Published in final edited form as:

Trends Parasitol. 2020 October ; 36(10): 850–863. doi:10.1016/j.pt.2020.07.009.

Humanized Mice and the Rebirth of Malaria Genetic Crosses

Katelyn M. Vendrely^{1,#}, Sudhir Kumar^{2,#}, Xue Li^{3,#}, Ashley M. Vaughan^{2,4,*}

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

²Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA, USA.

³Disease Intervention and Prevention Program, Texas Biomedical Research Institute, San Antonio, TX, USA

⁴Department of Pediatrics, University of Washington, Seattle, WA, USA

Abstract

The first experimental crosses carried out with the human malaria parasite *Plasmodium falciparum* played a key role in determining the genetic loci responsible for drug resistance, virulence, invasion, growth rate and transmission. These crosses relied on splenectomized chimpanzees to complete the liver stage of the parasite life cycle and the subsequent transition to asexual blood stage culture followed by cloning of recombinant progeny *in vitro*. Crosses can now be routinely carried out using human-liver chimeric mice infused with human erythrocytes to generate hundreds of unique recombinant progeny for genetic linkage mapping, bulk segregant analysis, and high throughput 'omics readouts. The high number of recombinant progeny should allow for unprecedented power and efficiency in the execution of a system genetics approach to study *P. falciparum* biology.

Keywords

malaria; genotype; phenotype; linkage; cross; recombinant progeny

In an era of modern genomics techniques, why use a classical genetics approach?

One of the primary goals of biological research is to determine the molecular mechanisms that link genotypes to phenotypes. In an era of new genomics technologies that are revolutionizing the way we understand the genetic basis of phenotypic traits [e.g., whole genome sequencing, CRISPR/Cas9 genome editing and **genome-wide association studies**

*Correspondence. ashley.vaughan@seattlechildrens.org (A. M. Vaughan).

#Contributed equally to this review

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

(GWAS, see **Glossary**)), more traditional methods such as **genetic linkage mapping** might be overlooked. For a recent review that highlights the advantages and disadvantages of genetic linkage mapping and GWAS, please see [1]. Ultimately, genetic linkage mapping is an approach that cannot be replaced. First, genetic linkage mapping is a powerful phenotype-driven approach that requires no *a priori* knowledge or hypothesis regarding the genetic basis of the phenotype. Thus, recombinant progeny clones generated from an experimental cross (hereafter referred to simply as ‘cross’ and ‘crosses’) contain the complete repertoire of genetic information responsible for observed phenotypes. Second, even though GWAS can be a powerful approach to provide insights into the **genetic architecture** of the trait, it is an *a posteriori* approach that requires a large number of historical samples which take time to acquire and is susceptible to both false positive and false negative results. Conversely, genetic linkage mapping can be used to study ongoing selection (including drug selection) in the field. Third, GWAS is not able to identify interactions between loci, while genetic linkage mapping can be used to uncover gene interactions that are responsible for complex phenotypes. Genetic linkage mapping is thus an unbiased approach and can reveal the genetic effects that control complex traits with complex genetic architectures. Therefore, with sufficient numbers of unique recombinant progeny, each with a distinct **genetic background**, genetic linkage mapping should be able to uncover genes that underlie any measurable phenotype of interest.

Modern genomic techniques and classic genetic approaches are complementary as advances in rapid and cost-effective sequencing have revolutionized the ability to map quantitative traits. These same advances have also increased the feasibility of GWAS for identifying loci of interest that are associated with a phenotype. Genetic linkage mapping can then be used for a deeper investigation of loci identified by GWAS, particularly the search for secondary loci and gene interactions that are responsible for complex phenotype. With the knowledge of these loci and their interactions, recombinant progeny from crosses with known genotypes can then be chosen for genetic editing, e.g. using CRISPR/Cas9, to confirm the genetic basis of the phenotype, an approach that is not possible with GWAS or gene manipulation studies alone.

Setting the stage for the first *P. falciparum* cross

The human malaria parasite *P. falciparum* is a powerful model system for generating crosses for downstream analysis. *P. falciparum* (i) has a well-annotated and small (~23 Mb) reference genome [3], (ii) has a vast database of supporting functional information and associated population data [4], (iii) is haploid throughout the majority of the life cycle which simplifies phenotyping through the cloning and amplification of asexual blood stage parasites, and (iv) has a high recombination frequency (Table 1) [5], compared to other eukaryotes [6] which is ideal for pinpointing causal gene locations. However, creating a cross in any organism is not trivial, especially in an organism with a complex life cycle that involves two hosts: female *Anopheles* mosquitoes and humans. The first crosses between malaria species were carried out in the rodent malaria species *Plasmodium yoelii* and *Plasmodium chabaudi* using rats. These crosses proved that recombination between gametes occurs in the mosquito, the blood stages of are haploid, and resistance to pyrimethamine and chloroquine (CQ) are due to point mutations in loci responsible for resistance [7, 8]. These

crosses thus revealed that conventional genetic analysis could be used for malaria parasites and set the stage for the first *P. falciparum* cross. Before the first *P. falciparum* cross, several major developments needed to occur. These included *in vitro* methods for *P. falciparum* asexual stage culture, allowing parasites to be maintained and manipulated indefinitely [9] which laid the foundation for all further *in vitro* *P. falciparum* laboratory work, including methods for isolating individual asexual parasite clones [10] and maturing *P. falciparum* sexual stages (gametocytes) in culture. This allowed for transmission to *Anopheles* mosquitoes to create sporozoites [11] and ultimately, the successful transmission of sporozoites from infected mosquitoes to splenectomized chimpanzees (*Pan troglodytes*) and South American owl monkeys (*Aotus trivirgatus*) [12]. Thus, the complete *P. falciparum* life cycle could be maintained in a laboratory setting without the use of human volunteers.

The past: crosses in non-human primates

Four crosses were conducted in non-human primates (NHPs) (Table 1) and analysis of their recombinant progeny have led to significant advances in malaria parasite biology.

1. 3D7 (The Netherlands, clone of NF54, African origin) × HB3 (Honduras)

This was the first controlled *P. falciparum* cross and demonstrated that progeny can be generated using an NHP model [13, 14]. Gametocytes were cultured from each parent and fed to *Anopheles freeborni* mosquitoes. Recombinant sporozoites from these mosquitoes were used to infect a splenectomized chimpanzee via both mosquito bite and intravenous injection. Blood stage parasites were adapted to *in vitro* human erythrocyte culture and the progeny were cloned. This early study was limited by the availability of genetic markers and relied upon restriction fragment length polymorphism (RFLP) for analysis [13]. Despite this cross only generating 15 unique recombinant progeny, genetic linkage mapping was used to provide evidence that resistance to the antimalarial drug pyrimethamine was due to a point mutation in the gene encoding dihydrofolate-reductase thymidylate synthase [14].

