nature portfolio

Peer Review File

Behavioral heterogeneity in host-seeking and post-feeding suppression among disease vector mosquitoes

Corresponding Author: Dr Laura Duvall

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)
This manuscript establishes an automated video tracking and machine learning classification pipeline to compare host-seeking behavior and post-blood meal suppression in six mosquito species across three major genera (Aedes, Anopheles, Culex). The authors find notable interspecific variation in baseline host-seeking repertoires and suppression timelines. The approach provides high temporal resolution and fine-scale posture analysis, which is valuable for comparative behavioral studies.
While the study is well-conceived and technically advanced, several conceptual, methodological, and interpretational issues need to be addressed before the manuscript can be considered for publication.

Major Comments
1) Control of biological variables
a) The authors do not state whether the strains were lab-adapted or field-derived, nor do they clearly indicate the generation in captivity. Long-term colonization can significantly alter host-seeking behavior.
b) Larval diet protocols differ slightly between Aedes/Culex (TetraMin) and Anopheles (Koi food). This is a confounding factor because larval nutrition is known to affect adult host-seeking and suppression. At minimum, the authors should justify why the diets were not standardized or discuss the potential bias this introduces.
c) Egg hatching conditions are not fully equivalent across species. For a comparative behavioral study, the developmental rearing environment should be standardized as far as possible.
2) Host cue choice and species ecology
a) The stimuli used (human odor + 10% CO_2) are optimal for anthropophilic Aedes and Anopheles gambiae, but suboptimal for Culex, especially ornithophilic species like Culex tarsalis. This might explain part of the behavioral differences rather than true physiological suppression differences. While the authors acknowledge this in Discussion, they should either include an avian odor control condition, or interpret Culex results with more caution.
b) The claim of a “unified behavioral framework” is weakened if stimulus relevance varies greatly by species.
3) Replicate numbers and statistical clarity
a) Sample sizes vary widely (n = 15–55 females) without clear rationale. This imbalance may affect statistical power. Authors should clarify how many replicates per species per timepoint were initially planned and whether any were excluded (e.g., due to non-feeding).
b) The manuscript uses multiple non-parametric and parametric tests interchangeably. It should be clear why certain tests were chosen for each dataset and whether normality was assessed.
4) Definition of “suppression” and baseline
a) Suppression is quantified as percent reduction from baseline, but baseline is defined as the pre-stimulus period in unfed controls. This assumes stability of unfed activity across days and individuals, which might not hold for all species. Baselines should ideally be measured in the same individuals before feeding.
b) Movement and probing are lumped as indicators of host-seeking, but high locomotion without probing may reflect oviposition searching rather than recovery of host-seeking.
5) Machine learning model generalizability
a) Separate DeepLabCut models were required for Anopheles, suggesting limited cross-species generalizability. If a “unified” approach is claimed, the authors should discuss how model choice might bias classification.
b) The validation accuracy (92–96%) is good, but it is based on genus-specific training sets. Cross-validation across genera (train on one genus, test on another) would better demonstrate general applicability.

6) Interpretation discrepancies with previous literature
a) For Anopheles gambiae, prolonged suppression up to 72 h conflicts with classic findings of early recovery. This discrepancy needs deeper discussion—whether due to strain origin, assay environment, or host cue type.

Minor comments
L184: Delete the second "of"
L370–376: How did you recognize the same individual after the mosquitoes were mixed for blood feeding? If identities were not preserved, it seems only the average blood-meal size of the batch was measured. In that case, the estimate would be a cohort mean and would not account for individual differences. Please clarify whether pre- and post-feeding weights were paired for each female (and how). L370–376: How did you recognize the same individual after the mosquitoes were mixed for blood feeding? If identities were not preserved, it seems only the average blood-meal size of the batch was measured. In that case, the estimate would be a cohort mean and would not account for individual differences. Please clarify whether pre- and post-feeding weights were paired for each female (and how).

Reviewer #2

(Remarks to the Author)
Summary:

Duvall et al. propose a new method for modeling and evaluating host seeking in mosquitoes as well as a comparison of post-blood meal host-seeking of different anthropophilic mosquito species, both of which are novel and interesting. However, the organization of the paper as well as the literature reviews (particularly in the discussion) require revision. Below are line-by-line suggestions to aid in revisions as well as some readings that may be useful to reference.

Line by line:

Introduction

Line 34: The citation punctuation differs here. Please revise.

Also Line 34: Vision can also play a role in host seeking. Please refer to Raji & DeGennero (Raji, J. I., & DeGennaro, M. (2017). Genetic analysis of mosquito detection of humans. Current opinion in insect science, 20, 34-38.)

Lines 51-53: For a more recent study on how larval nutrition can affect behavior: Vinauger, C., & Chandrasegaran, K. (2024). Context-specific variation in life history traits and behavior of Aedes aegypti mosquitoes. Frontiers in Insect Science, 4, 1426715.

Lines 88-91: This feels like it should cover a separate paragraph. This final intro paragraph talks about machine learning and how the authors will be applying it to their pipeline; these first couple sentences are a separate topic.

Line 91-92: This statement seems to be missing a verb. Please rephrase.

General comment: The preferred host of each species is mentioned in the discussion, but not in the introduction. It may be worth noting that some species of Culex, and even some strains within species, can prefer avian hosts. This is true for Aedes aegypti as well.
Tchouassi, D. P., Jacob, J. W., Ogola, E. O., Sang, R., & Torto, B. (2019). Aedes vector–host olfactory interactions in sylvatic and domestic dengue transmission environments. Proceedings of the Royal Society B, 286(1914), 20192136.

Results

Lines 113-117: This section should be in the introduction, since you are using these species throughout experiments.

Lines 113-117: Recent studies show the lineages may be older than that: Soghigian, J., Sither, C., Justi, S. A., Morinaga, G., Cassel, B. K., Vitek, C. J., ... & Wiegmann, B. M. (2023). Phylogenomics reveals the history of host use in mosquitoes. Nature Communications, 14(1), 6252.

Line 166: This is true of anthropophilic species, but these behaviours may not be conserved in mosquitoes that feed on other hosts.

Fig 2C-D: If the data is available, it might be nice to show a bit more of the activity from before stimulus delivery to even better highlight the difference in activity. For example, comparing 3 minutes of pre-stimulus activity vs. 8 minutes of post-stimulus activity rather than 1 minute vs. 8 minutes.

Figure 2F: Why are the PCAs separated? Putting them all together will help the reader to see the overlap for themselves. Or at least putting mosquito species of the same genus together would help to see the overlap, since that is the point being made.

Fig 2 - Do the authors have any hypotheses about the lack of flying in Cx. tarsalis or any other low level activity (extended data)? Please discuss these kinds of results in the discussion. Also please include data somewhere about the number of individuals removed from the analysis.

Lines 184-185: There is a typo “of is”

Discussion:

Line 227-228: This is a good point, which you could talk about more. What is this setup good for? What is the significance/novelty of it? What knowledge gaps could you fill with this setup? You don’t have direction, but you can measure response levels to known attractants/repellants, which could allow you to determine thresholds for detection of compounds (almost like a somewhat more accessible electroantennogram (EAG)?). Expand on why you think this system is important to share with the scientific community in the discussion.

Line 244-245: There may be variation within species as well, as different strains can vary in behaviour and life history traits, as demonstrated in the paper below. Some populations/strains are more anthropophilic than others as well, since some places in the world consider Cx. quinquefasciatus to be ornithophilic.
Carlassara, M., Khorramnejad, A., Oker, H., Bahrami, R., Lozada-Chávez, A. N., Mancini, M. V., ... & Bonizzoni, M. (2024). Population-specific responses to developmental temperature in the arboviral vector Aedes albopictus: Implications for climate change. Global change biology, 30(3), e17226.

Lines 268-271: This statement does not have much of an impact. Please consider revising or removing.

General comment: Overall, discussion needs more comparison of the results with known literature as well as explanations/hypotheses about the results.

Methods

Line 295: This sentence is a run-on and needs a comma before “and.”

Line 308: Please include a reference for the previous paper.

Line 307-322: It would be really helpful to have schematics of these to visualize your setup. This reviewer is unable to visualize beyond the 6 well plate.

Line 311: 10% CO_2 would be roughly 100,000ppm–humans only emit about 30,000ppm CO_2. Is there a reason why the researchers used such a high concentration? Or was it diluted? Please explain your reasoning for using such a high concentration.
Culex species will sometimes even respond to less, since they prefer avian hosts. For more information: Reinhold, J. M., Chandrasegaran, K., Oker, H., Crespo, J. E., Vinauger, C., & Lahondère, C. (2022). Species-specificity in thermopreference and CO_2-gated heat-seeking in Culex mosquitoes. Insects, 13(1), 92.

