Supporting Information

De Silva *et al*. 10.1073/pnas.0801993105

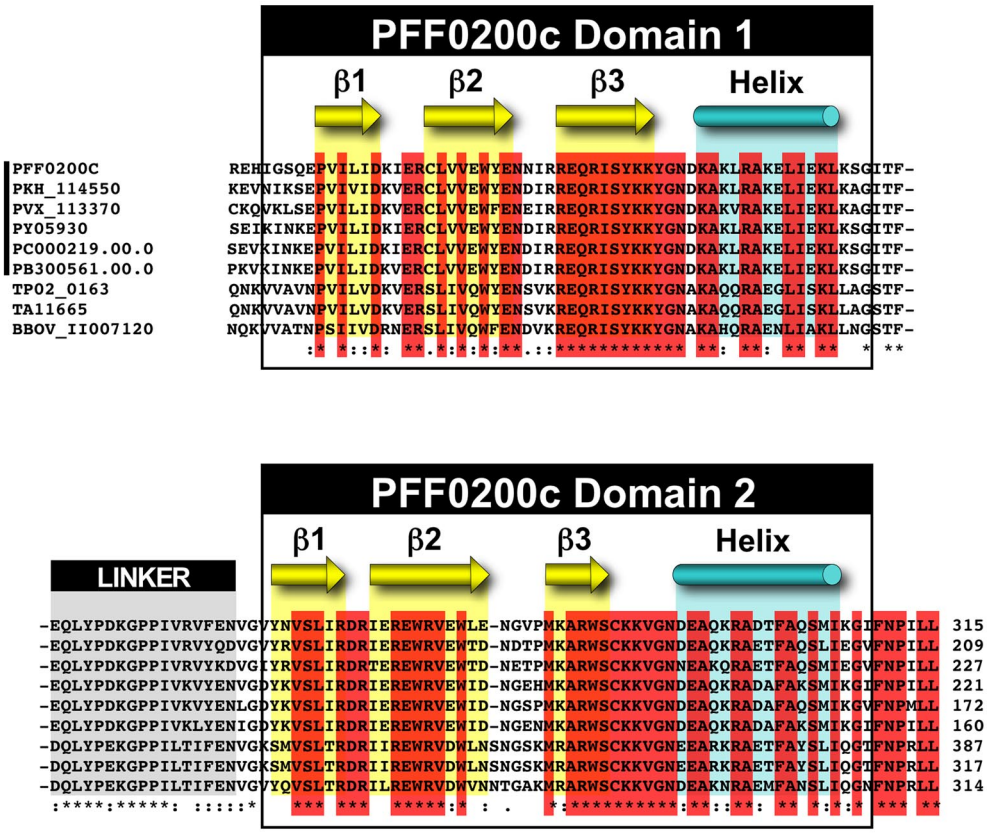

Fig. S1. Alignment of the AP2 domains (amino acids 177–311) from PFF0200c to orthologues in five additional *Plasmodium* spp. and three *Apicomplexan* species. The AP2 domain (boxed) is highly conserved across all species shown. Conservation of residues is most significant in the three β-strands (shaded yellow) of the AP2 domain, and is less significant in the α-helix (shaded blue). Absolutely conserved residues likely to be involved in DNA-binding are highlighted in red. Secondary structure predictions were predicted using Jnet [Cuff JA, Barton GJ (2000) Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins* 40:502–511]. There are no orthologues detected in the other sequenced Apicomplexan species *T. gondii* and the *Cryptosporidia*. PF = *Plasmodium falciparum*, PVX = *P. vivax*, PKH = *P. knowlesi*, PB = *P. berghei*, PY = *P. yoelii*, PC = *P. chabaudi*, TP = *Theileria parvum*, TA = *Theileria annulata*, BBOV = *Babesia bovis*.

