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

Elisabet Tintó-Font1, Lucas Michel-Todó1,7, Timothy J. Russell2,7, Núria Casas-Vila1, David J. Conway2, Zbynek Bozdech4, Manuel Llinás2,5 and Alfred Cortés1,6

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.
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 from other 3D7 stocks in our laboratory and the 3D7 reference genome, used a heat-shock survival assay consisting of heat-shock at 41.5 °C for 3 h (ref. 24) 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 10E 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...
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 at 35°C. 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_\Delta pfap2-hs and 10G_\Delta 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-HSΔ3D. 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-HSΔ3D) and 10E_\Delta 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_\Delta pfap2-hs. Cluster I comprises only three genes: a gene of unknown function (PF3D7_1421800), the cytoplasmic hsp70-1 (PF3D7_0708400), and hsp90 (PF3D7_124200; 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 14,25 (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 necessary for activation by PfAP2-HS. The strongest transcriptional response to heat shock is shown in the HS versus 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 log2-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. 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 log2-transformed values of the fold change in expression in the HS versus control (35 ºC) 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 chaperone14, presence of the G-box23 or tandem G-box (T_dGbox) in the upstream region, and log2-transformed fold-change values after HS from a previous study20. 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. d. Age progression during the HS assay, statistically estimated 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 log2-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 Δpfap2-hs and parental lines with a 3D7 genetic background at different temperatures (mean ± s.e.m. of n = 4 independent biological replicates). 10E_\Delta pfap2-hs: 37 versus 35 ºC, P = 2.3 × 10⁻³; 37.5 versus 35 ºC, P = 1.7 × 10⁻⁴; and 10G_\Delta pfap2-hs: 37 versus 35 ºC, P = 0.01); 37.5 versus 35 ºC, 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.
**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 *hs90*).

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 *hs90* 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_Δpfap2-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_Δpfap2-hs lines indeed showed a more pronounced deviation from a reference transcriptome\(^4\) 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_Δpfap2-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\(^4\) 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\(^3\) (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 *hs70* and most of its other target promoters under both basal and heat-shock conditions\(^1\). Enrichment for PfAP2-HS at the *hs90* 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* (Fig. 2e, 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_Δpfap2-hs and 10G_Δpfap2-hs) showed severe temperature-dependent
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_pFap2-hs line also had clearly reduced growth at 37 °C compared with 35 °C, whereas the HB3_pFap2-hs line did not (Fig. 3b). Both 3D7_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 stress29. 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 conditions9,31 but these are virtually absent from clinical isolates (in the www.malarigenen.net/data/pf3k-5 dataset32, 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 domain9 (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_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 snr044 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_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...
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 detrimental33, 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 damage34,35. Resistance to artemisinins is associated with mutations in the Kelch13 protein36,37 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 proteotoxic stress, such as by DHA or proteasome inhibition.

**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.
to the disruption of other factors involved in proteostasis maintenance. The 10E_Δpfap2-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 and in the cell is likely to participate in protein–protein interactions or dimerization. Other components of the protein-folding machinery necessary for heat-shock survival 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 HSFI—the conserved master regulator of the heat-shock response in most eukaryotes, from yeast to mammals—it serves an analogous role. HSFI regulates a compact transcriptional programme that includes the hsp70 and hsp90 genes. In yeast, the only essential role of this transcription factor is the activation of hsp70 and hsp90 (ref. 8), 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 HSFI (ref. 3). In other organisms, the heat-shock response mediates protection against different types of proteotoxic stress, in addition to high temperature. Here we report that parasites lacking PfAP2-HS have increased sensitivity to arteisinin, 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, 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.

Methods

Parasite cultures. The 3D7-A stock of the clonal P. falciparum line 3D7 (ref. 24), the 3D7-A subclones 10G, 1.2B, 10F, 4D, 6D, 1.2F, W4-1, W4-2, W4-4 and W4-5 (refs. 46–49), the HB3B10 (mosquito and chimpanzee-passaged HB3, provided by O. Kaneko, Ehime University, Japan) and D10 (ref. 21; provided by R. F. Anders, La Trobe University, Australia) clonal parasite lines, and the culture-adapted Line 1 from the Gambian40 have been 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 successive 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. Parasites were cultured in B8 erythrocytes at a 3% haematocrit under standard culture conditions in RPMI-based media containing Albumax II (without human serum), in an atmosphere of 5% CO2, 3% O2 and the balance N2 (except for cultures for ChIP–seq experiments, in which O2 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 (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 5’ end of the gene (positions 4,846–11,305). The guide was cloned into the donor plasmid 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, flankng the yfcu cassette, to generate the plasmid, pHR_C_pfap2hs_sgRNA_A.

