

Larval stage co-occurrence and ecological niche partitioning of invasive *Anopheles stephensi* relative to native malaria vectors in Kenya

Duncan K. Athinya

danathinya@gmail.com

University of Nairobi

Brian O. Polo

Kenya Medical Research Institute

Margaret Muchoki

Kenya Medical Research Institute

Edith Ramaita

Ministry of Health

Kibor Keitany

Ministry of Health

Teresa Bange

Kenya Medical Research Institute

Moureen Ekisa

Kenya Medical Research Institute

Sylvia Milanoi

Kenya Medical Research Institute

Cyprian Adala

Kenya Medical Research Institute

Miguel Okoko

Kenya Medical Research Institute

LidonMike Ochieng

Kenya Medical Research Institute

Wolfgang R. Mukabana

University of Nairobi

Florence Oyieke

University of Nairobi

Bernard Abong'o

Kenya Medical Research Institute

Vincent Moshi

Kenya Medical Research Institute

Luigi Sedda

Lancaster University

Samson Kiware

Ifakara Health Institute

Eric Ochomo

Kenya Medical Research Institute

Article

Keywords: Anopheles stephensi, Kenya, habitat characteristics, species co-occurrence

Posted Date: September 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10876468/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Anopheles stephensi was first detected in Kenya in 2022 with subsequent detections in 2023 and 2024. In response, the AnoSTEP Afrika Program conducted longitudinal surveillance using Community Health Promoters across Kenya's ecological and malaria transmission landscape to understand habitat characteristics associated with spread. Spread hotspots emerged in northern, northeastern, and Rift Valley Kenya, with newer invasion fronts appearing along coastal transport corridors. Univariate assessment signaled *An. stephensi* as a container breeder, preferring containers deeper than 50cm, with the most preferred habitat type being water storage tanks (OR = 2.74, $P < 0.001$), in both rural (OR = 1.61, $P = 0.002$) and urban (OR = 1.64, $P = 0.01$) areas. Multivariate niche partitioning using Redundancy analysis (RDA) confirmed water storage tanks within 100m of animal shelters as niche characteristics for *An. Stephensi*, with niche partitioning between *An. arabiensis* and *An. stephensi*. Affinity analysis confirmed negative co-occurrence ($\hat{\alpha} = -4.66$, $P < 0.001$) between *An. stephensi* and both *An. arabiensis* and *An. coluzzii* ($\hat{\alpha} = -1.011$, $P < 0.001$). No significant association was observed between *An. stephensi* and *An. gambiae* s.s. Targeting water storage containers, particularly near animal shelters, could disrupt *An. stephensi* spread trajectory in Kenya.

Introduction

Malaria continues to be one of the greatest global health challenges disproportionately affecting sub-Saharan Africa. The World Malaria Report of 2025 highlighted that out of the 282 million cases and 610,000 deaths estimated in 2024, 94% and 95%, respectively, were in Africa, with 11 countries in Africa contributing more than 60% of the burden [1]. Control and elimination efforts, including Insecticide-Treated Nets and Indoor Residual Spraying, have been in place and helped achieve significant gains over the years [2].

Malaria remains a persistent and highly adaptable disease. The *Plasmodium* parasites responsible for malaria transmission have co-evolved with *Anopheles* mosquitoes, which sustain their populations through rapid environmental adaptation despite control efforts [3]. Biological threats related to parasite and vector adaptations have plagued malaria control and elimination efforts in Africa. These include drug resistance in *Plasmodium* parasites, insecticide resistance in malaria vectors, and, more recently, the emergence of invasive vectors, including *Anopheles stephensi* [4].

Kenya accounts for 1.5% of global malaria cases and 1.9% of global malaria deaths [1]. Through strategic case management, vector control, and chemotherapies, Kenya has experienced sustained gains against malaria, including a 50% reduction in nationwide malaria prevalence between 2010 and 2020, and a 32% decline in inpatient malaria deaths between 2018 and 2023 [5]. Among the biological threats to malaria control and elimination, insecticide resistance and *An. stephensi* invasion pose the greatest risks for Kenya. Insecticide resistance is being addressed through the deployment of next-generation vector control tools, including pyrethroid-chlorfenapyr ITNs [6]. *An. stephensi*, by contrast,

poses a significant long-term risk because it is a recent arrival; therefore, its full impact on malaria transmission and how well current interventions will hold against it remain poorly understood [7, 8].

A recent study using whole-genome sequencing traced the spread of *An. stephensi* back to an initial lineage originating in India or Pakistan. In Africa, the species first became established in Djibouti, from where it spread along several distinct regional fronts, one being the Ethiopia-Kenya corridor. Invading populations appear well adapted to urban environments and carry metabolic detoxification genes conferring resistance to public health insecticides. The study also found that landscape features shape the pace of spread; complex, steep terrain such as the Ethiopian highlands acts as a natural barrier to gene flow, whereas open, flat landscapes such as those in Sudan permit rapid, uniform expansion over large distances [9].

The invasion of *An. stephensi* raises the risk of increased drug-resistant malaria and the expansion of *Plasmodium vivax*-related malaria cases in sub-Saharan Africa [10, 11], a practical concern for malaria control programs already affected by funding cuts. Understanding the habitat characteristics associated with its invasion trajectory could support the implementation of cost-effective control measures. In Kenya, following its first detection in Marsabit and Turkana in 2022 and subsequent detection in counties such as Mandera, Wajir, Isiolo, Elgeyo Marakwet in 2023 and 2024 [7, 8], the AnoSTEP Afrika Program [12] sought to conduct longitudinal surveillance to characterize the *An. stephensi* spread trajectory. Traditional vector surveillance often lacks the geographic reach and sampling frequency needed to monitor such expansion. This study addressed the gap by integrating Community Health Promoters (CHPs), already operating within the Ministry of Health structure, to conduct entomological collections. This extended spatial and temporal coverage while building sustainable local capacity for early detection and monitoring. A model-based adaptive sampling method was leveraged for site selection, ensuring effective targeting of *An. stephensi* across Kenya's ecological and malaria transmission landscape [13].

Using habitat and species data collected through this CHP-based surveillance process, this study assesses whether invasive and native *Anopheles* species show distinct environmental gradients and whether the expansion of *An. stephensi* in Kenya reflects directional ecological niche specialization. The resulting ecological findings translate into practical evidence for where resources can be deployed most effectively to break the current spread trajectory.

Methodology

Collection site mapping and identification

Sampling points were mapped across Kenya via a recently developed adaptive sampling design that leverages initial detection of *An. stephensi* in Kenya and environmental covariates, while stratifying for malaria intensity and urban/rural land cover. The surveillance was optimized to minimize uncertainty around the highest trapping probability for *An. stephensi* and identified 56 sites across Kenya [13]. Ground truthing was conducted, followed by replacement of unreachable sites (mostly due to insecurity)

and areas with limited CHP participation through a remapping exercise. This process reduced the risks of convenience sampling, which often concentrates sampling efforts near roads or research centers due to ease of access [14].

A total of 61 sites spread across 21 counties were adopted for longitudinal surveillance (Fig. 1). The larval collections were conducted between July 2025 and March 2026 across all 61 sites. In collaboration with the Kenya National Malaria Control Programme (NMCP) and the County Health Departments, local and resident Community Health Promoters (CHPs) in charge of the localities closest to the surveillance site were recruited, trained, and equipped to conduct larval habitat surveillance. Sampling was conducted within a 2km radius of each surveillance site to ensure the surveillance effort remained within walking distance. The CHPs conducted exhaustive larval collections at each habitat, using the standard dipping method as the primary technique for determining larval presence or absence. Because sampling spanned both urban and rural settings across Kenya, all water catchment and storage structures (including natural habitats) were sampled. These included dams, tanks, animal watering pans, drainage channels, garbage dumping sites, and other artificial containers.

