Supporting Information

Carballar-Lejarazú et al. 10.1073/pnas.1304722110

SI Materials and Methods

Mosquito Rearing and Maintenance. A colony of *An. stephensi* (gift of M. Jacobs-Lorena, John Hopkins University, Baltimore) bred in our insectary for >5 y was used in the experiments. The mosquitoes were maintained at 26 °C ± 1 °C with 77% humidity and 12 h day/night, 30 min dusk/dawn lighting cycle. Larvae were fed a diet of powdered fish food (Tetramin) mixed with yeast. Adults were provided ad libitum with water and either raisins or a 10% sucrose solution. Transgenic and wild-type (control, nontransgenic) mosquitoes were reared in parallel using standardized insectary procedures (1). All transgenic lines were outcrossed at every generation to maintain hemizygosity for each transgene insert.

Transformation Plasmid Assembly. The 3xP3-DsRed cassette with 387 base pairs (bp) of the *D. melanogaster gypsy* insulator sequences (GenBank accession no. AF242365, nucleotides 1381–1767, inclusive; pH Stinger vector) was constructed as inverted repeats flanking the marker gene. The synthetic expression cassette was obtained from GeneScript and subcloned into the PBac-attB plasmid described previously (2). The orientation of the cassette was directed using the restriction endonucleases, *Bst*BI and *Fse*I. The pBac-3xP3DsRed-attB plasmid described previously (3) was used to generate the control or uninsulated transgenic lines.

Microinjection Procedures. Microinjections of mosquito embryos were performed using two previously described procedures (4, 5): (*i*) *piggyBac* helper plasmid was used to generate random transgene integrations, and (*ii*) in vitro-transcribed *φC31* integrase used to generate site-specific transgene integrations. Embryos were injected with a solution consisting of 500 ng/μL of either GYP[3xP3-DsRed] or 3xP3-DsRed plasmid DNA and 300 ng/μL *piggyBac* helper plasmid, or 300 ng/μL of either GYP [3xP3-DsRed] or 3xP3-DsRed plasmid DNA and 400 ng/μL *φC31* integrase mRNA. G_0 males and females were outcrossed to wild-type mosquitoes in pools of ∼5 G_0 males or 15–30 G_0 females. G_1 progeny were screened as larvae for DsRed fluorescence under UV-fluorescence microscopy.

DNA/RNA Extraction, cDNA Synthesis, and Southern Blot Analyses. Total genomic DNA and RNA were extracted from single mosquitoes in the *piggy-Bac*–mediated transgene insertions using the ZR-Duet DNA/RNA MiniPrep kit (Zymo Research). Individual mosquitoes were used so the gene copy number could be determined precisely. Ten mosquitoes were used in each of the samples from the *φC31* integrase-mediated transgenes insertions because they all have the same copy number. A total of 0.5 μg of RNA was used for reverse transcription in reaction volumes of 20 μL using qScript cDNA SuperMix (Quanta Biosciences). Standard Southern blotting and hybridization techniques to detect transgene integration included digesting genomic DNA samples with *Eco*RI, membrane transfer, and hybridization with a ^{32}P-labeled probe complementary to the DsRed encoding DNA.

Real-Time Quantitative RT-PCR Analysis. DsRed mRNA quantification and genomic copy number determination were performed on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories). DsRed primers were designed using Prime3 software to amplify a 142 base pair (bp) product (nucleotides 399–540 of the ORF). Primer sequences are *RED-F* 5′-CCCCGTAATGCAGAAGAAGACT-3′ and *RED-R* 5′-GATTGACTTGAACTCCACCAGGT-3′. DsRed expression values and genomic copy numbers were normalized, respectively, to mRNA abundance levels and copy number of the *An. stephensi* ribosomal S7 gene with primers developed by Brown et al. (6). The primers S7-F 5′-GGTGCACCTGGATAAGAACCA-3′ and S7-R 5′-CGGCCAGTCAGCTTCTTGTAC-3′ amplify an 84 bp product. DsRed expression and genomic copy number were quantified using gene-specific primers in a 20 μL final reaction volume containing 10 μL of SsoFast EvaGreen SuperMix (Bio-Rad), 300 nM each forward and reverse primer, and 5.0 μL cDNA sample. The amplification protocol for DsRed expression consists of 30 s at 95 °C, followed by 40 cycles of amplification (95 °C for 5 s, 63 °C for 5 s, plate read of SYBR Green I fluorescence), after which a melting-curve reaction was conducted from 65 °C to 95 °C with plate readings every 0.5 °C. A template genomic DNA was used to quantify DsRed genome copy number with an amplification protocol consisting of 2 min at 98 °C, followed by 40 cycles of amplification (98 °C for 5 s, 63 °C for 5 s, plate read of SYBR Green I fluorescence), after which a melting-curve reaction was conducted from 65 °C to 95 °C with plate readings every 0.5 °C. Each run included negative controls consisting of wild-type cDNA and water instead of cDNA.

