

# Larval breeding sites of the mosquito *Aedes aegypti* in forest and domestic habitats in Africa and the potential association with oviposition evolution

Siyang Xia<sup>1</sup>, Hany Dweck<sup>1</sup>, Joel Lutomia<sup>2</sup>, Rosemary Sang<sup>2</sup>, Carolyn McBride<sup>3</sup>, Noah Rose<sup>3</sup>, Diego Ayala<sup>4</sup>, and Jeffrey Powell<sup>1</sup>

<sup>1</sup>Yale University

<sup>2</sup>Kenya Medical Research Institute

<sup>3</sup>Princeton University

<sup>4</sup>Institut de recherche pour le developpement France-Sud

March 31, 2021

## Abstract

Adaptations to anthropogenic domestic habitats contribute to the success of mosquito *Aedes aegypti* as a major global vector of several arboviral diseases. The species inhabited African forests before expanding into domestic habitats and spreading to the rest of the world. Despite a well-studied evolutionary history, how this species initially moved into human settlements in Africa remains unclear. During this initial habitat transition, *Ae. aegypti* switched from using natural containers like tree holes as larval breeding sites to using artificial containers like clay pots. Little is known about how these natural versus artificial containers differ in their environments, or whether *Ae. aegypti* in forest versus domestic habitats evolved any corresponding incipient behavioral divergence, such as in oviposition. To address these gaps, we first characterized physical characteristics, larval density, microbial density, bacterial composition, and volatile profiles of natural versus artificial containers used as mosquito larval breeding sites. We focused on two localities in Africa, La Lopé, Gabon and Rabai, Kenya. In both localities, our data showed that the two habitat-specific container types had significantly different characteristics. We then examined whether such containers differed in their attractiveness for oviposition, a key behavior affecting larval distribution. Forest *Ae. aegypti* readily accepted artificial containers in our field experiments, and laboratory choice experiments did not find distinct oviposition preference between forest and village *Ae. aegypti* colonies. These results suggested that African *Ae. aegypti* were likely generalists in their oviposition site choice. This flexibility to accept different containers might play a vital role during the initial domestication of *Ae. aegypti*, allowing the mosquitoes to use human-stored water as fallback breeding sites during dry seasons. Although ovipositional changes were not present initially, after longer domestic habitat breeding, the mosquitoes did evolve divergence oviposition preference, as suggested by previous comparisons of African *Ae. aegypti* and human-specialized non-African *Ae. aegypti*.

## Introduction

Species respond differently to anthropogenic habitats, such as villages and urban areas (McKinney, 2008; Otto, 2018; Szulkin et al., 2020). While generally these ecological changes have negative consequences, a handful of species have successfully exploited these novel human environments (Johnson & Munshi-South, 2017). Among them, one of the most notorious examples is the yellow fever mosquito, *Aedes aegypti*, a major vector of several arboviral diseases that cause millions of infections each year, including yellow fever, dengue, chikungunya, and Zika (Aubry et al., 2018; World Health Organization, 2014). The mosquitoes' high efficacy at transmitting diseases stems partly from their adaptation to human-made domestic habitats, assuring close contact with humans (Carvalho & Moreira, 2017; Fontenille & Powell, 2020).

The recent evolutionary history of *Ae. aegypti* is strongly linked to human activities (Powell et al., 2018; Powell & Tabachnick, 2013). *Ae. aegypti* originated on southwestern Indian Ocean islands and moved to continental Africa around 85,000 years ago before spreading across the continent in tropical forests (Soghian et al., 2020). Probably five to ten thousand years ago, this species invaded human settlements and likely evolved domestic adaptations (Crawford et al., 2017; Kotsakiozi et al., 2018). This domestication prepared them for later spread to the rest of the world a few hundred years ago, likely from West Africa and facilitated by human movements (Brown et al., 2014; Gloria-Soria et al., 2016; Powell & Tabachnick, 2013). Extant mosquitoes in and out of Africa show a clear genetic distinction (Gloria-Soria et al., 2016, but see exceptions in Kotsakiozi et al., 2018 and Rose et al., 2020), which roughly matches the two classical subspecies: *Ae. aegypti formosus* (*Aaf*) and *Ae. aegypti aegypti* (*Aaa*), respectively. Complexities exist in this subspecies definition (Powell & Tabachnick, 2013), but in this paper, we refer to them simply based on their geographic range (in or out of Africa). Non-African *Aaa* is mostly human specialized and lives only in domestic settings such as urban area (McBride, 2016; Powell & Tabachnick, 2013), except for a few forest-living populations in the Caribbean and Argentina (Chadee et al., 1998; Mangudo et al., 2015). On the other hand, African *Aaf* inhabits both forest and domestic habitats (Kotsakiozi et al., 2018; Paupy et al., 2014; Sylla et al., 2009), with the latter likely representing an intermediate step towards true human specialization outside Africa (i.e., *Aaa*).

Despite a well-characterized evolutionary path of *Ae. aegypti* deduced from genetic data, how this species initially moved into domestic habitats in Africa is not fully understood. In addition, most mosquito species (over 3,500 species) have not colonized domestic habitats, including many African mosquitoes that share the same forest habitats as *Ae. aegypti* (Clements, 1999). This raises the question of why *Ae. aegypti*, among only a few mosquito species, was able to invade domestic habitats (Carvalho & Moreira, 2017; Fontenille & Powell, 2020). Addressing these questions could help us further understand the unique evolutionary history of this epidemiologically important species, provide insights into mosquito control, and possibly predict other emerging disease vectors.

For *Ae. aegypti*, the transition from the ancestral forest habitat to human settlement involved two major behavioral changes: a preference for humans as a blood source (McBride, 2016) and using human-made containers for larval breeding (Day, 2016). Ancestral forest-living *Aaf* in Africa is a feeding generalist and bites wildlife for blood, while *Aaa* out of Africa specializes in biting humans (Powell et al., 2018). Recent studies have demonstrated the variations of blood-feeding preference across Africa in different habitats and between the two subspecies (McBride, 2016; McBride et al., 2014; Rose et al., 2020). They also identified dry season intensity and human population density as two main ecological drivers for the evolution of feeding preference for humans in Africa.

In comparison, larval breeding sites are relatively understudied, especially in Africa. *Aedes aegypti* lay eggs at the edge of small water containers, i.e., oviposition sites that becomes larval breeding sites (Christophers, 1960). Non-African *Aaa* uses various artificial containers, consistent with being a human specialist (Day, 2016; Swan et al., 2018; Vezzani, 2007; Yee, 2008). In Africa, *Aaf* in the forest and domestic habitats utilize different larval breeding sites: the former uses natural containers like rainwater-filled tree holes and rock pools (Lounibos, 1981), while the latter resembling *Aaa*, relies mostly on artificial containers, such as plastic buckets, tires, and clay pots (Leahy et al., 1978; McBride et al., 2014; Petersen, 1977). Some artificial containers hold human-stored water year-round and could provide valuable or even the only available larval habitats during the dry season when natural containers dry out. Therefore, it is hypothesized that seeking human water storage for oviposition during the dry season likely drove *Ae. aegypti* to enter domestic habitats, leading to the evolution of feeding preference for humans (Petersen, 1977; Powell et al., 2018; Rose et al., 2020). Despite the presumed key role of larval breeding habitats in the domestic adaptation of *Ae. aegypti*, few studies have characterized natural versus artificial larval breeding sites in Africa (Dickson et al., 2017).

If a substantial difference exists between natural and artificial containers, it could pose challenges to both female oviposition and larvae development. Oviposition preference and larvae performance are likely associated (Wong et al., 2012) but not always aligned (Albeny-Simoes et al., 2014; Refsnider & Janzen, 2010).

Given the large variety of larval habitats, it is possible that ancestral *Ae. aegypti* were generalist egg-layers, and the larvae can tolerate a wide range of conditions, which allows them to take advantage of artificial containers in harsh environments (Powell et al., 2018; Rose et al., 2020). Prolonged breeding in distinct containers could then lead to ecological divergence (Gimonneau et al., 2010; Schluter, 2000; Shafer & Wolf, 2013), resulting in adaptations and specializations to each container type. Such behavioral differentiation, in turn, could facilitate population segregation (Ayala et al., 2011; Nosil et al., 2009; Servedio et al., 2011). Indeed, a few studies have implied that oviposition divergence may have emerged between *Aaa* and *Aaf* (Leahy et al., 1978; Petersen, 1977). Larvae of the two subspecies also showed higher fitness in containers representing their respective preferred habitats (Saul et al., 1980). However, previous work only compared the two subspecies, i.e., two ends of the domestication history of *Ae. aegypti*. Whether divergence in oviposition or larval performance already exist within African *Aaf* living in forest versus domestic habitats remains largely unknown. Examining this potential incipient divergence could provide valuable insights on when, where, and how *Ae. aegypti* adapted to domestic habitats.

