

## RESEARCH ARTICLE SUMMARY

## MALARIA

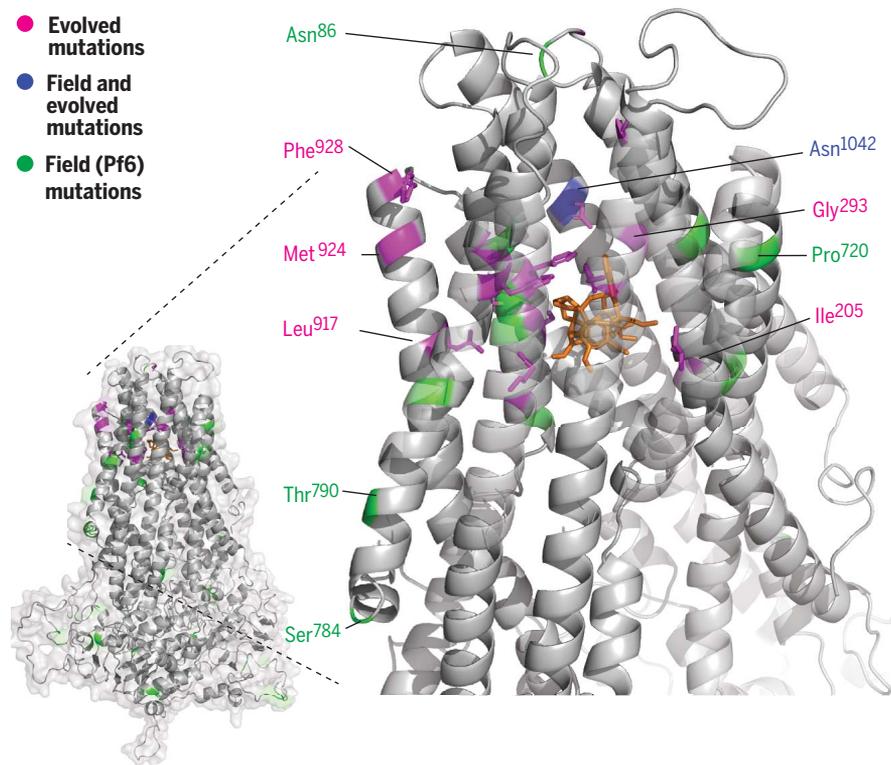
Systematic *in vitro* evolution in *Plasmodium falciparum* reveals key determinants of drug resistance

Madeline R. Luth<sup>†</sup>, Karla P. Godinez-Macias<sup>†</sup>, Daisy Chen<sup>†</sup>, John Okombo, Vandana Thathy, Xiu Cheng, Sindhu Daggupati, Heledd Davies, Satish K. Dhingra, Jan M. Economy, Rebecca C. S. Edgar, Maria G. Gomez-Lorenzo, Eva S. Istvan, Juan Carlos Jado, Gregory M. LaMonte, Bruno Melillo, Sachel Mok, Sunil K. Narwal, Tolla Ndiaye, Sabine Ottilie, Sara Palomo Diaz, Heekuk Park, Stella Peña, Frances Rocamora, Tomoyo Sakata-Kato, Jennifer L. Small-Saunders, Robert L. Summers, Patrick K. Tumwebaze, Manu Vanaerschot, Guoqin Xia, Tomas Yeo, Ashley You, Francisco-Javier Gamo, Daniel E. Goldberg, Marcus C. S. Lee, Case W. McNamara, Daouda Ndiaye, Philip J. Rosenthal, Stuart L. Schreiber, Gloria Serra, Jair Lage De Siqueira-Neto, Tina S. Skinner-Adams, Anne-Catrin Uhlemann, Nobutaka Kato, Amanda K. Lukens, Dyann F. Wirth, David A. Fidock, Elizabeth A. Winzeler\*

**INTRODUCTION:** Malaria parasites frequently evolve resistance to antimalarials in the laboratory and in the field. Because large-scale phenotyping of parasites for drug resistance is impractical, evaluation of the prevalence of molecular markers for drug resistance provides an early warning to inform region-specific malaria treatments. Thus, a major goal of sequencing clinical isolates is to identify the emergence of drug resistance markers and to scan for alleles

and regions of the genome under selection. A key challenge in these efforts is distinguishing functional variants that drive the observed phenotype from passenger mutations, which do not confer phenotype changes.

**RATIONALE:** *In vitro* evolution and whole-genome analysis, an early discovery method for identifying resistance mechanisms and drug targets, has yielded a rich dataset of mutations found in



#### Resistance-conferring missense mutations occur in conserved, well-ordered protein domains.

*P. falciparum* multidrug resistance protein 1 (PfMDR1) homology model bound to vincristine (colored orange), highlighting missense variants found in parasite samples worldwide (“field,” colored green), variants associated with resistance phenotypes from *in vitro* compound selection experiments (“evolved,” colored pink), and variants present in both datasets (colored blue).

*Plasmodium falciparum* parasites resistant to diverse antimalarial compounds. These samples reflect short-term selection, permitting the use of statistical methods to pinpoint mutations underlying resistance phenotypes. Insights into genetic determinants of antimalarial resistance from this dataset may enable *in silico* methods for identifying resistance-conferring mutations, which are needed to improve genomic surveillance of clinical drug resistance and accelerate target-based drug discovery of novel antimalarials.

**RESULTS:** Through comprehensive analysis of the whole-genome sequences of 724 *P. falciparum* clones evolved to resist one of 118 small-molecule growth inhibitors, we identify previously unknown resistance alleles and genes, highlight drivers of multidrug resistance, and show that *in vitro* evolved variants are more likely to (i) be missense or frameshift, (ii) involve bulky amino acid changes, and (iii) occur in conserved, ordered protein domains. Our data illustrate an evolutionary landscape in which each compound typically selects for driver mutations in only one or a few genes related to the compound’s mechanism of action, but multiple different mutations in a gene, ranging from substitutions near protein binding pockets to copy number amplification, can confer resistance. Copy number variants, in particular, frequently drove resistance by amplifying targets, such as tRNA synthetases, or drug efflux transporters such as PfABC13 and PfMDR1. Through network analysis, we also found that AP2 transcription factors were often mutated alongside known resistance drivers across selections with different compounds, suggesting roles in culture adaptation or multidrug resistance. By comparing compound susceptibility of parasites with *in vitro* evolved versus naturally occurring missense variants in the multidrug resistance genes *pfmdr1* and *pfcarl*, as well as a known target, PfATP4, we validated the roles of these *in vitro* evolved variants in resistance to the compound(s) with which they were selected and observed that protein structural localization is a key differentiator between driver and passenger mutations.

**CONCLUSION:** Our dataset provides a starting collection for algorithms that can identify genomic changes in clinical isolates that are likely associated with drug resistance in different species. It also presents insights for distinguishing functional from nonfunctional variants in forward genetic approaches. ■

The list of author affiliations is available in the full article online.

\*Corresponding author. Email: ewinzeler@health.ucsd.edu

<sup>†</sup>These authors contributed equally to this work.

Cite this article as M. R. Luth *et al.*, *Science* **386**, eadk9893 (2024). DOI: 10.1126/science.adk9893

**S** READ THE FULL ARTICLE AT  
<https://doi.org/10.1126/science.adk9893>

## RESEARCH ARTICLE

## MALARIA

# Systematic in vitro evolution in *Plasmodium falciparum* reveals key determinants of drug resistance

Madeline R. Luth<sup>1</sup>†, Karla P. Godinez-Macias<sup>1</sup>†, Daisy Chen<sup>1</sup>†, John Okombo<sup>2,3</sup>, Vandana Thathy<sup>2,3</sup>, Xiu Cheng<sup>4</sup>, Sindhu Daggupati<sup>1</sup>, Heledd Davies<sup>5</sup>, Satish K. Dhingra<sup>2,3</sup>, Jan M. Economy<sup>1</sup>, Rebecca C. S. Edgar<sup>6</sup>, Maria G. Gomez-Lorenzo<sup>7</sup>, Eva S. Istvan<sup>8,9</sup>, Juan Carlos Jado<sup>1</sup>, Gregory M. LaMonte<sup>1</sup>, Bruno Melillo<sup>10</sup>†, Sachel Mok<sup>3,11</sup>, Sunil K. Narwai<sup>2,3</sup>, Tolla Ndiaye<sup>2,3</sup>, Sabine Otilie<sup>1</sup>§, Sara Palomo Diaz<sup>7</sup>, Heekuk Park<sup>11</sup>, Stella Peña<sup>12</sup>, Frances Rocamora<sup>1</sup>, Tomoyo Sakata-Kato<sup>4,13</sup>, Jennifer L. Small-Saunders<sup>3,11</sup>, Robert L. Summers<sup>14,15</sup>, Patrick K. Tumwabeze<sup>16</sup>, Manu Vanaerschot<sup>2</sup>, Guoqin Xia<sup>17</sup>, Tomas Yeo<sup>2,3</sup>, Ashley You<sup>1</sup>, Francisco-Javier Gamo<sup>7</sup>, Daniel E. Goldberg<sup>8,9</sup>, Marcus C. S. Lee<sup>5,6</sup>, Case W. McNamara<sup>18</sup>, Daouda Ndiaye<sup>19</sup>, Philip J. Rosenthal<sup>20</sup>, Stuart L. Schreiber<sup>21</sup>, Gloria Serra<sup>12</sup>, Jair Lage De Siqueira-Neto<sup>22</sup>, Tina S. Skinner-Adams<sup>23</sup>, Anne-Catrin Uhlemann<sup>11</sup>, Nobutaka Kato<sup>4,13</sup>, Amanda K. Lukens<sup>14,15</sup>, Dyann F. Wirth<sup>14,15</sup>, David A. Fidock<sup>2,3,11</sup>, Elizabeth A. Winzeler<sup>1,22\*</sup>

Surveillance of drug resistance and the discovery of novel targets—key objectives in the fight against malaria—rely on identifying resistance-conferring mutations in *Plasmodium* parasites. Current approaches, while successful, require laborious experimentation or large sample sizes. To elucidate shared determinants of antimalarial resistance that can empower in silico inference, we examined the genomes of 724 *Plasmodium falciparum* clones, each selected in vitro for resistance to one of 118 compounds. We identified 1448 variants in 128 recurrently mutated genes, including drivers of antimalarial multidrug resistance. In contrast to naturally occurring variants, those selected in vitro are more likely to be missense or frameshift, involve bulky substitutions, and occur in conserved, ordered protein domains. Collectively, our dataset reveals mutation features that predict drug resistance in eukaryotic pathogens.

As drug resistance remains a major concern in controlling malaria, an infectious disease caused by protozoan *Plasmodium* parasites, detection of resistance is of vital importance. Because large-scale phenotyping of parasites for drug resistance is impractical and results of therapeutic efficacy studies are limited, a major goal of sequencing clinical isolates is to identify the emergence of genetic markers of drug resistance, thus providing an early warning to inform region-specific malaria treatment policies. As a result, surveillance efforts have now placed whole-genome sequences of >20,000 isolates of *Plasmodium falciparum*, the most common and virulent malaria parasite, into the public domain (1). A key challenge in these efforts is distinguishing functional variants, which drive the observed phenotype, from

passenger mutations, which do not confer phenotypic change. Although previous studies have identified some clinical resistance markers, the identification of novel resistance alleles either in laboratory-adapted or field isolates remains inefficient.

Since the earlier use of genetic crosses to determine key genes mediating parasite resistance to pyrimethamine, chloroquine, and other antimalarials, two alternative approaches for understanding drug resistance have emerged: population-based association studies of sequenced isolates, and in vitro evolution of resistant parasites (2). Leveraging whole-genome genotyping, these approaches have enabled the identification of mutations underlying resistance to many clinical drugs and new antimalarial compounds under development. For example, genome-

mapping studies first associated a >239-kb region surrounding the *kelch13* locus with partial resistance to artemisinins, the fast-killing components of current first-line treatments for *P. falciparum* malaria (3). Subsequent laboratory-based selection studies identified single-nucleotide variants (SNVs) in *kelch13* as alleles mediating resistance (4). However, both population-based association studies and in vitro resistance selection are laborious and time-consuming, limiting their utility for genomic surveillance. For example, evolving drug-resistant parasites may take several weeks or months, and the acquisition and phenotyping of the large quantity of clinical isolates needed for association studies poses considerable logistical challenges (1, 5). In addition, with genetic crosses (6) or genome-wide association studies (7), loci corresponding to dozens of potential phenotype-driving variants may be identified. An advantage of in vitro evolution studies is that because relatively few mutations emerge during drug selection, variants with a high probability of conferring resistance can be discovered through careful genome analysis of only a handful of independently derived drug-resistant clones. As a result, hundreds of laboratory-evolved organism genome sequences have been published. While studies of malaria parasites are among the most advanced for this type of work, the approach is also becoming widespread for other important microbes, including *Candida albicans* (8), *Mycobacterium tuberculosis* (9), and the model organism *Saccharomyces cerevisiae* (10).

To explore the potential of in silico approaches for predicting antimalarial resistance mediators by leveraging insights from in vitro evolution, we comprehensively analyzed whole-genome sequences of 724 *P. falciparum* compound-selected clones, including 210 created specifically for this work. We performed meta-analyses across compound-selected mutations to identify characteristics of functional variants associated with resistance. In addition to revealing previously unknown resistance alleles and genes, our analysis identified protein and genome features associated with drug resistance and specific protein domains that consistently yield resistance. We found that mutations in noncoding regions or the noncore genome [subtelomeric

<sup>1</sup>Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA. <sup>2</sup>Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY, USA. <sup>3</sup>Center for Malaria Therapeutics and Antimicrobial Resistance, Columbia University Irving Medical Center, New York, NY, USA. <sup>4</sup>Global Health Drug Discovery Institute, Beijing, China. <sup>5</sup>Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK. <sup>6</sup>Biological Chemistry and Drug Discovery, Wellcome Centre for Anti-Infectives Research, University of Dundee, Dundee, UK. <sup>7</sup>Global Health Medicines R&D, GSK, Tres Cantos, Madrid, Spain. <sup>8</sup>Department of Internal Medicine, Division of Infectious Diseases, Washington University School of Medicine, St. Louis, MO, USA. <sup>9</sup>Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA. <sup>10</sup>Chemical Biology and Therapeutics Science Program, Broad Institute, Cambridge, MA, USA. <sup>11</sup>Division of Infectious Diseases, Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA. <sup>12</sup>Química Farmacéutica, Departamento de Química Orgánica, Facultad de Química, Universidad de la República, Montevideo, Montevideo, Uruguay. <sup>13</sup>Department of Protozoology, Nekkens Institute for Tropical Medicine, Nagasaki University, Nagasaki, Japan. <sup>14</sup>Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA. <sup>15</sup>Infectious Disease and Microbiome Program, The Broad Institute, Cambridge, MA, USA. <sup>16</sup>Infectious Diseases Research Collaboration, Kampala, Uganda. <sup>17</sup>Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA. <sup>18</sup>Calibr, a division of The Scripps Research Institute, La Jolla, CA, USA. <sup>19</sup>Centre International de Recherche et de Formation en Génomique Appliquée et de Surveillance Sanitaire (CIGASS), Dakar, Senegal. <sup>20</sup>Department of Medicine, University of California, San Francisco, San Francisco, CA, USA. <sup>21</sup>Broad Institute of Harvard and MIT, Cambridge, MA, USA. <sup>22</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, USA. <sup>23</sup>Institute for Biomedicine and Glycomics and School of Environment and Science, Griffith University, Nathan, QLD, Australia.

