nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-021-00940-w

Check for updates

A heat-shock response regulated by the PfAP2-HS transcription factor protects human malaria parasites from febrile temperatures

Elisabet Tintó-Font[1], Lucas Michel-Todó[1,7], Timothy J. Russell[2,7], Núria Casas-Vila[1], David J. Conway[3], Zbynek Bozdech[4], Manuel Llinás[2,5] and Alfred Cortés[1,6]✉

Periodic fever is a characteristic clinical feature of human malaria, but how parasites survive febrile episodes is not known. Although the genomes of *Plasmodium* species encode a full set of chaperones, they lack the conserved eukaryotic transcription factor HSF1, which activates the expression of chaperones following heat shock. Here, we show that PfAP2-HS, a transcription factor in the ApiAP2 family, regulates the protective heat-shock response in *Plasmodium falciparum*. PfAP2-HS activates the transcription of *hsp70-1* and *hsp90* at elevated temperatures. The main binding site of PfAP2-HS in the entire genome coincides with a tandem G-box DNA motif in the *hsp70-1* promoter. Engineered parasites lacking PfAP2-HS have reduced heat-shock survival and severe growth defects at 37 °C but not at 35 °C. Parasites lacking PfAP2-HS also have increased sensitivity to imbalances in protein homeostasis (proteostasis) produced by artemisinin, the frontline antimalarial drug, or the proteasome inhibitor epoxomicin. We propose that PfAP2-HS contributes to the maintenance of proteostasis under basal conditions and upregulates specific chaperone-encoding genes at febrile temperatures to protect the parasite against protein damage.

A temperature increase of only a few degrees Celsius above the optimal growth temperature of any organism causes aberrant protein folding and aggregation, which contributes to an imbalance in protein homeostasis (proteostasis) that can lead to cell-cycle arrest or cell death[1]. To counteract the effect of high temperatures and other proteotoxic conditions, cells have a well-characterized heat-shock response that induces the expression of molecular chaperones that aid protein refolding and prevent non-specific protein aggregation[1,2]. In most eukaryotes, the immediate upregulation of chaperone-encoding genes during heat shock depends on the conserved transcription factor HSF1, whereas other transcriptional changes during thermal stress are driven by different transcription factors[3–6].

The human response to blood-stage infection with malaria parasites involves periodic fever episodes, which are the hallmark of clinical malaria[7–9]. Fever is an important part of the human innate immune response and may contribute towards reducing the total parasite burden[8–10]. Infection with *Plasmodium falciparum*, which causes the most severe forms of human malaria, results in fevers that typically occur on alternate days (tertian fever). Tertian fever reflects the approximately 48 h duration of the asexual intra-erythrocytic development cycle (IDC), during which parasites progress through the ring, trophozoite and multinucleated schizont stages. Fever episodes are triggered by schizont rupture and the release of invasive merozoites[8,9]. In vitro, febrile temperatures inhibit parasite growth, with maximal effect on trophozoites and schizonts[9,11,12], and induce conversion of asexual parasites into sexual forms that mediate transmission to mosquitoes[13].

Despite the importance of the heat-shock response for the survival of human malaria parasites during the host fever episodes, the regulation of the heat-shock response has not been characterized in these organisms. The genomes of the *Plasmodium* spp. lack an orthologue of HSF1 but they encode the main chaperone families described in other organisms[14], and several specific *P. falciparum* chaperones have been shown to be essential for heat-shock survival[15–19]. Phosphatidylinositol 3-phosphate and apicoplast-targeted pathways are also essential for parasite survival under thermal stress[18,19]. The transcript levels of over 300 genes are altered at febrile temperatures[20] but how these transcriptional changes are regulated is not known. Here we set out to characterize the regulation of the protective response of *P. falciparum* to increased temperature.

Results

A nonsense mutation in *pfap2-hs* is associated with low survival from heat shock. To understand the molecular basis of heat-shock resistance in *P. falciparum*, we first analysed parasite lines that had been previously selected with periodic heat shock (3 h at 41.5 °C) for five consecutive rounds of the IDC[21] (Fig. 1a). Although the parental parasite line (3D7-A) seemed to have lost the ability to withstand heat shock (approximately 30% heat-shock survival) during growth in vitro, it re-adapted to heat-shock pressure (>75% heat-shock survival) in only three generations[21], suggesting that this line contained a selectable subpopulation of parasites resistant to heat shock. To evaluate whether the heat-shock-resistance phenotype of 3D7-A had a genetic or an epigenetic basis, we first analysed the transcriptome across the full IDC under basal conditions (no heat shock) of two independently selected lines (3D7-A-HS r1 and r2) and non-selected cultures maintained in parallel (3D7-A r1 and r2). This analysis failed to identify any differences in the basal

[1]ISGlobal, Hospital Clínic—Universitat de Barcelona, Barcelona, Spain. [2]Department of Biochemistry and Molecular Biology and Huck Center for Malaria Research, Pennsylvania State University, University Park, PA, USA. [3]Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, UK. [4]School of Biological Sciences, Nanyang Technological University, Singapore, Singapore. [5]Department of Chemistry, Pennsylvania State University, University Park, PA, USA. [6]ICREA, Barcelona, Spain. [7]These authors contributed equally: Lucas Michel-Todó, Timothy J. Russell. ✉e-mail: alfred.cortes@isglobal.org

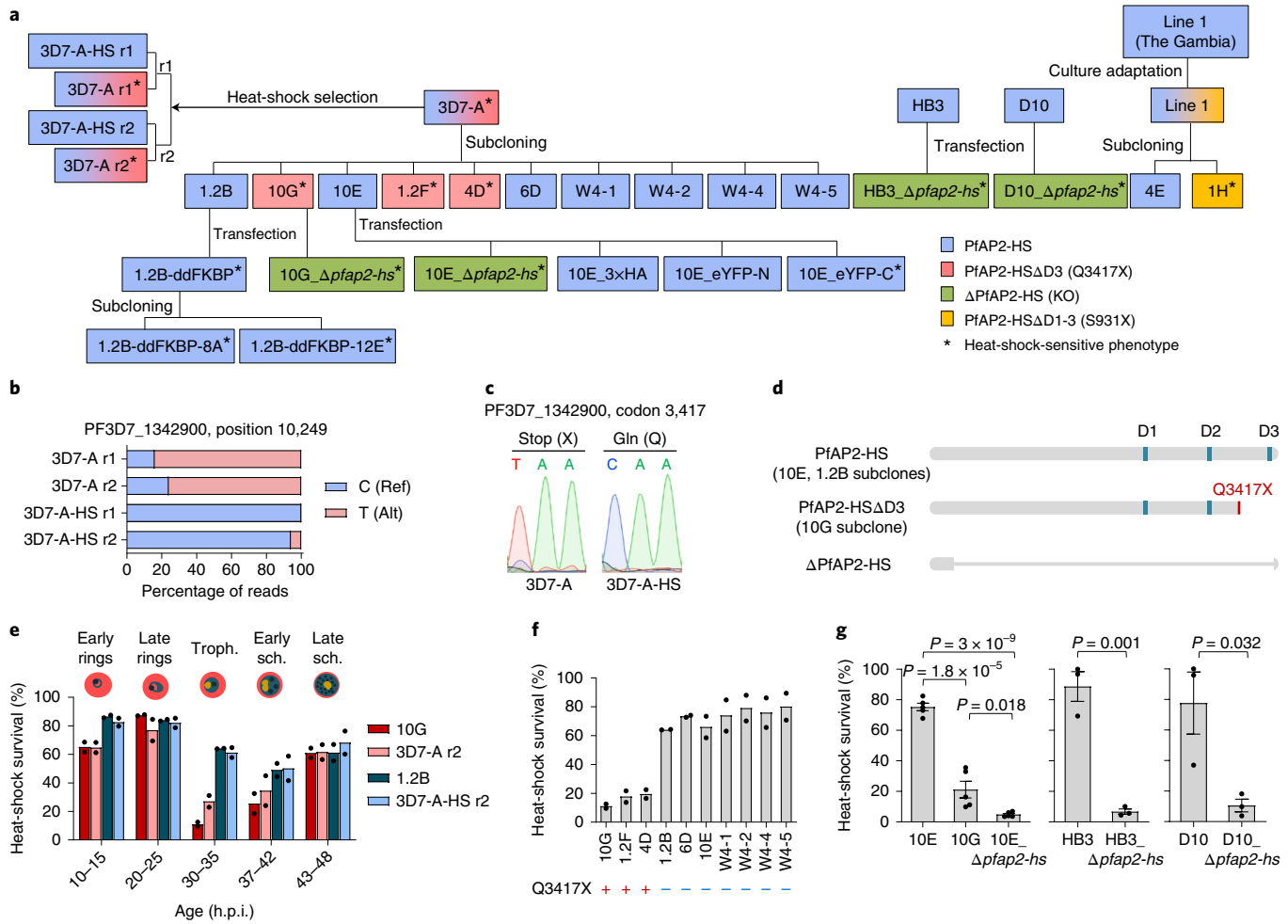
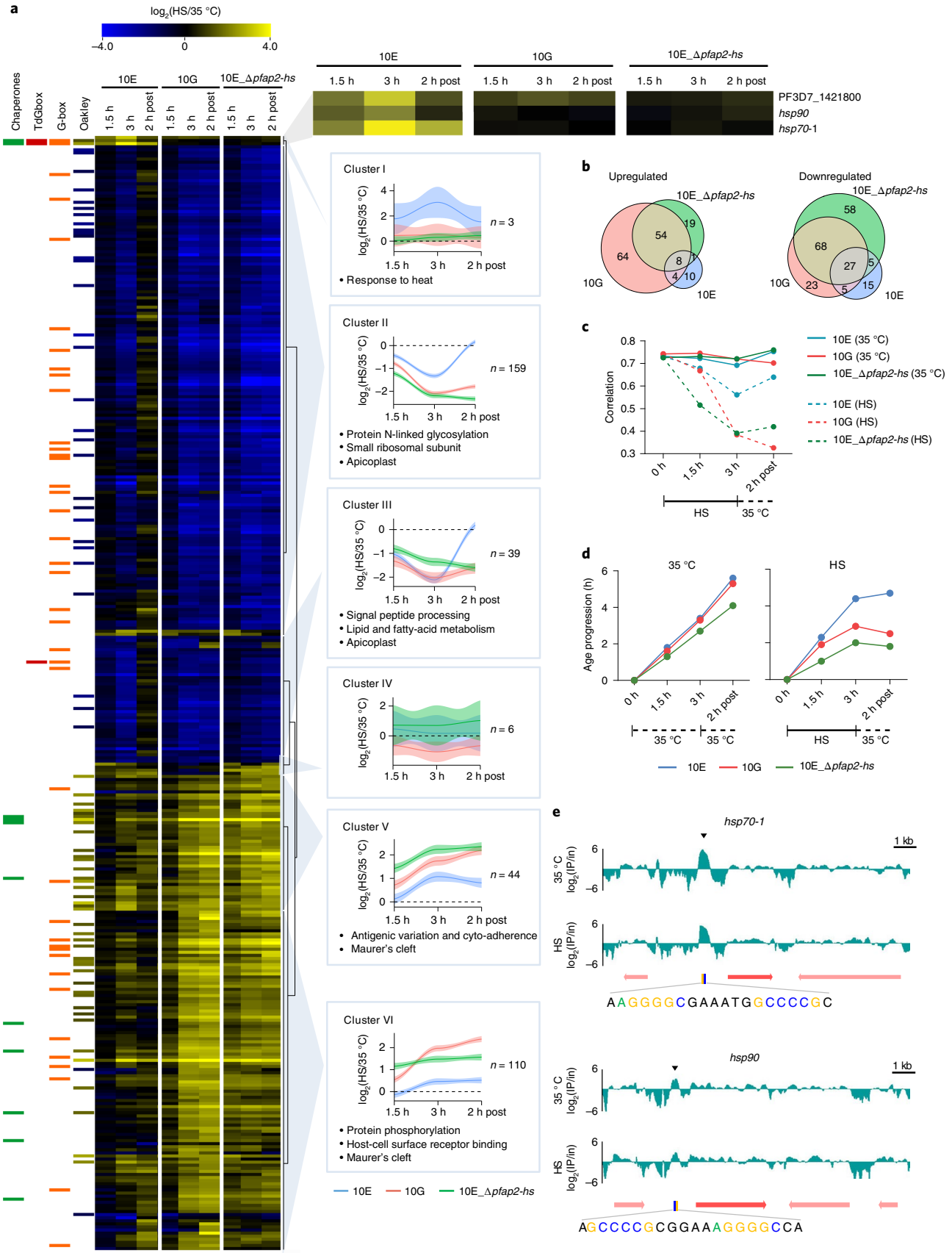


Fig. 1 | Mutations in PfAP2-HS and sensitivity to heat shock. a, Schematic of the parasite lines used in this study. Wild-type PfAP2-HS and truncated forms lacking AP2 D3 (ΔD3), the three AP2 domains (ΔD1–3) or virtually the full protein (KO) are indicated in different colours. The parasite lines shown with a colour gradient consist of a mixture of individual parasites expressing different versions of the protein. Asterisks indicate a heat-shock-sensitive phenotype, and r1 and r2 are independent replicates of the selection of 3D7-A with periodic heat-shock (3D7-A-HS r1 and r2 are the selected lines, whereas 3D7-A r1 and r2 are controls maintained in parallel at 37 °C). **b**, Proportion of Illumina reads with (Alt) or without (Ref) a nonsense mutation in *pfap2-hs* in two independently selected heat-shock-adapted cultures (3D7-A-HS r1 and r2) and their controls (3D7-A r1 and r2). **c**, Sanger-sequencing confirmation of the mutation (in the r1 replicate; representative of r1 and r2). **d**, Schematic of wild-type PfAP2-HS, PfAP2-HSΔD3 and ΔPfAP2-HS. The position of the AP2 domains is indicated (D1–3). **e**, Survival of tightly synchronized cultures exposed to heat shock at different ages (in hours post invasion, h.p.i.) for two heat-shock-sensitive (3D7-A r2 and 10G) and two heat-shock-resistant (3D7-A-HS r2 and 1.2B) lines (mean of $n=2$ independent biological replicates). Troph., trophozoites; and sch., schizonts. **f**, Heat-shock survival at the trophozoite stage of 3D7-A subclones with or without the Q3417X mutation (mean of $n=2$ independent biological replicates). **g**, Heat-shock survival of tightly synchronized cultures of parasite lines expressing wild-type or mutated PfAP2-HS. Values are the mean ± s.e.m. of $n=5$ (lines of 3D7 origin; left) or 3 (HB3 and D10 lines; middle and right, respectively) independent biological replicates. *P* values were calculated using a two-sided unpaired Student's *t*-test. **e**–**g**, Values are survival relative to control cultures maintained at 37 °C (**e**,**f**) or 35 °C (**g**).

transcriptomes that could explain the heat-shock resistance phenotypes (Extended Data Fig. 1). Therefore, we sequenced the genomes of these lines, which revealed a novel single nucleotide polymorphism (SNP) that was predominant in non-selected cultures but virtually absent after heat-shock selection (Fig. 1b,c and Supplementary Table 1). This mutation was also absent from two other 3D7 stocks in our laboratory and the 3D7 reference genome, indicating that it arose spontaneously in the 3D7-A stock during culture. This SNP results in a premature stop codon (Q3417X) in the gene *PF3D7_1342900*, which encodes a putative transcription factor of the ApiAP2 family[22–24] that we termed PfAP2-HS. PfAP2-HS has three AP2 domains (D1–3), and the Q3417X mutation results in a truncated protein that lacks D3 (PfAP2-HSΔD3; Fig. 1d). This result indicates that the adaptation of 3D7-A to heat shock involved the selection of parasites expressing full-length PfAP2-HS, consistent with a role for this protein in the heat-shock response. In support of this idea, the first AP2 domain (D1) of PfAP2-HS was previously reported to recognize in vitro a DNA motif termed G-box[23], which is enriched in the upstream region of some heat-shock protein (HSP) chaperone genes[25].

To test the involvement of PfAP2-HS in heat-shock resistance, we used a heat-shock survival assay consisting of heat-shock at 41.5 °C for 3 h (ref. [21]) at the mature trophozoite stage, because maximal survival differences between heat-shock-sensitive and -resistant parasite lines were observed when the parasites were exposed at this stage (Fig. 1e). The analysis of a collection of 3D7-A subclones revealed that all subclones with the Q3417X mutation (for example, the 10G subclone) have a heat-shock-sensitive phenotype, whereas subclones with the wild-type allele (for example, the 10E subclone) have a heat-shock-resistant phenotype (Fig. 1a,f).



