Ecological Predictors of Zoonotic Vector Status Among Dermacentor Ticks (Acari: Ixodidae): A Trait-Based Approach

Authors: Martin, Jessica T., Fischhoff, Ilya R., Castellanos, Adrian A., and Han, Barbara A.

**ENTOMOLOGICAL
SOCIETY OF AMERICA**
SHARING INSECT SCIENCE GLOBALLY

OXFORD

Vector-Borne Diseases, Surveillance, Prevention

Ecological Predictors of Zoonotic Vector Status Among *Dermacentor* Ticks (Acari: Ixodidae): A Trait-Based Approach

Jessica T. Martin,[1,3,] Ilya R. Fischhoff,[2] Adrian A. Castellanos,[2] and Barbara A. Han[2,]

[1]Department of Fish, Wildlife, and Conservation Ecology, New Mexico State University, 2980 South Espina Street, Las Cruces, NM 88003, USA, [2]Cary Institute of Ecosystem Studies, Box AB, Millbrook, NY 12545, USA, and [3]Corresponding author, E-mail: jmartin6@nmsu.edu

Abstract

Increasing incidence of tick-borne human diseases and geographic range expansion of tick vectors elevates the importance of research on characteristics of tick species that transmit pathogens. Despite their global distribution and role as vectors of pathogens such as *Rickettsia* spp., ticks in the genus *Dermacentor* Koch, 1844 (Acari: Ixodidae) have recently received less attention than ticks in the genus *Ixodes* Latreille, 1795 (Acari: Ixodidae). To address this knowledge gap, we compiled an extensive database of *Dermacentor* tick traits, including morphological characteristics, host range, and geographic distribution. Zoonotic vector status was determined by compiling information about zoonotic pathogens found in *Dermacentor* species derived from primary literature and data repositories. We trained a machine learning algorithm on this data set to assess which traits were the most important predictors of zoonotic vector status. Our model successfully classified vector species with ~84% accuracy (mean AUC) and identified two additional *Dermacentor* species as potential zoonotic vectors. Our results suggest that *Dermacentor* species that are most likely to be zoonotic vectors are broad ranging, both in terms of the range of hosts they infest and the range of ecoregions across which they are found, and also tend to have large hypostomes and be small-bodied as immature ticks. Beyond the patterns we observed, high spatial and species-level resolution of this new, synthetic dataset has the potential to support future analyses of public health relevance, including species distribution modeling and predictive analytics, to draw attention to emerging or newly identified *Dermacentor* species that warrant closer monitoring for zoonotic pathogens.

Key words: *Dermacentor*, tick-borne disease, host range, machine learning

Ticks vector the most diverse range of zoonotic pathogens and are responsible for the majority of vector-borne diseases in the United States (Eisen et al. 2017). Among ticks, the best studied species belong to the family Ixodidae (Acari), responsible for vectoring 40% of the 131 emerging vector-borne zoonotic diseases, including Lyme disease, Rocky Mountain spotted fever, and ehrlichiosis, among many others (Swei et al. 2020). Incidences of these diseases are increasing; for example, the number of tick-borne human disease cases doubled between 2004 and 2016 in the United States, primarily due to Lyme disease (Rosenberg et al. 2018). In northern temperate regions in the United States, some tick species are expanding their geographic range northward and westward in part due to climate

change, increasing the chances of zoonotic pathogen transmission to humans (Sonenshine 2018, Couper et al. 2021). While ticks and emerging tick-borne diseases are better studied in the United States, less information is available for Africa and South America (Swei et al. 2020). There is a high diversity of ticks in southeast Asia, yet relatively little is known about their ecology, distribution, and host–parasite relationships (Petney et al. 2019, Sharifah et al. 2020).

While similarly globally distributed and responsible for the transmission of multiple pathogens causing zoonotic disease in humans, ticks in the genus *Dermacentor* Koch 1844 have comparatively received less research attention than ticks in the genus *Ixodes* Latreille 1795 (Acari: Ixodidae), which over the past several decades have

become infamous as the primary vectors of Lyme disease in eastern North America. While studies reporting humans being bitten by *Dermacentor* ticks are relatively few compared to *Ixodes*, multiple *Dermacentor* species have been reported to parasitize humans (Guglielmone et al. 2020, Mathison and Sapp 2021). Diseases confirmed to be vectored by *Dermacentor* ticks include Rocky Mountain spotted fever, other rickettsioses (spotted fevers), tick-borne encephalitis, and tularemia (https://www.cdc.gov/ticks/). New *Dermacentor* tick species continue to be recognized (Apanaskevich et al. 2021, Lado et al. 2021).

The increasing incidence of tick-borne zoonoses worldwide has prompted questions about what causes some tick species to be more effective vectors of zoonotic pathogens compared to others. Identifying particular traits that confer this propensity may highlight mechanistic hypotheses about what drives vector competence across species and across different life stages within a species. To explore this, Yang and Han (2018) compiled data on the intrinsic traits of tick species in the genus *Ixodes* to predict their zoonotic vector status. The present work builds on that study by creating a global database of features for *Dermacentor* at the species level. We systematically compiled trait data, including morphological characteristics, host range, geographic distribution, and zoonotic diseases that they vector. We then trained a machine learning algorithm (generalized boosted regression) via the R package *gbm* (Greenwell et al. 2020, R Core Team 2022) on these species-level traits to predict zoonotic vector status. This modeling approach identified specific traits that had the highest influence on vector status, and also predicted additional *Dermacentor* species that may be undiscovered zoonotic vectors. Connecting the traits of individual tick species to their zoonotic vector status has implications for more targeted surveillance measures based on disease risk from vector species, contributing to disease prevention, mitigation, and management.

Materials and Methods

Data Collection

We established a set of species-level characteristics describing ticks, starting with those listed in Yang and Han (2018) and expanding the set of predictors to include more finely resolved geographic data variables (Supp Table S1 [online only]). For each *Dermacentor* tick species, we collected trait data from the published literature on life history (clutch size), morphological (biometric) characteristics, the range of hosts infested, the geographic distribution, and the frequency of human infestation (Supp Table S2 [online only]). We also determined which *Dermacentor* species are known or suspected vectors of zoonotic disease using published literature (Supp Table S3 [online only]), including references compiled within the Global Infectious Diseases and Epidemiology Network (GIDEON at https://www.gideononline.com, Berger 2005) and a recent review by Mathison and Sapp (2021).

Dermacentor Species

We used a standard reference text (Guglielmone et al. 2014) as our primary data source for currently recognized *Dermacentor* species, augmented by data from the literature on newly recognized *Dermacentor* tick species, those reinstated as valid species, and reconciled synonymies. Our final data set contained 44 species (Supp Table S3 [online only]).

