Mitochondrial heteroplasmy is responsible for Atovaquone drug resistance in *Plasmodium falciparum*

Sasha Siegel [1,2], Andrea Rivero [3], Swamy R. Adapa[3], ChengQi Wang [3], Roman Manetsch [4], Rays H.Y. Jiang[3*] and Dennis E. Kyle[2,3*]

Author affiliations

[1]Department of Molecular Medicine, College of Medicine, University of South Florida, Tampa, Florida, USA; [2]Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA, USA; [3]Department of Global Health, College of Public Health, University of South Florida, Tampa, Florida, USA; [4]College of Science and the Bouvé College of Health Sciences, Northeastern University, Boston, USA

Keywords

Mitochondria, Heteroplasmy, Malaria, Atovaquone, Drug resistance, *Plasmodium falciparum*

*Corresponding Authors: Dennis E. Kyle (dennis.kyle@uga.edu) and Rays H.Y. Jiang(Jiang2@health.usf.edu)

Abstract

Malaria is the most significant parasitic disease affecting humans, with 212 million cases and 429,000 deaths in 2015[1], and resistance to existing drugs endangers the global malaria elimination campaign. Atovaquone (ATO) is a safe and potent antimalarial drug that acts on cytochrome *b* (cyt. *b*) of the mitochondrial electron transport chain (mtETC) in *Plasmodium falciparum*, yet treatment failures result in resistance-conferring SNPs in cyt. *b*. Herein we report that rather than the expected *de novo* selection of resistance, previously unknown mitochondrial diversity is the genetic mechanism responsible for resistance to ATO, and potentially other cyt. *b* targeted drugs. We found that *P. falciparum* harbors cryptic cyt. *b.* Y268S alleles in the multi-copy (~22 copies) mitochondrial genome prior to drug treatment, a phenomenon known as mitochondrial heteroplasmy. Parasites with cryptic Y268S alleles readily evolve into highly resistant parasites with >95% Y268S copies under in vitro ATO selection. Further we uncovered high mitochondrial diversity in a global collection of 1279 genomes in which heteroplasmic polymorphisms were >3-fold more prevalent than homoplasmic SNPs. Moreover, significantly higher mitochondrial genome copy number was found in Asia (e.g., Cambodia) versus Africa (e.g., Ghana). Similarly, ATO drug selections *in vitro* induced >3-fold mitochondrial copy number increases in ATO resistant lines. Hidden mitochondrial diversity is a previously unknown mechanism of antimalarial drug resistance and characterization of mitochondrial heteroplasmy will be of paramount importance in combatting resistance to antimalarials targeting the electron transport chain.

Atovaquone (ATO), a napthoquinone, and the pyridones[2-4] acridones[5], acridinediones[6-8], tetrahydroacridines[6], and the 4(1H)-quinolones[9-12] potently inhibit the cytochrome *bc1* complex of the mitochondrial electron transport chain (mtETC) within the mitochondria, with disruption of pyrimidine biosynthesis and collapse of mitochondrial membrane potential leading to parasite death[13]. The first ATO treatment failures were observed in the Phase II clinical trials between 1991-1994 in Thailand[14]. These studies demonstrated that ATO monotherapy resulted in clinical treatment failures and subsequent recrudescence of infection, prompting the use of ATO in combination with proguanil (AP) for malaria prophylaxis and treatment. The clinical experience with ATO and numerous *in vitro* drug selection studies led to a hypothesis that ATO resistance arises readily *de* novo following treatment[15-17], however *in vitro* data are incongruent with clinical data of a single Y268S/N/C SNP as the genetic mechanism in two important aspects. First, clinical treatment failures are most commonly linked to an amino acid substitution at position Y268 in cyt. *b*[18] but *in vitro* drug pressure selects for a variety of different mutations except Y268 : M133I, M133V, P275T, K272R, G280D, L283I, V284K, L144S and F267V[18-20] (Extended Data Table 1). Second, drug susceptibility studies with ATO resistant Y268S mutants demonstrate a broad range of potency rather than a dichotomous response that would be expected from a single SNP. By using paired parasites collected upon patient admission and subsequent treatment failure from Phase II trials of ATO (Extended Data Table 2), we discovered three distinct *in vitro* ATO resistance phenotypes (Table 1). A recrudescent isolate (TM90-C6B) initially typed as Y268 (WT) exhibited low level ATO resistance, other isolates with Y268S from recrudescent isolates (e.g., TM90-C2B) demonstrated

moderate resistance to ATO and myxothiazol, and two isolates from ATO-pyrimethamine treatment failures showed extreme resistance (TM92-C1086 and TM92-C1088) (Table 1). Surprisingly, the extreme ATO resistance phenotype produced greatly reduced susceptibility to a broad range of mtETC inhibitors (Table 1), even though the parasites expressed the common Y268S/N SNPs and no apparent resistance-associated SNPs in candidate mtETC encoding genes (Extended Data Table 3). These data demonstrate that more than a single cyt. *b* SNP mediates the different ATO resistance phenotypes.

The mitochondria of asexual erythrocytic stage *P. falciparum* exist as a single organelle with a linear, non-recombining, tandemly-repeated 6-kb genome with approximately 22 copies present in each parasite, and is remarkably well-conserved compared to the nuclear and apicoplast genomes[21]. These data suggest there are inherent mechanisms in place to conserve the mitochondrial genome that are exclusive to the mitochondria and its three encoded genes: cyt*. b,* cytochrome c oxidase subunit I (*cox I*), and cytochrome c oxidase subunit III (*cox III*)[21]. The existence of multiple mtDNA copies creates the potential for mitochondrial allelic heterogeneity within a single parasite as well as the mtDNA pool at the population level, a phenomenon known as mitochondrial heteroplasmy[22]. Therefore, we hypothesized mitochondrial heteroplasmy as a genetic mechanism for ATO resistance and broader mitochondrial diversity.

If heteroplasmy is a source of mitochondrial diversity, parasites of different genetic backgrounds could possess variant SNPs at very low levels not detected by conventional Sanger sequence genotyping methods. To assess presence of

93 heteroplasmic SNPs that confer ATO resistance, we used a pyrosequencing assay for
94 quantitative analysis of cyt. *b*[23] to quantify the Y268S allele (Figure 1A) and applied it to
95 the earliest cryopreserved *P. falciparum* isolates from the Phase II studies (Extended
96 Data Table 1). These data indicate the presence of cryptic mutant alleles present in
97 admission parasite isolates (e.g., TM90-C2A and TM90-C40B2), with ~1-2% mutant
98 Y268S frequency detected (Figure 1B). Similarly, ATO resistant isolates (e.g.,TM90-
99 C2B and TM90-C50B5) from clinical failures possessed extremely high copies of
100 Y268S. The TM90-C6B isolate had low-level Y268S presence (Figure 1B); interestingly
101 this isolate was from a clinical failure initially identified as WT Y268 cyt. *b* (Extended
102 Data Table 1). In contrast, no Y268S heteroplasmy was found in *P. falciparum* clones
103 unable to develop ATO resistance (e.g., D6) [24] or found to develop alternative
104 cytochrome *b* mutations (e.g., W2; Extended Data Table 2), [18,20] confirming that the
105 observed Y268S allele is only present at low levels in some parasite mitochondrial
106 haplotypes.
107 To independently investigate the novel parasite heteroplasmy genetics, we next
108 studied the same set of isolates by Illumina sequencing with very deep coverage
109 (10,000-30,000x coverage of mtDNA) and determined the Y268S allele frequency (AF)
110 (Figure 2A). These data confirmed the presence of low-level Y268S heteroplasmy in *P.*
111 *falciparum* isolates prior to treatment and in a 2008 isolate from Cambodia (PL08-025).
112 The deep sequencing data for recrudescent ATO resistant isolates were consistent with
113 the pyrosequencing analysis, demonstrating very high, yet not purifying levels of Y268S
114 selection in the mtDNA. Interestingly, ATO resistant TM93-C1090 possessed low-level
115 Y268S although the majority of mitochondrial alleles were Y268N. The corresponding

admission isolate for this parasite (TM93-C0151) only showed low-level Y268S. It is plausible that this isolate had *de novo* selection of Y268N or possessed undetectable Y268N (< 0.5% AF) pre-treatment.

Overall, these findings sharply contrast with previous reports of *de novo* Y268S mutant selection in patients treated with ATO, since Y268S heteroplasmy demonstrates mutant alleles already were present in admission isolates prior to ATO exposure, but it required very high coverage, PCR-free Illumina sequencing of the mtDNA to uncover these low level heteroplasmic alleles. Therefore, we next aimed to experimentally demonstrate the potential for ATO resistance to result from low-level Y268S heteroplasmy *in vitro,* which to our knowledge, has never been successfully generated in prior studies. We selected for resistance to ATO in *P. falciparum* clones isolated from patients prior to ATO treatment (TM90-C2A and TM90-C40B2) and a 2008 clone from Cambodia (ARC08-88-8A; Figure 1C); each of these expressed low level Y268S heteroplasmy in pyrosequencing and deep sequencing analysis. Parasites were seeded at 10^8 per flask and exposed to 10x EC_{50} concentrations of ATO, and within 15-27 days all generated resistance with cyt. *b* Y268S mutations (Figures 1C and 2A). We then investigated if the Y268S mutation could be induced by another cyt. *b* acting drug, menoctone (MEN), a 2-hydroxy-3-(8-cyclohexyloctyl)-1,-4-naphthoquinone thought to have the same mechanism of action as ATO[25]. MEN at 10x EC_{50} (1.5 µM) selected for Y268S in TM90-C2A, but W2 developed the common M133I mutation associated with ATO exposure in other lab strains (Extended Data Figure 3) [18,20,26]. High coverage deep-sequencing results confirmed 99% Y268S and M133I frequencies in these drug selected mutants (Figure 2A). The MEN resistance phenotype of C2A+10xMEN-1A

displayed high-grade ATO resistance with an EC_{50} of 54 μM, similar to that of TM92-C1086. Conversely, W2+10xMEN-1A expressed moderate resistance (EC_{50} = 66 nM), presumably reflective of fitness differences between cyt. *b* Y268S and M133I. These data are consistent with other M133I selected mutants[18] (Extended Data Table 2). Obtaining the clinically relevant Y268S mutation in an *in vitro* setting using multiple drugs shows that pre-existing, low-level heteroplasmic SNPs are responsible for the development of resistance in these parasites.

P. falciparum dihydroorotate dehydrogenase (DHODH) is the only essential enzyme in the mtETC that is required to synthesize pyrimidine precursors, and we hypothesized that parasites with resistance to both DHOD and cyt. *b* inhibitors could recapitulate the extreme resistance phenotype. DSM1, a potent inhibitor of DHODH, induces mutations in DHODH as well as increased copies of the DHODH gene[19,27]. Unsurprisingly, 10x EC_{50} DSM1 pressure in the ATO sensitive TM90-C2A background readily generated DSM1 resistant parasites with copy number amplification in DHODH (Extended Data Table 4). However, there were two unusual selections in this background, with C2A+10xDSM-2B being a mixed genotype of 30% Y268S and 8% Y268N, a tri-allelic heteroplasmic parasite being selected by DSM1 alone (Figure 2A). In addition, this parasite was a partial R265G mutant in DHODH (34%); this mutation was only observed in one other DSM1 selected parasite from TM90-C2A (C2A+10xDSM1-3B), that was predominantly an R265G mutant (97%). This raises the possibility that DSM1 exposure can induce increased heteroplasmy in cyt. *b*., although in an unstable and inefficient manner.

