Title (full – 200 characters)

Measuring the impact of genetic heterogeneity and chromosomal inversions on the efficacy of CRISPR-Cas9 gene drives in different strains of *Anopheles gambiae*

Title (short – 70 characters)

Small polymorphisms and large chromosomal inversions do not inhibit gene drive homing in *Anopheles gambiae*

Authors

Poppy Pescod[1], Giulia Bevivino[2], Amalia Anthousi[1,3,4], Ruth Shelton[1], Molly Margiotta[5], Fabrizio Lombardo[2], Tony Nolan[1]*

*corresponding author

Affiliations

1 Vector Biology Department, Liverpool School of Tropical Medicine, Liverpool, UK

2 Department of Public Health and Infectious Diseases, Division of Parasitology, University of Rome "la Sapienza", Rome, Italy

3 Department of Biology, University of Crete, Vassilika Vouton, 71409 Heraklion, Crete, Greece

4 Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas, Heraklion, Greece

5 Department of Life Sciences, Imperial College London, UK

Abstract (300 words)

The human malaria vector *Anopheles gambiae* is becoming increasingly resistant to insecticides, necessitating the development of novel control strategies including genetic control. CRISPR/Cas9 gene drives allow the modification or suppression of a target population via a highly specific gene targeting system; a double-stranded break is made in the target gene, triggering copying of the gene drive into the cut site and thus its super-Mendelian inheritance. The DNA repair mechanism responsible for copying the gene drive requires homology between the donor and recipient chromosomes on either side of the cut site; this presents challenges for the invasion of lab-developed gene drives into non-homogenous wild target populations. *An. gambiae* s.l. has a variable genome, and despite the choice of conserved target regions for gene drive development there may be single-nucleotide polymorphisms and large chromosomal inversions with unknown impacts on the gene drive rate.

Two independent gene drives (*vas2*-5958 and *zpg*-7280) targeting genes with an essential role in female fertility (AGAP005958 and AGAP007280 respectively) were introduced into three alternate *An. gambiae* s.l. strains originally collected from different regions of Africa. These strains contained 7.1-8.3% sequence heterology around the cut site, which was used to assess the impact of variation on homing rate. The *vas2*-5958 gene drive element is located within the large 2La chromosomal inversion; drive rates were compared in individuals heterozygous and homozygous for the chromosomal inversion to assess the impact of inverted chromosomal recombination on the gene drive. No decrease in gene drive rate was seen in any of the alternate strains nor in the 2La inversion heterozygotes compared to control rates. This supports the amenability of *An. gambiae* to control by gene drives and demonstrates the applicability of gene drives developed in lab strains to the control of wild mosquito populations.

Author Summary (150-200 words, non-technical)

Although malaria cases have dropped dramatically in the last decade due to interventions, the emergence of insecticide resistance and lack of consistency in control strategies has slowed progress. Organisations working towards malaria elimination have mandated the investigation of new technologies to control the malaria vector mosquitoes in the *Anopheles* genus, including genetic control. One such genetic control strategy involves the use of selfish genetic elements called gene drives that can spread through a population to either suppress or modify it. The use of gene drives for vector control could reduce the intervention efforts required to control mosquitoes and/or the *Plasmodium* parasite, while circumventing many of the issues of insecticide use. In this study we

investigate the capability of gene drives to spread into different strains of a species containing both small (single base) and large (chromosomal rearrangement) genetic differences. We conclude that, in comparison to similar studies in other organisms, homing-based gene drives in *Anopheles gambiae* are tolerant of a larger amount of heterogeneity in the target gene than expected with no reduction in efficacy.

1 Introduction

1.1 Gene drives in vector control

Global control efforts have averted an estimated 1.5 billion cases of malaria in the last two decades but this progress has begun to slow, with 619 000 deaths reported in 2021 alone (1). Malaria transmission persistence has been attributed to a combination of stalling or inadequate control programs, insecticide resistance of the mosquito vector, and treatment resistance of the parasite (2). The World Health Organisation has stressed the importance of developing novel control strategies to meet its malaria elimination goals (3, 4).

Genetic control strategies can achieve population modification or suppression of a target species without collateral damage to non-targets or the environment by the modification of the target genome, making these strategies highly desirable alternatives to widespread insecticide use. One such strategy is the use of selfish genetic elements with super-Mendelian inheritance rates known as gene drives. Gene drives can deliver a genetic payload or disrupt an essential gene while overcoming any subsequent fitness cost by severely biasing its own inheritance, allowing its spread in a population (5, 6, 7, 8, 9, 10). Strategies using gene drives are being considered for the control of several pest species (11, 12), and have progressed to the successful development of CRISPR/Cas9-based gene drives in the primary vector of malaria in Africa, *Anopheles gambiae* (7).

The Cas9 protein guided by an sgRNA is capable of making highly-specific double-stranded breaks (DSBs) in a chromosome, allowing the introduction of an alternate sequence at the cut site using the cell's own DNA repair mechanism (7). DSB repair by the cell can involve the use of a homologous template strand, usually the paired chromosome, which is copied to accurately repair the break (13, 14). When a gene drive element is copied into the broken chromosome alongside the homologous template sequence, the gene drive is inserted at the breakpoint in a process known as homing. Homing from one chromosome to another in germline cells means the gene drive will be integrated in the majority of gametes, resulting in its super-Mendelian inheritance in the next generation. This mechanism can be exploited to bias the inheritance of a coupled effector gene through a population, such as antimicrobial peptides to impede malaria development (10), or to target a gene essential for fertility and therefore reduce the target population size (7).

1.2 Gene drive resistance

The emergence of resistance to gene drives has been demonstrated empirically in synthetic drive constructs (15, 16). CRISPR-based gene drive resistance occurs as small genetic differences at the cut

site, reducing gRNA binding and therefore the ability of the Cas9 enzyme to create a DSB. These cut site mutations can arise during the DSB repair process via alternate repair pathways such as non-homologous end-joining, which enzymatically repairs the cut without a template but with higher rates of error (15, 17, 18, 19). If these genetic differences produce a functional allele with a fitness cost less than that of the gene drive, they may be positively selected for in the population. Functional alternate alleles produced by gene drive-induced mutations can reduce CRISPR-gRNA binding enough to confer complete resistance to the gene drive (15).

