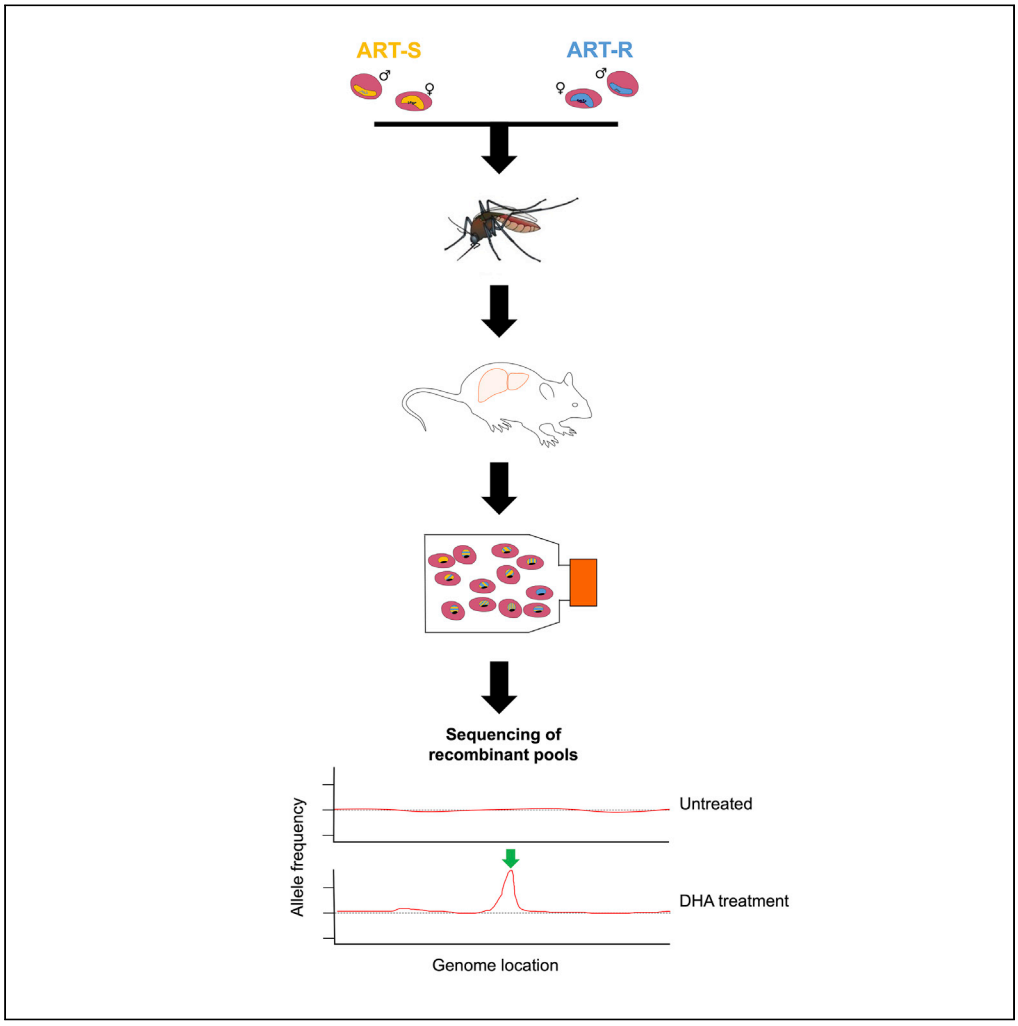
iScience

CellPress
OPEN ACCESS

Article

Optimizing bulk segregant analysis of drug resistance using *Plasmodium falciparum* genetic crosses conducted in humanized mice



Katelyn Vendrely Brenneman, Xue Li, Sudhir Kumar, ..., Ashley M. Vaughan, Michael T. Ferdig, Tim J.C. Anderson

ashley.vaughan@seattlechildrens.org (A.M.V.)
mferdig@nd.edu (M.T.F.)
tanderso@txbiomed.org (T.J.C.A.)

Highlights

Bulk segregant analysis (BSA) rapidly maps drug resistance loci in *P. falciparum*

BSA with dihydroartemisinin identifies *kelch13*, providing proof-of-principle

Synchronization of the bulk population is necessary when using stage-specific drugs

Cryopreserved progeny pools can be used for BSA

Brenneman et al., iScience *25*, 104095
April 15, 2022
https://doi.org/10.1016/j.isci.2022.104095

iScience

CellPress
OPEN ACCESS

Article

Optimizing bulk segregant analysis of drug resistance using *Plasmodium falciparum* genetic crosses conducted in humanized mice

Katelyn Vendrely Brenneman,[1,12] Xue Li,[2,12] Sudhir Kumar,[3,12] Elizabeth Delgado,[2] Lisa A. Checkley,[1] Douglas A. Shoue,[1] Ann Reyes,[2] Biley A. Abatiyow,[3] Meseret T. Haile,[3] Rupam Tripura,[4,5] Tom Peto,[4,5] Dysoley Lek,[6,7] Katrina A. Button-Simons,[1] Stefan H.I. Kappe,[3,8] Mehul Dhorda,[4,5] François Nosten,[5,9] Standwell C. Nkhoma,[10] Ian H. Cheeseman,[11] Ashley M. Vaughan,[3,8,*] Michael T. Ferdig,[1,13,*] and Tim J.C. Anderson[2,*]

SUMMARY

Classical malaria parasite genetic crosses involve isolation, genotyping, and phenotyping of progeny parasites, which is time consuming and laborious. We tested a rapid alternative approach—bulk segregant analysis (BSA)—that utilizes sequencing of bulk progeny populations with and without drug selection for rapid identification of drug resistance loci. We used dihydroartemisinin (DHA) selection in two genetic crosses and investigated how synchronization, cryopreservation, and the drug selection regimen impacted BSA success. We detected a robust quantitative trait locus (QTL) at *kelch13* in both crosses but did not detect QTLs at four other candidate loci. QTLs were detected using synchronized, but not unsynchronized progeny pools, consistent with the stage-specific action of DHA. We also successfully applied BSA to cryopreserved progeny pools, expanding the utility of this approach. We conclude that BSA provides a powerful approach for investigating the genetic architecture of drug resistance in *Plasmodium falciparum*.

INTRODUCTION

To understand the resistance mechanisms and track the spread of resistance alleles in infectious disease-causing organisms, including malaria parasites, it is important to identify the genetic changes that confer drug resistance. For malaria parasites, both association mapping and linkage analysis approaches have been used to understand the genetic mechanisms underlying drug resistance (Anderson et al., 2011). For example, the chloroquine resistance transporter (*crt*) conferring chloroquine (CQ) resistance in the malaria parasite *Plasmodium falciparum* was initially located on chromosome 7 through linkage analysis (Su et al., 1999; Wellems et al., 1990) and further identified by fine mapping (Fidock et al., 2000). Genome-wide association studies (GWAS) have also been applied to map genes associated with resistance to antimalarials (Wang et al., 2016) as well as additional loci arising on genetic backgrounds that carry resistance-associated alleles (Miotto et al., 2015). However, both linkage mapping and association studies have their limitations: linkage mapping requires laborious isolation, genotyping, and phenotyping of multiple progeny parasites; GWAS requires months or years of sample collection.

Bulk segregant analysis (BSA), a complementary alternative to traditional linkage methods, can provide a simple and rapid approach to identify loci that contribute to complex traits without the need to phenotype and genotype individual progeny. Instead, using pooled sequencing of progeny populations, BSA measures change in allele frequency following the application of different selection pressures (e.g., by antimalarial drugs). The BSA approach is simple; by using the complete pool of unique recombinants for analysis, it requires less labor and cost and has the potential for increased statistical power over traditional linkage mapping (Ehrenreich et al., 2010). Nevertheless, BSA cannot directly examine epistatic interactions or measure phenotypes that are not amenable to selection in bulk. The BSA approach was initially developed to study diseases in human and plants, in which DNA from individuals with different phenotypes were pooled and genotyped to identify loci that were enriched for different alleles in each pool (Michelmore et al., 1991). BSA has also been extensively used with genetic crosses of rodent malaria parasites (where it is termed

[1]Eck Institute for Global Health, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA

[2]Program in Disease Intervention and Prevention, Texas Biomedical Research Institute, San Antonio, TX, USA

[3]Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA, USA

[4]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

[5]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine Research Building, University of Oxford Old Road Campus, Oxford, UK

[6]National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia

[7]School of Public Health, National Institute of Public Health, Phnom Penh, Cambodia

[8]Department of Pediatrics, University of Washington, Seattle, WA, USA

[9]Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand

[10]BEI Resources, American Type Culture Collection (ATCC), Manassas, VA, USA

[11]Program in Host Pathogen Interactions, Texas Biomedical Research

Continued

linkage group selection (LGS)) to map genes determining blood stage multiplication rate, virulence, and immunity in *Plasmodium yoelii* (Abkallo et al., 2017) as well as mutations conferring artemisinin (ART) resistance and strain-specific immunity in *Plasmodium chabaudi* (Hunt et al., 2010; Martinelli et al., 2005). BSA has also been used in yeast (Ehrenreich et al., 2010), *C. elegans* (Burga et al., 2019), *Eimeria tenella* (Blake et al., 2011), and the human blood fluke *Schistosoma mansoni* (Chevalier et al., 2014). We have previously applied BSA to the human malaria parasite *P. falciparum* to identify genes that impact parasite fitness throughout the intra-erythrocytic stages of the lifecycle (Li et al., 2019) and polymorphisms in nutrient acquisition/metabolism pathways (Kumar et al., 2020).

ART combination therapies (ACTs) are the current frontline malaria treatments in most countries for the most lethal human malaria parasite, *P. falciparum*. Resistance to ART is mediated by mutations in *kelch13*, a gene required for parasite endocytosis (Birnbaum et al., 2020). Studies in Southeast Asia uncovered the independent emergence of *kelch13* resistance-associated alleles and their rapid spread (Amato et al., 2018; Malaria GEN Plasmodium falciparum Community Project, 2016). Furthermore, recent studies have also detected an increased allele frequency of *kelch13* mutations in clinically ART-resistant *P. falciparum* in Africa (Balikagala et al., 2021). Nonsynonymous polymorphisms in ferredoxin (*fd*), the apicoplast ribosomal protein S10 (*arps10*), multidrug resistance protein 2 (*mdr2*), and *crt* have also been identified and show strong associations with ART resistance (Miotto et al., 2015), while *in vitro* selection has revealed other loci, such as *pfcoronin*, that can modulate ART resistance (Demas et al., 2018).

