetic tools to study *P. ovale* spp. population genetics.

Microsatellite loci are useful molecular markers to study the population genetics and epidemiology of malaria because they exhibit high level of polymorphism, are generally neutral, and are frequently found in eukaryote genomes. While approximatively 30,000 microsatellites are found in their genomes [13,14], microsatellite typing protocols are missing for both *P. ovale* spp. species, contrarily to other human-infecting *Plasmodium* species [15–18].

In this study, we investigated relapse patterns in *P. ovale wallikeri* infections from a cohort of travelers who were exposed to the parasite in Sub-Saharan Africa and then experienced relapses after their return to France. Using a novel set of eight highly polymorphic

microsatellite markers, we genotyped 15 pairs of isolates from primary (D_0) and relapsing infections ($D_{relapse}$).

Methods

Ethical statement

No specific consent from patients was required since clinical and biological data were collected from the FNMRC database in accordance with the common public health mission of all National Reference Centers in France, in coordination with the Santé Publique France organization for malaria surveillance and care. The study was considered as non-interventional research according to article L1221–1.1 of the public health code in France and only requires the non-opposition of the patient (per article L1211–2 of the public health code).

Relapse definition & sample collection

A *P. ovale wallikeri* relapse infection ($D_{relapse}$) in the context of imported malaria was defined as asexual parasitaemia recurring 28 days or more after a primary *P. ovale wallikeri* infection (D_0) treated in France and without any reported trip to a malaria-endemic country between the two infections, thus excluding reinfections. Epidemiological, clinical and biological data were collected through an online patient form completed by French National Malaria Reference Center (FNMRC) correspondents. All isolates' pairs collected between 2013 and 2021 that fit those requirements were included in the study. We also included *P. ovale wallikeri* isolates from the FNMRC DNA bank to measure the genetic diversity of each microsatellite marker.

Genomic DNA was extracted from 200 µL of whole blood using MagNA Pure automaton (Roche diagnostics, United States of America) and eluted in 100 µL. *P. ovale* spp. infections were confirmed using the *Plasmodium* typage kit (Bio-Evolution, France). In-house qPCR-

High Resolution Melting (HRM) was used to differentiate *P. ovale wallikeri* from *P. ovale curtisi* [19].

Microsatellite development and analysis

We used the PowCR01 reference chromosomes (GenBank assembly accession: GCA_900090025.2) and the software Tandem Repeat Finder [20] to identify genome-wide microsatellites. The criteria for microsatellite selection were restricted to perfect or imperfect microsatellite with unit motifs of 2-10 base pairs (bp) and a minimum threshold of 10 repeats. Once the markers were selected *in silico*, several ones were experimentally tested and further selected based on (i) efficacy of the PCR; (ii) absence of nonspecific bands on agarose gel; (iii) location on different chromosomes; and (iv) easy-to-interpret fragment analysis profiles. Forward primers were marked in 5' with 6-FAM or HEX fluorophores (Eurogentec®) and when useful, reverse primers were pig-tailed to reduce non-templated addition of a nucleotide to the 3' end of the amplicon [21]. We used touchdown-PCR to increase sensibility and specificity of the amplification [22]. The PCR mix, cycle and microsatellite analysis are detailed in the Supplemental methods.

Measurement of genetic diversity

We measured allele frequencies at each marker using all the isolates ($D_{relapse}$ identical to D_0 excluded) that were successfully genotyped. We calculated the expected heterozygosity (H_E) for each marker using the following formula: $H_E = \left[\frac{n}{n-1}\right]\left[1 - \sum_{i=1}^{n} p^2\right]$ where p is the frequency of the ith allele and n the number of different alleles [23]. We counted the number of alleles (A) at each marker in each country for which more than ten isolates were successfully genotyped. The probability of having two isolates sharing the same allele(s) at all

markers was calculated by combining individual probabilities of each marker following this equation: $\pi p_i = P_1 \times P_2 \times (\ldots) \times P_i$ where $P_i = \sum p_i^2$ with the assumption of independent infections [17]. The discriminatory power of the microsatellite haplotype was also evaluated with the Simpson index of diversity (D): $D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{S} nj(nj-1)$ where N is the total number of isolates within the population, s is the total number of different haplotypes described and nj is the number of isolates belonging to the jth haplotype.

Polyclonal isolates were defined as isolates with at least two scored peaks for at least one microsatellite marker. Multiplicity of infection (MOI) was calculated for each isolate and averaged across markers.

Genetic relationship analysis

We built a matrix of the allele count (only the major peak was considered for polyclonal isolates) using the *adegenet* R package [24]. Principal Component Analyses (PCA) were made with *ade4* and plotted with the *factoextra* R package [25]. Phylogenetic tree was built by the neighbor-joining method with the *poppr* R package [26] from a matrix of Bruvo's distance [27] and plotted with the interactive tree of life (iTOL) software [28].

Comparison with microsatellite markers of primary (D_0) and relapse ($D_{relapse}$) infections

We compared the microsatellite haplotypes obtained for the primary and the relapse infections. We classified the two isolates within each pair according to three categories: (i) homologous isolates: all of the alleles detected in the primary infections are detected in the relapse infections; (ii) heterologous isolates: no shared allele(s) between the primary and relapse infections for at least one microsatellite marker; (iii) related isolates: when the primary infection is polyclonal, the D_0 and the $D_{relapse}$ isolates shared alleles at each marker but some extra-allele(s) is detected only in the primary isolate.

