nature communications

Article

https://doi.org/10.1038/s41467-024-54915-6

Population genomics and transcriptomics of *Plasmodium falciparum* in Cambodia and Vietnam uncover key components of the artemisinin resistance genetic background

Received: 18 October 2023

Accepted: 22 November 2024

Published online: 05 December 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

The emergence of *Plasmodium falciparum* parasites resistant to artemisinins compromises the efficacy of Artemisinin Combination Therapies (ACTs), the global first-line malaria treatment. Artemisinin resistance is a complex genetic trait in which nonsynonymous SNPs in *PfK13* cooperate with other genetic variations. Here, we present population genomic/transcriptomic analyses of *P. falciparum* collected from patients with uncomplicated malaria in Cambodia and Vietnam between 2018 and 2020. Besides the *PfK13* SNPs, several polymorphisms, including nonsynonymous SNPs (N1131I and N821K) in *PfRad5* and an intronic SNP in *PfWD11* (WD40 repeat-containing protein on chromosome 11), appear to be associated with artemisinin resistance, possibly as new markers. There is also a defined set of genes whose steady-state levels of mRNA and/or splice variants or antisense transcripts correlate with artemisinin resistance at the base level. In vivo transcriptional responses to artemisinins indicate the resistant parasite's capacity to decelerate its intraerythrocytic developmental cycle (IDC), which can contribute to the resistant phenotype. During this response, *PfRAD5* and *PfWD11* upregulate their respective alternatively/aberrantly spliced isoforms, suggesting their contribution to the protective response to artemisinins. *PfRAD5* and *PfWD11* appear under selective pressure in the Greater Mekong Sub-region over the last decade, suggesting their role in the genetic background of the artemisinin resistance.

Resistance of *Plasmodium falciparum* infections to artemisinins, the current first-line therapy for malaria as artemisinin-based combination therapies (ACTs), was first reported in western Cambodia in 2009[1,2]. In the last decade artemisinin resistance spread throughout the entire Greater Mekong Subregion (GMS)[3], and more recently also emerged in Eastern Africa[4,5]. In South America, several independent events also occurred that led to the emergence of resistance-associated markers, particularly in Guyana, as early as 2016[6,7]. The artemisinin resistance phenotype in vivo is characterized by an increased parasite clearance half-life (PC1/2 > 5 hours) after treatment with artemisinin alone or ACTs[8]. Artemisinin resistance is linked with nonsynonymous SNPs in the Kelch-repeat β-propeller domain of PfK13, and in vitro, it is characterized by an increased parasite survival of synchronized early ring stage parasites, assayed by the ring survival assay (RSA)[9]. Since their emergence, the *PfK13* SNP(s) have been the most reliable genetic markers of artemisinin resistance in clinical *falciparum* malaria in the GMS[8,10]. Independent emergence of *PfK13* mutations in Sub-Saharan Africa has accelerated since 2013 when the frequency of R561H increased from 7.4% to 13% and is now also associated with increases in PC1/2, albeit not past 5hr[4,11]. In Uganda, infections with *P. falciparum*

e-mail: arjen@tropmedres.ac; zbozdech@ntu.edu.sg

carrying the artemisinin resistance linked *PfK13* SNPs, A675V and C469Y, are also now spreading rapidly[12,13]. Occurrence of artemisinin resistance was also confirmed recently in other parts of Sub-Saharan Africa, such as Ethiopia[14], Uganda[4], Eritrea[15] and Tanzania[16]. Together with the independent emergence in South American Guyana reported earlier[6,7], these Sub-Saharan occurrences stress the fact that artemisinin resistance can emerge autonomously in essentially any other part of the malaria-endemic world[17,18]. The main concern is the fact that artemisinin resistance compromises the overall efficacies of essentially all ACT types, allowing the emergence of resistance alleles to the partner drugs, which was well documented for mefloquine[19] and piperaquine[20]. Given that all current worldwide malaria control and elimination programs rely on the ACTs; it is imperative to understand the nature of artemisinin resistance, particularly all components of its polygenetic molecular mechanism. It is feasible to expect that other, currently unknown, components of artemisinin resistance will be in a delicate balance driven by the positive selection of resistance on one side and negative impact on parasite fitness on the other, as it was demonstrated for several of the *PfK13* nonsynonymous SNPs such as C580Y and M579I[21].

As suggested earlier in the GMS, the *PfK13* SNPs mediate artemisinin resistance only in parasites with a specific genetic background[22,23]. This is also likely true for Sub-Saharan Africa, where the resistant infections exhibit more moderate but still significant shifts in PC1/2[11]. The genetic variations of the putative genetic background could either directly contribute to the resistance and/or aid *P. falciparum* growth and transmission. The latter may represent mutation that does not necessarily mediate the artemisinin resistance per se but instead help to alleviate the fitness cost due to the resistance mutation and/or boost gametocyte production to elevate transmission of the resistant parasite lines. Indeed, large segments on chromosomes 10, 13, and 14[24,25], and at least five specific nonsynonymous SNPs[23] were found to be associated with artemisinin resistance by earlier GWAS. Longitudinal analysis and replicate GWAS/random forest approaches identified an additional 7 and 8 new SNPs, respectively[17,26]. Population transcriptomics analyses (also termed transcription-phenotype association studies, TPAS), further supported the view of a complex genetic trait, revealing the artemisinin resistance-associated transcriptional profile (ARTP)[27–29]. The ARTP comprises genes associated with the PC1/2 by their expression level, and it underlines specific cellular physiological state(s) of the artemisinin-resistant *P. falciparum* parasites. The differential expression of most, if not all, individual genes of the ARTP is also associated with specific genetic polymorphism[30], and in one case, it was shown to be mediated by sequence length polymorphisms at the AT-rich tandem repeats in the gene promoter[31].

There is also a considerable number of genes associated with artemisinin resistance by functional inference (reviewed in refs. 32–35). The center of this is the PfK13 involvement in hemoglobin endocytosis, being located at the cytostomes and physically interacting with a set of proteins (K13-interacting candidates, KICs)[36]. The potential to mediate artemisinin resistance was demonstrated for at least two of these including KIC1 and KIC5 indicating their functional context[37,38]. Another example is PfCoronin, an *actin*-binding protein, carrying a WD40 domain, functionally and structurally related to the PfK13 Kelch-repeat β-propeller. Like PfK13, mutations selected under DHA drug pressure in the WD40 propeller domain in vitro mediate decreased susceptibility to artemisinins in the standard RSA[39]. Although naturally occurring SNPs in *Pfcoronin* do not exhibit an association with an increased PC1/2, for the time being, their rising allele frequencies and skewed distributions along the protein have been reported in Sub-Saharan Africa[40–42]. Like *PfK13*, the *PfCoronin* mutations depend largely on the genetic background of the *P. falciparum* parasites[43]. Based on multiple lines of evidence, Oberstaller, J. et al. l, 2021 compiled a putative set of functionalities and genes causing artemisinin resistance[37]. Apart from PfK13-linked hemoglobin endocytosis[36,44], these include vesicular trafficking (involving PIP5K)[45]; Redox metabolism (Thioredoxin)[46]; heat shock/unfolded protein responses (PfAP2-HS)[47]; cell cycle regulation (MYST)[48], and others. Indeed, most of these functionalities have been associated with increased PC1/2 in clinical samples in TPAS studies[27–29]. Complex involvement in artemisinin resistance of some of these genes was also confirmed by PiggyBac-transposon mutagenesis[49] and their regulatory "rewiring" in the PfK13 mutant parasites[50].

In this work, using GWAS/TPAS, we identified a novel set of genes associated with variation in PC1/2 in *P. falciparum* infections via genetic polymorphisms and/or differential transcription either at the baseline levels or as a response to artemisinin in natural malaria infections. Out of these, we single out the polymorphisms in *PfRAD5* (PF3D7_1343400) on chromosome 13 and a novel WD40 domain-containing protein on chromosome 11 (PfWD11, PF3D7_1138800) that associates with artemisinin resistance by exonic and intronic SNPs, respectively. Both genes also exhibited artemisinin-response/resistance-associated differential expression characteristics. As such *PfRAD5* and *PfWD11* represent potentially new genetic markers that can complement the clinical relevance of PfK13.

Results

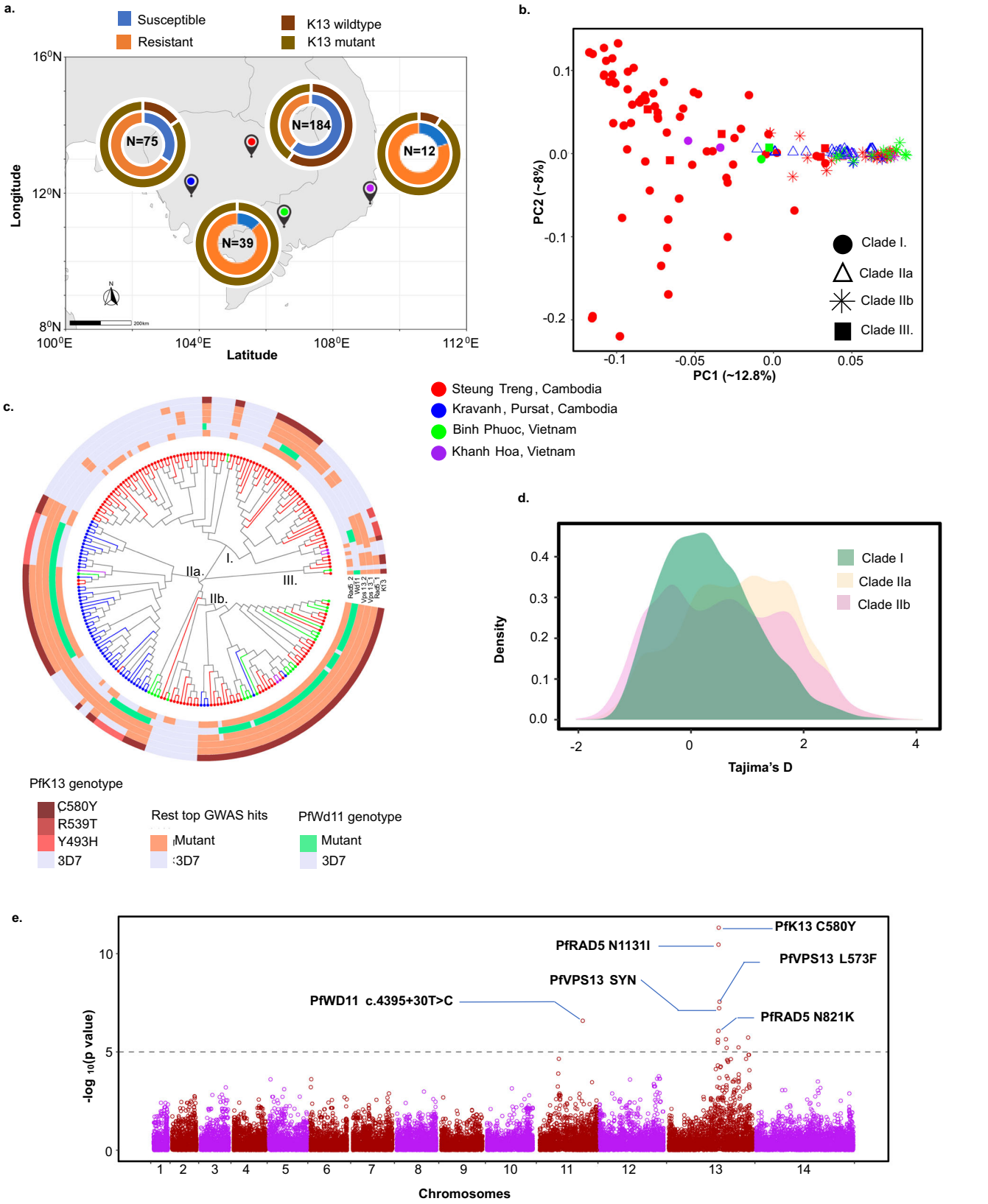
Data sets

The main goal of this study was to investigate the artemisinin resistance clinical phenotype of the delayed parasite clearance half-life (PC1/2)[23,51,52] in the context of a complex genetic trait. Specifically, we wished to identify additional molecular markers/mediators contributing to artemisinin resistance or other components of the genetic background. For that, we carried out population genomics and transcriptomics analyses of the *P. falciparum* parasite DNA and RNA collected from patients enrolled in a treatment study with the acronym TACT-CV similar to our previous published studies from the Tracking Resistance to Artemisinins Collaboration, TRACI & II[27,29,30]. TACT-CV was an open-label randomized clinical trial of triple artemisinin combination therapies (TACT) conducted in five hospitals or health centers in western and eastern provinces of Cambodia (wCambodia & eCambodia), including Kravanh, Pursat, and Stung Treng; and southern Vietnam (sVietnam): Binh Phuoc and Khanh Hoa between March 18, 2018, and January 30, 2020 (Fig. 1a, Supplementary Data 1)[52]. Reflecting the status of artemisinin resistance, the majority of the samples from wCambodia and sVietnam carried a *PfK13* mutant allele and came from infections with a slow clearance half-live (PC1/2 > 5 hours). In contrast, the set of samples collected in eCambodia was evenly split between artemisinin-resistant and sensitive parasites, the latter of which carried the ancestral wild-type *PfK13* allele. Out of the total 310 *P. falciparum*-infected erythrocyte samples, whole genome sequencing (WGS) could be performed successfully with 290 samples. Using RNA-Seq, we also generated high quality transcriptomes of these *P. falciparum* isolates at two distinct time points; *(i.)* just before the start of antimalarial treatment (0 hr samples, *baseline transcriptomes*, $n = 282$); and *(ii.)* six hours after the start of drug treatment (6 hr samples, *induced transcriptomes*, $n = 252$). Given the low transmission epidemiological settings, this sampling is highly representative of the overall transmission pattern of *P. falciparum* infections representing essentially all detectable infections in these western and eastern parts of the GMS between 2018 and 2020[53].

