Bioinformatic and cell-based tools for pooled CRISPR knockout screening in mosquitos

Raghuvir Viswanatha1,4, Enzo Mameli1,2,4, Jonathan Rodiger1, Pierre Merckaert1, Fabiana Feitosa-Suntheimer2, Tonya M. Colpitts2, Stephanie E. Mohr1, Yanhui Hu1 & Norbert Perrimon1,3

Mosquito-borne diseases present a worldwide public health burden. Current efforts to understand and counteract them have been aided by the use of cultured mosquito cells. Moreover, application in mammalian cells of forward genetic approaches such as CRISPR screens have identified essential genes and genes required for host-pathogen interactions, and in general, aided in functional annotation of genes. An equivalent approach for genetic screening of mosquito cell lines has been lacking. To develop such an approach, we design a new bioinformatic portal for sgRNA library design in several mosquito genomes, engineer mosquito cell lines to express Cas9 and accept sgRNA at scale, and identify optimal promoters for sgRNA expression in several mosquito species. We then optimize a recombination-mediated cassette exchange system to deliver CRISPR sgRNA and perform pooled CRISPR screens in an Anopheles cell line. Altogether, we provide a platform for high-throughput genome-scale screening in cell lines from disease vector species.

1 Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA 02115, USA. 2 Department of Microbiology, National Emerging Infectious Diseases Laboratories, Boston University School of Medicine, 620 Albany Street, Boston, MA 02118, USA. 3 HHMI, Harvard Medical School, Boston, MA 02115, USA. 4 These authors contributed equally: Raghuvir Viswanatha, Enzo Mameli. Email: ram@genetics.med.harvard.edu; perrimon@receptor.med.harvard.edu
Mosquito-borne diseases include a vast repertoire of viral, bacterial and parasitic diseases of medical and veterinary importance, with malaria alone causing nearly half a million human deaths each year. Current efforts to fight malaria and other mosquito-transmitted pathogens such as dengue, Zika, Chikungunya and West Nile viruses rely on control of vector populations, mostly by means of insecticides. These measures are hampered by ever-increasing insecticide resistance. Alternative strategies under current development include those based on the use of endosymbiotic bacteria such as Wolbachia, gene drives to suppress wild mosquito populations, or introduction of disease-refractory mosquitoes.

A key advantage of the introduction of CRISPR-Cas9 technology was the ability to generate large pools of sgRNAs and simultaneously test their effect in mammalian cells. This approach has transformed several areas of cell biology and revealed the function of previously unannotated genes. CRISPR screening in mammalian cells has already provided key insights into the entry and infection mechanisms of numerous viruses, parasites, bacteria and viruses, including mosquito-borne viruses. However, mosquito-borne viruses interact with a distinct set of host factors in the mammalian and insect host. Moreover, mosquito-borne viruses cause fewer cytopathological effects (CPE) in mosquito cells than in mammalian cells, and tend to develop persistent infections in mosquito cells but not in mammalian cells. Thus, to better understand the mosquito host genes involved in pathogen interactions, a method for unbiased genetic screening in mosquito cells is needed.

Roughly 20 mosquito cell lines from Aedes, Culex, and Anopheles genera have been established over the last 50 years. These cells are most widely used to propagate and characterize mosquito-borne viruses, including dengue, yellow fever, La Crosse, Japanese encephalitis virus, West Nile, Rift Valley, o’nyong-nyong, Sindbis, and Zika viruses. Studies using these cell lines have revealed dependencies, such as a need for low pH of endocytic compartments for infection and specific host factors. Immune-competent mosquito cell lines are also useful in dissecting the innate immune response and the unique mosquito cellular anti-viral response, which involves the somatic production of PIWI-interacting small RNAs. Cell lines also provide a platform to propagate viruses or intracellular pathogens, and permit in vitro characterization of mosquito-specific drugs, toxins, viruses, and Wolbachia.

Results

A unified resource for ortholog search and batch CRISPR guide design in mosquito species. To facilitate CRISPR-based genome engineering in mosquitoes and provide a batch-mode design resource for pooled CRISPR knockout (KO) screening targeting protein-coding genes, we developed CRISPR GuideXpress (https://www.lynnrai.org/tools/ly2mosquito/web/), an online resource with a number of features. First, CRISPR GuideXpress allows users to input genes from Drosophila, which as a model organism for diploids has a very well-annotated genome, and retrieve the closest mosquito orthologs. Orthology is mapped using an approach similar to DIOPT. Second, CRISPR GuideXpress also provides a cross-species reference when the same guide targets a homologous gene in one of the other supported species allowing, in some cases, inter-species targeting with the same reagents. For each mosquito species, the sgRNA designs cover ~92-99% of protein-coding genes, and at least ~62-93% of protein-coding genes are targeted by 6 or more high quality sgRNAs (i.e., designs with no predicted off-targets). The number of designs and relative coverage per gene for mosquito genomes is similar to the library used for CRISPR KO screening in Drosophila cells. Furthermore, for An. gambiae and An. coluzzii, we incorporated a variant database based on full genome sequences of hundreds of field samples from the Anopheles 1000 Genomes Project in order to allow for the selection of designs that would avoid common SNPs in wild populations.