Selective whole-genome amplification, whole-genome sequencing and variant calling analysis

To confirm the results obtained for relapses with microsatellite typing, genome-wide single nucleotide polymorphisms (SNPs) obtained from WGS were compared for four D_0-$D_{relapse}$ isolate pairs. Prior to WGS, *P. ovale wallikeri* DNA was enriched following a sWGA strategy and specific primers [12]. Both the sWGA protocol and the WGS procedure are fully detailed in the Supplemental methods.

After WGS, raw reads were mapped to the *P. ovale wallikeri* reference genome (GCA_900090025.2) and duplicate reads were removed using Picard MarkDuplicates (v. 2.26.10). Genome-wide SNPs were identified using BCFtools *mpileup* version 1.13 following previously published quality criteria [12]. We only kept positions with an allelic depth ≥ 5 for the reference and/or the alternative allele. We defined as homozygous a position with ≥ 70% of reads for the reference (homozygous wild-type) or the alternative allele (homozygous mutant) [29]. Otherwise, positions were defined as heterozygous. For genome-wide SNP comparisons, only homozygous SNPs located within the fourteen chromosomes of the nuclear genome were analyzed. Homozygous SNPs located on the contigs and heterozygous SNPs were filtered out.

Genome-wide SNP comparisons were done for each D_0 and $D_{relapse}$ pair and the proportion of identical genotypes between positions included for both isolates were calculated. We used several WGS data from previous published work [12] or previous unpublished sequencing as controls. The first control was an isolate sequenced twice (either with a sWGA or leukodepleted approach) and named “Matched control pair”. The two other controls were built with one isolate from a relapse pair and one isolate previously sequenced from the same country with no common history and named “Unmatched control pairs”. Finally, to increase

the number of comparisons, SNPs from D_0 and $D_{relapse}$ isolates from different pairs were compared and called "Unmatched infection pairs".

Data analysis

All statistical analyses and graphs were performed on R software v.4.0.3 [30]. Quantitative results were expressed as median [Interquartile Interval] or mean +/- Standard Deviation (SD). Spearman's rank test was used to assess correlations. Proportions were compared using the $\chi 2$ or Fisher's exact tests. We used the Mann-Whitney *U* test or the Wilcoxon signed-rank test to compare medians.

Results

Relapse epidemiology

Seventeen pairs of isolates from 2013 to 2021 met the requirements of relapse definition. Patients with relapse mainly lived in metropolitan France and experienced a primary malaria infection after a short travel to Africa (Table S1). Median time between the D_0 and the $D_{relapse}$ infections was 108 [60-145] days with three relapses that occurred less than 50 days after the primary infection. Delay between D_0 and $D_{relapse}$ was not different whether the primary infection occurred during high (from July to October) or low transmission malaria season (60 days [46-107] *vs* 121 days [86-142]; p=0.19, Mann-Whitney *U*-test) nor in chloroquine or non-chloroquine treated-patients (61 days [50-155] *vs* 128 days [66-142]; p=0.39, Mann-Whitney *U*-test). Parasitological follow-up as well as the treatment of the first infections are reported Table S2.

Sample collection

In order to evaluate the genetic diversity of the selected microsatellite markers, 52 isolates were added in addition to the isolates from relapsing pairs. Those "extra"-isolates were selected to cover several countries in Sub-Saharan Africa and large range of parasite density to evaluate the PCR success rate, especially for low parasite density infections. Sixty-nine isolates were included for microsatellite diversity evaluation ($D_{relapse}$ identical to D_0 were excluded). They were originated from West or Central Africa (16 different countries) and all were *P. ovale wallikeri* mono-infections (Figure 1). Three isolates were from patients who traveled through multiple African countries (not shown on the map). Ivory Coast (n=15 isolates), Cameroon (n=14 isolates) and Guinea (n=12 isolates) were the most represented countries. Median parasite density was 3,836 [555-9,474] parasites/µL (p/µL).

<u>Microsatellite marker development and validation</u>

We developed one-round PCR protocols for eight microsatellite markers, each located on a different chromosome of *P. ovale wallikeri* (Table 1 and Figure S1) and genotyped 69 *P. ovale wallikeri* isolates. Three markers (named 5.2, 11.1 and 14.2) displayed a sensitivity of 100% (no PCR failure; Table 2). The remaining five markers displayed a sensitivity ranging from 94 to 99%. Failed amplifications were associated with a lower parasite density compared to successful amplifications: 6.6% of PCR failures (n=9/136) for parasite densities ≤ 550 p/µL *vs* 1% (n=4/416) for parasite densities > 550 p/µL (p=0.001, Fisher's exact test). Sixty-one samples had a complete haplotype with the eight microsatellites that were successfully genotyped and the remaining had seven (n=5 isolates), six (n=1 isolates) or five (n=2 isolates) microsatellites successfully genotyped. Nonspecific amplification was observed for some markers when tested against *P. falciparum* (marker 2.1) and against *P. malariae* and *P. vivax* (markers 2.1, 8.3 and 14.2). It was significantly decreased when annealing temperature was increased or when DNA templates contained also *P. ovale wallikeri* (Figure S2). As mono-

specific *P. ovale wallikeri* infections were analyzed in this study, this caveat had no impact on our results.

