Supplementary Material for

Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis

Min Zhang, Chengqi Wang, Thomas D. Otto, Jenna Oberstaller, Xiangyun Liao, Swamy R. Adapa, Kenneth Udenze, Iraad F. Bronner, Deborah Casandra, Matthew Mayho, Jacqueline Brown, Suzanne Li, Justin Swanson, Julian C. Rayner,* Rays H. Y. Jiang,* John H. Adams*

*Corresponding author. Email: jadams3@health.usf.edu (J.H.A.); jiang2@health.usf.edu (R.H.Y.J.); jr9@sanger.ac.uk (J.C.R.)

Published 4 May 2018, *Science* **360**, eaap7847 (2017)
DOI: 10.1126/science.aap7847

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/360/6388/eaap7847/suppl/DC1)

Tables S1 to S9 as separate Excel files

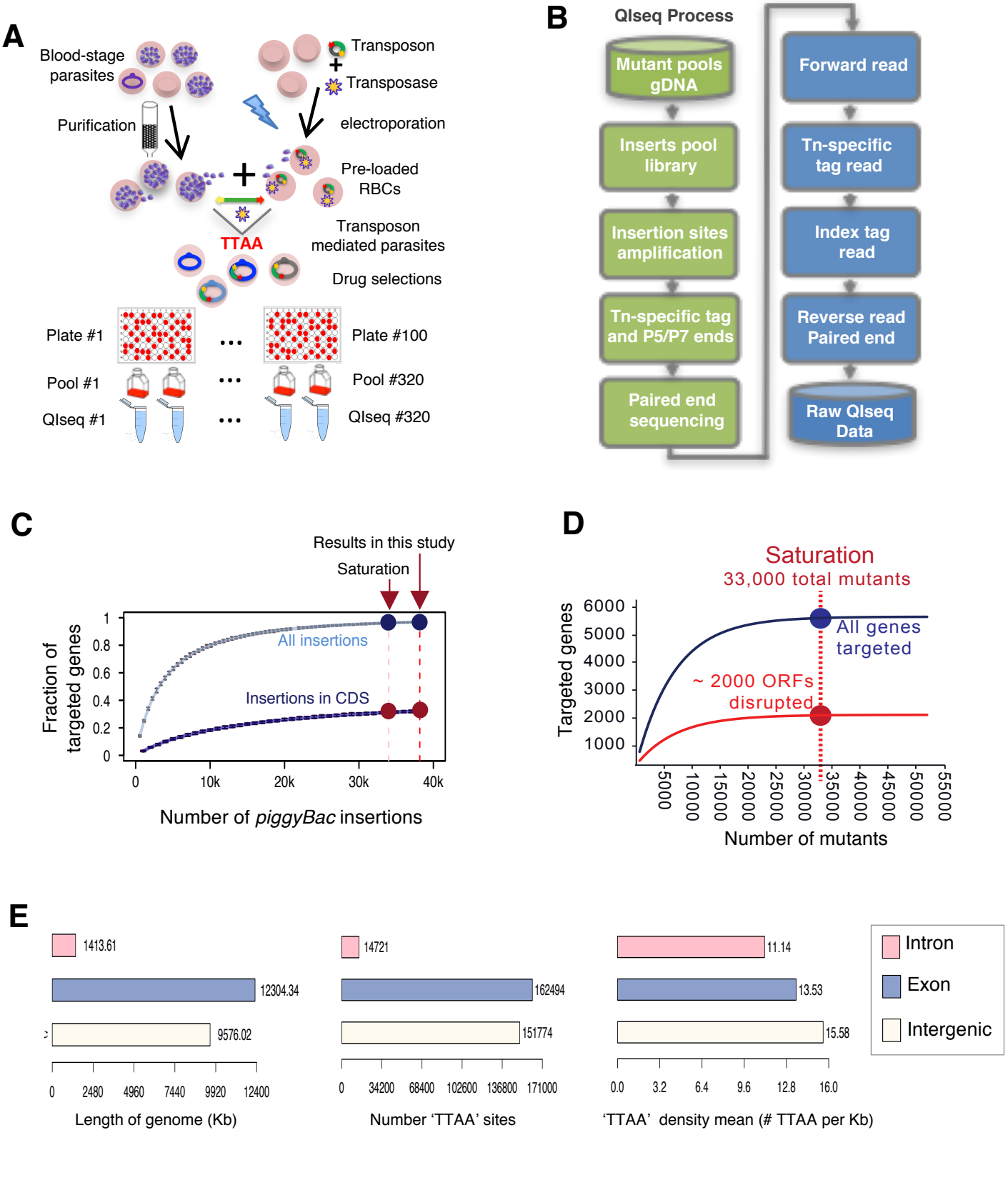

A
Blood-stage parasites
Purification
Transposon
Transposase
electroporation
Pre-loaded RBCs
TTAA
Transposon mediated parasites
Drug selections
Plate #1
Plate #100
Pool #1
Pool #320
QIseq #1
QIseq #320
B
QIseq Process
Mutant pools gDNA
Inserts pool library
Insertion sites amplification
Tn-specific tag and P5/P7 ends
Paired end sequencing
Forward read
Tn-specific tag read
Index tag read
Reverse read Paired end
Raw QIseq Data
C
Results in this study
Saturation
All insertions
Insertions in CDS
Fraction of targeted genes
Number of piggyBac insertions
D
Saturation
33,000 total mutants
All genes targeted
~ 2000 ORFs disrupted
Targeted genes
Number of mutants
E
1413.61
12304.34
9576.02
14721
162494
151774
11.14
13.53
15.58
Intron
Exon
Intergenic
Length of genome (Kb)
Number 'TTAA' sites
'TTAA' density mean (# TTAA per Kb)

Fig. S1. Methodology for *P. falciparum* whole genome saturation-level mutagenesis, estimate of saturation, and *piggyBac* insertion site localization distribution in the genome.

(**A**) The experiment approach for whole genome saturation mutagenesis was optimized to generate mutants with a single *piggyBac* insertion (12, 19). (**B**) Mutant parasite pools were harvested and insertion sites were identified by Quantitative Insertion-site Sequencing (QIseq)(17). (**C**) Plot shows fraction of the insertions of the total insertions sampled from this study. Whole genome saturation mutagenesis required >33,000 mutants to target all protein coding genes (dark blue line) and this study obtained 38,173 mutants (light blue line). (**D**) A negative binomial model is implemented here for predicting the necessary number of recovered insertions for achieving saturation-level mutagenesis. The total number of *P. falciparum* protein-coding genes was obtained from PlasmoDB v27. The mutagenesis process was modeled as a negative binomial distribution, as a function of the two parameters of total genes and given number of mutants. (**E**) The proportion of *P. falciparum* genome are shown. Exonic regions overall lacked insertions as compared to intergenic regions.

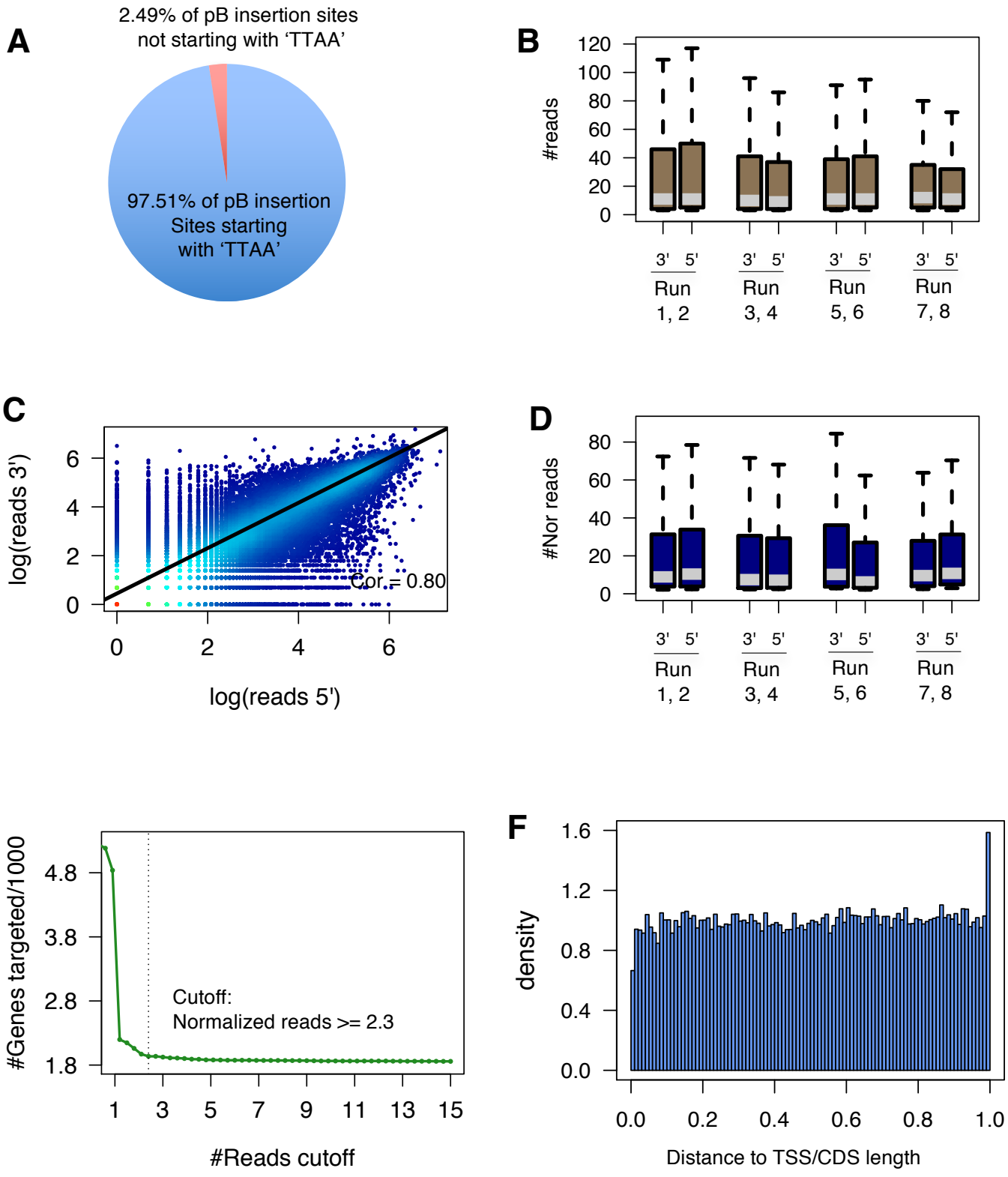

A
2.49% of pB insertion sites not starting with 'TTAA'
97.51% of pB insertion Sites starting with 'TTAA'
B
#reads
120
100
80
60
40
20
0
3' 5'
Run 1, 2
Run 3, 4
Run 5, 6
Run 7, 8
C
log(reads 3')
log(reads 5')
Cor. = 0.80
D
#Nor reads
E
#Genes targeted/1000
4.8
3.8
2.8
1.8
Cutoff:
Normalized reads >= 2.3
#Reads cutoff
F
density
1.6
1.2
0.8
0.4
0.0
Distance to TSS/CDS length

Fig. S2. Computational analyses were used to verify the sequence reads of the QIseq libraries, related to Fig. 1 and table S3.