Deletion of PfAP2-HS reduces survival from heat shock. To further characterize PfAP2-HS, we sought to disrupt the entire gene using clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR associated protein 9 (Cas9) technology. After several unsuccessful attempts with different 3D7 subclones at 37 °C (the physiological temperature for *P. falciparum*), we

Fig. 2 | Global transcriptional alterations in parasites exposed to heat shock. a, Hierarchical clustering of genes with altered transcript levels (≥fourfold change at any of the time points analysed) during (1.5 and 3 h) or 2 h after finishing (2 h post) heat shock (HS). The log_2-tranformed values of the fold change in expression in the HS versus control (35 °C) cultures are shown. Thirteen genes had values out of the range displayed (actual range: −4.78 to +4.93). For each cluster, mean values (lines) with the 95% confidence interval (shading) for the genes in the cluster and representative enriched Gene Ontology terms are shown. The columns on the left indicate annotation as chaperone[14], presence of the G-box[23] or tandem G-box (TdGbox) in the upstream region, and log_2-transformed fold-change values after HS from a previous study[20] (Oakley). **b**, Overlap in the genes that were altered (upregulated, left; and downregulated, right) in the three parasite lines following HS. **c**, Pearson's correlation of the genome-wide transcript levels of each culture versus the most-similar time point of a high-density time-course reference transcriptome[28]. **d**, Age progression during the HS assay, statistically estimated[61] from the transcriptomic data, under HS (right) or control (35 ºC, left) conditions. **e**, ChIP-seq analysis of HA-tagged PfAP2-HS, representative of $n=5$ and 3 independent biological replicates for 35 °C and HS, respectively. The log_2-transformed ChIP/input ratio (IP/in) at the *hsp70-1* (top) and *hsp90* (bottom) loci is shown. The position of the tandem G-box is indicated.

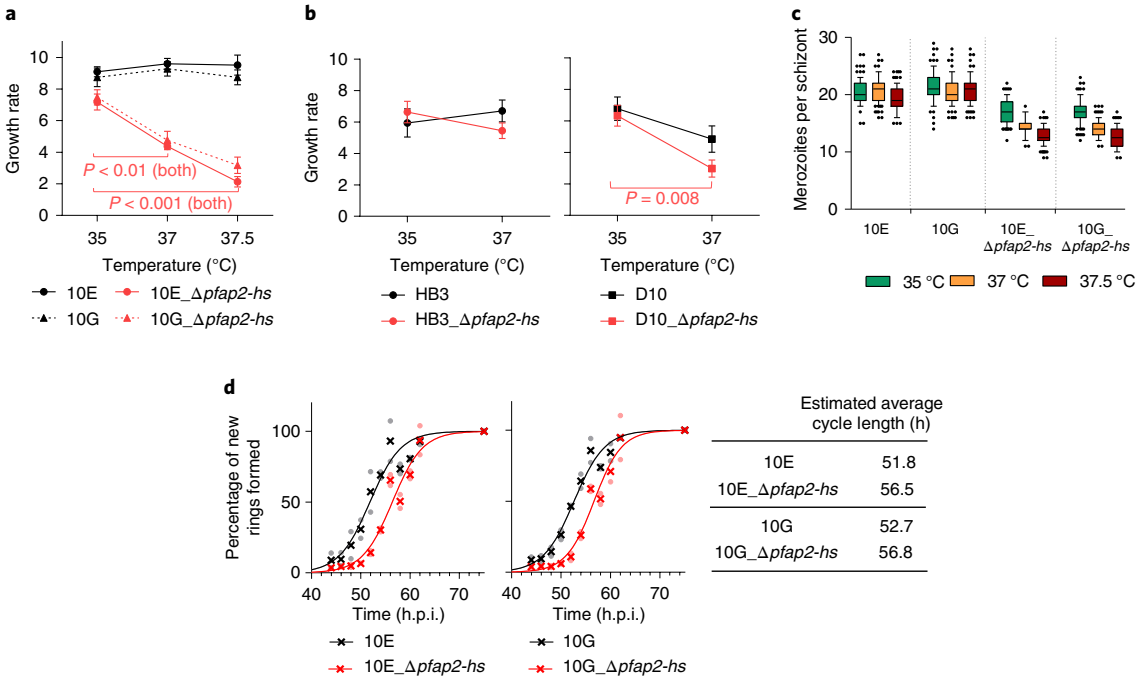


Fig. 3 | Phenotypes of parasite lines lacking PfAP2-HS. a, Growth rate of Δ*pfap2-hs* and parental lines with a 3D7 genetic background at different temperatures (mean ± s.e.m. of $n=4$ independent biological replicates). 10E_Δ*pfap2-hs*: 37 versus 35 °C, $P=2.3\times10^{-3}$; 37.5 versus 35 °C, $P=1.7\times10^{-4}$; and 10G_Δ*pfap2-hs*: 37 versus 35 °C, $P=0.011$; 37.5 versus 35 °C, $P=0.001$. **b**, Same as in **a** for parasite lines with a HB3 and D10 genetic background (mean ± s.e.m. of $n=4$ independent biological replicates). **a**,**b**, The *P* values were calculated using a two-sided unpaired Student's *t*-test. Only significant *P* values ($P<0.05$) are shown. **c**, Number of merozoites per schizont (horizontal line, median; box, quartiles; and whiskers, 10–90 percentile). Values were obtained from 100 schizonts for each parasite line and condition. **d**, Duration of the asexual blood cycle. The cumulative percentage of new rings formed at each time point (left; mean of $n=2$ independent biological replicates) and the estimated average length of the asexual blood cycle (right) are shown.

reasoned that PfAP2-HS may play a role in regulating the expression of chaperones under basal conditions in addition to being necessary for heat-shock survival. Therefore, we attempted to knock out the gene in cultures maintained at 35 °C, as mild hypothermia is expected to reduce protein unfolding and favour proteome integrity[26,27]. Knockout of *pfap2-hs* was indeed readily achieved at 35 °C in both the heat-shock-resistant 10E and the heat-shock-sensitive 10G subclones of 3D7-A (10E_Δ*pfap2-hs* and 10G_Δ*pfap2-hs* lines, respectively; Fig. 1a,d and Extended Data Fig. 2a). Deletion of *pfap2-hs* resulted in a substantial increase in the sensitivity to heat shock, with a level of heat-shock survival below that of parasites expressing PfAP2-HSΔD3. Deletion of the gene in two additional parasite lines with an unrelated genetic background, HB3 and D10, also resulted in a major reduction in heat-shock survival (Fig. 1g).

The PfAP2-HS-driven transcriptional response to heat shock is extremely compact. To define the PfAP2-HS-dependent and -independent heat-shock response, we carried out a time-course transcriptome analysis of the 10E (wild-type PfAP2-HS), 10G (PfAP2-HSΔD3) and 10E_Δ*pfap2-hs* lines during and after heat shock (Fig. 2a, Extended Data Fig. 3a and Supplementary Table 2). Hierarchical clustering based on changes in cultures exposed to heat shock compared with control cultures maintained in parallel without heat shock revealed one cluster of transcripts (cluster I) that are rapidly increased during heat shock in 10E but not in 10G or 10E_Δ*pfap2-hs*. Cluster I comprises only three genes: a gene of unknown function (*PF3D7_1421800*), the cytoplasmic *hsp70* (*hsp70-1*; *PF3D7_0818900*) and *hsp90* (*PF3D7_0708400*; Fig. 2a). The regulatory regions of these two chaperone-encoding genes contain the best two matches in the full genome for a tandem G-box[23,25] (Extended Data Fig. 3b). Although hundreds of genes in the *P. falciparum* genome contain a single G-box, only *hsp70-1* and *hsp90* showed PfAP2-HS-dependent activation during heat shock, suggesting that the tandem G-box arrangement may be needed for activation by PfAP2-HS. The strongest transcriptional response to heat

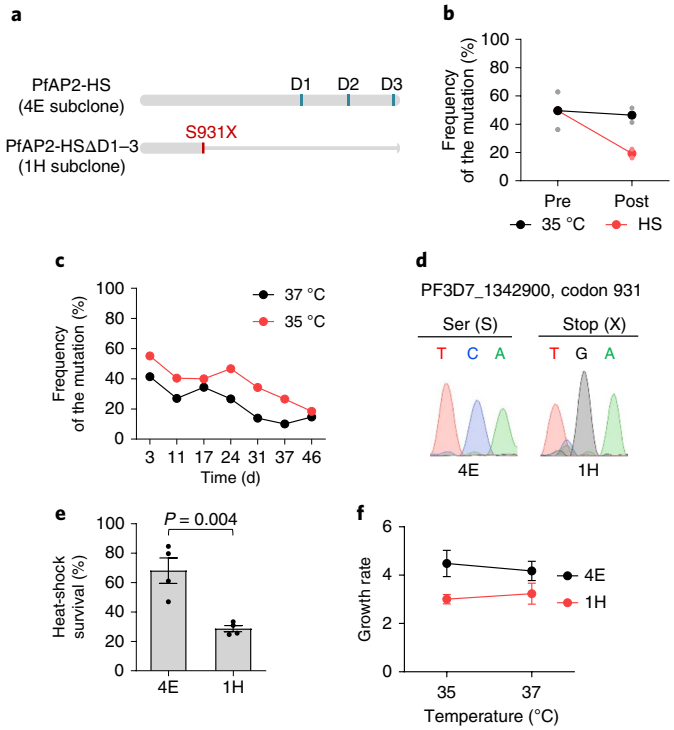


Fig. 4 | Characterization of a culture-adapted field isolate with mutations in *pfap2-hs*. **a**, Schematic of wild-type PfAP2-HS and PfAP2-HSΔD1–3 occurring in Line 1 from The Gambia after culture adaptation (C>G mutation at codon 931, S931X). The position of the AP2 domains is indicated (D1–3). **b**, Frequency of the mutation (as determined by Sanger sequencing) in culture-adapted Line 1 before (Pre) and after (Post) performing heat shock (HS) at the trophozoite stage and culturing for an additional cycle (mean of $n=2$ independent biological replicates). **c**, Frequency of the mutation during culture at different temperatures. The frozen stock from The Gambia (culture-adapted for 91 days) was placed back in culture on day 0. **d**, Sanger-sequencing determination of the presence or absence of the mutation at codon 931 in the Line 1 subclones 4E and 1H. **e**, Heat-shock survival of tightly synchronized 4E and 1H cultures relative to the control cultures maintained at 35 °C (mean ± s.e.m. of $n=4$ independent biological replicates). The *P* value was calculated using a two-sided unpaired Student's *t*-test. **f**, Growth rate of 4E and 1H at different temperatures (mean ± s.e.m. of $n=5$ independent biological replicates). No significant difference ($P<0.05$) was observed between growth at 35 and 37 °C (two-sided unpaired Student's *t*-test).

shock was observed for *hsp70-1* (approximately 16-fold increase versus about fourfold for *hsp90*).

To validate the observation that rapid activation of the cluster I genes following heat shock depends on PfAP2-HS and requires its D3, we analysed the heat-shock response of *pfap2-hs*-knockout parasite lines with different genetic backgrounds and several 3D7-A mutant subclones expressing PfAP2-HSΔD3 (Extended Data Fig. 4). The *hsp70-1* response to heat shock was delayed and of reduced magnitude in all of the knockout and mutant lines examined. These experiments also confirmed that the *hsp90* response is weaker than the *hsp70-1* response and is delayed in PfAP2-HS mutants.

PfAP2-HS-independent transcriptome alterations induced by heat shock. Genes in the other clusters (II–VI) of our transcriptomic analyses showed changes in expression during heat shock that were independent of PfAP2-HS; these changes were more pronounced in the mutant than wild-type lines (Fig. 2a). More genes with altered transcript levels were identified in the heat-shock-sensitive 10G and 10E_Δ*pfap2-hs* lines than in 10E (Fig. 2b). Furthermore, the alterations in the transcripts from clusters II–VI persisted 2 h after heat shock in both heat-shock-sensitive lines, whereas the majority of transcripts returned to basal levels in 10E (Fig. 2a). This suggests that many of these altered transcripts reflect unresolved cell damage or death. In 10E, the rapid PfAP2-HS-dependent response may protect cells from damage and enable rapid recovery from heat shock, thus limiting (in magnitude and duration) the changes in the expression of genes from clusters II–VI that reflect cell damage. After heat shock the transcriptome of the 10G and 10E_Δ*pfap2-hs* lines indeed showed a more pronounced deviation from a reference transcriptome[28] than 10E (Fig. 2c). Global transcriptional analysis also revealed that heat shock resulted in delayed IDC progression, which was again more pronounced in 10G and 10E_Δ*pfap2-hs* than 10E (Fig. 2d).

In addition to genes reflecting cell damage, clusters II–VI probably include some genes that participate in the PfAP2-HS-independent heat-shock response. In particular, clusters V and VI include several chaperone-encoding genes that were upregulated during heat shock, although at a later time point than the cluster I genes (Fig. 2a). However, the expression of the majority of known *P. falciparum* chaperones[14] was not altered by heat shock and, with the exception of cluster I genes, the alterations occurred mainly in the mutant lines (Extended Data Fig. 5). To provide a clearer view of the wild-type heat-shock response, we analysed heat-shock-induced changes in the 10E line alone. Overall, there was generally good concordance with the genes and the processes that are altered by heat shock described in a previous study using non-synchronized cultures[20] (Extended Data Fig. 6 and Supplementary Table 3). Altogether, we conclude that a number of genes are up- or downregulated during heat shock and some may contribute to heat-shock protection through PfAP2-HS-independent responses but in the absence of the rapid PfAP2-HS-dependent activation of cluster I genes, these responses are insufficient to ensure parasite survival at febrile temperatures.

Genome-wide analysis of PfAP2-HS binding sites. To determine the genome-wide occupancy of PfAP2-HS, we analysed a parasite line expressing haemagglutinin (HA)-tagged endogenous PfAP2-HS (Extended Data Fig. 2b) using chromatin immunoprecipitation (ChIP), followed by sequencing (ChIP–seq). The main binding site of PfAP2-HS coincides with the position of the tandem G-box in the upstream region of *hsp70-1* (Fig. 2e, Extended Data Fig. 7 and Supplementary Table 4). This is the only binding site with a median fold-enrichment value above ten (ChIP versus input) that was consistently detected in five independent ChIP–seq biological replicates, revealing an extremely restricted distribution of PfAP2-HS binding. Similar enrichment was observed between control and heat-shock conditions using ChIP–seq and ChIP with quantitative PCR (ChIP–qPCR; Extended Data Fig. 7), which suggests that PfAP2-HS binds constitutively to this site and is activated in situ by heat shock. This is reminiscent of yeast HSF1, which binds the promoter of *hsp70* and most of its other target promoters under both basal and heat-shock conditions[5]. Enrichment for PfAP2-HS at the *hsp90* promoter also coincided with the position of the tandem G-box, but was weaker, and a significant peak at this position was called in only one of the replicates (Fig. 2e, Extended Data Fig. 7 and Supplementary Table 4). No enrichment was observed at the promoter of the cluster I gene *PF3D7_1421800*, which lacks a G-box motif. The only other site consistently enriched for PfAP2-HS binding, albeit at much lower levels than *hsp70-1*, was the small nucleolar RNA *snoR04* (*PF3D7_0510900*) locus (Extended Data Fig. 7), which also lacks a G-box and was not upregulated during heat shock.

