RESEARCH ARTICLE

REVISED Targeted amplicon deep sequencing of ama1 and mdr1 to track within-host *P. falciparum* diversity throughout treatment in a clinical drug trial [version 4; peer review: 2 approved, 2 approved with reservations]

Kevin Wamae[1*], Leonard Ndwiga[1*], Oksana Kharabora[2], Kelvin Kimenyi[1], Victor Osoti[1], Zaydah de Laurent[1], Juliana Wambua[1], Jennifer Musyoki[1], Caroline Ngetsa[1], Peter Kalume[1], Gabriel Mwambingu[1], Mainga Hamaluba[1,3], Rob van der Pluijm[3,4], Arjen M. Dondorp[3,4], Jeffrey Bailey[5], Jonathan Juliano[2,6,7], Philip Bejon[1,3], Lynette Ochola-Oyier[1]

[1]Bioscience, KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya
[2]Division of Infectious Diseases, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599, USA
[3]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[4]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand
[5]Department of Pathology and Laboratory Medicine, Warren Alpert Medical School, Brown University, Providence, RI, 02903, USA
[6]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27516, USA
[7]Curriculum in Genetics and Molecular Biology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

* Equal contributors

v4 **First published:** 16 Mar 2022, **7**:95
https://doi.org/10.12688/wellcomeopenres.17736.1
Second version: 15 Dec 2022, **7**:95
https://doi.org/10.12688/wellcomeopenres.17736.2
Third version: 25 Oct 2023, **7**:95
https://doi.org/10.12688/wellcomeopenres.17736.3
Latest published: 14 Mar 2024, **7**:95
https://doi.org/10.12688/wellcomeopenres.17736.4

Abstract

Introduction

Antimalarial therapeutic efficacy studies are routinely conducted in malaria-endemic countries to assess the effectiveness of antimalarial treatment strategies. Targeted amplicon sequencing (AmpSeq) uniquely identifies and quantifies genetically distinct parasites within an infection. In this study, AmpSeq of *Plasmodium falciparum* apical membrane antigen 1 (*ama1*), and multidrug resistance gene 1 (*mdr1*),

Open Peer Review

Approval Status ? ✓ ✓ ?

	1	2	3	4
version 4 (revision) 14 Mar 2024				
version 3 (revision) 25 Oct 2023				
version 2 (revision) 15 Dec 2022		✓ view	✓ view	? view

were used to characterise the complexity of infection (COI) and drug-resistance genotypes, respectively.

Methods

P. falciparum-positive samples were obtained from a triple artemisinin combination therapy clinical trial conducted in 30 children under 13 years of age between 2018 and 2019 in Kilifi, Kenya. Nine of the 30 participants presented with recurrent parasitemia from day 26 (624h) onwards. The *ama1* and *mdr1* genes were amplified and sequenced, while *msp1*, *msp2 and glurp* data were obtained from the original clinical study.

Results

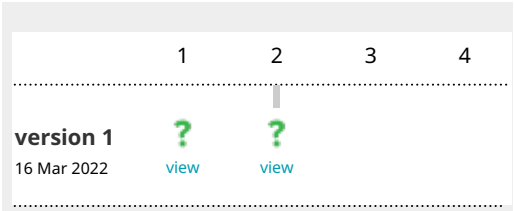
The COI was comparable between *ama1* and *msp1*, *msp2 and glurp*; overall, *ama1* detected more microhaplotypes. Based on ama1, a stable number of microhaplotypes were detected throughout treatment until day 3. Additionally, a recrudescent infection was identified with an *ama1* microhaplotype initially observed at 30h and later in an unscheduled follow-up visit. Using the relative frequencies of *ama1* microhaplotypes and parasitemia, we identified a fast (<1h) and slow (>5h) clearing microhaplotype. As expected, only two *mdr1* microhaplotypes (NF and NY) were identified based on the combination of amino acid polymorphisms at codons 86 and 184.

Conclusions

This study highlights AmpSeq as a tool for highly-resolution tracking of parasite microhaplotypes throughout treatment and can detect variation in microhaplotype clearance estimates. AmpSeq can also identify slow-clearing microhaplotypes, a potential early sign of selection during treatment. Consequently, AmpSeq has the capability of improving the discriminatory power to distinguish recrudescences from reinfections accurately.

Keywords
Artemisinin-based combination therapy, pfama1, Pfmdr1, artemisinin resistance, antimalarial resistance, targeted deep sequencing, deep sequencing, msp1, msp2, glurp

KEMRI wellcome This article is included in the KEMRI | Wellcome Trust gateway.



1. **Ian M. Hastings**, Liverpool School of Tropical Medicine, Liverpool, UK
2. **Mita Toshihiro**, Juntendo University, Tokyo, Japan
3. **Eniyou Cheryll Oriero**, Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia
4. **Aurel Holzschuh**, University of Notre Dame, Notre Dame, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Lynette Ochola-Oyier (LiOchola@kemri-wellcome.org)

Author roles: Wamae K: Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ndwiga L**: Formal Analysis, Investigation, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kharabora O**: Investigation, Validation, Writing – Review & Editing; **Kimenyi K**: Formal Analysis, Investigation, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Osoti V**: Data Curation, Formal Analysis, Investigation, Writing – Review & Editing; **de Laurent Z**: Data Curation, Formal Analysis, Investigation, Validation, Writing – Review & Editing; **Wambua J**: Project Administration, Writing – Review & Editing; **Musyoki J**: Project Administration, Writing – Review & Editing; **Ngetsa C**: Project Administration, Writing – Review & Editing; **Kalume P**: Formal Analysis, Investigation, Writing – Review & Editing; **Mwambingu G**: Formal Analysis, Investigation, Writing – Review & Editing; **Hamaluba M**: Conceptualization, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing; **van der Pluijm R**: Conceptualization, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing; **Dondorp AM**: Conceptualization, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing; **Bailey J**: Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – Review & Editing; **Juliano J**: Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – Review & Editing; **Bejon P**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ochola-Oyier L**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Developing Excellence in Leadership, Training and Science (DELTAS) Africa Initiative (grant number: DEL-15-003) and a Wellcome Trust Intermediate Fellowship awarded to L.I.O.-O (grant no. 107568/Z/15/Z). J.A.B and J.J.J received support from the National Institutes of Health (R01AI121558 to JAB and K24AI13499 and R01AI121558 to JJJ). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (107769/Z/10/Z) and the UK government. The views expressed in this publication are those of the authors and not necessarily those of AAS, NEPAD Agency, Wellcome Trust, or the UK government.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Wamae K, Ndwiga L, Kharabora O *et al.* **Targeted amplicon deep sequencing of ama1 and mdr1 to track within-host *P. falciparum* diversity throughout treatment in a clinical drug trial [version 4; peer review: 2 approved, 2 approved with reservations]** Wellcome Open Research 2024, **7**:95 https://doi.org/10.12688/wellcomeopenres.17736.4

First published: 16 Mar 2022, **7**:95 https://doi.org/10.12688/wellcomeopenres.17736.1

REVISED **Amendments from Version 3**

This version contains a revised Table S1 highlighting the parasitemia and DNA concentrations of the artificial DNA mixture used as sequencing controls.

Table S4 shows the occurrence of all microhaplotypes across all participants, while Table S5 shows the occurrence of rare microhaplotypes.

The term 'aliquot' has since been used to designate the multiple portions taken from the control samples. Similarly, the term 'duplicate' refers to the repeated measurements or observations taken from these control aliquots and samples to ensure consistency and reliability in our results.

The methods have been revised to provide more clarity, such as thresholds for read-depths used.

Any further responses from the reviewers can be found at the end of the article

Introduction

Artemisinin-based combination therapies (ACTs) have led to high cure rates for *P. falciparum* malaria (Bhatt *et al.*, 2015). Still, artemisinin (ART) resistance emerged and spread in Southeast (SE) Asia, evidenced by delayed parasite clearance following ACT treatment (Ashley *et al.*, 2014; Dondorp *et al.*, 2009; Noedl *et al.*, 2008; van der Pluijm *et al.*, 2019). Two recent studies have identified early signs of ART partial resistance in Rwanda (Uwimana *et al.*, 2021) and Uganda (Balikagala *et al.*, 2021), and this looming threat of widespread ACT resistance would be catastrophic in sub-Saharan Africa, where the burden of malaria is the most significant (WHO, 2021a).

To minimise the development of drug-resistant parasites and rescue a regimen with an already failing component of ACTs, novel chemotherapeutic strategies involving the roll-out of triple artemisinin-based combination therapies (TACTs) are being evaluated (van der Pluijm *et al.*, 2021). TACTs combine an established ACT with a second, slowly eliminated partner drug for additional antimalarial activity and protection of partner drug resistance. The potential advantage of TACTs is supported by evidence from Cambodia that artemisinin partner drugs may exert opposing selection pressures making it difficult to adapt to multiple partner drugs simultaneously (Parobek *et al.*, 2017). The safety and efficacy of this approach have been shown in clinical trials (Hamaluba *et al.*, 2021; van der Pluijm *et al.*, 2020), and the antimalarial therapeutic outcomes are assessed for a maximum of 42 days. Based on molecular methods, recurrent parasitemia during this period is classified as a new (reinfection) or recrudescent infection. The former is determined when genotyping methods find the recurrent parasites are distinguishable from those in the pre-treatment infection, and the latter when the parasites are indistinguishable. The standard genotyping method termed PCR correction examines three-length polymorphic markers in parasites, namely merozoite surface protein 1 (*msp1*), *msp2* and glutamate-rich protein (*glurp*) (Snounou & Beck, 1998). The amplicon sizes of these three markers are compared between pre-treatment and post-treatment parasites by either gel or capillary electrophoresis (Liljander *et al.*, 2009; WHO, 2003). However, the use of *msp1*/*msp2*/*glurp* genotyping is met with challenges such as the reliance on gel electrophoresis being limited to discriminating alleles of similar sizes (those with size differences less than 20 bp) and the inability to detect low-density parasite clones (Felger *et al.*, 2020; Jones *et al.*, 2021). Therefore, studies that rely on these markers may underestimate parasite diversity, are insensitive to low-abundant variants and are not quantitative for relative proportions of circulating parasite clones. Targeted amplicon sequencing, referred to as amplicon sequencing (AmpSeq) from here henceforth, offers high sensitivity in detecting minority parasite variants, quantifying the number of variants and their relative frequencies. AmpSeq also offers high-throughput sequencing of *P. falciparum* diversity and drug-resistance markers (Gruenberg *et al.*, 2019; Miller *et al.*, 2017). Apical membrane antigen 1 (*ama1*) is a highly polymorphic merozoite surface antigen (Polley & Conway, 2001) and serves as an excellent marker to explore parasite diversity within infections. On the other hand, mutations at codons N86Y and F184Y of the multidrug resistance gene (*mdr1*) modulate parasite susceptibility to ACT partner drugs such as amodiaquine, lumefantrine, piperaquine and mefloquine (Veiga *et al.*, 2016). Additionally, the rollout of ACTs has led to an increase in the *mdr1*-NFD microhaplotype (based on the combination of amino acid polymorphisms at codons 86, 184 and 1246) across several African studies, possibly due to ACT selection pressure (Okell *et al.*, 2018).

In this study, we examined samples from a TACT efficacy study in Kilifi (Hamaluba *et al.*, 2021). The administration of drugs was done under observation, and the patients were monitored in the hospital for three days of treatment. Frequent blood samples were obtained for pharmacokinetic analyses that were subsequently used for AmpSeq. The purpose of this study was to examine the utility of a genetic diversity marker (*ama1*) combined in a single deep sequencing run with a drug resistance marker (*mdr1*) to identify and track changes in the complexity of infections (COI), i.e., the number of *ama1* genotypes per infection as well as the *mdr1* wild-type and mutant genotypes throughout treatment.

Methods

Study design

P. falciparum positive samples were obtained from the TACT Kenya clinical trial conducted from 2018 to 2019 (ClinicalTrials.gov Identifier: NCT03452475) described in Hamaluba *et al.* (2021). The three-drug arms were arterolane-piperaquine (ART-PQ), arterolane-piperaquine + mefloquine (ART-PQ+MF) and artemether-lumefantrine (AL). A random sample of 30 individuals was selected, including all their sampling time-points: 0 hours (h), 0.5h, 1h, 2h, 3h, 4h, 6h, 8h, 12h, 18h, 24h, 30h, 36h, 42h, 48h, 72h (day 3), 168h (day 7), 336h (day 14), 504h (day 21), 672h (day 28), 840h (day 35), 1008h (day 42) and the hour of recurrent infection (REC, any sample taken during an unscheduled visit by the study participant, Table 1). 9/30 participants presented with recurrent parasitemia based on microscopy from 624h onwards, making a total of 609 individual samples. This study was approved by the Oxford Tropical Research Ethics Committee in the United Kingdom and the Kenya Medical Research Institute (KEMRI) -Scientific and Ethics Review Unit (SERU).

Table 1. Characteristics of the study participants.

Characteristic	Antimalarial regimen			*p-value*
	ART+PQ	ART+PQ+MF	AL	
Number of Participants [n=30]	11	11	8	0.41
Median age in years [Range]	6.7	5.7	10.65	0.51
	[2.7 – 10.3]	[2.1 – 11.9]	[6.0 - 12.6]	
Gender [Females]	3	7	3	0.12
Median Parasitemia per μl [Range]	142,236	78,274	90,158.5	0.7
	[8560 – 571,530]	[15,232 – 326,020]	[25,328 – 266,146]	

ART-PQ - arterolane-piperaquine, ART-PQ+MF arterolane-piperaquine + mefloquine and AL - artemether-lumefantrine.

DNA preparation and PCR from sequencing controls and clinical samples

Henceforth, we use the term "microhaplotype" to refer to the set of amino acid polymorphisms found on a single DNA amplicon. DNA was extracted from culture-adapted laboratory *P. falciparum* isolates, 3D7 and Dd2 (BEI Resources), and from 200µl of frozen patient blood samples using the QIAamp DNA Blood Mini Kit (Qiagen) according to the manufacturer's instructions. The DNA from 3D7 and Dd2 were mixed in the following ratios to generate artificial mixtures of sequencing controls: 1:1, 0.75:0.25, 0.85:0.15, 0.95:0.05 and 1:0 to determine the lowest limit of a microhaplotype detection. The level of parasitemia for the *P. falciparum* cultures from which 3D7 and Dd2 DNA were extracted is typically above 1% (www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-102.aspx and www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-150.aspx). These levels correspond to an approximate density of 50,000 parasites/µl, while the specific parasitemia for each isolate in the final mixture is presented in Table S1. The sequencing controls allowed for the detection of two ama1 (3D7 and Dd2) and three mdr1 (one from 3D7 and two from Dd2) microhaplotypes. Dd2 contained two mdr1 gene copies generated due to adaptation to in vitro culture. Amplicons spanning *ama1* (PF3D7_1133400, nucleotides 441–946) and *mdr1* (PF3D7_0523000, nucleotides 183–719) were generated in duplicate from each control and sample. This process was applied to each of the six aliquots prepared for every control, and to a single aliquot from each sample using primers designed in this study (Table S2) as follows: 1µl of template DNA (final amount <50ng), 0.2µl of Q5® High-Fidelity DNA Polymerase (final concentration 0.02U/µl, New England BioLabs), 1µl (10mM) forward primers tagged with Roche® molecular identifiers (MIDs, Table S2), and reverse primers, 0.4µl of 10mM dNTPs, 4µl of 5X Q5 reaction buffer, and 12.4µl of nuclease-free water. For both *ama1* and *mdr1,* the cycling conditions were: initial denaturation (98°C - 30 sec), followed by 30 cycles of denaturation (98°C - 10 sec), annealing (60°C - 30 sec), extension (72°C - 30 sec), and final extension (72°C - 2 min). PCR products were visualised on 1% agarose gels stained with RedSafe™ Nucleic Acid Staining Solution (iNtRON Biotechnology DR). Amplicon failures were repeated with 1.5µl of template DNA.