Engineering RMCE acceptor mosquito cell lines. To generate RMCE acceptor cell lines as a platform for CRISPR screens, we first chose well-characterized cell lines from three mosquito species that are susceptible to infection with biomedically important viruses or parasites, and for which genomic, transcriptomic, and small RNA sequencing data exist. From Anopheles coluzzii (formerly An. gambiense M form), NAMRU2-CQ-01-01 (also known as Hsu) from Culex quinquefasciatus, and C6/36 from Aedes albopictus. Our previous work in a Drosophila cell line showed that CRISPR screens can be conducted by first introducing constitutive Cas9 expression and then transfecting cells with donor sgRNA expression vectors that can integrate into the RMCE loci. This way, each cell stably integrates a small number of different sgRNA expression cassettes according to the number of RMCE insertion sites. To validate this approach in mosquito cells, we first constructed a series of RMCE lines using a Mimic vector mobilized from a plasmid in the host genome at low frequency. Modified cells are identified by the presence of an mCherry exon that becomes incorporated into a native gene. mCherry-expressing cells were isolated using fluorescence-activated cell sorting (FACS) and we selected a single, strong mCherry-positive derivative cell-line from each parental line: Sua-5B-IE8 (Anopheles), NAMRU2-CQ-01-1.7 (Culex), and C6/36-HE8 (Aedes). As expected, we observe different mCherry distributions in each clonal isolate.
Fig. 1 CRISPR GuideXpress: an online bioinformatics framework for CRISPR sgRNA design and analysis. a Features and sgRNA design workflow. Ortholog mapping, cell line-specific expression data, and sgRNA design for six supported mosquito species are integrated at one interface. Genes can be searched individually or in batch mode. Direct ortholog searching is available between An. gambiae and other mosquito species or Drosophila. After a gene name or ID is entered, the tool retrieves corresponding transcripts and displays precomputed sgRNAs and associated scores. The sgRNAs are computed as follows. The longest isoforms are identified from transcripts. Next, all possible PAMs and associated sgRNA designs on both strands are selected. Each design is then assigned a seed score based on uniqueness of the 12-15 nt 3’ sequence (excluding the PAM). For each guide, a BLAST search is used to define specificity (off-target score). Each guide is mapped to the genome and categorized based on the gene region targeted and the respective isoform coverage. All sgRNA designs are evaluated to yield multiple efficiency parameters: ‘Housden’ score, machine learning (ML) score, and distance from ATG. Additionally, sgRNA designs for An. gambiae and An. coluzzii are assigned a ‘wild population’ efficiency score calculated from the Ag1000 Genome project dataset (see methods). To optimize for use in An. coluzzii Sua-5B cells, the tool indicates if the sgRNA sequences fully match the Sua-5B whole-genome sequence. (b-d) Analysis of genome-wide CRISPR KO sgRNA designs targeting protein-coding genes in supported mosquito species. b Histogram representing total number of sgRNA designs in two categories: (green squares) “no OTE” (off-target effect), with minimal off-target effects, or (gray squares) “with OTE” within the criteria (see Methods). c Genome-wide sgRNA design coverage, showing the percentage of genes targetable by sgRNAs with minimal OTE (light yellow to green), targetable only by sgRNAs with potential OTE (gray), or untargetable (black). d Genome-wide sgRNA design coverage by gene (%) in wild populations sampled in the Ag1000 Genome project. % of genes targeted and ranking based on # of sgRNAs/gene, as specified above. For this analysis were considered only sgRNA designs matching ≥ 95% of the wild genome sequences sampled. Source data are provided as a Source Data file. Raw statistics can be found in Supplementary Data 4.
Identifying optimal U6 promoters for CRISPR KO in mosquito cells. An incompletely addressed challenge for CRISPR genome engineering in mosquitoes is the identification of optimal pol III promoters in mosquitoes is the identification of optimal pol III promoters for heterologous expression of sgRNAs. We performed a side-by-side evaluation of eleven pol III promoters from four mosquito species, as well as a consensus sequence, in each of the three mosquito cell lines. To choose promoters, we first used BLAST and multiple sequence alignments to identify orthologs of the Drosophila U6-2 (snRNA:U6:96Ab) promoter and chose eleven orthologous promoters from U6 snRNAs of Anopheles, Culex, or Aedes (Fig. 3a). When possible, we selected a minimum of three promoters per species, prioritizing U6 promoters that contain an intact pol III bipartite promoter motif and for which RNA-seq data suggests they are expressed in cell lines and in adult tissues (see Methods). These were synthesized and inserted into pLib6.440,44 to generate a suite of vectors for the expression of sgRNA under the control of different U6 promoters (Supplementary Data 1).

Mosquito cells with genomically-encoded mCherry allowed us to use a flow cytometry-based dual reporter assay to directly compare KO efficiency in cells expressing the same sgRNA from different U6 promoters (Fig. 3b, Supplementary Fig. 2, Supplementary Data 1). In this strategy, we test U6 promoter strength by measuring the ability of the downstream sgRNA to suppress mCherry. Specifically, we co-transfected mCherry expressing cells with a Cas9 expression vector and a plasmid containing a mCherry-targeting sgRNA driven by a variable U6 promoter. The U6 promoter plasmid co-expressed GFP as an indicator of transfection. After gating cells with GFP expression, the ratio of mCherry- cells is used to determine KO efficiency. An improvement of this approach over a plasmid-based dual reporter assay is that mCherry is genomically encoded rather than an episomal target, revealing repair outcomes that would be expected at a native gene. Although the number of insertions and expression level of mCherry vary between cell lines, this approach permits the comparison of different U6 promoters within the same cell line. In Anopheles cells, all mosquito promoters tested elicited measurable KO, whereas Drosophila promoters failed (Fig. 3c, Supplementary Fig. 2b). The native promoters (AGAP013695, AGAP013557) along with Culex CPIJ039596 and Ae. aegypti AAEI017774 showed the strongest activity, achieving approximately 75% KO efficiency relative to controls. In particular, AAEI017774 (mean = 81.3 SD ± 1.9) and AGAP013695 (mean = 76.6 SD ± 3) were the most efficient. The remaining promoters have moderate to low activity, and the mosquito consensus promoter performed similarly to the native promoters. In the Culex cell line, we observed a more uniform activity of mosquito U6
promoters, with an overall mean KO efficiency of about 30% (Fig. 3d, Supplementary Fig. 2b). Notably, the results for CPIJ039596 obtained using this assay were slightly lower but overall comparable to CRISPR allele editing efficiency as verified by deep sequencing for the same promoter in our previous study38. The most effective U6 promoters in Ae. albopictus C6/36-HE8 cells were the native promoters AALF029743-4 (mean = 28.6 SD ± 6.1; mean = 26.4 SD ± 4.9), Culex CPIJ039596 and Ae. aegypti AAEL017774, with about 27% mean KO efficiency (Fig. 3c, Supplementary Fig. 2b). Interestingly, Culex CPIJ039596,
CRISPR KO produces an observable phenotype in Anopheles cells. We next asked whether our *Anopheles* CRISPR screening platform results in penetrant, visible phenotypes. In Sua-5B-IE8-Act::Cas9-2A-Neo cells, we asked whether a visible phenotype can result from introducing a sgRNA expression cassette targeting Rho1 (AGAP005160), which is necessary for the completion of cytokinesis, driven by the optimal U6 promoter Agam_695 (Fig. 3c). Previous reports have shown that knockdown of Rho1 by RNAi in *Drosophila*57 or *Anopheles*25 results in a modest size increase (~2-fold) due to cell growth without division, and *Drosophila* cells expressing CRISPR sgRNAs targeting Rho1 become dramatically enlarged due to complete loss of Rho140. To test the novel *Anopheles* cell-based CRISPR system, we transfected sgRNAs targeting the *Anopheles* Rho1 ortholog AGAP005160 and observed transfected cells after several days of selection. We found that Rho1 sgRNA-expressing cells, but not control cells, became enlarged up to 6-fold (Fig. 4a, b). We used T7 Endonuclease I assays to confirm editing of the Rho1 locus (Fig. 4c). Enriching for the sgRNA-expressing cells resulted in greater editing, as would be expected if the editing frequency was limited by a low percentage of sgRNA transfection (Fig. 4c). These results clearly demonstrate that the Sua-5B-IE8-Act::Cas9-2A-Neo ‘CRISPR-ready’ cell line can yield highly penetrant phenotypes, suggesting that the system is compatible with CRISPR screening.