The 3’ guide was cloned into a modified version of the pDC2-Cas9-U6-hdhfr plasmid in which we previously removed the hdhfr expression cassette by digesting with Ncol and SacII, end blunting and re-ligation, and replaced the BbsI guide cloning site with a BglII site. The resulting plasmid was named pDC2_woo/hdhfr_pfp2hs_sgRNA_A. All guides were cloned into the In-Fusion system (Takara) as linearized plasmids, whereas homology regions were PCR-amplified from genomic DNA and cloned by ligation using the restriction sites Spel and AflII (HR1), and EcoRI and Ncol (HR2).

For the constructs aimed at C-terminal tagging of pfap2-hs using CRISPR–Cas9 (10E_pnap2-hs_eYFP-Cerm and 10E_pnap2-hs_3xHA-Cerm lines), we used a guide corresponding to positions 11,508–11,527 of the gene (Extended Data Fig. 2b,c and Supplementary Table 5). The guide was cloned into the pDC2-Cas9-U6-hdhfr plasmid to obtain pDC2_pfp2hs_sgRNA_C. The donor plasmid for tagging with enhanced yellow fluorescent protein (eYFP; pHR_C_pfp2hs_eYFP) was based on the plasmid pHRap2g-eYFP52, with the pfp2-hs homology region and the 3’ sequence of hsp90 replaced with pfp2-hs coding sequences generated with the HR1 guide. The HR1 region was generated with an HpaI restriction site 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 of the In-Fusion system. The HR2 region was a fragment containing positions +1 to +490 after the pfap2-hs stop codon. It was cloned into the Xhol–Aart restriction sites. The donor plasmid for 3xHA-C terminal tagging (pHR_C_pfp2hs_3xHA_hsp90-3’) was also based on the plasmid pHRap2g-eYFP52, with the eYFP coding sequence replaced by the 3xHA sequence (amplified from the plasmid pHH11w-pfp2-g-HA3x; ref. 41) and the same homology regions as in the plasmid pHR_C_pfp2hs_eYFP (but HR2 was cloned, using the In-Fusion system, into the EcoRI–Aart sites because in this construct the 3’ end of hsp90 in pHRap2g-eYFP was maintained).

For amino (N)-terminal tagging (10E_pfp2-hs_eYFP-Nterm line), we cloned the targeting the pfap2-hs positions 73–92 in the pDC2-Cas9-U6-hdhfrFCU plasmid to obtain the HR-N_eYFP cassette, which was cloned into a modified version of the pDC2-Cas9-U6-hdhfr plasmid (Extended Data Fig. 2d and Supplementary Table 5). The donor plasmid (pfp2hs_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–Ncol 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_pfp2hs_eYFP) was cloned using the In-Fusion system into the SpeI–Ncol sites.

To tag PfAP2-HS with a 2xHA-adDFKBP domain tag (1.2B_pfp2-hs_ddDFKBP line), we used a single homologous recombination approach (Extended Data Fig. 2e). To generate the pfp2hs_HA-ddDFKBP plasmid, we replaced the pfp2-hs homology region in the plasmid pFAP2-G-dDFKBP52 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 Ncol and Xhol. 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 transfected cultures were synchronized by culturing rounds 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 Pruv restriction enzyme (cleaving the ampicillin resistance gene of the donor plasmid). Transfected cultures were selected with 60 µg of the donor plasmid (HA-ddFKBP tagging) or with continuous WR92210 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 (Extended Data Fig. 1), as previously described12.

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 parasitaemia of sorbitol-synchronized cultures was transferred to an incubator at 41.5 °C for 3 h and then returned to 37 or 35 °C (mature trophozoite stage). For heat shock, the full incubation chamber was 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 (Extended Data Fig. 1), as previously described12.