We designed this study to optimize aspects of the BSA approach for the identification of drug resistance-associated loci in malaria parasites (Figure 1), using dihydroartemisinin (DHA), the active metabolite of ART, as our test drug. We chose DHA because (i) ART is one of the most widely used antimalarial drugs, (ii) *kelch13* is a known resistance locus, providing a good control for our methods, although other loci are also suspected (Demas et al., 2018; Mukherjee et al., 2017), (iii) identification of additional loci involved in drug resistance would be of great interest, and (iv) both drug activity and drug resistance are stage-specific, with maximal killing activity during the trophozoite stage (Klonis et al., 2013), and resistance only during the ring stage (Tilley et al., 2016). Stage specificity provides particular challenges for BSA because parasite cultures containing a mixture of stages are expected to show minimal differences in resistance. We aimed to answer three main questions. First, for stage-specific drugs such as DHA, can we detect QTLs in unsynchronized parasite pools, or is prior synchronization required? This is important because synchronization may reduce the size of progeny pools, and for successful BSA, it is critical to have progeny pools containing large numbers of recombinants. Second, how can we determine the optimal drug dose regimen to maximize signal? Third, can we use cryopreserved progeny pools for BSA experiments? This would greatly enhance the utility of BSA by allowing experiments to be conducted long after genetic crosses have been conducted, using archived progeny pools. However, as with synchronization, we were concerned that cryopreservation would reduce the diversity of progeny pools limiting our ability to detect QTLs.

We performed DHA BSA in two genetic crosses between ART-sensitive African and ART-resistant Southeast Asian parasites [Mal31×KH004 and NF54 (clone 3D7)×NHP1337]. These two crosses are of particular interest because SNPs in the *kelch13* locus, as well as SNPs in each of the four loci associated with ART-resistance in Miotto et al. (Miotto et al., 2015) genome-wide association study (GWAS) are segregating in both crosses. We detected strong and repeatable QTL at the *kelch13* locus on chromosome 13 in both crosses but found no evidence for involvement of other loci. Chromosome 13 QTLs were detected only in synchronized recombinant progeny pools. We also show that the application of eRRSA, a modification of the ring-stage survival assay, allows for a rational determination of optimal drug dose for use in BSA. Finally, we successfully applied BSA with cryopreserved progeny pools.

RESULTS

Genetic crosses

We employed *Anopheles stephensi* mosquitoes and human liver-chimeric FRG huHep mice as described in the study by Vaughan et al. (Vaughan et al., 2015), to generate two unique genetic crosses [Mal31×KH004 and NF54 (clone 3D7)×NHP1337] for this study (Figure 1A and Table 1) (Figure S1).

Mal31×KH004: Mal31 is a *kelch13* wild-type parasite isolated from a patient in Malawi in 2016; KH004 is a *kelch13* mutant parasite carrying the common C580Y mutation and was isolated from Western Cambodia

Institute, San Antonio, TX, USA

[12]These authors contributed equally

[13]Lead contact

*Correspondence: ashley.vaughan@seattlechildrens.org (A.M.V.), mferdig@nd.edu (M.T.F.), tanderso@txbiomed.org (T.J.C.A.)

https://doi.org/10.1016/j.isci.2022.104095

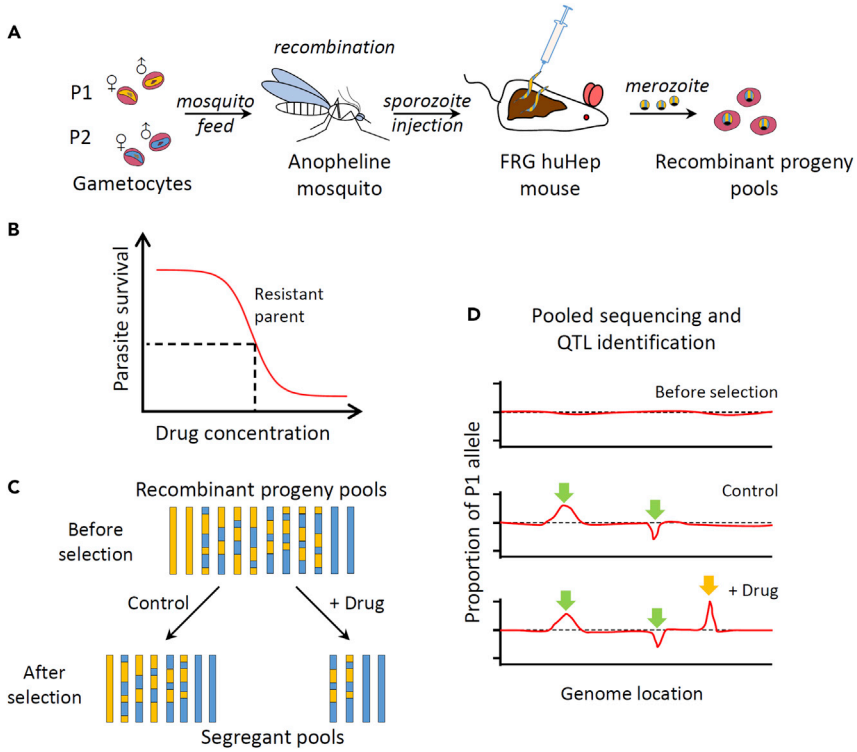


Figure 1. The principles of bulk segregant analysis with human malaria parasites

(A) Recombinant progeny pools are generated using *Anopheles stephensi* mosquitoes and FRG huHep mice.

(B) Determination of drug concentration for bulk segregant analysis (BSA). A standard approach is to measure the IC_{50} and IC_{90} value of the resistant parent; however, in this study, we used eRRSA to determine a discerning drug concentration for selection since classical IC_{50} and IC_{90} measurements using DHA do not correlate with the *in vivo* measure of ART resistance, clinical parasite clearance rate.

(C) An overview of drug BSA, selection is applied with drug of choice to half of the recombinant progeny pool; each bar indicates a representative recombinant progeny.

(D) Genomic sequencing measures changes in allele frequency of segregant pools. Green arrows indicate allele frequency changes detected from both pools with and without drug treatment, which are thus not related to drug resistance, while the orange arrow indicates a locus conferring drug resistance.

in 2016. There are 14,455 core-genome SNPs (defined in the study by (Miles et al., 2016)) distinguishing these two parental parasites. The gametocytes from both parental parasites were mixed in a ~50:50 ratio to infect approximately 450 mosquitoes. We generated four independent recombinant pools from this cross by using independent groups of 100 mosquitoes for the isolation of sporozoites and four individual FRG huHep mice that were infected with the sporozoites. Recombinants are generated after gametes fuse to form zygotes in the mosquito midgut (Figure 1A). Mitotic division of the four meiotic products ultimately leads to the generation of thousands of haploid sporozoites within each oocyst. The oocyst prevalence from the cross ranged from 71% to 91%, with a mean of 5 oocysts per mosquito (mean of median oocyst numbers from 3 cross cages) (Table 2). The estimated number of recombinants for each pool were 2,000 (5 oocysts × 4 recombinants × 100 mosquitoes) of which 50% (2,000/2 = 1,000) are expected to be outbred recombinants. The initial allele frequencies of Mal31 were 0.75–0.79 for bulk recombinant progeny pools directly from mice; the deviation from the expected even representation of the two parental alleles indicates the existence of Mal31 selfed progeny generated by fusion of male and female gametes of Mal31. Thus, the expected number of unique outbred recombinants per pool will be less than 1,000 for this cross. Two recombinant progeny pools were randomly selected for use in this study. For these two pools, the genome-wide allele frequencies shifted to 0.53 and 0.55, respectively, after 15 days of *in vitro* culture prior to DHA BSA experiments (Figure S2). The changes to a more even representation of the two parental genomes suggest selection against selfed progeny during *in vitro* blood culture, which we have detected previously in a different cross (Li et al., 2019).

Table 1. Segregating variation in two genetic crosses

		Cross 1		Cross 2	
		Cambodia	Malawi	Thailand	Africa
Gene name	Gene ID	KH004	Mal31	NHP1337	NF54
kelch13	PF3D7_13437000	C580Y	WT	C580Y	WT
fd	PF3D7_1318100	D193Y	WT	D193Y	WT
arps10	PF3D7_1460900	V127M,D128H	WT	V127M,D128H	WT
mdr2	PF3D7_1447900	T484I,F423Y,G299D,S208N	I492V,S208N	T484I,F423Y,G299D,S208N	WT
crt	PF3D7_0709000	DD2[a]+G367C	WT	DD2[1]	WT

In addition to a known ART-resistance-associated mutation in *kelch13*, three variants (underlined amino acid changes) showing strong associations in a genome-wide association analysis (Miotto et al., 2015) are also segregating in these crosses. See also Figure S1.

[a]Alleles carry same amino acid mutations as the Dd2 parasite line (M74I, N75E, K76T, A220S, Q271E, N326S, I356T, and R371I) when compared to reference (3D7).

NF54 (clone 3D7)×*NHP1337*: NF54 was isolated from a Dutch resident, who had not left the country (Delemarre and van der Kaay, 1979), has been postulated to be of African origin (Preston et al., 2014). For this study, we used a clone of NF54, 3D7 (hereafter referred to as NF54), which was isolated from a volunteer in a malaria clinical trial after controlled human malaria infection by 3D7-infected *Anopheles stephensi* mosquito bite (Ockenhouse et al., 1998). NHP1337 is a C580Y *kelch13* mutant parasite isolated from the Thai-Myanmar border in 2011 and has been used previously in a genetic cross (Button-Simons et al., 2021). There are 17,318 SNPs distinguishing the two parental parasites. The gametocytes from both parental parasites were mixed in a ~50:50 ratio to feed about 450 mosquitoes (150 per cage, three cages). We generated two independent recombinant pools for this cross by using different sets of mosquitoes. The mean oocyst prevalence for this cross was 90% (range: 80%–100%), with a burden of 18.75 oocysts per mosquito midgut (mean of median oocyst numbers from 4 cross cages) (Table 2). The estimated number of recombinants for each pool used were 7,500 (18.75 oocysts × 4 recombinants × 100 mosquitoes), of which 50% (7,500/2 = 3,750) are expected to be recombinants. The allele frequencies of NF54 were 0.41 and 0.46 for bulk pools before DHA BSA experiments (Figure S3).