\*Corresponding author. Email: ewinzeler@health.ucsd.edu

†These authors contributed equally to this work.

‡Present address: Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA.

§Present address: Calibr, a division of The Scripps Research Institute, La Jolla, CA, USA.

and internal regions composed of hypervariable gene families such as *var*, *rifin*, and *stevor* (11)] seldom contribute to resistance phenotypes. Our dataset provides a starting collection for algorithms that can identify genomic changes that are likely associated with drug resistance and presents insights for distinguishing functional from nonfunctional variants in forward genetic approaches.

## Results

### Compound-selected clones contain few mutations

We first sought to assess whether most of the nonsynonymous mutations in compound-selected clones were functional. Using laboratory resources of the Malaria Drug Accelerator (MaDA) Consortium (5), we compiled a set of 724 whole-genome sequences from in vitro-evolved *P. falciparum* strains and their isogenic parents (data S1). Blood-stage parasites were exposed to one of 118 small-molecule growth inhibitors, ranging from tool compounds identified by phenotypic screening as reviewed in (12) to licensed antimalarials to compounds in the developmental pipeline (data S2), for periods ranging from weeks to >2 years. These 118 compounds were collapsed into 95 groups on the basis of shared chemical groups. Although some datasets were previously published, this work contains the meta-analysis of our entire repository of previously reported samples, samples downloaded from the NCBI Sequence Read Archive (SRA), and 210 newly sequenced genomes of parasites resistant to 33 compounds (data S1).

Clones with phenotypic data ( $n = 448$ ) were, on average, 56-fold (ranging from 1.1- to 2654-fold) more resistant to their respective compound than the matched parent (data S1). Although sequenced in multiple locations, genome sequences were acquired using similar paired-end short-read whole-genome sequencing (WGS) methods on Illumina platforms, with an average coverage of  $135\times$ . Most clones were selected from the Dd2 ( $n = 340$ ), Dd2-pol $\delta$  ( $n = 35$ ), or 3D7 ( $n = 328$ ) strains, with a few derived from 7G8, NF54, or W2 (see methods). In some cases, multiple clones were isolated from the same selection flask and proved genetically identical; these were treated as technical replicates.

To standardize identification of mutations that evolved during selection, we aligned sequencing reads to the *P. falciparum* 3D7 reference genome. To identify high-quality variants, we applied hard filtering on SNVs and insertion-deletion variants (indels) according to Genome Analysis Toolkit (GATK) recommendations (13). Because the *P. falciparum* genome is haploid in asexual blood stages and most samples were expected to be clonal, we retained variants for which 90% or more of read calls mapped to alternate (nonreference) alleles. Variants were then assigned predicted effects using the an-

notation tool SnpEff (14) (fig. S1). Because each evolved sequence came with an isogenic parent, we subtracted variants present in the parent that had not emerged over the course of compound selection. Altogether, 5560 SNVs and indels were identified across the dataset (fig. S1), with an average of 17 mutated genes shared within individual compound chemotype groups. Of these variants, 4105 were in the “core” genome, which excludes genomic regions with repetitive sequences (11) (data S3). Each clone contained, on average, 7.65 mutations relative to its parent, of which an average of 1.2 were missense mutations in core regions. Given that there are 5247 core genes in the genome, the data show remarkable specificity and suggest that a large proportion of the identified mutations play some role in compound resistance.

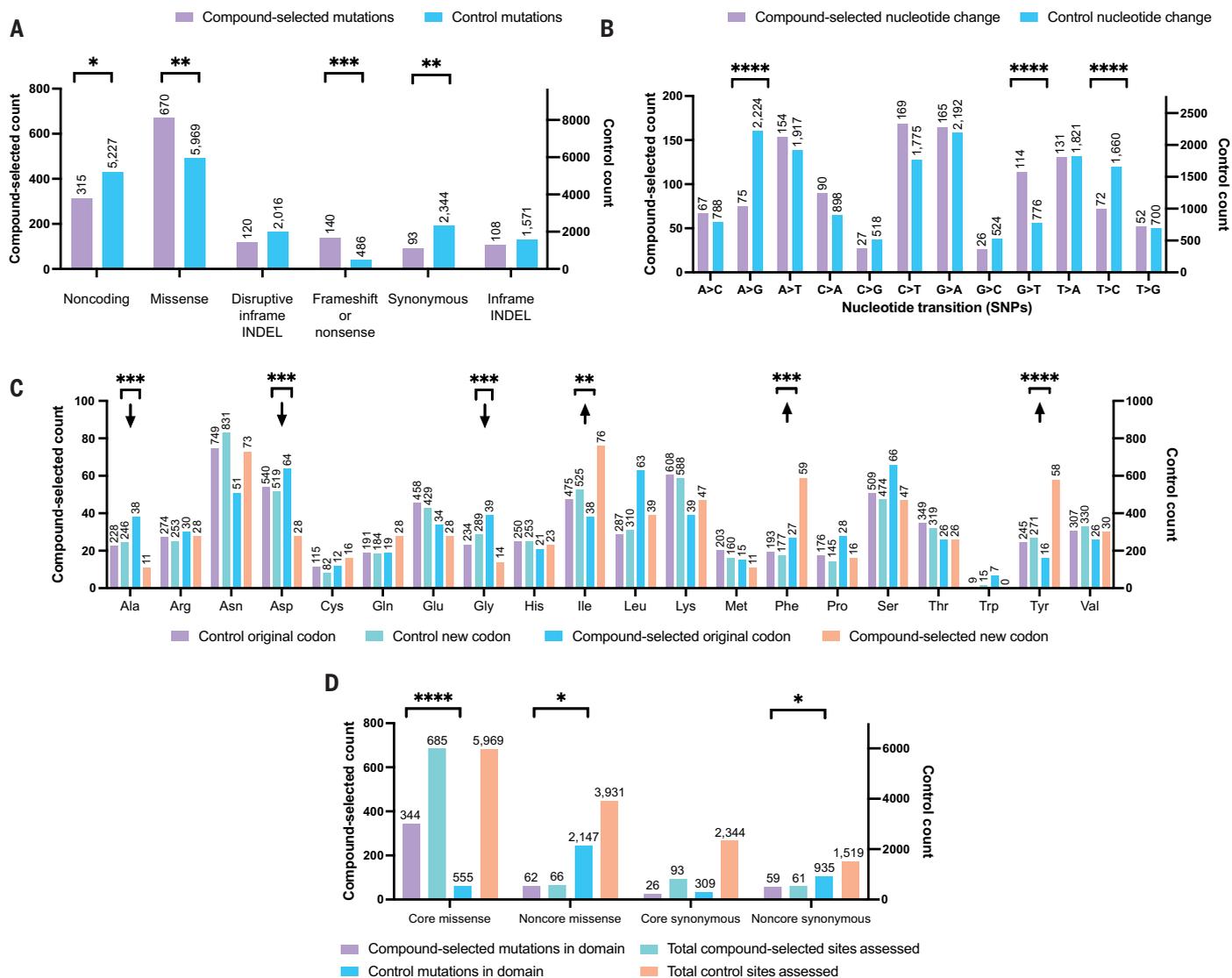
### Compound-selected mutations differ from passenger mutations

The characteristics of our compound-selected set of SNV and indel mutations were compared with a set of control SNVs and indels that differ between two non-compound-selected clonal strains. Removing technical duplicates from the compound-selected set reduced the number to 2628 independently derived mutations, of which 1448 were in the core genome (fig. S1). For the control set, we compared two laboratory strains, Dd2 and 3D7, originally derived from Southeast Asia and sub-Saharan Africa, respectively. Comparisons between Dd2 and 3D7 sequences using an identical pipeline and filters resulted in 48,364 variants, of which 38,064 were core variants. Overall, our compound-selected variants were proportionally more likely to be missense or frameshift mutations (Fig. 1A), whereas variants distinguishing the control clonal isolates were more likely to be synonymous. In terms of nucleotide changes, compound-selected mutations were more likely to be G-to-T transitions (Fig. 1B). Our data suggest that the compounds used in selections were generally not mutagenic, although there could be exceptions. The distribution of amino acid changes was also distinct from that of the control set; transitions were more likely to be from a small to a bulky amino acid such as phenylalanine or tyrosine (Fig. 1C). Finally, we examined whether core mutations were likely to be in a recognized protein domain. Around half of the core missense mutations ( $n = 685$ ) in the compound-selected dataset were located within an InterPro protein domain. By contrast, <10% of the 5969 core missense control mutations ( $n = 555$ ) were in a recognized protein domain (Fig. 1D). These differences in distributions were significant ( $P < 0.001$ , hypergeometric test), highlighting that a substantial proportion of compound-selected mutations are likely to play a functional role and are unlikely to be the result of random genetic drift.

### Copy number variants frequently drive compound resistance

Copy number variants (CNVs) have been shown to mediate clinically relevant drug resistance phenotypes in malaria parasites (15). To identify CNVs in our WGS dataset, we calculated differences in read coverage for core genes and compared coverage to that of a panel of untreated controls assembled from parent clones, depending on whether reads were derived from the 3D7 or Dd2 strain background (data S4 and fig. S2). To identify the amplified or deleted sequence region comprising the CNV, we identified groups of four or more genes with  $\log_2 > 0.4$ -fold change in coverage relative to their parent strain. This analysis resulted in 1168 potential amplification events. Using a Kruskal-Wallis test comparing copy ratios of genes within the amplification boundaries to those averaged across the corresponding panel of controls, we found that 271 of the 724 clones harbored at least one of 420 CNVs with  $P < 0.0001$ . These CNVs included amplifications of known targets, such as *pfpi4k*, *pfproRS*, and the acetyl-coenzyme A transporter (*pfat1*), as well as multidrug resistance genes (*pfabci3*, *pfmdr1*) (Fig. 2A). We also identified multiple CNVs, including one on chromosome 14, that had been noted as emerging after long-term exposure to artemisinin (16). Amplifications of guanosine triphosphate (GTP) cyclohydrolase1 (*pfgchl* on chromosome 12, PF3D7\_1224000) were not generally included in this set, as many strains have more than one copy of *pfgchl* (17).

Although large CNVs were identified with this method, we noticed multiple false-positive small CNVs, especially in clones that showed high levels of gene-to-gene read coverage variation. Because removing small CNV calls might miss functionally important CNVs, we used an orthogonal, non-coverage-based CNV validation approach that took advantage of paired-end sequencing. Paired-end reads show mapping inconsistencies when located near boundaries of structural variants and can give clues about inverted or tandem duplication events (fig. S3). Additionally, paired-end reads near recombination breakpoints may show variation in predicted inserts, with sizes similar to that of the CNV (e.g., 20 kb instead of 1 kb). Paired-end read support as measured by a heuristic algorithm (see methods) or strong copy ratio support, along with confirmation by the structural variant discovery tool DELLY (18), was found for 367 of the original 1168 amplification calls, of which 243 CNVs were independent. About half of the unsupported CNVs were in samples with the highest variability in  $\log_2$  copy ratio across genes (the average interquartile range of  $\log_2$  copy ratios for five samples in the MMV006901 selections was 0.41, compared with a median interquartile range of 0.18 across 3D7 samples containing at least one supported CNV). The other half tended to be in clones isolated from



**Fig. 1. Compound-selected SNVs and indels have different characteristics compared with naturally occurring variants.** (A to D) Comparison between the set of 1448 compound-selected variants (SNVs and indels) in the core genome from this study (left axis) and a set of 17,613 control variants obtained by aligning Dd2 to the 3D7 reference genome. Axes have been normalized to display differences in proportions of different categories for compound-selected (left) and control (right) variants. (B) Reference base and alternate base for

1141 different core SNVs (coding and noncoding) compared with 15,793 control variants. (C) Analysis of amino acid changes for 640 core missense and 6600 core missense control SNVs with arrows showing relative increases or decreases. (D) A total of 905 evolved and 13,763 control protein coding variants were analyzed for location within a defined InterPro domain obtained from PlasmoDB v61. \*\*\*\* $P < 0.0001$ ; \*\*\* $P < 0.001$ ; \*\* $P < 0.01$ ; calculated using a chi-square test.

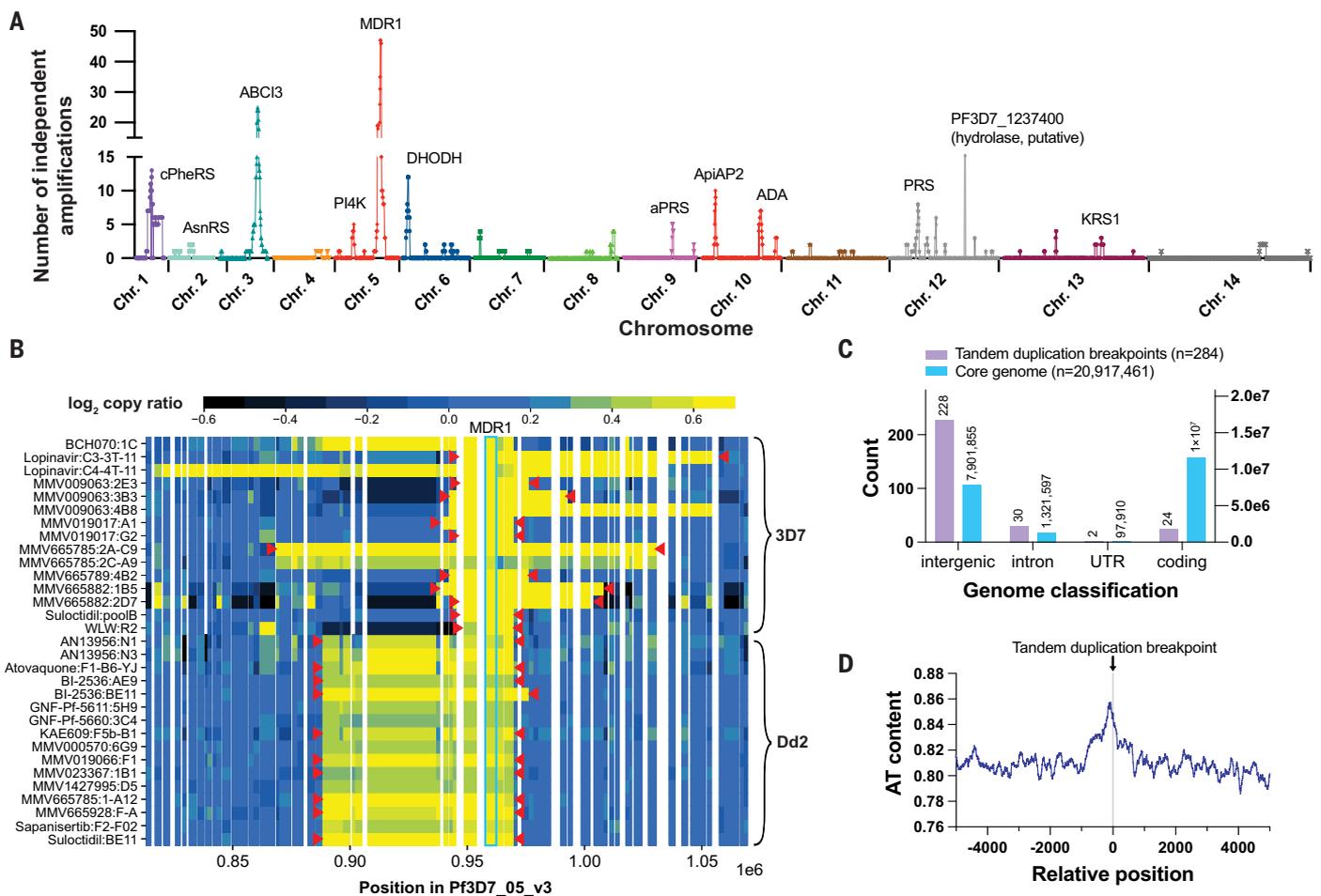
parents distinct from those used to denoise read counts in our CNV analysis pipeline, highlighting the need to keep a carefully matched parent clone for ratio-based methods.