Amplicon Library Preparation and Sequencing

PCR amplicons were purified using the Zymo ZR-96 DNA Clean & Concentrator-5 Kit (Zymo Research) and quantified using Quant-iT™ dsDNA Assay Kit, High Sensitivity (Invitrogen). Both procedures were done following the manufacturer's instructions. Subsequently, the PCR amplicons were normalised to equal amounts of 1ng each using EB Buffer (Qiagen) and mixed to create amplicon pools of non-overlapping 26 MIDs. The KAPA Dual-Indexed Adapter Kit and the KAPA Hyper Prep Kit (Roche) were used for library preparation, and the Agilent High Sensitivity D1000 ScreenTape System (5067-5584) confirmed adapter ligation. Eventually, the *ama1* and *mdr1* amplicon libraries were mixed to generate the final pool for paired-end sequencing (2x300bp chemistry) using MiSeq Reagent Kit v3 (Illumina).

Sequence data analysis

Data extraction, quality control processing, and microhaplotype clustering were performed using SeekDeep v3.0.0 (Hathaway *et al.*, 2018). We implemented SeekDeep's default threshold of 250 reads as the minimum required read-depth for each individual PCR replicate. Additionally, for a sample to be included in the analysis, it needed to have a combined total of at least 500 reads, summing the read counts from all its replicates. For samples that met this criterion, microhaplotypes were discarded if they did not occur in the two PCR duplicates and if their relative frequency was <5% (or less than 25 reads). A conservative cut-off of 5% was set based on the isolate with the least proportion in one of the artificial mixtures of sequencing controls (ratio of 0.95 3D7 to 0.05 Dd2) unless the microhaplotype was independently detected in other samples at >5%. Chimeric reads were considered PCR artefacts and discarded. The relative microhaplotype frequency in a sample was calculated as the number of

reads of each microhaplotype over the total number of reads per sample. COI was defined as the number of distinct *ama1* microhaplotypes (varying at the nucleotide level) in each sample, while codons 86 and 184 defined the *mdr1* microhaplotypes. *ama1* expected heterozygosity (H_e) was calculated using the formula below, where n is the sample size and p_i is the relative frequency of the ith microhaplotype in the population (Nei, 1978):

$$h = \frac{n}{n-1}(1 - \sum p_i^2)$$

msp1/msp2/glurp genotyping was performed according to the WHO-recommended method of gel electrophoresis (WHO, 2008). These data were obtained from the original study, and the following was ensured during the analysis: Each PCR product had to have well-defined and easy to visualise, bands had to be bright and sharp to be of sufficient quality for scoring, PCRs were repeated if bands appeared in the negative control, the interpretation of results did not include products with less than 100 bp and did not account for faint bands or bands that formed smile-shaped patterns on the gel (Hamaluba *et al.*, 2021). All statistical analyses were carried out in R v4.0.2, and all plots were generated using the R packages ggplot2 v3.3.1 and ggpubr v0.3.0 (Kassambara, 2020; Wickham, 2016).

Based on simulation studies of amplicon sequencing data analysis (Jones *et al.*, 2021), we set the lower sampling limit for a parasite (blood sampling limit) at ten parasites/µl. Finally, using the complexity of infection for each sample, we back-calculated each parasite isolate's parasitemia to determine which isolates were at risk of falling below the sampling limit.

Parasite clearance estimation

One of the early signs of slow clearing parasites is a clearance half-life greater than five hours (Ashley *et al.*, 2014). Therefore, parasite clearance half-lives were calculated using the Worldwide Antimalarial Resistance Network's (WWARN) parasite clearance estimator (Flegg *et al.*, 2011). This was done for the 30 participants by extrapolating the total parasitemia to each *ama1* microhaplotype per infection. An estimate of the parasitaemia for each *ama1* microhaplotype was calculated by multiplying the parasitaemia (based on microscopy) at each time point by the frequency of each microhaplotype (Mideo *et al.*, 2016) based on the number of reads per microhaplotype over the total number of reads per sample. These estimates were plotted as histograms.

Results

AmpSeq in artificial mixtures of sequencing controls

The expected microhaplotypes were successfully detected from the sequencing controls (two from *ama1* and three from *mdr1*). Additionally, the 3D7 and Dd2 *ama1* microhaplotypes were detected consistently across all mixtures and in the expected proportions. This provided evidence that the assay could detect mixed infections in clinical samples, but only when the minor microhaplotype was at a relative frequency of ≥5% (Figure S1).

AmpSeq of pre-and post-treatment samples

From the 30 individuals sampled, 11 were in both the ART+PQ and ART+PQ+MF drug arms, respectively, while 8 were in the AL drug arm. There was no difference in the baseline median parasitaemia between the drug arms (Table 1). From the available 608 samples (timepoints 0h–1008h), *ama1* and *mdr1* sequence data were successfully obtained from 330 and 233 samples, respectively (Figure S2A and B). Samples were grouped into three categories based on parasitemia: high (>5,000), moderate (100–5,000) and low (<100 parasites/µl). Many samples collected between 0h–12h had high parasitemia, those collected between 18h–30h had moderate parasitemia, while those collected after 30h were primarily low parasitemia (Figure S2C). The median read depth was 11,147 reads (range 580 –33,714) for *ama1* and 11,548 reads (range 1,022 – 55,664) for *mdr1*, consistent with decreasing parasitemia, post-treatment samples had lower sequencing success.

Pfama1 genetic diversity during and after treatment

Throughout treatment, the mean COI (Figure S3) and number of *ama1* microhaplotypes (Figure 1) were relatively stable and the expected heterozygosity for *ama1* was high, 0.96. The mean COI and 95% confidence interval (CI) at 0h, 0.5h – 72h (during treatment) and post-treatment (after 72h) was 2.44 [2.13-1.75], 2.37 [2.27-2.48] and 2.0 [1.28-2.7], respectively. Overall, 33 *ama1* microhaplotypes were detected from the 330 successfully sequenced samples (Table S3), and only 10 of these microhaplotypes were detected at frequencies >5%.

Two groups of individuals were identified based on COI, 11 individuals with monoclonal infections compared to 19 polyclonal infections. For participants with monoclonal infections and with sequence data only up to 72h, the same *ama1* microhaplotype was seen throughout. Only one exception, individual PID42 had data post-72h and a change in the *ama1* microhaplotype was detected before and after 72h (Figure 1A, Table S4).

Participants with polyclonal infections harboured more than one *ama1* microhaplotype at any one-time point. In individuals with sequence data up to 72h (n=11), the *ama1* microhaplotypes detected maintained relatively stable frequencies up to <24h, with most changes occurring after 24h. In five individuals (PIDs 10, 30, 32, 40, 59, 60 and 61), there were sporadic detections of rare *ama1* microhaplotypes (Figure 1B, Table S4), including one microhaplotype each in PID10 at 8h and 12h, PID30 at 4h, PID32 at 8h, PID38 at 30h, PID40 at 0h, PID49 at 42h, PID60 at 48h, PID63 at 30h and 72h, PID61 at 12h and PID65 at 6h. Rare microhaplotypes, except for one detected exclusively in PID30, were identified in multiple samples, often at relative frequencies above 5% (Table S5). Additionally, except for PID60, all such sporadic microhaplotypes were detected above the sampling limit of 10 parasites/µl (see Table S6). The analysis revealed no significant difference in read counts between samples harboring either monoclonal or polyclonal infections (p = 0.092, Wilcoxon signed-rank test).

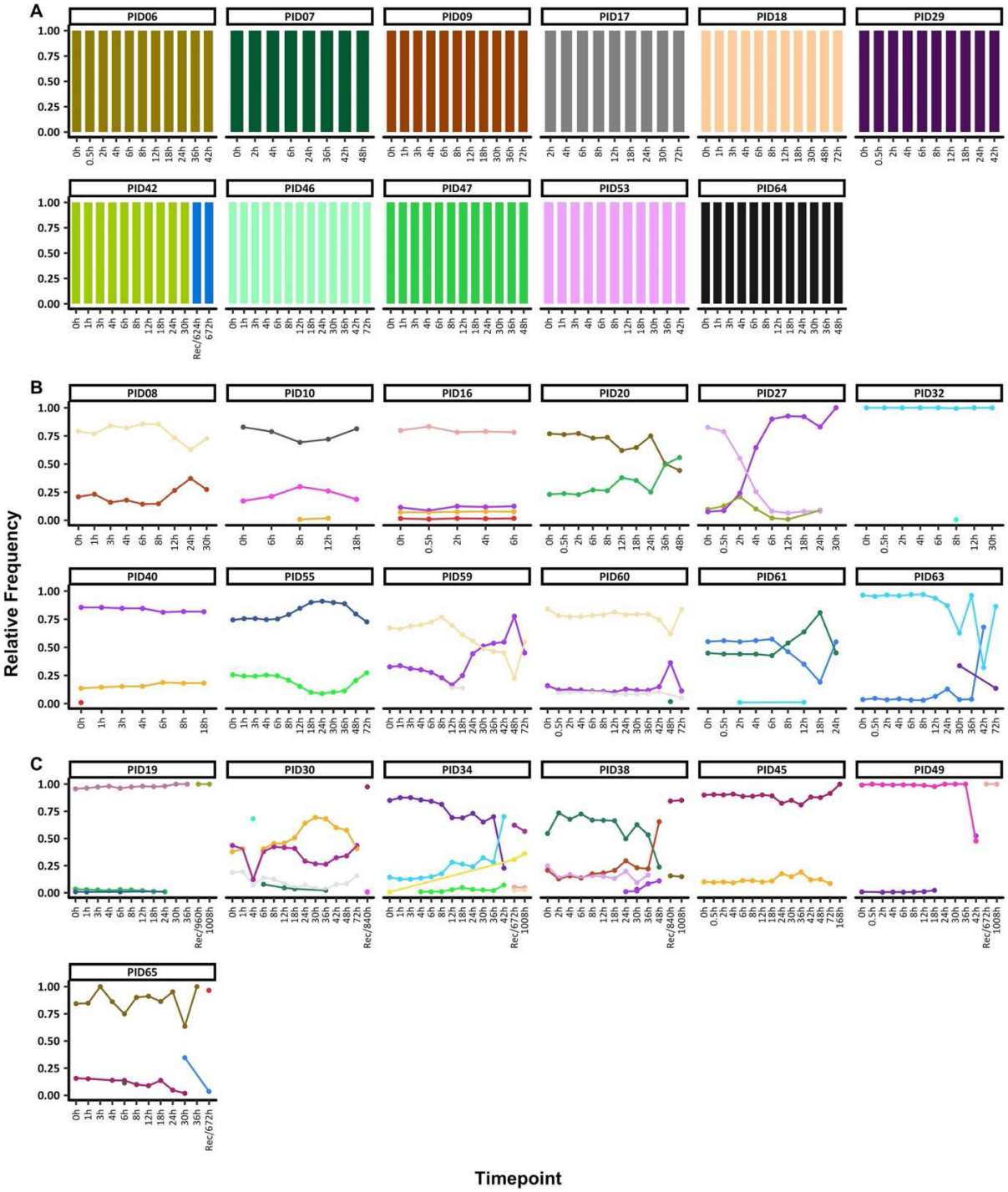


Figure 1. Temporal changes in *ama1* microhaplotypes throughout treatment. The figure highlights the individuals with **A**) monoclonal infections, if they had only one ama1 microhaplotype throughout treatment, **B**) polyclonal infections, if they had more than one *ama1* microhaplotype throughout treatment and sampled up to 72h and **C**) polyclonal infections, if they more than one *ama1* microhaplotype throughout treatment and sampled beyond 72h. Each coloured barplot represents a unique *ama1* microhaplotype, and matching colours represent the same *ama1* microhaplotype. Above each plot is the respective participant ID. The x-axis displays time points at which samples were collected, measured in hours. 'Rec' denotes samples taken during unscheduled visits due to the recurrence of parasites, although parasitemia data at these recurrence timepoints are not available, hence the line plot does not extend to these points. The primary y-axis corresponds to the bar charts, quantifying the relative distribution of ama1 microhaplotypes. The secondary y-axis (left) aligns with the dashed line graph, indicating the parasitemia levels on a log10 scale per microliter of blood. Logarithmic values of zero in the log10 parasitemia levels indicate samples that were determined to have zero parasitemia when examined using microscopy.

Participant PID27 experienced a drastic change in the relative microhaplotype frequencies by 4h post-treatment. The least dominant microhaplotype at 0h had become dominant by the 4h and remained this way up to the last timepoint with sequencing data (30h). All the remaining seven individuals with post-treatment data cleared their infections. However, the recurrent sample (672h) for PID65 contained an *ama1* microhaplotype present in the 30h sample, likely from a minor microhaplotype not detected at 0h (Figure 1C, Table S4). All participants with polyclonal infections also had polyclonal baseline samples (0h), except for PID32 who had only one polyclonal sample at 8h.

AmpSeq compared to msp1/msp2/glurp genotyping of recurrent infections

Based on microscopy data, all the nine individuals recurrences (two in ART-PPQ, two in ART-PPQ+MF, and five in the AL arm) were categorised as new infections based on *msp1/msp2/glurp* data (Hamaluba *et al.*, 2021). *msp1/msp2/glurp* identified a total of 13, 19 and 12 microhaplotypes, respectively, in the recurrent samples of these nine participants. Of these nine participants, *ama1* deep sequencing data were available for six participants and were also classified as new infections. In contrast to *msp1/msp2/glurp* genotyping, AmpSeq identified 21 *ama1* microhaplotypes, slightly more than the other markers. For the six individuals, paired *msp1/msp2/glurp* and AmpSeq data revealed that the mean COI and 95% CI was highest in *msp1* = 2.5 [0.98 – 3.5], *msp2* = 2.5 [0.99 – 3.7], *ama1* = 2.3 [0.81 – 3.3] and lowest in *glurp* = 1.2) during treatment (0h-72h), with the same trend post-treatment (>72h): *msp1* COI = 2.2, *msp2* (COI = 2), *ama1* (COI = 1.8) and *glurp* (COI = 1.2), and no significant difference was observed (p=0.2).

Parasite clearance estimates

Parasite clearance half-lives among all 30 study participants were below 5h except for PID59, with a parasite clearance half-life of 5.7h (Figure 2A). Nonetheless, the mean parasite clearance half-life was 2.8h for all participants. The extrapolation of the clearance rates to each *ama1* microhaplotype (based on the number of reads to quantify each microhaplotype) per infection demonstrated a similar clearance rate across most microhaplotypes. PID16 was excluded from any subsequent analysis since data was available for 6 hours only (Figure 2B). Participant PID32 exhibited rapid clearance of the sole detected *ama1* microhaplotype V9, with a clearance half-life of less than 1 hour. On the other hand, PID06, PID59, and PID65 each presented with at least one microhaplotype with a slower clearance half-life, ranging from 4.5 to 5 hours. Notably, PID59 harboured the V1 *ama1* microhaplotype with the longest clearance half-life in the study, recorded at 7 hours. There was no significant difference in the mean clearance half-lives when the microhaplotypes were grouped as major and minor (<5%) microhaplotypes (Welch two-sample t-test, $p = 0.61$).

Pfmdr1 genetic diversity pre-and post-treatment

Based on the combination of amino acid polymorphisms at codons N86Y and F184Y, only two microhaplotypes (NY and NF) were detected. Of the 30 individuals with baseline data, 22 had mixed infections with both mdr1 microhaplotypes, while 8 had monoclonal *mdr1* infections (either NF or NY) throughout treatment (Figure 3A). Of the individuals with mixed *mdr1* infections, 15/22 had sequence data up to 72h only, and they maintained *mdr1* microhaplotypes at stable frequencies, similar to *ama1*. However, in two individuals (PIDs 45 and 53), there were sporadic detections of rare *mdr1* microhaplotypes (Figure 3B).