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 before12. Briefly, a mixture of random primers and oligo (dT) were used for reverse transcription, and for qPCR, we used a 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_0717000), which shows relatively stable expression throughout the IDC.

Transcriptional 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 microarrays12. RNA was obtained from tightly synchronized cultures of sorbitol-staged cultures 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 previously12. 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 microarrays. The Affymetrix custom microarray design was based on Agilent design AMADID no. 037237 (ref. 12), but we modified it as previously described (new design AMADID no. 08456110). 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 sampled at 0–3 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. Sequencing of hybridization reactions and the microarray path were as described12. 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-oxygen area and initial data processing was performed using the GE2_1105, OpenArray Feature Extraction software and Bioconductor 2.11.1.

Agilent microarray data were analysed using Bioconductor in an R environment (R version 3.5.3). For each microarray, we calculated the Cy5 and Cy3 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(Cy5/Cy3) values, statistical estimation of parasite age11 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 described12. The log2-transformed value of the fold change in expression following heat shock was calculated, for each gene and time point, as the log2((Cy5/Cy3) of the heat-shock-exposed sample minus the log2((Cy5/Cy3) of the control sample at the same parasite age, calculated using linear interpolation in the log2((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)10, we calculated the Pearson’s correlation between each sample and the time point of the reference transcriptome time course. By which it had a high level of similarity 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. 10). Expression trend plots for each cluster were generated using ggplot2, with LOESS smoothing, and Venn diagrams were generated using the euler 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 software12 with the topology-elim method11. Gene set enrichment analysis was performed using GSEA v3.0 Preranked12.

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 ultracavitator 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 of the PowerQ primer (cutadapt algorithm), the sequence reads were mapped to the Plasmodb DB falciumparum 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 76–98 fold. To identify SNPs and small insertion–deletion mutations (indels), we followed the Genome Analysis Toolkit (GATK)
best-practices workflow, using SAMtools, Picard tools 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-VariantFilteration, 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 identify SNPs

For the analysis of publicly available genome sequences, we used the PGK project (2016) pilot data release 5 (www.malarilgen.net/data/pgk-5) containing the sequence of >2,500 field isolates. Only SNPs that passed all quality filters and did not fail within a region with multiple large insertions and deletions were considered to be high-confidence. Using these criteria, a single high-confidence polymorphism (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 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 released 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 using the cell-free HA antibody (1:500, Roche, cat. no. 1660) 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. Briefly, after trimming, quality control, mapping the remaining reads to the P. falciparum genome (PlasmidDB release 28) using BWA-MEM and filtering duplicated reads, peak calling was performed using MACS2 (ref. 68) with a peak calling quality (Phred QUAL > 20) and low read depth (DP < 25%) difference in any of the two replicates. In all cases, samples were obtained from independent cultures.

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, 6.6 has been deposited in the Gene Expression Omnibus (GEO) database under the accession code GSE149394. The genome sequencing and ChIP-seq data presented in Figs. 1b, 2c 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 PGK pilot data release 5 (www.malarilgen.net/pgk-5) 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/Corset/MalariaLab/PfAP2-HS_Tinto_etal_NatMicrobiol_2021, with https://doi.org/10.5281/zenodo.477598).

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. Portugalaza, H. P. et al. Artemisinin exposure at the ring or trophozoite stage impacts Plasmodium falciparum sexual conversion differently. eLife 9, e06058 (2020).
14. Pavithra, S. R., Kumar, R. & Tatu, U. Systems analysis of chaperone networks in the malaria parasite Plasmodium falciparum. PLoS Comput. Biol. 3, 1701–1707 (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. Kudya, H. M. et al. The endoplasmic reticulum chaperone PGrP170 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., Jey, 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, 3991–4004 (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. Jeninger, M. D., Quinn, J. E. & Petter, M. ApiAP2 transcription factors in apicomplexan parasites. Pathogens 8, 47 (2019).

25. Miliotello, 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., Roodbol, 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, e50 (2003).

29. Dogovski, C. et al. Targeting the cell stress response of Plasmodium falciparum to overcome artemisinin resistance. PLoS Biol. 13, e002132 (2015).

30. Claessens, A., Affara, M., Assese, 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 the evolution of the AP2-integrase DNA binding domains. Nucleic Acids Res 33, 318–336 (2005).

