Am. J. Trop. Med. Hyg., 102(2), 2020, pp. 392–402
doi:10.4269/ajtmh.19-0191

Predicting the Geographic Distribution of the *Bacillus anthracis* A1.a/Western North American Sub-Lineage for the Continental United States: New Outbreaks, New Genotypes, and New Climate Data

Anni Yang,[1,2] Jocelyn C. Mullins,[3] Matthew Van Ert,[1,2] Richard A. Bowen,[4] Ted L. Hadfield,[1,2] and Jason K. Blackburn[1,2]*

[1]*Spatial Epidemiology & Ecology Research Laboratory, Department of Geography, University of Florida, Gainesville, Florida;* [2]*Emerging Pathogens Institute, University of Florida, Gainesville, Florida;* [3]*University of Saint Joseph, West Hartford, Connecticut;* [4]*Animal Reproduction and Biotechnology Laboratory, Department of Biomedical Sciences, Colorado State University, Fort Collins, Colorado*

Abstract. *Bacillus anthracis*, the causative pathogen of anthrax, is a spore-forming, environmentally maintained bacterium that continues to be a veterinary health problem with outbreaks occurring primarily in wildlife and livestock. Globally, the genetic populations of *B. anthracis* include multiple lineages, and each may have different ecological requirements and geographical distributions. It is, therefore, essential to identify environmental associations within lineages to predict geographical distributions and risk areas with improved accuracy. Here, we model the ecological niche and predict the geography of the most widespread sublineage of *B. anthracis* in the continental United States using updated MERRA-derived (Modern Era Retrospective analysis for Research and Applications; the NASA atmospheric data reanalysis of satellite information with multiple data products) bioclimate variables (i.e., MERRAclim data) and updated soil variables. We filter the occurrence data associated with the A1.a/Western North American sub-lineage of *B. anthracis* from historical anthrax outbreaks using the multiple-locus variable-number tandem repeat system. In addition, we also incorporate recent cases associated with *B. anthracis* A1.a sub-lineage from 2008 to 2012 in Montana, Colorado, and Texas. Our results provide the predicted distribution of the A1.a sub-lineage of *B. anthracis* for the United States with better predictive accuracy and higher spatial resolution than previous estimates. Our prediction serves as an improved disease risk map to better inform anthrax surveillance and control in the United States, particularly the Dakotas and Montana where this sub-lineage is persistent.

INTRODUCTION

Anthrax is a zoonotic disease affecting animals and humans nearly worldwide.[1] The causative agent of anthrax, *Bacillus anthracis*, is a spore-forming, environmentally maintained bacterium, which is endemic to specific soil environments and can persist for years to decades under suitable conditions.[2] Several ecological niche modeling studies define the suitable habitat for *B. anthracis* spore persistence as grassland or steppe with a narrow range of moderate normalized difference vegetation index (NDVI; 0.2–0.5), limited annual precipitation, and high soil pH.[3–6] Anthrax is established in the United States and was likely introduced during the European colonization through cattle trading and animal production[7,8] (although it has also been hypothesized at least one lineage may have migrated across the Bering Land Bridge[9]). It was a significant problem to livestock and wildlife until the late 1950s when vaccination was introduced.[10,11] Although vaccination is available and inexpensive, it is often used as reactionary outbreak control rather than proactive disease prevention[4] in livestock; administration in wildlife is logistically untenable.[12,13] Presently, anthrax continues to occur in the historical enzootic zone of West Texas[14] and the re-emergent zone of southwestern Montana.[15,16] Current anthrax control and management strategies in wildlife focus on surveillance and carcass decontamination during the risk season.[16,17] Therefore, identifying potential locations where the pathogen might persist and quantifying the anthrax risk across the landscape would help to define priority areas for effective disease control and management. To precisely quantify the geographical distribution of *B. anthracis*, it is necessary to improve our understanding of the ecology of the pathogen.

Applications of ecological niche models (ENMs; i.e., species distribution models) to the disease systems remain an important tool to estimate disease distributions and risk areas. Multiple studies use ENMs to predict spatial distributions of the pathogens based on outbreak locations,[3,4,6,18] the presence of vectors,[19–21] and the occurrence of hosts or reservoirs.[22–24] Modeling a species' geographical distribution is primarily based on identification of nonrandom associations between the occurrence locations of a species and environmental suitability for its survival.[25,26] The ecological niche of a species was classically defined as an n-dimensional hypervolume of environmental covariates determining the ecological space of the species while maintaining the species population without immigration.[26,27] The exploration of environmental covariates for the ecological niche of a species and their environmental coverages (i.e., the range of environmental covariates suitable for species' survival) also helps capture the underlying biological mechanisms of the species responses to the environment, as well as the physiology and ecology of the species. In the case of *B. anthracis,* understanding the environmental factors associated with the outbreak locations or isolate locations can infer and predict the habitats promoting pathogen persistence, which are used to identify priory areas for targeted anthrax surveillance and management.

Recent studies suggest different genetic lineages of a pathogen have distinct ecological requirements for survival and geographical distributions.[6,28,29] To date, *B. anthracis* has broadly been defined by clades A–E (i.e., lineages; E also defined as Aβ or the West Africa Group [WAG][30]). Sub-lineages (or sub-branches) of the A clade, such as Ames sub-lineage (A3.b), Vollum sub-lineage (A4), and Aust94, are globally distributed,[31,32] whereas lineages B, C, and D are

* Address correspondence to Jason K. Blackburn, University of Florida, 330 Newell Dr., Gainesville, FL 32611. E-mail: jkblackburn@ufl.edu

spatially confined[33] and the WAG group appears limited to West Africa. A previous study in Kazakhstan suggested *B. anthracis* belonging to sub-lineage A1.a were associated with a broader environmental space than the larger genetic population comprising multiple A cluster sub-lineages.[6] This study highlighted the importance of understanding genetic-environmental associations for accurately modeling the disease distribution.[6] Moreover, the study on the ENMs of *B. anthracis* on three continents and their transferability among those countries suggested the ecological associations for *B. anthracis* within each country were different and reflected niche specialization across sub-lineages.[34]

Bacillus anthracis diversity in the United States includes three sub-lineages in A clade, the A1.a (Western North American [WNA]), A3.b (Ames like), and A4 (Vollum like),[4,31] with limited outbreaks of B and C clade isolates in recent decades.[33] Different lineages/sub-lineages broadly have different spatial patterns,[31] including subnational or regional distributions. For example, outbreaks in the North Central United States appear to be dominated by the A1.a sub-lineage, whereas the A3.b and A4 sub-lineages show constrained distributions in the southwest, especially Texas.[31,35] Although the potential geographic distribution of *B. anthracis* was mapped in the continental United States using historical outbreak records,[4,34] a stated limitation could be that outbreak data included multiple genotypes of the pathogen.[36,37] Although one study modeled the A1.a sub-lineage for the United States, the modeling was focused on the comparison of the United States with Italy and Kazakhstan using a limited set of coverages available for all three landscapes.[34] Re-emergence of anthrax in Colorado in 2012, an area without reported disease since the late 1970s, yielded a diagnostic strain for genotyping, thereby contributing to our understanding of the diversity and distribution of the A1.a sub-lineage in the United States.

