

# Supporting Information

Avraham et al. 10.1073/pnas.1214572109

## SI Materials and Methods

**Cell Cultures.** All parasites used were derivatives of the NF54 parasite line and were cultivated at 5% hematocrit in RPMI medium 1640, 0.5% Albumax II (Invitrogen), 0.25% sodium bicarbonate, and 0.1 mg/mL gentamicin. Parasites were incubated at 37 °C in an atmosphere of 5% (vol/vol) oxygen, 5% (vol/vol) carbon dioxide, and 90% (vol/vol) nitrogen. Parasite cultures were synchronized using Percoll/sorbitol gradient centrifugation as previously described (1, 2). Briefly, infected RBCs were layered on a step gradient of 40/70% (wt/vol) Percoll containing 6% (wt/vol) sorbitol. The gradients then were centrifuged at  $12,000 \times g$  for 20 min at room temperature. Highly synchronized late-stage parasites were recovered from the 40/70% (wt/vol) interphase, washed twice with complete culture medium, and placed back in culture. The level of parasitemia was calculated by counting three independent blood smears stained with Giemsa using light microscope.

**Plasmid Construction.** The 3' UTR of *Plasmodium berghiei* dhfr-thymidylate synthase (PbDT 3') was amplified from pHTK (3) using 5'-CCCAAGCTTGGATATGGCAGCTTAATG-3' and 5'-CGCGGATCCCTACCCTGAAGAAGAAAA-3' and was cloned into pVLhIdh (4) as the 3' UTR of either *luc* or the *dhfr* using HindIII/BamHI to create pVLhIdh and pVLhIdb, respectively. To make pV<sub>Δ102</sub>LhIdb, *luc*-hrp2 3' UTR was amplified from pVLhIdb using 5'-ATGTTAACATGCTGCATGAAGACG-3' and 5'-AACTGCAGTTAATAAATATGTTCTTATAT-3' and was cloned into pVLhIdb that was cut with HpaI/PstI (the HpaI restriction site is at -102 to ATG of *luc*). pVLhPEIdb was created by cutting pVLhIdb with XmaI/PstI and cloning a 33-bp pairing element (PE) that was created by hybridization of commercially synthesized oligos with 5' phosphate that enables direct ligation (forward: 5'-GCTGTGTATATGTATGTGTGCATTC-3'; reverse: 5'-CCGGGAATGCACACATACATATACACAGCTGCA-3'). pHLhIdb was made by cutting the *hrp2* promoter from pHLh (5) and cloning it to pV<sub>Δ102</sub>LhIdb using KpnI/HpaI. Similarly, pHPEIdb was made by inserting the same *hrp2* promoter to pVLhIdb cut with KpnI/HpaI. This restriction leaves the last 102 bp of the *var* promoter intact.

**Parasite Transfections and Selections.** Parasites were transfected as previously described (6). Briefly, 0.2-cm electroporation cuvettes were loaded with 0.175 mL of erythrocytes and ~100 μg of plasmid DNA in incomplete Cytomix solution. Stable transfectants were selected initially on 4 nM WR99210, and selection for specific *var* activation was done on 2 μg/mL of blasticidin S (Invitrogen).

**Genomic DNA Extraction, RNA Extraction, and cDNA Synthesis.** Genomic DNA was extracted as described (7). Briefly, infected RBCs were pelleted and lysed with saponin. RBC-free parasites from 20-mL cultures then were pelleted, washed with PBS, and taken up in 200 μL TSE buffer (100 mM NaCl, 50 mM EDTA, 20 mM Tris, pH 8), 40 μL of 10% (wt/vol) SDS, and 20 μL 6M NaClO<sub>4</sub>. This suspension was rocked for 12 h, and genomic DNA was extracted with phenol/chloroform. The resulting DNA was taken up in 100 μL dH<sub>2</sub>O.