Determining DHA concentrations for BSA experiments

Maximal inhibitory concentrations, such as IC_{50} and IC_{90}, would typically be used to determine the concentration of drugs used for BSA experiments (Figure 1B). However, for DHA, IC_{50} and IC_{90} values are not correlated with patient clearance half-life, the key clinical resistance readout (Witkowski et al., 2013). In this study, we used an assay we developed, eRRSA, to measure the DHA resistance level of the resistant parent KH004 and one of our recombinant progeny pools from the Mal31×KH004 cross. eRRSA relies on parasite growth fold change $[2^{(\text{average } ct \text{ drug treated} - \text{average } ct \text{ control})}]$ to quantify parasite resistance to DHA and this read out is strongly correlated to patient clearance half-life (Davis et al., 2020). We generated DHA eRRSA dose-response curves, i.e., one eRRSA for each of 10 different drug concentrations (Figure 2) to identify concentrations of DHA that would kill some but not all of the sensitive parasites. The eRRSA dose-response curves for resistant and pooled progeny were not significantly different ($F_{4,68} = 1.518$, $p = 0.207$). We chose 50 and 100 nM for our DHA BSA experiment: 50 and 100 nM correspond to an eRRSA fold change of 14 (46.1% parasite survival for KH004 and 52.7% parasite survival for the Mal31 × KH004 recombinant pool) and a fold change of 20 (32.4% parasite survival for KH004 and 34.1% parasite survival for the Mal31 × KH004 recombinant pool), respectively.

BSA with pooled progeny from the Mal31×KH004 cross

Malaria parasites show different tolerance levels to ART drugs at different lifecycle stages (ring, trophozoite, or schizont) (Cui et al., 2012). To remove the influence of parasites from different stages and maximize the power of DHA BSA, we synchronized samples from the recombinant progeny pools. Our method of synchronization selects for late-stage parasites (trophozoites and schizonts); therefore, to determine if synchronization would artificially select different progeny populations or reduce parasite population diversity, we took samples from the original asynchronous recombinant progeny pool at three timepoints (0, 18, or 36 h) prior to DHA BSA, and we also included an unsynchronized control (Figure 3A). High-depth Illumina sequence (>100×) profiles of these four progeny pools revealed no significant differences (Figure S4).

Table 2. Details of gametocytemia and mosquito infection for generating the crosses

Mosquito cage	% Gametocytemia	Post feed exflagellation centers/ field of view[a]	Oocyst numbers (Mean)	Oocyst numbers (Median)	% Prevalence	Sporozoites/ Mosquito
KH004 only infection	3.74	2.5	6.5	3	76	18,965
Mal31 only infection	2.19	3	8	7	84	45,500
KH004 × Mal31 cage 1	NA	3	8.3	7	91	40,832
KH004 × Mal31 cage 2	NA	3	7	3	83	26,590
KH004 × Mal31 cage 3	NA	3	11.5	5	71	44,541
NF54 only infection	3.83	4	42.45	36	91.5	15,866
NHP1337 only infection	3.28	2.6	23	15.5	79	21,437
NF54 × NHP1337 cage 1	NA	3.75	41	38.5	100	38,117
NF54 × NHP1337 cage 2	NA	3.75	21.6	14.5	78.6	32,850
NF54 × NHP1337 cage 3	NA	3.75	31	13	88	19,411
NF54 × NHP1337 cage 4	NA	3.75	14.8	9	92	26,000

NA, not applicable. See also Figures S2 and S3.

[a]average number of exflagellation center formation observed under ~8–10 random fields of view under bright-field microscope (40×).

We compared the Mal31 allele frequency from segregant pools with and without DHA treatment using genome-wide G' (Magwene et al., 2011) to measure the significance level of detected QTLs (Figures 1C and 1D). For each detected QTL (FDR <0.01), we calculated 95% confidence intervals to define the QTL boundaries (Table S1). For pools without synchronization, we were not able to detect any significant QTL. For pools that were synchronized, we identified a single strong QTL on chromosome 13 at the *kelch13* locus (Figures 3 and S5). The largest allele frequency difference between control and DHA-treated parasites was detected from pools synchronized at 0 h. We did not detect significant differences between 50 and 100 nM DHA treatments.

BSA with cryopreserved pooled progeny from the Mal31×KH004 cross

Generation of recombinant progeny pools for each BSA experiment requires *in vitro* culture of parasites, maintaining Anopheline mosquitoes, and infection of FRG huHep mice, and is extremely labor intensive and costly. An alternative strategy is to use cryopreserved progeny pools for BSA experiments. We compared allele frequencies of a Mal31×KH004 progeny pool before and after cryopreservation. In the absence of drug treatment, the allele frequency across the genome remained at the same level in progeny pools before and after synchronization. We observed a modest divergence in allele frequency at chromosome 2 and 14, with altered frequency of Mal31 alleles, which may indicate genetic drift resulting from a reduction in parasite population diversity (Figure S4B). We applied DHA BSA to this cryopreserved pool (Figure 3C) and compared results with BSA before cryopreservation (Figure 3B, panel 2). We detected a strong QTL at the *kelch13* locus, although the G prime values were approximately 5-fold less and the 95% CIs were approximately ten times larger as compared to the results from the original BSA analysis.

BSA with pooled progeny from the NF54×NHP1337 cross

Based on previously published eRRSA values for NHP1337 (Davis et al., 2020), we also applied DHA BSA to pooled progeny from the NF54×NHP1337 cross. For direct comparison of results, we used the same methodology and DHA concentrations as described for the Mal31×KH004 cross. As with the Mal31×KH004 cross, no significant difference in allele frequency was detected from pools with and without synchronization prior to BSA (Figure S3). Similarly, we detected a strong QTL at the *kelch13* locus from both biological replicates after synchronization, but not in the unsynchronized control (Figure S6). A second smaller QTL peak on chromosome 14 was detected in one of the two biological replicates.

DISCUSSION

Given the success of BSA approaches for identification of drug resistance loci in rodent malaria (Borges et al., 2011; Carter et al., 2007; Modrzynska et al., 2012), our aim was to develop a BSA method to identify drug resistance loci in *P. falciparum* to better exploit our capacity to conduct *P. falciparum* genetic crosses

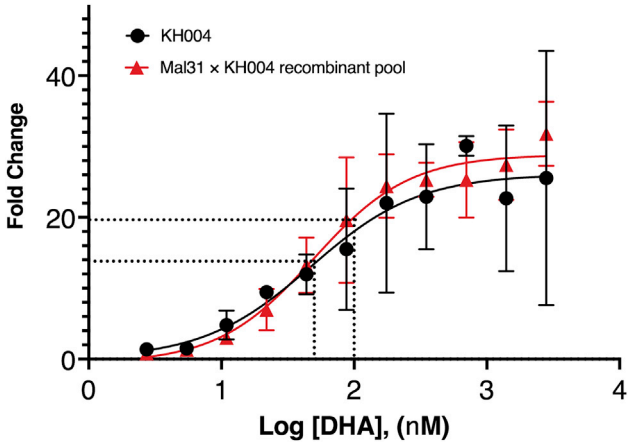


Figure 2. eRRSA dose-response curves for the DHA-resistant parent KH004 and a Mal31 × KH004 recombinant progeny pool

Synchronized ring stage parasites were treated with 10 DHA concentrations (in a two-fold dilution series starting at 2,800 nM) for 6 h. Higher fold change indicates increased parasite death in drug-treated versus untreated parasites using our eRRSA method. Both the KH004 parent (ART-resistant) and the Mal31×KH004 recombinant progeny pool were tested using three biological replicates to determine the optimal DHA concentration for bulk segregant DHA selection (data are represented as mean ± standard deviation). Dashed lines indicate 50 and 100 nM and were chosen as discerning doses for treatment of bulk populations to selectively enrich for resistant parasites.

in humanized mice (Vaughan et al., 2015). To develop our approach, we chose DHA, the active metabolite of ART, as a test drug. DHA shows highly stage-specific drug action against trophozoites (Klonis et al., 2013). Furthermore, there is a well known locus (*kelch13*) underlying ART resistance, a valuable positive control for assessing and optimizing our methods (Birnbaum et al., 2020). We specifically aimed to explore the impact of (i) parasite synchronization and (ii) parasite cryopreservation to develop strategies to choose optimal drug selection regimens for BSA. The power of BSA is strongly dependent on the number of recombinant parasites present in progeny pools. Therefore, methods for BSA with malaria parasites should aim to minimize time in culture prior to selection and avoid possible procedures that may reduce the diversity of recombinant pools.

Synchronization influences BSA success

Antimalarial drugs show a spectrum of specificity for different parasite asexual blood stages. Drugs such as chloroquine kill all stages, whereas cycloheximide and trichostatin (Wilson et al., 2013) and newly developed drugs like BCH070 (Clements et al., 2020) show strong specificity for early rings. Conversely, lumefantrine, mefloquine, and piperaquine target trophozoites (Hodel et al., 2016). Importantly, we determined that synchronization is essential for a successful DHA BSA: with synchronization, we were able to clearly detect the known *kelch13* peak and without synchronization we detected no peaks (Figure 3B). Thus, synchronizing parasites at the early ring stage allowed us to specifically target ring-stage drug resistance; the presence of other parasite stages in the culture is an additional variable that thus limits the success of selection. We suggest that synchronization should be used for BSA experiments with stage-specific drugs to ensure drugs act on their target stage.

Our experiments with Mal31×KH004 used synchronized parasite pools (0, 18, or 36 h) prior to BSA. Selection on all three pools resulted in a strong significant QTL peak at *kelch13*. Notably, the strongest QTL localizing to *kelch13* was in the population synchronized at 0 h and reduced in magnitude at 18 and 36 h. We suspect that the high proportion of ring stages in the 0 h synchronized progeny pool maintained higher parasite diversity than the 18 and 36 h synchronizations, which resulted in a stronger QTL. The lack of a detectable QTL in unsynchronized cultures likely results from presence of late-stage parasites that are not killed by DHA.

