Landsc Ecol (2025) 40:49
https://doi.org/10.1007/s10980-025-02059-3

RESEARCH ARTICLE

Larval habitat suitability and landscape genetics of the mosquito *Anopheles coluzzii* on São Tomé and Príncipe islands

Lisa Chamberland · Melina Campos · Maria Corrêa · João Pinto · Anthony J. Cornel · João Viegas · Gregory C. Lanzaro

Received: 5 November 2024 / Accepted: 6 February 2025 / Published online: 21 February 2025

Abstract

Context This study was conducted to contribute to the design of a field trial of a novel genetic strategy aimed at the elimination of malaria. The strategy involves the introduction, establishment and spread of a gene construct into natural populations of the malaria vector *Anopheles coluzzii* on the African islands of São Tomé and Príncipe (STP). The gene construct renders the mosquito incapable of transmitting the parasite. Understanding the ecology of this mosquito is an essential component of the STP trial design. Identifying landscape features that define the target mosquito's distribution, understanding connectivity among subpopulations and estimating population stability in the face of climate change are critical factors contributing to the field trial.

Objectives STP provides an ideal study site to isolate and identify the role of potentially influential environmental factors in mosquito vector distributions across heterogeneous landscapes, critical information for the design of a GEM field trial. In this study we aim to quantify the relative influence of biotic and abiotic environmental factors on *Anopheles coluzzii* larval habitat suitability and if environmental variables promote or restrict gene flow between local populations.

Methods We used an ecological niche modeling (ENM) approach to test the relationship between environmental variables and *A. coluzzii* larval occurrences within the islands of STP. We implemented high-resolution spatial models of both current and future larval distributions under a range of climate change scenarios. We assessed functional connectivity of *A. coluzzii* in STP using circuit theory-based approaches to identify environmental variables impeding or promoting gene flow.

Results Results from the ENM revealed higher habitat suitability in the northeastern regions of both islands, characterized by higher human population densities and lower elevations. Habitat suitability under future climate projection models predicted minimal range expansions on STP, even under a

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10980-025-02059-3.

L. Chamberland · M. Campos · M. Corrêa · J. Pinto · A. J. Cornel · G. C. Lanzaro (✉)
Vector Genetics Laboratory, University of California Davis, Davis, CA, USA
e-mail: gclanzaro@ucdavis.edu

J. Pinto
Global Health and Tropical Medicine, GHTM, LA-REAL, Instituto de Higiene E Medicina Tropical (IHMT), Universidade NOVA de Lisboa, Rua da Junqueira 100, 1349-008 Lisbon, Portugal

A. J. Cornel
Department of Entomology and Nematology, Mosquito Control Research Laboratory, UC Davis, Parlier, CA, USA

J. Viegas
Programa Nacional de Eliminação Do Paludismo, Centro Nacional de Endemias, São Tomé, São Tomé and Príncipe

'business-as-usual' model. There was a signal of isolation-by-resistance on São Tomé, with roads promoting gene flow and higher elevation restricting gene flow.

Conclusions This study provides a clearer understanding of the role of climate, topography and human activity on *A. coluzzii* larval habitat suitability and underscores the importance of considering both current and future climate projections to establish robust baseline data. Additionally, the role of roads in facilitating gene flow of this species will not only be essential to the design of GEM field trials, but will contribute to our understanding of malaria epidemiology in the islands and to improving ongoing vector control programs.

Keywords Ensemble model · Random forests · Climate change · Circuit theory · Isolation by resistance · Genetically modified mosquito

Introduction

Novel strategies for the control of malaria are currently under development. One of the most promising involves the release of genetically engineered mosquitoes (GEM) which employ a gene drive (Wang et al. 2021). The GEM under consideration in this paper contains a genetic construct that includes two anti-*Plasmodium* effector genes. These express single chain antibodies that destroy the malaria parasite at two different stages as it develops in the mosquito vector's body. The transgenes do not harm the GEM, at least in the laboratory setting (Carballar-Lejarazú et al. 2020). This approach is referred to as a population modification strategy. The goal is to modify natural vector populations in such a way that nearly one hundred percent of individuals in the target location carry the transgene and are therefore incapable of transmitting the malaria parasite. The population modification strategy targets the parasite, not the mosquito. An alternative approach, known as the population suppression strategy, utilizes a GEM that has, as its goal, the elimination of the mosquito vector; this strategy is not considered in this paper.

In order to establish and spread these effector genes following release, the genes are linked to a gene drive. The gene drive alters the inheritance of the transgene construct favoring its transmission to the next generation (Bier & Nizet 2021). Consequently, the transgene increases in frequency toward fixation following introduction into a wild-type target population. In addition, the transgene is expected to spread to subpopulations via normal mosquito dispersal and, due to the drive, once introduced it will increase in frequency toward fixation.

A GEM for population modification has been developed and laboratory evaluations suggest this approach can provide an effective, sustainable and cost-effective tool for the elimination of malaria (Carballar-Lejarazú et al. 2023). At this point the GEM is ready to be evaluated in a natural setting as a WHO Phase 2 field trial (WHO/TDR and FNIH 2014). Phase 2 field trials are intended to measure entomological endpoints including GEM survival, mating competitiveness, gene drive spread and construct stability; they do not involve measuring epidemiological outcomes, e.g. effects on malaria cases. Careful review of potential sites for a trial have been conducted and the islands of São Tomé and Príncipe (STP) have been identified as most suitable (Lanzaro et al. 2021). Criteria include geographic and genetic isolation, strong boundaries and anopheline mosquito diversity. Field trials involving the release of GEM in STP are currently being designed. These trials will target the mosquito, *Anopheles coluzzii,* which is a major malaria vector throughout Africa and is the sole vector of malaria in STP. Prior to the release of GEM into the environment of STP, baseline data is needed to plan both the release and post-release surveillance. Description of habitat suitability and landscape scale as they relate to *A. coluzzii* distribution in STP is essential to the development of an informed design plan for a GEM release. Initially this information will be used to determine the number, location, and timing of GEM releases to optimize prospects for GEM establishment and spread. Information on population connectivity will inform a post-release surveillance regime to assess efficacy of the drive by monitoring the spread of the transgene into non-release subpopulations within the islands. One of the advantages of the population modification strategy is that it does not remove the GEM mosquito population. Therefore, any wild-type individual that migrates into a GEM population will acquire the anti-parasite construct when it mates with a GEM. Conversely, when a GEM migrates into a wild-type population it serves as a "seed," introducing the transgene

