Supporting Information

Ikadai et al. 10.1073/pnas.1217712110

SI Materials and Methods

Quantitative RT-PCR of 16 Putatively Disrupted Transcripts in the Parental Line. Schizonts from the parental line were collected on a 40%/70% Percoll–sorbitol gradient and returned to culture at a 4% hematocrit; 3 h later, the parasites were sorbitol synchronized to ensure a highly synchronous population of parasites. RNA was collected at 0, 12, 24, 30, 34, 38, 44, and 48 h postinvasion. Giemsa smears were made to confirm parasite asexual stages. RNA was extracted (TRIzol; Invitrogen), DNase was treated (Promega), and cDNA was synthesized (SuperScript III; Invitrogen). cDNA concentration was determined on a Nanodrop (Thermo Scientific), and equal amounts of cDNA were used in each quantitative RT-PCR. Primers specific for the putatively disrupted transcripts were used to amplify ~150–200 bp of the coding region (Table S3). PF07_0073 (seryl-tRNA synthetase), which remains stable throughout the lifecycle of the parasite (1), was used as an internal control. Quantitative RT-PCR was performed using SYBR Green qPCR Mastermix (ABI). Results were collected on an ABI 7300 and analyzed using ABI 7300 software using the relative quantification by ΔΔCt method compared with the PF07_0073 gene reference (2).

1. Rovira-Graells N, et al. (2012) Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* 22(5):925–938.

2. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25(4):402–408.

Table S1. Southern blot and episomal PCR reveal that 22 of 29 IGMs contain single *piggyBac* insertions

Lane number	IGM	Transfection number	Generation number	Number of insertions	Episome? (PCR)
1	1A2	1	4	1	Yes
2	1B5	1	4	1	No
3	1A4	1	4	2	No
4	1E5	1	4	2	No
5	1G1	1	4	1	Yes
6	1C5	1	4	1	Yes
7	2G2	3	1	1	No
8	2F5	3	2	2	No
9	2F4	3	1	1	No
10	2D3	3	1	1	No
11	2B4	3	1	1	No
12	2A2	3	1	1	Yes
13	2E2	3	1	1	No
14	2D4	3	1	2	Yes
15	2G3	3	1	2	No
16	2E4	3	4	1	No
17	2H2	3	1	1	No
18	2H3	3	1	1	No
19	2H12	2	4	2	Yes
20	2G6	3	2	2	No
21	2D7	3	2	1	No
22	2G10	2	3	1	No
23	2H8	3	3	1	No
24	2A9	2	1	1	No
25	2G11	2	3	1	No
26	2F11	2	3	1	Yes
27	2A11	2	3	1	No
28	2F12	2	4	1	No
29	2F6	3	2	1	No

Summary of Southern blot results. Lane number and IGM indicate which IGM corresponds to each lane on the Southern blot of Fig. 1*C*. IGMs were recovered from three independent transfection experiments (transfection number). Generation number indicates how long the parasite was cultured posttransfection before cloning and drug pressure. A single band on the Southern blot of Fig. 1*C* is considered a single insertion. When digested with *EcoRI*, an episomal *piggyBac* plasmid will produce a 6.2-kb band. To verify whether this band corresponds to an epicome, PCR was performed using plasmid-specific primers (Fig. 1*B* and Table S3). IGMs that yielded two or more bands by Southern blotting and were not positive for an episome by PCR were classified as having more than one insertion.

Table S2. Inverse PCR reveals *piggyBac* insertions in 9 of 14 *Plasmodium falciparum* chromosomes