Genetic diversity

From the 69 isolates investigated, we counted from 9 to 17 different alleles per microsatellite (Table 2) and 65 eight-marker haplotypes were unique (Table S3). The expected heterozygosity H_E was calculated for each marker and varied from 0.764 to 0.940. For each microsatellite, allele frequencies were calculated in the total cohort (Figure 2 and Table S4) and in Cameroon, Guinea and Ivory Coast (Table S4). The combined likelihood that two random samples had the same eight-markers haplotype was 2.1×10^{-6} (Table S5). The Simpson index of diversity was equivalent (D=0.9987) using 8 or only 6 markers (2.1 and 10.1 markers excluded, the two loci with the lower H_E). D value remained high with the four markers with the highest H_E (D=0.9978; combination of the 8.3, 11.1, 12.1 and 14.2 markers), highlighting their combined high discriminatory power.

The diversity (based on the mean number of alleles, A, and on the mean expected heterozygosity, H_E) was found to be high in Cameroon and Guinea (Table S6). Interestingly, five of the 69 isolates displayed two peaks for at least one marker and were considered as polyclonal (Table S3**)**.

Population structure analysis

We next studied population structure between West (n=44) and Central African (n=21) isolates, and between isolates from Cameroon (n=14), Guinea (n=12) and Ivory Coast (n=15), respectively. In both analyses, PCA based on microsatellite allele sizes was not able to separate isolates according to their geographic origin (Figure S3). Similarly, the phylogenetic

tree reconstructed using eight-marker haplotypes was not able to separate the isolates from West and Central Africa or from Cameroon, Guinea and Ivory Coast (Figure S4 and S5).

Application of the microsatellite markers for relapse typing

Among the 17 D_0-$D_{relapse}$ isolate pairs, two were not analyzable because of PCR failures at more than four markers, likely due to very low parasite density (Table S7). Twelve pairs had homologous parasites (Table S7). Two pairs (pairs 14 and 17) had at least one marker with multiple peaks for the D_0 but a single shared peak for the $D_{relapse}$ and were then considered as related but not strictly homologous. For one pair (pair 14), the $D_{relapse}$ was monoclonal. For the other pair (pair 17), both D_0 and $D_{relapse}$ displayed multiple peaks for respectively 4 and 1 microsatellite markers. Finally, the pair 15 had heterologous parasites, with no shared allele for the marker 5.2 (allele sizes were 253 and 250 for D_0 and $D_{relapse}$ respectively, corresponding to one trinucleotide TAC motif).

Genome-wide SNP analysis

To complete the microsatellite results, we performed WGS for four homologous D_0-$D_{relapse}$ isolate pairs (Table S8) and counted the number of SNPs that differed within each pair. For the matched control pair (same isolate sequenced twice), 99.93% of the SNPs were identical, compared to 31.7% for the unmatched control pairs and 30.7% for the unmatched infection pairs (Figure 3 and Table S9). For the four D_0-$D_{relapse}$ pairs, (pairs 1, 2, 4 and 5; see Table S8) an average of 99.64% SNPs were identical (Figure 3) with very few different SNPs within pairs (Figure 4). No statistical difference in the count of identical SNPs was observed between the matched control pair and the relapses (p=0.93, $\chi 2$ test). In contrast, the number of identical SNPs statistically differed between the unmatched control pairs or the unmatched infection pairs and the relapses (p<0.001, $\chi 2$ test).

Discussion

Despite several epidemiological evidences of *P. ovale* spp. relapses during the last seventy years [6,8,11,31–33], biological realness of *P. ovale* spp. hypnozoites is still debated [9,34]. Indeed, proofs of hypnozoite existence are scarce, but Soulard *et al*. [35] described late-developing schizontes in humanized mice infected with *P. ovale* spp. sporozoites, consistent with a hypnozoite origin of those maturing forms [35].

Easy-to-use molecular tools are missing to study the genetic relatedness of *P. ovale* spp. isolates. We developed eight microsatellite markers that displayed very high H_E and high combined probability value (Table 2), suggesting a great discriminative potential for epidemiological studies. Interestingly, the combination of just four markers (8.3, 11.1, 12.1 and 14.2) remains highly informative. These could be used as a minimal set to genetically profile *P. ovale wallikeri* isolates for time and resource saving. We observed high diversity with 65 unique haplotypes and only 3 non-unique haplotypes each shared by two isolates from West Africa that did not have reported common history (Figure S4 and Table S3). More studies, in particular at a country level, are needed to better investigate *P. ovale wallikeri* population structure in West and Central Africa.