RNA extraction and cDNA synthesis were performed as described (8). Briefly, RNA was extracted from synchronized parasite cultures 36 h after invasion. RNA was extracted with the TRIZOL LS Reagent as described (9) and was purified on a PureLink column (Invitrogen) according to the manufacturer's protocol. Isolated RNA then was treated with recombinant DNase I (Takara) to degrade contaminating genomic DNA (gDNA).

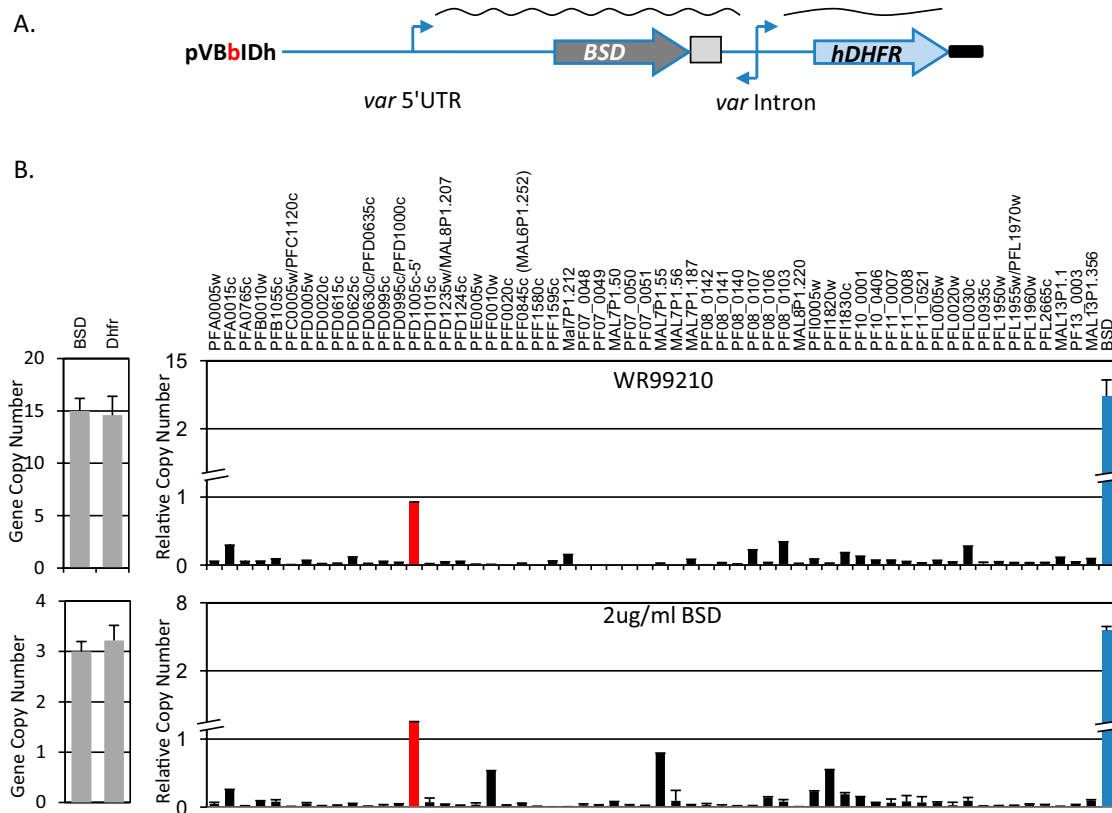
cDNA synthesis was performed from 750 ng total RNA with Primescript RT reagent (Takara) with oligo dT (Takara) as described by the manufacturer.