Growth defects under basal conditions in parasite lines mutated for PfAP2-HS. Both knockout lines of 3D7 origin (10E_Δ*pfap2-hs* and 10G_Δ*pfap2-hs*) showed severe temperature-dependent

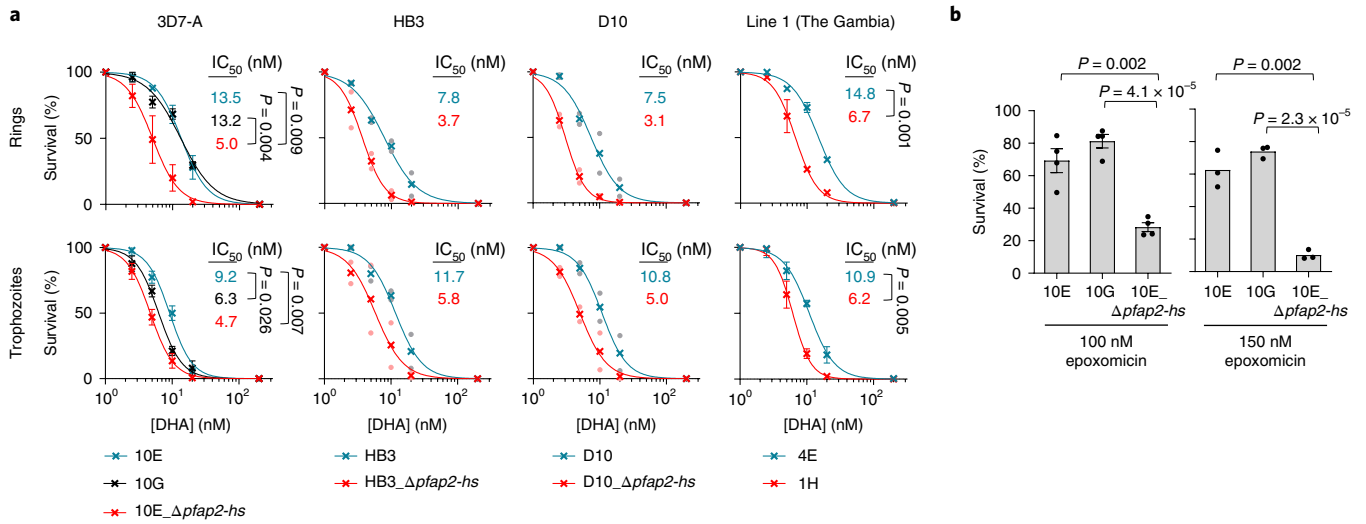


Fig. 5 | Sensitivity of parasites lacking PfAP2-HS to proteotoxic conditions. a, Survival, relative to the untreated control, after a 3 h DHA pulse at the ring (top) or trophozoite (bottom) stage. Values are the mean ± s.e.m. of $n = 3$ (3D7-A and Line 1 genetic backgrounds) or mean of $n = 2$ (HB3 and D10 genetic backgrounds) independent biological replicates. The mean half maximal inhibitory concentration (IC_{50}) for each line is shown (same colour code as the plots). **b**, Survival, relative to the untreated control, after a 3 h epoxomicin pulse at the trophozoite stage. Values are the mean ± s.e.m. of $n = 4$ (100 nM) and 3 (150 nM) independent biological replicates. **a**,**b**, The *P* values were calculated using a two-sided unpaired Student's *t*-test (only for experiments with $n \geq 3$). Only significant *P* values ($P < 0.05$) are shown.

growth defects in the absence of heat shock. They grew at similar rates to the parental lines at 35 °C but their growth was markedly reduced at 37 and 37.5 °C (Fig. 3a). The D10_Δ*pfap2-hs* line also had clearly reduced growth at 37 °C compared with 35 °C, whereas the HB3_Δ*pfap2-hs* line did not (Fig. 3b). Both 3D7 Δ*pfap2-hs* lines also showed a reduced number of merozoites per schizont, especially at 37 and 37.5 °C (Fig. 3c), which partly explains the reduced growth rate. In addition, even at 35 °C, the duration of the IDC was approximately 4 h longer in both knockout lines (Fig. 3d), which is reminiscent of the slower life-cycle progression observed in parasites under proteotoxic stress[29]. In contrast, no differences in growth rate or life-cycle duration were observed between the 10G (PfAP2-HSΔD3) and 10E (wild-type PfAP2-HS) lines, indicating that D3 is necessary for heat-shock survival but not for growth under non-stress (37 °C) conditions (Fig. 3a,c,d). Normal growth at 37 °C, but low heat-shock survival, was also observed in transgenic lines in which bulky carboxy (C)-terminal tags were added to the C terminus of endogenous PfAP2-HS, suggesting interference of the tag with the function of D3, which is located only 18 amino acids from the end of the protein (Fig. 1a and Extended Data Fig. 2).

Genome-wide sequence analysis has previously found that nonsense mutations arise in *pfap2-hs* during adaptation to culture conditions[30,31] but these are virtually absent from clinical isolates (in the www.malariagen.net/data/pf3k-5 dataset[32], only one of >2,500 isolates carries a high-confidence SNP resulting in a premature stop codon). The lack of mutations observed in clinical isolates suggests that there is a selection against loss-of-function mutations in PfAP2-HS during human infections, where parasites are frequently exposed to febrile conditions. We exposed a culture-adapted isolate in which approximately 50% of the parasites carried a mutation that results in the truncation of PfAP2-HS before its first AP2 domain[30] (monoclonal Gambian Line 1, PfAP2-HSΔD1–3) to one round of heat shock (41.5 °C, 3 h) and found that only about 20% of the parasites in the next generation carried the mutation. The frequency of the mutation remained stable in the control cultures maintained in parallel without heat shock (Fig. 4a,b). This result indicates strong selection by heat shock against parasites carrying the PfAP2-HS truncation. In contrast, there was relatively weak selection against mutants during culturing at either 35 or 37 °C, as the prevalence of the mutation only decreased from about 50% to approximately 20% after culturing for 23 generations at either temperature (Fig. 4c). Consistent with these results, a Gambian Line 1 subclone carrying the mutation (1H) was more sensitive to heat shock than a wild-type subclone (4E), but both showed no measurable difference between their growth at 35 and 37 °C (Fig. 4d–f). Together, these results indicate that PfAP2-HS is essential for heat-shock survival in all of the tested genetic backgrounds. However, PfAP2-HS is necessary for normal progression through the IDC at 37 °C in only specific genetic backgrounds (that is, 3D7 and D10), whereas in others (HB3 and the Gambian isolate) it is not essential.

Transcriptional alterations in parasites lacking PfAP2-HS under basal conditions. To gain insight on the molecular basis of the growth defects of some of the knockout lines, we compared the trophozoite transcriptome of 10E_Δ*pfap2-hs* with that of the parental 10E under basal (no heat shock) conditions. This revealed only a small set of genes with a fold decrease ≥ 2 in transcript levels, which included *hsp70-1*, the direct PfAP2-HS target *snoR04* RNA and several genes mainly involved in ribosome formation. The transcript levels for *hsp90* were also reduced (<twofold decrease) in the knockout line (Extended Data Fig. 8a,b and Supplementary Table 2). Reduced *hsp70-1* and *hsp90* transcript levels under basal conditions in 10E_Δ*pfap2-hs* mature trophozoites were independently confirmed by reverse transcription–qPCR and were also observed at the late ring stage and in the knockout lines with D10 and HB3 genetic backgrounds (Extended Data Fig. 8c). These results indicate that PfAP2-HS contributes to the regulation of the basal expression of the same chaperone-encoding genes that it activates following heat shock, among a few other genes. Together with the observation that the growth defect of the *pfap2-hs*-knockout lines is attenuated at 35 °C, this suggests that the knockout parasites have a reduced capacity for proteostasis, such that at 37 °C they are at the edge of proteostasis collapse. Parasite lines that can grow normally at 37 °C in spite of PfAP2-HS deletions including the three AP2 domains must therefore have alternative pathways active to ensure basal proteostasis. We hypothesise that mutant parasites expressing

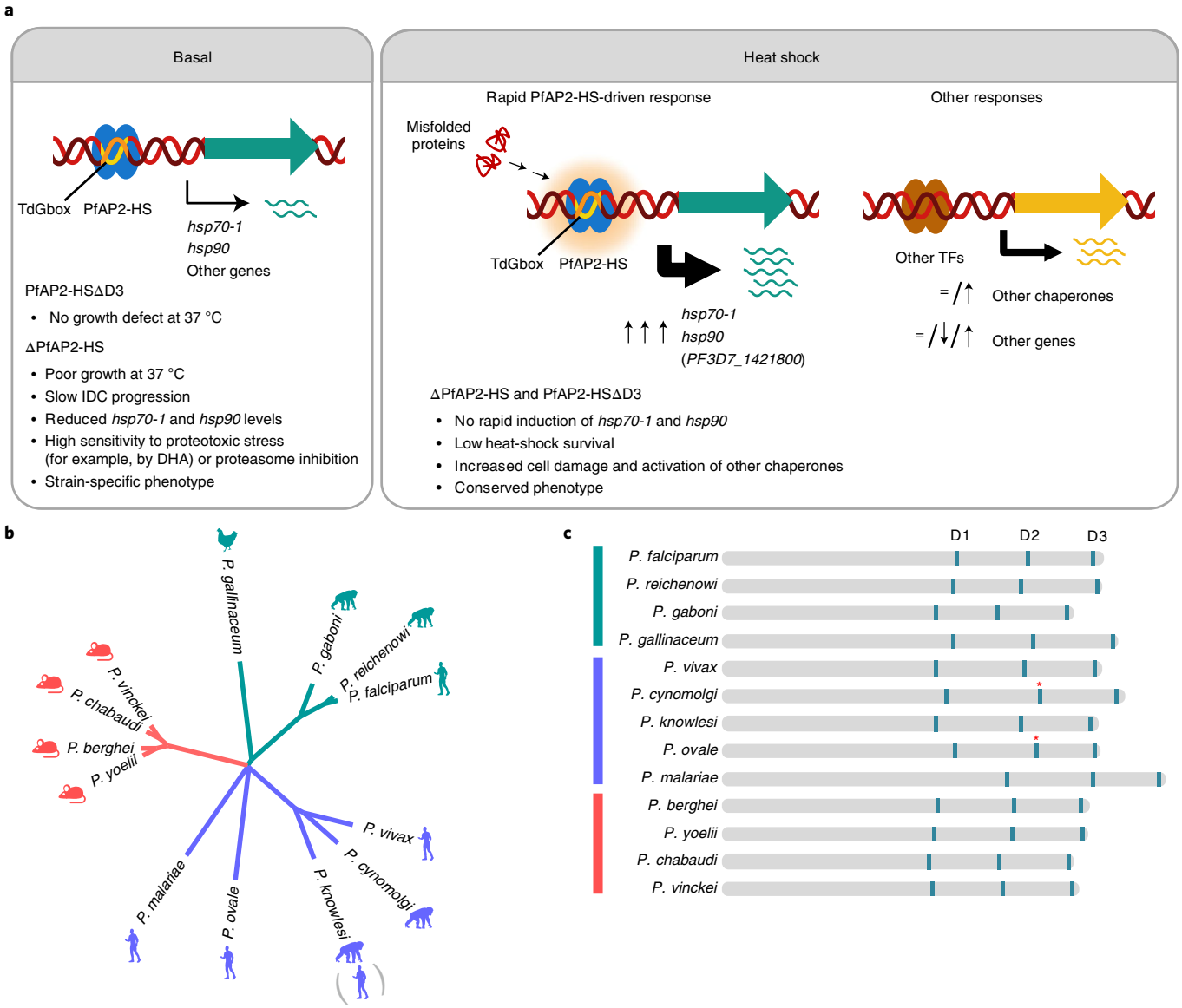


Fig. 6 | Model of the *P. falciparum* heat-shock response and phylogenetic analysis of AP2-HS. a, The *P. falciparum* heat-shock response involves rapid upregulation of the expression of a very restricted set of chaperones by PfAP2-HS. The *PF3D7_1421800* gene (in brackets) shows PfAP2-HS-dependent increased transcript levels following heat shock, but PfAP2-HS binding was not detected in its promoter and it lacks a G-box. The main defects associated with PfAP2-HS deletion or truncation, under heat shock (right) or basal (left) conditions, are listed. TFs, transcription factors. **b**, Phylogenetic analysis of the protein sequences of AP2-HS orthologues in *Plasmodium* spp. **c**, Schematic of the domain structure of the AP2-HS orthologues in *Plasmodium* spp. The position of the AP2 domains (D1–3) is based on domains identified in PlasmoDB release 50, except for those marked with an asterisk, which were annotated manually based on sequence alignments.

truncated PfAP2-HS are frequently selected under culture conditions because the truncations do not pose a fitness cost at 37 °C in the lines in which they appear and they prevent unnecessary activation of the heat-shock response, which can be detrimental[33], by unintended mild stress that may occur during culture.

PfAP2-HS-deficient parasites show increased sensitivity to artemisinin. Artemisinins are potent antimalarial drugs that kill parasites by causing general protein damage[34,35]. Resistance to artemisinin is associated with mutations in the Kelch13 protein[36,37] and involves cellular stress-response pathways such as the ubiquitin–proteasome system and the endoplasmic reticulum-based unfolded protein response (UPR)[34,35,38,39]. Given that PfAP2-HS regulates the expression of key chaperones, we tested the sensitivity of PfAP2-HS-deficient lines to dihydroartemisinin (DHA), the active metabolite of artemisinins. In all four different genetic backgrounds (3D7, D10, HB3 and Gambian Line 1), the knockout of *pfap2-hs* (or truncation before D1) resulted in a higher sensitivity to a pulse of DHA than in the lines with full PfAP2-HS (both at the ring or the trophozoite stage), whereas 10G showed increased sensitivity only when exposed at the trophozoite stage (Fig. 5a). These results indicate that deletion of PfAP2-HS renders parasites more sensitive to chemical proteotoxic stress, in addition to heat shock, probably as a consequence of basal defects in cellular proteostasis. We reasoned that if parasites lacking the PfAP2-HS protein bear constitutive proteome defects, they should have a low tolerance

to the disruption of other factors involved in proteostasis maintenance. The 10E_Δ*pfap2-hs* line was indeed more sensitive to the proteasome inhibitor epoxomicin than the parental 10E line or the 10G line (Fig. 5b). Furthermore, after heat shock there was greater accumulation of polyubiquitinated proteins in the knockout line than in 10E or 10G, reflecting higher levels of unresolved protein damage (Extended Data Fig. 9a). We also assessed the links between the PfAP2-HS-driven heat-shock response and the other main cell stress-response pathway—that is, the UPR. Using phosphorylation of eIF2α as a UPR marker, we found that the UPR does not depend on PfAP2-HS and is not directly activated by heat shock, because the marker was significantly elevated after heat shock only in the knockout line (Extended Data Fig. 9b).

Discussion

Our results show that the PfAP2-HS transcription factor is bound to the tandem G-box DNA motif in the promoter of the chaperone-encoding gene *hsp70-1* and rapidly upregulates the expression of this gene and, to a lesser extent, *hsp90* in response to febrile temperatures (Fig. 6a). Binding of PfAP2-HS to the G-box is mediated by D1 (ref. [23]) but rapid activation of *hsp70-1* and *hsp90* during heat shock requires D3, which is incapable of binding DNA in vitro[23] and in the cell is likely to participate in protein–protein interactions or dimerization[40]. Other components of the protein-folding machinery necessary for heat-shock survival[15–17,19] are either constitutively expressed or induced later, but the rapid PfAP2-HS-driven response is essential to avoid irreversible damage. Importantly, parasites lacking either the entire PfAP2-HS or its D3 cannot survive heat shock.

Although the sequence and domain organization of PfAP2-HS does not show any similarity with HSF1—the conserved master regulator of the heat-shock response in most eukaryotes, from yeast to mammals[3,6]—it serves an analogous role. HSF1 regulates a compact transcriptional programme that includes the *hsp70* and *hsp90* genes[4,5]. In yeast, the only essential role of this transcription factor is the activation of *hsp70* and *hsp90* (ref. [5]), the same chaperone-encoding genes activated by PfAP2-HS during heat shock. In addition to its role in the protective heat-shock response, PfAP2-HS is essential for growth at 37 °C in some *P. falciparum* genetic backgrounds. The function of PfAP2-HS under basal conditions is independent of its D3 domain. Several lines of evidence suggest that the role for PfAP2-HS under basal conditions involves proteostasis maintenance (Fig. 6a), similar to yeast HSF1 (ref. [5]). In other organisms, the heat-shock response mediates protection against different types of proteotoxic stress, in addition to high temperature[1,3]. Here we report that parasites lacking PfAP2-HS have increased sensitivity to artemisinin, and future research will be needed to establish the precise role of the *P. falciparum* heat-shock response in protection against different types of stress. We note that orthologues of *pfap2-hs* are present in all of the analysed *Plasmodium* spp. (Fig. 6b,c and Extended Data Fig. 10), including murine *Plasmodium* species that do not induce host fever. This observation suggests that, at least in these species, the heat-shock response regulated by AP2-HS may play a role in protection against different conditions.

Finally, while several ApiAP2 transcription factors regulate life-cycle transitions in malaria parasites[24,41–43], PfAP2-HS controls a protective response to a within-host environmental challenge. Our findings that the PfAP2-HS transcription factor regulates the activation of a protective heat-shock response settles the long-standing question of whether malaria parasites can respond to changes in within-host environmental conditions with specific transcriptional responses[44].