Measurements of mRNA abundance were taken in triplicate and their mean used for further analyses. A negative control (wild-type mosquito cDNA) and an RNA sample without a reverse transcriptase step (to determine genomic DNA contamination) were included in each run. GraphPad Prism software was used to calculate statistical significance using t tests. P values ≤0.05 were considered significant.

Reciprocal Blast Analyses of *gypsy*-Like Complex in Mosquitoes. Primary protein sequences of *D. melanogaster Su*(*Hw*) (AAF55044), *Mod*(*mdg4*)2.2 (AA041580), and *CP190* (AAF55159) were obtained from FlyBase and used to search VectorBase for mosquito orthologous genes. The sequence of the *Mod*(*mdg4*) isoform implicated in insulator-like function was the only one used here. Significant hits were verified manually using the reciprocal BLAST hit method, where two proteins from two genomes were considered orthologous if each protein used as a query returned the other protein as the highest scoring BLASTp match.

1. Benedict MQ (1996) Care and maintenance of anopheline mosquito colonies. *The Molecular Biology of Insect Disease Vectors: A Methods Manual*, eds Crampton JM, Beard CB, Louis C (Chapman & Hall, London), pp 3–12.
2. Isaacs AT, et al. (2011) Engineered resistance to *Plasmodium falciparum* development in transgenic *Anopheles stephensi*. *PLoS Pathog* 7(4):e1002017.
3. Horn C, Wimmer EA (2000) A versatile vector set for animal transgenesis. *Dev Genes Evol* 210(12):630–637.
4. Isaacs AT, et al. (2012) Transgenic *Anopheles stephensi* coexpressing single-chain antibodies resist *Plasmodium falciparum* development. *Proc Natl Acad Sci USA* 109(28): E1922–E1930.
5. Nimmo DD, Alphey L, Meredith JM, Eggleston P (2006) High efficiency site-specific genetic engineering of the mosquito genome. *Insect Mol Biol* 15(2):129–136.
6. Brown AE, Bugeon L, Crisanti A, Catteruccia F (2003) Stable and heritable gene silencing in the malaria vector *Anopheles stephensi*. *Nucleic Acids Res* 31(15):e85.

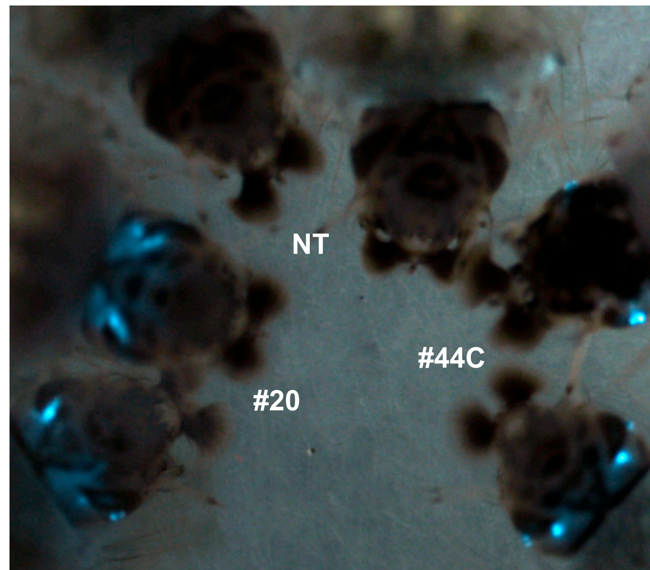

Fig. S1. Marker gene fluorescence phenotypes in larvae of docking-site lines 20 and 44C. Fourth instar larvae of lines 20 (*Left*) and 44C (*Right*) show qualitative differences in perceived brightness and distribution of cyan fluorescent protein (CFP) fluorescence. Nontransgenic larvae (NT; *Center*) have no discernible fluorescence in their eyes.