Strategies to reduce the likelihood of resistance developing have been suggested; modern gene drives will target genes which are haplosufficient (one functional copy is required for survival or fertility) and highly conserved, therefore making any mutations at the target site likely to result in unviability (8, 15, 20). This reduces the speed of gene drive resistance development but does not entirely prevent it; mutations produced during non-homologous end joining (NHEJ) will still eventually lead to resistance (15). NHEJ, and therefore related mutations conferring resistance, can be reduced by using more efficient germline-specific promoters with less accidental somatic expression of the Cas9 enzyme (21). Multiple target sites in different genes can be used in a single gene drive system by multiplexing gRNAs; homing can occur at all target sites, making independently-evolved resistance at all target sites necessary to prevent super-Mendelian inheritance of the gene drive (22, 23, 24).

Any intervention which applies a strong selection pressure will eventually produce a similarly strong and concomitant pressure to evolve resistance. Resistance has historically only been discovered after the implementation of a control strategy, and after the resistance has become a public health issue (25). By anticipating and investigating potential issues such as resistance during gene drive development we can reduce the impact on control strategies. Genetic variation at the target site, whether produced by Cas9-mediated NHEJ or naturally present in a target population, could act as a barrier to successful implementation of gene drives.

1.3 Genetic variation – a barrier to gene drive success?

1.3.1 Single nucleotide polymorphisms

Genetic differences around a gene drive target site, or target locus heterology (TLH), may occur naturally in a wild population even in highly functionally constrained genes. Differences within the gRNA target site have the most impact on drive efficiency (26), but due to the nature of DSB repair TLH will also potentially reduce the gene drive homing rate. Stringent regions of homology between the donor chromosome (containing the region to be copied) and the recipient chromosome (where

126 the DSB occurs) are required for homology-directed repair (HDR) in mammalian cells, where 1.2%
127 TLH within 1kb of the DSB causes a 22% reduction in the recombination required for HDR (27).
128 Similar dependence on homology has been noted in *Drosophila*, where 1.4% TLH suppressed
129 recombination by 32% (28); and *Aedes aegypti*, with 1.2% TLH created by experimental recoding
130 resulting in a 66% reduction in homing (29). Given the conserved nature of DNA repair mechanisms,
131 it is reasonable to expect that this sequence homology requirement would extend to *An. gambiae*,
132 which has an incredibly diverse genome including more than 57 million single-nucleotide
133 polymorphisms (SNPs) (30).

134

135 1.3.2 Chromosomal inversions

136 In addition to SNPs, the *An. gambiae* species complex contains over 120 chromosomal inversions
137 (31, 32). These inversions vary in their geographical and seasonal distribution and have been
138 associated with desiccation resistance, larval habitat, insecticide resistance, and malaria infection
139 rate (31, 33, 34, 35, 36, 37). The largest inversion in *An. gambiae* is the 2La/2L+^{a}, which spans
140 roughly half the length of chromosome 2L (38); the ancestral 2La form is implicated in anthropophilic
141 behaviour, aridity tolerance, and *Plasmodium* transmission (35, 39, 40, 41). Recombination of
142 inverted chromosomes in opposite orientations is reduced as the chromosomes are forced to form a
143 loop in order to align (42). Reduced recombination between inversion heterokaryotypes has been
144 empirically demonstrated during meiosis in multiple species, including *Drosophila* (7.7-fold decrease)
145 (43). The effect extends beyond the inversion breakpoints to suppress recombination in regions
146 close to the inversion and increase recombination at distant regions, known as the
147 interchromosomal effect (44), and can change the recombination landscape enough to suppress
148 recombination in homokaryotypes as well (45, 46).

149 A reduced recombination rate between inversion heterokaryotypes could theoretically interfere
150 with HDR in gene drive releases, leading to reduced spread of a gene drive situated within the
151 inversion into heterogenous wild populations. In an allelic drive system in *Drosophila*, inversion
152 heterokaryotypes had a drive rate a third lower than inversion homokaryotypes (47). Meiotic
153 recombination between 2La/2L+^{a} heterokaryotypes is at least 5-fold less than between 2L+^{a}
154 homokaryotypes (48). However, multiple gene drive systems have been developed within the 2La
155 inversion site in *An. gambiae* successfully, with super-Mendelian inheritance (76-98%) (49, 50). As
156 these were not developed with the 2La inversion karyotype in mind, or tested with different
157 karyotypes, the impact on recombination rate during gene drive homing in *An. gambiae* is still
158 unknown.

1.3.3 Variation outside of the target region

Genetic variation outside of the target region can also influence gene drive inheritance and resistance development. In a study of gene drive homing and resistance rates in different strains of *Drosophila*, all with identical target site sequences, inheritance rates ranged from 64.1-85.9%; increased inheritance was significantly associated with certain genotypes, but no SNPs were identified as contributing significantly (51). Moderate variation has been noted in gene drive homing in different genetic backgrounds of *Drosophila* despite little to no variation within 200 bp of the cut site (19). Background genetic variation can also influence the development of resistant alleles at the target site; no single gene was found to be significantly responsible for increased resistance development, indicating a combined effect of multiple genes (51). Differences in homing efficiency may be due to differences in a combination of genes, such as DNA repair mechanisms, DNA transcription or translation, or germline expression. In naturally occurring gene drives, suppressors can evolve to reduce the impact of the drive in the population; these are often unlinked to the gene drive, such as small RNAs or alterations in heterochromatin structure (52). Undoubtedly, the interaction between genetic variation and gene drive homing needs to be explored for their effective use in control strategies.

To assess the impact of TLH and inverted chromosomes on the homing of a gene drive element in *An. gambiae*, two lab-created gene drive strains *vas2*-5958 and *zpg*-7280 were crossed with three alternate *An. gambiae* wild type strains from across East, Central and West Africa (Kisumu, N'Gousso and Tiassale), all with TLH around the cut sites. The *vas2*-5958 gene drive element is located within the 2La inversion; gene drive homing rates were compared between 2La heterokaryotypes and homokaryotypes.

2 Results and discussion

2.1 *An. gambiae* gene drives are robust to TLH

The *vas2*-5958 and *zpg*-7280 *An. gambiae* gene drive lines, originally made in the G3 background and targeting haplosufficient female fertility genes, were crossed into three different strains to create F_1 hybrids which were backcrossed to wild type G3 to assess the F_1 hybrid homing rate (see Figure 1c). The TLH around the cut site was 7.1-8.3% between each strain and the gene drive background strain (G3). All F_1 hybrids for both gene drive colonies produced super-Mendelian inheritance rates of the gene drive element (*vas2*-5958: 81.8-100%; *zpg*-7280: 92.0-100%), with no significant difference from the control (Figure 2). No reduction in larval production was observed in *zpg*-7280 hybrids (Figure S2), suggesting no loss of fertility.

