

2023

## Estimating the Distribution of *Oryzomys palustris*, A Potential Key Host in Expanding Rickettsial Tick-Borne Disease Risk

Catherine A. Lippi

Samuel Canfield

Christina Espada  
Old Dominion University, cespa001@odu.edu

Holly D. Gaff  
Old Dominion University, hgaff@odu.edu

Sadie J. Ryan

Follow this and additional works at: [https://digitalcommons.odu.edu/biology\\_fac\\_pubs](https://digitalcommons.odu.edu/biology_fac_pubs)



Part of the [Animal Sciences Commons](#), [Parasitic Diseases Commons](#), and the [Population Biology Commons](#)

---

### Original Publication Citation

Lippi, C. A., Canfield, S., Espada, C., Gaff, H. D., & Ryan, S. J. (2023). Estimating the distribution of *Oryzomys palustris*, a potential key host in expanding rickettsial tick-borne disease risk. *Ecosphere*, 14(3), 1-15, Article e4445. <https://doi.org/10.1002/ecs2.4445>

This Article is brought to you for free and open access by the Biological Sciences at ODU Digital Commons. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

## ARTICLE

## Disease Ecology

# Estimating the distribution of *Oryzomys palustris*, a potential key host in expanding rickettsial tick-borne disease risk

Catherine A. Lippi<sup>1</sup>  | Samuel Canfield<sup>1</sup> | Christina Espada<sup>2</sup> |  
Holly D. Gaff<sup>2</sup>  | Sadie J. Ryan<sup>1</sup> 

<sup>1</sup>Department of Geography and Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA

<sup>2</sup>Department of Biology, Old Dominion University, Norfolk, Virginia, USA

**Correspondence**

Sadie J. Ryan

Email: [sjryan@ufl.edu](mailto:sjryan@ufl.edu)

**Funding information**

Centers for Disease Control and Prevention, Grant/Award Number: 1U01CK000510-01; National Institutes of Health, Grant/Award Number: NIH 1R01AI136035-01; National Science Foundation, Grant/Award Number: NSF 2016265

**Handling Editor:** Shannon L. LaDeau

**Abstract**

Increasingly, geographic approaches to assessing the risk of tick-borne diseases are being used to inform public health decision-making and surveillance efforts. The distributions of key tick species of medical importance are often modeled as a function of environmental factors, using niche modeling approaches to capture habitat suitability. However, this is often disconnected from the potential distribution of key host species, which may play an important role in the actual transmission cycle and risk potential in expanding tick-borne disease risk. Using species distribution modeling, we explore the potential geographic range of *Oryzomys palustris*, the marsh rice rat, which has been implicated as a potential reservoir host of *Rickettsia parkeri*, a pathogen transmitted by the Gulf Coast tick (*Amblyomma maculatum*) in the southeastern United States. Due to recent taxonomic reclassification of *O. palustris* subspecies, we reclassified geolocated collections records into the newer clade definitions. We modeled the distribution of the two updated clades in the region, establishing for the first time, range maps and distributions of these two clades. The predicted distribution of both clades indicates a largely Gulf and southeastern coastal distribution. Estimated suitable habitat for *O. palustris* extends into the southern portion of the Mid-Atlantic region, with a discontinuous, limited area of suitability in coastal California. Broader distribution predictions suggest potential incursions along the Mississippi River. We found considerable overlap of predicted *O. palustris* ranges with the distribution of *A. maculatum*, indicating the potential need for extended surveillance efforts in those overlapping areas and attention to the role of hosts in transmission cycles.

**KEYWORDS**

*Amblyomma maculatum*, *Oryzomys*, rickettsia, species distribution model, tick-borne diseases

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Ecosphere* published by Wiley Periodicals LLC on behalf of The Ecological Society of America.

## INTRODUCTION

Zoonotic vector-borne disease systems are ecologically complex due to the numerous interactions between pathogens, vectors, and hosts, in the context of the environment, that ultimately determine the risk of disease in humans. The presence of competent arthropod vectors is sometimes used as a proxy for pathogen exposure and transmission risk, particularly when transmission cycles and other risk factors are poorly understood or unknown (Lippi, Gaff, White, & Ryan, 2021; Lippi, Ryan, et al., 2021). This is the case for many tick-borne pathogen transmission systems, where there can be major knowledge gaps concerning wildlife reservoirs, environmental associations, or even the etiology of disease.

Leveraging tick presence to estimate geographic risk has logistical advantages, as existing data sources (e.g., surveillance programs, museum collections, online data repositories) can provide ample occurrence points for spatial risk modeling and are typically easier to obtain than protected human health data. Nevertheless, vector presence alone is not sufficient to maintain transmission cycles. Many medically important ticks are often habitat generalists, with broad geographic distributions exceeding known disease risk areas (Lippi, Gaff, White, St. John, et al., 2021). Additionally, ticks with expanding ranges can be associated with increases in the burden and diversity of pathogens in newly established locations compared with their historic range (Fornadel et al., 2011; Nadolny & Gaff, 2018; Ogden et al., 2013). These instances highlight the need to understand other components of transmission cycles that drive the risk of exposure beyond vector presence. Reservoir hosts are an important, yet in many instances underexplored, component of zoonotic tick-borne transmission cycles (Tomassone et al., 2018). When considering geographic risks of pathogen transmission, the distribution of wildlife hosts may serve as another important proxy for exposure risk. Therefore, while the use of ticks as proxies of exposure may be advisable in the absence of additional data, determining the distributions of reservoir hosts needed for pathogen amplification and spillover may serve as more nuanced indicators of risk and provide a focus for public health messaging and resources.

Tick-borne diseases have garnered increased attention from the medical community and public health practitioners over the past decade (Lippi, Ryan, et al., 2021; Rochlin & Toledo, 2020). This is largely a result of the increasing incidence of diseases such as Lyme borreliosis and spotted fever group (SFG) rickettsioses (CDC, 2021). Tick-borne infections are often underdiagnosed and go unreported, in part due to nondescript clinical presentations (e.g., febrile illness) and a lack of accurate diagnostic testing (Madison-Antenucci et al., 2020; Rosenberg et al., 2018).

Therefore, public health advisories aimed at preventing exposure to potentially infective tick bites are an important aspect of managing these diseases. The Gulf Coast tick (*Amblyomma maculatum*) is a species of medical and veterinary importance in the Americas (Paddock & Goddard, 2015). *Amblyomma maculatum* is a competent vector of *Rickettsia parkeri*, the causative agent of *Rickettsia parkeri* rickettsiosis, a SFG disease (Paddock et al., 2004; Paddock & Goddard, 2015). *Amblyomma maculatum* is a habitat specialist found in open fields dominated by grasses and shrubs, requiring high rainfall, temperature, and humidity (Nadolny & Gaff, 2018; Paddock & Goddard, 2015). In the United States, *A. maculatum* has been undergoing a northward range expansion from its historic range along the Gulf Coast, most recently establishing populations in Connecticut, Illinois, and New York (Molaei et al., 2021; Phillips et al., 2020; Ramírez-Garofalo et al., 2022). Importantly, studies have shown that the prevalence of *R. parkeri* infections is higher in *A. maculatum* collected from some newly established locations, a possible result of higher pathogen burden in new vertebrate host populations (Fornadel et al., 2011; Nadolny & Gaff, 2018). Therefore, identifying competent reservoir hosts that underpin transmission cycles and delineating their distributions could provide better estimates of risk in the face of vector range expansion.

The marsh rice rat (*Oryzomys palustris*) is a suspected reservoir host for *R. parkeri* (Cumbie et al., 2020) and a vertebrate host for *A. maculatum* (Nadolny & Gaff, 2018; Wilson & Durden, 2003). Recent field studies have indicated that *O. palustris* may be involved in the pathogen transmission cycle of *R. parkeri*; the rats themselves have tested positive for infection, and *A. maculatum* collected from rats in the field have also tested positive for the pathogen (Cumbie et al., 2020). The distribution of *O. palustris* has been previously described with observation-based range maps, which are limited to the extent of field observations (Wolfe, 1982). Coarse species range maps are useful in certain contexts but are prone to localized errors and provide no information on spatial variation in the probability of species occurrences (Maréchaux et al., 2017). Further, recently proposed taxonomic changes to the *O. palustris* species complex are relevant to our understanding of where potential reservoirs are located, yet are available only as range maps (Hanson et al., 2010). To our knowledge, species distribution models (SDMs) have not been used to estimate the suitable geographic range of marsh rice rats. Frequently used in ecology and conservation research, SDMs have become an increasingly utilized methodology in medical geography and public health for estimating the geographic distribution of vector species. A typical SDM workflow entails using occurrence records for the species of interest with environmental predictor

layers as inputs for an SDM algorithm, which produces spatially continuous outputs of potential habitat suitability throughout the study area. Here, we apply this technique to estimate the geographic range of the marsh rice rat to assess its potential role in tick-borne pathogen transmission.

