

Gene drive-based population suppression in the malaria vector *Anopheles stephensi*

Corresponding Author: Professor Jackson Chamber

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Xu et.al present the first CRISPR-Cas9 gene drive system targeting double-sex in *Anopheles stephensi*. They make use of regulatory elements (nos/vasa/U6/7SK) also used in other gene drive systems to target the female specific exon of dsx which they and others have previously targeted in other species (*D. suzukii* and *D. melanogaster*) after a high-profile publication in *Anopheles gambiae*. Those efforts have shown that transferring the seemingly immediate success in *A. gambiae* to other species is not straightforward. This manuscript shows much the same, but with limited follow-up work to identify the cause of the unexpected fitness costs and limited homing by nanos-Cas9.

In this reviewer's opinion they do an excellent job using reasoned insights from previous work (multiplexing/nanos) to attempt an improved dsx targeting gene drive system. The authors present an effective gRNA multiplexing strategy, which has generally been difficult to implement in practice (presumably also why their vasa-Cas9's gRNAs didn't work). However, when the performance is not as expected they do not do much to draw lessons that can be used for the next generations of drives.

The multiplexing system, and intermediate cut-rates of nanos-Cas9 are particularly interesting features to further investigate basic gene drive design criteria. In addition, I think it is essential that the authors make efforts to identify why male homozygote null mutants present as intersex (e.g., looking at splicing patterns as they did in their dsx Dm work).

There are also methodological issues that the authors resourcefully overcome (cross to vasa-Cas9 element with non-functional gRNAs and the same marker), but this should be bolstered with confirmatory (molecular) assays to ensure their assumption hold true. For example, showing that the mechanism of nanos-Cas9 rescue by vasa-Cas9 is as assumed. Confirming through small scale genotyping that they can indeed reliably discriminate between the two markers.

The computational modelling is of high quality and adds to the paper. In particular Fig3 usefully explores the impact of potential improvements to the design (male homozygous null fertility). One drawback is that the modelling is hard to place into context, even for the two genetic control systems described in the paper which are modelled in very different ways.

The authors do a good job in organising and making the raw data available. However, the statistical approach taken is very basic (i.e., not taking parent-to-parent/pool-to-pool variability into account) and at times produces non-intuitive results (i.e., significant difference between sexes in Fig1C).

In conclusion, this is novel work and a valuable contribution to the field, however, a substantial amount of data that can be found in related works is not included in the current manuscript. In particular, molecular investigation of splicing patterns, Cas9 expression, and cage trial data.

Specific points:

1. "Another promising target for population suppression is the sex determination pathway, with the female-specific exon of doublesex (dsx) emerging as a key target. Successful suppression in laboratory cage trials was achieved in *Anopheles gambiae*. However, functional resistance and fitness costs remain as potential problems in larger populations^{24,31}."

24 and 31 specifically mention no resistance was found and the drive performed well in cage populations. Text should be adapted to indicate your claims are an extrapolation and not specifically supported by those works.

2. "Although studies on modification drives have been published in *A. stephensi*, the major urban malaria vector in Asia that is becoming more invasive in East Africa, these lack efficiency and/or suitable cargo genes to be release candidates."

This is a strong general statement about others work which they may not agree with, I think more context/nuance is needed.

3. A description of efforts taken to confirm the sequence of the integration is as expected would be appreciated (given the lower-than-expected performance).

4. "Our findings revealed slightly higher drive inheritance rates for HSDdsx when G1 drive parents were males (74.85%±0.67%) compared to females (72.54%±0.75%, $p=0.022$, Fisher's Exact Test)."

The authors can find a statistically significant difference here due to a very large sample size. However, this does not take parent-to-parent batch effects into account. Usually this will not meaningfully change the conclusions (e.g., test for difference from 50% performed later), but I think it necessary if the authors wish to compare such small differences. They already show the large effect of these parent-to-parent batch effects in figure 1C and mention it regarding the fertility test.

5. Could you describe the criteria used to separate the different batches/pools?

6. Fertility tests for drive females, originating from either paternal or maternal drive, showed a significantly lower hatch rate when the G0 drive carrier was female (56.54%±3.59%, $p=0.0012$, Fisher's Exact Test), while no significant difference was observed between G0 drive males (74.08%±2.24%) and wild-type controls (85.75%±1.87%) ($p=0.0698$, Fisher's Exact Test). This suggests a potential fitness cost associated with maternal drive, though the fertility test could also be influenced by population density or batch effects (Data Set S2).

P values are much larger than I expect them to be for the given counts (and compared to p-values in the preceding paragraph described with the same statistical method). I cannot replicate those values by rerunning the analysis. Data Set S2 also lists p-values of 0.0001 for seemingly the same analysis (which is more in line with my expectations).

7. The phenotyping of G2 offspring revealed a significant boost in the HSDdsx drive inheritance to 100%, while the inheritance of vasa-Cas9 was 42.49%±1.78% (Data Set S4).

The inheritance rate of 100% provides a zero-width confidence interval (not shown in text, but listed in supplemental table and Fig4C) regardless of sample size. The authors may consider a more robust method of calculating the confidence intervals such as Wilson score interval (or even better a method that can also take parent-to-parent batch effects into account).

The way the zero-width confidence interval in Fig4C is shown may suggest it is based on the spread of the data from each batch (which indeed doesn't vary). However, that isn't the case and may confuse some readers to the approach taken.

8. Is deposition ignored in the model for Fig3?

9. Could you describe where the parameters in table S1 came from? In particular the female drive embryo resistance cut-rate and the female heterozygous fitness.