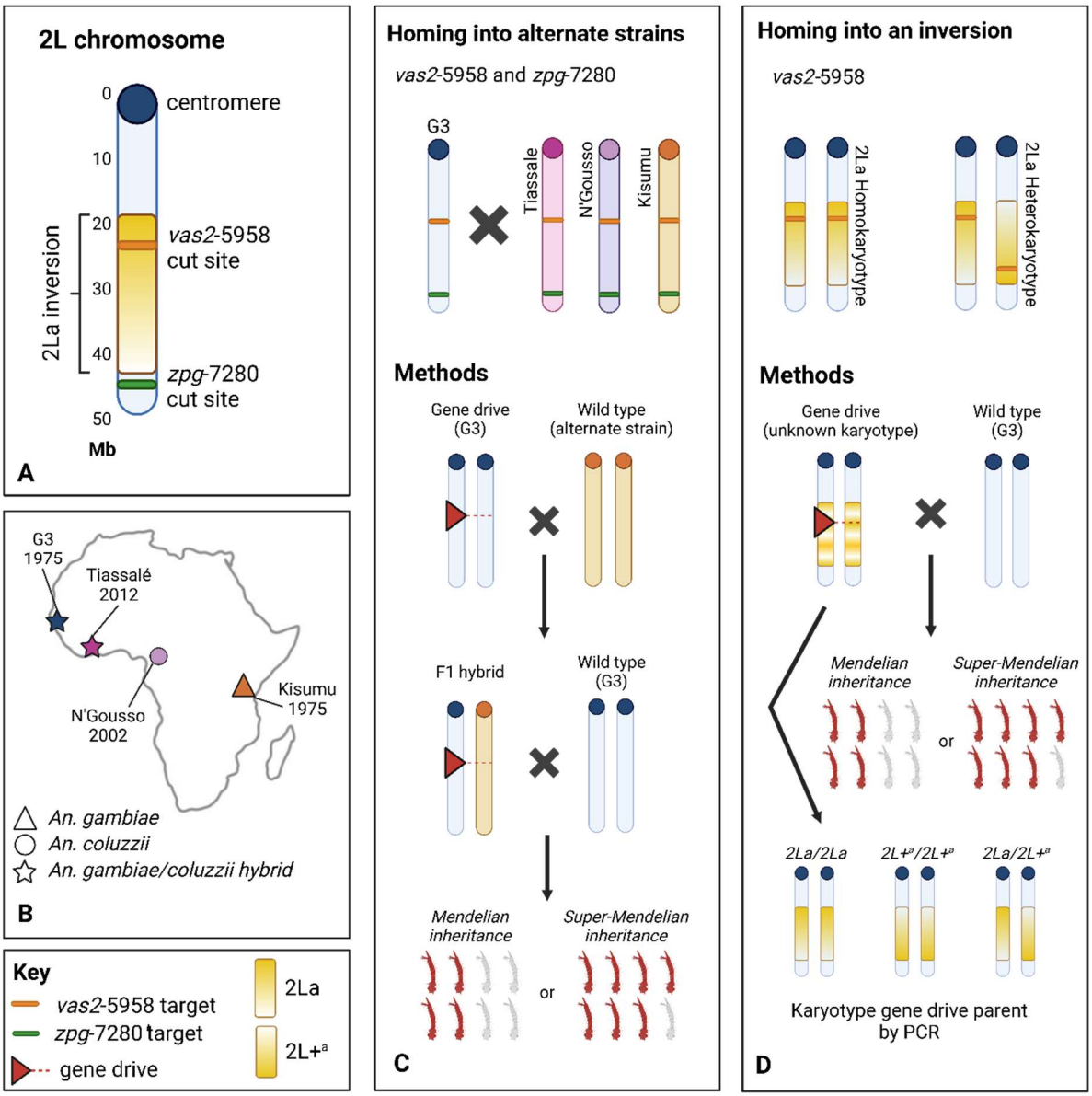

Figure 1 - Experimental homing of a gene drive into alternate genetic backgrounds.

A. The location of gene drive cassettes in the 2L chromosome of gene drive strains *vas2*-5958 and *zpg*-7280, relative to the centromere and the 2La inversion region.

B. The locations and dates of capture for each strain, and their species identification.

C. Methods for assessing gene drive homing into alternate strains. Gene drives were crossed with alternate strains to produce hybrid F_1 individuals; these were backcrossed to G3 and the F_2 offspring were screened for the RFP-tagged gene drive via fluorescent microscopy to determine the homing rate in the F_1 hybrid.

D. Methods for assessing gene drive homing in different inversion karyotypes. *vas2*-5958 gene drive individuals were crossed to wild type G3, the gene drive parent's 2La inversion karyotype was determined by PCR, and the offspring screened for inheritance of the gene drive to investigate the impact of chromosomal inversion on gene drive homing rate.