Cryopreserved recombinant progeny pools can be used for drug BSA experiments

Once bulk recombinant progeny pools are isolated from the FRG huHep mouse, they can be directly used in BSA experiments, and they also can be cryopreserved. Cryopreservation allows for multiple

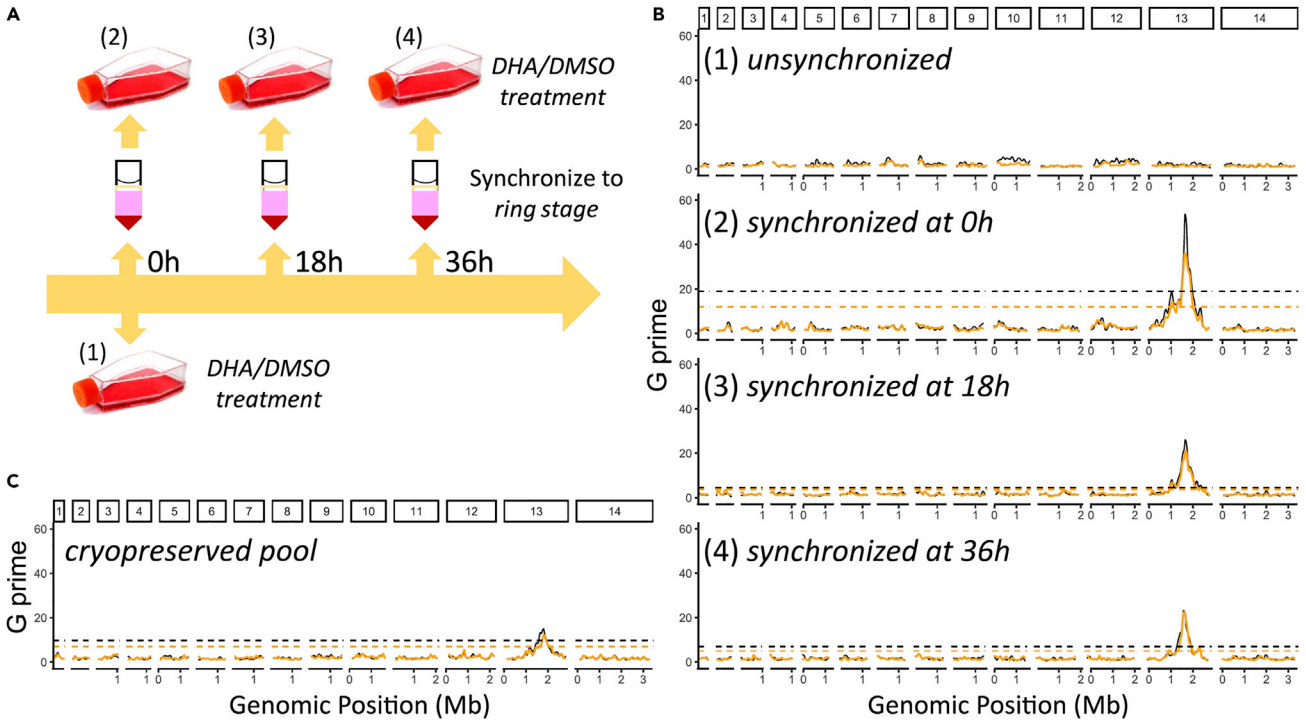


Figure 3. DHA bulk segregant analysis (BSA) on bulk progeny from the Mal31×KH004 cross

(A) Synchronization of parasites at the ring stage: (1) pools without synchronization; (2–4) pools synchronized at 0, 18, and 36 h, respectively.

(B) DHA BSA analyses of (1) the unsynchronized pool, (2–4) pools synchronized at 0, 18, and 36 h, respectively.

(C) DHA BSA with a cryopreserved recombinant pool synchronized at 0 h. Orange and black lines are G prime values comparing allele frequency of 50 nM or 100 nM DHA-treated pools with control pools; orange (50 nM) and black (100 nM) dashed lines are corresponding significance thresholds (FDR = 0.01). The numbers (1–14) above the x-axis in panels B and C indicate the 14 *P. falciparum* chromosomes. There was no threshold line plotted for panel B (1) unsynchronized as no G prime value passed the significance test. See also Figures S4, S5, and S6, and Table S1.

BSA experiments to be conducted using the same recombinant progeny pool at a later date. However, only ring stage parasites survive the freeze and thaw process. Owing to the asynchrony over time of the clones in the bulk recombinant progeny pools recovered from FRG huHep mice, the diversity of the population is likely to be diminished during the process of cryopreservation, thawing, and parasite regrowth *in vitro*. Nevertheless, our BSA experiments reveal the same QTL before and after cryopreservation (Figure 3C), confirming that sufficient genetic variation is maintained in cryopreserved stocks of recombinant progeny pools. However, we note that the chromosome 13 QTL peak detected, while strongly significant, is smaller than detected prior to cryopreservation, suggesting some loss of recombinants within pools. Methods that maximize size and diversity of progeny pools upon cryopreservation will increase the success of BSA experiments. These include ensuring that bulk progeny are cryopreserved when rings predominate and that culture time between parasite recovery from cryopreservation and BSA is minimized, since progeny can be outcompeted by other progeny over time *in vitro*, leading to decreased diversity in the population.

Determining drug concentrations for BSA selection experiments

The goal of a selective drug BSA is to kill only sensitive parasites in the sample exposed to drug. Therefore, the drug concentration chosen for selection is very important. If the dose is too high, few parasites in the treated pool will survive, while if the dose is too low, sensitive parasites will survive, and one will not see enrichment of alleles and QTLs resulting from drug selection. In yeast BSA experiments, particular strong selection (recovery of less than 1% of the total progeny) has been used as a rule of thumb to select progeny showing extreme phenotypes (Ehrenreich et al., 2010). This works well for yeast, where recombinant pools can contain more than 10^6 recombinants, but in our experience with *P. falciparum*, the use of such stringent drug doses for BSA can result in poor recovery of parasites in the treated pools.

To choose drug concentrations for BSA, we measured drug response in the resistant parent and one of our recombinant progeny pools. We examined dose response using a modified ring stage assay (eRRSA; Figure 2) (Davis et al., 2020) which examines parasite survival following a short pulse of drug exposure, to generate dose-response curves. We observed a slight shift in the eRRSA dose-response curve between the resistant parasite, KH004, and the Mal31 × KH004 recombinant progeny pool, although the two curves were not significantly different. We then chose two DHA concentrations that did not kill all sensitive parasites. Given the high value of stock samples of bulk progeny and the time and cost of the BSA experiment, it is important to conduct these initial experiments to enhance the chance of successful selections. Encouragingly, we found that a range of DHA doses revealed the same QTLs with near equivalent results. For the second cross (NF54 × NHP1337), we used the same DHA concentrations estimated for the Mal31×KH004 cross, and this worked well. Using the same concentrations allowed for a direct comparison between the Mal31×KH004 cross and the NF54×NHP1337 cross, suggesting that the bracketing approach provides a practical solution to determining drug doses.

Use of dose response data from uncloned progeny pools could be helpful when progeny drug responses fall outside the range of the parents (transgressive segregation) (Nogami et al., 2007) or in cases where there is no prior information about progeny drug responses. If the genetic architecture is simple and parent phenotypes are significantly different, parental drug responses can be measured, and two or three drug doses bracketing the IC_{50} (or equivalent metric) of the most resistant parent can be used for drug BSA. Indeed, the use of a well-chosen range of selection doses could reveal aspects of the genetic architecture that would not be revealed by a single dose.

Technical considerations & caveats

Statistical power and numbers of recombinants

The power of BSA is restricted by the number of unique recombinant progeny in a pool. To increase the number of recombinants, the number of mosquitoes used for sporozoite isolation per cross can be increased, thereby increasing the number of unique recombinants. In addition, the number of mice used per cross can also be increased; each mouse can be infected with unique sporozoite pools, providing biological replicates. Finally, to maximize the number of unique recombinant progeny, large volumes of recombinant progeny pools can be cryopreserved shortly after the transition to *in vitro* culture. Selfed progeny is another limitation for the statistical power of BSA. We have found large numbers of inbred progeny generated by mating between male and female gametes from the same parent, in some crosses (Button-Simons et al., 2021; Li et al., 2019). However, selfed progeny can be selectively removed during asexual growth (Li et al., 2019). Selfed progeny could also be removed by carrying out crosses with transgenic parental strains that only produce either functional male or female gametes (Kumar et al., 2021), to ensure that all progeny are recombinant. A more flexible approach could involve treatment with chemicals of antibody-based reagents to remove male gametes from one parent and female gametes from the other parent. Reagents that prevent production of male gametes, such as aphidicolin (Ramiro et al., 2015), are available, but we currently do not have reagents for removing female gametes.

Assessing the numbers of recombinants

We have been using the number of oocysts and sequences from cloned progeny to estimate the maximum number of recombinants in a recombinant pool. Measuring the number of recombinant progeny and the proportion of inbreeding in a pool before BSA could be particularly useful for future crosses. This can be done by laborious cloning from the same pools (Button-Simons et al., 2021). Alternatively, single-cell genome sequencing (Dia and Cheeseman, 2021) can be employed to analyze the diversity of the recombinant pool. Thus, rapid genotyping of sporozoites or the progeny pool could provide details on the dynamics of BSA experiments, which will then allow for further optimization approaches (Abkallo et al., 2017).

Balancing the experiment duration and culture volume

BSA can be used to examine the impact of multiple different selections in parallel. However, if multiple experiments are conducted at the same time, large numbers of parasites are needed, which requires a longer culture time. Population diversity decreases during *in vitro* culture due to parasite competition (Li et al., 2019). A solution to this is to use smaller culture volumes for BSA experiments without compromising

genetic diversity. We have used 3 mL (Kumar et al., 2020; Li et al., 2019) and 500 μL culture volumes (this study) at 1% parasitemia but have not been able to sample efficiently using smaller culture volumes such as 100 μL from 96-well plates.

Genetic architecture of DHA resistance

The primary goal of this project was to develop methodology for drug BSA experiments with *P. falciparum*. However, genetic crosses also provide information on the genetic architecture of ART resistance. Four lines of evidence suggest that loci other than *kelch13* may contribute to resistance: (i) GWAS studies have identified several loci (*ferredoxin*, *apicoplast ribosomal protein S10*, *multidrug resistance protein 2*, and *crt*) as well as *kelch13* to be associated with ART resistance (Miotto et al., 2015); (ii) parasites showing slow clearance from ART-treated patients but lacking mutations in *kelch13* have been described (Mukherjee et al., 2017); (iii) longitudinal analyses have shown that other loci show parallel changes in allele frequency over time (Cerqueira et al., 2017), suggesting that other loci may also contribute to slow clearance phenotypes; and (iv) *in vitro* selection has revealed other loci, such as *coronin*, that can modulate ART-R (Demas et al., 2018); similarly, *in vivo* selection in humanized mice resulted in high levels of resistance to DHA that was not associated with *kelch13* mutations (Tyagi et al., 2018), but which could be genetically mapped using the BSA approach described.