Methods

Parasite cultures. The 3D7-A stock of the clonal *P. falciparum* line 3D7 (ref. [45]), the 3D7-A subclones 10G, 1.2B, 10E, 4D, 6D, 1.2F, W4-1, W4-2, W4-4 and W4-5 (refs. [46,47]), the HB3B[48] (mosquito and chimpanzee-passaged HB3, provided by O. Kaneko, Ehime University, Japan) and D10 (ref. [49]; provided by R. F. Anders, La Trobe University, Australia) clonal parasite lines, and the culture-adapted Line 1 from The Gambia[30] have been previously described. The heat-shock-selected lines 3D7-A-HS r1 and r2 were derived from 3D7-A by exposing the cultures to 3 h of heat shock (41.5 °C) at the trophozoite stage for five consecutive cycles (each replicate, r1 and r2, is a fully independent selection from the 3D7-A stock), and the 3D7-A r1 and r2 lines are cultures that were maintained in parallel at 37 °C without heat shock[21]. Parasites were cultured in B^+ erythrocytes at a 3% haematocrit under standard culture conditions in RPMI-based media containing Albumax II (without human serum), in an atmosphere of 5% CO_2, 3% O_2 and the balance N_2 (except for cultures for ChIP–seq experiments, in which O^+ erythrocytes were used). Regular synchronization was performed using 5% sorbitol lysis, whereas tight synchronization (1, 2 or 5 h age window) was achieved by Percoll purification followed by sorbitol treatment 1, 2 or 5 h later. All cultures were regularly maintained at 37 °C, with the exception of the *pfap2-hs*-knockout lines that were maintained at 35 °C. For experiments performed in parallel with the knockout lines and other parasite lines, all cultures were maintained at 35 °C for at least one cycle before the experiment.

Generation of transgenic parasite lines. We used two single guide RNAs (hereafter referred to as sgRNA or guide) to knock out *pfap2-hs* (11,577 bp) using the CRISPR–Cas9 system[50] (Extended Data Fig. 2a and Supplementary Table 5). One guide targets a sequence near the 5′ end of the gene (positions 866–885 from the start codon), whereas the other recognizes a sequence near the 3′ end (positions 11,486–11,505). The 5′ guide was cloned into a modified pL6-*egfp* donor plasmid[50] in which the *yfcu* cassette had been removed by digestion with NotI and SacII, end blunting and re-ligation. The 5′ and 3′ homology regions (HR1, positions −2 to 808 of the gene; and HR2, from position +11520 of the gene to 490 bp after the stop codon, respectively) were also cloned in this plasmid, flanking the *hdhfr* expression cassette, to generate the plasmid pL7-pfap2hs_KO_sgRNA5′. The 3′ guide was cloned into a modified version of the pDC2-Cas9-U6-hdhfr[51] plasmid in which we previously removed the *hdhfr* expression cassette by digesting with NcoI and SacII, end blunting and re-ligation, and replaced the BbsI guide cloning site with a BtgZI site. The resulting plasmid was named pDC2_wo/hdhfr_pfap2hs_sgRNA3′. All guides were cloned using the In-Fusion system (Takara) as described[50], whereas homology regions were PCR-amplified from genomic DNA and cloned by ligation using the restriction sites SpeI and AflII (HR1), and EcoRI and NcoI (HR2).

For the constructs aimed at C-terminal tagging of *pfap2-hs* using CRISPR–Cas9 (10E_*pfap2-hs*_eYFP-Cterm and 10E_*pfap2-hs*_3×HA-Cterm lines), we used a guide corresponding to positions 11508–11527 of the gene (Extended Data Fig. 2b,c and Supplementary Table 5). The guide was cloned into the pDC2-Cas9-U6-hdhfr[51] plasmid to obtain pDC2_pfap2hs_sgRNA-C. The donor plasmid for tagging with enhanced yellow fluorescent protein (eYFP; pHR-C_pfap2hs_eYFP) was based on the plasmid pHRap2g-eYFP[52], with the *pfap2-g* homology regions and the 3′ sequence of *hsp90* replaced with *pfap2-hs* homology regions. The HR1 region was generated with a PCR-amplified fragment spanning from nucleotide 10964 to the sequence of the guide (recodonized) and a 47-bp fragment (generated by annealing two complementary oligonucleotides) consisting of a recodonized version of the remaining nucleotides to the end of the gene. The two fragments were cloned simultaneously into the SpeI–BglII sites using the In-Fusion system. The HR2 region was a PCR fragment spanning positions +1 to +590 after the *pfap2-hs* stop codon. It was cloned into the XhoI–AatII restriction sites. The donor plasmid for 3×HA C-terminal tagging (pHR-C_pfap2hs_3×HA_hsp90-3′) was also based on the plasmid pHRap2g-eYFP[52], with the eYFP coding sequence replaced by the 3×HA sequence (amplified from the plasmid pHH1inv-pfap2-g-HA×3; ref. [41]) and the same homology regions as in the plasmid pHR-C_pfap2hs_eYFP (but HR2 was cloned, using the In-Fusion system, into the EcoRI–AatII sites because in this construct the 3′ region of *hsp90* in pHRap2g-eYFP was maintained).

For amino (N)-terminal tagging (10E_*pfap2-hs*_eYFP-Nterm line), we cloned a guide targeting the *pfap2-hs* positions 73–92 in the pDC2-Cas9-U6-hDHFRyFCU[53] plasmid to obtain the plasmid pDC2_pfap2hs_sgRNA-N (Extended Data Fig. 2d and Supplementary Table 5). The donor plasmid (pfap2hs_HR-N_eYFP) consisted of a HR1 region including positions −366 to −1 relative to the *pfap2-hs* start codon, the *eYFP* gene and an in-frame HR2 spanning positions 4–756 of the gene (excluding the ATG) in which the nucleotides up to the position of the guide were recodonized. The HR1 and HR2 regions were cloned into the SacII–NcoI and SpeI–EcoRI sites, respectively. HR2 was amplified in two steps using a nested PCR approach to add the recodonized sequences. The eYFP fragment (PCR-amplified from plasmid pHR-C_pfap2hs_eYFP) was cloned using the In-Fusion system into the SpeI–NcoI sites.

To tag PfAP2-HS with a 2×HA-ddFKBP domain tag (1.2B_ *pfap2-hs*_ddFKBP line), we used a single homologous recombination approach (Extended Data Fig. 2e). To generate the pfap2hs_HA-ddFKBP plasmid, we replaced the *pfap2-g* homology region in the plasmid PfAP2-G-ddFKBP[41] with a PCR-amplified fragment including positions 9551–11574 of *pfap2-hs* in frame with the tag. The fragment was cloned using the restriction sites NotI and XhoI. All of

the oligonucleotides that were used to generate the plasmids are described in Supplementary Table 5. The relevant parts of all plasmids (that is, the new sequences incorporated) were sequenced before transfection.

The transfections were performed by electroporation of ring-stage cultures with 100 µg plasmid (HA-ddFKBP tagging) or a mixture of 12 µg linearized donor plasmid and 60 µg of circular Cas9 plasmid (CRISPR–Cas9 system). Linearization was achieved by digestion with the PvuI restriction enzyme (cleaving the *amp* resistance gene of the donor plasmid). Transfected cultures were selected with 10 nM WR99210 for four days as previously described[53] (transfections using the CRISPR–Cas9 system) or with continuous WR99210 pressure until parasites were observed, followed by three off–on drug cycles and subcloning by limiting dilution (transfections with the pfap2hs_HA-ddFKBP plasmid). In all cases, we used analytical PCR of genomic DNA (Extended Data Fig. 2) with specific primers (Supplementary Table 5) to assess correct integration.

Heat-shock resistance assay. Heat shock was always performed on cultures at the mature trophozoite stage unless otherwise indicated. To measure heat-shock survival, the cultures were tightly synchronized to a defined age window, diluted to 1% parasitaemia, split between two identical petri dishes (heat shock and control) maintained in independent air-tight incubation chambers and exposed to heat shock when the majority of parasites were at the mature trophozoite stage (typically about 30–35 h.p.i.; the Δ*pfap2-hs* lines were tightly synchronized 3 h earlier than the other lines but exposed to heat shock in parallel to account for their slower IDC progression). The exception was experiments to screen many subclones (that is, Fig. 1f) or to characterize transgenic parasite lines (that is, Extended Data Fig. 2) in which cultures were only sorbitol-synchronized and heat shock was performed approximately 20–25 h after sorbitol lysis (mature trophozoite stage). For heat shock, the full incubation chamber was transferred to an incubator at 41.5 °C for 3 h and then returned to 37 or 35 °C (the latter temperature was used for all lines in experiments including the *pfap2-hs*-knockout lines). The chamber with the control cultures was always maintained at 37 or 35 °C. After reinvasion (typically approximately 60–70 h after synchronization to ensure that all parasites had completed the cycle, including parasites subjected to heat shock that demonstrated delayed progression through the IDC), parasitaemia of the control and heat-shock-exposed cultures was measured by flow cytometry using a FACSCalibur flow cytometer (Becton Dickinson) and SYTO 11 to stain the nucleic acids (Supplementary Fig. 1), as previously described[54].

Phenotypic characterization. To determine the growth rate (increase in parasitaemia between consecutive cycles) at different temperatures, the parasitaemia of sorbitol-synchronized cultures was adjusted to 1% and then accurately determined by flow cytometry. The cultures were then split between two or three dishes and maintained in parallel in incubators at the different temperatures tested. Parasitaemia was again determined by flow cytometry at the next cycle to determine the growth rate. To measure the duration of the IDC (at 35 °C) in the different parasite lines, we used a recently described method based on synchronization to a 1 h age window achieved by Percoll purification of schizonts, followed by sorbitol lysis 1 h later[54]. The determination of the number of merozoites per fully mature schizont was based on light-microscopy analysis of Giemsa-stained smears from Percoll-purified schizonts[54]. Sensitivity to DHA (Sigma, cat. no. D7439) or epoxomicin (Selleckchem, cat. no. S7038) was measured after exposing tightly synchronized cultures (1% parasitaemia) at the ring (10–15 h.p.i.; DHA only) or trophozoite (30–35 h.p.i.; DHA or epoxomicin) stage to a 3 h pulse of the compounds at different concentrations (DHA, 2.5, 5, 10, 20 or 200 nM, which is lower than the approximately 700 nM plasma concentration after patient treatment that kills the vast majority of sensitive parasites[35]; epoxomicin, 100 or 150 nM, which is higher than the reported 7.7 nM IC_{50} after exposing parasites for 50 h (ref. [55]) and similar to the concentration used in previous studies with a 3 h pulse[29,38]). Parasitaemia was measured by light-microscopy analysis of Giemsa-stained smears at the next cycle (typically 70–80 h after Percoll–sorbitol synchronization). For these experiments, the Δ*pfap2-hs* lines were tightly synchronized 3 h earlier than the other lines but exposed to DHA or epoxomicin in parallel (13–18 or 33–38 h.p.i.) to account for their slower IDC progression. For the DHA experiments, the drug concentrations were log-transformed and the per cent survival data were fit to sigmoidal dose-response curves to calculate the IC_{50} values using GraphPad Prism.

Transcriptional analysis by reverse transcription–qPCR. RNA from tightly synchronized cultures exposed to heat shock and their controls was obtained using the TRIzol method, DNAse treated and purified essentially as described. Reverse transcription and qPCR analysis of the complementary DNA were also performed as described before[56,57]. Briefly, a mixture of random primers and oligo (dT) were used for reverse transcription, and for qPCR, we used the PowerSYBR Green master mix (Applied Biosystems) and the standard-curve method (each plate included a standard curve for each primer pair). All of the primers used are listed in Supplementary Table 5. Unless otherwise indicated, the transcript levels were normalized to *serine–tRNA ligase* (*PF3D7_0717700*), which shows relatively stable expression throughout the IDC.

Transcriptomic analysis using microarrays. To compare the transcriptome of control and heat-shock-adapted 3D7-A parasite lines across the IDC, we used previously described two-colour long-oligonucleotide-based glass microarrays[21]. RNA was obtained from tightly synchronized cultures (5 h age window) at 8–13, 16–21, 24–29, 32–37 and 40–45 h.p.i. All samples (Cy5-labelled) were hybridized together with a reference pool (Cy3-labelled) consisting of a mixture of equal amounts of cDNA from rings, trophozoites and schizonts from control and heat-shock-adapted lines. Comparative genome hybridization was used to determine which differences in transcript levels were potentially attributable to genetic deletions or duplications. Useful data were obtained from 5,142 genes. Sample preparation, microarray hybridization and data analysis were performed essentially as described previously[21].

To analyse the transcriptome of the 10E, 10G and 10E_Δ*pfap2-hs* parasite lines under control and heat-shock conditions, we used two-colour long-oligonucleotide-based custom Agilent microarrays[58]. The microarray design was based on Agilent design AMADID no. 037237 (refs. [58,59]), but we modified it as previously described (new design AMADID no. 084561)[60]. RNA was obtained from cultures synchronized to a 5 h age window at about 2.5% parasitaemia. Given the slower IDC progression of 10E_Δ*pfap2-hs*, cultures of this parasite line were synchronized to 0–5 h.p.i. 3 h earlier than the 10E and 10G cultures, such that all cultures were approximately at the same stage of IDC progression when heat shock was initiated (in parallel for all lines). Heat shock was started at 30–35 h.p.i. (33–38 h.p.i. for the 10E_Δ*pfap2-hs* line) and samples were collected before, during and after heat shock, as indicated. RNA was prepared using the TRIzol method. Sample preparation and microarray hybridization were performed essentially as described[59]. All samples (Cy5-labelled) were hybridized together with a reference pool (Cy3-labelled) consisting of a mixture of equal amounts of cDNA from rings, trophozoites and schizonts from 3D7-A. Microarray images were obtained using a DNA microarray scanner (Agilent Technologies, cat. no. G2505C) located in a low-ozone area and initial data processing was performed using the GE2 _1105_ Oct12 extraction protocol in the Agilent Feature Extraction 11.5.1.1 software.

Agilent microarray data were analysed using Bioconductor in an R environment (R version 3.5.3). For each microarray, we calculated the Cy3 and Cy5 background signal as the median of the 100-lowest signal probes for each channel, and probes with both Cy3 and Cy5 signals below three times the array background were excluded. Gene-level $\log_2$(Cy5/Cy3) values, statistical estimation of parasite age[61] and estimation of the average fold differences in expression across a time interval (for the comparison between parasite lines in the absence of heat shock) were performed as described[21]. The $\log_2$-transformed value of the fold change in expression following heat shock was calculated, for each gene and time point, as the $\log_2$(Cy5/Cy3) of the heat-shock-exposed sample minus the $\log_2$(Cy5/Cy3) of the control sample at the same parasite age, calculated using linear interpolation in the $\log_2$(Cy5/Cy3) versus estimated age plot. Visual inspection was used to exclude genes with apparent artefacts from further analyses. Genes that were missing data for ≥2 time points (or ≥1 for the comparison between parasite lines in the absence of heat shock across a time interval) or with values within the lowest 15th percentile of expression intensity (Cy5 sample channel) in all samples were also excluded from further analyses. Useful data were obtained from 4,964 genes.

To assess the level of similarity between the transcriptome of our samples and a reference non-stressed transcriptome with high temporal resolution (HB3 line)[28], we calculated the Pearson's correlation between each sample and the time point of the reference transcriptome with which it had a higher similarity. Heatmaps and hierarchical clustering based on Spearman's (Fig. 2) or Pearson's (Extended Data Fig. 6) correlation were generated using Multiple Experiment Viewer (MeV) 4.9 (ref. [62]). Expression trend plots for each cluster were generated using ggplot2, with LOESS smoothing, and Venn diagrams were generated using the eulerr package (both in an R environment). Motif analysis (5–8 bp) was performed using the MEME 5.0.3 software. Functional enrichment analysis using Gene Ontology terms annotated in PlasmoDB release 43 was performed using the Ontologizer 2.1 software[63] with the topology-elim method[64]. Gene set enrichment analysis was performed using GSEA v3.0 Preranked[65].

Whole-genome sequencing analysis, analysis of publicly available genome sequences from field isolates and phylogenetic analysis. To sequence the full genome of control and heat-shock-adapted 3D7-A lines (two biological replicates), we used PCR-free whole-genome Illumina sequencing. Briefly, genomic DNA was sheared to fragments of approximately 150–400 bp in length using a Covaris S220 ultrasonicator and analysed using an Agilent 2100 Bioanalyzer. For library preparation, we used the NEBNext DNA library prep master mix set for Illumina (cat. no. E6040S) using specific paired-end TruSeq Illumina adaptors for each sample. After a quality check by qPCR, we obtained more than six million 150-bp paired reads for each sample using an Illumina MiSeq sequencing system. After checking the quality of the reads (FastQC algorithm) and trimming adaptors (Cutadapt algorithm), the sequence reads were mapped to the PlasmoDB *P. falciparum* 3D7 reference genome release 24 (https://plasmodb.org/plasmo/) using the Bowtie2 local alignment algorithm and the alignment quality was assessed using the QualiMap platform. Average genome coverage was 76- to 98-fold. To identify SNPs and small insertion–deletion mutations (indels), we followed the Genome Analysis Toolkit (GATK)

best-practices workflow, using SAMtools, PicardTools and GATK algorithms. Variant calling was performed using GATK-UnifiedGenotyper. Variants with low calling quality (Phred QUAL < 20) and low read depth (DP < 10) were filtered out using GATK-VariantFiltration, and only variants present in both biological replicates were considered. Differences in the frequency of the SNPs or indels between the control and heat-shock-adapted lines were calculated for each SNP and indel using Microsoft Excel, and those showing <25% difference in any of the two replicates were filtered out. Genome Browse (Golden Helix) was used to visualize the alignments and variants.