2. HB3 (Honduras) × Dd2 (Laos)

The second cross was performed between CQ sensitive HB3 and CQ resistant Dd2 parasites to understand the genetic basis of CQ resistance [15]. The progeny were genotyped using RFLP and microsatellite (MS) analysis [2, 16]. The cross initially yielded 16 unique recombinant progeny and a plus-minus phenotype found linkage at chromosome 7 and demonstrated that CQ resistance was not linked to *mdr*-like genes, which had previously been hypothesized [15]. Later, after a massive cloning effort, a total of 35 (an additional 19) unique recombinant progeny were isolated and a high-resolution linkage map was generated using MS analysis [2, 16], narrowing the CQ resistance locus to 36 kb on chromosome 7. Further fine mapping of CQ resistance finally linked resistance to a mutation in a gene that was ultimately named the CQ resistance transporter, *PfCRT* [17]. The progeny from this cross supported further discoveries including linking sulfadoxine resistance to mutations in dihydropteroate synthetase [18], identification of recombination in telomeric clusters of *var* genes [19], linking multiple genes to low-level quinine resistance [20], identification of the role of *PfMDV1* in gametocyte membrane formation and integrity [21], and linking *Pfclag3* genes to nutrient uptake by the plasmodial surface anion channel [22]. These studies

demonstrate the power of classical genetics to discover the mechanisms by which genes govern critical phenotypes. Furthermore, the MS analysis from this cross laid the groundwork for the sequencing of the *P. falciparum* genome [3].

3. 7G8 (Brazil) × GB4 (Ghana)

The third cross was primarily carried out to understand erythrocyte invasion phenotypes [23]. The isolated progeny were genotyped using 285 MS markers [2] followed by high-density oligonucleotide arrays for higher resolution mapping and 32 unique recombinant progeny were obtained. From this cross, *P. falciparum* reticulocyte binding protein homologue 5 (*PfRh5*) was identified to be a key molecule for infection of *Aotus nancymae* erythrocytes [23] and *PfRh5* was later shown to bind the Basigin receptor on erythrocytes during human erythrocyte invasion [24]. Progeny from this cross also identified the role of *Pfs47* in the evasion of the mosquito immune system [25] and revealed that *PfCRT* and *PfMDR1* are involved in the modulation of parasite responses to the antimalarial drugs amodiaquine and CQ [26]. Additionally, linkage analysis on progeny from this cross demonstrated that the *Pfclag3* is dispensable under standard *in vitro* conditions for asexual blood stage replication but required for propagation under physiological conditions [27].

4. GB4 (Ghana) × 803 (Cambodia)

The fourth and final *P. falciparum* cross was carried out to investigate response phenotypes of the antimalarial drug artemisinin (ART) *in vitro* and *in vivo* in *Aotus* monkeys and, after the independent discovery that mutations in *PFK13* were associated with ART resistance [28, 29], to evaluate phenotypes attributed to the *PFK13* C580Y mutation [30]. In this cross, seven recombinant progeny pools were collected 18, 21, 22, 25, 28, 36, and 40 days post-inoculation of sporozoites and were genotyped using DNA microarrays. From these pools, a total of 425 progeny were cloned, resulting in the isolation of 27 unique recombinants. From these, 20 were used for *in vitro* IC₅₀ measurements of dihydroartemisinin and all showed similar low nanomolar levels, whereas, in a ring-stage survival assay, the C580Y-type progeny had 19.6-fold higher average counts [30].

5. Some key lessons learned from the four NHP crosses

These crosses have been invaluable in their ability to not only uncover genes involved in drug resistance [17, 20, 26, 31–34], but also genes contributing to parasite growth and invasion of erythrocytes [23, 35, 36], parasite transmission [21, 37], and loci affecting metabolite production [38] and gene expression [39], and revealing that indels, structural variation, and recombination drive genomic diversity in *P. falciparum* [5]. Much remains to be discovered using genetic linkage mapping but the success of this approach also depends on certain parameters that include the genetic backgrounds of the parents, the difference in the phenotype of interest between the two parents, the ability to accurately measure the phenotype, and the number of unique recombinant progeny; each of these will impact the ability to draw conclusions from mapping. Moving forward, the ability to create new *P. falciparum* crosses routinely and the generation of large numbers of recombinant progeny will likely allow this invaluable method to continue to connect genotypes to phenotypes.

The present: crosses using the FRG huHep mouse

A major obstacle in performing crosses in *P. falciparum* is the restriction of liver stage development to humans and NHPs [40]. However, after the USA National Institutes of Health and European research organizations banned chimpanzee research for ethical and financial reasons, a human hepatocyte-liver chimeric mouse model (the **FRG huHep mouse**) was adopted for the generation of recombinant *P. falciparum* progeny was developed (Figure 1) [41]. The FRG huHep mouse lacks the fumaryl acetoacetate hydrolase (FAH) gene (F designation) and this causes hepatocyte cell death, which is controlled by the drug nitisinone. The mouse also lacks Rag-2 (R Designation) and the interleukin 2 receptor common gamma chain (IL-2R γ , G designation) resulting in adaptive immune deficiency and thus, an inability to reject foreign tissue [42, 43]. Mice are injected with human hepatocytes and cycled on and off nitisinone [42], which allows the human hepatocytes to replace the mouse hepatocytes that die when nitisinone is removed. Human erythrocytes are injected into the mice to allow liver stage-to-blood stage transition and subsequent *in vitro* adaptation and downstream cloning of recombinant progeny.

To date, four crosses have been performed using the FRG huHep mouse [41, 44] and the ability to use multiple mice for a cross should ultimately allow for the recovery of 100's of unique recombinant progeny, allowing for robust downstream genetic linkage mapping, a huge advantage over chimpanzee crosses. The first two crosses [41], utilized parasites previously used in the chimpanzee crosses and the third and fourth crosses utilized parasites recently isolated from the field [41, 44].

1. NF54HT-GFP-luc (Africa origin) × GB4 (Ghana)

The first cross [41] was generated between CQ resistant GB4 and NF54HT-GFP-luc, which expresses a green fluorescent protein (GFP)-luciferase (luc) fusion cassette harboring a mutant dihydrofolate reductase, resulting in resistance to the antifolate WR99210 [41]. Gametocytes from both parents were mixed and fed to *An. stephensi* mosquitoes and sporozoites were injected into two mice intravenously. The GFP-luc cassette allows for liver stage development to be detected after luciferin injection into the mice and use of an *in vivo* imaging system. This cross was set up specifically to test if recombinant progeny could be isolated and thus transitioned blood stage parasites were selected with both CQ and WR99210 since only recombinant parasites could survive this drug pressure. A proportion of the parasites that survived the pressure were cloned and indeed shown to harbor both the GFP-luc cassette and the K76T mutation associated with CQ resistance. This cross supported using FRG huHep mice for the isolation of recombinant progeny.