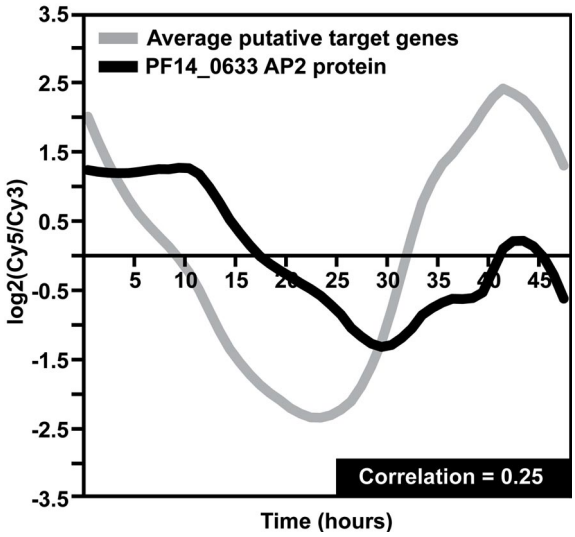

Fig. S2. Bloodstage gene expression profile of the PF14_0633 ApiAP2 protein, compared to the averaged expression profiles of putative target genes.

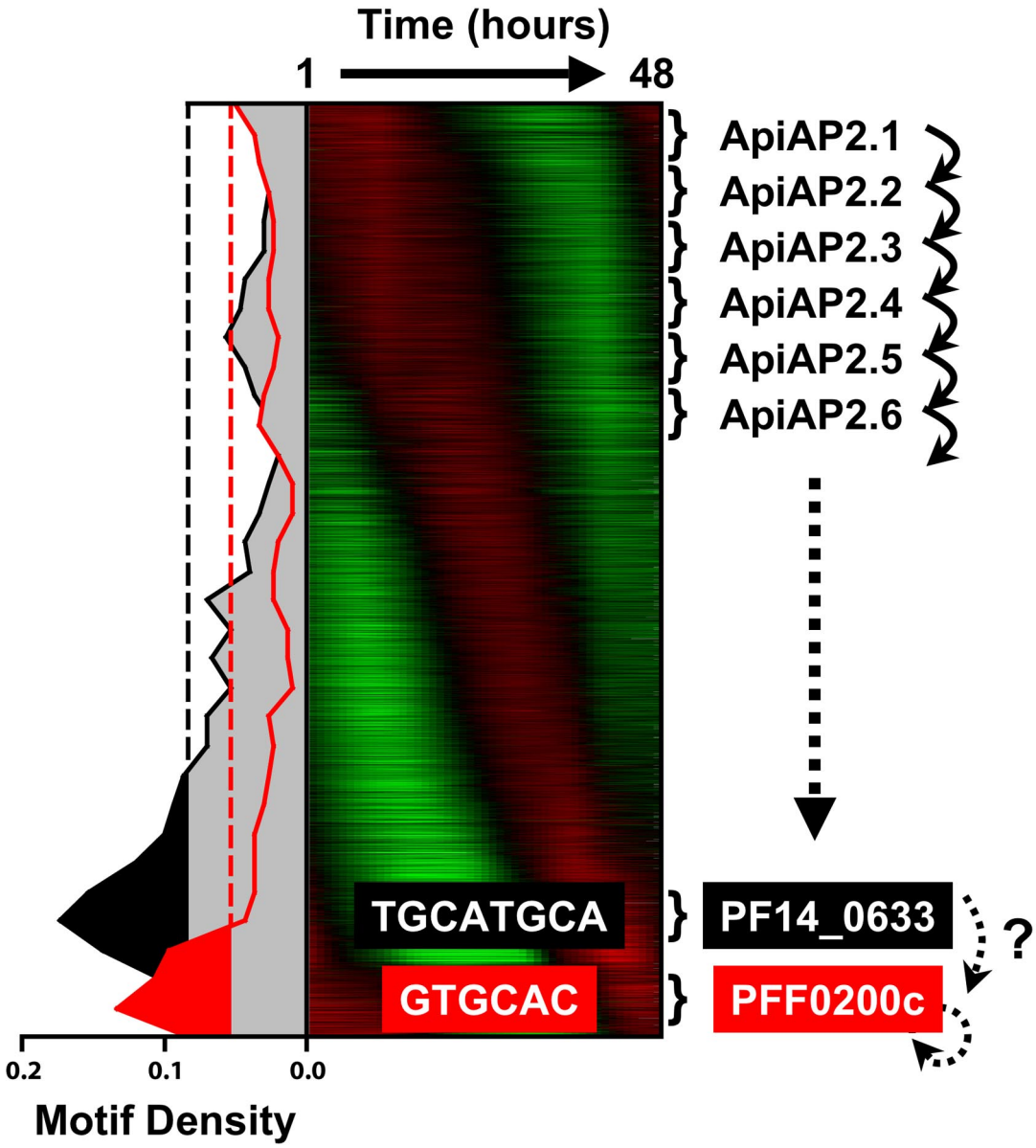

Fig. S3. Model of an ApiAP2 regulatory network throughout the asexual developmental stage of *Plasmodium* development. The 48-hour gene expression profiles are from a previous microarray study by Bozdech *et al.* [Bozdech Z, *et al.* (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum. PLoS Biol* 1:E5]. The predicted motif density is shown on the left for the two motifs identified in this study. Motif density was calculated as the fraction of genes in a window of 300 with an instance of the given motif. In black is the enrichment for the TGCATGCA motif recognized by PF14_0633, and in red the GTGCAC motif from PFF0200c. The dashed lines represent the significance cutoff (P value = 0.05) and were calculated using standard random permutation analysis repeated 10,000 times and calculating the densities at the 95th percentile. Values below the cutoff are shown in gray. On the right, we propose a model for a cascade of ApiAP2 regulators that includes PF14_0633 potentially regulating PFF0200c as well as autoregulation by PFF0200c.

Table S1. Oligonucleotide sequences used in gel-shift experiments with purified tagged AP2 domains

Putative regulator	Oligo name	Sequence (5′ to 3′)
PF14_0633	PFI0540w_site_fwd	ACTACACACCTAACTA**TGCATGCA**TCAATATAGTGGATTA
	PFI0540w_site_rev	TAATCCACTATATTGA**TGCATGCA**TAGTTAGGTGTGTAGT