Validation of the CRISPR screening platform in Anopheles cells. To directly test applications of the CRISPR screening platform at a large scale in mosquito cells, we first chose five genes that had previously been shown to be drug-resistance factors in *Drosophila* cell lines and used CRISPR GuideXpress to design a library targeting their orthologs in *Anopheles coluzzii*. Target genes included *Anopheles* orthologs of FKBP12 (AGAP012184), which encodes the cellular binding partner of the mTOR inhibitor rapamycin; EcR (AGAP028634) and usp (AGAP002095), which encode mediators of an antiproliferative transcriptional response to treatment with ecdysone; and PTP-ER (AGAP007118), which encodes a negative regulator of the mitogen-activated protein kinase (MAPK) signaling cascade that can be suppressed by treatment with the MEK inhibitor trametinib (Fig. 5a; Supplementary Data 3). In total, 3,487 sgRNAs were synthesized and cloned into pLib6.4-A gam_695 containing the strong *Anopheles* U6 promoter and transfected into *An. coluzzii* Sua-5B-IE8-Act::Cas9-2A-Neo cells in the presence of FC31 integrase to facilitate recombination, then selected for 16 days in puromycin-containing media with continuous passaging every four days. A theoretical copy number of 1000 cells per sgRNA was maintained during all passages. For the selection screens, the cells were grown for an additional 30 days in the presence of puromycin, ecdysone (20-hydroxyecdysone), or trametinib (Fig. 5b). Then, genomic DNA was collected, and the sgRNA-containing locus was PCR amplified, barcoded, and analyzed by next-generation sequencing (NGS). Guides targeting FKBP12 (AGAP012184) were clearly enriched by treatment with rapamycin but not in untreated, ecdysone-, or trametinib treatment conditions (Fig. 5c). Sequence analysis of the FKBP12 locus in the Sua-5B-IE8-Act::Cas9-2A-Neo cell line revealed a coding variant in the cells relative to the reference genome (AgamP4) that results in single-base mismatches between a subset of three sgRNAs designed to target the FKBP12 locus. Unlike no-mismatch guides, these mismatched guides were not selected in puromycin treatment conditions (Fig. 5d). Similar observations were made for the set of guides targeting usp (Supplementary Fig. 3; Supplementary Data 3). After observing these single nucleotide polymorphisms in specific genes, we referred to the whole-genome-sequence of the Sua-5B-IE8-Act::Cas9-2A-Neo ‘CRISPR-ready’ cell line and added a variant analysis to CRISPR GuideXpress, giving users the option to exclude these variants from sgRNA designs (Fig. 1a). Importantly, for all three screens, we found significant and selective enrichment for the orthologs of the expected genes: FKBP12 (*p < 4.99E-06 in rep # 1, p < 4.99E-06 in rep # 2) for rapamycin, EcR (*p < 1.50E-05 in rep # 1, p < 3.39E-06 in rep # 2) and usp (*p < 4.99E-06 in rep # 1, p < 4.32E-17 in rep # 2) for ecdysone, and PTP-ER (7.32E-17 in rep # 1, p < 1.14E-09 in rep # 2) in trametinib (Fig. 5e; Supplementary Fig. 3; Supplementary Data 3). This was driven by consistent enrichment of
multiple sgRNAs targeting each selected gene in each drug treatment regime: under rapamycin selection, **FKBP12** sgRNAs made up 9 of the top 20 enriched sgRNAs; under ecdysone selection, **EcR** made up 5 and **usp** made up 5 of the top 20 sgRNAs in the screen; under trametinib selection, **PTP-ER** made up 9 of the top 20 enriched sgRNAs (Supplementary Data 3). We note, however, that we did not observe enrichment for a candidate **Eci** ortholog, AGAP006638 (*p* < 0.78292 in rep # 1; *p* < 0.80726 in rep. # 2) (Supplementary Fig. 3; Supplementary Data 3). These results suggest that using the RMCE approach, optimized U6 expression, and CRISPR sgRNA design pipeline we have developed will make it possible to efficiently conduct massively parallel genetic screens in mosquito cells.
Discussion

Here we have developed tools for performing large-scale pooled CRISPR KO screens in mosquito cell lines and carried out a large-scale genetic screen in Anopheles cells. To establish the platform, we created a bioinformatic tool for batch sgRNA design; experimentally tested U6 promoters to identify those with high activity; cloned several CRISPR plasmid vectors and a large-scale sgRNA library targeting mosquito genes; modified mosquito cell lines for RMCE; and demonstrated, as expected from previous studies, that Rho1 KO in mosquito cell lines causes a strong, visible phenotype useful for assessing the efficiency of CRISPR modification. In a large-scale pooled CRISPR KO screen in Anopheles cells, we were successful in specifically enriching for genes that when knocked out were expected to provide resistance to one of three different experimental treatment conditions. Notably, the screen results validate the function of the predicted Anopheles orthologs of the Drosophila FKBP12, EcR, usp, and PTP-ER genes.

The online portal we introduce here, CRISPR GuideXpress, can be used to design single guides, genome-wide libraries, or focused libraries of variable size. As a tool for designing sgRNAs for individual CRISPR KO constructs, GuideXpress employs prioritization of guide designs by several parameters, including mismatches relative to cell lines or wild mosquito genomes. As a tool for batch sgRNA design, focused libraries, such as the one we created (Fig. 5), could have several immediate applications in mosquito research. Several studies have generated gene sets from proteomic or differential expression analyses (e.g., host proteins that interact with viral proteins), or genes up- or down-regulated in response to pathogen infection. Focused CRISPR screening based on these gene sets can provide functional validation of these data. The user can also provide high numbers of sgRNAs per gene, reducing noise particularly for challenging screens.

In addition, this study also addresses a broader lack of CRISPR tools for mosquito research. Basing CRISPR KO constructs on reagents optimized for use in Drosophila has worked for generating CRISPR KO mosquito cells. However, our results testing U6 promoters in this and previous work including our group suggest that species-specific optimization is worthwhile. Studies in mammalian cells have shown that empirically optimized sgRNAs lead to reduced off-targets and increased efficiency, and our studies enable the application of this strategy to mosquito genomes. Finally, our studies provide at least two high-expression U6 promoters for each mosquito species. Additional mosquito U6 promoters with lower activity could still be useful in applications where sgRNA dosage needs to be controlled. Furthermore, having multiple U6 promoters enables the combinatorial expression of sgRNAs in the same cell, reducing the chances of recombination between identical U6 promoter sequences. Interestingly, to our knowledge, we report the first comparison of mosquito U6 promoters in Drosophila cells, identifying at least three promoters with significant activity, expanding the array of tools for multiplexed CRISPR targeting in flies (Supplementary Fig. 2c). Finally, our work provides a platform that could accommodate the creation of single KO cell lines or KO pools of variable size, complementing and expanding substantially the tools currently available for KO studies in mosquito cell lines.