For the analysis of publicly available genome sequences, we used the Pf3K project (2016) pilot data release 5 (www.malariagen.net/data/pf3k-5) containing the sequence of >2,500 field isolates. Only SNPs that passed all quality filters and did not fall within a region with multiple large insertions and deletions were considered to be high-confidence. Using these criteria, a single high-confidence polymorphism (occurring in a single isolate) was identified at the *pfap2-hs* gene (producing the C3168X mutation that results in a truncated PfAP2-HS protein that lacks D3).

For sequence alignment and construction of the phylogenetic tree (using the neighbour-joining method), we used Clustal Omega[66] with the default parameters. The cladogram was generated (from the tree without distance corrections) using FigTree 1.4.4.

ChIP experiments and data analysis. For the ChIP experiments, synchronous 50 ml cultures at 2.5–5% parasitemia were harvested at the mid-trophozoite stage. For replicates in which ChIP was performed in parallel under heat-shock and control conditions, the cultures were split off from a single parent flask at the mid-trophozoite stage. Control flasks were immediately returned to 37 °C, whereas the heat-shock flasks were maintained at 41.5 °C for 3 h before harvesting for ChIP analysis.

ChIP, followed by qPCR or Illumina sequencing, was performed as described[67] using the 3F10 rat anti-HA antibody (1:500; Roche, cat. no. 11867423001) to immunoprecipitate the HA-tagged AP2-HS, with the following modification: total chromatin was diluted fivefold in dilution buffer following sonication. The Illumina HiSeq system was used to obtain 125-bp paired-end (replicates 1–3) or 150-bp single-end (replicates 4–5) reads.

Analysis of the ChIP–seq data was performed essentially as described[67]. Briefly, after trimming, quality control, mapping the remaining reads to the *P. falciparum* genome (PlasmoDB release 28) using BWA-MEM and filtering duplicated reads, peak calling was performed using MACS2 (ref. [68]) with a *q*-value cutoff of 0.01. Conversion to $\log_2$-transformed coverage of immunoprecipitate/input was performed using DeepTools BamCompare, selecting the paired-end parameter for all tools when analysing experiments including control and heat-shock conditions. Overlapping intervals within the called peaks for each dataset were determined using Bedtools MultiIntersect. The closest annotated gene coding sequence for each called peak was determined using Bedtools ClosestBed. To visualize the aligned data, we used the Integrative Genomics Viewer (IGV).

The ChIP samples were analysed by qPCR in triplicate wells using the primers described in Supplementary Table 5. All primer pairs were confirmed to have between 80 and 110% efficiency using sheared genomic DNA as a template control. The per cent input was calculated using the formula $100 \times 2^{(Ct\ \text{adjusted input} - Ct\ \text{IP})}$, where C_t adjusted input is the cycle threshold of the adjusted input and C_t IP is the cycle threshold of the immunoprecipitate.

Western blotting. Synchronized cultures at the mature trophozoite stage were exposed to a regular 3 h heat shock or a 1.5 h DHA pulse (10 or 100 nM, used as a positive control for a condition known to produce proteotoxic stress and induce the UPR)[29,38]. Parasites were obtained using saponin lysis (0.15% wt/vol saponin); the pellets were solubilized in 1×SDS–PAGE loading buffer with 4% β-mercaptoethanol and boiled at 95 °C for 5 min. Proteins were resolved by SDS–PAGE on 4–20% TGX Mini-PROTEAN gels (Bio-Rad) and transferred to nitrocellulose membranes (Bio-Rad). After blocking with 5% (wt/vol) bovine serum albumin (Biowest) in TBS-T (0.1% Tween 20 in Tris-buffered saline), the membranes were incubated at 4 °C overnight with the following primary antibodies: rabbit anti-ubiquitin (1:1,000; Cell Signaling Technology, cat. no. 3933), rabbit anti-phospho-eIF2α (1:1,000; Cell Signaling Technology, cat. no. 3398) and rabbit anti-histone H3 (1:1,000; Cell Signaling Technology, cat. no. 9715). After incubation with a goat anti-rabbit IgG–peroxidase (1:5,000; Millipore, cat. no. AP307P) secondary antibody, peroxidase was detected using the Pierce ECL western blotting substrate (Thermo Fisher Scientific) in an ImageQuant LAS 4000 imaging system. To control for equal loading, parts of the membranes corresponding to different molecular-weight ranges were separately hybridized with different antibodies. Signal quantification was performed using ImageJ.

Statistical analysis. Statistical analysis was performed using Microsoft Excel and GraphPad Prism. *P* values were calculated using a two-tailed Student's *t*-test (equal variance). No adjustment for multiple comparisons was made. Only significant *P* values ($P < 0.05$) are shown in the figures. No statistical analysis was performed for experiments involving only two replicates. In all cases, *n* indicates independent biological replicates (that is, samples were obtained from independent cultures).

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

Data availability

The microarray data presented in Fig. 2 and Extended Data Figs. 1,5,6,8 has been deposited to the Gene Expression Omnibus (GEO) database under the accession code GSE149394. The genome sequencing and ChIP–seq data presented in Figs. 1b, 2e and Extended Data Fig. 7 have been deposited to the Sequence Read Archive (SRA) database with the accession codes PRJNA626524 and PRJNA670721, respectively. The authors declare that all other relevant data generated or analysed during this study are included in the Article, the Extended Data or the Supplementary Information files. We used data from the Pf3k pilot data release 5 (www.malariagen.net/pf3k) and different releases of PlasmoDB (www.plasmodb.org) databases. The materials described in this article, including the *P. falciparum* transgenic lines, are available from the corresponding author on reasonable request. Source data are provided with this paper.

Code availability

The scripts used for the analysis of microarray and next generation sequencing data are available at github (https://github.com/CortesMalariaLab/PfAP2-HS_Tinto_etal_NatMicrobiol_2021, with https://doi.org/10.5281/zenodo.4775988).

Received: 12 March 2021; Accepted: 22 June 2021;
Published online: 16 August 2021

References

1. Richter, K., Haslbeck, M. & Buchner, J. The heat shock response: life on the verge of death. *Mol. Cell* **40**, 253–266 (2010).
2. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
3. Anckar, J. & Sistonen, L. Regulation of HSF1 function in the heat stress response: implications in aging and disease. *Annu. Rev. Biochem* **80**, 1089–1115 (2011).
4. Mahat, D. B., Salamanca, H. H., Duarte, F. M., Danko, C. G. & Lis, J. T. Mammalian heat shock response and mechanisms underlying its genome-wide transcriptional regulation. *Mol. Cell* **62**, 63–78 (2016).
5. Solis, E. J. et al. Defining the essential function of yeast Hsf1 reveals a compact transcriptional program for maintaining eukaryotic proteostasis. *Mol. Cell* **63**, 60–71 (2016).
6. Gomez-Pastor, R., Burchfiel, E. T. & Thiele, D. J. Regulation of heat shock transcription factors and their roles in physiology and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 4–19 (2018).
7. Milner, D. A. Jr. Malaria Pathogenesis. *Cold Spring Harb. Perspect. Med.* **8**, a025569 (2018).
8. Kwiatkowski, D. Malarial toxins and the regulation of parasite density. *Parasitol. Today* **11**, 206–212 (1995).
9. Oakley, M. S., Gerald, N., McCutchan, T. F., Aravind, L. & Kumar, S. Clinical and molecular aspects of malaria fever. *Trends Parasitol.* **27**, 442–449 (2011).
10. Gravenor, M. B. & Kwiatkowski, D. An analysis of the temperature effects of fever on the intra-host population dynamics of *Plasmodium falciparum*. *Parasitology* **117**, 97–105 (1998).
11. Kwiatkowski, D. Febrile temperatures can synchronize the growth of *Plasmodium falciparum* in vitro. *J. Exp. Med* **169**, 357–361 (1989).
12. Long, H. Y., Lell, B., Dietz, K. & Kremsner, P. G. *Plasmodium falciparum*: in vitro growth inhibition by febrile temperatures. *Parasitol. Res.* **87**, 553–555 (2001).
13. Portugaliza, H. P. et al. Artemisinin exposure at the ring or trophozoite stage impacts *Plasmodium falciparum* sexual conversion differently. *eLife* **9**, e60058 (2020).
14. Pavithra, S. R., Kumar, R. & Tatu, U. Systems analysis of chaperone networks in the malarial parasite *Plasmodium falciparum*. *PLoS Comput Biol.* **3**, 1701–1715 (2007).
15. Muralidharan, V., Oksman, A., Pal, P., Lindquist, S. & Goldberg, D. E. *Plasmodium falciparum* heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nat. Commun.* **3**, 1310 (2012).
16. Silva, M. D. et al. A role for the *Plasmodium falciparum* RESA protein in resistance against heat shock demonstrated using gene disruption. *Mol. Microbiol* **56**, 990–1003 (2005).
17. Kudyba, H. M. et al. The endoplasmic reticulum chaperone PfGRP170 is essential for asexual development and is linked to stress response in malaria parasites. *Cell. Microbiol.* **21**, e13042 (2019).
18. Lu, K. Y. et al. Phosphatidylinositol 3-phosphate and Hsp70 protect *Plasmodium falciparum* from heat-induced cell death. *eLife* **9**, e56773 (2020).
19. Zhang, M. et al. The apicoplast link to fever-survival and artemisinin-resistance in the malaria parasite. Preprint at *BioRxiv* https://doi.org/10.1101/2020.12.10.419788 (2021).

20. Oakley, M. S. et al. Molecular factors and biochemical pathways induced by febrile temperature in intraerythrocytic *Plasmodium falciparum* parasites. *Infect. Immun.* **75**, 2012–2025 (2007).
21. Rovira-Graells, N. et al. Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res.* **22**, 925–938 (2012).
22. Balaji, S., Babu, M. M., Iyer, L. M. & Aravind, L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* **33**, 3994–4006 (2005).
23. Campbell, T. L., De Silva, E. K., Olszewski, K. L., Elemento, O. & Llinas, M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog.* **6**, e1001165 (2010).
24. Jeninga, M. D., Quinn, J. E. & Petter, M. ApiAP2 transcription factors in apicomplexan parasites. *Pathogens* **8**, 47 (2019).
25. Militello, K. T., Dodge, M., Bethke, L. & Wirth, D. F. Identification of regulatory elements in the *Plasmodium falciparum* genome. *Mol. Biochem. Parasitol.* **134**, 75–88 (2004).
26. Dobson, C. M., Sali, A. & Karplus, M. Protein folding: a perspective from theory and experiment. *Angew. Chem. Int. Ed.* **37**, 868–893 (1998).
27. Masterton, R. J., Roobol, A., Al-Fageeh, M. B., Carden, M. J. & Smales, C. M. Post-translational events of a model reporter protein proceed with higher fidelity and accuracy upon mild hypothermic culturing of Chinese hamster ovary cells. *Biotechnol. Bioeng.* **105**, 215–220 (2010).
28. Bozdech, Z. et al. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, E5 (2003).
29. Dogovski, C. et al. Targeting the cell stress response of *Plasmodium falciparum* to overcome artemisinin resistance. *PLoS Biol.* **13**, e1002132 (2015).
30. Claessens, A., Affara, M., Assefa, S. A., Kwiatkowski, D. P. & Conway, D. J. Culture adaptation of malaria parasites selects for convergent loss-of-function mutants. *Sci. Rep.* **7**, 41303 (2017).
31. Stewart, L. B. et al. Intrinsic multiplication rate variation and plasticity of human blood stage malaria parasites. *Commun. Biol.* **3**, 624 (2020).
32. Manske, M. et al. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. *Nature* **487**, 375–379 (2012).
33. Lamech, L. T. & Haynes, C. M. The unpredictability of prolonged activation of stress response pathways. *J. Cell Biol.* **209**, 781–787 (2015).
34. Blasco, B., Leroy, D. & Fidock, D. A. Antimalarial drug resistance: linking *Plasmodium falciparum* parasite biology to the clinic. *Nat. Med.* **23**, 917–928 (2017).
35. Haldar, K., Bhattacharjee, S. & Safeukui, I. Drug resistance in *Plasmodium*. *Nat. Rev. Microbiol.* **16**, 156–170 (2018).
36. Ariey, F. et al. A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. *Nature* **505**, 50–55 (2014).
37. Birnbaum, J. et al. A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. *Science* **367**, 51–59 (2020).
38. Bridgford, J. L. et al. Artemisinin kills malaria parasites by damaging proteins and inhibiting the proteasome. *Nat. Commun.* **9**, 3801 (2018).
39. Mok, S. et al. Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance. *Science* **347**, 431–435 (2015).
40. Lindner, S. E., De Silva, E. K., Keck, J. L. & Llinas, M. Structural determinants of DNA binding by a *P. falciparum* ApiAP2 transcriptional regulator. *J. Mol. Biol.* **395**, 558–567 (2010).
41. Kafsack, B. F. et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248–252 (2014).
42. Modrzynska, K. et al. A knockout screen of ApiAP2 genes reveals networks of interacting transcriptional regulators controlling the *Plasmodium* life cycle. *Cell Host Microbe* **21**, 11–22 (2017).
43. Zhang, C. et al. Systematic CRISPR–Cas9-mediated modifications of *Plasmodium yoelii* ApiAP2 genes reveal functional insights into parasite development. *mBio* **8**, e01986-17 (2017).
44. Llora-Batlle, O., Tinto-Font, E. & Cortes, A. Transcriptional variation in malaria parasites: why and how. *Brief. Funct. Genomics* **18**, 329–341 (2019).
45. Cortés, A., Benet, A., Cooke, B. M., Barnwell, J. W. & Reeder, J. C. Ability of *Plasmodium falciparum* to invade Southeast Asian ovalocytes varies between parasite lines. *Blood* **104**, 2961–2966 (2004).
46. Cortés, A. A chimeric *Plasmodium falciparum Pfnbp2b*/*Pfnbp2a* gene originated during asexual growth. *Int. J. Parasitol.* **35**, 125–130 (2005).
47. Cortés, A. et al. Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog.* **3**, e107 (2007).
48. Walliker, D. et al. Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* **236**, 1661–1666 (1987).
49. Anders, R. F., Brown, G. V. & Edwards, A. Characterization of an S antigen synthesized by several isolates of *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **80**, 6652–6656 (1983).
50. Ghorbal, M. et al. Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR–Cas9 system. *Nat. Biotechnol.* **32**, 819–821 (2014).
51. Lim, M. Y. et al. UDP-galactose and acetyl-CoA transporters as *Plasmodium* multidrug resistance genes. *Nat. Microbiol* **1**, 16166 (2016).
52. Bancells, C. et al. Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*. *Nat. Microbiol.* **4**, 144–154 (2019).
53. Knuepfer, E., Napiorkowska, M., van Ooij, C. & Holder, A. A. Generating conditional gene knockouts in *Plasmodium*—a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Sci. Rep.* **7**, 3881 (2017).
54. Rovira-Graells, N., Aguilera-Simon, S., Tinto-Font, E. & Cortes, A. New assays to characterise growth-related phenotypes of *Plasmodium falciparum* reveal variation in density-dependent growth inhibition between parasite lines. *PLoS ONE* **11**, e0165358 (2016).
55. Prasad, R. et al. Blocking *Plasmodium falciparum* development via dual inhibition of hemoglobin degradation and the ubiquitin proteasome system by MG132. *PLoS ONE* **8**, e73530 (2013).
56. Crowley, V. M., Rovira-Graells, N., de Pouplana, L. R. & Cortés, A. Heterochromatin formation in bistable chromatin domains controls the epigenetic repression of clonally variant *Plasmodium falciparum* genes linked to erythrocyte invasion. *Mol. Microbiol.* **80**, 391–406 (2011).
57. Casas-Vila, N., Pickford, A. K., Portugaliza, H. P., Tintó-Font, E. & Cortés, A. Transcriptional analysis of tightly synchronized *Plasmodium falciparum* intraerythrocytic stages by RT–qPCR. *Methods Mol. Biol.* https://doi.org/10.1007/978-1-0716-1681-9_10 (2021).
58. Kafsack, B. F., Painter, H. J. & Llinas, M. New Agilent platform DNA microarrays for transcriptome analysis of *Plasmodium falciparum* and *Plasmodium berghei* for the malaria research community. *Malar. J.* **11**, 187 (2012).
59. Painter, H. J., Altenhofen, L. M., Kafsack, B. F. & Llinas, M. Whole-genome analysis of *Plasmodium* spp. utilizing a new agilent technologies DNA microarray platform. *Methods Mol. Biol.* **923**, 213–219 (2013).
60. Llorà-Batlle, O. et al. Conditional expression of PfAP2-G for controlled massive sexual conversion in *Plasmodium falciparum*. *Sci. Adv.* **6**, eaaz5057 (2020).
61. Lemieux, J. E. et al. Statistical estimation of cell-cycle progression and lineage commitment in *Plasmodium falciparum* reveals a homogeneous pattern of transcription in ex vivo culture. *Proc. Natl Acad. Sci. USA* **106**, 7559–7564 (2009).
62. Saeed, A. I. et al. TM4 microarray software suite. *Methods Enzymol.* **411**, 134–193 (2006).
63. Bauer, S., Grossmann, S., Vingron, M. & Robinson, P. N. Ontologizer 2.0–a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics* **24**, 1650–1651 (2008).
64. Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).
65. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
66. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
67. Josling, G. A. et al. Dissecting the role of PfAP2-G in malaria gametocytogenesis. *Nat. Commun.* **11**, 1503 (2020).
68. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