35. Birnbaum, J. et al. A Kelch13-defined endocytosis pathway mediates artemisinin resistance in malaria parasites. Nat. Commun. 9, 2012–2025 (2018).

36. Ariey, F. et al. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. Nature 505, 50–55 (2014).

37. Llorà-Batlle, O. et al. Conditional expression of PfAP2-G for controlled activation of the Pf3k project (www.malariagen.net/pf3k). This work was supported by grants from the NIH/NIGMS (grant no. T32 GM125592-01). This research is part of ISGlobal's Science and Innovation through the 'Centro de Excelencia Severo Ochoa 2019–2023' co-funded by the European Social Fund (ESF). T.J.R. was supported by a training grant from the Spanish Ministerio de Economía y Competitiveness (grant nos BES-2014-067901 and BES-2017-081079, respectively), and L.M.-T. were supported by fellowships from the Spanish Ministry of Economy and Innovation (grant nos SAF2013-43601-R, SAF2016-76190-R and PID2019-107323-2000). 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 PID2019-107323-2000 to A.C.). This work was supported by grants 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. CEX2018-000800-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 bioinformatic 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.
© The Author(s), under exclusive licence to Springer Nature Limited 2021
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 log₂ 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. 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 10G (heat-shock-sensitive subclone), 1.2B (heat-shock-resistant subclone) and 3D7-A (right panel) is from Rovira-Graells et al. 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 10G. 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 (10G, 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 10G 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, 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\textsuperscript{14}. Columns at the left indicate presence of the G-box\textsuperscript{23} or tandem G-box (TdGbox) in the upstream region, and log$_2$ fold-change during heat shock in a previous study\textsuperscript{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).
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 log2-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_Δpfap2-hs and 10E. Values are the log₂ 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. 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 UPR29,38. The Log2 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.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

- Agilent Feature Extraction 11.5.1.1
- Step One software v2.3
- 7900 SDS v2.4
- GenePix Pro 6.0
- BD CellQuest Pro 6.0
- GeneSnap 7.12.06
- ImageQuant LAS 4000 v1.2

Data analysis

- R version 3.5.3 (Bioconductor and other packages)
- Gene Set Enrichment Analysis (GSEA) 3.0
- MEME 5.0.3
- FIMO v5.0.5 (MEME suite)
- Multiple Experiment Viewer (MEV) 4.9
- Ontologizer 2.1
- FastQC 0.11.4 and Galaxy version 0.63
- Cutadapt 1.8
- Bowtie2-2.1.0
- SAMtools v1.18 and Galaxy version 1.1.2
- Qualimap_v2.1.1
- Picard-tools-1.138
- GenomeAnalysis Toolkit 3.4-46
- Genome Browse (Golden Helix) 2.0.7
- ggplot2 3.3.3
- eulerr 6.1.0
Image J 1.50i
Trimmmomatic (Galaxy version 0.32.3)
Microsoft Excel 2010
Graph Pad Prism v5 and v7
Integrated Genome Viewer (IGV)
BWA-MEM (Galaxy version 0.4.1)
MACS2 v.2.2.6
DeepTools v.3.2.1
BedTools v.2.28.0
Clustal Omega 1.2.4

Custom codes deposited in Github (https://github.com/CortesMalariaLab/PfAP2-HS_Tinto_etal_NatMicrobiol_2021), with doi: 10.5281/zenodo.4775988

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The microarray data presented in Fig. 2 and Extended Data Fig. 1, 5, 6 and 8 has been deposited to the Gene Expression Omnibus (GEO) database with accession code GSE149394. Genome sequencing and ChIP-seq data presented in Fig. 1b, Fig. 2e and Extended Data Fig. 7 have been deposited to the NCBI Sequence Read Archive (SRA) database with accession codes PRJNA626524 and PRJNA670721, respectively. The authors declare that all other relevant data generated or analyzed during this study are included in the Article, the Extended Data file, or the Supplementary Information files. We used data from the Pf3k pilot data release 5 (www.malariagen.net/pf3k) and PlasmoDB (www.plasmodb.org, different releases) databases.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences  ☐ Behavioural & social sciences  ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

This manuscript does not include any population-based or epidemiological research. For our lab-based experiments, we selected the number of independent biological replicates based on our previous experience with equivalent experiments, feasibility and level of variability between replicates. For the majority of experiments, we used between three and five biological replicates. If the results were consistent between three replicate experiments, we considered this number of replicates to be sufficient, but we performed additional experiments if large variability was observed between replicates. For experiments involving screening of many subclones, selection experiments, or microarray-based time-course experiments, we used less than three biological replicates, and the conclusions were validated by other techniques. For some experiments involving redundant comparisons (i.e., comparison of a trait between several pairs of mutant and wild type lines), we also performed less than three replicates for each line (if the conclusions were already clear and consistent between equivalent lines).