Here, we focused on improving the ENM-based prediction of the dominant sub-lineage of *B. anthracis* in the continental United States by combining data from new strains from recent field collections, multiple-locus variable-number tandem repeat (MLVA)-25 confirmation of lineage in three states, and using an updated bioclimatic variable set recently introduced to the modeling community and a new variable contribution estimation tool for Genetic Algorithm for Ruleset Prediction (GARP) to estimate optimal variable space.

MATERIALS AND METHODS

Anthrax occurrence data. A historical Geographical Information System (GIS) database of *B. anthracis* isolates from the continental United States was derived from previous studies,[4,9,14] and new strains from recent field collections in western Montana, Colorado, and Texas (Figure 1). The locations of those anthrax occurrence data were collected in the field with latitude/longitude matching carcass locations, farm front gate locations based on laboratory records or head-up digitized on high-resolution remote sensing imagery based on field reports or paper maps that defined to case locations.[38] The detailed information of the historical dataset has been summarized elsewhere.[4] The new strains from Montana were collected during the previously described 2008 outbreak[15] and included isolates from plains bison (*Bison bison bison*) and elk (*Cervus candensis*). One isolate was collected from the turbinates of a male bison skull in October 2010, ~27 months after it died in the 2008 outbreak. Strains from west Texas were collected from a domestic cattle/white-tailed deer (*Odcoileus virginianus*) outbreak in the summer of 2009 and an additional deer case from the same ranch in 2010. A strain from south Texas was recovered from a domestic cattle outbreak in 2008. The strain from Colorado was isolated during an outbreak in 2012 with at least 50 domestic cattle involved.[39] This outbreak occurred in an area that has not reported the disease in nearly 40 years.[39] Here, we filtered this database to include only isolates in the A1.a sub-lineage as defined by the MLVA system described in Keim et al.[40] (MLVA-8) and Lista et al.[41] (MLVA-25) (N = 160). The distribution of A1.a sub-lineage and the larger U.S. database are mapped in the Supplemental Figure.

Multiple-locus variable-number tandem repeat-25 genotyping. Strains from Montana, Colorado, and Texas were genotyped using the MLVA-25 described by Lista et al.,[41] with minor changes in PCR chemistry and adaptations in primer labeling to perform analyses on the Applied Biosystems (ABI, Foster City, CA) instruments. Multiplex PCR products were diluted 1:25 in molecular-grade water, and 0.5 µL of the diluted multiplex reactions was mixed with 9.5 µL of a formamide/LIZ 1200 (ABI) size standard mixture and denatured. Fragment sizing was performed on an ABI 3730 (Applied Biosystems) and variable-number tandem repeat (VNTR) sizes were determined using GeneMapper™ software (Applied Biosystems). We examined genetic relationships between samples in the context of global representatives from Lista et al.[41] using unweighted pair group method with arithmetic mean cluster analysis. Matrix distances were calculated in PAUP 4.0 (Sinauer Associates, Inc., Sunderland, MA) and imported into MEGA 5[42] to build phylogenetic trees based on MLVA-25. For this study, strains were used in modeling if they were closely related to GT39,[41] which is also the WNA lineage from Van Ert et al.[43]

Environmental data. We used 26 climatic and biophysical covariates to serve as the potential environmental coverages for modeling the distribution of the *B. anthracis* A1.a sub-lineage. We used the most recent bioclimatic dataset (i.e., MERRAclim) downloaded from https://datadryad.org/ as the set of temperature and moisture measurements (see Table 1).[44,45] Those 19 bioclimatic variables with a 2.5 arc minute (~4.5 × 4.5 km at the equator) resolution were built with data from the 2000s. Temperature-related layers were interpolated using hourly temperature data, and moisture-related layers were measured using specific humidity (i.e., kg of water/kg of air based on monthly mean values) rather than precipitation in millimeters like WorldClim.[45,46] Elevation data were accessed from WorldClim, which were derived from Shuttle Radar Topography Mission data. Two additional measurements of vegetation (NDVI) with 1 × 1 km spatial resolution were obtained from the Trypanosomiasis and Land Use in Africa research group (Oxford, United Kingdom),[47] and four 1-km soil variables were downloaded from the SoilGrids Web site (https://soilgrids.org/)(see Table 1). Also, the detailed description and source of all 26 environmental covariates are summarized in Table 1. These environmental layers were preprocessed using the rasterPrep function in "GARPTools" R-package, which uses a "bilinear" interpolation method to resample them to ~4.5 × 4.5-km resolution and then crops them to the same extent of the study area (i.e., the continental United States)[48] (available at https://github.com/cghaase/GARPTools).

FIGURE 1. Dendrogram of recent U.S. *Bacillus anthracis* A1.a/Western North American (WNA) lineage strains based on MLVA-25. Recent strains genotyped are color-coded and mapped along with spatially unique *B. anthracis* localities used to develop ecological niche model predictions (inset map). The dendrogram was generated using unweighted pair group method with arithmetic mean clustering. Gray genotypes are defined by Lista et al.[41] The red area in the map represents the enzootic anthrax zone in West Texas defined by Blackburn et al.[14] This figure appears in color at www.ajtmh.org.

Ecological niche modeling. Recently, a wide range of algorithms have been used to model the *B. anthracis* ecological niche. Several different models and algorithms, such as GARP,[3,16,34] boosted regression trees,[49] and random forests,[18] have been applied across spatial scales. Here, we used the GARP modeling approach to directly compare the ecological niche and predicted geographic distribution of the *B. anthracis* A1.a sub-lineage with previous studies in the continental Unitedtl States.[4] Genetic Algorithm for Ruleset Prediction is also one of the most common methods used to predict the distribution of *B. anthracis* in different regions.[3,4,6,16,34]

Genetic Algorithm for Ruleset Prediction is a presence-only modeling algorithm that iteratively searches for the non-random relationships between occurrence locality (outbreak locations) and environmental coverages (raster files of climatic or environmental data). Genetic Algorithm for Ruleset Prediction is one of the 13 key ecological niche modeling approaches available to niche modeling community[50] and has been widely applied in the literature to model species' geographic distributions.[19,51–53] The modeling system has been defined in detail elsewhere.[49,54–58] Briefly, GARP develops a series of if/then logic statements, called "rules" (range, negated range, atomic, or logit), to describe the presence or absence of the target species in ecological space. Rules are developed and tested internally using random draws of presence points from the input occurrences and pseudo-absences randomly generated by GARP from background environment. The quality of each rule is evaluated via a chi-square test on the comparison between the presence or absence prediction and the predefined proportion of input data. Rules in a GARP experiment can be accepted, modified, or deleted in a genetic fashion using deletions, insertions, crossovers, etc., to improve predictive accuracy. Once a ruleset (50 rules per model) is developed, it is projected onto the landscape to develop a presence/absence prediction describing the species' occurrence. Examples of the relationship between rulesets and geographic predictions can be found elsewhere.[38,55,59] Genetic Algorithm for Ruleset Prediction has good performance across different patterns of species distributions from endemic or disjunct to cosmopolitan and has been applied in the field of ecology, biogeography, conservation biology, evolution, and epidemiology.[21,55,60,61]

Model building and evaluation. Here, we used DesktopGARP (DG) version 1.1.3 to develop the GARP experiments.