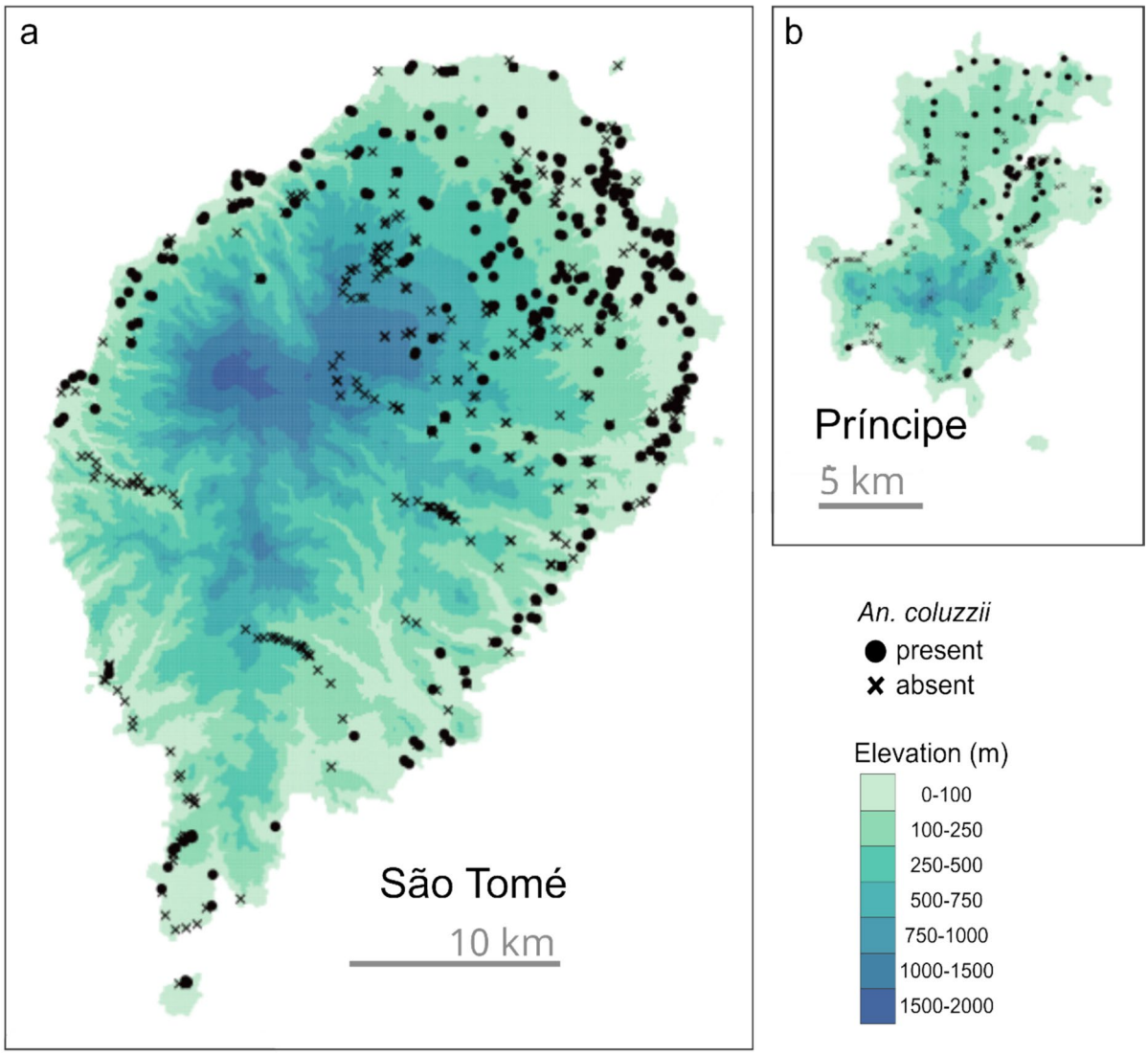


Fig. 1 Sampling localities of *A. coluzzii* larvae on **a** São Tomé and **b** Príncipe (2019–2024)

which will rapidly increase in frequency due to the drive. This is critical for both sustainability and cost because it reduces the number of releases required to effect the desired changes over a larger area. The size of this area depends on the extent of mosquito dispersal and landscape features that limit it. Evaluating the stability of target mosquito populations over both the short and long terms, for example in the face of climate change, is another important component of the field trial design.

Geologic history and regional ecosystems of São Tomé and Príncipe

The islands of STP are located in the Gulf of Guinea approximately 225 km from each other. These two islands are part of the Cameroon Volcanic Line, a volcanic mountain range comprising oceanic and continental peaks. This ~ 1600 km long chain of volcanoes spans from the tiny island of Annobón (area: ~ 17 km^2) to Lake Chad, partially transecting central Africa through Cameroon and Nigeria (Dedzo et al. 2015). Since emerging around 30 million years

Table 1 Environmental covariates used in the ENM

Classification	Covariate	Description	Source	Citation
Topography	Elevation	Digital Elevation Model (DEM)	U.S. Geological Survey	Verdin (2017)
	Slope	Computed from DEM using R package: *terra*	–	Hijmans et al. (2022)
Human Activity	Roads	Distance from R package: *osmdata* roads computed using R package: *qgis_process*	OpenStreetMap	Ooms (2023), Bivand and Mächler (2022)
	Buildings	Distance from building polygons computed using R package: *qgisprocess*	Google Research: Open Buildings	Dooley et al. (2020)
	Population	Population density	WorldPop	WorldPop (2021)
Ecoregions	Ecoregions[a]	Regional-scale ecoregions based on abiotic gradients (elevation, precipitation, distance to coast, and cloud frequency)	Biodiversity of the Gulf of Guinea Islands	Dauby et al. (2022)
Climate	Temperature Seasonality	Bioclimatic variable BIO4 (standard deviation × 100)	WorldClim2	Fick and Hijmans (2017)
	Precipitation Seasonality	Bioclimatic variable BIO15 (coefficient of variation)	WorldClim2	Fick and Hijmans (2017)
	Precipitation in coldest quarter	Bioclimatic variable BIO19 (Precipitation in the coldest quarter)	WorldClim2	Fick and Hijmans (2017)

[a]Categorical dataset

ago, the geologic histories of STP have been dynamic and diverse. For example, the islands did not reach their present-day topography until after the Last Glacial Maximum (LGM), approximately 7500 years ago (Ceríaco et al. 2020). During the LGM, sea levels were low, exposing more land surface on both islands and the coastal regions of Africa. As a result, a land-bridge once connected the continental shelf island of Bioko to mainland Africa as evidenced by both geologic (Ryan et al. 2009; Ali 2018) and biological (Blackburn 2010; Wagner et al. 2014; Wrobel 2006) data. Conversely, São Tomé, Príncipe, and Annobón were never connected to the mainland. Prior to human arrival, they had been primarily colonized by high dispersal ability and/or salt water tolerant plants and animals including birds (Peet and Atkinson 1994; Jones 1994; de Lima & Melo 2021), plants (Figueiredo et al. 1994; Stévart et al. 2022), frogs (Bell et al. 2015), and bats (Juste and Ibáñez 1994; Juste et al. 2023). Consequently, the islands are rich with endemic species (Jones 1994; Le Saout et al. 2013). Príncipe, the oldest of the Gulf of Guinea islands (~ 31 Ma), was designated a UNESCO Biosphere reserve in 2012 (Neto & Henriques 2023).

These young volcanic islands host impressive peaks that rise to elevations of 2024 m (Pico de São Tomé) and 942 m (Pico Príncipe) in the southern and interior portions of São Tomé and Príncipe respectively. The northeastern and coastal portions of each island taper off into flat regions with elevations at, or slightly above, sea level. *A. coluzzii* are generally found in coastal areas and where elevation is low and terrain less rugged (Bigoga et al. 2007; Lee et al. 2009; Simard et al. 2009; Tene Fossog et al. 2015); areas which are concentrated around the northeastern regions of both islands. We predict that these peaks will serve as important barriers to *A. coluzzii* movement on STP. Furthermore, although minute in size, São Tomé (50 km long, 30 km wide) and Príncipe (14 km long, 6 km wide) are mosaics of regional-scale ecosystems (Dauby et al. 2022; Senterre et al. 2021). Ranging from lowland rainforests to mountain grasslands, this diversity is, in-part, maintained by the physical geography of the islands. Dramatic elevation gradients play an important role in generating

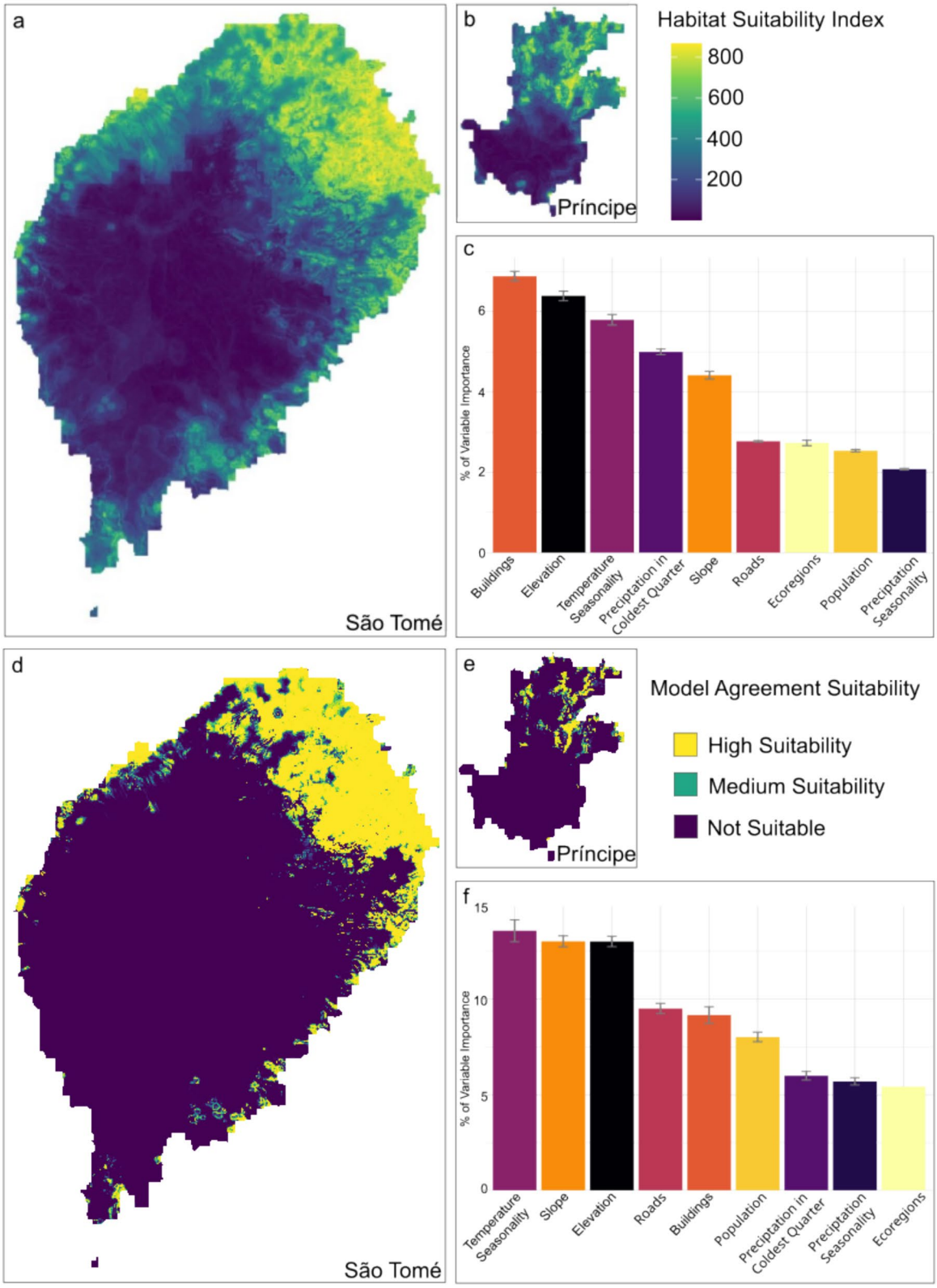
variable local and temporal climates throughout the islands. High relief intercepts, or mountain peaks, on STP coupled with prevailing, moist, south-westerly winds, create a rain-shadow on both islands (Bredero et al. 1977; Diniz & Matos 2002). As a result, both humidity and rainfall tend to be higher in the south-southwest. In addition to rainfall, these ecosystems are influenced by other abiotic factors including elevation, distance to the coast, and cloud frequency. The equatorial climate ensures year-round breeding of *A. coluzzii* on STP and they do not undergo aestivation, or dormancy during dry periods (Charlwood 2019). Still, STP experiences two distinct rainy seasons a year, a long humid season (mid-March to June) and a shorter rainy season (mid-September to mid-December) (Dauby et al. 2022), which corresponds with increased mosquito densities and consequently malaria transmission (Pinto et al. 2000; Lee et al. 2010) and even insecticide resistance (Correa et al. 2024). Although an integrated malaria control program on STP, including indoor residual spraying, insecticide nets, and community engagement has significantly reduced malaria prevalence on STP (from 37% in 2004 to 1% in 2018,) these low-transmission areas can be highly sensitive to seasonal increases in vector densities (Lee et al. 2010).

