Genetic dissociation of three antigenic genes in Plasmodium ovale curtisi and Plasmodium ovale wallikeri

Naowarat Saralamba, François Nosten, Colin Sutherland, Ana Paula Arez, Georges Snounou, Nicholas White, Nicholas Day, Arjen Dondorp, Mallika Imwong

To cite this version:
Naowarat Saralamba, François Nosten, Colin Sutherland, Ana Paula Arez, Georges Snounou, et al.. Genetic dissociation of three antigenic genes in Plasmodium ovale curtisi and Plasmodium ovale wallikeri. PLoS ONE, Public Library of Science, 2019, 14 (6), pp.e0217795. 10.1371/journal.pone.0217795. hal-02347227

HAL Id: hal-02347227
https://hal.archives-ouvertes.fr/hal-02347227
Submitted on 5 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Genetic dissociation of three antigenic genes in *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri*

Naowarat Saralamba\(^{1,2,*}\), François Nosten\(^3\), Colin J. Sutherland\(^4\), Ana Paula Arez\(^5\), Georges Snounou\(^6\), Nicholas J. White\(^2,7\), Nicholas P. J. Day\(^2,7\), Arjen M. Dondorp\(^2,7\), Mallika Imwong\(^1,2\)

\(^1\) Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand, \(^2\) Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand, \(^3\) Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Bangkok, Thailand, \(^4\) Immunology Unit, Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom, \(^5\) Global Health and Tropical Medicine, GHTM, Instituto de Higiene e Medicina Tropical, IHMT, Universidade Nova de Lisboa, UNL, Lisbon, Portugal, \(^6\) CEA-Université Paris Sud 11-INSERM U1184, Immunology of Viral Infections and Autoimmune Diseases (IMVA), IDMIT Department, IBFJ, DRF, Fontenay-aux-Roses, Paris, France, \(^7\) Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

* naowarat.sar@mahidol.ac.th

Abstract

*Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* are two sympatric human malaria species prevalent in Africa, Asia and Oceania. The reported prevalence of both *P. ovale* spp. was relatively low compared to other malaria species, but more sensitive molecular detection techniques have shown that asymptomatic low-density infections are more common than previously thought. Whole genome sequencing of both *P. ovale* spp. revealed genetic dissociation between *P. ovale curtisi* and *P. ovale wallikeri* suggesting a species barrier. In this study we further evaluate such a barrier by assessing polymorphisms in the genes of three vaccine candidate surface protein: circumsporozoite protein/thrombospondin-related anonymous-related protein (*ctrp*), circumsporozoite surface protein (*csp*) and merozoite surface protein 1 (*msp1*). The complete coding sequence of *ctrp* and *csp*, and a partial fragment of *msp1* were isolated from 25 *P. ovale* isolates and compared to previously reported reference sequences. A low level of nucleotide diversity (Pi = 0.02–0.10) was observed in all three genes. Various sizes of tandem repeats were observed in all *ctrp*, *csp* and *msp1* genes. Both tandem repeat unit and nucleotide polymorphism in all three genes exhibited clear dimorphism between *P. ovale curtisi* and *P. ovale wallikeri*, supporting evidence of non-recombination between these two species.

Introduction

*Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri* are two sympatric species of malaria parasites found across many malaria endemic countries in Africa, Asia and Oceania [1–3].
Although morphological features of *P. ovale curtisi* and *P. ovale wallikeri* are indistinguishable, these two *P. ovale* species are genetically distinct, and there is evidence of differences in latency and clinical presentation [4–6]. Nuclear genome sequences of *P. ovale curtisi* and *P. ovale wallikeri* were recently reported and revealed different expansion in some gene families [7]. Currently the target genes used for discriminating between *P. ovale curtisi* and *P. ovale wallikeri* are the SSU rRNA gene [8], tryptophan rich antigen (*potra*) [9], reticulocyte-binding protein 2 (*pORbp2*) [9], and some sexual stage proteins [9]. Sequence polymorphisms in the cell-surface associated proteins that are candidate targets for vaccine development have only been studied rarely. The current study assessed genetic diversity in a highly polymorphic region of the blood stage merozoite surface protein gene *msp1*, and in two genes encoding sexual stage and sporozoite proteins, *ctrp* and *csp* respectively, in *P. ovale curtisi* and *P. ovale wallikeri*.

CTRP is a member of the micronemal and cell-surface associated proteins. In *P. falciparum* disruption of the *ctrp* gene prevents oocyst development in the anopheline mosquito [10], indicating that CTRP is important for mosquito midgut development. For this reason CTRP has been proposed as a transmissions-blocking vaccine candidate. CSP is the major surface protein on the *Plasmodium* sporozoite. It is a candidate target for pre-erythrocytic stage vaccine development. Genetic polymorphism within the *csp* gene have been investigated in most human malaria species including *P. falciparum* [11, 12], *P. vivax* [13, 14], *P. malariae* [15, 16], and *P. knowlesi* [17], but not in *P. ovale*. MSP1 is one of the predominant antigen expressed in the erythrocytic stage of *Plasmodium* spp. The *msp1* gene is highly polymorphic and has been well characterized in *P. falciparum* [18, 19] and *P. vivax* [13, 14]. A study of *P. ovale* isolates from Thailand revealed low diversity in the *msp1* gene [20].

The current study evaluates sequence diversity of *ctrp*, *csp* and *msp1*, in a wider collection of *P. ovale* isolates collected from Thailand and African countries. Assessing diversity in these surface proteins is important for defining vaccine candidates, and to further assess the species barrier between *P. ovale curtisi* and *P. ovale wallikeri*. In the current era of malaria elimination, the better understanding of *P. ovale curtisi* and *P. ovale wallikeri* is essential to ensure success against all human malaria species.

### Materials and methods

#### Samples

Twenty-five samples of *P. ovale* (14 *P. ovale wallikeri* and 11 *P. ovale curtisi*) were collected from Thailand and African countries during 1995–2010 (S1 Table). All samples were obtained from patients enrolled in previous studies who gave written informed consent to blood sampling. Parasitaemia of these samples varied from 1 per 500 WBC to 198 per 500 WBC. The protocol for this study was reviewed and approved (reference number MUTM2001-049-04) by the ethics committee of the Faculty of Tropical Medicine, Mahidol University, Thailand.