Biometric (Morphological) Data

For each species, we searched the published literature for reported measurements of morphological features for larvae, nymphs, adult males, and adult females (Supp Table S4 [online only]). We targeted primary sources and, when multiple publications reported original data, we gave precedence to those whose results were based on the largest number of specimens and sources that included high-quality scaled illustrations of a representative individual of a species. When morphological values were not reported in the text, we used scaled illustrations to measure those characteristics via the application imageJ (Rasband 2011). If a scale was not provided with the illustration, we estimated it based on dimensions reported in the text. We also used imageJ if the reported values were based on endpoints that differed from our definitions (e.g., body length measured from the tip of the hypostome instead of from the scapular apices of the capitulum). For quality control, we compared morphologic data acquired from different sources in order to assess the reliability of the data we used in our study and measured parts of the body for which values were reported in the text in order to check whether our values were consistent. Supplementary Table S1 (online only) provides additional information about protocols for individual measurements (e.g., the endpoints used for measuring body length).

While these data represent the most complete compilation of data available for *Dermacentor*, there are some notable caveats. In contrast to data in published papers reporting the mean values of traits based on many specimens, some measurements from scaled illustrations and photos may not be representative of average trait values. Our measurements do not capture variation across the geographic range of a given species. Finally, we were unable to translate or obtain access to some primary sources of data for several species from Russia and China, namely, *D. marginatus*, *D. montanus*, *D. nuttalli*, *D. pavlovskyi*, *D. pomerantzevi*, *D. raskemensis*, *D. silvarum*, and *D. sinicus*.

Host Data

We collected data on the diversity of host species that are known to be parasitized by each *Dermacentor* tick species, including whether or not each tick species infests humans. An initial list of host orders and families was compiled starting with the references cited by Guglielmone et al. (2014), augmented by more recent publications (Supp Table S5 [online only]). Additional references included the original species description papers (especially for more recently recognized tick species) and data from the U.S. National Tick Collection. For host taxonomy, we used valid names from the Integrated Taxonomic Information System (ITIS 2021). We used Guglielmone et al. (2020) as our primary source to document whether or not a species is known to infest humans. We summarized these data for our analysis by tallying the number of mammalian host families and orders for each tick species, but there were not enough published data to assign host associations with particular life stages for each tick species. Explicit delineation of hosts for each life stage is the preferred approach because immature and adult ticks of the same tick species can feed on different host species. However, for our analysis, host families were included *in toto* for each species rather than separately for different life phases because host data for immature phases for the majority of tick species are incomplete to an unknown extent.

We calculated phylogenetic diversity of host species for each tick using the *pd* function from the *picante* package in R version 4.1.3 (Kembel et al. 2010, R Core Team 2022). For this analysis, we used the phylogeny of Mammalia from Upham et al. (2019). As the data reflecting the most comprehensive set of mammal species used interpolations of the placement of missing species, we used all 10,000 fossilized birth-death trees available and ran our estimations

of phylogenetic diversity with each tree. We then used the mean phylogenetic diversity across all 10,000 runs as a variable in our *gbm* modeling process.

Geographic Data

We collected geographic distribution data for each *Dermacentor* species by compiling a global database of georeferenced observations (Supp Table S6 [online]). Primary data sources for our geographic database included peer-reviewed journal publications, published georeferenced data sets (Rubel et al. 2014, Estrada-Peña and de la Fuente 2016, Zhang et al. 2019, Rubel and Brugger 2022), the Global Biodiversity Information Facility (GBIF 2022a,b), iNaturalist research-grade observations (iNaturalist 2022, see also GBIF 2022a), and VectorMap (2019; see Supp Table S6 [online only] for a list of the institutions holding the data records for each *Dermacentor* species). Our protocol for selecting, documenting, and excluding geographic data for this data set is described in Supp Table S6 (online only). We excluded records with unknown or insufficient geographic accuracy. To visualize the distributions of *Dermacentor* species, we created a global map of the coordinate data collected. Additionally, we rasterized the coordinate data at a resolution of one degree in order to examine the patterns of species richness for this genus. All maps were created using the *sf* and *raster* packages in R version 4.1.2 (Pebesma 2018, Hijmans 2020, R Core Team 2022).

Because ticks spend most or all of their time between blood meals as free-living organisms unattached to hosts (Sonenshine and Roe 2013), they are sensitive to abiotic factors such as temperature and relative humidity (Dergousoff et al. 2013). The Köppen–Geiger climate zone classification system ($n = 30$ climate zones) aggregates a set of climatic data variables to summarize the degree of variability in temperature and humidity (Beck et al. 2018). These climate zones are tied to the distribution of vegetation, which has been further characterized by biomes and ecoregions (Olson et al. 2001, Dinerstein et al. 2017). Biomes ($n = 14$) represent broad categories of plant and animal communities and the environments in which they occur, such as 'temperate coniferous forests' or 'deserts and xeric shrublands'. Nested within the terrestrial biomes are ecoregions ($n = 867$), which are delineated at a finer scale based on the distinct plant and animal assemblages they contain (Olson et al. 2001, Dinerstein et al. 2017). Although climate zones and ecoregions introduce similar types of geographic information, we included both variables in our model because the ability of a tick species to survive in a given area is controlled not only by abiotic factors, such as those represented by climate zones, but also by biotic factors, such as those represented by ecoregions.

In order to quantify the ability of a tick species to survive in variable environmental conditions, we summarized the geographic distribution of each *Dermacentor* tick species by counting the number of Köppen–Geiger climate zones, biomes, and ecoregions in which its presence has been recorded. For tallying ecoregions, we used a shapefile downloaded from The Nature Conservancy website (Dinerstein et al. 2017; https://www.gislounge.com/terrestrial-ecoregions-gis-data/). Biome and ecoregion counts were tallied in R version 4.1.3 using the *rgdal* and *sp* packages (Pebesma and Bivand 2005; Bivand et al. 2013, 2021; R Core Team 2022). Quality checks for biome and ecoregion counts were performed using ArcGIS layers from Dinerstein et al. (2017), which used updated geographic delineations of the biomes and ecoregions. Köppen–Geiger climate classifications were tallied using ArcGIS layers from Beck et al. 2018 (1-km resolution).

Study Effort

In general, vector species tend to be better studied than nonvector species. Therefore, it is possible that, rather than describing the features of zoonotic vectors among *Dermacentor*, our model would be unintentionally trained to distinguish well-studied from understudied *Dermacentor* species. To examine this possibility, we quantified study effort by performing searches on the Latin binomial for each tick species in the Web of Science database and tallied the number of citations for each species. We included these data as a feature in our model, and also compared our main model results to those from a separate model that we trained to distinguish well-studied *Dermacentor* species using traits.

