Malaria Journal

REVIEW

CrossMark

Molecular approaches to determine the multiplicity of *Plasmodium* infections

Daibin Zhong[1]*, Cristian Koepfli[1], Liwang Cui[2] and Guiyun Yan[1]*

Abstract

Multiplicity of infection (MOI), also termed complexity of infection (COI), is defined as the number of genetically distinct parasite strains co-infecting a single host, which is an important indicator of malaria epidemiology. PCR-based genotyping often underestimates MOI. Next generation sequencing technologies provide much more accurate and genome-wide characterization of polyclonal infections. However, complete haplotype characterization of multiclonal infections remains a challenge due to PCR artifacts and sequencing errors, and requires efficient computational tools. In this review, the advantages and limitations of current molecular approaches to determine multiplicity of malaria parasite infection are discussed.

Keywords: Malaria parasite, Multiplicity of infection, Molecular marker, Next generation sequencing, Amplicon ultra-deep sequencing, Single nucleotide polymorphism

Background

Individuals in malaria-endemic countries often carry several genetically distinct clones of the same parasite species. Multiplicity of infection (MOI), also termed complexity of infection (COI) is defined as the number of genetically distinct parasite strains co-infecting a single host [1]. Multiclonal infection can be the result of independent bites of infected mosquitoes (also called superinfection), or a single mosquito bite transmitting a genetically diverse sporozoite inoculum. Genetically-distinct malaria parasites in natural populations have an extremely high rate of genetic recombination during the sexual stages in a mosquito host, often resulting in multiple strains being transmitted simultaneously [2]. In the case of *Plasmodium vivax*, multiclonal infections also can be caused by the relapse of a liver-stage hypnozoite during ongoing blood-stage infection of a genetically distinct clone. The number of co-infections within a host might be an important indicator of transmission intensity, though data to date are inconclusive [3–9]. For example, a positive correlation between the rate of polyclonal infections and annual parasite incidence has been observed in

parasite populations from Indonesia [5] and Papua New Guinea [3], while in other studies, no correlation or negative correlation was found between the proportion of multiclonal infections and parasite prevalence [3, 6]. Several studies reported correlations between clinical symptoms and higher MOI [10–18] and others did not find any associations [19–21]. On the contrary, some studies reported that a reduced risk of clinical malaria was associated with polyclonal infections [22–24], while other studies reported that mono-infections and very common genotypes are more likely to develop severe malaria than polyclonal infections [21, 25]. MOI can also help to monitor vaccine efficacy. Vaccine-induced immune responses often show strain-transcending specificity [26, 27], depending on the polymorphic alleles of the vaccine candidate antigens [28]. Neafsey et al. found that the RTS, S vaccine has greater activity against malaria parasites with matched circumsporozoite protein allele than against mismatched parasite strains [29]. Hence, a reduction in MOI might be observed after vaccination, even if the risk of infection does not change. Understanding of the extent and dynamics of genetic diversity in vaccine antigens of all parasite strains is needed to guide rational vaccine design and to interpret the results of vaccine efficacy trials conducted in malaria endemic areas [30]. In drug trials, detecting all clones is important, as

*Correspondence: dzhong@uci.edu; guiyuny@uci.edu
[1] Program in Public Health, University of California, Irvine, CA 92617, USA
Full list of author information is available at the end of the article

Zhong *et al. Malar J* (2018) 17:172

Page 2 of 9

only one of them might be drug resistant and will result in recrudescence. If this clone is not detected at baseline, e.g., because it is present in a very low proportion of all clones, genotyping might falsely identify it as new infection and thus underestimate treatment failure. The molecular force of blood-stage infection (molFOB), i.e. the number of clones acquired over time, is an important parameter to estimate exposure at the individual level [31, 32]. Accurate estimates of molFOB require identification of all clones in multiclonal infections. However, sensitive identification of the within-host parasite component in multiclonal infection is often difficult due to lack of appropriate methods and analysis tools. In this review, the advantages and limitations of current molecular approaches to determine multiplicity of malaria parasite infection are assessed.

Size-polymorphic antigens and microsatellites

Over the past decade, hundreds of papers describing the multiplicity of *Plasmodium* infections have been published by typing size-polymorphic antigens or microsatellite markers [18, 21, 28, 33–45]. Size-polymorphic antigenic markers, such as *msp-1*, *msp-2*, *msp-3*, *glurp*, *csp* and *ama-1* in *Plasmodium falciparum* and *P. vivax*, can be amplified by PCR, and the size of the amplicon is determined by either gel or capillary electrophoresis. Protocols are comparably simple without requiring expensive equipment. In many populations, high diversity of size-polymorphic antigens has been observed in both *P. falciparum* [33, 34, 38, 43, 46–53] and *P. vivax* [54–63]. Antigens are likely under immune selection, which will affect the distribution of alleles in the population. While in most cases this will not affect estimates of MOI, antigen markers should not be included in studies assessing parasite population structure.

Microsatellite markers are highly polymorphic with tandem repeats of 2–6 bp. They are expected to be selection neutral with no phenotypic consequences [64–67]. Microsatellites are widely distributed throughout the genome and can be readily amplified by PCR [68]. A range of 5–14 microsatellite markers have been used to estimate MOI in *P. falciparum* [53, 69–71], *P. vivax* [11, 72–74], and *Plasmodium malariae* [75, 76]. After evaluating 42 microsatellite markers in *P. vivax* from 18 different global studies, Sutton et al. found that five (*3.27*, *8.504*, *PvMS11*, *MS16*, and *MS8*) markers could be used as candidates for MOI studies [77]. However, as stutter peaks and nonspecific amplification in the case of suboptimal PCR conditions have been frequently observed [78–83], the recommendations of markers also need to take into account the robustness of the PCR protocols. High levels of stutter peaks have been observed in different laboratories for markers Pv3.27, MS8, and MS16

(Koepfli, unpublished), thus these markers are no longer recommended. In *P. falciparum*, there are 12 microsatellite markers (*Polyα*, *TA42*, *TA81*, *TA1*, *TA109*, *TA87*, *TA40*, *2490*, *ARAII*, *pfG377*, *PfPk2* and *TA60*) from 7 chromosomes that are frequently used for genetic diversity and multiplicity of infection studies [35, 65, 84]. MOI is determined as number of alleles detected, and if several markers are combined, as the maximum number observed for any marker. Any highly polymorphic marker is suitable for estimating the MOI.