To infer the gene within a CNV that confers a selective advantage, we looked for evidence of a known target or resistance mechanism within the amplified segment or for recurring CNVs and identified the gene that was amplified most frequently across the dataset in overlapping CNVs. *Pfmdr1* was amplified in 47 independent CNV events (Fig. 2A) and harbored 19 independent SNVs, making it a clear driver of resistance. For smaller CNVs, there was more ambiguity. A set of five genes on chromosome 12 was ampli-

fied 11 times, but of this set, only *pfprs* (PF3D7\_1213800, proline-tRNA ligase) also contained SNVs. In several cases, there was no clear candidate resistance mediator, including a recurring CNV on chromosome 1 and one on chromosome 9. Not all CNVs are related to compound pressure. Many evolved strains contained amplification of an 86-kb region on chromosome 10 that was also found in some parent strains, including 3D7; it may contain unidentified fitness-conferring factors.

Characterization of amplification CNVs by structural variant type and breakpoint properties showed several interesting features. Paired-end read data showed that 179 of 243

independent amplifications were tandem duplications, as opposed to inverted or interchromosomal duplications. We did not find evidence that specific genomic locations are strongly preferred for amplification of driver genes, as many CNVs had different endpoints even for the same compound. For example, CNVs with varying size and breakpoint locations containing *pfmdr1* were found in resistant clones evolved from the 3D7 parent strain, which contains a single copy of *pfmdr1* (Fig. 2B). By contrast, all clones that had further amplification of *pfmdr1* compared with the Dd2 parent strain, which already has multiple copies, had similar breakpoints. Many of the amplifications



**Fig. 2. CNVs are a frequent driver of antimalarial compound resistance.**

(A) Manhattan plot showing the number of times each gene was amplified as part of an independent CNV in our dataset. Genes that are likely drivers of the selective advantage conferred by CNVs are annotated, including compound targets and multidrug resistance genes. (B) Heatmap visualization of amplifications containing *pfmdr1* for evolved clones (*pfmdr1* CNVs that independently evolved in the same compound selection were omitted for brevity). Denoised  $\log_2$  gene copy ratios are plotted for each clone; a contiguous segment of high  $\log_2$  copy ratio (yellow)

were tandem duplications, consistent with the original *pfmdr1* CNV in Dd2. CNV breakpoints were determined, with nearly single-base resolution, using discordant read pairs for 142 high-confidence tandem duplications. These breakpoints showed an enrichment of duplication recombination sites in intergenic regions, which typically have lower sequencing coverage (Fig. 2C). Tandem duplication recombination sites often occurred in highly AT-rich and/or repetitive regions, supporting the hypothesis that long A/T tracks present throughout the *P. falciparum* genome facilitate microhomology-mediated repair as a major mechanism of increasing copy number (19) (Fig. 2D).

#### Overrepresented genes affect compound resistance and culture adaptation

Experimental evolution is effective at associating compounds with targets or drug resistance

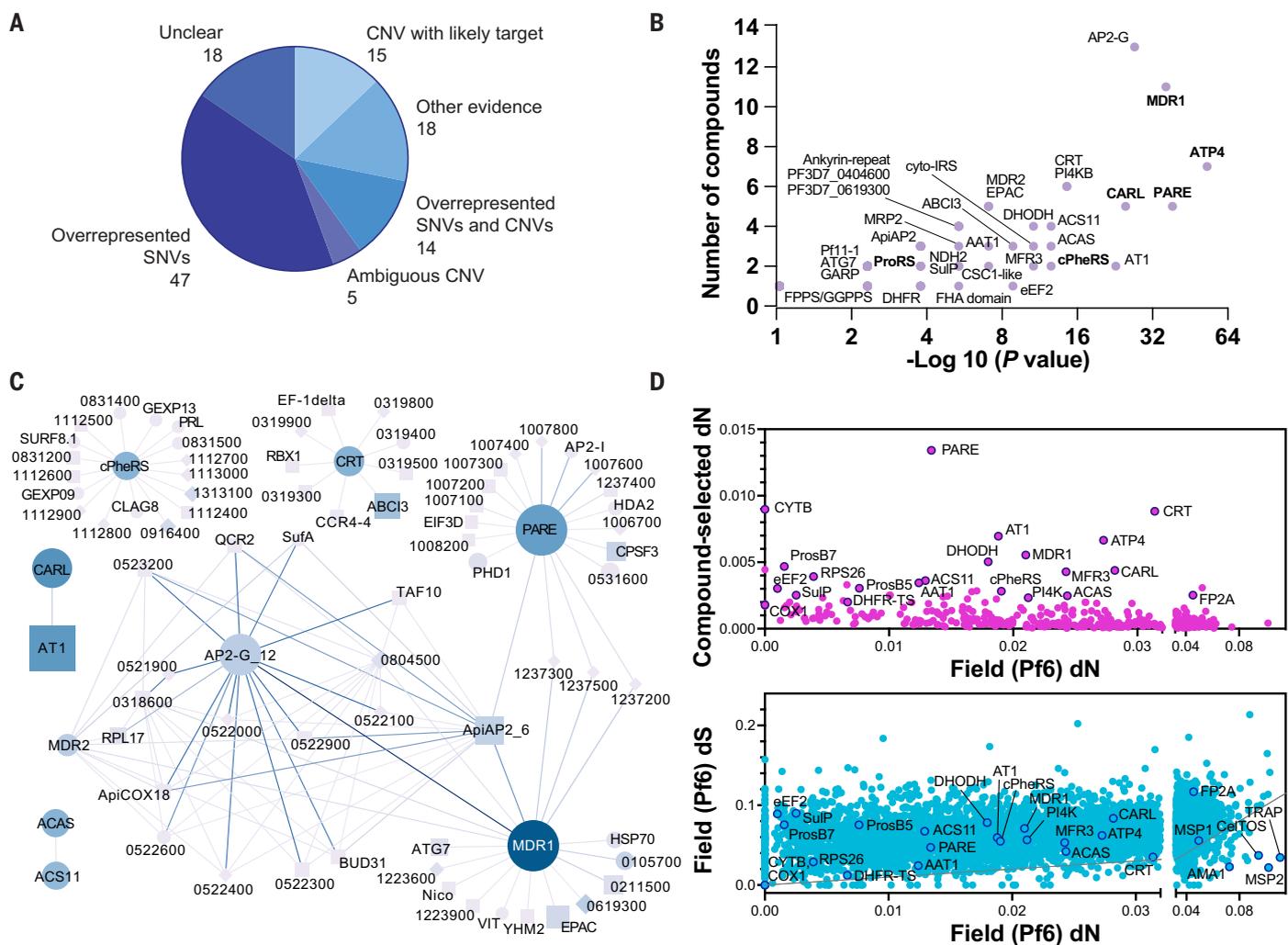
suggests an amplification including those genes. For CNVs identified as tandem duplications, more precise CNV boundaries are indicated by red arrowheads.

(C) Comparison of genomic classification distribution between 284 independent tandem duplication CNV breakpoints and all sites in the core genome. (D) AT content around tandem duplication breakpoints, averaged over the 284 independent tandem duplications. AT content for each relative position was computed as the proportion of bases that are A or T over all sequences beginning at that position with a sliding window of 5 bp. The plot is shown as a running average with a window of 100 bp.

genes if they appear at rates higher than expected by chance. Of the 118 compounds, 47 had a target or resistance mechanism suggested by overrepresented SNVs alone, 14 by both CNVs and SNVs, and 15 by CNVs with a clear driver gene alone (data S5). For 18 compounds, a resistance gene was determined by additional experimentation. In only 18 cases was no resistance gene identified, indicating a high success rate for experimental evolution (Fig. 3A). On a compound basis, the strongest signatures were for KAE609 (cipargamin), with 12 independent SNVs in *pfatp4* (non-SERCA-type  $\text{Ca}^{2+}$ -transporting P-ATPase) ( $P = 9.98 \times 10^{-41}$ ; data S5), and cladosporin, showing three independent CNVs amplifying lysine-tRNA ligase (KRS1) on chromosome 13.

An advantage of this dataset is that it permits the identification of genes that appear repeatedly across disparate compounds, sug-

gesting a multidrug resistance mechanism. Altogether, 128 genes were altered in more than one clone, 53 of which were identified three or more times. The list of statistically significant recurring genes (table S1) contained multiple antimalarial multidrug resistance genes, such as *pfmdr1* (20) (19 independently derived SNVs or indels and 47 CNVs), chloroquine resistance transporter *pfert* (nine independent variants) (20), and cyclic amine resistance locus *pfearl* (14 independent SNVs or indels and one CNV) (21). The list also contained clinically important antimalarial targets, known to acquire resistance mutations affecting inhibitor binding, such as *pfatp4* (22), *pfpi4k* (23), *pfeytb* (24), and *pfihodh* (totaling 14 CNVs and seven SNVs) (25). Another overrepresented gene was *pfap2-g* (PF3D7\_1222600), which encodes a transcription factor involved in commitment to gametocytogenesis (26), as well as the Rap



**Fig. 3. Enriched genes mutated in compound resistance selections.**

guanine nucleotide exchange factor *pfepac*, both known to mutate as a result of long-term culture in the absence of drug pressure (27). Several large, conserved proteins were over-represented, including PF3D7\_0619300, PF3D7\_1464500, and PF3D7\_0510100, each mutated three times independently. Many of the highly over-represented genes (7 of 14 genes mutated at least eight times independently), particularly drug targets, were also contained within CNVs. To identify genes likely to be involved in multidrug resistance, we plotted the likelihood of enrichment by chance against the number

of associated compounds (Fig. 3B and fig. S4). These data showed that genes such as *pfmdr1* and *pfap2-g* were mutated in selections with a wide variety of compounds. By contrast, targets such as *pfproRS* tended to be compound specific. This is not always the case, however, as some high-quality target genes, such as *pfcytb*, were identified in selections with a variety of scaffolds.

Although many of these examples have been published, we also identified previously unknown gene-compound associations. For example, desoxyepothilone B is an analog of epothilone B,

a naturally occurring macrolide isolated from the myxobacteria *Sorangium cellulosum*, which is known to have antitumor and antifungal activity (28). SNVs were identified in PF3D7\_1008700, tubulin beta chain (Ala231Asp, Thr274Ile,  $P = 8.9 \times 10^{-7}$ ). This finding supports published evidence that epothilone's mode of action is taxol-like, inhibiting cell proliferation through microtubule stabilization (29).

GNF-PF5611 is an intracellular copper chelator, also known as neocuproine (30). All resistant clones selected by this compound contained at least one SNV (Lys168Ile, Ser513Leu, and

Ser290\*) in PF3D7\_0915000, which encodes the type II NADH:ubiquinone oxidoreductase (NDH2) (data S3 and S5). NDH2 is one of five dehydrogenases in the *Plasmodium* mitochondrial electron transport chain that donates electrons to ubiquinone. It was considered an attractive antimalarial target until *P. falciparum* asexual blood stage (ABS) nonessentiality was shown (31). Given its proposed mechanism, NDH2 is unlikely to be a direct target of GNF-Pf-5611. Rather, NDH2 mutations may alter interactions with the cofactor flavin adenine dinucleotide (FAD), as they either truncate NDH2 (Ser290\*) or fall near its FAD-binding domains (Lys158Ile, Ser513Leu) (32) (fig. S5). These mutations could help parasites mitigate the effects of oxidation induced by copper chelation.

Selections with MMV008434 yielded six parasite clones with low-grade resistance phenotypes (1.2- to 2.3-fold relative to the parent) (data S1). Of these, three missense SNVs and two disruptive indels were identified in PF3D7\_0609100, putatively annotated as zinc transporter *pfzfp1* (data S3 and S5;  $P = 4.05 \times 10^{-19}$ ). A *PfZFP1* conditional knockout was generated (see methods) and assessed in dose-response assays against MMV008434 and its analog MMV011445. Results confirmed that *PfZFP1* dysfunction contributes to resistance to both compounds (fig. S6). MMV407834 is a compound from the Medicines for Malaria Venture Pathogen Box (33) with midnanomolar activity against ABS parasites (data S2). Selections in both Dd2 and 3D7 strains yielded highly resistant parasite clones [4- to 90-fold median inhibitory concentration ( $IC_{50}$ ) shifts] each containing one of four SNVs (Leu984Pro, Ser980Tyr, Asn36Ile, or Leu800Pro) in PF3D7\_1250200, the putative calcium-permeable stress-gated channel 1 (CSC1-like protein) (data S1, S3, and S5). This gene, which was nonmutable in a mutagenesis screen (34) and is expressed in ABS parasites (35), contains a Pfam domain (PF02714) inherent to a class of osmosensitive calcium-permeable cation channels that is conserved across eukaryotes (36).