In this study, we characterized the environmental conditions of *Ae. aegypti* larval breeding sites in forest and domestic habitats and examined whether oviposition divergence has evolved. We focused on two locations in Africa, La Lopé in Gabon and Rabai in Kenya. Mosquitoes in both locations are *Aaf* based on their morphology and broad genetic pattern (Kotsakiozi et al., 2018; Xia et al., 2020), but can be found in forest and village environments in close proximity. In each location, forest and village populations showed little genetic differentiation (Xia et al., 2020), suggesting local habitat expansion instead of external introduction. Therefore, *Ae. aegypti* in these two locations possibly exemplify the initial colonization step of domestic habitats. We first compare the physical characteristics, competition and predation, bacterial profiles, and chemical volatiles of natural and artificial containers used as larval breeding sites between habitats (forest and village). Many of these environmental variables have been shown to affect *Ae. aegypti* oviposition (Afify & Galizia, 2015; Harrington et al., 2008; Ponnusamy et al., 2008; Reiskind & Zarrabi, 2012; Zahiri & Rau, 1998). Therefore, we also investigated the oviposition choices of forest and domestic *Aaf* towards some variables that showed the greatest differences between natural and artificial containers. This allowed us to examine whether the mosquitoes in different habitats remained oviposition generalist or have developed behavioral specialization.

## Materials and methods

### Field study

We conducted field studies in La Lopé, Gabon in Central Africa from November to December 2016, and in Rabai, Kenya in East Africa from April to May 2017. The period of field study overlapped with the rainy season in each locality during which mosquitoes had a large population size. La Lopé has an extensive continuous tropical rainforest surrounding La Lopé village (Figure 1a). The forest in Rabai, on the other hand, is more fragmented, with several villages scattered around the forest patch (Figure 1b). In each location, we searched for water-holding containers as potential mosquito larval breeding sites in both the forests and nearby villages. A potential larval breeding site was defined as a container holding at least one mosquito larvae (not necessarily *Ae. aegypti*) at the time of sampling, which suggested that the site had been present long enough for a mosquito to lay eggs. We categorized larval breeding sites into three habitat groups: forest, peridomestic (outdoor containers in a village area), and domestic (indoor containers) (Table 1), according to their locations. We separated indoor and outdoor containers because classical studies from the 1970s reported that, at least in Rabai, Kenya, *Ae. aegypti* living indoor and outdoor showed distinct behavioral and genetic differences (Leahy et al., 1978; McBride et al., 2014; Petersen, 1977; Tabachnick et al., 1979; Trpis & Hausermann, 1975). Genetic analysis showed that these indoor mosquitoes in Rabai were likely descendent of non-African *Aaa* (Brown et al., 2011; Gloria-Soria et al., 2016). However, this previously described *Aaa*-like indoor form was no longer found during our field sampling (Rose et al., 2020; Xia et al., 2020).

In La Lope, we visited 60 larval breeding sites in seven forest locations and 38 sites in six village locations (Figure 1a, one village location, Kazamabika village, is further away from the other village locations). The

sampling locations separate by 5-17 km. Forest larval breeding sites were predominantly rock pools (n=49) around streams and tree holes accumulating rainwater (n=11). Previous studies have considered tree holes and rock pools as distinct mosquito larval habitat groups (Soghigian et al., 2017). However, we only found a few tree holes with complete data (n < 6 in all analyses), and comparing between tree holes and rock pools is beyond the scope of this study. Therefore, we grouped them as ‘natural containers.’ In the village, mosquito larvae were found in a variety of artificial containers, including construction bricks, tires, metal cans, and plastic containers. Because residents in the village rarely store water indoors, all village larval breeding sites were ‘peridomestic.’ In Rabai, Kenya, we sampled 31 larval breeding sites consisting of mainly plastic buckets, earthenware pots, and metal barrels in four villages. They were mostly indoor (i.e., ‘domestic’) containers. The 37 larval breeding sites in the Rabai forest were all tree holes holding rainwater (Figure 1b). We recorded the GPS coordinates of each sampling location (consist of multiple larval breeding sites) in La Lope, and of each larval breeding site in Rabai, Kenya (Figure 1).

Upon identifying a potential larval breeding site in any habitat, we measured 11-16 physical variables and collected water samples to analyze bacterial and volatile profiles. Sample sizes for each category of environmental variables were summarized in Table 1. Method details are described in the following sections and the Appendix. We also collected all mosquito larvae and pupae using pipets and reared them to adults in field stations, keeping collections from different larval breeding sites separate. Upon eclosion, adults were identified to genus and species based on taxonomic keys using a dissection microscope in the field (Rueda, 2004). We kept *Ae. aegypti* adults alive to establish lab colonies for later behavioral experiments.

We categorized each larval breeding site (i.e., container) as ‘*Ae. aegypti* present’ or ‘*Ae. aegypti* absent’ based on whether it held any *Ae. aegypti* larvae or pupae (Table 1). It is worth noting that the absence of *Ae. aegypti* did not necessarily suggest an avoidance. Some sites may be suitable for oviposition and larval development but not yet colonized by *Ae. aegypti*, and we also could not observe unhatched eggs. Bearing this potential caveat, we combined the three habitat categories and the two *Ae. aegypti* presence status to generate six ‘larval breeding site groups’ (Table 1). We focused on three comparisons for the analysis of environmental conditions: 1) across larval breeding site groups, 2) across habitat categories regardless of *Ae. aegypti* presence status, and 3) between *Ae. aegypti* present and absent sites regardless of habitats. In Rabai, almost all peridomestic and domestic larval breeding sites sampled were present with *Ae. aegypti*. The only peridomestic *Ae. aegypti* absent site was excluded from analyses comparing between larval breeding site groups, but retained in comparisons between habitats or between *Ae. aegypti* present vs. absent sites.

The fieldwork in La Lope was approved by the CENAREST with the authorization AR0013/16/MESRS/CENAREST/CG/CST/CSAR, and by the La Lope National Parks with the authorization AE16008/PR/ANPN/SE/CS/AEPN. The fieldwork in Rabai was approved by the Kenya Medical Research Institute Scientific and Ethical Review Unit with the authorization KEMRI/SERU/3433.

### *Characterizing larval breeding sites: physical variables*

We measured 11 physical variables for each larval breeding site in La Lope, Gabon, and five additional variables in Rabai, Kenya (Table S1 in Appendix). These variables describe the size of the larval breeding site (e.g., diameters, circumference, surface area, volume, container depth, water depth, etc.), ambient environments (temperature, relative humidity, and canopy coverage), and water characteristics (pH, conductivity, salinity, water temperature, and total dissolved solids). Methodological details were summarized in Table S1 in the Appendix.

After removing eight sites with excessive missing data, we performed a principal component analysis (PCA) to summarize all physical variables. The multivariate differences between larval breeding site groups, habitats, and *Ae. aegypti* presence status were tested by multiple response permutation procedure (MRPP) with 999 permutations. The p-values for multiple comparisons were adjusted using the Holm method. We also identified variables that are most differentiated in each comparison by ranking variable importance using a random forest algorithm in R package *randomForest* v4.6-14 (Liaw & Wiener, 2002). Random forest is a decision-tree based classification algorithm that works well with small sample size and correlated variables

(Qi, 2012).

#### *Characterizing larval breeding sites: competition and predation*

Competition and predation could influence larval development and female oviposition choice (Pamplona Lde et al., 2009; Soman & Reuben, 1970; Vonesh & Blaustein, 2010; Zahiri & Rau, 1998). To consider their effects, we measured the density of *Ae. aegypti* as well as all mosquito species. We also noted the presence of predatory larvae, predominately *Toxorhynchites* mosquitoes, and removed them immediately if found.

We first compared the density of *Ae. aegypti* (calculated by dividing mosquito numbers by water volume) between habitats, using only *Ae. aegypti*- present larval breeding sites. We then examined the mosquito density of all species and compared that between habitats, *Ae. aegypti* present vs. absent sites, and across larval breeding site groups. In La Lope, records of other mosquito species were only available for forest larval breeding sites, so we only compared sites present versus absent of *Ae. aegypti* within the forest. Density comparisons were performed using the Kruskal–Wallis test and *post hoc* pairwise Wilcoxon rank-sum test. Lastly, we analyzed the frequency of finding predators in different larval breeding site groups or habitats with chi-squared tests.

#### *Characterizing larval breeding sites: microbial density*

We examined the microbial profile in a subset of larval breeding sites, as previous studies have shown that the microbiome, particularly bacteria, affect *Ae. aegypti* oviposition choices (Arbaoui & Chua, 2014; Ponnusamy et al., 2015) as well as larval development (Correa et al., 2018; Dickson et al., 2017; Souza et al., 2019). In brief, we collected 15 mL (in La Lope) or 50 mL (in Rabai) water samples from each larval breeding site. To first examine the total microbial density, we fixed an aliquot of each water sample with formaldehyde and stained it with DAPI (4',6-diamidino-2-phenylindole). The number of microbial cells was counted using hemocytometers under a fluorescence microscope and back-calculated to its original density.