**Quantitative RT-PCR.** Transcript copy numbers were determined using the equation  $2^{-\Delta\Delta CT}$  as described in Applied Biosystems *User Bulletin 2* using NF54 gDNA as the calibrator. Specifically, relative copy number was calculated as a two-exponential negative [(Ct target gene in cDNA - Ct reference gene in cDNA) - (Ct target gene in gDNA - Ct target gene in gDNA)]. Relative copy numbers were calculated by comparison with the expression of the housekeeping gene arginyl-tRNA synthetase (PFL0900c). This gene and P60-seryl-tRNA synthetase (PF07\_0073), P61-fructose biphosphate aldolase (PF14\_0425), P100-actin (PFL2215), and glutamyl-tRNA synthetase (PF13\_0170) were used as control genes in all quantitative RT-PCR (qRT-PCR) assays. All qRT-PCR assays were performed at least in duplicate for each template with no apparent differences, and each experiment was completed three times in its entirety, again with no significant differences. The housekeeping and other control primers used for qRT-PCR were published previously (8).

**Luciferase Assays.** Luciferase activity was measured from 200 μL of culture containing tightly synchronized ring-stage parasites. Infected RBCs were pelleted by centrifugation and lysed in 50 μL Glo Lysis Buffer (Promega). Luciferase activity was measured immediately after the addition of 100 μL Bright-Glo luciferase reagent (Promega) in a FLUOROSKAN FL luminometer (Thermo Scientific). The luciferase activity of each clonal cell line was determined in at least three independent experiments and was normalized to 1% parasitemia and one copy number. The D10 parasite line (4) constitutively expressing *luciferase* from the endogenous *var* promoter was used as a positive control, and the DC-J parasite line (8) was used as a negative control.

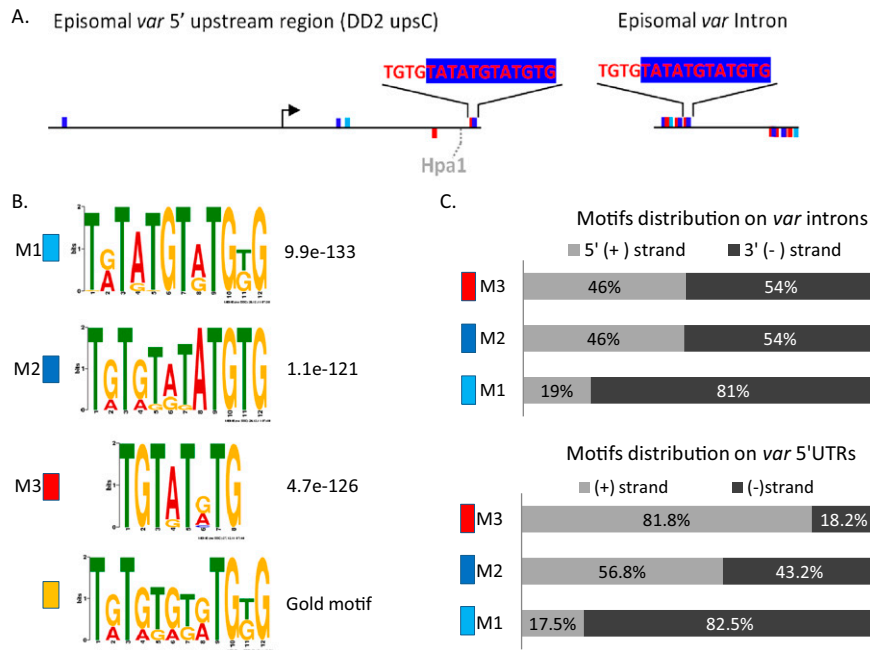
**EMSA.** To prepare nuclear and cytoplasmic protein extracts, parasites were released from RBCs using saponin lysis and then were fractionated by high-salt buffer as described (10). Probe labeling was done by end labeling of single-stranded oligonucleotides with [ $\gamma$ -<sup>32</sup>P] dATP and T4 polynucleotide kinase (M0201S; BioLabs) according to the supplier's instructions. Double-stranded radiolabeled ligands were obtained by incubating equimolar amounts of complementary oligos at 95 °C for 10 min followed by slow cooling to room temperature. Probes sequences for Int1 were 5'-TATGTGTATGTATAAAGTGTGTTTGTGTATATGTATGTGA-3' and for PbDT-3' were 5'-AATGTGTATGTTGTGTGCATATTTTTTTTGTGCATGC-3'. Sequences of Ups1-, hrp2-3', and Ups1-mutated competitors are indicated in Fig. 4C. For gel shift assays the 20-μL standard binding reaction mixture contained 50 ng poly(deoxyinosinic-deoxycytidylic) acid, 2.5% glycerol, 5 mM MgCl<sub>2</sub>, 0.05% Nonidet P-40, 50 mM KCl, 10 mM EDTA, 1× binding buffer (LightShift 20148 A; Thermo Scientific), and 12.5 fmol double-stranded 5'-<sup>32</sup>P-labeled ligands. Competitive unlabeled ligands were added in increasing concentrations (25 fmol, 125 fmol, and 312.5 fmol) as indicated. Parasite protein extracts were added last to ensure that the probe and its competitors had the same probability of binding the nuclear extract. The reaction mixtures were incubated on ice for 30 min; then samples were loaded on a 6% (wt/vol) native polyacrylamide gel in TAE buffer [6.7 mM Tris-acetate, 3.3 mM sodium acetate, 1 mM EDTA (pH 7.5)]. Electrophoresis was conducted at 2–4 °C and 16 V/cm for 1 h. Protein-DNA complexes were visualized and quantified by a Bio Imaging Analyzer (BAS1000; Fuji).