All larval habitats were geo-coded and photographed, with each record including the date of collection, geography (county, sub-county), and an automatically assigned unique barcode which was attached to the larvae storage tube for each collection. Environmental traits of the habitat were also recorded, including estimated size, water depth, permanence, clarity/turbidity, shade cover, algal growth, vegetation cover, and distance to the nearest human or animal shelter. To minimize the risk of sample cross-contamination, larvae collected from each habitat were preserved directly at the collection site in the sample tubes pre-filled with 80% ethanol, each having a unique barcode. The habitat collections focused on *Anopheles* larvae but also recorded the presence or absence of culicine larvae and potential predators. The assigned barcode linked each unique *Anopheles*-positive collection record to laboratory procedures for species identification, enabling assessment of larval co-occurrence and niche partitioning.

Each month, KEMRI-based field entomologists met with CHPs at designated collection hubs to collect larval samples, which were then transported to a central laboratory for molecular species identification. During the visits, field entomologists also supervised and supported CHPs, including targeted retraining for those having difficulty identifying and characterizing larval habitats.

Mosquito reception and PCR species identification

The larval specimens, stored in 80% ethanol in barcoded tubes, were transported to the Kenya Medical Research Institute Center for Global Health Research (KEMRI-CGHR) entomology and molecular laboratory in Kisumu, Kenya. Upon arrival, all *Anopheles* larvae from each tube were stratified by instar and divided into two batches: one for conventional PCR and one retained for future genomic studies. This two-batch approach (running the faster PCR assay while reserving material for future genomic studies) was adopted for samples collected between July and December 2025 to increase the likelihood

of more immediate detection of new sites with *An. stephensi* across the country while retaining material for future downstream genomic studies. To minimize reagent consumption, a subsampling strategy was subsequently adopted between January to March 2026 whereby, for habitats containing more than six *Anopheles* larvae, five larvae (in either the 3rd or 4th instar) were randomly sampled per habitat for both species identification with conventional PCR and storage of genetic material for future genomic studies. Whole larval DNA was extracted using ethanol precipitation [15]. Briefly, samples were homogenized in grinding buffer via bead-based tissue lysis, incubated at 65°C for 30 minutes, and then treated with 8M potassium acetate on ice to precipitate proteins. Following centrifugation, the supernatant was collected, and DNA was precipitated using cold absolute ethanol at -20°C, washed sequentially with 70% ethanol, air-dried, and reconstituted in TE buffer.

Anopheles stephensi was identified using a species-specific endpoint PCR assay targeting the ITS2 region, employing primers St-F, U5.8S-F, and UD2-R [16]. Reactions comprised Hot Start Taq 2X Master Mix, primers, nuclease-free water, and 0.15 µl DNA template, alongside positive and negative controls in each run. PCR products were resolved on a 2% agarose gel alongside a 100 bp DNA ladder and visualized using a gel documentation system, with an expected amplicon size of approximately 438 bp indicating a positive result. This PCR-based approach was adopted as the primary method of species identification, given that ethanol precipitation defaced the larval sample, making it impossible to identify mosquitoes based on taxonomy. Samples not identified as *An. stephensi* were further characterized using ribosomal DNA gene probe techniques to distinguish members of the *An. gambiae* complex [17]. Remaining samples were then amplified and identified using an endpoint PCR assay targeting *An. funestus* group species [18].

Data preparation and analysis

The recorded habitat characteristics, larval collection counts, and species identification results were imported into R version 4.5.1 for cleaning, using the data processing functions in Tidyverse [19]. The data tables (habitat characteristics, larval counts, and species identification) were left-joined using the unique habitat identifier, the sample tube barcode. Habitats without *Anopheles* larvae were excluded from the analysis.

Cleaning involved removing typos and standardizing the raw habitat characteristics into 11 habitat type categories based on prior knowledge of what the entries represented: Small Household Containers, Water Storage Tanks, Tap and Pipe Leakage, Drainage/Ditches/Waste Sites, Animal Prints/Troughs, Wells and Boreholes, Natural Water Bodies, Puddle and Stagnant Rainwater, Man-made Pits/Ponds/Dams, Agricultural Sites, and Domestic Water Sites.

Spread hotspot visualization

To visualize the *An. stephensi* temporal occurrence concentrations, detection data from this collection were combined with detections from previous collection efforts [7, 8] (from the year 2022–2024). The 2022–2024 collections were grouped into a single period bin, whereas collections within this AnoSTEP Afrika surveillance program were grouped by month. These were rasterized on a 2km raster grid covering Kenya, followed by Gaussian kernel smoothing with a 50km bandwidth to generate continuous cumulative detection densities using the functions in the terra package in R [20]. The 50km bandwidth was applied as a reasonable estimate of local movement, informed by metrics from a socio-cultural survey within the AnoSTEP Afrika project in Turkana County, Kenya. Smoothing was applied cumulatively to detections at each time step, with both past and current detections affecting each time step. This was done for all time steps except the first period (2022–2024 detections), where smoothing was applied to aggregate datasets in that period. The resulting continuous spatial density patterns were plotted alongside raw detection and non-detection points.

Univariate habitat association with *Anopheles stephensi*

Univariate analyses were conducted to probe the existence of an association between habitat characteristics and *An. stephensi* presence and absence before multivariate modeling. This used a binomial generalized linear model with Firth's bias reduction [21]. Categories with zero or very low counts were pooled into an "Other" category (Animal prints/Trough, Tire, Wells and Borehole) to address complete separation. Elevation was converted into a categorical variable using sample tertiles to create three equal-sized groups of observations: low elevation (< 387.3m), medium elevation (387.3-635.6m), and high elevation (> 635.6m). Additionally, categorical habitat characteristics (such as water tanks, dams, etc.) were converted to dummy-coded variables to assess the distinct effects of each habitat type.

Multivariate niche partitioning

Redundancy Analysis (RDA) was used to assess multivariate niche partitioning, evaluating *An. stephensi* and other *Anopheles* species identified by PCR against habitat characteristics from each collection effort. RDA was well suited to this evaluation, as it combines multivariate linear regression with principal component analysis to model how environmental factors shape species community structure, using a matrix of species counts per unique habitat and a matrix of environmental predictors for each habitat [22–24].

To prevent dominant species from skewing the linear analysis, we applied a Hellinger transformation to the species count matrix [22]. The categorical variables under habitat characteristics (such as water tanks, dams, etc.) were maintained as dummy variables to assess their distinct effects. Baseline multicollinearity was estimated using Variance Inflation Factors (VIF). Variables with VIF > 10 were considered indicative of high collinearity; however, these variables were not removed at this stage to avoid prematurely excluding biologically informative covariates.

The optimal subset of predictors in the final model was selected based on a forward selection process using the adjusted R^2 . The statistical significance of the final model, the retained environmental terms, and the individual canonical axes were assessed using Monte Carlo permutation tests with 1000 permutations. The analyses were conducted using the Vegan Package [25].

Species co-occurrence

Pairwise species associations across habitats were quantified using an affinity index. Associations were evaluated via maximum-likelihood estimation of the log-affinity parameter (α'), following the method of Mainali and Slud [26]. This approach provides a robust, scale-invariant measure of *An. stephensi*'s association with other species across habitats, accounting for sample size variation and using mid- P confidence intervals and associated p -values to assess statistical significance.

Results

Collection effort

Between July 2025 and March 2026, a total of 9,005 habitat surveys were conducted across 61 sites of which 25.75% (2,319) were positive for *Anopheles* larvae. In total, 6,821 *Anopheles* mosquitoes were identified to species level by PCR and are reported here (Fig. 2).

Anopheles arabiensis was the most abundant (72%) malaria vector among the PCR-identified collection, followed by *An. stephensi* at 22%. *Anopheles coluzzii*, recently identified in Kenya as an extension of its continuous geographic range rather than a single-point introduction [27], accounted for 5% of the collection, while its cryptic sibling, *An. gambiae* s.s. had an abundance of 1%. Eight *An. quadriannulatus* larvae were collected but not included in subsequent analysis due to low numbers.