Genomic DNA of all samples was confirmed for the present of *P. ovale*. Nested PCR of the SSU rRNA gene was performed with primer rPLU1/rPLU5 in the primary reaction and with primer rOVA1/rPLU2 in the secondary reaction [21]. A nested PCR protocol based on the linker region of dhfr-ts gene was applied with primer Pla-DHFR-F/Pla-TS-R in the primary reaction and with primer PO-Lin-F/PO-Lin-R in the secondary reaction [22]. In addition, a semi-nested PCR of *potra* gene was performed with a primer specific to both *P. ovale* spp. (*PoTRA*-F/*PoTRA* rev3) in the primary reaction and with specific *P. ovale curtisi* (*PoTRA*-F/*PocTRA*-R) and *P. ovale wallikeri* (*PoTRA*-F/*PowTRA*-R) primers in the secondary reaction [23].
Isolation of *poctrp* and *pocsp* gene

Specific primers targeting *poctrp* and *pocsp* genes were designed to obtain the full length of those two gene sequences (Table 1). A semi-nested PCR approach was used for amplification of each fragment with PCR conditions as presented in Table 1. All PCR reactions were performed with 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2 mM MgCl₂, 125 μM dNTPs, 250 nM of

Table 1. Primer sequences and PCR conditions for isolation of *poctrp*, *pocsp* and *pomsp1* genes.

| Target gene | Primer name | Sequences (5' to 3') | Annealing temperature (°C) | No. of PCR cycle | Product size (bp) |
|-------------|-------------|----------------------|-----------------------------|------------------|-------------------|
| *Pocsp*     | OCSP_F120   | CGTAGGAGCTGGGAATCAAG | 56                          | 30               | 1,500             |
|             | OCSP_R15k   | TTTCCCCCGATTCATATCA  |                             |                  |                   |
|             | OCSP_F120   | CGTAGGAGCTGGGAATCAAG | 58                          | 35               | 980               |
|             | OCSP_R10k   | ACTGCATGGTGGAAACGG   |                             |                  |                   |
|             | OCSP_F120   | CGTAGGAGCTGGGAATCAAG | 56                          | 30               | 1,500             |
|             | OCSP_R15k   | TTTCCCCCGATTCATATCA  |                             |                  |                   |
|             | OCSP_F700   | AGCTCCATGAAATGGGTTTG | 58                          | 35               | 800               |
|             | OCSP_R15k   | TTTCCCCCGATTCATATCA  |                             |                  |                   |
|             | OCSP_F700   | AGCTCCATGAAATGGGTTTG | 56                          | 30               | 1,300             |
|             | OCSP_R20k   | CCCACATGCAGTGAAATACC |                             |                  |                   |
|             | OCSP_F12k   | AGCGGAGCGCAATGGTAG   | 58                          | 35               | 750               |
|             | OCSP_R20k   | CCCACATGCAGTGAAATACC |                             |                  |                   |
|             | OCSP_F16k   | GGGAAAAATCAGATGCTCT  | 56                          | 30               | 1,400             |
|             | OCSP_R30k   | ATAACCGAGCAACACCAACC |                             |                  |                   |
|             | OCSP_F18k   | TTTCACCCCGACACCAACCG | 58                          | 35               | 1,160             |
|             | OCSP_R30k   | ATAACCGAGCAACACCAACC |                             |                  |                   |
|             | OCSP_F5start| AAATGCAGAGCAAAGACAAA | 57                          | 30               | 1,500             |
|             | OCSP_R15k   | TTTCCCCCGATTCATATCA  |                             |                  |                   |
|             | OCSP_F5start| AAATGCAGAGCAAAGACAAA | 59                          | 35               | 1,000             |
|             | OCSP_R10k   | ACTGCATGGTGGAAACGG   |                             |                  |                   |
|             | OCSP_F2700  | TTTCATGAAAGGCAATGAGA | 57                          | 30               | 1,600             |
|             | OCSP_R4300  | CAAACTGATATTGTTTCTTTTCAA |                         |                  |                   |
|             | OCSP_F2700  | TTTCATGAAAGGCAATGAGA | 59                          | 35               | 1,500             |
|             | OCSP_R4200  | TTTCATGAAAGGCAATGAGA |                             |                  |                   |
|             | OCSP_F4000  | TGGCCAACAGTATTCACATGT | 57                          | 30               | 1,400             |
|             | OCSP_R3500  | CACACAGGCAAGTCTGAGCA  |                             |                  |                   |
|             | OCSP_F4000  | TGGCCAACAGTATTCACATGT | 59                          | 35               | 1,400             |
|             | OCSP_R5400  | TCCACATCGAAATTTCAAGT |                             |                  |                   |
|             | OCSP_F5300  | GACGGAAGGAGGCCACCTTG | 57                          | 30               | 1,000             |
|             | OCSP_R3end  | AGACCGGAAATGGGCTATGAT |                             |                  |                   |
|             | OCSP_F5300  | GACGGAAGGAGGCCACCTTG | 59                          | 35               | 1,000             |
|             | OCSP_R3stop | GAAGAATCGAGCGGAGAAA  |                             |                  |                   |
| *Pocsp*     | PoCSP_F1    | ATGAGGAGCTGGGCAATT    | 50                          | 30               | 1,100             |
|             | PoCSP_R     | TTAATGGAAGAAATCTAGGAA |                             |                  |                   |
|             | PoCSP_F2    | GCCGTGACGCTTTTTTATT   | 52                          | 35               | 1,050             |
|             | PoCSP_R     | TTAATGGAAGAAATCTAGGAA |                             |                  |                   |
| *Pomsp1*    | OMSp1F1     | GATGAAAATCTATGCATTGAGG | 56                          | 30               | 1,000             |
|             | OMSp1R1     | CAT(C/T)ATACTTTTACTCTCCT |                         |                  |                   |
|             | OMSp1F1     | GATGAAAATCTATGCATTGAGG | 58                          | 35               | 900               |
|             | OMSp1R2     | CATCATC(A/G)TCCTCGGTTTCCC |                      |                  |                   |

https://doi.org/10.1371/journal.pone.0217795.t001
each primer and 4 unit of Taq Polymerase (Kapa biosystems, USA). PCR products were then purified by Gel/PCR purification kit (Favogen, Taiwan), before being submitted for DNA sequencing.

**Analysis of variable region in pomsp1 gene**

Twelve available pomsp1 sequences from both *P. ovale curtisi* and *P. ovale wallikeri* were retrieved from the NCBI database (accession number FJ824670, FJ824671, KC137340—KC137349) and multiple sequence alignments were performed. A highly polymorphic region within pomsp1 was observed between amino acid residues 700 to 1,000. The primers OMSP1.F1, OMSP1.R1, and OMSP1.R2, were designed for a semi-nested PCR approach to analyses this polymorphic domain in 25 *P. ovale* samples (Table 1). Positive PCR products were then purified by Gel/PCR purification kit (Favogen, Taiwan), before being submitted for DNA sequencing. All pomsp1 sequences obtained in this study were analyzed together with the previous reports.