Line 333: Product references like the one seen here are needed for other products mentioned (such as Hemotek feeders on line 297).

Line 335: You say that they were tested in either light or dark phase, depending on if the mosquitoes were nocturnal vs. diurnal. However, were they tested during their peak activity periods? Culex are typically more crepuscular than nocturnal and are more active around the time the light cycle switches from light to dark. Were peak activity times taken into account?

Some papers about daily activity patterns:
Menzato, M. M., Conceição, G. M. D. S., Câmara, D. C. P., Suesdek, L., & Lima-Camara, T. N. (2025). Effects of blood feeding and insemination on the locomotor activity of Culex quinquefasciatus (Diptera: Culicidae). Journal of Medical Entomology, 62(3), 633-640.

Mutebi, J. P., Wilke, A. B. B., Ostrum, E., Vasquez, C., Cardenas, G., Carvajal, A., ... & Beier, J. C. (2022). Diel activity patterns of two distinct populations of Aedes aegypti in Miami, FL and Brownsville, TX. Scientific reports, 12(1), 5315.

Gray, E. M., & Bradley, T. J. (2003). Metabolic rate in female Culex tarsalis (Diptera: Culicidae): age, size, activity, and feeding effects. Journal of medical entomology, 40(6), 903-911.

Line 364-366: Is there a figure or some data showing how many individuals were removed from the analysis?

References:

Make sure to properly format scientific names. Lines 425-6 need the genus names capitalized. All scientific names in the references need to be italicized, and none of them are.

Reviewer #3

(Remarks to the Author)
In this study Uehara and colleagues present a new automated method for measuring the onset and duration of host-seeking in mosquitoes and further aim to characterize phylogenetic variation in these responses. The key claims of the study are that they have successfully developed this pipeline and implemented it to reveal differences in the duration of suppression of host-seeking following a bloodmeal.

There would be interest among those interested in comparing patterns of host-seeking among mosquitoes in such a high throughput and automated method. I would not go so far as to agree that the work will influence thinking in the field, but it describes a method that many in the field would find useful. This usefulness is augmented by the author's provision of code and information that would allow others to repeat their method.
My main concerns with the study as it is currently presented relate to how well the data presented support the key claims. I believe this stems from a lack of clarity about the hypotheses being tested, the statistical analysis, and the relationship between the data presented and those hypotheses. Therefore, I have focused my comments on these aspects of the study.
The stated aim of the study is to identify behavioral features that are differentially expressed and to evaluate whether these differences correspond to evolutionary relationships among species.
The authors convincingly establish that the automated classification of mosquito movements into three behavioral categories (walking, probing, flying) works as well as manual behavioral classification.

They then move into a comparison of different time points after blood feeding for 6 species. The caption of figure and tables provided make me think that they have compared the "Pre" bloodmeal period with post blood meal time points (24, 48 72 hrs). I am slightly puzzled by this analysis. They have additionally measured non-blood fed controls (Line 338-340) and it actually a comparison of how these individuals respond compared to the blood fed treatment over time that would determine when suppression (a decrease from controls) is occurring. To assess that hypothesis a statistical analysis on the interaction between time point and blood feeding treatment using appropriate mixed effects to account for the repeated measures on individuals would seem most appropriate.

It would be very helpful to include a description of the statistical approach in the methods. The table provided gives statistical outputs but leave the reader to guess exactly how tests were implemented. Furthermore, these files should be annotated such that it is clear what each column contains. Formatting these in a way that can be easily used in statistical analyses (such that each column is a data category with rows representing each measurement) would greatly facilitate data reuse. It would also help to have columns labeled with species, treatments, and time point for each data point.

The authors identify an important limitation of this assay. It doesn't capture directional information and is not optimized for long-range flight. I agree and would like to see more discussion what this assay can and cannot be used to say about host-seeking. It would have been nice to have an independent validation of this assay against some of the more "traditional" assays like flight mills, but in absence of that carefully explaining to those who might want to use this method where it is and, importantly, is not appropriate to apply is important.

Reviewer #4

(Remarks to the Author)
Review of "Behavioral heterogeneity in host-seeking and post-feeding suppression among disease vector mosquitoes" by Uehara et al.

This study presents a framework to comparatively characterize the host-seeking behavior of multiple mosquito species from three genera (Aedes, Anopheles, Culex) under laboratory conditions. The method for visual quantification of mosquito behavior appears to be carefully developed for high accuracy and has strong potential for standardizing behavioral characterization across diverse species. It is particularly interesting that host-seeking behaviors (characterized by probing, flying, and walking) were distinct among genera, while post-feeding suppression tended to be species-specific. Overall, the manuscript is well written, and the results are presented clearly.

Major Concerns

Biological replicates
It was unclear whether the experiments were conducted with biological replicates. While the figures indicate the number of mosquitoes tested, it was not specified whether these mosquitoes originated from a single cohort or multiple cohorts. This distinction is important, especially given the wide variation in mosquito age (7–14 days, as noted in the method). If a single data point represents one cohort of a particular age, host-seeking behavior could be heavily influenced by age (an effect well documented in the literature) interacting with the main factors the authors intended to examine.

Time of day
The timing of the experiments was also unclear. Many studies have shown that mosquito host-seeking activity is strongly influenced by the time of day and that this sensitivity varies by species. Although the authors mention that recordings were

made under appropriate light/dark cycles for each species, explicitly controlling for (or at least reporting) the time of day would strengthen the study.

Minor Points

Mosquito rearing protocols differed, with Aedes and Culex reared on TetraMin fish food, while Anopheles were reared on Koi fish food. I assume the larval food was chosen for the best rearing condition for each species, but I wonder if the food type could be also standardized across species as larval nutrition can influence host-seeking behavior (as the authors noted). Similarly, was larval density controlled across all cohorts? Larval density is known to strongly affect mosquito physiology and host-seeking behavior.

Line 293: “one” should be corrected to “once.”

Extended Data Fig. 1: The Y-axis is labeled “proportion,” but the definition is unclear. What exactly does “proportion” represent here?

Extended Data Fig. 4: The Y-axis is labeled “blood meal size,” but the values appear to reflect mosquito body weight. Could the authors clarify this? Additionally, how was the “fold increase” calculated?

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have properly addressed my concerns. I don't have further comments.

Reviewer #2

(Remarks to the Author)
Uehara et al. notes: The authors did well responding to feedback and editing and have a paper that has the potential to make a positive impact in the field of mosquito ethology. The manuscript now seems to only have minor changes to make, detailed below.

Line By Line (based on the tracked changes document):

Line 68-69: More diseases could be added here if the authors want, such as eastern equine encephalitis, western equine encephalitis, elephantiasis, etc.

Line 69-70: The connection here is not clear: are you saying that there have been field studies on seasonal suppression, but studies on postprandial host seeking suppression are lacking? Consider modifying the first part of the sentence to make the connection more clear for smoother delivery.

Line 102: Maybe start this sentence with “For example,” and add a comma after Aedes aegypti (runon sentence).

Line 150-159: This still feels like introduction, specifically the last paragraph of the introduction.

Line 163: What constitutes “reliable”?

Line 201: Elaborate on times, perhaps?

Lines 244-246: This new sentence does not flow well at this position. It would work better if it was moved to be right after the first sentence of this paragraph.

Figure 3: While this reviewer appreciates the increased information provided by the box plots in this figure, I am surprised at the marked change in error bars from the original graph. Were stats rerun to account for these changes? Additionally, the error bars on G & H extend to -300: what does -300 suppression mean? Please explain or modify graphs.

This is minor, but why do some paragraphs have a space after them and some do not? Line 333 vs 345?

Lines 305-316: The connection between Culex host preference and their increased walking and probing activity is unclear.

This should be elaborated upon, rephrased, and/or decoupled. What is it about being ornithophilic that would require an increase in walking/probing and not flying?

361: Add a comma after Aedes.

Line 496: Add a reference to the figure containing this data.

Line 505: Were all the species fed on sheep blood? Since this is not the preferred host for Culex, that could have been causing some behavioral differences (i.e. searching for other blood meal?). Were there any reductions in fecundity for any of the strains?

Reviewer #3

(Remarks to the Author)
Overall, the authors have done an excellent job integrating the comments from all reviewers. In particular the edits to Figures and their legends has greatly improved the clarity of the statistical approach. Most of my concerns have been addressed by their edits.