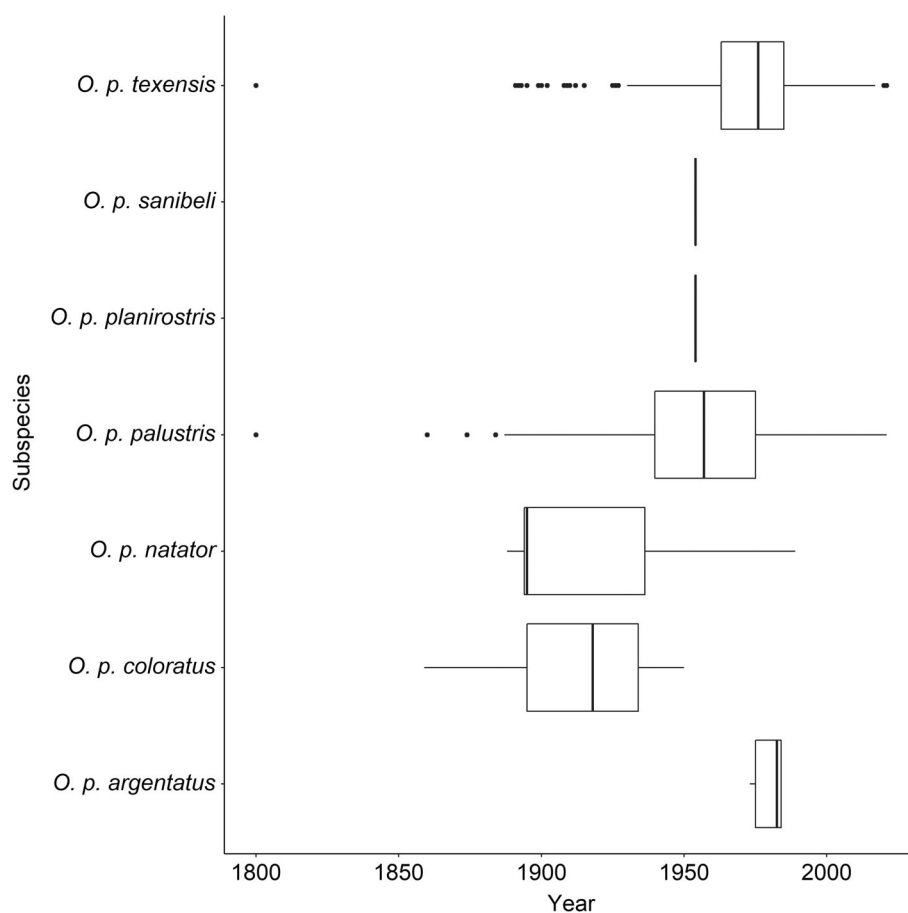
## METHODS

### Occurrence data

Occurrence records for *O. palustris* were downloaded from the Global Biodiversity Information Facility (GBIF) ([www.gbif.org](http://www.gbif.org)), an international network that collates museum records and information on where and when species were collected. The collection timeframe for GBIF records extended from 1800 to 2021, though most occurrences were collected between 1888 and 1999, with few records outside this range (Figure 1). The full dataset ( $n = 7276$ ) included many records that had not been

georeferenced ( $n = 2457$ ) (i.e., collection records with no associated latitude/longitude coordinates). The GEOLocate platform ([www.geo-locate.org](http://www.geo-locate.org)) was used to recover positional coordinates for records that had not been georeferenced but had locality data (i.e., a text description of the collection site was available) associated with records ( $n = 1727$ ) (Murphey et al., 2004). Returned point locations with low precision, where locality data could be matched to several potential sites, were not retained for model building. The spatial uncertainty of georeferenced data on GBIF and coordinates recovered with GEOLocate varied substantially, and we excluded records where spatial uncertainty exceeded 10 km to maintain comparable resolution with environmental covariates. A total of 819 locations were recovered using the GEOLocate workflow.

Morphological and molecular evaluations of the *O. palustris* species complex have led to proposed taxonomic revisions over the past decade. Notably, Hanson et al. (2010) conducted a taxonomic assessment of subspecies in the *O. palustris* and *O. couesi* species complexes, utilizing three genetic markers as lines of evidence. The subspecies



**FIGURE 1** Collection record date ranges for Global Biodiversity Information Facility records of *Oryzomys palustris* subspecies used in this study. Boxplots depict the distribution of collection record totals for each subspecies (i.e., minimum, mean, and maximum values, with outliers shown as points).

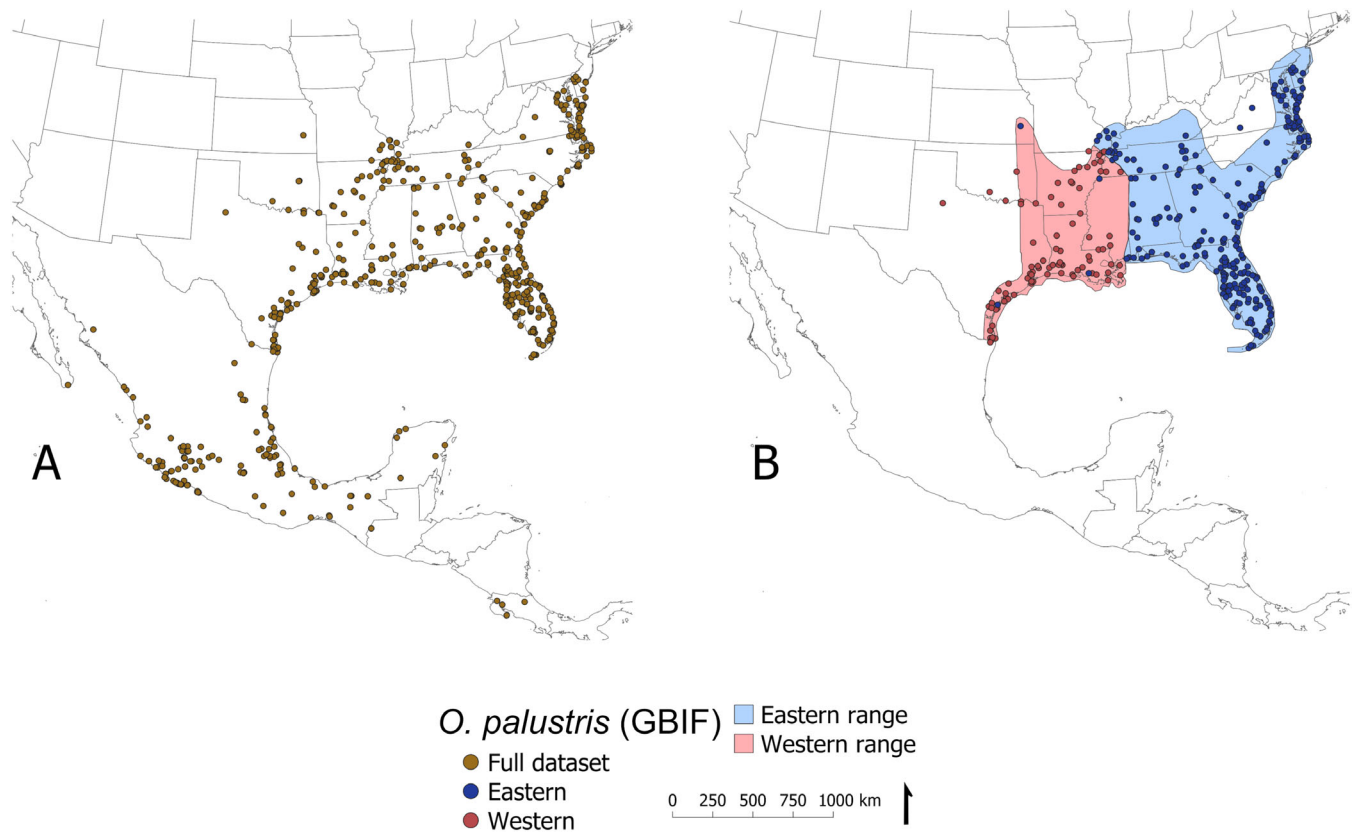
examined within the *O. palustris* clade were *O. p. texensis*, *O. p. palustris*, *O. p. coloratus*, *O. p. planirostris*, *O. p. sanibeli*, and *O. p. argentatus* (synonymous with *O. p. natator*; Baker et al., 2003; Humphrey & Setzer, 1989). Two clades within the *O. palustris* group showed considerable genetic divergence: a western clade consisting largely of *O. p. texensis* and an eastern clade comprising *Oryzomys palustris* and the remaining subspecies. Though Hanson et al. support the elevation of *O. p. texensis* to species following the Genetic Species Concept, these proposed taxonomic updates were not reflected in data obtained from GBIF, particularly for older collection events. Therefore, we created three separate spatial datasets of marsh rice rat occurrences for model building: (1) all geolocated GBIF records identified as *O. palustris* ( $n = 2195$ ), plus two subsets derived according to clades identified in Hanson et al. (2010); (2) the eastern *O. palustris* clade; and (3) the western *O. p. texensis* clade. The full dataset of *O. palustris* occurrences was partitioned into subsets by combining records with valid subspecies provided and spatial assignment of records without subspecific epithets. To conduct this spatial assignment, we rendered the published images of the geographic distribution of *O. palustris* and *O. p. texensis* as spatial layers, using QGIS (ver 3.24.0), and georeferenced the collection points,

yielding 1251 and 570 occurrence locations, respectively (Figure 2) (Hanson et al., 2010).

To prepare the geospatial occurrence data for model input, occurrence records in each of the three datasets were spatially thinned at a 10-km radius using spThin in R (Aiello-Lammens et al., 2015), removing duplicates and retaining one presence point per predictor pixel at the resolution of the study. The final datasets for model building included 336 occurrences in the full *O. palustris* dataset, 249 occurrences in the eastern dataset, and 90 occurrences in the western dataset. Each dataset was randomly partitioned into a 75% training/25% testing split for calculation of external model accuracy metrics.

## Environmental data

Environmental factors that have broad biological implications for landscape-level habitat suitability of small mammals were considered for initial model development, including indicators of precipitation, temperature, elevation, and soil conditions. A total of 22 environmental covariates at 5 min ( $\sim 10 \text{ km}^2$ ) spatial resolution were selected for variable screening (Table 1). Long-term



**FIGURE 2** (A) Map of the United States and Central America and the spatial distribution of Global Biodiversity Information Facility (GBIF) specimen data records recognized as *Oryzomys palustris*. (B) A 10-km spatial thinning of the eastern and western clades, following Hanson et al. (2010).

**TABLE 1** Variables from WorldClim or the International Soil Reference and Information Center (ISRIC) considered for model development (MOD), including those removed due to collinearity through variance inflation factor (R).