Population flux within the eastern GMS *P. falciparum* infections

Focusing on the WGS, the *P. falciparum* genomes were characterized with coverage ranging from 0.09X – 472.81X, with a mean of 149.54X (Fig. S1). After quality filtering (see material and methods), a total of 29,994 SNP with at least 1% minor allele frequency from 227 parasite isolates were considered. Principal component analysis (PCA) revealed that the first component (~ 12.8% of the overall variation) separated a

https://doi.org/10.1038/s41467-024-54915-6



subset of eCambodian samples from wCambodia and sVietnam (Fig. 1b). Indeed, the elbow analysis indicated that the first component underlines the main structure of the TACT-CV cohort and supports the K = 2 means clustering for stratification (Fig. S2). This specifically indicates that the eCambodian *P. falciparum* population has a distinct genetic structure that differs from that in wCambodia and is mainly dominated by the KEL1/PLA1 parasite lineage as a result of a selective sweep between 2010 and 2014 driven by DHA-piperaquine (DHA-PPQ) antimalarial treatment[53]. On the other hand, the second and third components (8% and ~7%) underline genetic diversity within the eCambodian subpopulation per se (Figs. 1b and S2). Indeed, neighbor-joining tree analysis based on a pairwise distance matrix supports the

Fig. 1 | Genetic features of the parasite cohort. a Geographic distribution of the clinical isolates along with their resistance status. The colors of the points agree with the point colors of the PCA plot and the node color of the neighbor-joining tree. **b** The first two principal components are showing the population structure of the parasites. The shape of the points corresponds to the clade classification defined in the tree. **c** Neighbor-joining tree separating parasites into three major clades. The circle encompassing the tree shows the allelic distribution of the top 6 GWAS hits. **d** Differential distribution of the Tajima's D for three clades defined by the neighbor-joining tree. **e** SNPs linked with $PC^{1/2}$ by genome-wide association study. The *P* values were obtained by fitting data in linear mixed model implemented in FastLMM packages. Presented *P* values are uncorrected. Apart from the locus at chromosome 13 already known to be linked with parasite resistance, a novel locus at chromosome 11 passed the statistical cut-off and comprises a single intronic variation in the WD repeat-containing gene (Pf3D7_1138800). Source data are provided as a Source Data file.

PCA findings, classifying the TACT-CV cohort into two distinct clades, I and II. Clade I consists nearly exclusively of eCambodian parasites, while clade II is split evenly between w and eCambodian parasites and thus can be further subdivided. Subclade IIa consists of wCambodian samples, while clade IIb includes eCambodian parasites, and the vast majority of samples are collected at the two sVietnam sites. The samples in clades IIa and IIb appear genetically related, which may indicate their common origin, likely represent the KEL1/PLA1 lineage that expanded in this region prior to the time of this study[53]. On the other hand, clade I parasites appear genetically distant, exhibiting a founder-like (sub)population structure (Fig. 1c). Clade I is characterized by a significant overrepresentation of the artemisinin sensitive phenotype, with most parasites carrying the *PfK13* wild-type allele. This contrasts subclades IIa and IIb, with most of the parasites being artemisinin resistant, carrying the C580Y or other PfK13 mutations. It is also important to note the existence of a small clade of 6 parasite isolates (clade III) that showed a high genetic distance from the TACT-CV cohort, whose origin is currently unknown.

Consistently, an admixture analysis identified a distinct ancestral subpopulation that is exclusive to eCambodia and comprised mainly of artemisinin-sensitive parasites (Fig. S3). This population is currently admixing with two distinct subpopulations of resistant parasites, one of which appears to be spreading from wCambodia. Tajima's D index value distribution is also consistent with this model (Fig. 1d). Clade I peak at 0, indicating an absence of any current selection pressure. On the other hand, Clade IIa and IIb exhibit complex distributions with multiple peaks, most of which shift towards the positive Tajima's D index values. This suggests a finer subgrouping of the parasite populations in Clade IIa and IIb with ongoing selection process(es). This is supported by the distribution of the Tajima's D values across the genome exhibiting a neutral profile for clade I but a highly variable pattern for IIa and IIb, with several subregions on chromosomes 2, 6, and 12 being particularly under a strong selection (Fig. S4, Supplementary Data 2). Finally, identity by descent (IBD) analysis revealed close genetic relatedness of Clade IIa and IIb, both of which are distant from Clade I (Fig. S5). Interestingly, parasites in Clade I exhibit broader genetic variability, possibly reflecting their diversion unaffected by drug selection. Clade I is, however, more related to Clade IIb, reflecting their common geographic distribution. Collectively, these results suggest that artemisinin-resistant parasites expanding from wCambodia are gradually admixing into the drug-sensitive population pre-existing in this region and/or spreading there from neighboring Laos. This admixture pattern is (at least partially) driven by the gradual change in first-line antimalarial treatment from DHA-PPQ to artesunate-mefloquine (AS-MFQ) in Cambodia between 2014 and 2017, implemented due to the rapid spreading of piperaquine resistance[54,55].

GWAS of the artemisinin resistance phenotype in the TACT-CV *P. falciparum* patient cohort

Taking the population structure of the TACT-CV cohort into consideration, next, we carried out a genome-wide association study (GWAS) to search polymorphic loci linked with artemisinin resistance. Here the artemisinin resistance phenotype, defined by the PC1/2, was used as a continuous variable. Here, we carried out GWAS with the Genetic Relatedness Matrix (GRM) that mitigates the population structure effect as implemented by the FastLMM program. The GRM method was chosen instead of the standard 10 principal component correction due to a lower inflation factor $\lambda^{GRM} = 1.05$ compared to $\lambda^{10PC} = 2.25$ (Fig. S6). Indeed, even this limited size cohort of ~300 samples of the TACT-CV can provide sufficient statistical power for genetic analysis of artemisinin resistance factors (Fig. S7). For that, however, we applied a lower statistical threshold of $p\text{Value} > 10^{-5,}$ instead of commonly used $10^{-8,}$ to accommodate for the sample scarcity as previously described[56]. By this approach, we identified C580Y in *PfK13* (PF3D7_1343700) as the top marker of artemisinin resistance (*p*-value 4.92E-12) (**1e**). In addition, we identified 11 SNPs significantly associated with PC1/2 ($p\text{Value} \sim 10^{-5}$), from which 10 were located at chromosome 13, several genes upstream and/or downstream of *PfK13*. All these 10 SNPs fell into a broader chromosome 13 region detected by previous GWAS[17,23,24]. These included missense mutations N1131I and N821K in DNA repair protein, *PfRAD5* (PF3D7_1343400), L573F in VPS13 domain-containing protein (PF3D7_1343800), and H1420Q in E1-E2 ATPase (PF3D7_1348800) (Supplementary Data 3). Curiously, the two nonsynonymous SNPs in *PfRAD5*, located three genes upstream of *PfK13*, were also previously found to be a part of the artemisinin resistance genetic architecture 2011-2013[30,57]. In addition to these, we identified an SNP (c.4395+30 T > C) at chromosome 11 that falls within the 2nd intron in WD40 repeat-containing protein (PF3D7_1138800) (*p*-value 2.57E-7). For further discussion, we term this *P. falciparum* WD40-domain protein on chromosome 11 (PfWD11) and the intronic SNP i2 + 30 T > C. Interestingly, the *PfWD11* falls in the regions with positive Tajima's D values in both clades II a and IIb, suggesting a balancing selection in both parts of Cambodia (Fig. S4). On the other hand, the region of chromosome 13 with *PfK13*, *PfRAD5*, and other GWAS hits appears to be under positive selection in clade IIa, representing the recent spread of artemisinin resistance in (e)Cambodia.

Transcriptome analysis of differential gene expression in resistant parasites

Next, we explored the transcriptome datasets (see above) to identify genes that are associated with the artemisinin resistance phenotype by their differential gene expression. Here, we maximized the detection of all RNA transcript classes, including the canonical protein-coding mRNA transcripts annotated in the *P. falciparum* genome (e.g., annotated Protein-Coding Transcripts, *PCT*) but also their alternatively spliced variants (*altPCT*), overlapping noncoding antisense transcripts (*asPCT)* and intergenic *ncRNA* transcripts. For that, we employed genome-guided transcriptome assembly in a strand-specific manner using a custom-designed bioinformatics pipeline with a specific focus on data filtering based on detection levels (materials and methods). Using a 70% cut-off threshold, we obtained quantitative measurements of 4542 transcripts from 282 0 hr samples (*baseline transcriptomes*) comprising 3289 *aPCTs*, 631 *altPCTs*, 341 *asPCTs* and 281 ncRNAs (Fig. 2a). From 252 of the 6 hr samples (*induced transcriptomes*), we measured an abundance of 2158 *aPCTs*, 299 *altPCTs*, 25 *asPCTs*, and 123 ncRNAs. There were comparable levels of the average expression of all four categories of the transcripts across the 0 hr and 6 hr sample sets. Also, 280 out of the 631 detected *altPCTs* showed differences at predicted splice sites within the protein-coding sequences, and 198 of these (including PfRAD5 and PfWD11, see below) could be classified as aberrant splicing by creating a frameshift that could lead to protein

https://doi.org/10.1038/s41467-024-54915-6

truncation/degradation (Fig. S8). Altogether, these results demonstrate the utility of the applied RNA-Seq methodology and bioinformatics pipeline that allows us to expand transcriptomic analyses of *P. falciparum* in natural infections beyond the canonical mRNA species that up until now were the sole subject of the previous studies[27,29].

Applying the IDC/gametocyte mixed model[29], we identified 137 transcripts, including 100 *aPCTs* and 37 non-a*PCTs*, whose differential expression correlated significantly with the PC1/2 (*p*-value range FDR < 0.05). From these, 113 were up- and 24 down-regulated in the resistant parasites, respectively (Fig. 2b, Supplementary Data 4, Fig. S9). GO Slim enrichment analysis of the 100 *aPCTs* revealed an

Fig. 2 | Parasite transcriptomes and their association with resistance. a The expression and detection levels of four different classes of transcripts were measured in both 0 hr *baseline* and 6 hr *induced* transcriptomes. **b** The Manhattan plot of the Transcriptome-Phenotype Association Study (TPAS) depicts upregulated and downregulated transcripts in resistant parasites along with their transcript classes as marked by the shape of the points. **c** An example of upregulated *asPCT* (anti-sense RNA, top panel, points in red) and down-regulated corresponding *aPCTs* (annotated protein coding transcripts, top panel, points in blue) in the resistant parasites. An example of an upregulated *altPCT* (alternatively spliced RNA, bottom panel, points in red) and down-regulated corresponding *aPCTs* (bottom panel, points in blue) in the resistant parasites. **d** The heatmap shows mean expression residuals of *non-PCT* transcripts found to be up-regulated or down-regulated in artemisinin resistant parasites and the expression profile of their corresponding *aPCTs*. Source data are provided as a Source Data file. Panels **B**, **C** present the uncorrected *P* values that were calculated by fitting data to the Generalized Additive Model.

overrepresentation of factors of nucleoside and Acetyl CoA metabolism and transcriptional regulations (with enrichment p-value < 0.05). Crucially, PfWD11 was also significantly upregulated *aPCTs* in the artemisinin resistant parasites (p-value = 2.6e-05). In addition, we identified 14 *altPCTs* and 23 *asPCTs* that correlated with the PC1/2. For 19 of these, we observed measurable levels of their corresponding *aPCTs* that in some cases exhibited the opposite correlation wed with PC1/2 (Fig. 2c, d). The marked example is an antisense transcript of RNA cytosine C(5)-methyltransferase (Pf3D7_1230600) and SET domain protein (Pf3D7_0629700) for which *asPCT* and *altPCT*, respectively, were upregulated in the resistant parasites, contrasting their *aPCTs* (Fig. 2c). Taken together, here we demonstrate that non-canonical mRNA transcripts, including alternatively spliced variants, antisense, and intergenic ncRNAs. The H0 transcriptomes analyzed in this study did not recapitulate the previously reported ARTP[29], identifying only a small portion of differentially expressed genes between the resistant and susceptible parasites. This is due to the limited statistical power of the TACT-CV samples set (n ~ 300), in which most differential gene expressions were below the statistical significance achieved with previous larger datasets of TRACI[27] and TRACII[29]. Nevertheless, it is feasible to suggest that the identified transcriptional events contribute to the artemisinin resistance genetic background either as direct factors of the resistance mechanism or by alleviating fitness costs and/or boosting gametocyte production (see discussion).