10. This reviewer would be interested to see a discussion/modelling on if it is necessary to increase the cut-rate of nanos-Cas9. Does it matter much if the resistance allele rate is also proportionally low?

11. The authors show that the combination of the nanos-Cas9 and vasa-Cas9 elements saturates inheritance rates, but they do not show if vasa-Cas9 itself is sufficient. nanos-Cas9 may be functioning perfectly well, but is being hampered by poor gRNA efficiency that is overcome with more gRNA:Cas9 complexes. Vasa-Cas9 may show the same deficiencies when used on its own.

I think their interpretations are likely correct, but I think the space of alternatives such as the one presented above are not so unlikely as to be dismissed without some additional investigation.

12. 294 are sterile, while for females,
Fort = for

13. 321 dsx gRNA target sites of six
Six = dsx?

14. This difference could be attributed to lower expression of Cas9, which may have been due to the limited size of our nos promoter, which could have been missing regulatory sequences.

Could you place it into context of the size of the nos promoter used in other homing drives? The promoter sequence you used is considerably larger than the two closest examples in *A. gambiae*.

15. This reviewer would be interested in a discussion/investigation on why the male HSDdsx;vasa-Cas9 double-hets are fertile while the homozygous HSDdsx males are not. Presumably due to somatic NHEJ mutations not disrupting the male dsx transcript, and hinting at what might be going wrong with HSDdsx.

(Remarks on code availability)

I was able to run one of the files provided on GitHub with the separately downloaded SLiM framework (found by googling). I did not have enough time to run it to completion.

No instructions were provided, but the process was remarkably straightforward in practice.

Reviewer #2

(Remarks to the Author)

This is an important paper by Xu et al. describing a new population suppression gene drive system in an important urban malaria vector, *Anopheles stephensi*. Previously, population replacement gene drive systems had been engineered in this species – notably AsMCRkh2 (Gantz et al., 2015) and Reckh (Adolfi et al., 2020) – however, as the authors note, no population suppression gene drive systems had been developed, despite suppression gene drives having been developed in *Anopheles gambiae* (most notably by Kyrou et al. (2018)). This paper therefore represents an important and novel contribution to the literature.

Enthusiasm is slightly muted since, as noted by the authors, the described HSDdsx construct is quite far from being field-ready – homing rates are only ~50%, which are much lower than rates for population modification constructs in *An. stephensi*, which tend to be above 90%. When a vasa-Cas9 allele is included, the “combination drive” system could potentially be used in an “SIT-like” strategy for suppression; but it is not clear that this combined system would be preferable over currently-available, self-limiting approaches such as SIT, RIDL or fsRIDL.

Despite these shortcomings, the research presents novel and important contributions to the field, which could form the basis for subsequent construct development eventually towards a field-ready product. In this sense, I am excited by this work. As a modeler, I will restrict my specific comments to this aspect of the paper. I only have minor comments in this area.

Minor comments:

1. What are the units for the low-density growth rate (i.e., the 2 and 6 values)? And are these based on values sourced from literature? Values for other anopheline species tend to be on the order of ~9 per generation. If this is the case, given model predictions for a growth rate of 6, are we to conclude that the drive is too weak to overcome this growth rate? This would make a literature search for the case of *An. stephensi* all the more pertinent.
2. Similarly, it seems that model predictions are sensitive to density-dependence at low larval densities, thus leading to the increased equilibrium population sizes in some HSDdsx simulations. As per low-density growth rates, it would be useful to have the parameterization of density-dependence be justified by a literature search (I didn't actually notice this parameterization reported in Table S1, so perhaps it could be added there, along with the low-density growth rates). If literature is unavailable, perhaps a modest sensitivity analysis could be conducted to see how general the prediction of increased equilibrium population size is under a variety of density-dependence model parameterizations.
3. gRNA multiplexing will not completely prevent the generation of functional r1 alleles, although it may reduce these rates significantly. At such low rates, functional r1 alleles may not be detected in small laboratory cage experiments; but in large field populations, their generation is a very real possibility. If the model is amenable, it would be interesting to know what rate of r1 resistance allele generation can be tolerated for suppression of population sizes considered here (~50,000 adult female mosquitoes). How large would a cage experiment need to be in order to detect r1 alleles generated at this rate?
4. In modeling the combination of HSDdsx and vasa-Cas9 for SIT-like population suppression, it is concluded that the results demonstrate the combined drive has “substantially higher efficiency than SIT and RIDL strategies,” although no evidence is included to back this claim up. It may be useful to simulate SIT and RIDL for similar assumptions and release sizes to demonstrate this with the approaches being simulated on a level playing field.

Copy edits:

There are a number of spelling and grammatical errors throughout the manuscript. It would benefit from a copy edit by the authors prior to resubmission.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I thank the authors for implementing the proposed revisions. The investigation of the splicing adds to the manuscript and helps contextualise the fertility data. The new statistical approach is an improvement, but has some wrinkles (see below). I think it is debatable if the revisions have added enough novelty to the work compared to prior dsx work in *gambiae*, but overall I recommend this work for publication.

Specific comments.

I don't want to be prescriptive, and I am happy for the authors to use their preferred analysis method. However, the reporting of the data could use another look to make it more consistent and easier to interpret.

Below some examples:

"Our findings revealed similar drive inheritance rates for HSDdsx from male ($74.6\% \pm 1.4\%$) and female ($73.6\% \pm 2.6\%$) drive parents ($p=0.741$, z-test). These correspond to drive conversion rates of $49.7\% \pm 1.3\%$ and $45.1\% \pm 1.5\%$, respectively."