With these eight newly developed microsatellite markers, we successfully genotyped 15 pairs of *P. ovale wallikeri* isolates that fitted with an epidemiological relapse definition. As our study is the first one that genetically analyzed *P. ovale wallikeri* relapses, we used *P. vivax* relapses as a reference. Our relapse definition is purely based on the epidemiology and history of cases as we did not have complete parasitemia follow-up for each patient neither formal evidence of perfect compliance to full treatment. As the World Health Organization recommends a 28 days parasitaemia follow-up for *P. vivax* [36], we hypothesized that infections with time to recurrence above this 28-days duration threshold were relapses. Of

note, even with precise molecular tools, a homologous relapse infections cannot be formally distinguished from a recrudescent infection caused by incomplete treatment clearance of the blood-stage primary infection. We decided to classify those second infections as homologous relapse and not recrudescence because: (i) in *P. vivax*, recrudescences seem to be much less frequent than relapses [37,38]; (ii) to the best we know, *P. ovale* spp. treatment failure is rarely reported in the literature (no failure reported in published clinical trials [39–43], except one persistent *P. ovale wallikeri* on day 28 in a *P. falciparum* co-infection in Gabon [44]); (iii) except for three pairs for which the time to recurrence was short (28, 33 and 39 days), this time was much longer for the other pairs (from 53 to 257 days) which is not in favor of a recrudescence. In addition, those patients were not living in endemic areas and therefore had likely not acquired sufficient immunity to maintain for a long time a low-level malaria infection with *P. ovale* spp. A prospective study, with WHO parasitological follow-up, directly observed treatment and plasma drug dosage is however needed to further confirm our conclusions.

Among the 15 analyzable pairs, 12 were scored as genetically homologous. For four out of these 12 genetically homologous pairs also tested by WGS, the very high genetic relatedness of D_0 and $D_{relapse}$ isolates was confirmed. Contrary to most of *P. vivax* relapses studied in Thailand where heterologous relapses are the norm [45] (exception made of the relapses observed in young infants that were mainly homologous [46]), we report high percentage of homologous relapses. This has to be interpreted in the light of two key features of our study: (i) we studied imported malaria cases; therefore patients mainly live in metropolitan France with short travel to Africa limiting the probabilities, compared to patients living in endemic area, of a previous *P. ovale wallikeri* infection or multiple infections (two ways to get heterologous relapses [47,48]); and (ii) according to the microsatellite results reported here, the polyclonal infections (second way to get heterologous relapses [45]) rate in *P. ovale*

wallikeri isolates from imported malaria appears to be much less frequent compared to the rate reported in *P. vivax* (10% of polyclonal infection for *P. ovale wallikeri* in our work *vs* a mean proportion of polyclonal infections of 43% for *P. vivax* in endemic zone [49]).

The genetic relatedness within the three last pairs (14, 15 and 17; Table S7) was more complex. For pairs 14 and 17, D_0 isolates were polyclonal while $D_{relapse}$ carried no new alleles, confirming their high genetic relatedness. Interestingly, $D_{relapse}$ from pair 14 was monoclonal contrary to relapse from pair 17 that was polyclonal. This may result from different ways of hypnozoite reactivation, stochastic for pair 14 (*i.e.* one hypnozoite reactivation) and external stimulus for pair 17 (*i.e.* at least two hypnozoites reactivation)[47]. Finally, the two isolates from pair 15 were both monoclonal and differed at a single marker by just one tandem repeat. Although initially classified as heterologous, the new allele in $D_{relapse}$ could be simply derived by mutation of the D_0 allele early during the course of the relapse infection. Slippage during the PCR was excluded since PCR was repeated.

In conclusion, using a novel set of eight microsatellite markers we genotyped for the first time *P. ovale wallikeri* isolates from primary and relapse infections in the context of imported malaria and low multiplicity of infections. As expected, most paired primary and relapse isolates appeared to be highly genetically related. WGS should help in the future to get a more complete genetic comparison of D_0-$D_{relapse}$ isolate pairs and also to look for putative genetic markers associated with *P. ovale wallikeri* dormancy.

Declarations

Data availability

The genome Illumina sequencing reads from the *P. ovale wallikeri* samples were deposited in the European Nucleotide Archive under the accession number PRJEB51041.

Authors' contributions

VJ, SH and JC designed the study. VJ and ECI performed the experimentations. VJ, ECI, EG and RC analyzed the data. VJ, EG and RC performed the bioinformatic analyses. VJ wrote the article and all authors reviewed the manuscript. SH supervised the study. There is no conflict of interests to disclose.

Acknowledgments

We thank all collaborating individuals for their participation in collecting materials and data (referred as the French National Reference Center for Imported Malaria Study Group), listed in alphabetical order (last name): Aboubacar Ahmed, Angebault Cécile, Angoulvant Adela, Argy Nicolas, Azjenberg Daniel, Belkadi Ghania, Bellanger Anne-Pauline, Bemba Dieudonné, Blaize Marion, Botterel Françoise, Bougnoux Marie-Elisabeth, Brun Sophie, Buret Bernadette, Chevrier Sylviane, Clauser Sylvain, Dahane Naima, Dannaoui Eric, Dard Cécile, Dardé Marie-Laure, de Gentile Ludovic, de Suremain Nathalie, Debourgogne Anne, Delaval Anne, Deleplancque Anne-Sophie, Desoubeaux Guillaume, Durand Rémy, Durieux Marie-Fleur, Dutoit Emmanuel, Eloy Odile, Fenneteau Odile, Gargala Gilles, Godineau Nadine, Guennouni Nadia, Guinard Jérôme, Hamane Samia, Herault Etienne, Larreché Sébastien, Lavergne Rose-Anne, Marteau Anthony, Mazars Edith, Moreno-Sabater Alicia, Morio Florian, Nourrisson Céline, Perraud-Cateau Estelle, Pons Denis, Pull Lauren, Quinio Dorothée, Raffenot Didier, Silva Muriel, Thellier Marc, Tielli Alexandra, Toubas Dominique.