I have a few outstanding queries which they may need to address:

1. Line 203: suggest to replace "enhanced" with "increased"

2. Line 499: I cannot access the Zenodo Files (set to restricted access), but trust the authors have provided the line by line for each mosquito used in the analysis here.

3. Line 514: It would extremely useful to include very brief description of the statistical analyses here. It do agree that the figure legends have the information, but it would be helpful for readers to have a short paragraph explaining the key comparisons, which tests were used for normality, and which test has been applied for each. For example, “We tested all data for normality using a X test. To compare the proportion of time active in the pre and post- stimulus periods (Fig 1B) we used a Wilcoxon rank-sum test”. It is repeating the information but it helps to have this laid out for the reader in the appropriate section.
Apologies in advance, if I am being dense, but I am unable to locate the Supplemental Table 1 in the reviewer files. If this is referring to the checklist this is not the full details of the statistical methods. If there is a supplemental Table that provides all of this information then that would also be fine, but I'd prefer a short paragraph that gives these key details in one place in the main text.

Reviewer #4

(Remarks to the Author)
The authors have addressed all my concerns in their revision. I have no further issues and comments.

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Responses to reviewers:

Reviewer #1 (Remarks to the Author):

This manuscript establishes an automated video tracking and machine learning classification pipeline to compare host-seeking behavior and post-blood meal suppression in six mosquito species across three major genera (Aedes, Anopheles, Culex). The authors find notable interspecific variation in baseline host-seeking repertoires and suppression timelines. The approach provides high temporal resolution and fine-scale posture analysis, which is valuable for comparative behavioral studies.

While the study is well-conceived and technically advanced, several conceptual, methodological, and interpretational issues need to be addressed before the manuscript can be considered for publication.

We appreciate your acknowledgment of our findings and will address the points raised.

Major Comments

1) Control of biological variables

a) The authors do not state whether the strains were lab-adapted or field-derived, nor do they clearly indicate the generation in captivity. Long-term colonization can significantly alter host-seeking behavior.

b) Larval diet protocols differ slightly between Aedes/Culex (TetraMin) and Anopheles (Koi food). This is a confounding factor because larval nutrition is known to affect adult host-seeking and suppression. At minimum, the authors should justify why the diets were not standardized or discuss the potential bias this introduces.

c) Egg hatching conditions are not fully equivalent across species. For a comparative behavioral study, the developmental rearing environment should be standardized as far as possible.

a) We apologize for omitting this important detail: all strains we used in this study are lab-adapted and obtained from research groups or the BEI stock center, and these strains are broadly used for behavioral assays on host-seeking and blood-feeding behaviors. We have now added these in the Methods as follows:

> "The following species and strains were used in this study: *Aedes aegypti* Liverpool, *Aedes albopictus* Foshan, *Anopheles gambiae, Anopheles stephensi, Culex tarsalis* (NR-43026), *Culex quinquefasciatus* (NR-43025) (BEI Resources). *Aedes aegypti* Orlando strain was used for training data. These strains are laboratory-reared and maintained for research purposes."

b) and c) We thank the reviewer for raising this important point. To ensure the proper maintenance of the animals tested in our behavioral setup, we carefully evaluated rearing protocols for each species, based on recent publications and consultations with the laboratories and stock centers from which the strains were obtained. These rearing standards take into account the species-specific nutritional requirements. We hypothesize that a sudden change in diet or rearing conditions could lead to malnutrition, which could affect performance in our behavioral assays. By adhering to these established protocols, we also ensure that the phenotype we report is comparable to data from other researchers in the field. A new statement has been added to the Methods Section to provide additional details and references for the rearing protocols used:

> "All mosquitoes were maintained and reared at 28°C, 70–80% relative humidity with a 12 h light: 12 h dark schedule. Feeding followed standard procedures appropriate for each species[62–65]. *Aedes* and *Culex* eggs were hatched using 1 L of TetraMin fish food (Tetra) suspension in deionized deoxygenated water (one tablet/L). Larvae were fed with TetraMin fish food tablets once daily as needed. Anopheles eggs were hatched in deionized deoxygenated water. Larvae were fed with Koi fish food tablets (Drs. Foster & Smith) once daily as needed. Larval density was adjusted to approximately 250 individuals per rearing pan (17 × 16 × 10 cm). Male and female mosquitoes were co-housed, and adult female mosquitoes used for experiments were between 7 and 14 days old, with ad libutum access to 10% sucrose solution. Females were blood-fed on defibrinated sheep blood (Quad5) with 1 mM ATP at 37°C using Hemotek artificial membrane feeders (SP6W1-3, Hemotek Ltd., UK)."

An additional statement has been added to the Discussion to acknowledge that it is possible that these differences in rearing may impact behavior:

> "In this study, we reared *Anopheles* species on a different diet than *Aedes* and *Culex* species according to standard laboratory protocols and, although all individuals were healthy, these differences may have influenced nutritional status and behavior.

2) Host cue choice and species ecology

a) The stimuli used (human odor + 10% CO_2) are optimal for anthropophilic Aedes and Anopheles gambiae, but suboptimal for Culex, especially ornithophilic species like Culex tarsalis. This might explain part of the behavioral differences rather than true physiological suppression differences. While the authors acknowledge this in Discussion, they should either include an avian odor control condition, or interpret Culex results with more caution.

b) The claim of a "unified behavioral framework" is weakened if stimulus relevance varies greatly by species.

Although *Culex* species are strongly attracted to avian odors and commonly feed on birds, the use of human odor in our experiments is important. Despite their avian preference, *Culex* mosquitoes also feed on humans, and their attraction to humans after an avian blood meal is critical for pathogen transmission. Our study focuses specifically on this anthropophilic component of their host preference and how it relates to blood-feeding status, given its direct relevance to human pathogen transmission. We have revised the text to state this explicitly:

> "These interspecific differences may arise from species-specific host ecology and sensory tuning. Our stimuli, human odor and CO_2, are highly effective for species such as *Aedes aegypti* and *Anopheles gambiae*[4,51–53] but may be less preferred for *Culex* species that often seek avian hosts[54], because birds emit distinct volatile blends compared with human, particularly in aldehyde and carboxylic acid composition and abundance[55,56]. Nevertheless, *Culex* mosquitoes do show attraction to humans, which contributes to their capacity to transmit pathogens to people, including avian-derived pathogens such as West Nile virus[57,58]."

At the same time, we confirmed that non-fed *Culex* females clearly responded to the stimuli used (human odor + 10% CO_2), which we consider to represent their baseline response. Therefore, although the stimuli may be not represent the most preferred host for these species, we believe this is a meaningful response from a pathogen transmission standpoint and supports our main argument for generalized host-seeking suppression.

3) Replicate numbers and statistical clarity
a) Sample sizes vary widely (n = 15–55 females) without clear rationale. This imbalance may affect statistical power. Authors should clarify how many replicates per species per timepoint were initially planned and whether any were excluded (e.g., due to non-feeding).
b) The manuscript uses multiple non-parametric and parametric tests interchangeably. It should be clear why certain tests were chosen for each dataset and whether normality was assessed.

a) We thank the reviewer for raising this point, we have now included the sample size across different analysis stages in our supplementary data. Varying sample size is a result of differential baseline activity level among species, which leads to more exclusions in certain species to ensure that we only report host-cue evoked, but not spontaneous, behavioral signatures. However, all comparisons were performed within species, and no multiple-comparison corrections were applied between species. Therefore, we do not expect this to affect statistical power.
b) Additionally, following the reviewer's request, we have expanded the Statistical Analysis section in Materials and Methods to clarify test selection and added exact N values for all experiments, in line with journal guidelines. We performed Shapiro–Wilk tests to assess normality prior to applying parametric methods. Although some data deviated from normality and were accordingly reanalyzed using non-parametric approaches (Wilcoxon Rank Sum tests for pairwise non-parametric comparisons and Kruskal–Wallis tests for multi-group comparisons), these changes did not alter our conclusions.