2. NF54HT-GFP-luc (Africa origin) × 7G8 (Brazil)

This cross [41] was set up to determine if increased numbers of recombinant parasites could be isolated and cloned. A limited cloning attempt isolated 44 clones, of which 15 were unique recombinants and interestingly, none of the clones were parental (Table 1), suggesting that outcrossing was favored in this cross. This result demonstrated the power of the FRG huHep mouse in the isolation of recombinant progeny.

3. NF54 (Africa origin) × NHP4026 (Thailand)

This cross [41, 44] was performed between NF54 and a recently isolated parasite, NHP4026, that cleared extremely slowly following ART treatment in a patient, but interestingly does not contain any mutations in the *PfK13* open reading frame [29]. This allopatric (parental parasites from different geographical regions) cross was initially carried out to investigate if genomic loci other than *PfK13* were linked to ART resistance. MS analysis of the initial 25 cloned progeny from a single FRG huHep mouse infection, revealed 20 independent unique progeny and few parental genotypes. To estimate the total number of progeny present in the parasite pool, **rarefaction** was used, a technique commonly used by ecologists to estimate the number of species present in a community [45]. Rarefaction estimated that 72 unique progeny could be isolated from 100 cloned parasites, far exceeding the number of progeny isolated from any cross using the chimpanzee. This cross was repeated, using increased numbers of sporozoites and three FRG huHep mice. Genomic sequencing of recombinant progeny confirmed that there was significant deviation from the Mendelian expectation of 1:1 inheritance of genomic regions on chromosomes 7, 12, 13 and 14 [44]. Cloning from the four total mice used for this cross has to date isolated 128 recombinant progeny of which 84 are unique. Interestingly, there was again, a strong bias for the isolation of recombinant progeny (Table 1) and 82% of the recombinants had skewed inheritance of their cytoplasmic genomes (mitochondria and apicoplast) from the NHP4026 parent [44]. Further studies are underway to understand the propensity for outcrossing and the skewed inheritance of the cytoplasmic genomes.

4. MKK2835 (Thailand) × NHP1337 (Thailand)

This sympatric cross [44] was performed between two recently isolated parasites from Southeast Asia to understand the fitness costs associated with the ubiquitous C580Y mutation in *PfK13* that is associated with ART resistance. MKK2835 is wildtype at *PfK13* and was isolated before heavy usage of ART and NHP1337 harbors the C580Y *PfK13* mutation. For this cross, sporozoites were initially injected into a single FRG huHep mouse and multiple rounds of cloning were initiated to maximize the number of recombinant clones obtained. 266 prospective progeny were identified during cloning and direct sequencing of clones was performed, leading to the identification of 60 unique recombinants. In contrast to the previous crosses carried using the FRG huHep mouse, cloning resulted in the isolation of a large number of selfed parasites, 144 selfed NHP1337 parents and five MKK2835 parents. In this cross however, parental alleles were inherited in statistically equal proportions across the genome in the recombinant progeny alongside an unbiased inheritance of the cytoplasmic genomes [44]. Progeny from this cross were also subjected to **Bulk Segregant Analysis** (see below) to understand parasite fitness during *in vitro* asexual blood stage culture [46].

Measuring phenotypes in *P. falciparum*

It is especially important to understand the genetic basis of antimalarial drug resistance since drug failure can lead to a massive loss of life. Indeed, the rapid spread of ART resistance in Southeast Asia and its possible migration to sub-Saharan Africa could derail malaria eradication efforts. Mutations in *PfK13* along with other permissive mutations in the genetic

background of resistant parasites have been implicated in ART resistance and the transcriptomic analysis of 1043 *P. falciparum* isolates indicates an association between decreased clearance times and upregulation of genes involved in the unfolded protein response [47]. However, global proteomics analyses of ART resistant and wildtype parasites have revealed that the basal level abundance and rate of turnover of unfolded protein response proteins is unchanged in resistant parasites [48, 49]. Furthermore, drug resistant parasite are often less fit [50, 51], compared to susceptible parasites when drug pressure is removed. Thus, drug resistance phenotypes are extremely complex.

Fortuitously, because crosses between haploid parental parasites generate haploid recombinant progeny, an integrative ‘omics approach that combines RNA-Seq, proteomics and metabolomics can be used to uncover a wide range of genetic mechanisms that underpin drug resistance. Although aspirational at present, we believe that using FRG huHep mice for crosses, it will be possible to design specific crosses to improve the power to detect associations for phenotypes even when a particular genetic variant only controls a small amount of variation. The ability to detect such loci will be greatly enhanced by isolating large numbers of unique recombinants for analysis.

To date, most of the phenotypes examined have focused on the asexual erythrocytic stages, as they can be easily measured *in vitro*. The FRG huHep mouse however, also enables investigation of liver stage phenotypes and has already been used to study liver stage growth [52], drug efficacies against liver stage parasites [53], as well as fitness costs across the whole lifecycle [46].

Genetic linkage mapping

Genetic linkage mapping (Figure 2) has been used for localizing determinants of phenotypic traits in organisms ranging from yeast to humans, as well as parasitic protozoa. As mentioned above, the four *P. falciparum* crosses generated using chimpanzees identified genes underlying numerous phenotypes of interest. Linkage analysis power is limited by the number of unique recombinant progeny but using by FRG huHep mice, the power of linkage mapping should be increased by the rapid generation of crosses using newly cloned clinical isolates (Figure 1), and the production of large numbers of unique recombinant progeny [44]. Indeed, 84 and 60 unique recombinant progeny were isolated from two separate crosses after only a limited amount of cloning, twice more than the previous chimpanzee crosses (15–35 per cross). With this many recombinants, it is possible to map phenotypes with small to modest **effect sizes** to smaller regions of the genome [44], which before had proven to be a big hurdle in the field. According to simulation data, for monogenic traits, with 30–40 recombinant progeny, loci with a large effect size (>0.5) can be confidently mapped while with 84 progeny, loci with a much smaller effect size (>0.2) can be mapped [44]. For phenotypes with complex genetic architectures, such as additive or epistatic interactions, 60 progeny are required to map loci with an effect size >0.4 [44]. Thus, the isolation of increased numbers of unique progeny greatly increases the power.

Bulk Segregant Analysis

Bulk Segregant Analysis (BSA, Box 1), also known as linkage group selection (LGS), is an additional strategy to genetic linkage mapping and utilizes selection-based approaches (e.g., with drugs) to enrich large progeny pools for different phenotypes. These pools are then compared using microarray or next generation deep sequencing approaches (Figure 3) and this allows for accurate measurement of allele frequencies. BSA has been used to study a number of rodent malaria phenotypes [54]. Specifically, the strategy has been used alongside crosses to map genes determining blood stage multiplication rate [55], virulence and immunity [56] in the rodent malaria *P. yoelii*, as well as mutations conferring ART resistance [57] and strain-specific immunity [58] in the rodent malaria *P. chabaudi*. In addition, BSA has demonstrated an association of the *P. vivax* chloroquine resistance transporter (*PvCRT*) with chloroquine resistance in this human malaria parasite [59]. BSA methods have also been used for mapping fitness throughout the *P. falciparum* lifecycle [46], using the MKK2835 (Thailand) × NHP1337 (Thailand) cross described above. After prolonged asexual blood stage culture, repeatable genome wide skews in blood stage parasites with strong against alleles from the NHP1337 parent were identified at loci on chromosomes 12 and 14. Thus, the BSA approach has the potential, in concert with classical genetic linkage mapping, to identify genes involved in multiple traits including drug resistance, compatibility with mosquito vectors, tolerance of temperature during blood stage episodic fever and selection by the human immune system.