- Aley SB, Sherwood JA, Howard RJ (1984) Knob-positive and knob-negative *Plasmodium falciparum* differ in expression of a strain-specific malarial antigen on the surface of infected erythrocytes. *J Exp Med* 160(5):1585–1590.
- Calderwood MS, Gannoun-Zaki L, Welles TE, Deitsch KW (2003) *Plasmodium falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem* 278(36):34125–34132.
- Duraisingh MT, Triglia T, Cowman AF (2002) Negative selection of *Plasmodium falciparum* reveals targeted gene deletion by double crossover recombination. *Int J Parasitol* 32(1):81–89.
- Frank M, et al. (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite *Plasmodium falciparum*. *J Biol Chem* 281(15):9942–9952.
- Wu Y, Kirkman LA, Welles TE (1996) Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. *Proc Natl Acad Sci USA* 93(3):1130–1134.
- Deitsch KW, Driskill CL, Welles TE (2001) Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res* 29(3):850–853.
- Dzikowski R, Deitsch KW (2008) Active transcription is required for maintenance of epigenetic memory in the malaria parasite *Plasmodium falciparum*. *J Mol Biol* 382(2):288–297.
- Dzikowski R, Frank M, Deitsch K (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathogens* 2:e22.
- Kyes S, Pinches R, Newbold C (2000) A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol* 105(2):311–315.
- Voss TS, Mini T, Jenoe P, Beck HP (2002) *Plasmodium falciparum* possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript. *J Biol Chem* 277(20):17493–17501.