TABLE 1
Environmental variables used for modeling the ecological niche of *Bacillus anthracis* sub-lineage A1.a

Environmental layer and description (unit)		Names	Source
Elevation (meter)		Alt	WorldClim*
Bioclimatic data (°C for temperature-related variables; kg of water/kg of air for water-related variables)	Annual mean temperature	Bio1	MERRAclim†
	Mean diurnal range	Bio2	
	Isothermality (Bio2/Bio7) (×100)	Bio3	
	Temperature seasonality (SD × 100)	Bio4	
	Max temperature of warmest month	Bio5	
	Min temperature of coldest month	Bio6	
	Temperature annual range (Bio5–Bio6)	Bio7	
	Mean temperature of most humid quarter	Bio8	
	Mean temperature of least humid quarter	Bio9	
	Mean temperature of warmest quarter	Bio10	
	Mean temperature of coldest quarter	Bio11	
	Annual mean specific humidity	Bio12	
	Specific humidity of most humid month	Bio13	
	Specific humidity of least humid month	Bio14	
	Specific humidity seasonality (coefficient of variation)	Bio15	
	Specific humidity mean of most humid quarter	Bio16	
	Specific humidity mean of least humid quarter	Bio17	
	Specific humidity mean of warmest quarter	Bio18	
	Specific humidity mean of coldest quarter	Bio19	
Mean NDVI (no unit)		wd0114a0	TALA‡
NDVI annual amplitude (no unit)		wd0114a1	TALA
Top soil pH (no unit)		pH	SoilGrids§
Sand fraction in top soil (% weight)		Sand fraction	SoilGrids
Calcium vertisols (% weight)		Calcium vertisol	SoilGrids
Top soil organic content (g per kg)		Organic content	SoilGrids

NDVI = normalized difference vegetation index.
* worldclim.org/.[48]
† https://datadryad.org/.[44,45]
‡ Trypanosomiasis and Land Use in Africa (TALA) research group.[47]
§ https://soilgrids.org/.

Readers can freely access this version of DG from GitHub (https://github.com/jkblackburn/DesktopGARP1.1.3). The model building and evaluation followed the framework proposed in a previous study.[48] DesktopGARP records the instance of the presence location in a raster cell, which has the same resolution as the environmental layers (~4.5 × 4.5 km at the equator) and, therefore, does not distinguish between multiple locations within a single cell. We used the spatially unique routine within "GARPTools" R-package to filter the occurrence dataset into 62 spatially unique points as the model input data.[59] We randomly split those spatially unique points into 75% training, 25% testing before inputting the occurrence data into DG. This external splitting procedure was also used to test the performance of the GARP experiments. The external testing sets (n = 15) were withheld from the GARP modeling experiments to calculate model accuracy,[62] whereas the remaining occurrence points (external training sets; n = 47) were used for model building. However, random splits may result in some spatial biases, particularly here, where reports of A1.a are limited in the southern states. It is plausible that training sets may ignore the southern points. To evaluate this, we developed a suite of 10 GARP experiments to evaluate the effect of input variability on model output, each with a different random subset of external training and testing data with 75%:25% ratio.[63]

Because the selection of appropriate predictors significantly affects the prediction accuracy of ENMs,[25] we adopted a new variable contribution estimation procedure typically designed for GARP[48] to get the optimal set of environmental variables. The variable contribution in GARP experiments (the Unimportance Index [UI]) is estimated based on the prevalence (i.e., frequency) of the variable being used to predict the presence of the species in the dominant presence rules in the best subset of GARP models and the scaled median range of variables.[48] We first incorporated all 26 candidate covariates in DG. We then calculated the UI for those covariates following the variable contribution evaluation procedures in a previous study.[48] In general, the variable contribution within a GARP experiment is estimated based on 1) the prevalence of each environmental variable in the dominant presence rules of the best subset of GARP model and 2) the range of the environmental coverage across the best subset.[48] A variable with higher prevalence (across rules) and narrower environmental coverage would indicate the variable is frequently used and the species' distribution is limited by this environmental condition.[48] The UI was designed to balance these two criteria. We used a threshold to select variables with UI less than 0.5 as the optimal variable set for *B. anthracis* A1.a sub-lineage and re-ran the GARP experiment using this reduced variable set.

All GARP experiments in this study were set up to run up to 200 models with a maximum of 1,000 iterations and a convergence limit of 0.01.[3,6] The external training data for model building were partitioned with a 75%:25% internal training testing split. We used the best subset procedure in DG-selected 10 top models under a 10% hard omission threshold and a 50% commission threshold for each of the experiment.[4]

TABLE 2

Multiple-locus variable-number tandem repeats (MLVA-25) for samples from Montana, Colorado, and Texas available for this study

	MLVA-25 repeats																								
Geographic location	01-vrrA	02-vrrB1	03-vrrB2	04-vrrC1	05-vrrC2	06-CG3	07-pXO1	08-pXO2	09-Bams1	10-Bams3	11-Bams5	12-Bams13	13-Bams15	14-Bams21	15-Bams22	16-Bams23	17-Bams24	18-Bams25	19-Bams28	20-Bams30	21-Bams31	22-Bams34	23-Bams44	24-Bams51	25-Bams53
West Texas; mixed livestock/wildlife outbreak 2009, 2010	9	16	7	57	21	1	8	8	13	29	7	30	49	10	16	11	10	13	14	72	64	9	12	9	8
South Texas; livestock outbreak 2008 (domestic cattle)	9	16	–1	57	21	1	8	8	13	27	7	30	49	10	16	11	10	13	14	–1	64	9	12	–1	8
Montana;	10	16	7	57	21	1	8	8	13	30	7	30	49	10	16	11	11	13	14	75	64	9	12	9	8
mixed bison, elk, deer outbreak 2008	10	16	7	57	21	1	8	8	13	30	4.1	30	49	10	16	11	11	13	14	75	64	9	12	9	8
Colorado; cattle outbreak in 2012	10	16	7	57	21	1	8	8	13	30	7	30	49	10	16	11	8	13	14	75	64	9	12	9	8

Missing alleles are coded as –1.

Finally, the best subset with 10 best presence–absence predictions in each experiment were summed and mapped on the landscape with model agreements indicating the likelihood of the species presences using the "GARPTools" R-package.

Predictive accuracies for the best subsets from 10 GARP experiments with the reduced variable set were evaluated using the area under the curve (AUC) in a receiver operating characteristics analysis based on the external testing dataset (i.e., the 25% of occurrence points withheld from the model building).[60,64] The AUC is used extensively in species distribution modeling and measures the ability of a model to discriminate between sites where a species is present, versus those where it is absent.[62,65] The AUC, although not an ideal metric for accuracy estimation,[66] is useful to identify models that perform well.[19,67] We also calculated omission and commission rates, including total omission, commission rates and average omission, and commission rates[62] (the definitions and calculations of those metrics can be found in Blackburn et al.[4]). We calculated accuracy metrics for 10 GARP experiments and ranked them by AUC, total omission, and total commission, selecting the experiment balancing high AUC, low omission, and relatively low commission. The best GARP experiment with the highest AUC value was selected to describe the ecological niche characteristics and map the potential geographic distribution of the *B. anthracis* sub-lineage A1.a in the continental United States.

RESULTS

Multiple-locus variable-number tandem repeat-25 results. Figure 1 illustrates the MLVA-25 genotypes of our isolates from Texas, Montana, and Colorado compared with the Lista et al.[41] diversity panel representing the global genetic population structure of *B. anthracis.* Multiple-locus variable-number tandem repeat analysis confirmed A1.a isolates from each Texas, Montana, and Colorado. Unique A1.a genotypes were identified in each of the three states, and repeat numbers at each locus for representatives of each genotype are reported in Table 2. All of the strains met the A1.a inclusion criteria to use strain locations in the ENM experiments.