The two crosses analyzed here show a single QTL indicating that *kelch13* is the largest genetic signal of this phenotype. We observed smaller signals meeting the significance threshold on chromosomes 14 for some replicates (Figure S6). However, as these reach significance in only some replicates, and are also seen in the control (untreated samples), they most likely reflect loci underlying differences in growth rate, rather than true resistance QTLs. This illustrates a further advantage of BSA: by conducting replicate BSA experiments using independent pools of progeny, we can minimize detection of false positive resistance QTLs.

Importantly, these crosses allow us to interrogate candidate loci. Both genetic crosses analyzed have segregating SNPs in *ferredoxin*, *apicoplast ribosomal protein S10*, *multidrug resistance protein 2*, and *crt*. Interestingly, while these loci show strong association with slow clearance in a GWAS, they were proposed to provide permissive background mutations on which ART-resistance mutations could arise (Miotto et al., 2015), perhaps by impacting parasite fitness. In our BSA experiments, we sampled parasite pools at several time points after drug exposure. This allows detection of QTLs underlying both drug resistance and fitness. We did not detect QTLs at any of these four loci, suggesting that these SNPs do not play a role in ART resistance, or contribute to fitness of ART-resistance parasites in these two crosses. These results are consistent with results from CRISPR/Cas9 editing of *ferrodoxin* and *mdr2* SNPs (Stokes et al., 2021), where no impact of these SNPs on ART resistance or fitness was found.

Undertaking more complex BSA analyses using an increased range of drug concentrations may help uncover further loci that play a role in resistance. In addition, further crosses may identify additional loci underlying ART resistance. For example, for 20 years, *crt* has been considered to be the primary determinant of CQ resistance. However, in parallel work on CQ resistance, we have identified a prominent chromosome 6 QTL that acts in concert with *crt* to determine CQ resistance in some crosses (unpublished data). Our ability to conduct genetic crosses in FRG huHep mice provides a direct approach to identify suspected non-*kelch13* mechanisms of ART resistance using parasites isolated from patients (Vendrely et al., 2020) or evolved *in vitro* (Tyagi et al., 2018).

BSA is broadly applicable to selectable phenotypic traits

We have demonstrated that BSA is a powerful approach in determining genes related to drug resistance in *P. falciparum*, paralleling findings from other organisms (Borges et al., 2011; Hunt et al., 2010; Khan et al., 2020; Park et al., 2014). In other work, we have used BSA to examine the impact of culture conditions (serum or AlbuMAX) (Kumar et al., 2020), and identifying genome regions that impact parasite fitness (Li et al., 2019). We anticipate that BSA can also be explored to study any phenotypic trait that can be measured by selection. For example, we can examine the selection on parasite populations at different temperatures, simulating malaria-induced fever (Gravenor and Kwiatkowski, 1998). We can control the pH of culture media, which can be altered in patients with acidosis caused by severe malaria (Geary et al., 1985). We can also study the selection of human immunity on progeny populations by culturing parasites with serum from recovered patients or people that have never had malaria, as was

shown for mouse malaria (Martinelli et al., 2005). BSA can also be used to study parasite-host interactions. Malaria parasites invade human red blood cells (RBCs) by a series of interactions between host and parasite surface proteins (Cowman et al., 2017). Hence, we can use different blood cell populations to examine genes underlying parasite invasion phenotypes. Furthermore, Anopheline mosquitoes are the intermediate host between parasite and human and the distribution of mosquitoes varies at different geographical locations. For example, *Anopheles gambiae* mosquitoes are common vectors in Africa, while *Anopheles dirus and Anopheles minimus* are the dominant vectors in Asia (Sinka et al., 2012). Hence, we can infect different mosquitoes with recombinant progeny populations to examine the selection at the mosquito stage, as pioneered by Molina-Cruz et al. (Molina-Cruz et al., 2013). Thus, BSA provides a powerful approach to investigate the genetic architecture of phenotypic traits across the complete *P. falciparum* life cycle.

Limitations of the study

Our BSA approach was well powered to identify major QTLs. However, BSA cannot determine interactions between loci in cases where two or more QTLs are identified. In such cases, isolation and characterization of small numbers of individual clones can be an effective approach to follow-up on initial BSA results and determine interlocus interactions underlying trait phenotypes. The power of BSA is strongly dependent on the numbers of independent recombinants in progeny pools, which is difficult to determine without cloning of a sample of individual progeny. Here, we have estimated maximal numbers of recombinants present in pools from numbers of oocysts in mosquito midguts. However, these are likely to be overestimates of actual numbers. A simple approach—perhaps using single cell genomics approaches—to empirically determine numbers of independent progeny within progeny pools would be valuable for ensuring reliability and determining the statistical power of BSA experiments. Generally, success of BSA to identify loci with minor impact on phenotypes will improve with larger numbers of independent recombinant progeny in the pool, less selfing (parental genotypes in the pool), more biological replication (infected mice), and refined drug selections that are possible using cryopreserved samples.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
  - Human subjects
  - *Plasmodium falciparum* parasite clones/strains
- METHOD DETAILS
  - Preparation of the genetic crosses
  - Cryopreservation and thawing of progeny bulks
  - Measuring DHA resistance level using the eRRSA
  - DHA BSA
  - Library preparation and sequencing
  - Mapping and genotyping
  - Bulk segregant analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104095.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH) program project grant P01 AI127338 (to MF) and by NIH grant R37 AI048071 (to TJCA). Work at Texas Biomedical Research Institute was conducted in

facilities constructed with support from Research Facilities Improvement Program grant C06 RR013556 from the National Center for Research Resources. SMRU is part of the Mahidol Oxford University Research Unit supported by the Wellcome Trust of Great Britain. The parental line, Mal31, used in the *Mal31×KH004* cross was sampled from a Malawian patient in 2016 as part of a cross-sectional study funded by the Wellcome Trust of Great Britain (Grant no. 099992/Z/12/Z to SCN). We thank the patients who provided parasites used in this work.

AUTHOR CONTRIBUTIONS

TJCA, MTF, and AMV conceived the project. KVB, XL, and SK conceptualized and planned the study. SN, FN, MD, RT, TP, and DL provided and characterized malaria parasites from infected patients from which parental parasites were chosen. SK, BAA, and MTH prepared all the genetic crosses. KVB, LAC, and DAS prepared BSA samples and DHA drug assay experiment. XL interpreted and visualized the data and performed formal analysis. ED and AR were involved in next-generation sequencing library preparations. KVB, XL, SK, AMV, MTF, and TJCA wrote the original manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 3, 2021
Revised: February 24, 2022
Accepted: March 11, 2022
Published: April 15, 2022

REFERENCES

Abkallo, H.M., Martinelli, A., Inoue, M., Ramaprasad, A., Xangsayarath, P., Gitaka, J., Tang, J., Yahata, K., Zoungrana, A., Mitaka, H., et al. (2017). Rapid identification of genes controlling virulence and immunity in malaria parasites. PLoS Pathog. *13*, e1006447. https://doi.org/10.1371/journal.ppat.1006447.

Amato, R., Pearson, R.D., Almagro-Garcia, J., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Drury, E., Stalker, J., Miotto, O., et al. (2018). Origins of the current outbreak of multidrug-resistant malaria in southeast Asia: a retrospective genetic study. Lancet Infect. Dis. *18*, 337–345. https://doi.org/10.1016/S1473-3099(18)30068-9.

Anderson, T., Nkhoma, S., Ecker, A., and Fidock, D. (2011). How can we identify parasite genes that underlie antimalarial drug resistance? Pharmacogenomics *12*, 59–85. https://doi.org/10.2217/pgs.10.165.

Balikagala, B., Fukuda, N., Ikeda, M., Katuro, O.T., Tachibana, S.I., Yamauchi, M., Opio, W., Emoto, S., Anywar, D.A., Kimura, E., et al. (2021). Evidence of artemisinin-resistant malaria in Africa. N. Engl. J. Med. *385*, 1163–1171. https://doi.org/10.1056/NEJMoa2101746.

Birnbaum, J., Scharf, S., Schmidt, S., Jonscher, E., Hoeijmakers, W.A.M., Flemming, S., Toenhake, C.G., Schmitt, M., Sabitzki, R., Bergmann, B., et al. (2020). A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. Science *367*, 51–59. https://doi.org/10.1126/science.aax4735.

Blake, D.P., Billington, K.J., Copestake, S.L., Oakes, R.D., Quail, M.A., Wan, K.L., Shirley, M.W., and Smith, A.L. (2011). Genetic mapping identifies novel highly protective antigens for an apicomplexan parasite. Plos Pathog. *7*, e1001279. https://doi.org/10.1371/journal.ppat.1001279.

Borges, S., Cravo, P., Creasey, A., Fawcett, R., Modrzynska, K., Rodrigues, L., Martinelli, A., and Hunt, P. (2011). Genomewide scan reveals amplification of mdr1 as a common denominator of resistance to mefloquine, lumefantrine, and artemisinin in Plasmodium chabaudi malaria parasites. Antimicrob. Agents Chemother. *55*, 4858–4865. https://doi.org/10.1128/AAC.01748-10.

Burga, A., Ben-David, E., Lemus Vergara, T., Boocock, J., and Kruglyak, L. (2019). Fast genetic mapping of complex traits in C. elegans using millions of individuals in bulk. Nat. Commun. *10*, 2680. https://doi.org/10.1038/s41467-019-10636-9.

Button-Simons, K.A., Kumar, S., Carmago, N., Haile, M.T., Jett, C., Checkley, L.A., Kennedy, S.Y., Pinapati, R.S., Shoue, D.A., McDew-White, M., et al. (2021). The power and promise of genetic mapping from Plasmodium falciparum crosses utilizing human liver-chimeric mice. Commun. Biol. *4*, 734. https://doi.org/10.1038/s42003-021-02210-1.

Carter, R., Hunt, P., and Cheesman, S. (2007). Linkage Group Selection–a fast approach to the genetic analysis of malaria parasites. Int. J. Parasitol. *37*, 285–293. https://doi.org/10.1016/j.ijpara.2006.11.013.

Cerqueira, G.C., Cheeseman, I.H., Schaffner, S.F., Nair, S., McDew-White, M., Phyo, A.P., Ashley, E.A., Melnikov, A., Rogov, P., Birren, B.W., et al. (2017). Longitudinal genomic surveillance of Plasmodium falciparum malaria parasites reveals complex genomic architecture of emerging artemisinin resistance. Genome Biol. *18*, 78. https://doi.org/10.1186/s13059-017-1204-4.

Chevalier, F.D., Valentim, C.L., LoVerde, P.T., and Anderson, T.J. (2014). Efficient linkage mapping using exome capture and extreme QTL in schistosome parasites. BMC Genomics *15*, 1–12.