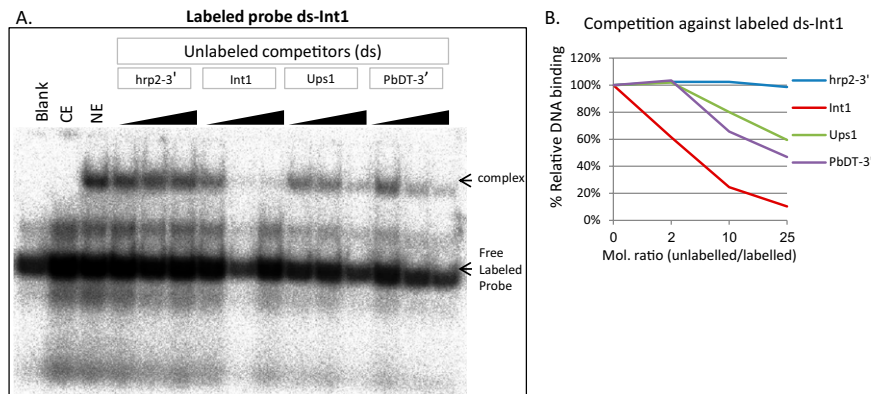


**Fig. S1.** An unregulated var promoter can silence the entire var gene family by titration. Clonal populations of NF54 parasite line (C3) predominantly expressing the “slow-switching” var gene PFD1005c (1–3) were transfected with the pVBbIDh plasmid, which is identical to pVLbIDh described in Fig. 1 except that luciferase is replaced by the bsd resistance gene. This plasmid contains a PbDT 3' between the var promoter and the var intron and is not recognized by the mechanism that controls mutually exclusive expression. (A) Schematic of the pVBbIDh plasmid. (B) Results of pVBbIDh selection on WR99210 only (resistance achieved by intron promoter activity; Upper) and on 2 ug/mL blasticidin (resistance achieved by var promoter; Lower) in the C3 parasite line. (Left) Quantification of the ratio of var promoter to intron by qRT-PCR. (Right) Steady-state mRNA levels of each individual var gene measured by qRT-PCR presented as copy number relative to the copy number of the housekeeping gene arginyl-tRNA synthetase (PFL0900c). All values presented are the average of at least two biological replicates. Error bars represent SE.

- Frank M, et al. (2006) Strict pairing of var promoters and introns is required for var gene silencing in the malaria parasite *Plasmodium falciparum*. *J Biol Chem* 281(15):9942–9952.
- Dzikowski R, Deitsch KW (2008) Active transcription is required for maintenance of epigenetic memory in the malaria parasite *Plasmodium falciparum*. *J Mol Biol* 382(2):288–297.
- Fastman Y, Noble R, Recker M, Dzikowski R (2012) Erasing the epigenetic memory and beginning to switch—the onset of antigenic switching of var genes in *Plasmodium falciparum*. *PLoS ONE* 7(3):e34168.



**Fig. S2.** Analysis of the TG-rich DNA motif found in *var* introns and 5' UTRs. (A) Schematic of the exact position on the *var*-regulatory regions of the PEs used in our constructs. The *var* promoter and intron have the same sequence, which is a fusion of the M2 and M3 motifs. (B) Alignment of the three T[G/A] motifs (M1-3) can create one gold motif (consensus motif). (C) Analysis of the distribution of motifs on *var* promoters and introns of the entire *var* gene family showing identical distribution of the M1 and M2 motifs on *var* promoters and introns.



**Fig. S3.** The *var* PEs form specific DNA-protein complexes. (A) EMSA of extracts using of radiolabeled ( $^{32}$ P) DNA ligand containing the PE of the *var* intron (Int1) shows the formation of a specific DNA-protein complex when incubated with nuclear extract (NE). Specific competition assays were performed with increasing concentrations of unlabeled DNA ligands (2 $\times$ , 10 $\times$ , and 25 $\times$  of the labeled ligand respectively) containing the PEs found in *var* 5' UTR (Ups1), *var* intron (Int1), and PbDT 3' UTR (PbDT-3'). The DNA sequence from *hrp2-3'* UTR was used as a nonspecific competitor (*hrp2-3'*). Blank, no protein was added; CE, cytoplasmic extract. (B) Phosphorimaging quantification of the EMSA data presented in A. The percentage of complex formation measured is presented relative to the value measured in the absence of competitive ligand, which was considered to be 100%.