Lane number	IGM	Number of insertions	Chromosome	Insertion position	ITR1 side sequence	ITR2 side sequence
1	1A2	1	5	1288428	ATCATATTACATTAA	TTAATTATAATGAAT
2	1B5	1	7	705469	TGGTAATCTTTTTAA	TTAAAATGGTATATA
3	1A4	2	7, 8	705469, 1308230	TGGTAATCTTTTTAA, AACAATGTTTATTAA	TTAAAATGGTATATA, TTAAAAAATTCTACG
4	1E5	2	5, ?	1142963	TTATAATATATTTAA	TTAAGATTGCATGAT
5	1G1	1	5	1288428	ATCATATTACATTAA	TTAATTATAATGAAT
6	1C5	1	14	2544860	ATAAATATCTTTTAA	TTAAATAACTCTTGG
7	2G2	1	9	1006267	AAAAATATATATTAA	TTAATATATATATAT
8	2F5	2	11, ?	1987528	ATATAATAATCTTAA	TTAATAATTATAATT
9	2F4	1	5	1319259	ATGTTAATCAATTAA	TTAATGATTATATAT
10	2D3	1	12	2166205	AATATATTAATTTAA	TTAAAATTATCAGTT
11	2B4	1	11	598906	CGTATGCTGCATTAA	TTAAAAAGTATATTC
12	2A2	1	3	213676	TAAAAACATAATTAA	TTAAATATAATAAAA
13	2E2	1	7	1379440	AAAAACGTAAATTAA	TTAAGAATAATTATA
14	2D4	2	14, ?	2285695	TTTTTTTTCTTTTAA	TTAAAGGAAATGTTG
15	2G3	1	5, ?	1288428	ATCATATTACATTAA	TTAATTATAATGAAT
16	2E4	1	13	724860	GGAGTTACCTATTAA	TTAATGCTAATCATA
17	2H2	1	12	1528986	TTTGTATTTAATTAA	TTAAAAAAAAAATTA
18	2H3	1	5	1142963	TTATAATATATTTAA	TTAAGATTGCATGAT
19	2H12	2	11, ?	1987528	ATATAATAATCTTAA	TTAATAATTATAATT
20	2G6	2	13	2403095	ATACAGCATATTTAA	TTAAAATATAATTAA
21	2D7	1	5	1288428	ATCATATTACATTAA	TTAATTATAATGAAT
22	2G10	1	5	1288428	ATCATATTACATTAA	TTAATTATAATGAAT
23	2H8	1	14	387148	TTATATTTCCTTTAA	TTAACAACTTTGAGT
24	2A9	1	12	1528986	TTTGTATTTAATTAA	TTAAAAAAAAAATTA
25	2G11	1	9	1006267	AAAAATATATATTAA	TTAATATATATATAT
26	2F11	1	14	2285695	TTTTTTTTCTTTTAA	TTAAAGGAAATGTTG
27	2A11	1	4	732679	ATCTATATTTTTTAA	TTAAAGGAAAGAAGA
28	2F12	1	13	724860	GGAGTTACCTATTAA	TTAATGCTAATCATA
29	2F6	1	9	823324	AATAGCATATGTTAA	TTAAATATAAAATAA

piggyBac insertion position for each of 29 IGMs identified by inverse PCR. In seven IGMs with multiple insertions, at least one insertion position was identified by inverse PCR. Insertion position indicates the nucleotide at the 5′ end of the insertion. In all cases, *piggyBac* inserted into the conical TTAA site at both the inverted terminal repeat 1 (ITR1) and ITR2 sides.