The role of humans in mosquito diversity and abundances on STP

The islands of São Tomé and Príncipe have a recorded 29 and 15 mosquito species respectively, that includes a remarkably high number of endemic mosquito species, (Loiseau et al. 2019). Both islands have a single primary human malaria vector, *A. coluzzii* (Pinto et al. 2000; 2002). The only other *Anopheles* species in STP, *Anopheles coustani*, a secondary malaria vector (Afrane et al. 2016), occurs only in São Tomé, but with a scattered distribution (Pinto et al. 2000). Malaria was introduced into STP with the arrival of Portuguese ships approximately 500 years ago (Baptista 1996). Sequence analysis of the *A. coluzzii* mitochondrial genome suggests that it arrived in São Tomé Island via two independent introductions at about the same time (Ditter et al. 2022) and that these originated from coastal west Africa (Campos et al. 2021). Introduction into the island of Príncipe appears to have originated from São Tomé Island and not directly from the mainland. Genetic analyses indicate little or no contemporary dispersal between the two islands or between them and the mainland (Campos et al. 2021; Ditter et al. 2022). Within the islands, *A. coluzzii* on STP have a flight range of around 3 to 7 km (Campos et al. 2024).

In addition to its initial introduction, humans continue to play an important role in the maintenance of *A. coluzzii* populations on the islands. The interconnectedness between people and anopheline mosquitoes can be attributed to two requirements in the mosquito lifecycle. First, natural and artificial standing water for their aquatic immature stages and second, sources of mammalian blood, required by adult females for successful egg development. *A. coluzzii* are anthropophilic and zoophilic, meaning they will blood feed on humans and other mammals. As there are relatively few native mammals in STP, humans and their domestic animals, including pigs and dogs, are the main sources of blood for *A. coluzzii* on the islands (Sousa et al. 2001). Volcanic soils, coupled with an equatorial climate and an elevational gradient have generated a landscape rich with opportunities for agriculture in these areas and as agricultural activities increased, so did populations of *A. coluzzii* (White et al. 2011; Gimonneau et al. 2012). These highly synanthropic malaria vectors thrive in environments altered by human activity, including agricultural fields, which often have direct or indirect association with human settlements, urban areas, and semi-urban settlements. *A. coluzzii* is a nocturnal species, with peak activity occurring during the night, and are endophagic, or prefer to eat indoors, and consequently, malaria transmission is often most effecting in regions where housing conditions allow close human-mosquito contact. Female *A. coluzzii* typically lay eggs in temporary and semi-permanent water bodies such as drainage ditches, puddles along dirt roads, and small ponds. In addition to providing potentially suitable habitats for *A. coluzzii*, we predict that humans may also play a role in mosquito movement providing blood meals and breeding sites along roads connecting the fragmented village populations.

Prior to a release of GEM, it is critical to identify distribution and dispersal of existing populations of *A. coluzzii* on STP in order to identify potential GEM release sites and optimize release strategies. Furthermore, predicting potential climate change-driven range shifts and assessing range stability are essential for the long-term management and evaluation of


a
São Tomé
b
Príncipe
Habitat Suitability Index
800
600
400
200
c
% of Variable Importance
0
2
4
6
Buildings
Elevation
Temperature Seasonality
Precipitation in Coldest Quarter
Slope
Roads
Ecoregions
Population
Precipitation Seasonality
d
São Tomé
e
Príncipe
Model Agreement Suitability
High Suitability
Medium Suitability
Not Suitable
f
% of Variable Importance
0
5
10
15
Temperature Seasonality
Slope
Elevation
Roads
Buildings
Population
Precipitation in Coldest Quarter
Precipitation Seasonality
Ecoregions

◂**Fig. 2** Random Forest ensemble model of larval habitat suitability of *A. coluzzii* larvae in São Tomé and Príncipe using the full dataset (9 covariates). Models were built using a **a–c** weighted mean (EMwmean) and **d–g** committee averaging (EMca) algorithm. Inset plots **c**, **f** indicate the variable importance in each model

GEM in this region. We aimed to assess how topography, climate, and anthropogenic influences affect the larval habitat suitability and landscape genetics of *A. coluzzii* on STP. This study employs an ecological niche modeling (ENM) approach to model contemporary and future projections of larval habitat suitability across STP and identify key environmental predictors. Given the dispersal capacities of mosquitoes and small size of the islands, we predicted that genetic structuring would not be characterized by isolation-by-distance alone; instead, we hypothesized an isolation-by-resistance pattern, where roads would promote gene flow, and higher elevations would restrict it.

Materials and methods

Ecological niche modeling

Sampling

Anopheles coluzzii larvae were sampled from a variety of potential aquatic breeding sites from 2019 to 2024 (Fig. 1). GPS coordinates were recorded, and sites were marked as absent or present for *A. coluzzii* larvae. Larvae were sampled using standard methods (Silver 2007) and total counts up to a maximum of 20 larvae were recorded for each site. Relative abundances and highest density regions (HDR) of kernel density estimation of STP islands based on the larval abundance at each coordinate were modeled using the *geom_density* density function in *ggdensity* (Wickham 2016). For the ENM, sites were classified as presence (≥ 1 larva) and absence (zero larvae). In order to avoid overfitting and model inflation due to spatial autocorrelation (Boria et al. 2014), these points were filtered so that there was one occurrence point per 10 m by 10 m raster grid cell using the R package *fuzzySim* (Kramer-Schadt et al. 2013; Barbosa 2015; Wieringa et al. 2021). Cells that had both absence and presence data recorded were filtered out so that only 'true' absences remained to reduce false negatives in the dataset.

Environmental variables

Inclusion or exclusion of explanatory variables can have an impact on model output and performance (Muscatello et al. 2021; Amiri et al. 2020). In order to reduce noise and bias and promote reproducibility in the models (Feng et al. 2019), we set criteria for selecting the environmental covariates. First, the covariate needed to be biologically and ecologically relevant to the focal species (Guisan and Zimmermann 2000; see Kujala et al. 2015). Second, the dataset needed to be both informative (i.e. have a relatively fine resolution) and accurate at the small spatial extent of STP. Third, open-source datasets were prioritized and ultimately selected. Fourth, highly correlated variables were excluded. For this, we computed the variance inflation factor (VIF) using a stepwise process in the R package *usdm* (Naimi 2017). Due to the mountains, equatorial climate, and small spatial dimension of the study area, variables with a VIF > 3 were removed to eliminate multicollinearity in the final dataset (Naimi & Araújo 2016, Table S1 of the of the Online Resource 1). Lastly, to avoid building a noisy model with poor predictive power, we also excluded variables that had low importance in initial model developments.