**Table S1. Genome-wide distribution of the specific PE sequence TGTATATGTATGTG in *var* gene-regulatory regions**

Gene type	Gene ID		Genomic location - intron		Genomic location - 5' UTR		
A	PFD0020c	R	Pf3D7_04_v3: 47482–47496	(+)			
			Pf3D7_04_v3: 47015–47029	(–)			
A/B	PFF1580c	R	Pf3D7_06_v3: 1355987–1356001	(+)			
			Pf3D7_06_v3: 1355949–1355963	(+)			
			Pf3D7_06_v3: 1355537–1355551	(–)			
A/B	PF08_0140	F	Pf3D7_08_v3: 48659–48673	(+)			
			Pf3D7_08_v3: 49468–49482	(–)			
A/B	PFL0020w	F	Pf3D7_12_v3: 40613–40627	(–)			
B	PFA0005w	F	Pf3D7_01_v3: 34869–34883	(+)	Pf3D7_01_v3: 28090–28104	(+)	ATG -1420
			Pf3D7_01_v3: 34907–34921	(–)			
			Pf3D7_01_v3: 35790–35804	(–)			
B	PFA0765c	R	Pf3D7_01_v3: 609491–609505	(+)	Pf3D7_01_v3: 616332–616346	(+)	ATG -1453
			Pf3D7_01_v3: 608870–608884	(–)			
			Pf3D7_01_v3: 608794–608808	(–)			
B	PFB0010w	F	Pf3D7_02_v3: 29080–29094	(+)			
B	PFB1055c	R	Pf3D7_02_v3: 918223–918237	(+)	Pf3D7_02_v3: 925015–925029	(+)	ATG -1381
			Pf3D7_02_v3: 917672–917686	(–)			
B	PFC1120c	R	Pf3D7_03_v3: 1032971–1032985	(+)	Pf3D7_03_v3: 1039658–1039672	(+)	ATG -1418
			Pf3D7_03_v3: 1032933–1032947	(+)			
B	PFC0005w	F	Pf3D7_03_v3: 42367–42381	(+)	Pf3D7_03_v3: 35564–35577	(+)	ATG -1402
			Pf3D7_03_v3: 43066–43080	(–)			
B	PFD1245c	R	Pf3D7_04_v3: 1174516–1174530	(–)			
B	PFD0005w	F	Pf3D7_04_v3: 35433–35447	(+)	Pf3D7_04_v3: 27294–27308	(+)	ATG -1412
B	PFE0005w	F	Pf3D7_05_v3: 26359–26373	(+)	Pf3D7_05_v3: 19501–19515	(+)	ATG -1428
			Pf3D7_05_v3: 26397–26411	(+)			
			Pf3D7_05_v3: 27115–27129	(–)			
B	PFF1595c	R	Pf3D7_06_v3: 1377261–1377275	(+)	Pf3D7_06_v3: 1383972–1383986	(+)	ATG -1359
B	MAL7P1.187	R	Pf3D7_07_v3: 1419707–1419721	(+)			
			Pf3D7_07_v3: 1419832–1419846	(+)			
B	MAL8P1.220	F	Pf3D7_07_v3: 25871–25885	(+)	Pf3D7_07_v3: 18798–18812	(+)	ATG -1509
			Pf3D7_07_v3: 25909–25923	(+)			
			Pf3D7_07_v3: 26647–26661	(–)			
B	MAL7P1.212	R		(+)	Pf3D7_08_v3: 1444841–1444855	(+)	ATG -1393
				(–)			
B	PF08_0142	F	Pf3D7_08_v3: 26412–26426	(+)	Pf3D7_08_v3: 19957–19971	(+)	ATG -1404
			Pf3D7_08_v3: 26458–26472	(+)			
B	PFI1830c	R	Pf3D7_09_v3: 1497759–1497773	(+)	Pf3D7_09_v3: 1504713–1504727	(+)	ATG -1391
B	PFI0005w	F	Pf3D7_09_v3: 25610–25624	(+)	Pf3D7_09_v3: 18706–18720	(+)	ATG -1374
			Pf3D7_09_v3: 26308–26322	(–)			
			Pf3D7_09_v3: 26426–26440	(–)			
			Pf3D7_09_v3: 26251–26265	(–)			
B	PF10_0406	R	Pf3D7_10_v3: 1644539–1644553	(+)	Pf3D7_10_v3: 1651322–1651336	(+)	ATG -1388
B	PF10_0001	F	Pf3D7_10_v3: 33935–33949	(+)	Pf3D7_10_v3: 27078–27092	(+)	ATG -1412
			Pf3D7_10_v3: 33998–34012	(+)			
			Pf3D7_10_v3: 34089–34103	(+)			
B	PF11_0007	F	Pf3D7_11_v3: 29522–29536	(+)	Pf3D7_11_v3: 22798–22812	(+)	ATG -1362
			Pf3D7_11_v3: 29560–29574	(+)			
B	PFL0005w	F		(+)	Pf3D7_12_v3: 15568–15582	(+)	ATG -1405
B	PFL2665c	R	Pf3D7_12_v3: 2242722–2242736	(–)	Pf3D7_12_v3: 2250334–2250348	(+)	ATG -1386
			Pf3D7_12_v3: 2243361–2243375	(+)			
B	PFL0935c	R	Pf3D7_12_v3: 768667–768681	(+)	Pf3D7_12_v3: 775641–775655	(+)	ATG -1459
B	MAL13P1.1	F	Pf3D7_13_v3: 26583–26597	(+)	Pf3D7_13_v3: 19936–19950	(+)	ATG -1428
B	MAL13P1.356	R	Pf3D7_13_v3: 2886940–2886954	(+)	Pf3D7_13_v3: 2893732–2893746	(+)	ATG -1406
			Pf3D7_13_v3: 2886902–2886916	(+)			
B/C	PFD0635c	R	Pf3D7_04_v3: 594152–594166	(+)			
B/C	MAL7P1.50	R	Pf3D7_07_v3: 545014–545028	(+)			
B/C	PF07_0050	R	Pf3D7_07_v3: 554036–554050	(+)			
B/C	MAL7P1.55	R	Pf3D7_07_v3: 583418–583432	(+)			
			Pf3D7_07_v3: 583454–583468	(+)			
B/C	PF08_0106	F	Pf3D7_08_v3: 445897–445911	(+)			
			Pf3D7_08_v3: 445935–445949	(+)			