Data exclusions

No data exclusions were applied, except for the microarrays analysis. In this case, as described in the Methods, we used visual inspection of the expression plots for the microarray data presented in Fig. 2 and Extended Data Fig. 8. One gene was excluded because it showed an aberrantly high signal in one particular sample (both in the sample channel and in the reference pool channel). Genes showing very low signal in all samples, or missing data for some samples, were also filtered according to our regular analysis pipeline (see Methods), because expression changes for such genes are of low confidence. The exclusion criteria were pre-established, as they are part of our regular analysis pipeline.

Replication

All attempts at replication were successful. The number of independent biological replicates for each experiment are provided in the Figure legends. In general, experiments were independently repeated at least three times, according to the criterion described above under "Sample size".

Randomization

Not relevant to this laboratory-based study. This study did not involve animals or human research. None of the experiments were susceptible to randomization.

Blinding

N/A. This study did not involve animals or human research. All experiments were laboratory-based. Blinding was not suitable for any of the experiments.
Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|

- **Antibodies**
- **Eukaryotic cell lines**
- **Palaeontology and archaeology**
- **Animals and other organisms**
- **Human research participants**
- **Clinical data**
- **Dual use research of concern**

### Methods

| n/a | Involved in the study |
|-----|-----------------------|

- **ChIP-seq**
- **Flow cytometry**
- **MRI-based neuroimaging**

### Antibodies

**Antibodies used**

- 3F10 rat anti-HA antibody (1:500; Roche no. 11867423001, lot 42155800); rabbit anti-ubiquitin (1:1000; Cell Signaling Technology no. 3933, lot 6); rabbit anti-phospho-eIF2α (1:1000; Cell Signaling Technology no. 3398, lot 9); rabbit anti-histone H3 (1:1000; Cell Signaling Technology no. 9715, lot 5); goat anti-rabbit IgG-peroxidase (1:5000; Millipore no. AP307P; batch number not available) secondary antibody.

**Validation**

The anti-HA primary antibody, recognizing an exogenous tag, has been extensively used in P. falciparum research, including use in ChIP-seq experiments (ref. PMID 32198457). In that study, the authors wrote the following statement: "In a Western blot (Supplementary Fig. 8d) with anti HA and anti GFP, no bands were seen in wildtype parasite lysate and only bands corresponding to proteins of the expected size were detected in AP2-G-DD::AP2-I-GFP parasite lysate."

The antibodies against phospho-eIF2α and ubiquitin have also been used before in P. falciparum research, i.e. ref. PMID: 30228310. While that article did not include a specific validation analysis, the results obtained using these antibodies were consistent with current knowledge in other organisms and with experiments using other reagents: the levels of ubiquitinated proteins were increased under proteotoxic stress conditions or with inhibitors of the proteasome, and the levels of phospho-eIF2α were elevated under conditions that induce the unfolded protein response (e.g. DHA, and DTT as a positive control).

The other primary antibody used in this study, anti-histone H3, was only used as a loading control and recognizes a band of the expected size. Even if it cross-reacted with other histones of a similar size, this would not affect any of our conclusions.

### Eukaryotic cell lines

**Policy information about** [cell lines](#)

**Cell line source(s)**

All Plasmodium falciparum lines used in this study are described in the first section of the Methods, with references to the articles in which they were described or details of how they were generated. Previously described lines were generated by us, except for the parental HB3B and D10 lines that were obtained from collaborators many years ago (source specified below and in the Methods).