195

196

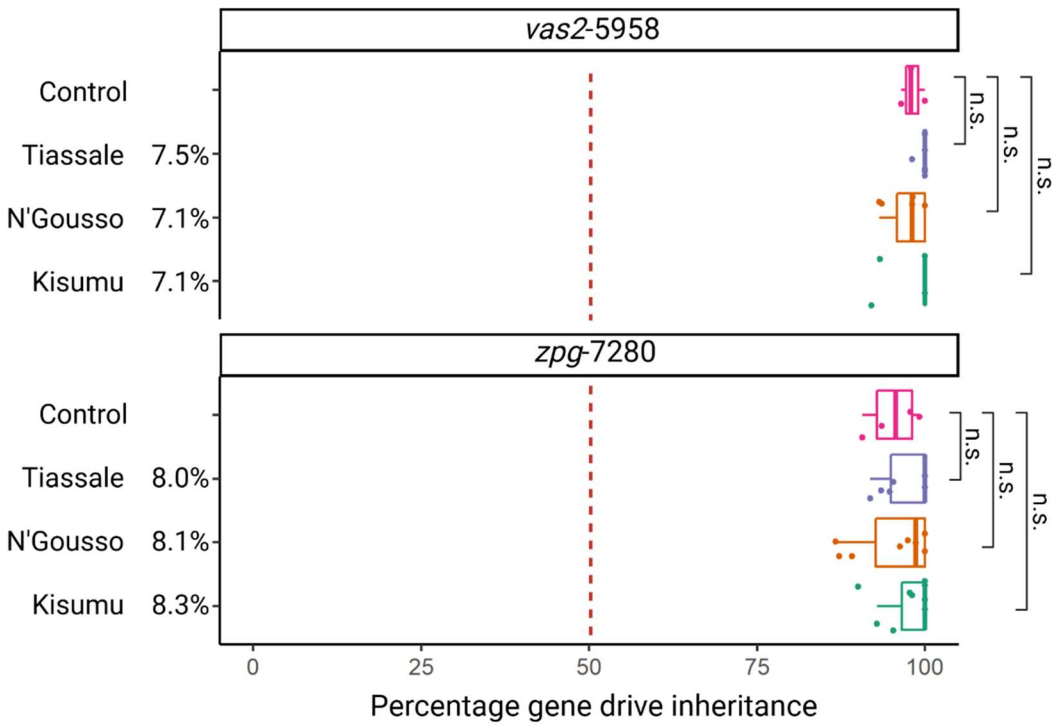

Figure 2 - The inheritance rate of two gene drive elements *vas2*-5958 and *zpg*-7280 in the offspring of F1 hybrids of three different strains, compared to the control rate of inheritance in the gene drive colony (G3 background). Target locus heterology ~600bp around the cut site between each strain and the G3 wild type is given in percentages next to strain names. Homing into alternate chromosomes produced drive rates which were not significantly different to the control drive rate (Pairwise Wilcoxon test, corrected for false discovery rate). n.s – non-significant.

The robustness of *An. gambiae* gene drive homing to variation has important consequences for its application in real-world vector control strategies. The development of gene drives in lab-bred mosquito strains allows for standardisation of the genetic background for easier study but has called into question their applicability to heterogenous wild populations. Despite the sensitivity of homing in other organisms to low amounts of TLH, our findings show no significant reduction in homing activity into multiple strains with up to 8.3% TLH in *An. gambiae*. The strains used in this experiment were collected from East, Central and West Africa across a span of 37 years, and are a mixture of *An. gambiae, An. coluzzi* and *An. gambiae/An. coluzzi* hybrids (Figure 1b). The demonstration of unimpeded gene drive homing into strains of this diversity represents the strong potential for gene drive implementation across members of the *An. gambiae* species complex that are able to produce fertile progeny.

2.2 No impact of 2La karyotype on gene drive homing

The *vas2*-5958 gene drive is located within the region covered by the 2La inversion (Figure 1a); homing rates for all three permutations of the 2La inversion were analysed (Figure 1d). There was no significant difference in homing rate between 2La inversion karyotypes (Figure 3).

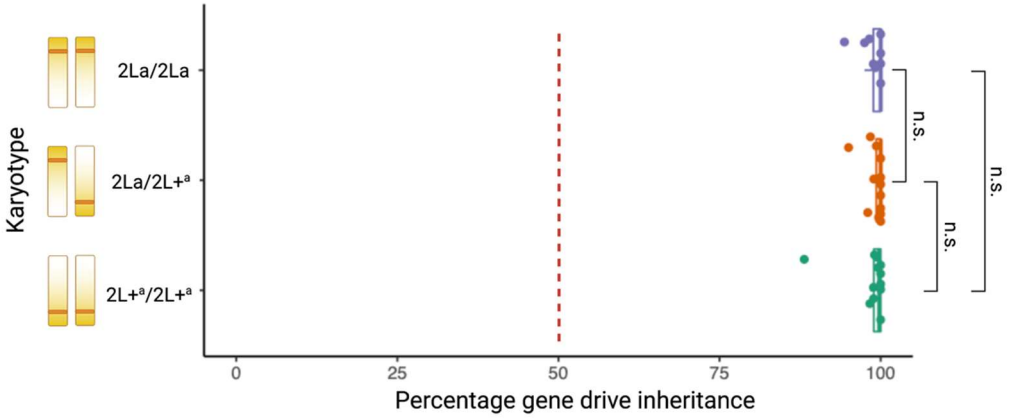

Figure 3 – The inheritance rate of the *vas2*-5958 gene drive element in the offspring of males either homozygous (2La/2La and 2L+ᵃ/2L+ᵃ) or heterozygous (2La/2L+a) for the 2La chromosomal inversion. There was no significant difference in gene drive inheritance between the three karyotypes (Wilcoxon test). n.s. – non-significant.

Despite previous observations of reduced gene drive conversion across inversions in *Drosophila* (47) and reduced meiotic recombination within the 2La inversion region in *An. gambiae* (48) we saw no evidence of reduced gene drive homing rate in 2La inversion heterokaryotypes. However, reduced recombination is not uniform across an inversion, and adjacent sequences external to the inversion can also show altered recombination rates. Meiotic recombination is slightly higher in the middle of the inversion compared to regions near the breakpoints, due to the increased ease of forming chiasmata between sister chromatids at the centre of the inversion loop (47, 48). The *vas2*-5958 gene drive target site is <4 Mb from the distal breakpoint of the 2La inversion (Figure 1) (7, 53); theoretically, recombination at this point should have been low, but this is not reflected in the gene drive rates we observed.

Adjacent to the inversion, the region between the proximal breakpoint of the 2La inversion and the centromere shows strong recombination reduction, with a less strong but still reduced recombination rate in the region distal to the centromere (48). The *zpg*-7280 gene drive target site is 2.8 Mb from the 2La distal breakpoint and is therefore located in a region with a known slight reduction in meiotic recombination (7, 53). Our results suggest that this is not sufficient to reduce

229 homing, but future work could explore if other targets within the inversion, or closer to the
230 breakpoint, may be affected.