Bibliography

1. World malaria report 2022 [Internet]. [cited 2022 Dec 11]. Available from: https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022

2. Hawadak J, Dongang Nana RR, Singh V. Global trend of Plasmodium malariae and Plasmodium ovale spp. malaria infections in the last two decades (2000–2020): a systematic review and meta-analysis. Parasites & Vectors. **2021**; 14(1):297.

3. Yman V, Wandell G, Mutemi DD, et al. Persistent transmission of Plasmodium malariae and Plasmodium ovale species in an area of declining Plasmodium falciparum transmission in eastern Tanzania. PLOS Neglected Tropical Diseases. Public Library of Science; **2019**; 13(5):e0007414.

4. Groger M, Lutete GT, Mombo-Ngoma G, et al. Effectiveness of pyronaridine-artesunate against Plasmodium malariae, Plasmodium ovale spp, and mixed-Plasmodium infections: a post-hoc analysis of the CANTAM-Pyramax trial. The Lancet Microbe. Elsevier; **2022**; 3(8):e598–e605.

5. Sutherland CJ, Tanomsing N, Nolder D, et al. Two nonrecombining sympatric forms of the human malaria parasite Plasmodium ovale occur globally. J Infect Dis. **2010**; 201(10):1544–1550.

6. Joste V, Bailly J, Hubert V, et al. Plasmodium ovale wallikeri and P. ovale curtisi Infections and Diagnostic Approaches to Imported Malaria, France, 2013-2018. Emerg Infect Dis. **2021**; 27(2).

7. Rojo-Marcos G, Rubio-Muñoz JM, Angheben A, et al. Prospective comparative multi-centre study on imported Plasmodium ovale wallikeri and Plasmodium ovale curtisi infections. Malar J. **2018**; 17(1):399.

8. Groger M, Veletzky L, Lalremruata A, et al. Prospective Clinical and Molecular Evaluation of Potential Plasmodium ovale curtisi and wallikeri Relapses in a High-transmission Setting. Clin Infect Dis. **2019**; 69(12):2119–2126.

9. Markus MB. Do hypnozoites cause relapse in malaria? Trends Parasitol. **2015**; 31(6):239–245.

10. Robinson LJ, Wampfler R, Betuela I, et al. Strategies for understanding and reducing the Plasmodium vivax and Plasmodium ovale hypnozoite reservoir in Papua New Guinean children: a randomised placebo-controlled trial and mathematical model. PLoS Med. **2015**; 12(10):e1001891.

11. Wångdahl A, Sondén K, Wyss K, et al. Relapse of Plasmodium vivax and Plasmodium ovale Malaria With and Without Primaquine Treatment in a Nonendemic Area. Clinical Infectious Diseases. **2022**; 74(7):1199–1207.

12. Joste V, Guillochon E, Clain J, Coppée R, Houzé S. Development and Optimization of a Selective Whole-Genome Amplification To Study Plasmodium ovale Spp. Microbiol Spectr. **2022**; :e0072622.

13. Rutledge GG, Böhme U, Sanders M, et al. Plasmodium malariae and P . ovale genomes provide insights into malaria parasite evolution. Nature. **2017**; 542(7639):101–104.

14. Mathema VB, Nakeesathit S, White NJ, Dondorp AM, Imwong M. Genome-wide microsatellite characteristics of five human Plasmodium species, focusing on Plasmodium malariae and P. ovale curtisi. Parasite. **2020**; 27:34.

15. Imwong M, Nair S, Pukrittayakamee S, et al. Contrasting genetic structure in Plasmodium vivax populations from Asia and South America. Int J Parasitol. **2007**; 37(8–9):1013–1022.

16. Greenhouse B, Myrick A, Dokomajilar C, et al. VALIDATION OF MICROSATELLITE MARKERS FOR USE IN GENOTYPING POLYCLONAL PLASMODIUM FALCIPARUM INFECTIONS. The American journal of tropical medicine and hygiene. NIH Public Access; **2006**; 75(5):836.

17. Mathema VB, Nakeesathit S, Pagornrat W, et al. Polymorphic markers for identification of parasite population in Plasmodium malariae. Malaria Journal. **2020**; 19(1):48.

18. Divis PCS, Singh B, Anderios F, et al. Admixture in Humans of Two Divergent Plasmodium knowlesi Populations Associated with Different Macaque Host Species. PLOS Pathogens. Public Library of Science; **2015**; 11(5):e1004888.

19. Joste V, Kamaliddin C, Kendjo E, Hubert V, Argy N, Houzé S. Distinction of Plasmodium ovale wallikeri and Plasmodium ovale curtisi using quantitative Polymerase Chain Reaction with High Resolution Melting revelation. Scientific Reports. **2018**; 8(1):1–8.

20. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **1999**; 27(2):573–580.

21. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques. **1996**; 20(6):1004–1006, 1008–1010.

22. Korbie DJ, Mattick JS. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat Protoc. Nature Publishing Group; **2008**; 3(9):1452–1456.

23. Sutton PL, Luo Z, Divis PCS, et al. Characterizing the genetic diversity of the monkey malaria parasite Plasmodium cynomolgi. Infect Genet Evol. **2016**; 40:243–252.

24. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. **2008**; 24(11):1403–1405.

25. Thioulouse J, Chessel D, Dole´dec S, Olivier J-M. ADE-4: a multivariate analysis and graphical display software. Statistics and Computing. **1997**; 7(1):75–83.

26. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. PeerJ Inc.; **2014**; 2:e281.

27. Bruvo R, Michiels NK, D'souza TG, Schulenburg H. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Molecular Ecology. **2004**; 13(7):2101–2106.

28. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research. **2021**; 49(W1):W293–W296.

29. Moser KA, Drábek EF, Dwivedi A, et al. Strains used in whole organism Plasmodium falciparum vaccine trials differ in genome structure, sequence, and immunogenic potential. Genome Medicine. **2020**; 12(1):6.

30. Wickham, H. and Henry, L. (2018). tidyr: Easily Tidy Data with “spread()” and “gather()” Functions. R package version 0.8.1. https://CRAN.R-project.org/package=tidyr.

31. Garnham PCC, Bray RS, Cooper W, Lainson R, Awad FI, Williamson J. The pre-erythrocytic stage of Plasmodium ovale. Trans R Soc Trop Med Hyg. Oxford Academic; **1955**; 49(2):158–167.

32. Austin SC, Stolley PD, Lasky T. The History of Malariotherapy for Neurosyphilis: Modern Parallels. JAMA. American Medical Association; **1992**; 268(4):516–519.

33. Collins WE, Jeffery GM. A retrospective examination of sporozoite-induced and trophozoite-induced infections with Plasmodium ovale: development of parasitologic and clinical immunity during primary infection. Am J Trop Med Hyg. **2002**; 66(5):492–502.

34. Richter J, Franken G, Mehlhorn H, Labisch A, Häussinger D. What is the evidence for the existence of Plasmodium ovale hypnozoites? Parasitol Res. **2010**; 107(6):1285–1290.

35. Soulard V, Bosson-Vanga H, Lorthiois A, et al. Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice. Nature Communications. **2015**; 6:7690.

36. Methods for surveillance of antimalarial drug efficacy [Internet]. [cited 2022 Dec 21]. Available from: https://www.who.int/publications-detail-redirect/9789241597531

37. Taylor AR, Watson JA, Chu CS, et al. Resolving the cause of recurrent Plasmodium vivax malaria probabilistically. Nat Commun. Nature Publishing Group; **2019**; 10(1):5595.

38. Popovici J, Pierce-Friedrich L, Kim S, et al. Recrudescence, Reinfection, or Relapse? A More Rigorous Framework to Assess Chloroquine Efficacy for Plasmodium vivax Malaria. J Infect Dis. **2019**; 219(2):315–322.

39. Visser BJ, Wieten RW, Kroon D, et al. Efficacy and safety of artemisinin combination therapy (ACT) for non-falciparum malaria: a systematic review. Malaria Journal. **2014**; 13(1):463.

40. Mombo-Ngoma G, Kleine C, Basra A, et al. Prospective evaluation of artemether-lumefantrine for the treatment of non-falciparum and mixed-species malaria in Gabon. Malar J. **2012**; 11:120.

41. Lwin KM, Phyo AP, Tarning J, et al. Randomized, double-blind, placebo-controlled trial of monthly versus bimonthly dihydroartemisinin-piperaquine chemoprevention in adults at high risk of malaria. Antimicrob Agents Chemother. **2012**; 56(3):1571–1577.

42. Siswantoro H, Russell B, Ratcliff A, et al. In Vivo and In Vitro Efficacy of Chloroquine against Plasmodium malariae and P. ovale in Papua, Indonesia. Antimicrobial Agents and Chemotherapy. American Society for Microbiology; **2011**; 55(1):197–202.

43. Groger M, Fischer HS, Veletzky L, Lalremruata A, Ramharter M. A systematic review of the clinical presentation, treatment and relapse characteristics of human Plasmodium ovale malaria. Malar J. **2017**; 16(1):112.

44. Groger M, Veletzky L, Lalremruata A, et al. Prospective Clinical Trial Assessing Species-Specific Efficacy of Artemether-Lumefantrine for the Treatment of Plasmodium malariae, Plasmodium ovale, and Mixed Plasmodium Malaria in Gabon. Antimicrob Agents Chemother. **2018**; 62(3):e01758-17.

45. Imwong M, Snounou G, Pukrittayakamee S, et al. Relapses of Plasmodium vivax infection usually result from activation of heterologous hypnozoites. J Infect Dis. **2007**; 195(7):927–933.

46. Imwong M, Boel ME, Pagornrat W, et al. The first Plasmodium vivax relapses of life are usually genetically homologous. J Infect Dis. **2012**; 205(4):680–683.

47. Popovici J, Friedrich LR, Kim S, et al. Genomic Analyses Reveal the Common Occurrence and Complexity of Plasmodium vivax Relapses in Cambodia. mBio. **2018**; 9(1):e01888-17.

48. Bright AT, Manary MJ, Tewhey R, et al. A High Resolution Case Study of a Patient with Recurrent Plasmodium vivax Infections Shows That Relapses Were Caused by Meiotic Siblings. PLoS Negl Trop Dis. **2014**; 8(6):e2882.

49. Lopez L, Koepfli C. Systematic review of Plasmodium falciparum and Plasmodium vivax polyclonal infections: Impact of prevalence, study population characteristics, and laboratory procedures. PLOS ONE. Public Library of Science; **2021**; 16(6):e0249382.