The major limitation of agarose gel electrophoresis is its limited discrimination of alleles of similar sizes. It is unable to discriminate allele size difference less than 20 bp [85], resulting in underestimates of diversity and multiplicity. This problem is more pronounced when using microsatellites with many alleles differing by as little as 2–6 bp, as compared to antigens, that often harbor larger repeat sizes. Also, it is difficult to detect minority clones that result in weak bands on gels. Polyacrylamide gels have a better discrimination power. Even better results are obtained by using primers labeled with fluorophores and sizing the PCR products by capillary electrophoresis, which can resolve sizes of 2 bp. The second major problem is that PCR preferentially amplifies shorter fragments [86], and artifacts, such as chimera, emerge during multi-template PCR [87]. Furthermore, irrespective of markers used, PCR can lead to amplification from nonspecific template and produce PCR artifacts differing in length from the main products [80]. To avoid non-specific amplification due to mispriming and/or undesirable interactions between primers, the PCR conditions should be more stringent such as lowest possible $MgCl_2$ concentration and the minimum number of amplification cycles.

High-throughput single nucleotide polymorphism (SNP) genotyping

As alternatives to size-polymorphic markers, several panels of genome-wide SNP makers have been identified and protocols for typing developed. One panel includes 24 *P. falciparum* SNPs selected from over 112,000 SNPs of 18 parasite genomes [88]. Typing is based on High Resolution Melting (HRM) assays, which can detect parasite with a minor allele frequency higher than 10–20%, similar to that of traditional length polymorphic markers *msp-1* and *msp-2* [86], but the SNP-based molecular barcode provides greater power to discriminate among strains than *msp-1* and *msp-2* genotyping because many more possible alleles exist among the 24 markers than length polymorphisms within these regions. In *P. vivax*, a panel of 42-SNP barcodes was developed based on clinical samples from parasite populations in South America (Brazil, French Guiana), Africa (Ethiopia) and Asia (Sri Lanka) [89], and typed also using HRM assays. The

Zhong *et al. Malar J* (2018) 17:172

Page 3 of 9

small standardized set or subset of the 42-SNP barcode not only can be used for analyses of genetic diversity and population differentiation, but also can be used for MOI analyses though calculation of minor allele frequency [90]. HRM genotyping analyses provide a robust, informative, and relatively low-cost method of identifying parasite infections. However, a large amount of template DNA is required, as typically a separate reaction for each SNP needs to be run, and typing of low-density infections (e.g. from asymptomatic carriers) is challenging.

High-throughput SNP genotyping can also be done on the Sequenom MassARRAY iPLEX platform, which is capable of multiplexing up to 40 SNPs per single reaction. SNP allele is determined by variations of MALDI-TOF mass spectrometry at SNP site [91]. The parasite genotype is determined by the SNP allelic intensity ratios, ranging between 0 and 1. The SNP allelic intensity ratios value nearing 0 and 1 indicate single parasite genotype infection, while intermediate values indicate mixed parasite genotype infections [92]. Galinsky et al. described a method for estimating MOI based on the data obtained from Sequenom platform using likelihood and developed the software package COIL as implementation [93]. COIL assumes that distinct parasite lineages in complex infections are unrelated and that genotyped loci do not exhibit significant linkage disequilibrium, which is very suitable to employ SNP barcodes for MOI estimation. Similarly, the THE REAL McCOIL package was developed, which does not require external allele frequency data [94]. However, although this approach can be used for inferring the number of strains (COIL), or simultaneously inferring the number of strains and their proportions, it does not infer haplotypes.

Large-scale SNP genotyping using a custom 384-SNP Illumina GoldenGate has been conducted in *P. falciparum* from laboratory strains and clinical samples from East and West Africa, Southeast Asia, and Oceania [95]. This genotyping assay has a high genotype calling specificity from artificial mixtures of laboratory clones with a moderate sensitivity to call minor frequency alleles. The advantage of these methods is the large number of SNPs typed simultaneously. It thus can provide a practical, rapid, robust, and inexpensive assay of genome-wide parasite genotyping with easy data interpretation. The GoldenGate platform relies on monoclonal infections to estimate allele frequencies, which is problematic when a large fraction of infections are polyclonal. While most SNPs are bi-allelic, in theory up to four alleles are possible. Effective computational tools and advanced bioinformatic analysis are required to distinguish more than two parasite clones within a host [94].

Nkhoma et al. used a set of 96 genome-wide SNPs and Illumina GoldenGate to examine changes in parasite population genetic parameters associated with decline in transmission in *P. falciparum* [96]. Polyclonal infection was determined by heterozygous base calls at >5% of the genotyped SNPs with an error rate of 4.13×10^{-4}. The clonal composition of parasite populations was measured by genotypic richness index (R) and the β-parameter of the Pareto distribution. R measures the proportion of unique genotypes (G) in the population and is estimated as: $R = (G-1)/(n-1)$, which varies from 0 when all n samples in a population have the same genotype, to 1.0, when all samples have a different genotype [97]. β-Parameter measures the slope of Power law (Pareto) distribution by resampling different numbers of samples (e.g. n = 100–1100) from the complete data set and recalculating genotypic data. The slope is highly skewed with a large number of rare genotypes and a few common ones [98]. A minimum of 25 SNP markers are required to exhaustively identify all distinct multilocus parasite genotypes present in infections [96]. However, the proportion of multiple-genotype infections is dependent on the sensitivity of genotyping methods used and is therefore difficult to compare between studies. Genotype richness and the β-parameter of the Pareto distribution can only be measured using single genotype infections.

Amplicon ultra-deep sequencing

With the advent of next-generation sequencing (NGS) technology, studies on population genetics, diversity, and multiplicity of *Plasmodium* infections have entered a new era. NGS is becoming the standard approach to generate population genomic data. Great progress has been made for NGS data analysis and MOI estimation with the development of various bioinformatic tools (Table 1). In highly polymorphic amplicons, several SNPs are concentrated within a locus of 100–200 bp. Amplicon ultra-deep sequencing of high polymorphic makers has high sensitivity and specificity to detect minority clones in multiclonal infections. A number of highly polymorphic amplicons have been identified [99, 100]. Using *pvmsp1* short amplicon (117 bp) deep sequencing, Lin et al. reported 67 unique haplotypes identified from 78 Cambodian *P. vivax* samples with an average of 3.6 MOI within each individual [100]. To increase the sensitivity of detecting minor alleles, three *P. falciparum* amplicons of 321 bp (*pf-csp*), 305 bp (*pf-ama1*), and 306 bp (*pf-k13*) have been used to conduct ultra-deep sequencing, which can quantitatively detect unique haplotypes comprising as little as 2% of a polyclonal infection [101].