Variable	Units	Source	Status
Annual Mean Temperature (Bio1)	°C	WorldClim ver 1.4	MOD
Temperature Seasonality (Bio4)	°C	WorldClim ver 1.4	MOD
Mean Temperature of Wettest Quarter (Bio8)	°C	WorldClim ver 1.4	MOD
Precipitation of Wettest Month (Bio13)	mm	WorldClim ver 1.4	MOD
Precipitation Seasonality (Bio15)	%	WorldClim ver 1.4	MOD
Elevation	m	WorldClim ver 2.1	MOD
Soil Organic Carbon Stock 30-cm depth	tons/ha	ISRIC	MOD
Soil Volumetric Water Content 33 kPa 0–5 cm depth	cm <sup>3</sup> /100 cm <sup>3</sup>	ISRIC	MOD
Maximum Temperature of Warmest Month (Bio5)	°C	WorldClim ver 1.4	R
Minimum Temperature of Coldest Month (Bio6)	°C	WorldClim ver 1.4	R
Temperature Annual Range (Bio7)	°C	WorldClim ver 1.4	R
Mean Temperature of Driest Quarter (Bio9)	°C	WorldClim ver 1.4	R
Mean Temperature of Warmest Quarter (Bio10)	°C	WorldClim ver 1.4	R
Mean Temperature of Coldest Quarter (Bio11)	°C	WorldClim ver 1.4	R
Annual Precipitation (Bio12)	mm	WorldClim ver 1.4	R
Precipitation of Driest Month (Bio14)	mm	WorldClim ver 1.4	R
Precipitation of Wettest Quarter (Bio16)	mm	WorldClim ver 1.4	R
Precipitation of Driest Quarter (Bio17)	mm	WorldClim ver 1.4	R
Precipitation of Warmest Quarter (Bio18)	mm	WorldClim ver 1.4	R
Precipitation of Coldest Quarter (Bio19)	mm	WorldClim ver 1.4	R

climate data were obtained from WorldClim (ver 1.4) (Hijmans et al., 2005), which included 19 bioclimatic variables for monthly temperature and rainfall values with biologically derived products to represent annual trends, seasonality, and extreme or limiting environmental factors. WorldClim 1.4 incorporates climate averages from the years 1960–1990 and was chosen to align with the species occurrence data obtained for this study. Bioclimatic variables considered for model building included Annual Mean Temperature (Bio1), Mean Diurnal Range (Bio2), Isothermality (Bio3), Temperature Seasonality (Bio4), Max Temperature of Warmest Month (Bio5), Min Temperature of Coldest Month (Bio6), Temperature Annual Range (Bio7), Mean Temperature of Wettest Quarter (Bio8), Mean Temperature of Driest Quarter (Bio9), Mean Temperature of Warmest Quarter (Bio10), Mean Temperature of Coldest Quarter (Bio11), Annual Precipitation (Bio12), Precipitation of Wettest Month (Bio13), Precipitation of Driest Month (Bio14), Precipitation Seasonality (Bio15), Precipitation of Wettest Quarter (Bio16), Precipitation of Driest Quarter (Bio17), Precipitation of Warmest Quarter (Bio18), and Precipitation of Coldest Quarter (Bio19). Gridded elevation data were also obtained from WorldClim (ver 2.1) (Fick & Hijmans, 2017), derived from the Shuttle Radar Topography Mission

(SRTM). Soil characteristics considered for model building were acquired from the International Soil Reference and Information Centre (ISRIC) and included organic carbon stock at a 30-cm depth (Hengl et al., 2017) and volumetric water content at 33 kPa in 0–5 cm range (Poggio et al., 2021). The timeframe during which most records were collected precluded the use of other descriptive land cover products in our analyses, as these products did not align with the collection period of most records, potentially misrepresenting historically suitable habitats.

Multicollinearity in environmental covariates can lead to known issues in SDMs, such as distortion of estimated effects of predictors (Dormann et al., 2013). This is a standard concern when modeling with numerous variables, particularly those derived from similar variables (e.g., bioclimatic variables derived from WorldClim data). While some SDM algorithms, such as boosted regression trees (BRTs), exhibit moderate resilience to predictor collinearity, these methods can still be prone to failure as collinearity increases (Dormann et al., 2013). To reduce collinearity, the variance inflation factor (VIF) technique was used, regressing predictors against each other. A VIF value of 1 signifies no collinearity, with increasing values indicating greater collinearity. The optimal threshold for VIF is debated, though common rules of thumb are

found in the literature (e.g.,  $VIF = 10$  is a commonly implemented cutoff) (O'Brien, 2007). In this study, we adopted a conservative cutoff ( $VIF = 3$ ) to minimize the potential impact of multicollinearity on model output. The resulting set of predictors used in baseline analyses, before further variable reduction procedures, were Annual Temperature (Bio1), Temperature Seasonality (Bio4), Mean Temperature of Wettest Quarter (Bio8), Precipitation of Wettest Month (Bio13), Precipitation Seasonality (Bio15), Elevation, Soil Organic Carbon Stock, and Soil Volumetric Water Content.

## Species distribution modeling

SDM for all *Oryzomys palustris*, eastern and western clades, was conducted over the contiguous United States using the BRT approach. Briefly, BRT modeling is a machine learning genetic algorithm that produces a boosted ensemble of simple regression trees for predicting species occurrence using environmental covariates and presence (i.e., species occurrences) and pseudo-absences (i.e., background points) (Elith et al., 2008). For BRTs, boosting is the process of developing an initial decision tree with inadequate predictive error (e.g., shallow tree) and then sequentially adding trees and their residuals together to improve predictive performance. Each tree ensemble is considered one model, which undergoes internal predictive evaluation by the  $k$ -fold cross-validation method that splits the occurrence dataset (i.e., presence and pseudo-absence localities) into fixed subsets utilized for model testing once, and training  $k - 1$  times. An area under the receiver operator characteristics (AUC-ROC) curve is produced to demonstrate model predictive performance by exhibiting the true positive rate (sensitivity) and false positive rate ( $1 - \text{specificity}$ ) of the model for predicting between presences (1) and pseudo-absences (0). Values of AUC exceeding 0.5 indicate that the model predicted presences better than random, with 1 being a perfect prediction. Furthermore, each model was evaluated using external and internal test data randomly selected from thinned presences and pseudo-absences prior to and during model development, respectively.

The selection of pseudo-absence locations (i.e., background data) impacts model development as much as species presence data. The use of target species for pseudo-absences is generally considered the optimal method, though a random selection of background data is acceptable when target species are not available (Cerasoli et al., 2017; Phillips et al., 2009; Stokland et al., 2011). It is suggested that the number of pseudo-absences should be approximately equal to the number of presence locations in BRT analyses (Barbet-Massin et al., 2012). Therefore,

pseudo-absences ( $n = 400$ ) were randomly selected for each model from a shapefile with buffered 10-km holes around each presence location (i.e., the Swiss-cheese method) to prevent known location resampling and for model convergence (Otieno et al., 2021). Replicate models were ensembled for averaging to further account for model stochasticity. These averaged models were then used to predict the probability of occurrence on the landscape.

The family of “gbm” functions in the R “dismo” package was used for BRT model production (“gbm.step”) and simplification (“gbm.simplify”). BRT simplification is backward elimination of low-contributing predictors using  $k$ -fold cross-validation ( $n$ -fold = 10) from an initial model, reducing the number of final model predictors. Fifty model replicates were generated for the simplification process, so predictors with the lowest average contribution were removed. Model replicate count and settings were preserved between the simplification step and final model processes. The “gbm” function incorporates numerous settings to customize model development, but the primary settings considered include model family, tree complexity (tc), learning rate (lr), bag fraction (bf), and maximum trees (mt). The “Bernoulli” family distribution was used because occurrence data were binary (i.e., presence and absence). Tree complexity refers to how many splits occur in a single regression tree, with each split considering increasing covariate interactions. At the same time, the learning rate determines the contribution amount of each tree to the ensembled model. Maximum trees generally reflect tree complexity and learning rate, as reducing these values requires additional trees to achieve model convergence. Bag fraction is the subset of training data randomly selected, without replacement, for tree construction, which implements stochastic gradient descent and reduces the opportunity for overfitting. Settings for the three SDMs were chosen to support model convergence and set as follows: tc = 3, lr = 0.0001, bf = 0.75, and mt = 100,000.

Relative variable influence (RVI) and partial dependency plots (PDPs) for predictor variables were used to visualize average variable contribution and relationships for model ensembles. The RVI is a measurement of whether a variable was selected in tree development and tree ensemble improvement (i.e., squared error decrease) through variable inclusion. All variables in the model contribute to tree construction, though some more than others. PDPs illustrate the marginal effects of the predictors on prediction outcomes of the model, including potential complex relationships between the predictors and response (Friedman, 2001; Molnar, 2020). The PDP curve for each variable is produced while holding the other variables at their average value, but can be skewed due to collinearity, which we reduced through the