Densities were log-transformed before statistical analysis. We compared the microbial density among larval breeding site groups in La Lope with the Kruskal–Wallis test and *post hoc* pairwise Wilcoxon rank-sum tests. The distribution of data in Rabai samples did not violate parametric test assumptions, so we performed the comparisons using analysis of variance (ANOVA) and *post hoc* Tukey tests.

#### *Characterizing larval breeding sites: bacterial community composition*

In addition to microbiome density, we performed 16s-rRNA gene amplicon sequencing to explore the bacterial community composition in most larval breeding sites (Table 1), inspired by previous studies that suggested different bacteria between habitats (Dickson et al., 2017). The sample processing and sequencing library preparation are described in the Appendix. In short, we collected cells from the water by centrifuge or filtering, extracted DNA, and amplified the 16s-rRNA gene V4 region using primers reported in Kozich et al. (2013). The amplicons from multiple samples were multiplexed and sequenced on Illumina MiSeq (Illumina, USA) at the Yale Center for Genome Analysis. We conducted amplicon sequencing for La Lope and Rabai samples separately.

We demultiplexed the sequencing reads using USEARCH v10.0.240 (Edgar, 2010) and followed the pipeline of DADA2 (v1.8.0) (Callahan et al., 2016) to determine the bacterial community composition. DADA2 estimates sequencing errors and infers the exact sequence variants (i.e., amplicon sequence variants, or ASVs), which are analog to the conventional operational taxonomic unit (OTU). ASVs were blasted to the Ribosomal Database Project (RDP) 16s-rRNA gene reference database (RDP trainset 16 and RDP species assignment 16) (Cole et al., 2014) for taxonomic assignment.

Using R package *phyloseq* (McMurdie & Holmes, 2013), we first calculated the alpha diversity of the bacteria community in each larval breeding site indicated by the Shannon index (Shannon, 1948), using raw read counts. We then compared the index across larval breeding site groups, habitats, and between *Ae. aegypti* present and absent sites. The community compositions were summarized by non-metric multidimensional scaling (NMDS) with the Bray-Curtis distance matrix. Similar to PCA, NMDS summarizes multivariate

data (each bacterial taxa as one variable) but is more appropriate for bacterial composition data (Ramette, 2007). Before NMDS analysis, we first removed samples with fewer than 5000 reads to avoid low-quality samples, and thinned each sample proportionally to the lowest read depth of all samples to control for uneven sequencing depth. Bacterial communities may show different assembly patterns at different taxonomic levels (Goldford et al., 2018). Therefore, we calculated the Shannon index and performed NMDS at four taxonomic levels: ASV, species, genus, and family. We also demonstrated the major bacterial groups at the family level using bar plots. Lastly, we used R package *DESeq2* to identify bacterial families that are most differentiated between habitats (Love et al., 2014).

To estimate the temporal stability of bacterial communities, we collected water samples more than once for five larval breeding sites in each habitat. The average interval between two consecutive collections ranges from 3 to 21 days, with an average of 8.4 days in La Lope and 17 days in Rabai. All temporal samples were sequenced, but only the first-day samples were included in the analyses described above. We performed a separate NMDS analysis to examine variation between temporal samples.

#### *Characterizing larval breeding sites: chemical volatiles in Rabai, Kenya*

Chemical volatiles released from a larval breeding site could act as olfactory cues for mosquitoes during oviposition (Afify & Galizia, 2015), yet the volatile profiles of wild larval breeding sites have rarely been examined. We attempted to describe the volatile profile in larval breeding sites in Rabai, Kenya. In brief, we collected water samples from a subset of larval breeding sites (Table 1) and extracted the volatiles into an absorbent with a steady airflow. The captured volatiles were examined by Gas Chromatography-Mass Spectrometry (GC-MS) at Yale West Campus Analytical Core. We then identified and quantified each compound by mapping the GC-MS peaks to a reference library. The technical details were described in the Appendix. Due to the sparsity of compounds in the final dataset, we did not perform statistical analysis but instead summarized the compound concentrations using a heatmap.

#### *Characterizing larval breeding sites: combining multiple environmental conditions*

To consolidate the analysis of different categories of environmental conditions, we performed another random forest model in La Lope and Rabai, respectively. The model included scores on the first three principal components (PCs) from the physical variable analysis and scores on the first two NMDS axes from the bacterial community composition analysis (i.e., predictive variables), and used them to classify the larval breeding site groups (i.e., the dependent variable). For the analysis from Rabai, we also added the microbial density and the density of all mosquito larvae. These two variables had many missing values in the La Lope dataset and thus were excluded. The model generated a confusion matrix, which displayed the number of samples correctly or wrongly assigned to each larval breeding site group. A lower proportion of misclassification between groups suggests a stronger distinction in their environmental conditions.

#### *Field oviposition choice experiments*

In order to examine if *Ae. aegypti* in the forest readily accept artificial containers as larval breeding sites, we performed a preliminary field experiment by placing artificial containers into the forest area. The containers used in La Lope included tires, plastic bottles, plastic bags, bricks, and metal cans, while in Rabai, Kenya, we used plastic buckets and earthenware pots. These containers were frequently found in the villages. In addition, we added bamboo segments, as African researchers have used them to collect forest mosquitoes (Kemp & Jupp, 1991), and bamboo plantations were sometimes found in villages. All containers were left in the forest area for 7-14 days and filled by either rainwater (in La Lope) or well water from the village (in Rabai). After retrieving these containers, we examined the existence of *Ae. aegypti* as described above. As a control, we also placed a similar set of experimental containers in the village (outdoor in La Lope and indoor in Rabai). For a subset of the experimental containers, we applied the 16S-rRNA gene amplicon sequencing to examine their bacterial community compositions.

#### *Laboratory oviposition assays*

To examine the interaction between larval breeding sites' conditions and oviposition evolution, we performed

laboratory oviposition assays in a common-garden setup. The goal was to examine whether forest and village *Ae. aegypti* have different oviposition preferences towards a subset of environmental variables that differed between forest and village larval breeding sites.

We established a forest colony and a peridomestic colony from La Lope using *Ae. aegypti* collected from larval breeding sites, supplemented with oviposition traps and human landing capture (approved by the National Research Ethics Committee of Gabon under the protocol 0031/2014/SG/CNE). In Rabai, we created six independent village colonies (four domestic colonies and two peridomestic colonies) from the four villages and four forest colonies from the Kaya Bomu forest. We blood-fed the mosquitoes in the field and brought the eggs (i.e., the second generation) back to our lab at Yale University and the McBride lab at Princeton University. The detailed information of the mosquito colonies and protocols for maintaining these colonies are in the Appendix. The Rabai forest colonies correspond to KBO1 and KBO2 in Rose et al. (2020). All laboratory oviposition assays were performed at Yale University. We used the fourth to the sixth generation of mosquitoes in these assays. For simplicity, we refer to the peridomestic and domestic colonies as village colonies.

$$OAI = \frac{N_1 - N_2}{N_1 + N_2},$$

where  $N_1$  and  $N_2$  are the number of eggs deposited in the two cups, respectively.  $OAI$  ranges from -1 to 1, representing a complete preference for the second choice to a complete preference for the first choice. We performed beta-binomial models in the R package *glmmTMB* (Brooks et al., 2017) to examine whether colonies differ in their oviposition preference, using the two egg counts in each cage as the dependent variable (Rose et al., 2020). We added the batch/trial IDs as random effects if data testing a specific condition were generated from more than one experimental batch. The statistical significance of colony or habitat effects were determined by comparing the full model with a null model that excludes colony or habitat (Table S9). We extracted mean  $OAI$  with a 95% confidence interval from the model using the R package *emmeans* (Lenth et al., 2018; Rose et al., 2020).

Using this assay, we compared the oviposition preference of forest versus village colonies from La Lopé and Rabai towards several environmental variables. We first focused on a pair of Rabai forest versus domestic colonies. The conditions tested include: 1) water samples collected from tree holes and artificial containers in Rabai, 2) pH, 3) shading, 4) larval density, 5) a combined effect of pH, conductivity, and shading, and 6) bacterial community composition. We selected these conditions as they showed significant differences between forest and village larval breeding sites in the field. In each assay, the two choices (i.e., two cups) roughly represent the median value of the focal variable measured in forest and village larval breeding sites (described in more detail in the Appendix). For example, in the experiment on water pH, we adjusted the pH in the two cups to the median pH values of all Rabai forest versus village larval breeding sites. Lastly, for the experiment of bacterial community composition, we expanded it to include all Rabai colonies as well as the two La Lopé colonies (Table S2 in the Appendix).