BSA methods can also be used as a standalone method of analysis. The pros of standalone BSA are the avoidance of the labor and cost intensive process of isolating individual recombinant progeny, as well as the increase in statistical power obtained by using the complete pool of unique recombinants for analysis. The number of unique recombinants can be increased by using high numbers of highly infected mosquitoes and using groups of mice that receive unique populations of sporozoites for the liver stage-to-blood stage transition. For instance, an average of 10 recombinant oocysts/mosquito can generate up to 10,000 unique recombinant progeny from one cage of 250 mosquitoes (10 recombinant oocysts × 4 unique recombinants per oocyst × 250 mosquitoes). The ability to easily increase the numbers of mice used per cross is a distinct advantage of FRG huHep mice over chimpanzees as an infection model and allows for replication to assess reproducibility of results. By using independent pools of mosquitoes to infect individual mice, one can multiply the numbers of recombinants generated, while also establishing true biological replicates for each experiment. The cons are that BSA does not provide the haplotype information needed to directly examine epistatic interactions. BSA also relies on phenotypes that are amenable to selection in bulk. Fortunately, *P. falciparum* has a key advantage over rodent malaria systems because parasites can be grown *in vitro* and cloned by limiting dilution. Hence, BSA and genetic linkage mapping can be used together to directly examine haplotypes carrying different allele combinations. Furthermore, BSA can be a valuable first step to direct the more intensive cloning and phenotyping efforts for a deeper dissection of genetic architecture with greater success.

Fine mapping and functional validation

Quantitative trait loci identified using genetic linkage mapping are typically large and span multiple genes but multiple approaches can be used to shorten the list of candidate genes and include: i) the prioritization of genes with non-synonymous mutations, indels or copy number variants between the two parents; ii) reference to candidate gene association studies, including GWAS [60], population selection signatures [61] and **selective sweeps** [62], and iii) the narrowing down of genes involved based on spatiotemporal expression patterns [63] or gene essentiality [64]. According to a simulation analysis, with 86 unique recombinant progeny, one can map large effect mutations to as small as 31 kb intervals – just three genes [44].

Once a small number of candidate genes have been identified, functional validation is necessary to confirm the target. An advantage for *P. falciparum* is that CRISPR/Cas9 editing is available. Ghorbal et al. [65] were the first group to use CRISPR/Cas9 in *P. falciparum* and deleted *kahrp* as well as introducing mutations into *orc1* and *PfK13*. Nair et al. [51] also used CRISPR/Cas9 editing to introduce *PfK13* substitutions associated with ART resistance to compare the fitness costs of different *PfK13* alleles. CRISPR/Cas9 editing has also been used to confirm mutations related to drug resistance [66], as well as in gene expression regulations [67]. These examples demonstrate the power of CRISPR/Cas9 editing to test hypotheses associated with the results from fine mapping.

The future: ongoing prospects for crosses

FRG huHep mice are now being used for crosses and provide power and efficiency [41, 46]. As an attractive extension of genetic linkage mapping, BSA has been applied to directly study multiple selectable traits [46]. By constructing a series of crosses using parasites recently isolated from patients that bear different drug resistance and fitness traits, it should be possible to gain insight into the molecular basis of malaria parasite drug resistance, evolution, and epidemiology. Improvements in modern genomic technologies will also be used to complement genetic linkage mapping, particularly combining CRISPR/Cas9 technology with crosses.

Moving forward, *P. falciparum* has the potential to be used as a model organism for **systems genetics** approaches in understanding the interaction between genotypes and phenotypes. This can be done by constructing multiple layered molecular datasets (e.g. transcriptomics, proteomics and metabolomics) with the parental parasites and recombinant progeny. Systems genetics analyses of organisms such as yeast and humans are revealing fundamental aspects of the interplay and feedback between genes, mRNA, proteins, and metabolites. We anticipate that linkage analyses of transcript, protein and metabolite levels will provide a powerful tool in investigating fundamental questions concerning gene regulation in *P. falciparum* and parasite evolution.

Improving the production success of recombinant progeny

The ultimate aim of a cross is hundreds of independent unique recombinant progeny for downstream analyses. This requires that both parental strains generate infectious

gametocytes, leading to well-infected mosquitoes harboring tens of thousands of recombinant sporozoites. For crosses to become routine, several factors should be considered including (i) the staging of staggered gametocyte cultures to ensure a high proportion of mature male and female gametocytes, (ii) the setup of gametocyte feeds to mosquitoes on consecutive days, (iii) using different proportions of gametocytes from each cloned parent for the feed, (iv) injecting salivary gland sporozoites from different mosquito cages into multiple FRG huHep mice per cross to ensure the liver stage-to-blood stage transition of unique progeny, (v) using different hepatocyte donors for FRG mouse repopulation, (vi) setting up cloning plates immediately after the liver stage-to-blood stage transition, and (vii) cryopreserving stocks for downstream analyses as soon as the *in vitro* culture of asexual blood stages is robust. In addition, long term *in vitro* asexual culture of parental asexual blood stage parasite leads to poor downstream infectious gametocyte production which can greatly affect the success of a cross. In this situation, carrying the parent throughout the complete life cycle and using ‘new’ parent progeny for cross set up will likely improve the outcome of a cross. Ultimately, because every cross is unique, it is unlikely that one methodology will fit all applications.

Generating crosses entirely *in vitro*

Performing crosses without the use of mosquitoes or vertebrate animals is fundamentally possible and should be a goal. This will alleviate the use of animals and enable more researchers to access the power of crosses. The challenge of providing *in vitro* culture conditions to allow for the transition of mature gametocyte to salivary gland sporozoite is profound when one considers the *in vivo* process takes up to three weeks and relies upon multiple life cycle stage transitions. However, sporozoites have been generated without mosquitoes using *in vitro* culture and various media compositions have been defined for different *Plasmodium* species [68–71] but due to the complex nature of the life cycle, much improvement is needed to establish this methodology routinely. Additionally, the development of microscale human liver platforms that combine primary human hepatocytes with supportive stromal cells have been described [72, 73] and can support infection with *P. falciparum* and *P. vivax* sporozoites, liver stage maturation and transition to blood stage [74]. Higher infection rates using these platforms and consistent transition to asexual blood stage replication will eventually eliminate the need for human liver-chimeric mice.