#### Network analysis reveals co-occurring mutations

To reveal gene interactions that arise in compound selections, we constructed a network with an edge between two genes if the combination of a specific mutation (SNV, indel, or CNV) in one gene and a specific mutation in the other gene was found in selections with different compounds (Fig. 3C and fig. S7; see methods). This network was, in some cases, able to group genes by function. We identified well-known interactions, such as those between *pfat1* and *pfear1*, or *pfacas* (PF3D7\_0627800) and *pfacs11* (PF3D7\_1238800) (37). The data also showed that three of six compounds that gave rise to *pfert* SNVs also selected for amplifications of *pfabc3* (Fig. 3C). There was a

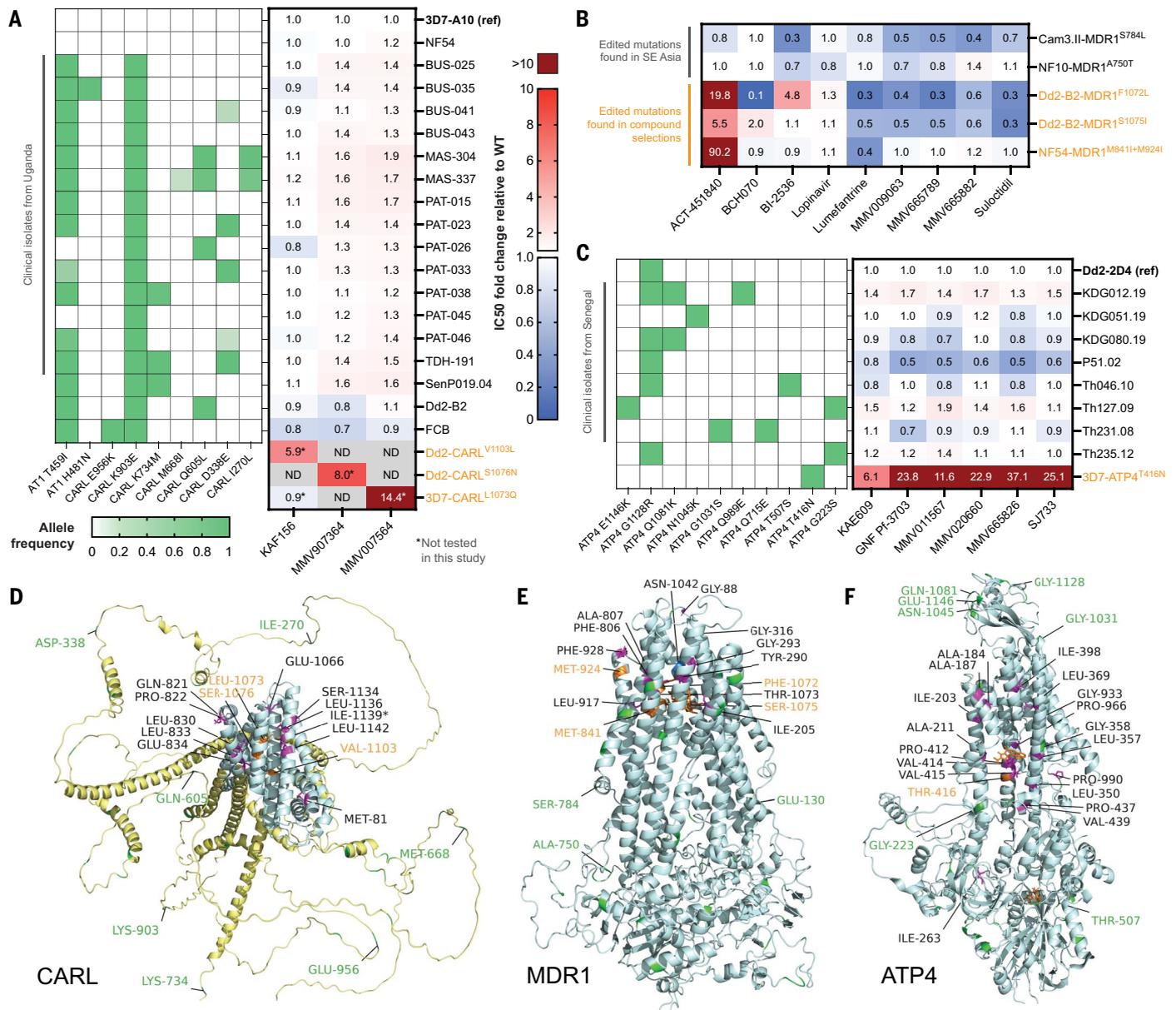
strong association between *pfap2-g* SNVs or indels and *pfmdr1* SNVs or CNVs. Of the 12 compounds yielding clones with *pfap2-g* mutations, seven also yielded *pfmdr1* mutations, and in five independent cases, both *pfap2-g* and *pfmdr1* mutations were found in the same clone. Notably, 11 of 14 different *pfap2-g* mutations were nonsense or frameshift, and only these loss-of-function *pfap2-g* mutations were found alongside *pfmdr1* amplification or missense SNVs in selections with the same compound. Furthermore, ~58% of genes on chromosome 12 with associations in the main network (7/12) were connected to both *pfmdr1* and *pfap2-g*. It may be that parasites cultured long term are likely to acquire loss-of-function mutations in *pfap2-g* because of pressure to eliminate genes causing gametocyte conversion in vitro, whereas *pfmdr1* point mutations arise infrequently. Another possibility is that *pfap2-g* regulates *pfmdr1*, and loss of *pfap2-g* function results in increased transcription of *pfmdr1*. Similarly, *pfapi-ap2* (PF3D7\_0420300), a gene mutated in artemisinin selections (16), had associations with the conserved protein PF3D7\_1324300. Another transcription factor, *pfapi-ap2* (PF3D7\_0613800, *ApiAP2\_6* in Fig. 3C and fig. S7), involved in the *Plasmodium* cell cycle and implicated in drug resistance evolution (38), had three associations with *pfpare* and 10 with *pfmdr1*. This finding further supports the role of *pfapi-ap2* in contributing to multidrug resistance and culture adaptation in *Plasmodium*. *Pfap2-i* (chromosome 10) was part of a frequent amplification event. Although there are few reports of transcription factors playing a role in antimalarial drug resistance, this is common in other species such as *S. cerevisiae* (10).

#### Compound-selected mutations differ from naturally occurring variants

An important question in antimalarial drug development is whether natural parasite populations can readily attain resistance because of existing genetic variants and the plasticity of the *Plasmodium* genome. We compared our mutational set to nucleotide and amino acid substitution rates in the Pf6 dataset (39), which includes sequences of 7113 *P. falciparum* isolates obtained from 28 countries (data S6). SNVs that arose at least twice independently in our selections were rarely present in the Pf6 dataset (6 of 53). Genes with higher  $d_N$  in the compound-selected set, measured as number of nonsynonymous SNVs normalized by nonsynonymous sites in the gene (see methods), tended to have low  $d_N$  in Pf6. This suggests that in vitro selections typically introduced novel evolutionary pressures in genes that are highly conserved in the field, such as the proteasome  $\beta 2$  subunit (PF3D7\_1328100), *pfdhfr*, and elongation factor 2 (PF3D7\_1451100) (Fig. 3D). We also show that multidrug resistance genes, such as *pfmdr1* or *pfert*, were less conserved in field

isolates than canonical enzymatic drug targets. Our data show that 4418 nuclear genes with SNVs (excluding singletons) in the Pf6 dataset had no nonsynonymous SNVs among the 724 evolved clones (excluding singletons). Among these are genes involved in antigenic variation and immune response, such as *pfesp* (PF3D7\_0304600), *pfama1* (PF3D7\_1133400), *pfmsp2* (PF3D7\_0206800), and *pfeltos* (PF3D7\_1216600), which have nonsynonymous variation in the field but were not mutated in our compound selections (40, 41) (Fig. 3D).

The large number of missense mutations in certain genes, such as *pfmdr1*, in both compound-selected and field samples offers the opportunity to explore whether compound-selected mutations are localized to specific protein domains. We examined the locations of compound-selected amino acid substitutions in PfCARL, PfMDR1, PfATP4, and PfCYTB (fig. S8), all of which had 10 or more independent mutations. PfCYTB and PfMDR1 mutations mediate resistance to clinical antimalarials, whereas mutations in PfATP4 and PfCARL mediate resistance to agents in late stages of development, namely cipargamin (KAE609) and ganaplacide (KAF156) (21, 42). We obtained protein models using either SWISS-MODEL (43) or AlphaFold (44) and mapped predicted ligands, including inhibitors, using AlphaFill. Because access to the PfMDR1 structure in Protein Data Bank (PDB) (45) was unavailable at the time of the study, we used a homology model. In the PfMDR1 homology model [SWISS-MODEL Template Library (SMTL) ID: 7a69.1] based on human ABCB1 bound to vincristine (46), all laboratory-derived mutations clustered in transmembrane domains that make up the predicted small-molecule binding site, and virtually all mutations were located in predicted alpha helices (fig. S8A and Fig. 4E). PfCARL is a predicted transmembrane protein, and although some domains are well conserved, thereby allowing high-confidence AlphaFold predictions (UniProt COH483), many regions cannot be effectively folded. Notably, all 12 *pfear1* mutations from this study are found in “ordered” domains (AlphaFold per-residue confidence score based on the IDDT structural similarity metric, pLDDT > 70), specifically predicted alpha helices (fig. S8C and Fig. 4D), as opposed to regions predicted to be unstructured under physiological conditions (44). Likewise, in the PfATP4 model, 17 of 19 mutated residues (with the exception of Pro<sup>990</sup> and Ile<sup>263</sup>) are found in alpha helices, with most located near the AlphaFill-predicted docking site of a close analog (PDB ligand ID: CZA) of cyclopiazonic acid, a potent inhibitor of the human ortholog of PfATP4, SERCA1a (Ca<sup>2+</sup>-ATPase of skeletal muscle sarcoplasmic reticulum) (47), which has 80.9% Tanimoto similarity to KAE609 (fig. S8B and Fig. 4F). The Pro<sup>990</sup> mutation is observed in concert with Ile398Phe and may be compensatory. PfCYTB mutations also localized



**Fig. 4. Compound susceptibility assays comparing compound-selected mutations and field variants in *pfcarl*, *pfmdr1*, and *pfatp4*.** (A) Heatmap showing mean  $IC_{50}$  fold changes relative to 3D7-A10 from dose-response experiments for four laboratory lines and 15 clinical isolates from Uganda and Senegal with *pfcarl* variants against three compounds that select for PfCARL mutations. *Pfcarl* and *pfat1* variants were confirmed by WGS; the left heatmap shows allele frequencies for the tested lines. For reference, fold changes based on previously reported  $IC_{50}$  values are shown for resistant lines with PfCARL substitutions, highlighted in orange. (B) Mean  $IC_{50}$  fold changes relative to isogenic parent for five edited lines with *pfmdr1* mutations. Parasite lines highlighted in orange have mutations identified from compound selections, whereas the rest occur in the field. (C) Mean  $IC_{50}$  fold changes relative to Dd2-

Dd4 for eight clinical isolates from Senegal and a KAE609-pressured resistant mutant, 3D7-ATP4<sup>T416N</sup>, highlighted in orange. *Pfpat4* allele frequencies confirmed by WGS are displayed in the left heatmap. (D to F) Ligand-filled models of PfCARL and PfATP4 were obtained from AlphaFill; PfMDR1 homology model was constructed with SWISS-MODEL (43) using 7a69 (see fig. S8 for details). Mutated residues in our compound-selected dataset [with the exception of PfCARL Ile-1139, which was found in (21)] are colored magenta; mutated residues in the field with global major allele frequency (gMAF) > 0.002 were obtained from (39) and are colored green, while those found in both sets are colored blue. Labels are colored orange and green, respectively, for compound-selected versus field mutations phenotyped in (A) to (C). Other compound-selected mutations are labeled in black.

near predicted binding pockets of heme and stigmatellin, a known inhibitor of the cytochrome bc<sub>1</sub> complex (48) (fig. S8D).

We next examined the hypothesis that compound-selected missense mutations could be spatially distinguished from substitutions

in *P. falciparum* field isolates. We considered SNVs with total read depth  $\geq 50$  and alternate allele frequency (AAF)  $\geq 0.5$  in at least one of the 7113 samples in Pf6 as well as global AAF (gAAF) > 0.002, as calculated by summing allele depths across all Pf6 samples (data S7). Placing

the 37 highest frequency PfMDR1 missense variants in Pf6 on the homology model showed that field isolate mutations were dispersed throughout the predicted structure. Only one residue was found among both compound-selected and high-confidence field isolate mutations: Asn<sup>1042</sup>,

a variant of clinical interest (in particular, Asn1042Asp) (49) with demonstrated functional importance in modulating parasite susceptibility to 4-aminoquinoline-based drugs (50). Among less prevalent alleles in the Pf6 dataset, a small population (14 samples with total depth > 20) of field parasites from Cambodia (collected 2011–2012) contained PfMDR1 Gly293Asp, the same residue mutated in selections with the HIV drug lopinavir (Gly293Val) and the kinase inhibitor BI-2536 (Gly293Cys) (fig. S8A and Fig. 4E). A single Cambodian sample collected in 2008 also contained a PfMDR1 Phe806Leu, which was selected for in vitro with MMV026596. In contrast to our compound-selected mutations, which were all nonsynonymous, 28 of 65 frequently observed mutations in the Pf6 dataset encode synonymous changes. Similar results were observed for PfCARL, with all except one (Ile<sup>1235</sup>) of the 32 high-confidence alleles from field isolates (gAAF > 0.002, at least one high-confidence sample) located in predicted disordered regions (pLDDT < 70) and none overlapping with the in vitro selected set (fig. S8C and Fig. 4D). For PfATP4, the 40 high-confidence field mutations also appeared to be randomly distributed throughout the model (fig. S8B and Fig. 4F). We noted a PfATP4 Gly223Ser mutation that was quite prevalent in African samples (51) and occurred at the same residue as Gly223Arg identified in selections with the spiroindolone KAE678 (22), which conferred an ~7-fold IC<sub>50</sub> increase (data S1).

#### Experimental variant validation in PfMDR1, PfCARL, and PfATP4

To compare the effect of missense mutations arising from compound selections to that of naturally occurring polymorphisms, we tested the sensitivity of different PfMDR1, PfCARL, and PfATP4 mutant parasites to panels of compounds. For PfCARL, 15 culture-adapted clinical isolates from Uganda and Senegal with variants relative to 3D7 and four laboratory lines (NF54, 3D7-A10, Dd2-B2, and FCB) were phenotyped against three compounds that selected SNVs in *pfcarl*: KAF156 (ganaplacide), MMV907364, and MMV007564 (Fig. 4A, fig. S9, and table S4). Whole-genome sequencing of the 19 lines showed various missense SNVs in *pfcarl*, *pfat1*, and *pfugt*, which are additional multidrug resistance genes implicated in KAF156 and GNF179 resistance (tables S2 and S3). Notably, all *pfcarl* mutations identified in the field isolates are in disordered regions of the protein's AlphaFold structure (Fig. 4D). Mean IC<sub>50</sub> fold changes compared with wild-type 3D7-A10 were small, reaching a maximum of 1.2-fold for KAF156 (for MAS-337), 1.6-fold for MMV907364 (for MAS-304, MAS-337, PAT-015, and SenP019.04), and 1.9-fold for MMV007564 (for MAS-304) (fig. S9 and table S4; metadata shown in table S2). In contrast, lines from resistance selections with the respective compounds with PfCARL substitutions (Val1103-

Leu, Ser1076Asn, and Leu1073Gln) yielded substantially higher IC<sub>50</sub> fold shifts of >5, on the basis of previously reported IC<sub>50</sub> values (52, 53) (Fig. 4A and data S1).

For PfMDR1, we tested the sensitivities of five lines with missense mutations introduced through gene editing, and their corresponding parental lines, to eight compounds that previously yielded *pfmdr1* SNVs and/or CNVs in selections: ACT-451840, MMV665789, MMV009063, BI-2536, lopinavir, sulcotidil, BCH070, and MMV665882. Lumefantrine, whose potency can be modulated by some *pfmdr1* variants, was also tested (tables S5 to S7). All edited *pfmdr1* mutations were confirmed using Sanger sequencing (fig. S10). Ala750Thr and Ser784Leu are naturally occurring variants previously described in field isolates from Western Cambodia and the Thai-Myanmar border (54); CRISPR-Cas9 or zinc-finger nuclease (ZFN)-based editing was used to create recombinant mutants with these PfMDR1 substitutions on the NF10 and Cam3.II<sup>C580Y</sup> backgrounds, respectively (see methods). Significant differences in mean IC<sub>50</sub> and IC<sub>90</sub> (90% inhibitory concentration) values across four or five biological replicates (Mann-Whitney *U* tests) between mutant lines and their isogenic parents indicate that PfMDR1 Ser784Leu increased parasite susceptibility to five of the eight tested compounds, whereas Ala750Thr did not influence the activity of any (Fig. 4B and fig. S11). PfMDR1 Met841Ile+Met924Ile, identified in ACT-451840 selections and edited into NF54 parasites (55), only altered susceptibility to ACT-451840. Phe1072Leu and Ser1075Ile were identified after selection with GNF-Pf-5660 and GNF-Pf-5668, respectively, and each mutation was edited into Dd2-B2 parasites using CRISPR-Cas9 (56). Phe1072Leu conferred resistance to BI-2536 and sensitization against BCH070; by contrast, Ser1075Ile did not. Both mutations conferred cross-resistance to ACT-451840 and are positioned next to vincristine in the PfMDR1 homology model (Fig. 4E). Furthermore, BCH070 selections yielded a substitution, Thr1073Ile, near the substrate-binding domain inside the transport channel of PfMDR1 in its inward-facing conformation (45). Consistent with the hypothesis that PfMDR1 binding-domain mutations modulate substrate recognition, both Phe1072Leu and Ser1075Ile increased susceptibility to compounds that select for PfMDR1 amplifications.