Visualizing *An. stephensi* spatial occurrence trends by combining previous observations from 2022–2024 [7, 8] with outcomes of this study revealed a steady rise in overall detections across Kenya, along with both persistent and newly emerging high-density clusters. Persistent clusters were concentrated in the northern, northeastern, and Rift Valley regions (Baringo, Turkana, Samburu, Marsabit, Garissa, Wajir, Mandera), while new detection clusters emerged in the central-eastern region (Kitui) and along the coastal transport corridor (Tana River, Kilifi, Lamu). These findings underscore both the alarming spread of *An. stephensi* in Kenya over a relatively short period and the value of a targeted sampling framework, such as the one applied here, for monitoring an invasive vector [13] (Fig. 3, Supplementary Table 1).

Univariate evaluation of habitat characteristics with *An. stephensi*

Among the recorded habitat characteristics, elevation, vegetation presence, urbanicity, habitat size, water depth, breeding site type, water source type, water appearance (turbidity), presence of algae, habitat shading and the season when collections were made showed significant associations with *An. stephensi* presence (Fig. 4).

Higher elevation was negatively associated with *An. stephensi* presence, with both high (OR = 0.23, 95% CI: 0.17–0.33, $P < 0.001$) and medium elevations (OR = 0.38, 95% CI: 0.28–0.53, $P < 0.001$) showing lower odds compared with low elevation. Habitats with vegetation were also less likely to have *An. stephensi* (OR = 0.58, 95% CI: 0.44–0.75, $P < 0.001$). Additionally, there was a lower likelihood of finding *An. stephensi* in a habitat if it was not shaded (OR = 0.65, 95% CI: 0.44–0.97, $P = 0.036$) or partially shaded (OR = 0.59, 95% CI: 0.38–0.9, $P = 0.0152$).

Other factors negatively associated with *An. stephensi* presence included: water depth less than 10 cm (OR = 0.16, 95% CI: 0.08–0.31, $P < 0.001$); turbid water (OR = 0.68, 95% CI: 0.52–0.88, $P = 0.004$); water from natural sources (OR = 0.61, 95% CI: 0.45–0.82, $P = 0.001$); temporary habitats such as rain puddles (OR = 0.71, 95% CI: 0.53–0.95, $P = 0.02$); presence of algae, particularly at moderate quantities (OR = 0.44, 95% CI: 0.22–0.87, $P = 0.02$); and collections conducted during the rainy season (OR = 0.74, 95% CI: 0.56–0.98, $P = 0.03$).

Conversely, *An. stephensi* was more likely to be in a habitat if the water depth was more than 50cm (OR = 1.81, 95% CI: 1.07–3.06, $P = 0.03$), in both rural and urban areas (urban OR = 1.64, 95% CI: 1.11–2.41, $P = 0.01$; rural OR = 1.61, 95% CI: 1.19–2.17, $P = 0.002$), with the habitat type as water storage tanks (OR = 2.74, 95% CI: 2.02–3.71, $P < 0.001$).

Multivariate niche partitioning

The RDA provided further insight beyond the univariate analyses. Forward stepwise selection based on adjusted R^2 retained eight variables: elevation, evaluated as a continuous variable (Variance = 0.0356, $F = 95.7$, $P < 0.001$); water storage tanks (Variance = 0.0196, $F = 52.6$, $P < 0.001$); distance to the nearest animal shelter (Variance = 0.005, $F = 6.66$, $P < 0.001$); algae quantity (Variance = 0.00751, $F = 6.7207$, $P < 0.001$); water depth (Variance = 0.0039, $F = 5.26$, $P = 0.005$); water appearance (Variance = 0.0025, $F = 6.69$, $P = 0.011$); and distance to the nearest house (Variance = 0.0023, $F = 3.02$, $P = 0.035$). Transect (rural or urban) was also retained, though as a marginally non-significant variable (Variance = 0.00185, $F = 2.4843$, $P = 0.076$). All retained variables had low VIF (< 6.62), indicating multicollinearity was not a substantial concern in the final model. Together, they explained 18.34% (0.078) of the variation in community structure across habitats.

The unexplained (residual) variance accounted for 81.66% (0.348) of the total variance. Residual variance in this range is common in similar ecological studies, reflecting unmeasured biotic factors and other stochastic processes [22]. The overall model was confirmed statistically significant by permutation test ($F = 15$, $P < 0.001$). The first canonical axis (RDA1) explained the great majority of the constrained variation in mosquito species community composition (variance = 0.076, 96.9% of constrained variance; $F = 203.55$, $P < 0.001$). The remaining canonical axes were not significant; only RDA2 had $F > 1$ (variance = 0.00213, 2.1% of constrained variance; $F = 5.78$, $P = 0.09$).

Strong evidence of niche partitioning was observed between *An. stephensi* and *An. arabiensis* along the first canonical axis (RDA1, Fig. 5). *An. stephensi* scored strongly positive (+ 1.41), while *An. arabiensis*

scored equally strongly negative (-1.41) on the same axis. Environmental variables loading positively on RDA1 - and thus aligned with *An. stephensi* occurrence - included water storage tanks (+ 0.6), water depth more than 50cm (+ 0.27), and breeding sites within 100 m of an animal shelter (+ 0.2). The environmental variables loading negatively on RDA1 - and thus aligned with *An. arabiensis* occurrence - included higher elevation (-0.68), habitats with water depth less than 10cm (-0.36), turbid water (-0.29), and moderate quantities of algae (-0.19).

Anopheles coluzzii and *An. gambiae* s.s. showed weak scores along RDA1 (+ 0.09 and + 0.01, respectively), placing both species near the origin. This may reflect either environmental plasticity along the habitats or that the habitat characteristics captured by this canonical axis were insufficient to resolve their specialized niches (Fig. 5).

Anopheles coluzzii showed some separation from the other species along the second canonical axis (RDA2). At a score of + 0.3 on this axis, it was more distinct than *An. stephensi* (-0.09), *An. arabiensis* (-0.07) and *An. gambiae* s.s. (-0.01), which clustered near zero. Variables loading positively on this axis included distance to animal shelter greater than 500 m (+ 0.34), moderate algae (+ 0.31), and small habitats under 10 m (+ 0.31). However, as noted above, RDA2 was not statistically significant ($P = 0.09$), and this pattern should therefore be investigated further.

Species co-occurrence

Anopheles stephensi and *An. arabiensis* showed a strong, significant negative habitat association ($\hat{\alpha} = -4.66$, 95% CI: -5.347 to - 4.06; $P < 0.001$), indicating that they co-occurred less frequently than expected by chance. *Anopheles coluzzii* also showed significant negative co-occurrence with *An. stephensi* ($\hat{\alpha} = -1.011$, 95% CI: -1.585 to - 0.478; $P < 0.001$). The habitat association between *An. stephensi* and *An. gambiae* s.s. was not statistically significant ($P = 0.4652$) (Fig. 6).

Anopheles arabiensis showed a marginally significant positive habitat association with *An. coluzzii* ($\hat{\alpha} = 0.61$, 95% CI: 0.052–1.217; $P = 0.04$), but no significant association with *An. gambiae* s.s. ($P = 0.17$). The strongest habitat association for *An. gambiae* s.s. was with *An. coluzzii* ($\hat{\alpha} = 1.669$, 95% CI: 0.585–2.668; $P = 0.004$), indicating that these two species co-occurred more frequently than expected by chance (Fig. 6).