Recently, the conserved *Plasmodium* membrane protein (*cpmp*) was described as a highly polymorphic marker, and an assay was developed with 100% sensitivity to detect all clones present at a frequency of >1% [102]. Ultra-deep sequencing of amplicons from ribosomal

Zhong *et al. Malar J* (2018) 17:172

Page 4 of 9

Table 1 The advantages and limitations of bioinformatics tools current available for analysis of polyclonal infections and estimation of MOI

Software/approach	Implementation	Data source	Advantage	Limitation	References
Fws	Fws statistics	SNPs	Provide a summary statistic of relative inbreeding and correlation with MOI	Does not provide a direct estimate of COI	Manske et al. [108]
estMOI	Perl scripts	TDS	estimate MOI, SNP combinations and proportion in a specific genomic region	Partial genomic region, sufficient density of SNPs in an amplicon is required	Assefa et al. [107]
COIL	Perl scripts web inter-face	SNPs	Use genome-wide SNPs to infer the number of strains and their proportions	Does not infer haplotypes; sequence read counts are assumed to be unbiased and the SNPs are unlinked	Galinsky et al. [93]
SeekDeep	Perl scripts	TDS	No reference sequence required, direct estimation of amplicon haplotype cluster and frequency	Partial genomic region, threshold for haplotype cluster calls needs to be determined by empirical methods in each study	Hathaway et al. [104]
HaplotypR	R package	TDS	Direct estimation of haplotype cluster and frequency based quality filtering of sequencing reads	Partial genomic region, enough high polymorphism in the amplicon is required and unable to deal with indel of amplicon	Lerch et al. [102]
pfmix	R package	SNPs	Infer inbreeding coefficients using WGS to assess within-sample parasite infections	Does not infer haplotypes, sequence read counts are assumed to be unbiased and SNPs need to be unlinked	O'Brien et al. [110]
THE REAL McCOIL	Perl scripts web inter-face	SNPs	Infer the number of strains and their proportions	Does not infer haplotypes, sequence read counts are assumed to be unbiased and SNPs need to be unlinked	Chang et al. [94]
DEploid	R package	SNPs	Estimate the number of strains, their relative propor-tions, and the haplotypes	Reference panel required, more reference strains are better, Multiallelic variant not considered	Zhu et al. [111]

SNPs single nucleotide polymorphism, *TDS* targeted deep sequencing, *WGS* whole genome sequencing, *MOI* multiplicity of infection

Zhong *et al. Malar J* (2018) 17:172

Page 5 of 9

genes and genes encoded by the mitochondrial and api-coplast genomes has been used to evaluate *Plasmodium* species diversity and MOI in symptomatic Gabonese patients [103]. It is, however, challenging to distinguish PCR and sequencing errors from true minority haplotypes. Recently, multiple packages have been developed for de novo clustering of haplotypes from this type of data allowing detection of low frequency variants, such as HaplotypR [102] and SeekDeep [104].

Several studies have evaluated long amplicon deep sequencing to examine multiplicity of infection by sequencing mixtures of multiple lab strains. For example, Patel et al. performed amplicon deep sequencing of 1.6 kb ID1-DBL2× region of *var2csa* gene in *P. falciparum* infected placental samples and found that parasite haplotypes can be recovered if they account for as little as 5% of the mixed template [105]. A long amplicon (~5 kb) of *var2csa* gene in *P. falciparum* was sequenced by PacBio deep sequencing and used to evaluate artificially created a mixture of laboratory strains [106]. De novo assembly of full-length amplicons was used to reconstruct *var2csa* haplotypes. It however identified minor alleles only if present at frequencies > 23% in the mixture, due to low read depth and coverage and high sequencing errors in homopolymer stretches [106]. In order to obtain highly accurate haplotype estimation from polyclonal clinical samples, Lerch et al. suggest that a minimum coverage of 10,000 high-quality sequencing reads with duplicate experiment is required to detect minority clones at 0.1% frequency [102]. The major advantage of single amplicon deep sequencing over a panel of genome-wide SNPs is that all SNPs occur within one amplicon. Thus, haplotypes can be directly identified without the need for multi-locus haplotype reconstruction. Software such as SeekDeep aids in the identification of haplotypes [100, 104]. After filtering of sequencing reads according to their base quality scores, they are clustered using a clustering algorithm based on k-mer distances. In contrast to microsatellites and SNP-panels, the haplotype frequency can be directly estimated from read counts with high accuracy. The main limitation is that only a small genomic region is amplified. The threshold for accurate genotype calls may be different between studies due to various sequencing error rate in different sequencing platforms and computational strategies. In order to exclude PCR or sequencing errors, it is recommended to perform experiments in duplicate and use appropriate controls with each study to help determine that no false calls are being made.

Whole genome sequencing

Whole genome sequencing of field isolates is rapidly becoming common, and a number of software packages have been developed to identify multiclonal infections and reconstruct haplotypes. The estMOI software package can be used to analyse deep sequencing data for estimation of the presence of multiple genotypes, the number of infections, as well as proportions by phasing information of each allele [107]. In other words, the software can be used to estimate minor allele (> 10% frequency) in multiclonal infections. The limitation of this software is that MOI is estimated by maximum number of haplotype in a specific genome region with a high coverage (> 30-fold), which is costly to generate at large scales by whole genome sequencing. In order to reduce cost, a list of 26 MOI informative genes with high levels of polymorphism has been recommended to use for genotyping to estimate multiplicity in clinical *P. falciparum* samples [107].

The F_{WS} (similar to Wright's inbreeding coefficient F_{IS}) statistics derived from read count data of whole genome sequencing at 86,158 SNP positions had been used to explore the number of within-host clones, their relative proportions and genetic divergence in *P. falciparum* [108, 109]. This approach is conducted by estimating levels of heterozygosity both within an individual sample (Hw) and within the local parasite population (Hs) using genome-wide SNP data. The F_{WS} values, where $F_{WS} = 1 - H_W/H_S$, observed in blood samples provide a proxy indicator of inbreeding rates in the parasite population. The F_{WS} ranges from 0 to 1, with 0 reflecting a mixture of highly unrelated clones and 1 reflecting a single clone. In other words, a low F_{WS} reflects a low risk of inbreeding/high risk of out-crossing and thus high within-host diversity. A threshold of $F_{WS} \geq 0.95$ was suggested as indication of an essentially clonal infection. A significant negative correlation was observed between the PCR-based *msp1* genotyping and F_{WS} statistics for calculating MOI. Although the F_{WS} can capture information of within-host heterozygosity relative to population diversity, it does not provide a direct estimate of MOI. To help make a connection of F_{WS} with MOI estimates, O'Brien et al. presented several new approaches for inferring inbreeding coefficients using read counts from WGS, which can be implemented in the R package [110].