Concluding remarks

Many questions remain regarding the genetic determinants of drug resistance, parasite adaptation, and parasite growth (see Outstanding Questions). Given that crosses can now be routinely carried out, crosses and their accompanying datasets will become an invaluable community resource to answer these questions. Metadata should be deposited in PlasmoDB to allow researchers to easily download datasets for their own analysis. The parents and progeny of the crosses should be shared with researchers to enable investigations of their own phenotypes of interest. As the cost of new technologies decreases and quality improves, existing phenotyping will become more precise and new phenotypes can be analyzed. By executing well thought-out crosses and sharing progeny as a community resource, these new progeny pools will, as the progeny have from the NHP crosses, be studied for many years to

come. The FRG huHep mouse model has led to the rebirth and enhancement of methods for genetic analyses in *P. falciparum*.

Acknowledgments

The authors' research on *P. falciparum* crosses using FRG huHep mice is supported by the National Institutes for Health P01 grant AI127338. We thank Ian H. Cheeseman, Timothy J.C. Anderson and Michael T. Ferdig for critical reading of the manuscript.

GLOSSARY

Bulk segregant analysis

A linkage mapping approach that compares allele frequencies from different population pools segregated by phenotype

Effect size

The contribution of a locus to the genetic variance of a given trait

FRG huHep mouse

An immunocompromised mouse engrafted with human hepatocytes that allows for *P. falciparum* liver stage development, and with transfusion of human erythrocytes, the transition from liver stage-to-blood stage

Genetic architecture

The genetic contribution to a phenotype. In the simplest case, a single major gene influences a phenotype. However, more typically, many genes influence a phenotype to varying degrees (each gene has a different effect size on the phenotype) and there are various interactions between these genes and the environment (genotype-by-genotype interactions and/or genotype-by-environment interactions)

Genetic background

The broader whole-genome context in which the phenotype of interest evolves and persists. The broader genetic context can include secondary (or modulatory) genes along with compensatory genes that play an important role in defining a phenotype

Genetic linkage mapping

This approach uses a genome-wide set of genetic markers across a set of recombinant progeny and parents to determine if there are genetic loci associated with the phenotype of interest. Loci that are associated with the phenotype emerge as significant peaks in quantitative trait locus mapping

Genome-wide association studies (GWAS)

This approach analyzes genome-wide variation in a set of unrelated individuals to search for genotypic variants that are associated with a specific phenotype

Rarefaction

A method in numerical ecology to estimate the species diversity in a community

Selective sweep

The process by which a newly beneficial mutation arises and becomes fixed in the population. This includes hard selective sweeps (mutations in the chloroquine resistance transporter, *PfCRT*, conferring chloroquine resistance) and soft selective sweeps (multiple point mutations in *PfK13* conferring artemisinin resistance)

Selective whole genome amplification (sWGA)

A method using short oligonucleotide probes that preferentially bind to a target genome to amplify and sequence genomes of interest from contaminating host tissues

Systems genetics

An approach to make sense of complex traits. It examines intermediate molecular phenotypes, such as transcript levels, protein and metabolite abundance, to attempt to bridge the gap DNA variation and the traits of interest. Of note, systems genetics methods can be integrated with GWAS to predict causal genes and their functions

References

1. Alqudah AM et al. (2020) GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley - A review. *J Adv Res* 22, 119–135. [PubMed: 31956447]
2. Su X et al. (1999) A genetic map and recombination parameters of the human malaria parasite *Plasmodium falciparum*. *Science* 286 (5443), 1351–3. [PubMed: 10558988]
3. Gardner MJ et al. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419 (6906), 498–511. [PubMed: 12368864]
4. Kissinger JC et al. (2002) The *Plasmodium* genome database. *Nature* 419 (6906), 490–2. [PubMed: 12368860]
5. Miles A et al. (2016) Indels, structural variation, and recombination drive genomic diversity in *Plasmodium falciparum*. *Genome Res* 26 (9), 1288–99. [PubMed: 27531718]
6. Stapley J et al. (2017) Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos Trans R Soc Lond B Biol Sci* 372 (1736).
7. Walliker D et al. (1971) Genetic recombination in malaria parasites. *Nature* 232 (5312), 561–2. [PubMed: 4937500]
8. Walliker D (1983) The genetic basis of diversity in malaria parasites. *Adv Parasitol* 22, 217–59. [PubMed: 6364735]
9. Trager W and Jensen JB (1976) Human malaria parasites in continuous culture. *Science* 193 (4254), 673–5. [PubMed: 781840]
10. Rosario V (1981) Cloning of naturally occurring mixed infections of malaria parasites. *Science* 212 (4498), 1037–8. [PubMed: 7015505]
11. Ifediba T and Vanderberg JP (1981) Complete in vitro maturation of *Plasmodium falciparum* gametocytes. *Nature* 294 (5839), 364–6. [PubMed: 7031476]
12. Campbell CC et al. (1982) *Plasmodium falciparum* gametocytes from culture in vitro develop to sporozoites that are infectious to primates. *Science* 217 (4564), 1048–50. [PubMed: 7051285]
13. Walliker D et al. (1987) Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* 236 (4809), 1661–6. [PubMed: 3299700]
14. Peterson DS et al. (1988) Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in falciparum malaria. *Proc Natl Acad Sci U S A* 85 (23), 9114–8. [PubMed: 2904149]
15. Wellems TE et al. (1990) Chloroquine resistance not linked to *mdr*-like genes in a *Plasmodium falciparum* cross. *Nature* 345 (6272), 253–255. [PubMed: 1970614]
16. Ferdig MT and Su XZ (2000) Microsatellite markers and genetic mapping in *Plasmodium falciparum*. *Parasitol Today* 16 (7), 307–12. [PubMed: 10858651]