The primary purpose of the transcriptomics analysis of the 6 hr clinical samples was to study in vivo transcriptional responses of *P. falciparum* to artemisinin, the fast-acting ACT component. Subtracting the expression level between the 0 hr and 6 h*r* samples, we identified significant transcriptional responses that were dominated by broad transcriptional downregulations (Fig. 3a, Supplementary Data 5). In the resistant parasites, dramatic downregulation of up to 555 *aPCTs* (FDR <1e-05), which is much less pronounced in the susceptible parasites, with only 105 downregulated *aPCTs* (Fig. 3a). A similar trend could be observed for the non-*PCTs*, with an additional 98 of these being downregulated in the resistant parasites compared to 12 down-regulated in the susceptible ones (Fig. S10. Functional enrichment analysis of the downregulated 555 *aPCTs* identified pathways and cellular processes related to parasites' lifecycle progression. (Fig. 3b). These included protein synthesis, hemoglobin degradation, host cytoplasm remodeling, antigenic presentation, and epigenetic regulation of gene expression. Approximately half (n = 299) of these genes exhibit their highest expression during mid to late ring stages (0-28 hours post-invasion, hpi) (Fig. S11). This suggests that the transcriptional downregulation induced by artemisinin in vivo reflects a general suppression of the ring development (hence IDC deceleration), which is considerably more extensive in the artemisinin-resistant parasites.

Differential expression of non-*aPCTs* of *PfRAD5* and *PfWD11* further strengthens their potential as novel artemisinin resistance markers

For at least 10 genes, we observed these divergent trends between the "*non-aPCTs*" and their canonical *PCT* with respect to response to artemisinin. A notable example is a nucleic acid binding protein, PF3D7_0202600, that exhibits induction of its *PCT* but significant downregulation of *asPCT* as a result to in vivo 6 hr artemisinin exposure (Fig. 3c). Similar diversion was observed for both PfRAD5 and PfWD11. The *altPCT* of PfRAD5 appears significantly more suppressed in the susceptible parasites with higher significance (mean expression residual change −2.87, FDR 4.01e-12) compared to resistant parasites (−2.05, p-value 2.34e-07) (Fig. 3d). Crucially, the *altPCT* of *PfRAD5* retains the second intron, presumably leading to a frameshift and, presumably, suppression of its biological activity. Indeed, both the baseline (H0) and inducible (H6) expression of the *PfRad5* altPCT is associated with the two SNPs with statistically significant p-value 0.02 (for N821K) and 0.03 (for N1131I) (Fig. S12a, b) both of which are also associated with artemisinin resistance (Fig. 1e). Interestingly, both *PfRAD5* SNPs appear to be covariates of the PfK13 SNP, increasing its frequency on the GMS regions in an essentially identical temporal profile (Fig. S12c) These results suggest that the two nonsynonymous *PfRAD5* SNPs are in epistatic relationship with its alternative splicing, both of which could contribute to the PfK13-driven artemisinin resistance phenotype.

Similar eQTL-like association was also observed between the transcript levels of PfWD11 and its intronic SNP i2 + 30 T > C (Fig. 4a). PfWD11 displayed a good correlation between the baseline level of *PCT* and PC1/2, making it a component of the putative ARTP (Fig. 4b). In the 6-hour transcriptome, however, we observed significant overall upregulation of the *altPCT* of PfWD11 (Fig. 4c). While there were no changes in transcript levels in the susceptible parasites, the significant differential expression of both *PCT* and *altPCT* of PfWD11 exhibited considerable transcriptional changes in the resistant parasites (Fig. S13). By performing paired analysis of samples taken before and after six hours of treatment, we observed significant downregulation of the PfWD11 *PCT* (p-value 2.07e-05) and, conversely, upregulation of the *altPCT* (p-value 2.09e-07) in the resistant parasites. These observations were subsequently validated by qPCR using specific oligonucleotide primers for both the *aPCT* and *altPCT* of PfWD11. In the in vitro cultured *P. falciparum* 3D7 strain, both transcripts can be detected in the early stages of the IDC, gradually declining with time (Fig. S14). qPCR of 24 paired (0hr-6hr) parasite isolates for both resistant (12) and susceptible (12) groups from the TACT-CV cohort further substantiate the RNA-Seq results detecting a statistically significant increase of PfWD11 *altPCT* in the resistant group (p-value 0.0008) (Fig. S15a). The *PCT* of PfWD11 followed a similar trend with lower statistical significance for fold-change when these 12 susceptible and 12 resistant parasites were compared (p-value 0.004). Moreover, a similar transcriptional trend can be observed in two culture-adapted artemisinin resistant Cambodian strains IPC4912 carrying $pfk13^{I543T}$ and $PfWD11^{wt}$ and IPC3445 carrying $pfk13^{C580Y}$ and $PfWD11^{i2+30\ T>C}$ genotypes, respectively. These two strains were collected by the Pasteur Institute, Cambodia, between 2010 and 2013[10] and are publicly available at www.beiresources.org. Here, we have observed a dose-dependent induction of both *aPCT* and *altPCT* isoforms in IPC3445, with the most significant values in the highest dose of the drug (Fig. 4d, 15b and Supplementary Data 6). This induction seems specific to the PfWD11i2 + 30 T > C -carrying IPC3445 strain, as no change in both transcript levels was observed in IPC4912 and 3D7. The alternatively spliced isoform retains the intron 5 (length=190 bp), causing a frameshift that first introduces 5 aberrant amino acids encoded by the first 15 intron nucleotides and a subsequent stop codon that truncates the proteins by 59 amino acids at

https://doi.org/10.1038/s41467-024-54915-6

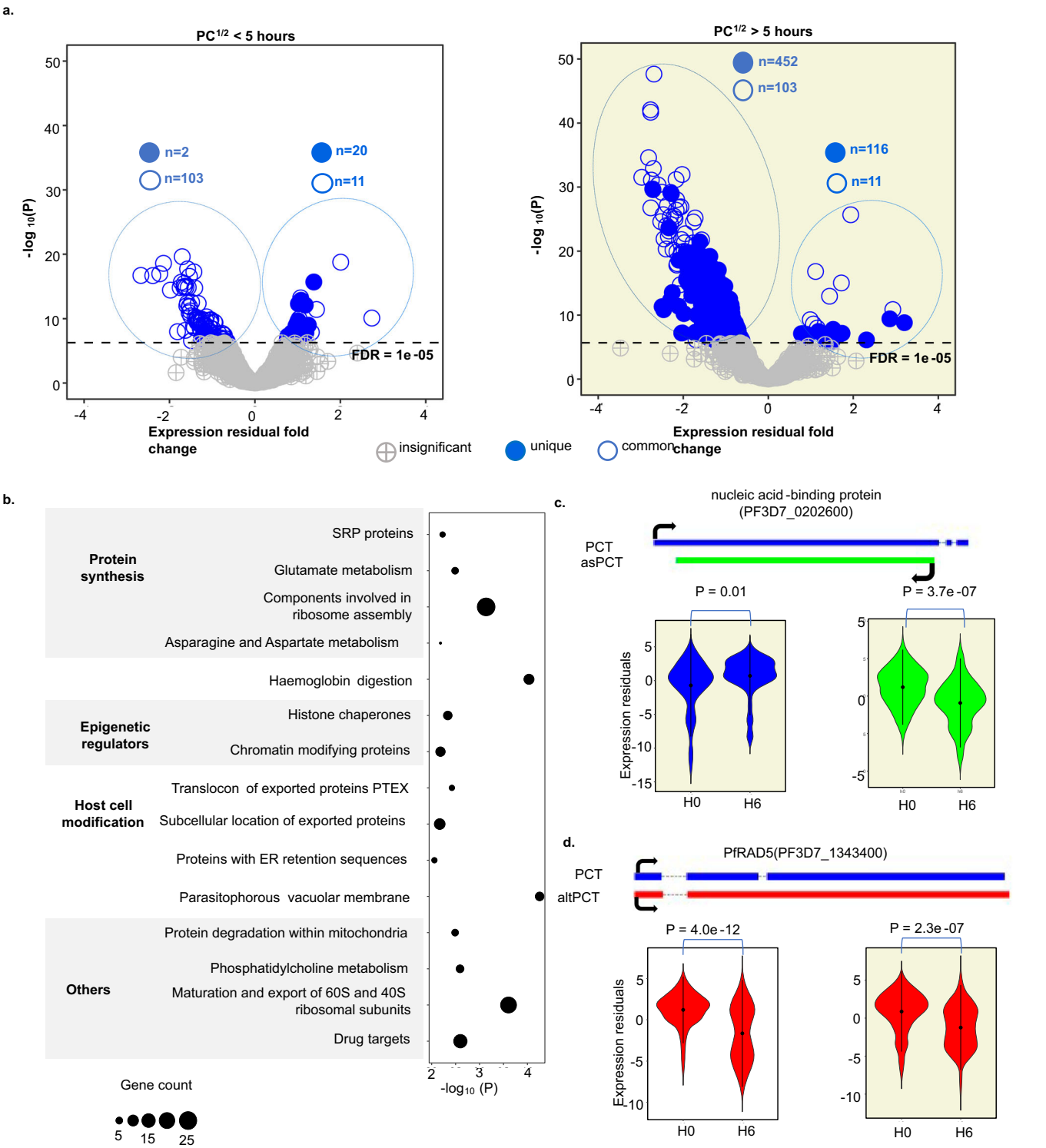
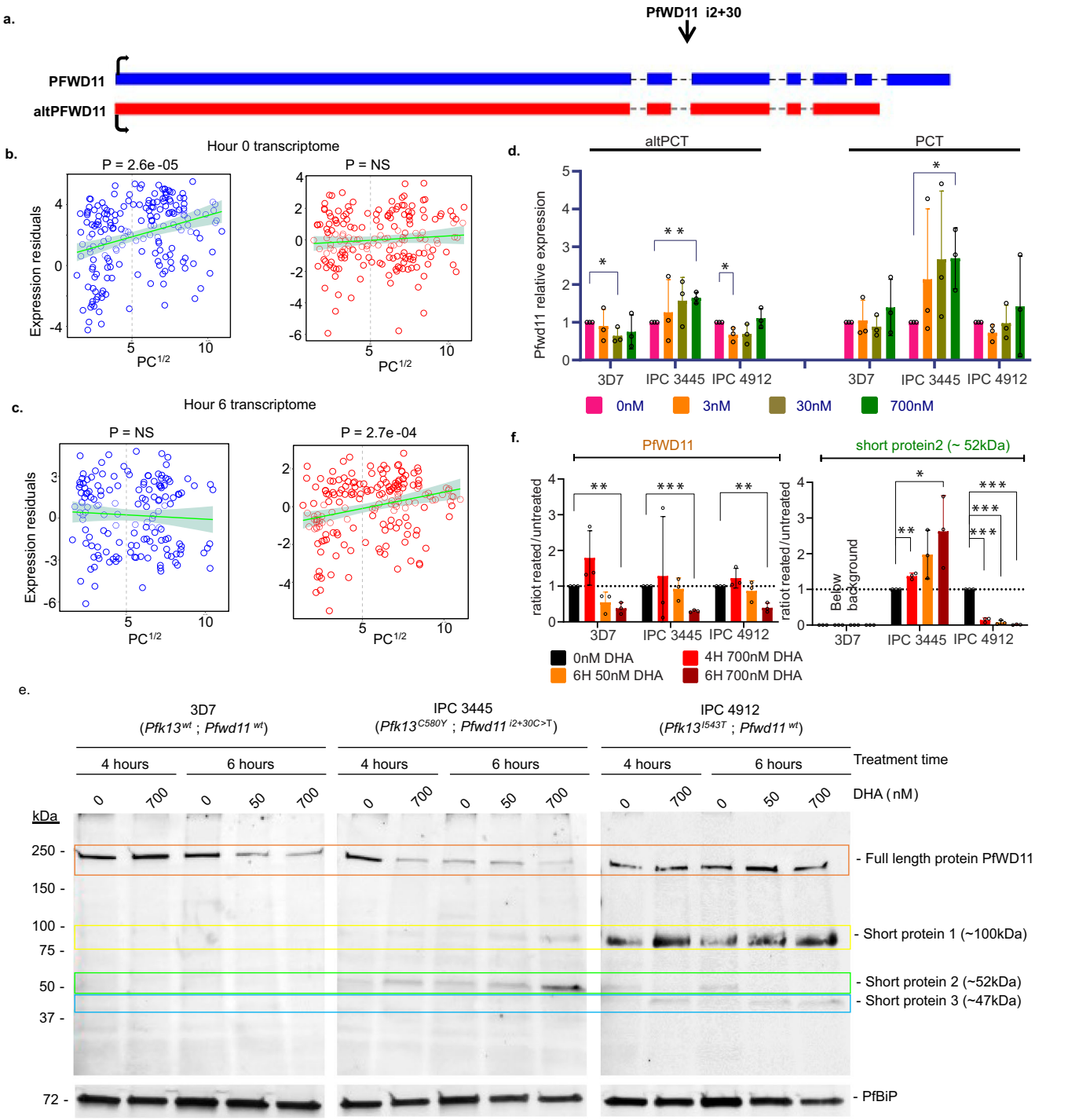