The drive conversion is generally a simple transformation of the inheritance rate ($2*(P - 0.5)$). However, in the revised manuscript the drive inheritance rate is estimated with the GLM method while the conversion rate still uses the non-GLM method. As such the estimates, and particularly the SEs don't match.

I think CIs instead of SE would be more meaningful, but not wrong. I attempted to check the approach with the updated code provided on github. However, the .R code requires .csv files that are not provided (presumably the same information as is in the source file, but reformatted).

"Fertility tests for drive females revealed no significant difference in hatch rates when the G0 drive carrier was either female ($65.3\% \pm 16\%$, $p = 0.302$, z-test) or male ($84.1\% \pm 11\%$, $p=0.940$, z-test) compared to wild-type controls ($86\% \pm 2.5\%$). However, it remains possible that females experienced somewhat lower fertility, which has been seen in several other suppression drives."

I would recommend rephrasing this section, and also plotting this data. I had to look at the raw data to understand what was meant. This could for example also be interpreted as meaning that the non-drive progeny from G0 drive carriers were compared to those that did inherit the drive.

"Additionally, no statistical variation was observed between paternal and maternal drive ($p=0.503$, z-test). However, the standard errors of these hatch rates were relatively high, likely due to significant batch variations (Data Set S2)."

Again, I had to look at the raw data to understand, as usually "paternal and maternal drive" means the inheritance rate (which was measured for this cross but is seemingly not used). However, drive !carriers! is meant.

To address this, we randomly collected mosquitoes carrying different markers for the three possible genotypes, and we designed primers to amplify fragments specifically indicating the presence of vasa-Cas9 or HSDdsx constructs, as recommended by the reviewer. Our result confirmed that most of our phenotyping was correct, including all identification of HSDdsx (both alone and together with vasa-Cas9). Identification of vasa-Cas9 only individuals was also successful. However, when identifying vasa-Cas9 in mosquitoes with HSDdsx, we indeed failed to visually find it in some sample carrying both constructs, which were misidentified as HSDdsx only by fluorescent phenotyping (# 16 in the top right of the below figure). We found that part of the larval tail showing fluorescence in vasa-Cas9 was fragile and sometimes was removed during phenotyping, leading to misidentification of HSDdsx+vasa-Cas9 vs. HSDdsx only phenotype. This could indeed explain why the inheritance ratio of vasa-Cas9 was lower than Mendelian in combination drive cross. We now mention this in the manuscript.

Thank you for performing this check. However, I did not see where this is mentioned in the revised manuscript. Seemingly, not in the relevant section of the results.

I have looked over the other changes made and I am happy with them.

(Remarks on code availability)

I attempted to check the GLM code provided on github. However, the .R code requires .csv files that are not provided (presumably the same information as is in the source file, but reformatted).

The GitHub depository would be improved by a "readme" instruction file.

Reviewer #2

(Remarks to the Author)

I am happy with the authors' responses to my previous comments, and have no further requests.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I am happy with the changes made.

(Remarks on code availability)

I have run the r code, this works on my machine and it is clear to me how this relates to the data presented in the manuscript.
I have not evaluated the SLIM code.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

We would like to express our sincere gratitude to both reviewers for their thorough consideration of our manuscript and for their insightful suggestions and comments. Their feedback has inspired us and significantly enhanced the quality of our work. We have carefully revised the manuscript in accordance with their recommendations and provided supporting evidence for our statements, with the hope of meeting the reviewers' expectations. All responses to the reviewers' comments are highlighted in green, and the changes the in main manuscript are marked with track changes.

Reviewer #1 (Remarks to the Author):

Xu et.al present the first CRISPR-Cas9 gene drive system targeting double-sex in *Anopheles stephensi*. They make use of regulatory elements (nos/vasa/U6/7SK) also used in other gene drive systems to target the female specific exon of *dsx* which they and others have previously targeted in other species (*D. suzukii* and *D. melanogaster*) after a high-profile publication in *Anopheles gambiae*. Those efforts have shown that transferring the seemingly immediate success in *A. gambiae* to other species is not straightforward. This manuscript shows much the same, but with limited follow-up work to identify the cause of the unexpected fitness costs and limited homing by nanos-Cas9.

In this reviewer's opinion they do an excellent job using reasoned insights from previous work (multiplexing/nanos) to attempt an improved *dsx* targeting gene drive system. The authors present an effective gRNA multiplexing strategy, which has generally been difficult to implement in practice (presumably also why their vasa-Cas9's gRNAs didn't work). However, when the performance is not as expected they do not do much to draw lessons that can be used for the next generations of drives.

The multiplexing system, and intermediate cut-rates of nanos-Cas9 are particularly interesting features to further investigate basic gene drive design criteria. In addition, I think it is essential that the authors make efforts to identify why male homozygote null mutants present as intersex (e.g., looking at splicing patterns as they did in their *dsx* *Dm* work).

We investigated the splicing patterns and relative expression levels of *dsx* in adult mosquitoes with various phenotypes. These results have been added on page 11 and supplementary Figure S2 in the manuscript.

In female drive homozygote, we found both male and female *dsx* variants with relatively lower expression levels than wild-type, which likely explains the intersex phenotype. Unfortunately, for males, the situation remains unclear despite our effort to investigate. Male drive homozygotes had successful expression of the male *dsx* isoform, and while perhaps had slightly lower expression, did not have a statistically significant difference based on real-time PCR. This defied our initial expectation of substantially lower male isoform expression in male homozygotes, and we are quite puzzled as to why. We suggest some possibilities in the discussion section covering this aspect, which has been expanded.