17. Fidock DA et al. (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell* 6 (4), 861–71. [PubMed: 11090624]
18. Wang P et al. (1997) Sulfadoxine resistance in the human malaria parasite *Plasmodium falciparum* is determined by mutations in dihydropteroate synthetase and an additional factor associated with folate utilization. *Mol Microbiol* 23 (5), 979–86. [PubMed: 9076734]
19. Freitas-Junior LH et al. (2000) Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407 (6807), 1018–22. [PubMed: 11069183]
20. Ferdig MT et al. (2004) Dissecting the loci of low-level quinine resistance in malaria parasites. *Mol Microbiol* 52 (4), 985–97. [PubMed: 15130119]
21. Furuya T et al. (2005) Disruption of a *Plasmodium falciparum* gene linked to male sexual development causes early arrest in gametocytogenesis. *Proc Natl Acad Sci U S A* 102 (46), 16813–8. [PubMed: 16275909]
22. Nguitragool W et al. (2011) Malaria parasite *clag3* genes determine channel-mediated nutrient uptake by infected red blood cells. *Cell* 145 (5), 665–77. [PubMed: 21620134]
23. Hayton K et al. (2008) Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of *Plasmodium falciparum* invasion. *Cell Host Microbe* 4 (1), 40–51. [PubMed: 18621009]
24. Crosnier C et al. (2011) Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. *Nature* 480 (7378), 534–7. [PubMed: 22080952]
25. Molina-Cruz A et al. (2013) The human malaria parasite Pf47 gene mediates evasion of the mosquito immune system. *Science* 340 (6135), 984–7. [PubMed: 23661646]
26. Sa JM et al. (2009) Geographic patterns of *Plasmodium falciparum* drug resistance distinguished by differential responses to amodiaquine and chloroquine. *Proc Natl Acad Sci U S A* 106 (45), 18883–9. [PubMed: 19884511]
27. Gupta A et al. (2020) Complex nutrient channel phenotypes despite Mendelian inheritance in a *Plasmodium falciparum* genetic cross. *PLoS Pathog* 16 (2), e1008363. [PubMed: 32069335]
28. Cheeseman IH et al. (2012) A major genome region underlying artemisinin resistance in malaria. *Science* 336 (6077), 79–82. [PubMed: 22491853]
29. Arie F et al. (2014) A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* 505 (7481), 50–5. [PubMed: 24352242]
30. Sa JM et al. (2018) Artemisinin resistance phenotypes and K13 inheritance in a *Plasmodium falciparum* cross and Aotus model. *Proc Natl Acad Sci U S A* 115 (49), 12513–12518. [PubMed: 30455312]
31. Wellems TE et al. (1991) Genetic mapping of the chloroquine-resistance locus on *Plasmodium falciparum* chromosome 7. *Proc Natl Acad Sci U S A* 88 (8), 3382–6. [PubMed: 1673031]
32. Su X et al. (1997) Complex polymorphisms in an approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell* 91 (5), 593–603. [PubMed: 9393853]
33. Yuan J et al. (2009) Genetic mapping of targets mediating differential chemical phenotypes in *Plasmodium falciparum*. *Nat Chem Biol* 5 (10), 765–71. [PubMed: 19734910]
34. Patel JJ et al. (2010) Chloroquine susceptibility and reversibility in a *Plasmodium falciparum* genetic cross. *Mol Microbiol* 78 (3), 770–87. [PubMed: 20807203]
35. Reilly HB et al. (2007) Quantitative dissection of clone-specific growth rates in cultured malaria parasites. *Int J Parasitol* 37 (14), 1599–607. [PubMed: 17585919]
36. Reilly Ayala HB et al. (2010) Quantitative trait loci mapping reveals candidate pathways regulating cell cycle duration in *Plasmodium falciparum*. *BMC Genomics* 11, 577. [PubMed: 20955606]
37. Vaidya AB et al. (1995) A genetic locus on *Plasmodium falciparum* chromosome 12 linked to a defect in mosquito-infectivity and male gametogenesis. *Mol Biochem Parasitol* 69 (1), 65–71. [PubMed: 7723789]
38. Lewis IA et al. (2014) Metabolic QTL analysis links chloroquine resistance in *Plasmodium falciparum* to impaired hemoglobin catabolism. *PLoS Genet* 10 (1), e1004085. [PubMed: 24391526]

39. Gonzales JM et al. (2008) Regulatory hotspots in the malaria parasite genome dictate transcriptional variation. *PLoS Biol* 6 (9), e238. [PubMed: 18828674]
40. Rodhain J and Jadin J (1964) [Transmission of *Plasmodium falciparum* in the Splenectomized Chimpanzee]. *Ann Soc Belges Med Trop Parasitol Mycol* 44, 531–5. [PubMed: 14259259]
41. Vaughan AM et al. (2015) *Plasmodium falciparum* genetic crosses in a humanized mouse model. *Nat Methods* 12 (7), 631–3. [PubMed: 26030447]
42. Azuma H et al. (2007) Robust expansion of human hepatocytes in *Fah*^{-/-}/*Rag2*^{-/-}/*Il2rg*^{-/-} mice. *Nat Biotechnol* 25 (8), 903–10. [PubMed: 17664939]
43. Bissig KD et al. (2007) Repopulation of adult and neonatal mice with human hepatocytes: a chimeric animal model. *Proc Natl Acad Sci U S A* 104 (51), 20507–11. [PubMed: 18077355]
44. Button-Simons KA et al. (2019) Surprising variation in the outcome of two malaria genetic crosses using humanized mice: implications for genetic mapping and malaria biology. *bioRxiv*, 2019.12.13.871830.
45. Colwell RK et al. (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. *Journal of Plant Ecology* 5 (1), 3–21.
46. Li X et al. (2019) Genetic mapping of fitness determinants across the malaria parasite *Plasmodium falciparum* life cycle. *PLoS Genet* 15 (10), e1008453. [PubMed: 31609965]
47. Mok S et al. (2015) Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science* 347 (6220), 431–5. [PubMed: 25502316]
48. Yang T et al. (2019) Decreased K13 Abundance Reduces Hemoglobin Catabolism and Proteotoxic Stress, Underpinning Artemisinin Resistance. *Cell Rep* 29 (9), 2917–2928 e5. [PubMed: 31775055]
49. Siddiqui G et al. (2017) Multi-omics Based Identification of Specific Biochemical Changes Associated With *PfKelch13*-Mutant Artemisinin-Resistant *Plasmodium falciparum*. *J Infect Dis* 215 (9), 1435–1444. [PubMed: 28368494]
50. Tirrell AR et al. (2019) Pairwise growth competitions identify relative fitness relationships among artemisinin resistant *Plasmodium falciparum* field isolates. *Malar. J.* 18 (1), 295. [PubMed: 31462253]
51. Nair S et al. (2018) Fitness Costs and the Rapid Spread of *kelch13*-C580Y Substitutions Conferring Artemisinin Resistance. *Antimicrob Agents Chemother* 62 (9), e00605–18. [PubMed: 29914963]
52. Foquet L et al. (2015) Assessment of parasite liver-stage burden in human-liver chimeric mice *In Malaria Vaccines*, pp. 59–68, Springer.
53. Flannery EL et al. (2018) Assessing drug efficacy against *Plasmodium falciparum* liver stages in vivo. *JCI Insight* 3 (1).
54. Culleton R et al. (2005) Linkage group selection: rapid gene discovery in malaria parasites. *Genome Res* 15 (1), 92–7. [PubMed: 15632093]
55. Pattaradilokrat S et al. (2009) Gene encoding erythrocyte binding ligand linked to blood stage multiplication rate phenotype in *Plasmodium yoelii yoelii*. *Proc Natl Acad Sci U S A* 106 (17), 7161–6. [PubMed: 19359470]
56. Abkhallo HM et al. (2017) Rapid identification of genes controlling virulence and immunity in malaria parasites. *PLoS Pathog* 13 (7), e1006447. [PubMed: 28704525]
57. Hunt P 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 (1), 499. [PubMed: 20846421]
58. Martinelli A et al. (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 (3), 814–9. [PubMed: 15640359]
59. Sa JM et al. (2019) *Plasmodium vivax* chloroquine resistance links to *pvcr1* transcription in a genetic cross. *Nat Commun* 10 (1), 4300. [PubMed: 31541097]
60. Miotto O et al. (2015) Genetic architecture of artemisinin-resistant *Plasmodium falciparum*. *Nat Genet* 47 (3), 226–34. [PubMed: 25599401]