Zoonotic Vector Status

Each *Dermacentor* species was assigned a binary score based on their zoonotic vector status. To accomplish this, we used the GIDEON database (https://www.gideononline.com/, Berger 2005) and a comprehensive article by Mathison and Sapp (2021) that consolidated information about which *Dermacentor* species bite humans, as well as the current consensus in the scientific community about their vectorial capacity. A *Dermacentor* tick species was classified as a zoonotic vector if it was explicitly stated as such in the GIDEON database and/or in the Mathison and Sapp's (2021) review.

We did not consider the detection of a pathogen in a tick species to constitute conclusive evidence of its ability to vector the pathogen to a human host although it serves as an indication that it has the potential to be a zoonotic vector species. We also did not categorize species as vectors if the transmission route is indirect, i.e., transmitted to humans through contact with blood from infected livestock.

Data Analysis

To identify which tick traits are the most important predictors of zoonotic disease status and, based on those traits, which *Dermacentor* species are most likely to vector disease, we used a machine learning technique, generalized boosted regression (package *gbm* in R version 4.1.3; Elith et al. 2008, Greenwell et al. 2020, Ridgeway 2020, R Core Team 2022). We chose this method following Yang and Han (2018), which utilized this technique for a similar model of *Ixodes* traits and which allowed us to directly compare model results for these two tick genera. Generalized boosted regression iteratively fits thousands of classification or regression trees by incorporating boosting, which combines multiple weakly predictive trees into an ensemble with superior predictive accuracy. This approach accommodates multiple data types, hidden and nonlinear interactions among multiple variables, and can also handle missing data by treating 'missingness' as a value for splitting decisions (Elith et al. 2008). The fitted model can be used to identify which variables are most important for accurate classification and to make predictions about novel vector species as new host and geographic data records become available. Our model classified tick species as vectors or nonvectors, a binary outcome.

To decrease the chance of overfitting on a small data set, we reduced the number of predictor variables and also employed 4-fold cross-validation within the gbm and generated a null model against which to compare gbm performance. To prune the number of predictor variables we excluded variables with near zero variance and/or less than 33% coverage across species. We then performed a grid search to tune the input parameters. The parameter values giving the highest training accuracy and stable deviance curve were learning rate = 0.0001, maximum tree depth = 3, minimum number of
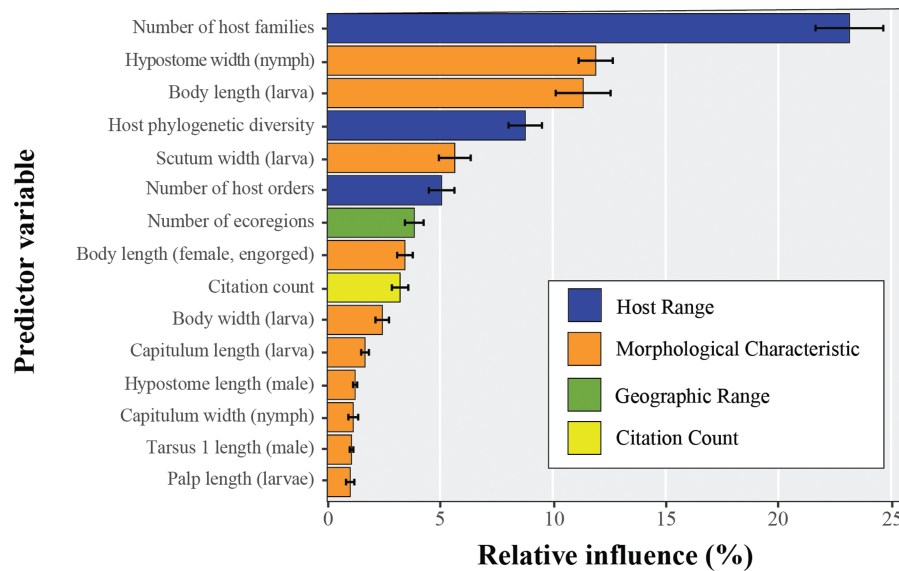
Fig. 1. Mean relative influence of the top predictor variables. Relative influence indicates the importance of each variable in reducing prediction error. Error lines represent ± 1.5 × IQR, where IQR is the interquartile range between the first and third quartiles, generated from 50 bootstrap runs of the generalized boosted regression model. Morphological measurements are given in millimeters. Predictor variables are defined in Supp. Table S1 (online only).

observations in node = 1, with a maximum number of trees = 80,000 and 70/30 split between training and test data. Boosted regression tree models often split data into a larger training dataset and smaller test set (Leathwick et al. 2006), especially when the dataset is small as in our case. Using these parameter values, we bootstrapped our gbm model (*n* = 50 runs), in each run using a randomly selected 70% of the data for training. To evaluate model performance, we computed the area under the receiver operating characteristic curve (AUC) for predictions made on the test data set (remaining 30% of data) for each bootstrap model run. To generate null models using this bootstrapping procedure (50 runs), binary labels (vector status) were randomly reassigned in each run (target shuffling, as in Fischhoff et al. 2021). We then corrected our mean test AUC by subtracting the difference between the mean test AUC of null models and the expected value for a model performing no better than chance (AUC = 0.5). Bootstrapped models give mean and standard error of the relative influence of each predictor variable and were also used to generate partial dependence plots that depict the importance of each feature for classification accuracy.

To assess the possible effects of sampling bias on classifying vector status, we also ran a *gbm* model with the number of citations for each species (citation count) as the response variable. This model enabled us to determine whether study effort is predictive of vector status among *Dermacentor* species, and if so, whether our models of vector status and study effort reflect similar trait profiles.

We further investigated the relationships between mammalian host range (number of host families, number of host orders, and host phylogenetic diversity) and geographic range (number of biomes, ecoregions, and climate zones) using scatterplots and Spearman's rank correlation tests for each pair of variables in R version 4.1.3 (R Core Team 2022).

Results

In sum, we collected data for 44 species of *Dermacentor* ticks for a total of 87 variables (Supp Table S2 [online only]). The following species were excluded from our *gbm* analysis due to lack of sufficient

accessible data: *D. asper*, *D. dispar*, and *D. pomerantzevi*. A full list of predictor variables, their definitions, and coverage is included in Supp Table S1 (online only). Of the 41 species included in our analysis, 10 species (24%) were reported to be zoonotic vectors by GIDEON and/or Mathison and Sapp (2021; Supp Table S3 [online only]).

Our compilation of *Dermacentor* species traits encompassed 76 morphological characteristics for all life stages (larvae, nymphs, and both male and female adults), such as body size, length and width of the capitulum (mouthparts), hypostome (used to anchor the tick during feeding), and tarsi (leg segment that bears a sensory organ used for locating hosts; Sonenshine and Roe 2013; Supp Table S1 [online only]). The number of mammalian host families and orders for individual tick species ranged from 1 to 21 (families) and 1 to 8 (orders); host families and orders for each species are provided in Supp Table S5 (online only). We compiled a total of 6,786 georeferenced occurrence points in our geographic database, of which 3,969 were used for tallying the number of biomes, ecoregions, and climate zones for each species (Supp Table S6 [online only]). We excluded geographic occurrence points from counts when they could not reliably be assigned to an ecoregion or climate zone.