Discussion

This survey and analyses add to the scale and ecological information of *An. stephensi* spread that must be taken into account for control efforts. It correlatively highlights specific habitat features and biotic interactions that appear to be shaping the persistent hotspots and new detection sites across Kenya, highlighted across Figs. 4–6. Three patterns emerge from the univariate, co-occurrence, and multivariate analyses: a distinct man-made, water-storage niche; strong habitat segregation from *An. arabiensis*; and weak or undetectable differentiation from *An. gambiae* s.s. and *An. coluzzii*. Taken together, these sketch a coherent picture of an invasive mosquito exploiting a relatively underused ecological niche.

Habitat features shaping *An. stephensi* establishment

The habitat characteristics showing statistically significant signals for linkage to *An. stephensi* presence from the univariate analysis (deeper (more than 10cm), without vegetation, with no algae, clear (not turbid), and most preferably if it's a water storage tank) line up closely with what has been reported for this species elsewhere in its invasive range [28, 29]. Its consistent preference for water storage tanks and man-made water sources is a useful reminder that this is fundamentally a container-breeding mosquito whose success depends largely on man-made infrastructure rather than the ephemeral pools, hoofprints, swamps, vegetated natural water source margins that the native *An. arabiensis*, *An. gambiae* s.s., *An. funestus* predominantly rely on [30, 31]. The negative association with elevation is unsurprising given that temperature constraints on larval development and adult survival are among the best-established limits on *Anopheles* species distribution [32]. The elevation related odds ratios reported here (0.23 at high elevation, 0.38 at medium) suggest the effect in Kenya is graded rather than an abrupt cutoff, with implications for how far up the country's highland gradient surveillance and control efforts may eventually need to extend. One univariate signal did not carry through to the multivariate model: collections made during the rainy season were less likely to be positive. The most likely explanation is confounding rather than a spurious univariate finding; once the other variables are added, its contribution may simply be small.

The more striking result is that *An. stephensi* was likely to be detected in both rural and urban transects. Much of the existing narrative around *An. stephensi* invasion has implied it as an urban problem [10, 33], an idea that made intuitive sense given its association with overhead tanks and domestic water storage, but these data suggest the framing may be too narrow for Kenya, where rural households and small trading centers often rely on similar storage infrastructure as towns. It's also noteworthy that the RDA analysis included urbanicity as a marginally non-significant factor but included habitats within 100m of animal shelters, which were associated with *An. stephensi* and not other species in the analysis, likely confirming a level of zoophilic behavior as identified from blood-meal analysis in other studies [34]. It is worth being explicit about how this balances off with the RDA result that the retained habitat variables explained only 18.34% of community-structure variance, leaving 81.66% unexplained. That low total is expected given the many unmeasured biotic and stochastic factors that shape larval communities, and it does not undercut the niche-partitioning conclusion: almost all of the explained variance is concentrated in the first canonical axis alone (96.9% of constrained variance, $P < 0.001$), and that axis places *An. stephensi* and *An. arabiensis* at opposite ends of the same gradient, consistent with the independent affinity-based co-occurrence result. Two methods converging on the same separation, supported by retention of biologically important variables via a reasonable forward selection criterion, provide a stronger basis for the niche-partitioning claim than either method alone.

These habitat preferences also help explain the spatial pattern described in Fig. 3. The persistent hotspots in the northern, northeastern, and Rift Valley counties sit within Kenya's arid and semi-arid pastoral zones, where households and livestock keepers depend heavily on stored water for both domestic and animal use - precisely the tank - and animal-shelter-associated niche identified here. The

newer detections along the coastal transport corridor (Tana River, Kilifi, Lamu) are harder to explain on habitat grounds alone. While localized water storage in these arid environments plays a role, this expanding front more plausibly reflects high introduction pressure via road traffic rather than just local habitat suitability. Taken together, the niche and spread-pattern results suggest that *An. stephensi*'s current geography in Kenya is being shaped by at least two processes - where suitable containers already exist, and where the species is still arriving – and that distinguishing between them will matter for prioritizing surveillance as the invasion continues.

Negative association with *An. arabiensis*

The strong negative habitat association between *An. stephensi* and *An. arabiensis* ($\hat{\alpha} = -4.66$, $P < 0.001$) was one of the clearest signals in the co-occurrence analysis, and it aligns with the multivariate result placing the two species at opposite ends of the first RDA axis. This should not be interpreted as indisputable evidence of active competitive exclusion at the larval stage. The affinity index used here captures a statistical pattern of co-occurrence rather than a mechanism [26]. The segregation could be an indirect consequence of the two species independently sorting along the same environmental gradient, rather than direct interference or exploitative competition between larvae sharing a container. Disentangling would need experimental work with co-rearing larvae under controlled conditions, or paired comparisons of habitats that are otherwise similar but differ in occupancy. Either way, the practical conclusion from this work is *An. stephensi* appears to be occupying a niche that is currently underused by Kenya's dominant native vectors, which could likely accelerate its establishment.

No clear signal for *An. gambiae* s.s.

Anopheles gambiae s.s. did not show a statistically significant habitat association with *An. stephensi* from the affinity analysis and sat close to the origin on the significant axis of the RDA biplot (RDA1). Two explanations are plausible and not mutually exclusive. The first is simple statistical power: *An. gambiae* s.s. made up only 1% of the identified collections, and co-occurrence estimates for rare species are inherently less precise regardless of the underlying biology. The second is a genuine ecological one: that it is a habitat generalist with respect to the variables measured here, and that whatever separates its niche from *An. stephensi* operates along axes this study did not capture. The significant positive co-occurrence between *An. gambiae* s.s. and *An. coluzzii* ($\hat{\alpha} = 1.669$, $P = 0.04$) is a useful indicator. It confirms that the affinity index can detect real structure among these species, similar to what has been documented in the literature [35]. Its absence for *An. stephensi* association is therefore more likely to reflect true habitat overlap or plasticity than a limitation of the method. Therefore, confirming this would require a broader sampling scheme that reflects true habitat overlap.

What about *An. coluzzii*?

From the RDA analysis, *An. coluzzii* did not show a clear niche of separation along the significant canonical axis (RDA1), unlike *An. stephensi*. Some niche separation was observed for *An. coluzzii* along the second canonical axis, though the axis was not statistically significant. The RDA outcomes suggest,

like what was observed for *An. gambiae* s.s., either habitat plasticity along the habitat characteristics mapped along RDA1 or insufficient statistical power, as *An. coluzzii* was composed of only 5% of the collection. Despite potential habitat plasticity, *An. coluzzii* rarely shared breeding sites with *An. stephensi* as the affinity analysis showed significant negative co-occurrence between the two species ($\alpha' = -1.011$, $P < 0.001$). *An. coluzzii* was highly likely to occur in the same habitats as *An. gambiae* s.s. ($\alpha' = 1.669$, $P = 0.004$) and *An. arabiensis* ($\alpha' = 0.61$, $P = 0.04$) as reported in other studies[27] further demonstrating that *An. stephensi* is occupying distinct habitats relative to the native vectors.

Implications for surveillance and control

For a vector still expanding its range in Kenya, the main message to public health authorities is that man-made habitats, particularly water storage tanks near animal shelters, deserve particular attention in early containment efforts. Unlike the habitats that sustain much of the native vector community [30, 31] that are difficult to modify at scale, tanks are relatively few and findable features owned by members of the at-risk population, making them a good target for screening, larvicide application, and other methods that target mosquitoes in their aquatic stages. The rural-urban detection argues against limiting these interventions to towns and cities, instead targeting any household or community with uncovered or poorly maintained water storage in all *An. stephensi* risk areas, not just in northern Kenya, but also along trade routes through the coastal parts of Kenya. The community health promoter network used to sustain this study is a practical surveillance model worth noting. It shows that a coordinated community-based approach can generate the kind of fine-grained, repeated habitat data that a purely research-driven effort would struggle to match at this geographic scale [36].