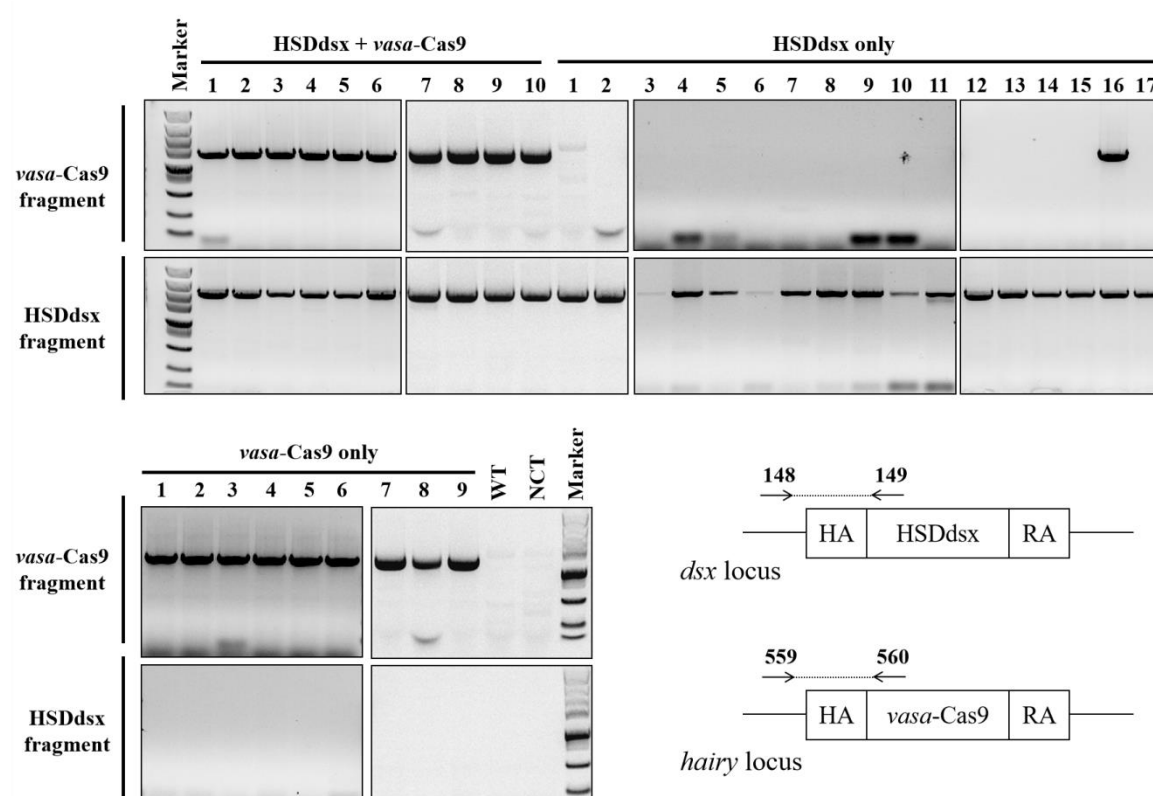
Although it will be certainly interesting and valuable to further identify factors that

influenced this intersex phenotype (e.g., differential expressions of *dsx* variants at other developmental stages or in specific tissues), this will likely require more expansive and complicated experimental designs. We hope that for now, our additional data collection is sufficient to support our main result in this manuscript.

There are also methodological issues that the authors resourcefully overcome (cross to *vasa*-Cas9 element with non-functional gRNAs and the same marker), but this should be bolstered with confirmatory (molecular) assays to ensure their assumption hold true. For example, showing that the mechanism of *nanos*-Cas9 rescue by *vasa*-Cas9 is as assumed. Confirming through small scale genotyping that they can indeed reliably discriminate between the two markers.

For this question, it's actually a fairly small change, even if we were unable to discriminate between the genotypes. With 100% green fluorescent among males and only about half inheriting the *vasa*-Cas9 allele based on other data, the expected number of males with *vasa*-Cas9 but not HSDdsx would be just a handful at most. Our visual phenotyping also seemed mostly reliable in the important direction - the eyes of HSDdsx were considerably brighter than *vasa*-Cas9, allowing easy identification of HSDdsx, even if *vasa*-Cas9 might sometimes be difficult to detect if HSDdsx was present, with the main distinguishing region being the tail (see Figure 4B). Still, this remains an important question, allowing us to solidify out 100% drive inheritance test while also potentially explaining the apparently slightly below Mendelian inheritance for *vasa*-Cas9.

To address this, we randomly collected mosquitoes carrying different markers for the three possible genotypes, and we designed primers to amplify fragments specifically indicating the presence of *vasa*-Cas9 or HSDdsx constructs, as recommended by the reviewer. Our result confirmed that most of our phenotyping was correct, including all identification of HSDdsx (both alone and together with *vasa*-Cas9). Identification of *vasa*-Cas9 only individuals was also successful. However, when identifying *vasa*-Cas9 in mosquitoes with HSDdsx, we indeed failed to visually find it in some sample carrying both constructs, which were misidentified as HSDdsx only by fluorescent phenotyping (# 16 in the top right of the below figure). We found that part of the larval tail showing fluorescence in *vasa*-Cas9 was fragile and sometimes was removed during phenotyping, leading to misidentification of HSDdsx+*vasa*-Cas9 vs. HSDdsx only phenotype. This could indeed explain why the inheritance ratio of *vasa*-Cas9 was lower than Mendelian in combination drive cross. We now mention this in the manuscript.