Our corrected mean test AUC was 0.84 (corrected test AUC = mean test AUC [0.95] – (mean test AUC for the null model [0.61] – 0.5)), indicating that our models trained on trait data correctly classified vector status among *Dermacentor* ticks with ~84% accuracy and a minimal degree of overfitting (correction factor = 0.11) lending confidence that variable pruning, cross-validation, and correction via target shuffling combined to effectively address the learning constraints of a small data set.

We found that host range (number of mammalian host families) had the highest relative influence on predicting vector status, followed by hypostome width of nymphal ticks and body length of larval ticks (Fig. 1). Additional variables with low relative importance (but >1%) included host phylogenetic diversity, scutum width of larval ticks, geographic range (number of ecoregions), citation count, body length of (engorged) female ticks, and several other morphological characteristics.

Some of the top predictors indicated by our model, such as the number of host families, subsume sampling and publication bias inherent in most ecological data sets. For instance, more hosts are documented for tick species that are better studied, and fewer hosts are known for ticks that are understudied. Other predictors, such as morphological characteristics, however, are not always subject to this same bias. The results of our model using traits to predict the number of citations for a given *Dermacentor* species indicated that tick traits poorly predict study effort (the number of citations), with a pseudo-R^2 of 0.05. These results suggest that while study effort does vary widely across this group, our model is learning and generating predictions using the features of zoonotic vectors rather than the features of species that are simply better studied.

In order to better understand the relationships between the top predictor variables and vector status, we generated partial dependence plots, which show the effect of a given trait on prediction accuracy while controlling for the effects of all other predictor variables in the model. We found that, while most tick species have a small number of host families (gray bars, Fig. 2), those that infest more than 10 host families are more likely to be zoonotic disease vectors (black lines, Fig. 2). *Dermacentor* species with larger hypostomes during their nymphal stage and species with smaller-bodied larvae were also more likely to be vectors, as were species found in a high number of ecoregions.

Our model correctly classified nine of the ten confirmed *Dermacentor* vector species, as determined by having over 50% probability of being a vector using predictions made with the full data set. Although *D. similis*, a known vector species described in 2021 by Lado et al., was not classified as a vector species by our model, it was close to our 50% probability cutoff (49% likely to be a vector). Other species near the cutoff include *D. niveus*, which our model assigned a 31% probability of being a vector based on its traits profile; and *D. parumapertus* (22% probability). Both species are not currently known to be vectors.

We used our geographic database (Supp Table S6 [online only]) to generate a global map of *Dermacentor* species distributions (Fig. 3). We also quantified *Dermacentor* species richness in three areas of interest—North America, Europe, and Asia—and found that areas with the greatest species richness occur in Asia, despite the relative scarcity of trait data for species in this region (Fig. 4).

We consistently found significantly positive relationships between mammalian host range (number of host families, number of host orders, and host phylogenetic diversity) and geographic range (number of biomes, ecoregions, and climate zones); all *p*-values were of <0.01. Results of Spearman's rank correlation tests are provided in Supp Table S7 (online only), and these relationships are visualized in Supp Fig. S1 (online only).

Discussion

We systematically reviewed the literature to compile a comprehensive, first of its kind data set for the genus *Dermacentor*, an understudied group of zoonotic pathogen-vectoring ticks. In particular, we collated and standardized detailed information on species' morphological characteristics, the diversity of host families they infest, and their geographic distribution, which represents a valuable ecoinformatic contribution to the knowledge base about this group. Our analyses suggest that *Dermacentor* species that are the most likely to be zoonotic vectors are broad ranging, both in terms of the range of hosts they infest and the range of ecoregions across which they are found, and also tend to have large hypostomes to anchor themselves to their hosts and be small-bodied as immatures.

The most important predictor of zoonotic vector status was host range, specifically, the number of mammalian host families a tick species is known to infest. Host phylogenetic diversity was also among the top predictor variables. Five mammalian families host at least half of all *Dermacentor* species (Bovidae, Suidae, Canidae, Muridae, and Leporidae). Bovidae, Suidae, and Canidae include domestic species (e.g., cattle, pigs, and dogs), and Muridae (rodents) are often human-associated species. A promising area for future research would be to compare the host families of vector and nonvector tick species by life stage. For pathogens in general, the ability to infect a broad range of hosts is associated with a high risk of emergence (Cleaveland et al. 2001, Woolhouse and Gowtage-Sequeria 2005). We postulate that ticks with a broad host range are also more likely to encounter a high diversity of pathogens, including zoonotic pathogens (Estrada-Peña et al. 2015), or they are more likely to encounter a single broadly distributed hyper-reservoir species (e.g., Han et al. 2015) that carries numerous zoonotic pathogens to which the tick is susceptible or for which it is a competent vector.

Mirroring broad host range, *Dermacentor* ticks found in a wide variety of ecoregions were more likely to be zoonotic vectors. We also found a positive correlation between host range and geographic range for *Dermacentor* ticks (Supp Fig. S1 [online only], Supp Table S7 [online only]). While species with broader host and geographic ranges tend to be those that are better studied, our analyses showed that this study bias did not extend to the intrinsic morphological traits distinguishing *Dermacentor* vectors from nonvectors in our study. An outstanding area for future work is to determine overlapping ranges of these zoonotic vectors with potential reservoir host species. To this end, our high-resolution geographic location database could be used to support species distribution modeling for *Dermacentor* ticks, which has thus far been limited to a few well-known zoonotic vector species on regional (Široký et al. 2011, Walter et al. 2016, Huercha et al. 2020) or continental (Minigan et al. 2018, Boorgula et al. 2020) scales. On a global scale, VectorMap (http://vectormap.si.edu/), one of the sources we used for our geographic data set, has mapped the distributions of and utilized predictive climate suitability models for several *Dermacentor* tick species. The global scope of our geographic data set enables augmentation with climatic data from sources such as TerraClimate (Abatzoglou et al. 2018) to facilitate more comprehensive species distribution models, generating suitability maps of *Dermacentor* species under the present and future climate scenarios. Taking seasonality into account, climate optima for different *Dermacentor* species could be determined, as has been done with thermal optima for mosquito vectors, to postulate the resulting implications for human disease incidence with a changing climate (Mordecai et al. 2019).