Zhu et al. developed a software package named *DEploid*, which can be used to analyse genome-wide SNP data for estimation of the number of strains, their relative proportions and the haplotypes present in a sample

Zhong *et al. Malar J* (2018) 17:172

Page 6 of 9

[111]. The software package uses haplotype structure within a reference panel of clonal isolates as a prior for haplotypes present in a multiclonal sample. The effective number of strains can be inferred without an assumption of the number of strains or the use of linkage disequilibrium information. The assumed haplotypes are inferred using read counts of the reference strain and alternative alleles. First, allele counts are extracted from the vcf file, and then the population allele frequencies of each allele are calculated from total read counts. Finally, inference of correct haplotypes and relative proportions can be completed by quality filtering of sequencing reads. Overall, this method can be used to infer mixed strains with proportions at > 20% with high accuracy, but it struggles with minor strains due to insufficient read coverage and if the minor strain carries the alternative allele. The computer program has been tested and confirmed to work well for the deconvolution of datasets with a mixture of samples from a single species in *P. falciparum* [111]. It is able to perform deconvolution of 372,884 variant sites of malaria parasites within one sample in just five and half hours. *DEploid* is the only software currently available that can reconstruct haplotypes as well as estimate COI using genome wide SNP data. The limitation is that it requires a panel of references, which should include enough different reference strains to cover all the haplotype structure representing the field population. However, this is not realistic for current computational technology.

Finally, sequencing coverage remains a main challenge when estimating MOI from whole genome sequencing data. By amplicon deep sequencing, a read depth of $1000–10,000\times$ coverage is easily obtained, resulting in a significantly increased power of detection of minority clones. At this sequencing depth, minority clones at a frequency of 0.1% have been successfully identified in experimental mixtures of strains [102]. A typical sequencing run on an Illumina NextSeq sequencer yields approximately 130 million reads of 150 base pairs each. Even if a single isolate was sequenced by a single run, and if even coverage across the genome of 23 million base pairs was obtained, coverage would be below $900\times$. Factoring in that field isolates always contain some human DNA, coverage would be even lower. Thus, when MOI is the main interest of a study, whole genome sequencing is not a cost-effective method at the current stage.

Conclusions

Traditional standard PCR-based genotyping methods using *msp-1*, *msp-2*, and microsatellite markers often underestimate the MOI. Next-generation amplicon sequencing technology has largely increased the sensitivity in detection of polyclonal infections, but targets only partial genomic regions, which may not represent complete polymorphism in mixed infections. Whole genome sequencing yields complete information for parasite infections, but requires extensive and efficient computation algorithms and mathematical models to reconstruct the complete haplotypes, and sensitivity to detect minority clones is limited due to low reads coverage for minor strains. Therefore, there is a need to develop an accurate and cost-effective method for detection of minority clones at the current stage. As the cost of whole genome sequencing continues to decrease and bioinformatics tools mature, fine-scale investigation of multiplicity of Plasmodium infections becomes possible in the near future.

Abbreviations
COI: complexity of infection; COIL: complexity of infection using likelihood; HRM: high resolution melting; LOD: limit of detection for minority haplotype; MOI: multiplicity of infection; SNPs: single nucleotide polymorphism; TDS: targeted deep sequencing; WGS: whole genome sequencing.

Authors' contributions
DZ and CK conducted the literature review and drafted the manuscript. CL participated in revising the draft manuscript. GY conceived of the study, designed the study, and participated in revising the draft manuscript. All authors read and approved the final manuscript.

Author details
[1] Program in Public Health, University of California, Irvine, CA 92617, USA. [2] Department of Entomology, Pennsylvania State University, University Park, PA 16802, USA.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
Data generated or analysed in this study are included within the article.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Kolakovich KA, Ssengoba A, Wojcik K, Tsuboi T, al-Yaman F, Alpers M, et al. *Plasmodium vivax*: favored gene frequencies of the merozoite surface protein-1 and the multiplicity of infection in a malaria endemic region. Exp Parasitol. 1996;83:11–9.

Zhong *et al. Malar J* (2018) 17:172

Page 7 of 9

2. Conway DJ, Roper C, Oduola AMJ, Arnot DE, Kremsner PG, Grobusch MP, et al. High recombination rate in natural populations of *Plasmodium falciparum*. Proc Natl Acad Sci USA. 1999;96:4506–11.

3. Fola AA, Harrison GLA, Hazairin MH, Barnadas C, Hetzel MW, Iga J, et al. Higher complexity of infection and genetic diversity of *Plasmodium vivax* than *Plasmodium falciparum* across all malaria transmission zones of Papua New Guinea. Am J Trop Med Hyg. 2017;96:630–41.

4. Arnot D. Unstable malaria in Sudan: the influence of the dry season. Clone multiplicity of *Plasmodium falciparum* infections in individuals exposed to variable levels of disease transmission. Trans R Soc Trop Med Hyg. 1998;92:580–5.

5. Noviyanti R, Coutrier F, Utami RA, Trimarsanto H, Tirta YK, Trianty L, et al. Contrasting transmission dynamics of co-endemic *Plasmodium vivax* and *P. falciparum*: implications for malaria control and elimination. PLoS Negl Trop Dis. 2015;9:e0003739.

6. Getachew S, To S, Trimarsanto H, Thriemer K, Clark TG, Petros B, et al. Variation in complexity of infection and transmission stability between neighbouring populations of *Plasmodium vivax* in southern Ethiopia. PLoS ONE. 2015;10:e0140780.

7. Lee SA, Yeka A, Nsobya SL, Dokomajilar C, Rosenthal PJ, Talisuna A, et al. Complexity of *Plasmodium falciparum* infections and antimalarial drug efficacy at 7 sites in Uganda. J Infect Dis. 2006;193:1160–3.