61. Mehlotra RK et al. (2001) Evolution of a unique *Plasmodium falciparum* chloroquine-resistance phenotype in association with pfcrt polymorphism in Papua New Guinea and South America. *Proc Natl Acad Sci U S A* 98 (22), 12689–94. [PubMed: 11675500]
62. Anderson TJ et al. (2017) Population Parameters Underlying an Ongoing Soft Sweep in Southeast Asian Malaria Parasites. *Mol Biol Evol* 34 (1), 131–144. [PubMed: 28025270]
63. Howick VM et al. (2019) The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* 365 (6455), eaaw2619. [PubMed: 31439762]
64. Zhang M et al. (2018) Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* 360 (6388), eaap7847. [PubMed: 29724925]
65. Ghorbal M et al. (2014) Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nature biotechnology* 32 (8), 819.
66. Ng CL et al. (2016) CRISPR-Cas9-modified pfmdr1 protects *Plasmodium falciparum* asexual blood stages and gametocytes against a class of piperazine-containing compounds but potentiates artemisinin-based combination therapy partner drugs. *Mol Microbiol* 101 (3), 381–93. [PubMed: 27073104]
67. Xiao B et al. (2019) Epigenetic editing by CRISPR/dCas9 in *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 116 (1), 255–260. [PubMed: 30584102]
68. Al-Olayan EM et al. (2002) Complete development of mosquito phases of the malaria parasite in vitro. *Science* 295 (5555), 677–9. [PubMed: 11809973]
69. Porter-Kelley JM et al. (2006) *Plasmodium yoelii*: axenic development of the parasite mosquito stages. *Exp Parasitol* 112 (2), 99–108. [PubMed: 16289466]
70. Warburg A and Miller LH (1992) Sporogonic development of a malaria parasite in vitro. *Science* 255 (5043), 448–50. [PubMed: 1734521]
71. Warburg A and Schneider I (1993) In vitro culture of the mosquito stages of *Plasmodium falciparum*. *Exp Parasitol* 76 (2), 121–6. [PubMed: 8454020]
72. Khetani SR and Bhatia SN (2008) Microscale culture of human liver cells for drug development. *Nat Biotechnol* 26 (1), 120–6. [PubMed: 18026090]
73. Roth A et al. (2018) A comprehensive model for assessment of liver stage therapies targeting *Plasmodium vivax* and *Plasmodium falciparum*. *Nat Commun* 9 (1), 1837. [PubMed: 29743474]
74. Gural N et al. (2018) Engineered Livers for Infectious Diseases. *Cell Mol Gastroenterol Hepatol* 5 (2), 131–144. [PubMed: 29322086]
75. Ehrenreich IM et al. (2010) Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature* 464 (7291), 1039–42. [PubMed: 20393561]
76. Burga A et al. (2019) Fast genetic mapping of complex traits in *C. elegans* using millions of individuals in bulk. *Nat Commun* 10 (1), 2680. [PubMed: 31213597]
77. Blake DP et al. (2011) Genetic mapping identifies novel highly protective antigens for an apicomplexan parasite. *PLoS Pathog* 7 (2), e1001279. [PubMed: 21347348]
78. Chevalier FD et al. (2014) Efficient linkage mapping using exome capture and extreme QTL in schistosome parasites. *BMC genomics* 15 (1), 617. [PubMed: 25048426]
79. Leichthy AR and Brisson D (2014) Selective whole genome amplification for resequencing target microbial species from complex natural samples. *Genetics* 198 (2), 473–81. [PubMed: 25096321]
80. Sundararaman SA et al. (2016) Genomes of cryptic chimpanzee *Plasmodium* species reveal key evolutionary events leading to human malaria. *Nat Commun* 7, 11078. [PubMed: 27002652]
81. Oyola SO et al. (2016) Whole genome sequencing of *Plasmodium falciparum* from dried blood spots using selective whole genome amplification. *Malaria journal* 15 (1), 597. [PubMed: 27998271]
82. Ranford-Cartwright LC and Mwangi JM (2012) Analysis of malaria parasite phenotypes using experimental genetic crosses of *Plasmodium falciparum*. *Int J Parasitol* 42 (6), 529–34. [PubMed: 22475816]
83. Jiang H et al. (2011) High recombination rates and hotspots in a *Plasmodium falciparum* genetic cross. *Genome Biol* 12 (4), R33. [PubMed: 21463505]

84. Davis SZ 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 (1), 54. [PubMed: 32005233]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

OUTSTANDING QUESTIONS

- ART resistance is, in part, mediated by mutations in *PfK13*, but ART resistant parasites carrying wildtype *PfK13* have been isolated. What are the genetic determinants of ART resistance in wildtype *PfK13* parasites and can crosses help answer this question?
- Linkage mapping power has been strengthened using crosses generated by FRG huHep mice since more unique progeny have been isolated. Is it now possible to perform high-throughput chemical screening with these progenies to unearth loci involved in the maintenance of drug resistance associated mutations?
- Malaria parasite transmission requires both female mosquitoes and human hosts. Can crosses help us understand the loci that underlie adaptation to different mosquito vectors, the human liver and the human immune system?
- Different culture conditions play a critical role in many *in vitro* assays, especially drug response assays. What genes determine parasite growth under these different culture conditions?
- Genome inheritance can be skewed, and outcrossing is preferred in multiple malaria genetic crosses. Is there cytoplasmic incompatibility and gamete recognition in *P. falciparum* genetic crosses? If so, what are the underlying genetic mechanisms?
- Can we use systems genetics within the framework of a genetic cross, to map genomic loci that determine levels of each transcript, protein and metabolite, and to understand more fully how phenotype is linked to genotype?

HIGHLIGHTS

- Historically, chimpanzees were used for experimental genetic crosses between cloned *Plasmodium falciparum* parasites and have played a key role in determining genetic loci responsible for drug resistance, virulence, invasion, growth rate, and transmission.
- The human hepatocyte chimeric FRG huHep mouse has replaced the chimpanzee and allows for the routine generation of *P. falciparum* genetic crosses and the recovery of hundreds and potentially thousands of unique recombinant progeny, strengthening the power of genetic linkage mapping.
- Classical linkage mapping, bulk segregant analysis, CRISPR/Cas9 technology, and high throughput 'omics techniques using recombinant progeny should allow for unprecedented power and efficiency to carry out a systems genetics approach to expand our understanding of *P. falciparum* biology.

Box 1.**Bulk segregant analysis (BSA)**

BSA or linkage group selection (LGS) is a method initially developed to study disease processes in humans and plants. Individuals with different phenotypes are pooled and genotyped to identify loci that segregate in different pools. In order to not separate and phenotype individual isolates, selection pressures such as drug treatment, immune sweeps, or growth competition can also be applied directly to the population to generate segregating pools. The genome-wide allele frequencies in each pool are then measured by microarray-based genotyping or next-generation sequencing.

