Population suppression of the malaria vector Anopheles gambiae by gene drive technology: A large-cage indoor study bridging the gap between laboratory and field testing

Andrew Hammond
Imperial College London https://orcid.org/0000-0002-1757-5009
Paola Pollegioni
National Research Council
Tania Persampieri
Polo d'Innovazione di Genomica, Genetica e Biologia
Ace North
https://orcid.org/0000-0002-4253-396X
Roxana Minuz
Polo d'Innovazione di Genomica, Genetica e Biologia
Alessandro Trusso
Polo d'Innovazione di Genomica, Genetica e Biologia
Alessandro Bucci
Polo d'Innovazione di Genomica, Genetica e Biologia
Kyros Kyrou
Imperial College London https://orcid.org/0000-0002-4919-2459
Ioanna Morianou
Imperial College of London
Alekos Simoni
Polo d'Innovazione di Genomica, Genetica e Biologia
Tony Nolan
Imperial College https://orcid.org/0000-0002-2982-8333
Ruth Müller (rmuller@itg.be)
Institute of Tropical Medicine Antwerp https://orcid.org/0000-0003-3909-3876
Andrea Crisanti
Imperial College London, University of Padova

Article

Keywords: Anopheles gambiae, population suppression, gene drive technology

Posted Date: April 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-411410/v1

Version of Record: A version of this preprint was published at Nature Communications on July 28th, 2021. See the published version at https://doi.org/10.1038/s41467-021-24790-6.

Population suppression of the malaria vector *Anopheles gambiae* by gene drive technology: A large-cage indoor study bridging the gap between laboratory and field testing

Andrew Hammond[1,2,¶], **Paola Pollegioni**[3,4¶], **Tania Persampieri**[3¶], Ace North[5], Roxana Minuz[3], Alessandro Trusso[3], Alessandro Bucci[3], Kyros Kyrou[1], Ioanna Morianou[1], Alekos Simoni[1,3], **Tony Nolan**[1,6,*], **Ruth Müller**[3,7,8*], **Andrea Crisanti**[1,*]

¶ shared first authorship

* shared last authorship, corresponding authors

[1] Department of Life Sciences, Imperial College London, London, UK

[2] Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA

[3] Polo d'Innovazione di Genomica, Genetica e Biologia, Genetics & Ecology Platform, Terni, Italy

[4] National Research Council, Research Institute on Terrestrial Ecosystems, Porano, Terni, Italy

[5] Department of Zoology, University of Oxford, Oxford, UK

[6] Department of Vector Biology, Liverpool School of Tropical Medicine, Liverpool, UK

[7] Goethe University, Institute of Occupational, Social and Environmental Medicine, Frankfurt am Main, Germany

[8] Department of Biomedical Sciences, Institute of Tropical Medicine Antwerp, Belgium

Abstract

CRISPR-based gene drives are self-sustaining genetic elements that have been recently generated in the laboratory with the aim to develop potent genetic vector control measures targeting disease vectors including *Anopheles gambiae*. We have shown that a gene drive directed against the gene *doublesex* (dsx) effectively suppressed the reproductive capability of mosquito populations reared in small laboratory cages. These experiments, though informative, do not recapitulate the complexity of mosquito behaviour in natural environments. Additional information is needed to bridge the gap between laboratory and the field to validate the vector control potential of the technology.

We have investigated the suppressing activity of dsx gene drive strain Ag(QFS)1 on age-structured populations of *Anopheles gambiae* in large indoor cages that provide a more challenging ecology by more closely mimicking natural conditions and stimulating complex mosquito behaviours. Under these conditions, the Ag(QFS)1 drive spreads rapidly from a single release to the indoor large-cage populations at low initial frequency, leading to full population suppression within one year and without inducing resistance to the gene drive.

Initial stochastic simulations of the expected population dynamics, as based on life history parameters estimated in small cages, did not fully capture the observed dynamics in the large cages. Thus, we used the method of approximate Bayesian computation to better estimate population dynamics in the more realistic ecological setting in large cages, allowing the mosquitoes to show a complex feeding and reproductive behaviour.

Together, these results establish a new paradigm for generating data to bridge laboratory and field studies, and form an essential component in the stepwise and sound development of gene drive based vector control tools.

Introduction

CRISPR-based gene drives are selfish genetic elements that can be used to modify entire populations of the malaria mosquito for sustainable vector control. First proposed in 2003, these elements use a mechanism of cut and paste ('homing') in the germline to facilitate their autonomous spread from a very low initial release frequency (Burt, 2003; Windbichler et al. 2011). One potentially powerful strategy aims to reduce the total number of mosquitoes by spreading a mutation that blocks female reproduction. To be effective for the control of malaria in sub-Saharan Africa, such a strain must be able to compete effectively with wild populations of *Anopheles gambiae*, and remain effective over the medium to long-term. To this end, we and others have adopted a step-wise approach to the development and testing of gene drives in progressively rigorous and challenging conditions.

First generation suppression drives failed to maintain their spread when tested in small, caged-population experiments within an Arthropod Containment Level 2 laboratory because of the creation and selection of drive-resistant alleles, sometimes exacerbated by unintended fitness costs in 'carrier' individuals (Hammond et al. 2016; Hammond et al. 2017; KaramiNejadRanjbar et al. 2018; Pham et al. 2019). One strategy to mitigate against the likelihood of target-site resistance arising is to target sequences that show high levels of functional constraint and can therefore not easily tolerate variant alleles. We recently demonstrated the success of this approach by developing a second generation gene drive, herein named Ag(QFS)1 (previously called $dsxF^{CRISPRh}$), that has been used to suppress entire populations of caged mosquitoes in proof-of-principle experiments (Kyrou et al. 2018). This gene drive is designed to target an ultra-conserved, essential sequence within the female-specific isoform of the gene *doublesex,* encoding a transcription factor that is the major regulator of sex determination in insects (Verhulst & Van de zande 2015; Kyrou et al 2018). Females homozygous for the gene drive display female-male sexual development (intersex) and cannot produce offspring. This strategy has proven effective for two independent gene drive designs, each tested by tracking invasion dynamics over time following single, low frequency introductions in six discrete-generation laboratory populations (Kyrou et al. 2018, Simoni et al. 2020).

Typically, the development of candidate gene drive strains for potential vector control involves assessment of basic parameters concerning both fitness and drive, such as the homing rate, life-span and fecundity, however these parameters are notoriously difficult to estimate and often context-specific. Promising strains are then tested to determine if the gene drive can spread in small caged populations, and to compare invasion dynamics with prediction. This initial testing is key for identifying promising candidate gene drive strains, however it provides information of limited predictive value as these experiments do not take into account age structured populations, complex mating behaviours, differing probabilities of finding food resources, oviposition sites, and mating opportunities. Indeed, previously developed genetically modified mosquito strains have shown strong

fitness costs when tested in large-cage or semi-field experiments - we refer to these herein as 'releases' since this is what they are designed to emulate, albeit they are performed in fully contained chambers that comply with appropriate arthropod containment guidelines - that were not observed in initial small cage testing (Aldersley et al. 2019), including severe mating disadvantages that precluded further testing of the strain (Facchinelli et al. 2013). Many of these fitness challenges and complex behaviours can be reproduced in large cages (Facchinelli et al. 2015) by allowing overlapping generations so as to reveal potential differences in life-span and fecundity over time that cannot be captured in discrete generation studies (Facchinelli et al. 2019; Pham et al. 2019; Pollegioni et al. 2020). As such, large-cage release experiments are now considered an essential bridge between laboratory and field testing within the tiered testing approach (Benedict et al. 2008; NASEM, 2016; James 2018; Facchinelli et al. 2019; James 2020).

The Ag(QFS)1 strain is designed to make homozygous gene drive females infertile, and so it is dependent upon high fitness in males and in heterozygous 'carrier' females (where the gene drive is designed to be active in the germline) to ensure it increases in frequency in the population. Initial testing of the strain revealed a reduction in the fertility of heterozygous females that is likely due to 'leaky' activity of the gene drive in the soma, leading to a mosaic pattern of knockout of the *doublesex* target gene (Kyrou et al. 2018). As *doublesex* plays a crucial role in the physiological development of females, this mosaicism may impact upon complex behaviours that are difficult or impossible to reproduce in small-cages, including swarming, food and oviposition site searching, and resting.