In vitro DSM1 drug resistance selection studies in an ATO resistant background

yielded intriguing results on the role of heteroplasmy and mitochondrial copy number. DSM1 drug selections with ATO resistant TM90-C2B were not successful at 10x EC_{50} (1.5 uM) concentrations, with three failed attempts (Extended Data Figure 4). Using lower DSM1 concentrations (2x EC_{50} = 300 nM) made resistance development possible. Deep-sequencing of C2B+10xDSM1 selected lines (Figure 2A) revealed that parasites thawed from earliest cryopreserved line and expanded for sequencing (45 days of total DSM1 exposure) largely lost their initial resistance to ATO during DSM1 exposure, with parasites gradually exchanging their Y268S allele in favor of copy number amplifications in DHODH, implying that these resistance mechanisms are antagonistic at low concentrations of DSM1, but completely incompatible at high concentrations. The initial ability to survive both DSM1 and atovaquone pressure could consist of immediate reactions to respirational stress, where DSM1 induces DHODH amplifications quickly and proportionately to DSM1 exposure, and Y268S is cryptically present and able to increase allele numbers at any time. Both of these mechanisms are attractive from the metabolic standpoint in that they are readily adaptable and can be appropriately tuned for maximum fitness. Similarly, the increased mitochondrial copy number observed with some selections (e.g, ARC08-88; Figure 2C) could contribute to an enhanced stress response and high-grade resistance phenotype. The initial phenotypic response and resultant fading high-grade ATO/DSM1 resistance seen in ATO, MEN, and DSM1 selections can all be attributed to a combination of these genetic plasticity features, leading to metabolic plasticity with clear advantages for the parasite.

We next sought to estimate the worldwide frequency and distribution of mitochondrial heteroplasmy. We analyzed the publicly available *P. falciparum* data from

the MalariaGEN Pf3k project (www.malariagen.net/pf3k)[28] to uncover heteroplasmic mitochondrial diversity globally. After removing samples of multiple infections, lab lines, duplicates, and progenies of genetic crosses, we calculated the copy number of mitochondria number $\hat{C}$ of 1279 genomes using the sequencing depth differences between nuclear genome and mitochondrial genomes, with sampling-computation based correction of the sequencing coverage bias impacted by GC content[29]. For validation of our method, we estimated the copy number of nuclear chromosome as controls; and obtained the median copy number value of 1.00 (Figure 2B), consistent with the haploid genome of *P. falciparum* [30]. In contrast, the median of mitochondrial copy number was 21.75 (Figure 2B). The majority of mitochondrial copy numbers from different samples ranged from 10 to 30, suggesting significant copy number variation in mitochondria globally (F-test p-value < 2.2e-16). The results are in agreement with recent reports of qPCR experiments[31] which only used *cytochrome b* to estimate copy number.

To further quantify the level of heteroplasmy in parasite populations, a statistical measurement of maximum likelihood (PL score) of SNP detection based on VCF (Variant Call Format) output in the GATK pipeline[32] (Figure 3A) was developed to detect allele frequency (AF). The lower AF value indicates lower levels of heteroplasmy, while AF = 1 indicates homoplasmy, either wild type or mutant. We scanned 1279 polymorphic isolates and observed heteroplasmy in the majority of polymorphic sites (n =1033, coverage > 200x) (Figure 3A & D). Our results showed that *P. falciparum* mt-diversity is currently underestimated by at least 3-fold (Poisson distributions, $p < 0.001$) without taking heteroplasmy into account. Geographically-specific polymorphisms exist

in both homoplasmic and heteroplasmic SNPs (Figure 3B), suggesting important phenotypes might be associated with heteroplasmy. Interestingly, Cambodian *P. falciparum* isolates had a much higher rate of heteroplasmic SNPs than isolates from Ghana (Figure 3B). In addition, the mitochondrial copy number varied geographically with the highest copy numbers found in SE Asia where antimalarial drug resistance repeatedly emerges (Figure 3C). In agreement with widespread heteroplasmy in the global population, our study of deep-sequencing Asian parasite collections also revealed heteroplasmy, with Y268S as a low frequency common allele (Figure 2A & 3E).

Finally, to provide physical evidence of Y268S heteroplasmy, we performed single molecule sequencing with Nanopore[33] on mtDNA of TM90-C2B. We obtained >500x coverage of mitochondrial genome and longest reads of >20 kb (Figures 3F and 3G). The single molecule sequencing resulted in the same levels of allele count with NGS sequencing with 90% Y268S in both systems (Figure 3H). Single reads consisting of both Y268S and references alleles were detected, thus supporting the heteroplasmy genetics of ATO resistance and mitochondrial diversity (Figure 3H).

Mitochondrial function in *P. falciparum* changes as the parasite transitions from asexual erythrocytic stages into gametocytes that are infectious to mosquitos. The increase in aerobic respiration is thought to block the transmission of cyt. *b* SNPs associated with ATO resistance[34], although other studies demonstrated successful transmission of cyt. *b* M133I mutants[25]. Regardless, our data suggest heteroplasmy is transmitted and maintained as shown in the population diversity data, since both homoplasmic and heteroplasmic SNPs could be seen circulating in local populations

(Figure 3B). Conceivably, heteroplasmy allows both neutral polymorphisms and alleles with fitness costs to be maintained at low frequencies and successfully transmitted. Additional fitness studies are required to further assess the transmission potential of the Y268S mutants with high copies of the mutant allele. The mechanism by which low-level diversity is maintained in organisms with similar mt-DNA structure/replication strategies has been studied extensively and is known as substoichiometric shifting. In plants, substoichiometric shifting produces unique subgenomic mitochondrial DNA molecules as the result of a recombination-based replication strategy, and can confer fitness advantages. It is unknown whether substoichiometric shifting can explain the maintenance of subgenomic mtDNA variants in *Plasmodium,* but it warrants further investigation.

In this study we provide multiple lines of evidence that mitochondrial heteroplasmy is the genetic mechanism underlying ATO resistance. These include *in vitro* drug selection, ATO resistance phenotype analysis, clinical isolate characterization, pyrosequencing analysis, mitochondrial copy number, and single molecule sequencing. Although current polymorphism studies represent a conservative estimate of heteroplasmy and mitochondrial copy numbers, our population analysis shows intriguing mitochondrial diversity exists in SE Asia, where resistance first emerged for many antimalarial drugs. Furthermore, increased mitochondrial copy number was observed in the same region, and also in parasites exposed to drug selection pressure *in vitro*. Our study revealed two novel aspects of malarial parasite mitochondrial genetics, i.e, heteroplasmy and copy number variations related to ATO resistance; these results will be crucial for developing other mitochondrial-targeting

drugs, as well as combating drug resistance for eventual malaria elimination.

Materials and Methods

Parasite lines and cell culture.

Admission and recrudescent parasite samples were collected from patients in the Phase II clinical trials upon admittance for treatment, and following failure of the treatment regimen (various dose regimens of atovaquone monotherapy or atovaquone/pyrimethamine combination therapy)[14]. The parasite history for paired admission and recrudescent isolates are outlined in Supplemental Table 1. Parasites were adapted to *in vitro* culturing and maintained according to the methods previously described by Trager and Jensen, with modifications first described by Webster *et al*.[35,36] Parasites were maintained at 2% hematocrit in human O+ erythrocytes in RPMI 1640 (Invitrogen) medium containing 25 mM HEPES, 28 mM $NaHCO_3$, 10% human type A positive plasma and incubated at 37°C in 5% O_2, 5% CO_2, and 90% N_2 atmospheric conditions. Cultures were sustained with media changes three times per week and kept below 5% parasitemia with sub-culturing.

Drugs and chemicals.

Atovaquone (ATOV; 2-hydroxynapthoquinone), was purchased from Sigma (St. Louis, MO). Menoctone was synthesized and purified by the Manetsch laboratory at the University of South Florida, Department of Chemistry. DSM1 was kindly provided by Pradipsinh Rathod at the University of Washington. All compounds were used following dissolution in DMSO with final solvent concentrations less than 0.5%.

Parasite EC_{50} determinations with hypoxanthine [^{3}H] incorporation assay.

The methods used were performed as described previously [24], with a modification of a 72-hour incubation period.

Selection of atovaquone or menoctone resistant parasites *in vitro*.

In order to evaluate whether the genetic cryptic heteroplasmy background of parasite strains is essential to the development of the Y268S mutation conferring atovaquone resistance, we assessed the resistance potential of admission isolate clones TM90-C2A-F6, TM90-C40B2, and ARC08-88-8A. TM90-C2A and TM90-C40 were taken from patients prior to treatment and later recrudesced with Y268S mutations in cyt. *b* following atovaquone monotherapy regimens (Supplementary Table 1). TM90-C2A, TM90-C40, and ARC08-88 were sub-cloned by limiting dilution[37] prior to any drug selections, and sub-clones TM90-C2A-F6, TM90-C40B2, and ARC08-88-8A were used for all drug selections. Sub-cloning the parasites prior to drug selections was necessary in order to provide an isogenic background as well as to maximize phenotypic stability, as many of these parasites experienced more widely fluctuating EC_{50} values to mitochondrial inhibitors (4-8 fold) than control parasites (> 3 fold). We hypothesize the phenotypic fluctuations are a function of the genotypic plasticity within the population as a whole, where successive replication rounds vary somewhat in Y268S frequency. ARC08-88 was originally obtained from the World Health Organization Global Plan Artemisinin Resistance Containment consortium, and was used to demonstrate the development of atovaquone resistance from a parasite outside the Phase II studies of atovaquone in Thailand that had cryptic Y268S heteroplasmy. TM90-C2A, TM90-C40B2, and ARC08-88-8A were grown from earliest available cryopreserves to 10^8 and seeded into 25 ml flasks in triplicate. The complete medium contained approximately

10x EC_{50} atovaquone (10 nM) or 10x EC_{50} menoctone (1.5 μM) with media changed twice per week, and split 1:2 with fresh erythrocytes every 10 days to maintain 2% hematocrit. Parasites were considered "recovered" from drug selection when parasite densities reached 2% parasitemia and sustained growth under continuous drug pressure. All parasites had the cytochrome *b* gene sequenced to look for possible mutations developed during drug pressure.

Parasite genomic DNA isolation.

P. falciparum-infected erythrocytes were treated with 0.05% saponin for 10 min and genomic DNA (gDNA) was extracted with the Qiagen DNeasy Kit according to manufacturer's protocols. Unless described otherwise, gDNA was harvested from parasites from earliest possible cryopreservation dates to best preserve phenotypes and genotypes.

PCR and cytochrome *b and* DHODH gene sequencing.

All *P. falciparum* cytochrome *b* PCR products were amplified using primers cytbFOR 5'—TGCCTAGACGTATTCCTG—3' and cytbREV 5'—GAAGCATCCATCTACAGC—3'. PCRs were amplified using Phusion HS II High-Fidelity PCR Master Mix (ThermoFisher Scientific) with ~20 ng parasite gDNA template, according to manufacturer's instructions with the following program: 98°C—30s initial denaturation step, then 35 cycles: (98°C—10s, 54°C—40s, 72°C—30s) and a final extension of 72°C for 7 min. PCR products were confirmed as a single, discrete band of 1,382 bp length on a 1% agarose gel then subsequently purified using the Qiagen PCR Purification Kit according to manufacturer's instructions. Purified PCR products were prepared for Sanger sequencing service at

Genewiz (Genewiz, South Plainfield, NJ) using the following sequencing primers: pf-cytb-SEQFOR1: 5'—GTGGAGGATATACTGTGAGTG—3', pf-cytb-SEQFOR2: 5'—TACAGCTCCCAAGCAAAC—3', pf-cytb-SEQREV1: 5'—GACATAACCAACGAAAGCAG—3', and pf-cytb-SEQREV2: 5'—GTTCCGCTCAATACTCAG—3'. The *dhodh* gene was sequenced similarly using primers described in Ross *et al*[19]. Sample sequences were analyzed and aligned using ApE (A Plasmid Editor) software, and mapped to the *Pf*-3D7 *cytochrome b* gene annotated on Plasmodb.org for mutation detection.