Ten parasite lines, including Dd2-2D4 (a clone of Dd2), 3D7-ATP4<sup>T416N</sup> [a KAE609-pressured resistant mutant (57)], and eight clinical isolates from Senegal (58) with distinct *pfatp4* genotypes, were phenotyped against 12 compounds, of which six have selected for resistant alleles in *pfatp4* (KAE609, SJ733, MMV665826, MMV020660, MMV011567, GNF-Pf-3703), whereas the rest are commonly used antimalarial drugs (artemether, piperazine, amodiaquine, mefloquine, atovaquone, chloroquine) (Fig. 4C, fig. S12, and table

S10). Whole-genome sequencing was performed on all lines and clinical isolates to profile missense mutations in *pfatp4* (tables S8 and S9). Only 3D7-ATP4<sup>T416N</sup> showed a substantial increase in IC<sub>50</sub> relative to Dd2-2D4 for the six *pfatp4*-associated compounds. While some of the field mutations are in ordered regions of the PfATP4 AlphaFold structure, only Thr416Asn is located near the AlphaFold-predicted binding site of the cyclopiazonic acid analog (PDB ligand ID: CZA), a known inhibitor of the human ortholog of PfATP4 (47) (Fig. 4F).

#### Discussion

Here, we present a comprehensive dataset of compound-selected resistance alleles for the *P. falciparum* malaria parasite. We anticipate that this resource and insights into shared characteristics of resistance-conferring alleles will—like MalariaGEN's latest Pf7 dataset of field isolate sequences—be useful for several applications in the discovery and deployment of antimalarial drugs.

The impressive breadth of resistance mechanisms across our dataset indicates that *P. falciparum* evolves resistance with relative ease. Indeed, half the compounds in our dataset, which encompasses much of the chemical space in the current portfolio of next-generation antimalarials, appear vulnerable to resistance acquisition. Drug development strategies should therefore minimize this liability by prioritizing resistance-refractory compounds and by identifying collateral sensitivity pathways, in which resistance to one antimalarial increases sensitivity to another, to inform combination therapies. Analysis of recent African *P. falciparum* isolates presumed to be multidrug-resistant showed that a region on chromosome 12 containing *pfap2-g* (PF3D7\_1222600) and *ap2-12* (PF3D7\_1222400) (59) is under strong selection. In vitro-evolved artemisinin-resistant parasites in our dataset had mutations in an AP2 transcription factor on chromosome 4 (PF3D7\_0420300) as well as a chromosome 14 amplification containing an AP2 transcription factor (PF3D7\_1456000) (16). PF3D7\_0420300 was also mutated in strains pressured with GNF179, closely related to ganaplacide, a compound in advanced stages of development that is likely to be licensed to treat malaria, whereas mutations in its 5' untranslated region were found in parasites pressured with atovaquone or the experimental compounds AN13762 and MMV024038. These data show the high level of interconnectivity among resistance mechanisms for both existing drugs and compounds under development.

As we show, in vitro evolution of compound resistance typically gives rise to few mutations over the course of compound selection, compared with the thousands of genetic variants that distinguish even slightly diverged isolates from the field. However, it remains the case that most SNV or indel mutations in our

dataset likely do not drive compound resistance and instead are neutral mutations or improve fitness in *in vitro* culture. Moreover, not all genes enriched for *in vitro* evolved mutations are drivers of compound resistance; some may play roles in culture adaptation, whereas multigene families in noncore hypervariable regions were also frequently mutated. Further experimental work is needed to validate the roles of these overrepresented genes.

Although transcription factor mutations appear to play a larger role in drug resistance in other microbial species, such as *S. cerevisiae* (10), the frequent appearance of *P. falciparum* mutations in ApiAP2 transcription factors indicates that transcriptional regulation may play an important role in stress adaptation in malaria parasites. We found mutations in AP2 transcription factors at higher rates than expected, especially when comparing sequences of laboratory-selected parasites to those of parasite field isolates. AP2 transcription factors were first identified in *Plasmodium* on the basis of their similarity to the AP2 transcription factor family in plants; these proteins are now known to be key regulators of various stress responses (38, 60). Mutations in ApiAP2 genes were frequently selected in this study. One example is PF3D7\_0613800, which had mutations after selections with nine different compounds but did not bear more than one independent mutation for any compound. While three of the selected PF3D7\_0613800 mutations are missense, there are also in-frame deletions, all in structurally disordered domains that are less likely to play a functional role. In *S. cerevisiae*, gain-of-function mutations in Zn<sub>2</sub>C<sub>6</sub> transcription factors *yrr1* and *yrm1* are also primarily found in distinct, structurally disordered but spatially conserved regions (10).

Our dataset can also serve to inform considerations of nonmalarial resistance mechanisms. For example, the dataset can be a resource for identifying potential drug resistance alleles in microbial pathogens or cancer cells collected from patients treated with dihydroorotate dehydrogenase inhibitors, which frequently target the same ubiquinone-binding pocket in a range of organisms (61). Our work suggests that drug resistance mechanisms are often conserved across species. As an example, cladosporin resistance has been associated with lysyl-tRNA synthetase mutations in yeast, and in *Plasmodium*, cladosporin selected for amplifications of lysyl-tRNA synthetase (62).

Consistent identification of specific mutations in independently evolved resistant parasites was usually sufficient to identify driver mutations, many of which were subsequently validated through genetic editing. However, in several cases, the variant(s) underlying resistance remained unclear. This challenge is pronounced for selections in which many nonsynonymous mutations arose after long selection times or were

identified in the hypermutable *P. falciparum* Dd2-pol $\delta$  strain. Secondary contributors to antimalarial resistance, variants sequenced with low allele frequency due to multiple gene copies or nonclonality, or variants with less obvious effects such as noncoding DNA mutations or balanced structural variants could be overlooked. Improved mutation calling methods could potentially uncover additional contributors to antimalarial resistance in this dataset, and an *in silico* approach to variant prioritization may help elucidate resistance-conferring alleles in unresolved selections.

While an advantage of our dataset is the diversity of compounds and resistance mechanisms included, certain compounds and drug targets are overrepresented. Thus, caution must be taken to ensure that models trained on this dataset do not learn compound-specific biases. The mutations we identified also should not be interpreted as capturing all possible resistance drivers for any given parasite strain and compound, as most selections yielded only a few independent resistant clones. Further experimental work, such as continuous directed evolution (63) and minimum inoculum of resistance (MIR) studies (64), will be required to fully characterize how resistance arises against specific compounds.

There are limitations to using *in vitro*-selected mutations to inform drug development decisions. For instance, *in vitro* models of *Plasmodium* infection cannot recapitulate the selective pressures applied by the host immune system and other host-pathogen, host-vector, or host-xenobiotic interactions. Additionally, most laboratory strains of *P. falciparum* are not derived from recent parasite isolates. Rather, these parasites likely contain multiple adaptations to long-term culture, and their genetic backgrounds may not be fully representative of current natural parasite populations. Despite differences in selective pressure in the laboratory versus the field, insights from experimentally controlled compound selections are valuable.

Finally, our work reveals how to distinguish phenotype-driving variants from passengers, a key challenge of forward genetics approaches. This dataset and others could empower future machine learning-based approaches to estimate variant functionality *in silico*.

## Materials and methods

### Data acquisition

Sequencing information for parasite samples with evolved resistance to 25 antimalarial compounds were downloaded from the NCBI SRA (<https://www.ncbi.nlm.nih.gov/sra>) or acquired through direct correspondence with the senior author(s) of published work. These compounds include the gold-standard antimalarial drug artemisinin (16); the benzoxaboroles AN13956, AN13762, and AN10248 (65); the triazolopyrimidine analog series including DSM1 (25), DSM265,

and DSM267 (66); the imidazolopiperazine GNF179 (52); halofuginone, a synthetic derivative of the natural quinazolinone alkaloid efbrifugine (67); the boronate human proteasomal inhibitor bortezomib (68); the pantothenamide CXP18.6-052 (69); the dihydroisoquinolones SJ733, SJ247, SJ311, and SJ279 (70); the 2,6-disubstituted quinoline-4-carboxamide DDD107498 (71); the antitubercular clinical candidate SQ109 (72); the peptide vinyl sulfones WLL-vs and WLW-vs (73); the primary sulfonamide glycoside PS-3 (74); the bis-1,2,4-triazine compound MIPS-0004373 (75); MMV019313 (76); the trisubstituted imidazole MMV030084 (77); the 3-hydroxypropanamide compound 22 (78); and the pyrazolopyrimidine sulfamate ML901 (79).

### Parasite strain information

*P. falciparum* 3D7 is a mostly drug-sensitive line, although it is partially resistant to sulfadoxine and contains an amplification of GTP cyclohydrolase (17). The 3D7-A10 clone has been used by MalDA in previous compound selection and sequencing efforts (15). Dd2 is a multidrug-resistant line originating from an Indochina III/CDC isolate that contains a mutant *pfprt* sequence and amplifications in *pfmdr1* and GTP cyclohydrolase that confer reduced susceptibility to chloroquine, mefloquine, lumefantrine, and pyrimethamine (80). Dd2-B2 and Dd2-2D4 are genetically homogeneous, independently isolated lines that were cloned from Dd2 by limiting dilution (56, 81). The Dd2-pol $\delta$  line is a Dd2 strain that was genetically modified at the DNA polymerase  $\delta$  subunit (Asp308Ala and Glu310Ala) to disrupt proofreading function and increase accumulation of mutations over successive replication events (82). W2 is an Indochinese chloroquine-resistant line, and 7G8 is a Brazilian line resistant to chloroquine, amodiaquine (partially), and pyrimethamine (83, 84). V1/S is a chloroquine-, pyrimethamine-, and cycloguanil-resistant strain from Vietnam (85). Isolates from Uganda were collected and cultured as previously described (86) and studied after culture adaptation and shipping to the United States. Isolates from Senegal were collected and culture-adapted as previously described (58).

### *In vitro* evolution of compound-resistant parasites

*P. falciparum* parasites were cultured in RPMI1640 media supplemented with 0.5% AlbuMAX II, 4.3% human serum, 25 mM HEPES, 25 mM NaHCO<sub>3</sub>, 0.36 mM hypoxanthine, and 100  $\mu$ g/ml gentamicin. Cultures were maintained in leukocyte-depleted red blood cells (RBCs) at 2.5% hematocrit (HCT) and incubated at 37°C in an atmosphere of 3% O<sub>2</sub>, 4% CO<sub>2</sub>, and 93% N<sub>2</sub>. In some cases, only AlbuMAX II supplemented RPMI media without gentamicin and with 0.15 mM hypoxanthine, or only human serum supplemented RPMI media with gentamicin and with 0.15 mM hypoxanthine, was used in standard culture

conditions. The human biological samples were sourced ethically, and their research use was in accord with the terms of the informed consents. Resistant parasites were generated using either a high-pressure pulse, ramp-up, stepwise, or constant method of compound exposure as previously described (15). Cultures were maintained under selection conditions until they demonstrated a reproducible IC<sub>50</sub> fold-shift of at least 3×, at which point parasites were cloned in 96-well plates by limiting dilution (87).

Compound IC<sub>50</sub> was assessed in dose-response format including using a SYBR Green-I based cell-proliferation assay as previously described (88). Parasites were incubated for 72 hours in 96- or 384-well plates with exposure to the compound of interest in a 12-point dilution series. An artemisinin dilution series was conducted in parallel as a positive control. Following the incubation, parasites were lysed, and DNA was stained using SYBR Green fluorescence and was measured at 535 nm on an Envision plate reader (Perkin Elmer, Waltham, MA) or SpectraMax iD5 plate reader (Molecular Devices, San Jose, CA) after excitation at 485 nm. For some lines, *P. falciparum* growth inhibition was determined using a modified *in vitro* [<sup>3</sup>H]-hypoxanthine incorporation method as previously described (24) or flow cytometry on an iQue Screener PLUS (Sartorius) with parasites stained with 1× SYBR Green and 100 nM MitoTracker Deep Red (56, 77). Dose-response curves were fitted, and log(IC<sub>50</sub>) values calculated using Prism (GraphPad Prism, La Jolla, CA) or Excel and Grafit 5 software. IC<sub>50</sub> fold-shift changes were calculated by comparing IC<sub>50</sub> of the resistant clones to that of the corresponding compound-sensitive parent line.

#### Whole-genome sequencing analysis and annotation of variants

Raw sequencing reads were aligned to the *P. falciparum* 3D7 reference genome (PlasmoDB v13.0) and preprocessed following standard GATK version 3.5 protocols (13). SNVs and indels were called with GATK HaplotypeCaller and filtered to retain high-quality variants. SNV calls were retained if they did not meet the following exclusion criteria: ReadPosRankSum > 8.0 or < -8.0, QUAL < 500, Quality by Depth (QD) < 2.0, Mapping Quality Rank Sum < -12.5, and filtered depth (DP) < 7. Indels were retained if they did not meet the following exclusion criteria: ReadPosRankSum < -20, QUAL < 500, QD < 2, and DP < 7. Allele fraction cut-offs for alternate alleles were ≥ 0.35 for bulk samples and ≥ 0.90 for clonal samples. Functional annotation of variants was carried out using SnpEff (14) with a custom database built from the 3D7 GFF from PlasmoDB (<https://plasmodb.org/plasmo/app>). Mutations that were considered background (native to the compound-sensitive parent) were removed, leaving a list of only high-quality mutations

that had evolved over the course of the compound selection process.

Mutations were assessed for whether they were positioned within annotated InterPro domains, low-complexity regions as defined by the PlasmoDB Genome Browser, or 3D7 5'/3' untranslated regions (UTRs) as defined by Chappell *et al.* (89).

CNVs were detected by calculating denoised log<sub>2</sub> copy ratios across gene intervals through the GATK 4.0 CNV pipeline. We constructed two separate panels of normals for Dd2 and 3D7, using 30 independently sequenced parent clones of each genetic background. Read counts for each sample were calculated across a predefined gene interval list where intergenic regions and the highly variable *pfvar*, *pfprifin*, *pfstevor*, and *pfsurfin* genes (90) were removed. After denoising against the strain-matched panel of normals, log<sub>2</sub> copy ratios were calculated for each gene interval. CNVs were retained if at least four sequential genes showed a denoised log<sub>2</sub> copy ratio of at least 0.4, indicating potential gene duplication.