8. Paganotti GM, Babiker HA, Modiano D, Sirima BS, Verra F, Konate A, et al. Genetic complexity of *Plasmodium falciparum* in two ethnic groups of Burkina Faso with marked differences in susceptibility to malaria. Am J Trop Med Hyg. 2004;71:173–8.

9. Talisuna AO, Langi P, Mutabingwa TK, Van Marck E, Speybroeck N, Egwang TG, et al. Intensity of transmission and spread of gene mutations linked to chloroquine and sulphadoxine–pyrimethamine resistance in falciparum malaria. Int J Parasitol. 2003;33:1051–8.

10. Pacheco MA, Lopez-Perez M, Vallejo AF, Herrera S, Arevalo-Herrera M, Escalante AA. Multiplicity of infection and disease severity in *Plasmodium vivax*. PLoS Negl Trop Dis. 2016;10:e0004355.

11. Pava Z, Handayuni I, Trianty L, Utami RAS, Tirta YK, Puspitasari AM, et al. Passively versus actively detected malaria: similar genetic diversity but different complexity of infection. Am J Trop Med Hyg. 2017;97:1788–96.

12. Beck S, Mockenhaupt FP, Bienzle U, Eggelte TA, Thompson WN, Stark K. Multiplicity of *Plasmodium falciparum* infection in pregnancy. Am J Trop Med Hyg. 2001;65:631–6.

13. Nicastri E, Paglia MG, Severini C, Ghirga P, Bevilacqua N. Disease severity, host characteristics and *Plasmodium falciparum* multiple infections in malaria affected travellers returning from Africa. Travel Med Infect Dis. 2007;5:405–6.

14. Kun JF, Schmidt-Ott RJ, Lehman LG, Lell B, Luckner D, Greve B, et al. Merozoite surface antigen 1 and 2 genotypes and rosetting of *Plasmodium falciparum* in severe and mild malaria in Lambarene, Gabon. Trans R Soc Trop Med Hyg. 1998;92:110–4.

15. Nicastri E, Paglia MG, Severini C, Ghirga P, Bevilacqua N, Narciso P. *Plasmodium falciparum* multiple infections, disease severity and host characteristics in malaria affected travellers returning from Africa. Travel Med Infect Dis. 2008;6:205–9.

16. Ranjit MR, Das A, Das BP, Das BN, Dash BP, Chhotray GP. Distribution of *Plasmodium falciparum* genotypes in clinically mild and severe malaria cases in Orissa, India. Trans R Soc Trop Med Hyg. 2005;99:389–95.

17. Rout R, Mohapatra BN, Kar SK, Ranjit M. Genetic complexity and transmissibility of *Plasmodium falciparum* parasites causing severe malaria in central-east coast India. Trop Biomed. 2009;26:165–72.

18. Kiwuwa MS, Ribacke U, Moll K, Byarugaba J, Lundblom K, Farnert A, et al. Genetic diversity of *Plasmodium falciparum* infections in mild and severe malaria of children from Kampala, Uganda. Parasitol Res. 2013;112:1691–700.

19. Mara SE, Silué KD, Raso G, N'Guetta SP, N'Goran EK, Tanner M, et al. Genetic diversity of *Plasmodium falciparum* among school-aged children from the Man region, western Côte d'Ivoire. Malar J. 2013;12:419.

20. Conway DJ, Greenwood BM, McBride JS. The epidemiology of multiple-clone *Plasmodium falciparum* infections in Gambian patients. Parasitology. 2009;103:1–5.

21. Durand R, Ariey F, Cojean S, Fontanet A, Ranaivo L, Ranarivelo LA, et al. Analysis of circulating populations of *Plasmodium falciparum* in mild and severe malaria in two different epidemiological patterns in Madagascar. Trop Med Int Health. 2008;13:1392–9.

22. al-Yaman F, Genton B, Reeder JC, Anders RF, Smith T, Alpers MP. Reduced risk of clinical malaria in children infected with multiple clones of *Plasmodium falciparum* in a highly endemic area: a prospective community study. Trans R Soc Trop Med Hyg. 1997;91:602–5.

23. Bereczky S, Liljander A, Rooth I, Faraja L, Granath F, Montgomery SM, et al. Multiclonal asymptomatic *Plasmodium falciparum* infections predict a reduced risk of malaria disease in a Tanzanian population. Microbes Infect. 2007;9:103–10.

24. Nielsen MA, Staalsoe T, Kurtzhals JA, Goka BQ, Dodoo D, Alifrangis M, et al. *Plasmodium falciparum* variant surface antigen expression varies between isolates causing severe and nonsevere malaria and is modified by acquired immunity. J Immunol. 2002;168:3444–50.

25. Amodu OK, Oyedeji SI, Ntoumi F, Orimadegun AE, Gbadegesin RA, Olumese PE, et al. Complexity of the *msp2* locus and the severity of childhood malaria, in south-western Nigeria. Ann Trop Med Parasitol. 2008;102:95–102.

26. Legutki JB, Johnston SA. Immunosignatures can predict vaccine efficacy. Proc Natl Acad Sci USA. 2013;110:18614–9.

27. Newbold CI, Schryer M, Boyle DB, McBride JS, McLean A, Wilson RJM, et al. A possible molecular basis for strain specific immunity to malaria. Mol Biochem Parasitol. 1984;11:337–47.

28. Patel P, Bharti PK, Bansal D, Raman RK, Mohapatra PK, Sehgal R, et al. Genetic diversity and antibody responses against *Plasmodium falciparum* vaccine candidate genes from Chhattisgarh, Central India: implication for vaccine development. PLoS ONE. 2017;12:e0182674.

29. Neafsey DE, Juraska M, Bedford T, Benkeser D, Valim C, Griggs A, et al. Genetic diversity and protective efficacy of the RTS, S/AS01 malaria vaccine. N Engl J Med. 2015;373:2025–37.

30. Takala SL, Plowe CV. Genetic diversity and malaria vaccine design, testing, and efficacy: preventing and overcoming "vaccine resistant malaria". Parasite Immunol. 2009;31:560–73.

31. Mueller I, Schoepflin S, Smith TA, Benton KL, Bretscher MT, Lin E, et al. Force of infection is key to understanding the epidemiology of *Plasmodium falciparum* malaria in Papua New Guinean children. Proc Natl Acad Sci USA. 2012;109:10030–5.

32. Koepfli C, Colborn KL, Kiniboro B, Lin E, Speed TP, Siba PM, et al. A high force of *Plasmodium vivax* blood-stage infection drives the rapid acquisition of immunity in papua new guinean children. PLoS Negl Trop Dis. 2013;7:e2403.