In addition to the above two-choice assays, we tested the oviposition preference of all mosquito colonies to five bacterial densities. This experiment was inspired by the large variation in bacterial density among field larval breeding sites (more than two orders of magnitude) and that previous laboratory experiments with *Ae. aegypti* found density-dependent ovipositional responses to bacteria (Ponnusamy et al., 2015; Ponnusamy et al., 2010). We used a similar experimental design as the two-choice assays but provided each cage of mosquito females five cups instead of two. The cups contained bacterial cultures at densities ranging from zero to nearly the maximal bacterial density in field larval breeding sites. The bacterial culture was generated from an even mixture of forest and domestic water samples (Table S2 in the Appendix). We counted the numbers of eggs laid in the five cups and fitted a negative-binomial model using the R package *lme4* (Bates et al., 2014), with bacterial density, habitat/colony, and their interactions as predictors. If mosquitoes from different colonies or habitat types have different oviposition choices, the interaction term would be significant, which was tested by comparing the full model with a null model excluding the interactive term (Table S9). We added cage ID

as a random effect. Lastly, we used the *emmeans* package to estimate the expected number of eggs in each bacterial density with 95% confidence intervals.

## Results

### *Characterizing larval breeding sites: physical characteristics*

PCA analysis summarizing the 11 physical variables in La Lopé showed that the four larval breeding site groups (two habitats  $\times$  *Ae. aegypti* present/absent) overlap extensively in the plane described by the first two principal components, which together account for 38% of the total variance (Figure 2a). However, forest and peridomestic sites appeared to differ slightly. In support of that, MRPP tests found a significant multivariate difference between habitats when including both *Ae. aegypti* present and absent sites ( $p = 0.001$ ). Such between-habitat difference was less significant when examining only *Ae. aegypti* present sites ( $p = 0.316$ ), possibly due to the small sample size (only five samples in the forest *Ae. aegypti* present group). *Ae. aegypti* present and absent sites did not differ significantly (all sites regardless of habitats:  $p = 0.311$ , within forest sites:  $p = 1$ , within peridomestic sites:  $p = 1$ ). The axis that differentiated forest and peridomestic larval breeding sites roughly aligned with the vectors of a few environmental variables, including ambient temperature and humidity, canopy coverage, container opening height, and water pH (Figure 2a). These variables also generally had high variable important ranks in the random forest analysis (Figure S2a), reflecting their difference between habitats.

In Rabai (Figure 2b, PCA summarizing 16 physical variables), similarly, forest and peridomestic sites were modestly but significantly different in their physical characteristics (MRPP,  $p = 0.001$ ). The differentiation between forest and domestic larval breeding sites was more apparent, while peridomestic and domestic sites shared similar conditions. *Ae. aegypti* present and absent sites in the forest did not show a strong difference. Consistent with PCA, MRPP found significant multivariate differences in most comparisons ( $p = 0.001$ ), except between forest sites with *Ae. aegypti* present vs. absent ( $p = 0.157$ ) and between domestic vs. peridomestic sites ( $p = 0.192$  for all sites and  $p = 0.147$  for *Ae. aegypti* present sites only). Forest and domestic breeding sites were separated primarily on the first PC, explained by container size (e.g., diameter, circumference, etc.), water volume, and water pH (Figure 2b). Domestic sites tended to be larger, contained more water and had higher pH (i.e., more alkaline). These variables were also identified as most influential in distinguishing habitats in the random forest analysis. In addition, canopy coverage, a measure of shading, also showed a high rank in variable importance in comparisons between habitats (Figure S2b). This is not surprising as domestic containers were under roof and peridomestic containers were exposed outdoor, while forest tree holes were partially shaded by the canopy.

### *Characterizing larval breeding sites: competition and predation*

The density of *Ae. aegypti* was similar in forest and peridomestic larval breeding sites in La Lopé (Figure 2c,  $W = 33.5$ ,  $p = 0.961$ ). Other mosquito species were recorded only in the La Lopé forest, with a similar density between *Ae. aegypti* present and absent site (Figure S3a,  $W = 121.5$ ,  $p = 0.723$ ). The frequency of finding predatory *Toxorhynchites* larvae (Table S3) did not differ among larval breeding site groups ( $\chi^2 = 3.67$ ,  $df = 3$ ,  $p = 0.30$ ), between habitats ( $\chi^2 = 1.14$ ,  $df = 1$ ,  $p = 0.29$ ), or between *Ae. aegypti* present and absent sites ( $\chi^2 = 0.41$ ,  $df = 1$ ,  $p = 0.52$ ).

In Rabai, *Ae. aegypti* density was significantly lower in domestic larval breeding sites (Figure 2d, forest vs. domestic:  $W = 19$ ,  $p < 0.001$ ; peridomestic vs. domestic:  $W = 140$ ,  $p = 0.013$ ). The density difference between domestic and forest containers was mainly driven by the larger size of the former. In contrast, the difference between domestic and peridomestic sites was due to the higher number of mosquitoes found in the latter (negative-binomial model, rate ratio = 2.89,  $p < 0.001$ ). Peridomestic and forest sites had similar *Ae. aegypti* density ( $W = 69$ ,  $p = 0.583$ ). When including other mosquito species, domestic breeding sites also had lower overall mosquito density than the other two habitats (Figure S3b, Table S5). Within the forest, *Ae. aegypti* present and absent sites had similar mosquito density.

### *Characterizing larval breeding sites: microbial density*



Microbial densities did not show significant differences among larval breeding site groups, habitats, or *Ae. aegypti* present and absent sites in La Lope (Figure 2e, Table S4). In Rabai, we found significantly lower microbial density in domestic sites than forest and peridomestic sites (Figure 2f, Table S5). Microbial densities were similar in forest and peridomestic breeding sites. Lastly, *Ae. aegypti* present and absent sites within each habitat had comparable levels of microbial density.

#### *Characterizing larval breeding sites: bacterial community composition*

In La Lope, peridomestic larval breeding sites had a higher Shannon index than forest sites at the species, genus, and family level (Figure S4, Table S4). Other pairwise comparisons were not significant. In Rabai, Shannon index was comparable between all larval breeding site groups (Figure S5, Table S5). *Ae. aegypti* present sites have lower alpha diversity than *Ae. aegypti* absent sites at the bacterial family level, but only when we ignored habitats.

NMDS analysis suggested that forest and village (including peridomestic and domestic) larval breeding sites had a very different bacterial community in both La Lope (Figure 2g) and Rabai (Figure 2h) at the ASV level. The forest-village divergence was less evident at higher taxonomic levels for the La Lope larval breeding sites (Figure S6), while Rabai samples retained the difference between forest and village at all four taxonomic levels (Figure S7). Within each habitat, larval breeding sites with *Ae. aegypti* present and absent shared similar bacterial community composition in both La Lope and Rabai.

When examining the most abundant bacterial families, we observed considerable variation among samples (Figure S8). Most larval breeding sites contained multiple families with no clear dominance. In La Lope, *Microbacteriaceae*, *Flavobacteriaceae*, and *Burkholderiaceae* showed higher abundance in forest breeding sites, while *Oxalobacteraceae* and *Sphingobacteriaceae* were more abundant in peridomestic sites (Figure S8a, Table S6). In Rabai, *Moraxellaceae* had an apparent dominance in domestic sites, but its abundance was not significantly different between habitats. *DESeq2* found a higher abundance of *Enterobacteriaceae*, *Xanthomonadaceae*, *Pseudomonadaceae*, and *Planococcaceae* in forest larval breeding sites than domestic and peridomestic sites (Figure S8b, Table S7). A full list of bacterial families that showed differential abundance between habitats are in Table S6 and S7 in the Appendix.

Lastly, NMDS analysis found that temporal samples collected from the same larval breeding sites did vary in their bacterial community composition, but the temporal changes did not exceed the variations observed within each habitat (Figure S9). That is, temporal samples from the forest clustered with the rest of forest sites instead of samples from the village, and vice versa. This result suggests that the strong divergence in bacterial communities between habitat was temporally stable.

#### *Characterizing larval breeding sites: chemical volatiles in Rabai, Kenya*

The volatile profiles of 31 larval breeding sites in Rabai were summarized in Figure 3. There were substantial variations in the chemical composition of samples, both within habitats and across habitats. GC-MS analysis identified a total of 29 chemical compounds. The majority of them were shared across different habitats, but a few chemicals were unique to either forest or domestic habitat across more than one larval breeding sites (Figure 3, top five rows and bottom five rows).

#### *Characterizing larval breeding sites: combining multiple environmental conditions*

The confusion matrices from random forest analysis suggested high accuracy in distinguishing larval breeding sites between forest and village (peridomestic and domestic) in both La Lope (Table 2) and Rabai (Table 3). Rabai peridomestic and domestic samples were less accurately classified, consistent with the finding that they shared similar physical characteristics and bacterial communities. Within each habitat, the classification algorithm distinguished the *Ae. aegypti* present versus absent sites poorly.

#### *Field oviposition choice experiments*

In La Lope, experimental containers placed in the forest produced 73 *Ae. aegypti*, with 51 from the bamboo containers and 22 from artificial containers. In parallel, bamboo and artificial containers placed in the village

yield 10 and 73 *Ae. aegypti*, respectively. In Rabai forest, from the containers used in the field experiment, we found 64 *Ae. aegypti* from bamboo containers and 645 from the plastic buckets and earthenware jars. In Rabai villages, bamboo and artificial containers produced 1 and 353 *Ae. aegypti*. When examining these experimental containers' bacterial communities, NMDS analysis found that regardless of the container type and the habitats in which they were located, all experimental containers clustered with the village (peridomestic and domestic) larval breeding sites (Figure S10).