(**A**) The pie chart shows the majority (> 97%) of *piggyBac* insertion candidates bearing an expected transposon insertion site 'TTAA', and those insertions not having the consensus TTAA (2.49%) were removed from the data set of *piggyBac* insertions. Here, the insertions with reads number ≥ 3 are considered. (**B**) Original sequencing reads number distribution in different QIseq runs. 3′ and 5′ represent reads number of 3′ and 5′ libraries. (**C**) Scatter plot indicates significant reads number correlation between 3' and 5' library. (Pearson's r = 0.8, P-value < 2.2e-16 compared with permutated data.). (**D**) Normalized reads number distribution in different QIseq runs. 3′ and 5′ are defined the same as in B. (**E**) Scatter plot shows the number of genes targeted by putative insertions based on specific reads cutoff. A significant amount of targeted genes show an increase from cutoff 2.3 to 2 (p-value < 0.05), which provides a statistical rationale of our chosen reads signal cutoff threshold. (**F**) Bar plot shows significantly higher density of insertions located on the very end of the gene CDS ($p < 0.05$), and those insertions located in distance <= 0.1 to TSS/CDS of the gene were removed from the data set.

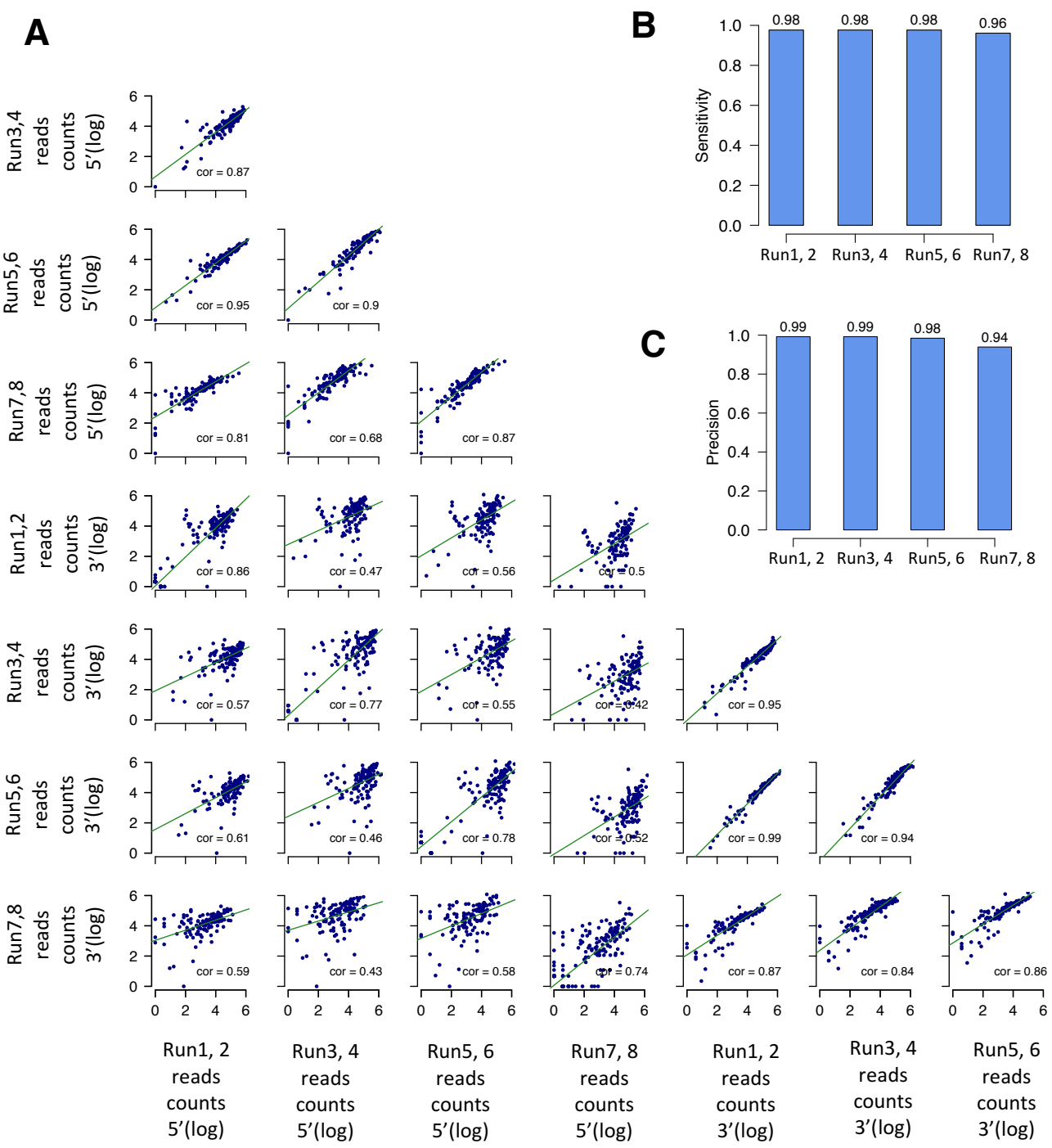
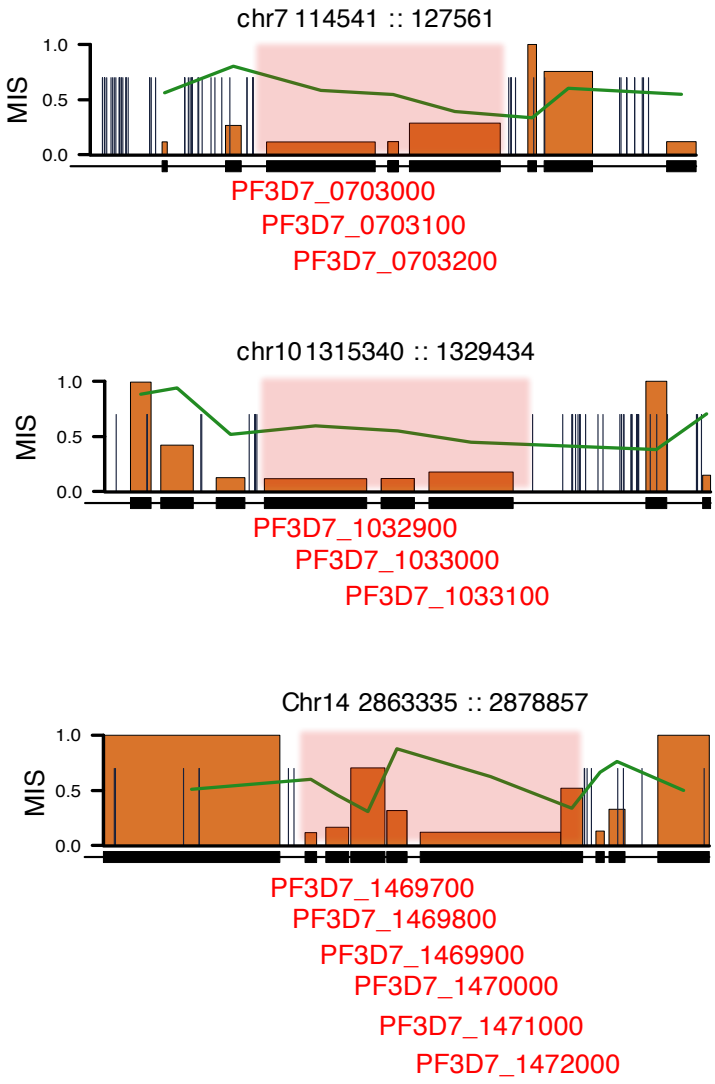

A
B
C
Run3,4 reads counts 5′(log)
Run5,6 reads counts 5′(log)
Run7,8 reads counts 5′(log)
Run1,2 reads counts 3′(log)
Run3,4 reads counts 3′(log)
Run5,6 reads counts 3′(log)
Run7,8 reads counts 3′(log)
cor = 0.87
cor = 0.95
cor = 0.9
cor = 0.81
cor = 0.68
cor = 0.87
cor = 0.86
cor = 0.47
cor = 0.56
cor = 0.5
cor = 0.57
cor = 0.77
cor = 0.55
cor = 0.42
cor = 0.95
cor = 0.61
cor = 0.46
cor = 0.78
cor = 0.52
cor = 0.99
cor = 0.94
cor = 0.59
cor = 0.43
cor = 0.58
cor = 0.74
cor = 0.87
cor = 0.84
cor = 0.86
Run1, 2 reads counts 5′(log)
Run3, 4 reads counts 5′(log)
Run5, 6 reads counts 5′(log)
Run7, 8 reads counts 5′(log)
Run1, 2 reads counts 3′(log)
Run3, 4 reads counts 3′(log)
Run5, 6 reads counts 3′(log)
Sensitivity
0.98
0.98
0.98
0.96
Run1, 2
Run3, 4
Run5, 6
Run7, 8
Precision
0.99
0.99
0.98
0.94
Run1, 2
Run3, 4
Run5, 6
Run7, 8

Fig. S3. Evaluation of QIseq accuracy, sensitivity and depth by QC samples, related to table S2.

(**A**) The QC samples of 128 previously characterized *piggyBac* mutant clones (19) were prepared in advance for testing QIseq accuracy. The matrix of scatter plots shows the reads number with overall high levels of correlation between different sequencing runs. These results also indicate highest correlation between same prime (3′ or 5′) and same runs. **(B, C)** Bar plots show (B) sensitivity and (C) precision of 4 QC samples with the number of false positives and false negatives recovered. The results verify the robustness and high sensitivity of QIseq.