We included four covariate classes: topography, climate, human activity, and ecoregions (Table 1). Using the digital elevation model Africa Land Surface Forms (DEM; Verdin 2017), we were able to extract elevation and compute slope using the R package *terra* (Hijmans et al. 2022). The former of which is an important limiting factor in *A. coluzzii* distributions (Adeogun et al. 2023), and the latter an important factor in maintaining standing water breeding sites. We extracted climate variables including precipitation, and minimum and maximum temperature from WorldClim version 2 at a spatial resolution of 30 arc seconds (1 km2) (Fick and Hijmans 2017). The R package *dismo* was used to extract Bioclimatic variables (Hijmans et al. 2024). Three climate variables remained after VIF filtering including temperature seasonality (BIO4), precipitation seasonality (BIO15), and precipitation in the coldest quarter (BIO19). We used a regional-scale ecosystem dataset from Dauby et al. 2022, which is based on abiotic

gradients of São Tomé including elevation, precipitation, distance to the coast, and cloud frequency to serve as a more accurate and precise measure of precipitation. Finally, given that *A. coluzzii* are synanthropic, we included landscape composition variables, including human population indices from the open-source global database WorldPop (WorldPop 2021). We also included buildings from Open Building dataset from Google Research (Dooley et al. 2020) and roads and paths from OpenStreetMap (Ooms 2023) using the R package *osmdata* (Padgham & Lovelace 2017). We computed distances from both roads and buildings using the r package *qgisprocess.*

Model calibration, evaluation, and fitting

Performance of seven different algorithms appropriate for presence/absence data were evaluated using the R package *Biomod2* version 4.2–5 (Thuiller et al. 2024). We assessed performance across regression algorithms, including general linear regression (GLM; McCullagh 1989), generalized additive regression (GAM; Hastie & Tibshirani 1990) and multivariate adaptive regression splines (MARS; Hastie et al. 2009). Performance of machine learning algorithms, including Random Forest (RF; Brieman 2001; , 2002; Elith et al. 2006), generalized boosting model (GBM; Ridgeway 1999, 2007), classification tree analysis (CTA; Breiman 1984) and Artificial Neural Network (ANN; McCulloch & Pitts 1943; Metropolis et al. 1953) were also evaluated. Model performance was assessed using the receiving operating characteristic curve (ROC; Fielding & Bell 1997) often referred to as the AUC (area under the curve) and the True Sensitivity Statistic (TSS; Allouche et al. 2006), which are both measures of model sensitivity (how well the model accurately identifies true presence sites) and specificity (how well the model accurately identifies true absence sites). We employed commonly accepted AUC thresholds to assess model performance (Bebber et al. 2019; Table S2 of the Online Resource 1). Six of the seven algorithms performed moderate to well, and one, ANN, failed under an ROC threshold of 0.8.

Ensemble models identified the RF algorithm as the best performing model (Fig. S1 of the Online Resource 1). In ecological problems, machine learning algorithms, including RFs, have been shown to perform better than simple models when handling complex datasets (De'ath and Fabricius 2000; Wagner and Fortin 2005; Cutler et al. 2007). K-fold cross-validation (k = 10) was employed to train models using 80 percent and 20 percent of the data for training and testing respectively. Single RF model predictions were combined using an ensemble modeling approach, which computes the average model prediction across all models and has been shown to outperform single models and improve prediction accuracy, especially with complex, ecological data (Araújo & New 2007; Marmion et al. 2009). We built ensemble models in *Biomod2* using averaging algorithms, including (1) the weighted mean (EMwmean) algorithm, which computes model averages and assigns weights to models based on model runs, and (2) committee averaging (EMca), which uses voting to summarize areas in which models are in high agreement. When the model is close to 1, this indicates high agreement or concordance among models of predicted presence points. Alternatively, when the values are close to 0, this indicates high agreement among models for absences localities. Single models with an ROC of < 0.8 were excluded from the ensemble model building. Finally, although there is a substantial rainfall gradient on both islands, remotely sensed data for precipitation is dubious on STP (Chou et al. 2020). Therefore, we calibrated models with and without precipitation data and compared evaluation scores following model runs and assessed model evaluation scores. Lastly, we assessed the relative contribution (importance) of the environmental covariates in *Biomod2*. The R package *ggplot2* (Wickham 2016) was used to generate barplots and maps using color pallets from the R package *viridisLite* (Garnier et al. 2023). Figures were finalized using Inkscape v1.4 (Bah 2011).

Landscape genetics

We performed resistance surface optimization to determine which individual resistance surface or combination of resistance surfaces best described the genetic structure of *A. coluzzii* on São Tomé. Initial model runs of Príncipe failed to converge due to lack of genetic structuring. This agreed with a recent genetic SNP dataset analysis indicating that *A. coluzzii* on this island exists as a single Mendelian population (Campos et al. 2024), and thus we excluded Príncipe from landscape genetic analyses.

Table 2 Single resistance surface optimization

Surface	AICc	ΔAICc	Marginal R^2	Conditional R^2	LL	k
Roads	−2736.37	0	0.396432	0.685829	1372.924	4
Elevation	−2731.96	−4.41	0.376898	0.66116	1370.72	4
Distance	−2706.16	−25.8	0.291926	0.624264	1355.287	2
Null	−2608.68	−97.48	0	0.438952	1305.409	1

Surfaces are ranked by Akaike Information Criterion corrected for sample size (AICc). Marginal R^2 explains the proportion of variance in the response variable that is explained by the fixed effects. Conditional R^2 explains the proportion of variance explained by both the fixed and random effects. Log-likelihood values and the number of parameters (k) are included

We used a maximum likelihood population effects (MLPE; Clarke et al. 2002) approach with the package *ResistanceGA* v4.0–14 (Peterman 2018; https://github.com/wpeterman/ResistanceGA) which has been shown to be more effective in landscape genetic analyses (Peterman et al. 2019). Using linearized F_{ST} genetic distances from 37 populations (Campos et al. 2024), single surface optimization was performed with CIRCUITSCAPE v4.0.5 (McRae et al. 2008; Anantharaman et al. 2019; Hall et al. 2021) in ResistanceGA. We included two resistance surfaces: (1) roads, in which resistance increases as distance from roads increases and (2) elevation. ResistanceGA also includes a null model and isolation-by-distance model. Single resistance surfaces were ranked by the Akaike Information Criterion corrected for finite sample size (AICc; Akaike; 1998) and resistance surfaces with a ΔAICc < 2 were considered indistinguishable. We ran multivariate optimization to assess whether a combination of roads and elevation could better explain genetic structuring compared to the individual resistance surfaces. Finally, we ran a 1000 bootstrap analysis in *ResistanceGA* subsampling 75% of sampling sites to compute the Akaike weight, which computes the relative likelihood of each resistance surface performing the best across single and multivariate surfaces.

Future projections

Future larval habitat suitability projections were built using Coupled Model Intercomparison Project (CMiP) under Shared Socio-economic Pathways (SSP) (Eyring et al. 2016) available on WorldClim 2. We used MPI-ESM1-2-HR (high resolution) future climate datasets with a resolution of 30 s (Mauritsen et al. 2019) because of the small spatial extent of the islands. Projections were built under an 'optimistic' scenario (SSP2 4.5) and 'pessimistic' or 'business-as-usual' (SSP5 8.5) scenario, which were models built under projected carbon dioxide emission data (O'Neill et al. 2016). We used the 'full' environmental covariate dataset to project larval distributions for two future time periods: 2021–2041 and 2041–2060 We computed the total areas and the percentage change of expansions, contractions, and shifts in *A. coluzzii* on STP in *Biomod2* by comparing current projections to future projections.