Clements, R.L., Streva, V., Dumoulin, P., Huang, W., Owens, E., Raj, D.K., Burleigh, B., Llinas, M., Winzeler, E.A., Zhang, Q., and Dvorin, J.D. (2020). A novel antiparasitic compound kills ring-stage Plasmodium falciparum and retains activity against artemisinin-resistant parasites. J. Infect. Dis. *221*, 956–962. https://doi.org/10.1093/infdis/jiz534.

Cowman, A.F., Tonkin, C.J., Tham, W.H., and Duraisingh, M.T. (2017). The molecular basis of erythrocyte invasion by malaria parasites. Cell Host Microbe *22*, 232–245. https://doi.org/10.1016/j.chom.2017.07.003.

Cui, L., Wang, Z., Miao, J., Miao, M., Chandra, R., Jiang, H., Su, X.z., and Cui, L. (2012). Mechanisms of *in vitro* resistance to dihydroartemisinin in Plasmodium falciparum. Mol. Microbiol. *86*, 111–128.

Davies, L., and Gather, U. (1993). The identification of multiple outliers. J. Am. Stat. Assoc. *88*, 782–792.

Davis, S.Z., Singh, P.P., Vendrely, K.M., Shoue, D.A., Checkley, L.A., McDew-White, M., Button-Simons, K.A., Cassady, Z., Sievert, M.A.C., Foster, G.J., et al. (2020). The extended recovery ring-stage survival assay provides a superior association with patient clearance half-life and increases throughput. Malar. J. *19*, 54. https://doi.org/10.1186/s12936-020-3139-6.

Delemarre, B.J., and van der Kaay, H.J. (1979). Tropical malaria contracted the natural way in The Netherlands. Ned Tijdschr Geneeskd *123*, 1981–1982.

Demas, A.R., Sharma, A.I., Wong, W., Early, A.M., Redmond, S., Bopp, S., Neafsey, D.E., Volkman, S.K., Hartl, D.L., and Wirth, D.F. (2018). Mutations in Plasmodium falciparum actin-binding protein coronin confer reduced artemisinin susceptibility. Proc. Natl. Acad. Sci. U S A. *115*, 12799–12804. https://doi.org/10.1073/pnas.1812317115.

Dia, A., and Cheeseman, I.H. (2021). Single-cell genome sequencing of protozoan parasites. Trends Parasitol. *37*, 803–814. https://doi.org/10.1016/j.pt.2021.05.013.

Ehrenreich, I.M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J.A., Gresham, D., Caudy, A.A., and Kruglyak, L. (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature *464*, 1039–1042.

Fidock, D.A., Nomura, T., Talley, A.K., Cooper, R.A., Dzekunov, S.M., Ferdig, M.T., Ursos, L.M., Naudé, B., Deitsch, K.W., and Su, X.-z. (2000). Mutations in the P. falciparum digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. Mol. Cell *6*, 861–871.

Geary, T.G., Divo, A.A., Bonanni, L.C., and Jensen, J.B. (1985). Nutritional requirements of Plasmodium falciparum in culture. III. Further observations on essential nutrients and antimetabolites. J. Protozool *32*, 608–613. https://doi.org/10.1111/j.1550-7408.1985.tb03087.x.

Gravenor, M., and Kwiatkowski, D. (1998). An analysis of the temperature effects of fever on the intra-host population dynamics of Plasmodium falciparum. Parasitology *117*, 97–105.

Hodel, E.M., Kay, K., and Hastings, I.M. (2016). Incorporating stage-specific drug action into pharmacological modeling of antimalarial drug treatment. Antimicrob. Agents Chemother. *60*, 2747–2756.

Hunt, P., Martinelli, A., Modrzynska, K., Borges, S., Creasey, A., Rodrigues, L., Beraldi, D., Loewe, L., Fawcett, R., Kumar, S., et al. (2010). Experimental evolution, genetic analysis and genome re-sequencing reveal the mutation conferring artemisinin resistance in an isogenic lineage of malaria parasites. BMC Genomics *11*, 499. https://doi.org/10.1186/1471-2164-11-499.

Khan, A.H., Lin, A., Wang, R.T., Bloom, J.S., Lange, K., and Smith, D.J. (2020). Pooled analysis of radiation hybrids identifies loci for growth and drug action in mammalian cells. Genome Res. *30*, 1458–1467. https://doi.org/10.1101/gr.262204.120.

Klonis, N., Xie, S.C., McCaw, J.M., Crespo-Ortiz, M.P., Zaloumis, S.G., Simpson, J.A., and Tilley, L. (2013). Altered temporal response of malaria parasites determines differential sensitivity to artemisinin. Proc. Natl. Acad. Sci. U S A. *110*, 5157–5162. https://doi.org/10.1073/pnas.1217452110.

Kumar, S., Haile, M.T., Hoopmann, M.R., Tran, L.T., Michaels, S.A., Morrone, S.R., Ojo, K.K., Reynolds, L.M., Kusebauch, U., Vaughan, A.M., et al. (2021). Plasmodium falciparum calcium-dependent protein kinase 4 is critical for male gametogenesis and transmission to the mosquito vector. mBio *12*, e0257521. https://doi.org/10.1128/mBio.02575-21.

Kumar, S., Li, X., McDew-White, M., Reyes, A., Sayeed, A., Haile, M., Kennedy, S., Carmago, N., Checkley, L., Vendrely, K., et al. (2020). Bulk segregant approaches to nutritional genomics in Plasmodium falciparum. Preprint at bioRxiv. https://doi.org/10.1101/2020.09.12.294736.

Li, H. (2011). A quick method to calculate QTL confidence interval. J. Genet. *90*, 355–360.

Li, X., Kumar, S., McDew-White, M., Haile, M., Cheeseman, I.H., Emrich, S., Button-Simons, K., Nosten, F., Kappe, S.H.I., Ferdig, M.T., et al. (2019). Genetic mapping of fitness determinants across the malaria parasite Plasmodium falciparum life cycle. Plos Genet. *15*, e1008453. https://doi.org/10.1371/journal.pgen.1008453.

Magwene, P.M., Willis, J.H., and Kelly, J.K. (2011). The statistics of bulk segregant analysis using next generation sequencing. Plos Comput. Biol. *7*, e1002255. https://doi.org/10.1371/journal.pcbi.1002255.

Malaria GEN Plasmodium falciparum Community Project (2016). Genomic epidemiology of artemisinin resistant malaria. Elife *5*, e08714. https://doi.org/10.7554/eLife.08714.

Mansfeld, B.N., and Grumet, R. (2018). QTLseqr: an R package for bulk segregant analysis with next-generation sequencing. Plant Genome *11*. https://doi.org/10.3835/plantgenome2018.01.0006.

Martinelli, A., Cheesman, S., Hunt, P., Culleton, R., Raza, A., Mackinnon, M., and Carter, R. (2005). A genetic approach to the de novo identification of targets of strain-specific immunity in malaria parasites. Proc. Natl. Acad. Sci. U S A. *102*, 814–819. https://doi.org/10.1073/pnas.0405097102.

Michelmore, R., Paran, I., and Kesseli, R. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci.U S A. *88*, 9828–9832.

Miles, A., Iqbal, Z., Vauterin, P., Pearson, R., Campino, S., Theron, M., Gould, K., Mead, D., Drury, E., O'Brien, J., et al. (2016). Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. Genome Res. *26*, 1288–1299. https://doi.org/10.1101/gr.203711.115.

Miotto, O., Amato, R., Ashley, E.A., MacInnis, B., Almagro-Garcia, J., Amaratunga, C., Lim, P., Mead, D., Oyola, S.O., Dhorda, M., et al. (2015). Genetic architecture of artemisinin-resistant Plasmodium falciparum. Nat. Genet. *47*, 226–234. https://doi.org/10.1038/ng.3189.

Modrzynska, K.K., Creasey, A., Loewe, L., Cezard, T., Borges, S.T., Martinelli, A., Rodrigues, L., Cravo, P., Blaxter, M., Carter, R., and Hunt, P. (2012). Quantitative genome re-sequencing defines multiple mutations conferring chloroquine resistance in rodent malaria. BMC Genomics *13*, 1–16.

Molina-Cruz, A., Garver, L.S., Alabaster, A., Bangiolo, L., Haile, A., Winikor, J., Ortega, C., van Schaijk, B.C., Sauerwein, R.W., Taylor-Salmon, E., and Barillas-Mury, C. (2013). The human malaria parasite Pfs47 gene mediates evasion of the mosquito immune system. Science *340*, 984–987.

Mukherjee, A., Bopp, S., Magistrado, P., Wong, W., Daniels, R., Demas, A., Schaffner, S., Amaratunga, C., Lim, P., Dhorda, M., et al. (2017). Artemisinin resistance without pfkelch13 mutations in Plasmodium falciparum isolates from Cambodia. Malar. J. *16*, 195. https://doi.org/10.1186/s12936-017-1845-5.

Nadaraya, E.A. (1964). On estimating regression. Theory Probab. Appl. *9*, 141–142. https://doi.org/10.1137/1109020.

Nair, S., Nkhoma, S.C., Serre, D., Zimmerman, P.A., Gorena, K., Daniel, B.J., Nosten, F., Anderson, T.J., and Cheeseman, I.H. (2014). Single-cell genomics for dissection of complex malaria infections. Genome Res. *24*, 1028–1038.

Nogami, S., Ohya, Y., and Yvert, G. (2007). Genetic complexity and quantitative trait loci mapping of yeast morphological traits. PLoS Genet. *3*, e31.

Ockenhouse, C.F., Sun, P.F., Lanar, D.E., Wellde, B.T., Hall, B.T., Kester, K., Stoute, J.A., Magill, A., Krzych, U., Farley, L., et al. (1998). Phase I/IIa safety, immunogenicity, and efficacy trial of NYVAC-Pf7, a pox-vectored, multiantigen, multistage vaccine candidate for Plasmodium falciparum malaria. J. Infect. Dis. *177*, 1664–1673. https://doi.org/10.1086/515331.

Park, Y., Gonzalez-Martinez, R.M., Navarro-Cerrillo, G., Chakroun, M., Kim, Y., Ziarsolo, P., Blanca, J., Canizares, J., Ferre, J., and Herrero, S. (2014). ABCC transporters mediate insect resistance to multiple Bt toxins revealed by bulk segregant analysis. BMC Biol. *12*, 46.

