Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/24/eaaz5057/DC1

Supplementary Materials for

Conditional expression of PfAP2-G for controlled massive sexual conversion in *Plasmodium falciparum*

Oriol Llorà-Batlle, Lucas Michel-Todó, Kathrin Witmer, Haruka Toda, Carmen Fernández-Becerra, Jake Baum, Alfred Cortés*

*Corresponding author. Email: alfred.cortes@isglobal.org

Published 10 June 2020, *Sci. Adv.* **6**, eaaz5057 (2020)
DOI: 10.1126/sciadv.aaz5057

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/6/24/eaaz5057/DC1)

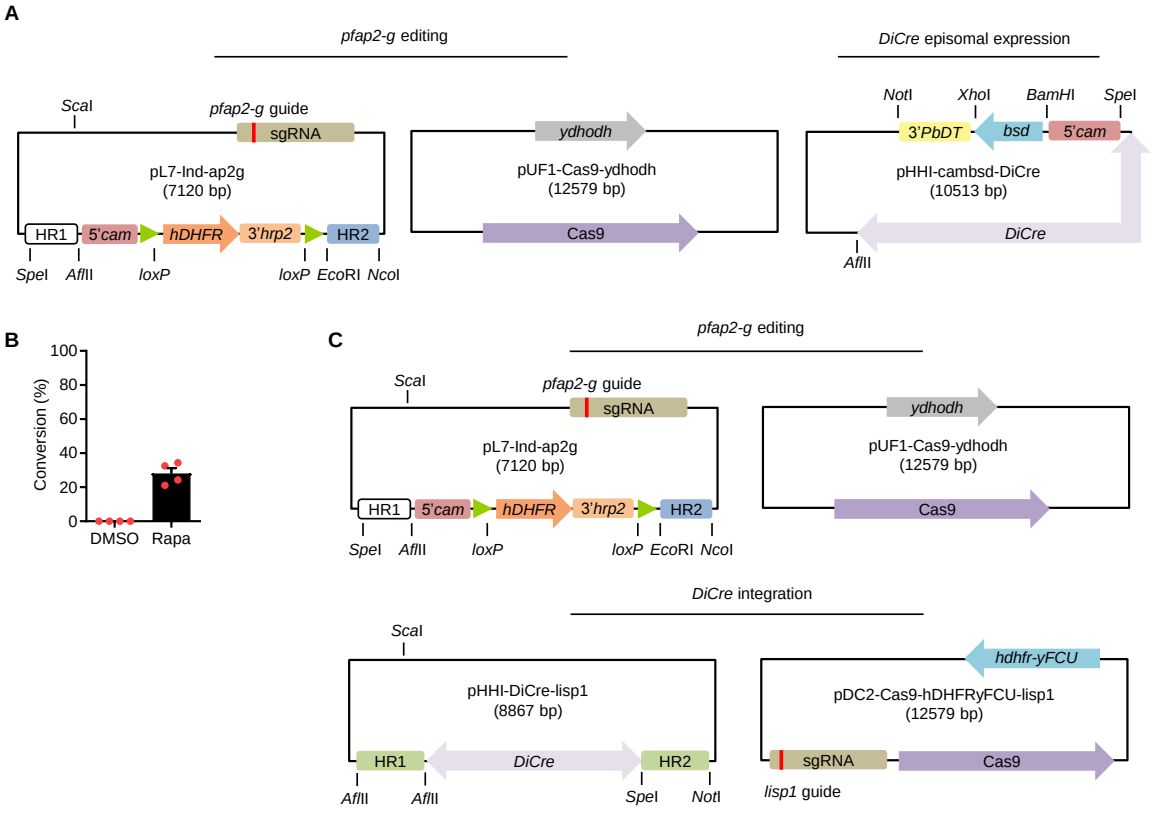

Fig. S1. Plasmids used to generate the transgenic lines and characterization of the line expressing DiCre episomally. (A) Schematic (not to scale) of the plasmids used for the *pfap2-g* conditional activation system with episomally-expressed DiCre. Restriction sites used for the generation of the plasmids are shown. **(B)** Sexual conversion levels in the transgenic line with the *pfap2-g* conditional activation system and episomally-expressed DiCre. Data are presented as the average and s.e.m. of four independent experiments. **(C)** Schematic (not to scale) of the plasmids used for the conditional activation system with the DiCre expression cassette integrated in the genome (at the *lisp1* locus). Plasmid pUF1-Cas9-ydhodh was not used to generate the E5ind line, for which the other three plasmids were simultaneously co-transfected (see Materials and Methods).

