Expansion of GTP cyclohydrolase I copy number in malaria parasites resistant to a pyrimidine biosynthesis inhibitor

Shiwei Liu[1], Emily R. Ebel[2], Jane Kim[1], Nnenna Ene[1], Thomas Werner Anthony Braukmann[3], Ellen Yeh[3], Elizabeth S. Egan[2], Jennifer L. Guler[1*]

[1]University of Virginia, Department of Biology, Charlottesville, VA, USA

[2]Stanford, Departments of Pediatrics and Microbiology & Immunology, Stanford, CA, USA

[3]Stanford University, Departments of Pathology and Microbiology & Immunology, Stanford, CA, USA

***Correspondence: Jennifer L. Guler**

Corresponding Author
jlg5fw@virginia.edu

Keywords: GTP cyclohydrolase I, Nanopore long reads, copy number variation, malaria, Dihydroorotate dehydrogenase, resistance

Abstract

Changes in the copy number of large genomic regions, termed copy number variations or CNVs, are an important adaptive strategy for malaria parasites. Numerous CNVs across the *Plasmodium falciparum* genome contribute directly to drug resistance or impact fitness of this protozoan parasite. CNVs that encompass the dihydroorotate dehydrogenase (DHODH) gene confer resistance to antimalarials that target this enzyme in the pyrimidine biosynthesis pathway (i.e. DSM1). During the characterization of DSM1 resistant parasite lines with DHODH CNVs, we detected selection of an additional CNV that encompasses 3 genes (~5 kb) including GTP cyclohydrolase I (GCH1 amplicon). While this locus has been implicated in increased fitness of antifolate resistant parasites, GCH1 CNVs had not previously been reported to contribute to resistance to other antimalarials. Here, we further explored the association between GCH1 and DHODH copy number. We visualized single long reads and directly quantified the number of tandem GCH1 amplicons in a parental line versus a DSM1-selected line. We found that the GCH1 amplicons share a consistent structure in all lines. However, we detected more reads that encompassed a higher number of amplicons in the resistant (up to 7 amplicons) compared to the parental line (3 amplicons). To better understand the implications of this result, we evaluated variation at this locus across multiple short- and long-read data sets collected from various parasite lines. Based on our analysis of parasites resistant to other DHODH inhibitors (DSM265, DSM267, and DSM705), GCH1 is not likely contributing directly to resistance; however, higher numbers of the GCH1 amplicon are associated with increased DHODH copies and may compensate for changes in metabolism of parasites. This is supported by the direct connection between folate and pyrimidine metabolism, which together contribute to nucleic acid biosynthesis. This study highlights the importance of studying clonal variation and potential biochemical connections as novel antimalarials move closer to clinical approval.

Introduction

Malaria is a disease caused by the protozoan *Plasmodium* parasite. *Plasmodium falciparum* is the leading cause of human malaria deaths (Rich et al., 2009). Due to lack of effective vaccines against malaria infection, antimalarial drugs are the primary approach for malaria treatment (Casares et al., 2010). However, drug efficacy is mitigated by the frequent emergence of antimalarial resistant parasites (Blasco et al., 2017).

Changes in the copy number of large genomic regions, termed copy number variations, or CNVs, are an important adaptive strategy for malaria parasites (Kidgell et al., 2006; Conway, 2007; Hyde, 2007; Ribacke et al., 2007; Nair et al., 2008; Cheeseman et al., 2009; Bopp et al., 2013; Guler et al., 2013; Menard and Dondorp, 2017). Numerous CNVs across the *P. falciparum* genome contribute directly to drug resistance or impact parasite fitness (Hyde, 2007; Ribacke et al., 2007; Nair et al., 2008; Guler et al., 2013). Amplification, one type of CNV with increased copy number, plays an essential role in the evolution of resistance to various antimalarials (Foote et al., 1989; Wilson et al., 1993; Hyde, 2007; Ribacke et al., 2007; Cheeseman et al., 2009; Guler et al., 2013; Heinberg et al., 2013; Osei et al., 2018). As one example, amplification of the dihydroorotate dehydrogenase (DHODH) gene in the *P. falciparum* genome confers resistance to DHODH inhibitors (i.e. DSM1) in parasites propagated in vitro (Guler et al., 2013). DHODH is an important enzyme in the *P. falciparum* pyrimidine biosynthesis pathway that contributes resources for nucleic acid synthesis (Phillips and Rathod, 2010; Mandt et al., 2019). DHODH amplicons presumably increase transcription and translation of the drug target to directly impact drug sensitivity (Guler et al., 2013).

In another example, amplification of the GTP cyclohydrolase 1 (GCH1) gene increases the fitness of clinical parasite populations that are antifolate resistant (i.e. pyrimethamine and sulfadoxine) (Kidgell et al., 2006; Ribacke et al., 2007; Nair et al., 2008; Osei et al., 2018). GCH1 is the first enzyme in the folate biosynthesis pathway and increased levels of this enzyme likely increases flux to compensate for the fitness costs of the resistance-conferring dihydropteroate synthase (DHPS) and dihydrofolate synthase (DHFS) mutations (Kidgell et al., 2006; Nair et al., 2008; Heinberg et al., 2013). Although the contribution of GCH1 amplification to antifolate resistance is well studied, this CNV has not been reported to contribute to resistance to antimalarials targeting other pathways (Heinberg et al., 2013; Kümpornsin et al., 2014; Heinberg and Kirkman, 2015).

Typically, gene copy number is studied using widely accessible high coverage short read sequencing (Guler et al., 2013; Herman et al., 2014; Manary et al., 2014; Cowell et al., 2018; Huckaby et al., 2018). However, this approach has limitations including non-unique mapping in repetitive regions, the inability to resolve complex genomic regions, and overall difficulty of detecting structural variations (Alkan et al., 2011; Treangen and Salzberg, 2012; Kosugi et al., 2019). These challenges are exacerbated by the high AT-content of the *P. falciparum* genome (Beghain et al., 2016; Miles et al., 2016). Long read technologies such as Oxford Nanopore sequencing have the potential to span low complexity and repetitive regions to better represent structural variation (Cretu Stancu et al., 2017; Sedlazeck et al., 2018a, 2018b; Ho et al., 2020). Moreover, the single molecule sequencing allows examination of clonal heterogeneity of a parasite population (Belikova et al., 2020; Rugbjerg et al., 2021).

In this study, we identified a potential positive association between GCH1 and DHODH copy number from previously acquired short read sequencing data. To explore this relationship further, we performed long read sequencing and directly observed the expansion of the GCH1 amplicon. Using single long read visualization, we also determined that the structure and orientation of the amplicon was preserved during expansion. When we evaluated short read data from parasite lines that were resistant to other DSM-based compounds, we did not detect increases in GCH1 copy number. This result indicates that GCH1 does not contribute directly to resistance; however, our observations, as well as the biochemical connection between folate and pyrimidine biosynthesis, suggest that increased copies of GCH1 may facilitate the acquisition of increased DHODH copy number under certain selection conditions. Further study of the relationship between these two genomic loci is important considering the imminent use of DHODH inhibitors to treat clinical malaria.

Materials and Methods

DSM1 and parasite clones

DSM1 is a triazolopyrimidine antimalarial that specifically and potently inhibits the *P. falciparum* dihydroorotate dehydrogenase enzyme of pyrimidine biosynthesis (Phillips et al., 2008). DSM1-resistant parasites were previously selected according to the scheme depicted in (**Figure 1**) (Guler et al., 2013). In this study, we simplified the naming scheme to represent low (L), moderate (M), and high (H) levels of resistance to DSM1 (**Figure 1A**).