33. Atroosh WM, Al-Mekhlafi HM, Mahdy MA, Saif-Ali R, Al-Mekhlafi AM, Surin J. Genetic diversity of *Plasmodium falciparum* isolates from Pahang, Malaysia based on MSP-1 and MSP-2 genes. Parasit Vectors. 2011;4:233.

34. Bin Dajem SM. Genetic diversity in the msp-1 and msp-2 alleles among *Plasmodium falciparum* field isolates from Jazan, Saudi Arabia. J Egypt Soc Parasitol. 2015;45:199–206.

35. Mohd Abd Razak MR, Sastu UR, Norahmad NA, Abdul-Karim A, Muhammad A, Muniandy PK, et al. Genetic diversity of *Plasmodium falciparum* populations in malaria declining areas of Sabah, East Malaysia. PLoS ONE. 2016;11:e0152415.

36. Niang M, Loucoubar C, Sow A, Diagne MM, Faye O, Faye O, et al. Genetic diversity of *Plasmodium falciparum* isolates from concurrent malaria and arbovirus co-infections in Kedougou, southeastern Senegal. Malar J. 2016;15:155.

37. Niang M, Thiam LG, Loucoubar C, Sow A, Sadio BD, Diallo M, et al. Spatio-temporal analysis of the genetic diversity and complexity of *Plasmodium falciparum* infections in Kedougou, southeastern Senegal. Parasit Vectors. 2017;10:33.

38. Ogouyemi-Hounto A, Gazard DK, Ndam N, Topanou E, Garba O, Elegbe P, et al. Genetic polymorphism of merozoite surface protein-1 and merozoite surface protein-2 in *Plasmodium falciparum* isolates from children in South of Benin. Parasite. 2013;20:37.

39. Oyebola MK, Idowu ET, Olukosi YA, Iwalokun BA, Agomo CO, Ajibaye OO, et al. Genetic diversity and complexity of *Plasmodium falciparum* infections in Lagos, Nigeria. Asian Pac J Trop Biomed. 2014;4:S87–91.

40. Etoka-Beka MK, Ntoumi F, Kombo M, Deibert J, Poulain P, Vouvoungui C, et al. *Plasmodium falciparum* infection in febrile Congolese children: prevalence of clinical malaria 10 years after introduction of artemisinin-combination therapies. Trop Med Int Health. 2016;21:1496–503.

Zhong *et al. Malar J* (2018) 17:172

Page 8 of 9

41. Kateera F, Nsobya SL, Tukwasibwe S, Mens PF, Hakizimana E, Grobusch MP, et al. Malaria case clinical profiles and *Plasmodium falciparum* parasite genetic diversity: a cross sectional survey at two sites of different malaria transmission intensities in Rwanda. Malar J. 2016;15:237.

42. Amoah LE, Nuvor SV, Obboh EK, Acquah FK, Asare K, Singh SK, et al. Natural antibody responses to *Plasmodium falciparum* MSP3 and GLURP(R0) antigens are associated with low parasite densities in malaria patients living in the Central Region of Ghana. Parasit Vectors. 2017;10:395.

43. Kumar D, Dhiman S, Rabha B, Goswami D, Deka M, Singh L, et al. Genetic polymorphism and amino acid sequence variation in *Plasmodium falciparum* GLURP R2 repeat region in Assam, India, at an interval of 5 years. Malar J. 2014;13:450.

44. Muhindo Mavoko H, Kalabuanga M, Delgado-Ratto C, Maketa V, Mukele R, Fungula B, et al. Uncomplicated clinical malaria features, the efficacy of artesunate-amodiaquine and their relation with multiplicity of infection in the Democratic Republic of Congo. PLoS ONE. 2016;11:e0157074.

45. Soe TN, Wu Y, Tun MW, Xu X, Hu Y, Ruan Y, et al. Genetic diversity of *Plasmodium falciparum* populations in southeast and western Myanmar. Parasit Vectors. 2017;10:322.

46. Bhattacharya PR. *Plasmodium falciparum*: genetic polymorphism of the merozoite surface antigen 2 gene of strains from India. Trans R Soc Trop Med Hyg. 1998;92:225–6.

47. Escalante AA, Grebert HM, Chaiyaroj SC, Magris M, Biswas S, Nahlen BL, et al. Polymorphism in the gene encoding the apical membrane antigen-1 (AMA-1) of *Plasmodium falciparum*. X. Asembo Bay Cohort project. Mol Biochem Parasitol. 2001;113:279–87.

48. Felger I, Tavul L, Kabintik S, Marshall V, Genton B, Alpers M, et al. *Plasmodium falciparum*: extensive polymorphism in merozoite surface antigen 2 alleles in an area with endemic malaria in Papua New Guinea. Exp Parasitol. 1994;79:106–16.

49. Quang ND, Hoa PT, Tuan MS, Viet NX, Jalloh A, Matsuoka H. Polymorphism at the apical membrane antigen 1 gene (AMA1) of the malaria parasite *Plasmodium falciparum* in a Vietnamese population. Biochem Genet. 2009;47:370–83.

50. Rajesh V, Singamsetti VK, Vidya S, Gowrishankar M, Elamaran M, Tripathi J, et al. *Plasmodium falciparum*: genetic polymorphism in apical membrane antigen-1 gene from Indian isolates. Exp Parasitol. 2008;119:144–51.

51. Verra F, Hughes AL. Evidence for ancient balanced polymorphism at the Apical Membrane Antigen-1 (AMA-1) locus of *Plasmodium falciparum*. Mol Biochem Parasitol. 2000;105:149–53.

52. Zhou YF, Zhang SY, Lin YY, Yang FZ, Xie HG, Xiao FZ. [Genetic polymorphism of the gene encoding the apical membrane antigen-1 of *Plasmodium falciparum*](in Chinese). Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 2014;32:344–7.

53. Zhong D, Afrane Y, Githeko A, Yang Z, Cui L, Menge DM, et al. *Plasmodium falciparum* genetic diversity in western Kenya highlands. Am J Trop Med Hyg. 2007;77:1043–50.

54. Grynberg P, Fontes CJ, Hughes AL, Braga EM. Polymorphism at the apical membrane antigen 1 locus reflects the world population history of *Plasmodium vivax*. BMC Evol Biol. 2008;8:123.

