Technical Appendix 1

Detailed description of PCR-based diagnostic assays and phylogenetic analyses of csp sequences

Microscopy

Slides were examined under microscopes 100× objective. Parasites were counted against 200 leukocytes. A slide was considered negative when no parasites were observed after counting over 100 microscopic fields. All slides were read in duplicate by two microscopists at the time of sample collection. In the case of discordance, the slides were examined by a third microscopist. The density of parasitemia was expressed as the number of asexual parasite per microliter of blood, assuming a leukocyte count of 8000 cells per microliter according to the WHO guidelines.

Nested PCR assay of Plasmodium species

Parasite DNA was extracted from half of a 50ul-dried blood spot by the Saponin/Chelex method (1). The final extracted volume was 200μl. A nested amplification of the 18S rRNA gene region of Plasmodium (P. falciparum, P. vivax, P. malariae and P. ovale) was used for parasite detection and species identification in all samples. DNA from P. falciparum isolates 7G8 (MR4-MRA-926) and HB3 (MR4-MRA-155), P. vivax Pakchong (MR4-MRA-342G) and Nicaragua (MR4-MRA-340G), P. malariae (MR4-MRA-179), as well as P. ovale (MR4-MRA-180) were used as positive controls in all amplifications. Water and uninfected samples were used as negative controls to ensure lack of contamination. Amplification was conducted in a 20ul reaction mixture containing 2ul of genomic DNA, 10ul of 2×DreamTaq™ Green PCR Master Mix (Fermentas), and 0.3uM primers. Reaction was performed in a BIORAD MyCycler thermal cycler following the published protocol (2). The amplified products were resolved electrophoretically on a 2% agarose gel in 0.5×Tris-borate (TBE) buffer and visualized under UV light.
Quantitative real-time PCR assay of Plasmodium species

Parasite DNA amount was estimated using the SYBR Green qPCR detection method with Plasmodium species-specific primers that targeted the 18S rRNA genes (3–5). Amplification was conducted in a 20µL reaction mixture containing 2µL of genomic DNA, 10µL 2×SYBR Green qPCR Master Mix (Thermo Scientific, USA), and 0.5µM primer. Reaction was performed in CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad), with an initial denaturation at 95°C for 3 min, followed by 45 cycles at 94°C for 30 sec, 55°C for 30 sec, and 68°C for 1 min with a final 95°C for 10 sec. This was then followed by a melting curve step of temperature that ranged from 65°C to 95°C with 0.5°C increment to determine the melting temperature of each amplified product. Melting curve analyses were performed for each amplified sample to confirm specific amplifications of the target sequence. For the measure of reproducibility of the threshold cycle number (Ct), the mean value and standard deviations were calculated from triplicates in two independent assays. A cut-off threshold of 0.02 fluorescence units that robustly represented the threshold cycle at the log-linear phase of the amplification and above the background noise was set to determine Ct value for each assay. Samples yielding Ct values higher than 40 (as indicated in the negative controls) were considered negative for Plasmodium species. The parasite gene copy number (GCN) in a sample was quantified based on the threshold cycle using the follow equation (6): GCNsample = e^{[E×ΔCt_{sample}]}, where GCN stands for gene copy number, ΔCt for the difference in threshold cycle between the negative control and the sample, and E for amplification efficiency. The amplification efficiency of primers was assessed on 10-fold serial dilutions ranging from 10^5 to 10^1 copies/µl of the control plasmids. DNA from *P. falciparum* isolates 7G8 (MR4-MRA-926) and HB3 (MR4-MRA-155), and *P. malariae* (MR4-MRA-179) isolate were used as positive controls. Water and uninfected samples were used as negative controls in all amplifications.