Parasite culture

We thawed erythrocytic stages of *P. falciparum* (*Dd2*, MRA-150; *3D7*, MRA-102, Malaria Research and Reference Reagent Resource Center, BEI Resources and DSM1 resistant clones as highlighted in **Figure 1A**, a generous gift from Pradipsinh Rathod, University of Washington) from frozen stocks and maintained them as previously described (Haynes et al., 1976). Briefly, we grew parasites at 37 °C in vitro at 3% hematocrit (serotype A positive human erythrocytes, Valley Biomedical, Winchester, VA) in RPMI 1640 medium (Invitrogen, USA) containing 24 mM NaHCO3 and 25 mM HEPES, and supplemented with 20% human type A positive heat inactivated plasma (Valley Biomedical, Winchester, VA) in sterile, sealed flasks flushed with 5% O_2, 5% CO_2, and 90% N_2 (Guler et al., 2013). We maintained the cultures with media changes every other day and sub-cultured them as necessary to keep parasitemia below 5%. We determined all parasitemia measurements using SYBR green-based flow cytometry (Bei et al., 2010). We routinely tested cultures using the LookOut Mycoplasma PCR Detection Kit (Sigma-Aldrich, USA) to confirm negative Mycoplasma status.

DNA extraction for long read sequencing

We lysed asynchronous *P. falciparum*-infected erythrocytes with 0.15% saponin (Sigma-Aldrich, USA) for 5 min at room temperature and washed them three times with 1× PBS (diluted from 10× PBS Liquid Concentrate, Gibco, USA). We then lysed parasites with 0.1% Sarkosyl Solution (Bioworld, bioPLUS, USA) in the presence of 1 mg/ml proteinase K (from *Tritirachium album*, Sigma-Aldrich, USA) overnight at 37 °C. We first extracted nucleic acids with phenol/chloroform/isoamyl alcohol (25:24:1) pH 8.0 (Sigma-Aldrich, USA) three times using 1.5 ml light Phase lock Gels (5Prime, USA), then further extracted nucleic acids with chloroform

twice using 1.5 ml light Phase lock Gels (5Prime, USA). Lastly, we precipitated the DNA with ethanol using the standard Maniatis method (Maniatis et al., 1989). To obtain high molecular weight genomic DNA, we avoided any pipetting during the extraction, transferred solutions by directly pouring it from one tube to another, and mixed solutions by gently inverting the tubes.

Oxford Nanopore long read sequencing and analysis

We subjected 1 µg of high molecular weight genomic DNA from each sample to library preparation for Oxford Nanopore sequencing following the Nanopore Native barcoding genomic DNA protocol (version: NBE_9065_v109_revAB_14Aug2019) with 1x Ligation Sequencing kit (SQK-LSK109, Oxford Nanopore Technologies, Oxford, UK). We performed DNA repair and end preparation using NEBNext FFPE DNA Repair Mix (New England Biolabs, Ipswich, MA, USA) and NEBNext End Repair/dA-Tailing Module (New England Biolabs, Ipswich, MA, USA). We cleaned the A-tailed fragments using 0.9X AMPure XP beads (Beckman Coulter, High Wycombe, UK). We then ligated barcodes to the end-prepped DNA using the Native Barcoding Expansion 1–12 kit (EXP-NBD104, Oxford Nanopore Technologies, Oxford, UK) and Blunt/TA Ligase Master Mix (New England Biolabs, Ipswich, MA, USA). We cleaned the barcoded samples using 0.9X AMPure XP beads. Then we pooled barcoded samples in equimolar ratios and subjected to an adaptor ligation step, using the Adapter Mix II from the Native Barcoding Expansion 1–12 kit and NEBNext Quick Ligation Reaction Buffer (New England Biolabs, Ipswich, MA, USA) as well as Quick T4 DNA Ligase (New England Biolabs, Ipswich, MA, USA). After adaptor ligation, we cleaned the library using AMPure XP beads. We quantified the adapter ligated and barcoded DNA using a Qubit fluorimeter (Qubit 1X dsDNA High Sensitivity Assay Kit, Life Technologies, Carlsbad, CA). We sequenced the WT1 and initial H2 libraries using a R9.4.1 flow cell (FLO-MIN106D, Oxford Nanopore Technologies, Oxford, UK) on MinION (Oxford Nanopore Technologies, Oxford, UK) and another H2 library using the R10 flow cell (FLO-MIN111) on MinION. To obtain the maximum number of reads, we ran both flow cells for 48 h (controlled and monitored using the MinKNOW software (3.6.5)).

For base calling and demultiplexing of the Nanopore sequencing reads, we used Guppy (version 3.4.5+fb1fbfb) with the parameter settings "-c dna_r9.4.1_450bps_hac.cfg --barcode_kits "EXP-NBD104" -x auto" for samples sequenced with R9.4.1 flow cell and "-c dna_r10_450bps_hac.cfg -x auto" for the sample sequenced with R10 flow cell. We checked the read length and read quality using "Nanoplot" (version 1.0.0) (see **Supplementary Table 1**). We trimmed the adapters with "qcat" (version 1.1.0) (Oxford Nanopore Technologies) and filtered the reads with a cutoff "length ≥ 500 and Phred value ≥ 10" using the program " filtlong version 0.2.0" (https://github.com/rrwick/Filtlong). To estimate the coverage of sequencing reads in each sample, we aligned the filtered reads to *Plasmodium falciparum* 3D7 reference genome using "minimap2" (version 2.17) (Li, 2018). "QualiMap" (version 2.2.1) (García-Alcalde et al., 2012) was used to calculate the coverage of the aligned reads (**see Supplementary Table 1**).

Shiny analysis of long reads

To visualize structural variants in the parasite genome, we used a custom R Shiny script to plot the arrangement of reference gene segments on individual Nanopore reads (Ebel et al., 2023). Briefly, we defined a target region in the 3D7 reference genome (chromosome 12: 932916 bp –

999275 bp) that contained 3 genes in GCH1 amplicon and 11 flanking genes. We extracted the reference sequences of these genes and subsequent intergenic regions, then split these sequences into fragments of 500-1000 base pairs. We compared these fragments to individual Nanopore reads using BLAST (Ye et al., 2006). We used the BLAST output as input for a custom R script, which drew rectangles representing homology between the defined genes (y-axis) and each individual read (x-axis). The percent identity required to draw a homologous rectangle was allowed to vary between reads, which varied in quality, using a slider in the Shiny app. To filter out long reads with potentially spurious hits to gene fragments, we also used BLAST to compare Nanopore reads to the reference genome and removed reads with <90% identity to chromosome 12. To compare the mean copy number of GCH1 amplicon between WT1 and H2 reads covering GCH1 as well as the read length of these reads between WT1 and H2, we performed a one way ANOVA in Microsoft Excel with Alpha value of 0.05.

Short read sequencing analysis and CNV detection

We analyzed CNVs in Illumina short read datasets of *P. falciparum* parasites selected by three DSM antimalarial drugs (DSM1, DSM265, DSM267, and DSM705, **Supplementary Table 2**) (Guler et al., 2013; Mandt et al., 2019; Palmer et al., 2021). We first processed and mapped the reads to reference genome as previously described (Huckaby et al., 2018; Liu et al., 2021). Briefly, we trimmed Illumina adapters from reads with BBDuk tool in BBMap (version 38.57) (Bushnell 2016). We aligned each fastq file to the *3D7 P. falciparum* reference genome with Speedseq (version 0.1.0) through BWA-MEM alignment (Chiang et al., 2015). We discarded the reads with low-mapping quality score (below 10) and removed duplicated reads using Samtools (version 1.10) (Li et al., 2009). We analyzed split and discordant reads from the mapped reads using LUMPY in Speedseq to determine the location and length of the previously reported GCH1, DHODH and multidrug resistance protein 1 (MDR1) CNVs (**Supplementary Table 2**) (Layer et al., 2014). For read-depth analysis, we further filtered the mapped reads using a mapping quality score of 30. To determine the copy number of the GCH1, DHODH and MDR1 CNVs, we used CNVnator (version 0.4.1) with a bin size of 100 bp; the optimal bin size was chosen to detect GCH1 CNVs in all analyzed samples (Abyzov et al., 2011).