**Sequence analysis and phylogenetic tree reconstruction**

Nucleotide polymorphisms of poctrp, pocsp and pomsp1 from *P. ovale curtisi* and *P. ovale wallikeri* were analyzed with ClustalW multiple alignment using BioEdit version 7.2.6.1 [24]. Nucleotide sequences of poctrp, pocsp and pomsp1 were translated to deduced amino acid sequences using BioEdit version 7.2.6.1 [24]. The sequences obtained from 25 samples of *P. ovale* spp. were analyzed in comparison with the previously reported sequences from the NCBI database (poctrp: accession number LT594512, LT594589, pocsp: accession number SBT72933, SBT84923, pomsp1: accession number LT594511, LT594588, KX672044, KX672045, FJ824670, FJ824671, KC137340—KC137349). Genetic variability including average pairwise nucleotide diversity (π), haplotype diversity, and sliding plot nucleotide diversity with a window length of 100 bp and 25 bp step size within poctrp, pocsp and pomsp1 from *P. ovale curtisi* and *P. ovale wallikeri* was obtained from DnaSP 6.10.4 [25]. The ratio of non-synonymous to synonymous (dN/dS) within each *P. ovale* spp. was measured by DnaSP 6.10.4 [25]. Tests for neutral evolution were assessed with Tajima’s D, Fu and Li’s D, and Fu and Li’s F tests using DnaSP 6.10.4 [25].

A neighbor-joining (NJ) phylogenetic tree was constructed from concatenated CTRP, CSP and MSP1 protein sequences to assess relationships between *P. ovale curtisi* and *P. ovale wallikeri*. A bootstrap test (1,000 replicates) was applied under the Jones-Taylor-Thornton (JTT) model of evolution using MEGA7 [26].

**Results**

**Isolation and analysis of poctrp**

The complete coding sequence of poctrp gene was obtained from 11 *P. ovale curtisi* and 14 *P. ovale wallikeri* isolates (accession number MK403987-MK404009). It revealed that the poctrp genes for both species has only one exon encoding for 2,007 to 2,047 amino acids. Sequence alignment of these 25 poctrp sequences, together with another two poctrp sequences (accession number LT594512 and LT594589) available in the NCBI database, and other ctrp sequences from the other *Plasmodium* spp. that infect humans, showed that poctrp is composed of a signal peptide, six vWA domains, seven TSP1 domains, transmembrane domain, and a cytoplasmic region (Fig 1). Alignment of the CTRP of all human *Plasmodium* spp. revealed highly conserved transmembrane (TM) and cytoplasmic regions (Fig 1). A conserved amino acid sequence YGYN/K for the tyrosine-based TM motif involved in cellular trafficking, and the
cytoplasmic domain tryptophan residue (Fig 1) which is the key interaction to drive parasite motility were conserved between all studied human Plasmodium spp. Multiple alignment of the full-length CTRP among all human-infecting Plasmodium spp. also identified a highly conserved region close to the C-terminus.

All poctrp sequences including the two reference sequences were translated to deduce their corresponding amino acids and analyzed for intra- and inter-specific sequence diversity at this locus. The deduced amino acid alignment of PoCTRP showed two prominent regions. The first region is located around 300–320 amino acids between the vWA1 and vWA2 domains. P. ovale curtisi isolates carries two amino acids repeats “PE” with 7–11 copies, while all P. ovale wallikeri isolates had 4 “PE” repeat units (Fig 2). The second region is located between codons 570 and 600, where a tandem repeat of six amino acids was identified. Three patterns of six amino acids repeats were observed: ENPDSS, EKPGSS, and ENPGSS. Different numbers of repeat units were presented in the P. ovale isolates (Fig 2). The repeat EKPGSS is the most frequent in both P. ovale curtisi and P. ovale wallikeri. This region showed a marked difference in
length, providing a potential additional genotypic marker to differentiate *P. ovale curtisi* from *P. ovale wallikeri*. Multiple sequence alignment of CTRP of all human *Plasmodium* spp. revealed species-specific regions for *P. ovale* spp. at codons 512–538 and codons 573–599. PCR amplification of *P. ovale* spp. with primers targeting those two regions are useful to distinguish *P. ovale curtisi* from *P. ovale wallikeri*.

The available *poctrp* genes were analysed in a sliding plot for nucleotide diversity between *P. ovale wallikeri* and *P. ovale curtisi* (Fig 3). *P. ovale curtisi* showed higher diversity around the first 1 kb where *P. ovale wallikeri* showed higher diversity at 4 kb—5 kb of *poctrp* (Fig 3). For this gene, the average nucleotide diversity of *P. ovale curtisi* is slightly lower than that of *P. ovale wallikeri*, and combined analysis of all 27 *P. ovale* sequences showed a higher diversity value than that calculated from each species alone (Table 2), which indicates distinct distributions of diversity across the *poctrp* locus in the two species.

**Isolation and analysis of pocsp**

The *pocsp* gene was successfully amplified from 14 *P. ovale wallikeri* and 11 *P. ovale curtisi* isolates (accession number MK404010-MK404031). The complete *pocsp* gene varied in size from 1,020 to 1,185 bp, and the size variation resulted from variable tandem repeats in the central repeat region. The *pocsp* sequences were analyzed together with two other sequences available in NCBI databases and those of the *csp* of the other human *Plasmodium* spp. The protein domain architecture of *pocsp* was determined based on homologous CSP proteins alignment with other human *Plasmodium* spp. The *pocsp* structure domain was similar to that of the *csp* from the other *Plasmodium* spp. Four domains in the conserved N-terminus domain (conserved region I) and in the conserved C-terminus domain (Th2R, conserved region II, and Th3R) were of particular interest. A summary of the amino acid patterns in each of these
domains is presented in Table 3. Overall, a higher number of haplotypes was observed in *P. ovale curtisi* as compared to *P. ovale wallikeri*. Whereas *P. ovale wallikeri* showed only one haplotype in conserved region I and two in conserved region II, three and six, respectively, were observed for *P. ovale curtisi*. A high number of haplotypes was observed in the Th2R and Th3R domains for both *P. ovale curtisi* and *P. ovale wallikeri*, but it is interesting to note that none were shared by both species. The central repeat region of *pocsp* was also analyzed. Several patterns of nine amino acid repeats were observed. A specific repeat unit was observed for *P. ovale wallikeri* (DPPAPVPQG), and for *P. ovale curtisi* NPPAPQEG, with the latter showing a higher diversity in repeat unit numbers (Fig 2).
The `pocsp` gene was evaluated for nucleotide diversity between *P. ovale curtisi* and *P. ovale wallikeri*. Sliding plots of nucleotide diversity revealed overall higher nucleotide diversity in *P. ovale curtisi* than in *P. ovale wallikeri* (Fig 3). The estimated synonymous (dS) and nonsynonymous (dN) substitution was also found at higher value in *P. ovale curtisi* than that of *P. ovale wallikeri* (Table 2). Combined analysis of both *P. ovale* spp. showed significantly positive values (*p*<0.05) for Fu and Li’s D and Fu and Li’s F tests, suggesting population bottlenecks or balancing selections in these two species (Table 2).