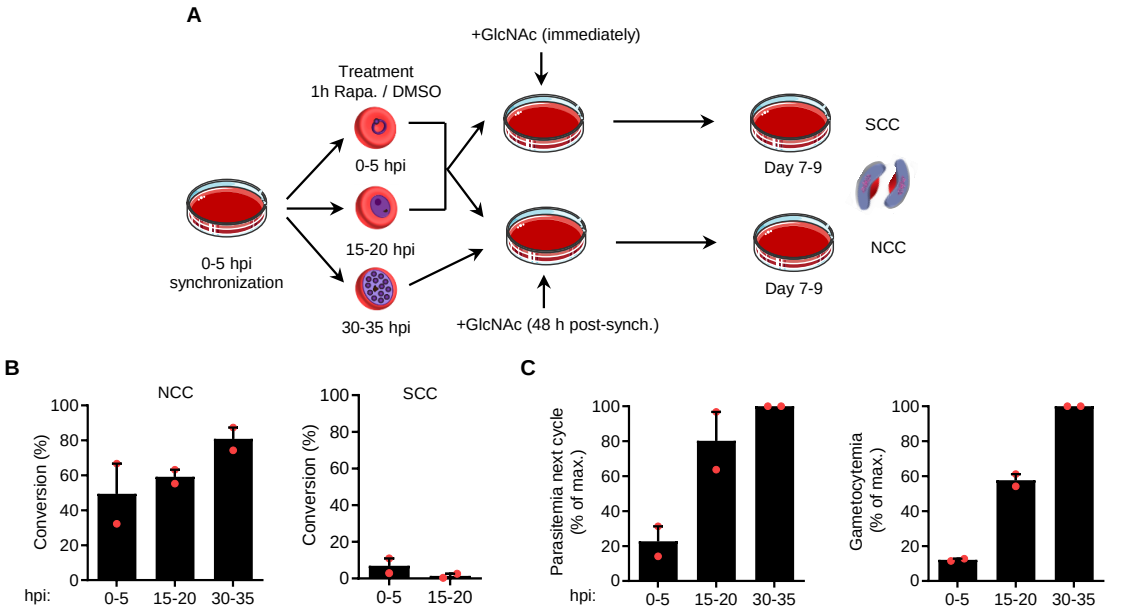

Fig. S2. Optimization of the protocol to induce sexual conversion in the E5ind line. (A) Schematic of the procedure used to determine the optimal time to induce conversion. E5ind cultures synchronized to a 5 h age window were treated with rapamycin at different times (in h post-invasion, hpi). GlcNac was added immediately after rapamycin treatment to measure only sexual conversion within the same cycle of *pfap2-g* activation (SCC route), or after reinvasion to measure conversion at the next cycle (NCC route). **(B)** Sexual conversion rate by the NCC and SCC routes after treating with rapamycin at different times. **(C)** Left, parasitemia at the cycle after rapamycin treatment, in cultures treated at different times, relative to the maximal parasitemia (obtained when treating at 30-35 hpi). Right, gametocytemia under NCC conditions in cultures induced with rapamycin at different times, relative to the maximal gametocytemia (obtained when inducing at 30-35 hpi). Induction at early stages results in reduced growth, which together with the conversion rate explains the lower number of gametocytes obtained. Data in panels **B** and **C** are presented as the average and s.e.m. of two independent experiments.