### Laboratory oviposition assays

Among all two-choice oviposition experiments, we found three significant preferences (Figure 4 and S11): Rabai Kwa Bendegwa village colony preferred forest water samples over village water samples, and preferred forest mosquito larval density over village larval density; La Lope forest colony preferred the bacterial culture started with peridomestic water samples over that started with forest water samples. However, there is substantial within-colony variation in most experiments. The beta-binomial models did not find any significant difference between colonies or between the habitat types in any assays (Table S8). We applied a negative-binomial model to analyze the oviposition assays testing multiple bacterial densities (Figure 5 and S12). Neither colonies nor the habitats of the colonies had a significant effect on the mosquito's preference for the five bacterial densities (Table S8). La Lope village colonies showed a weak preference for lower bacterial densities, but the trend was not statistically significant (ANOVA for the effect of oviposition choices:  $F = 1.56$ ,  $df = 4$ ,  $p = 0.200$ ).

## Discussion

In both La Lope and Rabai, our study found that mosquito larval breeding sites in the forest and villages (including peridomestic and domestic sites) had different physical and biological characteristics, though this between-habitat contrast varies among variables. Notably, bacterial community composition showed clear and consistent difference between habitats in both localities. Despite this environmental difference, behavioral investigations suggested that *Ae. aegypti* in the forest readily accepted artificial containers as oviposition and larval breeding sites. *Aedes aegypti* colonies derived from the forest and villages also showed similar weak oviposition preferences in the lab. These results are consistent with the hypothesis that *Ae. aegypti* are generalists in larval breeding site choice. This hypothesis was also supported by the indistinguishable conditions between *Ae. aegypti* present and absent larval breeding sites within each habitat, suggesting that the mosquitoes were likely not selective and can accept a wide range of larval habitats. Lastly, oviposition choices in the laboratory were highly heterogeneous, consistent with a lack of strong preference.

Being versatile in larval habitat allows *Ae. aegypti* to take advantage of novel artificial containers when natural breeding sites are scarce. This has been proposed as a key driver for this mosquito to move into domestic habitats in the first place (Brown et al., 2014; Powell et al., 2018; Rose et al., 2020). Consistent with this hypothesis, movement between habitats was suggested by genetic studies showing little genetic differentiation between forest and village *Ae. aegypti* populations in La Lope and Rabai (Kotsakiozi et al., 2018; Paupy et al., 2014; Rose et al., 2020; Xia et al., 2020). It should be noted that this genetic similarity in Rabai, as well as the lack of behavioral difference observed in this study and Rose et al. 2020, contrasts to studies in the 1970s and 2009 where Rabai forest and village *Ae. aegypti* showed significant genetic differences, feeding preference difference, and ovipositional difference (Brown et al., 2011; McBride et al., 2014; Petersen, 1977; Tabachnick et al., 1979). These strong differences found before 2017 resulted from the introduction of non-African *Aaa* to Rabai villages (Brown et al., 2011; Gloria-Soria et al., 2016). The exotic *Aaa* population was no longer found during our fieldwork in 2017, and the village mosquito population was likely originated from the local forest (i.e., *Aaf*).

Once the mosquito established themselves in the novel habitat (likely moving from forest to domestic habitat), ecological divergence could take place. However, this study was unable to detect evidence of consistent ovipositional divergence by the laboratory oviposition experiments. One possibility is that the habitat shifts in La Lope and Rabai happened recently and that the extensive connectivity in the local scale between habitats may hinder phenotypic divergence. Between more distant localities when gene flow is less frequent,

there may be more differences between mosquitoes from different habitats or places with different human population density, as found for host preference (Rose et al., 2020). A third possibility is that ecological divergence may happen in the immature stages, e.g., egg, larvae, and pupae (Saul et al., 1980). This study did not examine larval performance, but future investigations comparing eggs hatching and larval development in different water conditions that mimic either forest or village larval breeding sites could be insightful in this regard. For example, microbial density was significantly lower in Rabai domestic containers, which might pose selection pressure on larval starvation resistance (Barrera & Medialdea, 1996; Souza et al., 2019), leading to higher resistance in the domestic population.

In addition to these plausible ecological and evolutionary considerations, we cannot rule out the possibility that our laboratory oviposition experiments lacked the power to detect oviposition preference or differences between colonies, although the two-choice or multi-choice assays have been used widely to investigate *Ae. aegypti* oviposition preference (but see Singer (2004) for more discussion on measuring preference). Colonies may have also lost distinctive traits due to adaptation to laboratory conditions (Hoffmann & Ross, 2018). Moreover, the design of using five females per cage instead of one female might introduce some unknown complexity, for instance, interference between individuals (Allan & Kline, 1998). Lastly, the contrast of oviposition choices might not be of a magnitude detectable by female *Ae. aegypti*. However, the choices used in this study were informed by characteristics of natural oviposition sites, and therefore should be ecologically relevant for the mosquitoes. A recent study using the same *Ae. aegypti* colonies did find between-habitat ovipositional difference towards more extreme but unnatural conditions (Xia, 2021). The complexities regarding *Ae. aegypti* oviposition and experimental design warrant future studies to examine more environmental conditions or combinations of multiple variables, applying multiple preference measurements, and use younger colonies in a more natural setting (e.g. conducting choices assays in the field with mosquitoes collected from larval breeding sites).

Besides adding to our understanding of the domestication history of *Ae. aegypti*, this study also provided the first detailed physical and biological characterization of *Ae. aegypti* larval breeding sites, at least in Africa. Dickson et al. (2017) described the bacterial community composition in larval breeding sites in La Lope and found a strong difference between habitats, echoed in our study. Yee et al. (2012) found consistent differences between tree holes and tires in the U.S. although *Ae. aegypti* were not present in most containers. While our work provides useful baseline information for future studies on *Ae. aegypti* ecology and behavior, we acknowledge that some caveats still exist in our field sampling, so the results should be interpreted with caution. For example, the absence of *Ae. aegypti* in a larval breeding site did not necessarily reflect avoidance by the female nor that it is inhospitable for the larvae, especially as we could not inspect the existence of unhatched eggs. However, this does not affect our speculation that *Ae. aegypti* are not selective about larval breeding sites, as the *Ae. aegypti* present and absent sites had similar range of variation. We also only characterized larval breeding sites in a narrow temporal window during the rainy season. Future studies examining larval breeding sites throughout the year would be particularly relevant to the recent work suggesting the importance of seasonality in driving the domestication of *Ae. aegypti* (Rose et al., 2020). Furthermore, the sample sizes in our field study were relatively small. Field studies with larger sample sizes and more balanced sampling between different habitats and larval breeding site groups could further validate this study's results. We also grouped tree holes and rock pools as "natural" containers due to the limitation of sample sizes, yet previous studies have implied that they could be two distinct larval habitats (Soghigian et al., 2017). However, our preliminary analysis suggested that grouping or separating them did not affect the main findings from the field data. We also need to acknowledge that the chemical profiles of larval breeding sites in Rabai reported in this study were probably not complete and therefore calls for future studies with improved sample collection and analysis techniques. Lastly, in addition to the condition of each individual larval breeding sites, the local context could also be important, e.g., vegetation around the sites (Rey & O'Connell, 2014).

In summary, this study suggested that *Ae. aegypti* in Africa were likely generalists in their larval habitat choice, which allowed them to readily accept artificial containers as larval breeding sites and potentially facilitated their introduction into domestic habitats. Being flexible in oviposition and larval breeding site

choices could benefit *Ae. aegypti* by spreading the risk during reproduction and reduce larval competition. This is consistent with the observations that this mosquito has a bet-hedging ‘skip oviposition’ behavior (i.e., lay small batches of eggs in multiple containers) (Colton et al., 2003; Starrfelt & Kokko, 2012). However, outside of Africa, *Ae. aegypti* are closely associated with human communities and use almost exclusively artificial containers for larval breeding sites (Day, 2016; Swan et al., 2018; Vezzani, 2007; Yee, 2008), raising the interesting question of when and how this specialization on artificial containers evolved. A few recent studies suggested that human specialization may happen somewhere in West Africa, such as Sahel or Angola (Crawford et al., 2017; Powell et al., 2018; Rose et al., 2020). On the other hand, the human-specialized non-African *Aaa* could also move back to ancestral breeding sites, for instance, in the Caribbean (Chadee et al., 1998). It would be interesting to examine such processes and test whether the mosquitoes resumed generalist in larval breeding site choice during this process.