For *Dermacentor* tick nymphs, species with wider hypostomes were more likely to be zoonotic vectors. The hypostome is inserted into the host and serves to anchor the tick in place during blood feeding (Richter et al. 2013, Sonenshine and Roe 2013). Larger hypostomes could form a more secure anchor (Richter et al. 2013), enabling a longer period of attachment and increasing the probability of pathogen transmission (Kaufman 2010). We also found that smaller-bodied immature *Dermacentor* ticks were more likely to be vectors. It is possible that the results for body size of immature ticks was driven by the frequency distribution of this trait. In other words, the tendency to be classified as a vector species (Fig. 2, black line) tracks with the frequency distribution of this trait across *Dermacentor* species (Fig. 2, gray bars). Less information is available about immature *Dermacentor* ticks and although there were few records of human infestation by immature *Dermacentor* ticks (Supp Table S5 [online only]), we postulate that immature ticks could be
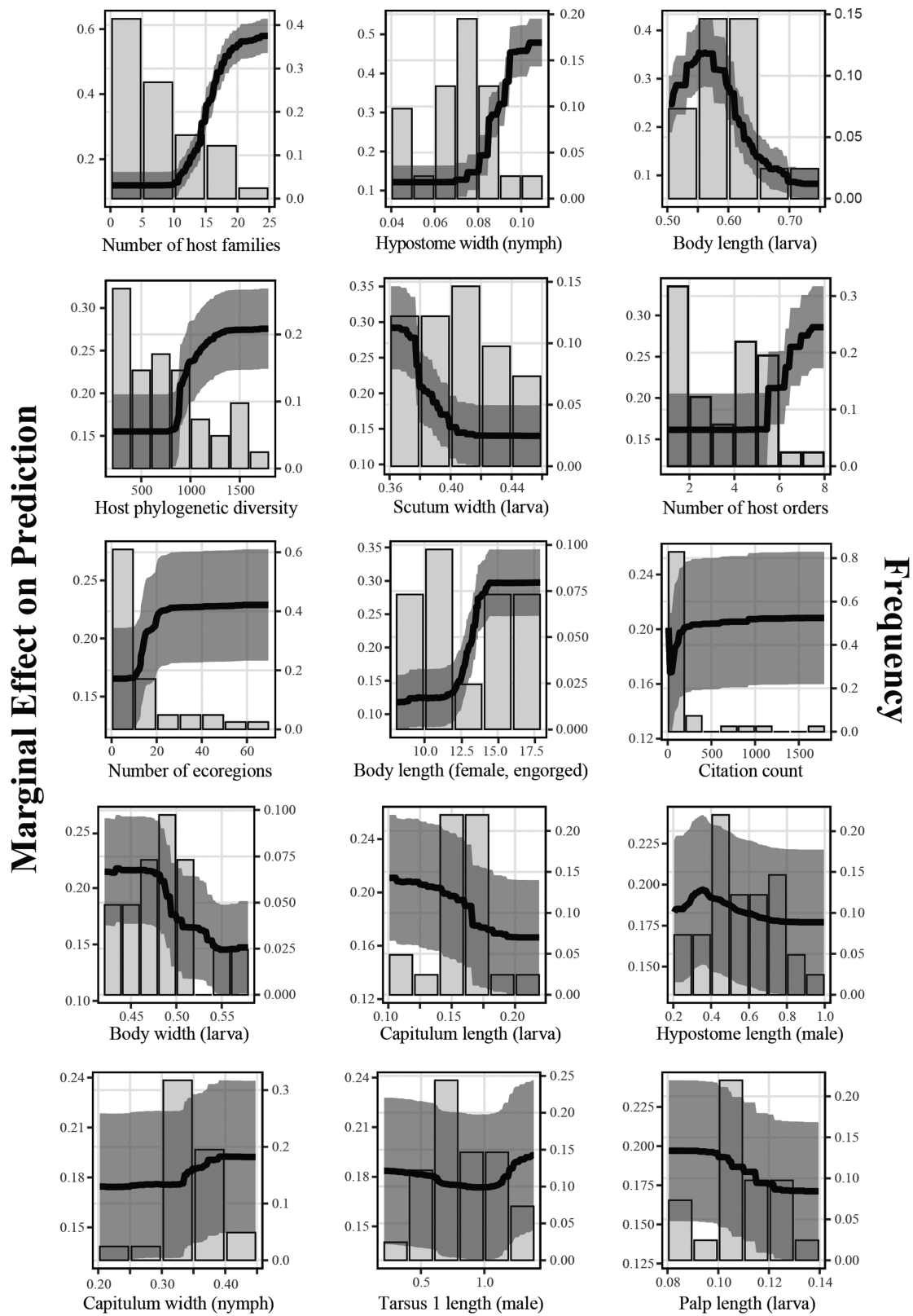
Fig. 2. Partial dependence plots for a selection of top predictor variables from the generalized boosted regression model used to predict the vector status of *Dermacentor* ticks. The black line represents the average marginal effect (*y*-axis, left) of a given trait (*x*-axis) on vector status after accounting for the average effect of all other predictor variables in the model. Gray bands represent 95% CI. The histograms show the relative frequency (*y*-axis, right) of tick species with a given value of each trait. Morphological measurements are given in millimeters.
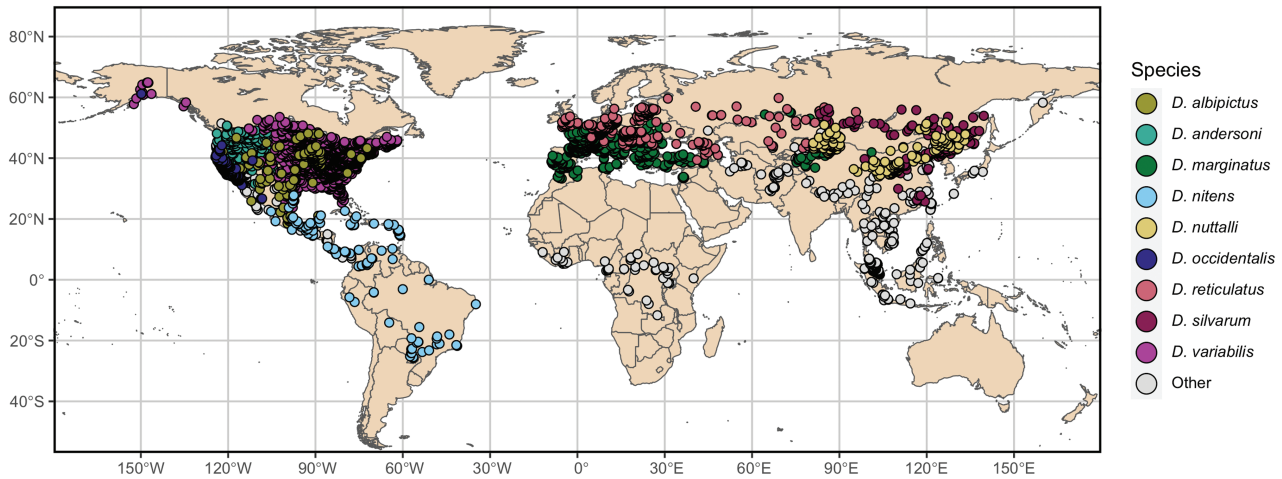
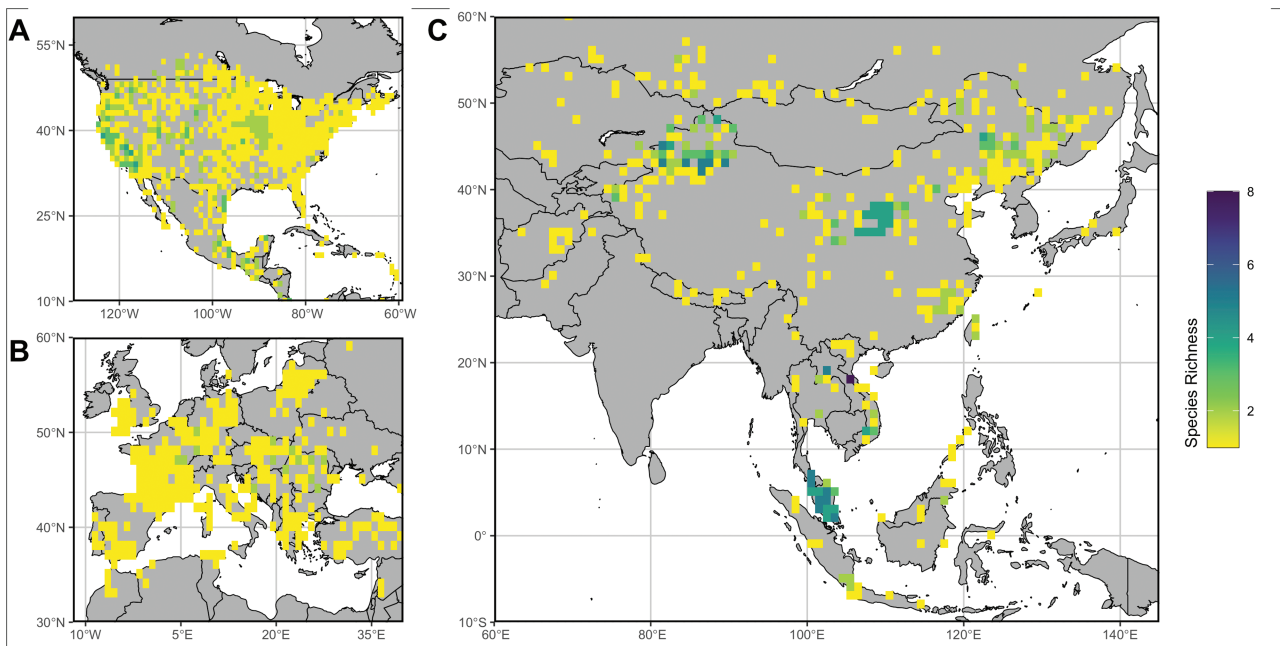**Fig. 3.** Global distribution of *Dermacentor* tick species.