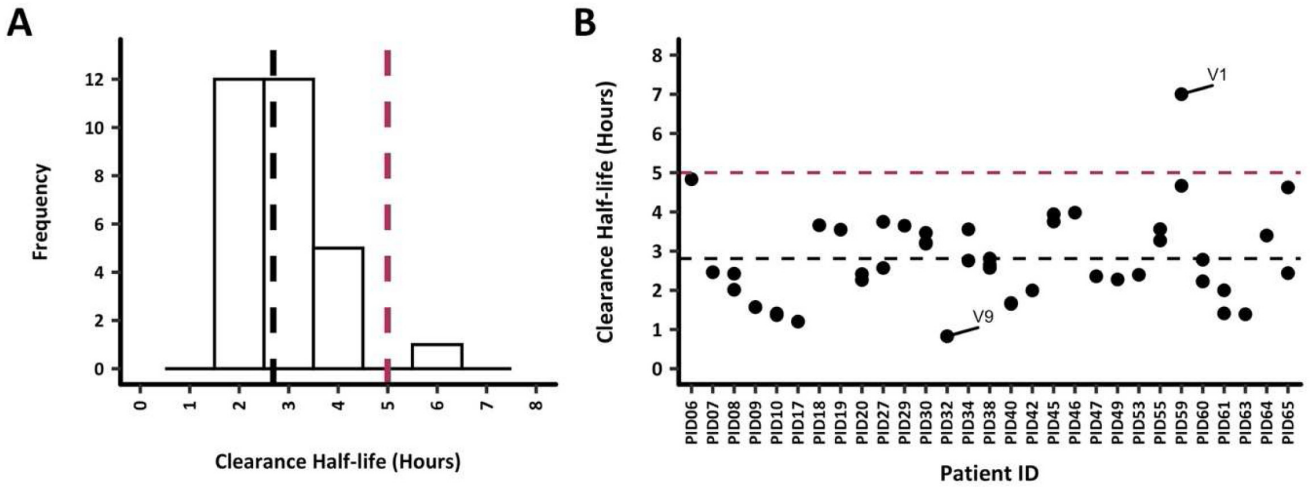


Figure 2. Parasite clearance estimates (PCE) for all study participants. Parasite clearance estimates for the 30 study participants are shown. The dotted lines represent the 5h clearance cut-off (red) and median clearance half-life, 2.6h (black). **A**) PCEs were calculated based on total parasitemia for each participant, the median clearance half-life was 2.7 hours. All participants had clearance half-lives <5h, however, PID59 had a clearance half-life of 5.7h. **B**) PCEs were calculated by extrapolating the clearance rates to each *ama1* microhaplotype (based on the parasitemia and number of reads to quantify each microhaplotype). The *ama1* microhaplotype V1 with the highest clearance half-life of 7h was from PID59. On the other hand, PID32 had the fastest clearing (0.8h) *ama1* microhaplotype V9.

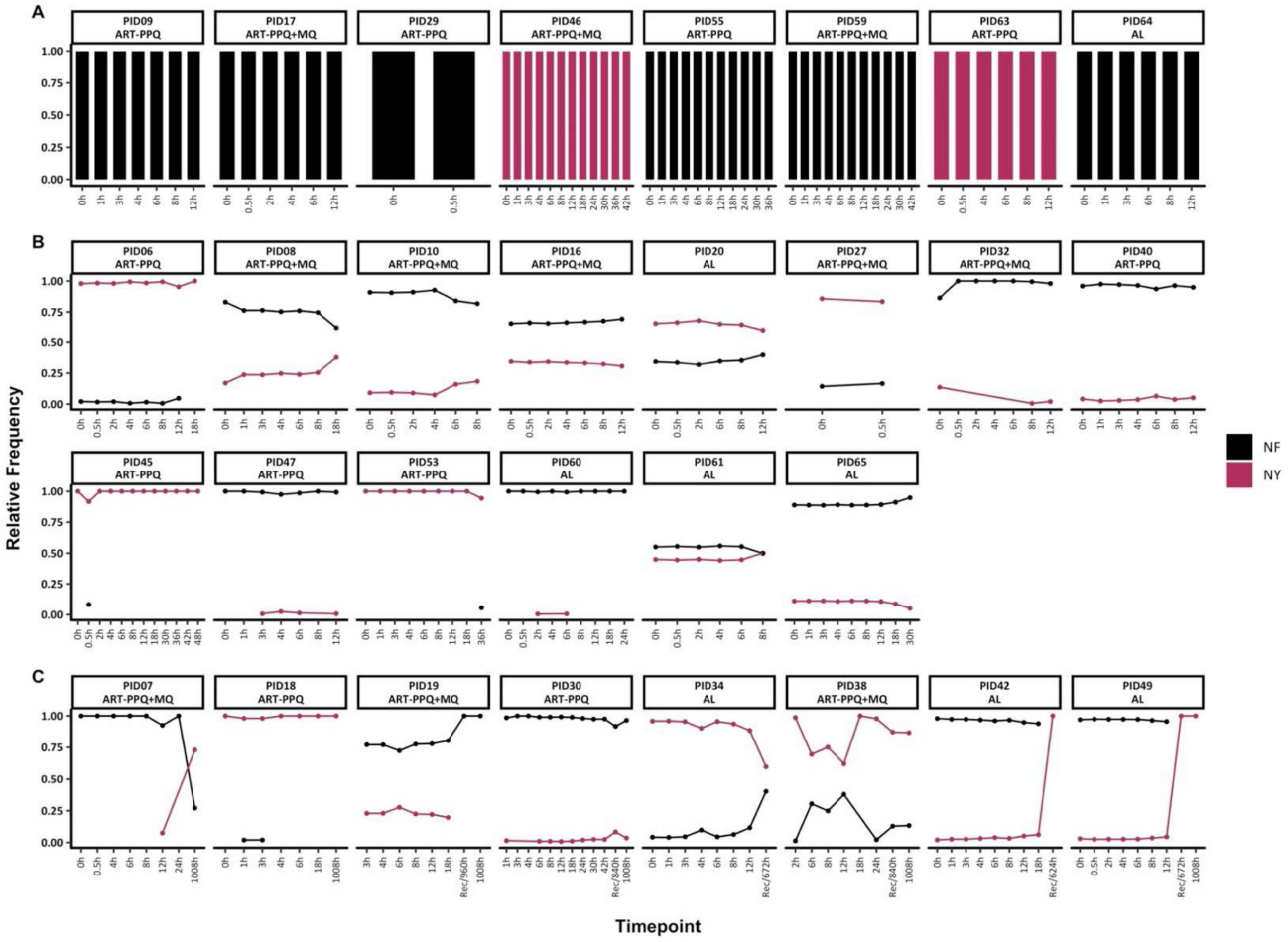


Figure 3. Temporal changes in *mdr1* microhaplotypes throughout treatment. The figure highlights the individuals with **A**) monoclonal infections, if they had only one mdr1 at any timepoint **B**) polyclonal infections, if they had more than one mdr1 microhaplotype throughout treatment and sampled up to 72h and **C**) polyclonal infections, if they more than one mdr1 microhaplotype throughout treatment and sampled beyond 72h. Each coloured barplot/point/line represents a unique mdr1 microhaplotype (black-NF and maroon-NY). Above each plot is the respective participant id. The x-axis represents the sampling timepoints at different hours and the recurrence time with the prefix "Rec" while the y-axis represents the relative proportions of mdr1 microhaplotypes.

In the remaining 8/22 individuals with mixed *mdr1* infections and with post-treatment (>72h) sequence data, three of these individuals switched from a predominance of the NF microhaplotype to the NY microhaplotype during follow-up either during recurrence infection (unscheduled visit) or on day 42 (1008h). Only PID18 maintained the same dominant microhaplotype (NY) throughout the treatment and follow-up period on day 42 (1008h) (Figure 3C). There was no significant difference in the frequencies of Y184 and 184F microhaplotypes in baseline samples vs samples collected from 12h onward (Chi-square test, $p = 0.52$).

Discussion

Children in this moderate-high malaria transmission setting in Kilifi maintained a stable homogeneous microhaplotype population throughout treatment until day 3 (72h). This is not surprising since febrile infections tend to occur with low COI even with high parasite densities (Beck *et al.*, 1997) and treatment reduces the establishment of new infections. Thus, the genetic homogeneity improves the confidence in determining reinfections post-treatment (day 28 onwards), such as the one individual identified with a microhaplotype at 30h and later in their recurrent infection (for an unscheduled visit at 672h). The sporadic observation of rare microhaplotypes at only a single time point potentially highlights the changes in parasite density, sequestration of parasites, lingering genetic material from dead parasites or the presence of parasites below the blood sampling limit (Jones *et al.*, 2021). Notably, PID27 had a COI of three at 0h, and while the relative frequencies of two *ama1* microhaplotypes decreased over time, one microhaplotype continued to increase in frequency. This patient did not experience a recurrence and had no microscopically detectable parasites by 72h. Therefore, the increase in one *ama1* microhaplotype may have originated from the nucleic

material of dead parasites. Most of these distinct changes in *ama1* microhaplotype frequencies occurred post-treatment from day 7 onwards, when reinfections are likely as the levels of drugs in the body continue to decrease. Importantly, our analysis of rare microhaplotypes was limited to a 5% cut-off based on sequencing controls to increase our confidence in calling mixed infections.

The clearance rates were similar within and between individuals irrespective of clonality, suggesting that the three antimalarial treatments were equally effective in clearing microhaplotypes and parasitaemia. Any significant deviations in microhaplotype clearance rates would signal emerging resistance during treatment if a microhaplotype was consistently cleared at a slower rate within and between individuals. This analysis identified one such individual with an estimated slow clearance of 5.7h. A closer examination of the two main microhaplotypes in this participant identified a slow clearing microhaplotype with an estimated clearance half-life of 7h. Thus, a drug-resistant microhaplotype circulating at a low frequency before treatment may survive and rapidly expand following treatment (Ecker *et al.*, 2012; Jafari *et al.*, 2004). In addition to the slow clearing microhaplotype, the sole fast and three slower clearing microhaplotypes indicate the variation in an individual's ability to clear an infection. Such infections should be interrogated to determine additional factors contributing to the slow clearing of parasites.

The *mdr1* genotype in codon 86 was 100% wild type (N86). This is consistent with previous findings (Wamae *et al.*, 2019) in the study area of a shift from 86Y to N86 by 2018 when this study began, and thus only two microhaplotypes were observed based on codon 184. Though the sample size was small, as expected, there was no selection of codon 184 with these drugs. This is in contrast to a study conducted in Tanzania between 2002 and 2004, when there was a higher frequency of *mdr1* mutant genotypes, that observed more N86 and 184F in post-treatment samples following artemether-lumefantrine treatment (Humphreys *et al.*, 2007).

Limitations of this study include the small sample size across the three treatment arms, which may have led to biases. For example, all but one of the recurrent infections were new infections with entirely different microhaplotypes from the pre-treatment sample by *ama1* AmpSeq. Therefore, in more extensive studies, it remains to be seen how sensitive AmpSeq will be in distinguishing new vs recurrent infections compared to msp1/mps2/glurp genotyping. The observations made from a single individual with a slow clearing microhaplotype require further validation as this was based on data from only one individual. Moreover, since this study was set up to provide a proof of concept, it is a scalable assay that allows for additional drug resistance markers, such as *k13*, to be monitored. The low parasitaemia following treatment minimised the generation of good quality AmpSeq data, impacting the sample size. Consequently, there were limited samples to examine the changes in microhaplotypes throughout the study period. However, the findings were similar to several drug trials that predominantly identified new rather than recrudescent infections when efficacious drugs are tested (Adegbite *et al.*, 2019; Davlantes *et al.*, 2018; Kakolwa *et al.*, 2018). Additionally, immunity may play a role in clearing infections; thus, new microhaplotypes are likely present in subsequent infections. Microhaplotypes detected post-treatment might have originated from circulating gametocytes, dormant or dead parasites. However, all participants were gametocyte negative throughout treatment except PID38 who had five gametocytes/µl at 72h but was negative thereafter, so it is unlikely that gametocytaemia biased our findings. As for genetic material originating from dormant or dead parasites, parasite mRNA can also be detected up to two weeks after successful treatment in microscopy-negative individuals (Mahamar *et al.*, 2021). However, additional work is needed to determine whether these originate from viable infections. Finally, future studies should include artificial sequencing controls of decreasing parasitaemia from, e.g., dilutions ranging from 10,000 to 1 parasite/µl to reliably determine the limit of detection, especially in post-treatment with low parasitemia.

Improved accuracy in distinguishing infections post-treatment is essential when considering the WHO recommendation of abandoning a drug if failure rates are >10% (WHO, 2008), determined by the number of recrudescent infections that is likely to vary based on genotyping methods used and genetic markers examined. The additional analyses of tracking variants throughout treatment improve the ability to identify dominant pre-treatment variants that appear as a new infection (in the follow-up period) and are misclassified as a recrudescent infection. *Ama1* yielded comparable COI to *msp1* and *msp2*, hence, the *ama1* locus genotyped in this study is a highly discriminatory marker with a larger number of variants at a prevalence of <5% and with a high expected heterozygosity value (0.96). Coupled with the high sensitivity of amplicon sequencing to capture minor variants, this assay provides a higher resolution to better distinguish reinfections from recrudescent infections in moderate to high transmission settings where polyclonal infections are common.

The assessment of genotypes throughout treatment follows the trajectory of infection, identifies the number of infecting microhaplotypes and rare variants per individual. This allows for the early detection of emerging resistant variants by identifying slow-clearing variants. Furthermore, examining drug resistance mutation frequencies during treatment can identify distinct shifts in occurrences likely to indicate directional selection of a rapidly rising variant (Henriques *et al.*, 2014; Mideo *et al.*, 2013). Given the high disease burden and high levels of reinfection in sub-Saharan Africa, there is a need for improved tools for conducting molecular assays to improve the interpretation of the genotyping outputs. In fact, a recent WHO consultation meeting concluded that AmpSeq provides the most robust and reliable genotyping method (WHO, 2021b). The AmpSeq assay, presented in this study, highlights a potential genotyping tool for PCR correction and monitoring drug resistance markers. The need for more genotyping reference labs in Africa remains to support this endeavour.

While this study serves as a proof of concept, the sample size employed was not aimed at establishing the clear superiority of AmpSeq over msp1/msp2/glurp genotyping for tasks such

as PCR correction and determining infection complexity. Moreover, the adoption of AmpSeq in low and middle-income countries (LMICs) is challenged by its current cost and the requisite molecular biology and bioinformatics expertise needed for effective implementation. Nevertheless, the constantly declining costs of next-generation sequencing, coupled with emerging technologies like the Oxford Nanopore platform that offer reduced buy-in and maintenance expenses, hint at a potentially more accessible AmpSeq in the future.

Regarding its performance, our study illustrates that AmpSeq yields results comparable to those obtained from msp1/msp2/glurp genotyping. However, AmpSeq boasts additional benefits over msp1/mps2/glurp genotyping. It demonstrates scalability advantages and heightened resolution based on sequence identity, surpassing the limitations of msp1/msp2/glurp genotyping, such as the reliance on labour-intensive gel electrophoresis-based band size analysis or fragment analysis that can be subject to interpretation errors and the inability to detect low-density parasite clones. An additional AmpSeq's strength lies in its potential for multiplexing, enabling the simultaneous identification of drug resistance markers alongside genotyping. This multiplexing capability enhances the assay's versatility and could offer valuable insights into the parasite's drug resistance profiles.

In light of these considerations, while the current prerequisites of expertise and cost might impede the immediate integration of AmpSeq in LMICs, the evolving landscape of next-generation sequencing technologies and cost reductions inspire optimism regarding AmpSeq's future accessibility. Therefore, this study recognises the prospective value and adaptability of AmpSeq for tasks like PCR correction and determining infection complexity, even though broader adoption may necessitate ongoing advancements in affordability and the availability of expertise.

Data availability

Underlying data

The raw fastq files have been deposited in Zenodo under: (Fastq Files) Amplicon sequencing of *ama1* and *mdr1* to track within-host P. falciparum diversity in Kilifi, KENYA (Version 1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6243929 (Wamae *et al.,* 2022a)

The nucleotide sequence data reported in this paper are available in the GenBank database under the accession numbers: *ama1* (MZ593448 - MZ593480) and *mdr1* (MZ593481 - MZ593484).

Extended data

Extended data tables and figures have been deposited in Zenodo under: (Extended Data) Amplicon sequencing of ama1 and mdr1 to track within-host P. falciparum diversity throughout treatment in a clinical drug trial (Version 1) [Data set]. DOI: https://doi.org/10.5281/zenodo.10801586 (Wamae *et al.*, 2022b).

This collection contains the following extended data:

- **Table S1: Concentration ratios and resulting parasitemia in artificial DNA mixtures of *P. falciparum* Lab Isolates 3D7 and Dd2.** This table presents the parasitemia for the artificial mixtures of P. falciparum lab isolates 3D7 and Dd2. Each mixture was prepared at varying ratios of 3D7 to Dd2, starting from equal proportions to a complete presence of only 3D7. The original concentration of each isolate was approximately 50,000 parasites per microliter (pf/µl), and the table displays the proportion of each strain in the mixture and the resulting total parasitemia concentration.