**Table 2. Nucleotide diversity and natural selection in *P. ovale* spp.**

|         | Species                  | No. of samples | Haplotype diversity | Pi     | dN/dS   | Tajima’s D | Fu and Li’s D | Fu and Li’s F |
|---------|--------------------------|----------------|---------------------|--------|---------|------------|---------------|---------------|
| **CTR** | *P. ovale wallikeri*     | 15             | 0.971               | 0.00473| 1.52976 | 0.17242    | 0.1958        | 0.10679       |
|         | *P. ovale curtisi*       | 12             | 1                   | 0.00416| 1.11169 | 0.42001    | 0.10998       | 0.03388       |
|         | *P. ovale*               | 27             | 0.991               | 0.01912| 0.28505 | 2.11247    | 1.53105**     | 2.02216**     |
| **CSP** | *P. ovale wallikeri*     | 15             | 0.93333             | 0.02529| 0.47695 | 1.08735    | 0.99957       | 1.1802        |
|         | *P. ovale curtisi*       | 12             | 1                   | 0.05958| 0.88465 | 0.588      | 0.82928       | 0.87353       |
|         | *P. ovale*               | 27             | 0.98006             | 0.1201 | 0.96289 | 1.45552    | 1.4117        | 1.68304*      |
| **MSP1**| *P. ovale wallikeri*     | 20             | 0.511               | 0.01309| 2.24763 | 1.28828    | 1.71084**     | 0.94791       |
|         | *P. ovale curtisi*       | 21             | 0.867               | 0.03582| 1.32754 | 0.80269    | 0.78703       | 0.92536       |
|         | *P. ovale*               | 41             | 0.817               | 0.11226| 0.89189 | 1.79705    | 1.58881**     | 1.98472**     |

* P<0.05, **P<0.02

https://doi.org/10.1371/journal.pone.0217795.t002

Genetic dissociation of *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri*

The `pocsp` gene was evaluated for nucleotide diversity between *P. ovale curtisi* and *P. ovale wallikeri*. Sliding plots of nucleotide diversity revealed overall higher nucleotide diversity in *P. ovale curtisi* than in *P. ovale wallikeri* (Fig 3). The estimated synonymous (dS) and nonsynonymous (dN) substitution was also found at higher value in *P. ovale curtisi* than that of *P. ovale wallikeri* (Table 2). Combined analysis of both *P. ovale* spp. showed significantly positive values (*p*<0.05) for Fu and Li’s D and Fu and Li’s F tests, suggesting population bottlenecks or balancing selections in these two species (Table 2).

**Genetic analysis of pomsp1**

The sequences for the variable regions within the `pomsp1` gene covering amino acids 710 to 1,020 were obtained from 14 *P. ovale wallikeri* and 11 *P. ovale curtisi* isolates (accession number MK404032-MK404049). Apart from this, sixteen sequences of `pomsp1` gene (accession number LT594511, LT594588, KX672044, KX672045, FJ824670, FJ824671, KC137340—KC137349) were available in the NCBI database. Taken together, 41 PoMSP1 sequences were used in the alignment. A clear dimorphic pattern was observed between *P. ovale curtisi* and *P. ovale wallikeri*. Amino acid tandem repeat patterns were found in *P. ovale* spp. The tandem repeats are characteristic for the two different *P. ovale* spp. There were three arrangement patterns of three 5-amino acid repeat units (PGAGG, PGAAG, and PGVPG) found exclusively in *P. ovale wallikeri* isolates. Whereas, nine arrangement patterns of six 4-amino acids repeat units (QAAT, QTAT, HAST, QATT, QVTT, QSAT) were observed specifically in the *P. ovale curtisi* isolates (S2 Table).

Analysis of gene diversity and haplotype diversity at the `pomsp1` locus showed that *P. ovale curtisi* has higher diversity than that of *P. ovale wallikeri* (Table 2). Sliding window plots showed higher overall nucleotide diversity in *P. ovale curtisi* than in *P. ovale wallikeri* (Fig 3). The ratio of synonymous (dS) and nonsynonymous (dN) substitutions was higher in *P. ovale wallikeri* than *P. ovale curtisi* (Table 2).

**Comparative analysis of *P. ovale curtisi* and *P. ovale wallikeri***

Genetic analysis of *P. ovale curtisi* and *P. ovale wallikeri* based on three surface protein genes revealed clear dissociation between these two species. Analysis within each species was performed though sequence diversity and amino acid patterns. The sequence polymorphism in
### Table 3. Sequence polymorphism in the conserved regions of *P. ovale* CSP.

| Haplotype | Amino acid | No. of *P. ovale curtisi* | No. of *P. ovale wallikeri* |
|-----------|------------|---------------------------|----------------------------|
| **Conserved region I** | | | |
| 1 | PVENKLKQG | 6 | 15 |
| 2 | PVENKLNQG | 1 | 0 |
| 3 | PVENNLNQG | 5 | 0 |
| **Th2R** | | | |
| 1 | PPSEDDIKKYIDKIRKD | 0 | 7 |
| 2 | PPSEDDIKKYIDKIRND | 2 | 0 |
| 3 | PPSEDDIKKYIDKIRRD | 0 | 1 |
| 4 | PPSEDDIKYLKDIDRND | 0 | 1 |
| 5 | PPSEDDIKYLDRIDRND | 0 | 1 |
| 6 | PPSEDDIKFIDKIRND | 1 | 0 |
| 7 | PPSEDDIKRYLDIDRND | 1 | 0 |
| 8 | PPSEDDIKSFIDKIRND | 3 | 0 |
| 9 | PPSEDDIKYIDKIRRD | 0 | 1 |
| 10 | PPSEDDIKROYLDIDRND | 1 | 0 |
| 11 | PPSEDDIRSFIDKIRND | 1 | 0 |
| 12 | PPSEDDIKKFLDIDRND | 0 | 2 |
| 13 | PPSEDDIKSFIDKIRND | 1 | 0 |
| 14 | PPSEDDKSFMDKIRND | 1 | 0 |
| 15 | PPSEDDIRYIDKIRKD | 0 | 1 |
| 16 | PPSEDDIRNFIDKIRND | 1 | 0 |
| 17 | PPSEDDLKKFLDIDRND | 0 | 1 |
| **Conserved region II** | | | |
| 1 | ITENWSPCRVTGC | 0 | 5 |
| 2 | ITENWSPCSVSCG | 1 | 0 |
| 3 | ITENWSPCSVSCV | 2 | 0 |
| 4 | ITENWSPCVTGC | 4 | 10 |
| 5 | ITENWSPCVTVC | 1 | 0 |
| 6 | LTENWSPCSVSCG | 2 | 0 |
| 7 | LTENWSPCVTGC | 2 | 0 |
| **Th3 R** | | | |
| 1 | KKAGANAKAQQKFTLSLE | 1 | 0 |
| 2 | KKAGANAKAQQKLTLSDFE | 1 | 0 |
| 3 | KKAGANAKAQQKFTLSDFE | 1 | 0 |
| 4 | KKAGANAKANELPINDVE | 0 | 3 |
| 5 | KKAGANAKANELTINDVE | 0 | 5 |
| 6 | KKAGASAKAKPKFTLSDLE | 1 | 0 |
| 7 | KKAGASAKAQELTSLSDLE | 3 | 0 |
| 8 | KKAGASAKAQKFTLSDE | 1 | 0 |
| 9 | KKAGASAKGPKLTLSDLE | 1 | 0 |
| 10 | KKAGASAKGQFKFTLSDE | 1 | 0 |
| 11 | RKAGASAKANELPINDVE | 0 | 1 |
| 12 | RKAGASAKANELTINDVE | 0 | 6 |
| 13 | RKAGASAKAQELTSLSDLE | 2 | 0 |