Acknowledgements

We thank J. J. López-Rubio (University of Montpellier) for the plasmid pL6-*egfp*, M. Lee (Wellcome Sanger Institute) for the plasmid pDC2-Cas9-U6-hdhfr and E. Knuepfer (The Francis Crick Institute) for the plasmid pDC2-Cas9-U6-hDHFRyFCU. We also thank O. Llorà-Batlle and C. Sànchez-Guirado (ISGlobal) for their assistance with the generation of the plasmids used in this study, N. Rovira-Graells (ISGlobal) and A. Gupta. (Nanyang Technological University) for assistance with the 3D7-A and 3D7-A-HS microarray experiments, O. Billker (Wellcome Sanger Institute) for the experiments attempted in *P. berghei* and H. Ginsburg (The Hebrew University of Jerusalem) for providing data from the Malaria Parasite Metabolic Pathways. This publication uses data generated by the Pf3k project (www.malariagen.net/pf3k). This work was supported by grants from the Spanish Ministry of Economy and Competitiveness (MINECO)/Agencia Estatal de Investigación (AEI) (grant nos SAF2013-43601-R, SAF2016-76190-R and PID2019-107232RB-I00 to A.C.), co-funded by the European Regional Development Fund (ERDF, European Union), and from the NIH/NIAID (grant no. 1R01 AI125565 to M.L.). E.T.-F. and L.M.-T. were supported by fellowships from the Spanish Ministry of Economy and Competitiveness (grant nos BES-2014-067901 and BES-2017-081079, respectively), co-funded by the European Social Fund (ESF). T.J.R. was supported by a training grant from the NIH/NIGMS (grant no. T32 GM125592-01). This research is part of ISGlobal's Program on the Molecular Mechanisms of Malaria, which is partially supported by the Fundación Ramón Areces. We acknowledge support from the Spanish Ministry of Science and Innovation through the 'Centro de Excelencia Severo Ochoa 2019–2023' Program (grant no. CEX2018-000806-S) and support from the Generalitat de Catalunya through the CERCA Program.

Author contributions

E.T.-F. performed all of the experiments, except for those presented in Extended Data Fig. 1 and the western blot and ChIP–seq experiments. L.M.-T., E.T.-F., T.J.R. and A.C. performed the bioinformatics analyses. N.C.-V. performed the western blot experiments. T.J.R. performed, and M.L. supervised, the ChIP–seq experiments. Z.B. provided microarray hybridizations for the experiments presented in Extended Data Fig. 1. D.J.C. provided advice on clinical isolates and provided Line 1 from The Gambia. E.T.-F. and A.C. conceived the project, designed and interpreted the experiments and wrote the manuscript (with input from all authors and major input from M.L. and D.J.C.).

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-021-00940-w.

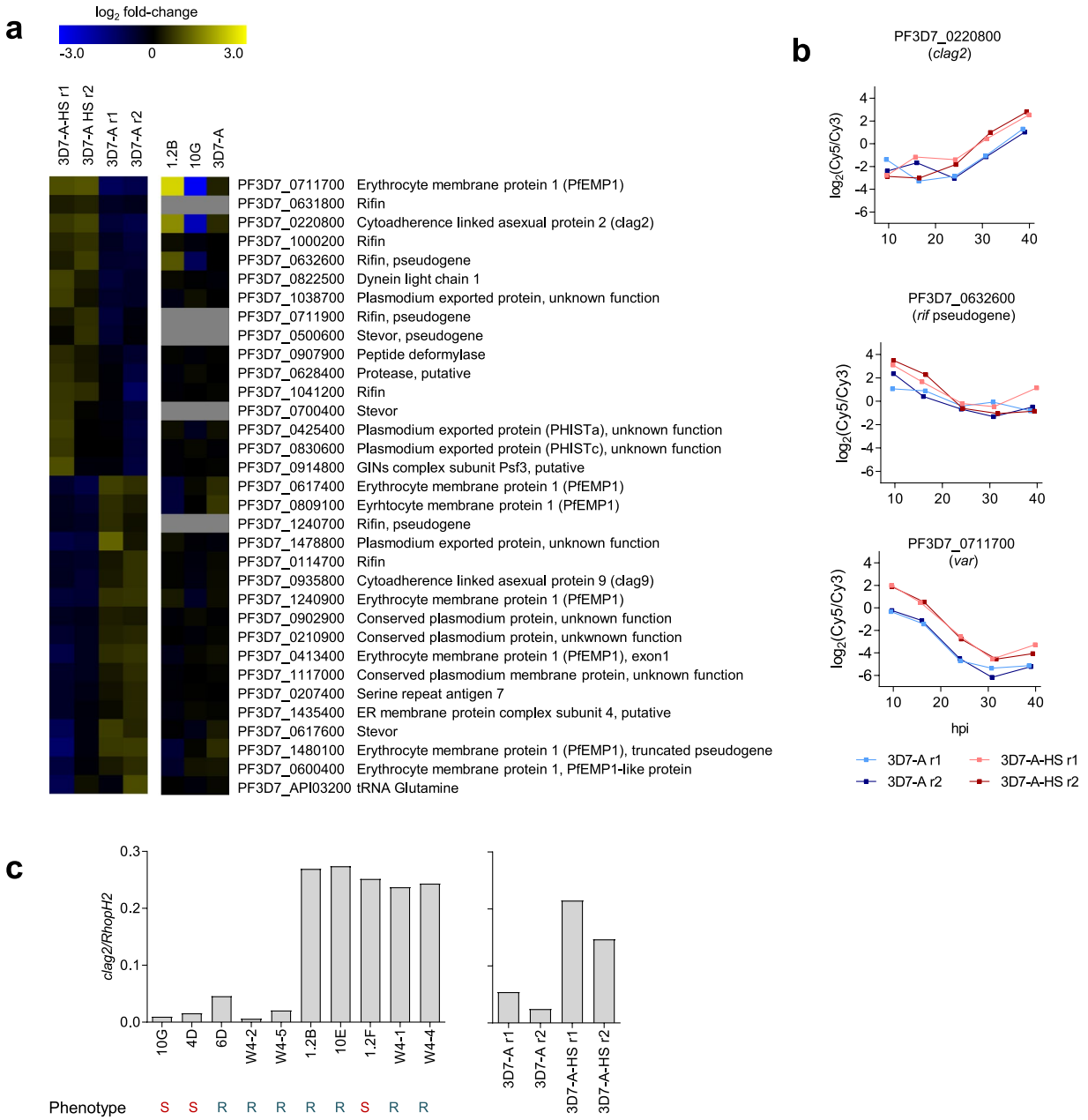
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-021-00940-w.

Correspondence and requests for materials should be addressed to A.C.

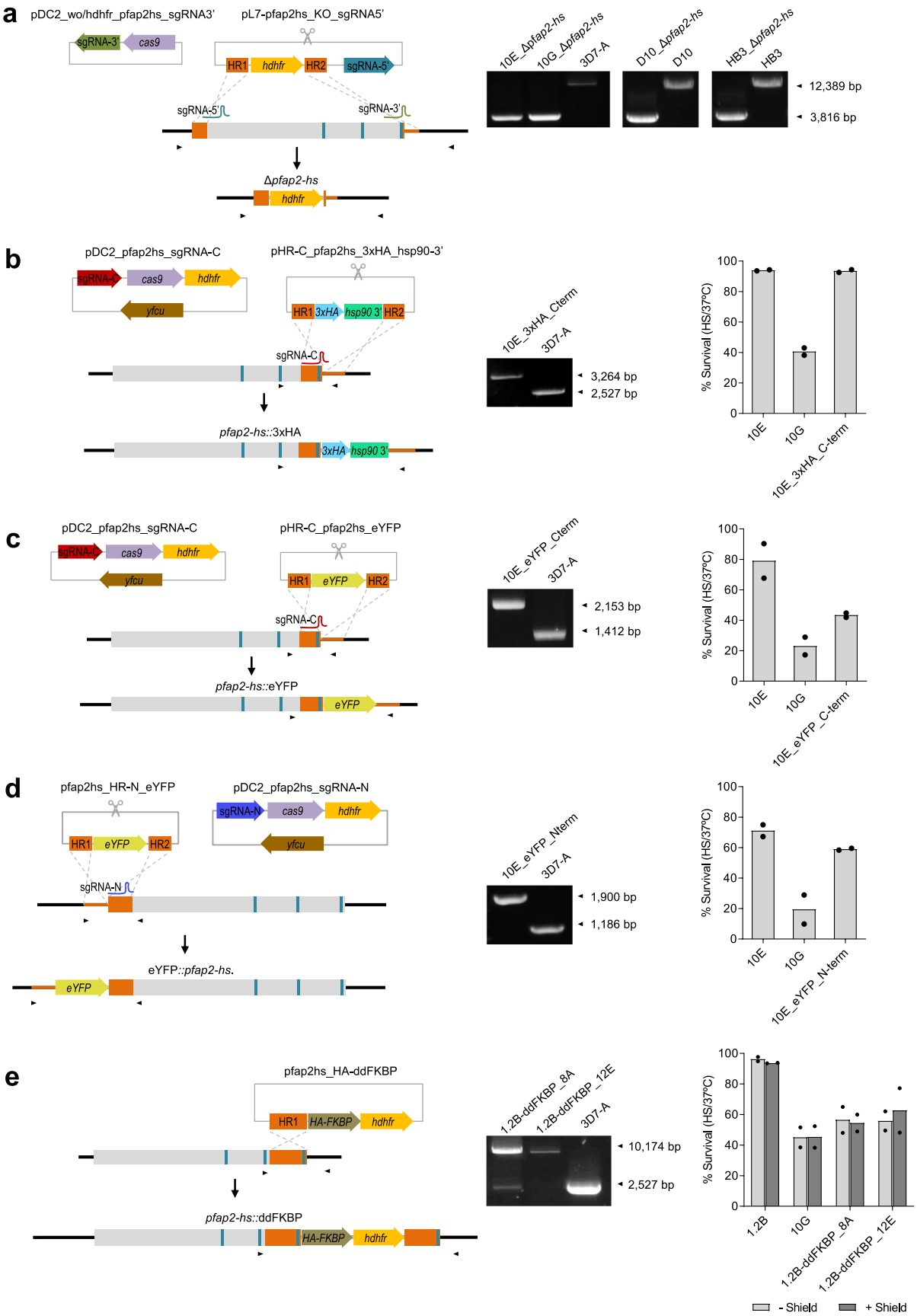
Peer review information *Nature Microbiology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

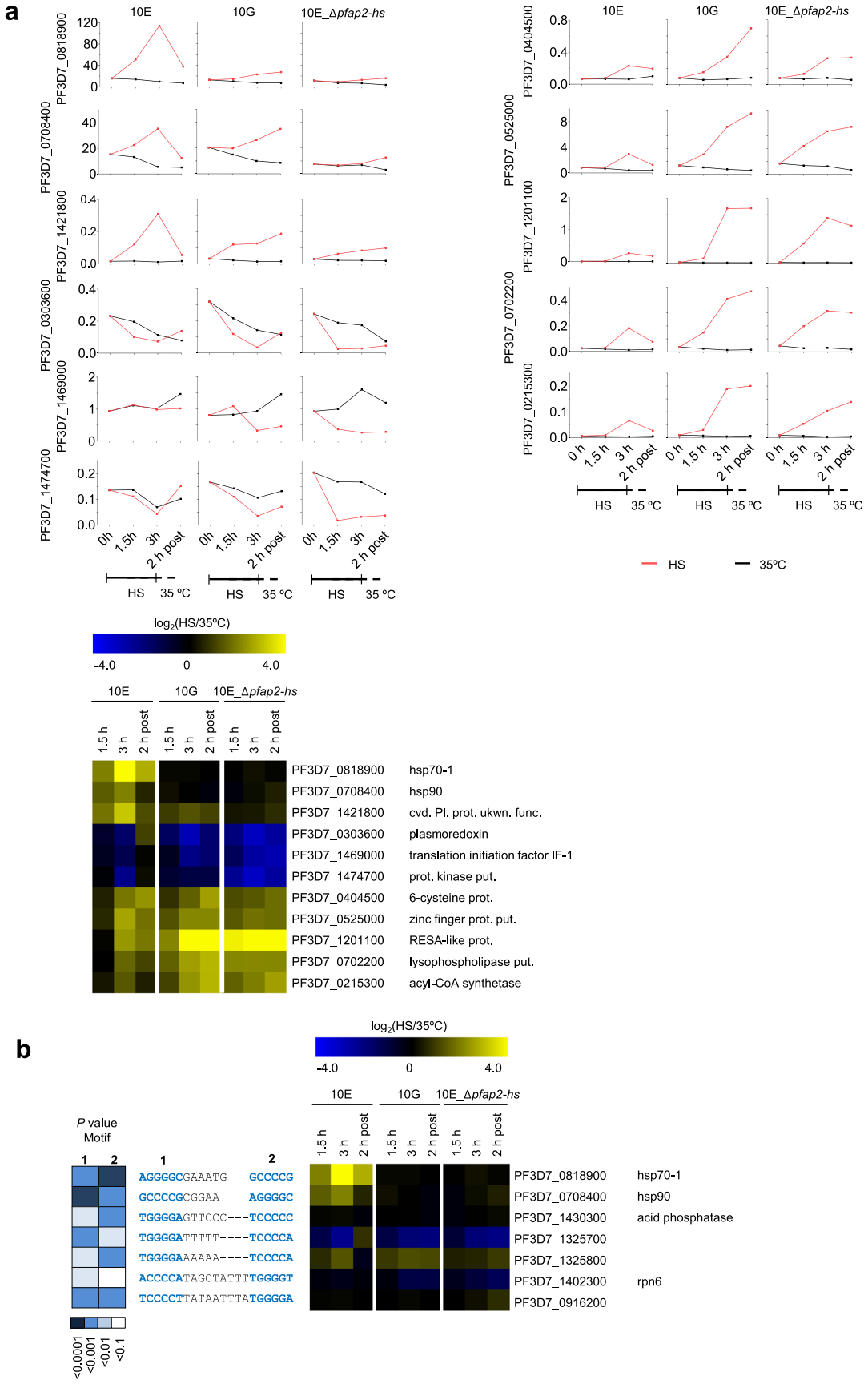


Extended Data Fig. 1 | Transcriptional analysis of heat-shock-adapted and control lines. a, Microarray-based transcriptomic comparison across the asexual blood cycle of 3D7-A cultures adapted to heat shock after five cycles of selection with a 3 h heat shock at 41.5 °C at the trophozoite stage (3D7-A-HS r1 and r2) and control parasite lines maintained in parallel without heat shock (3D7-A r1 and r2). Values are the $\log_2$ of the maximum expression fold-change (from the average of all lines compared) across a time interval corresponding to half the length of the asexual cycle, calculated using the aMAFC score as previously described[21]. Genes with a > 1.5-fold-change in expression in the two independent 3D7-A heat-shock-adapted lines (3D7-A-HS r1 and r2) relative to their respective controls (3D7-A r1 and r2) are shown. Data for parasite lines 10 G (heat-shock-sensitive subclone), 1.2B (heat-shock-resistant subclone) and 3D7-A (right panel) is from Rovira-Graells et al[21]. **b**, Time-course expression of genes in panel a that showed a concordant change in expression between heat-shock-adapted and control cultures, and between the heat-shock-resistant subclone 1.2B and the heat-shock-sensitive subclone 10 G. Based on the predicted function of the three genes, *clag2* was considered the most plausible candidate to play a role in heat-shock resistance. **c**, Expression of *clag2* is neither necessary nor sufficient for heat-shock resistance. RT-qPCR analysis of *clag2* transcript levels (normalized against *rhoph2*) in schizonts of heat-shock sensitive (S) and heat-shock resistant (R) 3D7-A subclones (see Fig. 1f), and of the heat-shock-adapted and control lines ($n = 1$ biological replicates).