Further development of the screening strategy is likely to improve the platform in the future. First, although ΦC31 RCME efficiency is high in Anopheles cells (~85%), the initial transfection efficiency is low compared with Drosophila cells. As a result, a larger number of cells must be transfected, using a larger amount of a costly transfection reagent to achieve a comparable screen. Optimizing transfection efficiency, such as by using electroporation, has the potential to reduce screening costs without changing screen outcomes. Second, the finding that there are 1 to 5 RMCE cassettes per cell following transfection of the pooled library raises the possibility of "passenger effects" during selection that could reduce the resolution of the screen. Even in conditions of high multiplicity delivery (up to 10 sgRNAs per cell), high-quality screens can still be conducted by applying recently developed strategies, mitigating these effects could be important, as this could reduce the false-discovery rate (FDR) and increase reproducibility. Third, although our screen results verify that our approaches to identification of mosquito orthologs of Drosophila genes and to sgRNA design are valid, there is room for improvement in sgRNA design. For example, we would like to incorporate cell line genome data for additional cell lines. Moreover, we and others have learned that large-scale screen data provide lists of 'good' and 'bad' sgRNA designs for genes that were positive in the screen data and as such, can be used to derive sgRNA design rules. Thus, as we accumulate more large-scale screen data, we expect to iteratively improve our CRISPR GuideXpress resource. Finally, the approach should be extensible to screens based on other CRISPR systems, including CRISPR activation and CRISPR interference.

Importantly, the new ability to perform pooled CRISPR KO screens in mosquito cell lines will facilitate screens for essential genes and for genes that confer sensitivity or resistance to any treatment that slows growth or results in cell death, including insecticides, biological toxins, and other agents, further contributing to functional annotation of mosquito genes. A screen for resistance to ecdysone-induced cell death in Drosophila cells, for example, revealed novel ecdysone pathway components, including a previously uncharacterized transporter for ecdysone. The information gained from these screens can also inform our ability to control mosquito populations or infection of mosquitoes with human pathogens. For example, the screening platform can be used to identify conserved or species-specific essential genes, which in turn could be used in the design of gene drives and for the development of new and potentially highly targeted insecticides. In addition, screens in mosquito cells have the potential to increase our knowledge of host-pathogen interactions. Genome-scale CRISPR and RNAi screens have been used in the past to investigate interactions between mosquito-borne viruses and mammalian cells. Similar work in mosquito cells has been limited to targeting a few genes using RNAi or drug treatments. The ability to perform large-scale CRISPR KO screens in mosquito cell lines opens the door to potentially novel findings regarding the interaction of mosquito host cells with viral and other pathogens and holds great potential for aiding the multifront effort to control mosquito-borne diseases.

Methods

Identification of mosquito U6 promoters. U6 snRNAs are conserved euarkaryotic non-coding RNAs that take part in the formation of the catalytic core of the spliceosome, while their promoters and associated pol II transcriptionsal machinery show divergence even between closely related species. We used the Drosophila U6.2 (CR32867) snRNA sequence as a query to perform a BLAST search of mosquito reference genomes, using the Vectorbase interface (now VEuPathDB at https://vectorbase.org/vectorbase/app). Next, all sequences of the identified orthologs within each species were subjected to multiple alignments (using ClustalW from www.ebi.ac.uk), including the full snRNA and 500 bp of upstream sequence, allowing us to visualize and exclude sequences lacking conserved portions of the pol II bipartite promoter motif (i.e., the PSEA and TATA). As a secondary criterion for selection, we relied on expression levels reported on RNA-seq data publicly available for each species, including tissue-specific RNA-seq of adult mosquitoes or cell lines (available on vectorbase/VEuPathDB, e.g., from the Arthropod Cell Line RNA Seq initiative at the Broad Institute). This data, although not fully reliable in reporting expression levels of small non-coding RNAs, was informative in narrowing our choice to putatively expressed snRNAs. Third, we consulted literature reports of the activity of these promoters. After selecting up to three U6 promoters for each species, we performed a second inter–species alignment and selected for each promoter a region upstream of the TSS of an arbitrary length ranging from 144 to 237 bp. In addition, we selected a 99 bp mosquito consensus sequence derived from the alignment. The length of the AGAP013695 promoter was chosen.
based on previous work from Konet et al. 35 in cells and Hammond et al. 36 in adult *Anopheles* mosquitoes. The alignment shown in Fig. 3a was obtained by aligning the 250 bp upstream of *Drosophila* U6-2 and U6-3 promoters to the selected plasmid sequences (only a 0.75 bp region is displayed). The mosquitocensus sequence is shown at the bottom. Jalview Version 2.70 was used to visualize the alignment and to infer the phylogenetic hierarchy based on the average sequence distance of the snRNA and the 75 bp upstream sequence. Additional information about the U6 promoter sequences used for library vector construction is available in Supplementary Data 1.

**Cloning procedures.** In order to build library vectors expressing gRNAs under mosquito U6 promoters we synthesized ~500 bp pigBlocks (Integrated DNA Technologies, Inc.) containing in order: the selected U6 promoter sequences; a Bsx cloning site (for gRNA insertion); the gRNA scaffold sequence including an at least 8' termination sequence71; 30 nucleotides of the native termination sequence, and an additional portion of the termination sequence derived from the 3' end of the *Drosophila* snRNA-U6-66A8 (CR32867). The gBlocks were sub-cloned into pCR*-Blunt II-TOPO® (Invitrogen) or directly digested with BstBI/KpnI and sub-cloned into plp6.b.4 (EGFP reporter) or plp6.b.4B (EBFP2 reporter) by replacing the entire gRNA expressing cassette. This resulted in 24 new library vectors harboring the mosquito gRNA expressing cassettes. An additional version of the pLib6.4 attB donor library vector, named plp6.b4.AB, was obtained by replacing EGFP with EBFP2 using overlap extension PCR with megaprimers as described in Bryksin et al. 75

**U6 promoter evaluation**

- In order to build library vectors expressing gRNAs under mosquito U6 promoters we synthesized ~500 bp pigBlocks (Integrated DNA Technologies, Inc.) containing in order: the selected U6 promoter sequences; a Bsx cloning site (for gRNA insertion); the gRNA scaffold sequence including an at least 8' termination sequence71; 30 nucleotides of the native termination sequence, and an additional portion of the termination sequence derived from the 3' end of the *Drosophila* snRNA-U6-66A8 (CR32867). The gBlocks were sub-cloned into pCR*-Blunt II-TOPO® (Invitrogen) or directly digested with BstBI/KpnI and sub-cloned into plp6.b.4 (EGFP reporter) or plp6.b.4B (EBFP2 reporter) by replacing the entire gRNA expressing cassette. This resulted in 24 new library vectors harboring the mosquito gRNA expressing cassettes. An additional version of the pLib6.4 attB donor library vector, named plp6.b4.AB, was obtained by replacing EGFP with EBFP2 using overlap extension PCR with megaprimers as described in Bryksin et al. 75