Droplet Digital PCR

Prior to Droplet Digital (dd) PCR, we digested DNA with restriction enzyme RsaI (Cut Site: GT/AC) following the manufacturer's instructions (New England Biolabs, Ipswich, MA, USA) in 37°C incubation for one hour. We selected the restriction enzyme RsaI to cut outside of the ddPCR amplified regions of desired genes and separate copies of CNVs to be distributed into droplets. We diluted the digested DNA for ddPCR reactions. We performed ddPCR using ddPCR Supermix for Probe (no dUTP, Bio-Rad Laboratories, California, USA) with DNA input 0.1 ng (in duplicate per sample), 0.025 ng (in duplicate per sample) as previously described (McDaniels et al., 2021). The primers and probes used in reactions are included in **Supplementary Table 3**. The PCR protocol for probe-based assay was 95°C for 10 min, followed by 40 rounds of 95°C for 30 sec and 60°C for 1 min. Seryl-tRNA synthetase and calcium-transporting ATPase (ATP6) served as a single copy reference genes on chromosome 7 and chromosome 1 respectively; dihydroorotate dehydrogenase (DHODH) and GTP cyclohydrolase 1 (GCH1) are multi-copy genes (**Supplementary Table 3**). We performed droplet generation and fluorescence readings per the manufacturer's instructions. For each

reaction, we required a minimum number of 10,000 droplets to proceed with analysis. We calculated the ratio of positive droplets containing a single- (ATP6) or multi-copy gene (GCH1, DHODH) versus a single-copy gene (Seryl-tRNA synthetase) using the Quantasoft analysis software (QuantaSoft Version 1.7, BioRad Laboratories) and averaged between independent replicates.

Results

Through expanded analysis of short read data from a family of parasites selected with DSM1, originally presented in (Guler et al., 2013) (**Figure 1A**), we noticed a potential positive association between GCH1 and DHODH copy number (**Figure 1B**). Using droplet digital PCR on analogous parasite lines that had recently been propagated in our laboratory, we confirmed that GCH1 copy number trended higher as DHODH copy number increased (**Table 1**). Due to the relationship between these parasite lines (**Figure 1A**), we were unable to perform statistics on these data.

We conducted long read sequencing to more precisely define GCH1 copy number and amplicon structure in the parental line versus one DSM1-selected line (WT1 and H2, **Supplementary Table 1**). We directly visualized single reads using an app that represents gene segments of individual Nanopore reads (Materials and Methods). Small amplicons, like those including GCH1, are especially conducive to this approach because long reads span multiple copies of the amplicon as well as both up and downstream regions. Read visualization showed that two 3-gene amplicons, separated by an inversion of the same gene set, were conserved between the parental and H2 line (**Figure 2A and B**). This amplicon structure was reported previously in the WT1 (*Dd2*) parental line (**Figure 2C**) (Kidgell et al., 2006).

Using read visualization, we also manually recorded the number of GCH1 amplicon units (depicted in **Figure 2C**) from both spanning and non-spanning reads (**Figure 3A, Supplementary Table 4**). We consistently detected 3 amplicon units per read from the parental line. However, we observed more reads that encompassed a higher number of GCH1 amplicon units in the H2 selected line (up to 7 units). If only considering spanning reads, that is those covering both upstream and downstream of the amplicon, the difference between WT1 and H2 mean copy number is significantly different ($F(1,10)=[75]$, $p = 5.84E-06$). There was not a significant difference in mean read lengths used in the GCH1 analysis ($F(1,34)=[1.07]$, $p = 0.31$, **Figure 3B**), indicating that a skewed read distribution was not contributing to this difference. All groups of amplicons from the H2 line began with a set of 3 amplicon units, as observed in the parental line, followed by groups of 2 amplicon units where one was inverted and the other was in normal orientation (**Figure 2C**).

To better understand whether variation in copy number of the GCH1 locus was common in other laboratory-adapted parasite lines, we evaluated several additional long read-based datasets (**Supplementary Table 4**, **Supplementary Methods**, Vembar et al., 2016). In general, different parasite lines exhibited different GCH1 amplicon sizes as expected (i.e. *Dd2* versus *3D7*, as reported previously (Kidgell et al., 2006)) but the amplicon copy number was relatively stable across diverse datasets; we detected ~10-25% of reads above the expected copy number for each respective line. This is in contrast to the H2 line that had >50% of reads that depicted amplicon copy numbers greater than the expected copy number for a *Dd2*-derived line.

To investigate whether an increase in GCH1 amplicons was contributing to resistance of DHODH inhibitors in general, we evaluated the copy number of GCH1 amplicons from parasites selected with other DSM derivatives (Mandt et al., 2019; Palmer et al., 2021). Contrary to DSM1-selected parasites, we did not detect increases in GCH1 amplicon number in parasites resistant to DSM265, DSM267, or DSM705 (compared to parental *Dd2* or *3D7* parasite lines, **Figure 4A-4C**, **Supplementary Table 5**). For this analysis, we only included resistant parasite lines that carried DHODH amplicons. When we compiled the data for each parasite line together, those that harbored more DHODH amplicons generally maintained higher numbers of GCH1 amplicons and this association was more prominent in *Dd2*-derived parasite lines then *3D7*-derived parasite lines (**Figure 4D** and **4E**).

Discussion

We found that GCH1 copy number is positively associated with DHODH copy number in parasites resistant to DSM1. This association was initially detected using both short-read sequencing as well as an accurate PCR-based method ddPCR (**Figures 1B and 4, Table 1**); however, limitations in each of these methods contributes to inaccuracies in these results. For example, because the GCH1 amplicon is small (~5kb), copy number analysis using short reads requires small stretches of the genome to be combined, or binned, together; smaller sized bins in general leads to higher levels of variation in CNV calling and potentially false differences. Furthermore, DNA fragmentation and restriction digestion can alter distribution of DNA fragments into oil droplets during ddPCR, thus limiting copy number quantification. Additionally, both of these methods result in an average value for the entire population of parasites.

We therefore utilized long-read sequencing combined with a custom visualization tool (Ebel et al., 2023) to directly quantify the copy number of this locus and assess the structure of the GCH1 amplicon (**Figures 2** and **3**). Each long read represents a DNA strand from a single parasite genome and therefore, this approach is an accurate way to visualize copy number heterogeneity across a population of haploid parasites. We considered whether our detection of more GCH1 amplicons on reads from the H2 line (**Figure 3A**) was related to Nanopore sample preparation or natural variation during in vitro culture. We excluded differences in sample preparation because the parental and H2 long read sequencing runs had similar N50s and read length distributions (**Supplementary Table 4**, **Figure 3B**). While we did observe a low level of variation at the GCH1 locus in different parasite lines that were grown independently (~10-25% of reads), the variation in copy number in the H2 line was well above this level (>50% of reads, **Supplementary Table 5**). This observation, combined with the visualization of up to 7 tandem amplicons in a single read (**Figure 2B**), provides direct evidence for GCH1 amplicon expansion in the DSM1 resistance context.

GCH1 amplicons have previously only been associated with antifolate resistance (Kidgell et al., 2006; Nair et al., 2008; Heinberg et al., 2013; Kümpornsin et al., 2014; Heinberg and Kirkman, 2015; Osei et al., 2018). An increase in flux through the folate biosynthesis pathway alleviates fitness effects of mutations that confer pyrimethamine and sulfadoxine resistance (Kidgell et al., 2006; Nair et al., 2008; Kümpornsin et al., 2014; Heinberg and Kirkman, 2015; Osei et al., 2018). A similar contribution to resistance to DHODH inhibitors would not be surprising, given the close connection of the folate and pyrimidine biosynthesis pathways (**Figure 5**); they both

contribute to nucleotide biosynthesis and converting dUMP to dTMP requires conversion of N^5, N^{10}-methylene-THF to DHF.

Although not required for resistance to all DHODH inhibitors (**Figure 4A-C**, **Supplementary Table 5**), we speculate that a change in GCH1 copy number arose serendipitously during DSM1 selection and further increases were beneficial for parasite fitness. Increased copies of GCH1 may facilitate the acquisition of increased DHODH copy numbers, especially under specific conditions. For example, DSM1 selections were uniquely performed in media supplemented with 20% human serum as opposed to AlbuMAX II (**Supplementary Table 5**) (Guler et al., 2013). It is possible that the presence of folate precursors in different media formulations changes the parasite's dependence on folate biosynthesis and thus, GCH1 flux. Levels of para-amino benzoic acid (pABA) can vary widely in human serum while the common AlbuMAX serum replacement contains no additional pABA (Chulay et al., 1984; Watkins et al., 1985; Wang et al., 1986; Salcedo-Sora et al., 2011; Valenciano et al., 2019). These environments likely exert different selective pressures on parasites during drug selection as reported previously (Kumar et al., 2021); higher pABA levels in human serum may drive positive selection for higher GCH1 copy numbers and hold extra benefits by contributing to both folate and pyrimidine biosynthesis (**Figure 5**). Of note, our observation of some level of variation in standard parasite lines (**Supplementary Table 4**) suggests that growth environment could drive changes at this locus.