Primers used for identifying the constructs:

148: AGCACAAGATTAGCATGACTGAAGTG

149: ATCAGCCATACCACATTTGTAGAGG

559: TGCTGCCATCTTTTGAGCAACC

560: TCCATAATGGGCTTATTCGAGCG

The computational modelling is of high quality and adds to the paper. In particular Fig3 usefully explores the impact of potential improvements to the design (male homozygous null fertility). One drawback is that the modelling is hard to place into context, even for the two genetic control systems described in the paper which are modelled in very different ways.

Our computational analysis of both systems used the same underlying model. Only the drive system and the way that it was used varies between them. We have tried to better explain this in the methods section. We have also added more details to the modeling in the results section, which will hopefully better orient readers.

The authors do a good job in organising and making the raw data available. However, the statistical approach taken is very basic (i.e., not taking parent-to-parent/pool-to-pool variability into account) and at times produces non-intuitive results (i.e., significant difference between sexes in Fig1C).

We thank the reviewer for the suggestion regarding batch effects. In fact, we already had an “alternate analysis” in our supplemental.xlsx sheet that took such batch effects into account and provided slightly different mean and SEM estimates. We have modified our alternative analysis method and can now generate statistical comparisons

from it, so we have decided to feature these results in our main manuscript. There is no longer a statistically significant difference between male and female drive performance after this adjustment.

In conclusion, this is novel work and a valuable contribution to the field, however, a substantial amount of data that can be found in related works is not included in the current manuscript. In particular, molecular investigation of splicing patterns, Cas9 expression, and cage trial data.

Specific points:

1. “Another promising target for population suppression is the sex determination pathway, with the female-specific exon of doublesex (*dsx*) emerging as a key target. Successful suppression in laboratory cage trials was achieved in *Anopheles gambiae*. However, functional resistance and fitness costs remain as potential problems in larger populations^{24,31}.”

24 and 31 specifically mention no resistance was found and the drive performed well in cage populations. Text should be adapted to indicate your claims are an extrapolation and not specifically supported by those works.

We have revised the corresponding sentences on page 3.

2. “Although studies on modification drives have been published in *A. stephensi*, the major urban malaria vector in Asia that is becoming more invasive in East Africa, these lack efficiency and/or suitable cargo genes to be release candidates.”

This is a strong general statement about others work which they may not agree with, I think more context/nuance is needed.

The statement has been revised on page 3. However, lacking cargo genes is a clear statement, and the earlier study (the only other one in *Anopheles stephensi* while other publications from the same group were in *Anopheles gambiae*) had a whole follow-up paper (Pham TB, Phong CH, Bennett JB, Hwang K, Jasinskiene N, Parker K, et al. (2019) Experimental population modification of the malaria vector mosquito, *Anopheles stephensi*. PLoS Genet 15(12): e1008440) where resistance stopped the drive or where the drive suppressed the population and resistance was avoided. Thus, while drive conversion efficiency was high, other overall efficiency aspects were clearly lacking, and this is what we were trying to convey. Our revision clarifies these.

3. A description of efforts taken to confirm the sequence of the integration is as expected would be appreciated (given the lower-than-expected performance).

We had confirmed the correct integration site by amplifying the fragments covering left and right junctions of the genome and drive construct (primers on either side of the homology arm of the drive plasmid). We also amplified several partial fragments inside the drives, confirming that all major components of the drive were present. A

description of these has been added on page 8.

4. “Our findings revealed slightly higher drive inheritance rates for HSDdsx when G1 drive parents were males ($74.85\% \pm 0.67\%$) compared to females ($72.54\% \pm 0.75\%$, $p=0.022$, Fisher’s Exact Test).” The authors can find a statistically significant difference here due to a very large sample size. However, this does not take parent-to-parent batch effects into account. Usually this will not meaningfully change the conclusions (e.g., test for difference from 50% performed later), but I think it necessary if the authors wish to compare such small differences. They already show the large effect of these parent-to-parent batch effects in figure 1C and mention it regarding the fertility test.

We have amended our analysis method to account for batch effects. There was no difference between males and females in the updated analysis. We have also corrected the description in the manuscript (page 9).

5. Could you describe the criteria used to separate the different batches/pools?

We identified a cross cage as an independent batch (these were pooled crosses with several drive and wild-type mosquitoes for the drive efficiency test and single pair crosses in the fertility test). Some of them were done in different generations, but their parental crosses were the same (e.g., drive male crossed to wild-type female) unless specifically stated in the supplementary Data Set S1.

6. Fertility tests for drive females, originating from either paternal or maternal drive, showed a significantly lower hatch rate when the G0 drive carrier was female ($56.54\% \pm 3.59\%$, $p=0.0012$, Fisher’s Exact Test), while no significant difference was observed between G0 drive males ($74.08\% \pm 2.24\%$) and wild-type controls ($85.75\% \pm 1.87\%$) ($p=0.0698$, Fisher’s Exact Test). This suggests a potential fitness cost associated with maternal drive, though the fertility test could also be influenced by population density or batch effects (Data Set S2).

P values are much larger than I expect them to be for the given counts (and compared to p-values in the preceding paragraph described with the same statistical method). I cannot replicate those values by rerunning the analysis. Data Set S2 also lists p-values of 0.0001 for seemingly the same analysis (which is more in line with my expectations).

We have revised the description (on page 9) based on our updated analysis with the new batch effects approach that supports p values. Female fertility is still reduced (though not significantly due to large error rates from highly variable batches), but fertility from male drive crosses is now the same as wild-type.

7. The phenotyping of G2 offspring revealed a significant boost in the HSDdsx drive inheritance to 100%, while the inheritance of vasa-Cas9 was $42.49\% \pm 1.78\%$ (Data Set S4).