**New lines generated in this study:**

- 10G_Δpfap2-hs
- 10E_Δpfap2-hs
- 10E_pfap2-hs_3xHA
- 10E_pfap2-hs_eYFP-Nterm
- 10E_pfap2-hs_eYFP-Cterm
- 1.2B_pfap2-hs_ddFKBP
- 1.2B_pfap2-hs_ddFKBP_8A
- 1.2B_pfap2-hs_ddFKBP_12E
- HB3_Δpfap2hs
- D10_Δpfap2hs
- 4E (subclone of Line 1 from The Gambia)
- 1H (subclone of Line 1 from The Gambia)

**Lines previously generated by the Cortés lab (or a stock maintained and characterised at this laboratory for many years):**

- 3D7-A (3D7 stock maintained and characterized at the Cortés lab for many years)
- 3D7-A r1
- 3D7-A r2
- 3D7-A-HS r1
- 3D7-A-HS r2
- 1.2B
- 10G
We did not directly authenticate all the parasite lines used, but correct edition of the transgenic parasite lines (all generated by us) was validated by analytical PCR. Some of the parasite lines are subclones of the same genetic background. However, their phenotypes and the mutations observed were consistently observed at multiple times. The genome of the 3D7-A r1, 3D7-A r2, 3D7-A-HS r1, 3D7-A-HS r2, 10G and 1.2B lines has been sequenced, which demonstrated that they are derived from 3D7. We analyzed the HB3B and D10 lines by comparative genome hybridization in a previous study (PMID: 22415456) and the results were consistent with the publicly available HB3 and D10 genome sequences.

We did not directly test all parasite lines used in this study for Mycoplasma, but many are subclones of a parasite stock (3D7-A) that tested negative by PCR. Mycoplasma contamination, even if it occurred, is not expected to affect any of the parameters analyzed in this study.

N/A. None of the cell lines used in this study (all P. falciparum) are listed in the database of commonly misidentified cell lines.

ChIP-seq

Confirm that both raw and final processed data have been deposited in a public database such as GEO.

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

https://dataview.ncbi.nlm.nih.gov/object/PRJNA670721?reviewer=9fu3kkauaf18d31m0rhbfddkjv
The called peaks in each individual replicate are also listed in Supplementary Table 4.

Control1_Input_R1.fastq
Control1_Input_R2.fastq
Control1_ip_R1.fastq
Control1_ip_R2.fastq
Control2_Input_R1.fastq
Control2_Input_R2.fastq
Control2_ip_R1.fastq
Control2_ip_R2.fastq
Control3_Input_R1.fastq
Control3_Input_R2.fastq
Control3_ip_R1.fastq
Control3_ip_R2.fastq
Control4_Input_R1.fastq
Control4_Input_R2.fastq
Control4_ip_R1.fastq
Control4_ip_R2.fastq
Control5_Input_R1.fastq
Control5_Input_R2.fastq
Control5_ip_R1.fastq
Control5_ip_R2.fastq
HS1_Input_R1.fastq
HS1_Input_R2.fastq
HS2_Input_R1.fastq
HS2_Input_R2.fastq
HS2_ip_R1.fastq
HS2_ip_R2.fastq
HS3_ip_R1.fastq
HS3_ip_R2.fastq
HS3_Input_R1.fastq
HS3_Input_R2.fastq
HS1_ip_R1.fastq
HS1_ip_R2.fastq
Control1_Input.bam
Control1_ip.bam
Control2_Input.bam
Control2_ip.bam
Control3_Input.bam
All the information from BED files of each experiment can be found in the Supplementary Table 3 (CTL1, CTL2, CTL3, CTL4, CTL5, HS1, HS2 and HS3 tabs).

Methodology

Replicates

Five independent biological replicates were preformed for the control condition (37°C), and three for the heat shock condition.

Sequencing depth

Control1-3 and HS1-3 were 125bp paired-end reads. Control4 and Control5 were 150bp single-end reads. SAMtools flagstat and GATK depth of coverage were used to obtain the stats information of each sample:

Control1-input: 22564099 QC-passed reads, 20282718 mapped reads (89.89%), 19611288 properly paired reads (86.99%), 98.26% mean coverage.
Control1-IP: 22941512 QC-passed reads, 11877122 mapped reads (51.77%), 11513274 properly paired reads (50.23%), 56.37% mean coverage.