- First, PCR was performed to amplify EBFP2 from pBFP2-Nuc (Addgene #14893) using primers 14-15. Second, overlap extension PCR was performed using as megaprimers pCR*-EBFP2 amplifying the EGFP construct in pCR*-EBFP2. A portion of the U6-3 promoter from pCFD3 (PMD: 25002478) was cloned into plp6.b4 to generate plp6.b6.6. The same strategy, using the primers, was also used to obtain plp6.b6.b from plp6.b6, as well as EBFP2 and EGFP versions of the psL1180-HR-PUB/EGFP plasmid (Addgene #47917). In order to generate the EGFP version of our library series by PCR, a fragment of EGFP was obtained by ligation cloning of the EGFP clone from pCR*-EBFP2 directly into plp6.b4 vector using the same primers (14-15). Insertion of the gRNA sequence targeting mCherry or Rho GTPase in plb vectors was performed following an established protocol 42. Donor library vector pLib6.4 vector was built from Ac5-STABLE1-Neo (Addgene # 32425) for expression of gRNA by Gibson assembly (New England Biolabs) resulting in pAePUbW destination vector, where the gRNA expression cassette was inserted. The plasmid contained the entire gRNA expression cassette encoding the gRNA scaffold sequence including a 30 nucleotides of the native termination sequence, and an additional portion of the termination sequence derived from the 3' end of the *Drosophila* snRNA-U6-66A8 (CR32867). The gRNA expression cassette was inserted downstream of the Ca9 expression cassette. Transfections in the Aedes aegypti cell line were performed with the pLib6.4 vector. The U6 promoter was used to drive the expression of gRNA targeting mCherry or Rho GTPase in plb vectors. The Aedes aegypti cell line was first centrifuged gRNA was delivered into the cell line using clonal populations were expanded and assessed based on signal intensity and subcellular localization of the mCherry reporter (example in Supplementary Fig. 1). Western blotting. Semiconfluent cells were washed twice in cold PBS and lysed directly in flasks using RIPA lysis buffer ( Pierce #89900) supplemented with protease inhibitor cocktail (Millipore Sigma #88830) and incubated on ice for 20 min. The supernatant of the cell line was first centrifuged at 20,000 g at 4°C for 10 min. Protein concentration was measured (Qubit, ThermoFisher) and 10% blocking buffer with mouse monoclonal anti-mCherry-Tag antibody (1:1000; St John's Laboratory #ST53473), rabbit polyclonal anti-flag (1:1000; Sigma # F7025), Washes were performed using PBS-T and secondary incubation was performed for 1 h at room temperature with goat anti-mouse Alexa Fluor Plus 800 (1:5000; Thermo Fisher Scientific, #A23720), goat anti-rabbit StarBright Blue 700 (1:5000; Bio-Rad, #12004161), human Fab anti-actin rhodamine-conjugated (1:10000; Bio-Rad, #12004164). Immune complexes were visualized using the ChemiDoc MP Imaging System and analyzed using Bio-Rad Image Lab (version 6.1).

**U6 promoter evaluation.** CRISPR KO efficiency of the stable mCherry reporter integrated into the mosquito cell line was used as a readout to test relative U6 promoter gRNA expression strength from different mosquito promoters, using transient transfection. Plasmid DNA for transfections was prepared using mini or midiprep kits (Qiagen) and quality was analyzed by spectrophotometry and agarose gel electrophoresis. Plasmid DNA concentration was measured (Qubit, ThermoFisher) and plasmid mix for transfection was normalized by copy number according to the M.W. of each plasmid and using puC19 (ThermoFisher) to normalize total DNA amounts as needed. All cells were transfected with 300 ng of plasmid mix (2 well format) using Effectene (Qiagen) and following the manufacturer protocol. Transfections in the *Anopheles* cell lines were performed with the pAct::Cas9-2A-Neo plasmid. Transfections in the *Anopheles* cell line were performed with the pAct::Cas9-2A-Neo and a second including the same control guide but lacking the Cas9-expressing plasmid. Transfections in the *Anopheles* cell line were performed with a different plasmid mix containing 150 ng (50%) of pAct::Cas9-2A-Neo, and 150 ng (50%) of 14 different copy-number balanced mCherry-sgRNA expressing plasmids was transfected. Additionally, two control transfections were performed: one including the “empty” plp6.b6.A-G4 (495 expression vector for sgControl), that effectively drives the expression of a “non-targeting guide” matching the sequence of the empty BbsI cloning cassette, and a second including the same control guide but lacking the Cas9-expressing plasmid. Transfections in the *Aedes* cell line were performed with a different plasmid mix containing 150 ng (50%) of pAePuC9::Cas9-2A-Neo as a source of Cas9 and 150 ng (50%) of 14 different copy-number balanced mCherry-sgRNA expressing plasmids and, in addition, 15 ng (5%) of psL1180-HR-PUB/EGFP-NLS to increase fluorescence output of the GFP reporter. For this cell line, we used the same sgControl plasmids, transfected with or without the Cas9 expressing plasmid. Transfections in the *Drosophila* cell line were performed with 495 expression vectors for each of one of 14 different copy-number balanced mCherry-sgRNA expressing plasmids and sgControl plasmid as a control. For this experiment since Cas9 was already expressed in the cell line transfection with Cas9 expressing
plasmid was not necessary. Additional information on plasmid vectors used is provided in Supplementary Data 1. Transfections were performed in 12-well format with 3 or more replicate wells per condition tested, and a total of three independent experiments were performed. C. elegans and A. bellus cells were detached from flasks using Accumax (Innovative Cell Technologies, Inc) and seeded onto 12-well plates 16–24 h before transfection. Anopheles cells were resuspended from flasks by pipetting and seeded 30 h before transfection. C. elegans and Aedes cells were transfected at ~70% confluence and Anopheles cells at ~90% confluence. 24 h after transfection, cells were transferred to new plates and cultured for 12 days, then either analyzed by flow cytometry or slowly frozen at ~80 °C in culture media supplemented with 10% DMSO (v/v) and stored until flow cytometry analysis. Anopheles and Aedes cells during the last passage preceding the experiment end-point were treated with a sub-lethal dose of puromycin (0.5 μg/ml) to increase the ratio of transfected cells and reduce the volume of cells analyzed with flow cytometry (puromycin resistance is conferred by the sgRNA expression plasmid).

Flow cytometry analysis. Flow cytometry analysis was performed on a FAC-Symphony analyzer (BD) (Harvard Medical School, Immunology Flow Core). Immediately prior to analysis, frozen cells were thawed quickly in a 30 °C water bath, washed twice in PBS, and resuspended in fresh culture media. Cells were analyzed by gating on a replicate well using Effective (QuantaSoft) software following the manufacturer’s instructions. Cells were transfected with a plasmid mixture containing 150 ng (50%) of pBS130 (Addgene #26290) encoding HSP70-βgal and 150 ng (50%) of plBb6-4-Agam695 "empty" (sgControl) or a sgRNA targeting Rho1 (sgRho1), and then cultured for 7 days in 4.5 μg/ml puromycin selective media. Cells were extracted using the QIAcube MiPO Prep kit (Qiagen). Each low-read sgRNA (those with fewer than 10 reads in the plasma library) were removed from the read count files. All subsequent read count and data analysis were performed using MaGeCK 0.5.7.