https://doi.org/10.1371/journal.pone.0217795.t003
poctrp, pocsp and pomsp1 showed more divergence in P. ovale curtisi than in P. ovale wallikeri (Fig 3, Table 2). The test for neutrality (Tajima’s D, Fu and Li’s D, and Fu and Li’s F tests) was applied to poctrp, pocsp and pomsp1 to compare observed polymorphism frequencies with expected frequencies. Significantly positive values were obtained from Fu and Li’s D and Fu and Li’s F when P. ovale spp. were analyzed as one group (Table 2). These statistics reflect higher than expected frequencies of alleles, which might have resulted from population bottlenecks or balancing selections. P. ovale curtisi had a higher number of different haplotypes in all conserved domains of the CSP (Table 3). This suggests that P. ovale curtisi is intrinsically more genetically diverse than P. ovale wallikeri, but may also represent limitations of our sample.

Some of the studied P. ovale curtisi and P. ovale wallikeri infections were mixed with other human malaria spp., which might have impacted the characteristics of CTRP, CSP and MSP1. Therefore the genetic analysis of CTRP, CSP and MSP1 was compared between single and mixed infections. There were four mixed infection found in P. ovale wallikeri, in which all four samples were collected from Thailand. For P. ovale curtisi, most samples were collected from Africa with four single infections and five mixed infections. There was no significant difference in average nucleotide diversity (Pi), haplotype diversity, and dN/dS between single and mixed infections. Statistical testing for neutrality (Tajima’s D, Fu and Li’s D, and Fu and Li’s F tests) was also not significantly different (S3 Table). In addition, the pattern of tandem repeats in CTRP, CSP and MSP1 in P. ovale spp. showed no difference either between single and mixed infections or between Asia and Africa isolates.

To infer genetic relationships of P. ovale curtisi and P. ovale wallikeri, a phylogenetic tree was reconstructed based on the three cell-surface associated proteins, CTRP, CSP and MSP1. The Neighbour-Joining method [27] was used to infer the evolutionary history of each species. Based on CTRP, CSP and MSP1 (Fig 4), P. ovale curtisi and P. ovale wallikeri clustered according to species, and the tree topologies inferred from each gene showed similar features of grouping into P. ovale curtisi and P. ovale wallikeri. This suggested that the genes of P. ovale curtisi and P. ovale wallikeri do not recombine and show distinct characteristics (Fig 4).

Discussion

In addition to the earlier described polymorphisms in pomsp1, the current study provides the genetic characterization for two more cell-surface associated proteins: ctp and csp. Analysis of the complete ctp gene from all human Plasmodium spp., including the P. ovale species presented here, revealed a strongly conserved region in the CTRP protein, likely related to its importance for parasite survival. The highly conserved transmembrane and cytoplasmic regions are likely associated with cellular trafficking and parasite development. The C-terminal part containing residues and domains crucial for CTRP function for all human Plasmodium spp. were also conserved. This protein could therefore be a candidate target for vaccine development. Analysis of the csp gene in P. ovale curtisi and P. ovale wallikeri revealed a similar gene structure compared to that of the other human malaria species. The amino acid haplotypes observed in the conserved region of the csp gene were nearly all specific to either one or other species, with only 1/3 and 1/7 shared for conserved region I and conserved region II, respectively. Interestingly, no overlap was observed for the 17 Th2R and 13 Th3R haplotypes detected. This could imply a species-specific immune interactions with these T helper epitopes, or indicate a distinct biologically functional constraint. As the two species harbor distinct pocsp repeat regions, NPPAPQGEG and DPPAPVPQG, respectively, these peptides may provide useful species-specific targets for the development of antibody reagents for serological distinction of sporozoites from the two ovale species.
The study also provided additional information on the cell-surface associated protein, MSP1. Analysis of the variable region within *pomsp1* of 25 *P. ovale* samples in this study supplemented with 16 *pomsp1* from previous reports showed a clear distinction between *P. ovale curtisi* and *P. ovale wallikeri*. Alignment of the MSP1 from all human *Plasmodium* spp. showed the interspecies conserved blocks corresponding to previous characterizations [28]. Sequence polymorphisms of CTRP, CSP and MSP1 from each *P. ovale* spp. can be used for determination of parasite evolutionary relationships. Phylogenetic tree reconstruction based on concatenated CTRP, CSP and MSP1 clearly showed that *P. ovale curtisi* and *P. ovale wallikeri* are cluster separately, consistent with previous reports [29, 30].

Analogies in the reported surface proteins in *P. ovale* with other human *Plasmodium* species could help selecting potential vaccine candidates. For instance, CTRP affects oocyst development of *P. falciparum* in *Anopheles* mosquitoes [10], and conserved regions within CTRP across human *Plasmodium* spp. could provide candidate targets for transmission-blocking vaccine. In MSP1, domain architectures are similar between all human *Plasmodium* spp., and our study of PoMSP1 revealed an interspecies conserved domains 6 (residues 812–911)
between the *Plasmodium* spp., which could be candidates for a trans-species malaria vaccine. Our data could also provide the basis for development of new serological reagents for distinguishing the two species, and for identifying individuals with a history of exposure to *P. ovale* spp. carrying species-specific serum antibodies. In addition to the genes evaluated in this study, other important polymorphic genes have been used for discrimination between the two *P. ovale* spp., including the *surfin* variant gene family and the *Plasmodium* interspersed repeat (pir) superfamily, which showed expansion in both *P. ovale* spp. [7]. Additional genes encoding potential targets for vaccine development warrant further study, including genes encoding reticulocyte binding proteins and tryptophan-rich domains [31].