Table S3. Primer sequences

Primer	Sequence
PFE1370wF	TGTGCACACGCTGATATTGA
PFE1370wR	ATGCCTTTTCAGCTGCTTTC
PF11_0167F	GCTCATCACTGCAGGACTTG
PF11_0167R	CCTTGGCACAAAGAGCATTT
PFE1615cF	TGTTGAAGAAGACCCAAGTGAA
PFE1615cR	GCGCTTCTTAAAAGATCAGCA
PFD0800cF	GATGCCCACGATCGTTTTAT
PFD0800cR	GGATTCAAAGGGTCCAAATG
PFL1770cF	GCCAAAGTGAAGAGCAAAGG
PFL1770cR	TTTCGGTTTTTGTGCCGTAT
PFL2550wF	CGGGAAGAAGAATTGACTGAA
PFL2550wR	CCACGATATATTTTCCCCACA
PF14_0532F	GGAAATATCCAGCCCACCAAA
PF14_0532R	ATTCGGTGCCATTAGGGTTA
PFI1215wF	GAGCCACGTTCGAAAAACAT
PFI1215wR	ACTTCTTGATCCGGCTTGAA
MAL7P1.171F	GAGGAACATTCCCCAGATGA
MAL7P1.171R	ATACCGTGGCTTTTTGATGC
PFC0200wF	TTGGAGGTCAAACAAACCA
PFC0200wR	CTGCTCCTCCCTTCTTTTTC
PFI0980wF	GCTTGAAACAGACGCATCAA
PFI0980wR	TCCACCGTAAAGCTCCATGT
PF14_0097F	TCCAACCATTTGTATCTTGGAA
PF14_0097R	GGTGCTAAGAATGCAGCAAA
PF07_0055F	CATGAATGGGCAGACAGAGA
PF07_0055R	TCTTTTGCTTTTTCTTTTTCAAGT
PF14_0595F	TGAATGCATCCAAACAAATGA
PF14_0595R	TTTTTCTTTCTGAAGAGCAGCA
PFE1570cF	AAACAAACCATTTTTGTTCCGAT
PFE1570cR	CTTTTGGGGGTTGGATTTT
PF13_0097F	AAAAGTTTTGGTGCGTCTTCA
PF13_0097R	CGTTATCCTCCGAAATGGAA
PF07_0073F	AAGTAGCAGGTCATCGTGGTT
PF07_0073R	TTGGGCACATTCTTCCATAA
PFE1365wF	TGAAATTAAAGAAAATGGCGAAG
PFE1365wR	CCCCATAAGCATACTGCACA
PFC0195wF	CGAAAAGGTGCTGATTCAAA
PFC0195wR	TCGATGGACATAGGACCGTA
PFD0805wF	CCTTTGGTATAAATTATAGTACCTG
PFD0805wR	TTCCACCTTTTCAAATAATATGG
PF14_0096F	CCCAGCACCATCGATACCTA
PF14_0096R	TCAATGGCCCTAAGGTTTCA
PFL1775cF	ATACCCCAATACGCATCCTG
PFL1775cR	CCCTCAATAAAACGGGAGTG
PF11_0166F	GAAAGAAGATCAAGTGTCAC
PF11_0166R	GAGACTTTTAATTCTGTCAGTAAAAG
PFL2545cF	TGATGCACAAGATGATTTTATG
PFL2545cR	CAACATTTTGACTGAACTTTTC
PFI1210wF	TAGGCCCAGACAAAGGTGAG
PF1210wR	TCGACCTTTCTTCGTTTCTTTC
MAL7P1.170F	TTGCTGTAAGCAACCCACAG
MAL7P1.170R	TGGCAATTACGTGAAGTGGA
PF07_0055F	TTCATCTTCACATATGACCC
PF07_0055R	AAAAGGTCTTCAGGTTTTTC
PFE1590wF	GAATGCCTAGTAAGCATGATTC
PFE1590wR	TTCACCTACGGAAACCTTG
Inverse F1	AGATGTCCTAAATGCACAGCGAC
Inverse F2	CTCCAAGCGGCGACTGAG
Inverse F3	CATTGACAAGCACGCCTCAC
Inverse F4	GCTATTTAGAAAGAGAGAGC
Inverse R1	GTCAATGCGGTAAGTGTCACTGA
Inverse R2	GTAAGTGTCACTGATTTTGAACTATAACG
Inverse R3	GACGCATGATTATCTTTTACGTGAC

Table S3. Cont.

Primer	Sequence
Inverse R4	TATCGATACCGTCGACCTCG
Episomal check F (T7)	TAATACGACTCACTATAGGG
Episomal check R (M13R)	GGAAACAGCTATGACCATG
hDHFRF	ATGCATGGTTCGCTAAACTGC
hDHFRR	TTAATCATTCTTCTCCATATAC
PFE1615c rescue F	CAAGGGCCCGATATGTGTGAAAAGTGGGAATG
PFE1615c rescue R	CAAGGTACCATTTATGTAGAGAGTGTTCTATG
PFD0800c rescue F	TAATGCGGCCGCGGTACCGAATGTAAATAAAATGTGCTCTG
PFD0800c rescue R	TAATACTAGTGGGCCCTTAACAAAGTGAGGAATCTTCATCA
PFI1215w rescue F	TAATGCGGCCGCTGGAGGAGATCCACAGAATG
PFI1215w rescue R	TAATACTAGTTGTATATACGGGTCGTCAAAAA
PFL2550w rescue F	TAATGCGGCCGCATTTAATGAAAAACGAAAAACCAT
PFL2550w rescue R	TAATACTAGTTTTCATTTTCACACAATCAAATCT
PFC0200w rescue F	CAAGGGCCCAGCAGGAAACCAAAGTAAGCA
PFC0200w rescue R	CAAGGTACCTGACTCCATTACATACAGTCACATAAA
Pfs230F(PFB0405w)	ACTCTTCCCCCTGTGGTACA
Pfs230R	TTCATCCAAGAAAGCACATCC
Pf11.1F (PF10_0374)	AAGGCTAAGCATCTAAAGGGTTC
Pf11.1R	AACTCCAGCCTGCTTTTCCT
Pfs48/45F(PF13_0247)	ATTTCTGCGAAAAAGGCTCA
Pfs48/45R	CCTACGTTCACGCATATCTGG
Pfs28F (PF10_0303)	TGAGGACACGTGTGGAAAAA
Pfs28R	TCTCGACAAACATTAGGAACACA
Pfg377F (PFL2405c)	CTGAACCATGGCCTCTTGAT
Pfg377R	AACCTTCGTTGTGGTCTTCG
PfsMR5F (PFB0400w)	ATGGATCACATGCATCCTCA
PfsMR5R	TTGTAGGGGCAGCTAAATCC
a-Tubulin IIF (PFD1050w)	TGCTGAGAAGGCATATCACG
a-Tubulin IIR	CGGCAGCATTAACATCCTTT
Pfs47F (PF13_0248)	TGGGTGTGATTTTACGAAACC
Pfs47R	TGCGAATCGATTTCTTTTTCA
Pfmdv1F (PFL0795c)	CCGTTTCTTCATTAGCATTTCC
Pfmdv1R	TAGGAGCAAAAGCAGGTGAT
Pfs16F (PFD0310w)	TTCTTCGCTTTTGCAAACCT
Pfs16R	AAAGGCATTTTGTCAGCAGAA
Pfg748F (PF14_0748)	CAATGGGAATTCTGATAAAGCTG
Pfg748R	GTCATTTCGTGGGGTATCGT
Pfg744F (PF14_0744)	TTCTTACAGGCGCCAGAACT
Pfg744R	CTTCCGTCAACAGAGCCATT
Pfg164F (PF10_0164)	GCGTTGAAGGATATCGACAGA
Pfg164R	GCGTTGAAGGATATCGACAGA
Pfg27F (PF13_0011)	GCCCTTGGATAAATTTGGAA
Pfg27R	GGATCCTTGCTAAGGGTCATC