Orthology mapping, sgRNA design, and variant analysis. Orthology mapping between Drosophila melanogaster (Dmel) and Anopheles gambiae (Agam). Orthology mapping was extracted from the following five different prediction algorithms: orthoMCL v5.55, eggNOG v5.0, InParanoid v8.0, orthoFinder (for which the orthoFinder server was locally running on a single 64-core server). Next, orthologous sgRNA designs were inputted and all computed sgRNAs were retrieved. Additionally, the second target effect) score; maximum ML (machine learning efficiency) score; Bsd-site bearing sgRNAs were culled from batch results, sgRNA sequences were placed within a 10-mer using specific tags and retrieved using dial-out PCR76. Each dial-out product was digested with Bsd and separated on a 20% non-denaturing polyacrylamide TBE gel (ThermoFisher Scientific). The product was then extracted using the crush-soak method and ligated into BsiI-digested pLb6-4-Agam695. The resulting PCR product was then cloned into pE. coli, digested with ApoI and TOPO Electrocompetent Cells (Lucigen) and plated onto ten 15 cm LB plates containing carbenicillin and grown overnight at 30 °C. A total of 100 bacterial colonies per construct were harvested into 25 ml of LB medium, mixed with an equal volume of 50% glycerol, and stored in 1 ml aliquots at ~80 °C. Prior to transfection, the plasmid library was prepared by miniprep (Qiagen).

Chemical-genetic CRISPR screening. Sua-5B-IE8:Act:Cas9-2A-Neon cells in log phase of growth were seeded at 6 × 10^6 cells per well in 6-well dishes in growth media containing antibiotics. They were transfected with a plasmid mixture equimolar of HSP70-βgalIntegrate plasmid (pBS130) and sgRNA donor plasmid library (plBb6-4-Agam695) using Effectene (Qiagen) according to the manufacturer’s base protocol (Quick-sgDNA MiniPrep kit (Zymo)). Transfected cells were kept for ~22–28% under these conditions, whereas the % of stably recombined cells after one month of passage without selection was found to be ~3% (Fig. 2c). To achieve ~150 cells per sgRNA, 3487 sgRNAs × 100 cells/sqRNA = ~8.1 × 10^4 were transfected in three wells of a 6-well dish. After 4 days, each well was expanded into a 15 cm dish containing 5 ug ml^-1 ambystoma (Sellek Chemical) was dissolved in DMSO and used at a final concentration of 40 nM; 20-hydroxyecdysone (Sigma) was dissolved in DMSO and used at a concentration of 100 ng/ml; trametinib (Sellek Chemical) was dissolved in DMSO and used at a final concentration of 400 nM. Cells were continually passed in the selective medium for 30 days with media changes or re-seeding. Re-seeding density was maintained above 1000 cells/sqRNA at all times. Following selection, cell pellets representing >1000 cells/ sqRNA were extracted using the Quick-sgDNA MiniPrep kit (Zymo). Next, the genomic DNA was subjected to 2-step PCR to introduce in-line barcodes, a variable fluorescent reporter, and Illumina sequencing primer and adapters. Amplicons were subjected to sequencing using a NextSeq500 at the Harvard Bioprocessors Facility at Harvard Medical School. Computational barcode removal was performed using in-house scripts (https://vectorbase.org/vectorbase/app/query-grid) local to the lab. The Bioart orthomapper algorithm was used for this study was subsequently replaced when Vectorbase merged with EuPathDB to become VEuPathDB. Orthology mapping is now available from the new user interface using a gene ID search strategy with an orthology transform step (https://vectorbase.org/vectorbase/app/query-grid). Genome reference and sgRNA design pipeline. Gene and genome sequence files available at the drop down menu on the CRISPR GuideXpress search page (https://www.flyrnai.org/tools/fly2mosquito/web/species). The species currently supported at CRISPR GuideXpress and their corresponding genome versions are as follows: Anopheles gambiae (AgamP4.12), Anopheles stephensi (Agst1.7), Aedes aegypti (AaegL2.5), Aedes albopictus (AaloF1.2 & C6/36), and Culex quinquefasciatus (Cqy2mosquito/store/species). The species currently supported at CRISPR GuideXpress and their corresponding genome versions are as follows: Anopheles gambiae (AgamP4.12), Anopheles coluzzi (AcMoL1.8), Anopheles stephensi (Agst1.7), Aedes aegypti (AaegL2.5), Aedes albopictus (AaloF1.2 & C6/36), and Culex quinquefasciatus (Cqy2mosquito/store/species). Genome sequence and annotation files were obtained from the following sources: vectorbase.org/vectorbase/app/downloads/Pre-VEuPathDB%20VectorBase%20files/. Input files used by the pipeline include <species_name>_BASEFEATURES_<species_version>.gtf, <species_name>_TRANSCRIPTS_<species_version>.fa, and <species_name>_CHROMOSOMES_<species_version>.fa.

The pipeline starts by using genome annotation to determine the transcript with the longest coding sequence (CDS) for each gene. In cases where there is a tie, the first gene transcript is chosen arbitrarily. All potential gene annotations within the CDS of the selected transcripts are identified and logged. Then, unique k-mers within the genome are identified for computing seed scores for each design. Design scores are BLASTed against the genome sequence and off-target scores are assigned based on off-target alignment hits with varying numbers of mismatches using UCSC genome tables. All off-target hits with ≥3 mismatches are weighted down by 70%. The OTE (Off-target effect) score was calculated based on the number of potential off-target sites at 3 different thresholds and defined as follows: "OTE score = a/bc" where (a) is added to the digit before the decimal point and is the number of off-target sites of the least stringent threshold (only considering off-target sites with 3 or fewer mismatches) added to the digit after the decimal point and (b) is the number of off-target sites at the moderate threshold (only considering off-target sites with 4 or fewer mismatches) (c) is added as the second digit after the decimal
place and is the number of off-target sites under the most stringent threshold (5 or fewer mismatches). The categorization of sgRNA designs in Fig. 1 b, c, e was defined as follows: absence of off-GTEx (least stringent criteria) = score ≥ 1; with OTE (least stringent criteria) = score ≥ 1.

For each sgRNA, a Housden efficiency score was computed using a position matrix, and a machine learning-based efficiency score is computed using the pipeline available here: https://github.com/PierreMks/Dmel-sgRNA-Efficiency-Prediction, which is based on Drosophila tools (Ghia et al. CRISPR screens). The comparison with SNPs associated with genome data from wild populations is calculated based on the Ag1000G dataset: https://www.malariagen.net/data/ag1000g-phase2-ar1 (available for AgamP4.12 and AcolM1.8). The sgRNA pipeline was written in Python 3 and Perl 5.24.0, and uses BLAST 2.6.0. Analysis of wild populations variants. Efficiency of sgRNAs for An. gambiae and An. coluzzi in wild populations was evaluated based on the datasets from Anopheles gambiae 1000 genomes project46,47 (Ag1000G) (https://www.malariagen.net/data/ag1000g-phase2-ar1). Percent efficiency in wild populations was calculated based on the percent of samples carrying the SNP in each of the sgRNA target sequence. The SNP analysis results were obtained from ftp://ngs.sanger.ac.uk/production/ag1000g/phase2/cas9_targets/77. The sample metadata was obtained from ftp://ngs.sanger.ac.uk/production/ag1000g/phase2/AR1/samples/45.