In summary, this study showed conserved domains in the *poctrp* and *pocsp* genes which code for potential targets for future vaccines. Quantifying polymorphism in nucleotide sequences and the tandem repeat diversity between *P. ovale curtisi* and *P. ovale wallikeri* showed absence of recombination, supporting their designation as distinct species. Within the three analysed genes, diversity was higher in *P. ovale curtisi* than in *P. ovale wallikeri*. However, this will need to be confirmed in a larger sample size with better comparison between the geographical areas where the strains were collected. In the current sample most *P. ovale curtisi* was collected from highly endemic African countries whereas most *P. ovale wallikeri* were collected in Thailand which has low endemicity.

### Supporting information

**S1 Table.** List of samples used in the study.

(XLSX)

**S2 Table.** Amino acid pattern of partial MSP1 in *P. ovale* spp.

(XLSX)

**S3 Table.** Comparative analysis of single and mixed infections *P. ovale* spp.

(XLSX)

### Acknowledgments

We would like to thank all the patients and the other support staff for the samples from Shoklo Malaria Research Unit, Tak, Thailand. This research project is supported by Mahidol University, and was part of the Wellcome Trust Mahidol University-Oxford Tropical Medicine Research Programme supported by the Wellcome Trust of Great Britain.

### Author Contributions

**Conceptualization:** Naowarat Saralamba, Francois Nosten, Colin J. Sutherland, Georges Snounou, Arjen M. Dondorp.

**Investigation:** Naowarat Saralamba.

**Resources:** Francois Nosten, Colin J. Sutherland, Ana Paula Arez, Georges Snounou.

**Supervision:** Nicholas J. White, Nicholas P. J. Day, Arjen M. Dondorp, Mallika Imwong.

**Writing – original draft:** Naowarat Saralamba.

**Writing – review & editing:** Francois Nosten, Colin J. Sutherland, Ana Paula Arez, Georges Snounou, Nicholas J. White, Nicholas P. J. Day, Arjen M. Dondorp, Mallika Imwong.
References

1. Oguike MC, Betson M, Burke M, Nolder D, Stothard JR, Kleinschmidt I, et al. Plasmodium ovale curtisi and Plasmodium ovale wallikeri circulate simultaneously in African communities. International Journal for Parasitology. 2011; 41(6):677–83. https://doi.org/10.1016/j.ijpara.2011.01.004 PMID: 21315074

2. Roucher C, Regier C, Sokhna C, Tall A, Trape JF. A 20-Year Longitudinal Study of Plasmodium ovale and Plasmodium malariae Prevalence and Mortality in a West African Population. Plos One. 2014; 9(2). https://doi.org/10.1371/journal.pone.0087169 PMID: 24520325

3. Sutherland CJ, Tanomsing N, Nolder D, Oguike M, Jennison C, Pukrittayakamee S, et al. Two Nonrecombinating Sympatric Forms of the Human Malaria Parasite Plasmodium ovale Occur Globally. Journal of Infectious Diseases. 2010; 201(10):1544–50. https://doi.org/10.1086/652240 PMID: 20380562

4. Nabarro LEB, Nolder D, Broderick C, Nadjm B, Smith V, Blaze M, et al. Geographical and temporal trends and seasonal relapse in Plasmodium ovale spp. and Plasmodium malariae infections imported to the UK between 1987 and 2015. Bmc Medicine. 2018; 16. https://doi.org/10.1186/s12916-018-1204-6 PMID: 30477484

5. Nolder D, Oguike MC, Maxwell-Scott H, Niyazi HA, Smith V, Chiodini PL, et al. An observational study of malaria in British travellers: Plasmodium ovale wallikeri and Plasmodium ovale curtisi differ significantly in the duration of latency. Bmj Open. 2013; 3(5). https://doi.org/10.1136/bmjopen-2013-002711 PMID: 23793668

6. Rojo-Marcos G, Rubio-Munoz JM, Angheben A, Jauregiquerry S, Garcia-Bujalance S, Tomasoni LR, et al. Prospective comparative multi-centre study on imported Plasmodium ovale wallikeri and Plasmodium ovale curtisi infections. Malaria Journal. 2018; 17. https://doi.org/10.1186/s12936-018-2544-6 PMID: 30376686

7. Ansari HR, Templeton TJ, Subudhi AK, Ramaprasad A, Tang JX, Lu F, et al. Genome-scale comparison of expanded gene families in Plasmodium ovale wallikeri and Plasmodium ovale curtisi with Plasmodium malariae and with other Plasmodium species. International Journal for Parasitology. 2016; 46(11):885–96. https://doi.org/10.1016/j.ijpara.2016.05.009 PMID: 27392654

8. Calderaro A, Piccò G, Perandini F, Gorrini C, Peruzzi S, Zuelli C, et al. Genetic polymorphisms influence Plasmodium ovale PCR detection accuracy. Journal of clinical microbiology. 2007; 45(5):1624–7. https://doi.org/10.1128/JCM.02316-06 PMID: 17360843.

9. Oguike MC, Sutherland CJ. Dimorphism in genes encoding sexual-stage proteins of Plasmodium ovale curtisi and Plasmodium ovale wallikeri. International Journal for Parasitology. 2015; 45(7):449–54. https://doi.org/10.1016/j.ijpara.2015.02.004 PMID: 25817462

10. Templeton TJ, Kaslow DC, Fidock DA. Developmental arrest of the human malaria parasite Plasmodium falciparum within the mosquito midgut via CTRP gene disruption. Mol Microbiol. 2000; 36(1):1–9. PMID: 10760158.

11. Le HG, Kang JM, Moe M, Jun H, Thai TL, Lee J, et al. Genetic polymorphism and natural selection of circumsporozoite surface protein in Plasmodium falciparum field isolates from Myanmar. Malaria Journal. 2018; 17. ARTN 361 https://doi.org/10.1186/s12936-018-2513-0 PMID: 30314440

12. Zeeshan M, Alam MT, Vinayak S, Bora H, Tyagi RK, Alam MS, et al. Genetic Variation in the Circumsporozoite Protein in India and Its Relevance to RTS,S Malaria Vaccine. Plos One. 2012; 7(8). https://doi.org/10.1371/journal.pone.0043430 PMID: 22912873

13. Bonilla JA, Valdivia C, Cummings R, Palmer CJ. Genetic diversity of Plasmodium vivax PVCSP and PV MSP1 in Guyana, south America. American Journal of Tropical Medicine and Hygiene. 2011; 41(6):677–83. https://doi.org/10.1016/j.ijpara.2011.01.004 PMID: 21315074