Here, we present the results of four large-caged release experiments designed to challenge the suppressing activity of Ag(QFS)1 in an environment that partially mimics natural conditions and can invoke complex behaviours. We use an overlapping generation study design that is more likely to reveal differences in general fitness, mating success, and fecundity over time that cannot be captured in discrete generation studies. Ag(QFS)1 males were introduced at approximately 12.5% or 25% initial frequency and key measurements of drive invasion and population fitness were monitored over time. We observed increases in frequency of the transgenic mosquitoes within the populations in all four cages initiated with the drive that led to complete elimination by 245-311 days after introduction. We compared these results to the output of a stochastic model using the method of Approximate Bayesian Computation, in order to infer key life history parameters that are difficult to measure in dedicated assays. Our findings represent the first successful demonstration of efficacy for a gene drive in the second phase of testing which focuses on acquiring information under ecological challenging conditions, provide a platform for generating key evidence to inform initial go/no-go operational decisions, and pave the way for the first field trials of gene drive technology.

Material and methods

Study design

Initially, we assessed life history traits of both Ag(QFS1) males and females as well as of the wild-type strain G3 of *Anopheles gambiae* and assessed their longevity under large cage conditions (4.7 m^3) in order to emulate more natural population dynamics (Pollegioni et al. 2020) (see Fig. 1, Supplement Material). Considering the initial Kaplan-Meier Survival estimate of wild-type G3 adult mosquitoes in 4.7 m^3 cages of 2 m x 1 m x 2.35 m size and the establishment of overlapping generations with bi-weekly introductions of 400 G3 pupae with a start-up population of 800 mosquitoes, we then analysed age-structured large cage (ASL) populations with an expected mean size of ~570 adult mosquitoes as 'receiving' populations for gene drive release experiments. To mimic field-like conditions absent in small cage conditions, the climate chambers were maintained under near-natural environmental conditions including simulated dusk, dawn and daylight, and each cage was equipped with proven swarming stimuli and a resting shelter (Facchinelli et al. 2015) (Fig. 1). Under these conditions male swarming, an important component of successful mating behaviour, was regularly induced. To mimic a hypothetical field gene drive release, we seeded Ag(QFS1) mosquitoes over a single week (two releases) into the established 'receiving' wild-type populations at two different starting frequencies, low (12.5% initial allele frequency) and medium (25% allele frequency), as well as control cages (0% gene drive release), all in duplicate (6 cages total). The ASL population dynamics and the potential selection of drive-resistant alleles were monitored in treated and control cages until wild-type populations were fully suppressed by the gene drive in the treatments. Finally, we constructed an individual-based stochastic simulation model of the experiment to better understand the observed dynamics of the gene drive frequency and population suppression.

Mosquito strains

Two *Anopheles gambiae* mosquito strains were used, the wild-type G3 strain (MRA-112) and Female Sterile Gene Drive strain, Ag(QFS)1, previously known as $dsxF^{CRISPRh}$ (Kyrou et al. 2018). This strain contains a Cas9-based homing cassette within the coding sequence of the female-specific exon 5 of the *dsx* gene (Supp. Fig. 1). The cassette includes a human codon-optimised *Streptococcus pyogenes Cas9* (*hSpCas9*) gene under the regulation of the *zero population growth (zpg)* promoter and terminator of *Anopheles gambiae* and a gRNA against exon 5 under the control of the *Anopheles gambiae* U6 snRNA promoter. The cassette also carries a dsRed fluorescent protein marker under the expression of the 3xP3 promoter.

Mosquito containment and maintenance

Anopheles gambiae mosquito strains were contained in a purpose-built Arthropod Containment Level 2 plus facility at Polo d'Innovazione di Genomica, Genetica e Biologia, Genetics & Ecology Research Centre, Terni, Italy. Mosquitoes were reared in cubical cages of 17.5 cm x 17.5 cm x 17.5 cm (BugDorm-4) as described in Valerio et al. (2016) at 28°C and 80% relative humidity (Suppl. Fig. 2). Larvae were maintained in trays (253 x 353 x 81 mm) at a density of 200 larvae per tray using 400

mL deionized water with sea salt at a concentration of 0.3 g/L and 5 mL of 2% w/v larval diet (Damiens et al. 2012) and screened for fluorescent markers *en masse* using a Complex Object Parametric Analyzer and Sorter (COPAS, Union Biometrica, Boston, USA).

Large cage environment

For experimental purposes, mosquitoes were housed in a large cage environment as described in Pollegioni et al. (2020). A single large climatic chamber was equipped with six 4.7 m^3 cages of 2 m x 1 m x 2.35 m (length, width, height) (Fig. 1) and maintained at 28ºC ±0.5°C and 80% ±5% relative humidity (Fig. 1, Suppl. Fig. 2). The climatic chamber was illuminated by three sets of three LEDs (3000K, 4000K and 6500K correlated colour temperatures) controlled by Winkratos software (ANGELANTONI Industries S.p.A, Massa Martana, Italy), allowing a gentle transition between light and dark sufficient to emulate dawn, and dusk. For the purpose of the current study, full light conditions (800 lux) were simulated using all LEDs and adjusted to last 11 hours and 15 minutes. Cages were additionally equipped with ambient lighting (3000K) designed to stimulate swarming, as described previously in Facchinelli et al. (2015), and a terracotta resting shelter moistened with a soaked sponge. Mosquitoes were fed on 10% sucrose and 0.1% methylparaben solution and blood-fed bi-weekly using defibrinated and heparinized sterile cow blood via the Hemotek membrane feeder (Discovery Workshops, Accrington, 34 UK). Oviposition sites consisted of a 12 cm diameter Petri dish with a wet filter paper strip introduced 2 days after the blood meal. Mosquito pupae, food, blood and water were introduced or removed through two openings, 12 cm in diameter, at the front of each cage with no operators entering the cage. No adult mosquitoes were removed from the large cages throughout the cage trials.

Measuring the life history parameters

To assess life history parameters of wild-type G3 and Ag(QFS)1 strains, standardized phenotypic assays were performed as described in Pollegioni et al. (2020). In brief, clutch size, hatching rate, larval, pupal and adult mortality rates, as well as the bias in transgenics among the offspring of heterozygous Ag(QFS)1 were measured in wild-type G3 and Ag(QFS)1 strains in triplicate in standard small laboratory cages (BugDorm-4). Ag(QFS)1 heterozygotes used in these assays had inherited the drive allele paternally and were therefore subject to paternal, but not maternal, effects of embryonic nuclease deposition that can lead to a mosaicism of somatic mutations at the doublesex locus and a resultant effect on fitness (Kyrou et al. 2018). 150 females and 150 males were mated to wild-type mosquitoes for 4 days, blood-fed, and their progeny counted as eggs using EggCounter v1.0 software (Mollahosseini et al. 2012). Hatching rate was evaluated 3 days post oviposition by visually inspecting 200 eggs under a stereomicroscope (Stereo Microscope M60, Leica Microsystems, Germany). Sex-specific larval mortality was calculated by rearing 200 larvae/tray and counting/sexing the number of surviving pupae.

Sex-specific adult survival was assessed by introducing 100 male and 100 female pupae of G3 and Ag(QFS)1 into either small (0.0049 m^3) or large cages (4.7 m^3) unsexed (Suppl. Fig. 3). Sex-specific survival of emerged adults was calculated from daily collections of dead adult mosquitoes from the respective cages and their sexing. Because homozygous Ag(QFS)1 do not show clear sex-specific phenotypes as pupae (Kyrou et al. 2018), 100 Ag(QFS)1 homozygotes were introduced into large cages unsexed (Suppl. Fig. 3a). The adult survival assays in large cages were performed twice, one before the large cage Ag(QFS)1 release experiment started and one after the large cage Ag(QFS)1 release experiment finished. For the latter adult survival assay, around 400 individual mosquitoes were collected from large cage populations at larval stage (before the cage populations declined, day 231 and 311 post-release for Ag(QFS)1 and G3 wild type, respectively), and kept in small cages until the start of the assay (Suppl. Fig. 3b).