Another factor that could contribute to GCH1 evolution is genetic background. The *Dd2* parasite line (and its parent line, *W2*) were isolated in Southeast Asia (Oduola et al., 1988; Wellems et al., 1990), where antifolates were widely used. Consequently, *W2* and *Dd2* carry 5 mutations in folate biosynthesis enzymes (DHPS and DHFR). The *3D7* line originated in Africa (Walliker et al., 1987) and is wild type at these loci. In our studies, we observed a positive association of copy numbers between DHODH and GCH1 in *Dd2* versus *3D7* parasite lines (**Figure 4D** and **4E**), indicating that mutant backgrounds may rely more heavily on the GCH1 CNV to alleviate fitness effects for resistance to both antifolates and DHODH inhibitors.

Importantly, DSM265 has entered clinical trial sfor the treatment of clinical malaria (Mandt et al., 2019). Antifolate resistance and GCH1 CNVs are widespread in clinical isolates (Kidgell et al., 2006; Ribacke et al., 2007; Nair et al., 2008; Osei et al., 2018), thus, necessitating further evaluation of the contribution of GCH1 in parasites that are resistant to DHODH inhibitors.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

S.L and J.L.G. designed the study and wrote the manuscript. S.L., N.E. and J.K performed the experiments and data analysis. E.R.E developed the Shiny application and provided instructions for using the application. E.R.E. and E.S.E. edited and reviewed the manuscript. E. Y. and T. B. conducted the experiment described in the Supplementary Methods and wrote the Supplementary Methods. All authors contributed to the article and approved the submitted version.

Funding

Research reported in this publication was supported by R01AI150856 (to J.L.G.), by 1R01AI141366, Burroughs Wellcome Fund, and Chan Zuckerberg Biohub (to E.Y.), and by DP2HL137186 (to E.S.E.). Additionally, discretionary research funds from Biology Distinguished Graduate Mentoring Award (to S.L.) and Small Research and Travel Grant (to N.E.) at the University of Virginia contributed support for sequencing efforts. E.R.E. is partially supported by a National Science Foundation (NSF) Postdoctoral Research Fellowship in Biology (Grant No. 2109851). These funding agencies were not involved in the experimental design, analysis, interpretation of the data, or writing of the manuscript.

Acknowledgments

Our thanks to Michelle Warthan and Noah Brown (University of Virginia) for experimental support, Martin Wu and Mercedes Campos-Lopez (University of Virginia) for Oxford Nanopore sequencing support, and the laboratories of Dr. William Petri Jr (University of Virginia) for their helpful discussions and insight. We are grateful for David Fidock (Columbia University), Dyann F. Wirth (Harvard University) for sharing sequencing data and information on selected cell lines.

Supplementary Material

The Supplementary Tables and Supplementary Methods for this article can be found in Supplementary files

Data Availability Statement

The nucleotide sequences obtained in the present study have been deposited in the Sequenced Read Archive (SRA) database under the accession number PRJNA820055.

Abbreviations

CNV, copy number variation; DHODH, dihydroorotate dehydrogenase; DSM1, 5-Methyl-N-(2-naphthyl)[1,2,4]triazolo[1,5-a]pyrimidin-7-amine; DHPS, dihydropteroate synthase; DHFS, dihydrofolate synthase; GCH1, GTP cyclohydrolase 1; MDR1, multidrug resistance protein 1; ddPCR, Droplet Digital PCR; ATP6, calcium-transporting ATPase; pABA, para-amino benzoic acid; SD, standard deviation; 50S RPL24, 50S ribosomal protein L24; YHM2, citrate/oxoglutarate carrier protein; Gln, Glutamine; DHO, Dihydroorotate; UMP, Uridine monophosphate; dUMP, Deoxyuridine monophosphate; dTMP, deoxythymidine monophosphate; GTP, Guanosine-5'-triphosphate; DHF, Dihydrofolate; THF, Tetrahydrofolate; HMDP-P2, 6-hydroxymethyl-7, 8-dihydropterin diphosphate; 5, 10-CH2-THF: 5,10-Methylenetetrahydrofolate.

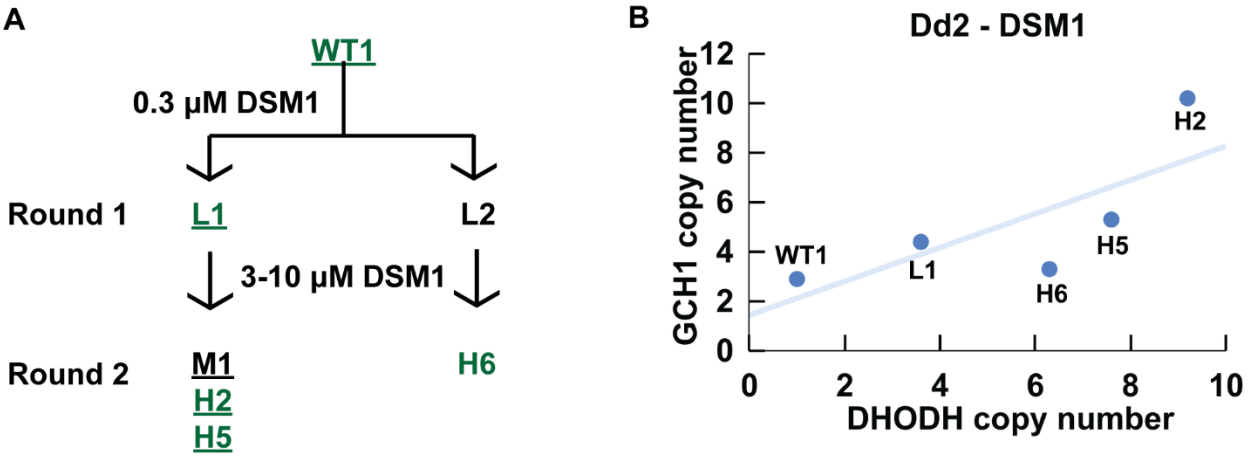
Tables

Table 1. Positive GCH1:DHODH association is validated using Droplet Digital PCR on modern parasite lines.

Line	Sample type	Average copy number (SD)*			GCH1 CN relative to parent	DHODH CN relative to parent
		GCH1	DHODH	ATP6		
WT1	Parent	2.5 (0.2)	0.8 (0.1)	1.1 (0.0)	1.0	1.0
L1	Round 1 selection	3.9 (0.3)	3.2 (0.1)	1.0 (0.1)	1.6	4
M1	Round 1 selection	3.9 (0.1)	5.2 (0.3)	1.0 (0.1)	1.6	6.5
H2	Round 2 selection	4.6 (0.2)	6.0 (0.3)	1.1 (0.1)	1.8	7.5
H5	Round 2 selection	4.3 (0.3)	7.0 (0.4)	1.0 (0.1)	1.7	8.8

*The average copy number is calculated by comparing ddPCR signal to a single copy gene signal (Seryl tRNA synthetase, PF3D7_0717700). N=4. ATP6 copy number is expected to be 1 in all parasite lines.

370 **Figures**



371 **Figure 1. GCH1 copy number increase is positively correlated with DHODH copy number**
372 **in one family of DSM1 resistant parasites. (A)** Schematic depicting DSM1 selections, as
373 presented previously (Guler et al. PLoS Pathogens 2013). Green: Illumina short read sequenced
374 lines. Underline: modern lines confirmed by ddPCR analysis. Wild-type (WT1, *Dd2*) *P.*
375 *falciparum* was selected with DSM1 in two steps; the first step selected for low-level (L)
376 resistance and the second step selected for moderate- (M) or high-level (H) resistance. DSM1
377 EC50 values are as follows: L1 (1 µM), L2 (0.9 µM), M1 (7.2 µM), H2 (85 µM), H5 (56 µM),
378 H6 (49 µM). All values were previously reported in and clone names adapted from (Guler et al.,
379 2013). **(B)** Relationship between GCH1 and DHODH copy number in DSM1 selected parasites
380 as quantified using short read data from Guler et al. 2013. A trendline was added to show the
381 relationship between GCH1 and DHODH copy numbers but a correlation coefficient could not
382 be calculated due to the small sample size (n=5) and dependence among the lines.