Fig. 3 | In vivo transcriptional response to TACT/ACT treatment. a Induction or repression of *PCTs* upon treatment with artemisinin combination therapy on parasite with PC1/2 < 5 hours (left) and PC1/2 < 5 hours resistant (right) parasite groups. Any transcripts with FDR <1e-05 considered as significant and commonly induced/repressed transcripts are marked as un-filled blue colored circles. The background and point color of the volcano plots are to differentiate the resistant parasites from susceptible parasites and *PCTs* from non-s respectively. White and beige background signifies the parasite groups as susceptible and resistant respectively, whereas blue circles here represent up-regulated/repressed *PCTs* with statistical significance. **b** Pathways linked with 555 PCTaPCTs suppression in the resistant parasites with $P < 0.05$ from the enrichment analysis. **c** An example of differentially regulated *asPCT* compared to corresponding PCT in the resistant parasites. Approximately scaled transcript structure of *PCT/asPCT* pair is shown above of violin plots. **d** Alternatively spliced PfRAD5 was found as repressed upon treatment in both susceptible and resistant parasite groups. Approximately scaled transcript structures shown above the violin plot reflect intron retention event in *altPCT* of PfRAD5. Source data are provided as a Source Data file. Panels **A**-**D** present the uncorrected *P* values that were calculated by fitting data to the Generalized Additive Model.

https://doi.org/10.1038/s41467-024-54915-6



the c-terminus that fall within its WD40 propeller domain (Fig. S16). Next, we carried out Westen blot analysis to test the effect of this aberrant splicing on the overall levels of PfWD11 protein (Fig. 4e, f, S17 and Supplementary Data 7). Using a polyclonal antibody targeting the PfWD11 protein region upstream of the putative truncation, we detected the full-length protein product (~ 250 kDa) in all three strains grown under normal conditions or after varying DHA treatments(s) (Fig. 4e). Curiously, 6 hr treatment with 700 nM DHA caused a suppression of the full-length PfWD11 levels in all three strains, with IPC3445 exhibiting a significantly more potent effect. Moreover, each strain exhibited a distinct pattern of truncated PfWD11 protein products, presumably reflecting the posttranslational processing and/or protein degradation (Fig. 4e). The abundance of all truncated protein products was insensitive to the DHA treatment, except a ~52 kDa protein band that exhibited dose and time-dependent induction upon DHA treatment (Fig. 4f). Altogether, these in vitro characterizations are consistent with the mutated form of PfWD11 exhibiting distinct gene expression characteristics that, in turn, could contribute to the artemisinin resistance given its strong association with this phenotype in vivo.

Indeed, in the resistant parasites, the upregulation of the *aPCT* in the 0 hr transcriptome and induction of *altPCT* in the 6 hr transcriptomes is covariant with the PfWD11 i2 + 30 T > C with statistical significance of 6.05e-09 and 1.26e-05, respectively. Moreover, in the TACT-CV cohort, PfWD11 i2 + 30 T > C is completely covariant, with the SNP *PfK13* mutations being particularly associated with the KEL1/

Fig. 4 | PfWD11 as a novel marker for artemisinin resistance. a Transcript structure of PRT/altPRT pair of *PfWD11* gene showing retention of intron 5 for the alternatively spliced isoform. The black arrow shows the position of intronic SNP associated with $PC^{1/2}$ in GWAS. **b** Transcriptome-phenotype association for hour 0 transcriptome shows up-regulation of *PCT* in resistant parasites (left, blue) compared to susceptible parasites, whereas no significant changes in the expression level were observed for altPCT (right, red). **c** Opposite to hour 0 transcriptome, in hour 6 transcriptome there were no significant changes in *PCT* expression observed between the resistant and susceptible group, but altPCT was found upregulated in resistant parasites. **d** Real-time PCR values for two artemisinin-resistant (IPC3445 and IPC4912) and one artemisinin-sensitive stain (3D7) after a 2-hour treatment with varying concentrations of Dihydroartemisinin (DHA). Values are shown as a ratio of PfWD11 expression compared to untreated control (0 nM DHA). DHA concentrations are indicated by different colors. Significance values were calculated using unpaired, two-tailed heteroscedastic *t*-test against untreated control. Values were obtained from three biological and three technical replicates. **e** Western blot images for the three tested strains are shown, illustrating the different levels of PfWD11 processing after 4-hour and 6-hour DHA treatments. In the 3D7 sensitive strain, only the full-length PfWD11 (orange) is present, with no significant additional bands detected over the background. In contrast, the resistant strains exhibit several smaller proteins, with a particular ~52 kDa band (Short protein 2, green) showing opposite regulation patterns in IPC3445 and IPC4912 in response to DHA treatment. **f** Graph showing the ratio of full-length PfWD11 abundance after DHA treatment compared to untreated control (left panel); Graph showing the ratio of the abundance of Short protein 2 after DHA treatment compared to its untreated control (right panel). Abundance was measured by densitometry in three biological replicates and values were normalized to PfBiP. Significance values were calculated using unpaired, two-tailed heteroscedastic *t*-test against untreated control. Source data are provided as a Source Data file. Panels **B**, **C** present the uncorrected *P* values that were calculated by fitting data to the Generalized Additive Model.

PLA1 genotype (Fig. 1e). This raises the question of whether PfWD11 i2 + 30 T > C is a product of recent emergence or originated at the onset of artemisinin resistance. Examination of the Catalogue of Genetic Variation in *P. falciparum* – v7.0 (Pf7)[58] showed that the majority *P. falciparum* population in most endemic areas exhibit ~2.9% ($n = 421$) minor allele frequency of PfWD11 i2 + 30 T > C from 14129 genotyped parasite isolates (Fig. 5a). The marked exceptions are eGMS and wGMS, where the PfWD11 i2 + 30 T > C frequency reached 37.7% and 11.3%, respectively. This is proportioned to the frequency of the mutant PfK13 genotype and the artemisinin resistance phenotype captured by the Pf7 samples set, with most *P. falciparum* infections outside the GMS still being fully sensitive to artemisinin. More specifically, ~80% of the mutant alleles in the Pf7 dataset originate from three countries from eGMS, including Cambodia, Vietnam, and Thailand, from 2005-2018. Indeed, the historical progression of the PfWD11 i2 + 30 T > C frequency in the GMS closely mirrors those of PfK13 such as C580Y (Fig. 5b). Before the emergence of artemisinin resistance (2005), PfWD11 i2 + 30 T > C and PfK13 C580Y exhibited low frequencies (< 5%) but were both rising through the following years to reach up to 80% in the GMS population by 2018. There was also a strong association between PfWD11 i2 + 30 T > C and the PC1/2 in TRAC I (2009-2011) and TRAC II (2016-2018) and finally TACT-CV (2018-2020) (Fig. 5c). In TRACI, conducted at the onset of artemisinin resistance in the GMS in 20011-2013, PfWD11 i2 + 30 T > C occurred in PfK13 wild-type background already in a small fraction of the parasite and it was not associated with PC1/2. However, already at that time, the vast proportion of the PfWD11 mutant alleles co-existed with PfK13 mutations and as such associated with PC1/2 > 5 hr. This trend progressed further from 2016 through 2020 (in TRACII and TACT-CV datasets), where the single mutant of PfWD11 nearly disappeared from the population and progressively larger proportions of parasites carried the PfWD11/PfK13 double mutant genotype; also associated with artemisinin resistance. Importantly, in the GMS population, PfWD11 i2 + 30 T > C was found to associate not only with C580Y, the historically most dominant genotype but also with at least two other PfK13 mutations, including R539T and Y493H SNP (Fig. 5d). To be specific, 72% of the parasite isolates carrying R539T and 22% of the parasite isolates carrying Y493H present in the Pf7 database also carry PfWD11 i2 + 30 T > C. These lines of evidence answer the initial question, suggesting that PfWD11 i2 + 30 T > C co-selected with resistance together with most if not all *PfK13* mutations since its emergence. As such *PfWD11* represents a key component of the artemisinin resistance phenotype.

Discussion

The main goal of this study was to expand our knowledge of molecular factors associated with the multifaceted artemisinin resistance mechanism and to monitor its evolution through its progressive spread within the GMS[59]. As opposed to previous GWAS/TPAS studies that surveyed the entire GMS region (and beyond), here we focused on the north-eastern region of eCambodia (Steung Treng) between 2018-2019 where the frequencies of artemisinin resistant *P. falciparum* infections were rising only gradually[52]. For that, we contrasted the eCambodian population with wCambodia and sVietnam, in which artemisinin resistant parasites fully dominated the *falciparum* malaria transmission, presumably as a result of the KEL1/PLA1 selective sweep[51,60]. To maximize the information content of this relatively limited size dataset (300 samples), we expanded the WGS to sequence polymorphisms in intragenic and intronic regions and the RNA-Seq, to capture alternatively spliced isoforms, antisense transcripts and ncRNAs. By capturing this broad spectrum of genomic/transcriptomics elements and correlating these with direct responses to artemisinin in vivo, we were able to make several key observations improving our understanding of artemisinin resistance including (i.) the ongoing admixture of resistant parasites into the drug-sensitive population as one of the key processes in spreading resistance; (ii.) at least two new genetic markers of artemisinin resistance, *PfRAD5* and *PfWD11*, the latter of which carries a mutation in its intronic region; (iii.) role of alternative splicing and antisense transcripts in in vivo adaptation of *P. falciparum* in its natural host including, drug resistance. We discuss these in the following section.

(i.) The genetic relatedness of the resistant parasite subpopulation in eCambodia indicates their origin in the KEL1/PLA1 lineage, hence, spreading from wCambodia and/or sVietnam[51,53,60]. In contrast, the drug susceptible eCambodia subpopulation exhibits a broad genetic diversity (Fig. 1b), unique genetic makeup (Fig. 1c), neutral selection status (Fig. 1d), and a distinct haplotype admixture structure (Fig. S3). This population is also characterized by a founder-like subpopulation structure previously observed in wCambodia, before the spread of artemisinin resistance[22]. This drug susceptible parasite population likely represents either an ancestral *P. falciparum* population (pre) existing in eCambodia and/or parasites that are spreading from the adjacent regions (such as southern Laos) where artemisinin resistance was less established[53,61]. Indeed, Artemether-lumefantrine (AL) is the first-line ACT in Laos that contrasts Cambodia, where Dihydroartemisinin-piperaquine (DHA-PPQ) was used up to 2014 and subsequently gradually changed to Artesunate-mefloquine (AS-MQ)[62]. From this perspective, TACT-CV represents a unique *P. falciparum* genetic cohort that illustrates how the main alleles admix into a drug susceptible population, presumably along with other parts of the artemisinin genetic background.