231

232 While homing does not appear to be reduced within the inversion in *An. gambiae*, other impacts of
233 the inversion on long-term control strategies should be considered. Reduced meiotic recombination
234 results in protection of the inverted regions and their accumulation within populations; a common
235 mechanism of speciation in *Anopheles* (54). Linked regions can result in persistence of deleterious
236 mutations or the spread of adaptive alleles for certain environments. In the case of gene drives,
237 regardless of the impact of recombination on the homing mechanism itself, inversions could impact
238 the penetrance of gene drives into wild populations indirectly, through reproductive isolation. That
239 said, unless this reproductive isolation is total, even rare cases of intra-strain hybridisation should
240 lead to the gene drive rapidly introgressing into the new karyotype. There is good precedent for this
241 in the adaptive introgression of insecticide resistance alleles between *An. gambiae* and *An. coluzzii*,
242 two separate species that are not fully isolated reproductively (55). The idea of 'forced'
243 introgression, whereby gene drives are backcrossed into wild populations before release, has been
244 suggested to reduce the introduction of novel chromosomal arrangements or variation into wild
245 populations (56).

246

247 2.3 Application of gene drives to wild populations

248 At first glance, the high level of variation in the *An. gambiae* species complex suggests that gene
249 drives developed in lab-bred colonies could struggle to spread in wild populations via HDR. Our
250 results suggest that this is not the case; with a highly conserved gRNA, variation in the surrounding
251 sequence or in the chromosomal structure had no impact on the gene drive constructs tested here.
252 The use of highly conserved gRNA sites is an important strategy for reducing the development of
253 gene drive resistance (8). The availability of deep sequencing data for *An. gambiae* via the Ag1000G
254 confirms the high variation within the species complex, but also greatly improves our ability to
255 choose gRNAs appropriately (57). Correspondingly, gRNA target sites must be chosen carefully to
256 confine gene drives to a particular strain; there are a variety of self-limiting strategies currently in
257 development that either combine non-autonomous elements or target alleles private to the target
258 population (58, 59, 60).

259 The specificity of the gRNA targeting system produces very low off target effects in *An. gambiae*,
260 making CRISPR/Cas9 gene drives resistant to unexpected homing outside of the target sequence

261 (61). However, there is potential for neighbouring sequences flanking the gene drive to be carried
262 over during HDR due to resection of the broken chromosome (28). This could result in tight allelic
263 linkage of neighbouring sequences to the gene drive and introgression of novel alleles into wild
264 strains, suggesting that gRNA target regions need to be chosen with the surrounding sequences in
265 mind. Future work will be able to determine the precise dynamics of genetic exchange between the
266 gene drive donor and recipient chromosome.

267 Regardless of TLH of up to 8.3%, gene drive strategies for *An. gambiae* control show promising
268 efficacy for malaria control in wild mosquito populations. The self-sufficiency of gene drives after
269 initial release has meant extra care is being taken to characterise how gene drives will function in
270 natural settings (62, 63). This work offers improved understanding of gene drive dynamics in wild
271 populations and demonstrates their potential for *Anopheles* control.

272

3 Materials and Methods

3.1 Mosquito rearing

All mosquitoes were reared under standard conditions of 26 ± 2°C and 70 ± 10% relative humidity, with a 12 hour light/dark cycle with one hour dusks/dawns. Larvae were fed on ground fish food flakes (TetraMin® tropical flakes) and adults were fed 10% sucrose solution *ad libitum*. Adults were allowed to mate for 5-10 days before blood feeding and egg collection.

3.2 Mosquito strains

Two G3 colonies containing gene drive elements were used, both created by Hammond *et al.* (7, 21). The *vas2*-5958 colony contains a CRISPR/Cas9 endonuclease construct within AGAP005958, an ortholog of the *Drosophila yellow-g* gene expressed in somatic ovarian follicle cells with an unknown function (64). The AGAP005958 gene is located within the 2La inversion (54), with a gRNA cut site within 4Mb of the distal breakpoint. The *zpg*-7280 colony contains a similar construct in AGAP007280, ortholog of the *Drosophila nudel* gene also expressed in follicle cells with a known role in dorsoventral patterning of the developing embryo (65). Both genes have a haplosufficient role in female fertility, making them useful targets for population modification or suppression gene drive strategies.

The inserted construct for both colonies consists of: a CRISPR/Cas9 protein under germline-only promotion (*zpg* in the *zpg*-7280 line and *vas2* in the *vas2*-5958 line), a gRNA sequence targeting the cut site for each line respectively under U6 (universal) promotion, and a red fluorescence protein marker with a 3xP3 promoter, all flanked by *attB* recombination sites to allow insertion into previously created docking lines via recombinase-mediated cassette exchange (66). The full sequence of vector p165 used to produce these two lines, with the only difference between them the gRNA sequence, is available on GenBank (accession ID: KU189142) (7).

The wild type strains used in crosses were taken from colonies kept at the Liverpool School of Tropical Medicine; details can be found in Table 1.

Table 1 - Wild type Anopheles strains crossed to gene drive lines

Strain	Species	Place and date of collection	Reference
G3	*An. gambiae/An. coluzzii* hybrid	The Gambia, 1975	(7)
N'Gousso	*An. coluzzii*	Yaoundé, Cameroon 2002	(67)
Tiassale	*An. gambiae/An. coluzzii* hybrid	Cote d'Ivoire, 2012	(68, 69)
Kisumu	*An. gambiae*	Kenya, 1975	(70)

301 ## 3.3 Crosses of gene drive strains into alternate backgrounds

302 The number and sex of mosquitoes used in each cross can be seen in Table 2. All F_1 hybrid adults
303 used in crosses were confirmed to be heterozygous for the gene drive element by screening for the
304 RFP marker via fluorescent microscopy during the larval stage. Females containing the *vas2*-5958
305 gene drive are sterile due to unintended somatic promotion of Cas9 under the *vas2* promoter (7);
306 therefore, in *vas2*-5958 crosses only males containing the gene drive were used. For *zpg*-7280 F_1
307 crosses female hybrids were used. F_1 cross females were forced to lay in single deposition and up to
308 50 offspring per female were screened for the presence of the RFP marker to determine the rate of
309 gene drive in the hybrid parent. Drive rates were compared to data from Hammond *et al.* (7, 21)
310 using a pairwise Wilcoxon test with false discovery rate correction (Table S1).