Results

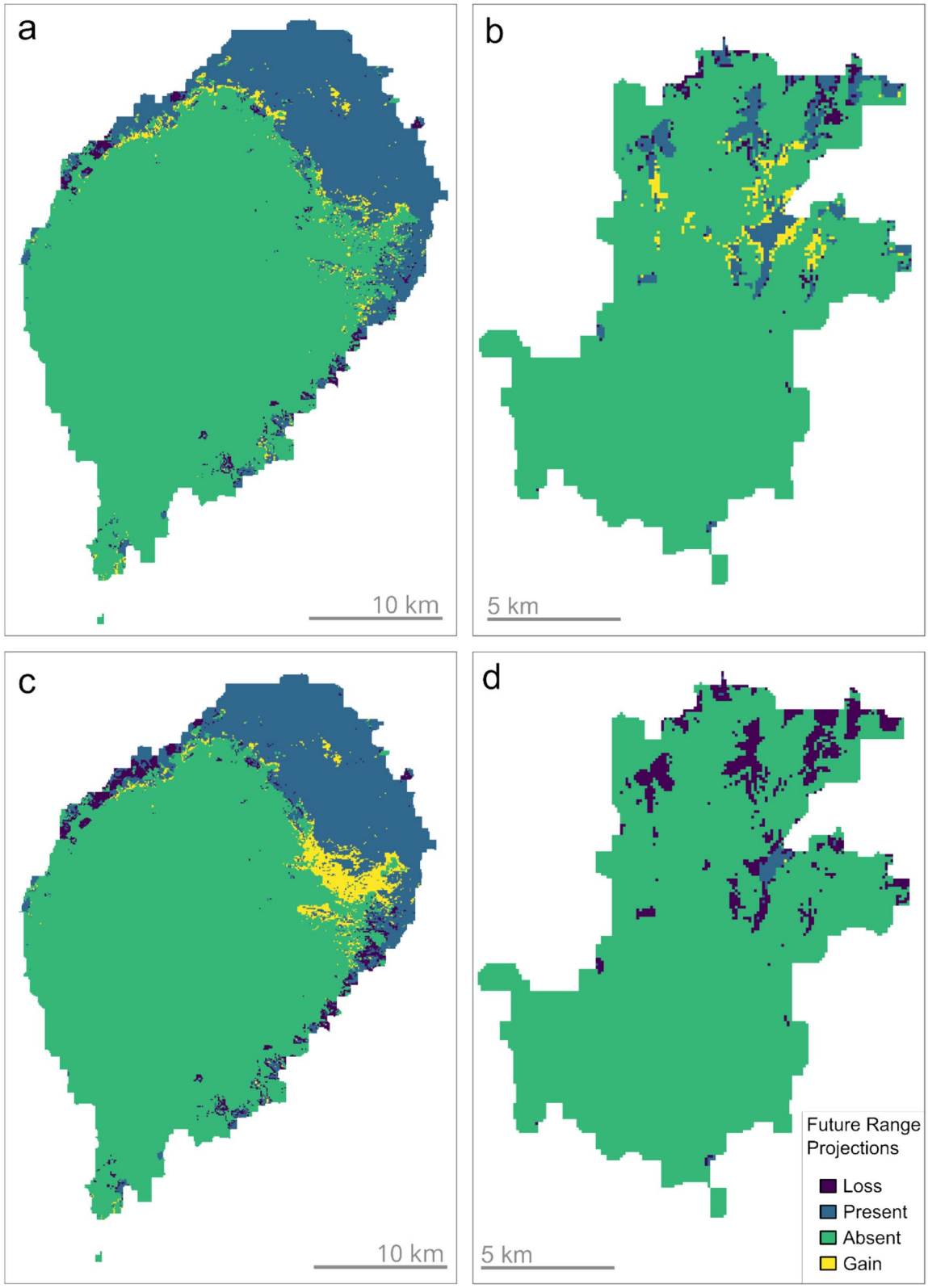
Larval habitat suitability

There were 429 and 71 present sites and 342 and 153 absent sites for larvae in São Tomé and Príncipe

Table 3 Bootstrap analysis of single and combined resistance surfaces

Surface	Average AIC weight	Average rank	% Top ranked	k
Roads	0.5175194232	1.6856	50.24	4
Elevation	0.4518486079	1.9048	45.78	4
Roads + Elevation	0.01315337186	2.4958	3.52	7
Distance	0.01747859709	3.9138	0.46	2

Average AIC weight and rank of each surface for 1000 bootstrap replicates. Percent each surface was selected as top ranked for per 1000 replicates. Surface parameters are included (k)


a
10 km
b
5 km
c
10 km
d
5 km
Future Range Projections
Loss
Present
Absent
Gain

◂**Fig. 3** Future range projections of *A. coluzzii* larval habitats under the SSP 8.5 model in **a** São Tomé for 2021–2040, **b** Príncipe 2021–2040, **c** São Tomé for 2041–2060, and **d** Príncipe 2041–2060. Color indicates future range loss, stable present, stable absent, and gain of *A. coluzzii* suitable habitat

respectively. (Fig. 1; Table S3 of the Online Resource 1). Both the 'no precipitation' and full models had high sensitivity and specificity (>0.90 ROC, TSS, and KAPPA), however, the full model performed better (Table S4 of the Online Resource 1). The EMca model had higher evaluation scores compared to EMwmean, although all scores were >0.90 (Table S4 of the Online Resource 1). Upon visual inspection larval density (Fig. S2 of the Online Resource 1), habitat suitability indices (Fig. 2a and b), and committee averaging agreement for high suitability (Fig. 2d and e), were highest in the northeastern regions of both STP islands and unsuitable regions in the interior and southwest regions of the islands. These northeastern regions are characterized by higher human population densities and lower elevation. Conversely, the interior of the islands, where the elevation is the highest and human population densities lowest, were unsuitable for larval habitat. For the full dataset, variables with the highest of mean importance (±standard deviation) included elevation (EMca = 13.00 ± 0.6%; EMwmean = 6.39 ± 0.9%), as well as temperature seasonality (EMca = 13.57 ± 1.28%; EMwmean = 5.80 ± 0.29%) and distance from buildings (EMca = 9.16 ± 0.97%; EMwmean = 6.88 ± 0.26%). Slope and distance from roads were ranked as one of the top important variables in the EMca model (slope = 13.02 ± 0.65%, roads = 9.51 ± 0.60%), but were lower ranked in the EMwmean model (slope = 4.41 ± 0.21%, roads = 2.77 ± 0.04%). Human population density (EMca = 8.02 ± 0.56%; EMwmean = 2.57 ± 0.0.06%), precipitation seasonality (EMca = 5.70 ± 0.41%; EMwmean = 2.07 ± 0.04%), precipitation in the coldest quarter (EMca = 6.00 ± 0.51%; EMwmean = 5.00 ± 0.16%), and ecoregions (EMca = 5.43 ± 0.57%; EMwmean = 2.73 ± 0.15%) were ranked lowest in importance (Fig. 2c and e; Table S5 of the Online Resource 1). Elevation, distance from buildings, and temperature seasonality also scored highest in the "without precipitation" model (Fig. S3 and Table S6 of the Online Resource 1).

Isolation-by-resistance

Continuous surfaces were transformed into resistance surfaces using an inverse-reverse transformation for distance from roads and a monomolecular transformation for elevation (Fig. S4 of the Online Resource 1). Results from the single resistance surface optimization ranked roads the top performing model in describing genetic structuring on São Tomé, followed by elevation (ΔAIC = −4.41; Table 2). Isolation-by-resistance was identified for both the transformed distance from roads surface (Fig. S5 of the Online Resource 1) and elevation surface (Fig. S6 of the Online Resource 1). Both resistance surfaces outperformed the isolation-by-distance (ranked third) and null (ranked last) models (Table 2). Bootstrapping analyses on single and combined surfaces also ranked roads as the top performing model (AICw = 0.52) selected for the top-ranking model 50.24% of the time and again elevation was ranked second (AICw = 0.45) (Table 3). Both the combined roads + elevation and distance models performed poorly with the lowest AIC weights and percent rankings (Table 3). The percent contribution of roads (63.39%) was greater than elevation (36.11%) on genetic structure (F_{ST} distances).

Future projections

Future range projections predicted no detectable range changes under an SSP 4.5 'optimistic' model and minimal range contractions, expansions, and shifts in both São Tomé and Príncipe under SSP 8.5 'business-as-usual' scenario (Fig. 3; Table 4). There is a minimal overall net gain in range (<15% of current range) in São Tomé for both the early and late time projections as well as relatively stable present and absent total areas. Instead of an extreme net loss or net gain of suitable habitat area, São Tomé is characterized by range shifts in both the early and late models, with a slight overall net gain in both (Table 4), which is most notable in the late projection, in which there is a shift towards the interior of the island (Fig. 3c). In Príncipe, there is a negligible range gain for both early and late models, and a ~ 74% and ~ 92% net loss projected in the early and late models respectively. Of the approximately 10 km^2 suitable larval habitat in Príncipe (EMca), 8% of the total island area, only 2.76 km^2 were

Table 4 Projected range expansions, contractions, and shifts in STP from 2021 to 2060

Island	Time period	Current present	Current absent	Stable present	Stable absent	Gain	Loss
São Tomé	2021–2040	196.37	637.13	190.47	629.64	7.5	5.9
	2041–2060	196.37	637.13	175.77	609.68	27.45	20.6
Príncipe	2021–2040	10.81	122.05	2.76	121.97	0.09	8.05
	2041–2060	10.81	122.05	0.08	122.04	0	10

Models were built under weighted mean (EMwmean) and committee (EMca) ensemble modeling algorithms. Area (km^2) on each island indicating a loss and gain of *A. coluzzii* occupancy, stable presence, stable absence. Total percent loss and gain are included as well as range change and current range size

predicted to be stable present for larval and 0.08 km^2 for the late model, which occurs in the only major populated city in the northeast of the island (Fig. 3d).

Discussion

The role of humans in mosquito habitat, movement, and connectivity

We found weak evidence for isolation-by-distance; however, results indicated that resistance outperformed distance in explaining *A. coluzzii* genetic structuring. Roads were most important in explaining genetic structuring, indicating roads facilitate connectivity among populations, while inversely, elevation is likely responsible for isolating populations. Consequently, human and associated *A. coluzzii* densities are highest in the northeast of the islands (Fig. S2 of the Online Resource 1).