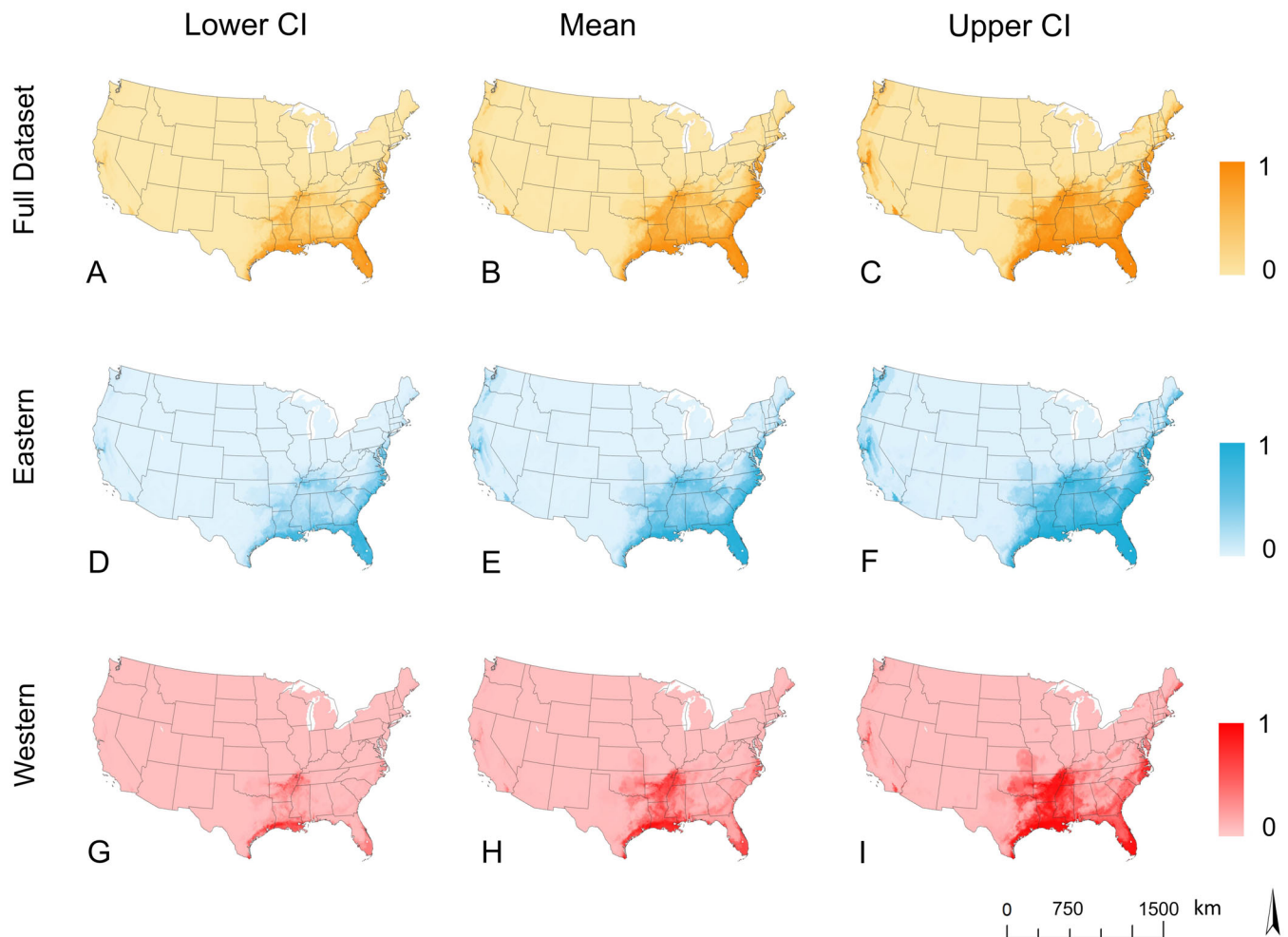
VIF process. The resulting RVIs and PDPs were averaged for each run of 50 ensemble models.

Averaged landscape-projected probabilities from each model run (i.e., full dataset, eastern clade, and western clade models) were recoded as binary geographic distributions to visualize and compare output across models. Rasters of continuous model output were dichotomized at three different probability thresholds of presence (0.5, 0.6, and 0.7) in ArcMap (ver 10.7.1), where pixels with probabilities exceeding a given threshold were recoded as present. Recoded distributions were combined using the “Raster Calculator” tool in ArcMap’s Spatial Analyst extension, allowing for visualization of range overlaps across models.

## RESULTS

The predicted geographic distribution of *O. palustris* primarily spans the southeastern United States and along the eastern Atlantic coastline. Mean landscape predictions for

the three models (Figure 3B,E,H) primarily occurred in four US regions: the West Coast, the Southeast, the Gulf Coast, and the East Coast. The West Coast, including Washington, Oregon, and California, was primarily represented in the full dataset model and eastern clade model with maximum probabilities of 0.817 and 0.821, but the western clade model projected lower probabilities (maximum probability = 0.379) on the West Coast. The western clade model also predicted a smaller area of suitability on the West Coast than the other models. In all three models, the predicted Southeast range spanned Tennessee, Arkansas, Mississippi, Louisiana, Alabama, and Florida, extending north to the southern tip of Illinois and West Virginia, and extending west into Texas, Oklahoma, and Kansas. The full dataset (maximum probability = 0.834) and western clade (maximum probability = 0.826) models had the highest predicted suitable habitat in southeastern states compared with the eastern clade (maximum probability = 0.744). Although the western clade model prediction probability does reduce eastward, the other



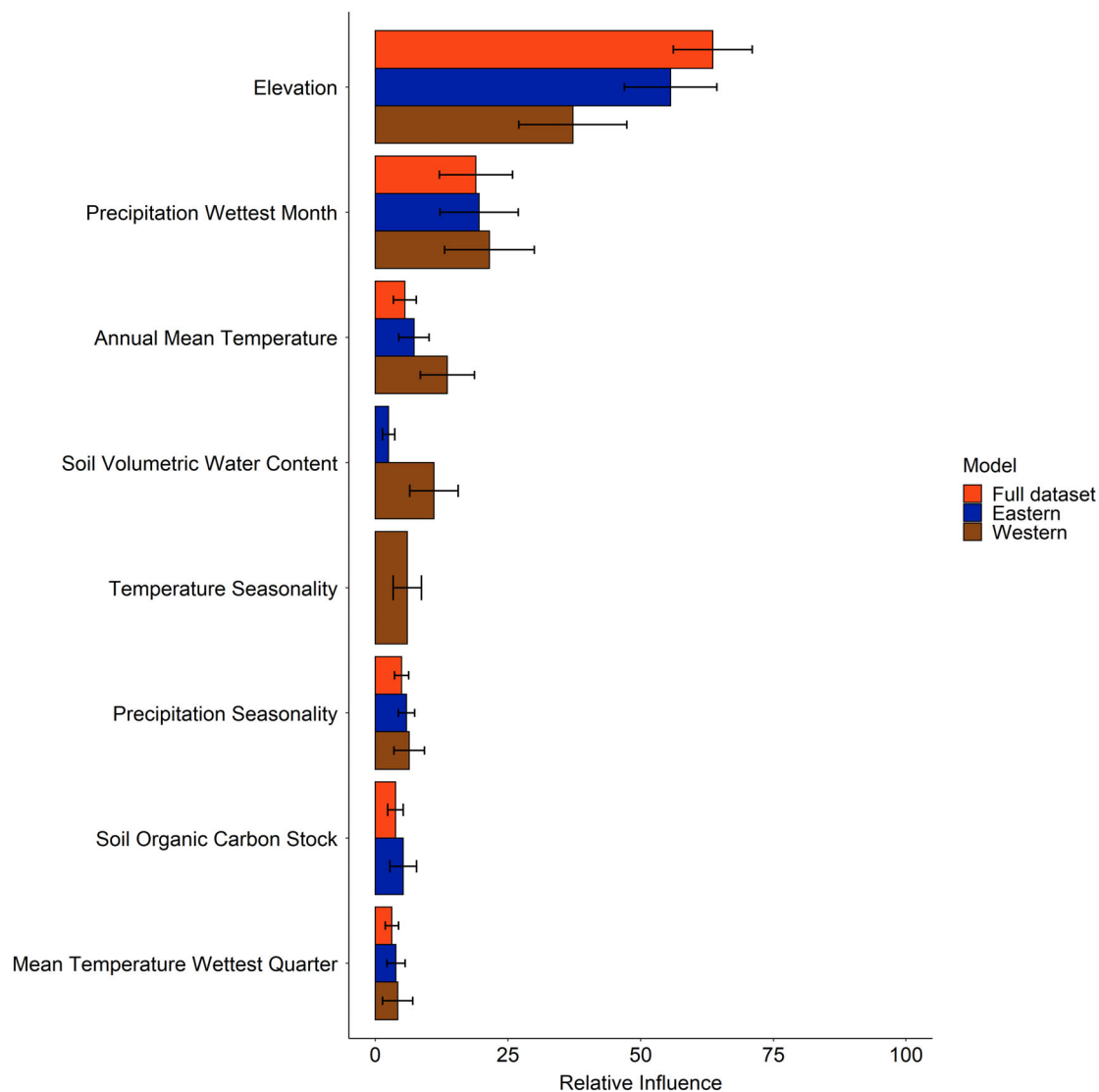
**FIGURE 3** Potential distribution prediction of *Oryzomys palustris* full dataset (A–C), eastern clade (D–F), and western clade (G–I) occurrence in the United States based on the mean prediction of an ensemble of 50 boosted regression tree experiments including lower (2.5%; left) and upper (97.5%; right) confidence intervals.