	PFI0540w_mutated_site_fwd	ACTACACACCTAACTA**TGCGCGC**ATCAATATAGTGGATTA
	PFI0540w_mutated_site_rev	TAATCCACTATATTGA**TGCGCGC**ATAGTTAGGTGTGTAGT
PFF0200c	MAL7P1.119_site_fwd	TAAGTGT**GTGCA**CATGTAACCACAAATCCA**GTGCA**CAACAG
	MAL7P1.119_site_rev	CTGTT**GTGCA**CTGGATTTGTGGTTACAT**GTGCA**CACACTTA
	MAL7P1.119_mutated_sites_fwd	TAAGTGT**GTATA**CATGTAACCACAAATCCA**GTATAC**AACAG
	MAL7P1.119_mutated_sites_rev	CTGTT**GTATAC**TGGATTTGTGGTTACAT**GTATACAC**ACTTA
	Random_oligo_fwd	TTAATTTATTACATATAATATTATATTATATATATATATA
	Random_oligo_rev	TATATATATATATAATATAATATTATATGTAATAAATTAA

Oligonucleotides correspond to sequences upstream of putative target genes PFI0540w and MAL7P1.119. Target motifs are in bold; mutated residues are bold and underlined.

Table S2. Nuclear localization signals predicted by PredictNLS for ApiAP2 proteins

Gene IDs	Predicted NLS amino acid sequence
MAL8P1.153	*KKKKKRKKKK*
PF11_0404	*KKKINTSEAYHEGKMNKKHKKKK*
PF13_0026	*RKKKRKEERIVNTGSAKRLEFFYPKKK*
PF13_0235	*KKKIMNNHHNNNKKNNKKKKK*
PF14_0079	*KKKKNQSNYYQQHNICDDHKSLYDDVKKK*
PF14_0271	*KKSRSKN*
PF14_0471	*KKRRNK*
PF14_0533	*KRKKPNHSKTNDNDQTEIYKKTKKMNNK*
PFD0200c	*KKKKKKKKKK*
PFD0985w	*KRRKQVDVVGDGSGTQALKRSKRSKSNSKYK*
PFF0200c	*KHKKKNIHDNNRKK*
PFF0670w	*KRKRTH*
PFF1100c	*KKRKSNEEEEEEEKKKKNNMDLPCCDKDNYIINKKLKKNHK*
PFL1075w	*KKKRNNIDEVRKNDNQEKKKKKRK*
PFL1085w	*KRKKYFDDCNMKDEHKKEGKNKQKNKKNNKNKKNNK*
PFL1900w	*KKKKNEDNEYINYKEQNNDERKKKMKDNNKGNKVDKKKK*

See Cokol M, Nair R, Rost B (2000) Finding nuclear localization signals. *EMBO Rep* 1: 411–415.

Other Supporting Information Files

Dataset S1 (XLS)

Dataset S2 (XLS)