Roads were identified as an important variable in promoting gene flow among populations of *A. coluzzii* in São Tomé. They were slightly more important in promoting gene flow compared to low elevations. Human activities associated with roads are well-known means of mosquito dispersal (Medley et al. 2015; Molina-Cruz et al. 2016; Sherpa et al. 2020). The roads are largely unpaved and so provide additional opportunities for breeding sites with puddles, especially since much of the island has dense forest cover. In STP, the southwesterly winds coupled with population density in the northeast make over-mountain dispersal from the northeast to the southwest unlikely. Instead, it is likely that dispersal into these sites may be facilitated by fisherman traveling by boat. This hypothesis would need to be explored in future studies with targeted monitoring of these seasonal villages.

The role of topography in mosquito habitat, movement, and connectivity

Elevation has been shown to play an important role in *A. coluzzii* distributions and population structuring. For example, continental populations of *A. coluzzii* are at the highest densities at elevations < 200 m in the Gulf of Guinea (Bigoga et al. 2007; Lee et al. 2009; Simard et al. 2009; Tene Fossog et al. 2015). Consequently, geographic barriers such as the Cameroon Mountain range have isolated populations of *A. coluzzii* (Pinto et al. 2013). Still, *A. coluzzii* has the capacity to exist at higher elevations, although previous studies found these high-elevation populations were restricted exclusively to urban areas. In this study, we found larvae in the village of São Nicolau on São Tomé Island at an elevation of 917 m, indicating this species has a capacity to occupy spaces across a broader range of elevations even in the absence of extensive human development (Cuamba et al 2006; Calzetta et al. 2008; Kamdem et al. 2012; Tene Fossog et al. 2015). Extreme topography and dense forests have played an important role in limiting human expansion and development on the islands, making future projections on humans challenging. Thus, one assumption in the climate projection model was that buildings, roads, and human densities all remained the same into the future. Future model calibrations may incorporate additional anthropogenic variables, especially since roads and buildings, were important in predicting habitat suitability, which is similar to other microhabitat studies (Kamdem et al. 2012), especially compared to continental scale distribution models of *A. coluzzii* where the signal is lost (Tene Fossog et al. 2015). One important limitation of this

study was the use of larvae for the ENM and genetic connectivity analysis. Relying solely on larvae for the ENM may underestimate the suitable habitat, as adult mosquitoes can occupy different areas when moving throughout the islands.

Current and future climate implications

Results from the ENM indicated that climate plays an important role in *A. coluzzii* larval habitat suitability on STP, which is surprising given the islands' mild equatorial climate. The highest habitat suitability was concentrated in the northeast of both islands, which is characterized by lower temperature and precipitation seasonality due to the island rain shadow. Despite the mild climate, temperature seasonality was an important predictor of habitat suitability across all models, which is consistent with other ENM of mosquitoes (Lippi et al. 2023). However, precipitation, which was identified as the second most important predictor variable across mosquito studies (Lippi et al. 2023), did not rank highly across all models in this study. In contrast to mainland sub-Saharan Africa, which experiences prolonged and extreme dry seasons during which *A. coluzzii* undergoes aestivation (Lehmann et al. 2010), STP's proximity to the Doldrums, where the Intertropical Convergence Zone (ITCZ) overlaps with the equator (Bredero et al. 1977), results in a milder climate with shorter, less severe dry seasons that support year-round larval habitat availability. Consequently, the relative importance of other environmental variables, including topography and anthropization, may be more important compared to regions within mainland Africa which experience greater weather variability.

Mosquito distributions were historically viewed to be static (Sallares 2002; Newfield 2017); however, recent analyses have found that range expansions into areas at increasing elevations are occurring rapidly, potentially due to shifting thermal limits, with anthropogenic factors as a significant contributor (Surendran et al. 2019; Carlson et al. 2023; Ryan et al. 2023). Although we expected to find similar shifts in *A. coluzzii* on STP, projections of larval habitat suitability indicated that range expansions and contractions were minimal on both São Tomé and Príncipe even under the "business-as-usual" climate model (SSP 8.5). This could be due to the small area of current suitable larval habitat in Príncipe (EMca), which is only around 10 km^2, approximately 8% of the total island area. Alternatively, late projections may be unable to effectively project on an area of such a small spatial extent. Still, on São Tomé, minimal shifts are projected in both early and late models. These results support STP as an ideal field site for GEM field trials due to the relative stability of larval habitat suitability, even under a 'pessimistic' climate scenario.

Conclusions

This work provides the most comprehensive georeferenced database of *A. coluzzii* larvae on STP to date. It provides insight into the role of environmental variables, including anthropization, in contemporary and predicted larval habitat suitability on STP and genetic structuring on the island of São Tomé. This was achieved by formally characterizing the dynamic interplay of environmental factors influencing the distribution of malaria vectors on STP. This information is critical in informing future vector control efforts, including identifying areas of potential occurrence to guide future surveillance efforts and also by identifying suitable areas for GEM release. The observed role of roads in promoting gene flow amongst São Tomé populations of *A. coluzzii* is important for supporting more effective and targeted vector management strategies to ensure mosquitoes are not transported through human activities. In the context of population replacement GEM release, roads would likely facilitate the spread of GEMs. Overall, this research provides essential baseline data for cataloging distributions and gene flow among existing populations of *A. coluzzii,* which is important to improving our understanding of the epidemiology of malaria in STP. It will also inform the current and future malaria control programs, including the potential release of GEM in this region.

Acknowledgements The authors would like to thank Ricardo de Lima and Filipa Soares for their insightful discussions on spatial modeling, regional ecosystems, and data availability in São Tomé and Príncipe. We would like to thank the National Malaria Control Programme (PNLP) of the Ministry of Health of São Tomé and Príncipe Field Entomology Team for their support in the larval collections. We would also like to thank Christine Coleman for her help editing the final manuscript.

Author contributions LC designed the study, performed data analyses and model building, wrote, evaluated, and edited

the R script, prepared figures and wrote the manuscript. MC, JV, JP, AC and MC collected, identified, and processed mosquito specimens and reviewed drafts of the manuscript text and analyses. GL designed and funded the research, wrote the manuscript, and reviewed the analyses.

Funding This work was supported by grants from the University of California Irvine Malaria Initiative (A17-0209) and from the Open Philanthropy Project (A22-2768-001).

Data availability No new sequence data was generated for this study. Sequence data was previously published from Campos et al., 2024 and are available under the Genbank SRA Codes SAMN25173698-SAMN25174013, SAMN38765460-SAMN38765466 under BioProject ID PRJNA779397. All scripts, raster files, and occurrence data necessary to run all analyses and generate all figures were deposited at FigShare.

Declarations

Competing interest The authors declare no competing interests.

References

Adeogun A, Babalola AS, Okoko OY, Oyeniyi T, Omotayo A, Izekor RT, Adetunji O et al (2023) Spatial distribution and ecological niche modeling of geographical spread of Anopheles gambiae complex in Nigeria using real time data. Sci Rep 13:13679

Afrane YA, Bonizzoni M, Yan G (2016) Secondary malaria vectors of Sub-Saharan Africa: threat to malaria elimination on the continent? INTECH. https://doi.org/10.5772/65359

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. Selected papers of Hirotugu Akaike. Springer, New York, pp 199–213

Ali JR (2018) Islands as biological substrates: continental. J Biogeogr 45:1003–1018

Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J Appl Ecol 43:1223–1232

Amiri M, Tarkesh M, Jafari R, Jetschke G (2020) Bioclimatic variables from precipitation and temperature records vs remote sensing-based bioclimatic variables: Which side can perform better in species distribution modeling? Ecol Inform 57:101060

Anantharaman R, Hall K, Shah V, Edelman A (2019) Circuitscape in Julia: high performance connectivity modelling to support conservation decisions. arXiv preprint arXiv:1906.03542. https://doi.org/10.48550/arXiv.1906.03542

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol 22:42–47

Bah T (2011) Inkscape: guide to a vector drawing program. Prentice Hall Press

Baptista JL (1996) Subsidio Para a História do Paludismo Em S. Tomé Acta Médica Portuguesa 9:259–265

Barbosa AM (2015) fuzzySim: applying fuzzy logic to binary similarity indices in ecology. Methods Ecol Evol 6:853–858

Bebber DP, Field E, Gui H, Mortimer P, Holmes T, Gurr SJ (2019) Many unreported crop pests and pathogens are probably already present. Global Change Biol 25:2703–2713

Bell RC, Drewes RC, Channing A, Gvoždík V, Kielgast J, Lötters S, Stuart BL, Zamudio KR (2015) Overseas dispersal of *Hyperolius* reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. J Biogeogr 42(1):65–75

Bier E, Nizet V (2021) Driving to safety: CRISPR-based genetic approaches to reducing antibiotic resistance. Trends Genet 37:745–757

Bigoga J, Manga L, Titanji V, Coetzee M, Leke R (2007) Malaria vectors and transmission dynamics in coastal south-western Cameroon. Malar J 6:5

Bivand R, Mächler M (2022) qgis_process: access QGIS processing algorithms from R. R package version 2.1.0. https://CRA.R-project.org/package=qgis_process