4) Definition of "suppression" and baseline

a) Suppression is quantified as percent reduction from baseline, but baseline is defined as the pre-stimulus period in unfed controls. This assumes stability of unfed activity across days and individuals, which might not hold for all species. Baselines should ideally be measured in the same individuals before feeding.

b) Movement and probing are lumped as indicators of host-seeking, but high locomotion without probing may reflect oviposition searching rather than recovery of host-seeking.

a) The baseline for the behavioral suppression assay was defined as the post-stimulus period in non-fed controls. "Pre" in Fig3 A-F indicated pre-feeding activity, not pre-stimulus activity. We have revised this to "non-fed" for clarity. As suggested, comparing behavioral suppression before and after feeding within the same individuals would be ideal. However, we reasoned that prior experience in the assay could influence subsequent behavioral phenotypes and multiple rounds of cold anesthesia and manual loading might further complicate the replicability of our assay. Therefore, we used comparisons between cohorts of different individuals rather than within-individual comparisons.

b) We apologize for this confusion. Please note that we did not intend to report host-seeking using locomotion, but to analyze the probing pattern comparatively with locomotor activity level. This is introduced in our results section when we first discuss the two metrics: "To differentiate general activity from host-seeking behavior, we evaluated both overall movement and probing, which is an indicator of active host-seeking."

We agree with the reviewer that high locomotion without probing might reflect oviposition behavior, as recently demonstrated in our preprint in gravid *Aedes aegypti* females (Dong et al. 2025 BioRxiv). Although this study was not designed to test for oviposition search, this could be a future application of the pipeline. We now discuss this possibility in the text:

> "Additionally, variations in the gonotrophic cycle, the time from blood-feeding to oviposition, may influence the duration of suppression. In *Anopheles*, oviposition generally occurs 48-60 hours after blood feeding[13] whereas oviposition may require more time in *Aedes* and these females can retain eggs and maintain host-seeking suppression if no suitable oviposition sites are available[61,62]. Notably, *Culex* females may require more than 5 days to oviposit after blood feeding[63] and these females exhibited high movement relative to probing at the 72-hour timepoint (Fig. 3H) suggesting that behavioral activation may represent a non-host-seeking state, potentially for oviposition."

5) Machine learning model generalizability

a) Separate DeepLabCut models were required for Anopheles, suggesting limited cross-species generalizability. If a "unified" approach is claimed, the authors should discuss how model choice might *bias classification.*

b) The validation accuracy (92–96%) is good, but it is based on genus-specific training sets. Cross-validation across genera (train on one genus, test on another) would better demonstrate general applicability.

We agree with the reviewer that the wording of a “unified” approach might be misleading. We initially attempted to develop a universal model for key-point tracking and pose classification but were unable to achieve sufficient cross-species accuracy, even after expanding the training dataset to 720 individuals. As a result, we shifted to species-specific models to ensure reliable tracking. Our use of the term “unified approach” refers to the overall analytical pipeline, which maintains a consistent framework while accommodating the need for genus-specific models. However, as the reviewer pointed out, genus specific tracking models are essential to accurate behavioral analysis. Therefore, we have removed the claim of establishing a “unified” behavioral framework from the abstract.

6) Interpretation discrepancies with previous literature

a) For Anopheles gambiae, prolonged suppression up to 72 h conflicts with classic findings of early recovery. This discrepancy needs deeper discussion—whether due to strain origin, assay environment, or host cue type.

We appreciate this important point. Indeed, previous studies report earlier recovery of host-seeking in *Anopheles gambiae*, while our study sees host-seeking suppression in a longer timeline. It is important to note that, however, earlier studies like Klowden and Briegel (1994) and Takken et al. (2001) reported host-seeking with population metrics (e.g. response percentage) in a group-based assay, while we house mosquitoes individually and analyzed postural signature. Due to the fundamental differences in the aspects of host-seeking behavior that were measured across these studies, it is hard to make direct comparisons about recovery timeline as a whole. However, we think these seemingly contradicting results might be a biologically relevant read out of the specific behavioral repertoire of a blood-fed mosquito when studied in a group vs in isolation, and in terms of flight paths vs postural signature.

Comparisons across previous studies reveal variation in the recovery time from host-seeking suppression for this species, suggesting that intraspecific differences may influence this timeline and we have expanded on our interpretation in the revised manuscript:

> “While all species examined exhibited host-seeking suppression after blood feeding, the onset, duration, and intensity of suppression varied. This suppression is believed to have evolved as an adaptive strategy to optimize reproductive success by reallocating energy from host-seeking toward egg development[39,59]. By temporarily ceasing host-seeking behavior, female mosquitoes

can also avoid potentially dangerous host defensive behaviors during this critical period. Interestingly, we observed differences in host-seeking suppression even among closely-related species: *Anopheles gambiae* exhibited prolonged suppression, while *Anopheles stephensi* showed a significant recovery of movement and probing by 72 hours post-feeding. This contrasts with previous studies suggesting that *Anopheles gambiae* females rapidly resume host-seeking[14], and suggests that significant variation likely exists within strains of the same species. It also suggests that housing conditions (group versus solitary) and the behavioral metric used (posture versus attraction) could influence experimental conclusions."

Minor comments
L184: Delete the second "of"

Thank you for catching this error – we have now corrected it.

L370–376: How did you recognize the same individual after the mosquitoes were mixed for blood feeding? If identities were not preserved, it seems only the average blood-meal size of the batch was measured. In that case, the estimate would be a cohort mean and would not account for individual differences. Please clarify whether pre- and post-feeding weights were paired for each female (and how).

We appreciate the reviewer's clarification on this point and have revised the manuscript accordingly. Different individuals from the same rearing group were used for meal-size comparisons, and an explanation has been added to the Statistical analysis section in the Materials and Methods: "All behavioral assays and blood meal size quantifications were performed with female mosquitoes from multiple rearing cohorts".

Reviewer #2 (Remarks to the Author):

Summary:
Duvall et al. propose a new method for modeling and evaluating host seeking in mosquitoes as well as a comparison of post-blood meal host-seeking of different anthropophilic mosquito species, both of which are novel and interesting. However, the organization of the paper as well as the literature reviews (particularly in the discussion) require revision. Below are line-by-line suggestions to aid in revisions as well as some readings that may be useful to reference.

We appreciate your time and effort in evaluating our study. We have incorporated all of your suggestions as much as possible.

Introduction

Line 34: The citation punctuation differs here. Please revise.

Also Line 34: Vision can also play a role in host seeking. Please refer to Raji & DeGennero (Raji, J. I., & DeGennaro, M. (2017). Genetic analysis of mosquito detection of humans. Current opinion in insect science, 20, 34-38.)

Lines 51-53: For a more recent study on how larval nutrition can affect behavior: Vinauger, C., & Chandrasegaran, K. (2024). Context-specific variation in life history traits and behavior of Aedes aegypti mosquitoes. Frontiers in Insect Science, 4, 1426715.

We thank the reviewer for this suggestion and have now added these references.

Lines 88-91: This feels like it should cover a separate paragraph. This final intro paragraph talks about machine learning and how the authors will be applying it to their pipeline; these first couple sentences are a separate topic.

Line 91-92: This statement seems to be missing a verb. Please rephrase.

We thank the reviewer for pointing out this disconnect. We intended to highlight that many behavioral modes are traditionally assessed using end-point measures (e.g., number of individuals attracted), whereas automated key-point tracking and pose classification pipeline allow behavioral assays to capture whole behavioral sequences. We have now updated this paragraph to more clearly make these connections:

> "Host-seeking is typically measured using endpoint assays that assess attraction to traps baited with host cues[39–41]. However, females often show increased flight activity before host-seeking fully resumes, potentially influenced by egg-laying behaviors that also elevate activity in the absence of active host-seeking[36,42]. As a result, general activity measures alone can produce false positives or inaccurate interpretations of behavioral state. Recent advances in ethology increasingly utilize machine learning pipelines that distinguish subtle postures and activity sequences, features that endpoint assays may miss and that can help differentiate true host-seeking from unrelated changes in activity[43]. By generating detailed postural signatures of host-seeking and its suppression after blood-feeding, these approaches can provide deeper insight into behavioral states beyond general activity levels[44,45].

General comment: The preferred host of each species is mentioned in the discussion, but not in the introduction. It may be worth noting that some species of Culex, and even some strains within species, can prefer avian hosts. This is true for Aedes aegypti as well.

We appreciate the reviewer's suggestion and agree that host preference is an important consideration. We have added new language in the Introduction:

> "While some species, including *Aedes aegypti* and *Anopheles gambiae*, strongly prefer humans, even more generalist feeders, including *Aedes albopictus* or ornithophilic *Culex* mosquitoes can contribute to pathogen transmission when they bite people, especially if they have previously fed on other hosts[4,10–12]."