Fig. 4. Species richness of *Dermacentor* ticks in North America (A), Europe (B), and Asia (C). Species richness ranges from a single species (lighter colors) to eight species (darker colors).

acquiring pathogens from wildlife (e.g., rodents) and transmitting them to humans as adults. Finally, we found that species with longer body sizes as engorged females were more likely to be zoonotic vectors, similar to Yang and Han (2018). This may serve as a rough proxy for fecundity and population abundance, if body length varies reliably with the maximum size of egg clutches.

Among known vector species, our model correctly classified all but one species, *D. similis*, and identified two additional potential vector species. Based on molecular and morphological traits, Lado et al. (2021) determined that the population of *D. variabilis* in the western United States constituted a new species, *D. similis*. We suspect that because it was recently documented, the diversity of hosts as well as the geographic distribution of this species may be underdescribed; for our data set, we excluded records where it could not be determined if a specimen was *D. variabilis* or *D. similis*. Our model identified *D. niveus* and *D. parumapertus*, based on trait

similarity to confirmed vector species, to be the highest probability species not currently classified as vectors. *Dermacentor niveus* has a broad host range (12 host families) and is known to infest humans on rare occasions (Guglielmone et al. 2020). The known distribution of this species is in mountainous regions of central Asia across six ecoregions. Crimean-Congo hemorrhagic fever virus was detected in *D. niveus* in Kazakhstan (Onishchenko et al. 2005) but definitive evidence that it vectors this virus to humans is currently lacking. *Dermacentor parumapertus* is found in North America across 16 ecoregions and is known to parasitize 10 host families. *Rickettsia parkeri* was detected in *D. parumapertus* ticks from black-tailed jackrabbits in Mexico (Sánchez-Montes et al. 2018) and the western United States (Paddock et al. 2017). Although *D. niveus* and *D. parumapertus* are known to parasitize humans, they do so rarely (Guglielmone et al. 2020). This suggests that while they share trait similarities with other known zoonotic vectors, host preference in

these species may confer some protection against transmission of zoonotic pathogens to humans, or that they are found in sparsely populated areas.

Mapping the distributions of known and predicted pathogen-vectoring ticks is a necessary step toward assessing high-risk areas for tick-borne disease (Eisen and Paddock 2020). Our finely resolved geographic location database that encompasses the entire genus of *Dermacentor* ticks has the potential to fill in key gaps in our knowledge of tick distributions for species whose geographic ranges have not been fully delineated. Taking advantage of this potential could generate critical predictions of high-risk areas in understudied regions such as southeast Asia and South America as anthropogenic climate change continues to shift the distributions of hosts and vectors. In particular, our global map of *Dermacentor* species richness underscores the need for additional research in Asia, a research gap that was also recognized in a recent review of mapping tick-borne disease (Lippi et al. 2021).

During a time when anthropogenic impacts such as climate and land use change appear to be favoring the continued spread of tick-borne diseases, knowledge about these vector species is even more crucial (Swei et al. 2020). Our database of *Dermacentor* species morphological characteristics, host range, and geographic distribution has relevance beyond the scope of our study. The information we compiled can be used for public health applications, such as species distribution modeling, to determine areas deserving increased surveillance measures.

Acknowledgments

Supplementary Data

Supplementary data are available at *Journal of Medical Entomology* online.

Supp Figure S1. Relationships between *Dermacentor* host range and geographic range.

Supp Table S1. List of predictor variables, including definitions, measurement protocols, and coverage.

Supp Table S2. *Dermacentor* master database.

Supp Table S3. Currently recognized *Dermacentor* species and their zoonotic vector status.

Supp Table S4. Biometric data.

Supp Table S5. *Dermacentor* host families and orders.

Supp Table S6. Geographic distribution data.