Primer sequences used in this work.

Fig. S1. RT-PCR analysis of the putatively disrupted genes in each of 22 insertional gametocyte-deficient mutants (IGMs) with single *piggyBac* insertions. mRNA abundance was assessed by semiquantitative RT-PCR using primers designed to amplify ~150 bp of the transcript coding region (Table S3). Seryl-tRNA synthetase (PF07_0073) was used as a loading control (CTRL) (1). The results are grouped according to the position of *piggyBac* insertion (*pB*) and the orientation of the flanking genes. (*A*) Analysis of mRNA abundance of IGMs where *piggyBac* inserted directly into the ORF. In all cases, mRNA abundance was reduced or completely absent, indicating gene knockdown or complete KO. (*B* and *C*) *piggyBac* insertions between two ORFs with the same transcriptional orientation. (*B*) Transcriptional orientation to the telomere. (*C*) Transcriptional orientation to the centromere. In all cases, mRNA abundance was reduced or absent in the downstream ORF. (*D*) *piggyBac* insertions between two ORFs with opposite transcriptional orientations. In all cases, mRNA was decreased in only one of the ORFs.

Fig. S1

1. Rovira-Graells N, et al. (2012) Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* 22(5):925–938.

Fig. S2. Changes of mRNA abundance for each of the putatively disrupted genes during the asexual life cycle. The patterns of mRNA expression for each of 16 putatively disrupted genes were measured in tightly synchronized parental parasites. RNA was prepared from samples collected at 0 (ER), 12 (LR), 24 (ET), 30 (T), 34 (LT), 38 (ES), 44 (S), and 48 (LS) h postinvasion. Quantitative RT-PCR was performed with the primers listed (Table S3). In all cases, the mRNA expression patterns matched those patterns estimated from microarray analysis (1). ER, early ring; ES, early schizont; ET, early trophozoite; LR, late ring; LS, late schizont; LT, late trophozoite; S, schizont; T, trophozoite.

Fig. S2

1. Le Roch KG, et al. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301(5639):1503–1508.

Fig. S3. Semiquantitative RT-PCR of gametocyte-specific transcripts for each IGM. Results of semiquantitative RT-PCR expression assays with two independent biological replicates of each IGM parasite as well as the parental line (P). The parasite source is indicated above each lane. Expression of each of 16 IGM genes as well as additional gametocyte-specific genes (1) was assayed. Results for the seryl-tRNA synthetase mRNA (loading control) are shown in the first row. The results were tabulated in a binary fashion. The gene was considered expressed if a band was present in either biological replicate. The gene was considered not expressed if both biological replicates were negative. A summary of these results is presented in Fig. 4.

Fig. S3

1. Young JA, et al. (2005) The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol* 143(1): 67–79.