This container-focused strategy assumes the niche described here holds as *An. stephensi* continues to establish, and that assumption deserves a caveat. Whole-genome evidence from elsewhere in its invasive range shows this species carries genes for insecticide resistance and adapts readily to new landscapes as it spreads [9]. The niche documented at this relatively early stage of the Kenyan invasion is not guaranteed to stay fixed, and control programs built around water storage tanks should be paired with continued monitoring for signs that the species is beginning to exploit other habitat types as competition, control pressure, or simple population growth change the calculus.

Limitations

The adaptive sampling framework used to prioritize the 61 study sites was built specifically to optimize surveillance for *An. stephensi* [13] and may not efficiently describe malaria vector composition and distribution. The habitat associations estimated here are best read as descriptive of where the species was searched for and found, rather than as an unbiased picture of its fundamental niche across the whole country. The method also limits the extrapolation of *An. coluzzii* and *An. gambiae* s.s. given that very few of them were collected. Finally, and as already noted, both the univariate and RDA results describe association rather than mechanism; they identify where *An. stephensi* is and is not being found, not why the underlying biology produces that pattern.

Conclusion

Taken together, these results reinforce that *An. stephensi* in Kenya is not simply a story of a new species arriving, but of it exploiting a specific niche that overlaps only loosely with the habitat used by the dominant native malaria vectors. This is important for implementation of control interventions, as container-focused source reduction can be realistically implemented using cost-effective community strategies. Whether that window of containment stays open depends on questions a single country snapshot cannot settle. This requires sustained, cross-country and community-backed surveillance and control activities.

Declarations

Ethical approval

This study was conducted in accordance with its protocol, which received ethical approval from the Scientific and Ethics Review Unit (SERU) of the Kenya Medical Research Institute (KEMRI) under protocol number KEMRI/SERU/CGHR/522/5105.

Funding

This work was supported, in whole, by ‘Comprehensive assessment of the biology and public health importance of *Anopheles stephensi* in Africa’ (funded by Wellcome Trust, 312634/Z/24/Z). The conclusions and opinions expressed in this work are those of the authors alone and shall not be attributed to the Wellcome Trust or authors’ affiliated organizations.

Author Contribution

DKA, BOP, MM, ER, WRM, FO, BA, VM, LS, SK, EOO conceived the study; BOP, MM, ER, KK, TB, ME, SM, CA, MO, LO, BA, VM, EOO contributed to sample collection, processing, and data collection. DKA, BOP, MM, WRM, FO, and EOO analysed the data and drafted the manuscript. All authors read and approved this manuscript.

Acknowledgement

We are grateful to Mathew Kipsum, Brian Morgan Odhiambo, Brenda Onyango, and Beverlyne Oduor of the KEMRI Molecular Laboratory for their dedication in receiving, sorting, and molecular processing of the larval samples for the AnoSTEP Project Kenya. We also thank Silas Koech, George Marube, and Duncan Ayata for their relentless travel to the study hubs to collect larval samples. We are indebted to the Community Health Promoters (CHPs) across all 61 study sites for their faithful and consistent

collection of larval samples throughout the study period. Finally, we acknowledge the Ministry of Health, through the National Malaria Control Program, and the county governments of the 21 counties where larval collection took place, for their support and collaboration throughout the AnoSTEP Project Kenya.

Data Availability

All data supporting the findings of this study are included in the article and its Supplementary Information files. Additionally, we supplied it to the World Health Organization (WHO), and it is accessible via WHO's Malaria Threats Map under "Invasive Vector Species".

<https://apps.who.int/malaria/maps/threats/>

No DNA sequence datasets were generated for the analyses presented in this manuscript. Some of the samples collected between 2022 and 2024 were previously sequenced and reported in a previous manuscript <https://www.nature.com/articles/s41598-025-04682-1>. DNA and mosquito tissues from the samples collected in 2025 and 2026 were stored for future genomic studies. PCR datasets from species ID are available upon request.

References

1. WHO. World malaria report 2025: addressing the threat of antimalarial drug resistance [Internet]. Geneva. (2025). <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2025>. Accessed 16 Feb 2026.
2. MAP. The changing impact of malaria control in Africa 2000–2025 [Internet]. 2025 [cited 2026 Feb 16]. <https://malariaatlas.org/project-resources/impact-of-control-2000-2025/>. Accessed 16 Feb 2026.
3. Loy, D. E. et al. Out of Africa: origins and evolution of the human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax*. *Int. J. Parasitol. Elsevier*. **47**, 87–97 (2017).
4. Platon, L., Lu, F. & Ménard, D. World malaria report 2025: Growing biological threats, shrinking resources. *Decoding Infect. Transmission Elsevier*. **4**, 100076 (2026).
5. Ministry of Health-National Malaria Control Programme. Kenya malaria strategy 2023–2027 [Internet]. Nairobi. (2023). <https://www.afro.who.int/sites/default/files/2025-03/Kenya%20Malaria%20Strategy%20%7C%202023-2027.pdf>. Accessed 2 Aug 2026.
6. Churcher, T. S. et al. The epidemiological benefit of pyrethroid–pyrrole insecticide treated nets against malaria: an individual-based malaria transmission dynamics modelling study. *Lancet Glob Health Elsevier*. **12**, e1973–e1983 (2024).
7. Ochomo, E. et al. Molecular surveillance leads to the first detection of *Anopheles stephensi* in Kenya. (2023). <https://doi.org/10.21203/rs.3.rs-2498485/v1>
8. Samake, J. N. et al. Spatial distribution and population structure of the invasive *Anopheles stephensi* in Kenya from 2022 to 2024. *Sci Rep. Nature Publishing Group UK London*; ;15:19878. (2025).

9. Dennis, T. P. W. et al. The origin, history, and resistance architecture of an invasive urban malaria mosquito in Africa. *Science American Association for the Advancement of Science*; 2026;393:eadx6925. (1979).
10. Emiru, T. et al. *Evidence for a role of Anopheles stephensi in the spread of drug-and diagnosis-resistant malaria in Africa*. *Nat Med* Vol. 29, 3203–3211 (Nature Publishing Group US New York, 2023).
11. de Santi, V. P. et al. Role of *Anopheles stephensi* mosquitoes in malaria outbreak, Djibouti, 2019. *Emerg. Infect. Dis.* **27**, 1697 (2021).
12. AnoSTEP Afrika [Internet]. [cited 2026 Aug 30]. <https://anostepafrika.net/>. Accessed 30 Aug 2026.
13. Sedda, L. et al. *Designing spatial adaptive surveillance for the emerging malaria vector Anopheles stephensi in Eastern and Horn of Africa*. *medRxiv* (Cold Spring Harbor Laboratory Press, 2026). 2023–6.
14. Sinka, M. E. et al. A global map of dominant malaria vectors. *Parasit. Vectors BioMed. Cent.* **5**, 1–11 (2012).
15. Green, M. R. & Sambrook, J. *Precipitation of DNA with ethanol*. *Cold Spring Harb Protoc* Vol. 2016, pdb–prot093377 (Cold Spring Harbor Laboratory Press, 2016).
16. Singh, O. P. et al. Molecular Tools for Early Detection of Invasive Malaria Vector *Anopheles stephensi* Mosquitoes. *Emerg. Infect. Dis. United States.* **29**, 36–44. <https://doi.org/10.3201/eid2901.220786> (2023).
17. Bass, C., Williamson, M. S. & Field, L. M. Development of a multiplex real-time PCR assay for identification of members of the *Anopheles gambiae* species complex. *Acta Trop. Elsevier.* **107**, 50–53 (2008).
18. Kokemoer, L. L., Kamau, L., Hunt, R. H. & Coetzee, M. A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *Am. J. Trop. Med. Hyg.* **66**, 804–811 (2002). [Lawrence, Kan., etc., Allen Press, etc.].
19. Wickham, H. et al. Welcome to the Tidyverse. *J. Open. Source Softw.* **4**, 1686 (2019).
20. Hijmans, R. J., Bivand, R., Dyba, K., Pebesma, E. & Sumner, M. Terra [R Package]. R Programming Language. ; (2023).
21. Heinze, G. & Schemper, M. A solution to the problem of separation in logistic regression. *Stat. Med. Wiley Online Libr.* **21**, 2409–2419 (2002).
22. Legendre, P. & Legendre, L. *Numerical ecology* (Elsevier, 2012).
23. Borcard, D., Gillet, F. & Legendre, P. *Numerical ecology with R* (Springer, 2011).
24. Capblancq, T. & Forester, B. R. Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods Ecol. Evol. Wiley Online Libr.* **12**, 2298–2309 (2021).
25. Oksanen, J. et al. Package ‘vegan.’ Community ecology package, version. ;2. (2018).
26. Mainali, K. P., Slud, E. & CooccurrenceAffinity An R package for computing a novel metric of affinity in co-occurrence data that corrects for pervasive errors in traditional indices. *PLoS One. Public Library of Science San Francisco, CA USA*; ;20:e0316650. (2025).