Blackburn DC (2010) A new squeaker frog (Arthroleptidae: Arthroleptis) from Bioko Island. Equatorial Guinea Herpetologica 66(3):320–334

Boria RA, Olson LE, Goodman SM, Anderson RP (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecol Model 275:73–77

Bredero JT, Heemskerk W, Toxopeus H (1977) Agriculture and livestock production in São Tomé and Príncipe (West Africa). SVP

Breiman L (1984) Classification and regression trees. Routledge

Breiman L (2001) Random forests. Mach Learn 45:5–32

Breiman L (2002) Using models to infer mechanisms. IMS Wald Lecture 2. http://www.oz.berkeley.edu/users/breiman/wald2002-2.pdf

Calzetta M, Santolamazza F, Carrara GC, Cani PJ, Fortes F, Di Deco MA, della Torre A et al (2008) Distribution and chromosomal characterization of the *Anopheles gambiae* complex in Angola. Am J Trop Med Hyg 78:169–175

Campos M, Hanemaaijer M, Gripkey H, Collier TC, Lee Y, Cornel AJ et al (2021) The origin of island populations of the African malaria mosquito. Anopheles Coluzzii Commun Biol 4:630

Campos M, Rašić G, Viegas J, Cornel AJ, Pinto J, Lanzaro GC (2024) Patterns of gene flow in *Anopheles coluzzii* populations from two African Oceanic Islands. Evol Appl 17:e70044

Carballar-Lejarazú R, Ogaugwu C, Tushar T, Kelsey A, Pham TB, Murphy J, Schmidt H, Lee Y, Lanzaro GC, James AA (2020) Next-generation gene drive for population modification of the malaria vector mosquito, *Anopheles gambiae*. Proc Natl Acad Sci USA 117:22805–22814

Carballar-Lejarazú R, Dong Y, Pham TB, Tushar T, Corder RM, Mondal A, Sánchez HM, Lee HF, Marshall JM, Dimopoulos G (2023) Dual effector population modification gene-drive strains of the African malaria mosquitoes, *Anopheles gambiae* and *Anopheles coluzzii*. Proc Natl Acad Sci. https://doi.org/10.1073/pnas.2221118120

Carlson CJ, Bannon E, Mendenhall E, Newfield T, Bansal S (2023) Rapid range shifts in African *Anopheles* mosquitoes over the last century. Biol Lett 19:20220365

Ceríaco LMP, Bernstein J, Sousa AC et al (2020) The reptiles of Tinhosa Grande islet (Gulf of Guinea): a taxonomic update and the role of Quaternary sea level fluctuations in their diversification. Afr J Herpetol 69(2):200–216

Charlwood JD (2019) The ecology of malaria vectors, 1st edn. CRC Press

Chou SC, de Arruda LA, Gomes JL, Rodriguez DA, Alves Martins M, Costa Resende N et al (2020) Downscaling projections of climate change in Sao Tome and Principe Islands, Africa. Clim Dyn 54:4021–4042

Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. J Agric Biol Environ Stat 7:361–372

Correa M, Lopes J, Sousa CA, Rocha G, Oriango R, Cardetas A et al (2024) The status of insecticide resistance of *Anopheles coluzzii* on the islands of São Tomé and Príncipe, after 20 years of malaria vector control. Malar J 23(1):390

Cuamba NK, Choi K, Townson H (2006) Malaria vectors in Angola: distribution of species and molecular forms of the *Anopheles gambiae c*omplex, their pyrethroid insecticide knockdown resistance (kdr) status an*d Plasmodium falciparum* sporozoite rates. Malar J 5:2

Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J et al (2007) Random forests for classification in ecology. Ecology 88:2783–2792

Dauby G, Stévart T, Barberá P, Benitez L, do Céu Madureira M, Soares FC et al (2022) Classification, distribution, and biodiversity of terrestrial ecosystems in the Gulf of Guinea oceanic islands. Biodiversity of the Gulf of Guinea Oceanic Islands: science and conservation. Springer International Publishing, pp 37–69

de Lima RF, Melo M (2021) A revised bird checklist for the oceanic islands of the Gulf of Guinea (Príncipe, São Tomé and Annobón). Bull Brit Ornithol Club 141(2):179–198

De'ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192. https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

Dedzo MG, Kamgang P, Njonfang E, Tefogoum GZ (2015) Hazards from pyroclastic density currents at Mount Bamenda (West-Cameroon, Cameroon Volcanic Line): mapping and assessment by AMS method. Current Research in Geoscience 5(2):26–39

Diniz AC, Matos GC (2002) Carta da zonagem agro-ecológica e da vegetação de S. Tomé e Príncipe. Garcia De Orta, Série Botânica 15(2):1–72

Ditter RE, Campos M, Pinto J, Cornel AJ, Rompão H, Lanzaro GC (2022) Mitogenome analyses reveal limited introduction of *Anopheles coluzzii* into the Central African Islands of São Tomé and Príncipe. Front Trop Dis 3:855272

Dooley CA, Tatem AJ, Bondarenko M (2020) Gridded maps of building patterns throughout sub-Saharan Africa, version 1.0. University of Southampton: Southampton. Source of building Footprints "Ecopia Vector Maps Powered by Maxar Satellite Imagery

Elith J, Graham CP, Anderson RP, Dudik M, Ferrier S, Guisan A et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9:1937–1958

Feng X, Park DS, Walker C et al (2019) A checklist for maximizing reproducibility of ecological niche models. Nat Ecol Evol 3:1382–1395

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37:4302–4315

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49

Figueiredo E, Paiva J, Martins ES (1994) Endemic plants of São Tomé and Príncipe. Cour Forschungsinst Senck 169:231–243

Garnier S, Ross N, Rudis R, Camargo PA, Sciaini M, Scherer C (2023) viridisLite: colorblind-friendly color maps (Lite version). R package version 0.4.2

Gimonneau G, Dossou-Yovo J, Gouteux JP et al (2012) Larval habitat segregation between the molecular forms of the mosquito *Anopheles gambiae* in a rice field area of Burkina Faso, West Africa. Med Vet Entomol 26:9–17

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Hall KR, Anantharaman R, Landau VA, Clark M, Dickson BG, Jones A, Platt J, Edelman A, Shah VB (2021) Circuitscape in Julia: empowering dynamic approaches to connectivity assessment. Land 10:301

Hastie T, Tibshirani R (1990) Generalized additive models. Monographs on statistics and applied probability, vol 43. Chapman and Hall, London, p 335

Hastie T, Tibshirani R, Friedman J (2009) MARS: multivariate adaptive regression splines. The elements of statistical learning: data mining, inference, and prediction. Springer, New York, pp 321–329

Hijmans RJ, Bivand R, Forner K, Ooms J, Pebesma E, Sumner MD (2022) Package 'terra.' Maintainer, Vienna

Hijmans RJ, Phillips S, Leathwick J, Elith J (2024) dismo: Species Distribution Modeling. R package version 1.3–15. Available at: https://github.com/rspatial/dismo

Jones PJ (1994) Biodiversity in the Gulf of Guinea: an overview. Biodivers Conserv 3:772–784

Juste B, Ibáñez C (1994) Bats of the Gulf of Guinea islands: faunal composition and origins. Biodivers Conserv 3:837–850

Juste J, Torrent L, Méndez-Rodríguez A, Howard K, García-Mudarra JL, Nogueras J, Ibáñez C (2023) A new Pipistrelle bat from the oceanic Island of Príncipe (Western Central Africa). J Mammal 104(2):361–371

Kamdem C, Tene Fossog B, Simard F, Etouna J, Ndo C, Kengne P, Bousses P, Kengne P, Costantini C (2012) Anthropogenic habitat disturbance and ecological divergence between incipient species of the malaria mosquito *Anopheles gambiae*. PLoS ONE. https://doi.org/10.1371/journal.pone.0039453

Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V, Stillfried M et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. Divers Distrib 19:1366–1379

Kujala H, Whitehead AL, Wintle BA (2015) Identifying conservation priorities and assessing impacts and trade-offs of potential future development in the Lower Hunter Valley in New South Wales. Report by the NERP Environmental Decisions Hub

Lanzaro GC, Campos M, Crepeau M, Cornel A, Estrada A, Gripkey H, Haddad Z, Kormos A, Palomares S (2021) Selection of sites for field trials of genetically engineered mosquitoes with gene drive. Evol Appl 14:2147–2161

Le Saout S, Hoffmann M, Shi Y, Hughes A, Bernard C, Brooks TM, Bertzky B, Butchart SHM, Stuart SN, Badman T, Rodrigues ASL (2013) Protected areas and effective biodiversity conservation. Science 342:803–805

Lee Y, Cornel AJ, Meneses CR, Fofana A, Andrianarivo AG, McAbee RD, Fondjo E (2009) Ecological and genetic relationships of the Forest-M form among chromosomal and molecular forms of the malaria vector *Anopheles gambiae* sensu stricto. Malar J 8:75