Preston, M.D., Campino, S., Assefa, S.A., Echeverry, D.F., Ocholla, H., Amambua-Ngwa, A., Stewart, L.B., Conway, D.J., Borrmann, S., Michon, P., et al. (2014). A barcode of organellar genome polymorphisms identifies the geographic origin of Plasmodium falciparum strains. Nat. Commun. *5*, 4052. https://doi.org/10.1038/ncomms5052.

Ramiro, R.S., Khan, S.M., Franke-Fayard, B., Janse, C.J., Obbard, D.J., and Reece, S.E. (2015). Hybridization and pre-zygotic reproductive barriers in Plasmodium. Proc. Biol. Sci. *282*, 20143027.

Sinka, M.E., Bangs, M.J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., Mbogo, C.M., Hemingway, J., Patil, A.P., Temperley, W.H., et al. (2012). A global map of dominant malaria vectors. Parasit Vectors *5*, 69. https://doi.org/10.1186/1756-3305-5-69.

Stokes, B.H., Dhingra, S.K., Rubiano, K., Mok, S., Straimer, J., Gnadig, N.F., Deni, I., Schindler, K.A., Bath, J.R., Ward, K.E., et al. (2021). Plasmodium falciparum K13 mutations in Africa and Asia impact artemisinin resistance and parasite fitness. Elife *10*, e66277. https://doi.org/10.7554/eLife.66277.

Su, X.-z., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., Wootton, J.C., and Wellems, T.E. (1999). A genetic map and recombination

parameters of the human malaria parasite Plasmodium falciparum. Science *286*, 1351–1353.

Tilley, L., Straimer, J., Gnadig, N.F., Ralph, S.A., and Fidock, D.A. (2016). Artemisinin action and resistance in Plasmodium falciparum. Trends Parasitol. *32*, 682–696. https://doi.org/10.1016/j.pt.2016.05.010.

Tyagi, R.K., Gleeson, P.J., Arnold, L., Tahar, R., Prieur, E., Decosterd, L., Perignon, J.L., Olliaro, P., and Druilhe, P. (2018). High-level artemisinin-resistance with quinine co-resistance emerges in P. falciparum malaria under *in vivo* artesunate pressure. BMC Med. *16*, 181. https://doi.org/10.1186/s12916-018-1156-x.

Vaughan, A.M., Pinapati, R.S., Cheeseman, I.H., Camargo, N., Fishbaugher, M., Checkley, L.A., Nair, S., Hutyra, C.A., Nosten, F.H., Anderson, T.J., et al. (2015). Plasmodium falciparum genetic crosses in a humanized mouse model. Nat. Methods *12*, 631–633. https://doi.org/10.1038/nmeth.3432.

Vendrely, K.M., Kumar, S., Li, X., and Vaughan, A.M. (2020). Humanized mice and the rebirth of malaria genetic crosses. Trends Parasitol. *36*, 850–863. https://doi.org/10.1016/j.pt.2020.07.009.

Wang, Z., Cabrera, M., Yang, J., Yuan, L., Gupta, B., Liang, X., Kemirembe, K., Shrestha, S., Brashear, A., Li, X., et al. (2016). Genome-wide association analysis identifies genetic loci associated with resistance to multiple antimalarials in Plasmodium falciparum from China-Myanmar border. Sci. Rep. *6*, 33891. https://doi.org/10.1038/srep33891.

Wellems, T.E., Panton, L.J., Gluzman, I.Y., Do Rosario, V.E., Gwadz, R.W., Walker-Jonah, A., and Krogstad, D.J. (1990). Chloroquine resistance not linked to mdr-like genes in a Plasmodium falciparum cross. Nature *345*, 253–255.

Wilson, D.W., Langer, C., Goodman, C.D., McFadden, G.I., and Beeson, J.G. (2013). Defining the timing of action of antimalarial drugs against Plasmodium falciparum. Antimicrob. Agents Chemother. *57*, 1455–1467.

Witkowski, B., Amaratunga, C., Khim, N., Sreng, S., Chim, P., Kim, S., Lim, P., Mao, S., Sopha, C., Sam, B., et al. (2013). Novel phenotypic assays for the detection of artemisinin-resistant Plasmodium falciparum malaria in Cambodia: in-vitro and ex-vivo drug-response studies. Lancet Infect. Dis. *13*, 1043–1049. https://doi.org/10.1016/S1473-3099(13)70252-4.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
O+ human red blood cells	Interstate Blood Bank (Memphis, TN) & Valley Biomedical (Winchester, VA)	N/A
NHP1337 [cloned *P. falciparum* sample]. Originally cloned from a blood sample collected from uncomplicated *P. falciparum* malaria patient on the Thailand-Myanmar border.	Shoklo Malaria Research Unit, Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Mae Sot, Thailand.	N/A
KH004 [cloned *P. falciparum* sample]. Originally cloned from a blood sample collected from uncomplicated *P. falciparum* malaria patient from Cambodia.	Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand	N/A
MAL31 [cloned *P. falciparum* sample] Originally cloned from a blood sample collected from uncomplicated *P. falciparum* malaria patient from Malawi	Malawi-Liverpool Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Blantyre, Malawi	N/A
Chemicals, peptides, and recombinant proteins		
DHA (dihydroartemisinin)	Sigma Aldrich	Cat# D7439
RPMI 1640 with L-glutamine	Gibco, Life Technologies	Cat# 31800022
50 mg/L hypoxanthine	Millipore, Fisher Scientific	Cat# 4010CBC
25 mM HEPES	Corning, Fisher Scientific	Cat# 61034RO
0.5% Albumax II	Gibco, Life Technologies	Cat# 11021045
10 ug/mL gentamycin	Gibco, Life Technologies	Cat# 15710072
0.225% Sodium bicarbonate	Corning, VWR	Cat# 25-035-CI
Glycerolyte	Medline	Cat# FWL4A7831
1X PBS (phosphate buffer saline)	Fisher Scientific	Cat# PB665-1
70% Percoll	Sigma	Cat# P1644
13.3% sorbitol	Sigma	Cat# S1876
SYBR Green I	Life Technologies	Cat# S785
SYTO 61	Life Technologies	Cat# S11343
Phusion Blood Direct PCR kit	Life Technologies	Cat# F547L
KAPA HyperPlus Kit	Roche	Lot 0000123512
NEXTflex Dual-Indexed DNA Barcodes (1-96)	PerkinElmer	Cat# NOVA-514160
KAPA Pure Beads	Roche	Cat# 0798328001
Quant-iT PicoGreen dsDNA Assay Kit	Invitrogen	Cat# Q33120
KAPA Library Quant Kit (Illumina)	Roche	Cat# 07960140001
D1000 ScreenTape	Agilent Technologies	Cat# 5067-5584
Human Serum	Valley Biomedical	HS10040PM
Deposited data		
Raw sequencing data	NABI Sequence Read Archive	PRJNA524855
Experimental models: Parasite cell lines		
P. falciparum NF54	Delemarre and van der Kaay, 1979 Ockenhouse et al., 1998	https://pubmed.ncbi.nlm.nih.gov/390409/ https://pubmed.ncbi.nlm.nih.gov/9607847/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Prism v9	GraphPad	https://www.graphpad.com/
R	R-Project	https://www.r-project.org
QTLseqr	R package	https://github.com/bmansfeld/QTLseqr

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Michael T. Ferdig (mferdig@nd.edu).

Materials availability

All unique reagents generated in this study are available from the lead contact without restriction.

Data and code availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the supplemental information. All raw sequencing data have been deposited at the NABI Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra) under the project number of PRJNA524855 and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request. All original code has been deposited at github (https://github.com/emilyli0325/BSA.optimization) and is publicly available.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

FRG NOD huHep mice with human chimeric livers were used in this study to generate *Plasmodium falciparum* genetic crosses. The study was performed in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health (NIH), USA. To this end, the Seattle Children's Research Institute (SCRI) has an Assurance from the Public Health Service (PHS) through the Office of Laboratory Animal Welfare (OLAW) for work approved by its Institutional Animal Care and Use Committee (IACUC). All the work carried out in this study was specifically reviewed and approved by the SCRI IACUC.

Human subjects

Blood was obtained from healthy (no diagnosis of disease) from Interstate Blood Bank (Memphis, TN) or Valley Biomedical (Winchester, VA) through protocols approved by the relevant institutional review boards.

Plasmodium falciparum parasite clones/strains

The four clones used to generate genetic crosses were KH004, Mal31, NHP1337, and NF54 (clone 3D7). KH004 was isolated from a patient in Western Cambodia in 2016, Mal31 was isolated from a Malawian patient in 2016, and NHP1337 was isolated from a patient in Thailand in 2011. KH004, Mal31 and NHP1337 were each cloned by limiting dilution to obtain genetically homogenous clones prior to being used as parents in the genetic crosses. For this study, we used a clone of NF54, 3D7, which was isolated in 1995 from a volunteer in a malaria clinical trial after controlled human malaria infection by 3D7-infected *Anopheles stephensi* mosquito bite (Ockenhouse et al., 1998). KH004 and Mal31 were cloned at Notre Dame in October 2019, NHP1337 was cloned at Seattle Children's Research Institute in August 2018 (Figure S1).

METHOD DETAILS

Preparation of the genetic crosses

Crosses were generated using FRG NOD huHep mice with human chimeric livers and *A. stephensi* mosquitoes as described by Vaughan et al., (Vaughan et al., 2015). Two genetic crosses were generated for this study: Mal31×KH004 and NF54×NHP1337. Three individual recombinant pools were generated for

each cross by using different cages of infected mosquitoes. To start each cross, gametocytes from both parental parasite strains were diluted to 0.5% gametocytemia in a human serum erythrocyte mix, to generate infectious blood meals (IBMs). IBMs from each parent were mixed at equal proportions and fed to three cages of mosquitos (150 per cage).

The mosquito infection rate and oocyst number per infected mosquito were determined 7 days post-feeding by randomly choosing and dissecting 15 mosquitoes under microscopy. Sporozoites were isolated from infected mosquito salivary glands and 2-4 million sporozoites from each independent recombinant pool of mosquitoes were injected intravenously into three FRG huHep mice (Mal31×KH004) and two FRG huHep mice (NF54×NHP1337). To allow the liver stage-to-blood stage transition, mice are infused with human erythrocytes six and seven days after sporozoite injection. Four hours after the second infusion, the mice were euthanized and exsanguinated to isolate the circulating ring stage *P. falciparum*-infected human erythrocytes. The parasites from each mouse constitute the initial recombinant pools of recombinant progeny for genetic mapping experiments. Initial recombinant progeny pools were maintained in AlbuMAX (Gibco, Life Technologies) supplemented RPMI media; we genome sequenced aliquots from each pool to check allele frequencies from both parents. Two recombinant pools from each cross were used for DHA BSA experiments.