383

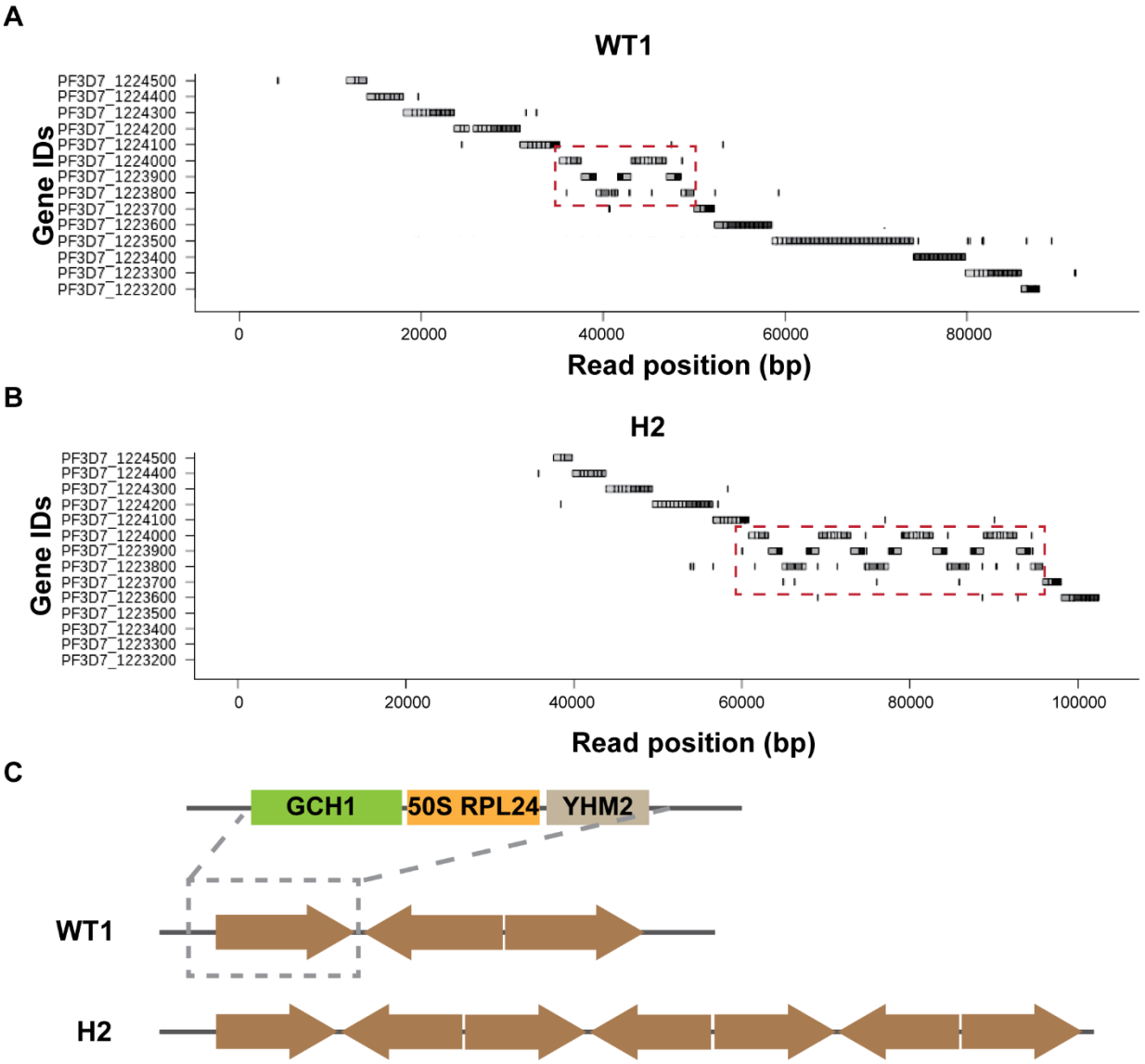


Figure 2. Long-read visualization shows GCH1 amplicon has the same boundaries and structure in a DSM1 resistant parasite line. (A) and **(B)** Representative images from the Shiny app comparing the GCH1 amplicon in WT1 and H2 reads (alignment to the *3D7* reference genome with no GCH1 amplicon represented). Red dashed square: GCH1 amplicon. Each gene sequence was split into <=500 bp fragments and blasted against individual Nanopore reads (darker: genic regions, lighter: intergenic regions). **(C)** Diagram of amplicon orientation in GCH1 amplicon for WT1 and H2 parasite lines. The three genes within the GCH1 amplicon unit include PF3D7_1224000 (GTP cyclohydrolase I, GCH1), PF3D7_1223900 (50S ribosomal protein L24, putative, 50S RPL24), and PF3D7_1223800 (citrate/oxoglutarate carrier protein, putative, YHM2).

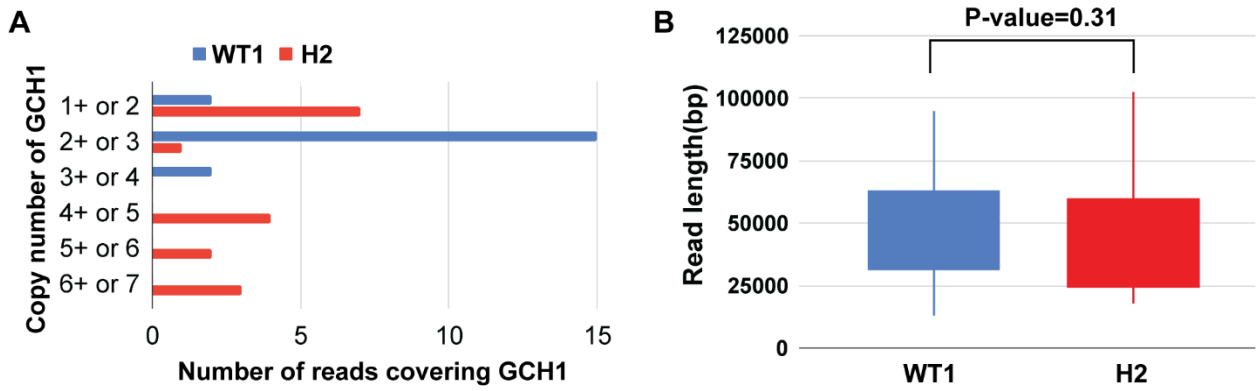


Figure 3. Quantification of long reads displays greater variability and increase in GCH1 amplicon number in DSM1 resistant parasites. Copy number **(A)** and read length **(B)** distributions from Nanopore long reads (>=5kb) covering the GCH1 amplicon in WT1 and H2 parasite line. A one-way ANOVA revealed that there was not a significant difference in mean read length between WT1 and H2 (F (1,34)=[1.07], p = 0.31).