We have also expanded our discussion of this point and clarified out design to use human host stimuli as interaction with humans in all of these species is critical for driving pathogen transmission:

> "Host-seeking females of all species tested exhibited walking, probing, and flight behaviors following stimulus onset, whereas *Culex* species showed a marked bias toward walking and probing. These interspecific differences may arise from species-specific host ecology and sensory tuning. Our stimuli, human odor and CO_2, are highly effective for species such as *Aedes aegypti* and *Anopheles gambiae*[4,51–53] but may be less preferred for *Culex* species that often seek avian hosts[54], because birds emit distinct volatile blends compared with human, particularly in aldehyde and carboxylic acid composition and abundance[55,56]. Nevertheless, *Culex* mosquitoes do show attraction to humans, which contributes to their capacity to transmit pathogens to people, including avian-derived pathogens such as West Nile virus[57,58]."

Results

Lines 113-117: This section should be in the introduction, since you are using these species throughout experiments.

We appreciate the reviewer's suggestion and agree that information on the species and their evolutionary relationships provides valuable context for the study, however we ultimately feel that moving this section to the Introduction would shift the focus away from the main themes of that section: host-seeking suppression, underlying physiological mechanisms, and advances in ethological analysis. To maintain clarity and structure, we have therefore kept this material in the Results section, where it directly supports the presentation of our comparative analyses. Nonetheless, we have updated our Introduction to ensure that it more clearly motivates the use of multiple species and the rationale for a comparative approach.

Lines 113-117: Recent studies show the lineages may be older than that: Soghigian, J., Sither, C., Justi, S. A., Morinaga, G., Cassel, B. K., Vitek, C. J., ... & Wiegmann, B. M. (2023). Phylogenomics reveals the history of host use in mosquitoes. Nature Communications, 14(1), 6252.

We appreciate this suggestion and have now incorporated this reference.

Line 166: This is true of anthropophilic species, but these behaviours may not be conserved in mosquitoes that feed on other hosts.

We agree that this is an important point. However, we confirmed that non-fed individuals of all species clearly responded to the stimuli used (human odor + 10% CO_2), including species that typically feed on other hosts. We therefore used this response as their baseline response for each species. Although, the stimuli may be non-preferred for some species, we believe that our results on behavioral comparison still reflect distinct host-seeking strategies across species.

Fig 2C-D: If the data is available, it might be nice to show a bit more of the activity from before stimulus delivery to even better highlight the difference in activity. For example, comparing 3 minutes of pre-stimulus activity vs. 8 minutes of post-stimulus activity rather than 1 minute vs. 8 minutes.

We appreciate the reviewer's suggestion to include a longer pre-stimulus window for comparison. Unfortunately, we did not collect pre-stimulus data to extend this period in the figures. However, based on our preliminary experiments, the representative traces included, comparing 1 minute of pre-stimulus activity with 8 minutes of post-stimulus activity, are consistent with activity patterns observed at earlier and later time points. We therefore believe that the current presentation accurately captures the key differences in activity relevant to this analysis.

Figure 2F: Why are the PCAs separated? Putting them all together will help the reader to see the overlap for themselves. Or at least putting mosquito species of the same genus together would help to see the overlap, since that is the point being made.

We appreciate the reviewer's suggestion to combine the PCA plots to facilitate direct visual comparison. This was indeed our first choice when we originally plotted these out. However, density maps overlaid on the two-dimensional coordinates was difficult to interpret when multiple species are plotted together. For this reason, we opted for separate panels to allow the patterns within each species to remain easily interpretable. We have also considered overlaying the two species from the same genus, but we remained careful not to overlay species based on our subjective judgement after seeing these plots. We think it might introduce bias when comparing species within the same genus is made easier than across genera.

Fig 2 - Do the authors have any hypotheses about the lack of flying in Cx. tarsalis or any other low level *activity (extended data)? Please discuss these kinds of results in the discussion. Also please include data somewhere about the number of individuals removed from the analysis.*

We apologize for this potentially confusing point. Due to high individual variation in flight response and the journal requirement to show data distribution, the y-axis in Figure 2D is not matched between species to highlight differences within the same species across the recording timeline. To compensate for this, we included panel E to help visualize these behavioral types on the same scale, which does reveal species-specific difference in activity level. We propose that these differences in host-seeking responses may arise from species-specific host ecology and sensory tuning and note that the design of our arena presents limitations in quantifying aspects of navigation and long-range flight. We have added text to the Discussion to address this:

> "The machine learning pipeline enabled automated, observer-independent tracking ands classification, capturing behavioral dynamics through metrics such as repertoire, flight distance, and probing time, rather than relying on traditional endpoint measurements. However, this setup was not optimized for long-range flight, limiting our ability to assess directional navigation. Thus, while it is well suited for quantifying the patterns, duration, and frequency of behavioral responses to specific stimuli, it should be complemented by traditional approaches, such as wind tunnel assays, to evaluate orientational responses, including attractiveness and repellency."

In addition, we now include a detailed supplementary table (Supplementary Table S1: "Exclusion by Species and Condition") which fully tracks the number of individuals removed from the analysis.

Lines 184-185: There is a typo "of is"

Thank you for catching this error, we have now corrected this.

Discussion:
Line 227-228: This is a good point, which you could talk about more. What is this setup good for? What is the significance/novelty of it? What knowledge gaps could you fill with this setup? You don't have direction, but you can measure response levels to known attractants/repellants, which could allow you to determine thresholds for detection of compounds (almost like a somewhat more accessible electroantennogram (EAG)?). Expand on why you think this system is important to share with the scientific community in the discussion.

We appreciate the reviewer's encouraging comment and further input. We have now added additional text in the Discussion to address this and more clearly highlight the novelty of this analysis:

"Because the system captures continuous behavioral output at high resolution, it can be used to estimate detection thresholds and response sensitivities to attractants or repellents offering an accessible complement to electrophysiological approaches. By enabling standardized, cross-species comparisons of fine-scale behavioral responses, this fills an important methodological gap and provides a resource for the community to study the mechanistic basis of host seeking with greater precision, laying the groundwork for future studies investigating the molecular and neural mechanisms behind behavioral suppression and how they vary between species. The approach can also be expanded to study responses to different hosts or environmental cues."

Line 244-245: There may be variation within species as well, as different strains can vary in behaviour and life history traits, as demonstrated in the paper below. Some populations/strains are more anthropophilic than others as well, since some places in the world consider Cx. quinquefasciatus to be ornithophilic.

Regarding this point, we have added text to discuss potential differences in host preferences (related to response above to host preference):

"Host-seeking females of all species tested exhibited walking, probing, and flight behaviors following stimulus onset, whereas *Culex* species showed a marked bias toward walking and probing. These interspecific differences may arise from species-specific host ecology and sensory tuning. Our stimuli, human odor and CO_2, are highly effective for species such as *Aedes aegypti* and *Anopheles gambiae*[4,51–53] but may be less preferred for *Culex* species that often seek avian hosts[54], because birds emit distinct volatile blends compared with human, particularly in aldehyde and carboxylic acid composition and abundance[55,56]. Nevertheless, *Culex* mosquitoes do show attraction to humans, which contributes to their capacity to transmit pathogens to people, including avian-derived pathogens such as West Nile virus[57,58]."

Lines 268-271: This statement does not have much of an impact. Please consider revising or removing.

We have now removed this statement.

General comment: Overall, discussion needs more comparison of the results with known literature as well as explanations/hypotheses about the results.

We agree with the reviewer and we have substantially revised our discussion to provide additional interpretation and context for our findings related to host preference, species differences in reproductive cycle duration and timing, as well as candidates for differential mechanosensory or neuropeptide regulation.

Methods

Line 295: This sentence is a run-on and needs a comma before "and."

Thank you – we have now corrected this.

Line 308: Please include a reference for the previous paper.

We thank the reviewer for catching this and have now included this reference.

Line 307-322: It would be really helpful to have schematics of these to visualize your setup. This reviewer is unable to visualize beyond the 6 well plate.

The overall behavioral platform was developed and described in a recent publication from our laboratory (Dong L, Hormigo R, Barnett JM, Greppi C, Duvall LB. (2025) Circadian modulation of mosquito host-seeking persistence by Pigment-Dispersing Factor impacts daily biting patterns. Proc Natl Acad Sci USA. 122 (46) e2520826122) and we now refer readers to this publication for details about assay setup and design.

Line 311: 10% CO2 would be roughly 100,000ppm–humans only emit about 30,000ppm CO2. Is there a reason why the researchers used such a high concentration? Or was it diluted? Please explain your reasoning for using such a high concentration.
Culex species will sometimes even respond to less, since they prefer avian hosts.