To account for a subset of samples yielding false positives because of noisy or altered coverage profiles, we further filtered the candidate CNV list obtained from segmentation of genic copy ratios on the basis of three additional criteria: statistical significance of difference in copy ratio vectors, discordant read-pair support for tandem duplications, and overlap with an independent CNV calling method, DELLY (18). DELLY is an integrated structural variant discovery method that uses both paired-end and split-read analysis to call CNVs. To assign confidence to CNVs on the basis of the copy ratio data, a Kruskal-Wallis test was performed on the copy ratio vectors of each candidate CNV and averaged copy ratio across all parent samples in the corresponding panel of normals. The resulting *P* values were used to quantify how strongly a candidate CNV's copy ratio vector deviates from parents. As orthogonal evidence for the presence of duplications, we wrote an algorithm that searches for discordant paired-end reads within a 10-kb window of candidate CNV boundaries, excluding non-core regions. This approach was effective for validating tandem duplications that produce a discordant read-pair signal at both boundaries, which is unlikely to occur multiple times by chance. Although DELLY alone yielded a large number of false positives as well, sufficiently high-quality DELLY calls were useful for validating candidate CNVs with strong discordant read-pair and split-read support that were not necessarily tandem duplications.

On the basis of manual inspection of select CNVs, we decided on a set of heuristics for filtering candidate CNVs to produce the final list. Samples were classified on the basis of interquartile range (IQR) of core genome-wide gene copy ratios; stricter support thresholds were

applied to those with high (>0.3) or intermediate (0.12 to 0.3) IQR. Adjacent candidate CNVs were also combined. To optimize sensitivity and specificity on a manually validated test set, candidate CNVs were retained if they either had both low *P* value and sufficient overlap with DELLY calls, or both tandem duplication support (not near noncore regions) and some overlap with DELLY calls. The specific thresholds used in the filtering step are described in data S4.

#### Gene network construction

A subset of the variants was generated using SNV or indel (data S3) and CNV (data S4) data filtered for core genes and mutation types, that is, missense, disruptive in-frame insertion or deletion (indel), frameshift, start loss, stop gain or loss, splice region variant, and combinations of these. The list was processed to create pairs of compound-gene(s) and compound-gene-variant type. The network was generated using these pairs, having genes as nodes, and edges where two genes have at least once instance of a pair of compounds yielding resistant samples with the same SNV/indel (gene 1)–SNV/indel (gene 2) or SNV/indel (gene 1)–CNV (gene 2) variant pair. The network was visualized using Cytoscape v.3.9.1 organic layout (91). The node color represents the score calculated on the basis of the variant type, adjusting for nonsense (light blue, lower score) to missense (dark blue, higher score); the total of variants is shown by the node's size adjusting for CNV (smaller) versus SNV/indel (bigger). Edge intensities show the total number of mutations combined by the compound pair from light gray (less mutated) to blue (highly mutated). Highly variable multigene families (*pfvar*, *pfprifin*, *pfstevor*, and *pfsurfin*) were removed from the final network.

#### Parasite culture

For the experiments testing for susceptibility to compounds (KAF156, MMV907364, and MMV007564) that select for PfCARL mutations (Fig. 4A, fig. S9, and tables S2 to S4), ABS parasites, including laboratory lines and culture-adapted field isolates, were cultured at 2% HCT in human O+ or A+ RBCs in RPMI-1640 media, supplemented with 25 mM HEPES (Fisher), 50 mg/liter hypoxanthine (Sigma Aldrich), 2 mM L-glutamine (Cambridge Isotope Laboratories, Inc.), 0.21% sodium bicarbonate (Sigma Aldrich), 0.5% (w/v) AlbuMAX II (Invitrogen), 8% filtered, heat-inactivated, pooled off-the-clot AB+ human serum (Innovative Research, Inc.), and 10 µg/ml gentamicin (Fisher). Cultures were propagated in tissue culture flasks gassed with a mixture of 5% O<sub>2</sub>, 5% CO<sub>2</sub> and 90% N<sub>2</sub> and maintained at 37°C.

For the experiments phenotyping PfMDR1 mutant lines (Fig. 4B, fig. S10 to S11, and tables S5 to S7), clonal parasite lines were thawed from frozen stocks and put into culture where

they were maintained at 3% HCT in human O+ RBCs in RPMI-1640 media as listed above. Once parasites attained 3% parasitemia, infected blood pellet was collected, and DNA was extracted for polymerase chain reaction (PCR) and sequencing of the *pfmdr1* loci of interest to ascertain presence of the mutations at the correct codon position.

For the experiments testing for susceptibility to compounds (KAE609, SJ733, MMV665826, MMV020660, MMV011567, GNF Pf-3703) that select for PfATP4 mutations (Fig. 4C, fig. S12, and tables S8 to S10), ABS parasites, including laboratory lines and culture-adapted field isolates, were cultured at 5% HCT in RPMI-1640 media as listed above, with a lower O<sub>2</sub> gas mixture (1% O<sub>2</sub>, 4.1% CO<sub>2</sub>, and 94.9% N<sub>2</sub>).

#### Targeted *pfmdr1* sequencing and analysis of mutations

The *pfmdr1* gene was PCR-amplified from parasite genomic DNA using primers flanking the edited locus in PF3D7\_0523000. Genomic DNAs of parental strains were used as a control. The PCR conditions to amplify a 812-bp fragment using primer pair p9160+p9161 encompassing the Ala750Thr SNV in NF10, Ser784Leu SNV in Cam3.II, or Met841Ile+Met924Ile SNVs in NF54 were: 95°C for 3 min, 35 rounds of 95°C for 30 s, 58°C for 50 s, and 62°C for 1 min, with a final extension of 3 min at 68°C. PCR conditions to amplify a 1130-bp fragment encompassing the Phe1072Leu and Ser1075Ile SNVs in Dd2-B2 using primer pair p7823+p7923 and KAPA HiFi HotStart ReadyMix (2X) (Roche) were: 95°C for 3 min, 35 rounds of 95°C for 30 s, 52°C for 1 min, and 62°C for 1 min, with a final extension of 3 min at 68°C. Sequences of the primers are given in table S6.

#### Compound susceptibility assays for lines with *pfcarl* or *pfmdr1* mutations

To define the 50% (IC<sub>50</sub>) and 90% (IC<sub>90</sub>) growth-inhibitory KAF156, MMV907364, and MMV007564 concentration values for parasites, asynchronous ABS cultures at 0.2 to 0.8% parasitemia and 1% HCT in 0.5% (w/v) AlbuMAX II/8% serum-containing media were exposed for 72 hours to a range of 10 compound concentrations that were twofold (MMV007564) or threefold (KAF156 and MMV907364) serially diluted in duplicates along with compound-free controls. For the lines with edited *pfmdr1* mutations, asynchronous parasites at 0.2% parasitemia and 1% HCT in 0.5% (w/v) AlbuMAX II-containing media were exposed for 72 hours to a range of 10 compound concentrations that were twofold serially diluted in duplicates along with compound-free controls. Parasite survival was assessed by flow cytometry on an IntelliCyt iQue Screener PLUS (Sartorius) or a BD FACSCelesta (BD Biosciences) using 1× SYBR Green (Invitrogen) and 200 nM MitoTracker Deep Red FM (Life Technologies) as nuclear stain and vital dyes, re-

spectively. IC<sub>50</sub> and IC<sub>90</sub> values were derived from growth inhibition data using nonlinear regression (Prism 10, GraphPad) or linear interpolation (assays for lines with *pfcarl* mutations) as means ± SEM from four or five independent biological repeats with two technical replicates. Statistical significance of IC<sub>50</sub> and IC<sub>90</sub> shifts compared with wild-type reference line (3D7-A10 for *pfcarl*) or isogenic parents (*pfmdr1* edited lines) was determined using two-tailed Mann-Whitney *U* tests (Prism 10, GraphPad).

#### Compound susceptibility assays for lines with *pfatp4* mutations

Compound IC<sub>50</sub> was assessed in dose-response format including using a SYBR Green-I based cell-proliferation assay as previously described (88). In brief, cultures synchronized by sorbitol at 1% ring-stage parasitemia and 1% HCT in 0.5% (w/v) AlbuMAX II-supplemented media (described above) were incubated for 72 hours in 384-well plates with exposure to the compound of interest in a 12-point dilution series with technical triplicates. A fixed concentration of dihydroartemisinin was used as a positive control. Compounds were dispensed with an HP D300 Digital Dispenser (Hewlett Packard, Palo Alto, CA, USA). Following the incubation, parasites were lysed, and DNA was stained using SYBR Green fluorescence and was measured at 535 nm on a SpectraMax iD5 plate reader (Molecular Devices, San Jose, CA) after excitation at 485 nm. Data were archived and analyzed using the CDD Vault from Collaborative Drug Discovery (Burlingame, CA; [www.collaborativedrug.com](http://www.collaborativedrug.com)). IC<sub>50</sub> fold changes were calculated by comparing IC<sub>50</sub> of the *pfatp4* mutant isolates to that of the compound-sensitive Dd2-2D4 parasite line.

#### Whole-genome sequencing of lines phenotyped against *pfcarl* or *pfatp4* mutation-selecting compounds and variant analysis

Parasite genomic DNA was extracted upon completion of the dose-response assays. Cultures at 5 to 10% parasitemia were initially lysed with 0.15% saponin and washed twice with 1× phosphate-buffered saline. Genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen). WGS was performed using a Nextera Flex DNA library kit and multiplexed on a MiSeq flow cell to generate 300 bp paired-end reads. Sequences were aligned to the Pf3D7 reference genome (PlasmoDB, version 48) using the Burrows-Wheeler Aligner (BWA). PCR duplicates and unmapped reads were filtered out using Samtools and Picard. Base quality scores were recalibrated using GATK BaseRecalibrator. GATK Haplotype Caller (version 4.1.8) was used to identify all possible single nucleotide variants (SNVs) in test parasite lines filtered on the basis of quality scores (variant quality as function of depth QD > 1.5, mapping quality > 40, min base

quality score > 18, read depth > 5) to obtain high quality SNVs that were annotated using SnpEff. Comparative SNV analyses between the laboratory or field isolate genomes and the reference Pf3D7 genome were performed to generate the final list of SNVs. BIC-Seq was used to discover CNVs against the reference strain using the Bayesian statistical model. SNVs and CNVs were visually inspected and confirmed using Integrative Genome Viewer (IGV). Low-frequency SNVs (<70% alternate allele frequency) were quantified by visual inspection of the reads covering individual loci using IGV. Alternate allele frequencies were derived from observed alternate reads / sum of alternate and reference reads.

#### PfZIP1 conditional knockout

For the pZIP1-cKO repair plasmid, a recodonized sequence encompassing amino acids 242-358 (C-term) of PfZIP1 and a barcode cassette (100 bp) were ordered as gBlocks (IDT). The 5' homology region (540 bp) and 3' homology region (496 bp) were PCR-amplified from parasite genomic DNA. The first loxP intron and a sequence containing a 3xV5 tag, 2A skip peptide, neomycin resistance cassette, and second loxP intron were PCR-amplified from a previously used plasmid (92). Fragments were inserted by Gibson assembly into a pUC19 vector with an ampicillin resistance cassette. The guide RNA sequence (TGTCCTAACACTATACCCAG) was inserted into a double BbsI site within the pDC2-coCas9-gRNA plasmid (93). The repair plasmid (60 µg), the guide RNA plasmid (30 µg), and 100 µl packed NF54 DiCre (94) parasites were suspended in cytomix (120 mM KCl, 0.15 mM CaCl<sub>2</sub>, 2 mM EGTA, 5 mM MgCl<sub>2</sub>, 10 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, 25 mM HEPES, pH 7.6) and transfected using a Biorad GenePulser II electroporator (310 V, 950 µF, 200 Ω) in 2 mm electroporation cuvettes (Biorad). After 24 hours the transfected parasite lines were treated with 2.5 nM WR99210 for 5 days to select for the Cas9 plasmid. Once parasites recovered these were treated for 2 weeks with 225 µg/ml G418 to select for correct integration into the PfZIP1 locus. Parasites were cloned by limiting dilution, and correct integration and rapamycin-induced excision verified by PCR.

#### Pfmdr1 gene editing

The *pfmdr1* Ala750Thr and Ser784Leu mutations were introduced into NF10 and Cam3. II<sup>C580Y</sup> parasites, respectively, using a T7-based CRISPR-Cas9 two-plasmid system described previously (55). Site-directed mutagenesis using primer pairs p6985+p6986 and p6987+p6988 was used to reverse Met841Ile+Met924Ile mutations to the wild-type residues in the pT7pol-donor-bsd plasmid harboring the Met841Ile+Met924Ile *pfmdr1* donor sequence, respectively. A second round of site-directed mutagenesis (SDM) on this donor plasmid using primer

pairs p6989+6990 or p6916+p6917 resulted in two donor plasmids harboring mutations Ser784Leu or Ala750Thr, respectively. Donor plasmids were coelectroporated with the pCas9-gRNA2-hours*dhfr* plasmid into parasites that were cultured in the presence of 2.5 nM WR99210 and 2.0 µg/ml blasticidin. *pfmdr1*-edited parasites were identified by PCR and Sanger sequencing and cloned by limiting dilution. All primers used are listed in table S6.