311

312 Table 2 - Details on the number, sex and strain of each F0 and F1 cross

F_0 crosses	**F_1 crosses**
10 ♂ *vas2*-5958 x 25 ♀ N'Gousso	40 ♂ *vas2*-5958/N'Gousso F_1 x 20 ♀ G3
10 ♂ *vas2*-5958 x 25 ♀ Kisumu	40 ♂ *vas2*-5958/Kisumu F_1 x 20 ♀ G3
10 ♂ *vas2*-5958 x 42 ♀ Tiassale	40 ♂ *vas2*-5958/Tiassale F_1 x 20 ♀ G3
10 ♂ *zpg*-7280 x 25 ♀ N'Gousso	20 ♀ *zpg*-7280/N'Gousso F_1 x 40 ♂ G3
10 ♂ *zpg*-7280 x 25 ♀ Kisumu	20 ♀ *zpg*-7280/Kisumu F_1 x 40 ♂ G3
10 ♂ *zpg*-7280 x 42 ♀ Tiassale	20 ♀ *zpg*-7280/Tiassale F_1 x 40 ♂ G3

313

314 ## 3.4 Target site sequence heterology

315 DNA was extracted from pools of 33-37 larvae, one pool from each strain, using a Wizard® genomic
316 DNA purification kit (Promega) and a region of ~690bp spanning the gRNA cut sites for both *vas2*-
317 5958 and *zpg*-7280 gene drives was amplified in two fragments either side of the gene drive insert.
318 Fragments were amplified by PCR using Phusion Hot Start II High-Fidelity DNA polymerase (Thermo
319 Scientific™), with forward and reverse primers at a final concentration of 0.5 µM (Table S2) and 1 µl
320 genomic DNA in a 50 µl reaction. PCR conditions were: an initial denaturation step at 98°C for 30
321 seconds; followed by 30 cycles of denaturation at 98°C for 30 seconds, 30 seconds at the annealing
322 temperature (Table S2), and extension at 72°C for 15 seconds; and a final extension step of 10
323 minutes at 72°C.

324 PCR products were sequenced by Illumina MiSeq sequencing; reads were quality filtered and aligned
325 against an amplicon containing all SNP variants present in G3 deep sequencing data (15) using
326 CRISPResso (71). Alleles present at >1% relative abundance were aligned to G3 sequences in

327 Benchling to determine the percentage of small polymorphisms between G3 and each strain at the
328 homing sites (Table S3) (raw files accession: PRJNA914102).

329 3.5 Homing rate analysis in alternative 2La karyotypes

330 Mosquitoes from the *vas2*-5958 colony were backcrossed to G3 and offspring were screened for the
331 gene drive marker; 65 F1 males were mated individually to three G3 females, with eggs collected
332 from each group and screened for the gene drive element. Each male parent was karyotyped for the
333 2La inversion as previously described (38), and drive rates in heterozygotes and homozygotes of
334 both karyotypes were compared using a Wilcoxon test.

335

336 4 Supporting information captions

337 **S1 Appendix: Analysis of a fourth strain, Busia.**

338 **Figure S1: Gene drive inheritance rate analysis of Kisumu, N'Gousso, Tiassale and Busia hybrids**
339 **with *zpg*-7280 and *vas2*-5958 gene drives.** TLH is given as a percentage next to each strain name;
340 statistical analysis was conducted using a pairwise Wilcoxon test with false discovery rate correction.
341 n.s. – non-significant.

342 **Table S1: Pairwise Wilcoxon test p values with false discovery rate correction for gene drive**
343 **inheritance rate.** Comparisons were performed for hybrids with all four strains of both gene drives.
344 No comparisons were statistically significant.

345 **Table S2: Primers used for amplicon sequencing of TLH regions**. Sequences in brackets indicate
346 Illumina adapter sequences (not included in annealing temperature calculation). WT = wild type, GD
347 = gene drive.

348 **Table S3: TLH for Kisumu, N'Gousso, Tiassale and Busia compared to G3.** TLH was calculated by
349 comparing combined sequences from each strain to wild type G3 sequences and calculating the
350 percentage of base pairs differing from the wild type. For the additional fourth strain, Busia, only the
351 309 bp sequence to the left of the *vas2*-5958 cut site was used to calculate TLH.

352 **Figure S2: Comparison of larvae number produced by single females from hybrids of *zpg*-7280 and**
353 **four different strains**. Control larvae numbers were from Hammond *et al.* (2021) (n=66). Tiassale
354 n=10, N'Gousso n=11, Kisumu n=11, Busia n=11. No significant difference in larval production was
355 found between any strain (pairwise t test, $p > 0.05$ for all comparisons), suggesting that there was no
356 reduction in fertility in gene drive/alternate strain hybrids.

5 References

1. World Health Organization. World Malaria Report 2022. Geneva; 2022.
2. Rodriguez MH. Residual Malaria: Limitations of Current Vector Control Strategies to Eliminate Transmission in Residual Foci. The Journal of Infectious Diseases. 2021;223(Supplement_2):S55-S60.
3. World Health Organization. Evaluation of genetically modified mosquitoes for the control of vector-borne diseases. Geneva; 2020.
4. World Health Organization. World malaria report 2020: 20 years of global progress and challenges. Geneva; 2020.
5. Burt A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. Proceedings of the Royal Society of London Series B: Biological Sciences. 2003;270(1518):921-8.
6. Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. Proceedings of the National Academy of Sciences. 2015;112(49):E6736-E43.
7. Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. Nature Biotechnology. 2016;34(1):78-83.
8. Kyrou K, Hammond AM, Galizi R, Kranjc N, Burt A, Beaghton AK, et al. A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes. Nature Biotechnology. 2018;36(11):1062-6.
9. Windbichler N, Menichelli M, Papathanos PA, Thyme SB, Li H, Ulge UY, et al. A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. Nature. 2011;473(7346):212-5.
10. Hoermann A, Habtewold T, Selvaraj P, Del Corsano G, Capriotti P, Inghilterra MG, et al. Gene drive mosquitoes can aid malaria elimination by retarding *Plasmodium* sporogonic development. Science Advances. 2022;8(38):eabo1733.
11. Legros M, Marshall JM, Macfadyen S, Hayes KR, Sheppard A, Barrett LG. Gene drive strategies of pest control in agricultural systems: Challenges and opportunities. Evolutionary Applications. 2021;14(9):2162-78.
12. Scott MJ, Gould F, Lorenzen M, Grubbs N, Edwards O, O'Brochta D. Agricultural production: assessment of the potential use of Cas9-mediated gene drive systems for agricultural pest control. Journal of Responsible Innovation. 2018;5(sup1):S98-S120.
13. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. Trends in Cell Biology. 2016;26(1):52-64.
14. Xue C, Greene EC. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends in Genetics. 2021;37(7):639-56.
15. Hammond AM, Kyrou K, Bruttini M, North A, Galizi R, Karlsson X, et al. The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito. PLOS Genetics. 2017;13(10):e1007039.
16. Beaghton AK, Hammond A, Nolan T, Crisanti A, Burt A. Gene drive for population genetic control: non-functional resistance and parental effects. Proceedings of the Royal Society B: Biological Sciences. 2019;286(1914):20191586.
17. Chang HH, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nature Reviews Molecular Cell Biology. 2017;18(8):495-506.
18. Unckless RL, Clark AG, Messer PW. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. Genetics. 2017;205(2):827-41.