**Table S1. Cont.**

Gene type	Gene ID		Genomic location - intron		Genomic location - 5' UTR		
B/C	PF08_0103	F	Pf3D7_08_v3: 464517–464531	(+)			
			Pf3D7_08_v3: 464713–464727	(+)			
B/C	PFL1955w	F	Pf3D7_12_v3: 1710054–1710068	(+)			
C	PFD0625c	R	Pf3D7_04_v3: 563669–563683	(+)	Pf3D7_04_v3: 569433–569447	(+)	ATG -104
			Pf3D7_04_v3: 563707–563721	(+)			
			Pf3D7_04_v3: 563052–563066	(–)			
C	PFD0630c	R	Pf3D7_04_v3: 579013–579027	(+)	Pf3D7_04_v3: 584758–584772	(+)	ATG -104
C	PFD0995c	R		(–)	Pf3D7_04_v3: 941912–941926	(+)	ATG -50
C	PFD1000c	R	Pf3D7_04_v3: 947514–947527	(–)	Pf3D7_04_v3: 953809–953823	(+)	ATG -50
C	PFD1015c	R	Pf3D7_04_v3: 971159–971173	(+)	Pf3D7_04_v3: 976629–976643	(+)	ATG -52
			Pf3D7_04_v3: 970473–970487	(–)			
C	PFF0845c	R	Pf3D7_06_v3: 724928–724942	(+)			
			Pf3D7_06_v3: 725219–725233	(+)			
C	PF07_0048	R	Pf3D7_07_v3: 514031–514045	(+)			
C	PF07_0049	R	Pf3D7_07_v3: 528993–529007	(–)			
			Pf3D7_07_v3: 528833–528847	(–)			
C	PF07_0051	R	Pf3D7_07_v3: 568765–568779	(+)			
			Pf3D7_07_v3: 568956–568970	(+)			
C	MAL7P1.56	R	Pf3D7_07_v3: 591649–591663	(–)	Pf3D7_07_v3: 597771–597785	(+)	ATG-52
			Pf3D7_07_v3: 592329–592343	(+)			
C	PF08_0107	F	Pf3D7_08_v3: 436823–436837	(+)	Pf3D7_08_v3: 431076–431090	(+)	ATG-89
			Pf3D7_08_v3: 436785–436799	(+)			
C	PFL1960w	F	Pf3D7_12_v3: 1725397–1725411	(+)	Pf3D7_12_v3: 1719469–1719483	(+)	ATG-105
C	PFL1970w	F	Pf3D7_12_v3: 1741135–1741149	(+)	Pf3D7_12_v3: 1735491–1735505	(+)	ATG-52
			Pf3D7_12_v3: 1741743–1741757	(–)			
			Pf3D7_12_v3: 1741796–1741810	(–)			
			Pf3D7_12_v3: 1741846–1741860	(–)			
Pseudo	PF14_0001	F	Pf3D7_14_v3: 3172–3186	(+)			
			Pf3D7_14_v3: 3804–3818	(–)			
Pseudo	PF14_0773	F			Pf3D7_14_v3: 3290898–3290912	(+)	ATG-55