- **Table S2. List of PCR and deep sequencing primers**. This table shows the list of forward and reverse primers used for deep sequencing. In boldface are the MID tags, while in the regular face are the forward primers

- **Table S3. The relative frequencies of each ama1 variant and the number of samples with each variant**. The relative frequencies (%) of the 33 AMA1 variants in pre- and post-treatment samples (n = 330) are shown as a 33 amino acid sequence. The frequencies were calculated by dividing the number of reads of each microhaplotype by the total number of reads obtained per sample (116,187,131).

- **Table S4.** Distribution of microhaplotypes among samples. This table shows the occurrence of microhaplotypes across all participants, both with monoclonal and multiclonal ama1 infections. It presents the *ama1* clonality – monoclonal or multiclonal (column 1) - participant IDs (column 2), microhaplotype IDs (column 3), and the relative frequencies of these microhaplotypes across timepoints from 0 to 1008 hours (day 42) (column 3). Dashes represent time points where microhaplotypes were missing or were not detected.

- **Table S5.** Distribution of rare microhaplotypes among samples. This table shows the occurrence of rare microhaplotypes in various samples. It presents participant IDs (column 1), microhaplotype IDs (column 2), and the relative frequencies of these microhaplotypes across time points from 0 to 1008 hours (day 42) (column 3). Samples containing rare microhaplotypes - specifically from PID10, PID32, PID38, PID40, PID49, PID60, PID63, and PID65 - are shown in orange, along with the corresponding rare microhaplotypes and their time points of occurrence. Furthermore, participants are categorised by shared microhaplotypes to indicate instances of rarity and commonality. Except for one microhaplotype unique to PID30, rare microhaplotypes were detected in several samples, frequently exceeding a 5% relative frequency. Dashes represent time points where microhaplotypes were missing or were not detected.

- **Table S6. The parasitemia levels associated with each *ama1* microhaplotype per timepoint**. This table shows the parasitemia for each *ama1* microhaplotype per timepoint and each participant. "Patient ID" represents the patient ID, "AMA1 COI at 0h" represents the complexity

of infection (COI) for each participant at baseline, based on *ama1* while subsequent columns represent the parasitemia for each *ama1* microhaplotype from timepoint 0h to 1008h. Parasitemia was back-calculated using the COI and total parasitemia for each time point. For time points with a COI > 1, parasitemia for the respective *ama1* microhaplotypes are separated by commas, cells in red indicate timepoints without sequencing data (ND = not determined). In contrast, cells in grey indicate time points where microhaplotypes were detected below 10 parasites/µl, hence at risk of falling below the sampling limit.

- **Figure S1. Performance of AmpSeq in the sequencing controls.** Six aliquots were prepared for each control set to ensure sufficient control data in case of PCR or sequencing failure. The median read depth in the lab controls was 5,658 (range 4,310 – 12,603) and 704 (291 – 1,676). The x-axis represents the aliquot identifier across the five mixtures, starting from 1 to 6, while the y-axis represents the proportions of each variant across all aliquots. For *ama1* (A), two variants (3D7 and Dd2) were detected, whereas in *mdr1* (B), two variants were detected YY, FY and NY following amplification of Dd2 Copy I, Dd2 Copy II and 3D7, respectively. For *ama1*, sequencing failed for aliquot 6 of control set 1, while for *mdr1*, sequencing failed for aliquot 2 and 6 of control set 3, aliquots 1 and 6 of control set 4 and aliquots 1 and 5 of control set 5. Under the *mdr1* control set 4, the Dd2 copy II (86F, 184Y) was not identified, possibly due to having very low concentrations that were not picked up in this aliquot. Based on our control mixtures, the minimum variant frequency we could detect was 5%.

- **Figure S2. Heatmaps of the successfully PCR amplified and sequenced samples for *ama1* (A) and *mdr1* (B).** The rows represent the study participants, while the columns represent time in hours. Successfully sequenced samples are shown in blue, those that failed PCR are shown in red and those that failed sequencing are in black. The timepoint "*Rec*" represents unscheduled visits where a recurrent sample was collected. The unshaded areas with "-" are time points where samples were not collected. For each time point, the number of samples successfully sequenced (n Successful) is indicated in the last row of each panel. The table in panel C shows the groupings of samples based on parasitemia, high (> 5,000), moderate (100-5,000) and low (< 100 parasites per microlitre). Many samples collected between 0h-12h had high parasitemia, samples collected between 18h–30h had moderate parasitemia, while samples collected after 30h were primarily of low parasitemia.

- **Figure S3. The mean complexity of infection (COI) by AMA1 throughout treatment.** The mean COI (red diamonds) appeared to be stable (between 1.5 - 2) from baseline (0h) up to 72h and thereafter fluctuated due to the small sample sizes (<5) in the post-treatment samples. The black dots represent the COI per sample.

Acknowledgements

We thank the staff of the Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand, for their contribution to the PCR correction of recurrent infections. Additionally, the following reagents were obtained through BEI Resources, NIAID, NIH: Genomic DNA from *Plasmodium falciparum*, Strain 3D7, MRA-102G, contributed by Daniel J. Carucci and genomic DNA from *Plasmodium falciparum*, Strain Dd2, MRA-150G, contributed by David Walliker.

References

Adegbite BR, Edoa JR, Honkpehedji YJ, *et al.*: **Monitoring of efficacy, tolerability and safety of artemether-lumefantrine and artesunate-amodiaquine for the treatment of uncomplicated *Plasmodium falciparum* malaria in Lambaréné, Gabon: an open-label clinical trial.** *Malar J.* 2019; **18**(1): 424.
PubMed Abstract | Publisher Full Text | Free Full Text

Ashley EA, Dhorda M, Fairhurst RM, *et al.*: **Spread of Artemisinin Resistance in *Plasmodium falciparum* Malaria.** *N Engl J Med.* 2014; **371**(5): 411–423.
PubMed Abstract | Publisher Full Text | Free Full Text

Balikagala B, Fukuda N, Ikeda M, *et al.*: **Evidence of Artemisinin-Resistant Malaria in Africa.** *N Engl J Med.* 2021; **385**(13): 1163–1171.
PubMed Abstract | Publisher Full Text

Beck HP, Felger I, Huber W, *et al.*: **Analysis of Multiple *Plasmodium falciparum* Infections in Tanzanian Children during the Phase III Trial of the Malaria Vaccine SPf66.** *J Infect Dis.* 1997; **175**(4): 921–926.
PubMed Abstract | Publisher Full Text

Bhatt S, Weiss DJ, Cameron E, *et al.*: **The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015.** *Nature.* 2015; **526**(7572): 207–211.
PubMed Abstract | Publisher Full Text | Free Full Text

Davlantes E, Dimbu PR, Ferreira CM, *et al.*: **Efficacy and safety of artemether-lumefantrine, artesunate-amodiaquine, and dihydroartemisinin-piperaquine for the treatment of uncomplicated *Plasmodium falciparum* malaria in three provinces in Angola, 2017.** *Malar J.* 2018; **17**(1): 144.
PubMed Abstract | Publisher Full Text | Free Full Text

Dondorp AM, Nosten F, Yi P, *et al.*: **Artemisinin Resistance in *Plasmodium falciparum* Malaria.** *N Engl J Med.* 2009; **361**(5): 455–467.
PubMed Abstract | Publisher Full Text | Free Full Text

Ecker A, Lehane AM, Clain J, *et al.*: **PfCRT and its role in antimalarial drug resistance.** *Trends Parasitol.* 2012; **28**(11): 504–514.
PubMed Abstract | Publisher Full Text | Free Full Text

Felger I, Snounou G, Hastings I, *et al.*: **PCR correction strategies for malaria drug trials: updates and clarifications.** *Lancet Infect Dis.* 2020; **20**(1): e20–e25.
PubMed Abstract | Publisher Full Text

Flegg JA, Guerin PJ, White NJ, *et al.*: **Standardizing the measurement of parasite clearance in *falciparum* malaria: the parasite clearance estimator.** *Malar J.* 2011; **10**: 339.
PubMed Abstract | Publisher Full Text | Free Full Text

Gruenberg M, Lerch A, Beck HP, *et al.*: **Amplicon deep sequencing improves**

Plasmodium falciparum **genotyping in clinical trials of antimalarial drugs.** *Sci Rep.* 2019; **9**(1): 17790.
PubMed Abstract | Publisher Full Text | Free Full Text

Hamaluba M, van der Pluijm RW, Weya J, *et al.*: **Arterolane-piperaquine-mefloquine versus arterolane-piperaquine and artemether-lumefantrine in the treatment of uncomplicated *Plasmodium falciparum* malaria in Kenyan children: a single-centre, open-label, randomised, non-inferiority trial.** *Lancet Infect Dis.* 2021; **21**(10): 1395–1406.
PubMed Abstract | Publisher Full Text | Free Full Text

Hathaway NJ, Parobek CM, Juliano JJ, *et al.*: **SeekDeep: single-base resolution *de novo* clustering for amplicon deep sequencing.** *Nucleic Acids Res.* 2018; **46**(4): e21.
PubMed Abstract | Publisher Full Text | Free Full Text

Henriques G, Hallett RL, Beshir KB, *et al.*: **Directional selection at the *pfmdr1*, *pfcrt*, *pfubp1*, and *pfap2mu* loci of *Plasmodium falciparum* in Kenyan children treated with ACT.** *J Infect Dis.* 2014; **210**(12): 2001–2008.
PubMed Abstract | Publisher Full Text | Free Full Text

Humphreys GS, Merinopoulos I, Ahmed J, *et al.*: **Amodiaquine and Artemether-Lumefantrine Select Distinct Alleles of the *Plasmodium falciparum mdr1* Gene in Tanzanian Children Treated for Uncomplicated Malaria.** *Antimicrob Agents Chemother.* 2007; **51**(3): 991–997.
PubMed Abstract | Publisher Full Text | Free Full Text

Jafari S, Le Bras J, Bouchaud O, *et al.*: ***Plasmodium falciparum*** **clonal population dynamics during malaria treatment.** *J Infect Dis.* 2004; **189**(2): 195–203.
PubMed Abstract | Publisher Full Text

Jones S, Kay K, Hodel EM, *et al.*: **Should Deep-Sequenced Amplicons Become the New Gold Standard for Analyzing Malaria Drug Clinical Trials?** *Antimicrob Agents Chemother.* 2021; **65**(10): e0043721.
PubMed Abstract | Publisher Full Text | Free Full Text

Kakolwa MA, Mahende MK, Ishengoma DS, *et al.*: **Efficacy and safety of artemisinin-based combination therapy, and molecular markers for artemisinin and piperaquine resistance in Mainland Tanzania.** *Malar J.* 2018; **17**(1): 369.
PubMed Abstract | Publisher Full Text | Free Full Text

Kassambara A: **ggpubr: 'ggplot2' Based Publication Ready Plots.** 2020.
Reference Source

Liljander A, Wiklund L, Falk N, *et al.*: **Optimization and validation of multi-coloured capillary electrophoresis for genotyping of *Plasmodium falciparum* merozoite surface proteins (*msp1* and *2*).** *Malar J.* 2009; **8**: 78.
PubMed Abstract | Publisher Full Text | Free Full Text

Mahamar A, Lanke K, Graumans W, *et al.*: **Persistence of mRNA indicative of *Plasmodium falciparum* ring-stage parasites 42 days after artemisinin and non-artemisinin combination therapy in naturally infected Malians.** *Malar J.* 2021; **20**(1): 34.
PubMed Abstract | Publisher Full Text | Free Full Text

Mideo N, Bailey JA, Hathaway NJ, *et al.*: **A deep sequencing tool for partitioning clearance rates following antimalarial treatment in polyclonal infections.** *Evol Med Public Health.* 2016; **2016**(1): 21–36.
PubMed Abstract | Publisher Full Text | Free Full Text

Mideo N, Kennedy DA, Carlton JM, *et al.*: **Ahead of the curve: next generation estimators of drug resistance in malaria infections.** *Trends Parasitol.* 2013; **29**(7): 321–8.
PubMed Abstract | Publisher Full Text | Free Full Text

Miller RH, Hathaway NJ, Kharabora O, *et al.*: **A deep sequencing approach to estimate *Plasmodium falciparum* complexity of infection (COI) and explore apical membrane antigen 1 diversity.** *Malar J.* 2017; **16**(1): 490.
PubMed Abstract | Publisher Full Text | Free Full Text

Nei M: **Estimation of average heterozygosity and genetic distance from a small number of individuals.** *Genetics.* 1978; **89**(3): 583–590.
PubMed Abstract | Publisher Full Text | Free Full Text

Noedl H, Se Y, Schaecher K, *et al.*: **Evidence of artemisinin-resistant malaria in western Cambodia.** *N Engl J Med.* 2008; **359**(24): 2619–20.
PubMed Abstract | Publisher Full Text

Okell LC, Reiter LM, Ebbe LS, *et al.*: **Emerging implications of policies on malaria treatment: genetic changes in the *Pfmdr-1* gene affecting susceptibility to artemether–lumefantrine and artesunate–amodiaquine in Africa.** *BMJ Glob Health.* 2018; **3**(5): e000999.
PubMed Abstract | Publisher Full Text | Free Full Text

Parobek CM, Parr JB, Brazeau NF, *et al.*: **Partner-drug resistance and population substructuring of artemisinin-resistant *plasmodium falciparum* in Cambodia.** *Genome Biol Evol.* 2017; **9**(6): 1673–1686.
PubMed Abstract | Publisher Full Text | Free Full Text

Polley SD, Conway DJ: **Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene.** *Genetics.* 2001; **158**(4): 1505–12.
PubMed Abstract | Publisher Full Text | Free Full Text

Snounou G, Beck HP: **The use of PCR genotyping in the assessment of recrudescence or reinfection after antimalarial drug treatment.** *Parasitol Today.* 1998; **14**(11): 462–467.
PubMed Abstract | Publisher Full Text

Uwimana A, Umulisa N, Venkatesan M, *et al.*: **Association of *Plasmodium falciparum kelch13* R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study.** *Lancet Infect Dis.* 2021; **21**(8): 1120–1128.
PubMed Abstract | Publisher Full Text

van der Pluijm RW, Amaratunga C, Dhorda M, *et al.*: **Triple Artemisinin-Based Combination Therapies for Malaria - A New Paradigm?** *Trends Parasitol.* 2021; **37**(1): 15–24.
PubMed Abstract | Publisher Full Text

van der Pluijm RW, Imwong M, Chau NH, *et al.*: **Determinants of dihydroartemisinin-piperaquine treatment failure in *Plasmodium falciparum* malaria in Cambodia, Thailand, and Vietnam: a prospective clinical, pharmacological, and genetic study.** *Lancet Infect Dis.* 2019; **19**(9): 952–961.
PubMed Abstract | Publisher Full Text | Free Full Text

van der Pluijm RW, Tripura R, Hoglund RM, *et al.*: **Triple artemisinin-based combination therapies versus artemisinin-based combination therapies for uncomplicated *Plasmodium falciparum* malaria: a multicentre, open-label, randomised clinical trial.** *Lancet.* 2020; **395**(10233): 1345–1360.
PubMed Abstract | Publisher Full Text | Free Full Text

Veiga MI, Dhingra SK, Henrich PP, *et al.*: **Globally prevalent PfMDR1 mutations modulate *Plasmodium falciparum* susceptibility to artemisinin-based combination therapies.** *Nat Commun.* 2016; **7**: 11553.
PubMed Abstract | Publisher Full Text | Free Full Text

Wamae K, Ndwiga L, Kharabora O, *et al.*: **(Fastq Files) Amplicon sequencing of ama1 and mdr1 to track within-host P. falciparum diversity in Kilifi, KENYA (Version 1).** [Data set]. Zenodo. 2022a.
http://www.doi.org/10.5281/zenodo.6243929

Wamae K, Ndwiga L, Kharabora O, *et al.*: **(Extended Data) Amplicon deep sequencing of ama1 and mdr1 to track within-host P. falciparum diversity throughout treatment in a clinical drug trial (Version 1).** [Data set]. Zenodo. 2022b.
http://www.doi.org/10.5281/zenodo.6253570

Wamae K, Okanda D, Ndwiga L, *et al.*: **No evidence of *P. falciparum* K13 artemisinin conferring mutations over a 24-year analysis in Coastal Kenya, but a near complete reversion to chloroquine wild type parasites.** *Antimicrob Agents Chemother.* 2019; **63**(12): e01067–19.
PubMed Abstract | Publisher Full Text | Free Full Text

WHO: **Assessment and monitoring of antimalarial drug efficacy for the treatment of uncomplicated falciparum malaria.** World Health Organization (Ref.No# WHO/HTM/RBM/2003.50), 2003.
Reference Source

WHO: **Informal consultation on methodology to distinguish reinfection from recrudescence in high malaria transmission areas**. 2021b.
Reference Source

WHO: **Methods and techniques for clinical trials on antimalarial drug efficacy: Genotyping to identify parasite populations.** World Health Organization, 2008.
Reference Source

WHO: **World Malaria Report 2021**. 2021a.
Reference Source

Wickham H: **ggplot2: Elegant Graphics for Data Analysis**. 2016.
Reference Source

Open Peer Review

Current Peer Review Status: ? ✓ ✓ ?