Lee PW, Liu CT, do Rosario VE, de Sousa B, Rampao HS, Shaio MF (2010) Potential threat of malaria epidemics in a low transmission area, as exemplified by São Tomé and Príncipe. Malaria J 9:1–12

Lehmann T, Dao A, Yaro AS, Adamou A, Kassogue Y, Diallo M, Traoré S, Coscaron-Arias C (2010) Aestivation of the African malaria mosquito, *Anopheles gambiae* in the Sahel. Am J Trop Med Hyg 83:601

Lippi CA, Mundis SJ, Sippy R et al (2023) Trends in mosquito species distribution modeling: insights for vector surveillance and disease control. Parasites Vectors 16:302

Loiseau C, Melo M, Lee Y, Pereira H, Hanemaaijer MJ, Lanzaro GC, Cornel AJ (2019) High endemism of mosquitoes on São Tomé and Príncipe Islands: evaluating the general dynamic model in a worldwide island comparison. Insect Conserv Divers 12:69–79

Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W (2009) Evaluation of consensus methods in predictive species distribution modelling. Divers Distrib 15:59–69

Mauritsen T, Gutjahr O, Haak H et al (2019) The MPI-ESM1.2 global coupled climate model based on the CMIP6 experimental design. J Adv Model Earth Syst 11:998–1038

McCullagh P (1989) Generalized linear models, 2nd edn. Routledge

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 5:115–133

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724

Medley KA, Jenkins DG, Hoffman EA (2015) Human-aided and natural dispersal drive gene flow across the range of an invasive mosquito. Molec Ecol 24(2):284–295

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. J Chem Phys 21:1087–1092

Molina-Cruz A, Zilversmit MM, Neafsey DE, Hartl DL, Barillas-Mury C (2016) Mosquito vectors and the globalization of Plasmodium falciparum malaria. Annu Rev Genet 50(1):447–465

Muscatello A, Elith J, Kujala H (2021) How decisions about fitting species distribution models affect conservation outcomes. Conserv Biol 35:1309–1320

Naimi B, Araújo MB (2016) sdm: a reproducible and extensible R platform for species distribution modelling. Ecography 39:368–375

Naimi B (2017) usdm: uncertainty analysis for species distribution models. R package version 1.1–18. https://CRA.R-project.org/package=usdm

Neto K, Henriques MH (2023) Geoheritage of the Príncipe UNESCO World Biosphere Reserve (West Africa): selected geosites. Geoheritage 15:118

Newfield TP (2017) Malaria and malaria-like disease in the early Middle Ages. Early Medieval Europe 25:251–300

O'Neill BC, Tebaldi C, Van Vuuren DP, Eyring V, Friedlingstein P, Hurtt G, Knutti R, Kriegler E, Lamarque J-F, Lowe J, Meehl GA, Moss R, Riahi K, Sanderson BM (2016) The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci Model Dev 9:3461–3482

Ooms J (2023) osmdata: download and manipulate OpenStreetMap Data. R package version 0.6.0. https://CRA.R-project.org/package=osmdata

Padgham M, Lovelace R (2017) osmdata: import 'OpenStreetMap' data as simple features or spatial objects. R package version 0.1.1. https://crA.r-project.org/package=osmdata

Peet NB, Atkinson PW (1994) The biodiversity and conservation of the Gulf of Guinea islands: São Tomé and Príncipe. Biodivers Conserv 3(9):854–867

Peterman WE (2018) ResistanceGA: an R package for the optimization of resistance surfaces using genetic algorithms. Methods Ecol Evol 9:1638–1647

Peterman WE, Winiarski KJ, Moore CE, da Silva CC, Gilbert AL, Spear SF (2019) A comparison of popular approaches to optimize landscape resistance surfaces. Landscape Ecol 34:2197–2208

Pinto J, Sousa CA, Gil V, Ferreira C, Gonçalves L, Lopes D et al (2000) Malaria in São Tomé and Príncipe: parasite prevalences and vector densities. Acta Trop 76:185–193

Pinto J, Donnelly MJ, Sousa CA, Gil V, Ferreira C, Elissa N et al (2002) Genetic structure of *Anopheles Gambiae* (Diptera: Culicidae) in São Tomé and Príncipe (West Africa): implications for malaria control. Mol Ecol 11:2183–2187

Pinto J, Egyir-Yawson A, Vicente J, Gomes B, Santolamazza F, Moreno M, Charlwood JD (2013) Geographic population structure of the African malaria vector *Anopheles gambiae* suggests a role for the forest-savannah biome transition as a barrier to gene flow. Evol Appl 6:910–924

Ridgeway GK (1999) Generalization of boosting algorithms and applications of Bayesian inference for massive datasets. University of Washington. http://www.stat.washington.edu/tech.reports/1999/tr267.pdf

Ridgeway G (2007) Generalized boosted models: a guide to the GBM package. Update 1:2007. https://CRA.R-project.org/package=gbm

Ryan WB, Carbotte SM, Coplan JO, O'Hara S, Melkonian A, Arko R et al (2009) Global multi-resolution topography synthesis. Geochem Geophys Geosyst. https://doi.org/10.1029/2008GC002332

Ryan SJ, Lippi CA, Villena OC, Singh A, Murdock CC, Johnson LR (2023) Mapping current and future thermal limits to suitability for malaria transmission by the invasive mosquito *Anopheles stephens*i. Malaria J 22:104

Sallares R (2002) Malaria and Rome: a history of malaria in ancient Italy. Oxford University Press, Oxford

Senterre B, Lowry PP II, Bidault E, Stévart T (2021) Ecosystemology: a new approach toward a taxonomy of ecosystems. Ecol Complex 47:100945

Sherpa S, Renaud J, Guéguen M et al (2020) Landscape does matter: disentangling founder effects from natural and human-aided post-introduction dispersal during an ongoing biological invasion. J Anim Ecol 89(9):2027–2042

Silver JB (2007) Mosquito ecology: field sampling methods. Springer Science & Business Media

Simard F, Ayala D, Kamdem GC, Etouna J, Ose K, Fotsing JM, Fontenille D (2009) Ecological niche partitioning between the M and S molecular forms of *Anopheles gambiae* in Cameroon: the ecological side of speciation. BMC Ecol 9:17

Sousa CA, Pinto J, Almeida PG, Ferreira C, do Rosário VE, Charlwood JD (2001) Dogs as a favored host choice of *Anopheles gambiae sensu* stricto (Diptera: Culicidae) of São Tomé, West Africa. J Med Entomol 38:122–125

Stévart T, Dauby G, Ikabanga DU, Lachenaud O, Barberá P, de Oliveira F, Benitez L, Madureira M (2022) Diversity of the vascular plants of the Gulf of Guinea oceanic islands. Biodiversity of the Gulf of Guinea oceanic islands. Springer, Cham, pp 249–272

Surendran SN, Sivabalakrishnan K, Sivasingham A, Jayadas TT, Karvannan K, Santhirasegaram S, Ramasamy R (2019) Anthropogenic factors driving recent range expansion of the malaria vector Anopheles stephensi. Front Public Health 7:53

Tene Fossog B, Ayala D, Acevedo P, Kengne P, Ngomo Abeso Mebuy I, Makanga B, Costantini C (2015) Habitat segregation and ecological character displacement in cryptic African malaria mosquitoes. Evol Appl 8:326–345

Thuiller W, Georges D, Engler R, Breiner F (2024) biomod2: ensemble platform for species distribution modeling. R package version 4.2-5, 3: r539. https://biomodhub.github.io/biomod2/

Verdin KL (2017) Hydrologic Derivatives for Modeling and Applications (HDMA) database: U.S. Geological Survey data release. https://doi.org/10.5066/F7S180ZP

Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. Ecology 86:1975–1987

Wagner P, Leaché AD, Fujita MK (2014) Description of four new West African forest geckos of the *Hemidactylus fasciatus* Gray, 1842 complex, revealed by coalescent species delimitation. Bonn Zool Bull 63(1):1–14

Wang GH, Gamez S, Raban RR, Marshall JM, Alphey L, Li M, Rasgon JL, Akbari OS (2021) Combating mosquito-borne diseases using genetic control technologies. Nat Commun 12:4388

White BJ, Collins FH, Besansky NJ (2011) Evolution of *Anopheles gambiae* in relation to humans and malaria. Annu Rev Ecol Evol Syst 42:111–132

WHO/TDR, FNIH (2014) The guidance framework for testing genetically modified mosquitoes. Framework. https://www.who.int/tdr/publications/year/2014/Guidance_framework_mosquitoes.pdf

Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer-Verlag, New York. https://ggplot2.tidyverse.org

Wieringa JG, Carstens BC, Gibbs HL (2021) Predicting migration routes for three species of migratory bats using species distribution models. PeerJ. https://doi.org/10.7717/peerj.11177

WorldPop (2021) WorldPop project: population data for the globe. https://www.worldpop.org

Wrobel M (2006) Elsevier's dictionary of mammals, 1st edn. Elsevier B.V, Amsterdam