## Literature cited

- Afify, A., & Galizia, C. G. (2015). Chemosensory cues for mosquito oviposition site selection. *J. Med. Entomol.*, *52* (2), 120-130. doi:10.1093/jme/tju024
- Albeny-Simoes, D., Murrell, E. G., Elliot, S. L., Andrade, M. R., Lima, E., Juliano, S. A., & Vilela, E. F. (2014). Attracted to the enemy: *Aedes aegypti* prefers oviposition sites with predator-killed conspecifics. *Oecologia.*, *175* (2), 481-492. doi:10.1007/s00442-014-2910-1
- Allan, S. A., & Kline, D. L. (1998). Larval Rearing Water and Preexisting Eggs Influence Oviposition by *Aedes aegypti* and *Ae. albopictus* (Diptera: Culicidae). *J. Med. Entomol.*, *35* (6), 943-947. doi:10.1093/jmedent/35.6.943
- Arbaoui, A., & Chua, T. (2014). Bacteria as a source of oviposition attractant for *Aedes aegypti* mosquitoes. *Trop Biomed.*, *31* (1), 134-142.
- Aubry, F., Martynow, D., Baidaliuk, A., Merkling, S., Dickson, L., Romero-Vivas, C., . . . Paupy, C. (2018). A worldwide survey of *Aedes aegypti* susceptibility to Zika virus sheds light on the african exception to Zika emergence. *Am. J. Trop. Med. Hyg.*, *99* (4), 72-72.
- Ayala, D., Fontaine, M. C., Cohuet, A., Fontenille, D., Vitalis, R., & Simard, F. (2011). Chromosomal inversions, natural selection and adaptation in the malaria vector *Anopheles funestus*. *Mol. Biol. Evol.*, *28* (1), 745-758.
- Barrera, R., & Medialdea, V. (1996). Development time and resistance to starvation of mosquito larvae. *J. Nat. Hist.*, *30* (3), 447-458. doi:10.1080/00222939600770231
- Bates, D., Machler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., . . . Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, *9* (2), 378-400.
- Brown, J. E., Evans, B. R., Zheng, W., Obas, V., Barrera-Martinez, L., Egizi, A., . . . Powell, J. R. (2014). Human impacts have shaped historical and recent evolution in *Aedes aegypti*, the dengue and yellow fever mosquito. *Evolution*, *68* (2), 514-525. doi:10.1111/evo.12281
- Brown, J. E., McBride, C. S., Johnson, P., Ritchie, S., Paupy, C., Bossin, H., . . . Powell, J. R. (2011). Worldwide patterns of genetic differentiation imply multiple ‘domestications’ of *Aedes aegypti*, a major vector of human diseases. *Proc Biol Sci*, *278* (1717), 2446-2454. doi:10.1098/rspb.2010.2469
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods*, *13* (7), 581.

- Carvalho, F. D., & Moreira, L. A. (2017). Why is *Aedes aegypti* Linnaeus so Successful as a Species? *Neotrop. Entomol.*, *46* (3), 243-255. doi:10.1007/s13744-017-0520-4
- Chadee, D. D., Ward, R. A., & Novak, R. J. (1998). Natural habitats of *Aedes aegypti* in the Caribbean—a review. *J. Am. Mosq. Control Assoc.*, *14* (1), 5-11.
- Christophers, S. R. (1960). *Aedes aegypti* (L.) the yellow fever mosquito: its life history, bionomics and structure : Cambridge University Press, Cambridge, England.
- Clements, A. N. (1999). *The biology of mosquitoes. Volume 2: sensory reception and behaviour* : CABI publishing.
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., . . . Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.*, *42* (D1), D633-D642.
- Colton, Y., Chadee, D., & Severson, D. (2003). Natural skip oviposition of the mosquito *Aedes aegypti* indicated by codominant genetic markers. *Med. Vet. Entomol.*, *17* (2), 195-204.
- Correa, M. A., Matusovsky, B., Brackney, D. E., & Steven, B. (2018). Generation of axenic *Aedes aegypti* demonstrate live bacteria are not required for mosquito development. *Nature Communications*, *9* (1), 4464. doi:10.1038/s41467-018-07014-2
- Crawford, J. E., Alves, J. M., Palmer, W. J., Day, J. P., Sylla, M., Ramasamy, R., . . . Jiggins, F. M. (2017). Population genomics reveals that an anthropophilic population of *Aedes aegypti* mosquitoes in West Africa recently gave rise to American and Asian populations of this major disease vector. *BMC Biol.*, *15* (1), 16.
- Day, J. F. (2016). Mosquito oviposition behavior and vector control. *Insects*, *7* (4), 65.
- Dickson, L. B., Jiolle, D., Minard, G., Moltini-Conclois, I., Volant, S., Ghazlane, A., . . . Lambrechts, L. (2017). Carryover effects of larval exposure to different environmental bacteria drive adult trait variation in a mosquito vector. *Science Advances*, *3* (8), e1700585. doi:10.1126/sciadv.1700585
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26* (19), 2460-2461.
- Fontenille, D., & Powell, J. R. (2020). From Anonymous to Public Enemy: How Does a Mosquito Become a Feared Arbovirus Vector? *Pathogens*, *9* (4), 265.
- Gimonneau, G., Bouyer, J., Morand, S., Besansky, N. J., Diabate, A., & Simard, F. (2010). A behavioral mechanism underlying ecological divergence in the malaria mosquito *Anopheles gambiae*. *Behav. Ecol.*, *21* (5), 1087-1092. doi:10.1093/beheco/arq114
- Gloria-Soria, A., Ayala, D., Bheecarry, A., Calderon-Arguedas, O., Chadee, D. D., Chiappero, M., . . . Kamal, H. A. (2016). Global genetic diversity of *Aedes aegypti* . *Mol. Ecol.*, *25* (21), 5377-5395.
- Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., . . . Sanchez, A. (2018). Emergent simplicity in microbial community assembly. *Science*, *361* (6401), 469-474. doi:10.1126/science.aat1168
- Harrington, L. C., Ponlawat, A., Edman, J. D., Scott, T. W., & Vermeulen, F. (2008). Influence of container size, location, and time of day on oviposition patterns of the dengue vector, *Aedes aegypti*, in Thailand. *Vector Borne Zoonotic Dis.*, *8* (3), 415-423. doi:10.1089/vbz.2007.0203
- Hoffmann, A. A., & Ross, P. A. (2018). Rates and patterns of laboratory adaptation in (mostly) insects. *J. Econ. Entomol.*, *111* (2), 501-509.
- Johnson, M. T. J., & Munshi-South, J. (2017). Evolution of life in urban environments. *Science*, *358* (6363), eaam8327. doi:10.1126/science.aam8327

- Kemp, A., & Jupp, P. (1991). Potential for dengue in South Africa: Mosquito ecology with particular reference to *Aedes aegypti*. *J. Am. Mosq. Control Assoc.*, 7 (4), 574-583.
- Kotsakiozi, P., Evans, B. R., Gloria-Soria, A., Kamgang, B., Mayanja, M., Lutwama, J., . . . Badolo, A. (2018). Population structure of a vector of human diseases: *Aedes aegypti* in its ancestral range, Africa. *Ecol. Evol.*, 8 (16), 7835-7848.
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.*, 79 (17), 5112-5120.
- Kramer, W. L., & Mulla, M. S. (1979). Oviposition attractants and repellents of mosquitoes: oviposition responses of *Culex* mosquitoes to organic infusions. *Environ. Entomol.*, 8 (6), 1111-1117.
- Leahy, S. M., VandeHey, R., & Booth, K. (1978). Differential response to oviposition site by feral and domestic populations of *Aedes aegypti* (L.) (Diptera: Culicidae). *Bull. Entomol. Res.*, 68 (03), 455-463.
- Lenth, R., Singmann, H., & Love, J. (2018). Emmeans: estimated marginal means, aka least-squares means. *R package version*, 1 (1).
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2 (3), 18-22.
- Lounibos, L. (1981). Habitat segregation among African treehole mosquitoes. *Ecol. Entomol.*, 6 (2), 129-154.
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15 (12), 550.
- Mangudo, C., Aparicio, J. P., & Gleiser, R. M. (2015). Tree holes as larval habitats for *Aedes aegypti* in urban, suburban and forest habitats in a dengue affected area. *Bull. Entomol. Res.*, 105 (6), 679-684.
- McBride, C. S. (2016). Genes and odors underlying the recent evolution of mosquito preference for humans. *Curr. Biol.*, 26 (1), R41-R46.
- McBride, C. S., Baier, F., Omondi, A. B., Spitzer, S. A., Lutomiah, J., Sang, R., . . . Vosshall, L. B. (2014). Evolution of mosquito preference for humans linked to an odorant receptor. *Nature*, 515 (7526), 222-227. doi:10.1038/nature13964
- McKinney, M. L. (2008). Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosyst.*, 11 (2), 161-176. doi:10.1007/s11252-007-0045-4
- McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, 8 (4).
- Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.*, 18 (3), 375-402.
- Otto, S. P. (2018). Adaptation, speciation and extinction in the Anthropocene. *Proceedings of the Royal Society B: Biological Sciences*, 285 (1891), 20182047. doi:doi:10.1098/rspb.2018.2047
- Pamplona Lde, G., Alencar, C. H., Lima, J. W., & Heukelbach, J. (2009). Reduced oviposition of *Aedes aegypti* gravid females in domestic containers with predatory fish. *Trop. Med. Int. Health*, 14 (11), 1347-1350. doi:10.1111/j.1365-3156.2009.02377.x
- Paupy, C., Brengues, C., Kamgang, B., Hervé, J.-P., Fontenille, D., & Simard, F. (2014). Gene flow between domestic and sylvan populations of *Aedes aegypti* (Diptera: Culicidae) in North Cameroon. *J. Med. Entomol.*, 45 (3), 391-400.
- Petersen, J. L. (1977). *Behavioral Differences in Two Subspecies of Aedes aegypti (L.) (Diptera: Culicidae) in East Africa*. (Ph.D. dissertation), University of Notre Dame, Indiana,