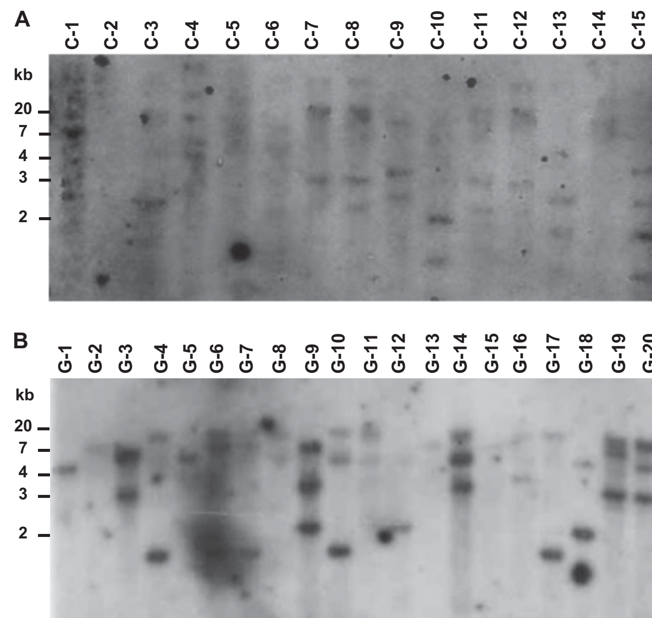

Fig. S2. Southern blot analyses of uninsulated and insulated transgenes inserted randomly into the *Anopheles stephensi* genome. Southern blot analyses were used to verify transgene insertions and to determine the number of integrated constructs in each line. Genomic DNA samples were digested with *Eco*RI and probed with DsRed. (*A*) Samples from 3xP3-DsRed control insertions. (*B*) GYP[3xP3-DsRed] insertion samples. Approximate fragment lengths in kilobase pairs (kb) are indicated to the left of each image. C, control samples; G, GYP[3xP3-DsRed] insertion samples.

Table S1. Statistical analyses of differences in DsRed accumulation between G_2 and G_5 of insulated transgenic lines*

Transgenic line	G_2 vs. G_5
G-1	0.4616
G-2	0.844982
G-3	0.78452
G-4	0.503123
G-5	0.652212
G-6	0.74025
G-7	0.328367
G-8	0.804168
G-9	0.762184
G-10	0.942616
G-11	0.953766
G-12	0.9846
G-13	0.838593
G-14	0.215043
G-15	0.209354
G-16	0.304843
G-17	0.011976
G-18	0.185253
G-19	0.117397
G-20	0.421883

Generations 2 and 5 (G_2, G_5) following strain isolation.
**P* values calculated with a *t* test.

Table S2. Statistical analyses of differences in DsRed accumulation between G_2 and G_5 of uninsulated (control) transgenic lines*

Transgenic line	G_2 vs. G_5
C-1	0.0015332
C-2	0.028164
C-3	0.026006
C-4	0.003461
C-5	0.034218
C-6	0.001749
C-7	0.163027
C-8	0.101812
C-9	0.460561
C-10	0.357412
C-11	0.013936
C-12	0.002395
C-13	0.027814
C-14	0.00278
C-15	0.003112

Generations 2 and 5 (G_2, G_5) following strain isolation.
**P* values calculated with a *t* test.

Table S3. Statistical analyses of differences in DsRed mRNA abundance in G_2 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #P12 (g^-)	44C #12 (g^+)	44C #P2 (g^+)
44C #28 (g^-)	0.069471	0.000052	0.000298
44C #P12 (g^-)		0.000005	0.000078
44C #12 (g^+)			0.088028
Docking site 20 transgenic lines	20 #P1 (g^+)		
20 #P3 (g^-)	0.006522		

P values calculated with a *t* test. (g^-), control lines without gypsy; (g^+), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct.

Table S4. Statistical analyses of differences in DsRed mRNA abundance in G_4 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #P12 (g^-)	44C #12 (g^+)	44C #P2 (g^+)
44C #28 (g^-)	0.375305	0.008415	0.033715
44C #P12 (g^-)		0.000698	0.005203
44C #12 (g^+)			0.607108
Docking site 20 transgenic lines	20 #P1 (g^+)		
20 #P3 (g^-)	0.06979		

P values calculated with a *t* test. (g^-), control lines without gypsy; (g^+), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct. Pink-shaded values are exceptional results.

Table S5. Statistical analyses of differences in DsRed mRNA abundance in G_6 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #P12 (g^-)	44C #12 (g^+)	44C #P2 (g^+)
44C #28 (g^-)	0.437162	0.004283	0.004406
44C #P12 (g^-)		0.026926	0.0277276
44C #12 (g^+)			0.581383
Docking site 20 transgenic lines	20 #P1 (g^+)		
20 #P3 (g^-)	0.05127		

P values calculated with a *t* test. (g^-), control lines without gypsy; (g^+), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct. Pink-shaded values are exceptional results.