Ecological niche modeling: estimation of variable contributions. Table 3 reports the estimation of variable contributions for all 26 candidate variables tested to model the potential distribution of *B. anthracis* A1.a sub-lineage. The lower the rescaled UI, the more the variable contributes to the prediction. We selected 13 variables with UI less than 0.5, including the seasonality of temperature and moisture, elevation, mean NDVI, seasonality of NDVI, organic contents, pH, and sand fractions of soil, as the optimal environmental affinities to describe the ecological niche of the *B. anthracis* A1.a sub-lineage.

Model accuracy and geographic distribution. Table 4 summarizes the AUC values, omission, and commission rates for the best subsets of 10 GARP experiments based on the selected 13 variables. The AUC values of 10 experiments varied from 0.831 to 0.938. The predicted geographic distribution of *B. anthracis* A1.a is illustrated in Figure 2 as the summation of the 10-model best subset for experiment 6 (see Table 4). The total and average omission rates of this best subset were 0 (suggesting all test data were predicted by all models in the best subset), and the total and average commission rates were 7.82% and 12.41%, respectively.

The predicted distribution of the *B. anthracis* A1.a sub-lineage was concentrated across the Dakotas, Minnesota, most of Montana, eastern Idaho, Nebraska and Iowa, northwestern Wyoming, southern Wisconsin, northern Illinois, North Central Kansas, and southwestern Texas (Figure 2). Some sporadic parts in western California, northern Utah, and southern Arizona were identified with the relatively high model agreement in the best subset. Figure 3 illustrates the scaled median ranges and coverages of variables in the dominant rules of the best subset in GARP experiment 6 with the reduced variable set. Organic content had the narrowest median range (5–118 g per kg soil), whereas NDVI annual amplitude had the widest median range varying from 0.006 to 0.594. Soil pH (6.23–7.75), mean NDVI (–0.15–0.42), and Bio5 (35.3–44.25°C) also had relatively narrow coverages.

DISCUSSION

This study examined the relationship of *B. anthracis* genotypes from recent outbreaks and ideal environmental coverages to better predict the geographic distribution of a single sub-lineage of *B. anthracis* (A1.a), the most geographically widespread sub-lineage in the continental United States. We identified 13 of 26 covariates of high importance in defining the ecological niche of *B. anthracis* A1.a, which included the seasonality of temperature and moisture, soil pH, organic contents, sand fraction, and vegetation index. The geographic prediction of *B. anthracis* A1.a sub-lineage was primarily constrained in North Central United States and southwestern Texas.

The distribution of *B. anthracis* A1.a sub-lineage was more conservative when compared with predictions in previous studies using confirmed anthrax outbreaks regardless of genetic lineage.[4] In the Blackburn et al.,[4] *B. anthracis* was predicted to be restricted to a north–south corridor from the Dakotas, eastern Montana, and western Minnesota, southward through western Oklahoma, central Kansas, central Nebraska, to southwestern Texas. Those estimates also predicted areas in eastern Washington, Oregon, western California, and southern New Mexico. A more recent study describing the UI function in GARPTools[48] applied the tool to reexamine variable importance in modeling the original Blackburn et al.[4] outbreak data. Again, the distribution from the Dakotas to Texas was more continuous than the predictions here. However, extensive efforts to examine genetic diversity in the United Studies using MLVA-25 reveal frequent circulation of A4 (Vollum) and A3.b (Ames-like) strains in Texas,[35,68] particularly in the recently defined enzootic zone,[14] with fewer A1.a outbreaks in Texas. Restricting the occurrence data to A1.a strains produced disjunct predictions along the north–south corridor compared with the earlier studies. A similar disjunct distribution was noted in a study focused on the genetics of Ames-like isolates.[31]

We hypothesize that these differences in predictions are driven by different ecological affinities across *B. anthracis* sub-lineages. Whereas the northern states are dominated by A1.a isolates, West Texas has a long history of anthrax with annual sporadic cases and frequent outbreaks.[69] More than 179 isolates were spread across 39 counties in Texas from 1970 to 2000.[31] In recent decades, outbreaks were continuously reported in the enzootic zones in southwestern Texas. The genetic populations of *B. anthracis* in West Texas are

TABLE 3
Estimation of variable contributions in Genetic Algorithm for Ruleset Prediction experiment

Variables	Descriptions	Prevalence	Median range	Rescaled UI
Bio5	Maximum temperature of warmest month	0.857	0.382	0
Mean NDVI	Mean NDVI	0.857	0.385	0.001
Sand fraction	Sand fraction	0.857	0.479	0.037
pH	Soil pH	0.743	0.332	0.082
Altitude	Elevations	0.8	0.477	0.109
Organic contents	Organic contents in top soil	0.629	0.26	0.112
Bio4	Temperature seasonality	0.829	0.669	0.160
Bio2	Mean diurnal range	0.629	0.383	0.234
Bio15	Specific humidity seasonality	0.686	0.491	0.266
NDVI amplitude	NDVI amplitude	0.714	0.573	0.292
Bio7	Temperature annual range	0.657	0.558	0.365
Bio8	Mean temperature of most humid quarter	0.543	0.464	0.420
Bio12	Annual mean specific humidity	0.686	0.768	0.498
Calcium vertisols	Calcium vertisols	0.543	0.551	0.527
Bio14	Specific humidity of least humid month	0.743	1	0.541
Bio13	Specific humidity of most humid month	0.664	0.768	0.543
Bio3	Isothermality	0.627	0.696	0.548
Bio10	Mean temperature of warmest quarter	0.457	0.488	0.562
Bio17	Specific humidity mean of least humid quarter	0.714	1	0.618
Bio11	Mean temperature of coldest quarter	0.6	0.795	0.704
Bio18	Specific humidity mean of warmest quarter	0.543	0.71	0.721
Bio1	Annual mean temperature	0.486	0.667	0.770
Bio16	Specific humidity mean of most humid quarter	0.514	0.73	0.802
Bio6	Minimum temperature of coldest month	0.457	0.668	0.823
Bio9	Mean temperature of least humid quarter	0.543	0.81	0.843
Bio19	Specific humidity mean of coldest quarter	0.571	1	1

NDVI = normalized difference vegetation index; UI = unimportance index. Here, we use 0.5 as the threshold to select variables that are relatively more important to *Bacillus* anthracis A1.a sub-lineage based on rescaled UI.

diverse with at least three sub-lineages of *B. anthracis* A clades included. A localized ENM to study the distributions of different strains may, therefore, be needed to accurately model the anthrax risk in Texas.