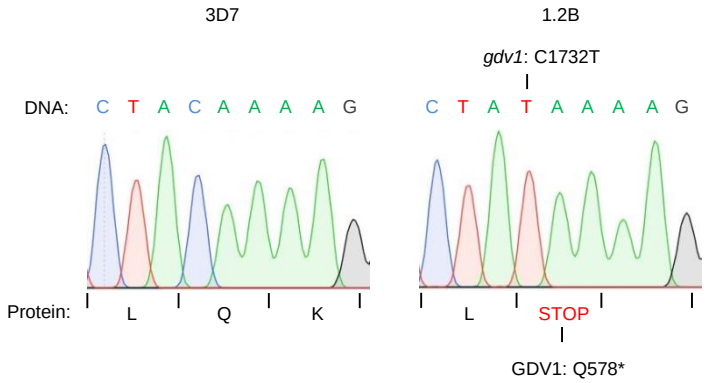

Fig. S3. The 1.2B parasite line has a premature stop codon in *gdv1*. Sanger sequencing validation of the mutation in *gdv1* showing the presence of a C to T mutation at position 1,732 of the coding sequence, resulting in a premature STOP codon (Q578*).

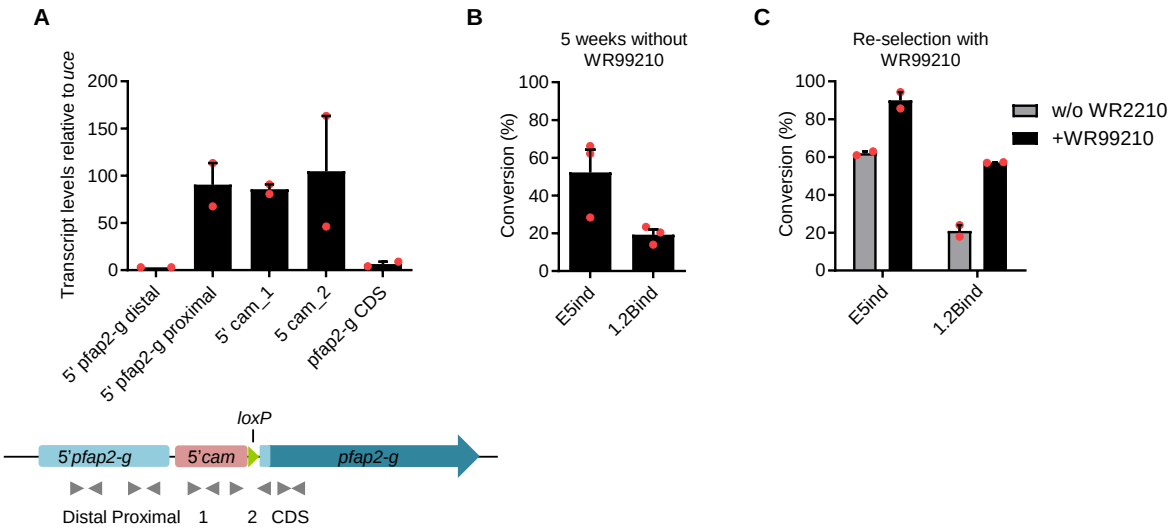

Fig. S4. Transcriptional and epigenetic events at the *pfap2-g* locus in the inducible lines. (A) Reverse transcriptase-quantitative PCR (RT-qPCR) analysis of transcript levels containing different elements of the *pfap2-g* locus in induced schizonts of the E5ind line (40-45 h post-invasion). The primers used for the analysis (arrowheads in the scheme, not to scale) were located at positions -1412 to -1302 (5' *pfap2-g* distal) and -449 to -351 (5' *pfap2-g* proximal) of the endogenous *pfap2-g* upstream region (relative to the position of the *pfap2-g* start codon in wild type parasites), -699 to -581 (5'cam_1) and –201 to -68 (5'cam_2, using a reverse primer located in the region between the *5'cam* and the *pfap2-g* coding sequence) of the 5'*cam* promoter (relative to the position of the *pfap2-g* start codon in the inducible line after recombination), and +259 to +382 of the *pfap2-g* coding sequence (*pfap2-g* CDS). Transcript levels at the coding sequence were similar to levels measured with primers located further downstream within the coding sequence (Fig. 2). Transcript levels are normalized against ubiquitin-conjugating enzyme (*uce*). Data are presented as the average and s.e.m. of two independent experiments. **(B)** Sexual conversion rates in rapamycin-treated E5ind and 1.2Bind cultures maintained for 5 weeks in the absence of WR99210 before induction. Data are presented as the average and s.e.m. of three independent experiments (each experiment corresponds to a separate culture maintained in the absence of WR99210 before induction). **(C)** Sexual conversion rates upon rapamycin induction of E5ind and 1.2Bind cultures reselected with WR99210 for 6-

8 days after 5 weeks without the drug (+WR99210). Conversion rates upon induction were also measured in cultures maintained in parallel without WR99210 (w/o WR99210) during the 6-8 days. Data are presented as the average and s.e.m. of two independent experiments.