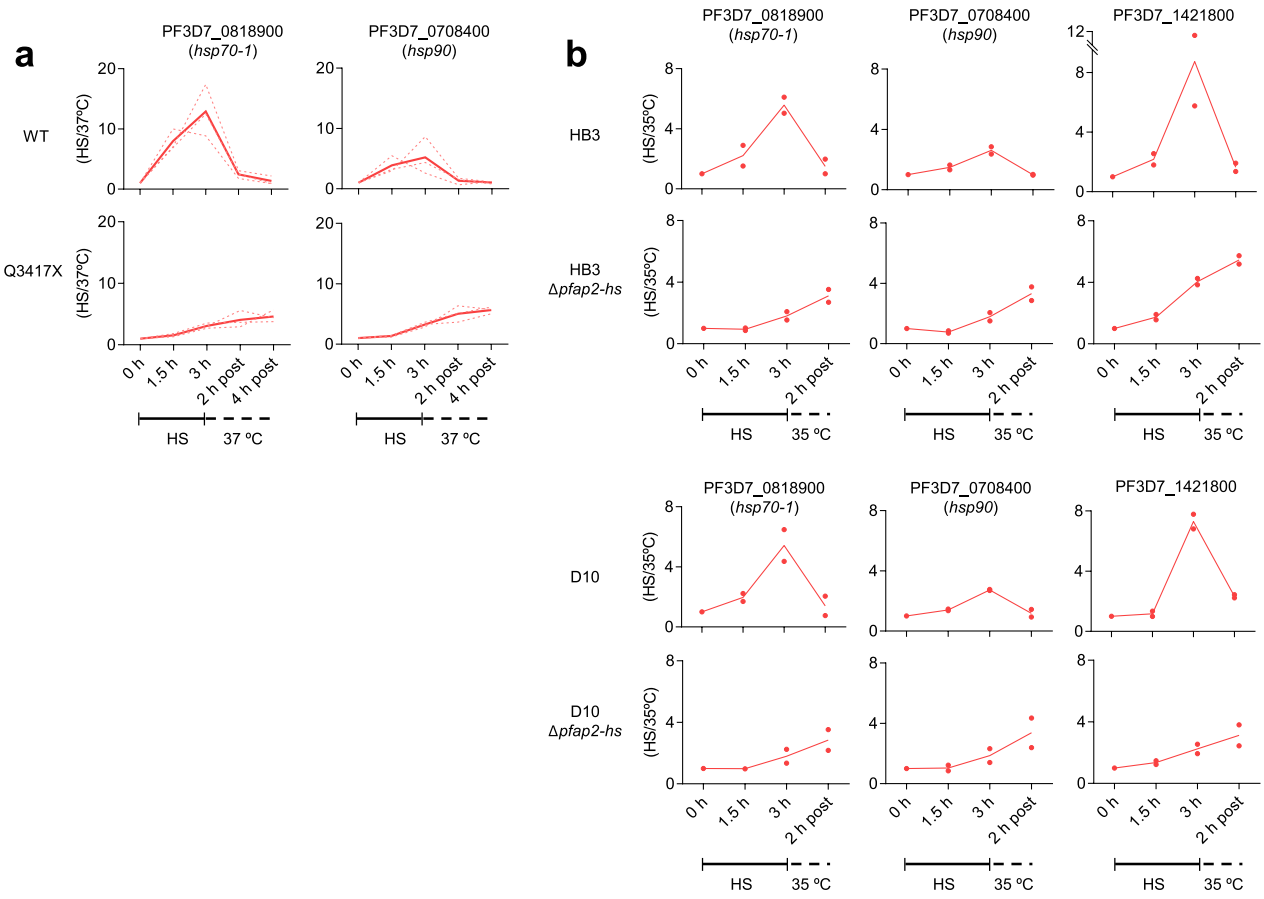
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Generation and characterization of transgenic parasite lines edited at the *pfap2-hs* locus. a, Schematic of the CRISPR–Cas9 strategy used to knockout *pfap2-hs*, using two guide RNAs. **b**–**d**, Tagging of endogenous PfAP2-HS using CRISPR–Cas9 technology. The tags used were a C-terminal 3xHA (**b**), a C-terminal eYFP (**c**) and an N-terminal eYFP (**d**). **e**, C-terminal tagging of endogenous PfAP2-HS by single homologous recombination with a tag consisting of a 2xHA epitope and an FKBP destabilization domain (DD domain). In all panels, the position of the primers used for analytical PCR (arrowheads), guide RNA and AP2 domains (blue vertical bars) is indicated. The electrophoresis images at the right are the analytical PCR validation of the genetic edition (single genomic DNA extraction and PCR analysis), showing correct edition and absence of wild-type locus in all cases except for the 8 A subclone of 1.2B-ddFKBP (8 A and 12E are subclones obtained after drug cycling). The bar charts at the right show the level of survival (mean of $n = 2$ independent biological replicates) of sorbitol-synchronized cultures of the transgenic lines upon heat shock (HS) exposure at the trophozoite stage, with heat-shock-resistant (10E) and heat-shock-sensitive (10 G, expressing PfAP2-HSΔD3) subclones as controls. Addition of a C-terminal eYFP or HA-FKBP tag did not affect growth at 37 °C but resulted in high heat-shock sensitivity, similar to the 10 G line. In contrast, C-terminal addition of the smaller 3xHA tag or addition of an N-terminal eYFP did not affect growth at 37 °C or heat-shock sensitivity. In all cases, tagged PfAP2-HS was not detectable by immunofluorescence or Western blot analysis, probably as a consequence of the very low abundance of this transcription factor (see proteomic data in www.PlasmoDB.org).

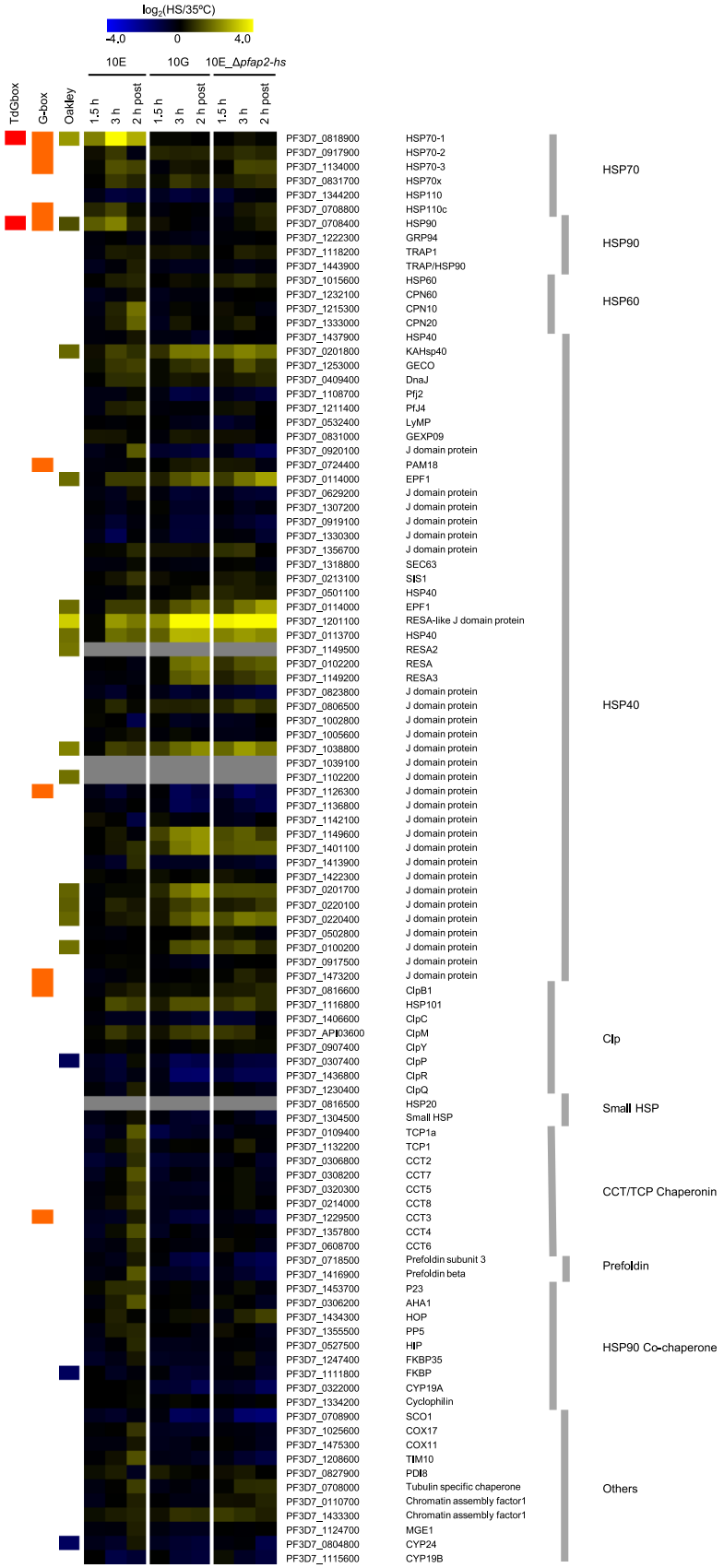


Extended Data Fig. 3 | See next page for caption.

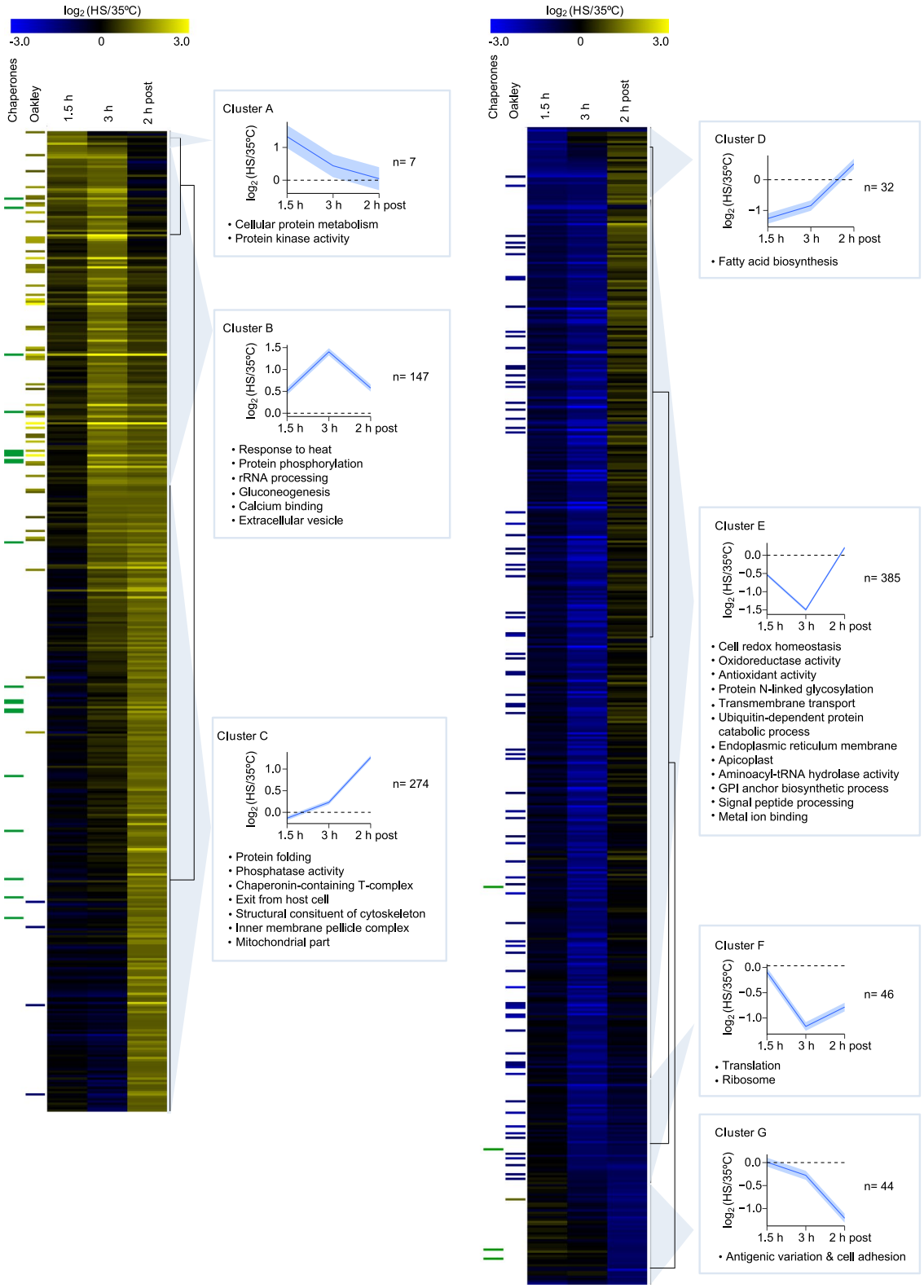
Extended Data Fig. 3 | Validation of the transcriptomic changes upon heat shock and distribution of the tandem G-box motif. a, RT-qPCR analysis of transcript levels (normalized against *serine–tRNA ligase*) of the genes selected for validation, using biological samples independent from the samples used for microarray analysis. Values are the average of triplicate reactions. The $\log_2$ expression fold-change [heat shock (HS) relative to control (35 °C) conditions] for these genes in the microarray analysis (Fig. 2a) is shown in a heatmap to facilitate comparison. **b**, Genes in the *P. falciparum* genome containing tandem arrangements (maximum distance between the two: 9 nucleotides) of the G-box [(A/G)NGGGG(C/A)] motif in their regulatory regions (defined as the 2 kb upstream of the start codon or until the neighbour gene, when it is closer). The sequence of the G-box in each gene is shown in blue, and the level of concordance with the consensus G-box motif is expressed as the *P* value of the match (determined using the FIMO v5.0.5 function in the MEME suite). Expression changes upon heat shock for these genes are shown as in panel **a**.



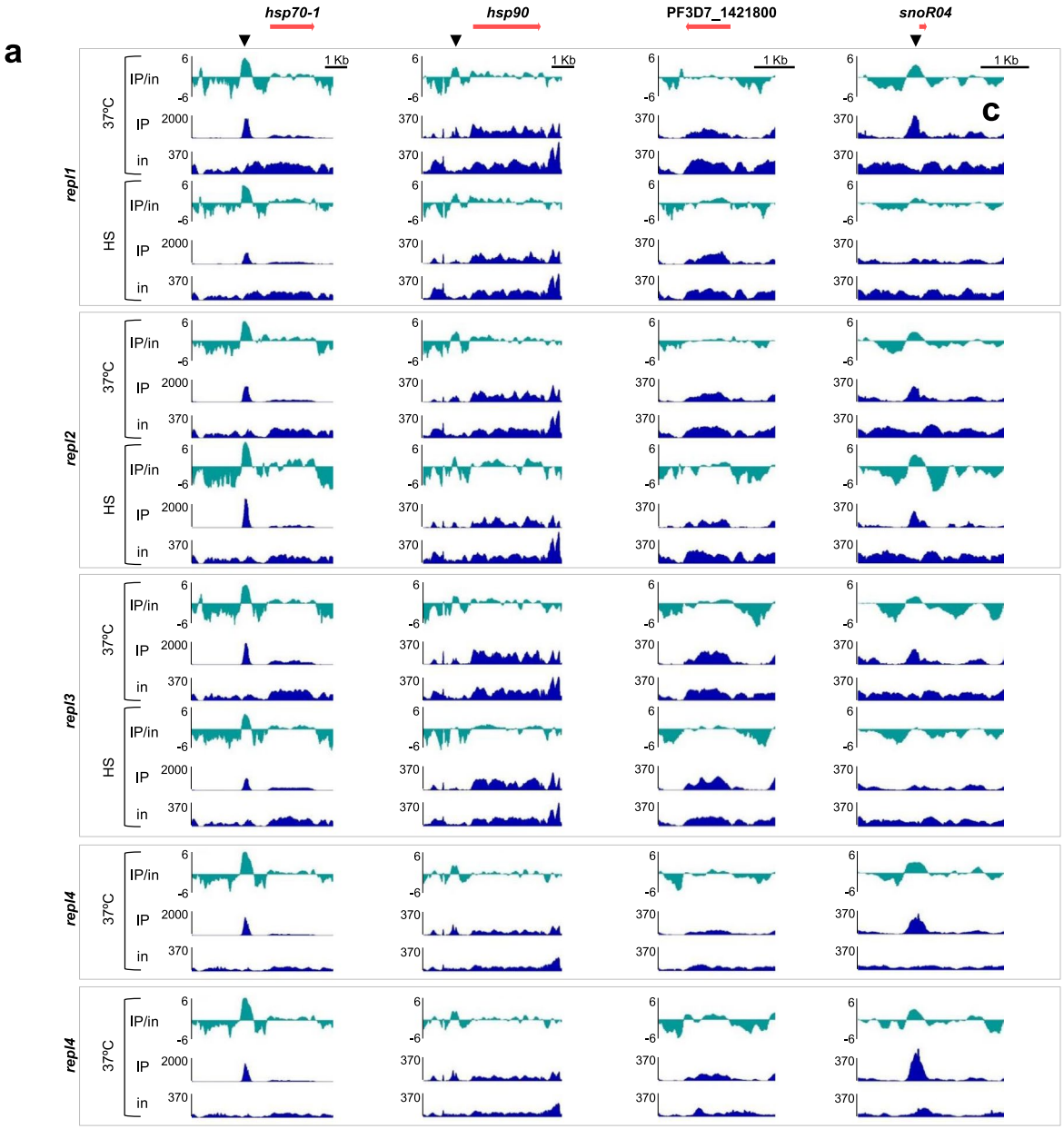
Extended Data Fig. 4 | Changes in *hsp70-1, hsp90* and PF3D7_1421800 transcript levels in parasites lacking the entire PfAP2-HS or D3. Fold increase in transcript levels (determined by RT-qPCR, normalized against *serine–tRNA ligase*) during and after heat shock (HS) starting at 33-35 (**a**) or 30-35 (**b**) h.p.i., relative to cultures maintained in parallel without heat shock (37 or 35 °C). In panel **a**, values for three individual 3D7-A subclones carrying or not the Q3417X mutation are shown as dotted lines, whereas the average of the three subclones is shown as a continuous line. In panel **b**, the mean of $n = 2$ independent biological replicates is shown.