Establishment, maintenance and monitoring of age-structured large cage (ASL) populations

To test the suppressive potential of Ag(QFS)1, we first established stable ASL populations of *An. gambiae* (G3 strain) housed in a purpose-built climatic chamber. Each population was initiated and maintained at the maximum rearing capacity through bi-weekly introductions of 400 G3 pupae (200 males and 200 females) over a period of 21 days ('establishment'), estimated to sustain a mean adult population of 574 mosquitoes based on the initial Kaplan-Meier estimate (Suppl Fig. 3a). After this initial period only progeny of these populations were used to repopulate the cages twice weekly ('restocking') for a period of 53 days ('pre-release', 74 days total), or supplemented with wild type reared separately when progeny numbers were too low. Each ASL population was considered stabilised after retrieving a sufficiently large and stable number of eggs to restock the population over four consecutive weeks. In detail, the receiving populations in all six cages were stabilised to produce a similar number of eggs in the 31 days before Ag(QFS)1 release, with an average egg production per cage ranging from 2262-5334. Bi-weekly blood meals were initiated at dusk and extended for a period of 5 hours, and oviposition sites were illuminated with blue light for egg collection 2 days later. Eggs were removed from the cages, counted, and allowed to hatch in a single tray within the climatic test chamber. For re-stocking the cage populations with wild-type pupae, a maximum of 400 randomly selected pupae were collected at the peak of pupation, manually sexed and screened, and introduced to their respective cage twice per week.

Ag(QFS)1 release experiments in large cages

To assess invasion dynamics of the Ag(QFS)1 strain in ASL populations of *Anopheles gambiae*, we performed duplicate releases designed to randomly seed ASL populations at low (12.5%, cages 2 & 5) or medium (25%, cages 3 & 6) allelic frequencies. After 74 days pre-release initiation period, heterozygous Ag(QFS)1 males were released into duplicate cages in addition to the regular re-stocking of the ASL populations with wild-type pupae. Releases took place on two consecutive

restocking occasions, representing 15.2% (71 & 72) or 26.3% (142 & 143) of pupae introduced that week (943 and 1085, respectively), equivalent to 25% or 50% of the estimated mean pre-released adult population (on average 574 mosquitoes were present in large cages). No further releases were carried out and indoor ASL populations were maintained through restocking of 400 pupae twice per week. From then, the ASL populations were maintained in the same way we established the receiving population, with the same constant re-stocking rate from offspring. No adult mosquitoes were removed from the cages. Duplicate control cages were similarly maintained, but without release of Ag(QFS)1.

While not statistically significant (Kruskal-Wallis Test $P = 0.06^{ns}$), there was some variation in reproductive output amongst the six cages due to random effects (cage 1: mean egg number = 4265.77, CI95% = 1550.36; cage 2: mean egg number = 2691.73, CI95% = 790.41; cage 3: mean egg number = 2517.46, CI95% = 889.66; cage 4: mean egg number = 1799.18, CI95% = 573.18; cage 5: mean egg number = 2350.82, CI95% = 745.44; cage 6: mean egg number = 2060.05, CI95% = 767.77). To control for random effects that could affect reproductive capacity of the population independently of the effect of the gene drive, we chose as control populations those cages with reproductive output at the upper and lower end of the distribution (cages 1 & 4). Replicate gene-drive release cages were distributed to cages 2 and 5 (12.5% allelic frequency) and cages 3 and 6 (25% allelic frequency) to mitigate against potential local environmental position effects (Fig. 1).

Key indicators of population fitness and drive invasion were monitored for the duration of the experiment, including total egg output, hatching rate, pupal mortality, and the frequency of transgenics amongst L1 offspring and the pupal cohorts used for restocking. Total larvae were counted and screened for RFP fluorescence linked to Ag(QFS)1 using the COPAS larval sorter, and 1000 randomly selected to rear at a density of 200 per tray. Pupae positive for the gene drive element could be identified by expression of the RFP marker gene that is contained within the genetic element. Triplicate samples of up to 400 L1 larvae were stored in absolute ethanol at -80°C for subsequent analysis.

Modelling

A stochastic model was set up to replicate the experimental design with respect to twice-weekly egg-laying, the initiation phase, the transgene introductions, and the subsequent monitoring phase. A full model description is given in the Supplementary Methods. In brief, daily changes to the population result from egg laying, deaths, and matings, and are assumed to occur with probabilities that may be genotype specific. Adult longevity parameters were estimated from the large cage survival assays that were performed before the gene-drive release experiments began, and after the gene-drive dynamics had run their course. We compared the data to model simulations using a suite of summary statistics (Csilléry et al. 2010; Supplementary Methods) to infer three parameters representing female fertility

costs associated to the drive allele. In addition, we inferred two parameters that determined the egg production of unaffected (wildtype) females, and one parameter that determined the rate of R2 allele creation. We obtained a posterior distribution for all six parameters by retaining the 200 best fitting parameter combinations from 200,000 parameter samples generated by a Monte-Carlo algorithm (Fig. Suppl 4, Table 1).

Pooled amplicon sequencing and analysis

We previously developed a strategy to detect and quantify target site resistance based upon targeted amplicon sequencing using pooled samples of larvae (Hammond et al. 2017), and found no evidence for resistance to Ag(QFS)1 in small caged release populations (Kyrou et al. 2018). To further investigate resistance in the large caged release experiment, we analysed mutations found at the genomic target of Ag(QFS)1 in samples collected at early and late timepoints. Genomic DNA (gDNA) was extracted *en masse* from triplicate samples of 400 L1 larvae, or 50-300 larvae where larval numbers were limiting, that were collected after blood meals given on days 4 and 193 from all 6 cages, and on day 235 where sufficient larvae were available.

gDNA extractions were performed using the DNeasy Blood & Tissue kit (Qiagen). 100 ng of extracted gDNA was used to amplify a 291 bp region spanning the target site of Ag(QFS)1 in *doublesex*, using the KAPA HiFi HotStart Ready Mix PCR kit (Kapa Biosystems) and primers containing Illumina Genewiz AmpEZ partial adaptors (underlined): Illumina-AmpEZ-4050-F1 ACACTCTTTCCCTACACGACGCTCTTCCGATCTACTTATCGGCATCAGTTGCG and Illumina-AmpEZ-4050-R1 GACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTGAATTCCGTCAGCCAGC. PCR reactions were performed under non-saturating conditions and run for 25 cycles, as in Hammond et al. (2017), to maintain proportional representation of alleles from the extracted gDNA in the PCR products.

Pooled amplicon sequencing reads, averaging approximately 1.5 million per condition, were analysed using CRISPResso2 (Clement et al. 2019), using an average read quality threshold of 30. Insertions and deletions were included if they altered a window of 20 bp surrounding the cleavage site that was chosen on the basis of previously observed mutations at this locus (Kyrou et al. 2018). Individual allele frequencies were calculated based upon their total frequency in triplicate samples. A threshold frequency of 0.25% per mutant allele was set to distinguish putative resistant alleles from sequencing error (Pfeiffer et al. 2018).

Results

Ag(QFS)1 spreads rapidly through ASL populations

After stabilising the receiving wild type populations in the large cages, we seeded the cages, in duplicate, with gene drive mosquitoes at 12.5% and 25% allelic frequencies of the estimated pre-released adult population size. We also kept two cages unseeded as controls. We were able to track the inheritance of the gene drive allele by virtue of the dominant RFP marker gene. We observed substantial variability in the rise in frequency of gene drive-positive mosquitoes, regardless of starting frequency (Fig. 2g, Fig. 2h, Supplementary Data 1). We also observed an apparent 'phasing' pattern of transgene frequency between consecutive re-stockings, persisting for up to 200 days post release, that could be related to (but not only) the two phased single week releases. The spread of the Ag(QFS)1 followed a sigmoidal pattern of invasion, increasing in frequency slowly for the first 100-150 days, followed by a rapid period of invasion, and finally slowing as the drive approached fixation between 220-276 days after introduction in the low frequency release cages (Fig. 2g) and between 224-241 days after introduction in the medium frequency release cages (Fig. 2h). No gene-drive positive individuals were detected in control cages, consistent with the cages being fully isolated from one another (Supplementary Data 1).