Version 2

Reviewer Report 25 July 2023

https://doi.org/10.21956/wellcomeopenres.20663.r61598

?

Aurel Holzschuh
[1] University of Notre Dame, Notre Dame, Indiana, USA
[2] University of Notre Dame, Notre Dame, Indiana, USA
[3] University of Notre Dame, Notre Dame, Indiana, USA

In this study, the researchers investigated the potential of AmpSeq for monitoring the within-host diversity of parasites before and after ACT treatment. They examined the presence or absence of *ama1* and *mdr1* microhaplotypes in blood samples collected at regular intervals from patients treated with three different regimens. This study shows the potential of AmpSeq as a tool for tracking parasite microhaplotypes throughout treatment. However, I think that there is some clarification needed regarding the methods and results. I have some comments for the authors to address:

Methods:
1. DNA preparation and PCR from sequencing controls and clinical samples
The authors noted that they could detect the minority allele in a mixture of 95%:5% and regarded any reads less than 5% of the total as noise. However, I could not find any information regarding the parasite densities of the artificial control mixtures. I assume the different ratios were all at the same total parasite density? Reliable detection of a 5% minority clone at e.g., 1,000 parasites/uL (i.e., 50 parasites/uL at 5%) does not guarantee detection of a 5% minority clone at 100 or 10 parasites/uL. It seems like the authors have chosen to include all samples with a "blood sampling limit" of ≥10 parasites/uL, do you have any control that can confirm the detection of 5% minority clones at this parasite density? If you look at your control mixtures, what are the number of reads for those? In control set 4, Rep_4, marker mdr1, did you obtain fewer reads for this sample and thus missed the YY microhaplotype (maybe it was only below the 5% threshold)?

2. DNA preparation and PCR from sequencing controls and clinical samples
Regarding the sequencing controls, you mention the following: "Amplicons spanning *ama1* (PF3D7_1133400, nucleotides 441–946) and *mdr1* (PF3D7_0523000, nucleotides 183–719) were generated in duplicate from each control and sample,...". However, Figure S1 indicates 6 replicates for each control set. Under sequence data analysis you again mention: "Microhaplotypes were

discarded if they did not occur in the two PCR replicates and if their combined relative frequency was <5%." Please specify.

3. Sequence data analysis
Was there a minimum number of reads needed to call a microhaplotype, e.g., 10 reads? The lower the total reads for a sample, the more I expect that whether or not certain microhaplotypes were detected was a matter of chance, especially for low-abundant minority clones. As an example, replicate 1 has 1,000 total reads and two clones, one with 960 reads and the second with 40 reads. You would discard the minor clone because of the >5% threshold. Replicate 2 has 500 total reads and also two clones. One with 460 reads and the second with 40 reads. You would consider this a polyclonal infection because the minor clone is above 5%. Could you give a brief explanation if only the 5% threshold was considered to call a microhaplotype or if the number of reads was also considered?

Results:
4. AmpSeq in artificial mixtures of sequencing controls
The authors state: "This provided evidence that the assay could detect mixed infections in clinical samples, but only when the minor microhaplotype was at a relative frequency of ≥5%." This is very much density-dependent. What was the parasite density of the controls? Please add the density of the controls, as evidence is only there for the parasite densities used (which was?). With decreasing parasitemia, the likelihood of detecting minority clones at 5% also declines (becomes a chance event at one point).

Figure S1. Control set 2 and 3 are either mixed up, or the observed proportions do not really match the expected proportions. In the results section the authors state "...the 3D7 and Dd2 *ama1* microhaplotypes were detected consistently across all mixtures and in the expected proportions." Looking at control set 2, the minority clone seems to be around 25%, and not at the expected 15%. Similarly, for control set 3, the minority clone is around 15% whereas it would be expected at around 15%. Please check.

Figure S1. In the legend, the authors state "Based on our control mixtures, the minimum variant frequency we were able to detect was 0.5%.". I think this should be 5% and not 0.5%.

5. AmpSeq of pre-and post-treatment samples
Figure S2. The legend states "Successfully sequenced samples are shown in green, those that failed PCR in red, and those that failed sequencing are in yellow." I do not see any green or yellow color in the heatmap, I see blue, red, and black. Please correct. Similarly, the authors state "For each time-point, the number of samples successfully sequenced, *n*, is indicated in the last row of each panel". I do not see any *n* in either of the panels.

6. Pfama1 genetic diversity during and after treatment
Figure 1. Could the authors maybe provide the parasite densities for each sample at each time point instead of only stating high and low? This could explain some of the rare haplotypes popping up potentially. Where rare haplotypes occurred, did the authors see if the same rare haplotype was detected at other time points but possibly below the 5% threshold (check if reads were present)? Was there a difference in the number of reads between the monoclonal samples and polyclonal samples?
Maybe it would be good if some of those sporadic microhaplotypes were present but below either

the 5% BIC or the 10 parasites/µl blood sampling limit (could also be provided as supplementary information).

Table S2. All 33 ama1 variants reported seem to be nonsynonymous. Is it true that the authors only observed nonsynonymous mutations in the *ama1* microhaplotype leading to 33 different microhaplotypes at the AA level? What about SNPs that are synonymous? In the Methods section under "Sequence data analysis" you state the following: "COI was defined as the number of distinct *ama1* microhaplotypes (varying at the nucleotide level) in each sample".

Except for PID60, all these sporadic microhaplotypes were detected above the sampling limit of 10 parasites/µl (Table S3). So why is this sporadic microhaplotype shown for PID60 if it fell below the 10 parasites/µl? In the methods section you specify this threshold, thus should be removed. Were there other sporadic microhaplotypes that fell below the 10 parasites/ul threshold but were excluded?

7. Parasite clearance estimates
"Participant PID32 had one ama1 microhaplotype, V9, that was cleared quite rapidly at <1h…". Looking at Figure 1B, I do not see any microhaplotype that was cleared <1h. I only see a microhaplotype (light blue) that persists throughout and a sporadic microhaplotype at 8h). Am I missing something here? Without any parasite densities for the samples/time points, it's impossible to tell based on the proportions of Figure 1 (as we do not see any drop in parasitemia if only the proportions of each microhaplotype are shown). Again, consider adding this information somewhere.

8. AmpSeq compared to msp1/msp2/glurp genotyping of recurrent infections
Might be helpful to also calculate expected heterozygosity for *msp1/msp2/glurp* since you provide this metric for *ama1* and mention in the discussion "the *ama1* locus genotyped in this study is a highly discriminatory marker".

9. Pfmdr1 genetic diversity pre-and post-treatment
Interestingly, some of the COI based on ama1 and mdr1 are not concordant. PID06, PID07, PID18, PID42, PID47, and PID53 are all monoclonal based on *ama1* but polyclonal based on *mdr1*. I would expect the other way around, where several different ama1 microhaplotypes likely share the same mdr1 microhaplotype. For most of them, the mdr1 minor microhaplotype is present at very low proportions. PID42 is somewhat concordant (when the 2nd clone takes over at 642h) as you detect the second clone but miss the first clone (possibly just below the 5% threshold?). Any explanation for this? Maybe take a closer look at the blood sampling limit or the 5% BIC? *ama1* generally seemed to have worked better (e.g., generated a higher number of reads and higher sequencing success).

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular diagnostics and genomic epidemiology of malaria.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 10 Mar 2024

Kevin Wamae

In this study, the researchers investigated the potential of AmpSeq for monitoring the within-host diversity of parasites before and after ACT treatment. They examined the presence or absence of *ama1* and *mdr1* microhaplotypes in blood samples collected at regular intervals from patients treated with three different regimens. This study shows the potential of AmpSeq as a tool for tracking parasite microhaplotypes throughout treatment. However, I think that there is some clarification needed regarding the methods and results. I have some comments for the authors to address:

Methods: ***1. DNA preparation and PCR from sequencing controls and clinical samples***

The authors noted that they could detect the minority allele in a mixture of 95%:5% and regarded any reads less than 5% of the total as noise. However, I could not find any information regarding the parasite densities of the artificial control mixtures. I assume the different ratios were all at the same total parasite density? Reliable detection of a 5% minority clone at e.g., 1,000 parasites/uL (i.e., 50 parasites/uL at 5%) does not guarantee detection of a 5% minority clone at 100 or 10 parasites/uL. It seems like the authors have chosen to include all samples with a "blood sampling limit" of ≥10 parasites/uL, do you have any control that can confirm the detection of 5% minority clones at this parasite density? If you look at your control mixtures, what are the number of reads for those? In control set 4, Rep_4, marker mdr1, did you obtain fewer reads for this sample and thus missed the YY microhaplotype (maybe it was only below the 5% threshold)? *The level of parasitemia for the P. falciparum cultures from which 3D7 and Dd2 DNA were extracted is typically above 1% (* *www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-102.aspx* *and* *www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-150.aspx*). *This translates to approximately 50,000 parasites/µl. Hence, the table below (included as supplementary table 1) illustrates the varying levels of parasitemia in the different artificial ratios we generated and indicates that reliable detection of a 5% minority variant was achieved at a parasitemia level of*

2,500 parasites/µl. However, we did not include controls for our blood sampling limit (≥10 parasites/*µl), and this was acknowledged as a limitation in our discussion section, as indicated below: "…Finally, future studies should include artificial sequencing controls of decreasing parasitaemia from, e.g., dilutions ranging from 10,000 to 1 parasite/µl to reliably determine the limit of detection, especially in post-treatment samples with low parasitemia…"* **ratio 3D7 Dd2** 1:1 25,000 pf/µl 25,000 pf/µl 0.75:0.25 37,500 pf/µl 12,500 pf/µl 0.85:0.15 42,500 pf/µl 7,500 pf/µl 0.95:0.05 47,500 pf/µl 2,500 pf/µl 1:0 50,000 pf/µl 0 pf/µl

- *The median read depth in the lab controls was 5,658 (range 4,310 – 12,603) and 658 (range 291 - 1,676) for ama1 and mdr1, respectively. In contrast, the read depth for samples was 11,147 reads (range 580 –33,714) for ama1 and 11,548 reads (range 1,022 – 55,664) for mdr1. The read depths for control set 4, aliquot 2, 3, 4 and 5 (the term replicate has since been replaced with aliquot for clarity) were 574, 1009, 310 and 565. Therefore, the low read depth for aliquot-4 might have contributed to the lack of detection of the YY microhaplotype.*

2. DNA preparation and PCR from sequencing controls and clinical samples

Regarding the sequencing controls, you mention the following: "Amplicons spanning *ama1* (PF3D7_1133400, nucleotides 441–946) and *mdr1* (PF3D7_0523000, nucleotides 183–719) were generated in duplicate from each control and sample,…". However, Figure S1 indicates 6 replicates for each control set. Under sequence data analysis you again mention: "Microhaplotypes were discarded if they did not occur in the two PCR replicates and if their combined relative frequency was <5%." Please specify.

- *The term 'aliquot' has since been used to designate the multiple portions taken from the control samples. Similarly, the term 'duplicate' refers to the repeated measurements or observations taken from these control aliquots and samples to ensure consistency and reliability in our results. These will provide clarity and accuracy in describing our methodology. Additionally, the methods section has been revised to read:*

"…Amplicons spanning ama1 (PF3D7_1133400, nucleotides 441–946) and mdr1 (PF3D7_0523000, nucleotides 183–719) were generated in duplicate from each control and sample. This process was applied to each of the six aliquots prepared for every control, and to a single aliquot from each sample using primers designed in this study…"

3. Sequence data analysis

Was there a minimum number of reads needed to call a microhaplotype, e.g., 10 reads? The lower the total reads for a sample, the more I expect that whether or not certain microhaplotypes were detected was a matter of chance, especially for low-abundant minority clones. As an example, replicate 1 has 1,000 total reads and two clones, one with 960 reads and the second with 40 reads. You would discard the minor clone because of the >5% threshold. Replicate 2 has 500 total reads and also two clones. One with 460 reads and the second with 40 reads. You would consider this a polyclonal infection because the minor clone is above 5%. Could you give a brief explanation if only the 5% threshold were considered to call a microhaplotype or if the number of reads was also considered?

- *The first paragraph under "Sequence data analysis" in the methods section has been rewritten to include the read-depth cut-off as indicated in the text below:*

"…We implemented SeekDeep's default threshold of 250 reads as the minimum required read-depth for each individual PCR replicate. Additionally, for a sample to be included in the analysis, it needed to have a combined total of at least 500 reads, summing the read counts from all its replicates. For samples that met this criterion, microhaplotypes were discarded if they did not occur in the two PCR duplicates and if their relative frequency was <5% (or less than 25 reads)…"

- *Concerning the example provided above, since the minority variant was not detected above 5% frequency in both sample replicates, it would be discarded*

Results:

4. AmpSeq in artificial mixtures of sequencing controls

The authors state: "This provided evidence that the assay could detect mixed infections in clinical samples, but only when the minor microhaplotype was at a relative frequency of ≥5%." This is very much density-dependent. What was the parasite density of the controls? Please add the density of the controls, as evidence is only there for the parasite densities used (which was?). With decreasing parasitemia, the likelihood of detecting minority clones at 5% also declines (becomes a chance event at one point).

- *The methods section under "DNA preparation and PCR from sequencing controls and clinical samples" has been revised to include the parasite densities (also supplementary table 1, refer to comment 1 above):*

"...The level of parasitemia for the P. falciparum cultures from which 3D7 and Dd2 DNA were extracted is typically above 1% (www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-102.aspx and www.beiresources.org/Catalog/BEIParasiticProtozoa/MRA-150.aspx). These levels correspond to an approximate density of 50,000 parasites/µl, while the specific parasitemia for each isolate in the final mixture is presented in Table S1..." Figure S1. Control set 2 and 3 are either mixed up, or the observed proportions do not really match the expected proportions. In the results section the authors state "...the 3D7 and Dd2 *ama1* microhaplotypes were detected consistently across all mixtures and in the expected proportions." Looking at control set 2, the minority clone seems to be around 25%, and not at the expected 15%. Similarly, for control set 3, the minority clone is around 15% whereas it would be expected at around 15%. Please check.

- *There was a mix-up in the labelling of the different facets for Controls Set 2 and 3 in Figure S1. These have now been amended to reflect the correct labels.*

Figure S1. In the legend, the authors state "Based on our control mixtures, the minimum variant frequency we were able to detect was 0.5%.". I think this should be 5% and not 0.5%.

- *This section has been revised to read 5% and not 0.5%*

5. AmpSeq of pre-and post-treatment samples

Figure S2. The legend states "Successfully sequenced samples are shown in green, those that failed PCR in red, and those that failed sequencing are in yellow." I do not see any green or yellow color in the heatmap, I see blue, red, and black. Please correct. Similarly, the authors state "For each time-point, the number of samples successfully sequenced, *n*, is indicated in the last row of each panel". I do not see any *n* in either of the panels.

- *The problem with colours has been corrected to match the respective colours as indicated below:*

"...Successfully sequenced samples are shown in blue, those that failed PCR are shown in red, and those that failed sequencing are in black..."

- *Additionally, Figure S2 has been revised to include an additional row at the bottom of panels A and B, indicating the number of samples successfully sequenced (n) per time point.*

6. Pfama1 genetic diversity during and after treatment

Figure 1. Could the authors maybe provide the parasite densities for each sample at each time point instead of only stating high and low? This could explain some of the rare haplotypes popping up potentially. Where rare haplotypes occurred, did the authors see if

the same rare haplotype was detected at other time points but possibly below the 5% threshold (check if reads were present)? Was there a difference in the number of reads between the monoclonal samples and polyclonal samples?
Maybe it would be good if some of those sporadic microhaplotypes were present but below either the 5% BIC or the 10 parasites/µl blood sampling limit (could also be provided as supplementary information).