The inheritance rate of 100% provides a zero-width confidence interval (not shown in text, but listed in supplemental table and Fig4C) regardless of sample size. The authors may consider a more robust method of calculating the confidence intervals such as Wilson score interval (or even better a method

that can also take parent-to-parent batch effects into account).

The way the zero-width confidence interval in Fig4C is shown may suggest it is based on the spread of the data from each batch (which indeed doesn't vary). However, that isn't the case and may confuse some readers to the approach taken.

We thank the reviewer for pointing out this useful statistical method. Most of our new analysis uses the new batch method, but for the 100% inheritance rate, we now use a Wilson score interval method for determined the SEM (we couldn't figure out how to combine these methods, but it should be fine for just showing residual error in one data experiment). The Wilsonian confidence interval is 0.995-1, so even a somewhat wider interval on the order of other differences between our batch method and pooled analysis would not substantially affect our conclusions.

8. Is deposition ignored in the model for Fig3?

Female deposition is included in this model. We set up the embryo resistance cut rates of 0.01, as a pessimistic estimate (if it was this high, we would likely have found more resistance alleles by phenotyping). We now explain this in the results section when introducing the model.

9. Could you describe where the parameters in table S1 came from? In particular the female drive embryo resistance cut-rate and the female heterozygous fitness.

We now clarify this in the first paragraph of the results modeling section (page 12).

10. This reviewer would be interested to see a discussion/modelling on if it is necessary to increase the cut-rate of nanos-Cas9. Does it matter much if the resistance allele rate is also proportionally low?

A high drive cut rate is essential as it establishes the foundation for drive conversion. However, if conversion does not occur, it can lead to the development of resistance alleles. A higher cut rate would thus likely somewhat increase the rate of resistance allele formation. Nonfunctional resistance alleles can hinder a suppression drive, and this can sometimes be critical (in our recent work, reduction of embryo resistance alleles allowed a drive to succeed - in another study, we showed a different type of design that can be helped a lot by germline resistance alleles).

However, even with zero resistance allele formation, it is possible to have a drive with low suppressive power. Low drive conversion and fitness costs and result in this. In fact, in our model, eliminating the 1% resistance allele formation rate would only have marginally increased the suppressive power of the drive. This concept has been known for a while for some time in suppression drive models, but this knowledge is surprisingly uncommon among even gene drive experts today.

We have added some additional discussion about this result in the results section, next to the modeling result itself.

11. The authors show that the combination of the nanos-Cas9 and vasa-Cas9 elements saturates inheritance rates, but they do not show if vasa-Cas9 itself is sufficient. nanos-Cas9 may be functioning perfectly well, but is being hampered by poor gRNA efficiency that is overcome with more gRNA:Cas9 complexes. Vasa-Cas9 may show the same deficiencies when used on its own. I think their interpretation is likely correct, but I think the space of alternatives such as the one presented above are not so unlikely as to be dismissed without some additional investigation.

We agree that it is uncertain whether *vasa*-Cas9 alone will efficiently achieve drive conversion. Previous research by Anthony James' group demonstrated a high conversion rate of the *vasa* promoter in *Anopheles stephensi*, suggesting that this promoter may also be effective for the *dsx* drive. Nonetheless, if our gRNAs are less effective, this could be an explanation. However, one of our gRNA sequences is the same as that used in *Anopheles gambiae*, a different system, but one where the gRNA was effective. We have incorporated both these possibilities into the second paragraph of the discussion.

12. 294 are sterile, while fort females,
Fort = for

Revised.

13. 321 dsx gRNA target sites of six
Six = dsx?

Sorry for this unclear statement. We aimed to say that six out of seven double heterozygous males showed mutations in both gRNA target sites, confirming the high somatic expression of Cas9. We have revised the sentence.

14. This difference could be attributed to lower expression of Cas9, which may have been due to the limited size of our nos promoter, which could have been missing regulatory sequences.

Could you place it into context of the size of the nos promoter used in other homing drives? The promoter sequence you used is considerably larger than the two closest examples in *A. gambiae*.

The size of our *nanos* promoter was 3489 bp, which is indeed longer than that in *A. gambiae*. However, there may be some potential contribution of distal regulatory elements, which likely vary among species. Size has been added in the context.

15. This reviewer would be interested in a discussion/investigation on why the male HDSdsx;vasa-Cas9 double-hets are fertile while the homozygous HDSdsx males are not. Presumably due to

somatic NHEJ mutations not disrupting the male dsx transcript, and hinting at what might be going wrong with HDSdsx.

We thank the reviewer for this suggestion. We have expanded the third paragraph of the discussion to cover this and some other related details. Part of this paragraph comparing the result to previous studies has been split off into its own paragraph immediately afterward.

Reviewer #1 (Remarks on code availability):

I was able to run one of the files provided on GitHub with the separately downloaded SLiM framework (found by googling). I did not have enough time to run it to completion.

No instructions were provided, but the process was remarkably straightforward in practice.

Reviewer #2 (Remarks to the Author):

This is an important paper by Xu et al. describing a new population suppression gene drive system in an important urban malaria vector, *Anopheles stephensi*. Previously, population replacement gene drive systems had been engineered in this species – notably AsMCRkh2 (Gantz et al., 2015) and Reckh (Adolfi et al., 2020) – however, as the authors note, no population suppression gene drive systems had been developed, despite suppression gene drives having been developed in *Anopheles gambiae* (most notably by Kyrou et al. (2018)). This paper therefore represents an important and novel contribution to the literature.