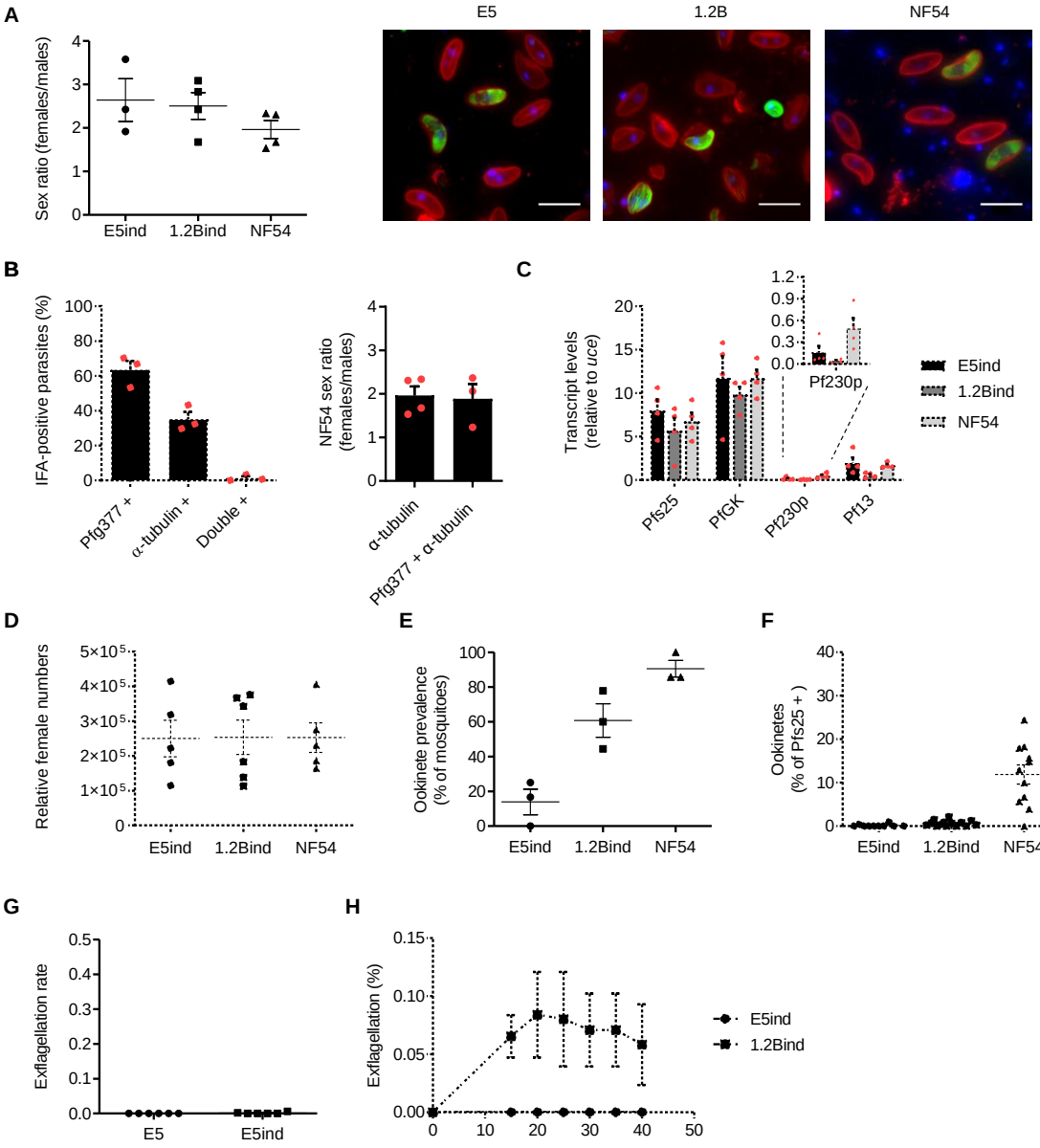

Fig. S5. Characterization of E5ind and 1.2Bind mature gametocytes. (A) Sex ratio determined by IFA of stage V gametocytes using antibodies against α-tubulin (green, male-specific at stage V) and glycophorin A (red, stains the red cell membrane of all gametocytes). DAPI (blue) stains nuclei. Male and female

gametocytes were identified as α-tubulin-positive and -negative stage V gametocytes, respectively (immature gametocytes were excluded). Data are presented as the average and s.e.m. of three (E5ind) or four (1.2Bind and NF54) independent experiments. The total number of parasites scored was >2,000 for each parasite line. Representative IFA images are shown. Scale bar, 10 μm. **(B)** Validation of α-tubulin as a male-specific marker in stage V gametocytes. Left, quantification of IFA co-staining with anti-α-tubulin and anti-Pfg377 (a female-specific marker) antibodies in the NF54 line. Double-positive parasites were virtually not seen. Right, sex ratios determined by only using α-tubulin or both α-tubulin and Pfg377 antibodies. Data are presented as the average and s.e.m. of three or four independent experiments. The total number of parasites scored was >500. **(C)** Transcriptional analysis of female (*pfs25*, *pfGK*) and male (*pf230p*, *pf13*) specific markers (*52*) in day 12-14 gametocyte cultures. Transcript levels are normalized against ubiquitin-conjugating enzyme (*uce*). Data are presented as the average and s.e.m. of four independent experiments. **(D)** Number of Pfs25-positive