(ii.) In our GWAS, we found no associations between the PC1/2 and most of the SNPs identified in earlier studies, including that characterizing the genetic architecture of artemisinin resistance during its emergence between 2009 and 2013[23]. The exceptions are N1131I and N821K in PfRAD5, which were identified by this study and thus appear

https://doi.org/10.1038/s41467-024-54915-6

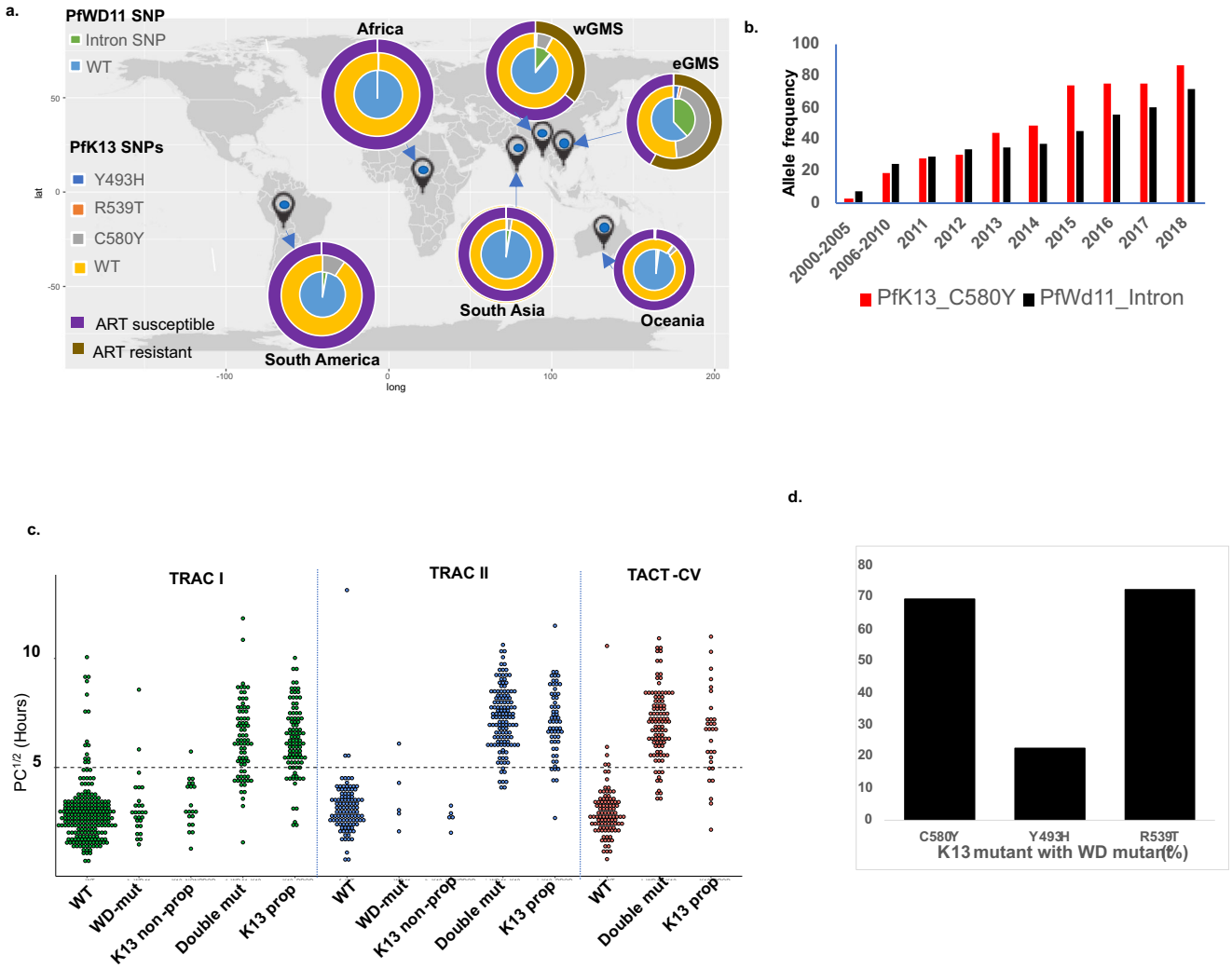


Fig. 5 | Allelic features of PfWD11 i2 + 30 T > C. a World-wide distribution of PfWD11 i2 + 30 C allele (mutant) on the background of the three PfK13 propeller domain mutations (found in this study) and resistance status from the Pf7 dataset. **b** Temporal changes of PfK13 C580Y and PfWD11 i2 + 30 C allele frequencies in the three countries of wGMS (Cambodia, Vietnam, and Thailand) showed an increasing trend. **c** PC1/2 plotted along with the different combinations of PfK13 - PfWD11 genotypes for three large-scale epidemiological studies: TRAC I, TRAC II, and TACT-CV. (WT = wild type, K13 non-prop= mutations in PfK13 but not in propeller domain, WD-mut= PfWD11 i2 + 30 T > C but no K13 mutations, Double_mut= PfWD11 i2 + 30 T > C along with K13 mutations in propeller domain, K13 prop = K13 mutations in propeller domain but PfWD11 i2 + 30 T > C absent). **d** Pf7 database shows that the PfWD11 i2 + 30 T > C allele highly coexisted with PfK13 C580Y and R539T alleles for >70% and Y493H allele carrying isolates. Source data are provided as a Source Data file.

as the only components of the artemisinin resistance genetic architecture that has persisted in the GMS for over a decade. Canonically, PfRAD5 represents one of the key factors of the DNA repair machinery playing roles in post-replication repair and template switching[63]. Indeed, DNA repair was previously linked with artemisinin resistance by in vitro analyses and genetic polymorphisms in several DNA repair factors were also detected in the natural infections[50,64–67]. The human homolog of Rad5 is known to inhibit autophagy-induced cell death by repairing reactive oxygen species (ROS)-induced DNA damage[68]. Curiously, artemisinin resistant *P. falciparum* is believed to escape a programmed cell death-like process[69] that is at least partially induced by ROS-mediated stress as a part of artemisinin MOA[70]. The results from this study now provide further evidence for PfRAD5 to play a part in the artemisinin resistance mechanism.

PfWD11 i2 + 30 T > C represents another SNP associated with artemisinin resistance that was not detected by any previous studies. Nevertheless, a retrospective analysis showed that PfWD11 i2 + 30 T > C withstood the test of time, being associated with essentially all *PfK13* alleles regardless of their origin since their emergence in the GMS. It is feasible that the i2 + 30 T > C mutation causes the intron retention directly by inactivating a putative intronic splicing enhancer and/or activating an intronic splicing silencer, as previously shown in other eukaryotic systems[71]. Based on a saturation mutagenesis by transposon study, *PfWD11* gene is essential for parasite intraerythrocytic development, playing a role in stress responses, gametocyte production and potentially artemisinin resistance phenotype[72]. Moreover multiomics study by Mok et al. et al 2021 shown downregulation of overall PfWD11 protein levels in CAM3.II isogenic PfK13 R539T and C580Y mutants at the trophozoites stages when compared to their isogenic wild-type controls[50]. Three most recent PiggyBac transposon studies revealed that a disruption in the intergenic region between *PfKIC5* 3′UTR (downstream neighboring gene) and *PfWD11* 5′UTR (774 bp and 734 bp distance from each gene, respectively) sensitized parasites to heat shock, DHA, oxidative stress and at the same time increased gametocyte conversion rates[73]. Another single disruption identified by this study occurred in a long non-coding RNA downstream from the *PfWD11* transcription start site sensitizing parasite to DHA, lumefantrine, heat shock, oxidative stress and proteasome inhibitor bortezomib. Although the precise biological function of PfWD11 remains unknown, there is evidence of it being a part of the TAF

epigenetic reader complex along with the TAF1/BDP5 chromatin-binding protein[74]. The key role of the TAF complex in many eukaryotes is in the regulation of cell cycle progression by transcriptional initiation[75]. Disruption of the TAF complex by TAF1/BDP5 knockdown was shown to lead to a general deceleration of the IDC and, at the same time, derepression of genes involved in sexual development[74]. Both phenomena were previously observed in artemisinin resistant *P. falciparum* natural infections[27,28,50,76].

Taken together, unlike *PfK13*, for which different alleles emerged, spread, and subsequently disappeared, being replaced by others, the SNPs in both and *PfWD11* remained constant throughout the last 15 years during which artemisinin resistance spread through the GMS. Indeed, PfRAD5 N1131I and PfWD11 i2 + 30 T > C show strong associations with PfK13 mutat phenotype when studied by GWAS of linkage disequilibrium (LD) studies in the TACT-CV cohort (Fig. S18a, b and Supplementary Data 8). Moreover, the temporal profile of their frequency progression in the GMS since 2009 also shows a high correlation with PfK13 C580Y, which classifies both mutations into a small group of outlier mutations with a high link with artemisinin resistance (Fig. S18c, d). Pending validations of their biological relevance, these may now represent more suitable markers of artemisinin resistance in the future. Currently, little is known about the putative genetic background of the *PfK13* mutant parasites emerging in Sub-Saharan Africa[11,77]. It will be particularly interesting to investigate if these alleles also contribute to artemisinin resistance in African parasite lines; owing to otherwise low frequencies of PfK13 β-propeller mutations in the African populations in general[21,78].

(iii.) One of the key observations made in this study is the role of *asPCTs* and *altPCTs (non-aPCTs)* and possibly ncRNAs in *P. falciparum* natural infections, probably driving the parasite's adaptation process, including drug resistance. These types of transcripts were previously shown by in vitro studies to play regulatory roles in the parasite life cycle progression and other *Plasmodium* traits[79–81], albeit still poorly understood. The discovery of *the non-PCT* and ncRNAs in *P, falciparum* dates to as early as 1991 and subsequently 2001 respectively[82,83]. It was, however, not until the advancement of RNA-Seq that these RNA species were captured at the genome-wide level, with a subset of them found to be developmentally regulated in *Plasmodium*[84,85]. Since then, alternative splicing was characterized also in *P. vivax*[86], and sub-species of *P. yoelii* and along with pre/post-exposure to mefloquine[87]. In *P. yoleii* particularly, alternative splicing was shown to regulate of surface antigen genes[88], while in *P. beghei* it is involved in the differentiation of male gametocytes and oocysts[89]. In *P. falciparum, lncRNAs* have been linked to several biological functions, including sexual conversion through antisense silencing[90], immune evasion by antigenic switching, mutual exclusiveness of CLAG3.1 and CLAG3.2 expression[91] and several other biological processes[92,93]. With the advances in long-read and single-molecule sequencing, it is clear now that the complexity of the *P. falciparum* transcriptome goes far beyond protein-coding mRNA[79,94]. Here, we show that both non-aPCTs and ncRNAs exhibit marked variations in *P. falciparum* parasites during their natural infections and, as such, can play key roles in host environment adaptation, including drug resistance. It is unclear whether alternative slicing in *P. falciparum* directly impacts the proteome diversity or if it has a regulatory role by modifying protein levels by (over)production of aberrant transcripts[81,95]. The alternatively spliced PfWD11 and PfRAD5 may represent such aberrant transcripts that could lead to speculated reductions at the protein level.

In future in vitro studies, exploring the casualty of the identified genetic factors for artemisinin resistance will be crucial. In particular, it will be crucial to assess whether these factors contribute directly to the multifaceted mechanism of artemisinin resistance[96] or contribute to its spread in the clinical/epidemiological setting. The latter may involve factors that can alleviate the fitness cost caused by different *PfK13* mutations in the asexual parasite growth[97,98] or boost of sexual stage induction observed even in the initial study of artemisinin resistance in the GMS[1]. The natural limitation of this study, similar to any association studies, is the incapability to distinguish these two categories of the artemisinins resistance from each other. In turn, however, here we identify new genetic factors that contribute to artemisinin resistance of *P. falciparum* malaria infection in real-life clinical settings, which, in our view, represent the main strength of this work.

Methods

Ethics statement

All samples used for this study were collected with written informed consent from the patients or their legal guardians. The study was approved by the National Ethics Committee for Health Research (NECHR; Phnom Penh, Cambodia); the Ethical Committee, Hospital for Tropical Diseases (Ho Chi Minh City, Vietnam; and Oxford University Tropical Research Ethics Committee (OXTREC; Oxford, UK). Oxford University was the study sponsor. Monitoring was done by the Mahidol-Oxford Tropical Medicine Research Unit (MORU; Bangkok, Thailand) and the Oxford University Clinical Research Unit (Ho Chi Minh City, Vietnam). Ethics approval was obtained from the NECHR, Cambodia (NECHR 0042); the Ethical Committee, Hospital for Tropical Diseases, Vietnam (1096); and OXTREC

In vivo sample collection

All samples used in this study were collected from 310 patients involved in a multi-site clinical trial at five hospitals or health centers in three locations (western Cambodia, eastern Cambodia, and Vietnam). The trial took place between March 18, 2018, and Jan 30, 2020. All details regarding sample collection, site locations, inclusion criteria, parasitemia assessment, and given treatments were previously published[52].

Total RNA extraction and measurements

8 volumes of TRIzol reagent (Invitrogen) (v/v) were added to 0.5 ml infected packed red blood cells. Samples were vortexed at 7000 RPM for 1 min. Chloroform was added in a 1:5 volumetric ratio, vortexed and samples were centrifuged for 15 min to separate phases as previously described[99,100]. The supernatant was aspirated and transferred to the new tubes. An equal volume of 100% ethanol was added to the supernatant and RNA was extracted using ZYMO DirectZol-96 RNA kit (ZYMO) following manufacturer instructions. RNA was eluted in 17 ul of RNAse/DNAse free water, and its purity was assessed by spectrometry on Nanodrop (Thermo Scientific). RNA concentration was estimated using RNA-specific Qubit fluorometric assay (Invitrogen). RNA integrity was assessed on Bioanalyzer RNA Nano Chip (Agilent).

gDNA extraction and measurements

gDNA was extracted from infected packed red blood cells by Mahidol Oxford Tropical Medicine Research Unit (Bangkok, Thailand) as described previously[20] and purified gDNA was shipped to Nanyang Technological University (Singapore) for further processing. Briefly, all DNA samples have been extracted from white blood cell-depleted samples using the QiaSymphony Investigator kit with an extraction volume of 200 uL. gDNA concentration was estimated using DNA-specific Quant-IT High Sensitivity DNA and Quant-IT Broad Range DNA fluorometric assays (Invitrogen). gDNA integrity was assessed on 0.5% agarose gel.

gDNA library preparation

gDNA sequencing libraries have been prepared from 1 ng of previously purified gDNA using a DNA Prep kit (Illumina) following the manufacturer's instructions with modification to the protocol described below. In brief, volumes of all reagents in the entire protocol have been scaled down four times to accommodate a larger number of samples. As recommended by the manufacturer due to low concertation of DNA in the samples manual normalization of gDNA input (1 ng gDNA) was performed using DNA-specific Quant-IT High Sensitivity DNA and Quant-IT

Broad Range DNA fluorometric assays (Invitrogen). 10 cycles of PCR have been used for final index incorporation. Unique Dual Indexes (Illumina) were used for multiplexing. The resulting libraries were eluted in 15 μl of the elution buffer and verified by real-time PCR, fluorometric DNA concentration assay (Invitrogen), and Bioanalyzer DNA High-Sensitivity chip (Agilent Technologies) by Novogene Singapore.