Enthusiasm is slightly muted since, as noted by the authors, the described HSDdsx construct is quite far from being field-ready – homing rates are only ~50%, which are much lower than rates for population modification constructs in *An. stephensi*, which tend to be above 90%. When a vasa-Cas9 allele is included, the “combination drive” system could potentially be used in an “SIT-like” strategy for suppression; but it is not clear that this combined system would be preferable over currently-available, self-limiting approaches such as SIT, RIDL or fsRIDL.

Despite these shortcomings, the research presents novel and important contributions to the field, which could form the basis for subsequent construct development eventually towards a field-ready product. In this sense, I am excited by this work. As a modeler, I will restrict my specific comments to this aspect of the paper. I only have minor comments in this area.

Thank you for these kind comments. Please see our responses below for details.

Minor comments:

1. What are the units for the low-density growth rate (i.e., the 2 and 6 values)? And are these based on values sourced from literature? Values for other anopheline species tend to be on the order of ~9

per generation. If this is the case, given model predictions for a growth rate of 6, are we to conclude that the drive is too weak to overcome this growth rate? This would make a literature search for the case of *An. stephensi* all the more pertinent.

These are surprisingly tricky considerations. Ultimately, we could not find a reliable value to use for this. While the low-density growth rate can be seasonally measured, this may not be representative of the situation year-round, when growth of predators and competing species could make it considerably easier to eliminate the population. This has not been well-explored in ecological studies, so these are just plausible values that show the different possibilities for the drives. We go into this in more detail in another modeling study on competing species, which we now reference when describing our parameters with more detail (we have also added some explanation to the methods).

2. Similarly, it seems that model predictions are sensitive to density-dependence at low larval densities, thus leading to the increased equilibrium population sizes in some HSDdx simulations. As per low-density growth rates, it would be useful to have the parameterization of density-dependence be justified by a literature search (I didn't actually notice this parameterization reported in Table S1, so perhaps it could be added there, along with the low-density growth rates). If literature is unavailable, perhaps a modest sensitivity analysis could be conducted to see how general the prediction of increased equilibrium population size is under a variety of density-dependence model parameterizations.

We used a linear density-dependent growth curve, one of two common options (the other being a Beverton-Holt curve). This can increase population size a little if there are somewhat fewer larvae. This phenomenon is called overcompensation and has been observed in some mosquito populations, though it is unclear how prevalent it may be. With a Beverton-Holt curve, equilibrium population size from failed suppression would be considerably lower, and the repeated-release self-limiting systems would all do much better (though still keep roughly the same proportion of release size). We have better explained these in the methods and in the modeling results sections, including mentioning our specific density growth curve shape, which we neglected to do in the original draft.

We are hesitant to do a sensitivity analysis for this because the relationship between drive power (genetic load) and population size has already been explored in other studies, and this specific drive is not being considered a "release candidate", reducing priority for lots of additional specific analysis (we certainly hope to be doing this sort of analysis in future systems). For self-limiting drives, our recent preprint (which we now reference) does an adequate job at covering this topic (it also includes more sensitivity analysis for various relevant parameters):

Han Y, Champer J. A comparative assessment of self-limiting genetic control strategies for population suppression. *bioRxiv*, 2024.09.23.614516, 2024.

For self-sustaining suppression gene drives, there have been several studies already

exploring the genetic load of drives with different performance. How it relates to equilibrium population size is explored in this nice review:

Dhole S, Lloyd AL, Gould F. Gene drive dynamics in natural populations: The importance of density dependence, space, and sex. *Annu Rev Ecol Evol Syst*, 51, 505–531, 2020.

4. gRNA multiplexing will not completely prevent the generation of functional r1 alleles, although it may reduce these rates significantly. At such low rates, functional r1 alleles may not be detected in small laboratory cage experiments; but in large field populations, their generation is a very real possibility. If the model is amenable, it would be interesting to know what rate of r1 resistance allele generation can be tolerated for suppression of population sizes considered here (~50,000 adult female mosquitoes). How large would a cage experiment need to be in order to detect r1 alleles generated at this rate?

While these are interesting questions, they are probably beyond the scope of this study (the low efficiency of our drives means that a thorough evaluation is lower priority). However, we have examined very similar situations in earlier work using similar methods. We have referenced these when discussing multiplexed gRNAs in the manuscript, and the reviewer can get potential answers therein:

Champer SE, Oh SY, Liu C, Wen Z, Clark AG, Messer PW, Champer J. Computational and experimental performance of CRISPR homing gene drive strategies with multiplexed gRNAs. *Sci Adv*, 6, eaaz0525, 2020.

Yang E, Metzloff M, Langmüller AM, Xu X, Clark AG, Messer PW, Champer J. A homing suppression gene drive with multiplexed gRNAs maintains high drive conversion efficiency and avoids functional resistance alleles. *G3 Genes|Genomes|Genetics*, 2022.

These additional recent modeling studies by us and others also examine functional resistance in suppression drives on a more general level and may also be of interest (though some use a simplified gRNA multiplexing model):

Champer J, Kim IK, Champer SE, Clark AG, Messer PW. Suppression gene drive in continuous space can result in unstable persistence of both drive and wild-type alleles. *Mol Ecol*, 30, 1086–1101, 2021.

Liu Y, Champer J. Modelling homing suppression gene drive in haplodiploid organisms. *Proc R Soc B*, 289, 2022.

Khatri BS, Burt A. A theory of resistance to multiplexed gene drive demonstrates the significant role of weakly deleterious natural genetic variation. *Proc Natl Acad Sci U S A*, 119, e2200567119, 2022.