Sequencing to identify SNPs in mitochondrial genes. We sequenced multiple candidate genes in the mtETC to determine if additional SNPs were associated with the drug resistance spectrum phenotypes. All PCR reactions were set up similarly to the sequencing of the cytochrome *b* gene above. Candidate genes included PF3D7_0915000 (NDH2), PF3D7_0603300 (DHODH), PF3D7_0523100 (Core 1), PF3D7_093360 (Core 2), PF3D7_1462700 (cyt. c_1), PF3D7_1439400 (Rieske), PF3D7_1426900 (QCR6), PF3D7_1012300 (QCR7), and the three mt-encoded genes: MAL_MITO_3 (cyt. *b*), MAL_MITO_1 (*coxIII*), MAL_MITO_2 (*coxI*). For the primers in Extended Data Table 5, those labeled PCR FOR and PCR REV were used in PCR amplifications, and SEQ PR denotes primers used to sequence the amplification in its entirety.

Pyrosequencing of Y268S allele.

The Pyromark Q96 ID system was used for the detection of single-nucleotide polymorphism (SNP) for Y268S detection in *Pf*-cytochrome *b*, with Qiagen Pyromark Gold Q96 reagents and buffers along with streptavidin sepharose beads (GE

Healthcare). All template and reaction components were prepared according to manufacturer's protocols. Pyrosequencing primers were designed using Pyromark Assay Design Software. Primers for the initial PCR reaction were amplified with PFcytb_pyro_Biotin_FOR 5'—Biotin-ACCATGGGGTCAAATGAGTTAT—3' and PFcytb_pyro_REV 5'—AGCTGGTTTACTTGGAACAGTTTT—3' as 50 µL reactions with 25 µL 2X Phusion Hot Start II HF PCR Master Mix, 0.2 µM primer concentrations, ~10-50 ng template gDNA, brought to 50 µL total volume with nuclease-free water, with the following thermocycling conditions: initial denaturation of 98°C for 30s, 55 cycles of 98°C for 30s, 53°C for 5s, and 72°C for 8s. All parasites resulting from *in vitro* drug selections had gDNA harvested immediately following parasite recovery to 2% parasitemia, which provided insight into the early period of resistance where more extreme phenotypes were observed in EC_{50}. Subsequent PCRs were run on 1.5% agarose gels to confirm a single discrete band without excess primer present, as unconsumed primer has been shown to interact with pyrosequencing primers to contribute to a background signal in no template controls, and was minimized by using low primer concentrations and using a high cycle number to exhaust primers. The Pyromark pyrosequencing assay was performed according to manufacturer's protocols with pyrosequencing primer PFcytb_seq_assay_REV 5'—TGGAACAGTTTTTAACATTG—3'. Each parasite gDNA sample was initially amplified independently in triplicate, and had two technical replicates per reaction (25 µL PCR per pyrosequencing reaction) on the Pyromark Q96 ID for a total of at least 6 pyrosequencing runs per parasite gDNA template. Allele frequencies were analyzed by Pyromark ID software in allele quantification mode.

Pyrosequencing Y268S assay standard curve. Since all Y268S mutant genotypes still contained some small quantities of wild-type allele, we chose to use the parasite with the highest percentage Y268S mutant, TM90-C50B5 gDNA (99.52% mutant) and D6 (0% mutant) were mixed at 10% increments from 0% wild type gDNA + 100% wild type gDNA, adding in additional increment mixtures at the lower 5% and upper 95%, with 1% increments to look at the sensitivity of detection. These ratio wild type:mutant gDNA mixtures were made independently three times and then used in subsequent PCR reactions and pyrosequencing reactions to generate the standard curve in Figure 1A.

DHODH Copy Number Variation Quantitative PCR (qPCR). *Pf*-DHODH copy number was determined using the DHODH qPCR primers previously described by Guler *et al.*[27] and the LDH-T1 FOR/REV control primers from Chavchich *et al.*[38] using Brilliant II/III SYBR Green Master Mix with ROX and the Mx3005P qPCR machine (Applied Biosystems). The relative copy number of DHODH was determined for 0.1 ng of gDNA and normalized to the LDH gene using the $\Delta\Delta C_T$ method[39].

Deep sequencing analysis and data mining. For deep sequencing of clinical and *in vitro* selected parasite genomes, we ran samples on the Illumina HiSeq 3000 with the 150 cycle protocol (PCR-free) and used >1 ug purified DNA as starting material, and reached 10,000–30,000 x coverage of mtDNA. Each allele must have a minimum of 100 high quality reads (QC score > 60) mapped to the loci to be identified as a rare allele. As controls, we examined all loci with the same stringency, and our method did not recover any low frequency heteroplasmy in the vast majority of the loci (n=5954), other than a few loci including Y268S.

For computational mining of the sites with heteroplasmy in the global parasite populations, we developed a maximum likelihood based method to differentiate high confidence heteroplasmic sites from background error count. We use a statistical measurement of maximum likelihood (PL score) of SNP detection based on VCF (variant call format) output in GATK pipeline[32] to detect allele frequency (AF). First, we obtained the genotype likelihood (GL) for the given loci, with GL defined as LOG10 scaled likelihoods for all possible genotypes given the set of alleles defined in the REF and ALT fields, as specified by GATK protocols[40,41]. Then, the phred-scaled genotype likelihood (PL) score is obtained. Only the PL score of being heteroplasmic vs being homoplasmic, at a given locus, larger than 6 (i.e, Pvalue < 0.000001), is considered a true heteroplasmic site. The lower AF value indicates lower level of heteroplasmy, while AF = 1 indicates homoplasmy, either wild type or mutants.

To independently confirm our computational mining methods, we examined 100 isolates from the population data sets. We retrieved the raw data mitochondrial BAM files, and examined that the heteroplasmic sites are correctly identified by the maximum likelihood method, as shown by more than a hundred high quality reads in the mapping alignments.

We used the sequencing depth differences to estimate the copy number of query chromosome from sequencing signal. Here, we assume the PCR amplification level is same for each chromosome copy. It has been reported the copy number of nuclear chromosome is 1 during asexual stage in *Plasmodium*[30]. Therefore, the sequencing depth difference between input and reference nuclear chromosome is the estimated copy number of query chromosome (equation 1).

$$\widehat{C_{query}} = \frac{\widehat{r_{query}}}{\widehat{r_{nuclear}}} \quad (1)$$

where $\widehat{C}$ is the estimated copy number and $\hat{r}$ is the estimated sequencing signal. To find the best estimation of sequencing signal, we assume that the true sequencing reads signal of a whole specific chromosome follows a uniform distribution. Then we calculated the optimal sequencing depth with

$$\hat{r} = \arg\min \sum_{i=1}^{n} (\hat{r} - r_i)^2 \quad (2)$$

Here, $i(n = \sum i)$ is a given segment on the chromosome; and the sequencing signal r_i is calculated as

$$r_i = \frac{R_i}{L_i} \quad (3)$$

where R_i is the total reads mapped on segment i, L_i is the segment length. The final estimation $\hat{r}$ from equation 2 is

$$\hat{r} = \frac{\sum_{i=1}^{n} r_i}{n}$$

Nanopore mtDNA sequencing.

Isolation of mitochondrial DNA in TM90-C2B

Parasites were grown to 10% parasitemia in 50 mL culture volume, then synchronized with 5% sorbitol twice, spaced 4 hours apart, followed by MACS column purification to remove any residual trophozoites and schizonts so that only ring stages were collected to obtain non-replicative mtDNA molecule forms only. Red blood cells

were then lysed using 0.1% saponin in 1x PBS, washed 3x in PBS, and mitochondrial DNA was isolated from parasite material using the Abcam Mitochondrial DNA Isolation Kit using the manufacturer's instructions, with the exception of using needle passage to homogenize cells. Isolated mitochondrial DNA was used for subsequent Nanopore library preparation and sequencing.

Library Preparation and Sequencing

1 µg of parasite mitochondrial DNA (RNAse treated) was subsequently treated with protease and mtDNA eluted in 45 µl nuclease-free water (NFW). End-repair and dA-tail of mtDNA was performed by adding 7 µl Ultra II End-Prep buffer, 3 µl Ultra II End-Prep enzyme mix (NEBNext Ultra II End-Repair/dA-tailing Module, New England BioLabs), and 5 µl NFW. The mix was incubated for 5 minutes at 20 °C and 5 minutes at 65 °C using a thermocycler. The end-prep reaction cleanup was performed by adding 60 µl of resuspended AMPure XP beads and mtDNA was eluted in 31 µl NFW. A 1 µl aliquot from the elute was quantified by fluorometry (Qubit) to ensure ≥700 ng end-prepped mtDNA was retained.

Adapter ligation was performed by adding 10 µl of Adapter Mix, 2 µl HP Adapter (SQK-NSK007 Nanopore sequencing Kit, Oxford Nanopore Technologies), 50 µl NEB Blunt/TA Master Mix (NEB, cat no M0367), and 8 µl NFW to 30 µl dA-tailed mtDNA, mixing gently and incubating at room temperature for 10 minutes. 1 µl of HP Tether (SQK-NSK007) was added to the mix and incubated for 10 minutes (RT).

Library purification of the adapted and tethered mtDNA was performed by adding MyOne C1 Streptavidin Beads, incubated for 5 minutes at room temperature and resuspended the pellet in 150 µl Bead Binding Buffer (SQK-NSK007). The purified-

ligated mtDNA pellet was resuspended in 25 µl Elution Buffer (SQK-NSK007), incubated for 10 minutes at 37 °C and the 25 µl elute (Pre-sequencing Mix) was transferred into a new tube. A 1 µl aliquot was quantified by fluorometry (Qubit) to ensure ≥ 500 ng of adapted and tethered mtDNA was retained.

The MinION Flow Cell (R9) was primed twice prior to loading sample with 500 µl flow cell priming mix made of 1:1 ratio of Running Buffer with Fuel Mix 1 (RBF1) and NFW. The library for loading was prepared by adding 12 µl adapted and tethered library, 75 µl RBF1 and 63 µl NFW. Using a P-1000 tip set to 150 µl, the library was loaded into the flow cell keeping the pipette vertical. The flow cell was run on the 48□h 2D protocol.

Data Analysis

Nanopore sequencing reads were processed using the Metrichor cloud platform 2D workflow. Only the reads (Fast5 files) that passed Metrichor quality cutoffs were converted into fastq format using poretools (v0.5)[42]. The reads were aligned against the Plasmodium falciparum 3D7 (PlasmoDB v28)[43] reference genome using the BWA-MEM (v0.7.15)[44] with settings “-x ont2d”. The aligned data was used to generate a .bam file using SAMtools (v1.3.1)[45] and BCFtools (v1.3.1). Finally, alignments were visually inspected using IGV viewer[46] in an attempt to trace back the % Y268S mutation in the parasite.