Stranded total RNA library preparation

RNA sequencing libraries have been prepared from 100 ng of purified total RNA using ZYMO-Seq RiboFree Total RNA Library Kit (ZYMO) following the manufacturer's instructions with some modifications to the protocol described below. This protocol allows for the generation of stranded, rRNA, and human globin-depleted libraries. The whole protocol has been scaled down to use half of the recommended volume of each reagent. The depletion step of rRNA (both human and the parasite) and globin has been adjusted to 4 hours at 68 °C. 14 cycles of PCR have been used for final index incorporation. The resulting libraries have been eluted in 15ul of the elution buffer and verified by real-time PCR, fluorometric DNA concentration assay (Invitrogen), and Bioanalyzer DNA High-Sensitivity chip (Agilent Technologies) by Novogene Singapore.

Library sequencing

96 libraries (either gDNA or total RNA) have been equimolarly pooled and sequenced per one lane of Novaseq S4 sequencing technology (Illumina) generating approximately 750GB of data per lane (Novogene, Singapore). Technical replicates of 3D7 lab strain gDNA/RNA were also included on each sequencing lane to control for library preparation and sequencing efficiency.

Variant detection and filtering

Raw sequencing reads were trimmed using the Trimgalore package specifying: -q 30 –phred33 -a adapter1 -a2 adapter2 –stringency 5 –trim-n -e 0.2 –length 75[101]. Trimmed reads were aligned to the *Plasmodium falciparum* 3D7 reference genome by the BWA package[102]. PCR duplicates were removed using PicardTools and bases were recalibrated using GATK BQSR using variants from genetic crosses described by - Miles et al., 2016[103]. Short variants (SNPs and Indels) were called using GATK HaplotypeCaller[104]. Samples with at least 1.5 million paired reads aligned to the 3D7 reference genome and with a minimum of 10X genome coverage are considered for further analysis. GATK hard filters with values MQ < 60.0 FS > 60.0 or QD < 2.0 were used to filter the data. SNPs detected in the core genome were retained further by excluding SNPs detected in the sub-telomere repeat regions. Samples with a missing genotype rate <0.1 were discarded. Further using Plink only bi-allelic SNPs with minor allele frequency > 0.01 and missing genotype rate <0.1 were retained for analysis[105].

Population structure analysis and GWAS

To investigate the population structure three individual analyses were performed using filtered SNP data. Principle component analysis was performed at first using Plink to visualize the variability among the isolates. Next, the pairwise distances between individual isolates were calculated and the Neighbour-Joining tree (NJ Tree) was plotted using APE packages[106]. Based on the PCA and Neighbour-joining tree results parasite populations were grouped into three sub-population/clades. For each population, the Tajima's D value was estimated for 5Kb window by VCFtools[107] and visualized as a density plot. Before performing the Tajima's D analysis isolates were filtered according to FWS value as estimated by the R package moimix (https://bahlolab.github.io/moimix/). Lastly, the ancestry of different populations inferred from PCA and NJ- tree were revisited by performing the Admixture analysis with ancestral component K = 3[108]. The value of K was inferred from the groups of isolates from PCA analysis and the division of clades in the NJ tree. Admixture-estimated ancestral shares for three populations were then linked with parasite clearance half-life and k13 mutation (a known marker for resistance). To link the SNP genotype with phenotype – parasite clearance half-life, the Fast-LMM package was used[109]. The first 10 principal components were used in this study as a covariate to correct for population structure and unknown confounders. Any SNP with a *p*-value of association less than 1×10^{-5} was examined.

Transcriptome assembly

Sequencing reads for h0 transcriptome and h6 transcriptome were first trimmed using Trimgalore package specifying: -q 20 --phred33 --clip_R2 10 -a adapter1 -a2 adapter2 --trim-n --length 75[101]. High-quality reads were aligned to the *Plasmodium falciparum* 3D7 reference genome (annotation version 59) with Hisat2[110]. Reads mapped to a maximum of three genomic regions were extracted and assembled using the Stringtie package with options: -p 50 --fr -m 150 -j 3 -g 20 for each sample[111]. To achieve a high-quality transcriptome, we first merged samples with (i) at least 6 M reads aligned ($n = 358$; including *Plasmodium falciparum* life cycle transcriptome with 2-hour resolution), (ii) isoform fraction 0.1, (iii) minimum transcript length 150. This transcriptome was then used for abundance estimation for each transcript and sample. Any sample with at least 70% transcript detected and any transcript present in at least 70% of samples were considered for further analysis.

Transcriptome phenotype association

The age of each parasite isolates (both h0 and h6) was first predicted by comparing it with IDC time points using Spearman's rank-order correlation in R (stats package cor.test function). A transcriptome-based approach to detect gametocyte proportion developed by Zhu et al., 2022 was implemented next to determine the overrepresentation of the gametocyte populations within each isolate[29]. We adapted the identical method described by Zhu et al., 2022[29] to detect transcriptomic markers for drug-resistant parasites from h0 transcriptome data. In brief, using r package - 'gam' (generalized additive model) expression of each transcript was associated with resistance for isolates with predicted hpi within 8-12. The gam function was designed to have two components (a) loess function fitted predicted HPI, (b) PC1/2. Using this method age as a confounding factor for expression change, up-regulated and down-regulated transcript in resistance parasites were detected. The r-script used to perform this analysis will be made available upon request. The H6 transcriptome data set was filtered using the same condition applied to the H0 transcriptome.

In-vivo transcriptome response

To examine parasite response through changes in transcriptional level, resistant (PC1/2 > 5 hours) were separated from susceptible parasites (PC1/2 < 5 hours). In total 251 susceptible parasites' transcriptomes (h0 = 137 and h6 = 114) and 284 resistant parasites' transcriptomes (h0 = 147 and h6 = 137) were used in this study. Again, we adapted the method described by Zhu et al., 2022[29] to detect transcriptomic response upon drug treatment for the detection of induced or repressed transcripts in the h6 transcriptome compared to the h0 transcriptome. In short, a generalized additive model in R was designed with three components (a) loess function fitted predicted HPI, (b) PC1/2, and (c) treatment group i.e., pre-treatment or post-treatment. Results were obtained for both resistant and susceptible parasite groups independently. Selected gene sets were examined for enriched pathways with a *p*-value of 0.05 by an in-house developed script. Both R-scripts (a) in-vivo response study, and (b) pathway enrichment study will be made available upon request.

qPCR detection of altPFWD11 in IDC and field isolates

We aimed to validate the presence of the alternative-spliced isoform of the WD-repeat-containing protein (PF3D7_1138800) in both laboratory-cultured PF3D7 strain and field samples utilized for RNA-Seq analysis. To achieve this, total RNA was treated with the TURBO

DNA-free™ Kit (Invitrogen) to eliminate any contaminating genomic DNA. Subsequently, cDNA was synthesized from the treated RNA using SuperScript™ II Reverse Transcriptase (Thermo Scientific). The qPCR reactions were then performed in triplicate, and the uniqueness of the PCR products was verified by examining the melting curve within the temperature range of 60–95 °C using the StepOne™ Real-Time PCR System (Applied Biosystems). Data were analyzed using the comparative critical threshold (ΔΔCt) method, in which the amount of target RNA was compared with that of Tubulin 2 (PF3D7_0422300), serving as an internal control. The expression levels of a target between groups of interest were compared using a *t*-test. To perform PCR amplification, the same primers as those used for qPCR were used with 2X DreamTaq Green PCR Master Mix (Thermo Scientific). The resulting PCR and qPCR products were verified using 2% agarose gel electrophoresis and Sanger sequencing. More comprehensive protocols and primer sequences are outlined in S1 Protocols.

Real-time quantitative PCR of in vitro drug treatments

Three different strains were used: 3D7, IPC3445 and IPC4912 (Bei Resources) and their genomes sequenced using Illumina short read technology as described above to verify genomic sequence variations. Cultures (0.2 ml pRBC, 2% hematocrit, 5% parasitemia, late ring stage approx. 20hpi) were treated for 2 hours with 0 nM, 3 nM, 30 nM, and 700 nM DHA. After treatment, infected pRBCs have been lysed with TRIZol and RNA extraction was performed as described above. Total RNA was treated with the TURBO DNA-free™ Kit (Invitrogen) to eliminate any contaminating genomic DNA. Subsequent quantitative PCR was performed as described in the section above. All experiments have been performed in three biological and three technical replicates.

Western blot analysis of PfWD11 protein

Western blots were performed following standard BioRad protocols for Trans-Blot Turbo System (available on manufacturer's website) with some modifications. Three different strains were used: 3D7, IPC3445 and IPC4912 (Bei Resources) and their genomes sequenced using Illumina short read technology as described above. Cultures (0.7 ml pRBC, 5% hematocrit, 10% parasitemia, late ring stage approx. 20hpi) were treated for 4 hours with 0 nM DHA and 700 nM DHA or for 6 hours with 0 nM, 50 nM, and 700 nM DHA. After treatment parasites were liberated with ice-cold saponin (Sigma) (0.1%) in 1x PBS (1st Base) and washed three times with ice-cold 1x PBS with protease inhibitor cocktail (1:10000). Parasite pellet was lysed in RIPA buffer supplemented 1:1000 (v/v) with EDTA-free protease inhibitor (Nacalai Tesque). Protein concertation of the clarified lysate was measured using Pierce BCA Protein Assay Kit (Thermo Scientific). An identical amount of total protein (25 µg) was mixed with Laemmeli Sample Buffer (6X) (Thermo Scientific) supplemented with 2-mercaptoethanol (10%) (Sigma) incubated at 4°C for 1 hour. Normalized volumes of identical amounts of protein lysates were separated using premade Mini-PROTEAN TGX Precast Stain-free 4-20% gradient SDS-PAGE gel (Bio-Rad). Gels were activated for 45 seconds under UV light. Proteins were transferred to the PVDF membrane (BioRad) using Trans-blot Turbo Transfer System using manufacturer's original reagents (BioRad) for 15 minutes at 2.5 A constant current. Transfer efficiency was examined under UV light. Membranes were blocked with 5% (w/v) solution of Bovine Serum Albumin (BSA) in TBS-T (Sigma). For PfWD11 a custom made polyclonal anti-mouse antibody has been synthetized by Genscript Singapore. Antibody was designed to target protein region upstream of the possible truncation event encompassing the C-terminal end of predicted TAF5-like domain and N-terminal end of WD40 domain (Fig. S17). Antigen used for immunization was expressed in *E.coli* for subsequent mouse immunization (ENFKNQYLKKFS AYEPVTKFLLTCVSLTQIFEISLIRFKKKNSKHIVHMTKLGKKSLLQYLSVYE GIPLYNIITTKIKIIEIDDTKRNFNFFYAFVSSNFFYNVKLTFPVQWNLPSIYLT NDEIVEDENKKSLTEKEQNNYVPNENNDAYYYYKHILKTQATNRLRVTKK NIPSILYYCLNNCNDLTCAELSGYDGSLIATAHSNNIIKLWNLKKSEMNKV MNEKRDIDEINEYMDDVR). Antibody was affinity purified and tested for recognizing epitope in denaturing conditions by using purified antigen. Antibody was diluted 1:500 (3.18 µg/µl) and probing was carried on in 3% BSA TBS-T solution overnight at 4°C. Anti-PfBiP (rabbit, 1:5000, BEI Resources) was used as an internal control. Primary antibodies were probed using anti-mouse Highly Cross Absorbed Alexa Fluor 647 1:10000 (Thermo Scientific, #A32728, WE322197) and anti-rabbit Highly Cross Absorbed Alexa Fluor 488 1:10000 (Jackson Immunoresearch, # 115-545-062). All antibodies have been tested for cross-reactivity and signal leakage. Resulting fluorescence signal was acquired using ChemiDoc MP (Bio-Rad) and analyzed using Image Lab (Bio-Rad). Complete images of all western blot membranes can be found in Supplementary materials.