Genome variants in Sua-5B cell line. The Sua-5B cell line was sequenced by BGI Genomics (https://www.bgi.com/global/) and processed as follows. The partner-end sequencing data was QC’d using fastqc 0.11.8 and multqc 1.9. Pre-processing was performed according to the Broad Institute’s best practices workflow for data preprocessing for variant discovery (https://gatk.broadinstitute.org/hc/en-us/articles/3600385533912-Data-pre-processing-for-variant-discovery). First, the raw reads were aligned to the Anopheles gambiae (AgamP4) genome sequence or separately, to the Anopheles coluzzi (AcolM1.8) genome sequence using the Burrows-Wheeler Aligner BWA-MEM algorithm (http://bio-bwa.sourceforge.net/). Next, MarkDuplicatesSpark from the Broad’s Genome Analytics Toolkit Genome (GATK) was used to identify read pairs that likely originated from the same original DNA fragments due to artifacts (https://gatk.broadinstitute.org/hc/en-us/articles/360095814112-MarkDuplicatesSpark). Variants were identified by following the Broad Institute’s best practices workflow for germine short variant discovery (https://gatk.broadinstitute.org/hc/en-us/articles/360035553932-Germine-short-variant-discovery-SNPs-Indels). Variant calling was performed using the GATK HaplotypeCaller (https://gatk.broadinstitute.org/hc/en-us/articles/360037225632-HaplotypeCaller). Finally, hard filtering was performed in accordance with the Broad Institute’s recommendations for generic hard filtering (https://gatk.broadinstitute.org/hc/en-us/articles/360035389471-Hard-filtering-germine-short-variants). GATK version 4.1.8.1 was used for the analysis.

To obtain RNA-seq expression data from Sua-5B cells, the raw sequencing files were first obtained from VectorBase (https://vectorbase.org/vectorbase/app/record/dataset/DS_1c16f776df). The sample metadata was obtained from ftp://ngs.sanger.ac.uk/production/ag1000g/phase2/AR1/samples/45. (https://www.flyrnai.org/tools/fly2mosquito/web/download) as well as the pilot CRISPR screen result. Source data are provided with this paper. All plasmids and cell lines created in this work are available in repositories listed in Supplementary Data 1. Source data are provided with this paper.

Received: 18 April 2021; Accepted: 2 November 2021 Published online: 24 November 2021

References

1. World Health Organization. World Malaria Report 2020: 20 years of global progress and challenges. Geneva: World Health Organization; Licence: CC BY-NC-SA 3.0 IGO. (2020).
2. Bhatt, S. et al. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. Nature 526, 207–211 (2015).
3. Wilson, A. L. et al. The importance of vector control for the control and elimination of vector-borne diseases. PLoS Negl. Trop. Dis. 14, e0007831 (2020).
4. Ranson, H. & Lissenden, N. Insecticide resistance in African Anopheles mosquitoes: A worsening situation that needs urgent action to maintain malaria control. Trends. Parasitol. 32, 187–196 (2016).
5. Dusford, I. et al. Management of insecticide resistance in the major Aedes vectors of arboviruses: Advances and challenges. PLoS Negl. Trop. Dis. 13, 1–22 (2019).
6. Flores, H. A. & O’Neill, S. L. Controlling vector-borne diseases by releasing modified mosquitoes. Nat. Rev. Microbiol. 16, 508–518 (2018).
7. Shaw, W. R. & Catteruccia, F. Vector biology meet disease control: using basic research to fight vector-borne diseases. Nat. Microbiol. 4, 20–34 (2019).
8. Iturbe-Ormaetxe, I., Walker, T. & O’Neill, S. L. Wolbachia and the biological control of mosquito-borne disease. EMBO Rep. 12, 508–518 (2011).
9. Utarini, A. et al. Efficacy of Wolbachia-Infected Mosquito Deployment for the Control of Dengue. N. Engl. J. Med. 384, 2177–2186 (2021).
10. Hammond, A. et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. Nat. Biotechnol. 34, 78–83 (2015).
11. Kypour, K. et al. A CRISPR-Cas9 gene drive targeting doublesexes causes complete population suppression in caged Anopheles gambiae mosquitos. Nat. Biotechnol. 36, 1062–1071 (2018).
12. Gantz, V. M. et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proc. Natl. Acad. Sci. U. S. A. 112, E8736–E8743 (2015).
13. Adolli, A. et al. Efficient population modification gene-drive rescue system in the malaria mosquito Anopheles stephensi. Nat. Commun. 11, 5553 (2020).
14. Carballar-Lejarrazú, R. et al. Next-generation gene drive for population modification of the malaria vector mosquito, Anopheles gambiae. Proc. Natl. Acad. Sci. U. S. A. 117, 22805–22814 (2020).
15. Ford, K. Mcdonald, D. & Halli, P. Functional genomics via CRISPR–Cas. J. Mol. Biol. 431, 48–65 (2019).
16. Doerflinger, M., Forsyth, W., Ebert, G., Pellegrini, M. & Herold, M. J. CRISPR/Cas9—The ultimate weapon to battle infectious diseases? Cell. Microbiol. 19, 1–10 (2017).
17. Puschkin, A. S., Majzoub, K., Ooi, Y. S. & Carette, J. E. A CRISPR toolbox to study virus-host interactions. Nat. Rev. Microbiol. 15, 351–364 (2017).
18. Zhang, R. et al. A CRISPR screen defines a salt peptide processing pathway required by flaviviruses. Nature 535, 164–168 (2016).
19. Ma, H. et al. A CRISPR-based screen identifies genes essential for west-nile-virus-induced cell death. Cell. Rep. 12, 673–683 (2015).
20. Filomatori, C. V. et al. Dengue virus genomic variation associated with mosquito adaptation defines the pattern of viral non-coding RNAs and fitness in human cells. PLoS Pathog. 13, e1006294 (2017).
21. Blair, C. D. Mosquito RNAi is the major innate immune pathway controlling arbovirus infection and transmission. Future Microbiol. 6, 265–277 (2011).
22. Tassetto, M. et al. Control of RNA viruses in mosquito cells through the acquisition of dVNA and endogenous viral elements. Elife 8, 1–29 (2019).
23. Jose, J., Taylor, A. B. & Kuhn, R. J. Spatial and Temporal Analysis of arbovirus infection and transmission in the United Kingdom. Parasites Vectors 7, 8 (2014).
24. Acosta, E. G., Castilla, V. & Dameron, E. B. Infectious dengue-1 virus entry into mosquito C6/36 cells. Virus Res. 160, 173–179 (2011).
25. Londono-Renteria, B. et al. Dengue virus infection of aedes aegypti requires a putative cysteine rich venom protein. PLoS. Pathog. 11, 1–23 (2015).
27. Lombardo, F., Ghan, Y., Kafatos, F. C. & Christophides, G. K. Comprehensive Genetic Dissection of the Hemoocyte Immune Response in the Malaria Mosquito Anopheles gambiae. PLoS Pathog. 9, e1003145 (2013).