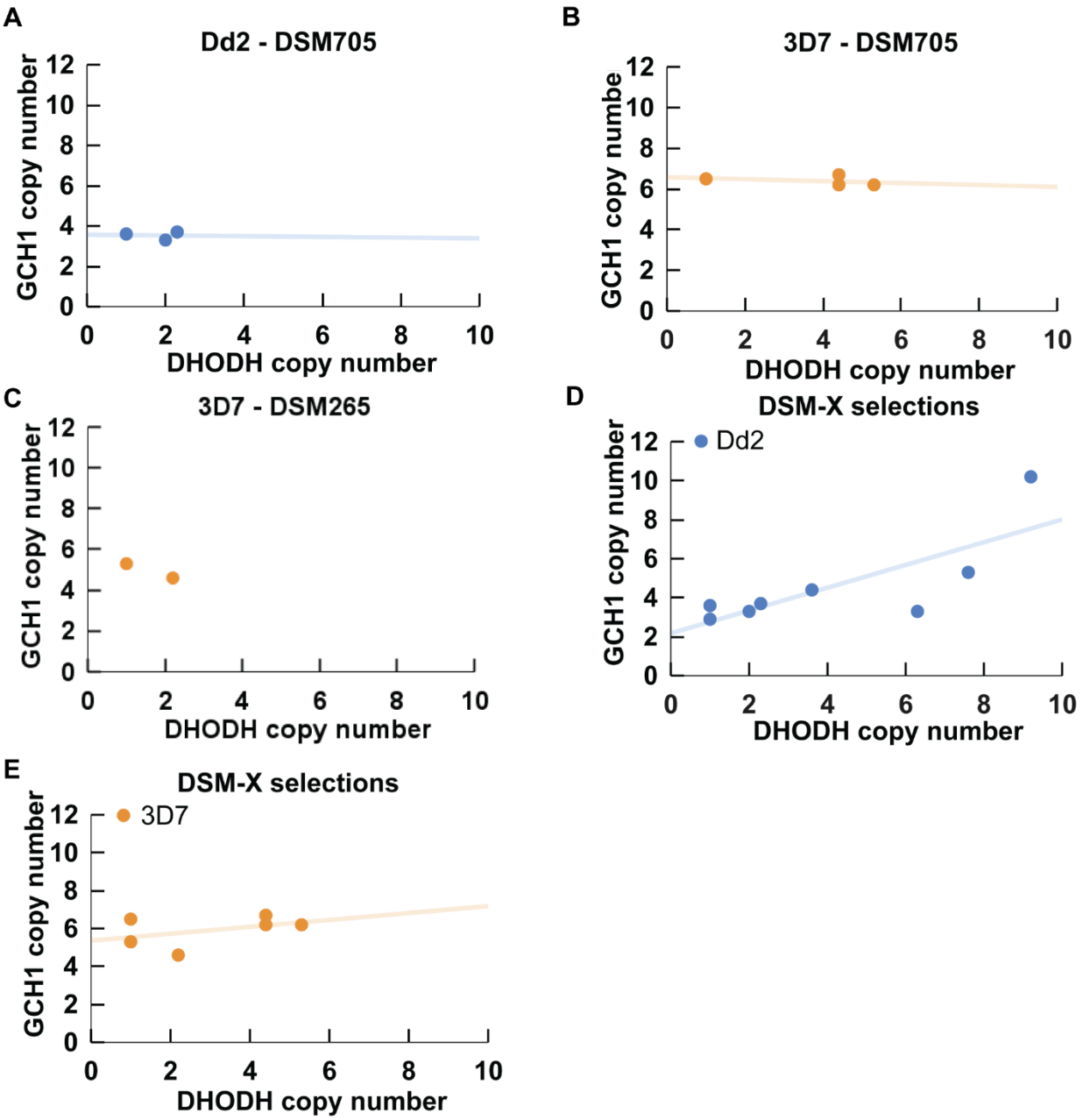


Figure 4. Parasite lines with more DHODH amplicons generally have more GCH1 amplicons. **(A)** Relationship between GCH1 and DHODH copy number in *Dd2* parent line and DSM705 selected lines. **(B)** Relationship between GCH1 and DHODH copy number in *3D7* parent line and DSM705 selected lines. **(C)** Relationship between GCH1 and DHODH copy number in *3D7* parent line and DSM265 selected lines. **(D)** Relationship between GCH1 and DHODH copy number in *Dd2* parent line and DSM1, DSM705, DSM265-selected lines. Data from DSM1 selected lines are represented in Figure 1B as well. **(E)** Relationship between GCH1 and DHODH copy number in *3D7* parent line and DSM1, DSM705, DSM265-selected lines. Correlation coefficients are not calculated due to small sample sizes and dependence among DSM1 selected lines. Trendlines show the relationship between GCH1 and DHODH copy numbers in available lines.

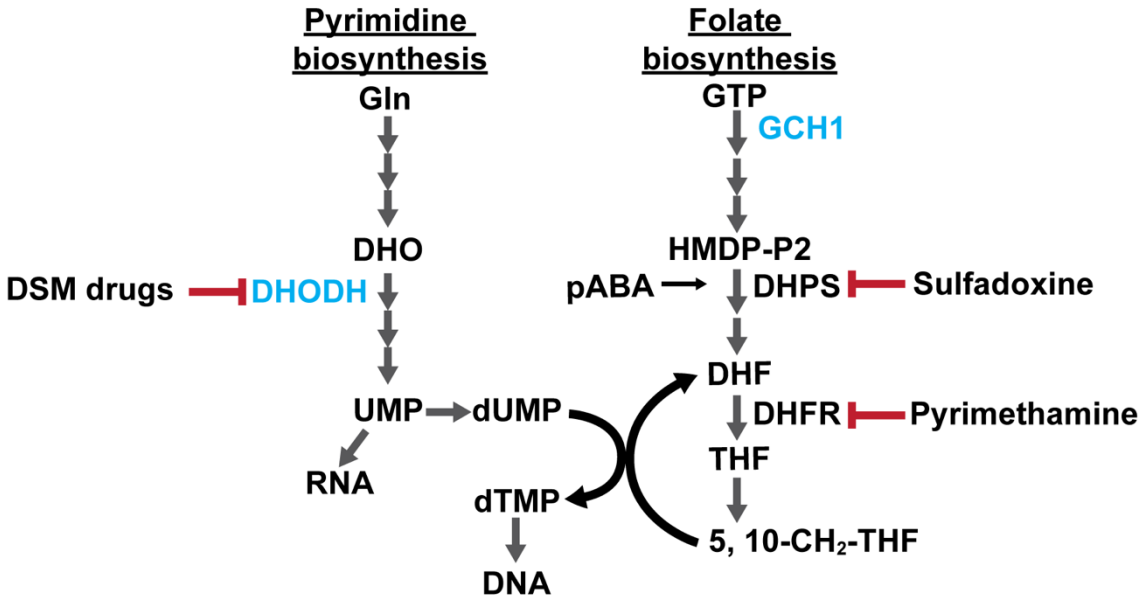


Figure 5. The connection between pyrimidine and folate biosynthesis pathways. Enzymes with gene copy number variations are indicated in blue. Gln: Glutamine; DHO: Dihydroorotate; UMP: Uridine monophosphate; dUMP: Deoxyuridine monophosphate; dTMP: deoxythymidine monophosphate; GTP: Guanosine-5'-triphosphate; DHPS: Dihydropteroate synthase; DHFR: Dihydrofolate reductase; DHF: Dihydrofolate; THF: Tetrahydrofolate; HMDP-P2: 6-hydroxymethyl-7, 8-dihydropterin diphosphate; pABA: para-amino-benzoic acid; 5, 10-CH2-THF: 5,10-Methylenetetrahydrofolate.

References

Abyzov, A., Urban, A. E., Snyder, M., and Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 21, 974–984. doi:10.1101/gr.114876.110.

Alkan, C., Coe, B. P., and Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics* 12, 363–376. doi:10.1038/nrg2958.

Beghain, J., Langlois, A.-C., Legrand, E., Grange, L., Khim, N., Witkowski, B., et al. (2016). Plasmodium copy number variation scan: gene copy numbers evaluation in haploid genomes. *Malaria Journal* 15, 206. doi:10.1186/s12936-016-1258-x.

Bei, A. K., Desimone, T. M., Badiane, A. S., Ahouidi, A. D., Dieye, T., Ndiaye, D., et al. (2010). A flow cytometry-based assay for measuring invasion of red blood cells by Plasmodium falciparum. *Am J Hematol* 85, 234–237. doi:10.1002/ajh.21642.

Belikova, D., Jochim, A., Power, J., Holden, M. T. G., and Heilbronner, S. (2020). "Gene accordions" cause genotypic and phenotypic heterogeneity in clonal populations of Staphylococcus aureus. *Nature Communications* 11, 3526. doi:10.1038/s41467-020-17277-3.

435 Blasco, B., Leroy, D., and Fidock, D. A. (2017). Antimalarial drug resistance: linking
436 Plasmodium falciparum parasite biology to the clinic. *Nature medicine* 23, 917–928.
437 doi:10.1038/nm.4381.