Control2-input: 18698296 QC-passed reads, 17006893 mapped reads (90.95%), 16471186 properly paired reads (88.16%), 81.94% mean coverage.
Control2-IP: 16830580 QC-passed reads, 10269434 mapped reads (61.02%), 9935082 properly paired reads (59.1%), 49.73% mean coverage.

Control3-input: 19025160 QC-passed reads, 16381876 mapped reads (86.11%), 15841842 properly paired reads (83.34%), 79.3% mean coverage.
Control3-IP: 23817077 QC-passed reads, 13430413 mapped reads (56.39%), 13149850 properly paired reads (55.23%), 64.79% mean coverage.

Control4-input: 7545789 QC-passed reads, 6134160 mapped reads (81.29%), 35.01% mean coverage.
Control4-IP: 7941529 QC-passed reads, 4695094 mapped reads (59.12%), 27.68% mean coverage.

Control5- input: 7642248 QC-passed reads, 6010966 mapped reads (78.65%), 32.22% mean coverage.
Control5-IP: 7655816 QC-passed reads, 6047105 mapped reads (78.99%), 35.5% mean coverage.

HS1-input: 19579753 QC-passed reads, 17029300 mapped reads (86.97%), 16479338 properly paired reads (84.25%), 82.01% mean coverage.
HS1-IP: 17332846 QC-passed reads, 9716160 mapped reads (56.06%), 9408510 properly paired reads (54.34%), 46.92% mean coverage.

HS2-input: 18936095 QC-passed reads, 16967050 mapped reads (89.6%), 16402392 properly paired reads (86.69%), 81.14% mean coverage.
HS2-IP: 14845754 QC-passed reads, 7788892 mapped reads (52.47%), 7624136 properly paired reads (51.37%), 36.88% mean coverage.

HS3-input: 16137833 QC-passed reads, 13273821 mapped reads (82.25%), 12846862 properly paired reads (79.67%), 64.44% mean coverage.
HS3-IP: 15464452 QC-passed reads, 11975804 mapped reads (77.44%), 11717650 properly paired reads (75.8%), 59.8% mean coverage.

Antibodies

3F10 rat anti-HA antibody (1:500; Roche no. 11867423001, lot 42155800)

Peak calling parameters

Reads were mapped to the Plasmodium falciparum genome (PlasmoDB release 28) using BWA-MEM and filtered using SAMtools to remove duplicates. Peak calling was performed using MACS2 with a q-value cut-off of 0.01. All these steps were done using Galaxy utility.

Conversion to log2 coverage of immunoprecipitate/input was performed using DeepTools BamCompare, selecting the paired end parameter for all tools when analyzing experiments including control and HS conditions.

Overlapping intervals within called peaks for each dataset were determined using Bedtools MultiIntersect. The closest annotated gene coding sequence for each called peak was determined by Bedtools ClosestBed.

Data quality

Quality of reads was assessed by FastQC and adapter sequences were trimmed using Trimmomatic. After removing duplicates (SAMtools), peak calling was performed with a q-value cut-off of 0.01 and the following peaks were obtained: 73 (Control1), 404
From them, only a small subset of peaks presented a fold-enrichment above 5: 2 (Control1), 1 (Control2), 1 (Control3), 2 (Control4), 0 (Control5), 1 (HS1), 2 (HS2) and 1 (HS3).

Only peaks present in ≥3 out of 5 replicate ChIP-seq control experiments (37°C) or ≥2 out of 3 replicate heat shock experiments (HS) and with a MACS score >100 in each positive replicate were included in the final analysis.

Software

- FastQC
- BWA-MEM
- SAMtools
- MACS2
- DeepTools
- Bedtools

Flow Cytometry

Plots

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Plasmodium falciparum cultures were stained with Syto 11 (0.5 uM) to detect parasite DNA and quantify parasitaemia (% of erythrocytes infected) as described in the Methods.

Instrument

FACScalibur, Becton Dickinson.

Software

BD Cell Quest Pro 6.0

Cell population abundance

No sorting was performed.

Gating strategy

Erythrocytes were gated using standard FSC/SSC plots.

NOTE: There is no main figure or Extended Data figure displaying flow cytometry data in this article. We only used flow cytometry to measure parasitaemia (after staining parasite DNA). Supplementary Fig. 1 exemplifies the gating strategy.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.