females/ml (determined using an anti-Pfs25 antibody coupled with Cy3) normalized by gametocytemia, in day 12-14 gametocyte cultures. Data are presented as the average and s.e.m. of five (E5ind and NF54) or six (1.2Bind) independent experiments. **(E)** Ookinete prevalence (percentage of mosquito midguts with at least one ookinete) 24h after feeding. Data are presented as the average and s.e.m. of three independent experiments. The number of mosquitoes analyzed between the three experiments was: E5ind, 20 midguts; 1.2Bind, 23 midguts; NF54: 18 midguts. **(F)** Ookinete intensity (percentage of ookinetes relative to the total Pfs25-positive cells present in each mosquito midgut). Data are presented as the average and s.e.m of 10 midguts (E5), 14 midguts (1.2B), or 11 midguts (NF54), derived from two independent feeding experiments. **(G)** Exflagellation rate (percentage of exflagellation normalized by gametocytemia, see Materials and Methods) of E5ind and parental E5 gametocyte cultures at day 12-14. Data are presented as the average and s.e.m. of six independent experiments. **(H)** Time-course percentage of exflagellation analysis. To assess the possibility that in E5ind and 1.2Bind DNA replication and exflagellation proceed more slowly rather than

failing to occur, exflagellation centers were first scored 15 min after activation and then every 5 min. The decreasing trend in the percentage of exflagellation after 20 min excludes this possibility. Data are presented as the average and s.e.m. of two independent experiments.