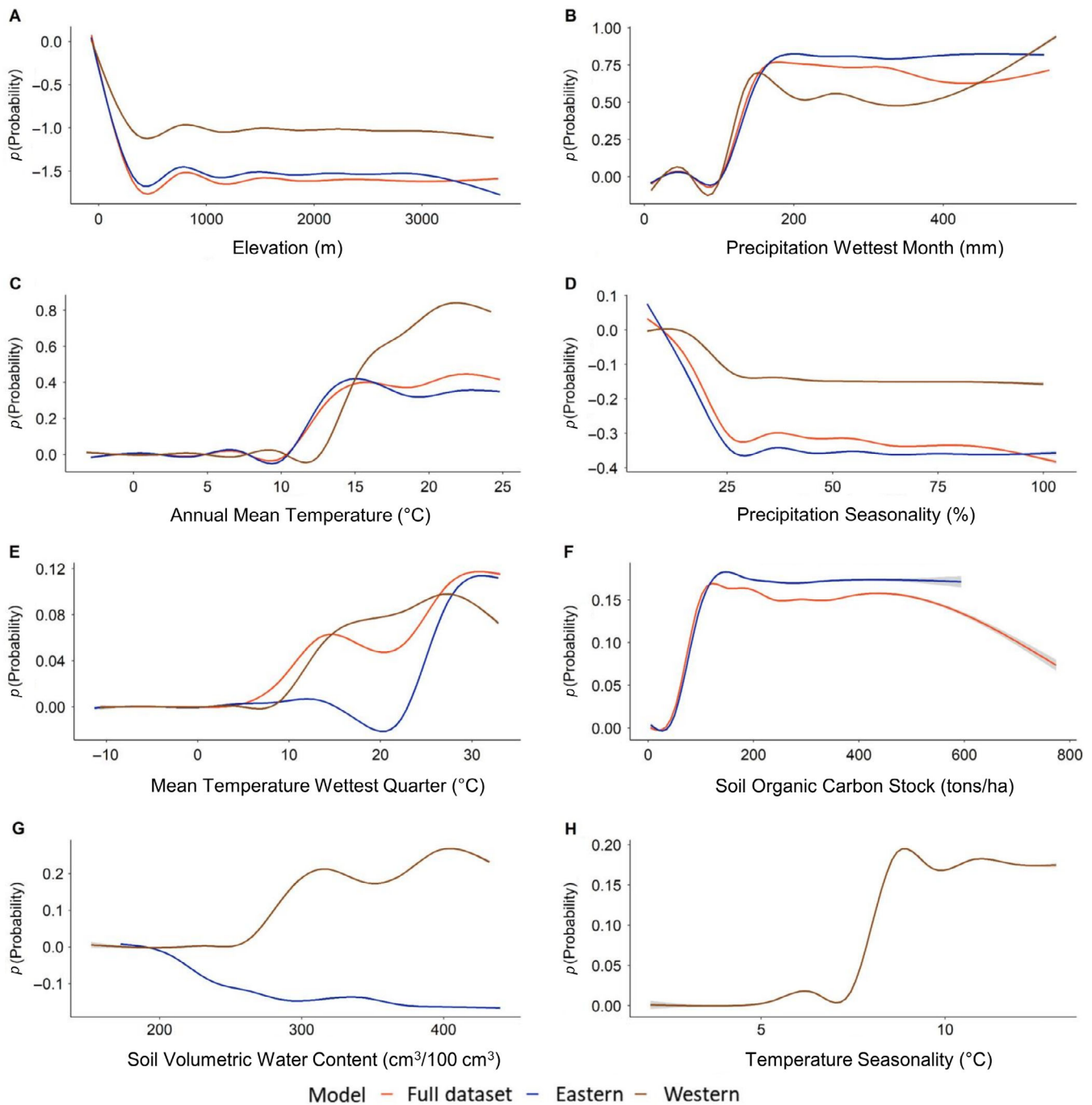
models maintain predictive consistency across the region. Maximum predicted probabilities for the full dataset (0.917) and eastern (0.916) models are the same for the Gulf Coast and East Coast regions. These models also have consistent prediction probabilities from the Gulf Coast of Texas to Maryland and taper off further north. The western clade model predicts the Texas and Louisiana sections of the Gulf Coast with a maximum probability of 0.928, but probabilities are not consistent towards Florida as in the other models, and the maximum probability for the East Coast is 0.708.

All three models had similar responses to the covariates consistent with occurrence data partitioning and overlap between models (Figure 4); however, differences in the magnitude of relationships between predicted probabilities

and environmental predictors were apparent, especially between the western clade and the other two models (Figure 5). Elevation was the top contributing environmental predictor in all three models; prediction probability was highest at sea level and rapidly decreased with increasing elevation. The western clade model plateaued in prediction probability reduction earlier than the other models. Precipitation of the wettest month and annual mean temperature were the next highest contributors. All models generally exhibited positive relationships with increasing precipitation and temperature beginning around 100 mm and 10°C, respectively. There were some distinctions in the relationship with soil volumetric water content between the eastern and western models, as the former maintained a negative relationship while the



**FIGURE 4** Variable relative influence for the final variable set used to model the distribution for the full *Oryzomys palustris* dataset, eastern clade, and western clade in the United States using boosted regression tree (BRT) experiments. Error bars represent variability across an ensemble of 50 BRT experiments.



**FIGURE 5** Partial dependency plots showing marginal effects on the prediction probability of the potential distribution of *Oryzomys palustris* models of the full dataset, eastern clade, and western clade by each variable across the 50 boosted regression tree experiments.

latter had a positive relationship. However, this variable contributed minimally to the eastern model. The BRT simplification process excluded soil volumetric water content from the full dataset model and soil organic carbon stock from the western model (Figure 4). Temperature seasonality was excluded from the full dataset and eastern models (Figure 4).

All three models performed well with cross-validation, internal AUC values, and AUC values from external test

data that remained above 0.9. The mean value for the full dataset model cross-validation AUC was 0.957 (SD = 0.007), internal test data AUC was 0.959 (SD = 0.015), and external test data AUC was 0.939 (SD = 0.005). The mean AUC of the eastern clade for cross-validation AUC was 0.960 (SD = 0.008), internal test AUC was 0.958 (SD = 0.013), and external test data AUC was 0.983 (SD = 0.004). The western clade model did have marginally higher standard deviations with AUCs of 0.965

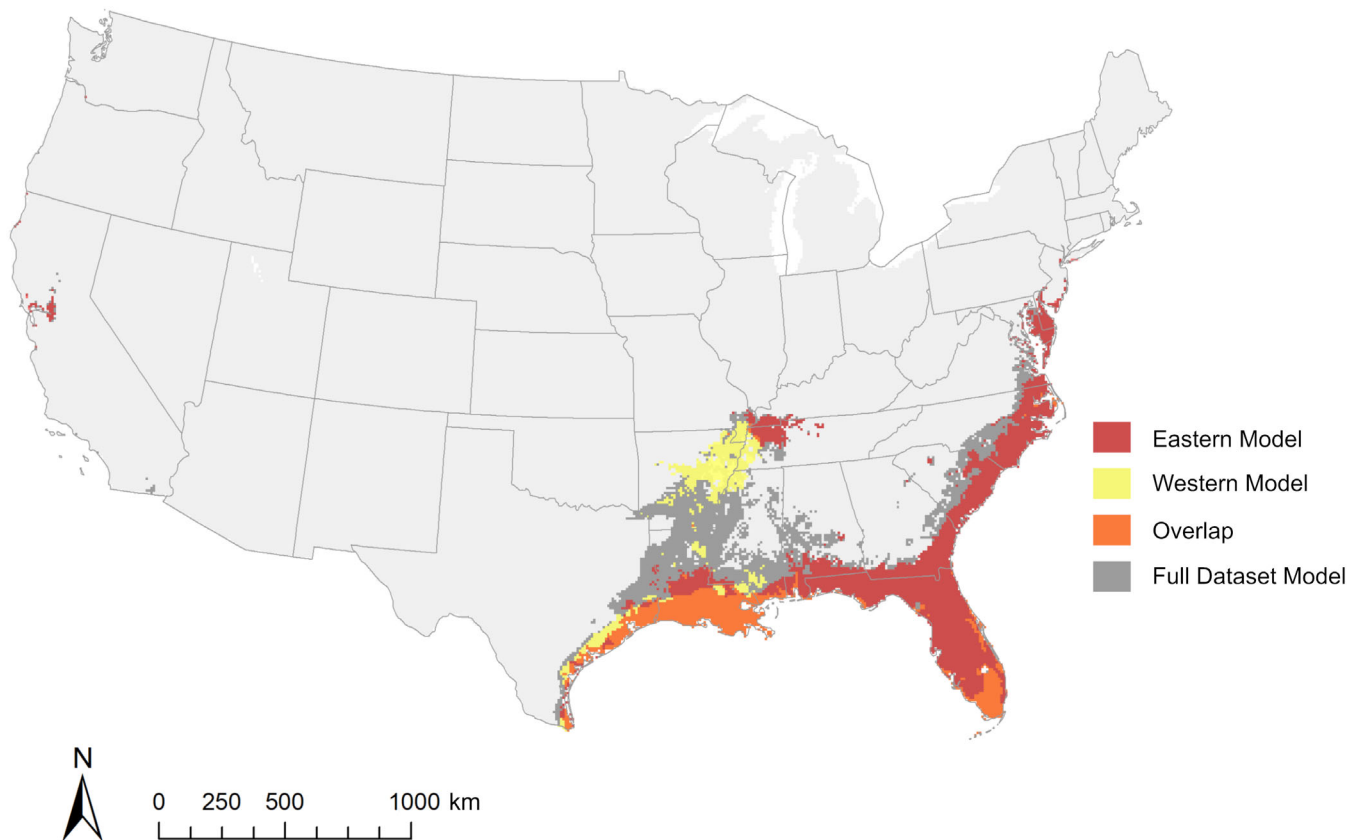
(SD = 0.009) for cross-validation, 0.964 (SD = 0.018) for internal test data, and 0.989 (SD = 0.011) for external test data. The accuracy metrics for all three models were similar, although the AUC value for external testing data from the full dataset model was the lowest of the three models.

Dichotomized probabilities (at 0.5, 0.6, and 0.7) of mean landscape predictions for rice rat habitat suitability indicate the East Coast region extending from Maryland to Florida, peninsular Florida, the lower Mississippi River Basin, and the Gulf Coast region extending from Florida to Texas have the highest potential for rice rat species occurrence in all models (Figure 6; Appendix S1: Figures S1 and S2). The full dataset and eastern clade models predict occurrence across this entire area, with predictions of suitability extending further inland to Arkansas, Tennessee, Missouri, and Kentucky, and a small area on the West Coast (Figure 6; Appendix S1: Figures S1 and S2). Modeled distribution of the western clade is more restricted, by comparison, with predicted suitability primarily centered in Louisiana, Texas, and along the Mississippi River (Figure 6; Appendix S1: Figures S1 and S2). Distribution models for each type of data partitioning (full dataset, eastern clade, and western clade) overlap in mean prediction along the western

portion of the Gulf Coast, primarily in Texas, Louisiana, and southern peninsular Florida (Figure 6; Appendix S1: Figures S1 and S2).

## DISCUSSION

This work is the first effort to apply SDM techniques to estimate the geographic distribution of the *O. palustris* species complex. The predicted distributions of all three taxonomic groupings were found to be predominantly restricted to coastal areas in the southeastern United States. Additionally, the estimated suitable habitat for both *O. palustris* taxonomic groupings (i.e., the full dataset and the eastern subset) extends into the southern portion of the Mid-Atlantic region, with a discontinuous, limited area of suitability in coastal California. In contrast, habitat suitability for the western clade is primarily highest along the Gulf Coast, extending west into coastal Louisiana and Texas. Further, output across models indicates a corridor of suitable habitat that extends north, throughout Louisiana, and into Arkansas, Mississippi, Tennessee, Kentucky, and southeastern Missouri, an area that coincides with the lower Mississippi River Basin.