We appreciate this comment. The inflow of air containing 10% CO_2 was controlled in pulses and introduced into the odor box, where it mixed with ambient air (approximately 400 ppm CO_2). As a result, the CO_2 concentration gradually increased and was estimated to reach 40,000–70,000 ppm after about 3 minutes. While this level is slightly higher than that emitted by humans, it was used to ensure a robust consistent behavioral response across trials.

Line 333: Product references like the one seen here are needed for other products mentioned (such as Hemotek feeders on line 297).

Thank you for catching this oversight, we have now included product references and identified the suppliers of the major materials used in this study.

Line 335: You say that they were tested in either light or dark phase, depending on if the mosquitoes were nocturnal vs. diurnal. However, were they tested during their peak activity periods? Culex are typically more crepuscular than nocturnal and are more active around the time the light cycle switches from light to dark. Were peak activity times taken into account?

This is an important point. Although we did not determine the locomotor activity of each strain, we ensured that the experiments were performed during the range of their expected peak activity times. Specifically, experiments for diurnal species were conducted 3–6 hours after lights turned on, whereas experiments for nocturnal species were conducted 1–4 hours after lights turned off. We have added this explanation to the Materials and Methods:

> "Behaviors on diurnal (*Aedes aegypti* and *Aedes albopictus*) and nocturnal (*Anopheles gambiae*, *Anopheles stephensi*, *Culex tarsalis*, and *Culex quinquefasciatus*) species were recorded in light and dark phase, respectively. Experiments for diurnal species were conducted 3–6 hours after lights turned on, whereas experiments for nocturnal species were conducted 1–4 hours after lights turned off, corresponding to the range of their peak activity times."

Line 364-366: Is there a figure or some data showing how many individuals were removed from the analysis?

We thank the reviewer for raising this point. We have now included a detailed supplementary table (Supplementary Table S1: "Exclusion by Species and Condition") which fully tracks the number of individuals removed from the analysis. All other group details are provided in Supplementary Table S1.

References:
Make sure to properly format scientific names. Lines 425-6 need the genus names capitalized. All scientific names in the references need to be italicized, and none of them are.

We apologize for this oversight. We have now updated our formatting in references as suggested.

Reviewer #3 (Remarks to the Author):

In this study Uehara and colleagues present a new automated method for measuring the onset and duration of host-seeking in mosquitoes and further aim to characterize phylogenetic variation in these *responses. The key claims of the study are that they have successfully developed this pipeline and implemented it to reveal differences in the duration of suppression of host-seeking following a bloodmeal. There would be interest among those interested in comparing patterns of host-seeking among mosquitoes in such a high throughput and automated method. I would not go so far as to agree that the work will influence thinking in the field, but it describes a method that many in the field would find useful. This usefulness is augmented by the author's provision of code and information that would allow others to repeat their method.*

My main concerns with the study as it is currently presented relate to how well the data presented support the key claims. I believe this stems from a lack of clarity about the hypotheses being tested, the statistical analysis, and the relationship between the data presented and those hypotheses. Therefore, I have focused my comments on these aspects of the study.

We appreciate your recognition of the significance of this paper. The reviewers' concerns are addressed as follows.

The stated aim of the study is to identify behavioral features that are differentially expressed and to evaluate whether these differences correspond to evolutionary relationships among species. The authors convincingly establish that the automated classification of mosquito movements into three behavioral categories (walking, probing, flying) works as well as manual behavioral classification.

They then move into a comparison of different time points after blood feeding for 6 species. The caption of figure and tables provided make me think that they have compared the "Pre" bloodmeal period with post blood meal time points (24, 48 72 hrs). I am slightly puzzled by this analysis. They have additionally measured non-blood fed controls (Line 338-340) and it actually a comparison of how these individuals respond compared to the blood fed treatment over time that would determine when suppression (a decrease from controls) is occurring. To assess that hypothesis a statistical analysis on the interaction between time point and blood feeding treatment using appropriate mixed effects to account for the repeated measures on individuals would seem most appropriate.

We appreciate the reviewer's insightful comment. Our analysis was designed to compare the distance traveled and the time probing before blood feeding and at post-blood feeding time points (24, 48, and 72 hours). Because the sample size for each group was relatively small (N = 15–55) and the data did not meet the assumptions required for mixed-effects modeling, we used non-parametric tests for these comparisons.

It would be very helpful to include a description of the statistical approach in the methods. The table provided gives statistical outputs but leave the reader to guess exactly how tests were implemented.

Furthermore, these files should be annotated such that it is clear what each column contains. Formatting these in a way that can be easily used in statistical analyses (such that each column is a data category with rows representing each measurement) would greatly facilitate data reuse. It would also help to have columns labeled with species, treatments, and time point for each data point.

We thank the reviewer for these important suggestions regarding transparency and data reuse. We have addressed the concerns as follows:

1. Statistical Methods and Annotation
We have fully revised the Methods section and the figure legends to clearly detail the specific statistical tests, the handling of multiple comparisons, and the software used. Furthermore, we agree that clarity is paramount. We have improved the legends and annotations for all Supplementary data to explicitly define all column headings, abbreviations, treatments, species, and time points, thus ensuring the tables are fully understandable.

2. Data Format and Reuse
We appreciate the suggestion to format the data in a "tidy" manner for facilitated reuse. We have addressed this by:
Raw Data Availability: The complete raw data, suitable for statistical re-analysis, has been uploaded to the public repository at https://doi.org/10.5281/zenodo.15478199. This data is provided in a standard, machine-readable format that is easily convertible to tidy data, fully satisfying the requirement for data reuse.
Supplementary Data Purpose: The Supplementary Tables are specifically designed to present summary statistics and analyzed outputs to support the manuscript's findings. For this purpose, we maintain the current format, as it maximizes human readability and immediate understanding, while the raw data in the repository facilitates computational reuse.

The authors identify an important limitation of this assay. It doesn't capture directional information and is not optimized for long-range flight. I agree and would like to see more discussion what this assay can and cannot be used to say about host-seeking. It would have been nice to have an independent validation of this assay against some of the more "traditional" assays like flight mills, but in absence of that carefully explaining to those who might want to use this method where it is and, importantly, is not appropriate to apply is important.

We agree with the reviewer that it is important to clarify the capabilities and limitations of this assay. The machine learning pipeline enables automated, observer-independent tracking and classification, capturing behavioral dynamics through metrics such as repertoire, flight distance, and probing time, rather than relying on traditional endpoint measurements. However, this assay does not capture directional information and is not optimized for long-range flight. We have added this point and limitations to the Discussion to guide readers on the appropriate use of this method.

> "The machine learning pipeline enabled automated, observer-independent tracking and classification, capturing behavioral dynamics through metrics such as movement repertoire, flight distance, and probing time, rather than relying on traditional endpoint measurements. However, this setup was not optimized for long-range flight, limiting our ability to assess directional navigation. Thus, while it is well suited for quantifying the patterns, duration, and frequency of behavioral responses to specific stimuli, it should be complemented by traditional approaches, such as wind tunnel assays, to evaluate orientational responses, including attractiveness and repellency."

Reviewer #4 (Remarks to the Author):

Review of "Behavioral heterogeneity in host-seeking and post-feeding suppression among disease vector mosquitoes" by Uehara et al.
This study presents a framework to comparatively characterize the host-seeking behavior of multiple mosquito species from three genera (Aedes, Anopheles, Culex) under laboratory conditions. The method for visual quantification of mosquito behavior appears to be carefully developed for high accuracy and has strong potential for standardizing behavioral characterization across diverse species. It is particularly interesting that host-seeking behaviors (characterized by probing, flying, and walking) were distinct among genera, while post-feeding suppression tended to be species-specific. Overall, the manuscript is well written, and the results are presented clearly.

We appreciate your recognition of the value of the framework we proposed and we hope to have addressed your remaining concerns in our revised manuscript.

Major Concerns
Biological replicates
It was unclear whether the experiments were conducted with biological replicates. While the figures indicate the number of mosquitoes tested, it was not specified whether these mosquitoes originated from a single cohort or multiple cohorts. This distinction is important, especially given the wide variation in mosquito age (7–14 days, as noted in the method). If a single data point represents one cohort of a particular age, host-seeking behavior could be heavily influenced by age (an effect well documented in the literature) interacting with the main factors the authors intended to examine.