Increase in frequency of the gene drive allele causes elimination of ASL mosquito populations

As Ag(QFS)1 approached fixation there was a rapid decline in the fraction of fertile females as the growing proportion of gene drive homozygotes, lacking a functional copy of the female isoform of the *doublesex* gene, develop into sterile "intersex" adults (Fig. 2d and Fig. 2e). As the formation of homozygotes is a requirement for population suppression, a strong and unambiguous reduction in egg output occurred only after the frequency of the gene drive allele rose above 90%, culminating in complete elimination 245-311 days after release of Ag(QFS)1 in the low frequency cages (Fig. 2a) and by days 266-276 in the medium release cage (Fig. 2b). By comparison, the mosquito population in the control cages maintained a stable sex ratio (Fig. 2f) and an average of more than 10,000 eggs over the final month of the experiment (Fig. 2c), while cages seeded with Ag(QFS)1 collapsed.

Adult longevity increases over the course of the large cage release experiment

No significant differences in adult survival between Ag(QFS)1 and wild-type strains were detected in large cages ($P = 1.0$, Kruskal-Wallis test), with 50% median mortality at day 6 (95% CI = 5-6 days) and day 11 (95% CI wild-type = 9-13 days, 95% CI Ag(QSF)1 = 11-12 days) at the beginning and the end of the large caged release experiment, respectively (Supp. Fig. 3 and Supplementary Data 2). Overall, survival in large cages is substantially lower than in small cages maintained under similar environmental and rearing conditions, where 50% mean mortality occurred at 20 days. In agreement with Pollegioni et al. (2020), our data suggest that females survive longer than males when housed in large cages.

We observed an increased adult longevity in the large cages after the year-long experiment compared to before the release (median of 11 days and 6 days, respectively; $P = 0.032$, Kruskal-Wallis test)

irrespectively of the genotype. Individuals reared in the small cages tested in the same conditions (after the year-long experiment) showed the same adult survival than those collected from the ASL populations (for both G3 wild type and Ag(QFS)1 transgenics), suggesting the difference is due to the micro-environmental conditions of the large cages and not due to strain adaptation or the genotypes.

Parameter inference reveals drive allele female fertility costs in age-structured mosquito populations

The ASL caged populations showed a similar trend of increasing egg output over time prior to the suppressive effect of the drive (Fig. 2a-c) that may be explained by a general increase in adult survival that was observed between the start and end of the population experiment (Supp. Fig. 3). To account for these changes in the stochastic model, we assumed a small increase in adult survival over time, irrespective of genotype, based on experimental data (Supp. Fig 3). The posterior distribution of our stochastic model is summarised in Table 1. We were particularly interested in the drive allele fertility costs, because these are potentially important to drive allele dynamics in natural populations (Beaghton et al. 2019, North et al. 2020). Fertility costs may arise from paternal and maternal effects of Cas9 deposition into the sperm or egg, or from ectopic activity of Cas9 in the soma (Kyrou et al. 2018).

The full posterior distribution indicated the presence of fertility costs, yet did not allow the relative roles of deposition and ectopic activity to be disentangled; the posterior probabilities for each factor strongly covary (Supp. Fig. 4). We therefore determined posterior estimates of transformed parameters that summarise the fertility costs of transgenic females depending on whether they had a transgenic father, mother, or if both parents were transgenic (Supplementary Methods).

The posterior mean density for the fertility cost to transgenic females whose father was transgenic was 0.35 (indicating a 35% reduction in egg output relative to wildtype females), with a 95% credible interval of (0.18-0.56) (Fig. 2i). This increased slightly to 0.39 if instead the mother was transgenic, with a much wider credible interval (0.02-0.85), and reduced to 0.18 (0.01-0.36) if both parents were transgenic. The overlap in the parent-specific estimates means we cannot determine whether the sex of the transgenic parent makes a difference to the fertility of transgenic female offspring on the basis of this data.

The posterior densities indicated that females typically lay around 117 eggs per batch (54-219), and around 13% of mated females laid eggs at each twice-weekly opportunity (7-20%). The posterior mean density for the fraction of non-homed gametes produced by heterozygous individuals becoming non-functional resistance alleles was around one half (49%; 27-81%).

Stochastic simulations capture dynamics of spread and suppression

Simulations of the cage dynamics using parameters drawn at random from the posterior distribution gave a close correspondence to the observed trends in the frequency of drive-carrying individuals (Fig. 2g-h). This is expected, since the posterior distribution was inferred from the data, yet it gives confidence that the model captures much of the biology of the cage population. The simulations performed less well in replicating the variability in egg laying in the control cages, suggesting the model does not incorporate all the sources of this variation (Fig. 2c). We ran 1000 simulations of the posterior informed model to predict the range of potential cage dynamics. All simulations ended with complete population suppression within 560 days, and 95% of the simulations reached this state within 399 or 329 days for the low and high frequency releases, respectively (Supp. Fig. 5).

Drive-resistant alleles were not generated in large cage releases of Ag(QFS)1

To investigate whether drive-resistant alleles had been generated or selected as the gene drive allele increases in frequency in the populations, we performed pooled amplicon sequencing around the gRNA target site on samples of the larval progeny (150-1200/cage) collected at early and late timepoints after release (Fig. 3). These alleles can take two forms: functional resistant alleles that restore a viable gene product, and non-functional resistant alleles that do not. Resistant alleles may be pre-existing in the population or generated by the gene drive itself as a result of error-prone end-joining. In spite of the incredible selective pressure exerted by Ag(QFS)1, no mutant alleles were generated that could conceivably code for a functional DSX protein.

We identified three putative end-joining mutations present above the threshold frequency of 0.25% in any of the four release cages. All three alleles introduce a frameshift mutation that would disrupt the female isoform of *doublesex*, including a 5-bp insertion that was uniquely identified in this study and two deletions (1 bp and 11 bp in length) that were previously identified in small caged testing of Ag(QFS)1 (Kyrou et al. 2018). The failure of any of these alleles to spread above 1% frequency amongst non-drive alleles would suggest they are highly deleterious and undergo no positive selection as the gene drive allele increases in frequency.

Discussion

In this study we provide evidence that the dsx targeting gene drive strain, Ag(QFS)1, is able to effectively suppress age-structured populations reared in an environment that recapitulates some parameters typical of natural conditions and induces some mosquito behaviours observed in the field. This gene drive has previously been demonstrated to spread effectively through populations of wild-type *Anopheles gambiae* mosquitoes maintained in small cages (0.0156 m^3) with non-overlapping generations (Kyrou et al. 2018). We observed similar dynamics of spread in duplicate large cages (4.7 m^3) cages initiated with low or medium frequency of the drive, leading to complete population

suppression within 245-311 days. Compared with previous discrete generation testing we find a strong phasing of drive frequency over time, suggesting that perhaps interbreeding between young and old cohorts of cohabiting adults is rare under these conditions. Nevertheless, we find that both stochasticity and dynamics of spread can be fully explained by modelling predictions based upon comprehensive characterisation of the life history traits Ag(QFS)1 (Supp. Fig. 6).

Retrospective inference of life-history parameters from cage population data allows a deeper insight into the phenotypic effects of transgenes, beyond what one can learn from small cage studies alone (Liu et al. 2019, Pollegioni et al. 2020). This analysis suggests that female fertility is the most important parameter that determines the dynamics of this gene drive. The simulations based on small cage data alone (from Kyrou et al. 2018) corresponded to the observations almost as well as the retrospective informed simulations (Supp Fig 5), probably because the single generation measurements of female fertility gave similar results to the inference from the large cage data (Fig. 2). This in itself suggests that the costs to female fertility conferred, at least by this gene drive targeting the female isoform of *doublesex*, may be quite stable within the environmental conditions in which mosquitoes are reared. Moreover, the accuracy of the prior simulations indicate that this drive allele confers few, if any, fitness effects in the semi-field environment that were overlooked by the small cage studies (with the exception of adult survival). Whether this holds for future gene drive designs, and which aspect of the resulting phenotype conferred has the largest effect on the veracity of predicting its trajectory in a population, will depend on the nature of the gene drive element and its molecular target.