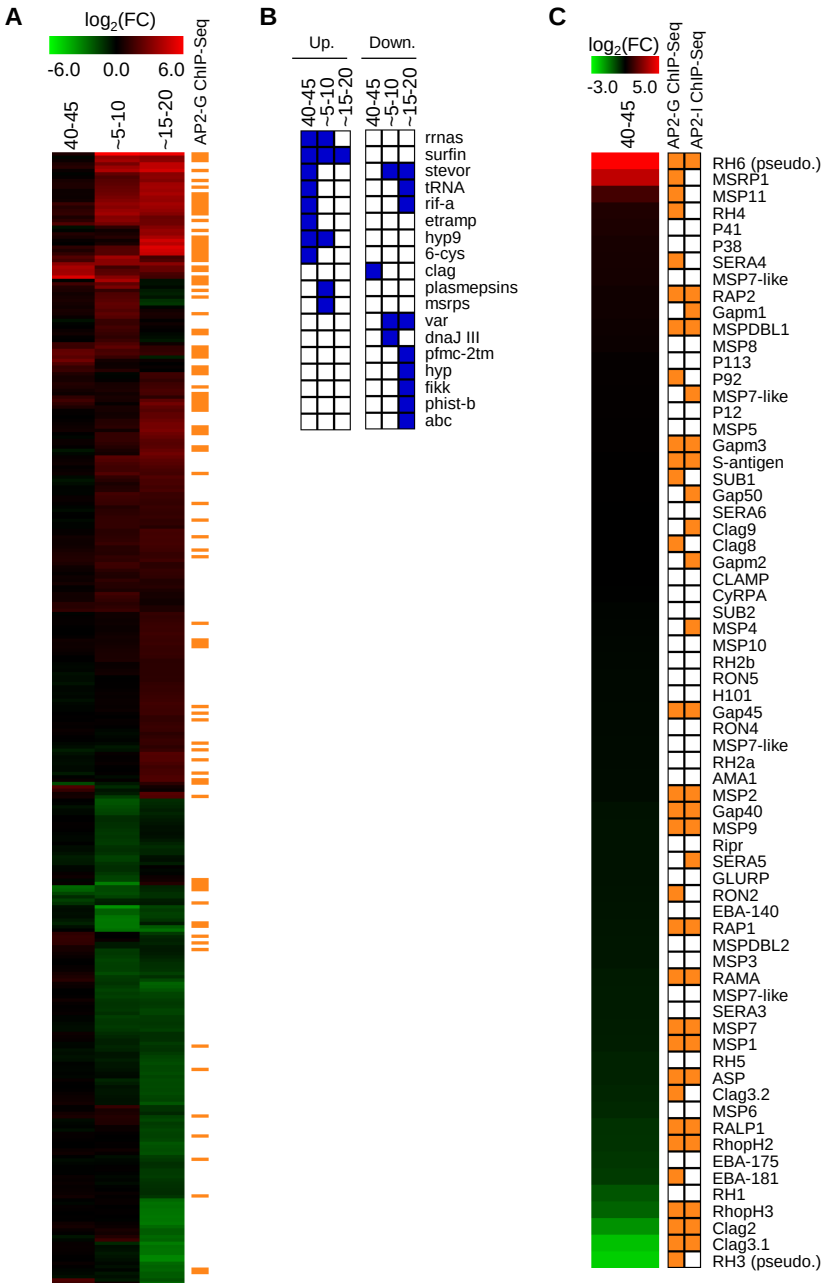

Fig. S6. Transcriptomic alterations upon *pfap2-g* activation. (A) Hierarchical clustering of genes differentially expressed in a time-course analysis of DMSO-treated (uninduced control) and rapamycin-treated (induced) E5ind cultures. Values are the average of the $\log_2$ of the expression fold-change (FC) (induced/non-induced) in the two replicates. Genes with a $\log_2$(FC)>1 in the two independent experiments at any of the time-points analyzed are shown. Samples were collected at 40-45 h post-invasion (hpi) of the cycle of induction or ~5-10 and ~15-20 hpi of the next cycle. Genes bound by PfAP2-G according to ChIP-Seq experiments are indicated in orange (*29*). **(B)** Gene set enrichment analysis (GSEA) of gene families and other gene sets. Only gene sets significantly upregulated (Up.) or downregulated (Down.) in induced vs non-induced cultures are shown (FDR<0.05). **(C)** Expression FC (ordered by FC intensity) of known genes involved in erythrocyte invasion. Genes bound by PfAP2-G and PfAP2-I according to ChIP-Seq data are shown in orange (*29, 47*).

Table S1. Oligonucleotide sequences. List of oligonucleotides used in this study, classified according to their use: cloning procedures, diagnostic PCR or quantitative PCR (qPCR). In lowercase: non-annealing part of the oligonucleotide; in bold: restriction sites included in the oligonucleotide; underlined: loxP sequence; dashed underlined: guide sequences.