**FIGURE 6** Overlap between mean dichotomized predictions of *Oryzomys palustris* suitability at a threshold for presence of 60% for models produced with the three datasets used in the study (i.e., full dataset, eastern clade, and western clade).

These spatial trends in estimated habitat suitability are perhaps unsurprising, as *O. palustris* is a semiaquatic rodent species that is primarily associated with both freshwater and estuarine wetlands, including tidal marshes, watersheds, and emergent wetlands (Eubanks et al., 2011; Kruchek, 2004). Rice rats are heavily dependent on the vegetation typically associated with coastal wetlands, including common wetland plants such as sedges (*Carex* spp.), rushes (*Juncus* spp.), cattails (*Typha* spp.), panic grass (*Panicum* spp.), and salt grass (*Spartina alterniflora*) (Eubanks et al., 2011; Rose & McGurk, 2006; Svihla, 1931). Suitable vegetation is a critical microhabitat feature for *O. palustris*, directly providing essential resources as both cover and food (Eubanks et al., 2011; Kruchek, 2004; Rose & McGurk, 2006). We were limited in this study by the inability to directly include environmental predictors related to land cover, as few *O. palustris* collections were made after the year 2000 in our datasets. Future work that ties contemporary rice rat observations with current land cover will enable us to assess reliance on vegetation type with greater resolution.

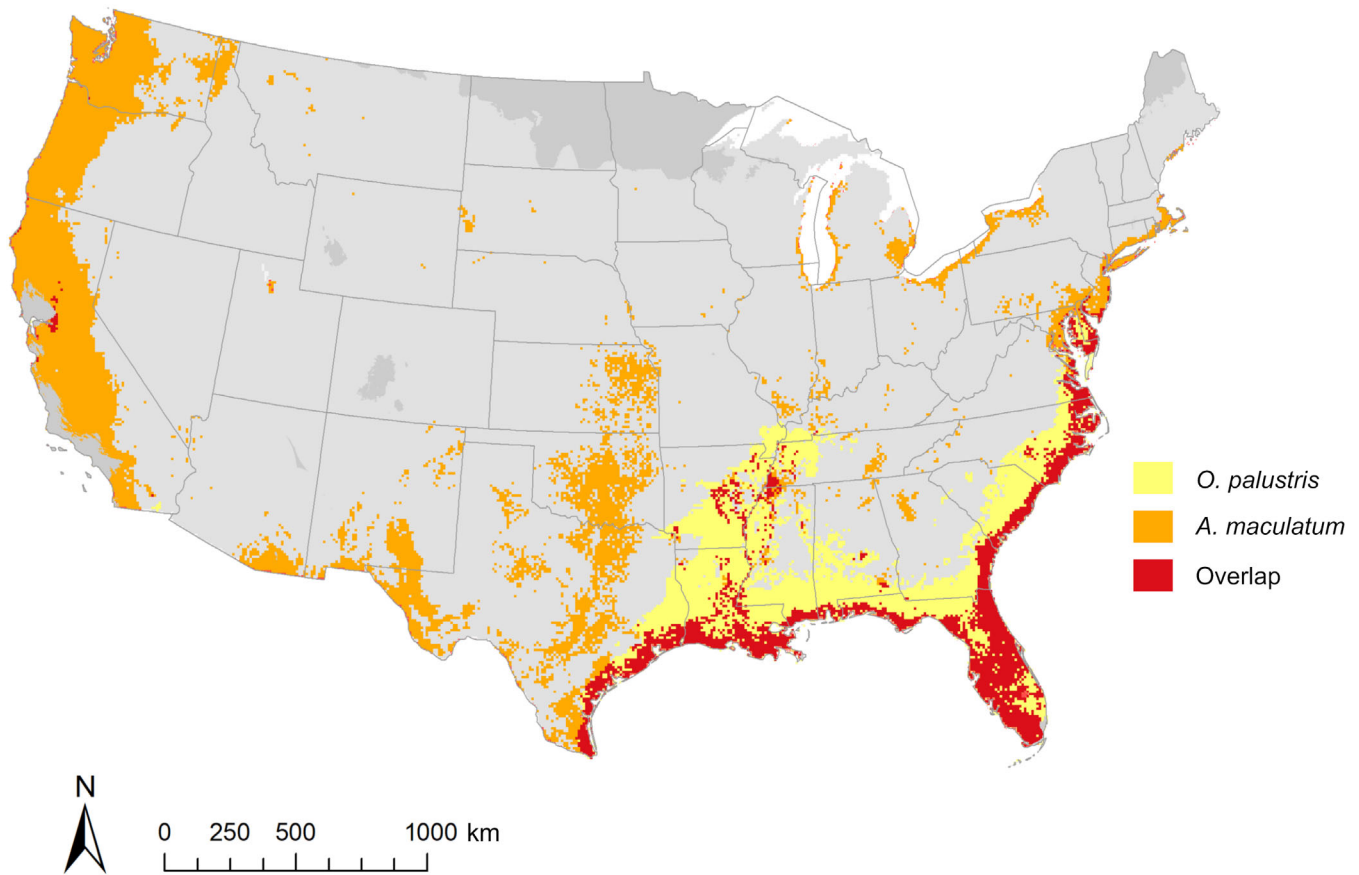
Elevation was the environmental predictor that contributed the most to estimated rice rat distributions, regardless of taxonomic clade. The highest probabilities of occurrence coincide with areas of low elevation at or near sea level, predominantly along coastlines in the Southeast and Gulf states. Precipitation of the wettest month and annual mean temperature were also influential factors in all three models, where higher precipitation values and temperatures exceeding  $\sim 13^{\circ}\text{C}$  were associated with increasing habitat suitability. Together, elevation, precipitation, and temperature are likely capturing qualities of low-elevation marsh and wetland habitats that are necessary to maintain wetland plant communities and support rice rat populations. When considering other potential environmental drivers of rice rat populations, we start to see differences between the models for different taxonomic groupings. Low precipitation seasonality and soil organic carbon stock exceeding  $\sim 100$  tons/ha are associated with increased probability of occurrence in the full *O. palustris* and eastern clade models. Again, these are factors that may support stable and robust plant communities in wetlands. In contrast, precipitation seasonality did not cause the same suitability drop-off in the western clade model. Instead, soil volumetric water content and temperature seasonality become important environmental predictors, indicating that water availability and temperature are more limiting factors for western rice rat populations.

Estimating the range of *O. palustris* is an important step forward in delineating the general distribution of potential hosts for *A. maculatum* and the rickettsial pathogens they transmit. Reservoir hosts involved in tick-borne transmission cycles are targets for control efforts and may serve as

better indicators of risk than vector presence alone, particularly for ticks with widespread distributions. However, vertebrate hosts involved in many tick-borne transmission cycles are poorly understood or unknown. This is particularly true for SFG pathogens, many of which have only recently garnered increased scrutiny from medical and public health authorities, relative to other tick-borne diseases such as Lyme disease, where pathogen–host–vector dynamics are extensively studied (Halsey et al., 2018; Ostfeld, 2012). Components of the transmission cycle for *R. parkeri* are not currently known, but several small mammal species are suspected to play a role in the pathogen's ecological maintenance, based on field surveys and laboratory experiments (Cumbie et al., 2020; Moraru et al., 2013). Investigation of the potential role of *O. palustris* as a reservoir host for *R. parkeri* is warranted. In addition to detection of *R. parkeri* in field-collected *O. palustris*, other closely related species of the genus *Oryzomys* have been established as competent hosts of rickettsial pathogens (Krawczak & Labruna, 2018; Padilha et al., 2013). Although the marsh rice rat has suspected involvement, further work is needed to establish the competency of *O. palustris* as a host capable of supporting the *R. parkeri* transmission cycle.

There is considerable geographic overlap with *O. palustris* and *A. maculatum* in the southeastern United States. Comparing the model of our full *O. palustris* dataset to an SDM for *A. maculatum* produced by Flenniken et al. (2022), most of the overlap seen in estimated distributions occurs along the Gulf Coast, peninsular Florida, and along the East Coast into Virginia, Maryland, and New Jersey, with some incursion along the lower Mississippi River Basin (Figure 7; Appendix S1: Figures S3 and S4). The availability of suitable hosts may be an important element of successful range expansion of *A. maculatum*, which has drastically expanded the northern extent of its range in recent years (Molaei et al., 2021; Nadolny & Gaff, 2018; Wojan et al., 2022). Notably, much of this northward expansion has occurred along the East Coast, a potential indicator of the role that underlying host species like *O. palustris* play in the habitat suitability of tick vectors. The ranges of both *O. palustris* and *A. maculatum* also coincide with much of the estimated distribution of *R. parkeri* in North America, as described in Moo-Llanes et al. (2021). Moo-Llanes et al. built SDMs with records of ticks, comprised of several *Amblyomma* species, that tested positive for *R. parkeri* bacteria. The resulting model shows high predicted suitability in North America along the Gulf of Mexico and portions of the East Coast, with moderate to high predicted suitability extending inland throughout the eastern United States.

In addition to *O. palustris*, two other rodents in the family Cricetidae have also been implicated as hosts of *A. maculatum*, with potential involvement in the



**FIGURE 7** Overlap between mean dichotomized (threshold of 60%) full dataset model and predicted distribution *Amblyomma maculatum*, adapted from Flenniken et al. (2022).

transmission cycle for *R. parkeri*: the hispid cotton rat (*Sigmodon hispidus*) and the eastern meadow vole (*Microtus pennsylvanicus*) (Cumbie et al., 2020). The geographic range of *S. hispidus* overlaps with much of our estimated distribution for *O. palustris*, extending further north into Nebraska and west into California, whereas the range of *M. pennsylvanicus* has some overlap with the northern extent of the *O. palustris* distributions, but extends north into Canada and Alaska (Jackson & Cook, 2020; USGS, 2018). Although *S. hispidus* has been documented as an important blood-meal host for immature *A. maculatum*, lab experiments have demonstrated that this species quickly clears *R. parkeri* infection, making them unlikely reservoir hosts (Baker et al., 2003; Johnson et al., 2022; Moraru et al., 2013). Likewise, *M. pennsylvanicus* is not a likely reservoir host of *R. parkeri*. While infected *A. maculatum* have been collected from *M. pennsylvanicus*, the ranges of the rodent and tick have minimal historical overlap (Cumbie et al., 2020; Nadolny & Gaff, 2018). Nevertheless, these three rodent species collectively span a vast area, potentially facilitating the expansion of ticks that utilize these species as hosts and providing opportunities for pathogen spillover events.