This is an important point. We have added a Statistical Analysis section in the revised manuscript and to clarify this point. We ensured that all data were collected from multiple rearing cohorts. Regarding age, we carefully ensured that mosquitos were neither too young nor too old for the experiments. Behavioral responses were recorded from 0 to 3 days post-feeding; for example, if recordings began when flies were 7 days old, they continued until 10 days old. Therefore, we consider the age window used in this study to be appropriate.

Time of day

The timing of the experiments was also unclear. Many studies have shown that mosquito host-seeking activity is strongly influenced by the time of day and that this sensitivity varies by species. Although the authors mention that recordings were made under appropriate light/dark cycles for each species, explicitly controlling for (or at least reporting) the time of day would strengthen the study.

We ensured that the experiments were performed during the range of their expected peak activity times. Specifically, experiments for diurnal species were conducted 3–6 hours after lights turned on, whereas experiments for nocturnal species were conducted 1–4 hours after lights turned off. We have added this explanation to the Materials and Methods:

> "Behaviors on diurnal (*Aedes aegypti* and *Aedes albopictus*) and nocturnal (*Anopheles gambiae, Anopheles stephensi, Culex tarsalis*, and *Culex quinquefasciatus*) species were recorded in light and dark phase, respectively. Experiments for diurnal species were conducted 3–6 hours after lights turned on, whereas experiments for nocturnal species were conducted 1–4 hours after lights turned off, corresponding to the range of their peak activity times."

Minor Points

Mosquito rearing protocols differed, with Aedes and Culex reared on TetraMin fish food, while Anopheles were reared on Koi fish food. I assume the larval food was chosen for the best rearing condition for each species, but I wonder if the food type could be also standardized across species as larval nutrition can influence host-seeking behavior (as the authors noted). Similarly, was larval density controlled across all cohorts? Larval density is known to strongly affect mosquito physiology and host-seeking behavior.

We thank the reviewer for raising this point. Regarding differences in larval diet, we considered that the observed behavior might not reflect the mosquitos' natural tendencies if unhealthy adults were reared on inappropriate diets. Therefore, in this study, larvae of each species were reared according to standard protocols. To ensure the proper maintenance of the animals tested in our behavioral setup, we carefully evaluated rearing protocols for each species, based on recent publications and consultations with the laboratories and stock centers from which the strains were obtained. To address the possibility that differences in larval diet could influence host-seeking behavior, we have added a new statement to the text: "*In this study, we reared Anopheles species on a different diet than Aedes and Culex species according to standard laboratory protocols and, although all individuals were healthy, these differences may have influenced nutritional status and behavior."* Regarding larval density, we adjusted it to approximately 250 individuals per rearing pan (17 × 16 × 10 cm) and we have added this detail to our Methods section.

Line 293: "one" should be corrected to "once."

We have revised the manuscript accordingly.

Extended Data Fig. 1: The Y-axis is labeled "proportion," but the definition is unclear. What exactly does "proportion" represent here?

Thank you for catching this error. We have now revised the Y-axis label to "Average confidence score" (values 0–1, indicating prediction reliability) instead of "proportion," and added explanatory text to the figure legend as follows:

> "Scores are calculated for every video frame from model-generated heatmaps, then averaged per key point. These values represent the model's certainty in body-part localization, where higher means more reliable tracking."

Extended Data Fig. 4: The Y-axis is labeled "blood meal size," but the values appear to reflect mosquito body weight. Could the authors clarify this? Additionally, how was the "fold increase" calculated?

Since body size differs among species, the absolute meal size is not a suitable indicator for comparison. Therefore, we used the fold increase, calculated as the post-blood feeding body weight divided by the pre-blood feeding body weight. We have now added these details in out Methods Section.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

We appreciate the time and effort of the Early Career Researcher who co-reviewed our manuscript and thank them for their evaluation of our work.

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The authors have properly addressed my concerns. I don't have further comments.
We thank the reviewer for their comments and are happy to hear that we addressed all concerns.

Reviewer #2 (Remarks to the Author):

Uehara et al. notes: The authors did well responding to feedback and editing and have a paper that has the potential to make a positive impact in the field of mosquito ethology. The manuscript now seems to only have minor changes to make, detailed below.

Line By Line (based on the tracked changes document):

Line 68-69: More diseases could be added here if the authors want, such as eastern equine encephalitis, western equine encephalitis, elephantiasis, etc.
We have taken this suggestion and have updated this sentence:

> "While most research has focused on *Aedes* and *Anopheles*, *Culex* mosquitoes also feed on humans and can transmit pathogens such as St. Louis Encephalitis, Eastern Equine Encephalitis, Western Equine Encephalitis, and West Nile viruses as well as the parasites that cause Lymphatic Filasriasis[24–26]."

Line 69-70: The connection here is not clear: are you saying that there have been field studies on seasonal suppression, but studies on postprandial host seeking suppression are lacking? Consider modifying the first part of the sentence to make the connection more clear for smoother delivery.
Thank you for catching this, we have now updated this sentence to clarify this point:

> "Although seasonal modulation of host seeking has been documented in some *Culex* species, and several field and laboratory studies have addressed this phenomenon, laboratory studies specifically examining postprandial suppression of host-seeking in *Culex* are lacking[27]."

Line 102: Maybe start this sentence with "For example," and add a comma after Aedes aegypti (runon sentence).
We appreciate this comment and have incorporated this change.

Line 150-159: This still feels like introduction, specifically the last paragraph of the introduction.
We thank the reviewer's detailed comment. During manuscript writing, we collected internal feedback on the flow of our writing and we received the feedback that the technical analysis details were harder to place in context without direct link to species tested. We decided to place this information here for easier understanding of our technical workflow and hope the reviewer understand our consideration here.

Line 163: What constitutes "reliable"?
We thank the reviewer for raising this point. After unsupervised behavioral clustering (VAME), we were able to visually assess the resulting behavioral classes and identify these key mosquito postural signatures. We have modified this sentence to improve clarity:

> "This approach resulted in the identification of 6–10 key postures related to walking, flying, and probing, confirmed with visual inspection."

Line 201: Elaborate on times, perhaps?
We have added these details to our experimental methods section:

> "Experiments for diurnal species were conducted 3–6 hours after lights turned on, whereas experiments for nocturnal species were conducted 1–4 hours after lights turned off, corresponding to the range of their peak activity times."

Lines 244-246: This new sentence does not flow well at this position. It would work better if it was moved to be right after the first sentence of this paragraph.
We appreciate this suggestion and have now moved this information earlier in the paragraph as suggested:

> "We next allowed females to blood feed to repletion, confirmed that females consumed meals that approximately doubled their body weights in all species tested (Extended Data Fig. 4), and then assessed their behavioral responses to host cues 24, 48, and 72 hours after a blood meal."

Figure 3: While this reviewer appreciates the increased information provided by the box plots in this figure, I am surprised at the marked change in error bars from the original graph. Were stats rerun to account for these changes? Additionally, the error bars on G & H extend to -300: what does -300 suppression mean? Please explain or modify graphs.
We appreciate the reviewer's comment. We updated our data representation according to journal requirements to show data distribution with box-plots. In the previous manuscript, this was represented as mean and SEM. In the current format, whiskers were used to show 5–95 percentile, which covers a larger range given the nature of some variability in individual behaviors. Additionally, we did not perform any outlier identification for data exclusion at this stage since this level of behavioral variability is expected and relevant in mosquito behavioral studies. The underlying dataset and statistical analyses were not rerun and remain unchanged. All raw data are also included in supplemental files. Similarly, the whiskers in g & h include individuals that shows variable behavioral signatures. In the case of whiskers extending to -300% suppression, this reflects that individuals at the 5 or 95 percentiles may show about 3 times more activity than a non-blood-fed female, while most of the females in the interquartile range (25–75 percentile) show less activity compared to a non-blood-fed female.

This is minor, but why do some paragraphs have a space after them and some do not? Line 333 vs 345?
Thank you for catching this, we have now corrected this.

Lines 305-316: The connection between Culex host preference and their increased walking and probing activity is unclear. This should be elaborated upon, rephrased, and/or decoupled. What is it about being ornithophilic that would require an increase in walking/probing and not flying?
Thank you for pointing this out, this was unclear as originally written and we have removed this sentence to avoid overinterpretation and make it clear that, while the strategies are different, we cannot make any claims about how this relates to host preference.

361: Add a comma after Aedes.
We have now corrected this.

Line 496: Add a reference to the figure containing this data.
We have now corrected this to reference Figure 2f.