A



PF3D7_0703000
PF3D7_0703100
PF3D7_0703200
conserved *Plasmodium* protein, unknown function

PF3D7_1032900
RNA polymerase II-associated protein 1, putative

PF3D7_1033000
conserved *Plasmodium* protein, unknown function

PF3D7_1033100
S-adenosylmethionine decarboxylase/ornithine decarboxylase

PF3D7_1469800
PF3D7_1469900
PF3D7_1470100
PF3D7_1470200
conserved *Plasmodium* protein, unknown function

PF3D7_1470000
autophagy-related protein 12, putative

PF3D7_1469700 mediator of RNA polymerase II transcription subunit 6, putative

B

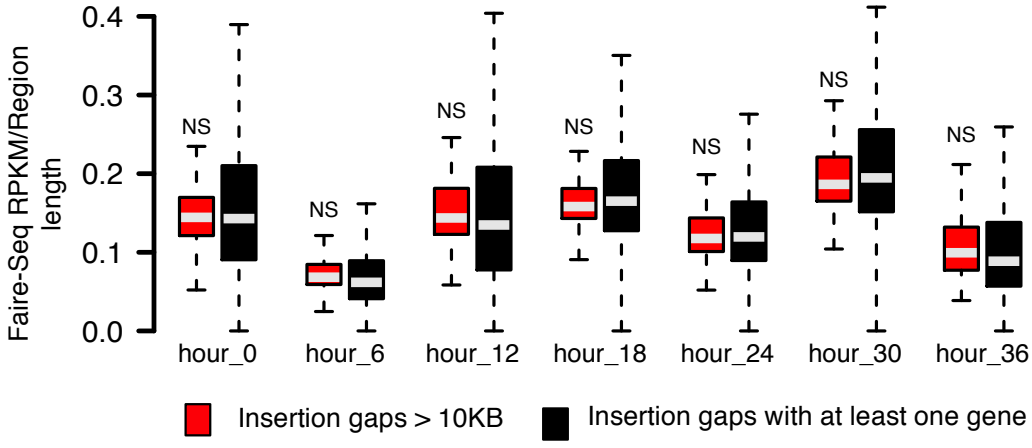


Fig. S4. Some genes were completely devoid of insertions even in the surrounding regions, related to Fig. 1C and table S4.

(**A**) 2.9% genes were completely devoid of insertions even in the surrounding intergenic regions. High-resolution maps of several essential gene cluster regions show ~16 KB essential gene regions. The example region of chromosome 14 depicts a region (~15Kb, 6 genes, pink) without any insertions. Five conserved genes (PF3D7_1469700, PF3D7_1469800, PF3D7_1469900, PF3D7_1470100, PF3D7_1470200) in this region have no functional annotation, while the gene (PF3D71470000) is annotated as putative autophagy-related protein 12. **(B**) The peaks of FAIR-seq indicate open and accessible chromatin status. The results indicate large transposon insertion gaps are not due to the insertion occlusions by the hetrochromatin along the regions. Overlaying Faire-seq signals distribution between the large transposon insertion gaps (length > 10KB) and other insertion gaps with at least one gene ('NS' means non-significant).

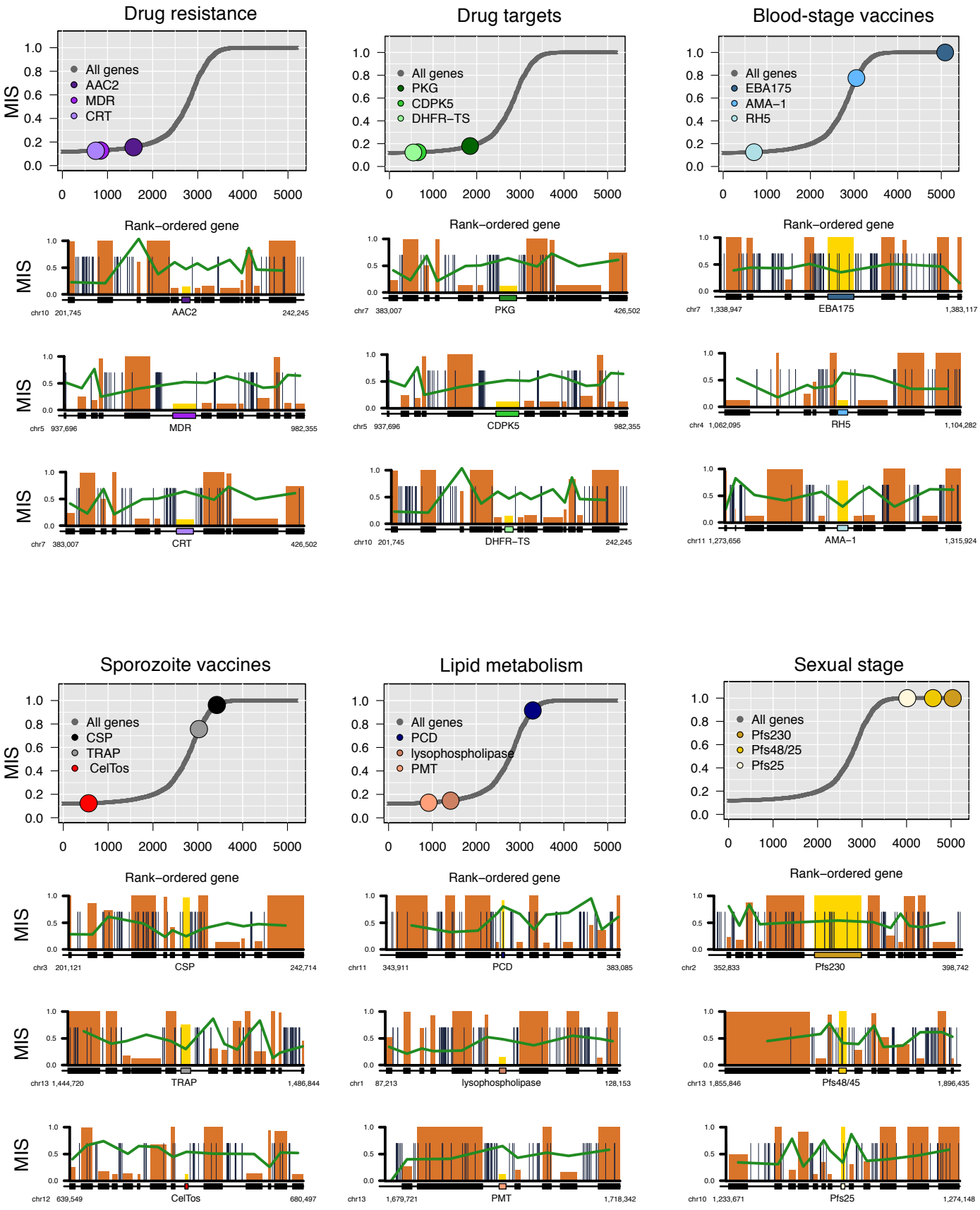

Drug resistance
MIS
All genes
AAC2
MDR
CRT
Rank−ordered gene
chr10 201,745
AAC2
242,245
chr5 937,696
MDR
982,355
chr7 383,007
CRT
426,502
Drug targets
All genes
PKG
CDPK5
DHFR−TS
chr7 383,007
PKG
426,502
chr5 937,696
CDPK5
982,355
chr10 201,745
DHFR−TS
242,245
Blood-stage vaccines
All genes
EBA175
AMA−1
RH5
chr7 1,338,947
EBA175
1,383,117
chr4 1,062,095
RH5
1,104,282
chr11 1,273,656
AMA−1
1,315,924
Sporozoite vaccines
All genes
CSP
TRAP
CelTos
chr3 201,121
CSP
242,714
chr13 1,444,720
TRAP
1,486,844
chr12 639,549
CelTos
680,497
Lipid metabolism
All genes
PCD
lysophospholipase
PMT
chr11 343,911
PCD
383,085
chr1 87,213
lysophospholipase
128,153
chr13 1,679,721
PMT
1,718,342
Sexual stage
All genes
Pfs230
Pfs48/25
Pfs25
chr2 352,833
Pfs230
398,742
chr13 1,855,846
Pfs48/45
1,896,435
chr10 1,233,671
Pfs25
1,274,148

Fig. S5. Identification of dispensable and essential genes through Mutagenesis Index Score (MIS).

MIS plots and high-resolution chromosome maps highlighting important genes of interest for malaria research, drug and vaccine candidate development, and critical biological processes [-20kb, +20kb].