The underlying distribution of reservoir hosts is suspected to play a role in the geographic risk of tick-borne disease transmission. Tick-borne pathogens can exhibit geographically restricted ranges that are “nested” within the larger distribution of competent vectors. For example, within the eastern portion of the United States, foci for *Borrelia burgdorferi*, the spirochete that causes Lyme disease, are restricted within the range of its vectors, primarily *Ixodes scapularis* (Fleshman et al., 2021). The existence of an “infected” niche of ticks that tested positive for *R. montanensis* infections within the full geographic range of its vector, *Dermacentor variabilis*, has also been supported with SDMs (Lippi, Gaff, White, St. John, et al., 2021). The absence of pathogen transmission can be due to many factors, and certainly, vector presence alone is not sufficient for transmission. Absence of pathogens could be a result of undersampling or a lack of robust surveillance for tick-borne disease. Differences in suitable environmental conditions for tick presence and pathogen transmission may also contribute to range discrepancies (Galletti et al., 2016). However, the importance of reservoir hosts in maintaining vector-borne transmission cycles is widely acknowledged, yet in the case of many tick-borne diseases, not well understood. Even so, there is increasing

evidence to support the key role that reservoir and amplifying hosts play in the localized density of infected ticks, and risk of transmission for an array of tick-borne diseases (Krawczyk et al., 2020). This has notably been the case with Lyme disease, where increasing burden and range expansion of the disease have been linked to range shifts and population changes in vertebrate hosts (Levi et al., 2012; Thompson et al., 2001).

## CONCLUSION

Conclusive identification of reservoir and amplifying hosts is still needed to establish transmission cycles for many tick-borne diseases, including the majority of those caused by rickettsial agents. Given that many species of medically important ticks have generalized habitat requirements, or are actively expanding their geographic ranges, identifying areas where tick presence overlaps with suitable reservoir habitat may be necessary for refining risk maps and public health advisories derived solely from tick presence. In this study, we estimated the geographic range of *O. palustris*, a suspected reservoir host of *R. parkeri*, and found considerable overlap with the distribution of *A. maculatum*. Though future work is needed to investigate the definitive role of *O. palustris* in transmission cycles, our findings highlight the potential need for extended surveillance efforts in those overlapping areas.

## ACKNOWLEDGMENTS

Catherine A. Lippi, Holly D. Gaff, Samuel Canfield, and Sadie J. Ryan were funded by NIH 1R01AI136035-01 as part of the joint NIH-NSF-USDA Ecology and Evolution of Infectious Diseases program. Catherine A. Lippi, Samuel Canfield, and Sadie J. Ryan were additionally funded by CDC grant 1U01CK000510-01: Southeastern Regional Center of Excellence in Vector-Borne Diseases: The Gateway Program. Catherine A. Lippi and Sadie J. Ryan were also funded by NSF 2016265. We thank J. Matthew Flenniken for providing the Gulf Coast tick distribution model output used in comparisons. This publication was supported by the cooperative agreement number above from the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The occurrence data used to build species distribution models in this study are openly available on the Global

Biodiversity Information Facility ([www.gbif.org](http://www.gbif.org)), where we pulled North American records of *Oryzomys palustris* (<https://doi.org/10.15468/dl.v9j4x5>).

## ORCID

Catherine A. Lippi  <https://orcid.org/0000-0002-7988-0324>

Holly D. Gaff  <https://orcid.org/0000-0002-4034-2684>

Sadie J. Ryan  <https://orcid.org/0000-0002-4308-6321>

## REFERENCES

- Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. "spThin: An R Package for Spatial Thinning of Species Occurrence Records for Use in Ecological Niche Models." *Ecography* 38: 541–5.
- Baker, R. J., L. C. Bradley, R. D. Bradley, J. W. Dragoo, M. D. Engstrom, R. S. Hoffmann, C. A. Jones, F. Reid, D. W. Rice, and C. Jones. 2003. "Revised Checklist of North American Mammals North of Mexico, 2003." *Occasional Papers, Museum of Texas Tech University* 229: 1–23.
- Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. "Selecting Pseudo-Absences for Species Distribution Models: How, where and how Many? How to Use Pseudo-Absences in Niche Modelling?" *Methods in Ecology and Evolution* 3: 327–38.
- CDC. 2021. "Lyme and Other Tickborne Diseases Increasing." Centers for Disease Control and Prevention. <https://www.cdc.gov/ncezid/dvbd/media/lyme-tickborne-diseases-increasing.html>.
- Cerasoli, F., M. Iannella, P. D'Alessandro, and M. Biondi. 2017. "Comparing Pseudo-Absences Generation Techniques in Boosted Regression Trees Models for Conservation Purposes: A Case Study on Amphibians in a Protected Area." *PLoS One* 12: e0187589.
- Cumbie, A. N., C. D. Espada, R. M. Nadolny, R. K. Rose, R. D. Dueser, W. L. Hynes, and H. D. Gaff. 2020. "Survey of *Rickettsia parkeri* and *Amblyomma maculatum* Associated with Small Mammals in Southeastern Virginia." *Ticks and Tick-borne Diseases* 11: 101550.
- Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, et al. 2013. "Collinearity: A Review of Methods to Deal with it and a Simulation Study Evaluating their Performance." *Ecography* 36: 27–46.
- Elith, J., J. R. Leathwick, and T. Hastie. 2008. "A Working Guide to Boosted Regression Trees." *Journal of Animal Ecology* 77: 802–13.
- Eubanks, B. W., E. C. Hellgren, J. R. Nawrot, and R. D. Bluett. 2011. "Habitat Associations of the Marsh Rice Rat (*Oryzomys palustris*) in Freshwater Wetlands of Southern Illinois." *Journal of Mammalogy* 92: 552–60.
- Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37: 4302–15.
- Flenniken, J. M., H. C. Tuten, H. Rose Vineer, V. C. Phillips, C. M. Stone, and B. F. Allan. 2022. "Environmental Drivers of Gulf Coast Tick (Acari: Ixodidae) Range Expansion in the United States." *Journal of Medical Entomology* 59(5): 1625–35.
- Fleshman, A. C., C. B. Graham, S. E. Maes, E. Foster, and R. J. Eisen. 2021. "Reported County-Level Distribution of Lyme