A previous study found that the fitness of drive-heterozygous females was dependent on which parent contributed the drive allele (Kyrou et al. 2018), and two explanations were given. The cost may be due to paternal and maternal effects of Cas9 deposition into the sperm or egg, or it may result from ectopic activity of Cas9 in the soma, rather than the germline. Both possibilities, which are not mutually exclusive, will lead to suboptimal fitness due to a mosaic pattern of disruption of the *doublesex* gene. However, they may have subtly different ramifications to the potential spread of the drive allele in natural populations, since parental deposition will affect all offspring of heterozygous parents while ectopic activity will only affect offspring with the gene drive (Beaghton et al. 2019, North et al. 2020). We were unable to identify the potential causes of fertility cost from our analysis of the large cage observations, which reflects the relatively modest differences in their effects. Such disentanglement is perhaps easier to achieve from small cage studies where specific genotypes are crossed; by this method, Kyrou et al. (2018) found that females had lower fitness if descended from a transgenic father than transgenic mother, indicating an important role of paternal Cas9 deposition.

Our estimates were more precise in assessing the fertility costs to females with a transgene inherited from father rather than mother. This is because paternal inheritance is more common than maternal

423 inheritance, due to both homozygous female sterility and also the heterozygous fitness costs
424 themselves, meaning the cage dynamics are thus more sensitive to paternal effects. This shows how,
425 unlike experiments focussed on specific parameters, inference from population trends gives
426 information on the importance of parameters to the observed trends. Moreover, retrospective
427 inference is most effective at estimating the most important parameters. More generally, we have
428 shown that both methods of parameter estimation that we have used here – inference from population
429 trends and measurements from small scale experiments – are complimentary, and provide valuable
430 input to investigations of how these kinds of gene-drive products may impact natural vector
431 populations (Beaghton et al. 2019, North et al. 2020).

432 As with other forms of vector control, gene drives designed for population suppression will exert a
433 strong selection for resistance (Hammond et al. 2017). The force of selection for resistant mutations is
434 proportional to the fitness cost imposed by the gene drive itself but it can apply even to population
435 modification gene drives that are intended to drive an anti-parasitic effector gene into a vector
436 population, with the intention of changing its competency to transmit pathogens (Adolfi et al. 2020).
437 The most likely form of resistance is a change in the target sequence that can prevent cleavage by the
438 nuclease. Various strategies exist for reducing the probability of resistance arising against both
439 population suppression and population modification gene drives. In the case of Ag(QFS)1 the gene
440 drive is deliberately designed to target a region of its *doublesex* target gene that is under high
441 functional constraint and cannot readily generate or accommodate sequence variants that confer
442 functional resistance.

443 We previously showed that this strategy greatly improves the resilience of this strain to resistance,
444 failing to select for any resistance mutations (Kyrou et al. 2018, Simoni et al. 2020). The gene drive
445 release into large caged age-structured populations presented here provide an even greater selective
446 pressure for resistance, by starting with low release rates that ensure a long duration of the study (245-
447 311 days after initial release), and by the potential to reveal additional fitness costs such as complex
448 mating and oviposition behaviours that would be undetected in small cage testing. In spite of this
449 pressure and a concerted effort to identify resistant alleles, none were found to be capable of restoring
450 the function of *doublesex*.

451 Indeed, we identified just three mutant alleles that were each unable to encode a functional DSX
452 protein and present at low frequency (<1% amongst non-drive alleles). Somewhat surprisingly, fewer
453 non-functional mutant alleles were detected in our large semi-field cages than in the previous small
454 caged release experiments (Fig 3). This may be due to the harsher environment of the large cages that
455 results in a stronger purifying selection against non-functional resistant alleles, or it may simply
456 reflect differences in the effective population size, which have a similar effect in reducing the variety
457 of available alleles. Though these non-functional resistant alleles cannot completely displace a gene

drive, modelling suggests that under specific permissive conditions they can compete to reach a stable equilibrium (that nonetheless results in a strong and sustained population suppression) (Beaghton et al. 2019), an outcome we found neither in caged releases of Ag(QFS)1 nor in 1000 stochastic simulations (Suppl. Fig. 5). Large population sizes and low release rates increase the probability of these equilibriums forming; conversely, high frequency releases and multiplexed/combined drives can mitigate against it. Further studies must specifically address the probability of resistance, either naturally occurring or generated by the nuclease, to predict the potential spread, suppression and operational lifetime of Ag(QFS)1.

This study is the first successful test of gene drive technology in age-structured populations in an environment that mimics natural conditions and can invoke complex behaviours, and thus represents an essential intermediate step to move gene drive technology from laboratory studies to the field. Our data generated in the more realistic ecological setting in large cages, allowing the mosquitoes to show a complex feeding and reproductive behaviour, can inform go/no-go decisions by reducing uncertainty on the efficiency of gene-drive modified mosquitoes and better estimating post-release population dynamics.

In accordance with the Code of Ethics for Gene Drive Research (Annas et al. 2021), we have established a paradigm for generating data that help to bridge lab and field studies. Indeed, the pathway to deployment of gene drive mosquitoes recommends that prior to outdoor or open release testing, gene drive-modified mosquitoes are secondarily evaluated in large, overlapping generation indoor cages designed to mimic more closely the native ecological conditions (NASEM, 2016; James et al. 2018). The Ag(QFS)1 strain is the first gene drive strain to pass this essential intermediate step within a tiered testing approach, and whilst comprehensive resistance testing and environmental risk assessment will be needed ahead of field trials (Benedict et al. 2008), gene-drive modified mosquitoes show great promise as a tool for vector control.

Acknowledgements

We would like to thank Austin Burt, Silke Fuchs, John Mumford and John B. Connolly for their valuable comments and suggestions to improve the quality of the paper. The authors of the manuscript have no conflict of interest to declare.

Funding

Funded by a Grant from the Foundation for the National Institutes of Health through the Vector-Based Control of Transmission Discovery Research (VCTR) program of the Grand Challenges in Global Health Initiative of the Bill & Melinda Gates Foundation.

Author contributions

Conceptualization: AH, PP, TN, RMu, AC; Methodology: AH, PP, TP, AN, AS, RMu, TN, AC; Investigation: PP, TP, AC, RMi, AT, AB, KK, IM, AS, and RMu; Formal analysis AH, PP, TP, IM, AN, AS, TN and RMu; Project Administration: TN and RMu; Writing – original draft: AH, PP, AS, AN, TN, RMu. Supervision: TN, RMu and AC; Writing – review & editing: AH, PP, AS, AN, TN, RMu, AC; Resources: AC; Funding Acquisition: AC.

References

Adolfi, Adriana, Valentino M. Gantz, Nijole Jasinskiene, Hsu Feng Lee, Kristy Hwang, Gerard Terradas, Emily A. Bulger, et al. 2020. "Efficient Population Modification Gene-Drive Rescue System in the Malaria Mosquito *Anopheles stephensi*." *Nature Communications* 11 (1). https://doi.org/10.1038/s41467-020-19426-0.

Aldersley, Andrew, Arissara Pongsiri, Kamonchanok Bunmee, Udom Kijchalao, Wachiraphan Chittham, Thanyalak Fansiri, Nattaphol Pathawong, et al. 2019. "Too 'Sexy' for the Field? Paired Measures of Laboratory and Semi-Field Performance Highlight Variability in the Apparent Mating Fitness of *Aedes aegypti* Transgenic Strains." *Parasites and Vectors* 12 (1). https://doi.org/10.1186/s13071-019-3617-2.