Tables

Table 1 – List of the microsatellite markers for *P. ovale wallikeri* genotyping.

Name	Chr	Primer sequence (5'-3')[$]	Position : end-start*	Repeat sequence*
2.1	2	F: 6FAM-TTTTAAAGTTGCCGCATTGTC-3' R: gtttcttATACCAAAAGCGCTGTCTCC	103,121-103,270 intergenic	$(GTT)_{12}(GCT)_9$
5.2	5	F: HEX-ACACCCGAAGTGGTATCAGC R: ACGGATGCGCTATAATTTGG	664,493-664,733 POWCR01_050018800	$(TAC)_{10}$
8.3	8	F: HEX-CCCCTTGCAAACGTGTAAAT R: GTGTTTGTGCACATGGAAGG	855,305-855,486 intergenic	$(CTAC)_{16}$
10.1	10	F: HEX-GATTCATCGGTGCTTTCTCC R: GGAAGAAACGGAAGAAGTGG	432,312-432,527 POWCR01_100015100	$(CGT)_{12}$
11.1	11	F: HEX-AAGGGGTGGGGGAGTTTAC	539,226-539,442	$(GTT)_{16}$

		R: gtttcttCAGGCACGTACTAGCATGGA	intergenic	
12.1	12	F: FAM-ACATGCAACGCAACTTTCAG R: GTGCGCCTCCACTTTTTACT	1,540,305-1,540,570 intergenic	$(CATA)_{14}$
13.3	13	F: FAM-CGAATGAAGTAGCACAGTTAAGC R: GCTCAATATCGCTCTTTTCCA	268,664-268,882 intergenic	$(TTC)_{10}$
14.2	14	F: FAM-GGCATTTTTCCATTTCGTCA R: AAAATTGATAAGGCGCGAGA	1,141,610-1,141,858 intergenic	$(ATAC)_{11}(AT)_5$

[$]F: forward, R: reverse, gtttctt: pig-tailed primer, Chr: chromosome.

*on the PowCR01 reference genome (GenBank assembly accession: GCA_900090025.2)

Table 2 – Characteristics of the microsatellite markers developed for *P. ovale wallikeri* (n=69 isolates).

Microsatellite	PCR length range (bp)	PCR success rate	Allelic diversity, A	Expected heterozygosity (He)	MOI[$] mean (max*)
2.1	132-159	99%	10	0.764	1.04 (2)
5.2	233-263	100%	9	0.849	1.01 (2)
8.3	127-608	99%	15	0.902	1.04 (2)
10.1	203-239	97%	10	0.786	1.01 (2)
11.1	185-279	100%	17	0.932	1.03 (2)
12.1	227-356	94%	13	0.906	1.06 (2)
13.3	183-267	94%	14	0.834	1 (1)
14.2	227-344	100%	14	0.940	1.01 (2)

[$]MOI: multiplicity of infection.

*max: maximum of peaks detected in one sample

Figures

Figure 1 – Geographic origins of the *P. ovale wallikeri* isolates included in the study. Countries with included isolates are colored in blue. Number of isolates per country is indicated in salmon circle. * symbols indicate the countries of origin of the relapsing isolates.

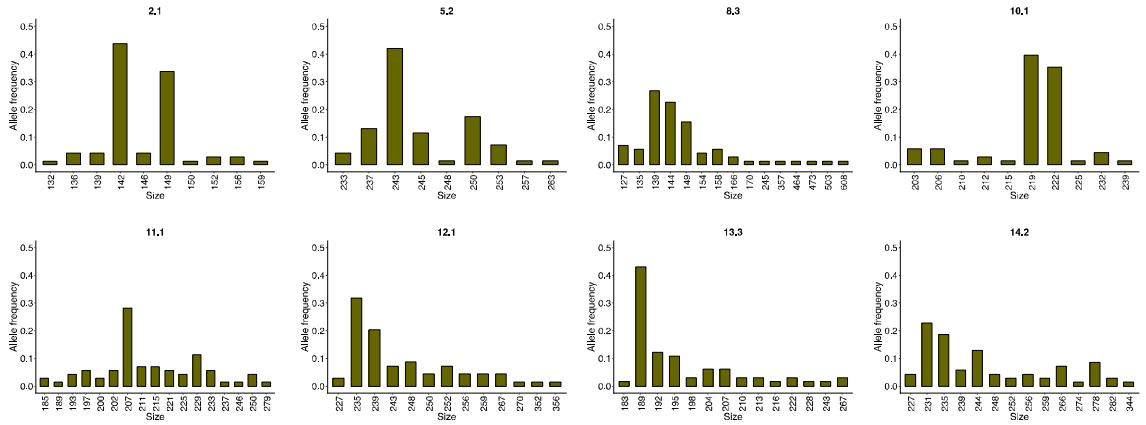


Figure 2 - Number of alleles and their frequencies for the eight microsatellite markers.