## REFERENCES AND NOTES

- M. M. Abdel Hamid *et al.*, Pf7: an open dataset of *Plasmodium falciparum* genome variation in 20,000 worldwide samples. *Wellcome Open Res.* **8**, 22 (2023). doi: [10.12688/wellcomeopenres.18681.1](https://doi.org/10.12688/wellcomeopenres.18681.1); PMID: 36864926
- F. Rocamora, E. A. Wenzler, Genomic approaches to drug resistance in malaria. *Annu. Rev. Microbiol.* **74**, 761–786 (2020). doi: [10.1146/annurev-micro-012220-064343](https://doi.org/10.1146/annurev-micro-012220-064343); PMID: 32905749
- S. Takala-Harrison *et al.*, Genetic loci associated with delayed clearance of *Plasmodium falciparum* following artemisinin treatment in Southeast Asia. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 240–245 (2013). doi: [10.1073/pnas.1211205110](https://doi.org/10.1073/pnas.1211205110); PMID: 23248304
- F. Ariey *et al.*, A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* **505**, 50–55 (2014). doi: [10.1038/nature12876](https://doi.org/10.1038/nature12876); PMID: 24352242
- T. Yang *et al.*, MalDA, accelerating malaria drug discovery. *Trends Parasitol.* **37**, 493–507 (2021). doi: [10.1016/j.pt.2021.01.009](https://doi.org/10.1016/j.pt.2021.01.009); PMID: 33648890
- K. A. Button-Simons *et al.*, The power and promise of genetic mapping from *Plasmodium falciparum* crosses utilizing human liver-chimeric mice. *Commun. Biol.* **4**, 734 (2021). doi: [10.1038/s42003-021-02210-1](https://doi.org/10.1038/s42003-021-02210-1); PMID: 34127785
- C. Timmann *et al.*, Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature* **489**, 443–446 (2012). doi: [10.1038/nature11334](https://doi.org/10.1038/nature11334); PMID: 22895189
- L. E. Cowen *et al.*, Evolution of drug resistance in experimental populations of *Candida albicans*. *J. Bacteriol.* **182**, 1515–1522 (2000). doi: [10.1128/JB.182.6.1515-1522.2000](https://doi.org/10.1128/JB.182.6.1515-1522.2000); PMID: 10692355
- T. M. Smith *et al.*, Rapid adaptation of a complex trait during experimental evolution of *Mycobacterium tuberculosis*. *eLife* **11**, e78454 (2022). doi: [10.7554/eLife.78454](https://doi.org/10.7554/eLife.78454); PMID: 35726854
- S. Ottilie *et al.*, Adaptive laboratory evolution in *S. cerevisiae* highlights role of transcription factors in fungal xenobiotic resistance. *Commun. Biol.* **5**, 128 (2022). doi: [10.1038/s42003-022-03076-7](https://doi.org/10.1038/s42003-022-03076-7); PMID: 35149760
- A. Miles *et al.*, Indels, structural variation, and recombination drive genomic diversity in *Plasmodium falciparum*. *Genome Res.* **26**, 1288–1299 (2016). doi: [10.1101/gr.203711.115](https://doi.org/10.1101/gr.203711.115); PMID: 27531718
- J. L. Siqueira-Neto *et al.*, Antimalarial drug discovery: Progress and approaches. *Nat. Rev. Drug Discov.* **22**, 807–826 (2023). doi: [10.1038/s41573-023-00772-9](https://doi.org/10.1038/s41573-023-00772-9); PMID: 37652975
- G. A. Van der Auwera *et al.*, From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33 (2013). doi: [10.1002/0471250953.b1110s43](https://doi.org/10.1002/0471250953.b1110s43); PMID: 25431634
- P. Cingolani *et al.*, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012). doi: [10.4161/fly.19695](https://doi.org/10.4161/fly.19695); PMID: 22728672
- A. N. Cowell *et al.*, Mapping the malaria parasite druggable genome by using *in vitro* evolution and chemogenomics. *Science* **359**, 191–199 (2018). doi: [10.1126/science.aan4472](https://doi.org/10.1126/science.aan4472); PMID: 29326268
- F. Rocamora *et al.*, Oxidative stress and protein damage responses mediate artemisinin resistance in malaria parasites. *PLoS Pathog.* **14**, e1006930 (2018). doi: [10.1371/journal.ppat.1006930](https://doi.org/10.1371/journal.ppat.1006930); PMID: 29538461
- C. Kidgell *et al.*, A systematic map of genetic variation in *Plasmodium falciparum*. *PLoS Pathog.* **2**, e57 (2006). doi: [10.1371/journal.ppat.0020057](https://doi.org/10.1371/journal.ppat.0020057); PMID: 16789840
- T. Rausch *et al.*, DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012). doi: [10.1093/bioinformatics/bts378](https://doi.org/10.1093/bioinformatics/bts378); PMID: 22962449
- A. C. Huckaby *et al.*, Complex DNA structures trigger copy number variation across the *Plasmodium falciparum* genome. *Nucleic Acids Res.* **47**, 1615–1627 (2019). doi: [10.1093/nar/gky1268](https://doi.org/10.1093/nar/gky1268); PMID: 30576466
- S. H. Shafiq, S. N. Richards, B. Corry, R. E. Martin, Mechanistic basis for multidrug resistance and collateral drug sensitivity conferred to the malaria parasite by polymorphisms in PfMDR1 and PfCRT. *PLoS Biol.* **20**, e3001616 (2022). doi: [10.1371/journal.pbio.3001616](https://doi.org/10.1371/journal.pbio.3001616); PMID: 35507548
- G. LaMonte *et al.*, Mutations in the *Plasmodium falciparum* cyclic amine resistance locus (PfCARL) confer multidrug resistance. *mBio* **7**, e00696-16 (2016). doi: [10.1128/mBio.00696-16](https://doi.org/10.1128/mBio.00696-16); PMID: 27381290
- M. Rottmann *et al.*, Spiroindolones, a potent compound class for the treatment of malaria. *Science* **329**, 1175–1180 (2010). doi: [10.1126/science.1193225](https://doi.org/10.1126/science.1193225); PMID: 20813948
- C. W. McNamara *et al.*, Targeting *Plasmodium* Pf(4)K to eliminate malaria. *Nature* **504**, 248–253 (2013). doi: [10.1038/nature12782](https://doi.org/10.1038/nature12782); PMID: 24284631
- M. Korsinczyk *et al.*, Mutations in *Plasmodium falciparum* cytochrome b that are associated with atovaquone resistance are located at a putative drug-binding site. *Antimicrob. Agents Chemother.* **44**, 2100–2108 (2000). doi: [10.1128/AAC.44.8.2100-2108.2000](https://doi.org/10.1128/AAC.44.8.2100-2108.2000); PMID: 10898682
- J. L. Guler *et al.*, Asexual populations of the human malaria parasite, *Plasmodium falciparum*, use a two-step genomic strategy to acquire accurate, beneficial DNA amplifications. *PLoS Pathog.* **9**, e1003375 (2013). doi: [10.1371/journal.ppat.1003375](https://doi.org/10.1371/journal.ppat.1003375); PMID: 23717205
- G. A. Josling *et al.*, Dissecting the role of PfATP2-G in malaria gametocytogenesis. *Nat. Commun.* **11**, 1503 (2020). doi: [10.1038/s41467-020-15026-0](https://doi.org/10.1038/s41467-020-15026-0); PMID: 32198457
- A. Claessens, M. Affara, S. A. Assefa, D. P. Kwiatkowski, D. J. Conway, Culture adaptation of malaria parasites selects for convergent loss-of-function mutants. *Sci. Rep.* **7**, 41303 (2017). doi: [10.1038/srep41303](https://doi.org/10.1038/srep41303); PMID: 28117431
- K. Gerth, N. Bedorf, G. Höfle, H. Irschik, H. Reichenbach, Epothilons A and B: Antifungal and cytotoxic compounds from *Sorangium cellulosum* (Myxobacteria). Production, physico-chemical and biological properties. *J. Antibiot. (Tokyo)* **49**, 560–563 (1996). doi: [10.7164/antibiotics.49.560](https://doi.org/10.7164/antibiotics.49.560); PMID: 8698639
- P. F. Muhlradt, F. Sasse, Epothilone B stabilizes microtubuli of macrophages like taxol without showing taxol-like endotoxin activity. *Cancer Res.* **57**, 3344–3346 (1997). doi: [10.1158/0008-5472.CCR-97-0292](https://doi.org/10.1158/0008-5472.CCR-97-0292); PMID: 9269992
- J. G. De Man, T. G. Moreels, B. Y. De Winter, A. G. Herman, P. A. Pelckmans, Neocuproine potentiates the activity of the nitric neurotransmitter but inhibits that of S-nitrosothiols. *Eur. J. Pharmacol.* **381**, 151–159 (1999). doi: [10.1016/S0014-2999\(99\)00564-6](https://doi.org/10.1016/S0014-2999(99)00564-6); PMID: 10554883
- H. Ke *et al.*, Mitochondrial type II NADH dehydrogenase of *Plasmodium falciparum* (PfNDH2) is dispensable in the asexual blood stages. *PLoS ONE* **14**, e0214023 (2019). doi: [10.1371/journal.pone.0214023](https://doi.org/10.1371/journal.pone.0214023); PMID: 30964863
- Y. Yang *et al.*, Target elucidation by cocrystal structures of NADH-ubiquinone oxidoreductase of *Plasmodium falciparum* (PfNDH2) with small molecule to eliminate drug-resistant malaria. *J. Med. Chem.* **60**, 1994–2005 (2017). doi: [10.1021/acs.jmedchem.6b01733](https://doi.org/10.1021/acs.jmedchem.6b01733); PMID: 28195463
- S. Duffy *et al.*, Screening the Medicines for Malaria Venture Pathogen Box across multiple pathogens reclassifies starting points for open-source drug discovery. *Antimicrob. Agents Chemother.* **61**, e00379-17 (2017). doi: [10.1128/AAC.00379-17](https://doi.org/10.1128/AAC.00379-17); PMID: 28674055
- M. Zhang *et al.*, Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* **360**, eaap7847 (2018). doi: [10.1126/science.aap7847](https://doi.org/10.1126/science.aap7847); PMID: 29724925
- V. M. Howick *et al.*, The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* **365**, eaaw2619 (2019). doi: [10.1126/science.aaw2619](https://doi.org/10.1126/science.aaw2619); PMID: 31439762
- C. Hou *et al.*, DUF221 proteins are a family of osmosensitive calcium-permeable cation channels conserved across eukaryotes. *Cell Res.* **24**, 632–635 (2014). doi: [10.1038/cr.2014.14](https://doi.org/10.1038/cr.2014.14); PMID: 24503647
- L. E. de Vries, M. Lunghi, A. Krishnan, T. W. A. Kooij, D. Soldati-Favre, Pantothenate and CoA biosynthesis in Apicomplexa and their promise as antiparasitic drug targets. *PLoS Pathog.* **17**, e1010124 (2021). doi: [10.1371/journal.ppat.1010124](https://doi.org/10.1371/journal.ppat.1010124); PMID: 34969059
- M. D. Jeninga, J. E. Quinn, M. Petter, ApiAP2 transcription factors in apicomplexan parasites. *Pathogens* **8**, 47 (2019). doi: [10.3390/pathogens8020047](https://doi.org/10.3390/pathogens8020047); PMID: 30959972
- A. Ahouidi *et al.*, An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples. *Wellcome Open Res.* **6**, 42 (2021). doi: [10.12688/wellcomeopenres.16168.1](https://doi.org/10.12688/wellcomeopenres.16168.1); PMID: 33824913
- H. G. Lê *et al.*, Changing pattern of the genetic diversities of *Plasmodium falciparum* merozoite surface protein-1 and merozoite surface protein-2 in Myanmar isolates. *Malar. J.* **18**, 241 (2019). doi: [10.1186/s12936-019-2879-7](https://doi.org/10.1186/s12936-019-2879-7); PMID: 31311565
- J. M. Kang *et al.*, Genetic polymorphism and natural selection of apical membrane antigen-1 in *Plasmodium falciparum* isolates from Vietnam. *Genes* **12**, 1903 (2021). doi: [10.3390/genes12121903](https://doi.org/10.3390/genes12121903); PMID: 34946853
- E. L. Flannery *et al.*, Mutations in the P-type cation-transporter ATPase 4, PfATP4, mediate resistance to both aminopyrazole and spiroindolone antimalarials. *ACS Chem. Biol.* **10**, 413–420 (2015). doi: [10.1021/cb500616x](https://doi.org/10.1021/cb500616x); PMID: 25322084
- A. Waterhouse *et al.*, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018). doi: [10.1093/nar/gky427](https://doi.org/10.1093/nar/gky427); PMID: 29788355
- M. Varadi *et al.*, AlphaFold Protein Structure Database: Massively expanded the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022). doi: [10.1093/nar/gkab1061](https://doi.org/10.1093/nar/gkab1061); PMID: 34791371
- K. Si *et al.*, The structure of *Plasmodium falciparum* multidrug resistance protein 1 reveals an N-terminal regulatory domain. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2219905120 (2023). doi: [10.1073/pnas.2219905120](https://doi.org/10.1073/pnas.2219905120); PMID: 37527341
- K. Nosol *et al.*, Cryo-EM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 26245–26253 (2020). doi: [10.1073/pnas.2010264117](https://doi.org/10.1073/pnas.2010264117); PMID: 33020312
- M. Takahashi, Y. Kondou, C. Toyoshima, Interdomain communication in calcium pump as revealed in the crystal structures with transmembrane inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5800–5805 (2007). doi: [10.1073/pnas.0700979104](https://doi.org/10.1073/pnas.0700979104); PMID: 17389383
- B. Gurung, L. Yu, C. A. Yu, Stigmatellin induces reduction of iron-sulfur protein in the oxidized cytochrome *bc<sub>1</sub>* complex. *J. Biol. Chem.* **283**, 28087–28094 (2008). doi: [10.1074/jbc.M804229200](https://doi.org/10.1074/jbc.M804229200); PMID: 18701458
- M. I. Veiga *et al.*, Globally prevalent PfMDR1 mutations modulate *Plasmodium falciparum* susceptibility to artemisinin-based combination therapies. *Nat. Commun.* **7**, 11553 (2016). doi: [10.1038/ncomms11553](https://doi.org/10.1038/ncomms11553); PMID: 27189525
- A. B. Sidhu, S. G. Valderramos, D. A. Fidock, *pfmdr1* mutations contribute to quinone resistance and enhance mefloquine and artemisinin sensitivity in *Plasmodium falciparum*. *Mol. Microbiol.* **57**, 913–926 (2005). doi: [10.1111/j.1365-2958.2005.04729.x](https://doi.org/10.1111/j.1365-2958.2005.04729.x); PMID: 16091034
- O. Kretzfeldt *et al.*, Associations between varied susceptibilities to PfATP4 inhibitors and genotypes in Ugandan *Plasmodium falciparum* isolates. *Antimicrob. Agents Chemother.* **65**, e0077121 (2021). doi: [10.1128/AAC.00771-21](https://doi.org/10.1128/AAC.00771-21); PMID: 34339273
- S. Meister *et al.*, Imaging of *Plasmodium* liver stages to drive next-generation antimalarial drug discovery. *Science* **334**, 1372–1377 (2011). doi: [10.1126/science.1211936](https://doi.org/10.1126/science.1211936); PMID: 22096101
- P. A. Magistrado *et al.*, *Plasmodium falciparum* cyclic amine resistance locus (PfCARL), a resistance mechanism for two distinct compound classes. *ACS Infect. Dis.* **2**, 816–826 (2016). doi: [10.1021/acsinfecdis.6b00025](https://doi.org/10.1021/acsinfecdis.6b00025); PMID: 27933786
- M. I. Veiga *et al.*, Novel polymorphisms in *Plasmodium falciparum* ABC transporter genes are associated with major ACT antimalarial drug resistance. *PLoS ONE* **6**, e20212 (2011). doi: [10.1371/journal.pone.0020212](https://doi.org/10.1371/journal.pone.0020212); PMID: 21633513
- C. L. Ng *et al.*, 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**, 381–393 (2016). doi: [10.1111/mmi.13397](https://doi.org/10.1111/mmi.13397); PMID: 27073104
- M. Vanarschot *et al.*, Hexahydroquinolines are antimalarial candidates with potent blood-stage and transmission-blocking activity. *Nat. Microbiol.* **2**, 1403–1414 (2017). doi: [10.1038/s41564-017-0007-4](https://doi.org/10.1038/s41564-017-0007-4); PMID: 28808258
- A. H. Lee, D. A. Fidock, Evidence of a mild mutator phenotype in Cambodian *Plasmodium falciparum* malaria parasites. *PLoS ONE* **11**, e0154166 (2016). doi: [10.1371/journal.pone.0154166](https://doi.org/10.1371/journal.pone.0154166); PMID: 27100094