- *Figure 1 has been revised to include a line plot of the log10 parasitaemia/µl per timepoint as a secondary y-axis to the right.*
- *Additionally, we have included Table S5 in the supplementary materials to illustrate the distribution of rare microhaplotypes. This table provides a detailed view of the rare microhaplotypes' occurrences, highlighting their rarity and prevalence within the samples analysed. Notably, it reveals instances where rare haplotypes are present as dominant (>5% frequency) microhaplotypes in other samples.*
- *Our findings revealed no significant discrepancies when examining the variance in read-depth between monoclonal and multiclonal infections. These results have been included in the third paragraph of the section titled 'Pfama1 genetic diversity during and after treatment':*

"…The analysis revealed no significant difference in read counts between samples harbouring either monoclonal or polyclonal infections (p = 0.092, Wilcoxon signed-rank test)…" Table S2. All 33 ama1 variants reported seem to be nonsynonymous. Is it true that the authors only observed nonsynonymous mutations in the *ama1* microhaplotype leading to 33 different microhaplotypes at the AA level? What about SNPs that are synonymous? In the Methods section under "Sequence data analysis" you state the following: "COI was defined as the number of distinct *ama1* microhaplotypes (varying at the nucleotide level) in each sample".

- *Indeed, it may appear unusual, but our observations confirmed that all 33 variants identified exhibited variation at both the nucleotide and amino acid levels, resulting in nonsynonymous changes in the ama1 microhaplotype.*

Except for PID60, all these sporadic microhaplotypes were detected above the sampling limit of 10 parasites/µl (Table S3). So why is this sporadic microhaplotype shown for PID60 if it fell below the 10 parasites/µl? In the methods section you specify this threshold, thus should be removed. Were there other sporadic microhaplotypes that fell below the 10 parasites/ul threshold but were excluded?

- *Table S3 has been updated to Table S6: The detection threshold for rare haplotypes was set at a level of 5%, as outlined in the methods section. This threshold was based on the lowest proportion observed in the control mixtures during sequencing. Despite PID60's microhaplotype falling below the standard sampling limit of 10 parasites/µl, it was included in our analysis because it was concurrently identified in other participants (PID30, PID34, and PID61) at higher frequencies. For example, in PID61, this microhaplotype surpassed the 5% relative frequency mark, justifying its inclusion in our analysis.*

7. Parasite clearance estimates
"Participant PID32 had one ama1 microhaplotype, V9, that was cleared quite rapidly at <1h…". Looking at Figure 1B, I do not see any microhaplotype that was cleared <1h. I only see a microhaplotype (light blue) that persists throughout and a sporadic microhaplotype at 8h). Am I missing something here? Without any parasite densities for the samples/time points, it's impossible to tell based on the proportions of Figure 1 (as we do not see any drop in parasitemia if only the proportions of each microhaplotype are shown). Again, consider adding this information somewhere.

- *This section has been revised to clarify that "clearance" refers to clearance half-life, as calculated by the Worldwide Antimalarial Resistance Network's (WWARN) parasite clearance estimator. It now reads:*

"…Participant PID32 exhibited rapid clearance of the sole detected ama1 microhaplotype V9, with a clearance half-life of less than 1 hour. On the other hand, PID06, PID59, and PID65 each presented with at least one microhaplotype with a slower clearance half-life, ranging from 4.5 to 5 hours. Notably, PID59 harboured the V1 ama1 microhaplotype with the longest clearance half-life in the study, recorded at 7 hours…"

- *We have also included the temporal changes in the relative frequencies of all microhaplotypes across all samples in Table S4.*

8. AmpSeq compared to msp1/msp2/glurp genotyping of recurrent infections

Might be helpful to also calculate expected heterozygosity for *msp1/msp2/glurp* since you provide this metric for *ama1* and mention in the discussion "the *ama1* locus genotyped in this study is a highly discriminatory marker".

- *We acknowledge the value this would add, especially given the high discriminative power of the ama1 locus highlighted in our study. However, we did not have access to the requisite data from the parent study to perform these calculations during our study.*

9. Pfmdr1 genetic diversity pre-and post-treatment

Interestingly, some of the COI based on ama1 and mdr1 are not concordant. PID06, PID07, PID18, PID42, PID47, and PID53 are all monoclonal based on *ama1* but polyclonal based on *mdr1*. I would expect the other way around, where several different ama1 microhaplotypes likely share the same mdr1 microhaplotype. For most of them, the mdr1 minor microhaplotype is present at very low proportions. PID42 is somewhat concordant (when the 2nd clone takes over at 642h) as you detect the second clone but miss the first clone (possibly just below the 5% threshold?). Any explanation for this? Maybe take a closer look at the blood sampling limit or the 5% BIC? *ama1* generally seemed to have worked better (e.g., generated a higher number of reads and higher sequencing success).

- *Thank you for raising these observations regarding the concordance of COI values between the ama1 and mdr1 genes. The expectation that multiple ama1 microhaplotypes might share the same mdr1 microhaplotype is based on ama1's higher expected heterozygosity than mdr1.*
- *Indeed, we conducted further analysis and found that the COI between the two genes varied significantly in paired samples ($p < 0.001$, Wilcoxon signed-rank test) even though both genes had a median COI of 2 and a range of 1 to 4. The read-depth analysis also revealed statistical differences ($p = 0.002$), with PFMDR1 having a median read-depth of 11,676 (range: 1,022-55,664) and PFAMA1 having a median read-depth of 13,391 (range: 4,673-33,714). Overall, ama1 showed greater read-depth and COI values in paired samples than mdr1.*
- *You make a valid point about the discrepancy in COI values for PIDs 06, 07, 18, 42, 47, and 53. The discrepancy in COI between ama1 and mdr1 likely arises from inherent differences in the levels of polymorphisms between the two markers. Nonetheless, ama1 was used to determine COI, while mdr1 was used to track drug resistance markers through treatment. Since they are under different selective pressures, concordance in their COI values is not necessarily expected.*

Competing Interests: No competing interests were disclosed.

Reviewer Report 14 July 2023

https://doi.org/10.21956/wellcomeopenres.20663.r61601

Eniyou Cheryll Oriero
[1] Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia
[2] Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia
[3] Medical Research Council Unit The Gambia at LSHTM, Fajara, The Gambia

The manuscript evaluated the use of amplicon sequencing to simultaneously determine complexity of *Plasmodium falciparum* infection, PCR correction and drug resistance following administration of triple combination ACTs. There are a few minor comments for the authors to address:

1. Parasite clearance: The authors have shown half-life graphs of parasite clearance estimates, adding a kaplan-meier curve of parasite clearance by microscopy and PCR (if data is available) will support parasite genotyping at the 72hr time point post-treatment, with evidence of PCR detectable parasites at that time point.

2. Though the manuscript is a proof-of-concept study with added advantage of simultaneous drug resistance genotyping, it has not sufficiently shown superiority of AmpSeq over *msp1/msp2/glurp* genotyping for PCR correction and determining complexity of infection. The authors may want to comment further on the cost-effectiveness and availability of expertise (e.g bioinformatics) for AmpSeq in LMICs.

3. The authors used artificial mixtures of two parasite strains to generate sequencing controls with artificial microhaplotypes similar to natural mixed infections. The authors should comment if they were able to determine copy number variation (particularly for *mdr1*) using read counts obtained from the AmpSeq data.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular diagnostics, Genomic epidemiology, non-falciparum malaria

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Sep 2023

Kevin Wamae

We want to thank Dr. Eniyou for agreeing to review this article, and we respond to her questions below:

1. ***Parasite clearance: The authors have shown half-life graphs of parasite clearance estimates, adding a Kaplan-Meier curve of parasite clearance by microscopy and PCR (if data is available) will support parasite genotyping at the 72hr time point post-treatment, with evidence of PCR detectable parasites at that time point:***

   The information presented in Figure S2C demonstrates that all individuals had eliminated the parasites by day 7 (hour 168). Therefore, the Kaplan-Meier (KM) analysis does not give us more information beyond this and is better suited to the parent study that included a larger sample size to conduct the KM analysis. It's noteworthy that parasites were undetectable by microscopy beyond 72 hours, and we acknowledge that we encountered challenges with PCR and sequencing failures from this point onward, which we discussed as one of the limitations in our study's discussion section.

   "*...The low parasitaemia following treatment minimised the generation of good quality AmpSeq data, impacting the sample size. Consequently, there were limited samples to examine the changes in microhaplotypes throughout the study period...*"

2. **Though the manuscript is a proof-of-concept study with added advantage of simultaneous drug resistance genotyping, it has not sufficiently shown superiority of AmpSeq over msp1/msp2/glurp genotyping for PCR correction and determining complexity of infection. The authors may want to comment further on the cost-effectiveness and availability of expertise (e.g. bioinformatics) for AmpSeq in LMICs:**

   We have amended the manuscript to acknowledge this observation in the last three

paragraphs of the discussion, as indicated below:

"...While this study serves as a proof of concept, the sample size employed was not aimed at establishing the clear superiority of AmpSeq over msp1/msp2/glurp genotyping for tasks such as PCR correction and determining infection complexity. Moreover, the adoption of AmpSeq in low and middle-income countries (LMICs) is challenged by its current cost and the requisite molecular biology and bioinformatics expertise needed for effective implementation. Nevertheless, the constantly declining costs of next-generation sequencing, coupled with emerging technologies like the Oxford Nanopore platform that offer reduced buy-in and maintenance expenses, hint at a potentially more accessible AmpSeq in the future. Regarding its performance, our study illustrates that AmpSeq yields results comparable to those obtained from msp1/msp2/glurp genotyping. However, AmpSeq boasts additional benefits over msp1/mps2/glurp genotyping. It demonstrates scalability advantages and heightened resolution based on sequence identity, surpassing the limitations of msp1/msp2/glurp genotyping, such as the reliance on labour-intensive gel electrophoresis-based band size analysis or fragment analysis that can be subject to interpretation errors and the inability to detect low-density parasite clones. An additional AmpSeq's strength lies in its potential for multiplexing, enabling the simultaneous identification of drug resistance markers alongside genotyping. This multiplexing capability enhances the assay's versatility and could offer valuable insights into the parasite's drug resistance profiles. In light of these considerations, while the current prerequisites of expertise and cost might impede the immediate integration of AmpSeq in LMICs, the evolving landscape of next-generation sequencing technologies and cost reductions inspire optimism regarding AmpSeq's future accessibility. Therefore, this study recognises the prospective value and adaptability of AmpSeq for tasks like PCR correction and determining infection complexity, even though broader adoption may necessitate ongoing advancements in affordability and the availability of expertise..."

3. **The authors used artificial mixtures of two parasite strains to generate sequencing controls with artificial microhaplotypes similar to natural mixed infections. The authors should comment if they were able to determine copy number variation (particularly for mdr1) using read counts obtained from the AmpSeq data:**

   In our study, we employed AmpSeq, a method characterised by deep sequencing of targeted regions but with shorter read lengths than whole-genome sequencing. While AmpSeq excels in the depth of sequencing within specified zones, the comprehensive detection of CNVs demands a wider genomic scan to pinpoint duplications or deletions. Concerning the *mdr1* gene, our sequencing strategy was not tailored for CNV identification. As such, while AmpSeq might provide preliminary indications of CNVs through read-depth analysis, robust and precise CNV assessments necessitate other techniques, and would require the inclusion of controls that include mdr1 genes with CNVs for comparison. This was not an objective of our study, and we primarily focused on characterising microhaplotypes for drug resistance.

Competing Interests: No competing interests were disclosed.

Reviewer Report 20 January 2023

https://doi.org/10.21956/wellcomeopenres.20663.r53729

Mita Toshihiro
[1] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan
[2] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan
[3] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan

I read the detailed response and finally found that revision was satisfactory.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 31 May 2022

https://doi.org/10.21956/wellcomeopenres.19627.r49308

?

Mita Toshihiro
[1] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan
[2] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan
[3] Department of Tropical Medicine and Parasitology, Juntendo University, Tokyo, Japan

The authors investigated the usefulness of Targeted amplicon deep sequencing (TADS) as a molecular marker to track within-host parasite diversity during and after the ACT treatments. They constructed ama1 haplotypes based on the deep sequence data and evaluated the presence/absence of haplotypes in the blood samples frequently obtained from treated patients using three different regimens. Despite that sample numbers were not so large, this study showed important findings.

Major

1. Since dead-parasites' DNA can persist in the human blood after treatment, the evaluate amount (also presence) of haplotypes may not have directly reflected the real situation. This may become a potential confounding factor. Could you comment on this?

2. Could you discuss the potential artifact of parasite DNA from persistent gametocytes after treatment?

3. In PID27, one clone overtook the other two clones and became dominant 6 hour after treatment. This clone has been regarded as slow clearing haplotype in the manuscript. What did the authors think the cause of slow clearing? I am interested in the discussion about this observation. I think there is a possibility that DNA from dead-sequestered parasites may have affected the result. If a particular clone tends to sequester and that DNA of such clone releases into blood stream after treatment, it will be expected that relative frequency of such parasite increases after treatment. I feel that, from your data, these may be unlikely to occur in the studied subject except from PID 27.

4. Minor haplotypes<5% were excluded from the analysis in this study. This cut-off was set based on "the lowest sequencing control's mixture (95% 3D7 vs 5% Dd2)". This is a bit unclear for me because next to this control mixture was 0% vs 100%. If control mixture of 99% 3D7 and 1% Dd2 shows almost same ratio in the TADS, will the authors take this cut-off?

5. In the abstract, they described that "This study highlights TADS as a sensitive tool for tracking parasite haplotypes throughout treatment ~". However, it might be difficult to say "sensitive tool" without detection limit of this deep-sequencing method. Furthermore, I am very interested in whether TADS is relatively useful compared to traditional methods in

terms of sensitivity. However, I could not find the data that compared the parasite detection ability between TADS and molecular methods (msp-1, msp-2 and glurp) recommended by WHO. Would it be possible to show further data?

Minor

1. Page 4; “Nonetheless, artemisinin (ART) resistance emerged and spread in Southeast (SE) Asia, evidenced as delayed parasite clearance following ACT treatment.” Emergence of artemisinin resistance has been already evidenced in our study in Northern Uganda (Balaikagara et al. NEJM 2021)[1] and Rwanda (Uwimana et al. Lancet Infect Dis, 2021).[2] These can be included in the Introduction.

2. Page 4; “The DNA from 3D7 and Dd2, were mixed as follows to come up with sequencing controls: 100%:100%, 75%:25%, 85%:15%, 95%:5% and 100%:0% to determine the lowest limit of haplo-type detection.” “Does “100%:100%” mean “50%:50%”?

3. Page 4; How much blood volume did you use for deep sequencing?

4. Page 5; The word “haplotype” is defined as “The paired consensus reads for each sample were trimmed and clustered to estimate the frequency of clusters (henceforth referred to as “haplotypes”)”. However, “haplotype” was also used when it implies parasite clone, e.g. “TADS can also identify slow clearing haplotypes, a potential early sign of selection during treatment.” I understand that strikingly they are not clones, but haplotypes used here are different from the aforementioned definition.

5. Page 5; “There was successful detection of the two expected ama1 haplotypes from the 3D7 and Dd2 laboratory isolates, as well as three mdr1 haplotypes, one from 3D7 and the 2 copies of Dd2”. I think some readers might feel curious about two haplotypes in the laboratory clone (Dd2). Could add some explanation?

6. Page 6; “There was only one exception as individual PID42 had data post-72h and there was a change in the ama1 haplotype detected before and after 72h.” How did you interpret this case?

7. Figure S1; There are inconsistency of color between the figure and legend. It seems to be successfully sequenced samples in blue and failed sequencing in black. Furthermore, I could not find the number of samples successfully sequenced in the last row in each panel.