Table S6. Statistical analyses of differences in CFP mRNA abundance in G_2 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #28 (g^-)	44C #P12 (g^-)	44C #12 (g^+)	44C #P2 (g^+)
44C (g^-)	0.330941	0.131921	0.001143	0.000037
44C #28 (g^-)		0.666577	0.00565	0.000172
44C #P12 (g^-)			0.00565	0.00026
44C #12 (g^+)				0.807394
Docking site 20 transgenic lines	20 #P3 (g^-)	20 #P1 (g^+)		
20 (g^-)	0.01916	0.00259		
20 #P3 (g^-)		0.000907		

P values calculated with a *t* test. (g^-), control lines without gypsy; (g^+), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct.

Table S7. Statistical analyses of differences in CFP mRNA abundance in G_4 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #28 (g^-)	44C #P12 (g^-)	44C #12 (g^+)	44C #P2 (g^+)
44C (g^-)	0.124536	0.26638	0.000916	0.000297
44C #28 (g^-)		0.091453	0.00872	0.001519
44C #P12 (g^-)			0.000624	0.000041
44C #12 (g^+)				0.416148
Docking site 20 transgenic lines	20 #P3 (g^-)	20 #P1 (g^+)		
20 (g^-)	0.000014	0.00025		
20 #P3 (g^-)		0.000356		

P values calculated with a *t* test. (g^-), control lines without gypsy; (g^+), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct.

Table S8. Statistical analyses of differences in CFP mRNA abundance in G_6 mosquitoes of 44C- and 20-insulated and control transgenic lines arising from *φC31*-mediated site-specific recombination

Docking site 44C transgenic lines	44C #28 (g⁻)	44C #P12 (g⁻)	44C #12 (g⁺)	44C #P2 (g⁺)
44C (g⁻)	0.420387	0.266897	0.000271	0.001116
44C #28 (g⁻)		0.747312	0.001095	0.009644
44C #P12 (g⁻)			0.000231	0.001199
44C #12 (g⁺)				0.285122
Docking site 20 transgenic lines	20 #P3 (g⁻)	20 #P1 (g⁺)		
20 (g⁻)	0.000632	0.000015		
20 #P3 (g⁻)		0.000402		

P values calculated with a *t* test. (g⁻), control lines without gypsy; (g⁺), lines with gypsy. *P* values for interstrain comparisons are unshaded. Green-shaded values are the comparisons of the two independently derived transgenic lines for each construct.

Table S9. Reciprocal BLAST analyses for genes encoding putative *gypsy* insulator complex proteins

	Query: *D. melanogaster**		
	Su(*Hw*)	*Mod*(*mdg4*)2.2	*CP190*
Species	AAF55044	AA041580	AAF55159
Anopheles gambiae	36.8% (AGAP003111-PA)	44.7% (AGAP003439-PA)	53.0% (AGAP001794-PA)
An. stephensi	36.8% (ASTM008993-PA)	35.0% (ASTM009688-PA)	47.9% (ASTM005466-PA)
Aedes aegypti	37.5% (AAEL002145-RA)	44.4% (AAEL010576)	47.6% (AAEL002771-RA)
Culex quinquefasciatus	35.0% (CPIJ007394-RA)	40.2% (CPIJ001652-RA)	45.5% (CPIJ013368-RA)
	Query: *An. gambiae**		
	Su(*Hw*)	*Mod*(*mdg4*)2.2	*CP190*
	AGAP003111-PA	AGAP003439-PA	AGAP001794-PA
D. melanogaster	36.8% (AAF55044)	44.7% (AA041580)	53.0% (AAF55159)
An. stephensi	81.5% (ASTM008993-PA)	66.0% (ASTM009688-PA)	78.2% (ASTM005466-PA)
Ae. aegypti	66.0% (AAEL002145-RA)	59.3% (AAEL010576)	59.3% (AAEL002771-RA)
C. quinquefasciatus	63.7% (CPIJ007394-RA)	56.9% (CPIJ001652-RA)	50.8% (CPIJ013368-RA)

*The query sequences were used to search the most recent Vectorbase genome versions for each mosquito species and Flybase for the fruit fly. Percent similarity in primary amino acid sequence is followed by the accession numbers of each gene.