27. Kamau, L. et al. The *Anopheles coluzzii* range extends into Kenya: detection, insecticide resistance profiles and population genetic structure in relation to conspecific populations in West and Central Africa. *Malar. J. Springer*. **23**, 122 (2024).
28. Zhou, G., Zhong, D., Yewhalaw, D. & Yan, G. *Anopheles stephensi* ecology and control in Africa. *Trends Parasitol. Elsevier*. **40**, 102–105 (2024).
29. Yared, S. et al. *Anopheles stephensi* larval habitat superproductivity and its relevance for larval source management in Ethiopia. *Malar. J. Springer*. **24**, 357 (2025).
30. Imbahale, S. S. et al. A longitudinal study on *Anopheles* mosquito larval abundance in distinct geographical and environmental settings in western Kenya. *Malar. J. Springer*. **10**, 81 (2011).
31. Ondiba, I. M., Oyieke, F. A., Athinya, D. K., Nyamongo, I. K. & Estambale, B. B. A. Larval species diversity, seasonal occurrence and larval habitat preference of mosquitoes transmitting Rift Valley fever and malaria in Baringo County, Kenya. *Parasit. Vectors Springer*. **12**, 295 (2019).
32. Agyekum, T. P. et al. A systematic review of the effects of temperature on *Anopheles* mosquito development and survival: implications for malaria control in a future warmer climate. *Int. J. Environ. Res. Public. Health MDPI*. **18**, 7255 (2021).
33. Taylor, R. et al. Invasive *Anopheles stephensi* in Africa: insights from Asia. *Trends Parasitol. Elsevier*. **40**, 731–743 (2024).
34. Carter, T. E. et al. *Tracking of Anopheles stephensi in Ethiopia using mitochondrial DNA reveals pattern of spread*. *BioRxiv* 2021–2024 (Cold Spring Harbor Laboratory, 2021).
35. De Marco, C. M. et al. Habitat drivers and predicted distribution shifts of *Anopheles coluzzii* under climate change: Results from the systematic review. *Sci. Total Environ. Elsevier*. **992**, 179939 (2025).
36. Abong'o, B. et al. Evaluation of community-based vector surveillance system for routine entomological monitoring under low malaria vector densities and high bed net coverage in western Kenya. *Malar. J. Springer*. **22**, 203 (2023).

Figures

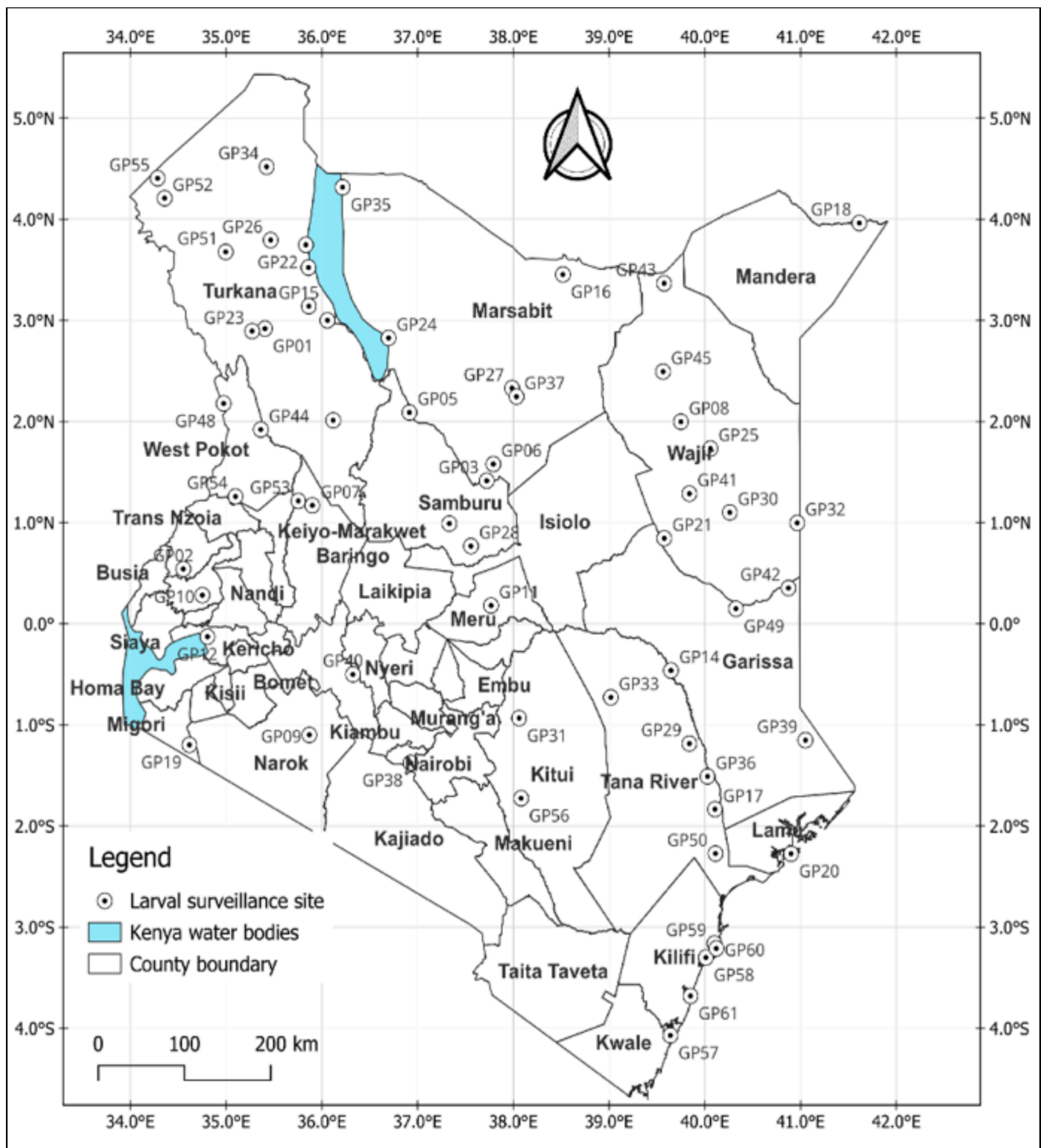


Figure 1

Sampling sites across Kenya. Larval habitat surveillance was conducted within a 2km radius of each of the 61 sites (GP1-61).