- Disease Spirochetes, *Borrelia burgdorferi sensu stricto* and *Borrelia mayonii* (Spirochaetales: Spirochaetaceae), in Host-Seeking *Ixodes scapularis* and *Ixodes pacificus* Ticks (Acari: Ixodidae) in the Contiguous United States.” *Journal of Medical Entomology* 58: 1219–33.
- Fornadel, C. M., X. Zhang, J. D. Smith, C. D. Paddock, J. R. Arias, and D. E. Norris. 2011. “High Rates of *Rickettsia parkeri* Infection in Gulf Coast Ticks (*Amblyomma maculatum*) and Identification of “*Candidatus* *Rickettsia Andeanae*” from Fairfax County, Virginia.” *Vector-Borne and Zoonotic Diseases* 11: 1535–9.
- Friedman, J. H. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” *The Annals of Statistics* 29(5): 1189–232.
- Galletti, M. F. B. M., A. Fujita, R. D. Rosa, L. A. Martins, H. S. Soares, M. B. Labruna, S. Daffre, and A. C. Fogaça. 2016. “Virulence Genes of *Rickettsia rickettsii* Are Differentially Modulated by either Temperature Upshift or Blood-Feeding in Tick Midgut and Salivary Glands.” *Parasites & Vectors* 9: 331.
- Halsey, S. J., B. F. Allan, and J. R. Miller. 2018. “The Role of *Ixodes scapularis*, *Borrelia burgdorferi* and Wildlife Hosts in Lyme Disease Prevalence: A Quantitative Review.” *Ticks and Tick-borne Diseases* 9(5): 1103–14.
- Hanson, J. D., J. L. Indorf, V. J. Swier, and R. D. Bradley. 2010. “Molecular Divergence within the *Oryzomys palustris* Complex: Evidence for Multiple Species.” *Journal of Mammalogy* 91: 336–47.
- Hengl, T., J. Mendes de Jesus, G. B. Heuvelink, M. Ruiperez Gonzalez, M. Kilibarda, A. Blagotić, W. Shangguan, M. N. Wright, X. Geng, and B. Bauer-Marschallinger. 2017. “SoilGrids250m: Global Gridded Soil Information Based on Machine Learning.” *PLoS One* 12: e0169748.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. “Very High Resolution Interpolated Climate Surfaces for Global Land Areas.” *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25: 1965–78.
- Humphrey, S. R., and H. W. Setzer. 1989. “Geographic Variation and Taxonomic Revision of Rice Rats (*Oryzomys palustris* and *O. argentatus*) of the United States.” *Journal of Mammalogy* 70: 557–70.
- Jackson, D. J., and J. A. Cook. 2020. “A Precarious Future for Distinctive Peripheral Populations of Meadow Voles (*Microtus pennsylvanicus*).” *Journal of Mammalogy* 101: 36–51.
- Johnson, C. R., L. Ponnusamy, A. L. Richards, and C. S. Apperson. 2022. “Analyses of Bloodmeal Hosts and Prevalence of *Rickettsia parkeri* in the Gulf Coast Tick *Amblyomma maculatum* (Acari: Ixodidae) from a Reconstructed Piedmont Prairie Ecosystem, North Carolina.” *Journal of Medical Entomology* 59: 1382–93.
- Krawczak, F. S., and M. B. Labruna. 2018. “The Rice Rat *Euryoryzomys russatus*, a Competent Amplifying Host of *Rickettsia parkeri* Strain Atlantic Rainforest for the Tick *Amblyomma ovale*.” *Ticks and Tick-borne Diseases* 9: 1133–6.
- Krawczyk, A. I., G. L. A. van Duijvendijk, A. Swart, D. Heylen, R. I. Jaarsma, F. H. H. Jacobs, M. Fonville, H. Sprong, and W. Takken. 2020. “Effect of Rodent Density on Tick and Tick-Borne Pathogen Populations: Consequences for Infectious Disease Risk.” *Parasites & Vectors* 13: 34.
- Kruchek, B. L. 2004. “Use of Tidal Marsh and Upland Habitats by the Marsh Rice Rat (*Oryzomys palustris*).” *Journal of Mammalogy* 85: 569–75.
- Levi, T., A. M. Kilpatrick, M. Mangel, and C. C. Wilmers. 2012. “Deer, Predators, and the Emergence of Lyme Disease.” *Proceedings of the National Academy of Sciences of the United States of America* 109: 10942–7.
- Lippi, C. A., H. D. Gaff, A. L. White, and S. J. Ryan. 2021. “Scoping Review of Distribution Models for Selected *Amblyomma* Ticks and Rickettsial Group Pathogens.” *PeerJ* 9: e10596.
- Lippi, C. A., H. D. Gaff, A. L. White, H. K. St. John, A. L. Richards, and S. J. Ryan. 2021. “Exploring the Niche of *Rickettsia montanensis* (Rickettsiales: Rickettsiaceae) Infection of the American Dog Tick (Acari: Ixodidae), Using Multiple Species Distribution Model Approaches.” *Journal of Medical Entomology* 58: 1083–92.
- Lippi, C. A., S. J. Ryan, A. L. White, H. D. Gaff, and C. J. Carlson. 2021. “Trends and Opportunities in Tick-Borne Disease Geography.” *Journal of Medical Entomology* 58: 2021–9.
- Madison-Antenucci, S., L. D. Kramer, L. L. Gebhardt, and E. Kauffman. 2020. “Emerging Tick-Borne Diseases.” *Clinical Microbiology Reviews* 33(2): e00083–18.
- Maréchaux, I., A. S. L. Rodrigues, and A. Charpentier. 2017. “The Value of Coarse Species Range Maps to Inform Local Biodiversity Conservation in a Global Context.” *Ecography* 40: 1166–76.
- Molaei, G., E. A. H. Little, N. Khalil, B. N. Ayres, W. L. Nicholson, and C. D. Paddock. 2021. “Established Population of the Gulf Coast Tick, *Amblyomma maculatum* (Acari: Ixodidae), Infected with *Rickettsia parkeri* (Rickettsiales: Rickettsiaceae), in Connecticut.” *Journal of Medical Entomology* 58: 1459–62.
- Molnar, C. 2020. *Interpretable Machine Learning: A Guide for Making Black Box Methods Interpretable*. Munich: Christoph Molnar.
- Moo-Llanes, D. A., A. C. Montes de Oca-Aguilar, D. Romero-Salas, and S. Sanchez-Montes. 2021. “Inferring the Potential Distribution of Emerging Rickettsiosis in America: The Case of *Rickettsia parkeri*.” *Pathogens* 10(5): 592.
- Moraru, G. M., J. Goddard, C. D. Paddock, and A. Varela-Stokes. 2013. “Experimental Infection of Cotton Rats and Bobwhite Quail with *Rickettsia parkeri*.” *Parasites & Vectors* 6: 70.
- Murphey, P. C., R. P. Guralnick, R. Glaubitz, D. Neufeld, and J. A. Ryan. 2004. “Georeferencing of Museum Collections: A Review of Problems and Automated Tools, and the Methodology Developed by the Mountain and Plains Spatio-Temporal Database-Informatics Initiative.” *PhyloInformatics* 3: 1–29.
- Nadolny, R. M., and H. D. Gaff. 2018. “Natural History of *Amblyomma maculatum* in Virginia.” *Ticks and Tick-borne Diseases* 9: 188–95.
- O’Brien, R. M. 2007. “A Caution Regarding Rules of Thumb for Variance Inflation Factors.” *Quality & Quantity* 41: 673–90.
- Ogden, N. H., S. Mechai, and G. Margos. 2013. “Changing Geographic Ranges of Ticks and Tick-Borne Pathogens: Drivers, Mechanisms and Consequences for Pathogen Diversity.” *Frontiers in Cellular and Infection Microbiology* 29(3): 46.
- Ostfeld, R. S. 2012. *Lyme Disease: The Ecology of a Complex System*. New York: Oxford University Press.
- Otieno, F. T., J. Gachohi, P. Gikuma-Njuru, P. Kariuki, H. Oyas, S. A. Canfield, J. K. Blackburn, M. K. Njenga, and B. Bett. 2021. “Modeling the Spatial Distribution of Anthrax in Southern Kenya.” *PLoS Neglected Tropical Diseases* 15: e0009301.

- Paddock, C. D., and J. Goddard. 2015. "The Evolving Medical and Veterinary Importance of the Gulf Coast Tick (Acari: Ixodidae)." *Journal of Medical Entomology* 52: 230–52.
- Paddock, C. D., J. W. Sumner, J. A. Comer, S. R. Zaki, C. S. Goldsmith, J. Goddard, S. L. F. McLellan, C. L. Tamminga, and C. A. Ohl. 2004. "*Rickettsia parkeri*: A Newly Recognized Cause of Spotted Fever Rickettsiosis in the United States." *Clinical Infectious Diseases* 38: 805–11.
- Padilha, A. F., C. L. Mafra, B. S. Milagres, M. B. Labruna, M. A. M. Galvão, R. N. Freitas, C. E. Montandon, D. H. Walker, and R. Pacheco. 2013. "Spotted Fever Group Rickettsia in Small Rodents from Areas of Low Endemicity for Brazilian Spotted Fever in the Eastern Region of Minas Gerais State, Brazil." *The American Journal of Tropical Medicine and Hygiene* 88: 937–9.
- Phillips, S. J., M. Dudík, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. "Sample Selection Bias and Presence-Only Distribution Models: Implications for Background and Pseudo-Absence Data." *Ecological Applications* 19: 181–97.
- Phillips, V. C., E. A. Ziemann, C.-H. Kim, C. M. Stone, H. C. Tuten, and F. A. Jiménez. 2020. "Documentation of the Expansion of the Gulf Coast Tick (*Amblyomma maculatum*) and *Rickettsia parkeri*: First Report in Illinois." *Journal of Parasitology* 106(1): 9–13.
- Poggio, L., L. M. De Sousa, N. H. Batjes, G. Heuvelink, B. Kempen, E. Ribeiro, and D. Rossiter. 2021. "SoilGrids 2.0: Producing Soil Information for the Globe with Quantified Spatial Uncertainty." *Soil* 7: 217–40.
- Ramírez-Garofalo, J. R., S. R. Curley, C. E. Field, C. E. Hart, and S. Thangamani. 2022. "Established Populations of *Rickettsia parkeri*- Infected *Amblyomma maculatum* Ticks in New York City, New York, USA." *Vector-Borne and Zoonotic Diseases* 22: 184–7.
- Rochlin, I., and A. Toledo. 2020. "Emerging Tick-Borne Pathogens of Public Health Importance: A Mini-Review." *Journal of Medical Microbiology* 69: 781–91.
- Rose, R. K., and S. W. McGurk. 2006. "Year-Round Diet of the Marsh Rice Rat, *Oryzomys palustris*, in Virginia Tidal Marshes." *Virginia Journal of Science* 57: 115–21.
- Rosenberg, R., N. P. Lindsey, M. Fischer, C. J. Gregory, A. F. Hinckley, P. S. Mead, G. Paz-Bailey, et al. 2018. "Vital Signs: Trends in Reported Vectorborne Disease Cases—United States and Territories, 2004–2016." *MMWR. Morbidity and Mortality Weekly Report* 67: 496–501.
- Stokland, J. N., R. Halvorsen, and B. Støa. 2011. "Species Distribution Modelling—Effect of Design and Sample Size of Pseudo-Absence Observations." *Ecological Modelling* 222: 1800–9.
- Svihla, A. 1931. "Life History of the Texas Rice Rat (*Oryzomys palustris texensis*)." *Journal of Mammalogy* 12: 238–42.
- Thompson, C., A. Spielman, and P. J. Krause. 2001. "Coinfecting Deer-Associated Zoonoses: Lyme Disease, Babesiosis, and Ehrlichiosis." *Clinical Infectious Diseases* 33: 676–85.
- Tomassone, L., A. Portillo, M. Nováková, R. de Sousa, and J. A. Oteo. 2018. "Neglected Aspects of Tick-Borne Rickettsioses." *Parasites & Vectors* 11: 263.
- USGS. 2018. "Hispid Cotton Rat (*Sigmodon Hispidus*) mHCRAX\_CONUS\_2001v1 Habitat Map." United States Geological Survey. <https://doi.org/10.5066/F7M9071N>.
- Wilson, N., and L. A. Durden. 2003. "Ectoparasites of Terrestrial Vertebrates Inhabiting the Georgia Barrier Islands, USA: An Inventory and Preliminary Biogeographical Analysis: Ectoparasites on the Georgia Barrier Islands." *Journal of Biogeography* 30: 1207–20.
- Wojan, C., T. Thrasher, E. Lacey, and K. Clay. 2022. "Distribution, Dynamics, and Diversity of Questing Ticks in the Lower Midwest." *Journal of Medical Entomology* 59: 273–82.
- Wolfe, J. L. 1982. "*Oryzomys palustris*." *Mammalian Species* 176: 1–5.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Lippi, Catherine A., Samuel Canfield, Christina Espada, Holly D. Gaff, and Sadie J. Ryan. 2023. "Estimating the Distribution of *Oryzomys palustris*, a Potential Key Host in Expanding Rickettsial Tick-Borne Disease Risk." *Ecosphere* 14(3): e4445. <https://doi.org/10.1002/ecs2.4445>