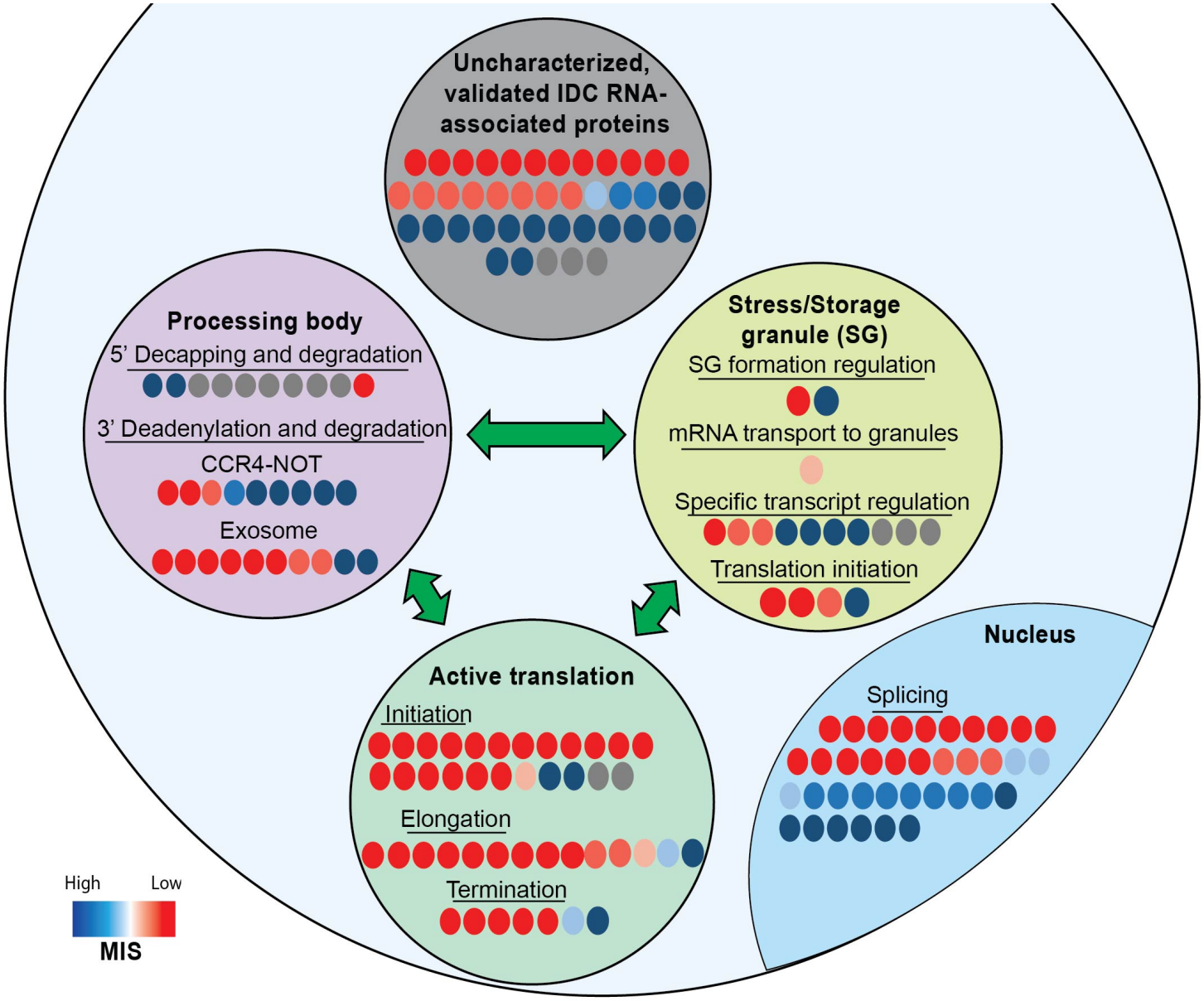

Uncharacterized, validated IDC RNA-associated proteins
Processing body
5' Decapping and degradation
3' Deadenylation and degradation
CCR4-NOT
Exosome
Stress/Storage granule (SG)
SG formation regulation
mRNA transport to granules
Specific transcript regulation
Translation initiation
Active translation
Initiation
Elongation
Termination
Nucleus
Splicing
High
Low
MIS

Fig. S6. Essentiality across select aspects of *P. falciparum* RNA metabolism.
Genes involved in each process are indicated as a circle colored by MIS. Grey circles indicate genes that have 0 gene-body insertions, but have low TTAA density (<7) and thus are not scored. RNA is generally subjected to any of three fates once it has been transcribed and has left the nucleus. RNA can be transported to active centers of translation, where it is immediately transcribed. RNA can be translationally paused, or repressed, and stored for later translation in storage granules (also termed "stress granules" as these granules often form in response to some stress, such as pH or temperature changes that might accompany the parasite's host transition). RNA can be shuttled to another type of RNA granule, the Processing Body (or P-body), for decapping and possible degradation. The composition of these RNA granules has some overlap, and the compartments can also physically associate. Some transcript-specific regulators are represented in the Storage Granules as they have an unspecified, punctate-cytoplasmic localization that may indicate RNA granules during at least some point during the IDC. The post-transcriptional regulatory landscape is highly dynamic and mRNAs may be shuttled between any of these compartments for active translation/storage as protein products are required before finally being degraded. Genes involved at each of these stops along the course of an mRNA's fate are highly likely to be essential, as well as a significant number of functionally uncharacterized proteins confirmed to associate with RNAs during the IDC. See Table S7 for represented genes.

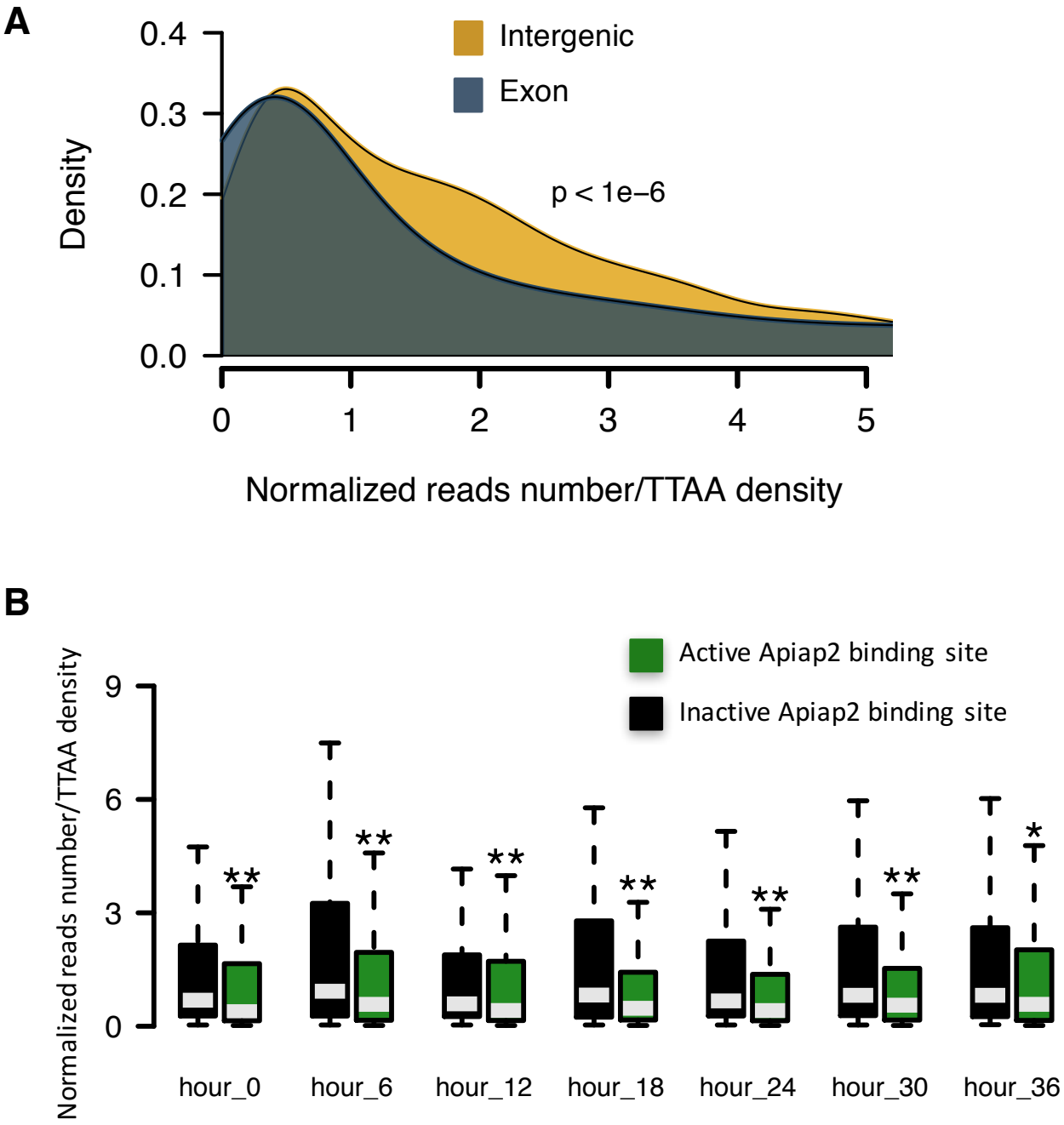

A
Intergenic
Exon
p < 1e−6
Density
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
5
Normalized reads number/TTAA density
B
Active Apiap2 binding site
Inactive Apiap2 binding site
Normalized reads number/TTAA density
9
6
3
0
**
**
**
**
**
**
*
hour_0
hour_6
hour_12
hour_18
hour_24
hour_30
hour_36

Fig. S7. Large scale mutagenesis assay can identify functional region on intergenic regions.

(**A**) Higher MFS of mutants having insertions in intergenic regions compared to MFS of exons indicates less fitness penalty of competitive growth in the intergenic regions. (Wilcoxon test $p < 1e\text{-}6$). **(B)** Boxplot indicates that active ApiAp2 binding sites in intergenic regions exhibit significantly higher MFS. Green boxes indicate ApiAp2 motifs with top quartile Faire-seq signal in each time stage, while black ones indicate motifs with bottom quartile Faire-seq signal. ApiAp2 binding candidates are defined as a possible binding region with a core ApiAp2 motif [±50bp]. ('**' indicates Wilcoxon test $p < 0.01$, '*' indicates $p < 0.05$).