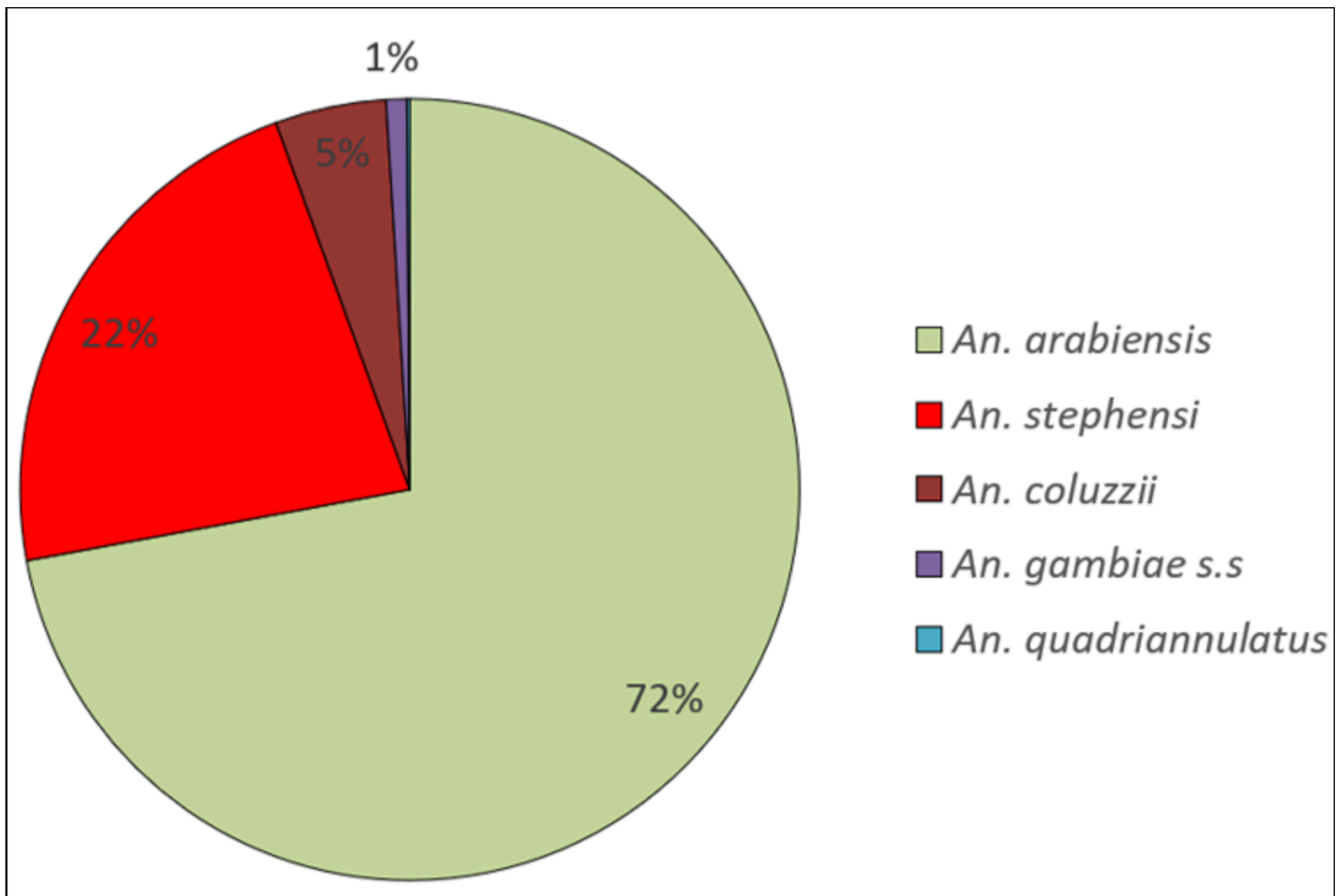


Figure 2

The abundance of different *Anopheles* species using PCR identification shows *An. arabiensis* as being the dominant vector across the study sites.

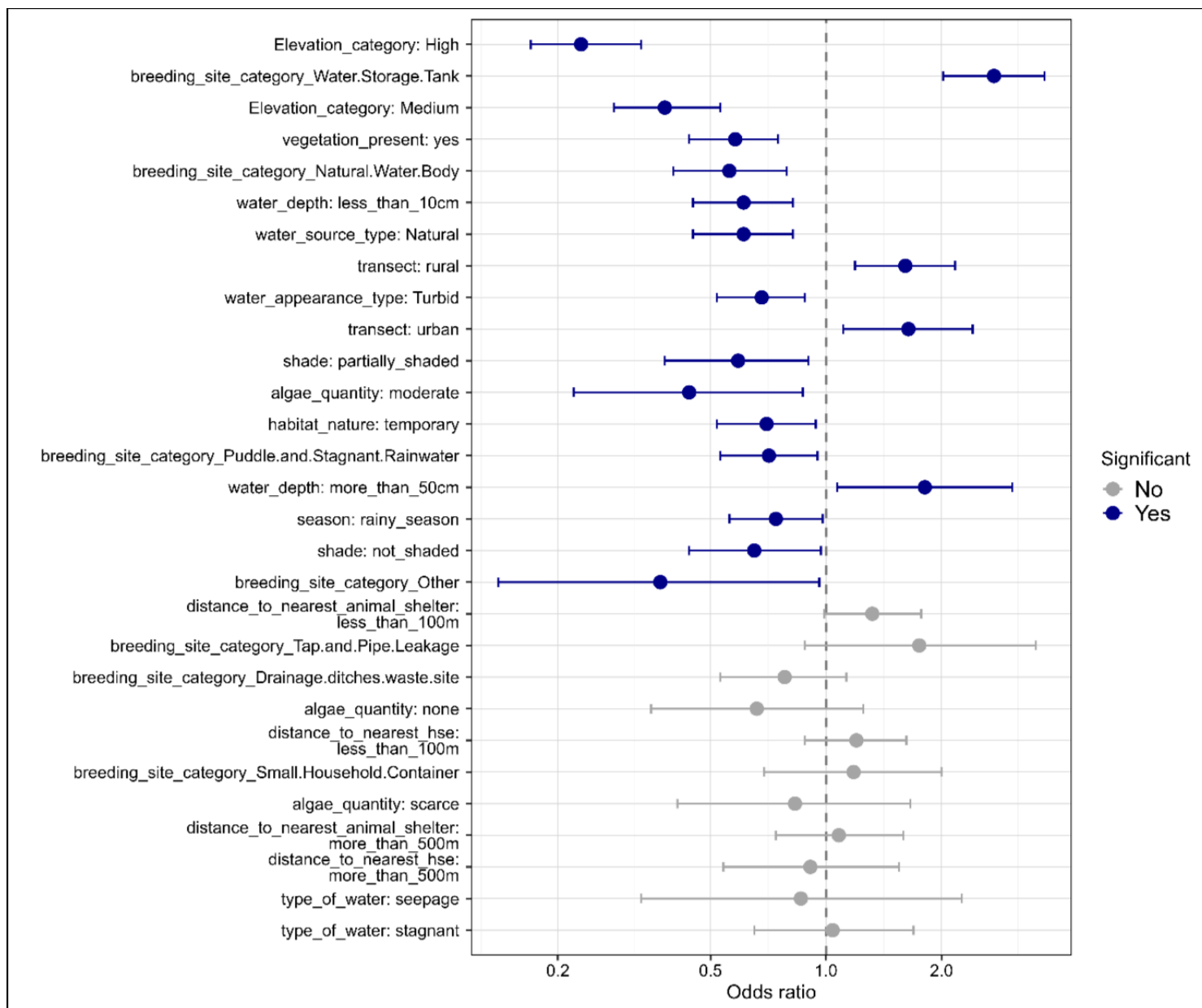


Figure 4

Odds ratios with 95% confidence intervals from univariate logistic regression analyses of each recorded habitat characteristic and its association with *An. stephensi* positivity. Blue indicates a statistically significant association ($P < 0.05$).