Cryopreservation and thawing of progeny bulks

We expanded recombinant progeny pools by growing in AlbuMAX supplemented RPMI for 7 days after the mouse exsanguination prior to being set up in DHA BSA. To test the effect of cryopreservation on DHA BSA, when recombinant progeny pools were 3% parasitemia and at least 70% ring-stage, recombinant progeny pool stocks were cryopreserved in glycerolyte (Medline, Mundelein IL). Each 0.5 mL cryopreserved stock contained ~100 μL packed red cells. Cryopreserved stocks of the KH004 cloned parent and the Mal31×KH004 uncloned bulk recombinant progeny pool and were thawed and grown in complete media (CM) [RPMI 1640 with L-glutamine (Gibco, Life Technologies), 50 mg/L hypoxanthine (Millipore, Fisher Scientific), 25 mM HEPES (Corning, VWR), 0.5% Albumax II (Gibco, Life Technologies), 10ug/mL gentamycin (Gibco, Life Technologies), 0.225% sodium bicarbonate (Corning, VWR) at 5% hematocrit in O^+ red blood cells (RBC) (Interstate Blood Bank, Memphis TN). Cultures were grown in separate flasks and maintained at constant pH, 7.0–7.5, temperature, 37°C, and atmosphere, 5% CO_2/5% O_2/90% N_2. Cultures were kept below 2% parasitemia with media changes every 48 h. Recombinant progeny pools were thawed from cryopreserved stocks and cultured for 9 days prior to set-up in BSA experiments.

Measuring DHA resistance level using the eRRSA

Parasites were synchronized to late stage schizonts using a single-layer 70% Percoll density gradient (Davis et al., 2020). 350 μL of packed, infected erythrocytes at high schizogony (>50% schizonts) was suspended in 2 mL of RPMI. Cultures were layered over a single 70% Percoll layer (Sigma-Aldrich, St. Louis MO) in 1× RPMI and 13.3% sorbitol in phosphate buffer saline (PBS) (Fisher Scientific, Hanover Park IL) and centrifuged (1561×*g* for 10 min, no brake). The top layer (infected late stage schizonts) was removed and washed with 10 mL of RPMI twice. Cultures were suspended in 2 mL of complete media (CM) at 2% hematocrit and placed in culture flasks on a shaker in a 37°C incubator for 4 h to allow for re-invasion. Four hours post-synchronization, parasitemia and stage were quantified using flow cytometry. 80 μL of culture and a 2% hematocrit RBC control were stained with SYBR Green I (Life Technologies, Grand Island NY) and SYTO 61 (Life Technologies) and measured on a Guava easyCyte HT (Luminex Co.). 50,000 events were recorded to determine relative parasitemia and stage. Analysis was performed with guavaSoft version 3.3 (Luminex Co.). When at least 70% of parasites were in the ring-stage, the culture was diluted to 2% hematocrit and 0.5% parasitemia and parasites were aliquoted into a 96-well plate. Ten DHA (Sigma-Aldrich) concentrations were tested with a 2-fold dilution series starting with 2800 nM DHA. eRRSA assay set-up and qPCR amplification was done as previously described in (Davis et al., 2020) to calculate a fold change value between the untreated and treated sample for three biological replicates for each parasite. Briefly, qPCR was performed using the Phusion Blood Direct PCR kit (ThermoFisher, cat #F547L), supplemented with 1× SYBR. 3 μL of culture was used in a 10 μL reaction and amplified using forward and reverse primers of the *Pfcrt* gene. PCR amplification was measured using the fast mode of the ABI 7900HT, with a 20 s denaturation at 95°C, followed by 30 cycles of 95°C for 1 s, 62.3°C for 30 s, and 65°C for 15 s. Cycle threshold (Ct) values were calculated using the ABI SDS 2.4.1. Fold change ($2^{\Delta Ct}$) was calculated by determining the average ΔCt Ct for the three technical replicates for the untreated and treated samples by applying the following equation: Fold change = $2^{(\text{average Ct of treated sample}-\text{average Ct of untreated sample})}$. GraphPad Prism

version 9 was used to perform all statistics and generate figures. This data generated an eRRSA dose-response curve based on 10 DHA concentrations.

DHA BSA

Two recombinant pools from each cross (Mal31×KH004 or NF54×NHP1337) were used for DHA BSA experiments. For each pool of recombinants, one sample was taken from the original recombinant progeny pool and was not synchronized (unsynchronized control). To test the effect of synchronization on the recombinant pools (across the 48-h *P. falciparum* life cycle), three samples were taken from the original recombinant progeny pool over time and synchronized. Notably, the original recombinant progeny pool was an asynchronous culture. The first sample was taken from the original asynchronous recombinant progeny pool and synchronized when more than 50% of the pool was late-stage parasites (called 0 h), at the same time as the unsynchronized control (which was taken from the original asynchronous recombinant progeny pool and not synchronized). 18 h later, a second sample was taken from the original asynchronous recombinant pool and synchronized (18 h). And 36 h after the first synchronization (18 h after the second synchronization), a third and final sample was taken from the original asynchronous recombinant pool and synchronized (36 h) (Figure 3A). For each synchronization a single 70% Percoll density gradient was used (as described above for the eRRSA, this density gradient selects late-stage parasites) and allowed late-stage parasites to reinvade for 4 h after synchronization to ring stages before determining parasitemia and stage using flow cytometry and setting up the DHA BSA experiment. According to flow cytometry of the progeny pool synchronized at 0 h, the population was 91% ring stage; at 18 h 89% of the synchronized population was ring stage, and at 36 h 65% of the synchronized population was ring stage, and the unsynchronized population was 68% ring stage. All progeny pools (0, 18, and 36 h synchronized and unsynchronized control) were diluted to 1% parasitemia and were treated with DMSO (control), 50 nM DHA, or 100 nM and maintained as 0.5 mL cultures in 48-well plates. Six hours after assays were treated with DMSO or DHA, infected RBCs were washed three times with RPMI to remove any residual DHA or DMSO and then resuspended with new culture media to allow surviving parasites to grow. Samples for library prep and sequencing were collected right before drug treatment (day 0) and 5 days later (day 5) when there was enough parasite material for DNA extraction.

Library preparation and sequencing

Qiagen DNA mini kits were used to extract and purify the genomic DNA and Quant-iT™ PicoGreen® Assay (Invitrogen) was used to quantify the amount of DNA. For samples with less than 50 ng DNA obtained, whole genome amplification (WGA) was performed before NGS library preparation. WGA reactions were performed following Nair et al. (Nair et al., 2014). Each 25 μL reaction contained at least 5 ng of *Plasmodium* DNA, 1× BSA (New England Biolabs), 1 mM dNTPs (New England Biolabs), 3.5 μM of Phi29 Random Hexamer Primer, 1× Phi29 reaction buffer (New England Biolabs), and 15 units of Phi29 polymerase (New England Biolabs). We used a PCR machine (SimpliAmp, Applied Biosystems) programmed to run a "stepdown" protocol: 35°C for 10 min, 34°C for 10 min, 33°C for 10 min, 32°C for 10 min, 31°C for 10 min, 30°C for 6 h then heating at 65°C for 10 min to inactivate the enzymes prior to cooling to 4°C. Samples were cleaned with AMPure XP Beads (Beckman Coulter) at a 1:1 ratio.

Next generation sequencing (NGS) libraries were constructed using 50–100 ng DNA or WGA product following the KAPA HyperPlus Kit protocol with 3-cycle of PCR. All libraries were sequenced at 150 bp pair-end using Illumina Novaseq S4 or Hiseq X sequencers. All bulk were sequenced to a minimum coverage of 100×.

Mapping and genotyping

We individually mapped whole-genome sequencing reads for each library against the *P. falciparum* 3D7 reference genome (PlasmoDB, release32) using the alignment algorithm BWA mem (http://bio-bwa.sourceforge.net/) under the default parameters. The resulting alignments were then converted to SAM format, sorted to BAM format, and deduplicated using picard tools v2.0.1 (http://broadinstitute.github.io/picard/). We used Genome Analysis Toolkit GATK v3.7 (https://software.broadinstitute.org/gatk/) to recalibrate the base quality score based on a set of verified known variants (Miles et al., 2016).

After alignment, we excluded the highly variable genome regions (subtelomeric repeats, hypervariable regions, and centromeres) and only performed genotype calling in the 21 Mb core genome (defined in (Miles et al., 2016)). We called variants for each sample using HaplotypeCaller and calls from every 100 samples

were merged using CombineGVCFs with default parameters. Variants were further called at all sample-level using GenotypeGVCFs, with parameters: –max_alternate_alleles 6 –variant_index_type LINEAR –variant_index_parameter 128000 –sample_ploidy 2 -nt 20. We further filtered the variants calls by calculating the recalibrated variant quality scores (VQSR) of genotypes from parental parasites. Loci with VQSR less than 1 or not distinguishable between two parents were removed from further analysis.

The variants in VCF format were annotated for predicted functional effect on genes and proteins using snpEff v4.3 (https://pcingola.github.io/SnpEff/) with 3D7 (PlasmoDB, release32) as the reference.

Bulk segregant analysis

Only single-nucleotide polymorphism (SNP) loci with coverage >30× that differed between the two parents were used for bulk segregant analysis. We counted reads with genotypes of each parent and calculated allele frequencies at each variable locus. Allele frequencies of one of the parents (NF54 or Mal31 in this study) were plotted across the genome, and outliers were removed following Hampel's rule (Davies and Gather, 1993) with a window size of 100 loci. We performed the BSA analyses using the R package QTLseqr (Mansfeld and Grumet, 2018). We first calculated the G value for each of the SNP locus following equation in (Magwene et al., 2011). We then tricube-smoothed (Nadaraya, 1964) the G statistic (G prime) with a window size of 100 kb to remove noise. FDRs (Benjamini-Hochberg adjusted p-values) were calculated for each SNP based on the distribution of G prime, and extreme-QTLs were defined when FDR were less than 0.01 (Magwene et al., 2011). Once a QTL was detected, we calculated the approximated 95% confidence interval using Li's method (Li, 2011) to localize causative genes.

QUANTIFICATION AND STATISTICAL ANALYSIS

All quantification and statistical analyses were performed as described in the Mapping and genotyping section and Bulk segregant analysis section of the STAR Methods.