Figure legends

Figure 1. Pyrosequencing assessment of Y268S allele in patient isolates and drug selection strategy. A. Y268S cytochrome *b* pyrosequencing assay standard curve reliably detects frequencies of wild-type and mutant alleles. Using wt and mutant gDNA mixtures from 0% to 100% wt, the standard curve shows the correlation between the percentage of wt DNA expected in each mixture and the pyrosequencing assay's ability to detect that percentage, with a correlation coefficient of R^2=0.9946. B. Pyrosequencing analysis of the Y268S allele in admission and recrudescent isolates from Phase II studies in Thailand. Pyrosequencing analysis of control lab strains that lack the Y268S allele indicate that there is a very low false positive detection rate using this assay. Admission isolates show low level Y268S heteroplasmy, as well as "wt" treatment failure TM90-C6B, and recrudescent parasites show high levels of the Y268S allele. C. *In vitro* drug selections with clinical isolates harboring cryptic Y268S alleles. Admission isolates from the Phase II studies of ATO in Thailand (TM90-C2A and TM90-C40B2) and isolate from Thailand in 2008 (ARC08-22-4G) all had low level Y268S heteroplasmy, and rapidly developed majority Y268S mtDNA after exposure to 10xEC50 ATO, recovering after 15-32 days.

Figure 2. Using mitochondrial deep sequencing to identify low frequency heteroplasmy sites in patient isolates and from in vitro drug selections. A. Deep sequencing shows heteroplasmy/homoplasmy status of patient admission/recrudescent isolates and *in vitro* drug selected parasites at known ATO resistance alleles. Parasites were taken from earliest cryopreserved isolates and mtDNA was sequenced at > 10,000x coverage to

preserve respective phenotypes/genotypes. B. Histogram plot shows the distribution of estimated copy number of nuclear and mitochondria genomes (yellow line indicates the median value). For 1279 *P. falciparum genome* samples, we extracted median coverage (DP) for each reported sites in chromosome 2 (*N,* median value 1.00) and mitochondria genome (*M*, median value 21.75). C. Heteroplasmy and copy number from clinical isolates and *in vitro* drug selection studies. Both Y268S heteroplasmy and copy number variations were found in *in vitro* drug selection studies.

Figure 3. Analysis of the *P. falciparum* mitochondrial genetic diversity with data from the MalariaGEN Pf3K[28] project and single molecule sequencing. A. Analysis of the heteroplasmic mitochondrial diversity in the 1279 parasite genome collection. A total of 1033 high-confidence SNPs with hundreds of mapped reads for support were studied. B. Geographical specific polymorphisms are found in both homoplasmy and heteroplasmy SNPs in Cambodia and Ghana. Cambodian isolates have higher diversity and higher mitochondrial copy numbers. C. Boxplot shows the copy number distribution of mitochondria in different regions. The regions are ranked based on the median value of copy number. The data from Nigeria is not shown here due to low sample number (n = 3). D. In the global population of 1279 parasites, number of SNPs per isolate is estimated with a Poisson distribution (Difference between homoplasmy and heteroplasmy pvalue < 0.001). E: Low level Y268S heteroplasmy is a common, low level SNP found in all patient admission isolates from the Phase II clinical trial, in contrast to high frequency, common heteroplasmic loci. F. Nanopore sequencing covers the nuclear genome at > 6x and mitochondrial genome at >500x. G Average reads

length of Nanopore sequencing is 5 kb; and the longest read is 40 kb. H. Nanopore sequencing and NGS found the same allele frequencies. Y268S hetereplasmy is shown on a 20 kb single molecule reads with two tandemly arranged mitochondrial DNA units.

References:

1 World Malaria Report 2016. (World Health Organization, 2016).
2 Bueno, J. M. *et al.* Potent antimalarial 4-pyridones with improved physico-chemical properties. *Bioorg Med Chem Lett* **21**, 5214-5218, doi:10.1016/j.bmcl.2011.07.044 (2011).
3 Yeates, C. L. *et al.* Synthesis and structure-activity relationships of 4-pyridones as potential antimalarials. *J Med Chem* **51**, 2845-2852, doi:10.1021/jm0705760 (2008).
4 Bueno, J. M. *et al.* Exploration of 4(1H)-pyridones as a novel family of potent antimalarial inhibitors of the plasmodial cytochrome bc1. *Future medicinal chemistry* **4**, 2311-2323, doi:10.4155/fmc.12.177 (2012).
5 Winter, R. W. *et al.* Evaluation and lead optimization of anti-malarial acridones. *Exp Parasitol* **114**, 47-56, doi:10.1016/j.exppara.2006.03.014 (2006).
6 Cross, R. M. *et al.* Optimization of 1,2,3,4-tetrahydroacridin-9(10H)-ones as antimalarials utilizing structure-activity and structure-property relationships. *Journal of medicinal chemistry* **54**, 4399-4426, doi:10.1021/jm200015a (2011).
7 Biagini, G. A. *et al.* Acridinediones: selective and potent inhibitors of the malaria parasite mitochondrial bc1 complex. *Mol Pharmacol* **73**, 1347-1355, doi:10.1124/mol.108.045120 (2008).
8 Raether, W. *et al.* Antimalarial activity of new floxacrine-related acridinedione derivatives: studies on blood schizontocidal action of potential candidates against P. berghei in mice and P. falciparum in vivo and in vitro. *Parasitol Res* **75**, 619-626 (1989).
9 Biagini, G. A. *et al.* Generation of quinolone antimalarials targeting the Plasmodium falciparum mitochondrial respiratory chain for the treatment and prophylaxis of malaria. *Proc Natl Acad Sci U S A* **109**, 8298-8303, doi:10.1073/pnas.1205651109 (2012).
10 Cross, R. M. *et al.* Synthesis, antimalarial activity, and structure-activity relationship of 7-(2-phenoxyethoxy)-4(1H)-quinolones. *Journal of medicinal chemistry* **54**, 8321-8327, doi:10.1021/jm200718m (2011).
11 Cross, R. M. *et al.* Endochin optimization: structure-activity and structure-property relationship studies of 3-substituted 2-methyl-4(1H)-quinolones with antimalarial activity. *Journal of medicinal chemistry* **53**, 7076-7094, doi:10.1021/jm1007903 (2010).
12 Winter, R. *et al.* Optimization of endochin-like quinolones for antimalarial activity. *Exp Parasitol* **127**, 545-551, doi:10.1016/j.exppara.2010.10.016 (2011).

13 Srivastava, I. K., Rottenberg, H. & Vaidya, A. B. Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. *J Biol Chem* **272**, 3961-3966 (1997).

14 Looareesuwan, S. *et al.* Clinical studies of atovaquone, alone or in combination with other antimalarial drugs, for treatment of acute uncomplicated malaria in Thailand. *The American journal of tropical medicine and hygiene* **54**, 62-66 (1996).

15 Schwartz, E., Bujanover, S. & Kain, K. C. Genetic Confirmation of Atovaquone-Proguanil–Resistant Plasmodium falciparum Malaria Acquired by a Nonimmune Traveler to East Africa. *Clinical Infectious Diseases* **37**, 450-451, doi:10.1086/375599 (2003).

16 Musset, L., Bouchaud, O., Matheron, S., Massias, L. & Le Bras, J. Clinical atovaquone-proguanil resistance of Plasmodium falciparum associated with cytochrome b codon 268 mutations. (2006).

17 Musset, L., Le Bras, J. & Clain, J. Parallel evolution of adaptive mutations in Plasmodium falciparum mitochondrial DNA during atovaquone-proguanil treatment. *Mol Biol Evol* **24**, 1582-1585, doi:10.1093/molbev/msm087 (2007).

18 Korsinczky, M. *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).

19 Ross, L. S. *et al.* In vitro resistance selections for Plasmodium falciparum dihydroorotate dehydrogenase inhibitors give mutants with multiple point mutations in the drug-binding site and altered growth. *J Biol Chem* **289**, 17980-17995, doi:10.1074/jbc.M114.558353 (2014).

20 Bopp, S. E. *et al.* Mitotic evolution of Plasmodium falciparum shows a stable core genome but recombination in antigen families. *PLoS Genet* **9**, e1003293, doi:10.1371/journal.pgen.1003293 (2013).

21 Preston, M. D. *et al.* A barcode of organellar genome polymorphisms identifies the geographic origin of Plasmodium falciparum strains. *Nat Commun* **5**, 4052, doi:10.1038/ncomms5052 (2014).

22 Cottrell, G., Musset, L., Hubert, V., Le Bras, J. & Clain, J. Emergence of resistance to atovaquone-proguanil in malaria parasites: insights from computational modeling and clinical case reports. *Antimicrob Agents Chemother* **58**, 4504-4514, doi:10.1128/aac.02550-13 (2014).

23 Conaway, J. L. *Detection and Epidemiological Patterns of Drug Resistance Mutations in NYS Plasmodium falciparum Clinical Data Specimens*, University at Albany, (2012).

24 Rathod, P. K., McErlean, T. & Lee, P. C. Variations in frequencies of drug resistance in Plasmodium falciparum. *Proc Natl Acad Sci U S A* **94**, 9389-9393 (1997).

25 Blake, L. D. *et al.* Menoctone Resistance in Malaria Parasites Is Conferred by M133I Mutations in Cytochrome b That Are Transmissible through Mosquitoes. LID - e00689-17 [pii] LID - 10.1128/AAC.00689-17 [doi]. (2017).

26 Schwobel, B., Alifrangis, M., Salanti, A. & Jelinek, T. Different mutation patterns of atovaquone resistance to Plasmodium falciparum in vitro and in vivo: rapid

detection of codon 268 polymorphisms in the cytochrome b as potential in vivo resistance marker. *Malar J* **2**, 5 (2003).

27 Guler, J. L. *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, doi:10.1371/journal.ppat.1003375 (2013).

28 The Pf3K Project: pilot data release 5. (2016).

29 Benjamini, Y. & Speed, T. P. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res* **40**, e72-e72, doi:10.1093/nar/gks001 (2012).

30 Gerald, N., Mahajan, B. & Kumar, S. Mitosis in the human malaria parasite plasmodium falciparum. *Eukaryotic Cell* **10**, 474-482 (2011).

31 Beghain, J. *et al.* Plasmodium copy number variation scan: gene copy numbers evaluation in haploid genomes. *Malaria journal* **15**, 206, doi:10.1186/s12936-016-1258-x (2016).

32 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303 (2010).

33 Jain, M., Olsen, H. E., Paten, B. & Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome biology* **17**, 239 (2016).

34 Goodman, C. D. *et al.* Parasites resistant to the antimalarial atovaquone fail to transmit by mosquitoes. *Science* (2016).

35 Trager, W. & Jensen, J. B. Human Malaria Parasites in Continuous Culture. *Journal of Parasitology* **91**, 484-486, doi:10.1645/0022-3395(2005)091[0484:hmpicc]2.0.co;2 (2005).

36 Webster, H. K., Boudreau, E.F., Pavanand, K., Youngvanithchit, K., Pang, L.W. Antimalarial drug susceptibility testing of Plasmodium falciparum in Thailand using a microdilution radioisotope method. *Am J Trop Med Hyg* **34**, 228-235 (1985).

37 Rosario, V. Cloning of naturally occurring mixed infections of malaria parasites. *Science* **212**, 1037 (1981).

38 Chavchich, M. *et al.* Role of pfmdr1 amplification and expression in induction of resistance to artemisinin derivatives in Plasmodium falciparum. *Antimicrob Agents Chemother* **54**, 2455-2464, doi:10.1128/aac.00947-09 (2010).

39 Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* **29**, e45-e45 (2001).

40 Van der Auwera, G. A. *et al.* From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics*, 11.10. 11-11.10. 33 (2013).

41 Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993 (2011).