14. Imwong M, Pukrittayakamee S, Gruner AC, Renia L, Letourneur F, Looareesuwan S, et al. Practical PCR genotyping protocols for Plasmodium vivax using Pvs and Pvmsp1. Malaria Journal. 2005; 4. https://doi.org/10.1186/1475-2875-4-20 PMID: 15854233

15. Lo E, Nguyen K, Nguyen J, Hemming-Schroeder E, Xu JB, Etemesi H, et al. Plasmodium malariae Prevalence and csp Gene Diversity, Kenya, 2014 and 2015. Emerging Infectious Diseases. 2017; 23(4):601–10. https://doi.org/10.3201/eid2304.161245 PMID: 28322694

16. Saralamba N, Mayxay M, Newton PN, Smithuis F, Nosten F, Arчасакsan L, et al. Genetic polymorphisms in the circumsporozoite protein of Plasmodium malariae show a geographical bias. Malaria Journal. 2018; 17. https://doi.org/10.1186/s12936-018-2413-3 PMID: 30012172

17. Fong MY, Ahmed MA, Wong SS, Lau YL, Sitam F. Genetic Diversity and Natural Selection of the Plasmodium knowlesi Circumsporozoite Protein Nonrepea t Regions. Plos One. 2015; 10(9). https://doi.org/10.1371/journal.pone.0137734 PMID: 26379157

18. Chen JT, Li J, Zha GC, Huang G, Huang ZX, Xie DD, et al. Genetic diversity and allele frequencies of Plasmodium falciparum msp1 and msp2 in parasite isolates from Bioko Island, Equatorial Guinea. Malaria Journal. 2018; 17. https://doi.org/10.1186/s12936-018-2611-2 PMID: 30526609
19. Mwingira F, Nkwengu Gila G, Schoepflin S, Sumari D, Beck HP, Snounou G, et al. Plasmodium falciparum msp1, msp2 and glurp allele frequency and diversity in sub-Saharan Africa. Malaria Journal. 2011; 10. https://doi.org/10.1186/1475-2875-10-79 PMID: 21470428

20. Putaporntip C, Hughes AL, Jongwutiwes S. Low level of sequence diversity at merozoite surface protein-1 locus of Plasmodium ovale curtisi and P. ovale wallikeri from Thai isolates. PloS one. 2013; 8(3): e58962. https://doi.org/10.1371/journal.pone.0058962 PMID: 23536840

21. Snounou G, Singh B. Nested PCR analysis of Plasmodium parasites. Methods in molecular medicine. 2002; 72:189–203. https://doi.org/10.1385/1-59259-271-6:189 PMID: 12125116.

22. Tanomsing N, Imwong M, Theppabutr S, Pukrittayakamee S, Day NP, White NJ, et al. Accurate and sensitive detection of Plasmodium species in humans by use of the dihydrofolate reductase-thymidylate synthase linker region. Journal of clinical microbiology. 2010; 48(10):3735–7. https://doi.org/10.1128/JCM.00898-10 PMID: 20702666.

23. Tanomsing N, Imwong M, Sutherland CJ, Dolecek C, Hien TT, Nosten F, et al. Genetic marker suitable for identification and genotyping of Plasmodium ovale curtisi and Plasmodium ovale wallikeri. Journal of clinical microbiology. 2013; 51(12):4213–6. https://doi.org/10.1128/JCM.01527-13 PMID: 24068009.

24. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. NuclAcidsSympSer. 1999; 41:95–8.

25. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution. 2017; 34(12):3299–302. https://doi.org/10.1093/molbev/msx248 PMID: 29029172

26. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 2016; 33(7):1870–4. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

27. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution. 1987; 4(4):406–25. https://doi.org/10.1093/oxfordjournals.molbev.a040454 PMID: 3447015.

28. del Portillo HA, Longacre S, Khouri E, David PH. Primary structure of the merozoite surface antigen 1 of Plasmodium vivax reveals sequences conserved between different Plasmodium species. Proceedings of the National Academy of Sciences of the United States of America. 1991; 88(9):4030–4. https://doi.org/10.1073/pnas.88.9.4030 PMID: 2023952.

29. Escalante AA, Ayala FJ. Phylogeny of the malarial genus Plasmodium, derived from rRNA gene sequences. Proceedings of the National Academy of Sciences of the United States of America. 1994; 91(24):11373–7. https://doi.org/10.1073/pnas.91.24.11373 PMID: 7972067.

30. Escalante AA, Freeland DE, Collins WE, Lal AA. The evolution of primate malaria parasites based on the gene encoding cytochrome b from the linear mitochondrial genome. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95(14):8124–9. https://doi.org/10.1073/pnas.95.14.8124 PMID: 9653151.

31. Garrido-Cardenas JA G-C L, Manzano-Agugliaro F, Mesa-Valle C. Plasmodium genomics: an approach for learning about and ending human malaria. Parasitol Res. 2019; 118(1):1–27. https://doi.org/10.1007/s00436-018-6127-9 PMID: 30402656
| No. | Sample code | Area                | collected year | Species identification by nested PCR               |
|-----|-------------|---------------------|----------------|---------------------------------------------------|
| 1   | PoW1        | Thailand            | 1995           | P. ovale wallikeri                                 |
| 2   | PoW2        | Thailand            | 2004           | P. ovale wallikeri                                 |
| 3   | PoW3        | Thailand            | 2006           | P. ovale wallikeri                                 |
| 4   | PoW4        | Thailand            | 2006           | P. ovale wallikeri + P. vivax                     |
| 5   | PoW5        | Thailand            | 2006           | P. ovale wallikeri                                 |
| 6   | PoW6        | Thailand            | 2006           | P. ovale wallikeri + P. vivax + P. falciparum     |
| 7   | PoW7        | Thailand            | 2006           | P. ovale wallikeri                                 |
| 8   | PoW8        | Thailand            | 2006           | P. ovale wallikeri                                 |
| 9   | PoW9        | Thailand            | 2007           | P. ovale wallikeri + P. vivax                     |
| 10  | PoW10       | Thailand            | 2007           | P. ovale wallikeri                                 |
| 11  | PoW11       | Thailand            | 2007           | P. ovale wallikeri + P. vivax                     |
| 12  | PoW12       | Thailand            | 2007           | P. ovale wallikeri                                 |
| 13  | PoW13       | Africa (Guinea Bissau) | 1997     | P. ovale wallikeri                                 |
| 14  | PoW14       | Africa (Ghana)      | 2010           | P. ovale wallikeri                                 |
| 15  | PoC1        | Thailand            | 2006           | P. ovale curtisi                                   |
| 16  | PoC2        | Thailand            | 2007           | P. ovale curtisi                                   |
| 17  | PoC3        | Africa (Guinea Bissau) | 1999     | P. ovale curtisi                                   |
| 18  | PoC4        | Africa (Guinea Bissau) | 1995     | P. ovale curtisi + P. falciparum                   |
| 19  | PoC5        | Africa (Guinea Bissau) | 1995     | P. ovale curtisi                                   |
| 20  | PoC6        | Africa (Guinea Bissau) | 1995     | P. ovale curtisi                                   |
| 21  | PoC7        | Africa (Guinea Bissau) | 1995     | P. ovale curtisi + P. falciparum + P. malariae    |
| 22  | PoC8        | Africa (Guinea Bissau) | 1997     | P. ovale curtisi + P. falciparum                   |
| 23  | PoC9        | Africa (Guinea Bissau) | 1995     | P. ovale curtisi + P. falciparum                   |
| 24  | PoC10       | Africa (Guinea Bissau) | 1995     | P. ovale curtisi + P. falciparum                   |
| 25  | PoC11       | Africa (Sierra Leone) | 2010     | P. ovale curtisi                                   |
Amino acid pattern of partial MSP1 in *P. ovale* spp.