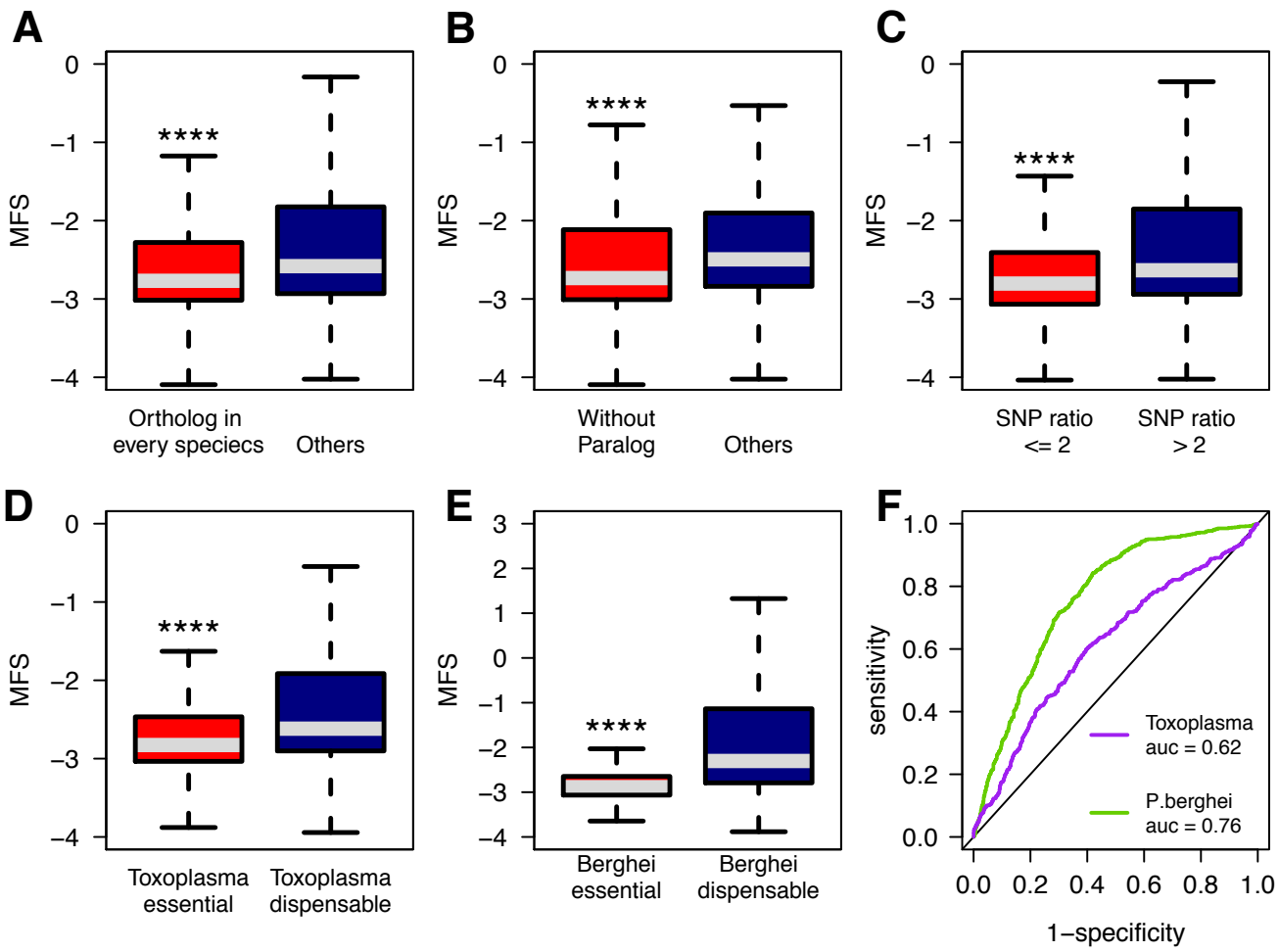
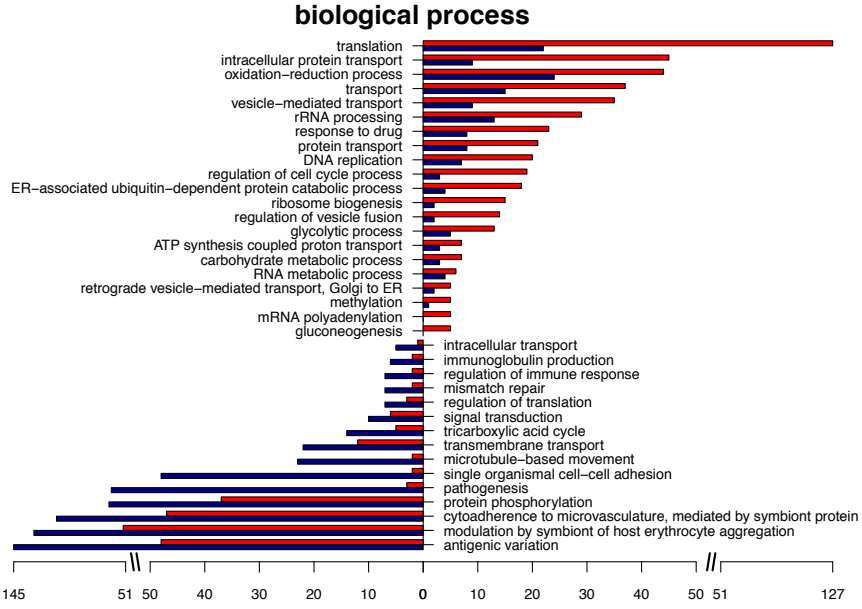

A
B
C
D
E
F
MFS
****
Ortholog in every speciecs
Others
Without Paralog
SNP ratio <= 2
SNP ratio > 2
Toxoplasma essential
Toxoplasma dispensable
Berghei essential
Berghei dispensable
sensitivity
1−specificity
Toxoplasma auc = 0.62
P.berghei auc = 0.76

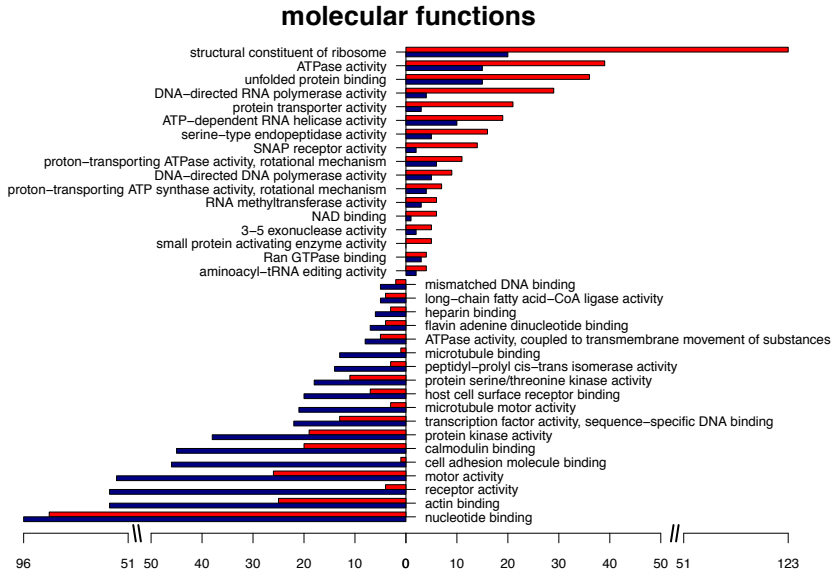
Fig. S8. Distinct biological process and evolutionarily conservation segregate the tendency of dispensable and essential genes by MFS.

(**A)** Evolutionarily conserved genes with significantly lower MFS enriched in the gene set retained 1:1 ortholog counts; (**B**) have very few paralogs; and (**C**) show reduced non-synonymous to synonymous single nucleotide polymorphism rates. Bars indicate the group median ('****' indicates Wilcoxon P < 2.2e-16). The genes enriched with low MIS are more likely to be essential in multiple Plasmodium species. **(D, E)** The orthologs of essential genes reported in (D) *Toxoplasma* and (E) *P. berghei* showed significantly lower MFS in PB-QIseq mutagenesis screen of *P. falciparum* ('****' indicates Wilcoxon P < 2.2e-16). (**F)** Two Receiver operating characteristic (ROC) curve which shows the level of retention of essential genes across species. The area under the ROC curve (AUC) indicates that MIS from the *P. falciparum piggyBac*-QIseq mutagenesis screen has stronger correlation with the deduced *P. berghei* genome essentiality than that of *Toxoplasma*.

A


biological process
translation
intracellular protein transport
oxidation−reduction process
transport
vesicle−mediated transport
rRNA processing
response to drug
protein transport
DNA replication
regulation of cell cycle process
ER−associated ubiquitin−dependent protein catabolic process
ribosome biogenesis
regulation of vesicle fusion
glycolytic process
ATP synthesis coupled proton transport
carbohydrate metabolic process
RNA metabolic process
retrograde vesicle−mediated transport, Golgi to ER
methylation
mRNA polyadenylation
gluconeogenesis
intracellular transport
immunoglobulin production
regulation of immune response
mismatch repair
regulation of translation
signal transduction
tricarboxylic acid cycle
transmembrane transport
microtubule−based movement
single organismal cell−cell adhesion
pathogenesis
protein phosphorylation
cytoadherence to microvasculature, mediated by symbiont protein
modulation by symbiont of host erythrocyte aggregation
antigenic variation
145
51
50
40
30
20
10
0
10
20
30
40
50
51
127

B


molecular functions
structural constituent of ribosome
ATPase activity
unfolded protein binding
DNA−directed RNA polymerase activity
protein transporter activity
ATP−dependent RNA helicase activity
serine−type endopeptidase activity
SNAP receptor activity
proton−transporting ATPase activity, rotational mechanism
DNA−directed DNA polymerase activity
proton−transporting ATP synthase activity, rotational mechanism
RNA methyltransferase activity
NAD binding
3−5 exonuclease activity
small protein activating enzyme activity
Ran GTPase binding
aminoacyl−tRNA editing activity
mismatched DNA binding
long−chain fatty acid−CoA ligase activity
heparin binding
flavin adenine dinucleotide binding
ATPase activity, coupled to transmembrane movement of substances
microtubule binding
peptidyl−prolyl cis−trans isomerase activity
protein serine/threonine kinase activity
host cell surface receptor binding
microtubule motor activity
transcription factor activity, sequence−specific DNA binding
protein kinase activity
calmodulin binding
cell adhesion molecule binding
motor activity
receptor activity
actin binding
nucleotide binding
96
51
50
40
30
20
10
0
10
20
30
40
50
51
123

C

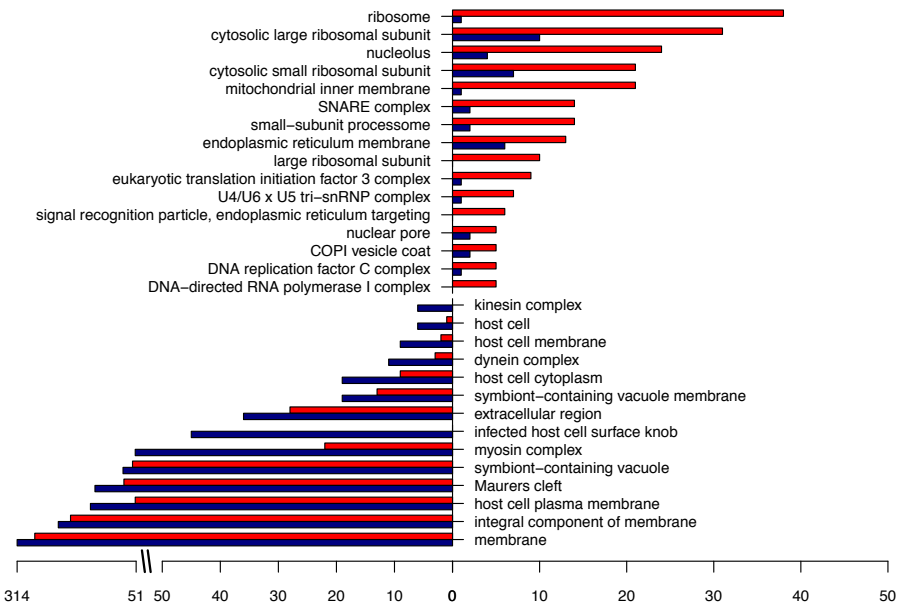

Cellular components
ribosome
cytosolic large ribosomal subunit
nucleolus
cytosolic small ribosomal subunit
mitochondrial inner membrane
SNARE complex
small−subunit processome
endoplasmic reticulum membrane
large ribosomal subunit
eukaryotic translation initiation factor 3 complex
U4/U6 x U5 tri−snRNP complex
signal recognition particle, endoplasmic reticulum targeting
nuclear pore
COPI vesicle coat
DNA replication factor C complex
DNA−directed RNA polymerase I complex
kinesin complex
host cell
host cell membrane
dynein complex
host cell cytoplasm
symbiont−containing vacuole membrane
extracellular region
infected host cell surface knob
myosin complex
symbiont−containing vacuole
Maurers cleft
host cell plasma membrane
integral component of membrane
membrane
314
51
50
40
30
20
10
0
10
20
30
40
50

Fig. S9. Gene numbers associated with individual GO category biological process, molecular functions, and cellular components.