19. Champer J, Reeves R, Oh SY, Liu C, Liu J, Clark AG, et al. Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. PLoS Genetics. 2017;13(7):e1006796.
20. Champer J, Liu J, Oh SY, Reeves R, Luthra A, Oakes N, et al. Reducing resistance allele formation in CRISPR gene drive. Proceedings of the National Academy of Sciences. 2018;115(21):5522-7.
21. Hammond A, Karlsson X, Morianou I, Kyrou K, Beaghton A, Gribble M, et al. Regulating the expression of gene drives is key to increasing their invasive potential and the mitigation of resistance. PLOS Genetics. 2021;17(1):e1009321.
22. Marshall JM, Buchman A, Sánchez C HM, Akbari OS. Overcoming evolved resistance to population-suppressing homing-based gene drives. Scientific Reports. 2017;7(1):3776.
23. Champer J, Yang E, Lee E, Liu J, Clark AG, Messer PW. A CRISPR homing gene drive targeting a haplolethal gene removes resistance alleles and successfully spreads through a cage population. Proceedings of the National Academy of Sciences. 2020;117(39):24377-83.
24. Khatri BS, Burt A. A theory of resistance to multiplexed gene drive demonstrates the significant role of weakly deleterious natural genetic variation. Proceedings of the National Academy of Sciences. 2022;119(32):e2200567119.
25. Denholm I, Pickett JA, Devonshire AL, McKenzie JA, Batterham P. Predicting insecticide resistance: mutagenesis, selection and response. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences. 1998;353(1376):1729-34.
26. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34(2):184-91.
27. Elliott B, Jasin M. Repair of Double-Strand Breaks by Homologous Recombination in Mismatch Repair-Defective Mammalian Cells. Molecular and Cellular Biology. 2001;21(8):2671-82.
28. Do AT, Brooks JT, Le Neveu MK, LaRocque JR. Double-Strand Break Repair Assays Determine Pathway Choice and Structure of Gene Conversion Events in *Drosophila melanogaster*. G3 Genes|Genomes|Genetics. 2014;4(3):425-32.
29. Ang JXD, Nevard K, Ireland R, Purusothaman D-K, Verkuijl SA, Shackleford L, et al. Considerations for homology-based DNA repair in mosquitoes: Impact of sequence heterology and donor template source. PLoS genetics. 2022;18(2):e1010060.
30. Clarkson CS, Miles A, Harding NJ, Lucas ER, Battey C, Amaya-Romero JE, et al. Genome variation and population structure among 1142 mosquitoes of the African malaria vector species *Anopheles gambiae* and *Anopheles coluzzii*. Genome Research. 2020;30(10):1533-46.
31. Coluzzi M, Sabatini A, Petrarca V, Di Deco M. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Transactions of the Royal Society of tropical Medicine and Hygiene. 1979;73(5):483-97.
32. Coluzzi M, Sabatini A, della Torre A, Di Deco MA, Petrarca V. A polytene chromosome analysis of the *Anopheles gambiae* species complex. Science. 2002;298(5597):1415-8.
33. Gray EM, Rocca KA, Costantini C, Besansky NJ. Inversion 2La is associated with enhanced desiccation resistance in *Anopheles gambiae*. Malaria Journal. 2009;8(1):1-12.
34. Manoukis NC, Powell JR, Touré MB, Sacko A, Edillo FE, Coulibaly MB, et al. A test of the chromosomal theory of ecotypic speciation in *Anopheles gambiae*. Proceedings of the National Academy of Sciences. 2008;105(8):2940-5.
35. Petrarca V, Beier JC. Intraspecific chromosomal polymorphism in the *Anopheles gambiae* complex as a factor affecting malaria transmission in the Kisumu area of Kenya. The American Journal of Tropical Medicine and Hygiene. 1992;46(2):229-37.
36. Brooke B, Hunt R, Chandre F, Carnevale P, Coetzee M. Stable chromosomal inversion polymorphisms and insecticide resistance in the malaria vector mosquito *Anopheles gambiae* (Diptera: Culicidae). Journal of Medical Entomology. 2002;39(4):568-73.