Shifts in allele frequencies in populations of potentially thousands of progeny generated by crosses provide locus-specific readouts of competitive fitness. This approach has been successfully applied to the studies of yeast [75], *Caenorhabditis elegans* [76], *Eimeria tenella* [77], *Schistosoma mansoni* [78], the rodent malaria parasite *P. yoelii* [54, 56] as well as *P. falciparum* [46].

For malaria parasites, which have complicated life cycles and are difficult to separate from their hosts, it is extremely challenging to accurately measure genome-wide allele frequencies since parasite DNA can account for a very small proportion of the total DNA that can be extracted from mosquito and mammalian samples. To overcome this challenge, **selective whole genome amplification (sWGA)**, originally described by Leitchy and Brisson [79], has been successfully applied to *P. falciparum* [80, 81].

Figure 1. Laboratory maintenance of the complete *Plasmodium falciparum* life cycle using FRG huHep mice.

(A) The FRG huHep mouse lacks the fumaryl acetoacetate hydrolase (FAH) gene (F designation), causing hepatocyte cell death, which is controlled by the drug nitisinone. The mouse also lacks Rag-2 (R designation) and the interleukin 2 receptor common gamma chain (IL-2R γ , G designation), resulting in adaptive immune deficiency and thus, an inability to reject foreign tissue. Injection with human hepatocytes and cycling on and off nitisinone, allows the human hepatocytes to replace the mouse hepatocytes. Human serum albumin levels are measured to check liver humanization. The FRG huHep mouse is injected with human erythrocytes on days 6 and 7 post injection with sporozoites to allow for the liver stage-to-blood stage transition. (B) Experimental genetic crosses use sporozoites from infected female *Anopheles stephensi* mosquitoes and FRG huHep mice as described by Vaughan et al. [41]. To initiate a new cross, gametocytes are enriched from independent cultures of the two parental parasite clones (yellow and blue). Day 0 is defined as the day of the gametocyte feed to the mosquitoes. Male and female gametes fuse in the blood meal within the mosquito midgut and form a zygote. Recombination of parasite genomes (striped yellow and blue) occurs during meiosis in the presence of genetically distinct gametes although fusion between male and female gametes of the same genotype (selfing) is also possible. The zygote transforms into an oocyst, within which hundreds of sporozoites mature that ultimately invade the mosquito salivary gland and in a natural setting, lie in wait for transmission to the next vertebrate host. Sporozoites are isolated from infected mosquito salivary glands and injected intravenously into FRG huHep mice (Day 14). The sporozoites migrate to the liver, invade human hepatocytes and develop into exo-erythrocytic merozoites. To allow for the liver stage-to-blood stage transition, mice are infused with human erythrocytes on Days 20 and 21 (six and seven days after sporozoite injection). Four hours after the second infusion (Day 21), the mice are euthanized and exsanguinated to isolate the circulating ring stage *P. falciparum*-infected human erythrocytes. The parasites now initiate haploid asexual multiplication during *in vitro* culture by repeatedly invading erythrocytes every 48 hours. It takes approximately five weeks in total to generate recombinant pools: two weeks for gametocyte maturation, two weeks for sporozoite maturation and one week for the liver stage-to-blood stage transition. Cloning and expansion of individual progeny allows for genotyping and downstream phenotyping and mapping.

Figure 2. The principles of gene mapping.

Two approaches are often used to map determinant genes: (i) genetic linkage mapping and (ii) Bulk Segregant Analysis (BSA). For genetic linkage mapping, parasites are cloned using limiting dilution. Cultures are maintained for as long as eight weeks for single cloned parasites to amplify. All cloned cultures are whole genome sequenced to exclude non-clonal, selfed and repeat parasites from further phenotyping. Thereafter, multiple layers of quantitative and heritable traits can then be measured in the unique recombinant progeny and used for genetic linkage mapping. Methods used to measure a heritable traits include: IC₅₀'s, the ring-stage survival assay (RSA), the extended recovery ring-stage survival assay

(eRRSA [84]), parasite fitness and multiple ‘omics approaches. For BSA, recombinant pools are segregated through selection pressure such as drugs, followed by pooled sequencing to determine genome-wide allele frequencies and identify genomic regions under selection. Candidate genes can be further verified using CRISPR/Cas9-based gene editing.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Bulk segregant analysis (BSA).

The BSA approach uses pooled sequencing of progeny populations to measure changes in allele frequency following specific interventions (such as antimalarial drug selection of *in vitro* blood stage culture). The frequency of resistant-associated alleles is expected to increase after drug treatment, while the frequency of sensitive alleles decreases. Without any intervention, parasite competition also exists during [46], and thus side-by-side control experiments are necessary with and without interventions. In addition, there are likely both recombinant (yellow and blue) and parental parasites (yellow or blue) in the segregant pools and these must be considered. The green arrows indicate allele frequency changes caused by parasite competition during *in vitro* culture and the magenta arrow indicates a locus/gene conferring drug resistance.

Table 1.Status of current *Plasmodium falciparum* experimental crosses

Cross Date [reference]	Parental Parasite Origin	Progeny Cloned (% recombinants)	# Unique Recombinants	Genetic Map Length (cM)	Recombination Rate	Genotyping method
1985 (chimpanzee) [5]	3D7 (West Africa, clone of NF54) × HB3 (Honduras)	113 (89%)	15	nd	11.0 kb/cM [82]	RFLP, whole genome sequencing
1990 (chimpanzee) [2]	HB3 (Honduras) × Dd2 (Laos)	76 (21%)	16	1556	12.1 kb/cM [2]	RFLP, microsatellite, whole genome sequencing
2008 (chimpanzee) [83]	7G8 (Brazil) × GB4 (Ghana)	>200 (86%)	32	1655	12.8 kb/cM [83]	Microsatellite, oligonucleotide array, whole genome sequencing
2018 (chimpanzee) [30]	GB4 (Ghana) × 803 (Cambodia)	>400 (?)	27	nd	13.3 kb/cM [30]	Oligonucleotide array
2013 (FRG huHep mouse) [41]	NF54HT-GFP-luc (West Africa) × GB4 (Ghana)	23 (83%)	5	nd	nd	Microsatellite
2013 (FRG huHep mouse) [41]	NF54HT-GFP-luc (West Africa) × 7G8 (Brazil)	44 (100%)	15	nd	nd	Microsatellite
2013 & 2019 (FRG huHep mouse) [41], [44]	NF54(HT-GFP-luc) × NHP4026 (Thailand)	166 (77%)	84	1521	13.7 kb/cM [44]	Microsatellite, whole genome sequencing
2019 (FRG huHep mouse) [44]	MKK2835 (Thailand) × NHP1337 (Thailand)	266 (23%)	60	1626	12.8 kb/cM [44]	Whole genome sequencing