References

1. Balikagala B, Fukuda N, Ikeda M, Katuro O, et al.: Evidence of Artemisinin-Resistant Malaria in Africa. *New England Journal of Medicine*. 2021; **385** (13): 1163-1171 Publisher Full Text
2. Uwimana A, Umulisa N, Venkatesan M, Svigel S, et al.: Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *The Lancet Infectious Diseases*. 2021; **21** (8): 1120-1128 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Malaria drug resistance

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 27 Sep 2022

Kevin Wamae

Major

1. Since dead-parasites' DNA can persist in the human blood after treatment, the evaluate amount (also presence) of haplotypes may not have directly reflected the real situation. This may become a potential confounding factor. Could you comment on this?*We respond to this under point 2.*

2. Could you discuss the potential artifact of parasite DNA from persistent gametocytes after treatment?

- *It's possible that some of the signals detected originated from gametocytes, dead or dormant parasites and we have included this as one of the study's limitations in the discussion:*

  "Microhaplotypes detected post-treatment might have originated from circulating gametocytes, dormant or dead parasites. However, all participants were gametocyte negative throughout treatment except PID38 who had five gametocytes/µl at 72h but was negative thereafter, so it is unlikely that gametocytaemia biased our findings. As for genetic material originating from dormant or dead parasites,

parasite mRNA can also be detected up to two weeks after successful treatment in microscopy-negative individuals (Mahamar et al. 2021). However, additional work is needed to determine whether these originate from viable infections"

3. In PID27, one clone overtook the other two clones and became dominant 6 hour after treatment. This clone has been regarded as slow clearing haplotype in the manuscript. What did the authors think the cause of slow clearing? I am interested in the discussion about this observation. I think there is a possibility that DNA from dead-sequestered parasites may have affected the result. If a particular clone tends to sequester and that DNA of such clone releases into blood stream after treatment, it will be expected that relative frequency of such parasite increases after treatment. I feel that, from your data, these may be unlikely to occur in the studied subject except from PID 27.

- *We have included a sentence in the discussion to highlight this individual and we hypothesize that the increase in one ama1 haplotype may have come from DNA of dead parasites:*

  "Notably, PID27 had a COI of three at 0h, and while the relative frequencies of two ama1 microhaplotypes decreased over time, one microhaplotype continued to increase in frequency. This patient did not experience a recurrence and had no microscopically detectable parasites by 72h. Therefore, the increase in one ama1 microhaplotype may have originated from the nucleic material of dead parasites."

4. Minor haplotypes <5% were excluded from the analysis in this study. This cut-off was set based on "the lowest sequencing control's mixture (95% 3D7 vs 5% Dd2)". This is a bit unclear for me because next to this control mixture was 0% vs 100%. If control mixture of 99% 3D7 and 1% Dd2 shows almost same ratio in the TADS, will the authors take this cut-off?

- *The control mixture 0% vs 100% was basically a clonal mixture of 3D7 and this helped to evaluate whether we would detect false positives since we expect to find only one microhaplotype. Conversely, the control's mixture 95% 3D7 vs 5% Dd2 was multiclonal to determine the ability of the assay to pick up multiclonal infections as well as multiclonal infections with minor variants.*

  To respond to the last question, yes, if we included an artificial mixture of 3D7 and Dd2 with lower than 5% frequency and detected haplotypes, we would have taken that cut-off.

In the abstract, they described that "This study highlights TADS as a sensitive tool for tracking parasite haplotypes throughout treatment ~". However, it might be difficult to say "sensitive tool" without detection limit of this deep-sequencing method. Furthermore, I am very interested in whether TADS is relatively useful compared to traditional methods in terms of sensitivity. However, I could not find the data that compared the parasite detection ability between TADS and molecular methods (msp-1, msp-2 and glurp) recommended by WHO. Would it be possible to show further data?

- *Unfortunately, we do not have this data and we acknowledge this limitation in the discussion:*

  "...Finally, future studies should include artificial sequencing controls of decreasing parasitaemia from, e.g., dilutions ranging from 10,000 to 1 parasite/µl to reliably

determine the limit of detection, especially in post-treatment with low parasitemia…"

Minor

Page 4; "Nonetheless, artemisinin (ART) resistance emerged and spread in Southeast (SE) Asia, evidenced as delayed parasite clearance following ACT treatment." Emergence of artemisinin resistance has been already evidenced in our study in Northern Uganda (Balaikagara et al. NEJM 2021)[1] and Rwanda (Uwimana et al. Lancet Infect Dis, 2021).[2] These can be included in the Introduction.

- *We have included this in the introduction:*

  "…Two recent studies have identified early signs of ART resistance in Rwanda (Uwimana et al. 2021) and Uganda (Balikagala et al. 2021) and this looming threat of widespread ACT resistance would be catastrophic in sub-Saharan Africa, where the burden of malaria is the most significant (WHO 2021a)…"

Page 4; "The DNA from 3D7 and Dd2, were mixed as follows to come up with sequencing controls: 100%:100%, 75%:25%, 85%:15%, 95%:5% and 100%:0% to determine the lowest limit of haplotype detection." "Does "100%:100%" mean "50%:50%"?

- *To make it easier for the reader to understand the mixtures, we have switched from percentages to ratios of 0.5:0.5, 0.75:0.25, 0.85:0.15, 0.95:0.05 and 1:0.*

Page 4; How much blood volume did you use for deep sequencing?

- *We have included the volume in the methods:*

  "…DNA was extracted from P. falciparum laboratory reference clones, 3D7 and Dd2 (BEI Resources), and from 200µl of frozen patient blood samples using the QIAamp DNA Blood Mini Kit (Qiagen) according to the manufacturer's instructions…"

Page 5; The word "haplotype" is defined as "The paired consensus reads for each sample were trimmed and clustered to estimate the frequency of clusters (henceforth referred to as "haplotypes")". However, "haplotype" was also used when it implies parasite clone, e.g. "TADS can also identify slow clearing haplotypes, a potential early sign of selection during treatment." I understand that strikingly they are not clones, but haplotypes used here are different from the aforementioned definition.

- *We acknowledge how confusing this was. Hence, we have clarified the use of "haplotype" by settling on the term "microhaplotype" to mean the set of amino acid polymorphisms found on a single sequence.*

Page 5; "There was successful detection of the two expected ama1 haplotypes from the 3D7 and Dd2 laboratory isolates, as well as three mdr1 haplotypes, one from 3D7 and the 2 copies of Dd2". I think some readers might feel curious about two haplotypes in the laboratory clone (Dd2). Could add some explanation?

- We have included the following statement in the methods:
- ***"…The sequencing controls allowed for the detection of two ama1 (3D7 and Dd2) and three mdr1 (one from 3D7 and two from Dd2) microhaplotypes. Dd2 contained two mdr1 gene copies generated due to adaptation to in vitro culture…"***

Page 6; "There was only one exception as individual PID42 had data post-72h and there was a change in the ama1 haplotype detected before and after 72h." How did you interpret this case?

- *To present the data on changes of ama1 microhaplotypes throughout treatment, we grouped participants into those with monoclonal infections and those with multiclonal infections. PID42 had a monoclonal infection since like all individuals in this group, only one ama1 haplotype was observed per timepoint. However, unlike all monoclonal infections where the same clone at baseline was seen throughout treatment, PID42 had a change in clones at the time of recurrence.*

Figure S2; There are inconsistencies of color between the figure and legend. It seems to be successfully sequenced samples in blue and failed sequencing in black. Furthermore, I could not find the number of samples successfully sequenced in the last row in each panel.

- *Thank you for pointing out the inconsistencies in colours. We have updated the figure legend for match the colours depicted in the figure. The last row indicating the number of successfully sequenced samples was also included in Figure. S2.*

Competing Interests: No competing interests were disclosed.

Reviewer Report 31 March 2022

https://doi.org/10.21956/wellcomeopenres.19627.r49311

?

Ian M. Hastings
[1] Liverpool School of Tropical Medicine, Liverpool, UK
[2] Liverpool School of Tropical Medicine, Liverpool, UK
[3] Liverpool School of Tropical Medicine, Liverpool, UK

We recently published a paper investigated the use of deep sequenced amplicons in malaria drug clinical trials, in which we used extensive computer simulations of parasite dynamics post-treatment to identify how best to incorporate and interpret deep sequenced amplicon data: Jones, S., et al. (2021).[1]

The aim was to identify best practice for their application and to identify traps and errors likely to arise in in their initial applications. That paper was focused on molecular correction, but a

prerequisite for this is good quality genotyping data and we spent a lot of time considering this.

The authors of the current paper appear unaware of this work which is unfortunate as they appear to have fallen into some of the pitfalls we anticipate, in particular (i) the necessity of estimating a blood sampling limit (ie. the minimum number or density of parasites in a patient that is likely to be detected by the sequencing) and (ii) ability of deep sequencing to detect genetic material in circulating gametocytes. These are discussed below.

One of our recommendations was to be very careful using the word “haplotype” as it can mean at least three distinct things in the current context (part 1 of SI in (Jones et al)[1]

- A unique genetic variant at the amplified locus. This terminology has slipped into the bio-informatics jargon but in genetics is known as an “allele”.
- The genotype along a whole contiguous stretch of chromosome which may contain several distant alleles
- The entire malaria clonal genotype (because it is haploid in the human stages.)

So in their Abstract they use “haplotype” to mean two different things. In the first part it means allele (as in “overall *ama1* detected more haplotypes)”. Whereas later it means the whole malaria genotype e.g. “we identified a fast (<1h) and slow (>5h) clearing haplotype.” and “clone” would be much better here than “haplotype”.

This is not just being pedantic: using the same term to mean different thing is confusing even for authors and it is important to establish some sort of unambiguous, consensus terminology early in the development of field.

I’ll enumerate comment for ease of cross reference.

(1) Please be accurate and consistent with use of the term “haplotype” and consider revising the manuscript accordingly.

(2) Note that in our publication [1]we used the abbreviation “AmpSeq”, the current authors used “TADS” to mean the same thing. It would be nice to establish a consistent terminology noting the latest WHO recommendation also uses “AmpSeq” i.e.
https://www.who.int/publications/i/item/9789240038363

If “TADS” is already in the literature I would request (but not demand) they consider changing TADS to Ampseq. Like the use of “haplotype” it is not essential but if a field starts off by using different anacronyms for the same thing it tends to persist which is in nobody’s long-term interest. If “TADS” is not already in use in the literature it would be hard to justify using a new terms when one already exist.

Methods.

(3) I have sat through seeming interminable technical discussions on amplicon sequencing, most of which went over my head. I do know enough to realise it is not as straightforward as most people think so suggest an expert on genotyping also reviews this manuscript. The authors noted that they could detect the minority allele in mixture of 95%:5% and regarded any reads less than 5% of total as noise. We previously found this threshold to be extremely important (i.e. [1]) and termed it the Bioinformatics cut-off (BIC), and identified it as a vital piece of information to be

included in publications, so it is nice to see it clearly defined here. However, on page 8 it states "our analysis of rare haplotypes was limited to a 5% cut-off based on sequencing controls" so some clarification is needed. Is it 5% of total reads or 5% of controls? If the latter, this could be explained much more clearly.

(4) It looks like the authors extracted their DNA from frozen blood samples. One of things we previously identified as important was the "blood sampling limit"[1] i.e. the number or density of parasites below which they were unlikely to be detected because the blood volume entering the genotyping assay was so small that parasites of a given genotype were statistically unlikely to be present in the sample. We made an estimate based on blood extracted from dried blood spots on filter paper. It looks like the authors need to the same calculations based on the volume of blood they used. This is important for understanding the likely impact of gametocytes, and to interpret observed patterns (i.e. their figure 1) post treatment; see comment (10) below. They have total parasitaemia and relative frequencies of ama1 alleles in each patient so they can back-calculate likely number/density of each clone and hence figure out whether it is in danger of falling below the sampling limit. The calculations tend to be estimates but they need to do this before interpreting their figure 1 i.e. they need to be confident that all the ama1 signals in the patient are being detected and there is no danger that, at later timepoints, one ama1 signal becoming so low (because the clone(s) containing it has fallen to very low density) that it is no longer detected in the assay despite its continued persistence in the patient

The blood sampling limit is also one plausible explanation for their observation (page 6) that" In five individuals (PIDs 10, 32, 40, 59 and 60), there were sporadic detections of rare *ama1* haplotypes " i.e. because parasitaemias of the clone containing that ama1 allele were around the sampling limit so that whether or not they were detected was a matter of chance. The second plausible reason is that it fluctuated around the 5% BIC so sporadically exceed the BIC. Discussion around these points would be good and it also emphasises why it is important to establish the sampling limit before interpreting the results

(5) I am worried about these statements "The frequency of each haplotype in the population was calculated using the total number of samples that contained the haplotype over the total number of samples genotyped." This will calculate prevalence *not* frequencies (this has caused immense confusion in the past). The problem can be illustrated by a simple example: assume there are only two samples: sample #1 has alleles A,B and C, and sample #2 has A, B, and D. Then according to their statement the "frequencies" of A,B, C and D will be 100%, 100%, 50% and 50% respectively which is obviously wrong as in this context the allele frequencies should sum to 100%. I assume the statement is written incorrectly and that they counted total number of clones in the two samples i.e. 3+3=6 then calculated the frequencies of A,B, C and D as 2/6. 2/6, 1/6 and 1/6 respectively i.e. ensuring frequencies sum to 100%. They then need, for reasons given below, to calculate expected heterozygosity (He) to quantify genetic variability ie.the probably that two randomly selected clones have different alleles

I also worry about their statement on page 10 when discussing table S2 i.e. "The frequencies were calculated by dividing the number of reads of each haplotype by the total number of reads obtained (116,187,131)." These "frequencies" are likely biased towards alleles in clones in high-parasitaemia patients. So I think the calculation of frequency suggested above is better, especially because it can be used to calculate He.

(6) I assume the genotyping at msp1/msp2/glurp was the standard WHO-recommended method of gel electrophoresis. As above, the sensitivity (equivalent to the BIC) used in their assay should be stated. Labs typically regard genotyping peaks less than 25% or 30% of the largest peak as noise but there is no consistency in this (and some labs apparently do it "informally") but this needs to be stated.

Interpretation

(7) My main concern is that the authors seem to take no account of the possible impact of detecting alleles in gametocytes. We identified this as a key consideration as illustrated in Figure 2 of our main text and extensively discussed in our Supplementary Material Part 3: "Simulating the potential impact of gametocytes on the accuracy of molecular correction".[1] The problem is that current drugs clear the asexual form of falciparum but not the gametocytes which continue to circulate and the increased sensitivity of deep sequencing allows detection of alleles in the gametocytes. Suppose a clone initially has 1% gametocytes then after treatment their proportion increases (roughly) 2 fold with every half-life of the asexual forms i.e. gametocytes increase (approx.) from 1% to 2% to 4% to 8% etc after 1,2,3,4, etc half-lives. The observed clearance half-life will therefore gradually decrease as the drug-insensitive gametocytes make up an increasing proportion of the clone. Whether the gametocytes will be observable in the initial test slide depends on the microscopy and on the clone size. So it's a two-stage calculation Firstly, the extent to which any putative gametocytes will gradually reduce observed clearance rates post-treatment. Secondly, the likely impact of gametocyte signals will also depend on the blood sampling limit (i.e. whether gametocytes are of sufficient density to de detectable) which I have suggested the authors calculate (cf Fig 2 of Jones *et al*.)

(8) First paragraph of section "Pfama1 genetic diversity during and after treatment". The slight decrease in COI is presumably a consequence of clones dropping below the blood sampling limit (see above), or below the 5% BIC, and being lost. Given that monoclonal infections cannot lose diversity, it would be better to exclude these infections from the calculations to get a better idea of the drop in COI

(9) I am extremely curious about the assertion made in the caption to figure 2 i.e. "PID32 had the fastest clearing (0.8h) *ama1* haplotype V9.". Looking it Figure 1, it appears that PID32 was largely monoclonal except for a single detection of another ama1 allele at time 8hours which may well be spurious. More explanation is needed.

(10) So once background calculation have been made for blood sampling limit and impact of gametocytes there are (as far as I can see) three plausible patterns post treatment.

1. Immediate and sustained difference in clearance rates between clones in the same patient. PID27 therefore looks the most interesting to me as is suggested significant differences in clearance between the two clones immediately post-treatment (Figure 1). Figure 2B also shows that the two clearance rates in this patient are clearly distinguishable on the Y axis unlike the other patients . Note this could be due to differences in drug sensitivity or, less likely, to differential impact of immunity acting on the two clone.

2. An impact of gametocytes. All clones would retain roughly the same proportions immediately post-treatment, but one eventually rises. The one that increases in proportion could arise as gametocytes become a significant proportion of its biomass (see point 7

above) which are not cleared by the drug . PID08 seems to show this pattern.

3. Impact of a sampling limit. All clones maintain same proportions until the lower one falls below the sampling limit and disappears, hence the 100% has to be retained by increased proportions on those remaining above the sampling limit. This would be an artefact. Interesting, I can't see evidence of this of Figure 1 so I'm guessing sampling limit was not a big issue (the requested background calculation should confirm this.