(**A**) Biological processes. (**B**) Molecular functions. (**C**) Cellular components.

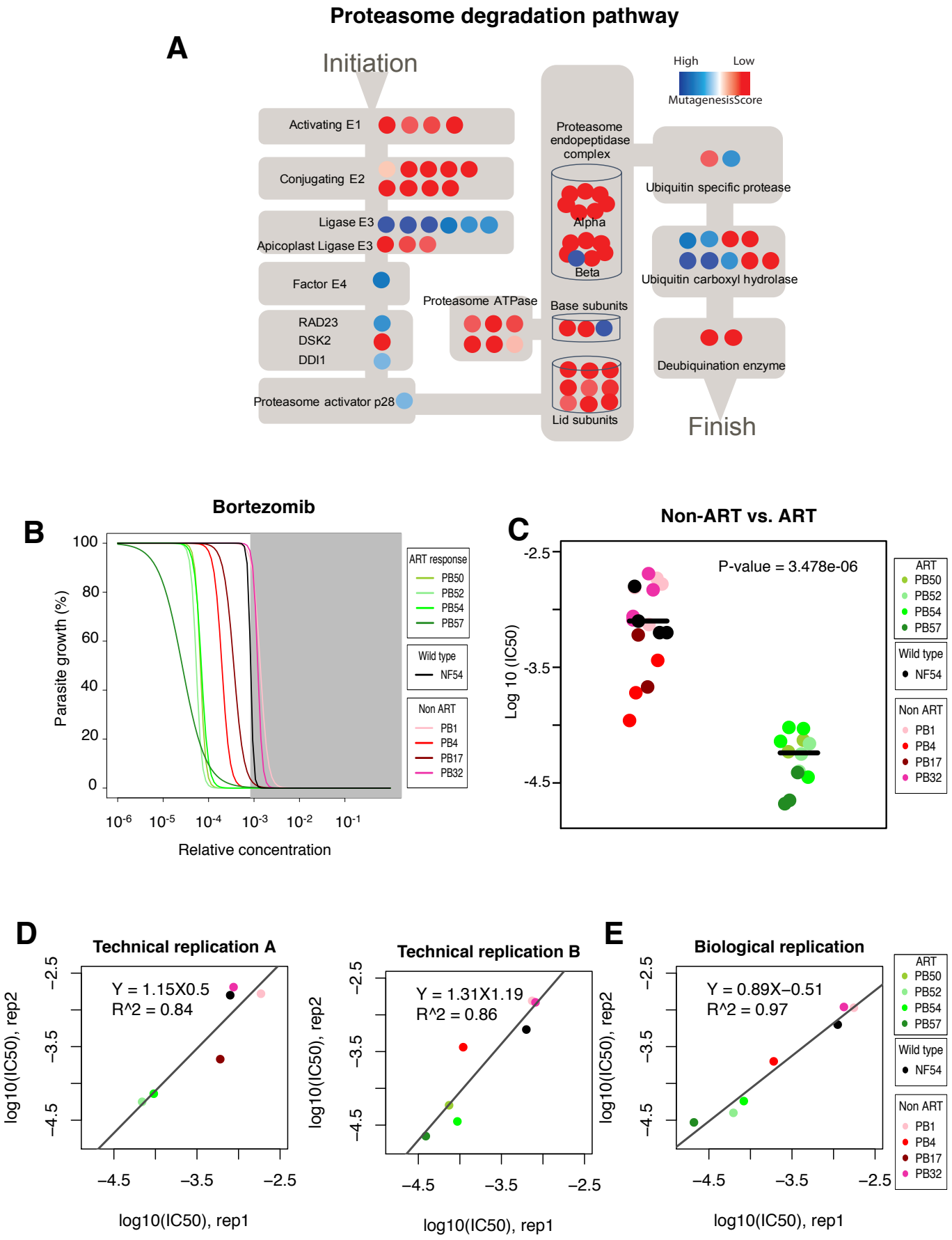
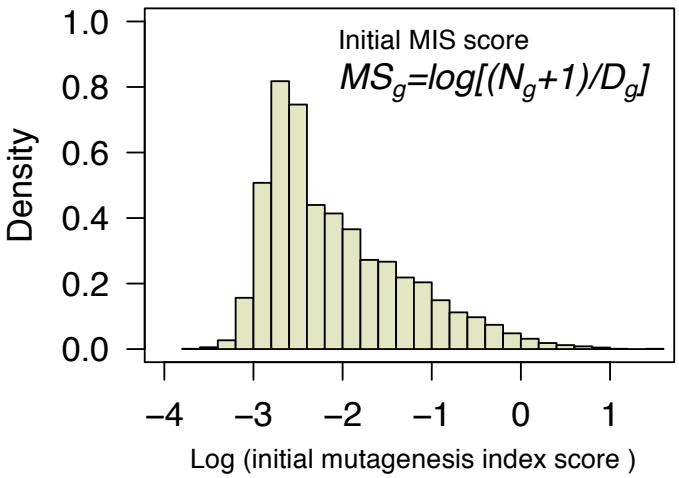

Proteasome degradation pathway
A
Initiation
Activating E1
Conjugating E2
Ligase E3
Apicoplast Ligase E3
Factor E4
RAD23
DSK2
DDI1
Proteasome activator p28
Proteasome ATPase
Proteasome endopeptidase complex
Alpha
Beta
Base subunits
Lid subunits
High
Low
MutagenesisScore
Ubiquitin specific protease
Ubiquitin carboxyl hydrolase
Deubiquination enzyme
Finish
B
Bortezomib
Parasite growth (%)
Relative concentration
ART response
PB50
PB52
PB54
PB57
Wild type
NF54
Non ART
PB1
PB4
PB17
PB32
C
Non-ART vs. ART
P-value = 3.478e-06
Log 10 (IC50)
D
Technical replication A
Y = 1.15X0.5
R^2 = 0.84
log10(IC50), rep2
log10(IC50), rep1
Technical replication B
Y = 1.31X1.19
R^2 = 0.86
E
Biological replication
Y = 0.89X−0.51
R^2 = 0.97

Fig. S10. Chemogenomic profiling of *piggyBac* mutant clones of the ART sensitivity cluster, related to Fig. 2.

(**A**) The proteasome-mediated degradation pathway shows a high level of essentiality. MIS is plotted as a heatmap of red (Low MIS essential) to blue (High MIS, dispensable). Most of the genes related in this pathway are characterized as essential (red). (**B**) Chemogenomic profiles of *piggyBac* mutants to the proteasome inhibitor Bortezomib. Mutants of the ART sensitivity cluster (16) also have increased sensitivity to Bortezomib compared to mutants with chemogenomic profiles negatively correlated with ART mechanism of action (40). (**C**) ART response cluster shows significantly lower IC50 (higher sensitivity) to Bortezomib. Phenotype reproducibility of chemogenomic profiles was estimated by technical and biological replicates. (**D**) Technical replicates correlation plot for essay A and B. (**E**) Biological replicates A and B correlation plot.



Bayesian mixture of Gaussians

$$(MS_g) = \sum_{\pi_g=1,0} P(\pi_g)N(MS_g|\mu_{\pi_g}, \sigma_{\pi_g})$$

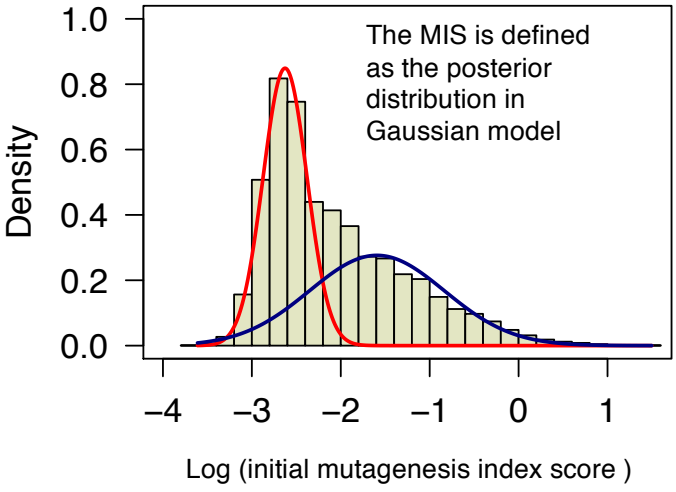


Fig. S11. Mutagenesis Index Score MIS development.

The initial mutagenesis score (#insertions/Density of TTAA) of the *P. falciparum* genes can be characterized by two posterior Gausssian models, which belong to essential and dispensable groups.

Supplemental Tables Caption:

Table S1. *piggyBac* insertion sites identified from preliminary study.

Quantitative Insertion-site Sequencing (QIseq) identified 3651 *piggyBac* insertion sites distributed across all fourteen chromosomes, combined with 326 previous published mutagenesis results shown on this table.

Table S2. QC samples in each of QIseq run.

128 pooled mutants with known *pB* insertion sites identified in previous study were prepared in advance as QC samples for monitoring accuracy and depth of each QIseq run.

Table S3. *piggyBac* insertion sites identified from this study.

34,522 *piggyBac* insertion sites identified from this study. A total of 38,173 *piggyBac* insertion sites were used for determining saturation-level mutagenesis scores (MIS and MFS) including preliminary study data (Table S1).

Table S4: The non-mutable genes located in essential blocks.

We observed that some non-disrupted genes were completely devoid of insertions even in the surrounding intergenic regions. 143 non-mutable genes (representing 2.9% of genes) are located in essential blocks, which are defined as having an insertion-free gap > 10 kb.

Table S5: MIS and MFS identify essentiality of the genes.

Mutagenesis index score (MIS) was calculated based on the susceptibility of the ORF in each transcriptional unit to being disrupted (Fig. 2A). Mutagenesis Fitness Score (MFS) which is calculated by the normalized reads number (Fig. 2E). Two measuring models (MIS and MFS) were used to validate essentiality of the genes of *Plasmodium falciparum* genome.

Table S6: RNA metabolism, related to Fig. 2 C, D and fig. S6.

RNA metabolism genes by compartment or process as displayed in figure S7. MIS is indicated from 0 to 1. Genes in red text indicate those with TTAA density <7 or length <400bp (that are thus not scored) that have 0 recovered gene-body insertions. Translation and splicing-related genes of interest were identified via GO term. RNA granule-related genes of interest are as classified in Reddy et al. 2015 and as reviewed (41). Unknown, validated RNA-associated genes are proteins identified as being mRNA-bound as per (42) that have no or little functional annotation as per PlasmoDB.