Supp Table S7. Correlations between *Dermacentor* host range and geographic range variables.

References Cited

Abatzoglou, J. T., S. Z. Dobrowski, S. A. Parks, and K. C. Hegewisch. 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* 5: 1–12.

Apanaskevich, D. A., M. A. Apanaskevich, W. Nooma, A. Ahantarig, and W. Trinachartvanit. 2021. Reinstatement of *Dermacentor tricuspis* (Schulze, 1933) n. comb., n. stat. (Acari: Ixodidae) as a valid species, synonymization of *D. atrosignatus* Neumann, 1906 and description of a new species from Indonesia, Malaysia and Thailand. *Syst. Parasitol*. 98: 207–230.

Beck, H. E., N. E. Zimmermann, T. R. McVicar, N. Vergopolan, A. Berg, and E. F. Wood. 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* 5: 180214.

Berger, S. A. 2005. GIDEON: a comprehensive Web-based resource for geographic medicine. *Int. J. Health Geogr*. 4: 1–11.

Bivand, R. S., E. Pebesma, and V. Gomez-Rubio. 2013. *Applied spatial data analysis with R*, 2nd ed. Springer, New York.

Bivand, R., T. Keitt, and B. Rowlingson. 2021. rgdal: bindings for the 'geospatial' data abstraction library. R package version 1.5-27. (https://cran.r-project.org/web/packages/rgdal/index.html).

Boorgula, G. D. Y., A. T. Peterson, D. H. Foley, R. R. Ganta, and R. K. Raghavan. 2020. Assessing the current and future potential geographic distribution of the American dog tick, *Dermacentor variabilis* (Say) (Acari: Ixodidae) in North America. *PLoS One* 15: e0237191.

Cleaveland, S., M. K. Laurenson, and L. H. Taylor. 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci*. 356: 991–999.

Couper, L. I., A. J. MacDonald, and E. A. Mordecai. 2021. Impact of prior and projected climate change on US Lyme disease incidence. *Glob. Change Biol*. 27: 738–754.

Dergousoff, S. J., T. D. Galloway, L. R. Lindsay, P. S. Curry, and N. B. Chilton. 2013. Range expansion of *Dermacentor variabilis* and *Dermacentor andersoni* (Acari: Ixodidae) near their northern distributional limits. *J. Med. Entomol*. 50: 510–520.

Dinerstein, E., D. Olson, A. Joshi, C. Vynne, N. D. Burgess, E. Wikramanayake, N. Hahn, S. Palminteri, P. Hedao, and R. Noss, *et al*. 2017. An ecoregion-based approach to protecting half the terrestrial realm. *BioScience* 67: 534–545.

Eisen, R. J., and C. D. Paddock. 2020. Tick and tickborne pathogen surveillance as a public health tool in the United States. *J. Med. Entomol*. 58: 1490–1502.

Eisen, R. J., K. J. Kugeler, L. Eisen, C. B. Beard, and C. D. Paddock. 2017. Tick-borne zoonoses in the United States: persistent and emerging threats to human health. *ILAR J*. 58: 319–335.

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. *J. Anim. Ecol*. 77: 802–813.

Estrada-Peña, A., and J. de la Fuente. 2016. Species interactions in occurrence data for a community of tick-transmitted pathogens. *Sci. Data* 3: 160056.

Estrada-Peña, A., J. de la Fuente, R. S. Ostfeld, and A. Cabezas-Cruz. 2015. Interactions between tick and transmitted pathogens evolved to minimise competition through nested and coherent networks. *Sci. Rep*. 5: 10361.

Fischhoff, I. R., A. A. Castellanos, J. P. G. L. M. Rodrigues, A. Varsani, and B. A. Han. 2021. Predicting the zoonotic capacity of mammals to transmit SARS-CoV-2. *Proc. R. Soc. B Biol. Sci*. 288: 20211651.

(GBIF) Global Biodiversity Information Facility. 2022a. (https://www.GBIF.org) (8 April 2022). GBIF Occurrence Download (https://doi.org/10.15468/dl.dqb4y6).

(GBIF) Global Biodiversity Information Facility. 2022b. (https://www.GBIF.org) (9 April 2022). GBIF Occurrence Download (https://doi.org/10.15468/dl.xmacpp).

Greenwell, B., B. Boehmke, J. Cunningham, and GBM Developers. 2020. gbm: Generalized Boosted Regression Models [R package 'gbm' version 2.1.8]. (https://cran.r-project.org/web/packages/gbm/index.html).

Guglielmone, A. A., R. G. Robbins, D. A. Apanaskevich, T. N. Petney, A. Estrada-Peña, and I. G. Horak. 2014. *The hard ticks of the world (Acari: Ixodida: Ixodidae)*. Springer, Heidelberg, Germany.

Guglielmone, A. A., T. N. Petney, and R. G. Robbins. 2020. Ixodidae (Acari: Ixodoidea): Descriptions and redescriptions of all known species from 1758 to December 31, 2019. *Zootaxa* 4871: 1–322.

Han, B. A., J. P. Schmidt, S. E. Bowden, and J. M. Drake. 2015. Rodent reservoirs of future zoonotic diseases. *Proc. Natl. Acad. Sci. USA* 112: 201501598.

Hijmans, R. J. 2020. raster: Geographic data analysis and modeling [R package 'raster' version 3.3-13]. (https://cran.r-project.org/web/packages/raster/index.html).

Huercha, R. Song, Y. Ma, Z. Hu, Y. Li, M. Li, L. Wu, C. Li, E. Dao, X. Fan, Y. Hao, and C. Bayin. **2020**. MaxEnt modeling of *Dermacentor marginatus* (Acari: Ixodidae) distribution in Xinjiang, China. *J. Med. Entomol.* 57: 1659–1667.

iNaturalist. **2022**. The iNaturalist data were downloaded from GBIF. (https://www.inaturalist.org).

(ITIS) Integrated Taxonomic Information System. **2021**. *Species 2000 & ITIS Catalogue of Life, 2020-09-01 Beta* (Y. Roskov, G. Ower, T. Orrell, D. Nicolson, N. Bailly, P. M. Kirk, T. Bourgoin, R. E. DeWalt, W. Decock, E. van Nieukerken, and L. Penev, eds.). Species 2000: Naturalis, Leiden, The Netherlands. (https://www.catalogueoflife.org/).

Kaufman, W. R. **2010**. Ticks: physiological aspects with implications for pathogen transmission. *Ticks Tick Borne Dis.* 1: 11–22.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. **2010**. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.

Lado, P., M. G. Glon, and H. Klompen. **2021**. Integrative taxonomy of *Dermacentor variabilis* (Ixodida: Ixodidae) with description of a new species, *Dermacentor similis* n. sp. *J. Med. Entomol.* 58: 2216–2227.