Our findings here suggest genotype-specific models focusing on the association between a single genetic group and their response to the environment improves model accuracy and predictive power. The best experiment in this study showed high model accuracy with an AUC value of 0.93, which indicates a reliable model output and prediction. Notably, all of the experiments performed well (Table 4). The model accuracy for *B. anthracis* A1.a sub-lineage increased significantly, compared with previous U.S. modeling efforts.[4,48]

Here, we identified the optimal variable set for the *B. anthracis* A1.a sub-lineage as the combination of the seasonality of temperature (Bio2, Bio4, Bio5, Bio7, and Bio8) and moisture (Bio12 and Bio15), elevation, mean NDVI, seasonality of NDVI, organic contents, pH, and sand fractions based on our estimation of variable contributions. This selection correlates well with previous knowledge about the ecology of *B. anthracis.*[37] Blackburn et al.[4] used a different and limited set of covariates from other sources to develop the original *B. anthracis* prediction in the United States, including the annual trend of climate (i.e., mean annual temperature and precipitation), elevation, mean NDVI, soil moisture, and pH. Recently, we used an exhaustive method, examining UI, to select 12 of 26 variables to describe the ecological niche of *B. anthracis* with multiple sub-lineages in the A Clade, with calcium vertisols and Bio1 being selected rather than the Bio4, Bio7, and Bio12 selected in this study.[48] Although the randomness in GARP algorithms are hard to avoid, the employment of the same modeling approach, variable selection

TABLE 4
Model accuracy metrics of GARP experiments with the optimal variable set

GARP experiment	Area under the curve	SE	*z*-score	Total omission	Average omission	Total commission	Average commission
6	0.9379	0.0433	6.78	0	0	12.41	7.82
3	0.9008	0.0531	6.36	0	6.67	12.83	7.82
4	0.8980	0.0538	6.55	0	0	20.4	7.82
8	0.8973	0.0539	6.44	0	0.67	18.78	7.82
2	0.8960	0.0542	6.63	0	0	20.79	7.82
9	0.8841	0.0566	6.57	0	0	23.18	7.82
10	0.8704	0.0611	4.02	0.07	7.29	15.56	7.82
7	0.8566	0.0622	3.79	0.07	17.96	9.65	7.82
1	0.8378	0.0675	3.85	0.13	15.24	12.15	7.82
5	0.8310	0.0641	3.85	0.07	13.9	19.93	7.82

GARP = Genetic Algorithm for Ruleset Prediction.

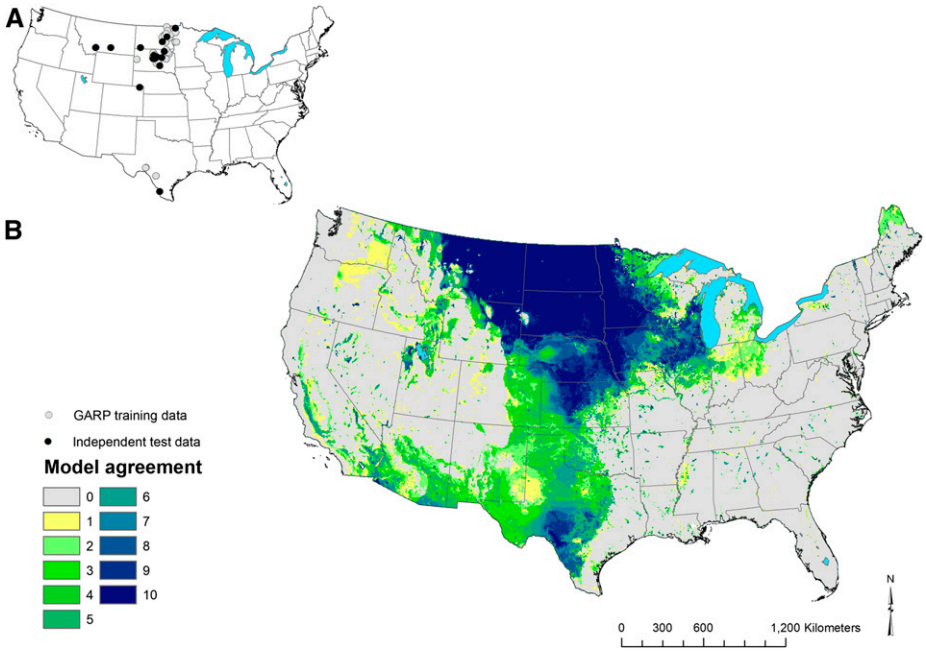


FIGURE 2. The prediction of *Bacillus anthracis* A1.a sub-lineage in the continental United States from the best performing best subset in the Genetic Algorithm for Ruleset Prediction (GARP) experiment using the selected variable set. The value of each pixel is the model agreement. It shows how many models from the 10 best models in the best subset of the GARP experiment 6 predict the presence of the species in this pixel, which indicates the likelihood of the species occurrence. Areas of darker blue represent areas of greatest model agreement. This figure appears in color at www.ajtmh.org.

framework, and candidate environmental layers (26 variables) between this study and Yang et al.[48] allows us to directly compare the two model results to assess the differences of the ecological niche within A Clade in the continental United States.

The GARP experiments in this study described the unique ecological niche of the *B. anthracis* A1.a sub-lineage, one of three major sub-lineages in the United States and the most geographically widespread. Given the exploration of environmental coverages of GARP outputs, the presence of the *B. anthracis* A1.a sub-lineage is predicted to be found in areas with high soil pH (6.23–7.75), low sand fraction (18.79–52.84%), organic contents ranging between 5 and 118 g/kg soil, mean NDVI from –0.13 to 0.42, NDVI amplitude from 0.01 to 0.59, relatively low elevation (151–1,755 m), and some highly varied climate conditions (detailed in Figure 3). Comparing with the environmental coverages suggested in Yang et al.[48] for multiple genetic lineages, some environmental covariates have similar coverages, such as mean NDVI and Bio2. However, other covariates exhibit narrower coverages for the *B. anthracis* A1.a sub-lineage, like organic contents, sand fraction, Bio15, Bio8, and Bio5. This may be because most of the occurrence data for the *B. anthracis* A1.a sub-lineage is the subset of the data used in multi-strain studies. There were also some covariates suggesting unique ranges with the larger envelopes for the A1.a sub-lineage. For example, NDVI annual amplitude was 0.01–0.38 in the previous study, but here we found a wider range (0.006–0.594); elevation was 134.8–1,321.95 m in the previous study, whereas here we found the range of 150.84–1,755.1 m. For soil pH, the previous study of multiple strains suggested the optimal median range was 6.52–8.19, whereas soil pH for A1.a sub-lineage was 6.23–7.75. These findings likely result from the incorporation of new occurrence data in Montana and Colorado. The two bison cases in Montana were found in the areas with soil pH of 6.24 and 6.52, elevations of 1,755.15 and 1,717.93 m, and NDVI annual amplitude of 0.43 and 0.28. Because the soil pH in the continental United States ranges from 4.02 to 8.87, the suitable soil pH for A1.a sub-lineage are still considered to be relatively alkaline. These larger ecological envelopes support the hypothesis that *B. anthracis* A1.a sub-lineage has broad environmental tolerances that contribute to its broad geographical distributions globally.[6,33] Our results highlight the importance of understanding the environmental affinities of individual lineages or sub-lineages of *B. anthracis* to describe the ecology of the disease and estimate the disease risk on the landscape.[6]

Here, we show improved model accuracy for mapping the extent of *B. anthracis* A1.a in the continental United States using a combination of new genetic subtyping data, modern coverages, updated outbreak information, and novel ENM tools. Integrating MLVA-25 data improved mapping efforts and suggests different lineages are adapted to different ecological conditions. The use of new climatic data improves resolution to delineate areas of the Dakotas and west Texas. The changes in prediction in Texas using the A1.a model supports the concept that the A4 and A3b lineages, both frequently identified as the primary lineages in enzootic outbreaks in the areas, likely have different environmental factors driving persistence. Ongoing efforts to model these lineages are underway using these new tools and up-to-date outbreak data. This approach is necessary for improving mapping efforts to support near-term disease surveillance and control and to guide longer term planning and decision-making. This disease remains a high priority across the American West, a vast and difficult-to-traverse landscape.[4]