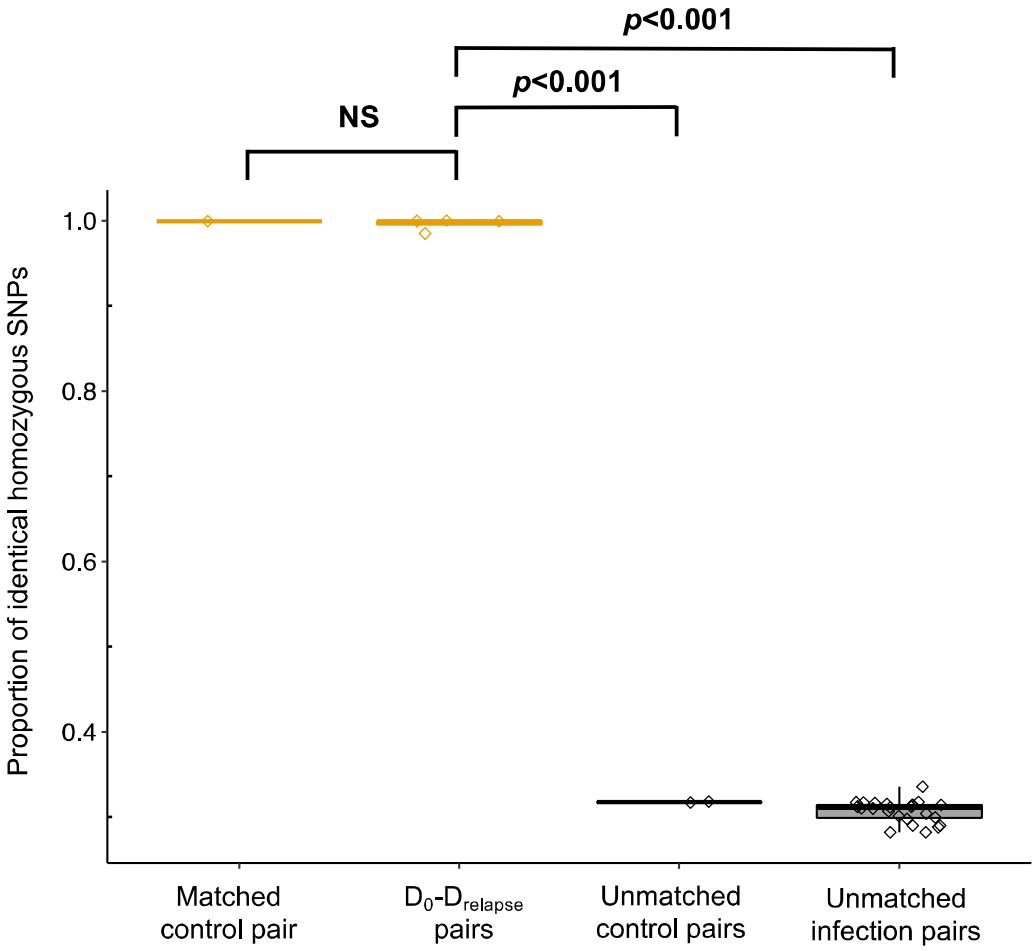
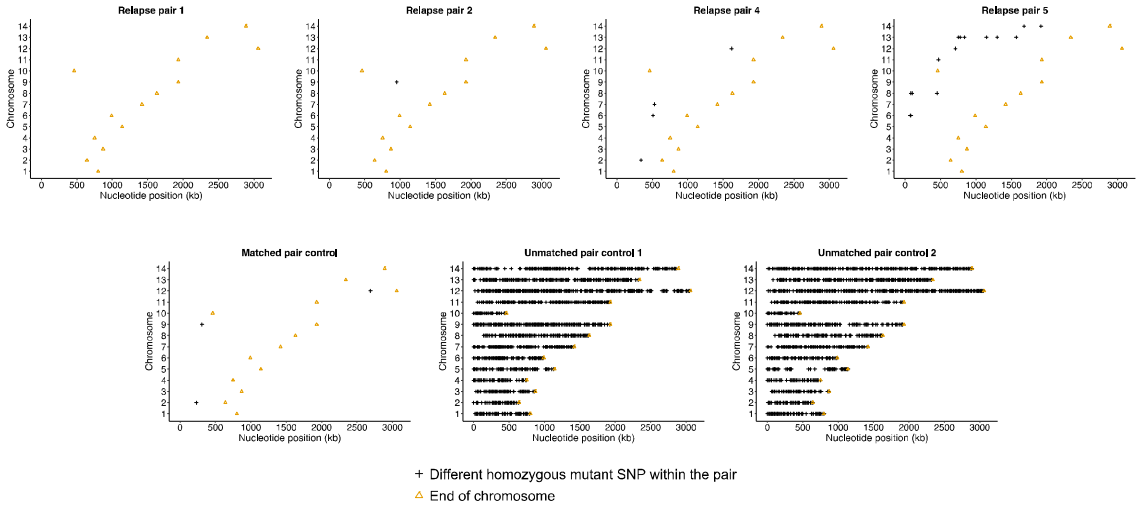


Figure 3 - Proportion of identical genome-wide SNPs within infection pairs.

Matched control pair and D_0-$D_{relapse}$ pairs are in light brown. Unmatched control pairs and unmatched infection pairs are in gray. Each dot represents one pair. NS stands for non-significant ($\chi 2$ tests).

Figure



4 – Genetic relatedness within matched or unmatched infection pairs.

The distribution of the different alleles within each D0-Drelapse pair is shown throughout the fourteen chromosomes of the nuclear genome. Each cross represents a 10 kb window containing at least one differing SNP, represented according to its genomic position (x axis, nucleotide position in kilobase pair; y axis, chromosome). Beige triangle shows the end of each chromosome.