PEs with a single base-pair replacement from guanine to cytosine [G/C] in either position 8 or 12 are shaded in light gray and dark gray, respectively. Gene chromosomal orientation is marked as forward (F) or reverse (R). + and – refer to the sequence orientation on the DNA strand.

**Table S2. Genome-wide distribution of the specific PE sequence TGTATATGTATGTG in regions other than var genes**

Gene type	Gene ID		Genomic location	
Conserved Plasmodium protein, unknown function	PFD0160w	F	Pf3D7_04_v3: 193615–193629	(–) Ex2
Conserved Plasmodium protein, unknown function	PFE0990w	F	Pf3D7_05_v3: 819231–819245	(–)/(+) 3' UTR
Conserved Plasmodium protein, unknown function	PFE0995c	R		
Cysteine desulfurase, putative (SufS)	PF07_0068	F	Pf3D7_07_v3: 727682–727696	(+) 3' UTR
Conserved Plasmodium protein, unknown function	PF07_0069	F		5' UTR
Methionine aminopeptidase 1c, putative (MetAP1c)	MAL8P1.140	F	Pf3D7_08_v3: 249646–249660	(–) 3' UTR
Conserved Plasmodium protein, unknown function	MAL8P1.139	F		5' UTR
Transcription factor with AP2 domain(s) (ApiAP2)	PFL1085w	F	Pf3D7_12_v3: 918636–918649	(–) 3' UTR
Glideosome-associated protein 45 (GAP45)	PFL1090w	F		5' UTR
Sterile transcript non -rotein coding	PF3D7_1363900	F	Pf3D7_13_v3: 2562250–2562264	(–) ncRNA
Conserved Plasmodium membrane protein, unknown function	MAL13P1.320	F		3' UTR
Conserved Plasmodium protein, unknown function	MAL13P1.321	F		5' UTR
Conserved protein, unknown function	PF13_0106	F	Pf3D7_13_v3: 791981–791995	(+)/(–) 3' UTR
Conserved Plasmodium protein, unknown function	PF13_0107	R		
NOT family protein, putative	PF14_0170	F	Pf3D7_14_v3: 703766–703780	(+) Ex1
Conserved Plasmodium protein, unknown function	MAL7P1.102	R	Pf3D7_07_v3: 853284–853298	(+) Ex1
Serine/threonine protein phosphatase (PP7)	PF14_0224	F	Pf3D7_14_v3: 938906–938920	(+) Int9

PE with a single base-pair replacement from guanine to cytosine [G/C] in position 12 is shaded in light gray. Gene chromosomal orientation is marked as forward (F) or reverse (R).+ and – refer to the sequence orientation on the DNA strand. ncRNA, noncoding RNA.