Primers used for cloning procedures		
Number	**Name**	**Sequence (5'-3')**
p1	LoxP_site_BamHI_ends_F	**gatcc**ATAACTTCGTATAATGTATGCTATACGAAGTTAT**g**
p2	LoxP_site_BamHI_ends_R	**gatcc**ATAACTTCGTATAGCATACATTATACGAAGTTAT**g**
p3	HR2_ap2g_-73_LoxP_EcoRI_F	tggaga**gaattc**ATAACTTCGTATAATGTATGCTATACGAAGTTATAACAGTTTTATATCGGACTAAC
p4	HR2_ap2g_+311_NcoI_R	tggtgtt**ccatgg**GTTGATAATCGTATCTTCGAG
p5	HR1_ap2g_-449_SpeI_F	tggaga**actagt**ATATGTCCTATAGGTGTCAAAC
p6	HR1_ap2g_-160_AflII_R	tggtgtt**cttaag**GGAGATATTTGAATGTACCTAC
p7	ap2-g_-158_guide_F	TTCTAGCTCTAAAACTTATATTGGCACTAATTTAGAATATTATATACTTA
p8	ap2-g_-139_guide_R	TAAGTATATAATATTCTAAATTAGTGCCAATATAAGTTTTAGAGCTAGAA
p9	pHLHbsdR_bsd_+1_BamHI_F	tgttggt**ggatcc**tATGGCACCTTTGTCTCAAGAAG
p10	pHLHbsdR_PbDT3'_+22_R	CGAACATTAAGCTGCCATATCC
p11	pHH1eba140_7305_AflII_SpeI_F	tggc**aagctt**g**actagt**tgttc**ttaag**GCGATCCATATAATTATTAATAGGT
p12	pHH1eba140_7387_R	GACGGCCAGTGAATTGTAATA
p13	LISP1_HR2_+5891_F_SpeI	tgttgg**actagt**TACCTATAGAGGATAAGGAGAA
p14	LISP1_HR2_+6235_R_NotI	tggttg**gcggccgc**GTAAGTGTTGTGGGTATGCTT
p15	LISP1_HR1_+5169_F_Infusion	tatggatcgccttaaAAGACAATGGGAAATGGTGTTA
p16	LISP1_HR1_+5523_R_Infusion	tttttattaagttccTTCTGGGACGACATTTATTGTT
p17	LISP1_Guide_+5526_F	TAAGTATATAATATTGAGGAACTGGGAACATGTAGGTTTTAGAGCTAGAA
p18	LISP1_Guide_+5545_R	TTCTAGCTCTAAAACCTACATGTTCCCAGTTCCTCAATATTATATACTTA