438 Bopp, S. E. R., Manary, M. J., Bright, A. T., Johnston, G. L., Dharia, N. V., Luna, F. L., et al.
439 (2013). Mitotic Evolution of Plasmodium falciparum Shows a Stable Core Genome but
440 Recombination in Antigen Families. *PLOS Genetics* 9, e1003293.
441 doi:10.1371/journal.pgen.1003293.

442 Bushnell B. BBMap. Available at: http://sourceforge.net/projects/bbmap (University of
443 California, Berkeley, 2016) [Accessed December 1, 2021].

444 Casares, S., Brumeanu, T.-D., and Richie, T. L. (2010). The RTS,S malaria vaccine. *Vaccine* 28,
445 4880–4894. doi:10.1016/j.vaccine.2010.05.033.

446 Cheeseman, I. H., Gomez-Escobar, N., Carret, C. K., Ivens, A., Stewart, L. B., Tetteh, K. K. A.,
447 et al. (2009). Gene copy number variation throughout the Plasmodium falciparum genome. *BMC*
448 *Genomics* 10, 353–353. doi:10.1186/1471-2164-10-353.

449 Chiang, C., Layer, R. M., Faust, G. G., Lindberg, M. R., Rose, D. B., Garrison, E. P., et al.
450 (2015). SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nat Methods* 12, 966–
451 968. doi:10.1038/nmeth.3505.

452 Chulay, J. D., Watkins, W. M., and Sixsmith, D. G. (1984). Synergistic Antimalarial Activity of
453 Pyrimethamine and Sulfadoxine against Plasmodium falciparum In Vitro. *The American Journal*
454 *of Tropical Medicine and Hygiene* 33, 325–330. doi:10.4269/ajtmh.1984.33.325.

455 Conway, D. J. (2007). Molecular epidemiology of malaria. *Clin Microbiol Rev* 20, 188–204.
456 doi:10.1128/CMR.00021-06.

457 Cowell, A. N., Istvan, E. S., Lukens, A. K., Gomez-Lorenzo, M. G., Vanaerschot, M., Sakata-
458 Kato, T., et al. (2018). Mapping the malaria parasite druggable genome by using in vitro
459 evolution and chemogenomics. *Science* 359, 191–199. doi:10.1126/science.aan4472.

460 Cretu Stancu, M., van Roosmalen, M. J., Renkens, I., Nieboer, M. M., Middelkamp, S., de Ligt,
461 J., et al. (2017). Mapping and phasing of structural variation in patient genomes using nanopore
462 sequencing. *Nature Communications* 8, 1326. doi:10.1038/s41467-017-01343-4.

463 Ebel, E. R., Kim, B. Y., McDew-White, M., Egan, E. S., Anderson, T. J. C., and Petrov, D. A.
464 (2023). Antigenic diversity in malaria parasites is maintained on extrachromosomal DNA.
465 *bioRxiv*, 2023.02.02.526885. doi: 10.1101/2023.02.02.526885.

466 Foote, S. J., Thompson, J. K., Cowman, A. F., and Kemp, D. J. (1989). Amplification of the
467 multidrug resistance gene in some chloroquine-resistant isolates of P. falciparum. *Cell* 57, 921–
468 930. doi:10.1016/0092-8674(89)90330-9.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., et al. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678–2679. doi:10.1093/bioinformatics/bts503.

Guler, J. L., Freeman, D. L., Ahyong, V., Patrapuvich, R., White, J., Gujjar, R., et al. (2013). Asexual Populations of the Human Malaria Parasite, Plasmodium falciparum, Use a Two-Step Genomic Strategy to Acquire Accurate, Beneficial DNA Amplifications. *PLOS Pathogens* 9, e1003375. doi:10.1371/journal.ppat.1003375.

Haynes, J. D., Diggs, C. L., Hines, F. A., and Desjardins, R. E. (1976). Culture of human malaria parasites Plasmodium falciparum. *Nature* 263, 767–769. doi:10.1038/263767a0.

Heinberg, A., and Kirkman, L. (2015). The molecular basis of antifolate resistance in Plasmodium falciparum: looking beyond point mutations. *Ann N Y Acad Sci* 1342, 10–18. doi:10.1111/nyas.12662.

Heinberg, A., Siu, E., Stern, C., Lawrence, E. A., Ferdig, M. T., Deitsch, K. W., et al. (2013). Direct evidence for the adaptive role of copy number variation on antifolate susceptibility in Plasmodium falciparum. *Molecular microbiology* 88, 702–712. doi:10.1111/mmi.12162.

Herman, J. D., Rice, D. P., Ribacke, U., Silterra, J., Deik, A. A., Moss, E. L., et al. (2014). A genomic and evolutionary approach reveals non-genetic drug resistance in malaria. *Genome Biol* 15, 511–511. doi:10.1186/PREACCEPT-1067113631444973.

Ho, S. S., Urban, A. E., and Mills, R. E. (2020). Structural variation in the sequencing era. *Nature Reviews Genetics* 21, 171–189. doi:10.1038/s41576-019-0180-9.

Huckaby, A. C., Granum, C. S., Carey, M. A., Szlachta, K., Al-Barghouthi, B., Wang, Y.-H., et al. (2018). Complex DNA structures trigger copy number variation across the Plasmodium falciparum genome. *Nucleic Acids Research* 47, 1615–1627. doi:10.1093/nar/gky1268.

Hyde, J. E. (2007). Drug-resistant malaria - an insight. *FEBS J* 274, 4688–4698. doi:10.1111/j.1742-4658.2007.05999.x.

Kidgell, C., Volkman, S. K., Daily, J., Borevitz, J. O., Plouffe, D., Zhou, Y., et al. (2006). A Systematic Map of Genetic Variation in Plasmodium falciparum. *PLOS Pathogens* 2, e57. doi:10.1371/journal.ppat.0020057.

Kosugi, S., Momozawa, Y., Liu, X., Terao, C., Kubo, M., and Kamatani, Y. (2019). Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. *Genome Biology* 20, 117. doi:10.1186/s13059-019-1720-5.

Kumar, S., Li, X., McDew-White, M., Reyes, A., Delgado, E., Sayeed, A., et al. (2021). Bulk segregant analysis reveals environment × genotype interactions determining malaria parasite growth. *bioRxiv*, 2020.09.12.294736. doi: 10.1101/2020.09.12.294736.

Kümpornsin, K., Modchang, C., Heinberg, A., Ekland, E. H., Jirawatcharadech, P., Chobson, P., et al. (2014). Origin of Robustness in Generating Drug-Resistant Malaria Parasites Single-

molecule nanopore sequencing reveals extreme target copy number heterogeneity in arylomycin-resistant mutants. *Molecular Biology and Evolution* 31, 1649–1660. doi:10.1093/molbev/msu140.

Layer, R. M., Chiang, C., Quinlan, A. R., and Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology* 15, R84. doi:10.1186/gb-2014-15-6-r84.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi:10.1093/bioinformatics/bty191.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

Liu, S., Huckaby, A. C., Brown, A. C., Moore, C. C., Burbulis, I., McConnell, M. J., et al. (2021). Single-cell sequencing of the small and AT-skewed genome of malaria parasites. *Genome Medicine* 13, 75. doi:10.1186/s13073-021-00889-9.

Manary, M. J., Singhakul, S. S., Flannery, E. L., Bopp, S. E. R., Corey, V. C., Bright, A. T., et al. (2014). Identification of pathogen genomic variants through an integrated pipeline. *BMC Bioinformatics* 15, 63. doi:10.1186/1471-2105-15-63.

Mandt, R. E. K., Lafuente-Monasterio, M. J., Sakata-Kato, T., Luth, M. R., Segura, D., Pablos-Tanarro, A., et al. (2019). In vitro selection predicts malaria parasite resistance to dihydroorotate dehydrogenase inhibitors in a mouse infection model. *Science Translational Medicine* 11, eaav1636. doi:10.1126/scitranslmed.aav1636.

Maniatis, T., Sambrook, J., and Fritsch, E. F. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

McDaniels, J. M., Huckaby, A. C., Carter, S. A., Lingeman, S., Francis, A., Congdon, M., et al. (2021). Extrachromosomal DNA amplicons in antimalarial-resistant Plasmodium falciparum. *Mol Microbiol* 115, 574–590. doi:10.1111/mmi.14624.