| No. | PoW/KoC | PGAGG | PGAAG | PGVPG | QAAT | QTAT | HAST | QATT | QVTT | QSAT |
|-----|---------|-------|-------|-------|------|------|------|------|------|------|
| 1   | PoW_KC137340 | 1     | 2     | 3     | 0    | 0    | 0    | 0    | 0    | 0    |
| 2   | PoW_KC137341 | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 3   | PoW_KC137342 | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 4   | PoW_KC137344 | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 5   | PoW_KC137345 | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 6   | PoW_LT594511 | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 7   | PoW1      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 8   | PoW2      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 9   | PoW3      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 10  | PoW4      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 11  | PoW5      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 12  | PoW6      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 13  | PoW7      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 14  | PoW8      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 15  | PoW9      | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 16  | PoW10     | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 17  | PoW11     | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 18  | PoW12     | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 19  | PoW13     | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 20  | PoW14     | 1     | 1     | 2     | 0    | 0    | 0    | 0    | 0    | 0    |
| 21  | PoC_FJ824670 | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 0    |
| 22  | PoC_FJ824671 | 0     | 0     | 0     | 4    | 2    | 3    | 1    | 1    | 1    |
| 23  | PoC_KC137343 | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 24  | PoC_KC137346 | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 25  | PoC_KC137347 | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 26  | PoC_KC137348 | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 27  | PoC_KC137349 | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 28  | PoC_KK672044 | 0     | 0     | 0     | 5    | 1    | 2    | 1    | 1    | 2    |
| 29  | PoC_KK672045 | 0     | 0     | 0     | 5    | 1    | 2    | 1    | 1    | 2    |
| 30  | PoC_LT594588 | 0     | 0     | 0     | 5    | 1    | 2    | 1    | 1    | 2    |
| 31  | PoC1      | 0     | 0     | 0     | 4    | 1    | 2    | 1    | 1    | 2    |
| 32  | PoC2      | 0     | 0     | 0     | 3    | 2    | 1    | 1    | 1    | 2    |
| 33  | PoC3      | 0     | 0     | 0     | 4    | 2    | 1    | 1    | 1    | 0    |
| 34  | PoC4      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 1    |
| 35  | PoC5      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 1    |
| 36  | PoC6      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 1    |
| 37  | PoC7      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 1    |
| 38  | PoC8      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 0    |
| 39  | PoC9      | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 0    |
| 40  | PoC10     | 0     | 0     | 0     | 4    | 2    | 2    | 1    | 1    | 1    |
| 41  | PoC11     | 0     | 0     | 0     | 4    | 2    | 3    | 1    | 1    | 0    |
Comparative analysis of single and mixed infections *P. ovale* spp.

| CSP Po spp. infection (area) | No. of samples | Haplotype diversity | Pi   | dN/dS  | Tajima's D | Fu and Li's D | Fu and Li's F |
|-----------------------------|----------------|---------------------|------|--------|------------|--------------|--------------|
| PoW single infection (Thailand) | 8 | 0.964 | 0.02233 | 1.71143 | 0.23227 | 0.35607 | 0.35972 |
| PoW single infection (Africa) | 2 | 0 | 0 | NA | NA | NA | NA |
| PoW mixed infection (Thailand) | 4 | 0.833 | 0.01641 | 0.8373 | 0.27601 | 0.18438 | 0.2147 |
| PoC single infection (Thailand) | 2 | 1 | 0.01693 | 1.42434 | NA | NA | NA |
| PoC single infection (Africa) | 4 | 1 | 0.07926 | 1.61314 | 0.20165 | 0.27026 | 0.27689 |
| PoC mixed infection (Africa) | 5 | 1 | 0.04328 | 1.76945 | 0.0391 | 0.01422 | 0.00386 |

| MSP1 Po spp. infection (area) | No. of samples | Haplotype diversity | Pi   | dN/dS  | Tajima's D | Fu and Li's D | Fu and Li's F |
|-----------------------------|----------------|---------------------|------|--------|------------|--------------|--------------|
| PoW single infection (Thailand) | 8 | 0.464 | 0.00068 | 0.32624 | 1.31009 | 1.4098 | 1.51361 |
| PoW single infection (Africa) | 2 | 0 | 0 | NA | NA | NA | NA |
| PoW mixed infection (Thailand) | 4 | 0.833 | 0.00135 | NA | 0.7099 | 0.7099 | 0.60427 |
| PoC single infection (Thailand) | 2 | 1 | 0.01624 | 0.37 | NA | NA | NA |
| PoC single infection (Africa) | 4 | 1 | 0.20957 | 1.21774 | 0.38932 | 0.38932 | 0.41228 |
| PoC mixed infection (Africa) | 5 | 1 | 0.02211 | 0.99476 | 0.01648 | 0.01648 | 0.0178 |

| CTRP Po spp. infection (area) | No. of samples | Haplotype diversity | Pi   | dN/dS  | Tajima's D | Fu and Li's D | Fu and Li's F |
|-----------------------------|----------------|---------------------|------|--------|------------|--------------|--------------|
| PoW single infection (Thailand) | 8 | 0.929 | 0.00423 | 1.39076 | 0.73607 | 0.88945 | 0.9524 |
| PoW single infection (Africa) | 2 | 1 | 0.00179 | 0.47854 | NA | NA | NA |
| PoW mixed infection (Thailand) | 4 | 0.833 | 0.00352 | 1.68995 | 0.63894 | 0.63894 | 0.6763 |
| PoC single infection (Thailand) | 2 | 1 | 0.00314 | 1.44635 | NA | NA | NA |
| PoC single infection (Africa) | 4 | 1 | 0.00486 | 0.76254 | 0.29479 | 0.29479 | 0.31326 |
| PoC mixed infection (Africa) | 5 | 1 | 0.00319 | 1.17857 | 0.35839 | 0.35839 | 0.38759 |