**CSP sequencing and phylogenetic analyses**

Four internal primers were designed specifically on the *P. malariae csp* gene region and used together with the published primers (7) (Technical Appendix Table) to unambiguously amplify the three segments, the N-terminal, the central repeat, and the C-terminal regions of the csp gene. A total of 37 *P. malariae* isolates were amplified and sequenced. Amplification was conducted in a 20µL reaction mixture containing 2µL of genomic DNA, 10µL 2××DreamTaq™ Green PCR Master Mix (Fermentas) and 0.5µL of 10 µM primers. PCR cycles included an initial
denaturing step at 95°C for 3 min, 40 cycles of 95°C for 30 sec, 48-50°C for 30 sec, and 72°C for 2 min, followed by an additional extension at 72°C for 5 min in a Bio-Rad MyCycler Thermal Cycler. PCR products were visualized on 1% agarose gel and then purified and sequenced from both ends with BigDye Terminator v3.1 Cycle Sequencing Kit on an ABI 3130xl sequencer.

All obtained sequences were blasted against NCBI GenBank database for verification. They were translated into protein sequences and analyzed together with all available csp protein sequences of *P. malariae* as well as *P. brasilianum* retrieved from the GenBank database. Due to potential alignment errors associated with gaps in the nucleotide sequences, translated amino acid sequences with unambiguous indels were used in phylogenetic analyses. Sequences were aligned with MUSCLE v3.7 (8) using default settings followed by manual editing in Sequence Alignment Editor v1.d1 (9). A phylogenetic tree was reconstructed using the maximum likelihood method implemented in the PhyML program v3.0 (10). The WAG (Whelan And Goldman) substitution model, which assumes an estimated proportion of invariant sites and four gamma-distributed rate categories to account for rate heterogeneity across sites, was selected. Resulted trees were visualized in FigTree v1.4.2.

Sequence diversity including measures of evolutionary distances and average pairwise divergence were estimated using SSE v1.2 (11). The matrix-normalized distances based on the standard PAM model (12) that accounts for the probability of change from one amino acid to another were calculated. In addition, a similarity scan was performed between and within sequences using the standard PAM-Dayhoff matrix to normalize a matching score. Regions that meet or exceed the set criteria of the number of matches were plotted on a dot plot graph. *Plasmodium malariae* of East Africa (Kenya and Uganda), Central/West Africa (Cameroon and Cote d’Ivoire), and South America (Brazil and Venezuela) were compared for level of sequence diversity among gene regions.

References


Technical Appendix Table. Primer sequences and PCR conditions of the circumsporozoite protein (csp) gene

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
<th>Expected size (bp)</th>
<th>Annealing temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>csp-F*</td>
<td>ATGAAGAAGTTATCTGTCTTAGCAATATCC</td>
<td>280</td>
<td>50°C</td>
</tr>
<tr>
<td>csp-280R</td>
<td>CCGGGGGGTGTTTCAATTATTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csp-280F</td>
<td>GGTGTTGAAAAATTAAATGAAACAACC</td>
<td>700-800</td>
<td>48°C</td>
</tr>
<tr>
<td>csp-1070R</td>
<td>CCACCTTTATTATCCTTTTTGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>csp-1070F</td>
<td>GCGAAAAATAGGATAATAAGTGG</td>
<td>400</td>
<td>50°C</td>
</tr>
<tr>
<td>csp-R*</td>
<td>TTAGTGAAAGAGTTAGCTAAAAAC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Primer published in (7).

Technical Appendix Figure 1. Scatter plot showing the significant correlation of parasite gene copy number measured by quantitative real-time PCR and parasitemia by microscopy of *Plasmodium malariae* isolates.
Technical Appendix Figure 2. Scatter plots showing (A) parasite gene copy number of *P. malariae* and *P. falciparum* ranked from low to high *P. malariae* parasite gene copy number and (B) parasitemia of *P. malariae* and *P. falciparum* ranked from low to high *P. malariae* parasitemia of co-infected samples.
Technical Appendix Figure 3. Scatter plot showing the non-significant correlation of *P. malariae* gene copy number against age among samples.