42 Loman, N. J. & Quinlan, A. R. Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics* **30**, 3399-3401 (2014).

655 43 Aurrecoechea, C. *et al.* PlasmoDB: a functional genomic database for malaria
656 parasites. *Nucleic acids research* **37**, D539-D543 (2008).
657 44 Li, H. Aligning sequence reads, clone sequences and assembly contigs with
658 BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
659 45 Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics*
660 **25**, 2078-2079 (2009).
661 46 Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer
662 (IGV): high-performance genomics data visualization and exploration. *Briefings in*
663 *bioinformatics* **14**, 178-192 (2013).
664

665

Table 1.

Drug susceptibility (IC_{50}s, nM) of *Plasmodium falciparum* isolates and clones to mitochondrial electron chain inhibitors reveal three atovaquone (ATO) resistance phenotypes*

			Complex III								
	NDH2 enzyme	DHODH enzyme	Qo site (cyt *b*)		Qi site (cyt *b*)			unknown			
	HDQ	DSM-1	ATO	MYX	ANT	ELQ 300	GSK121	ICI56,780	P4Q-391	Cyt. *b* genotype	Admission/ Recrudescent
W2	55.0	52.4	0.41	10.8	165.5	2.72	3.39	0.03	11.1	WT	--
TM90-C6B	ND	24.5	109	7.45	125	7.67	36.4	0.50	33.0	WT	R
TM90-C2A	14.8	79.4	3.09	154	72.0	0.69	20.2	0.04	1.92	WT	A
TM90-C2B	146	57.4	5290	428	152	4.61	77.5	14.3	55.5	Y268S	R
TM90-C40B2	830	47.2	1.53	45.1	72.9	4.02	7.84	0.04	22.1	WT	A
TM90-C50B5	ND	ND	3940	3010	ND	6.45	122	13.4	32.4	Y268S	R
TM92-C1086	180	770	31400	4160	>18200	16300	2470	882	14160	Y268S	R
TM90-C1088	252	540	29100	3990	>18200	21000	3470	219	>20200	Y268S	R

*Atovaquone (ATO) Resistance phenotypes: low (light grey), moderate (grey), and extreme (dark grey); Drugs: HDQ – 1-hydroxy-2-dodecyl-4=(1H)-quinolone; MYX – myxothiazol; DSM-1 – triazolopyrimidine; ANT – antimycin A; GSK121 – GSK932121A, 4(1H)-pyridone; ICI56,780 – phenoxyethoxy-4(1H)-quinolone; P4Q-391 – 4(1H)-quinolone

Extended Data Table 1.

Cytochrome b mutations induced by drug selection with atovaquone (ATO) and menoctone (MEN) in vitro

Parasite	Drug	Cyt. *b* mutation(s)	Source
D6	ATO	No resistance	Rathod *et al.* 1997
3D7		M133I M133I & P275T M133I & K272R M133I & G280D L283I & V284K	Korsinczky *et al.* 2000
3D7		M133V M133I M133I & L144S F267V	Bopp *et al.* 2013
K1		M133I M133I & G280D	Schwobel *et al.* 2003
AT200		M133I M133I & L271F	Schwobel *et al.* 2003
TM90-C2A		Y268S	This study
TM90-C40B2		Y268S	This study
ARC08-88-8A		Y268S	This study
W2	MEN	M133I	This study
TM90-C2A		Y268S	This study

Extended Data Table 2.

Treatment regimens and cytochrome b (cyt b) genotypes of admission (A) and recrudescent (R) isolates of *Plasmodium falciparum* from Phase II clinical studies with atovaquone (ATO) alone or in combination with pyrimethamine (PYR)[14]

Treatment Regimen	Patient No.	Admission/Recrudescent	Isolate	Cyt. *b* mutation
ATO 750 mg q8h x 4	2	A R	C2A C2B	-- Y268S
	6	A R	C6A C6B	-- --
ATO 750 mg q8h x 21	29	A R	C40 C50	-- Y268S
	32	R	C32B	Y268N
ATO 1000 mg plus PYR 25 mg q24h x 3	210	A R	C1028 C1086	-- Y268S
	207	A R	C1051 C1090	-- Y268N
	206	R	C1088	Y268S

Extended Data Table 3.

Sanger sequence results of candidate mtETC resistance genes in patient isolates and reference clones of *P. falciparum*

	Parasite names (abbreviated)											
Gene ID	W2	D6	C2A	C2B	C6A	C6B	C40B2	C50B5	C1051	C1090	C1086	C1088
PF3D7_0915000 (NDH2)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_0603300 (DHODH)	--	--	--	--	--	--	--	--	--	--	--	--
MAL_MITO_1 (COXIII, Complex IV)	--	I239V	I239V	I239V	--	--	I239V	I239V	I239V	--	--	--
MAL_MITO_2 (COXI, Complex IV)	--	--	--	--	--	--	--	--	--	--	--	--
MAL_MITO_3 (Cyt. *b*)	--	--	--	Y268S	--	--	Y268S	Y268S	--	Y268N	Y268S	Y268S
PF3D7_0523100 (Core 1, Complex III)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_093360 (Core 2, Complex III)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_1439400 (Rieske, Complex III)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_1426900 (QCR6, Complex III)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_1012300 (QCR7, Complex III)	--	--	--	--	--	--	--	--	ND	ND	--	--
PF3D7_1462700 (Cyt. c_1)	--	--	--	--	--	--	--	--	ND	ND	ND	--

-- Indicates wild type sequence; ND – not determined

Extended Data Table 4.

Genotypes of DSM1 drug selections immediately following recovery

DSM1-selected populations	*dhodh* copy number	PFF_0160C (*dhodh*)	MALMITO_3 (cyt. *b*)	MALMITO_1 (*coxIII*)
C2A+10xDSM1-1A	3	WT	WT	I239V
C2A+10x DSM1-2B	1	R265G	Y268S/WT	I239V
C2B+2xDSM-1A	2	WT	Y268S/WT	I239V
C2B+2xDSM1-1B	2	WT	Y268S/WT	I239V
C2B+2xDSM1-3B	1.5	WT	Y268S/WT	I239V
C2B+2xDSM1-2C	1	WT	Y268S/WT	I239V

Extended Data Table 5.

PCR Primers and Programs Used in mtETC Sequencing

PCR FOR PCR REV SEQ PR SEQ PR SEQ PR SEQ PR	0120PCRFOR 0120PCRREV 0120FORINT 0120E1REV 0120E1FOR 0120REVINT	5'—CTAACCGCGTTTGTCCTAACC—3' 5'—CTGGTGGTATCGTGTCATC—3' 5'—ATTCAGCTCCAAGCCTGTTC—3' 5'—TAAGAGCACCATATGAGAGATGG—3' 5'—CAAGGAGATATAGAAGGATGTTAAGAGGAAC—3' 5'—AGCAGCCATACCTCATTC—3'	2733 bp	94°-1:00 32 cycles: 94°-0:20 52°-0:20 58°-4:00
PCR FOR PCR REV SEQ PR SEQ PR SEQ PR SEQ PR	0735PCRFOR 0735PCRREV 0735FORINT 0735REVINT 0735FORINT2 0735REVINT2	5'—ACCCTAATTCGCCTGCTC—3' 5'—GGTTCCTCCAAATCACATGC—3' 5'—GTTCAGGAAATGTGGACAAG—3' 5'—CAAATGGTATGGGCGTCCTT—3' 5'—TATGGTCTTCTTATCTGGGCTAGTGG—3' 5'—CATGTAGCTGTTGTAGGAGGAGGTC—3'	5038 bp	94°-1:00 32 cycles: 94°-0:20 55°-0:20 58°-4:00
PCR FOR PCR REV SEQ PR SEQ PR	0248PCRFOR 0248PCRREV 0248FORINT 0248REVINT	5'—CTTGACACATTCACCTGAAC—3' 5'—ACAGTACATTCTTGTGGGAC—3' 5'—GCAGTCAAATGTGTAAGACCAG—3' 5'—ACAGTACATTCTTGTGGGAC—3'	2707 bp	94°-1:00 32 cycles: 94°-0:20 52°-0:20 58°-4:00
PCR FOR PCR REV SEQ PR SEQ PR SEQ PR SEQ PR	1155PCRFOR 1155PCRREV 1155REVINT 1155E1REV 1155E2REV 1155E2FOR	5'—AGCATAGCACTGAGAACAAG—3' 5'—ACGGACAAGAGTTGATACTG—3' 5'—GAACCATCGAATACCTCTG—3' 5'—GCTTCACGTTTACCTATCGAACAC—3' 5'—GTTCTTAAATGAGATAAATGTGCCGTACTATG**—3''** 5'—GAATCATGTATGGCCTTTAGTACTCAGCATTCAG—3'	3209 bp	94°-1:00 32 cycles: 94°-0:20 55°-0:20 58°-3:30
PCR FOR PCR REV SEQ PR SEQ PR SEQ PR SEQ PR	1625PCRFOR 1625PCRREV 1625FORINT 1625FORINT2 1625REVINT 1625REVINT2	5'—TCCTGCCCTCTTCATTTG—3' 5'—CGAGCAATACAAACGGAC—3' 5'—CGAGCAATACAAACGGAC—3' 5'—TATGTGCCGTTGGTGATG—3' 5'—TGATGACTCAGGTCCAAATG—3' 5'—TCAGTACATCGACCTCAG—3'	3511 bp	94°-1:00 32 cycles: 94°-0:20 55°-0:20 58°-3:30
PCR FOR PCR REV SEQ PR	MM2PCRFOR MM2PCRREV MM2REVINT	5'—CTGGCCTACACTATAAGAAC—3' 5'—GAGAATTATGGAGTGGATGGTG—3' 5'—GGTATGATACACAGCTCTTC—3'	1809 bp	98°-0:30 32 cycles: 98°-0:10 53°-0:30 72°-7:00 72°- 0:30FE
PCR FOR PCR REV	MM1PCRFOR MM1PCRREV	5'—TGCGATGAGACGACATGGAG—3' 5'—GCTATCAAATGGCGAGAAGGGAAG—3'	1008 bp	98°-0:30 32 cycles: 98°-30s 61°-0:30 72°-0:15 72°- 3:00FE
PCR FOR PCR REV SEQ PR SEQ PR	MM3PCRFOR MM3PCRREV MM3-2B-REV MM3-2A-FOR	5'—TGCCTAGACGTATTCCTG—3' 5'—GCTGTAGATGGATGCTTC—3' 5'—CTGAGTATTGAGCGGAAC—3' 5'—GTGGAGGATATACTGTGAGTG—3'	1382 bp	98°-0:30 32 cycles: 98°-0:10 54°-0:40 72°-0:30 72°- 7:00FE
PCR FOR PCR REV SEQ PR	DHODHFOR DHODHREV DHODHINT	5'—GATCCCTAGGATGATCTCTAAATTGAAACCTCAATTTATG—3' 5'—GATACTCGAGTTAACTTTTGCTATGCTTTCGGCCAATG—3' 5'-CATTATTTGGATTATATGGGTTTTTTTGAATCTTATAATCCTG—3'	1774 bp	94°-1:00 32 cycles: 94°-0:20 55°-0:20 58°-3:30
PCR FOR PCR REV SEQ PR	0597PCRFOR 0597PCRREV 0597FORINT 0597REVINT2	5'—AAAAATGGCTGGTGGGGGAG—3' 5'—CCAACGTCCAAAAATAAGAAACTAATCCA—3' 5'—TTCCTTGTCCACTGTGTAG—3' 5'—GGCAAAGATTCTTCTGGAC—3'	1428 bp	94°-1:00 32 cycles: 94°-0:20 56°-0:20 58°-2:10