58. S. F. Schaffner *et al.*, Malaria surveillance reveals parasite relatedness, signatures of selection, and correlates of transmission across Senegal. *Nat. Commun.* **14**, 7268 (2023). doi: [10.1038/s41467-023-43087-4](https://doi.org/10.1038/s41467-023-43087-4); pmid: 37949851
59. A. Amambua-Ngwa *et al.*, Major subpopulations of *Plasmodium falciparum* in sub-Saharan Africa. *Science* **365**, 813–816 (2019). doi: [10.1126/science.aav5427](https://doi.org/10.1126/science.aav5427); pmid: 31439796
60. Z. Xie, T. M. Nolan, H. Jiang, Y. Yin, AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in *Arabidopsis*. *Front. Plant Sci.* **10**, 228 (2019). doi: [10.3389/fpls.2019.00228](https://doi.org/10.3389/fpls.2019.00228); pmid: 30873200
61. Y. Zhou *et al.*, DHODH and cancer: Promising prospects to be explored. *Cancer Metab.* **9**, 22 (2021). doi: [10.1186/s40170-021-00250-z](https://doi.org/10.1186/s40170-021-00250-z); pmid: 33971967
62. D. Hoepfner *et al.*, Selective and specific inhibition of the *plasmodium falciparum* lysyl-tRNA synthetase by the fungal secondary metabolite cladospirin. *Cell Host Microbe* **11**, 654–663 (2012). doi: [10.1016/j.chom.2012.04.015](https://doi.org/10.1016/j.chom.2012.04.015); pmid: 22704625
63. G. Rix *et al.*, Scalable continuous evolution for the generation of diverse enzyme variants encompassing promiscuous activities. *Nat. Commun.* **11**, 5644 (2020). doi: [10.1038/s41467-020-19539-6](https://doi.org/10.1038/s41467-020-19539-6); pmid: 33159067
64. M. Duffey *et al.*, Assessing risks of *Plasmodium falciparum* resistance to select next-generation antimalarials. *Trends Parasitol.* **37**, 709–721 (2021). doi: [10.1016/j.pt.2021.04.006](https://doi.org/10.1016/j.pt.2021.04.006); pmid: 34001441
65. K. M. V. Sindhe *et al.*, *Plasmodium falciparum* resistance to a lead benzoxaborole due to blocked compound activation and altered ubiquitination or sumoylation. *mBio* **11**, e02640-19 (2020). doi: [10.1128/mBio.02640-19](https://doi.org/10.1128/mBio.02640-19); pmid: 31992618
66. R. E. K. Mandt *et al.*, In vitro selection predicts malaria parasite resistance to dihydroorotate dehydrogenase inhibitors in a mouse infection model. *Sci. Transl. Med.* **11**, eaav1636 (2019). doi: [10.1126/scitranslmed.aav1636](https://doi.org/10.1126/scitranslmed.aav1636); pmid: 31801884
67. J. D. Herman *et al.*, A genomic and evolutionary approach reveals non-genetic drug resistance in malaria. *Genome Biol.* **15**, 511 (2014). doi: [10.1186/s13059-014-0511-2](https://doi.org/10.1186/s13059-014-0511-2); pmid: 25395010
68. S. C. Xie *et al.*, Target validation and identification of novel boronate inhibitors of the *Plasmodium falciparum* proteasome. *J. Med. Chem.* **61**, 10053–10066 (2018). doi: [10.1021/acs.jmedchem.8b01161](https://doi.org/10.1021/acs.jmedchem.8b01161); pmid: 30373366
69. J. Schalkwijk *et al.*, Antimalarial pantothenamide metabolites target acetyl-coenzyme A biosynthesis in *Plasmodium falciparum*. *Sci. Transl. Med.* **11**, eaas9917 (2019). doi: [10.1126/scitranslmed.aas9917](https://doi.org/10.1126/scitranslmed.aas9917); pmid: 31534021
70. M. B. Jiménez-Díaz *et al.*, (+)-SJ733, a clinical candidate for malaria that acts through ATP4 to induce rapid host-mediated clearance of *Plasmodium*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E5455–E5462 (2014). doi: [10.1073/pnas.1414221111](https://doi.org/10.1073/pnas.1414221111); pmid: 25453091
71. B. Baragaña *et al.*, A novel multiple-stage antimalarial agent that inhibits protein synthesis. *Nature* **522**, 315–320 (2015). doi: [10.1038/nature14451](https://doi.org/10.1038/nature14451); pmid: 26085270
72. K. Tahlan *et al.*, SQ109 targets MmpL3, a membrane transporter of trehalose monomycolate involved in mycolic acid donation to the cell wall core of *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **56**, 1797–1809 (2012). doi: [10.1128/AAC.05708-11](https://doi.org/10.1128/AAC.05708-11); pmid: 22252828
73. B. H. Stokes *et al.*, Covalent *Plasmodium falciparum*-selective proteasome inhibitors exhibit a low propensity for generating resistance *in vitro* and synergize with multiple antimalarial agents. *PLOS Pathog.* **15**, e1007722 (2019). doi: [10.1371/journal.ppat.1007722](https://doi.org/10.1371/journal.ppat.1007722); pmid: 31170268
74. G. M. Fisher *et al.*, The key glycolytic enzyme phosphofructokinase is involved in resistance to antiparasitoid glycosides. *mBio* **11**, e02842-20 (2020). doi: [10.1128/mBio.02842-20](https://doi.org/10.1128/mBio.02842-20); pmid: 33293381
75. K. M. Ellis *et al.*, The novel bis-1,2,4-triazine MIPS-0004373 demonstrates rapid and potent activity against all blood stages of the malaria parasite. *Antimicrob. Agents Chemother.* **65**, e0031121 (2021). doi: [10.1128/AAC.00311-21](https://doi.org/10.1128/AAC.00311-21); pmid: 34460304
76. J. E. Gisselberg, Z. Herrera, L. M. Orchard, M. Llinás, E. Yeh, Specific inhibition of the bifunctional farnesyl/geranylgeranyl diphosphate synthase in malaria parasites via a new small-molecule binding site. *Cell Chem. Biol.* **25**, 185–193.e5 (2018). doi: [10.1016/j.chembiol.2017.11.010](https://doi.org/10.1016/j.chembiol.2017.11.010); pmid: 29276048
77. M. Vanaerschoot *et al.*, Inhibition of resistance-refractory *P. falciparum* kinase PKG delivers prophylactic, blood stage, and transmission-blocking antiparasitoid activity. *Cell Chem. Biol.* **27**, 806–816.e8 (2020). doi: [10.1016/j.chembiol.2020.04.001](https://doi.org/10.1016/j.chembiol.2020.04.001); pmid: 32359426
78. T. C. Knaab *et al.*, 3-Hydroxy-propanamides, a new class of orally active antimalarials targeting *Plasmodium falciparum*. *J. Med. Chem.* **64**, 3035–3047 (2021). doi: [10.1021/acs.jmedchem.0c01744](https://doi.org/10.1021/acs.jmedchem.0c01744); pmid: 33666415
79. D. A. Fidock *et al.*, Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol. Cell* **6**, 861–871 (2000). doi: [10.1016/S1097-2765\(05\)00077-8](https://doi.org/10.1016/S1097-2765(05)00077-8); pmid: 11090624
80. T. E. Wellemes *et al.*, Chloroquine resistance not linked to *mdr*-like genes in a *Plasmodium falciparum* cross. *Nature* **345**, 253–255 (1990). doi: [10.1038/345253a0](https://doi.org/10.1038/345253a0); pmid: 1970614
81. M. A. Tye *et al.*, Elucidating the path to *Plasmodium* prolyl-tRNA synthetase inhibitors that overcome halofuginone resistance. *Nat. Commun.* **13**, 4976 (2022). doi: [10.1038/s41467-022-32630-4](https://doi.org/10.1038/s41467-022-32630-4); pmid: 36008486
82. K. Kumpornsin *et al.*, Generation of a mutator parasite to drive resistome discovery in *Plasmodium falciparum*. *Nat. Commun.* **14**, 3059 (2023). doi: [10.1038/s41467-023-38774-1](https://doi.org/10.1038/s41467-023-38774-1); pmid: 37244916
83. S. Pelleau *et al.*, Adaptive evolution of malaria parasites in French Guiana: Reversal of chloroquine resistance by acquisition of a mutation in *pfprt*. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11672–11677 (2015). doi: [10.1073/pnas.1507142112](https://doi.org/10.1073/pnas.1507142112); pmid: 26261345
84. S. G. Valdeiramos *et al.*, Identification of a mutant PfCRT-mediated chloroquine tolerance phenotype in *Plasmodium falciparum*. *PLOS Pathog.* **6**, e1000887 (2010). doi: [10.1371/journal.ppat.1000887](https://doi.org/10.1371/journal.ppat.1000887); pmid: 20485514
85. D. S. Peterson, W. K. Milhous, T. E. Wellemes, Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3018–3022 (1990). doi: [10.1073/pnas.87.8.3018](https://doi.org/10.1073/pnas.87.8.3018); pmid: 2183222
86. P. K. Tumwebaze *et al.*, Drug susceptibility of *Plasmodium falciparum* in eastern Uganda: A longitudinal phenotypic and genotypic study. *Lancet Microbe* **2**, e441–e449 (2021). doi: [10.1016/S2666-5247\(21\)00085-9](https://doi.org/10.1016/S2666-5247(21)00085-9); pmid: 34553183
87. I. D. Goodyer, T. F. Taraschi, *Plasmodium falciparum*: A simple, rapid method for detecting parasite clones in microtiter plates. *Exp. Parasitol.* **86**, 158–160 (1997). doi: [10.1006/expr.1997.4156](https://doi.org/10.1006/expr.1997.4156); pmid: 9207746
88. D. Plouffe *et al.*, In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9059–9064 (2008). doi: [10.1073/pnas.0802982105](https://doi.org/10.1073/pnas.0802982105); pmid: 18579783
89. L. Chappell *et al.*, Refining the transcriptome of the human malaria parasite *Plasmodium falciparum* using amplification-free RNA-seq. *BMC Genomics* **21**, 395 (2020). doi: [10.1186/s12864-020-06787-5](https://doi.org/10.1186/s12864-020-06787-5); pmid: 32513207
90. N. G. Metwally *et al.*, Characterisation of *Plasmodium falciparum* populations selected on the human endothelial receptors P-selectin, E-selectin, CD9 and CD151. *Sci. Rep.* **7**, 4069 (2017). doi: [10.1038/s41598-017-04241-3](https://doi.org/10.1038/s41598-017-04241-3); pmid: 28642573
91. P. Shannon *et al.*, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003). doi: [10.1101/gr.1239303](https://doi.org/10.1101/gr.1239303); pmid: 14597658
92. H. Davies *et al.*, An exported kinase family mediates species-specific erythrocyte remodelling and virulence in human malaria. *Nat. Microbiol.* **5**, 848–863 (2020). doi: [10.1038/s41564-020-0702-4](https://doi.org/10.1038/s41564-020-0702-4); pmid: 32284562
93. S. Adjalley, M. C. S. Lee, CRISPR/Cas9 editing of the *Plasmodium falciparum* genome. *Methods Mol. Biol.* **2470**, 221–239 (2022). doi: [10.1007/978-1-0716-2189-9\\_17](https://doi.org/10.1007/978-1-0716-2189-9_17); pmid: 35881349
94. M. Tibúrcio *et al.*, A novel tool for the generation of conditional knockouts to study gene function across the *Plasmodium falciparum* life cycle. *mBio* **10**, e01170-19 (2019). doi: [10.1128/mBio.01170-19](https://doi.org/10.1128/mBio.01170-19); pmid: 31530668
95. D. Chen, Justawayx/malaria\_resistome\_2023: P. falciparum in vitro evolved mutations meta-analysis, version 1, Zenodo (2024); <https://doi.org/10.5281/zenodo.13766513>
- PFATP4 (entry AOA143ZZK9), PfCARL (entry COH483), and PfCYTB (entry Q7HP03). The authors thank Nosol *et al.* for making available the structure of reconstituted human ABCB1 in complex with vincristine (PDB ID: 7a69) used for PfMDR1 homology model. All data in this study are free to access and use. **Funding:** E.A.W. is supported by the Target Discovery for Antimalarials grant from the Bill and Melinda Gates Foundation (OPP1054480), M.C.S.L. is supported by the CRISPR platform for MaIDA target validation and drug discovery grant from the Bill and Melinda Gates Foundation (INV-045906), M.R.L. was supported in part by a Ruth L. Kirschstein Institutional National Research Award from the National Institutes for General Medical Sciences (T32 GM008666). The National Institutes of Health supported this work with grants to D.C. (T32 GM139790), E.A.W. (R01 AI169892, R01 AI154777, R01 AI172066, and R01 AI152533), and D.A.F. (R01 AI109023, R37 AI50234, R01 AI124678, and R01 AI185559). This publication includes data generated at UC San Diego IGM Genomics Center with an Illumina NovaSeq 6000 that was purchased with funding from a National Institutes of Health SIG grant (S10 OD026929). X.C. declares the work conducted at the Global Health Drug Discovery Institute (GHDDI) was supported by the Bill and Melinda Gates Foundation (grant INV-008249) and by funding from the Beijing Municipal Government. **Author contributions:** E.A.W. conceived of experiments, provided funding, provided supervision, wrote and reviewed the manuscript, constructed supplemental files, performed analysis, and designed figures. M.R.L., K.P.G.-M., and D.C. wrote the manuscript, constructed supplemental files, performed analysis, and constructed figures, with feedback from V.T. and D.A.F. Figure 1 was prepared by M.R.L. and E.A.W. Figure 2 was prepared by D.C. and E.A.W. Figure 3 was prepared by K.P.G.-M., D.C., and E.A.W. Figure 4 was prepared by D.C., K.P.G.-M., M.R.L., and E.A.W. E.S.I., T.S.-K., F.R., R.L.S., M.V., J.O., X.C., G.M.L., J.C.J., M.G.G.-L., S.P.D., S.P., and J.L.S.-S. performed *in vitro* selections. J.O., V.T., S.K.N., and T.N. provided phenotypic data. H.D. and R.C.S.E. generated transgenic conditional knockout lines and performed phenotypic assays. T.Y., S.M., H.P., and A.-C.U. provided additional sequence analysis. A.Y. and S.K.D. performed validation experiments. G.S., G.X., S.L.S., and T.S.S.-A. sourced compounds. All authors read and approved the submission. S.O. and J.L.D.S.-N. provided project administration. D.A.F., D.E.G., D.F.W., A.K.L., N.K., C.W.M., B.M., F.-J.G., and M.C.S.L. supervised the work, provided project administration, and acquired funding. J.M.E. helped write the manuscript and provided laboratory support. P.J.R. and S.L.S. reviewed the manuscript. **Competing interests:** H.D. is also affiliated with the Department of Pathology, University of Cambridge, Cambridge CB2 1QP, UK. F.-J.G. is a GSK employee and owns shares of the company. S.L.S. served on the scientific advisory board of Eisai Co., Ltd. The other authors declare that they have no competing interests. **Data and materials availability:** The raw sequencing data generated in this study have been submitted to the NCBI Sequence Read Archive database (<https://www.ncbi.nlm.nih.gov/sra/>) under accession number PRJNA1022010. Previously published sequencing files were downloaded from the following SRA accession numbers: PRJNA385508, PRJNA504044, PRJNA560380, PRJNA299203, PRJNA253899, PRJNA226625, PRJNA308112, PRJNA315690, PRJNA167166, and SRP012591. All other data described in this study are in the supplementary materials. Code created for this study is available in Zenodo (95). Evolved lines that were created for this study are available from the originating lab as listed in data S1. **License information:** Copyright © 2024 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>. This research was funded in whole or in part by the Bill and Melinda Gates Foundation (OPP1054480, INV-045906, INV-008249), a cOAlition S organization. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

## SUPPLEMENTARY MATERIALS

[science.org/doi/10.1126/science.adk9893](https://science.org/doi/10.1126/science.adk9893)

Figs. S1 to S12

Tables S1 to S10

References (96–100)

MDAR Reproducibility Checklist

Data S1 to S7

Submitted 2 October 2023; resubmitted 21 May 2024

Accepted 7 October 2024

10.1126/science.adk9893