Leathwick, J. R., J. Elith, M. P. Francis, T. Hastie, and P. Taylor. **2006**. Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees. *Marine Ecol. Prog. Ser.* 321: 267–281.

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. *J. Med. Entomol.* 58: 2021–2029.

Mathison, B. A., and S. G. H. Sapp. **2021**. An annotated checklist of the eukaryotic parasites of humans, exclusive of fungi and algae. *Zookeys* 1069: 1–313.

Minigan, J. N., H. A. Hager, A. S. Peregrine, and J. A. Newman. **2018**. Current and potential future distribution of the American dog tick (*Dermacentor variabilis*, Say) in North America. *Ticks Tick Borne Dis.* 9: 354–362.

Mordecai, E. A., J. M. Caldwell, M. K. Grossman, C. A. Lippi, L. R. Johnson, M. Neira, J. R. Rohr, S. J. Ryan, V. Savage, and M. S. Shocket, *et al.* **2019**. Thermal biology of mosquito-borne disease. *Ecol. Lett.* 22: 1690–1708.

Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D. Amico, I. Itoua, H. E. Strand, and J. C. Morrison, *et al.* **2001**. Terrestrial ecoregions of the world: a new map of life on Earth. *BioScience* 51: 933–938.

Onishchenko, G. G., I. Tumanoya, O. I. Vyshemirskiǐ, J. Kuhn, S. V. Seregin, G. I. Tiunnikov, I. D. Petrova, F. K. Tishkova, K. S. Ospanov, and S. V. Kazakov, *et al.* **2005**. ELISA and RT-PCR-based research of viruses in the ticks collected in the foci of Crimean-Congo fever in Kazakhstan and Tajikistan in 2001–2002. *Vopr. Virusol.* 50: 23–26.

Paddock, C. D., M. E. J. Allerdice, S. E. Karpathy, W. L. Nicholson, M. L. Levin, T. C. Smith, T. Becker, R. J. Delph, R. N. Knight, and J. M. Ritter, *et al.* **2017**. Unique strain of *Rickettsia parkeri* associated with the hard tick *Dermacentor parumapertus* Neumann in the western United States. *Appl. Environ. Microbiol.* 83: 1–15.

Pebesma, E. **2018**. Simple features for R: standardized support for spatial vector data. *R J.* 10: 439.

Pebesma, E. J., and R. S. Bivand. **2005**. Classes and methods for spatial data in R. *R News* 5/2. (https://cran.r-project.org/doc/Rnews/).

Petney, T. N., W. Saijuntha, N. Boulanger, L. Chitimia-Dobler, M. Pfeffer, C. Eamudomkarn, R. H. Andrews, M. Ahamad, N. Putthasorn, and S. V. Muders, *et al.* **2019**. Ticks (Argasidae, Ixodidae) and tick-borne diseases of continental Southeast Asia. *Zootaxa* 4558: 1–89.

R Core Team. **2022**. *R: a language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rasband, W. S. **2011**. National Institutes of Health, Bethesda, MD. (http://imagej.nih.gov/ij/).

Richter, D., F.-R. Matuschka, A. Spielman, and L. Mahadevan. **2013**. How ticks get under your skin: insertion mechanics of the feeding apparatus of *Ixodes ricinus* ticks. *Proc. R. Soc. B* 280: 20131758.

Ridgeway, G. **2020**. Generalized boosted regression models: a guide to the gbm package. (https://cran.r-project.org/web/packages/gbm/index.html).

Rosenberg, R., N. P. Lindsey, M. Fischer, C. J. Gregory, A. F. Hinckley, P. S. Mead, G. Paz-Bailey, S. H. Waterman, N. A. Drexler, and G. J. Kersh, *et al.* **2018**. Vital signs: trends in reported vectorborne disease cases – United States and Territories, 2004–2016. *Morb. Mortal. Wkly. Rep.* 67: 496–501.

Rubel, F., and K. Brugger. **2022**. Maps of ticks (Acari: Argasidae, Ixodidae) for Austria and South Tyrol, Italy. *Exp. Appl. Acarol.* 86: 211–233.

Rubel, F., K. Brugger, M. Monazahian, B. Habedank, H. Dautel, S. Leverenz, and O. Kahl. **2014**. The first German map of georeferenced ixodid tick locations. *Parasit. Vectors* 7: 1–5.

Sánchez-Montes, S., A. M. López-Pérez, C. Guzmán-Cornejo, P. Colunga-Salas, I. Becker, J. D. Mora, J. D. Licona-Enríquez, D. Delgado-de la Mora, S. E. Karpathy, C. D. Paddock, and G. Suzán. **2018**. *Rickettsia parkeri* in *Dermacentor parumapertus* ticks, Mexico. *Emerg. Infect. Dis.* 24: 1108–1111.

Sharifah, N., C. C. Heo, J. Ehlers, J. Houssaini, and D. Tappe. **2020**. Ticks and tick-borne pathogens in animals and humans in the island nations of Southeast Asia: a review. *Acta Trop.* 209: 105527.

Široký, P., M. Kubelová, M. Bednář, D. Modrý, Z. Hubálek, and E. Tkadlec. **2011**. The distribution and spreading pattern of *Dermacentor reticulatus* over its threshold area in the Czech Republic—how much is range of this vector expanding?. *Vet. Parasitol.* 183: 130–135.

Sonenshine, D. **2018**. Range expansion of tick disease vectors in North America: implications for spread of tick-borne disease. *Int. J. Environ. Res. Public Health* 15: 478.

Sonenshine, D. E., and R. M. Roe (eds.). **2013**. *Biology of ticks*, Vol. 1. Oxford University Press, Oxford, United Kingdom.

Swei, A., L. I. Couper, L. L. Coffey, D. Kapan, and S. Bennett. **2020**. Patterns, drivers, and challenges of vector-borne disease emergence. *Vector Borne Zoonotic Dis.* 20: 159–170.

Upham, N. S., J. A. Esselstyn, and W. Jetz. **2019**. Inferring the mammal tree: species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biol.* 17: e3000494.

VectorMap. **2019**. (http://vectormap.si.edu). Accessed 8 August 2019.

Walter, M., K. Brugger, and F. Rubel. **2016**. The ecological niche of *Dermacentor marginatus* in Germany. *Parasitol. Res.* 115: 2165–2174.

Woolhouse, M. E. J., and S. Gowtage-Sequeria. **2005**. Host range and emerging and reemerging pathogens. *Emerg. Infect. Dis.* 11: 1842–1847.

Yang, L. H., and B. A. Han. **2018**. Data-driven predictions and novel hypotheses about zoonotic tick vectors from the genus *Ixodes*. *BMC Ecol.* 18: 1–6.

Zhang, G., D. Zheng, Y. Tian, and S. Li. **2019**. A dataset of distribution and diversity of ticks in China. *Sci. Data* 6: 105.