Annas, G.J., Beisel, C.L., Clement, K., Crisanti, A., Francis, S., Galardini, M., Galizi, R., Grünewald, J., Immobile, G., Khalil, A.S., Müller, R., Pattanayak, V., Petri, K., Paul, L., Pinello, L., Simoni, A., Taxiarchi, C., Joung, J.K. (2021) A Code of Ethics for Gene Drive Researchers. The CRISPR Journal 4(1) 1-6

Beaghton, Andrea K., Andrew Hammond, Tony Nolan, Andrea Crisanti, and Austin Burt. 2019. "Gene Drive for Population Genetic Control: Non-Functional Resistance and Parental Effects." *Proceedings of the Royal Society B: Biological Sciences* 286 (1914): 20191586. https://doi.org/10.1098/rspb.2019.1586.

Benedict, M., D'Abbs, P., Dobson, S., Gottlieb, M., Harrington, L., Higgs, S., James, A., James, S., Knols, B., Lavery, J., O´Neill, S., Scott, T., Takken, W. and Toure, Y. (2008). Guidance for contained field trials of vector mosquitoes engineered to contain a gene drive system: recommendations of a scientific working group. *Vector-Borne and Zoonotic Diseases*, *8*(2), 127-166.

Burt, A. 2003. "Site-Specific Selfish Genes as Tools for the Control and Genetic Engineering of Natural Populations." *Proc Biol Sci* 270 (1518): 921–28. https://doi.org/10.1098/rspb.2002.2319.

Clement, Kendell, Holly Rees, Matthew C. Canver, Jason M. Gehrke, Rick Farouni, Jonathan Y. Hsu, Mitchel A. Cole, et al. 2019. "CRISPResso2 Provides Accurate and Rapid Genome Editing Sequence Analysis." *Nature Biotechnology* 37 (3): 224–26. https://doi.org/10.1038/s41587-019-0032-3.

Csilléry, Katalin, Michael G.B. Blum, Oscar E. Gaggiotti, and Olivier François. 2010. "Approximate Bayesian Computation (ABC) in Practice." *Trends in Ecology and Evolution*. https://doi.org/10.1016/j.tree.2010.04.001.

Damiens, D., M. Q. Benedict, M. Wille, and J. R.L. Gilles. 2012. "An Inexpensive and Effective Larval Diet for *Anopheles arabiensis* (Diptera: Culicidae): Eat like a Horse, a Bird, or a Fish?" *Journal of Medical Entomology* 49 (5): 1001–11. https://doi.org/10.1603/ME11289.

Facchinelli, Luca, Laura Valerio, Janine M. Ramsey, Fred Gould, Rachael K. Walsh, Guillermo Bond, Michael A. Robert, et al. 2013. "Field Cage Studies and Progressive Evaluation of Genetically-Engineered Mosquitoes." *PLoS Neglected Tropical Diseases* 7 (1). https://doi.org/10.1371/journal.pntd.0002001.

Facchinelli, Luca, Laura Valerio, Rosemary S. Lees, Clelia F. Oliva, Tania Persampieri, C. Matilda Collins, Andrea Crisanti, Roberta Spaccapelo, and Mark Q. Benedict. 2015. "Stimulating *Anopheles gambiae* Swarms in the Laboratory: Application for Behavioural and Fitness Studies." *Malaria Journal* 14 (1). https://doi.org/10.1186/s12936-015-0792-2.

Facchinelli, Luca, Ace R. North, C. Matilda Collins, Miriam Menichelli, Tania Persampieri, Alessandro Bucci, Roberta Spaccapelo, Andrea Crisanti, and Mark Q. Benedict. 2019. "Large-Cage Assessment of a

Transgenic Sex-Ratio Distortion Strain on Populations of an African Malaria Vector." *Parasites and Vectors* 12 (1): 70. https://doi.org/10.1186/s13071-019-3289-y.

Hammond, Andrew, Roberto Galizi, Kyros Kyrou, Alekos Simoni, Carla Siniscalchi, Dimitris Katsanos, Matthew Gribble, et al. 2016. "A CRISPR-Cas9 Gene Drive System Targeting Female Reproduction in the Malaria Mosquito Vector *Anopheles gambiae*." *Nature Biotechnology* 34 (1): 78–83. https://doi.org/10.1038/nbt.3439.

Hammond, Andrew M., Kyros Kyrou, Marco Bruttini, Ace North, Roberto Galizi, Xenia Karlsson, Nace Kranjc, et al. 2017. "The Creation and Selection of Mutations Resistant to a Gene Drive over Multiple Generations in the Malaria Mosquito." *PLoS Genetics* 13 (10): e1007039. https://doi.org/10.1371/journal.pgen.1007039.

James, Stephanie, Frank H. Collins, Philip A. Welkhoff, Claudia Emerson, H. Charles J. Godfray, Michael Gottlieb, Brian Greenwood, et al. 2018. "Pathway to Deployment of Gene Drive Mosquitoes as a Potential Biocontrol Tool for Elimination of Malaria in Sub-Saharan Africa: Recommendations of a Scientific Working Group †." *The American Journal of Tropical Medicine and Hygiene* 98 (6_Suppl): 1–49. https://doi.org/10.4269/ajtmh.18-0083.

James, Stephanie L., John M. Marshall, George K. Christophides, Fredros O. Okumu, and Tony Nolan. 2020. "Toward the Definition of Efficacy and Safety Criteria for Advancing Gene Drive-Modified Mosquitoes to Field Testing." *Vector-Borne and Zoonotic Diseases*. https://doi.org/10.1089/vbz.2019.2606.

KaramiNejadRanjbar, Mohammad, Kolja N. Eckermann, Hassan M. M. Ahmed, Héctor M. Sánchez C., Stefan Dippel, John M. Marshall, and Ernst A. Wimmer. 2018. "Consequences of Resistance Evolution in a Cas9-Based Sex Conversion-Suppression Gene Drive for Insect Pest Management." *Proceedings of the National Academy of Sciences*115 (24): 6189–94. https://doi.org/10.1073/pnas.1713825115.

Kyrou, Kyros, Andrew M. Hammond, Roberto Galizi, Nace Kranjc, Austin Burt, Andrea K. Beaghton, Tony Nolan, and Andrea Crisanti. 2018. "A CRISPR-Cas9 Gene Drive Targeting *Doublesex* Causes Complete Population Suppression in Caged *Anopheles gambiae* Mosquitoes." *Nature Biotechnology* 36 (11): 1062–66. https://doi.org/10.1038/nbt.4245.

Liu, J., Champer, J., Langmüller, A. M., Liu, C., Chung, J., Reeves, R., ... & Messer, P. W. (2019). Maximum likelihood estimation of fitness components in experimental evolution. *Genetics*, *211*(3), 1005-1017.

Mollahosseini, Ali, Marie Rossignol, Cédric Pennetier, Anna Cohuet, António Dos Anjos, Fabrice Chandre, and Hamid Reza Shahbazkia. 2012. "A User-Friendly Software to Easily Count *Anopheles* Egg Batches." *Parasites and Vectors* 5 (1). https://doi.org/10.1186/1756-3305-5-122.