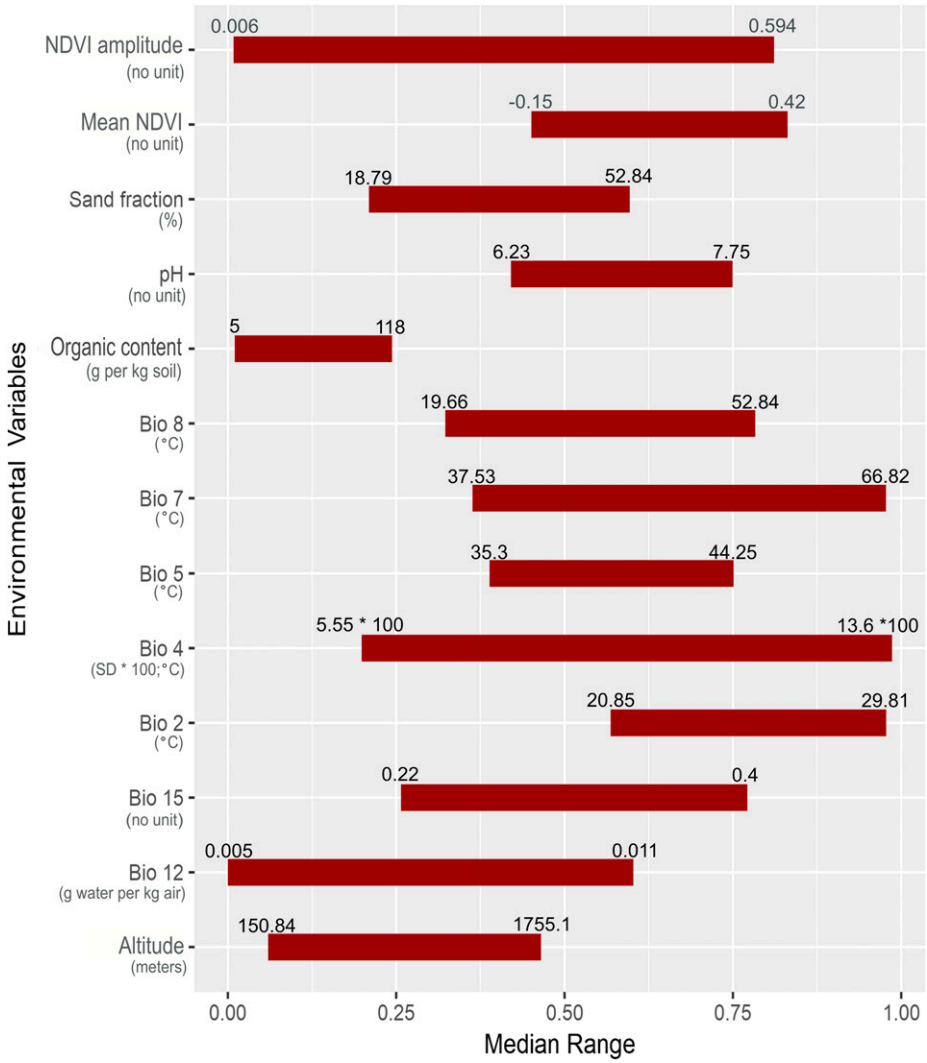


FIGURE 3. The scaled median range of the covariates from the best performing best subset in the Genetic Algorithm for Ruleset Prediction experiment 6 using the selected variable set. Values above each bar describe the unscaled coverage with true values for each covariate. This figure appears in color at www.ajtmh.org.

Control efforts for wildlife continue to rely on rapid identification of carcasses and burning; identifying where to best expend resources for these efforts remains a critical priority. With the prediction maps serving as the proxy of anthrax risk, these efforts will improve the efficacy of disease control, carcass surveillance, and disease management in the continental United States.[6,16]

Received March 6, 2019. Accepted for publication October 23, 2019.

Published online December 2, 2019.

Note: Supplemental figure appears at www.ajtmh.org.

Financial support: This study was supported by the National Institutes of Health [grant number 1R01GM117617] to J. K. B.

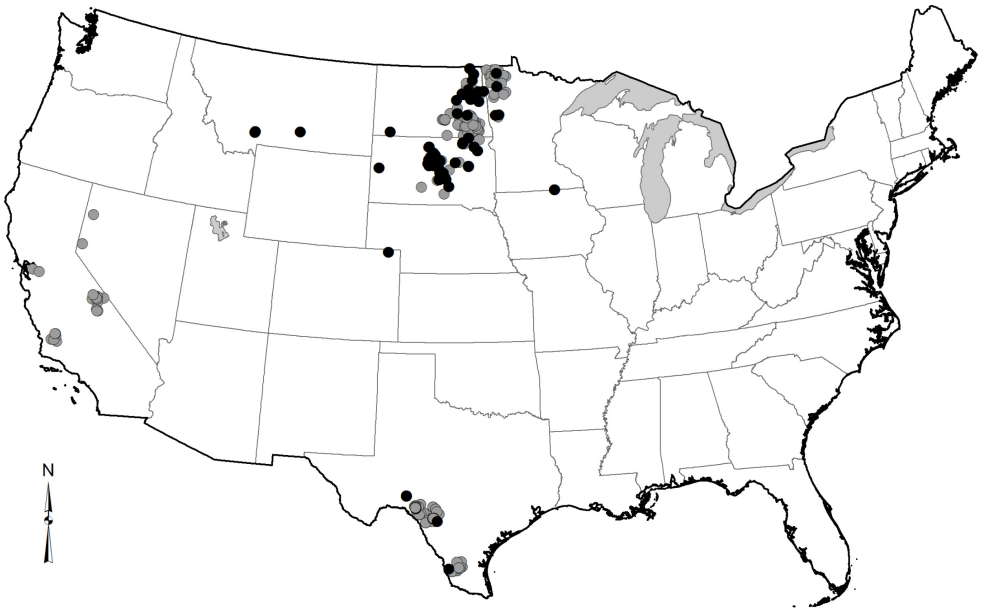
Authors' addresses: Anni Yang, Ted L. Hadfield, and Jason K. Blackburn, Spatial Epidemiology & Ecology Research Laboratory, Department of Geography, University of Florida, Gainesville, FL, and Emerging Pathogens Institute, University of Florida, Gainesville, FL, E-mails: yangann1@ufl.edu, thadfield@ufl.edu, and jkblackburn@ufl.edu. Jocelyn C. Mullins, University of Saint Joseph, West Hartford, CT, E-mail: jcmvet36@gmail.com. Matthew Van Ert, Van Ert Nemoto Consulting, LLC, Vero Beach, FL, E-mail: vanert_associatesllc@yahoo.com. Richard A. Bowen, Animal Reproduction and Biotechnology Laboratory, Department of Biomedical Sciences, Colorado State University, Fort Collins, CO, E-mail: rbowen@rams.colostate.edu.