Line 505: Were all the species fed on sheep blood? Since this is not the preferred host for Culex, that could have been causing some behavioral differences (i.e. searching for other blood meal?). Were there any reductions in fecundity for any of the strains?
We thank the reviewer for raising this point, our choice of sheep blood here is a balance between sufficient nutrients and accessibility. The sheep blood source is commercially available to us with reliable re-stocking so we can ensure freshness of blood meals and quality consistency. We also note that sheep blood generally has a higher total protein content than avian blood and should therefore provide sufficient protein, as all species took meals that approximately doubled their body weight. We did not observe any reductions in fecundity during regular passaging.

Reviewer #3 (Remarks to the Author):

Overall, the authors have done an excellent job integrating the comments from all reviewers. In particular the edits to Figures and their legends has greatly improved the clarity of the statistical approach. Most of my concerns have been addressed by their edits.

I have a few outstanding queries which they may need to address:

1. Line 203: suggest to replace "enhanced" with "increased"
We updated this wording as suggested.

2. Line 499: I cannot access the Zenodo Files (set to restricted access), but trust the authors have provided the line by line for each mosquito used in the analysis here.
We thank the reviewer for this comment. We note that access to the Zenodo files was previously restricted. However, the dataset has now been made publicly available.

3. Line 514: It would extremely useful to include very brief description of the statistical analyses here. It do agree that

the figure legends have the information, but it would be helpful for readers to have a short paragraph explaining the key comparisons, which tests were used for normality, and which test has been applied for each. For example, "We tested all data for normality using a X test. To compare the proportion of time active in the pre and post- stimulus periods (Fig 1B) we used a Wilcoxon rank-sum test". It is repeating the information but it helps to have this laid out for the reader in the appropriate section.
Apologies in advance, if I am being dense, but I am unable to locate the Supplemental Table 1 in the reviewer files. If this is referring to the checklist this is not the full details of the statistical methods. If there is a supplemental Table that provides all of this information then that would also be fine, but I'd prefer a short paragraph that gives these key details in one place in the main text.
We thank the reviewer for their detailed input. We did include all test details in Supplemental Table 1 in the following format (using figure 2D as an example). We apologize if this file was made unavailable to the reviewer.

Aedes aegypti						
Flight						
Kruskal-Wallis test						
P value	<0.0001					
Exact or a	Approximate					
P value su	****					
Do the me	Yes					
Number of	9					
Kruskal-W	39.19					
Data summary						
Number of	9					
Number of	387					
Number of	1					
Number of	8					
Alpha	0.05					
Dunn's mul	Mean rank	Significant?	Summary	Adjusted P	A-?	
-1–0 vs. 0	-99.63	Yes	****	<0.0001	B	0–1
-1–0 vs. 1	-28.19	No	ns	>0.9999	C	1–2
-1–0 vs. 2	-69.38	Yes	**	0.0077	D	2–3
-1–0 vs. 3	-91.99	Yes	****	<0.0001	E	3–4
-1–0 vs. 4	-85.56	Yes	***	0.0004	F	4–5
-1–0 vs. 5	-44.45	No	ns	0.2749	G	5–6
-1–0 vs. 6	-42.24	No	ns	0.3549	H	6–7
-1–0 vs. 7	-43.08	No	ns	0.3226	I	7–8
Test details	Mean rank	Mean rank	Mean rank	n1	n2	Z
-1–0 vs. 0	137.9	237.6	-99.63	43	43	4.742
-1–0 vs. 1	137.9	166.1	-28.19	43	43	1.342
-1–0 vs. 2	137.9	207.3	-69.38	43	43	3.302
-1–0 vs. 3	137.9	229.9	-91.99	43	43	4.378
-1–0 vs. 4	137.9	223.5	-85.56	43	43	4.072
-1–0 vs. 5	137.9	182.4	-44.45	43	43	2.116
-1–0 vs. 6	137.9	180.2	-42.24	43	43	2.011
-1–0 vs. 7	137.9	181	-43.08	43	43	2.05

Probe						
Kruskal-Wallis test						
P value	0.0006					
Exact or a	Approximate					
P value su	***					
Do the me	Yes					
Number of	9					
Kruskal-W	27.58					
Data summary						
Number of	9					
Number of	387					
Number of	1					
Number of	8					
Alpha	0.05					
Dunn's mul	Mean rank	Significant?	Summary	Adjusted P	J-?	
-1–0 vs. 0	-82.37	Yes	**	0.003	K	0–1
-1–0 vs. 1	-65.24	Yes	*	0.0387	L	1–2
-1–0 vs. 2	-69.73	Yes	*	0.0208	M	2–3
-1–0 vs. 3	-111.1	Yes	****	<0.0001	N	3–4
-1–0 vs. 4	-80.7	Yes	**	0.0039	O	4–5
-1–0 vs. 5	-49.84	No	ns	0.2509	P	5–6
-1–0 vs. 6	-58.42	No	ns	0.093	Q	6–7
-1–0 vs. 7	-49.63	No	ns	0.2566	R	7–8
Test details	Mean rank	Mean rank	Mean rank	n1	n2	Z
-1–0 vs. 0	131	213.4	-82.37	43	43	3.558
-1–0 vs. 1	131	196.2	-65.24	43	43	2.818
-1–0 vs. 2	131	200.7	-69.73	43	43	3.012
-1–0 vs. 3	131	242.1	-111.1	43	43	4.797
-1–0 vs. 4	131	211.7	-80.7	43	43	3.485
-1–0 vs. 5	131	180.8	-49.84	43	43	2.153
-1–0 vs. 6	131	189.4	-58.42	43	43	2.523
-1–0 vs. 7	131	180.6	-49.63	43	43	2.143

Walk						
Kruskal-Wallis test						
P value	0.0002					
Exact or a	Approximate					
P value su	***					
Do the me	Yes					
Number of	9					
Kruskal-W	29.96					
Data summary						
Number of	9					
Number of	387					
Number of	1					
Number of	8					
Alpha	0.05					
Dunn's mul	Mean rank	Significant?	Summary	Adjusted P	S-?	
-1–0 vs. 0	-85.22	Yes	**	0.0012	T	0–1
-1–0 vs. 1	-36.05	No	ns	0.8702	U	1–2
-1–0 vs. 2	-69.07	Yes	*	0.017	V	2–3
-1–0 vs. 3	-94.83	Yes	***	0.0002	W	3–4
-1–0 vs. 4	-79.08	Yes	**	0.0035	X	4–5
-1–0 vs. 5	-32.01	No	ns	>0.9999	Y	5–6
-1–0 vs. 6	-43.87	No	ns	0.4076	Z	6–7
-1–0 vs. 7	-38.13	No	ns	0.7186	AA	7–8
Test details	Mean rank	Mean rank	Mean rank	n1	n2	Z
-1–0 vs. 0	140.9	226.1	-85.22	43	43	3.791
-1–0 vs. 1	140.9	176.9	-36.05	43	43	1.604
-1–0 vs. 2	140.9	209.9	-69.07	43	43	3.073
-1–0 vs. 3	140.9	235.7	-94.83	43	43	4.219
-1–0 vs. 4	140.9	219.9	-79.08	43	43	3.518
-1–0 vs. 5	140.9	172.9	-32.01	43	43	1.424
-1–0 vs. 6	140.9	184.7	-43.87	43	43	1.952
-1–0 vs. 7	140.9	179	-38.13	43	43	1.696

Regardless, we think the reviewer's suggestion here increases transparency and helps the readers to understand our overall statistical choice. We have therefore added this in the revised manuscript:

> "Data were assessed for normality using the Shapiro-Wilk test. Since the datasets did not follow a normal distribution, non-parametric tests were used throughout the study. For comparisons between two groups, we applied the Wilcoxon signed-rank test for paired samples (e.g., percentage of time active pre- vs. post-stimulus in Fig 2b) and the Mann-Whitney U test for unpaired samples (e.g., total movement and behavioral indices non-fed vs. post meal in Fig 3a-f). For experiments involving multiple groups, such as percent of time spent on each behavior (Fig 2d), the Kruskal-Wallis test followed by Dunn's multiple comparison test was used. Statistical analyses were conducted using GraphPad Prism, R, or custom Python code. All statistical details, including sample sizes and exact p-values, can be found in Supplementary Table 1."

Reviewer #4 (Remarks to the Author):

The authors have addressed all my concerns in their revision. I have no further issues and comments.
We thank the reviewer for their comments and are happy to hear that we addressed all concerns.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.
We thank the reviewer for their comments and are happy to hear that we addressed all concerns.