Population-level analysis of PfWD11 i2 + 30 T > C mutation

To investigate the PfWD11 i2 + 30 T > C mutation from a global perspective we first downloaded the genome-wide variant presented in Pf7 database[58]. We classified the parasites into 6 classes by geographical regions (a) wGMS (Pf7 identifier AS-SE-W) (b) eGMS (Pf7 identifier AS-SE-E) (c) Africa (Pf7 African isolates) (d) South America (Pf7 identifier SA) (e) Oceania (Pf7 identifier OC-NG) and, (f) South Asia (Pf7 south Asian isolates). Allele frequencies were calculated for three PfK13 mutations detected in this study as well as PfWD11 i2 + 30 T > C mutation disregarding the missing genotype information. Frequency distribution for PfK13 C580Y and PfWD11 i2 + 30 T > C was specifically plotted for three countries in eGMS over time.

Further relation between allelic combination and PC1/2 from other two large-scale epidemiological studies TRAC I and TRAC II were visualized in the context of TACT-CV study.

In-silico structure prediction of altPfWD11

AlphaFold predicted structure of PfWD11 was downloaded from the PlasmoDB database[112]. The native form of PfWD11 was found to have a Kelch13 propeller-like domain at the c-terminal which was expected to get disrupted in altPfWD11 because of retention of intron 5 which leads to early truncation of the protein product as predicted. Hence the structure of the C-terminal altPfWD11 covering the propeller-like domain was predicted using AlphaFold[113].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw genome sequence generated in this study have been deposited in the NCBI SRA database under the bio-project accession code PRJNA1011501 (https://www.ncbi.nlm.nih.gov/bioproject/1011501). The RNA Sequence data generated in this study have been deposited in the NCBI GEO database under accession code GSE24254. Source data are provided with this paper.

Code availability

All the codes used in the analysis, including the packages in R, are provided via github https://github.com/sourav-bioinfo/tact_cv_analysis_protocol.

References

1. Dondorp, A. M. et al. Artemisinin resistance in Plasmodium falciparum malaria. *N. Engl. J. Med* **361**, 455–467 (2009).
2. Noedl, H. et al. Evidence of artemisinin-resistant malaria in western Cambodia. *N. Engl. J. Med* **359**, 2619–2620 (2008).
3. Jacob, C. G. et al. Genetic surveillance in the Greater Mekong subregion and South Asia to support malaria control and elimination. *Elife* **10** https://doi.org/10.7554/eLife.62997 (2021).

4. Conrad, M. D. et al. Evolution of Partial Resistance to Artemisinins in Malaria Parasites in Uganda. *N. Engl. J. Med* **389**, 722–732 (2023).
5. Fidock, D. A. & Rosenthal, P. J. Artemisinin resistance in Africa: How urgent is the threat? *Med* **2**, 1287–1288 (2021).
6. Chenet, S. M. et al. Independent Emergence of the Plasmodium falciparum Kelch Propeller Domain Mutant Allele C580Y in Guyana. *J. Infect. Dis.* **213**, 1472–1475 (2016).
7. Mathieu, L. C. et al. Local emergence in Amazonia of Plasmodium falciparum k13 C580Y mutants associated with in vitro artemisinin resistance. *Elife* **9** https://doi.org/10.7554/eLife.51015 (2020).
8. Ashley, E. A. et al. Spread of artemisinin resistance in Plasmodium falciparum malaria. *N. Engl. J. Med* **371**, 411–423 (2014).
9. Witkowski, B. et al. Novel phenotypic assays for the detection of artemisinin-resistant Plasmodium falciparum malaria in Cambodia: in-vitro and ex-vivo drug-response studies. *Lancet Infect. Dis.* **13**, 1043–1049 (2013).
10. Straimer, J. et al. Drug resistance. K13-propeller mutations confer artemisinin resistance in Plasmodium falciparum clinical isolates. *Science* **347**, 428–431 (2015).
11. Uwimana, A. et al. Association of Plasmodium falciparum kelch13 R561H genotypes with delayed parasite clearance in Rwanda: an open-label, single-arm, multicentre, therapeutic efficacy study. *Lancet Infect. Dis.* **21**, 1120–1128 (2021).
12. Asua, V. et al. Changing Prevalence of Potential Mediators of Aminoquinoline, Antifolate, and Artemisinin Resistance Across Uganda. *J. Infect. Dis.* **223**, 985–994 (2021).
13. Balikagala, B. et al. Evidence of Artemisinin-Resistant Malaria in Africa. *N. Engl. J. Med* **385**, 1163–1171 (2021).
14. Ndwiga, L. et al. A review of the frequencies of Plasmodium falciparum Kelch 13 artemisinin resistance mutations in Africa. *Int J. Parasitol. Drugs Drug Resist* **16**, 155–161 (2021).
15. Mihreteab, S. et al. Increasing Prevalence of Artemisinin-Resistant HRP2-Negative Malaria in Eritrea. *N. Engl. J. Med* **389**, 1191–1202 (2023).
16. Juliano, J. J. et al. Country wide surveillance reveals prevalent artemisinin partial resistance mutations with evidence for multiple origins and expansion of high level sulfadoxine-pyrimethamine resistance mutations in northwest Tanzania. *medRxiv* https://doi.org/10.1101/2023.11.07.23298207 (2023).
17. Takala-Harrison, S. et al. Independent emergence of artemisinin resistance mutations among Plasmodium falciparum in Southeast Asia. *J. Infect. Dis.* **211**, 670–679 (2015).
18. Woodrow, C. J. & White, N. J. The clinical impact of artemisinin resistance in Southeast Asia and the potential for future spread. *FEMS Microbiol Rev.* **41**, 34–48 (2017).
19. Phyo, A. P. et al. Declining Efficacy of Artemisinin Combination Therapy Against P. Falciparum Malaria on the Thai-Myanmar Border (2003-2013): The Role of Parasite Genetic Factors. *Clin. Infect. Dis.* **63**, 784–791 (2016).
20. van der Pluijm, R. W. et al. Determinants of dihydroartemisinin-piperaquine treatment failure in Plasmodium falciparum malaria in Cambodia, Thailand, and Vietnam: a prospective clinical, pharmacological, and genetic study. *Lancet Infect. Dis.* **19**, 952–961 (2019).
21. Stokes, B. H. et al. Plasmodium falciparum K13 mutations in Africa and Asia impact artemisinin resistance and parasite fitness. *Elife* **10** https://doi.org/10.7554/eLife.66277 (2021).
22. Miotto, O. et al. Multiple populations of artemisinin-resistant Plasmodium falciparum in Cambodia. *Nat. Genet* **45**, 648–655 (2013).
23. Miotto, O. et al. Genetic architecture of artemisinin-resistant Plasmodium falciparum. *Nat. Genet* **47**, 226–234 (2015).
24. Cheeseman, I. H. et al. A major genome region underlying artemisinin resistance in malaria. *Science* **336**, 79–82 (2012).
25. Takala-Harrison, S. et al. Genetic loci associated with delayed clearance of Plasmodium falciparum following artemisinin treatment in Southeast Asia. *Proc. Natl Acad. Sci. USA* **110**, 240–245 (2013).
26. Cerqueira, G. C. et al. Longitudinal genomic surveillance of Plasmodium falciparum malaria parasites reveals complex genomic architecture of emerging artemisinin resistance. *Genome Biol.* **18**, 78 (2017).
27. Mok, S. et al. Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science* **347**, 431–435 (2015).
28. Mok, S. et al. Artemisinin resistance in Plasmodium falciparum is associated with an altered temporal pattern of transcription. *BMC Genomics* **12**, 391 (2011).
29. Zhu, L. et al. Artemisinin resistance in the malaria parasite, Plasmodium falciparum, originates from its initial transcriptional response. *Commun. Biol.* **5**, 274 (2022).
30. Zhu, L. et al. The origins of malaria artemisinin resistance defined by a genetic and transcriptomic background. *Nat. Commun.* **9**, 5158 (2018).
31. Kucharski, M. et al. Short tandem repeat polymorphism in the promoter region of cyclophilin 19B drives its transcriptional upregulation and contributes to drug resistance in the malaria parasite Plasmodium falciparum. *PLoS Pathog.* **19**, e1011118 (2023).
32. Rosenthal, M. & Ng, C. Parasite proteostasis and artemisinin resistance. *Res Sq.* https://doi.org/10.21203/rs.3.rs-2926003/v1 (2023).
33. Rosenthal, M. R. & Ng, C. L. Plasmodium falciparum Artemisinin Resistance: The Effect of Heme, Protein Damage, and Parasite Cell Stress Response. *ACS Infect. Dis.* **6**, 1599–1614 (2020).
34. Talman, A. M., Clain, J., Duval, R., Menard, R. & Ariey, F. Artemisinin Bioactivity and Resistance in Malaria Parasites. *Trends Parasitol.* **35**, 953–963 (2019).
35. Tilley, L., Straimer, J., Gnadig, N. F., Ralph, S. A. & Fidock, D. A. Artemisinin Action and Resistance in Plasmodium falciparum. *Trends Parasitol.* **32**, 682–696 (2016).
36. Birnbaum, J. et al. A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. *Science* **367**, 51–59 (2020).
37. Oberstaller, J. et al. Integration of population and functional genomics to understand mechanisms of artemisinin resistance in Plasmodium falciparum. *Int J. Parasitol. Drugs Drug Resist* **16**, 119–128 (2021).
38. Simmons, C. F. et al. Protein KIC5 is a novel regulator of artemisinin stress response in the malaria parasite Plasmodium falciparum. *Sci. Rep.* **13**, 399 (2023).
39. Demas, A. R. et al. Mutations in Plasmodium falciparum actin-binding protein coronin confer reduced artemisinin susceptibility. *Proc. Natl Acad. Sci. USA* **115**, 12799–12804 (2018).
40. Delandre, O. et al. Absence of association between polymorphisms in the pfcoronin and pfk13 genes and the presence of Plasmodium falciparum parasites after treatment with artemisinin derivatives in Senegal. *Int J. Antimicrob. Agents* **56**, 106190 (2020).
41. Sharma, A. I., Demas, A. R., Hartl, D. L. & Wirth, D. F. Reply to Velavan et al.: Polymorphisms of pfcoronin in natural populations: Implications for functional significance. *Proc. Natl Acad. Sci. USA* **116**, 12613–12614 (2019).
42. Velavan, T. P., Nderu, D., Agbenyega, T., Ntoumi, F. & Kremsner, P. G. An alternative dogma on reduced artemisinin susceptibility: A new shadow from east to west. *Proc. Natl Acad. Sci. USA* **116**, 12611–12612 (2019).
43. Sharma, A. I. et al. Genetic background and PfKelch13 affect artemisinin susceptibility of PfCoronin mutants in Plasmodium falciparum. *PLoS Genet* **16**, e1009266 (2020).
44. Gnadig, N. F. et al. Insights into the intracellular localization, protein associations and artemisinin resistance properties of Plasmodium falciparum K13. *PLoS Pathog.* **16**, e1008482 (2020).