REFERENCES

1. Alexander KA, Lewis BL, Marathe M, Eubank S, Blackburn JK, 2012. Modeling of wildlife-associated zoonoses: applications and caveats. *Vector Borne Zoonotic Dis 12:* 1005–1018.
2. Van Ness GB, 1971. Ecology of anthrax. *Science 172:* 1303–1307.
3. Barro AS, Fegan M, Moloney B, Porter K, Muller J, Warner S, Blackburn JK, 2016. Redefining the Australian anthrax belt: modeling the ecological niche and predicting the geographic distribution of *Bacillus anthracis*. *PLoS Negl Trop Dis 10:* e0004689.
4. Blackburn JK, McNyset KM, Curtis A, Hugh-Jones ME, 2007. Modeling the geographic distribution of *Bacillus anthracis*, the causative agent of anthrax disease, for the contiguous United States using predictive ecologic niche modeling. *Am J Trop Med Hyg 77:* 1103–1110.
5. Joyner TA, Lukhnova L, Pazilov Y, Temiralyeva G, Hugh-Jones ME, Aikimbayev A, Blackburn JK, 2010. Modeling the potential distribution of *Bacillus anthracis* under

multiple climate change scenarios for Kazakhstan. *PloS One 5:* e9596.
6. Mullins J, Lukhnova L, Aikimbayev A, Pazilov Y, Van Ert M, Blackburn JK, 2011. Ecological niche modelling of the *Bacillus anthracis* A1. a sub-lineage in Kazakhstan. *BMC Ecol 11:* 32.
7. Stein CD, 1945. The history and distribution of anthrax in livestock in the United States. *Vet Med 40:* 340–349.
8. Sternbach G, 2003. The history of anthrax. *J Emerg Med 24:* 463–467.
9. Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Ravel J, Hoffmaster AR, Trim CP, Chung W-K, Beaudry JA, 2009. Pre-columbian origins for North American anthrax. *PLoS One 4:* e4813.
10. Stein CD, Van Ness GB, 1955. A ten year survey of anthrax in livestock with special reference to outbreaks in 1954. *Vet Med 50:* 579–588.
11. Hugh-Jones ME, De Vos V, 2002. Anthrax and wildlife. *Rev Sci Tech Off Int Epizoot 21:* 359–384.
12. Watson R, 2012. Public wildlife on private land: unifying the split estate to enhance trust resources. *Duke Environ Law Policy Forum 23:* 291.
13. Morris LR, Blackburn JK, 2016. Predicting disease risk, identifying stakeholders, and informing control strategies: a case study of anthrax in Montana. *Ecohealth 13:* 262–273.
14. Blackburn JK, Hadfield TL, Curtis AJ, Hugh-Jones ME, 2014. Spatial and temporal patterns of anthrax in white-tailed deer, *Odocoileus virginianus*, and hematophagous flies in west Texas during the summertime anthrax risk period. *Ann Assoc Am Geogr 104:* 939–958.
15. Blackburn JK, Asher V, Stokke S, Hunter DL, Alexander KA, 2014. Dances with anthrax: wolves (*Canis lupus*) kill anthrax bacteremic plains bison (*Bison bison bison*) in southwestern Montana. *J Wildl Dis 50:* 393–396.
16. Morris LR, Proffitt KM, Asher V, Blackburn JK, 2016. Elk resource selection and implications for anthrax management in Montana. *J Wildl Manag 80:* 235–244.
17. Bellan SE, Gimenez O, Choquet R, Getz WM, 2013. A hierarchical distance sampling approach to estimating mortality rates from opportunistic carcass surveillance data. *Methods Ecol Evol 4:* 361–369.
18. Kracalik IT, Kenu E, Ayamdooh EN, Allegye-Cudjoe E, Polkuu PN, Frimpong JA, Nyarko KM, Bower WA, Traxler R, Blackburn JK, 2017. Modeling the environmental suitability of anthrax in Ghana and estimating populations at risk: implications for vaccination and control. *PLoS Negl Trop Dis 11:* e0005885.
19. Sloyer K, Burkett-Cadena ND, Yang A, Corn JL, Vigil SL, McGregor BL, Wisely SM, Blackburn JK, 2019. Ecological niche modeling the potential geographic distribution of four *Culicoides* species of veterinary significance in Florida, USA. *PloS One 14:* e0206648.
20. Porretta D, Mastrantonio V, Amendolia S, Gaiarsa S, Epis S, Genchi C, Bandi C, Otranto D, Urbanelli S, 2013. Effects of global changes on the climatic niche of the tick Ixodes ricinus inferred by species distribution modelling. *Parasites Vectors 6:* 271.
21. Sweeney AW, Beebe NW, Cooper RD, 2007. Analysis of environmental factors influencing the range of anopheline mosquitoes in northern Australia using a genetic algorithm and data mining methods. *Ecol Model 203:* 375–386.
22. de Oliveira SV, Escobar LE, Peterson AT, Gurgel-Gonçalves R, 2013. Potential geographic distribution of hantavirus reservoirs in Brazil. *PLoS One 8:* e85137.
23. González C, Wang O, Strutz SE, González-Salazar C, Sánchez-Cordero V, Sarkar S, 2010. Climate change and risk of leishmaniasis in North America: predictions from ecological niche models of vector and reservoir species. *PLoS Negl Trop Dis 4:* e585.
24. Peterson AT, Sánchez-Cordero V, Beard CB, Ramsey JM, 2002. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. *Emerg Infect Dis 8:* 662.
25. Araujo MB, Guisan A, 2006. Five (or so) challenges for species distribution modelling. *J Biogeogr 33:* 1677–1688.
26. Hutchinson GE, 1957. *Concluding Remarks*. In: Cold Spring Harbor symposia on quantitative biology, Vol 22, 415–427. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
27. Grinnell J, 1917. The niche-relationships of the California thrasher. *Auk 34:* 427–433.
28. Nakazawa Y, Williams RA, Peterson AT, Mead PS, Kugeler KJ, Petersen JM, 2010. Ecological niche modeling of *Francisella tularensis* subspecies and clades in the United States. *Am J Trop Med Hyg 82:* 912–918.
29. Fisher MC, Hanage WP, De Hoog S, Johnson E, Smith MD, White NJ, Vanittanakom N, 2005. Low effective dispersal of asexual genotypes in heterogeneous landscapes by the endemic pathogen *Penicillium marneffei*. *PLoS Pathog 1:* e20.
30. Blackburn JK, Odugbo MO, Van Ert M, O'Shea B, Mullins J, Perrenten V, Maho A, Hugh-Jones M, Hadfield T, 2015. *Bacillus anthracis* diversity and geographic potential across Nigeria, Cameroon and Chad: further support of a novel west African lineage. *PLoS Negl Trop Dis 9:* e0003931.
31. Simonson TS, Okinaka RT, Wang B, Easterday WR, Huynh L, U'Ren JM, Dukerich M, Zanecki SR, Kenefic LJ, Beaudry J, 2009. *Bacillus anthracis* in China and its relationship to worldwide lineages. *BMC Microbiol 9:* 71.
32. Khmaladze E, Birdsell DN, Naumann AA, Hochhalter CB, Seymour ML, Nottingham R, Beckstrom-Sternberg SM, Beckstrom-Sternberg J, Nikolich MP, Chanturia G, 2014. Phylogeography of *Bacillus anthracis* in the country of Georgia shows evidence of population structuring and is dissimilar to other regional genotypes. *PloS One 9:* e102651.
33. Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, Zanecki SR, Pearson T, Simonson TS, U'Ren JM, 2007. Global genetic population structure of *Bacillus anthracis*. *PloS One 2:* e461.
34. Mullins JC, Garofolo G, Van Ert M, Fasanella A, Lukhnova L, Hugh-Jones ME, Blackburn JK, 2013. Ecological niche modeling of *Bacillus anthracis* on three continents: evidence for genetic-ecological divergence? *PloS One 8:* e72451.
35. Blackburn JK, Mullins JC, Van Ert M, Hadfield T, O'Shea B, Hugh-Jones ME, 2014. The necrophagous fly anthrax transmission pathway: empirical and genetic evidence from a wildlife epizootic in west Texas 2010. *Vector Borne Zoonotic Dis 14:* 576–583.
36. Blackburn JK, 2010. Integrating geographic information systems and ecological niche modeling into disease ecology: a case study of *Bacillus anthracis* in the United States and Mexico. In: O'Connell KP, Sulakvelidze A, Bakanidze L, eds. *Emerging and Endemic Pathogen: Advance in Surveillance, Detection, and Identification*. New York, NY: Springer, 59–88.
37. Hugh-Jones M, Blackburn J, 2009. The ecology of *Bacillus anthracis*. *Mol Asp Med 30:* 356–367.
38. Blackburn JK, 2006. *Evaluating the Spatial Ecology of Anthrax in North America: Examining Epidemiological Components Across Multiple Geographic Scales Using a GIS-Based Approach*. Baton Rouge, LA: Louisiana State University, 141.
39. ProMED-Mail, 2012. Anthrax, Bovine–USA (03): (CO). Archive Number: 20120809.1236375.
40. Keim P, Price L, Klevytska A, Smith K, Schupp J, Okinaka R, Jackson P, Hugh-Jones M, 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J Bacteriol 182:* 2928.
41. Lista F, Faggioni G, Valjevac S, Ciammaruconi A, Vaissaire J, Le Doujet C, Gorgé O, De Santis R, Carattoli A, Ciervo A, 2006. Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiol 6:* 33.
42. Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol 24:* 1596–1599.
43. Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, Zanecki SR, Pearson T, Simonson TS, U'Ren JM, 2007. Global genetic population structure of *Bacillus anthracis*. *PloS One 2:* e461.
44. Vega G, Pertierra L, Olalla-Tárraga M, 2016. Data from: MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Dryad Digital Repository*. Available at: https://doi.org/10.5061/dryad.s2v81.2. Accessed by July 21, 2018.
45. Vega GC, Pertierra LR, Olalla-Tárraga MÁ, 2017. MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling. *Sci Data 4:* 170078.

46. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol 25:* 1965–1978.
47. Hay SI, Tatem AJ, Graham AJ, Goetz SJ, Rogers DJ, 2006. Global environmental data for mapping infectious disease distribution. *Adv Parasitol 62:* 37–77.
48. Yang A, Gomez JP, Blackburn JK, 2019. Exploring environmental coverages of species: a new variable selection methodology for rulesets from the genetic algorithm for ruleset prediction. *bioRxiv ■:* 531079.
49. Carlson CJ, Kracalik IT, Ross N, Alexander KA, Hugh-Jones ME, Fegan M, Elkin BT, Epp T, Shury TK, Zhang W, 2019. The global distribution of *Bacillus anthracis* and associated anthrax risk to humans, livestock and wildlife. *Nat Microbiol 4:* 1337–1343.
50. Ahmed SE, McInerny G, O'Hara K, Harper R, Salido L, Emmott S, Joppa LN, 2015. Scientists and software–surveying the species distribution modelling community. *Divers Distrib 21:* 258–267.
51. Padalia H, Srivastava V, Kushwaha S, 2014. Modeling potential invasion range of alien invasive species, *Hyptis suaveolens* (L.) Poit. in India: comparison of MaxEnt and GARP. *Ecol Inf 22:* 36–43.
52. Qin Z, Zhang J, DiTommaso A, Wang R, Wu R, 2015. Predicting invasions of *Wedelia trilobata* (L.) Hitchc. with Maxent and GARP models. *J Plant Res 128:* 763–775.
53. Chikerema SM, Gwitira I, Murwira A, Pfukenyi DM, Matope G, . Comparison of GARP and Maxent in modelling the geographic distribution of *Bacillus anthracis* in Zimbabwe. *Zimb Vet J* 2017; *35:* 1–6.
54. Stockwell D, 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *Int J Geogr Inf Sci 13:* 143–158.
55. Stockwell DR, Beach JH, Stewart A, Vorontsov G, Vieglais D, Pereira RS, 2006. The use of the GARP genetic algorithm and internet grid computing in the Lifemapper world atlas of species biodiversity. *Ecol Model 195:* 139–145.
56. Martinez-Meyer E, Peterson AT, Servín JI, Kiff LF, 2006. Ecological niche modelling and prioritizing areas for species reintroductions. *Oryx 40:* 411–418.
57. Peterson A, Cohoon K, 1999. Sensitivity of distribution prediction algorithms to geographic completeness. *Ecol Model 117:* 159–164.
58. Lippi CA, Stewart-Ibarra AM, Loor MFB, Zambrano JED, Lopez NAE, Blackburn JK, Ryan SJ, 2019. Geographic shifts in *Aedes aegypti* habitat suitability in Ecuador using larval surveillance data and ecological niche modeling: implications of climate change for public health vector control. *PLoS Negl Trop Dis 13:* e0007322.
59. Joyner TA, 2010. *Ecological Niche Modeling of a Zoonosis: A Case Study Using Anthrax Outbreaks and Climate Change in Kazakhstan*. Gainesville, FL: University of Florida.
60. Peterson AT, Papeş M, Eaton M, 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography 30:* 550–560.
61. Raxworthy CJ, Martinez-Meyer E, Horning N, Nussbaum RA, Schneider GE, Ortega-Huerta MA, Peterson AT, 2003. Predicting distributions of known and unknown reptile species in Madagascar. *Nature 426:* 837.
62. McNyset K, 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. *Ecol Freshw Fish 14:* 243–255.
63. Blackburn JK, Matakarimov S, Kozhokeeva S, Tagaeva Z, Bell LK, Kracalik IT, Zhunushov A, 2017. Modeling the ecological niche of *Bacillus anthracis* to map anthrax risk in Kyrgyzstan. *Am J Trop Med Hyg 96:* 550–556.
64. Lim B, Klein KJ, 2006. Team mental models and team performance: a field study of the effects of team mental model similarity and accuracy. *J Organ Behav 27:* 403–418.
65. Wiley EO, McNyset KM, Peterson AT, Robins CR, Stewart AM, 2003. Niche modeling perspective on geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography 16:* 120–127.
66. Lobo JM, Jiménez-Valverde A, Real R, 2008. AUC: a misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr 17:* 145–151.
67. Hanley JA, McNeil BJ, 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology 143:* 29–36.
68. Mullins J, 2013. *Combining Genetic Diversity and Spatio-Temporal Data to Characterize the Spatial Ecology of Anthrax across Multiple Scales*. University of Florida.
69. Blackburn JK, Goodin DG, 2013. Differentiation of springtime vegetation indices associated with summer anthrax epizootics in west Texas, USA, deer. *J Wildl Dis 49:* 699–703.



- A1.a isolates modeled in this study
- Comfirmed outbreaks from other or untyped lineages (Blackburn et al. 2007 [4]; Mullins et al. 2013 [58])

Lakes

0 375 750 1,500 Kilometers