- Ponnusamy, L., Schal, C., Wesson, D. M., Arellano, C., & Apperson, C. S. (2015). Oviposition responses of *Aedes* mosquitoes to bacterial isolates from attractive bamboo infusions. *Parasites Vectors*, *8*, 486. doi:10.1186/s13071-015-1068-y
- Ponnusamy, L., Wesson, D. M., Arellano, C., Schal, C., & Apperson, C. S. (2010). Species composition of bacterial communities influences attraction of mosquitoes to experimental plant infusions. *Microb. Ecol.*, *59* (1), 158-173.
- Ponnusamy, L., Xu, N., Nojima, S., Wesson, D. M., Schal, C., & Apperson, C. S. (2008). Identification of bacteria and bacteria-associated chemical cues that mediate oviposition site preferences by *Aedes aegypti*. *Proc. Natl. Acad. Sci. U. S. A.*, *105* (27), 9262-9267. doi:10.1073/pnas.0802505105
- Powell, J. R., Gloria-Soria, A., & Kotsakiozi, P. (2018). Recent history of *Aedes aegypti* : vector genomics and epidemiology records. *Bioscience*, *68* (11), 854-860.
- Powell, J. R., & Tabachnick, W. J. (2013). History of domestication and spread of *Aedes aegypti* -A Review. *Mem. Inst. Oswaldo Cruz*, *108*, 11-17.
- Qi, Y. (2012). Random forest for bioinformatics. In *Ensemble machine learning* (pp. 307-323): Springer.
- Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.*, *62* (2), 142-160.
- Refsnider, J. M., & Janzen, F. J. (2010). Putting Eggs in One Basket: Ecological and Evolutionary Hypotheses for Variation in Oviposition-Site Choice. *Annu Rev Ecol Evol Syst*, *41* (1), 39-57. doi:10.1146/annurev-ecolsys-102209-144712
- Reiskind, M. H., & Zarrabi, A. A. (2012). Water surface area and depth determine oviposition choice in *Aedes albopictus* (Diptera: Culicidae). *J. Med. Entomol.*, *49* (1), 71-76.
- Rey, J. R., & O'Connell, S. M. (2014). Oviposition by *Aedes aegypti* and *Aedes albopictus*: Influence of congeners and of oviposition site characteristics. *J. Vector Ecol.*, *39* (1), 190-196. doi:https://doi.org/10.1111/j.1948-7134.2014.12086.x
- Rose, N. H., Sylla, M., Badolo, A., Lutomia, J., Ayala, D., Aribodor, O. B., . . . McBride, C. S. (2020). Climate and Urbanization Drive Mosquito Preference for Humans. *Curr. Biol.*, *30* (18), 3570-3579.e3576. doi:https://doi.org/10.1016/j.cub.2020.06.092
- Rueda, L. M. (2004). Pictorial keys for the identification of mosquitoes (Diptera: Culicidae) associated with Dengue Virus Transmission. *2004*, *589* (1), 60. doi:10.11646/zootaxa.589.1.1
- Saul, S. H., Novak, R. J., & Ross, Q. E. (1980). The role of the preadult stages in the ecological separation of two subspecies of *Aedes aegypti*. *Am. Midl. Nat.*, 118-134.
- Schluter, D. (2000). *The ecology of adaptive radiation* : OUP Oxford.
- Servedio, M. R., Van Doorn, G. S., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: 'magic' but not rare? *Trends Ecol. Evol.*, *26* (8), 389-397.
- Shafer, A., & Wolf, J. B. (2013). Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecol. Lett.*, *16* (7), 940-950.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27* (3), 379-423.
- Singer, M. C. (2004). *Oviposition preference: its definition, measurement and correlates, and its use in assessing risk of host shifts*. Paper presented at the Proceedings of the XI International Symposium on Biological Control of Weeds.

- Soghigian, J., Andreadis, T. G., & Livdahl, T. P. (2017). From ground pools to treeholes: convergent evolution of habitat and phenotype in *Aedes* mosquitoes. *BMC Evol. Biol.*, *17* (1), 262. doi:10.1186/s12862-017-1092-y
- Soghigian, J., Gloria-Soria, A., Robert, V., Le Goff, G., Failloux, A. B., & Powell, J. R. (2020). Genetic evidence for the origin of *Aedes aegypti*, the yellow fever mosquito, in the southwestern Indian Ocean. *Mol. Ecol.*, *29* (19), 3593-3606. doi:10.1111/mec.15590
- Soman, R. S., & Reuben, R. (1970). Studies on the preference shown by ovipositing females of *Aedes aegypti* for water containing immature stages of the same species. *J. Med. Entomol.*, *7* (4), 485-489.
- Souza, R. S., Virginio, F., Riback, T. I. S., Suesdek, L., Barufi, J. B., & Genta, F. A. (2019). Microorganism-Based Larval Diets Affect Mosquito Development, Size and Nutritional Reserves in the Yellow Fever Mosquito *Aedes aegypti* (Diptera: Culicidae). *Frontiers in Physiology*, *10* (152). doi:10.3389/fphys.2019.00152
- Starrfelt, J., & Kokko, H. (2012). Bet-hedging—a triple trade-off between means, variances and correlations. *Biol. Rev.*, *87* (3), 742-755.
- Swan, T., Lounibos, L., & Nishimura, N. (2018). Comparative oviposition site selection in containers by *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) from Florida. *J. Med. Entomol.*, *55* (4), 795-800.
- Sylla, M., Bosio, C., Urdaneta-Marquez, L., Ndiaye, M., & Black IV, W. C. (2009). Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLoS Negl. Trop. Dis.*, *3* (4), e408.
- Szulkin, M., Munshi-South, J., & Charmantier, A. (2020). *Urban evolutionary biology* : Oxford University Press, USA.
- Tabachnick, W. J., Munstermann, L. E., & Powell, J. R. (1979). Genetic distinctness of sympatric forms of *Aedes aegypti* in East Africa. *Evolution*, *33*, 287-295.
- Trpis, M., & Hausermann, W. (1975). Demonstration of differential domesticity of *Aedes aegypti* (L.) (Diptera, Culicidae) in Africa by mark-release-recapture. *Bull. Entomol. Res.*, *65* (02), 199-208.
- Vezzani, D. (2007). Artificial container-breeding mosquitoes and cemeteries: a perfect match. *Trop. Med. Int. Health*, *12* (2), 299-313.
- Vonesh, J. R., & Blaustein, L. (2010). Predator-induced shifts in mosquito oviposition site selection: a meta-analysis and implications for vector control. *Isr J Ecol Evol*, *56* (3-4), 263-279.
- Wong, J., Morrison, A. C., Stoddard, S. T., Astete, H., Chu, Y. Y., Baseer, I., & Scott, T. W. (2012). Linking oviposition site choice to offspring fitness in *Aedes aegypti*: consequences for targeted larval control of dengue vectors. *PLoS Negl. Trop. Dis.*, *6* (5), e1632. doi:10.1371/journal.pntd.0001632
- World Health Organization. (2014). *A global brief on vector-borne diseases* . WHO, Geneva, Switzerland.
- Xia, S. (2021). Laboratory Oviposition Choice of *Aedes aegypti* (Diptera: Culicidae) From Kenya and Gabon: Effects of Conspecific Larvae, Salinity, Shading, and Microbiome. *J. Med. Entomol.* doi:10.1093/jme/tjaa285
- Xia, S., Cosme, L. V., Lutomiah, J., Sang, R., Ngangue, M. F., Rahola, N., . . . Powell, J. R. (2020). Genetic structure of the mosquito *Aedes aegypti* in local forest and domestic habitats in Gabon and Kenya. *Parasites Vectors*, *13* (1), 417. doi:10.1186/s13071-020-04278-w
- Yee, D. A. (2008). Tires as habitats for mosquitoes: a review of studies within the eastern United States. *J. Med. Entomol.*, *45* (4), 581-593.
- Yee, D. A., Allgood, D., Kneitel, J. M., & Kuehn, K. A. (2012). Constitutive Differences Between Natural and Artificial Container Mosquito Habitats: Vector Communities, Resources, Microorganisms, and Habitat Parameters. *J. Med. Entomol.*, *49* (3), 482-491. doi:10.1603/me11227



Zahiri, N., & Rau, M. E. (1998). Oviposition attraction and repellency of *Aedes aegypti* (Diptera: Culicidae) to waters from conspecific larvae subjected to crowding, confinement, starvation, or infection. *J. Med. Entomol.*, 35, 782-787.