Primers for diagnostic PCR		
Number	**Name**	**Sequence (5'-3')**
p19	LISP1_+5088_F	TATGAAGAATATATTGAACGAATC
p20	LISP1_+6006_R	AGTATACCCAGGAGTGGATAA
p21	Pfap2-g_5'UTR_-460_F	GCTTCTTTAATGTTGTATGTATG
p22	Pfap2-g_CDS_+345_R	GATACATTCTCGTTACTCTGC

Primers for qPCR		
Number	**Name**	**Sequence (5'-3')**
p23	PF3D7_1222600_F1 (pfap2-g)	AACAACGTTCATTCAATAAATAAGG
p24	PF3D7_1222600_R1 (pfap2-g)	ATGTTAATGTTCCCAAACAACCG
p25	hdhfr_qPCR_F	AGTAGAAGGTAAACAGAATCTG
p26	hdhfr_qPCR_R	GGCATCATCTAGACTTCTGG
p27	Calmodulin_qPCR_F	GATGGAACTATAACAACTAAGG
p28	Calmodulin_qPCR_R	CATTAAGGTTAGAAATTCGGGA
p29	Seryl_qPCR_F (serrs)	AAGTAGCAGGTCATCGTGGTT
p30	Seryl_qPCR_R (serrs)	TTCGGCACATTCTTCCATAA
p31	PfUCE_F	GGTGTTAGTGGCTCACCAATAGGA
p32	PfUCE_R	GTACCACCTTCCCATGGAGTA
p33	5'pfap2-g_distal_F	TTAATAATACGTATGCTTGTGGA
p34	5'pfap2-g_distal_R	AAGAAAGAACTACAACATTCTG
p35	5'pfap2-g_proximal_F	ATATGTCCTATAGGTGTCAAAC
p36	5'pfap2-g_proximal_R	ACAAATACTTATTATATAAATTGTAC
p37	5'cam_1_F	ATTTTTAAATGCTTACTTAATTATCT
p38	5'cam_1_R	TACATTGTTTAATACTACTACATGT
p39	5'cam_2_F	TTTTAAACTAGAAAAGGAATAACTA
p40	5'pfap2-g_-45_R	TATGTTAGTCCGATATAAAACTG
p41	Pfap2-g_CDS_+259_F	GAAGAGAGCATGCAATGAAGT
p42	Pfap2-g_CDS_+382_R	TTGTCCATGCAACTATTCGATA
p43	Pfs25_F1	AATGCGAAAGTTACCGTGGA
p44	Pfs25_R1	TACATTCCAAATGACCACTC
p45	PfGK_F	GTCAGAACATGTTCCGGACT
p46	PfGK_R	TAGCTTGACCAATGCATGCA
p47	Pf230p_F1	CCCAACTAATCGAAGGGATGAA
p48	Pf230p_R2	TGTTGTTCGATTCCAGTTGGT
p49	Pf13_F	ACGAATATGCTCGAGAACGA
p50	Pf13_R	GCCTTTTCATCTGACACGT
p51	RAP1_+479_F	TCAGCTAGTCCACATGGTGAA
p52	RAP1_+619_R	TTTTTGGTGCAGGAGGTGCTT
p53	RAMA_+1735_F	GACCTTCTAGATGTCCAACAA
p54	RAMA_+1855_R	ATGCATGTACGTCGCTAACAT
p55	RhopH2_2F	TGTTGCTGTCCATATTTAGTTTT
p56	RhopH2_2R	AATATATCGCTACATAACTTCGT
p57	RhopH3 _+1289_F	TATCTGTTCAATGCCCAACTTA
p58	RhopH3_+1430_R	CTGCAGAAGGGTGTTTACTTT
p59	clag3.1_clag3.2_6F	TAGTAATGAGAATTAGTTGGACA
p60	clag3.1_6R	ATAAATATTTGGATGCTTCAGCA

Data file S1. Processed microarray data.

Data file S2. Data from GSEA analysis.

Data file S3. Description of AMADID-084561 and 085763 microarray designs.