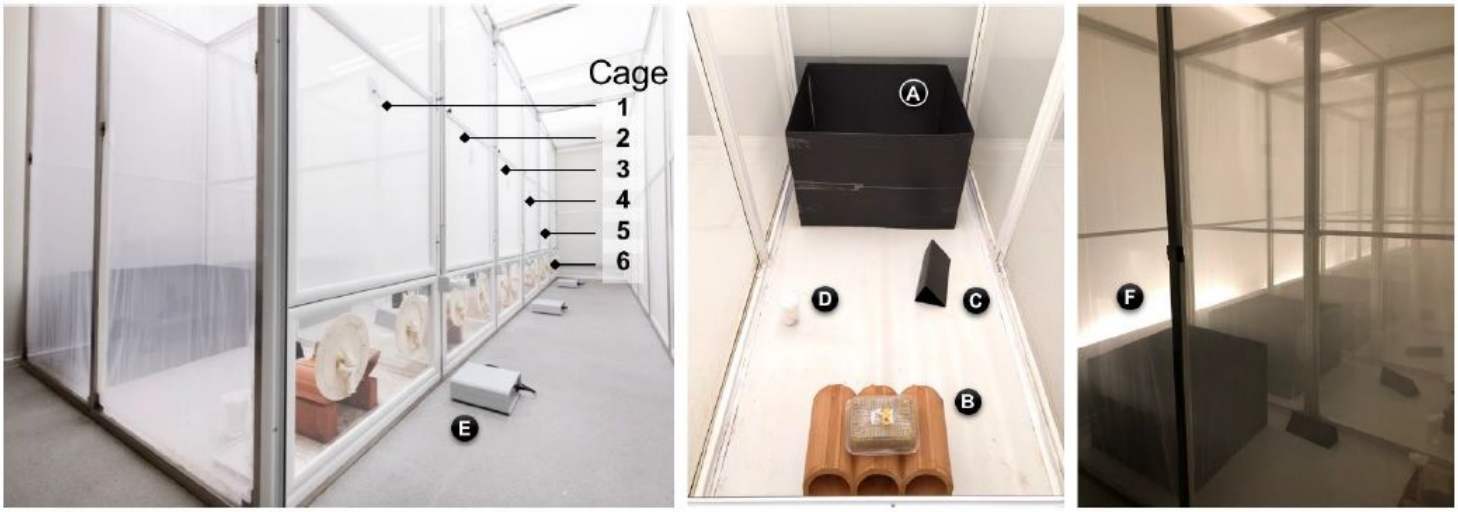
National Academies of Sciences, Engineering, and Medicine. 2016. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values*. Washington, DC: The National Academies Press. https://doi.org/10.17226/23405.

North, Ace R., Austin Burt, and H. Charles J. Godfray. 2020. "Modelling the Suppression of a Malaria Vector Using a CRISPR-Cas9 Gene Drive to Reduce Female Fertility." *BMC Biology* 18 (1). https://doi.org/10.1186/s12915-020-00834-z.

Pfeiffer, Franziska, Carsten Gröber, Michael Blank, Kristian Händler, Marc Beyer, Joachim L. Schultze, and Günter Mayer. 2018. "Systematic Evaluation of Error Rates and Causes in Short Samples in Next-Generation Sequencing." *Scientific Reports* 8 (1). https://doi.org/10.1038/s41598-018-29325-6.

577 Pham, Thai Binh, Celine Hien Phong, Jared B. Bennett, Kristy Hwang, Nijole Jasinskiene, Kiona Parker,
578 Drusilla Stillinger, John M. Marshall, Rebeca Carballar-Lejarazú, and Anthony A. James. 2019.
579 "Experimental Population Modification of the Malaria Vector Mosquito, *Anopheles stephensi*." *PLoS*
580 *Genetics*. Vol. 15. https://doi.org/10.1371/journal.pgen.1008440.
581 Pollegioni, Paola, Ace R. North, Tania Persampieri, Alessandro Bucci, Roxana L. Minuz, David Alexander
582 Groneberg, Tony Nolan, Philippos Aris Papathanos, Andrea Crisanti, and Ruth Müller. 2020. "Detecting
583 the Population Dynamics of an Autosomal Sex Ratio Distorter Transgene in Malaria vector Mosquitoes."
584 *Journal of Applied Ecology* 57 (10): 2086–96. https://doi.org/10.1111/1365-2664.13702.
585 Simoni, Alekos, Andrew M. Hammond, Andrea K. Beaghton, Roberto Galizi, Chrysanthi Taxiarchi, Kyros
586 Kyrou, Dario Meacci, et al. 2020. "A Male-Biased Sex-Distorter Gene Drive for the Human Malaria
587 Vector *Anopheles gambiae*." *Nature Biotechnology* 38 (9): 1054–60. https://doi.org/10.1038/s41587-020-
588 0508-1.
589 Valerio, Laura, C. Matilda Collins, Rosemary Susan Lees, and Mark Q. Benedict. 2016. "Benchmarking Vector
590 Arthropod Culture: An Example Using the African Malaria Mosquito, *Anopheles gambiae* (Diptera:
591 Culicidae)." *Malaria Journal* 15 (1). https://doi.org/10.1186/s12936-016-1288-4.
592 Verhulst, Eveline C., and Louis Van de zande. 2015. "Double Nexus-*Doublesex* Is the Connecting Element in
593 Sex Determination." *Briefings in Functional Genomics* 14 (6): 396–406.
594 https://doi.org/10.1093/bfgp/elv005.
595 Windbichler, N, M Menichelli, P A Papathanos, S B Thyme, H Li, U Y Ulge, B T Hovde, et al. 2011. "A
596 Synthetic Homing Endonuclease-Based Gene Drive System in the Human Malaria Mosquito." *Nature*
597 473 (7346): 212–15. https://doi.org/10.1038/nature09937.

Figures



(A) Swarming arena. (B) Wet resting site (bricks). (C) Dry resting site. (D) Glucose feeder. (E) Hemotek blood feeding system. (F) Sunset simulation

Figure 1

Design of the large cages used in this study. Images of the six large cages (numbered) within the climatic chamber (left panel) with the typical arrangement (central panel) of the swarming arena (A), wet (B) and dry (C) resting sites and sugar source (D). The six cages were seeded with control (cage 1 and 4), low frequency Ag(QFS)1 (cages 2 & 5) and medium frequency Ag(QFS)1 (cages 3 & 6). Also shown the Hemotek feeding system (E) and the black horizon marker to emulate sunset (F, panel on the right).