55. Moon SU, Na BK, Kang JM, Kim JY, Cho SH, Park YK, et al. Genetic polymorphism and effect of natural selection at domain I of apical membrane antigen-1 (AMA-1) in *Plasmodium vivax* isolates from Myanmar. Acta Trop. 2010;114:71–5.

56. Putaporntip C, Jongwutiwes S, Grynberg P, Cui L, Hughes AL. Nucleotide sequence polymorphism at the apical membrane antigen-1 locus reveals population history of *Plasmodium vivax* in Thailand. Infect Genet Evol. 2009;9:1295–300.

57. Thakur A, Alam MT, Bora H, Kaur P, Sharma YD. *Plasmodium vivax*: sequence polymorphism and effect of natural selection at apical membrane antigen 1 (*PvAMA1*) among Indian population. Gene. 2008;419:35–42.

58. Yan J, Zhang S, Cu W. [Polymorphism of circumsporozoite protein and merozoite surface antigen of *Plasmodium vivax*](in Chinese). Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi. 1998;16:138–42.

59. Zakeri S, Sadeghi H, Mehrizi AA, Djadid ND. Population genetic structure and polymorphism analysis of gene encoding apical membrane antigen-1 (AMA-1) of Iranian *Plasmodium vivax* wild isolates. Acta Trop. 2013;126:269–79.

60. Zhong D, Bonizzoni M, Zhou G, Wang G, Chen B, Vardo-Zalik A, et al. Genetic diversity of *Plasmodium vivax* malaria in China and Myanmar. Infect Genet Evol. 2011;11:1419–25.

61. Soares LA, Evangelista J, Orlandi PP, Almeida ME, de Sousa LP, Chaves Y, et al. Genetic diversity of MSP1 Block 2 of *Plasmodium vivax* isolates from Manaus (central Brazilian Amazon). J Immunol Res. 2014;2014:671050.

62. Karunaweera ND, Ferreira MU, Hartl DL, Wirth DF. Fourteen polymorphic microsatellite DNA markers for the human malaria parasite *Plasmodium vivax*. Mol Ecol Notes. 2007;7:172–5.

63. Koepfli C, Mueller I, Marfurt J, Goroti M, Sie A, Oa O, et al. Evaluation of *Plasmodium vivax* genotyping markers for molecular monitoring in clinical trials. J Infect Dis. 2009;199:1074–80.

64. Alam MT, Agarwal R, Sharma YD. Extensive heterozygosity at four microsatellite loci flanking *Plasmodium vivax* dihydrofolate reductase gene. Mol Biochem Parasitol. 2007;153:178–85.

65. Anderson TJ, Su XZ, Bockarie M, Lagog M, Day KP. Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. Parasitology. 1999;119(Pt 2):113–25.

66. de Souza AM, de Araujo FC, Fontes CJ, Carvalho LH, de Brito CF, de Sousa TN. Multiple-clone infections of *Plasmodium vivax*: definition of a panel of markers for molecular epidemiology. Malar J. 2015;14:330.

67. Schlötterer C. Genome evolution: are microsatellites really simple sequences? Curr Biol. 1998;8:R132–4.

68. Su X, Wellems TE. Toward a high-resolution *Plasmodium falciparum* linkage map: polymorphic markers from hundreds of simple sequence repeats. Genomics. 1996;33:430–44.

69. Mwangi JM, Omar SA, Ranford-Cartwright LC. Comparison of microsatellite and antigen-coding loci for differentiating recrudescing *Plasmodium falciparum* infections from reinfections in Kenya. Int J Parasitol. 2006;36:329–36.

70. Sutton PL, Torres LP, Branch OH. Sexual recombination is a signature of a persisting malaria epidemic in Peru. Malar J. 2011;10:329.

71. Morlais I, Nsango SE, Toussile W, Abate L, Annan Z, Tchioffo MT, et al. *Plasmodium falciparum* mating patterns and mosquito infectivity of natural isolates of gametocytes. PLoS ONE. 2015;10:e0123777.

72. Koepfli C, Ross A, Kiniboro B, Smith TA, Zimmerman PA, Siba P, et al. Multiplicity and diversity of *Plasmodium vivax* infections in a highly endemic region in Papua New Guinea. PLoS Negl Trop Dis. 2011;5:e1424.

73. Delgado-Ratto C, Gamboa D, Soto-Calle VE, Van den Eede P, Torres E, Sanchez-Martinez L, et al. Population genetics of *Plasmodium vivax* in the Peruvian Amazon. PLoS Negl Trop Dis. 2016;10:e0004376.

74. Hong NV, Delgado-Ratto C, Thanh PV, Van den Eede P, Guetens P, Binh NT, et al. Population genetics of *Plasmodium vivax* in four rural communities in central Vietnam. PLoS Negl Trop Dis. 2016;10:e0004434.

75. Bruce MC, Macheso A, Galinski MR, Barnwell JW. Characterization and application of multiple genetic markers for *Plasmodium malariae*. Parasitology. 2007;134:637–50.

76. Bruce MC, Macheso A, McConnachie A, Molyneux ME. Comparative population structure of Plasmodium malariae and *Plasmodium falciparum* under different transmission settings in Malawi. Malar J. 2011;10:38.

77. Sutton PL. A call to arms: on refining *Plasmodium vivax* microsatellite marker panels for comparing global diversity. Malar J. 2013;12:447.

78. Fernando P, Evans BJ, Morales JC, Melnick DJ. Electrophoresis artefacts—a previously unrecognized cause of error in microsatellite analysis. Mol Ecol Notes. 2001;1:325–8.

79. Manrique P, Hoshi M, Fasabi M, Nolasco O, Yori P, Calderón M, et al. Assessment of an automated capillary system for *Plasmodium vivax* microsatellite genotyping. Malar J. 2015;14:326.

80. Hosseinzadeh-Colagar A, Haghighatnia MJ, Amiri Z, Mohadjerani M, Tafrihi M. Microsatellite (SSR) amplification by PCR usually led to polymorphic bands: evidence which shows replication slippage occurs in extend or nascent DNA strands. Mol Biol Res Commun. 2016;5:167–74.

81. Havryliuk T, Ferreira MU. A closer look at multiple-clone *Plasmodium vivax* infections: detection methods, prevalence and consequences. Mem Inst Oswaldo Cruz. 2009;104:67–73.