Table S7: Phenotype screen of competitive growth assay, related to Fig. 3 A to C.

Phenotype screen of competitive growth assay identifies growth winners and losers from ~400 mutants in four individual pools which provides evidence to validate MIS and MFS scoring of each gene's essentiality.

Table S8: Genes corresponding to GO enrichment.

Functional annotation of biological processes, molecular function and cellular component are represented by the p-value and the fraction of the genes with MIS > 0.5. Each GO term is assigned a p-value to represent the tendency to be essential or dispensable.

Table S9: Sample pool ID and accession number for this study, related to Fig. 1, 2 and table S3.

Sample pool ID and accession number from all QIseq runs are shown on this table. Insertion sites identified from each pool are shown on table S3.

References and Notes

1. E. A. Ashley, M. Dhorda, R. M. Fairhurst, C. Amaratunga, P. Lim, S. Suon, S. Sreng, J. M. Anderson, S. Mao, B. Sam, C. Sopha, C. M. Chuor, C. Nguon, S. Sovannaroth, S. Pukrittayakamee, P. Jittamala, K. Chotivanich, K. Chutasmit, C. Suchatsoonthorn, R. Runcharoen, T. T. Hien, N. T. Thuy-Nhien, N. V. Thanh, N. H. Phu, Y. Htut, K.-T. Han, K. H. Aye, O. A. Mokuolu, R. R. Olaosebikan, O. O. Folaranmi, M. Mayxay, M. Khanthavong, B. Hongvanthong, P. N. Newton, M. A. Onyamboko, C. I. Fanello, A. K. Tshefu, N. Mishra, N. Valecha, A. P. Phyo, F. Nosten, P. Yi, R. Tripura, S. Borrmann, M. Bashraheil, J. Peshu, M. A. Faiz, A. Ghose, M. A. Hossain, R. Samad, M. R. Rahman, M. M. Hasan, A. Islam, O. Miotto, R. Amato, B. MacInnis, J. Stalker, D. P. Kwiatkowski, Z. Bozdech, A. Jeeyapant, P. Y. Cheah, T. Sakulthaew, J. Chalk, B. Intharabut, K. Silamut, S. J. Lee, B. Vihokhern, C. Kunasol, M. Imwong, J. Tarning, W. J. Taylor, S. Yeung, C. J. Woodrow, J. A. Flegg, D. Das, J. Smith, M. Venkatesan, C. V. Plowe, K. Stepniewska, P. J. Guerin, A. M. Dondorp, N. P. Day, N. J. White, Tracking Resistance to Artemisinin Collaboration (TRAC), Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* **371**, 411–423 (2014). doi:10.1056/NEJMoa1314981 Medline

2. C. J. Woodrow, N. J. White, The clinical impact of artemisinin resistance in Southeast Asia and the potential for future spread. *FEMS Microbiol. Rev.* **41**, 34–48 (2017). doi:10.1093/femsre/fuw037 Medline

3. M. J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. A. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman, C. Newbold, R. W. Davis, C. M. Fraser, B. Barrell, Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002). doi:10.1038/nature01097 Medline

4. M. Ghorbal, M. Gorman, C. R. Macpherson, R. M. Martins, A. Scherf, J.-J. Lopez-Rubio, Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat. Biotechnol.* **32**, 819–821 (2014). doi:10.1038/nbt.2925 Medline

5. J. C. Wagner, R. J. Platt, S. J. Goldfless, F. Zhang, J. C. Niles, Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*. *Nat. Methods* **11**, 915–918 (2014). doi:10.1038/nmeth.3063 Medline

6. S. M. Sidik, D. Huet, S. M. Ganesan, M.-H. Huynh, T. Wang, A. S. Nasamu, P. Thiru, J. P. J. Saeij, V. B. Carruthers, J. C. Niles, S. Lourido, A Genome-wide CRISPR Screen in toxoplasma identifies essential apicomplexan genes. *Cell* **166**, 1423–1435.e12 (2016). doi:10.1016/j.cell.2016.08.019 Medline

7. T. F. de Koning-Ward, P. R. Gilson, B. S. Crabb, Advances in molecular genetic systems in malaria. *Nat. Rev. Microbiol.* **13**, 373–387 (2015). doi:10.1038/nrmicro3450 Medline

8. S. T. Thibault, M. A. Singer, W. Y. Miyazaki, B. Milash, N. A. Dompe, C. M. Singh, R. Buchholz, M. Demsky, R. Fawcett, H. L. Francis-Lang, L. Ryner, L. M. Cheung, A. Chong, C. Erickson, W. W. Fisher, K. Greer, S. R. Hartouni, E. Howie, L. Jakkula, D. Joo, K. Killpack, A. Laufer, J. Mazzotta, R. D. Smith, L. M. Stevens, C. Stuber, L. R. Tan, R. Ventura, A. Woo, I. Zakrajsek, L. Zhao, F. Chen, C. Swimmer, C. Kopczynski, G. Duyk, M. L. Winberg, J. Margolis, A complementary transposon tool kit for *Drosophila melanogaster* using P and piggyBac. *Nat. Genet.* **36**, 283–287 (2004). doi:10.1038/ng1314 Medline

9. M. J. Fraser, G. E. Smith, M. D. Summers, Acquisition of host cell DNA sequences by baculoviruses: Relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. *J. Virol.* **47**, 287–300 (1983). Medline

10. M. J. Fraser, J. S. Brusca, G. E. Smith, M. D. Summers, Transposon-mediated mutagenesis of a baculovirus. *Virology* **145**, 356–361 (1985). doi:10.1016/0042-6822(85)90172-2 Medline

11. L. C. Cary, M. Goebel, B. G. Corsaro, H.-G. Wang, E. Rosen, M. J. Fraser, Transposon mutagenesis of baculoviruses: Analysis of *Trichoplusia* ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology* **172**, 156–169 (1989). doi:10.1016/0042-6822(89)90117-7 Medline

12. B. Balu, D. A. Shoue, M. J. Fraser Jr., J. H. Adams, High-efficiency transformation of Plasmodium falciparum by the lepidopteran transposable element piggyBac. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16391–16396 (2005). doi:10.1073/pnas.0504679102 Medline

13. B. Balu, S. P. Maher, A. Pance, C. Chauhan, A. V. Naumov, R. M. Andrews, P. D. Ellis, S. M. Khan, J. W. Lin, C. J. Janse, J. C. Rayner, J. H. Adams, CCR4-associated factor 1 coordinates the expression of *Plasmodium falciparum* egress and invasion proteins. *Eukaryot. Cell* **10**, 1257–1263 (2011). doi:10.1128/EC.05099-11 Medline

14. H. Ikadai, K. Shaw Saliba, S. M. Kanzok, K. J. McLean, T. Q. Tanaka, J. Cao, K. C. Williamson, M. Jacobs-Lorena, Transposon mutagenesis identifies genes essential for *Plasmodium falciparum* gametocytogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E1676–E1684 (2013). doi:10.1073/pnas.1217712110 Medline

15. B. Balu, C. Campbell, J. Sedillo, S. Maher, N. Singh, P. Thomas, M. Zhang, A. Pance, T. D. Otto, J. C. Rayner, J. H. Adams, Atypical mitogen-activated protein kinase phosphatase implicated in regulating transition from pre-S-Phase asexual intraerythrocytic development of *Plasmodium falciparum*. *Eukaryot. Cell* **12**, 1171–1178 (2013). doi:10.1128/EC.00028-13 Medline

16. A. Pradhan, G. H. Siwo, N. Singh, B. Martens, B. Balu, K. A. Button-Simons, A. Tan, M. Zhang, K. O. Udenze, R. H. Y. Jiang, M. T. Ferdig, J. H. Adams, D. E. Kyle, Chemogenomic profiling of Plasmodium falciparum as a tool to aid antimalarial drug discovery. *Sci. Rep.* **5**, 15930 (2015). doi:10.1038/srep15930 Medline

17. I. F. Bronner, T. D. Otto, M. Zhang, K. Udenze, C. Wang, M. A. Quail, R. H. Y. Jiang, J. H. Adams, J. C. Rayner, Quantitative insertion-site sequencing (QIseq) for high throughput phenotyping of transposon mutants. *Genome Res.* **26**, 980–989 (2016). doi:10.1101/gr.200279.115 Medline

18. T. Hart, M. Chandrashekhar, M. Aregger, Z. Steinhart, K. R. Brown, G. MacLeod, M. Mis, M. Zimmermann, A. Fradet-Turcotte, S. Sun, P. Mero, P. Dirks, S. Sidhu, F. P. Roth, O. S. Rissland, D. Durocher, S. Angers, J. Moffat, High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515–1526 (2015). doi:10.1016/j.cell.2015.11.015 Medline

19. B. Balu, C. Chauhan, S. P. Maher, D. A. Shoue, J. C. Kissinger, M. J. Fraser Jr., J. H. Adams, piggyBac is an effective tool for functional analysis of the *Plasmodium falciparum* genome. *BMC Microbiol.* **9**, 83 (2009). doi:10.1186/1471-2180-9-83 Medline

20. Q. Zhang, T. N. Siegel, R. M. Martins, F. Wang, J. Cao, Q. Gao, X. Cheng, L. Jiang, C.-C. Hon, C. Scheidig-Benatar, H. Sakamoto, L. Turner, A. T. R. Jensen, A. Claes, J. Guizetti, N. A. Malmquist, A. Scherf, Exonuclease-mediated degradation of nascent RNA silences genes linked to severe malaria. *Nature* **513**, 431–435 (2014). doi:10.1038/nature13468 Medline

21. R. Ménard, A. A. Sultan, C. Cortes, R. Altszuler, M. R. van Dijk, C. J. Janse, A. P. Waters, R. S. Nussenzweig, V. Nussenzweig, Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. *Nature* **385**, 336–340 (1997). doi:10.1038/385336a0 Medline

22. A. M. Bobenchik, W. H. Witola, Y. Augagneur, L. Nic Lochlainn, A. Garg, N. Pachikara, J.-Y. Choi, Y. O. Zhao, S. Usmani-Brown, A. Lee, S. H. Adjalley, S. Samanta, D. A. Fidock, D. R. Voelker, E. Fikrig, C. Ben Mamoun, *Plasmodium falciparum* phosphoethanolamine methyltransferase is essential for malaria transmission. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18262–18267 (2013). doi:10.1073/pnas.1313965110 Medline