45. Leber, W. et al. A unique phosphatidylinositol 4-phosphate 5-kinase is activated by ADP-ribosylation factor in Plasmodium falciparum. *Int J. Parasitol.* **39**, 645–653 (2009).
46. Rocamora, F. et al. Oxidative stress and protein damage responses mediate artemisinin resistance in malaria parasites. *PLoS Pathog.* **14**, e1006930 (2018).
47. Tinto-Font, E. et al. A heat-shock response regulated by the PfAP2-HS transcription factor protects human malaria parasites from febrile temperatures. *Nat. Microbiol* **6**, 1163–1174 (2021).
48. Miao, J. et al. The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite Plasmodium falciparum. *Mol. Microbiol* **78**, 883–902 (2010).
49. Zhang, M. et al. The apicoplast link to fever-survival and artemisinin-resistance in the malaria parasite. *Nat. Commun.* **12**, 4563 (2021).
50. Mok, S. et al. Artemisinin-resistant K13 mutations rewire Plasmodium falciparum's intra-erythrocytic metabolic program to enhance survival. *Nat. Commun.* **12**, 530 (2021).
51. Hamilton, W. L. et al. Evolution and expansion of multidrug-resistant malaria in southeast Asia: a genomic epidemiology study. *Lancet Infect. Dis.* **19**, 943–951 (2019).
52. Peto, T. J. et al. Triple therapy with artemether-lumefantrine plus amodiaquine versus artemether-lumefantrine alone for artemisinin-resistant, uncomplicated falciparum malaria: an open-label, randomised, multicentre trial. *Lancet Infect. Dis.* **22**, 867–878 (2022).
53. Imwong, M. et al. The spread of artemisinin-resistant Plasmodium falciparum in the Greater Mekong subregion: a molecular epidemiology observational study. *Lancet Infect. Dis.* **17**, 491–497 (2017).
54. Amato, R. et al. Genetic markers associated with dihydroartemisinin-piperaquine failure in Plasmodium falciparum malaria in Cambodia: a genotype-phenotype association study. *Lancet Infect. Dis.* **17**, 164–173 (2017).
55. Witkowski, B. et al. A surrogate marker of piperaquine-resistant Plasmodium falciparum malaria: a phenotype-genotype association study. *Lancet Infect. Dis.* **17**, 174–183 (2017).
56. Wendler, J. P. et al. A genome wide association study of Plasmodium falciparum susceptibility to 22 antimalarial drugs in Kenya. *PLoS One* **9**, e96486 (2014).
57. Kang, Y. J. et al. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res* **45**, W12–W16 (2017).
58. MalariaGen et al. Pf7: an open dataset of Plasmodium falciparum genome variation in 20,000 worldwide samples. *Wellcome Open Res* **8**, 22 (2023).
59. Muller, O., Lu, G. Y. & von Seidlein, L. Geographic expansion of artemisinin resistance. *J Travel Med* **26** https://doi.org/10.1093/jtm/taz030 (2019).
60. van der Pluijm, R. W. et al. Triple artemisinin-based combination therapies versus artemisinin-based combination therapies for uncomplicated Plasmodium falciparum malaria: a multicentre, open-label, randomised clinical trial. *Lancet* **395**, 1345–1360 (2020).
61. Imwong, M. et al. Molecular epidemiology of resistance to antimalarial drugs in the Greater Mekong subregion: an observational study. *Lancet Infect. Dis.* **20**, 1470–1480 (2020).
62. Global Malaria Programme, T. W. H. O. *Artemisinin resistance and artemisinin-based combination therapy efficacy*, <https://www.who.int/docs/default-source/documents/publications/gmp/who-cds-gmp-2018-26-eng.pdf> (2018).
63. Elserafy, M., Abugable, A. A., Atteya, R. & El-Khamisy, S. F. Rad5, HLTF, and SHPRH: A Fresh View of an Old Story. *Trends Genet* **34**, 574–577 (2018).
64. Gibbons, J. et al. Altered expression of K13 disrupts DNA replication and repair in Plasmodium falciparum. *BMC Genomics* **19**, 849 (2018).
65. Goyal, M. et al. An SR protein is essential for activating DNA repair in malaria parasites. *J Cell Sci* **134** https://doi.org/10.1242/jcs.258572 (2021).
66. Gupta, D. K., Patra, A. T., Zhu, L., Gupta, A. P. & Bozdech, Z. DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. *Sci. Rep.* **6**, 23603 (2016).
67. Maneekesorn, S. et al. Deletion of Plasmodium falciparum ubc13 increases parasite sensitivity to the mutagen, methyl methanesulfonate and dihydroartemisinin. *Sci. Rep.* **11**, 21791 (2021).
68. Piao, S. et al. ALDH1A1 and HLTF modulate the activity of lysosomal autophagy inhibitors in cancer cells. *Autophagy* **13**, 2056–2071 (2017).
69. Dogovski, C. et al. Targeting the cell stress response of Plasmodium falciparum to overcome artemisinin resistance. *PLoS Biol.* **13**, e1002132 (2015).
70. Meshnick, S. R. Artemisinin: mechanisms of action, resistance and toxicity. *Int J. Parasitol.* **32**, 1655–1660 (2002).
71. Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev. Biochem* **84**, 291–323 (2015).
72. Zhang, M. et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360** https://doi.org/10.1126/science.aap7847 (2018).
73. Chawla, J. et al. Phenotypic Screens Identify Genetic Factors Associated with Gametocyte Development in the Human Malaria Parasite Plasmodium falciparum. *Microbiol Spectr.* **11**, e0416422 (2023).
74. Hoeijmakers, W. A. M. et al. Epigenetic reader complexes of the human malaria parasite, Plasmodium falciparum. *Nucleic Acids Res* **47**, 11574–11588 (2019).
75. Davidson, I., Kobi, D., Fadloun, A. & Mengus, G. New insights into TAFs as regulators of cell cycle and signaling pathways. *Cell Cycle* **4**, 1486–1490 (2005).
76. Witmer, K. et al. Transmission of Artemisinin-Resistant Malaria Parasites to Mosquitoes under Antimalarial Drug Pressure. *Antimicrob Agents Chemother* **65** https://doi.org/10.1128/AAC.00898-20 (2020).
77. Uwimana, A. et al. Emergence and clonal expansion of in vitro artemisinin-resistant Plasmodium falciparum kelch13 R561H mutant parasites in Rwanda. *Nat. Med* **26**, 1602–1608 (2020).
78. Kayiba, N. K. et al. Spatial and molecular mapping of Pfkelch13 gene polymorphism in Africa in the era of emerging Plasmodium falciparum resistance to artemisinin: a systematic review. *Lancet Infect. Dis.* **21**, e82–e92 (2021).
79. Li, Y., Baptista, R. P. & Kissinger, J. C. Noncoding RNAs in Apicomplexan Parasites: An Update. *Trends Parasitol.* **36**, 835–849 (2020).
80. Simantov, K., Goyal, M. & Dzikowski, R. Emerging biology of non-coding RNAs in malaria parasites. *PLoS Pathog.* **18**, e1010600 (2022).
81. Yeoh, L. M., Lee, V. V., McFadden, G. I. & Ralph, S. A. Alternative Splicing in Apicomplexan Parasites. *mBio* **10** (2019).
82. Knapp, B., Nau, U., Hundt, E. & Kupper, H. A. Demonstration of alternative splicing of a pre-mRNA expressed in the blood stage form of Plasmodium falciparum. *J. Biol. Chem.* **266**, 7148–7154 (1991).
83. Patankar, S., Munasinghe, A., Shoaibi, A., Cummings, L. M. & Wirth, D. F. Serial analysis of gene expression in Plasmodium falciparum reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol. Biol. Cell* **12**, 3114–3125 (2001).
84. Otto, T. D. et al. New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq. *Mol. Microbiol* **76**, 12–24 (2010).
85. Sorber, K., Dimon, M. T. & DeRisi, J. L. RNA-Seq analysis of splicing in Plasmodium falciparum uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Res* **39**, 3820–3835 (2011).

86. Zhu, L. et al. New insights into the Plasmodium vivax transcriptome using RNA-Seq. *Sci. Rep.* **6**, 20498 (2016).
87. Li, J. et al. UTR introns, antisense RNA and differentially spliced transcripts between Plasmodium yoelii subspecies. *Malar. J.* **15**, 30 (2016).
88. Fonager, J. et al. Transcription and alternative splicing in the yir multigene family of the malaria parasite Plasmodium y. yoelii: identification of motifs suggesting epigenetic and post-transcriptional control of RNA expression. *Mol. Biochem Parasitol.* **156**, 1–11 (2007).
89. Yeoh, L. M. et al. Alternative splicing is required for stage differentiation in malaria parasites. *Genome Biol.* **20**, 151 (2019).
90. Filarsky, M. et al. GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science* **359**, 1259–1263 (2018).
91. Mira-Martinez, S. et al. Expression of the Plasmodium falciparum Clonally Variant clag3 Genes in Human Infections. *J. Infect. Dis.* **215**, 938–945 (2017).
92. Gadalla, N. B. et al. Alternatively spliced transcripts and novel pseudogenes of the Plasmodium falciparum resistance-associated locus pfcrt detected in East African malaria patients. *J. Antimicrob. Chemother.* **70**, 116–123 (2015).
93. Subudhi, A. K. et al. Natural antisense transcripts in Plasmodium falciparum isolates from patients with complicated malaria. *Exp. Parasitol.* **141**, 39–54 (2014).
94. Shaw, P. J. et al. Transcriptomic complexity of the human malaria parasite Plasmodium falciparum revealed by long-read sequencing. *PLoS One* **17**, e0276956 (2022).
95. Bunnik, E. M. et al. Polysome profiling reveals translational control of gene expression in the human malaria parasite Plasmodium falciparum. *Genome Biol.* **14**, R128 (2013).
96. Kucharski, M., Nayak, S., Gendrot, M., Dondorp, A. M. & Bozdech, Z. Peeling the onion: how complex is the artemisinin resistance genetic trait of malaria parasites? *Trends Parasitol.* **40**, 970–986 (2024).
97. Behrens, H. M. et al. Impact of different mutations on Kelch13 protein levels, ART resistance, and fitness cost in Plasmodium falciparum parasites. *mBio* **15**, e0198123 (2024).
98. Mok, S. et al. Mapping the genomic landscape of multidrug resistance in Plasmodium falciparum and its impact on parasite fitness. *Sci. Adv.* **9**, eadi2364 (2023).
99. Kucharski, M. et al. A comprehensive RNA handling and transcriptomics guide for high-throughput processing of Plasmodium blood-stage samples. *Malar. J.* **19**, 363 (2020).
100. Chomczynski, P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* **15**, 536–537 (1993).
101. Krueger, F. TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).
102. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
103. Miles, A. et al. Indels, structural variation, and recombination drive genomic diversity in Plasmodium falciparum. *Genome Res* **26**, 1288–1299 (2016).
104. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* **43**, 11 10 11–11 10 33 (2013).
105. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
106. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
107. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
108. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* **19**, 1655–1664 (2009).
109. Lippert, C. et al. FaST linear mixed models for genome-wide association studies. *Nat. Methods* **8**, 833–835 (2011).
110. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
111. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
112. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444 (2022).
113. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Acknowledgements

The authors thank the funding agencies for supporting this work, including the Singapore Ministry of Education (grant # MOE2019-T3-1-007) (Z.B.) and The Singapore National Medical Research Council (grant # MOH-001107) (Z.B.).

Author contributions

S.N. designed and performed all the computational analyses for this study, prepared the figures, and drafted the manuscript. T.J.P. supervised all field-related work to obtain samples. M.K. performed all lab experiments, including the design of the experimental protocols, and necessary validations, and helped to analyze the data and draft the manuscript and figures. Z.B. designed the overall study, performed the analysis, supervised the analysis, and drafted the manuscript, obtained the funding to perform this study. A.M.D. participated in the design and supervised the project and gave vital comments to the discussion. D.T.Q.H. helped to perform qPCR, Real-time quantitative PCR, and Western Blot experiments. X.R. gave vital comments on the RNA-related work. M.G. provided the three *Plasmodium falciparum* strains to perform the validation experiments. R.T., J.J.C, D.L., H.D.T.N., R.W.P., N.D., L.T.L., R.V., H.R., N.H.C. contributed to field-related work to obtain the *Plasmodium falciparum* samples. O.M., M.M., M.D., L.S., M.I., N.P.J.D., N.J.W. participated in the design, supervised the sample collection and organization process, and gave vital comments to the draft. The manuscript was reviewed by all the authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-54915-6.

Correspondence and requests for materials should be addressed to Arjen M. Dondorp or Zbynek Bozdech.

Peer review information *Nature Communications* thanks Mike Ferdig, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Sourav Nayak[1,14], Thomas J. Peto[2,3,14], Michal Kucharski[1,4,14], Rupam Tripura[2,3], James J. Callery[2,3], Duong Tien Quang Huy[1], Mathieu Gendrot[1], Dysoley Lek[5,6], Ho Dang Trung Nghia[7,8], Rob W. van der Pluijm[2,3,9], Nguyen Dong[10], Le Thanh Long[11], Ranitha Vongpromek[2,12], Huy Rekol[4], Nguyen Hoang Chau[6], Olivo Miotto[2,3], Mavuto Mukaka[2,3], Mehul Dhorda[2,3,12], Lorenz von Seidlein[2,3], Mallika Imwong[2,13], Xavier Roca[1], Nicholas P. J. Day[2,3], Nicholas J. White[2,3], Arjen M. Dondorp[2,3,8] ✉ & Zbynek Bozdech[1] ✉

[1]School of Biological Sciences, Nanyang Technological University, Singapore, Singapore. [2]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [3]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom. [4]Amsterdam UMC, University of Amsterdam, Department of Global Health, Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands. [5]Centre for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia. [6]National Institute for Public Health, Phnom Penh, Cambodia. [7]Oxford University Clinical Research Unit, Hospital for Tropical Diseases, Ho Chi Minh City, Vietnam. [8]Pham Ngoc Thach University of Medicine, Ho Chi Minh City, Vietnam. [9]Institut Pasteur, Université Paris Cité, G5 Infectious Disease Epidemiology and Analytics, Paris, France. [10]Khanh Hoa Hospital for Tropical diseases, Ho Chi Minh City, Khanh Hoa province, Vietnam. [11]Phuoc Long Hospital, Ho Chi Minh City, Binh Phuoc province, Vietnam. [12]WorldWide Antimalarial Resistance Network - Asia-Pacific Regional Centre, Bangkok, Thailand. [13]Department of Molecular Tropical Medicine and Genetics, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [14]These authors contributed equally: Sourav Nayak, Thomas J. Peto, Michal Kucharski. ✉e-mail: arjen@tropmedres.ac; zbozdech@ntu.edu.sg