Fig 1

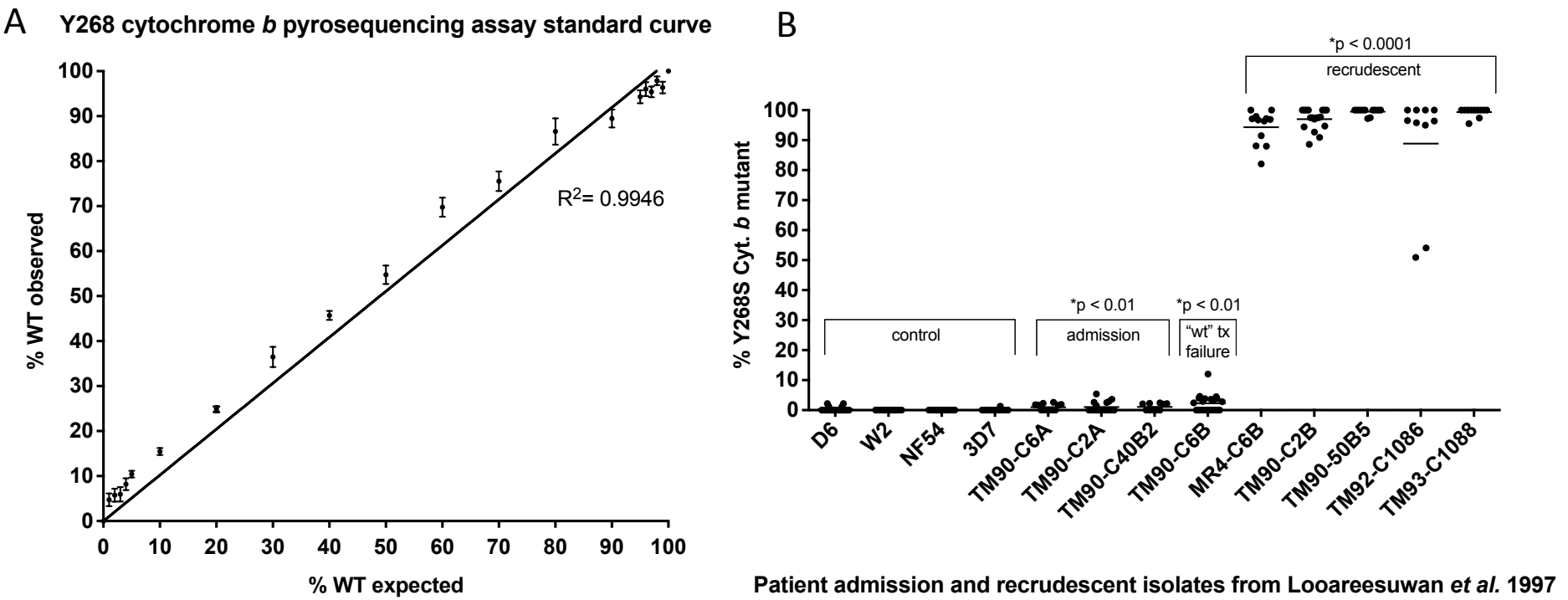

A
Y268 cytochrome b pyrosequencing assay standard curve
% WT observed
% WT expected
R2= 0.9946
B
% Y268S Cyt. b mutant
*p < 0.0001
recrudescent
*p < 0.01
admission
*p < 0.01
"wt" tx failure
control
D6
W2
NF54
3D7
TM90-C6A
TM90-C2A
TM90-C40B2
TM90-C6B
MR4-C6B
TM90-C2B
TM90-50B5
TM92-C1086
TM93-C1088
Patient admission and recrudescent isolates from Looareesuwan et al. 1997
C
TM90-C2A-F6
10x EC50 ATOV
5/9 flasks recovered
(10^8 parasites)
27-32 days
Y268S
TM90-C40B2
10x EC50 ATOV
9/9 flasks recovered
(10^8 parasites)
24-28 days
Y268S
ARC08-88-8A
10x EC50 ATOV
3/3 flasks recovered
(10^8 parasites)
15-17 days
Y268S

Fig 2

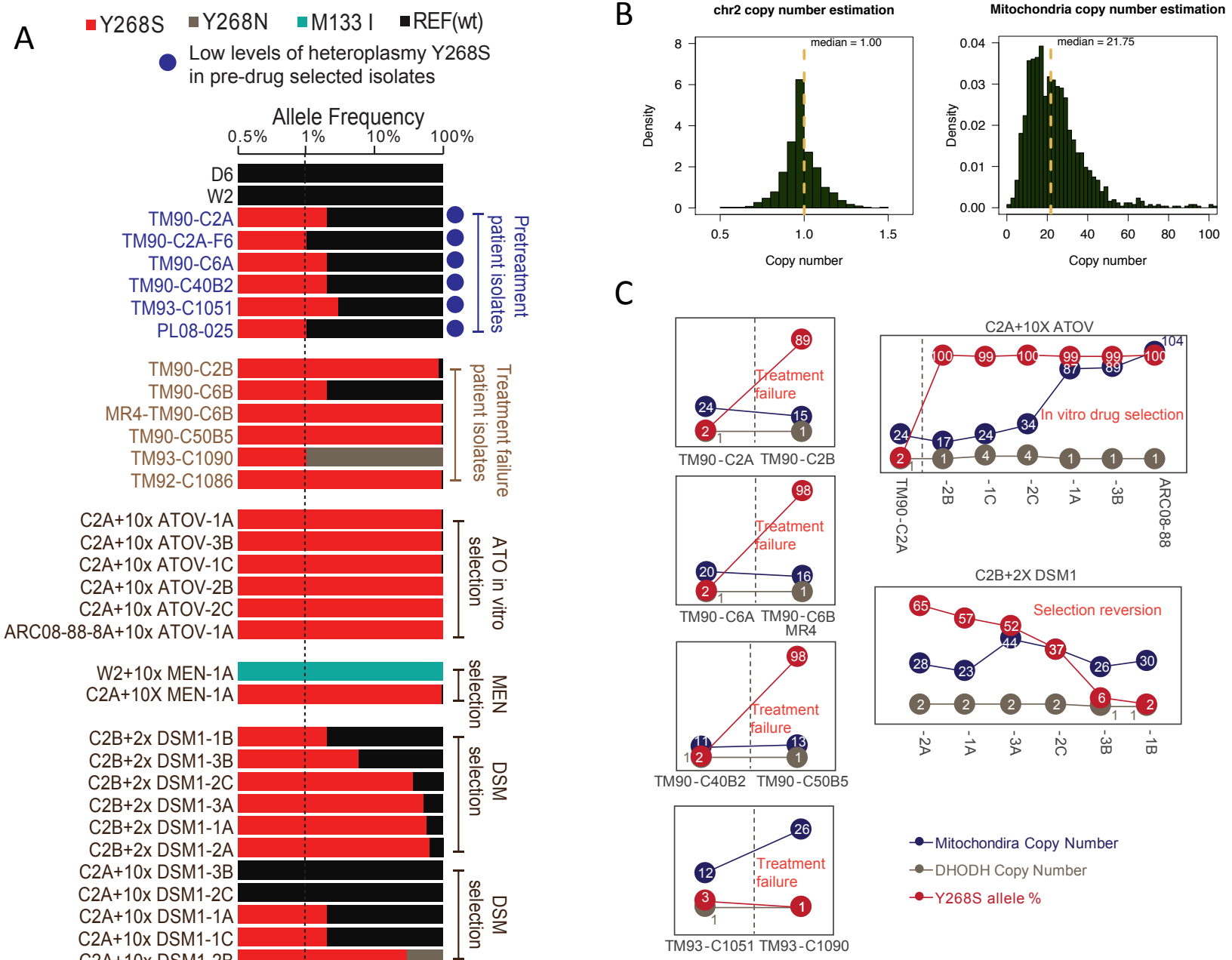
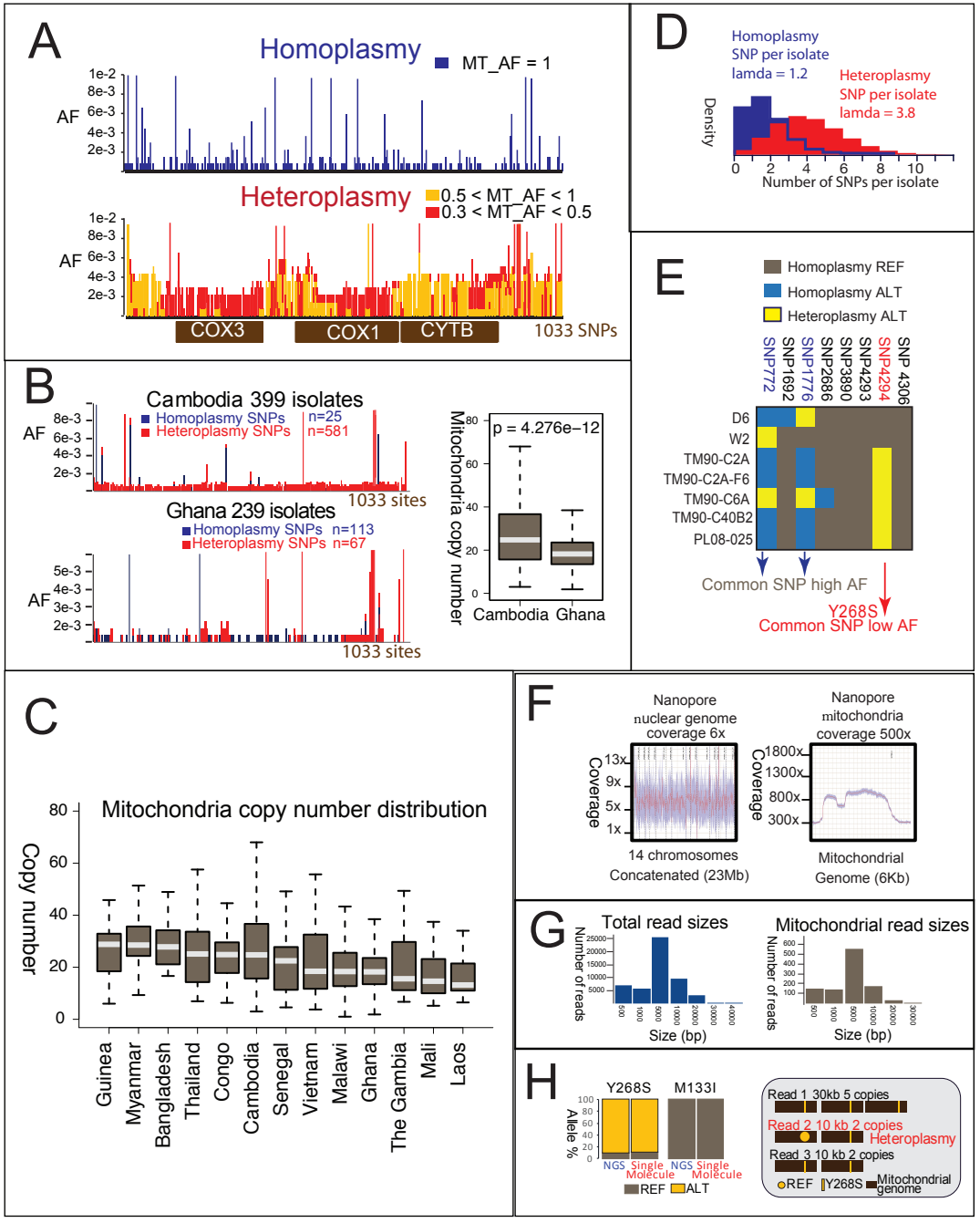
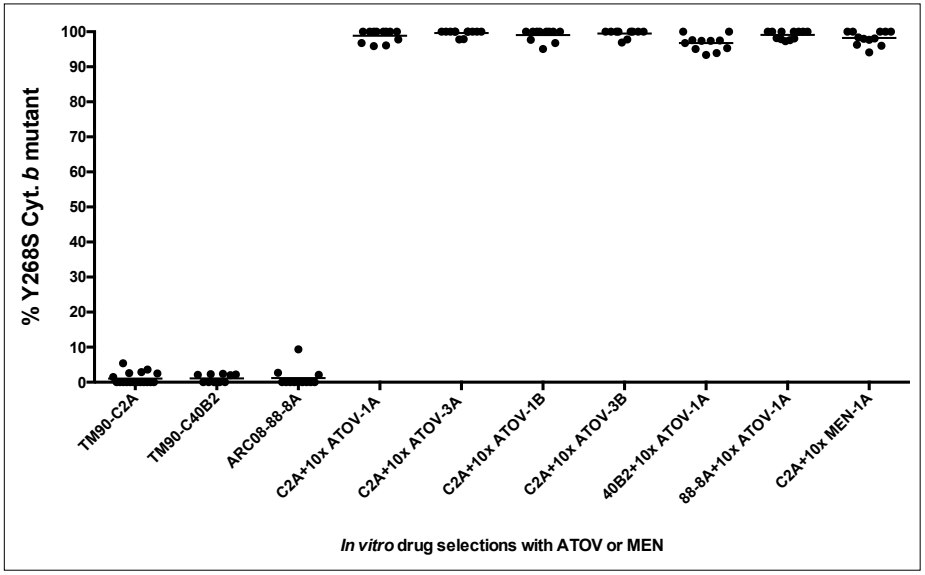