Extended Data Fig. 5 | Transcript level changes upon heat shock in chaperone-encoding genes. Log_2 expression fold-change [heat shock (HS) relative to control (35 °C) conditions, as in Fig. 2a] for all chaperone-encoding genes described by Pavithra and colleagues[14]. Columns at the left indicate presence of the G-box[23] or tandem G-box (TdGbox) in the upstream region, and log_2 fold-change during heat shock in a previous study[20] (Oakley).

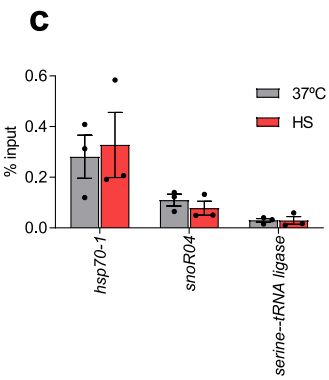


Extended Data Fig. 6 | Transcriptomic characterization of the heat-shock response in parasites expressing complete PfAP2-HS (10E line). Log_2 expression fold-change [heat shock (HS) relative to control (35 °C) conditions, as in Fig. 2a] in the wild-type 10E line determined by microarray analysis. Genes with a fold-change ≥2 at any of the time points analysed are shown. The mean log_2 expression fold-change (with 95% confidence interval) and representative enriched GO terms are shown for each cluster. Columns at the left indicate fold-change during heat shock in a previous study[20] (Oakley), and annotation as chaperone[14]. Ten genes had values out of the range displayed (actual range: -3.89 to +4.03).



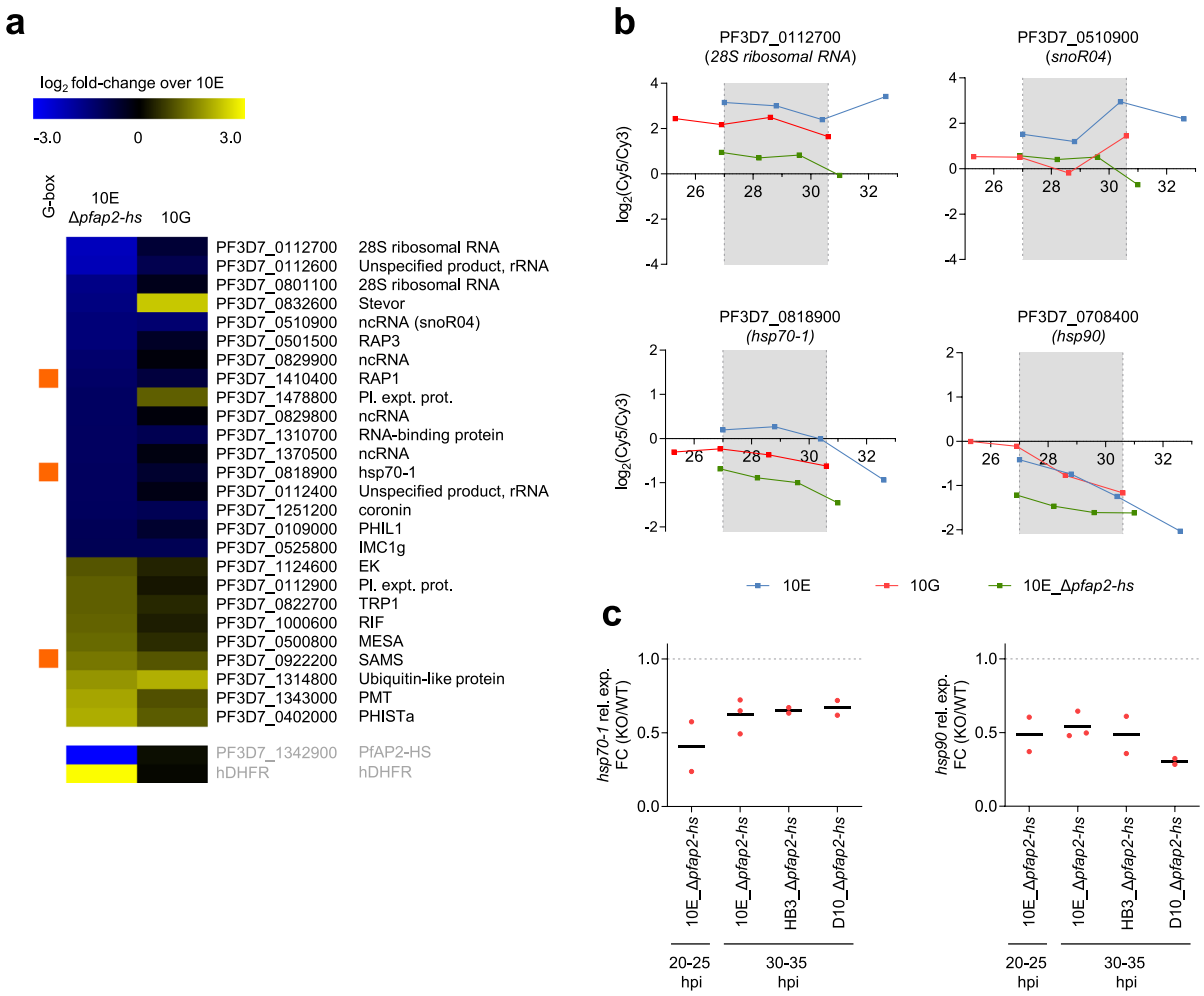
b

	Chrom.	Start	End	Median MACS2 Score	Median Fold Enrich.	Closest Gene	Gene Name	Location
Control	8	859109	859513	3155	14.6	PF3D7_0818900	*hsp70-1*	5'
	5	465548	465850	429	4.2	PF3D7_0510900	*snoR04*	5' & gene body
	13	2895119	2897998	150	2.3	Telomere	-	–
HS	8	859123	859514	3410	21.8	PF3D7_0818900	*hsp70-1*	5'
	10	1437164	1437430	226	3.5	PF3D7_1036400	*lsa1*	Gene body
	13	2909970	2912982	181	2.9	Telomere	-	-

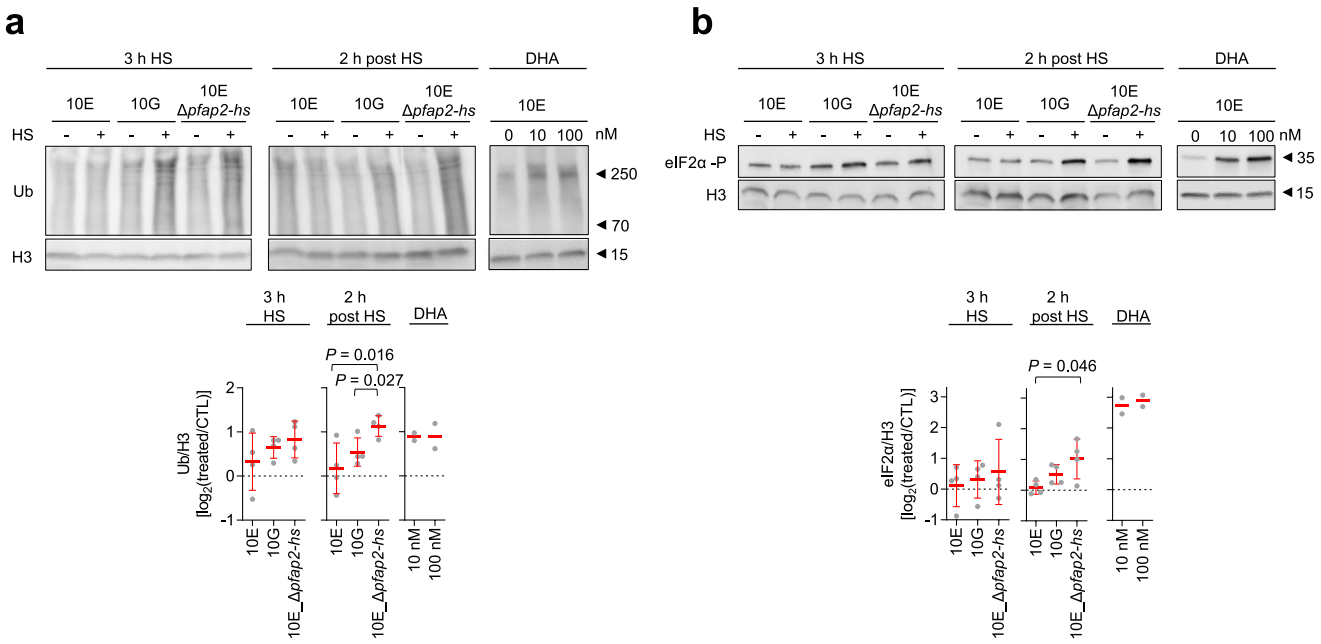


Extended Data Fig. 7 | See next page for caption.

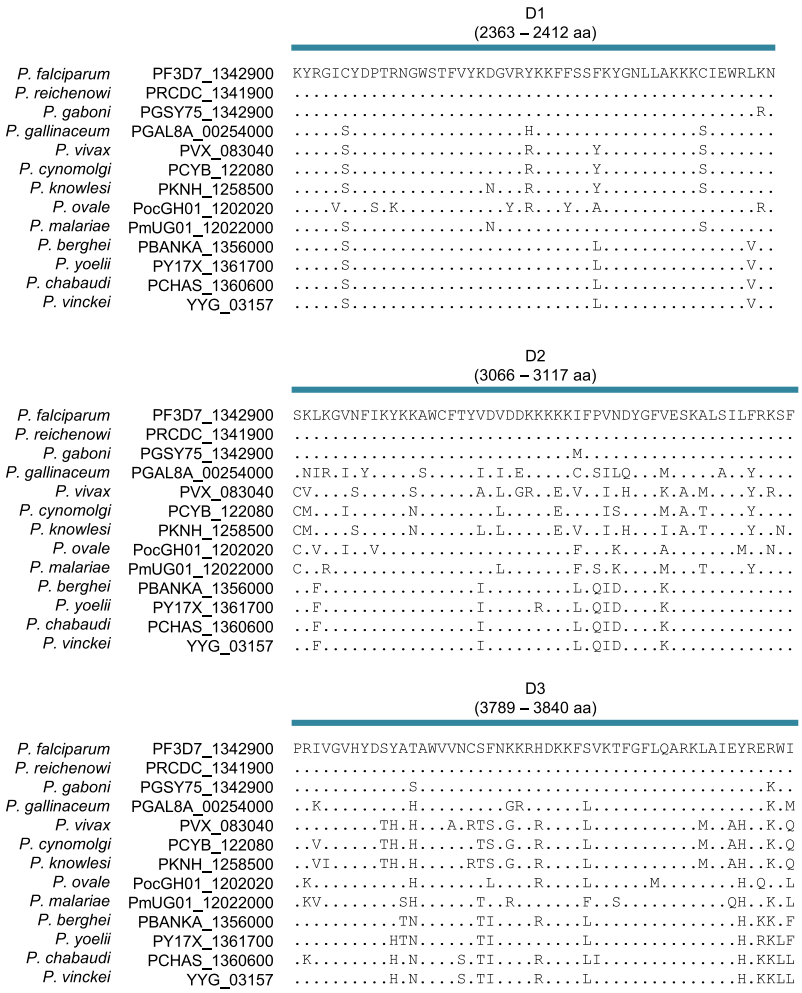
Extended Data Fig. 7 | ChIP analysis of the chromosomal distribution of PfAP2-HS. a, ChIP–seq analysis of HA-tagged PfAP2-HS. Number of reads of ChIP (IP) and input (in) tracks, and $\log_2$-transformed ChIP/input ratio tracks (IP/in) for five independent biological replicates [three including heat shock (HS) and 37 °C conditions, two including only the 37 °C condition]. Snapshots are shown for the three genes in cluster I (Fig. 2a) and *snoR04*. Binding at the *hsp70-1* and *hsp90* promoters coincides with the position of a tandem G-box motif, whereas PF3D7_1421800 and *snoR04* lack a G-box. The positions of the relevant peaks are indicated by an arrowhead. **b**, Peaks present in ≥3 out of 5 replicate ChIP–seq experiments (37 °C) or ≥2 out of 3 replicate experiments (heat shock) and with a MACS score >100 in each positive replicate. **c**, ChIP–qPCR analysis of HA-tagged PfAP2-HS binding at selected loci, in cultures exposed to heat shock (HS) or control (37 °C) conditions (mean and s.e.m. of % input in $n = 3$ independent biological replicates). No significant difference ($P < 0.05$) was observed between 37 °C and heat shock using a two-sided unpaired Student's *t*-test.



Extended Data Fig. 8 | Transcriptional changes associated with PfAP2-HS deletion under basal (no heat shock) conditions. a, Changes in transcript levels in the absence of heat shock for genes with an average expression fold-change >2 between 10E_Δ*pfap2-hs* and 10E. Values are the $\log_2$ of the average expression fold-change relative to 10E across the time period compared (~27–30.5 hpi). Genes artificially modified or introduced in the knockout line, which serve as controls, are shown at the bottom (their values are out of the range displayed). The column at the left indicates the presence of the G-box[23]. **b**, Expression plots for selected genes under basal conditions. Expression values are plotted against statistically estimated parasite age, expressed in h post-invasion (h.p.i.). Grey shading marks the interval used to calculate the average expression fold-change. **c**, RT-qPCR analysis of *hsp70-1* and *hsp90* transcript levels in *pfap2-hs* knockout (KO) lines compared to their wild-type (WT) controls (the parental line for each knockout line) under basal conditions. Expression values are normalized against *serine–tRNA ligase*, and expressed as the fold-change (FC) in the knockout versus control lines. The mean of $n = 3$ (10E, 30-35 hpi) or $n = 2$ (others) independent biological replicates is shown.



Extended Data Fig. 9 | Analysis of proteome stress and UPR markers in *pfap2-hs* mutants. a, **b**, Western blot analysis (representative of $n = 4$) of polyubiquitinated proteins (Ub) (**a**) or phosphorylated eIF2α (eIF2α-P) (**b**) immediately after a 3 h heat shock (3 h HS) and 2 h later (2 h post HS). Histone H3 is a loading control. DHA was used as a positive control, as it is a known inducer of the UPR[29,38]. The Log_2 of histone H3-normalized signal in heat shock or DHA-treated cultures versus control cultures is shown at the bottom (mean and s.e.m. of $n = 4$, except for the DHA control mean of $n = 2$ independent biological replicates). *P* values were calculated using a two-sided unpaired Student's *t*-test. Only significant *P* values ($P < 0.05$) are shown. The position of molecular weight markers is shown (in kDa).



Extended Data Fig. 10 | Sequence alignment of the three AP2 domains (D1–3) present in AP2-HS orthologues in *Plasmodium* spp. Dots indicate identity with the amino acid in the first sequence.