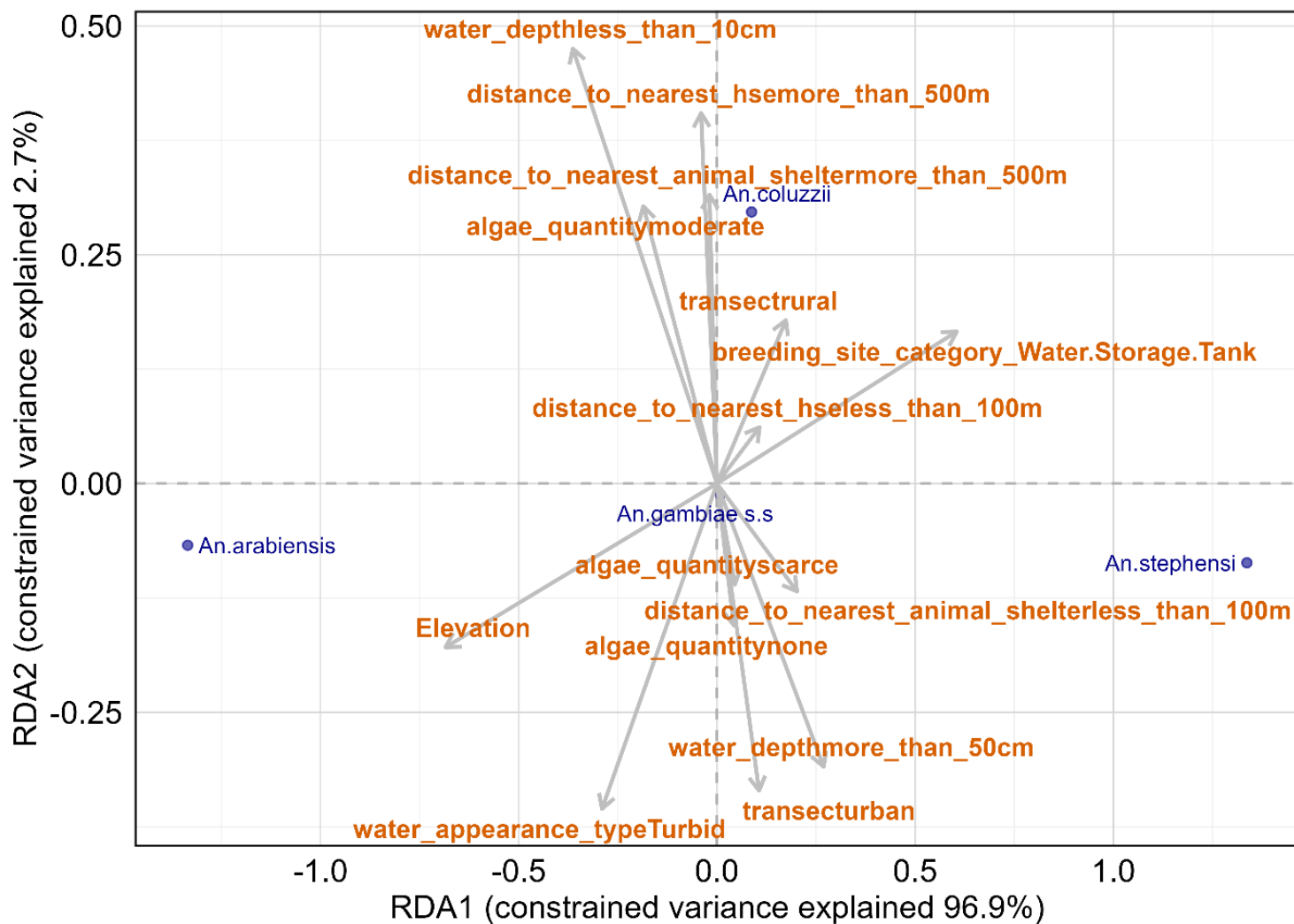


Figure 5

RDA biplot with niche partitioning along the first and second canonical axes. The greatest proportion of constrained species community variation was explained by RDA 1 (96.9%).

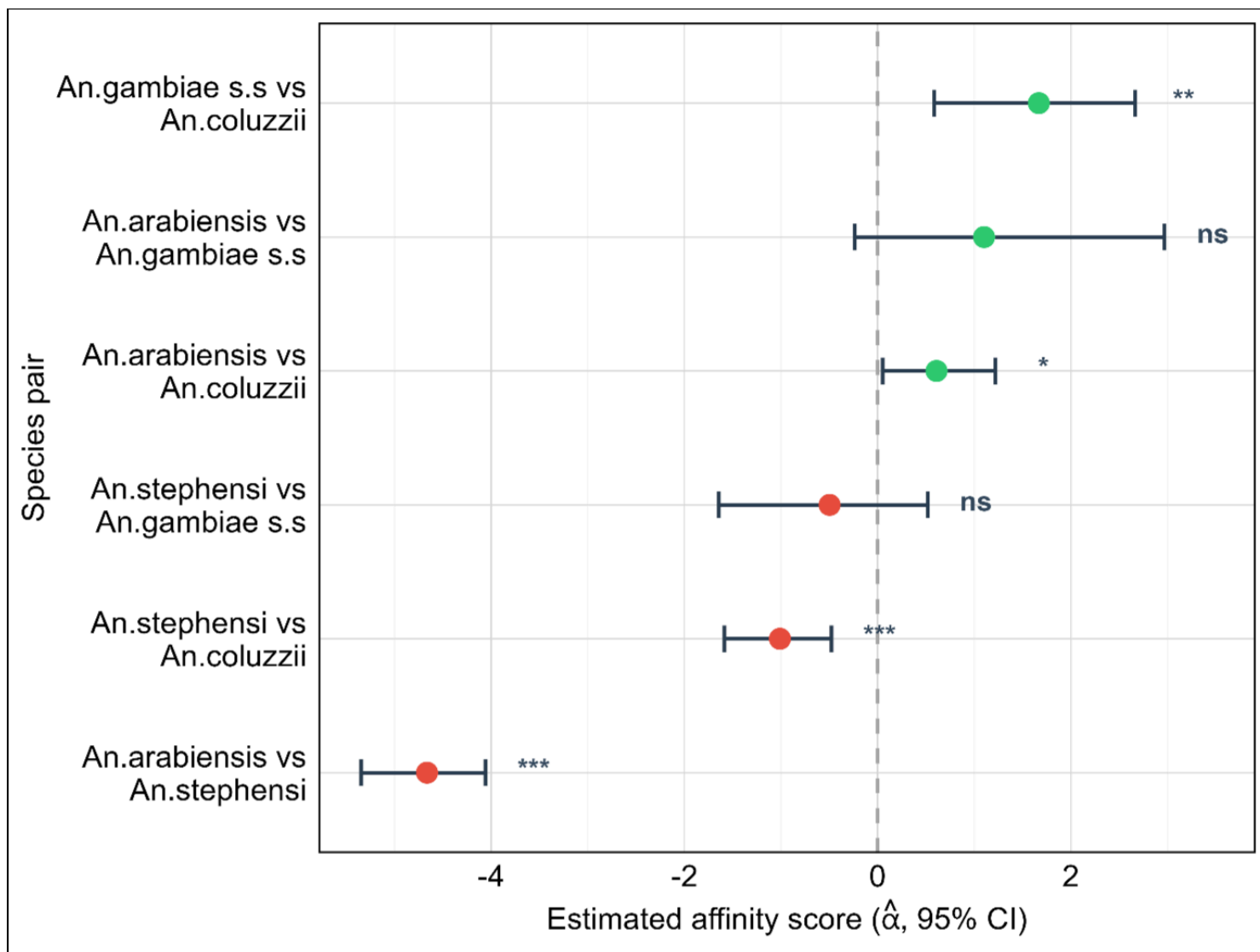


Figure 6

$\hat{\alpha}$ and 95% mid-P confidence intervals for species-pair habitat associations. Associated P-values are denoted by stars: $P < 0.001 = \text{"***"}$, $P < 0.01 = \text{"**"}$, $P < 0.05 = \text{"*"}$. Green points represent positive co-occurrence; red points represent negative co-occurrence between species pairs.

Supplementary Files

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

- [SupplementaryTable1.docx](#)