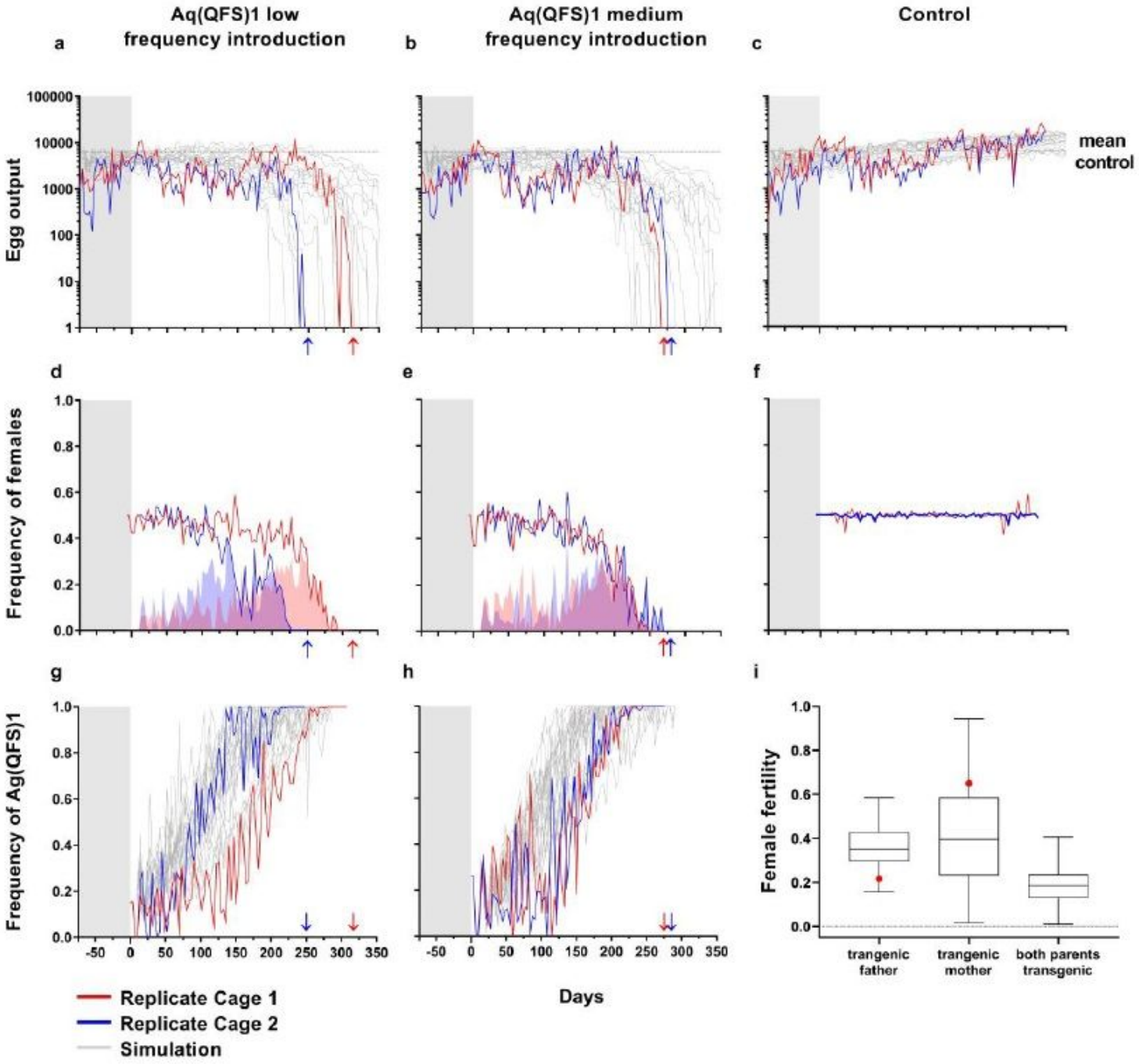


Figure 2

Kinetics of spread of Ag(QFS)1 in age-structured large cages. Age-structured large (ASL) cages were established over a period of 74 days (shaded grey) and seeded in duplicate with Ag(QFS)1 heterozygous males at low (12.5%, panels a, d, g) and medium (25%, panels b, e, h) allelic frequency, whereas two control cages were maintained without introduction of the Ag(QFS)1 gene drive (c, f). The total egg output (a, b, c), the frequency of putatively fertile non-intersex females (i.e. wild-type and heterozygous) (d, e, f), and the frequency of Ag(QFS)1 alleles (g, h) were monitored over time (red and blue lines for replicate cages). Mean egg output of the control is indicated by a dashed line (a, b, c). Red and blue

shaded areas indicate the fraction of non-intersex females carrying the gene drive in heterozygosity (d, e). Arrows indicate the point at which no further eggs were recovered, the point at which populations were considered eliminated. A total of 20 stochastic simulations of the egg output and the frequency of Ag(QFS)1 (grey lines) were modelled using parameter estimates drawn at random from the posterior distribution (Supp Fig 4) and superimposed to experimental data for the control and gene drive introductions (a, b, c, g, h). The 1posterior distribution of the relative fertility of Ag(QFS)1 heterozygous females that putatively received deposited nuclease paternally, maternally, or from both parents, as compared to the average fertility of wild-type females (i). Shown in red are the estimates of female fertility from experimental observation in Kyrou et al. (2018).

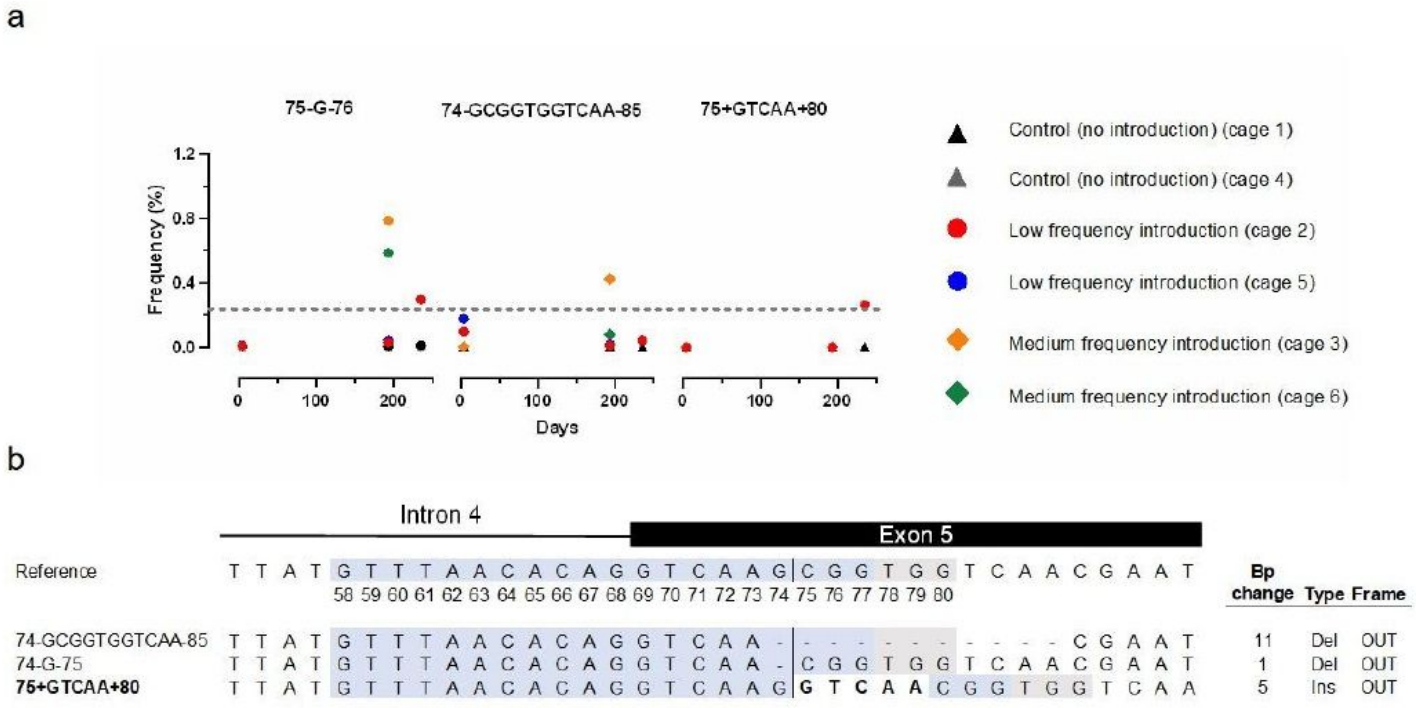


Figure 3

Drive-resistant mutations do not come under positive selection as Ag(QFS)1 spreads in 3 semi-field populations. (A) The % frequency of three putative non-restorative resistant alleles (R2) (75-G-76, 74-GCGGTGGTCAA-85, and 75+GTCAA+80) detected above the threshold frequency of 0.25% (Pfeiffer et al. 2018), in at least one cage at a single point in time, amongst all non-drive alleles, is shown over time. Samples were taken on days 4 and 193 for all cages, and on day 235 where the number of mosquitoes exceeded the restocking requirement. The naming of each mutation indicates the base pairs inserted (‘+’) or deleted (‘-’), and its location relative to the Cas9/gRNA cleavage site between position 74 and 75, depicted in panel B. Low frequency introduction cages 2 & 5 were initiated at a maximum Ag(QFS)1 allelic frequency of 12.5%, whilst medium frequency introduction cages 3 & 6 were initiated at 25%. Wild-type control cages 1 & 4 did not contain Ag(QFS)1. (B) The position of the three R2 alleles detected is shown, and compared to the reference An. coluzzii and An. gambiae sequence of the intron4/exon5 junction of the doublesex gene. Highlighted nucleotides indicate the gRNA binding site (blue) and PAM

sequence (grey). Inserted nucleotides are shown in bold. The number of base pairs inserted or deleted and the effect on the resulting allele (in-frame (IN), or out-of-frame (OUT)) is shown to the right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 05MSHammondetalTable.pdf
- 12MSHammondetalLargeCageTrialSupplMethodsmodelling.pdf
- 06MSHammondetalSupplFigure1.pdf
- 07MSHammondetalSupplFigure2.pdf
- 08MSHammondetalSupplFigure3.pdf
- 09MSHammondetalSupplFigure4.pdf
- 10MSHammondetalSupplFigure5.pdf
- 11MSHammondetalSupplFigure6.pdf
- 13MSHammondetalSupplementarydatasetX1merged.pdf
- 14MSHammondetalSupplementarydatasetX2merged.pdf