82. Greenhouse B, Myrick A, Dokomajilar C, Woo JM, Carlson EJ, Rosenthal PJ, et al. Validation of microsatellite markers for use in genotyping

Zhong *et al. Malar J (2018) 17:172*

Page 9 of 9

polyclonal *Plasmodium falciparum* infections. Am J Trop Med Hyg. 2006;75:836–42.

83. Van den Eede P, Erhart A, Van der Auwera G, Van Overmeir C, Thang ND, le Hung X, et al. High complexity of *Plasmodium vivax* infections in symptomatic patients from a rural community in central Vietnam detected by microsatellite genotyping. Am J Trop Med Hyg. 2010;82:223–7.

84. Ruybal-Pesantez S, Tiedje KE, Rorick MM, Amenga-Etego L, Ghansah A, Ro A, et al. Lack of geospatial population structure yet significant linkage disequilibrium in the reservoir of *Plasmodium falciparum* in Bongo District, Ghana. Am J Trop Med Hyg. 2017;97:1180–9.

85. Brody JR, Calhoun ES, Gallmeier E, Creavalle TD, Kern SE. Ultra-fast high-resolution agarose electrophoresis of DNA and RNA using low-molarity conductive media. Biotechniques. 2004;37:598–602.

86. Messerli C, Hofmann NE, Beck HP, Felger I. Critical evaluation of molecular monitoring in malaria drug efficacy trials and pitfalls of length-polymorphic markers. Antimicrob Agents Chemother. 2016;61:e01500–16.

87. Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). J Biosci Bioeng. 2003;96:317–23.

88. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, et al. A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking. Malar J. 2008;7:223.

89. Baniecki ML, Faust AL, Schaffner SF, Park DJ, Galinsky K, Daniels RF, et al. Development of a single nucleotide polymorphism barcode to genotype *Plasmodium vivax* infections. PLoS Negl Trop Dis. 2015;9:e0003539.

90. Diez Benavente E, Ward Z, Chan W, Mohareb FR, Sutherland CJ, Roper C, et al. Genomic variation in *Plasmodium vivax* malaria reveals regions under selective pressure. PLoS ONE. 2017;12:e0177134.

91. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr Protoc Hum Genet. 2009;2:2–12.

92. Omedo I, Mogeni P, Bousema T, Rockett K, Amambua-Ngwa A, Oyier I, et al. Micro-epidemiological structuring of *Plasmodium falciparum* parasite populations in regions with varying transmission intensities in Africa. Wellcome Open Res. 2017;2:10.

93. Galinsky K, Valim C, Salmier A, de Thoisy B, Musset L, Legrand E, et al. COIL: a methodology for evaluating malarial complexity of infection using likelihood from single nucleotide polymorphism data. Malar J. 2015;14:4.

94. Chang HH, Worby CJ, Yeka A, Nankabirwa J, Kamya MR, Staedke SG, et al. THE REAL McCOIL: a method for the concurrent estimation of the complexity of infection and SNP allele frequency for malaria parasites. PLoS Comput Biol. 2017;13:e1005348.

95. Campino S, Auburn S, Kivinen K, Zongo I, Ouedraogo J-B, Mangano V, et al. Population genetic analysis of *Plasmodium falciparum* parasites using a customized illumina goldengate genotyping assay. PLoS ONE. 2011;6:e20251.

96. Nkhoma SC, Nair S, Al-Saai S, Ashley E, McGready R, Phyo AP, et al. Population genetic correlates of declining transmission in a human pathogen. Mol Ecol. 2013;22:273–85.

97. Dorken ME, Eckert CG. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). J Ecol. 2001;89:339–50.

98. Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA. Standardizing methods to address clonality in population studies. Mol Ecol. 2007;16:5115–39.

99. Juliano JJ, Porter K, Mwapasa V, Sem R, Rogers WO, Ariey F, et al. Exposing malaria in-host diversity and estimating population diversity by capture-recapture using massively parallel pyrosequencing. Proc Natl Acad Sci USA. 2010;107:20138–43.

100. Lin JT, Hathaway NJ, Saunders DL, Lon C, Balasubramanian S, Kharabora O, et al. Using amplicon deep sequencing to detect genetic signatures of *Plasmodium vivax* relapse. J Infect Dis. 2015;212:999–1008.

101. Levitt B, Obala A, Langdon S, Corcoran D, O'Meara WP, Taylor SM. Overlap extension barcoding for the next generation sequencing and genotyping of *Plasmodium falciparum* in individual patients in Western Kenya. Sci Rep. 2017;7:41108.

102. Lerch A, Koepfli C, Hofmann NE, Messerli C, Wilcox S, Kattenberg JH, et al. Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections. BMC Genomics. 2017;18:864.

103. Lalremruata A, Jeyaraj S, Engleitner T, Joanny F, Lang A, Belard S, et al. Species and genotype diversity of *Plasmodium* in malaria patients from Gabon analysed by next generation sequencing. Malar J. 2017;16:398.

104. Hathaway NJ, Parobek CM, Juliano JJ, Bailey JA. SeekDeep: single-base resolution de novo clustering for amplicon deep sequencing. Nucleic Acids Res. 2018;46:e21.

105. Patel JC, Hathaway NJ, Parobek CM, Thwai KL, Madanitsa M, Khairallah C, et al. Increased risk of low birth weight in women with placental malaria associated with *P. falciparum* VAR2CSA clade. Sci Rep. 2017;7:7768.

106. Dara A, Travassos MA, Adams M, Schaffer DeRoo S, Drábek EF, Agrawal S, et al. A new method for sequencing the hypervariable *Plasmodium falciparum* gene *var2csa* from clinical samples. Malar J. 2017;16:343.

107. Assefa SA, Preston MD, Campino S, Ocholla H, Sutherland CJ, Clark TG. estMOI: estimating multiplicity of infection using parasite deep sequencing data. Bioinformatics. 2014;30:1292–4.

108. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, et al. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. Nature. 2012;487:375–9.

109. Auburn S, Campino S, Miotto O, Djimde AA, Zongo I, Manske M, et al. Characterization of within-host *Plasmodium falciparum* diversity using next-generation sequence data. PLoS ONE. 2012;7:e32891.

110. O'Brien JD, Amenga-Etego L, Li R. Approaches to estimating inbreeding coefficients in clinical isolates of *Plasmodium falciparum* from genomic sequence data. Malar J. 2016;15:473.

111. Zhu SJ, Almagro-Garcia J, McVean G. Deconvolution of multiple infections in *Plasmodium falciparum* from high throughput sequencing data. Bioinformatics. 2018;34:9–15.