So these are three paradigms. There may be variation caused by factors such as sequestration, but given that the authors chose to present and discuss the profiles, I strongly suggest they improve their discussion to include the putative/likely impact of gametocytes and sampling limits.

(11) Page 9 they state "Hence, *ama1* is a highly discriminatory marker with a larger number of variants with a prevalence of <5%, providing a higher resolution to better distinguish new from recrudescent infections in moderate to high transmissions settings where polyclonal infections are common." This is probably incorrect and definitely incomplete. Msp1, msp2 and glurp are all highly variable, possible more so than ama1: they would need to report He values for ama1 to substantiate this assertion of greater genetic diversity. The most important point about deep sequencing vs msp1/msp2/glurp is sequencing has much better ability to detect minor clones... they need to cite our previous work (i..e Jones *et al*.) here as we explicitly discussed the impact of improved sensitivity in detail.

Recommendations.

The manuscript is, in my opinion, methodologically and genetically naive in parts. I realize it will appear arrogant to insist they go and read our previous investigations of this (i.e. Jones *et al*), but there was a team of us who spent a lot of time and effort anticipating likely problems (and how to mitigate them) and this is highly relevant to the work presented here. I would recommend major revisions to the data interpretation and discussion. The underlying approach and data seem robust to me.

Minor comments.

(a) Page 4. The discussion of limitations of current WHO-recommended method is a bit dated and it would be useful to cite our two recent quantitative investigations i.e.. Felger, I., et al. (2020), and Jones, S., et al. (2019). [ref-2, ref-3]

Also note that WHO recently updated their genotyping recommendations to acknowledge the future importance of amplicon sequencing
https://www.who.int/publications/i/item/9789240038363.

(b) I see no point in parts (A) of figures 1 and 3.... If an infection is monoclonal the one allele will inevitable be present at 100% i.e. the plots tell us nothing.

(c) Page 8. The statement "The clearance rates were similar within and between individuals irrespective of clonality, suggesting that the three antimalarial treatments were equally effective in clearing haplotypes and parasitaemia." Is a bit naive in my opinion. Differences in artemisinin sensitivity putatively result in differences in clearance rate post-treatment. However there is no

suggestion this occurs for the partner drugs which mainly kill parasites when sequestered: resistance to the partner drug will likely first show as a reduced parasite reduction ratio not as reduced clearance.

References

1. Jones S, Kay K, Hodel E, Gruenberg M, et al.: Should Deep-Sequenced Amplicons Become the New Gold Standard for Analyzing Malaria Drug Clinical Trials?. *Antimicrobial Agents and Chemotherapy*. 2021; **65** (10). Publisher Full Text
2. Felger I, Snounou G, Hastings I, Moehrle J, et al.: PCR correction strategies for malaria drug trials: updates and clarifications. *The Lancet Infectious Diseases*. 2020; **20** (1): e20-e25 Publisher Full Text
3. Jones S, Kay K, Hodel EM, Chy S, et al.: Improving Methods for Analyzing Antimalarial Drug Efficacy Trials: Molecular Correction Based on Length-Polymorphic Markers msp-1, msp-2, and glurp.*Antimicrob Agents Chemother*. **63** (9). PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
No

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Malaria drug treatment and resistance.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 27 Sep 2022

Kevin Wamae

1. Please be accurate and consistent with use of the term “haplotype” and consider revising the manuscript accordingly.

- *We have clarified the use of "haplotype" by settling on the term "microhaplotype" to mean the set of amino acid polymorphisms found on a single DNA amplicon. The term microhaplotype is now a preferred term in similar work.*

***Reference**:*

- *Tessema, S. K., et. al. (2020). Sensitive, Highly Multiplexed Sequencing of Microhaplotypes From the Plasmodium falciparum Heterozygome. The Journal of Infectious Diseases, 1–11.*
- *LaVerriere, E., et. al. (2022). Design and implementation of multiplexed amplicon sequencing panels to serve genomic epidemiology of infectious disease: A malaria case study. Molecular Ecology Resources, 22(6), 2285–2303.*
- *Taylor, A. R., Jacob, P. E., Neafsey, D. E., & Buckee, C. O. (2019). Estimating Relatedness Between Malaria Parasites. Genetics, 212(4), 1337–1351.*
- *WHO: Informal consultation on methodology to distinguish reinfection from recrudescence in high malaria transmission areas. https://www.who.int/publications/i/item/9789240038363*

2. Note that in our publication1 we used the abbreviation "AmpSeq", the current authors used "TADS" to mean the same thing. It would be nice to establish a consistent terminology noting the latest WHO recommendation also uses "AmpSeq" i.e. https://www.who.int/publications/i/item/9789240038363

If "TADS" is already in the literature I would request (but not demand) they consider changing TADS to Ampseq. Like the use of "haplotype" it is not essential but if a field starts off by using different anacronyms for the same thing it tends to persist which is in nobody's long-term interest. If "TADS" is not already in use in the literature it would be hard to justify using a new terms when one already exist.

- *Throughout the manuscript, we have switched to using AmpSeq.*

Methods.

2. I have sat through seeming interminable technical discussions on amplicon sequencing, most of which went over my head. I do know enough to realise it is not as straightforward as most people think so suggest an expert on genotyping also reviews this manuscript. The authors noted that they could detect the minority allele in mixture of 95%:5% and regarded any reads less than 5% of total as noise. We previously found this threshold to be extremely important (i.e. 1) and termed it the Bioinformatics cut-off (BIC), and identified it as a vital piece of information to be included in publications, so it is nice to see it clearly defined here. However, on page 8 it states "our analysis of rare haplotypes was limited to a 5% cut-off based on sequencing controls" so some clarification is needed. Is it 5% of total reads or 5% of controls? If the latter, this could be explained much more clearly.

- *We have clarified that this 5% is based on the proportion of the minor clone in the artificial mixture of the two laboratory clones, that is 95% 3D7 and 5% Dd2.*

4. It looks like the authors extracted their DNA from frozen blood samples. One of the things we previously identified as important was the "blood sampling limit"1 i.e. the number

or density of parasites below which they were unlikely to be detected because the blood volume entering the genotyping assay was so small that parasites of a given genotype were statistically unlikely to be present in the sample. **We made an estimate based on blood extracted from dried blood spots on filter paper**. It looks like the authors need to the same calculations based on the volume of blood they used. This is important for understanding the likely impact of gametocytes, and to interpret observed patterns (i.e., their figure 1) post-treatment; see comment (10) below. **They have total parasitaemia and relative frequencies of ama1 alleles in each patient so they can back-calculate likely number/density of each clone and hence figure out whether it is in danger of falling below the sampling limit**. The calculations tend to be estimates but they need to do this before interpreting their figure 1 i.e. they need to be confident that all the ama1 signals in the patient are being detected and there is **no danger that, at later time points, one ama1 signal becoming so low (because the clone(s) containing it has fallen to very low density) that it is no longer detected in the assay despite its continued persistence in the patient**.

The blood sampling limit is also one plausible explanation for their observation (page 6) that" In five individuals (PIDs 10, 32, 40, 59 and 60), there were sporadic detections of rare *ama1* haplotypes " i.e. because parasitaemias of the clone containing that ama1 allele were around the sampling limit so that whether or not they were detected was a matter of chance. The second plausible reason is that it fluctuated around the 5% BIC so sporadically exceed the BIC. Discussion around these points would be good and it also emphasises why it is important to establish the sampling limit before interpreting the results.

- *We have responded to this in comment 10, below.*

5. I am worried about these statements "The frequency of each haplotype in the population was calculated using the total number of samples that contained the haplotype over the total number of samples genotyped." This will calculate prevalence *not* frequencies (this has caused immense confusion in the past). The problem can be illustrated by a simple example: assume there are only two samples: sample #1 has alleles A,B and C, and sample #2 has A, B, and D. Then according to their statement the "frequencies" of A, B, C and D will be 100%, 100%, 50% and 50% respectively which is obviously wrong as in this context the allele frequencies should sum to 100%. I assume the statement is written incorrectly and that they counted total number of clones in the two samples i.e. 3+3=6 then calculated the frequencies of A,B, C and D as 2/6. 2/6, 1/6 and 1/6 respectively i.e. ensuring frequencies sum to 100%. They then need, for reasons given below, to calculate expected heterozygosity (He) to quantify genetic variability ie. the probably that two randomly selected clones have different alleles.

- *We agree with the suggestion. However, we have dropped the statement "The frequency of each haplotype in the population was calculated using the total number of samples that contained the haplotype over the total number of samples genotyped." This population frequency statistic was not used anywhere in the paper and was erroneously included.*

I also worry about their statement on page 10 when discussing table S2 i.e. "The frequencies were calculated by dividing the number of reads of each haplotype by the total number of reads obtained (116,187,131)." These "frequencies" are likely biased towards alleles in clones in high-parasitaemia patients. So, I think the calculation of frequency suggested above is better, especially because it can be used to calculate H_e.

- *This has been edited to "The frequencies were calculated by dividing the number of reads of each microhaplotype by the **total number of reads obtained per sample**".*

- *Therefore, we did not use just total reads but total reads per sample. Sequencing libraries from all samples were normalised to equimolar concentrations before pooling and sequencing to avoid having more input DNA from higher parasitaemia samples. Additionally, one of our cut-offs for identifying minor variants required that they were identified in other samples so this reduces bias due to high parasitaemia samples.*

6. I assume the genotyping at *msp1/msp2/glurp* was the standard WHO-recommended method of gel electrophoresis. As above, the sensitivity (equivalent to the BIC) used in their assay should be stated. Labs typically regard genotyping peaks less than 25% or 30% of the largest peak as noise but there is no consistency in this (and some labs apparently do it "informally") but this needs to be stated.

- *We have included the following statement in the methods section:*

  "msp1/msp2/glurp genotyping was performed according to the WHO-recommended method of gel electrophoresis (WHO 2008). These data were obtained from the original study, and the following was ensured during the analysis: Each PCR product had to have well-defined and easy to visualise, bands had to be bright and sharp to be of sufficient quality for scoring, PCRs were repeated if bands appeared in the negative control, the interpretation of results did not include products with less than 100 bp and did not account for faint bands or bands that formed smile-shaped patterns on the gel (Hamaluba et al. 2021)"

Interpretation

7. My main concern is that the authors seem to take no account of the possible impact of detecting alleles in gametocytes. We identified this as a key consideration as illustrated in Figure 2 of our main text and extensively discussed in our Supplementary Material Part 3: "Simulating the potential impact of gametocytes on the accuracy of molecular correction".1 The problem is that current drugs clear the asexual form of falciparum but not the gametocytes which continue to circulate and the increased sensitivity of deep sequencing allows detection of alleles in the gametocytes. Suppose a clone initially has 1% gametocytes then after treatment their proportion increases (roughly) 2 fold with every half-life of the asexual forms i.e. gametocytes increase (approx.) from 1% to 2% to 4% to 8% etc after 1,2,3,4, etc half-lives. The observed clearance half-life will therefore gradually decrease as the drug-insensitive gametocytes make up an increasing proportion of the clone. Whether the gametocytes will be observable in the initial test slide depends on the microscopy and on the clone size. So it's a two-stage calculation Firstly, the extent to which any putative gametocytes will gradually reduce observed clearance rates post-treatment. Secondly, the likely impact of gametocyte signals will also depend on the blood sampling limit (i.e. whether gametocytes are of sufficient density to de detectable) which I have suggested the authors calculate (cf Fig 2 of Jones *et al*.)

- *This has been included as one of our limitations. None of the 30 individuals was gametocyte on day 0 and they remained negative up to day 42. One participant, however, was gametocyte positive on day 3 but negative thereafter so it's unlikely that gametocytaemia biased our findings.*

8. First paragraph of section "Pfama1 genetic diversity during and after treatment". The slight decrease in COI is presumably a consequence of clones dropping below the blood sampling limit (see above), or below the 5% BIC, and being lost. Given that monoclonal infections cannot lose diversity, it would be better to exclude these infections from the calculations to get a better idea of the drop in COI.

- *After excluding monoclonal infections, we still see a decrease in COI from before treatment, at 0h the mean COI was 2.44, 0.5h to 72h the mean COI was 2.37 and 2 thereafter. These new values have been included in the text.*

9. I am extremely curious about the assertion made in the caption to figure 2 i.e. "PID32 had the fastest clearing (0.8h) *ama1* haplotype V9.". Looking at Figure 1, it appears that PID32 was largely monoclonal except for a single detection of another ama1 allele at time 8hours which may well be spurious. More explanation is needed.

- *Regarding the single ama1 microhaplotype detected at 8h, we have shown in extended data table S3 that it was detected above the blood sampling limit. Still, we have no reason to believe that PID32's infection being clonal would have biased our estimate of determining parasite clearance.*

10. So once background calculation have been made for blood sampling limit and impact of gametocytes there are (as far as I can see) three plausible patterns post treatment.

1. Immediate and sustained difference in clearance rates between clones in the same patient. PID27 therefore looks the most interesting to me as is suggested significant differences in clearance between the two clones immediately post-treatment (Figure 1). Figure 2B also shows that the two clearance rates in this patient are clearly distinguishable on the Y axis unlike the other patients . Note this could be due to differences in drug sensitivity or, less likely, to differential impact of immunity acting on the two clone.
    - *We did not see a difference in the sensitivity of the drugs; hence we hypothesize that acquired immunity may have had a larger role to play in the difference in clearing times for PID27.*

10. An impact of gametocytes. All clones would retain roughly the same proportions immediately post-treatment, but one eventually rises. The one that increases in proportion could arise as gametocytes become a significant proportion of its biomass (see point 7 above) which are not cleared by the drug . PID08 seems to show this pattern.
As mentioned above, all participants in the current study were gametocyte negative throughout treatment, thus it's likely that gametocytaemia biased our findings.
Impact of a sampling limit. All clones maintain same proportions until the lower one falls below the sampling limit and disappears, hence the 100% has to be retained by increased proportions on those remaining above the sampling limit. This would be an artefact. Interesting, I can't see evidence of this of Figure 1 so I'm guessing sampling limit was not a

big issue (the requested background calculation should confirm this).

- *We set our arbitrary blood sampling limit to 10 parasites/µl based on the lower limit used in the PMID: 34252299. In multiclonal infections, some microhaplotypes started to fall below this limit, however, for the seven individuals with microhaplotypes that were detected sporadically, only one (PID60) had a microhaplotype below this limit.*

- *We have included this comment in the discussion, and we recommended future studies to include artificial sequencing-controls with decreasing parasitaemia, e.g. 10,000 down to 1 parasite/µl to determine their limit of detection.*

So these are three paradigms. There may be variation caused by factors such as sequestration, but given that the authors chose to present and discuss the profiles, I strongly suggest they improve their discussion to include the putative/likely impact of gametocytes and sampling limits.

- *We had mentioned earlier that the sporadic detections of rare microhaplotypes could have originated from changes in parasite densities, sequestered parasites, and lingering nucleic material from dead parasites.*

- *We have also included in our limitations that these rare microhaplotypes could have originated from gametocytes too. However, this was unlikely in this study since all participants were gametocyte negative throughout treatment.*

11. Page 9 they state “Hence, *ama1* is a highly discriminatory marker with a larger number of variants with a prevalence of <5%, providing a higher resolution to better distinguish new from recrudescent infections in moderate to high transmissions settings where polyclonal infections are common.” This is probably incorrect and definitely incomplete. Msp1, msp2 and glurp are all highly variable, possible more so than ama1: they would need to report H_e values for ama1 to substantiate this assertion of greater genetic diversity. The most important point about deep sequencing vs msp1/msp2/glurp is sequencing has much better ability to detect minor clones… they need to cite our previous work (i..e Jones *et al*.) here as we explicitly discussed the impact of improved sensitivity in detail.

- *This section has been amended to include the high H_e we obtained for the ama1 locus we genotyped. Additionally, we have added that amplicon sequencing is more sensitive than msp1/msp2/glurp genotyping because it can better capture minority variants. The section now reads as follows:*

  “hence, the ama1 locus genotyped in this study is a highly discriminatory marker with a larger number of variants with a prevalence of <5% and a high expected heterozygosity value (0.96). Coupled with the high sensitivity of amplicon sequencing to capture minor variants, this assay provides a higher resolution to better distinguish reinfections from recrudescent infections in moderate to high transmissions settings where polyclonal infections are common.”

Competing Interests: No competing interests were disclosed.