455 37. Ayala D, Zhang S, Chateau M, Fouet C, Morlais I, Costantini C, et al. Association mapping
456 desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles*
457 *gambiae*. Molecular Ecology. 2019;28(6):1333-42.
458 38. White BJ, Santolamazza F, Kamau L, Pombi M, Grushko O, Mouline K, et al. Molecular
459 karyotyping of the 2La inversion in *Anopheles gambiae*. The American Journal of Tropical Medicine
460 and Hygiene. 2007;76(2):334-9.
461 39. Powell J, Petrarca V, Della Torre A, Caccone A, Coluzzi M. Population structure, speciation,
462 and introgression in the *Anopheles gambiae* complex. Parassitologia. 1999;41(1-3):101-13.
463 40. Sharakhov IV, White BJ, Sharakhova MV, Kayondo J, Lobo NF, Santolamazza F, et al.
464 Breakpoint structure reveals the unique origin of an interspecific chromosomal inversion (2La) in the
465 *Anopheles gambiae* complex. Proceedings of the National Academy of Sciences. 2006;103(16):6258-
466 62.
467 41. Fouet C, Gray E, Besansky NJ, Costantini C. Adaptation to aridity in the malaria mosquito
468 *Anopheles gambiae*: chromosomal inversion polymorphism and body size influence resistance to
469 desiccation. PloS ONE. 2012;7(4):e34841.
470 42. Kirkpatrick M. How and why chromosome inversions evolve. PLoS Biology.
471 2010;8(9):e1000501.
472 43. Crown KN, Miller DE, Sekelsky J, Hawley RS. Local Inversion Heterozygosity Alters
473 Recombination throughout the Genome. Current Biology. 2018;28(18):2984-90.e3.
474 44. Lucchesi JC, Suzuki DT. The interchromosomal control of recombination. Annual Review of
475 Genetics. 1968;2(1):53-86.
476 45. Hunter CM, Huang W, Mackay TFC, Singh ND. The Genetic Architecture of Natural Variation
477 in Recombination Rate in *Drosophila melanogaster*. PLOS Genetics. 2016;12(4):e1005951.
478 46. Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. Variation in recombination
479 frequency and distribution across eukaryotes: patterns and processes. Phil Trans R Soc B.
480 2017;372(1736):20160455.
481 47. Guichard A, Haque T, Bobik M, Xu X-RS, Klanseck C, Kushwah RBS, et al. Efficient allelic-drive
482 in Drosophila. Nature Communications. 2019;10(1):1640.
483 48. Stump A, Pombi M, Goeddel L, Ribeiro J, Wilder J, Torre A, et al. Genetic exchange in 2La
484 inversion heterokaryotypes of *Anopheles gambiae*. Insect Molecular Biology. 2007;16(6):703-9.
485 49. Hoermann A, Tapanelli S, Capriotti P, Del Corsano G, Masters EKG, Habtewold T, et al.
486 Converting endogenous genes of the malaria mosquito into simple non-autonomous gene drives for
487 population replacement. eLife. 2021;10:e58791.
488 50. Ellis DA, Avraam G, Hoermann A, Wyer CAS, Ong YX, Christophides GK, et al. Testing non-
489 autonomous antimalarial gene drive effectors using self-eliminating drivers in the African mosquito
490 vector Anopheles gambiae. PLOS Genetics. 2022;18(6):e1010244.
491 51. Champer J, Wen Z, Luthra A, Reeves R, Chung J, Liu C, et al. CRISPR Gene Drive Efficiency and
492 Resistance Rate Is Highly Heritable with No Common Genetic Loci of Large Effect. Genetics.
493 2019;212(1):333-41.
494 52. Courret C, Chang C-H, Wei KH-C, Montchamp-Moreau C, Larracuente AM. Meiotic drive
495 mechanisms: lessons from *Drosophila*. Proc R Soc B. 2019;286(1913).
496 53. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, et al. The Genome
497 Sequence of the Malaria Mosquito *Anopheles gambiae*. Science. 2002;298(5591):129-49.
498 54. Lee Y, Collier TC, Sanford MR, Marsden CD, Fofana A, Cornel AJ, et al. Chromosome
499 Inversions, Genomic Differentiation and Speciation in the African Malaria Mosquito *Anopheles*
500 *gambiae*. PLoS ONE. 2013;8(3):1-12.
501 55. Clarkson CS, Weetman D, Essandoh J, Yawson AE, Maslen G, Manske M, et al. Adaptive
502 introgression between *Anopheles* sibling species eliminates a major genomic island but not
503 reproductive isolation. Nature Communications. 2014;5(1):4248.

504 56. Pollegioni P, Persampieri T, Minuz RL, Bucci A, Trusso A, Martino SD, et al. Introgression of a
505 synthetic sex ratio distortion transgene into different genetic backgrounds of *Anopheles coluzzii*.
506 Insect Molecular Biology. 2023;32(1):56-68.
507 57. The *Anopheles gambiae* 1000 Genomes Consortium. Ag1000G. In: MalariaGEN, editor.
508 https://www.malariagen.net/mosquito/ag1000g2014.
509 58. Noble C, Min J, Olejarz J, Buchthal J, Chavez A, Smidler AL, et al. Daisy-chain gene drives for
510 the alteration of local populations. Proceedings of the National Academy of Sciences.
511 2019;116(17):8275-82.
512 59. Dhole S, Lloyd AL, Gould F. Tethered homing gene drives: A new design for spatially
513 restricted population replacement and suppression. Evolutionary Applications. 2019;12(8):1688-702.
514 60. Willis K, Burt A. Double drives and private alleles for localised population genetic control.
515 PLOS Genetics. 2021;17(3):e1009333.
516 61. Garrood WT, Kranjc N, Petri K, Kim DY, Guo JA, Hammond AM, et al. Analysis of off-target
517 effects in CRISPR-based gene drives in the human malaria mosquito. Proceedings of the National
518 Academy of Sciences. 2021;118(22):e2004838117.
519 62. James SL, Marshall JM, Christophides GK, Okumu FO, Nolan T. Toward the definition of
520 efficacy and safety criteria for advancing gene drive-modified mosquitoes to field testing. Vector-
521 Borne and Zoonotic Diseases. 2020;20(4):237-51.
522 63. Long KC, Alphey L, Annas GJ, Bloss CS, Campbell KJ, Champer J, et al. Core commitments for
523 field trials of gene drive organisms. Science. 2020;370(6523):1417-9.
524 64. Claycomb JM, Benasutti M, Bosco G, Fenger DD, Orr-Weaver TL. Gene Amplification as a
525 Developmental Strategy: Isolation of Two Developmental Amplicons in *Drosophila*. Developmental
526 Cell. 2004;6(1):145-55.
527 65. Hong CC, Hashimoto C. The maternal nudel protein of *Drosophila* has two distinct roles
528 important for embryogenesis. Genetics. 1996;143(4):1653-61.
529 66. Bateman JR, Lee AM, Wu C-T. Site-specific transformation of *Drosophila* via φC31 integrase-
530 mediated cassette exchange. Genetics. 2006;173(2):769-77.
531 67. Harris C, Lambrechts L, Rousset F, Abate L, Nsango SE, Fontenille D, et al. Polymorphisms in
532 Anopheles gambiae Immune Genes Associated with Natural Resistance to Plasmodium falciparum.
533 PLOS Pathogens. 2010;6(9):e1001112.
534 68. Edi CV, Koudou BG, Jones CM, Weetman D, Ranson H. Multiple-insecticide resistance in
535 Anopheles gambiae mosquitoes, Southern Côte d'Ivoire. Emerg Infect Dis. 2012;18(9):1508-11.
536 69. Williams J, Flood L, Praulins G, Ingham VA, Morgan J, Lees RS, et al. Characterisation of
537 Anopheles strains used for laboratory screening of new vector control products. Parasites & Vectors.
538 2019;12(1):522.
539 70. Reid JA. Pupal differences between species A and B of the *Anopheles gambiae* group from
540 Kisumu, East Africa. Mosq Syst. 1975;7:1-7.
541 71. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides
542 accudate and rapid genome editing sequence analysis. Nature Biotechnology. 2019;37(3):224-6.

543

544