23. N. M. B. Brancucci, J. P. Gerdt, C. Wang, M. De Niz, N. Philip, S. R. Adapa, M. Zhang, E. Hitz, I. Niederwieser, S. D. Boltryk, M.-C. Laffitte, M. A. Clark, C. Grüring, D. Ravel, A. Blancke Soares, A. Demas, S. Bopp, B. Rubio-Ruiz, A. Conejo-Garcia, D. F. Wirth, E. Gendaszewska-Darmach, M. T. Duraisingh, J. H. Adams, T. S. Voss, A. P. Waters, R. H. Y. Jiang, J. Clardy, M. Marti, Lysophosphatidylcholine regulates sexual stage differentiation in the human malaria parasite *Plasmodium falciparum*. *Cell* **171**, 1532–1544.e15 (2017). doi:10.1016/j.cell.2017.10.020 Medline

24. T. W. Kooij, J. M. Carlton, S. L. Bidwell, N. Hall, J. Ramesar, C. J. Janse, A. P. Waters, A Plasmodium whole-genome synteny map: Indels and synteny breakpoints as foci for species-specific genes. *PLOS Pathog.* **1**, e44 (2005). doi:10.1371/journal.ppat.0010044 Medline

25. J. D. DeBarry, J. C. Kissinger, Jumbled genomes: Missing Apicomplexan synteny. *Mol. Biol. Evol.* **28**, 2855–2871 (2011). doi:10.1093/molbev/msr103 Medline

26. T. D. Otto, D. Wilinski, S. Assefa, T. M. Keane, L. R. Sarry, U. Böhme, J. Lemieux, B. Barrell, A. Pain, M. Berriman, C. Newbold, M. Llinás, New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol. Microbiol.* **76**, 12–24 (2010). doi:10.1111/j.1365-2958.2009.07026.x Medline

27. V. A. Blomen, P. Májek, L. T. Jae, J. W. Bigenzahn, J. Nieuwenhuis, J. Staring, R. Sacco, F. R. van Diemen, N. Olk, A. Stukalov, C. Marceau, H. Janssen, J. E. Carette, K. L. Bennett, J. Colinge, G. Superti-Furga, T. R. Brummelkamp, Gene essentiality and

synthetic lethality in haploid human cells. *Science* **350**, 1092–1096 (2015). doi:10.1126/science.aac7557 Medline

28. T. Wang, K. Birsoy, N. W. Hughes, K. M. Krupczak, Y. Post, J. J. Wei, E. S. Lander, D. M. Sabatini, Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015). doi:10.1126/science.aac7041 Medline

29. E. Bushell, A. R. Gomes, T. Sanderson, B. Anar, G. Girling, C. Herd, T. Metcalf, K. Modrzynska, F. Schwach, R. E. Martin, M. W. Mather, G. I. McFadden, L. Parts, G. G. Rutledge, A. B. Vaidya, K. Wengelnik, J. C. Rayner, O. Billker, Functional profiling of a *Plasmodium* genome reveals an abundance of essential genes. *Cell* **170**, 260–272.e8 (2017). doi:10.1016/j.cell.2017.06.030 Medline

30. S. S. Vembar, D. Droll, A. Scherf, Translational regulation in blood stages of the malaria parasite *Plasmodium* spp.: Systems-wide studies pave the way. *Wiley Interdiscip. Rev. RNA* **7**, 772–792 (2016). doi:10.1002/wrna.1365 Medline

31. C. Dogovski, S. C. Xie, G. Burgio, J. Bridgford, S. Mok, J. M. McCaw, K. Chotivanich, S. Kenny, N. Gnädig, J. Straimer, Z. Bozdech, D. A. Fidock, J. A. Simpson, A. M. Dondorp, S. Foote, N. Klonis, L. Tilley, Targeting the cell stress response of Plasmodium falciparum to overcome artemisinin resistance. *PLOS Biol.* **13**, e1002132 (2015). doi:10.1371/journal.pbio.1002132 Medline

32. A. Mbengue, S. Bhattacharjee, T. Pandharkar, H. Liu, G. Estiu, R. V. Stahelin, S. S. Rizk, D. L. Njimoh, Y. Ryan, K. Chotivanich, C. Nguon, M. Ghorbal, J.-J. Lopez-Rubio, M. Pfrender, S. Emrich, N. Mohandas, A. M. Dondorp, O. Wiest, K. Haldar, A molecular mechanism of artemisinin resistance in *Plasmodium falciparum* malaria. *Nature* **520**, 683–687 (2015). doi:10.1038/nature14412 Medline

33. H. Li, A. J. O'Donoghue, W. A. van der Linden, S. C. Xie, E. Yoo, I. T. Foe, L. Tilley, C. S. Craik, P. C. A. da Fonseca, M. Bogyo, Structure- and function-based design of *Plasmodium*-selective proteasome inhibitors. *Nature* **530**, 233–236 (2016). doi:10.1038/nature16936 Medline

34. W. C. Van Voorhis, J. H. Adams, R. Adelfio, V. Ahyong, M. H. Akabas, P. Alano, A. Alday, Y. Alemán Resto, A. Alsibaee, A. Alzualde, K. T. Andrews, S. V. Avery, V. M. Avery, L. Ayong, M. Baker, S. Baker, C. Ben Mamoun, S. Bhatia, Q. Bickle, L. Bounaadja, T. Bowling, J. Bosch, L. E. Boucher, F. F. Boyom, J. Brea, M. Brennan, A. Burton, C. R. Caffrey, G. Camarda, M. Carrasquilla, D. Carter, M. Belen Cassera, K. Chih-Chien Cheng, W. Chindaudomsate, A. Chubb, B. L. Colon, D. D. Colón-López, Y. Corbett, G. J. Crowther, N. Cowan, S. D'Alessandro, N. Le Dang, M. Delves, J. L. DeRisi, A. Y. Du, S. Duffy, S. Abd El-Salam El-Sayed, M. T. Ferdig, J. A. Fernández Robledo, D. A. Fidock, I. Florent, P. V. T. Fokou, A. Galstian, F. J. Gamo, S. Gokool, B. Gold, T. Golub, G. M. Goldgof, R. Guha, W. A. Guiguemde, N. Gural, R. K. Guy, M. A. E. Hansen, K. K. Hanson, A. Hemphill, R. Hooft van Huijsduijnen, T. Horii, P. Horrocks, T. B. Hughes, C. Huston, I. Igarashi, K. Ingram-Sieber, M. A. Itoe, A. Jadhav, A. Naranuntarat Jensen, L. T. Jensen, R. H. Y. Jiang, A. Kaiser, J. Keiser, T. Ketas, S. Kicka, S. Kim, K. Kirk, V. P. Kumar, D. E. Kyle, M. J. Lafuente, S. Landfear, N. Lee, S. Lee, A. M. Lehane, F. Li, D. Little, L. Liu, M. Llinás, M. I. Loza, A. Lubar, L. Lucantoni, I. Lucet, L. Maes, D. Mancama, N. R. Mansour, S. March, S. McGowan, I. Medina Vera, S. Meister, L.

Mercer, J. Mestres, A. N. Mfopa, R. N. Misra, S. Moon, J. P. Moore, F. Morais Rodrigues da Costa, J. Müller, A. Muriana, S. Nakazawa Hewitt, B. Nare, C. Nathan, N. Narraidoo, S. Nawaratna, K. K. Ojo, D. Ortiz, G. Panic, G. Papadatos, S. Parapini, K. Patra, N. Pham, S. Prats, D. M. Plouffe, S.-A. Poulsen, A. Pradhan, C. Quevedo, R. J. Quinn, C. A. Rice, M. Abdo Rizk, A. Ruecker, R. St Onge, R. Salgado Ferreira, J. Samra, N. G. Robinett, U. Schlecht, M. Schmitt, F. Silva Villela, F. Silvestrini, R. Sinden, D. A. Smith, T. Soldati, A. Spitzmüller, S. M. Stamm, D. J. Sullivan, W. Sullivan, S. Suresh, B. M. Suzuki, Y. Suzuki, S. J. Swamidass, D. Taramelli, L. R. Y. Tchokouaha, A. Theron, D. Thomas, K. F. Tonissen, S. Townson, A. K. Tripathi, V. Trofimov, K. O. Udenze, I. Ullah, C. Vallieres, E. Vigil, J. M. Vinetz, P. Voong Vinh, H. Vu, N. A. Watanabe, K. Weatherby, P. M. White, A. F. Wilks, E. A. Winzeler, E. Wojcik, M. Wree, W. Wu, N. Yokoyama, P. H. A. Zollo, N. Abla, B. Blasco, J. Burrows, B. Laleu, D. Leroy, T. Spangenberg, T. Wells, P. A. Willis, Open source drug discovery with the malaria box compound collection for neglected diseases and beyond. *PLOS Pathog.* **12**, e1005763 (2016). doi:10.1371/journal.ppat.1005763 Medline

35. S. P. Maher, M. Zhang, B. Balu, J. H. Adams, in *Methods in Malaria Research*, K. Moll, A. Kaneko, A. Scherf, M. Wahlgren, Eds. (EVIMalaR, MR4/BEI Resources, 2013), chap. VII, pp. 391–396.

36. C. Aurrecoechea, A. Barreto, J. Brestelli, B. P. Brunk, S. Cade, R. Doherty, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O. S. Harb, M. Heiges, S. Hu, J. Iodice, J. C. Kissinger, E. T. Kraemer, W. Li, D. F. Pinney, B. Pitts, D. S. Roos, G. Srinivasamoorthy, C. J. Stoeckert Jr., H. Wang, S. Warrenfeltz, EuPathDB: The eukaryotic pathogen database. *Nucleic Acids Res.* **41** (D1), D684–D691 (2013). doi:10.1093/nar/gks1113 Medline

37. M. J. López-Barragán, J. Lemieux, M. Quiñones, K. C. Williamson, A. Molina-Cruz, K. Cui, C. Barillas-Mury, K. Zhao, X. Z. Su, Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum*. *BMC Genomics* **12**, 587 (2011). doi:10.1186/1471-2164-12-587 Medline

38. N. Ponts, E. Y. Harris, J. Prudhomme, I. Wick, C. Eckhardt-Ludka, G. R. Hicks, G. Hardiman, S. Lonardi, K. G. Le Roch, Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.* **20**, 228–238 (2010). doi:10.1101/gr.101063.109 Medline

39. T. L. Campbell, E. K. De Silva, K. L. Olszewski, O. Elemento, M. Llinás, Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLOS Pathog.* **6**, e1001165 (2010). doi:10.1371/journal.ppat.1001165 Medline

41. E. M. Bunnik, G. Batugedara, A. Saraf, J. Prudhomme, L. Florens, K. G. Le Roch, The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. *Genome Biol.* **17**, 147 (2016). doi:10.1186/s13059-016-1014-0 Medline