Menard, D., and Dondorp, A. (2017). Antimalarial Drug Resistance: A Threat to Malaria Elimination. *Cold Spring Harb Perspect Med* 7, a025619. doi:10.1101/cshperspect.a025619.

Miles, A., Iqbal, Z., Vauterin, P., Pearson, R., Campino, S., Theron, M., et al. (2016). Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. *Genome Res* 26, 1288–1299. doi:10.1101/gr.203711.115.

Nair, S., Miller, B., Barends, M., Jaidee, A., Patel, J., Mayxay, M., et al. (2008). Adaptive Copy Number Evolution in Malaria Parasites. *PLOS Genetics* 4, e1000243. doi:10.1371/journal.pgen.1000243.

539 Oduola, A. M. J., Weatherly, N. F., Bowdre, J. H., and Desjardins, R. E. (1988). Plasmodium
540 falciparum: Cloning by single-erythrocyte micromanipulation and heterogeneity in vitro.
541 *Experimental Parasitology* 66, 86–95. doi:10.1016/0014-4894(88)90053-7.

542 Osei, M., Ansah, F., Matrevi, S. A., Asante, K. P., Awandare, G. A., Quashie, N. B., et al.
543 (2018). Amplification of GTP-cyclohydrolase 1 gene in Plasmodium falciparum isolates with the
544 quadruple mutant of dihydrofolate reductase and dihydropteroate synthase genes in Ghana.
545 *PLOS ONE* 13, e0204871. doi:10.1371/journal.pone.0204871.

546 Oxford Nanopore Technologies qcat (version 1.1.0). Available at:
547 https://github.com/nanoporetech/qcat [Accessed June 1, 2021].

548 Palmer, M. J., Deng, X., Watts, S., Krilov, G., Gerasyuto, A., Kokkonda, S., et al. (2021). Potent
549 Antimalarials with Development Potential Identified by Structure-Guided Computational
550 Optimization of a Pyrrole-Based Dihydroorotate Dehydrogenase Inhibitor Series. *Journal of*
551 *Medicinal Chemistry* 64, 6085–6136. doi:10.1021/acs.jmedchem.1c00173.

552 Phillips, M. A., Gujjar, R., Malmquist, N. A., White, J., El Mazouni, F., Baldwin, J., et al.
553 (2008). Triazolopyrimidine-based dihydroorotate dehydrogenase inhibitors with potent and
554 selective activity against the malaria parasite Plasmodium falciparum. *J Med Chem* 51, 3649–
555 3653. doi:10.1021/jm8001026.

556 Phillips, M. A., and Rathod, P. K. (2010). Plasmodium dihydroorotate dehydrogenase: a
557 promising target for novel anti-malarial chemotherapy. *Infect Disord Drug Targets* 10, 226–239.
558 doi:10.2174/187152610791163336.

559 Ribacke, U., Mok, B. W., Wirta, V., Normark, J., Lundeberg, J., Kironde, F., et al. (2007).
560 Genome wide gene amplifications and deletions in Plasmodium falciparum. *Molecular and*
561 *Biochemical Parasitology* 155, 33–44. doi:10.1016/j.molbiopara.2007.05.005.

562 Rich, S. M., Leendertz, F. H., Xu, G., LeBreton, M., Djoko, C. F., Aminake, M. N., et al. (2009).
563 The origin of malignant malaria. *Proc Natl Acad Sci U S A* 106, 14902–14907.
564 doi:10.1073/pnas.0907740106.

565 Rugbjerg, P., Dyerberg, A. S. B., Quainoo, S., Munck, C., and Sommer, M. O. A. (2021). Short
566 and long-read ultra-deep sequencing profiles emerging heterogeneity across five platform
567 Escherichia coli strains. *Metabolic Engineering* 65, 197–206. doi:10.1016/j.ymben.2020.11.006.

568 Salcedo-Sora, J. E., Ochong, E., Beveridge, S., Johnson, D., Nzila, A., Biagini, G. A., et al.
569 (2011). The Molecular Basis of Folate Salvage in Plasmodium falciparum:
570 CHARACTERIZATION OF TWO FOLATE TRANSPORTERS*. *Journal of Biological*
571 *Chemistry* 286, 44659–44668. doi:10.1074/jbc.M111.286054.

572 Sedlazeck, F. J., Lee, H., Darby, C. A., and Schatz, M. C. (2018a). Piercing the dark matter:
573 bioinformatics of long-range sequencing and mapping. *Nature Reviews Genetics* 19, 329–346.
574 doi:10.1038/s41576-018-0003-4.

575 Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., et al.
576 (2018b). Accurate detection of complex structural variations using single-molecule sequencing.
577 *Nat Methods* 15, 461–468. doi:10.1038/s41592-018-0001-7.

578 Treangen, T. J., and Salzberg, S. L. (2012). Repetitive DNA and next-generation sequencing:
579 computational challenges and solutions. *Nature Reviews Genetics* 13, 36–46.
580 doi:10.1038/nrg3117.

581 Valenciano, A. L., Fernández-Murga, M. L., Merino, E. F., Holderman, N. R., Butschek, G. J.,
582 Shaffer, K. J., et al. (2019). Metabolic dependency of chorismate in Plasmodium falciparum
583 suggests an alternative source for the ubiquinone biosynthesis precursor. *Scientific Reports* 9,
584 13936. doi:10.1038/s41598-019-50319-5.

585 Vembar, S. S., Seetin, M., Lambert, C., Nattestad, M., Schatz, M. C., Baybayan, P., et al. (2016).
586 Complete telomere-to-telomere de novo assembly of the Plasmodium falciparum genome
587 through long-read (>11 kb), single molecule, real-time sequencing. DNA Res 23, 339–351.
588 doi:10.1093/dnares/dsw022.

589 Walliker, D., Quakyi, I. A., Wellems, T. E., McCutchan, T. F., Szarfman, A., London, W. T., et
590 al. (1987). Genetic Analysis of the Human Malaria Parasite Plasmodium falciparum. *Science*
591 236, 1661–1666. doi:10.1126/science.3299700.

592 Wang, P., Brobey, R. K. B., Horii, T., Sims, P., Hyde, J. E., Schapira, A., et al. (1986).
593 Utilization of exogenous folate in the human malaria parasite Plasmodium falciparum and its
594 critical role in antifolate drug synergy The Susceptibility of Plasmodium Falciparum to
595 Sulfadoxine and Pyrimethamine: Correlation of in Vivo and in Vitro Results Synergistic
596 Antimalarial Activity of Pyrimethamine and Sulfadoxine against Plasmodium falciparum In
597 Vitro. *Molecular Microbiology* 32, 239–245. doi:10.4269/ajtmh.1986.35.239.

598 Watkins, W. M., Sixsmith, D. G., Chulay, J. D., and Spencer, H. C. (1985). Antagonism of
599 sulfadoxine and pyrimethamine antimalarial activity in vitro by p-aminobenzoic acid, p-
600 aminobenzoylglutamic acid and folic acid. *Molecular and Biochemical Parasitology* 14, 55–61.
601 doi:10.1016/0166-6851(85)90105-7.

602 Wellems, T. E., Panton, L. J., Gluzman, I. Y., do Rosario, V. E., Gwadz, R. W., Walker-Jonah,
603 A., et al. (1990). Chloroquine resistance not linked to mdr-like genes in a Plasmodium
604 falciparum cross. *Nature* 345, 253–255. doi:10.1038/345253a0.

605 Wilson, C. M., Volkman, S. K., Thaithong, S., Martin, R. K., Kyle, D. E., Milhous, W. K., et al.
606 (1993). Amplification of pfmdr1 associated with mefloquine and halofantrine resistance in
607 Plasmodium falciparum from Thailand. *Molecular and Biochemical Parasitology* 57, 151–160.
608 doi:10.1016/0166-6851(93)90252-S.

609 Ye, J., McGinnis, S., and Madden, T. L. (2006). BLAST: improvements for better sequence
610 analysis. *Nucleic Acids Research* 34, W6–W9. doi:10.1093/nar/gkl164.

611