A
Y268S
Y268N
M133 I
REF(wt)
Low levels of heteroplasmy Y268S in pre-drug selected isolates
Allele Frequency
0.5%
1%
10%
100%
D6
W2
TM90-C2A
TM90-C2A-F6
TM90-C6A
TM90-C40B2
TM93-C1051
PL08-025
Pretreatment patient isolates
TM90-C2B
TM90-C6B
MR4-TM90-C6B
TM90-C50B5
TM93-C1090
TM92-C1086
Treatment failure patient isolates
C2A+10x ATOV-1A
C2A+10x ATOV-3B
C2A+10x ATOV-1C
C2A+10x ATOV-2B
C2A+10x ATOV-2C
ARC08-88-8A+10x ATOV-1A
ATO in vitro selection
W2+10x MEN-1A
C2A+10X MEN-1A
MEN selection
C2B+2x DSM1-1B
C2B+2x DSM1-3B
C2B+2x DSM1-2C
C2B+2x DSM1-3A
C2B+2x DSM1-1A
C2B+2x DSM1-2A
DSM selection
C2A+10x DSM1-3B
C2A+10x DSM1-2C
C2A+10x DSM1-1A
C2A+10x DSM1-1C
C2A+10x DSM1-2B
DSM selection
B
chr2 copy number estimation
median = 1.00
Density
Copy number
Mitochondria copy number estimation
median = 21.75
Density
Copy number
C
Treatment failure
TM90-C2A TM90-C2B
Treatment failure
TM90-C6A TM90-C6B MR4
Treatment failure
TM90-C40B2 TM90-C50B5
Treatment failure
TM93-C1051 TM93-C1090
C2A+10X ATOV
In vitro drug selection
TM90-C2A
-2B
-1C
-2C
-1A
-3B
ARC08-88
C2B+2X DSM1
Selection reversion
-2A
-1A
-3A
-2C
-3B
-1B
Mitochondira Copy Number
DHODH Copy Number
Y268S allele %

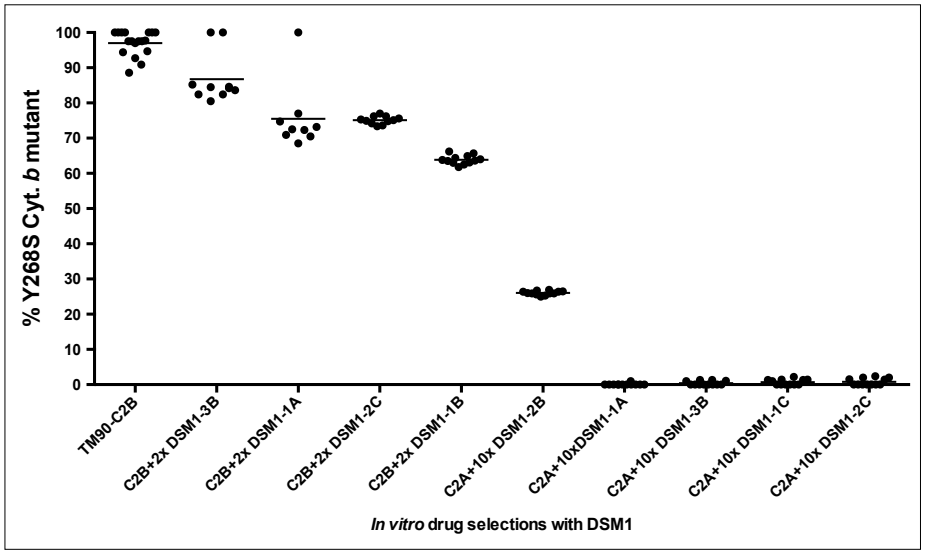
Fig 3


A
Homoplasmy
MT_AF = 1
AF
1e-2
8e-3
6e-3
4e-3
2e-3
Heteroplasmy
0.5 < MT_AF < 1
0.3 < MT_AF < 0.5
COX3
COX1
CYTB
1033 SNPs
B
Cambodia 399 isolates
Homoplasmy SNPs n=25
Heteroplasmy SNPs n=581
1033 sites
Ghana 239 isolates
Homoplasmy SNPs n=113
Heteroplasmy SNPs n=67
1033 sites
p = 4.276e−12
Mitochondria copy number
Cambodia
Ghana
C
Mitochondria copy number distribution
Copy number
Guinea
Myanmar
Bangladesh
Thailand
Congo
Cambodia
Senegal
Vietnam
Malawi
Ghana
The Gambia
Mali
Laos
D
Homoplasmy
SNP per isolate
lamda = 1.2
Heteroplasmy
SNP per isolate
lamda = 3.8
Density
Number of SNPs per isolate
E
Homoplasmy REF
Homoplasmy ALT
Heteroplasmy ALT
SNP772
SNP1692
SNP1776
SNP2686
SNP3890
SNP4293
SNP4294
SNP 4306
D6
W2
TM90-C2A
TM90-C2A-F6
TM90-C6A
TM90-C40B2
PL08-025
Common SNP high AF
Y268S
Common SNP low AF
F
Nanopore
nuclear genome
coverage 6x
Coverage
14 chromosomes
Concatenated (23Mb)
Nanopore
mitochondria
coverage 500x
Mitochondrial
Genome (6Kb)
G
Total read sizes
Number of reads
Size (bp)
Mitochondrial read sizes
H
Y268S
M133I
Allele %
NGS Single Molecule
REF
ALT
Read 1 30kb 5 copies
Read 2 10 kb 2 copies
Heteroplasmy
Read 3 10 kb 2 copies
REF
Y268S
Mitochondrial genome

Extended Figures

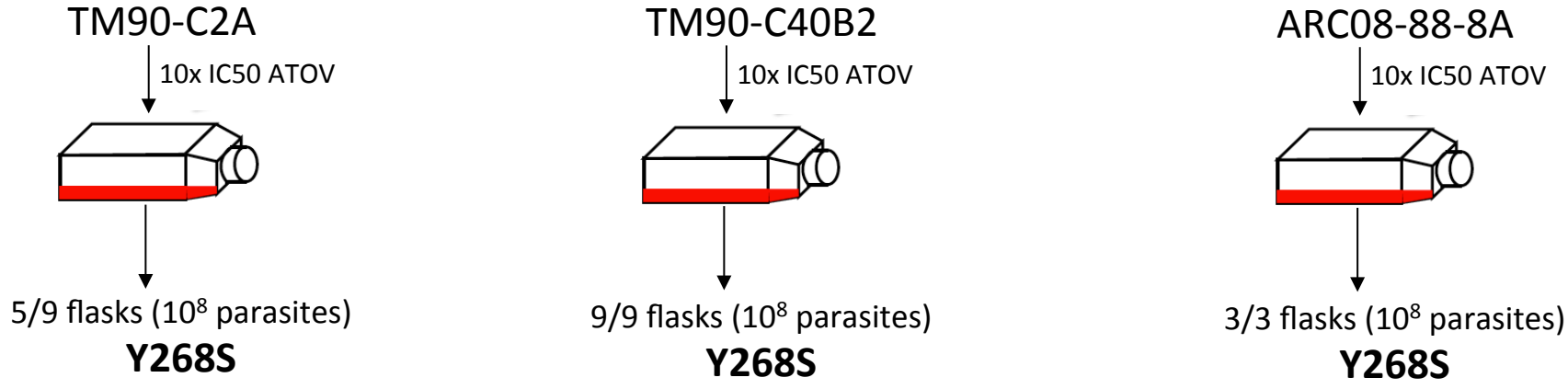
Extended Data Figure 1





Extended Data Figure 2

Y268S mutation is readily selected *in vitro*

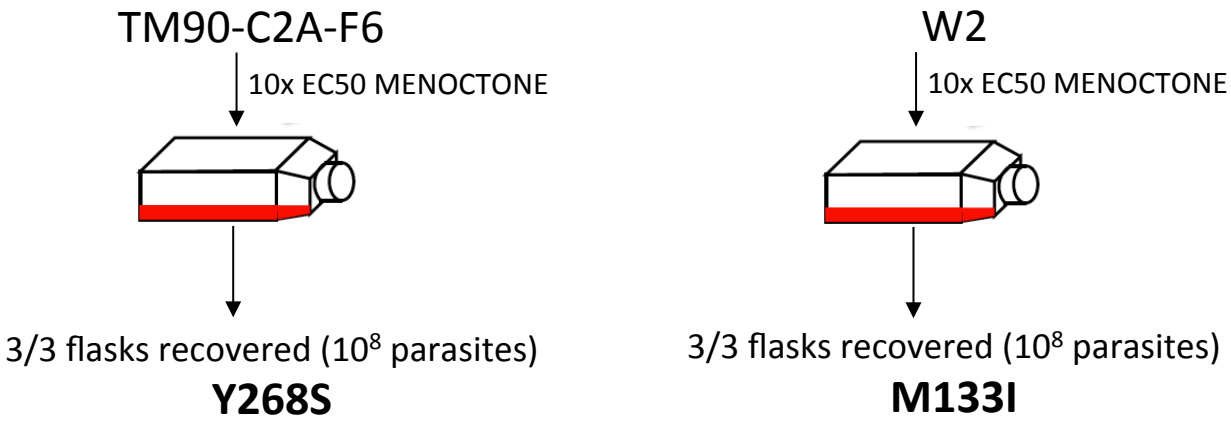


Initial resistance to mitochondrial electron transport chain inhibitors in atovaquone-selected populations of *P. falciparum* (EC_{50}, μM) and cytochrome *b* genotypes

Parasite	Qo site inhibitors		DHODH inhibitor	cyt. *b* genotype
	ATOV	MYX	DSM-1	
ARC08-88-8A	0.0076	ND	0.10	
TM90-C2A-F6	0.0013	0.094	0.12	--
TM90-C2B-A3	12	1.2	0.040	Y268S
C2A-F6+10x ATOV-1A	26.7	1.27	0.078	Y268S
C2A-F6+10x ATOV-3A	5.43	4.32	0.033	Y268S
C2A-F6+10x ATOV-1B	28.2	1.99	0.094	Y268S
C2A-F6+10x ATOV-2B	4.16	0.280	0.029	Y268S
C2A-F6+10x ATOV-3B	4.21	0.287	0.031	Y268S
ARC08-88-8A+10xATOV-1A	63	ND	0.067	Y268S