Prowse TAA, Cassey P, Ross J V, Pfitzner C, Wittmann TA, Thomas P. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. *Proc R Soc B Biol Sci*, 284, 20170799, 2017.

5. In modeling the combination of HSDdsx and vasa-Cas9 for SIT-like population suppression, it is concluded that the results demonstrate the combined drive has “substantially higher efficiency than

SIT and RIDL strategies,” although no evidence is included to back this claim up. It may be useful to simulate SIT and RIDL for similar assumptions and release sizes to demonstrate this with the approaches being simulated on a level playing field.

We appreciate the reviewer’s comment regarding the absence of a comparison between our combination drive and existing self-limiting approaches. In response, we have added a new figure (Figure S5) to show the modeling of SIT and fsRIDL for comparison. Our results indicate that the release ratios of these two strategies were significantly higher, thereby demonstrating the superior performance of the self-limiting approach proposed in this study.

Copy edits:

There are a number of spelling and grammatical errors throughout the manuscript. It would benefit from a copy edit by the authors prior to resubmission.

We have re-reviewed the whole manuscript and hopefully eliminated at least most of these.

We sincerely thank both reviewers for their valuable suggestions and positive feedback. In response, we have revised the manuscript accordingly and adjusted our files in GitHub. All changes made in the main manuscript are highlighted using track changes for clarity.

Reviewer #1 (Remarks to the Author):

I thank the authors for implementing the proposed revisions. The investigation of the splicing adds to the manuscript and helps contextualise the fertility data. The new statistical approach is an improvement, but has some wrinkles (see below). I think it is debatable if the revisions have added enough novelty to the work compared to prior dsx work in *gambiae*, but overall I recommend this work for publication.

Specific comments.

I don't want to be prescriptive, and I am happy for the authors to use their preferred analysis method. However, the reporting of the data could use another look to make it more consistent and easier to interpret.

Below some examples:

"Our findings revealed similar drive inheritance rates for HSDdsx from male ($74.6\% \pm 1.4\%$) and female ($73.6\% \pm 2.6\%$) drive parents ($p=0.741$, z-test). These correspond to drive conversion rates of $49.7\% \pm 1.3\%$ and $45.1\% \pm 1.5\%$, respectively."

The drive conversion is generally a simple transformation of the inheritance rate ($2*(P - 0.5)$). However, in the revised manuscript the drive inheritance rate is estimated with the GLM method while the conversion rate still uses the non-GLM method. As such the estimates, and particularly the SEs don't match.

This has been removed in the main text, and fixed in the supplemental xlsx file. We now use the batch method for all reporting in the text for both mean and SEM.

I think CIs instead of SE would be more meaningful, but not wrong. I attempted to check the approach with the updated code provided on github. However, the .R code requires .csv files that are not provided (presumably the same information as is in the source file, but reformatted).

We now include the csv files and better instructions on use of the code.

"Fertility tests for drive females revealed no significant difference in hatch rates when the G0 drive carrier was either female ($65.3\% \pm 16\%$, $p = 0.302$, z-test) or male ($84.1\% \pm 11\%$, $p=0.940$, z-test) compared to wild-type controls ($86\% \pm 2.5\%$). However, it remains possible that females

experienced somewhat lower fertility, which has been seen in several other suppression drives."

I would recommend rephrasing this section, and also plotting this data. I had to look at the raw data to understand what was meant. This could for example also be interpreted as meaning that the non-drive progeny from G0 drive carriers were compared to those that did inherit the drive.

We have revised this section and the methods too, and we plot relevant data to clarify this result (Figure S6).

"Additionally, no statistical variation was observed between paternal and maternal drive ($p=0.503$, z-test). However, the standard errors of these hatch rates were relatively high, likely due to significant batch variations (Data Set S2)."

Again, I had to look at the raw data to understand, as usually "paternal and maternal drive" means the inheritance rate (which was measured for this cross but is seemingly not used). However, drive !carriers! is meant.

As described above, we have revised this section.

To address this, we randomly collected mosquitoes carrying different markers for the three possible genotypes, and we designed primers to amplify fragments specifically indicating the presence of vasa-Cas9 or HSDdsx constructs, as recommended by the reviewer. Our result confirmed that most of our phenotyping was correct, including all identification of HSDdsx (both alone and together with vasa-Cas9). Identification of vasa-Cas9 only individuals was also successful. However, when identifying vasa-Cas9 in mosquitoes with HSDdsx, we indeed failed to visually find it in some sample carrying both constructs, which were misidentified as HSDdsx only by fluorescent phenotyping (# 16 in the top right of the below figure). We found that part of the larval tail showing fluorescence in vasa-Cas9 was fragile and sometimes was removed during phenotyping, leading to misidentification of HSDdsx+vasa-Cas9 vs. HSDdsx only phenotype. This could indeed explain why the inheritance ratio of vasa-Cas9 was lower than Mendelian in combination drive cross. We now mention this in the manuscript.

Thank you for performing this check. However, I did not see where this is mentioned in the revised manuscript. Seemingly, not in the relevant section of the results.

We have added this result on page 13-14.

I have looked over the other changes made and I am happy with them.

Reviewer #1 (Remarks on code availability):

I attempted to check the GLM code provided on github. However, the .R code requires .csv files that are not provided (presumably the same information as is in the source file, but reformatted).

The GitHub depository would be improved by a "readme" instruction file.

This has been added, together with csv files that can be easily used.

Reviewer #2 (Remarks to the Author):

I am happy with the authors' responses to my previous comments, and have no further requests.

We thank Reviewer #2 for the positive feedback.