## Tables

**Table 1.** Number of oviposition sites measured for different environmental variables

Field site location	Habitat	<i>Aedes aegypti</i>	Physical characteristics	Larval density	Microbial density	Bacteria composition	Volatile profile
La Lopé, Gabon	Forest	Present	5	5	5	5	na*
		Absent	48	55	10	33	na*
	Peridomestic (Village)	Present	13	13	10	10	na*
		Absent	24	25	12	23	na*
Rabai, Kenya	<b>Total</b>		90	98	37	71	na*
	Forest	Present	15	15	15	15	7
		Absent	22	22	11	22	12
	Peridomestic (Village)	Present	8	8	8	8	5
		Absent	1	1	1	1	1
	Domestic (Village)	Present	22	22	22	22	17
		<b>Total</b>	68	68	57	68	42

\* Headspace volatiles were not collected in La Lopé, Gabon.

**Table 2.** Classification of La Lopé larval breeding sites by random forest

	Forest +	Forest -	Peridomestic +	Peridomestic -	Total
Forest +	<b>0</b>	5	0	0	5
Forest -	0	<b>28</b>	0	0	28
Peridomestic +	0	0	<b>2</b>	8	10
Peridomestic -	0	4	3	<b>13</b>	20

Note: the row names indicate the actual larval breeding site groups, and the column names indicate the assigned larval breeding site groups by the random forest classification. “+” and “-” denote *Ae. aegypti* present and absent sites, respectively. The number of sites assigned to the correct groups is marked in bold.

**Table 3.** Classification of Rabai larval breeding sites by random forest

	Forest +	Forest -	Peridomestic +	Peridomestic -	Domestic +	Total
Forest +	<b>9</b>	4	0	0	0	13
Forest -	8	<b>1</b>	0	0	1	10
Peridomestic +	0	0	<b>3</b>	0	5	8
Peridomestic -	0	0	1	<b>0</b>	0	1
Domestic +	0	0	2	0	<b>18</b>	20

Note: the row names indicate the actual larval breeding site groups, and the column names indicate the assigned larval breeding site groups by the random forest classification. “+” and “-” denote *Ae. aegypti* present and absent sites, respectively. The number of sites assigned to the correct groups is marked in bold.

## Figure legends

**Figure 1.** Sampling locations in (a) La Lopé, Gabon, and (b) Rabai, Kenya. The inset in each graph shows the location of the field site in continental Africa. In (a), each point represents a sampling location where one to multiple larval breeding sites were found. In (b), each point represents a single larval breeding site. The color of the point indicates the habitat category: red - domestic (village indoor), yellow - peridomestic (village outdoor), green - forest. The satellite images were from (a) Google Satellite and (b) Bing Satellite in QGIS.

**Figure 2.** Environmental conditions of larval breeding sites in La Lopé (a, c, e, g) and Rabai (b, d, f, h). (a-b) Principal component analysis (PCA) of all physical variables. The first two PCs are shown, and the variance explained by each PC was indicated in the axis label. Each point represents a single larval breeding site. Colors and point shapes indicate habitat and whether *Ae. aegypti* were found in the sites, respectively. An eclipse was drawn for each larval breeding site group with a 75% confidence level. The colors of the eclipses represent habitat types and match the colors of the points. The solid and dashed eclipses correspond to *Ae. aegypti* present and absent sites. The original variables were overlaid on the PC1-PC2 plate with major variables labeled. (c-d) Comparison of *Ae. aegypti* density between habitats. “+” and “-” denote *Ae. aegypti* presence and absence, respectively. The boxplots show the minimum, 25% quartile, median, 75% quartile, and maximum. Asterisks indicate statistical significance ( $p < 0.05$ ). (e-f) Comparison of microbial density across larval breeding site groups. (g-h) NMDS analysis of bacterial community compositions. The color and shape of points and eclipses are the same as in the PCA (panel a-b).

**Figure 3.** Chemical profile of volatile samples collected from Rabai larval breeding sites. Each row represents a compound, and each column represents a larval breeding site. The five columns of points between the compound names and the heatmap summarize whether the compounds were present in each of the five larval breeding site groups. The color and shape of the points are the same as in Figure 2. The color of each cell in the heatmap quantifies the concentration on a log scale. Gray cells indicate that the compound was not found in the larval breeding sites, according to the GC-MS results. The inset Venn diagram shows the total numbers of compounds unique in each habitat or shared between different habitats.

**Figure 4.** Two-choice laboratory oviposition assays testing preference for field-collected waters, pH, shading, a combination of water pH, salinity, and shading, *Ae. aegypti* larval density, and bacterial culture. Colony-wise results are shown in Figure S11 in the Appendix. The two choices in each assay are described in detail in Table S2 in the Appendix. Higher OIA implies a preference for the forest condition. Each point represents the OAI of one cage with five gravid females. The mean and 95% confidence interval (CI) were estimated by beta-binomial models. The asterisks and ‘ns’ above each colony indicates whether the 95% CI excludes zero. No significant differences were found between habitats or between colonies in any experiments.

**Figure 5.** Five-choice laboratory oviposition assays testing preference for bacterial density. Five cups were provided in each cage with increasing bacterial density at 0,  $2 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ ,  $2.5 \times 10^7$  cells/mL (details in Table S2), which correspond to the five columns (left to right) in each panel. Each line connects the five egg counts in one cage. Colors represent the habitats of the colonies. Multiple colonies from the same habitat in Rabai were combined in this figure, and colony-wise results are shown in Figure S12 in the Appendix. A negative binomial model was used to fit the results of each oviposition assay. The model estimates the mean number of eggs in each bacterial density and a 95% confidence interval, shown by the open circles and the error bars, respectively.

## Data Accessibility Statement

The datasets that describe the basic information, physical characteristics, larval density, predator presence, microbial density, and chemical profile of oviposition sites in La Lopé and Rabai are archived in Dryad:

doi:10.5061/dryad.7m0cfxprg (La Lopé) and doi:10.5061/dryad.3tx95x6cz (Rabai), respectively. The 16S-rRNA gene amplicon sequencing data were deposited in the NCBI SRA database with ID SUB7716639 (La Lopé samples) and SUB7719551 (Rabai samples).

## Competing Interests Statement

The authors declare that they have no competing interests.

## Author contributions

SX and JRP designed and conceptualized the study. SX, DA, RS, JL, CSM, NHR, and JRP coordinated the fieldwork. SX, DA, and JL conducted field sampling of oviposition sites and field oviposition experiments. SX and HD designed the volatile collection in Rabai and performed the GC-MS analysis. SX performed the lab work to generate the data on microbial density and bacterial community composition. SX, CSM, and NHR established mosquito colonies, and SX performed the laboratory oviposition assays. SX wrote the first draft of the manuscript. JRP provided funding, coordinated the entire study, and interpreted results with SX. All authors provided critical feedback on the manuscript.

## Acknowledgments

We appreciate the collaboration and the support from Institut de Recherche pour le Développement (IRD) and the research Unit ESV-GAB at the Centre International de Recherches médicales de Franceville (CIRMF) in Gabon, and Kenya Medical Research Institute (KEMRI) in Kenya during the fieldwork. We are grateful to all the field assistants and scientists in the field, especially Nil Rahola and Marc F. Ngangue in La Lopé and Rotich Gilbert in Rabai. In addition, we thank Andrew Goodman and his lab for providing primers for bacterial amplicon sequencing and helpful guidance in library preparation. We also thank Nanxi Lu for the instructions on bioinformatic analysis of the sequencing results. We received a lot of technical support and training from the Yale Center for Genome Analysis (YCGA) on Illumina sequencing, from the West Campus Analytic Core on GC-MS, and from the Yale West Campus Imaging Core on fluorescent microscopes, and we are grateful for all the support. The design of the laboratory experiments benefited greatly from the helpful discussions with Luciano Cosme, Ryan Joseph, and Lisa Baik. We appreciate all the useful discussions, suggestions, and feedback from Gisella Caccone, Tom Chiodo, Benjamin Evans, Stephen Gaughran, Andrea Gloria-Soria, Evelyn Jensen, Panagiota Kotsakiozi, Joshua Miller, Evelyn Pless, Maud Quinzin, Norah Saarman, Samuel Snow, and John Soghigian. We also want to thank members of the McBride lab at Princeton University for valuable feedback and discussions, and we are grateful for the advice and guidance from Stephen Stearns, John Carlson, and Alvaro Sanchez. Lastly, we thank the advice and comments from the editor and the reviewers that improved this manuscript.

This work was supported by NIH RO1 AI101112 to JRP and YIBS Small Grants Program, Doctoral Dissertation Improvement Awards to SX.