28. Schnettler, E. et al. Knockdown of piRNA pathway proteins results in enhanced semilak forest virus production in mosquito cells. J. Gen. Virol. 94, 1680–1689 (2013).

29. Ma, Q. et al. A mosquito small RNA genomics resource reveals dynamic evolution and host responses to viruses and transponders. Genome Res. 26, 1527–1537 (2016).

30. Sakamoto, J. M. & Azad, A. F. Propagation of arthropod-borne Rickettsia spp. in two mosquito cell lines. Appl. Environ. Microbiol. 73, 6637–6643 (2007).

31. O’Neill, M. A., Posner, B. A., Coates, C. J. & Abrams, J. M. A cell-based screening platform identifies novel mosquitoquiloid toxins. J. Biomed. Screen. 18, 688–694 (2013).

32. Kaudkutilan, A., Smith, D. R. & Berry, C. Investigation of the cry4B-prohibitin interaction in Aedes aegypti cells. Curr. Microbiol. 65, 446–454 (2012).

33. Ohlund, P., Lunden, H. & Blomström, A. L. Insect-specific virus evolution and potential effects on vector competence. Virus Genes. 55, 127–137 (2019).

34. Hughes, G. L. et al. Wolbachia Infections in Anopheles gambiae Cells: Transcriptomic Characterization of a Novel Host-Symbiont Interaction. PLoS Pathog. 7, e1002296 (2011).

35. Rozen-Gagnon, K., Yi, S., Jacobson, E., Novack, S. & Rice, C. M. A selectable, mutant system for engineering Drosophila melanogaster genes. BMC Biotechnol. 12, e144 (2012).

36. Azlan, A., Obeidat, S. M., Das, K. T., Yunus, M. A. & Azzam, G. Genome-wide RNAi with dengue and zika virus infection. Elife 8, e00144 (2019).

37. Lombardo, F., Ghani, Y., Kafatos, F. C. & Christophides, G. K. Comprehensive ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-27129-3 | www.nature.com/naturecommunications

38. Ma, Q. et al. A mosquito small RNA genomics resource reveals dynamic evolution and host responses to viruses and transponders. Genome Res. 26, 1527–1537 (2016).

39. Schmierer, B. et al. CRISPR/Cas9 screening using unique molecular identifiers. Mol. Syst. Biol. 13, 945 (2017).

40. Koen, D. S. et al. Short-hairpin RNA expressed from polymerase III promoters mediates RNA interference in mosquito cells. Infect. Mol. Biol. 16, 199–206 (2007).

41. Li, M. et al. Development of a configurable gene drive system in the human disease vector aedes aegypti. Elife 9, 1–22 (2020).

42. Rogers, S. L. & Rogers, G. C. Culture of Drosophila S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy. Nat. Protoc. 3, 606–611 (2008).

43. Shah, P. S. et al. Comparative Flavivirus-Host Protein Interaction Mapping Reveals Mechanisms of Dengue and Zika Virus Pathogenesis. Cell 1931–1945 (2018).

44. Taramoto, T., Huang, X., Armbuster, P. A. & Padmanabhan, R. Infection of Aedes albopictus Mosquito C6/36 Cells with the wMelPop Strain of Wolbachia Modulates Dengue Virus-Induced Host Cellular Transcripts and Induces Critical Sequence Alterations in the Dengue Viral Genome. J. Virol. 93, 1–21 (2019).

45. Zhao, J. et al. Transcriptome analysis of Aedes aegypti Aag2 cells in response to dengue virus-2 infection. Parases and Vectors 13, 1–14 (2020).

46. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. 34, 184–191 (2016).

47. Gebre, M., Nomburg, J. L. & Gewurz, B. E. Crispr–Cas9 genetic analysis of virus–Host interactions. Viruses 10, 1–22 (2018).

48. Oberhofer, G., Irving, T. & Bailey, A. Behavior of homing endonuclease gene drives targeting genes required for viability or female fertility with multiplexed guide RNAs. Proc. Natl. Acad. Sci. U. S. A. 115, E9343–E9352 (2018).

49. Port, F., Chen, H.-M., Lee, T. & Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proc. Natl. Acad. Sci. 111, E2967–E2976 (2014).

50. Okamoto, N. et al. A Melanogaster Transporter Is Required for Steroid Hormone uptake in Drosophila. Dev. Cell. 47, 294–305.e7 (2018).

51. Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. Elife 3, e03401 (2014).

52. Sessions, O. M. et al. Discovery of insect and human dengue virus host factors. Nature 458, 1047–1050 (2009).

53. Yasunaga, A. et al. Genome-wide RNAi screen identifies broadly-acting host factors that inhibit arbovirus infection. PLoS Pathog. 10, e1003914 (2014).

54. Madeira, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641 (2019).

55. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2:A multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191 (2009).

56. Gao, Z., Herrera-Carrillo, E. & Berkhour, B. Delination of the exact transcription termination signal for Type 3 Polymerase III. Mol. Ther. - Nucleic Acids 10, 36–44 (2018).

57. Bryant, A. V. & Matsumura, I. Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids. Biotechniques 48, 463–465 (2010).

58. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308 (2013).

59. Anderson, M. A. E., Gross, T. L., Myles, K. M. & Adelman, Z. N. Validation of novel promoter sequences derived from two endogenous ubiquitin genes in transgenic Aedes aegypti. Insect Mol. Biol. 19, 441–449 (2010).

60. Santolamazza, F. et al. Insertion polymorphisms of SINE200 retrotransposons within speciation islands of Anopheles gambiae. Malar. J. 7, 1–10 (2008).

61. Schwartz, J. J., Lee, C. & Shendure, J. Accurate gene synthesis with tag-directed retrieval of sequence-verified DNA molecules. Nat. Methods 9, 913–915 (2012).

62. Clarkson, C. S. et al. Genome variation and population structure among 1142 mosquitoes of the African malaria vector species Anopheles gambiae and Anopheles coluzzii. Genome. Res. 30, 1533–1546 (2020).

Acknowledgements
We thank Flammas Catteruccia and Nelson Lau for cell lines. We thank Daniela Silva Ayala for helpful discussions. The research reported in this paper was supported by NIH NIGMS grant P41GM132087 to N.P. N.P is an investigator of Howard Hughes Medical Institute.

Author contributions
RV and EM contributed equally to this work. RV, EM, YH, SEM and NP conceived the project. RV, EM, YH, SEM, TMC and NP contributed to the design of the experiments. JR, PM and YH developed the bioinformatic tools. RV, EM and FJS performed the experiments and contributed to the collection and analysis of data. EM wrote the first draft of the paper with input from RV, SEM and NP. All authors edited and approved the paper.

Competing interests
The authors declare no competing interests.