Extended Data Figure 3

Menoctone generates cytochrome *b* mutants identical to atovaquone resistance

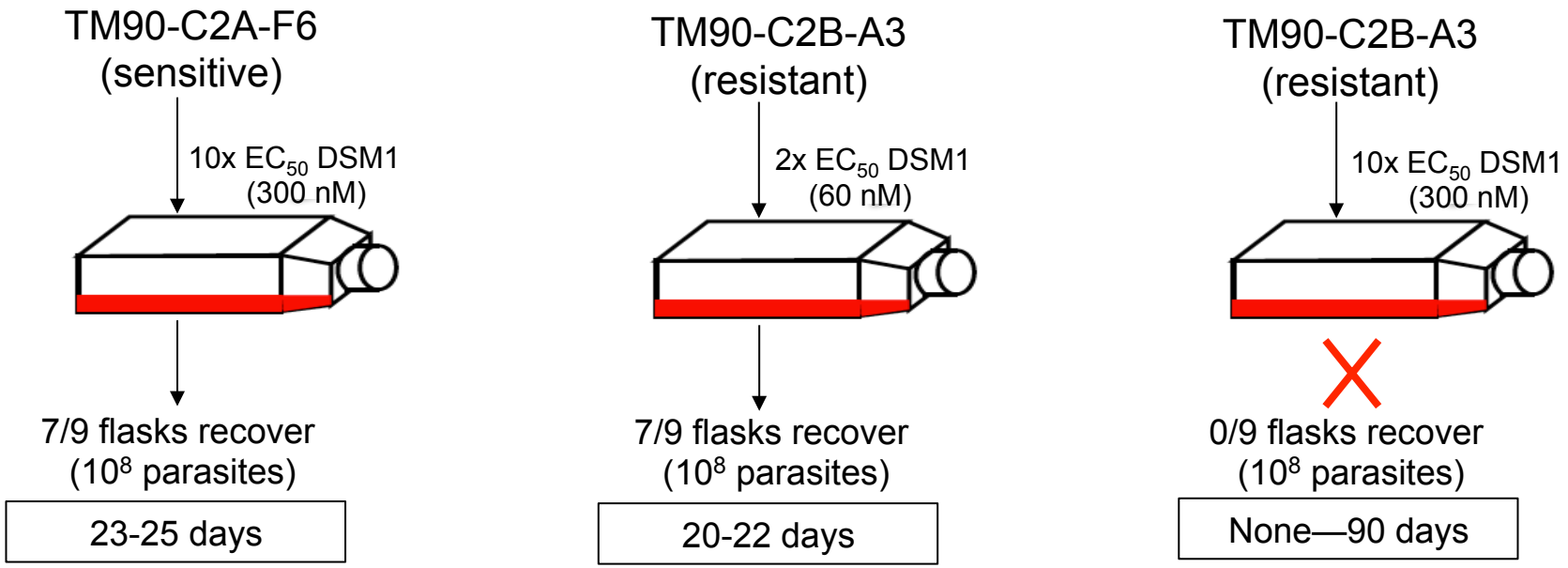


Resistance to mtETC inhibitors in menoctone-selected populations of *P. falciparum* (EC_{50} μM), SEM

Parasite	Q_o site inhibitors		DHODH inhibitor	cyt. *b* genotype
	ATOV	MEN	DSM-1	
W2	0.0023 ± 0.00051	0.12 ± 0.031	0.11 ± 0.0018	
TM90-C2A-F6	0.0013 ± 0.00030	0.070 ± 0.0085	0.067 ± 0.042	--
TM90-C2B-A3	12 ± 3.0	19 ± 1.4	0.040 ± 0.0042	Y268S
C2A+10xMEN-1A	54 ± 8.9	23 ± 2.06	0.052 ± 0.0023	Y268S
W2+10xMEN-1A	0.066 ± 0.022	18 ± 6.07	0.16 ± 0.028	M133I

Extended Data Figure 4

DSM1-Selection in atovaquone-sensitive and resistant parasites



Earliest drug susceptibility profiles of DSM1-resistant *P. falciparum* (μM)

DSM1-selected populations	ATOV	MYX	DHODH
TM90-C2A-F6	0.0013	0.094	0.067
TM90-C2B-A3	12	1.2	0.040
C2A+10xDSM-1A	0.0018	0.010	0.43
C2A+10xDSM-2B	5.8	0.56	0.22
C2B+2xDSM-1A	7.0	1.7	0.20
C2B+2xDSM-1B	31	0.62	0.55
C2B+2xDSM-3B	29	5.6	0.31
C2B+2xDSM-2C	19	1.7	0.37

Extended Data Table 4

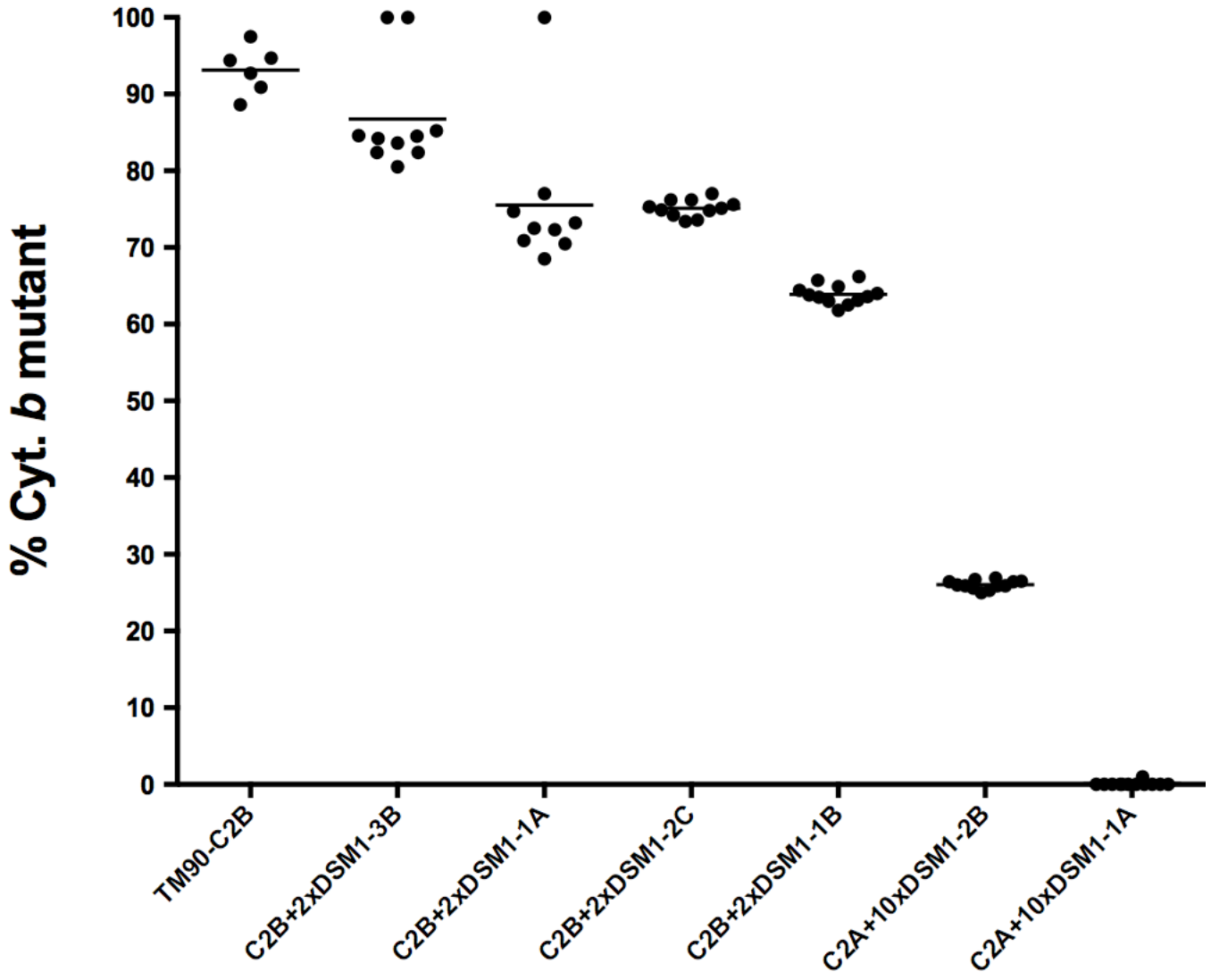
DSM1 pressure induces DHODH copy number amplifications, DHODH mutations, and Y268S heteroplasmy in cytochrome *b*

Genotypes of DSM1 drug selections immediately following recovery

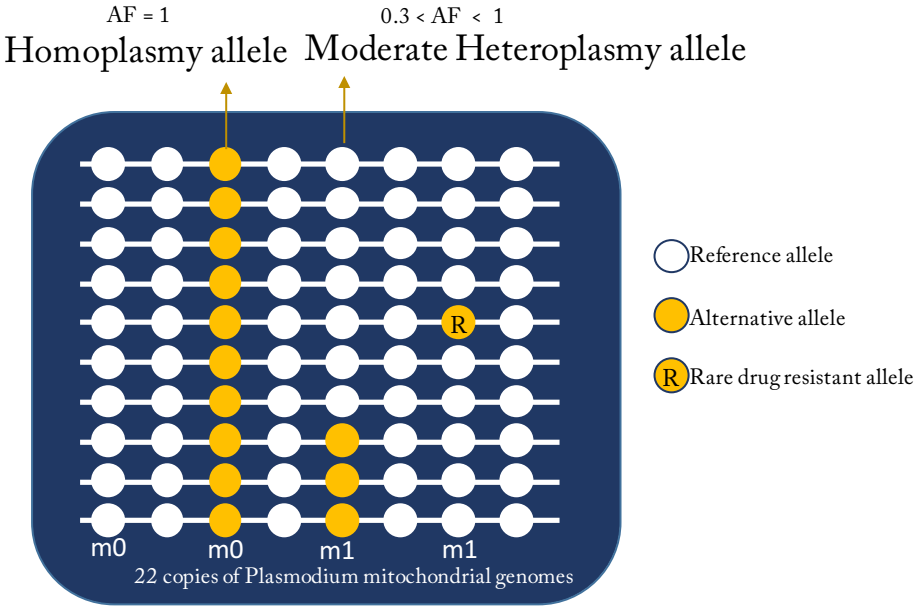
DSM1-selected populations	*dhodh* copy number	PFF_0160C (*dhodh*)	MALMITO_3 (cyt. *b*)	MALMITO_1 (*coxIII*)
C2A+10xDSM1-1A	3	WT	WT	I239V
C2A+10x DSM1-2B	1	R265G	Y268S/WT	I239V
C2B+2xDSM-1A	2	WT	Y268S/WT	I239V
C2B+2xDSM1-1B	2	WT	Y268S/WT	I239V
C2B+2xDSM1-3B	1.5	WT	Y268S/WT	I239V
C2B+2xDSM1-2C	1	WT	Y268S/WT	I239V

Extended Data Figure 5

Y268S heteroplasmy is responsible for combined atovaquone/DSM1 resistant phenotypes



Extended Data Figure 6



maximum likelihood estimation f_{m1} , f_{m0} can be calculated based on pre-classified potential heteroplasmy(m1) and homoplasmy (m0) data. The $\log(L(f_{m1}) / L(f_{m0}))$ was used for finding significant heteroplasmy site