Range-wide Determinants of Plague Distribution in North America

Sean P. Maher,* Christine Ellis, Kenneth L. Gage, Russell E. Enscore, and A. Townsend Peterson

Department of Ecology and Evolutionary Biology and Biodiversity Institute, University of Kansas, Lawrence, Kansas; College of Veterinary Medicine, Kansas State University, Manhattan, Kansas; Centers for Disease Control and Prevention, National Center for Infectious Diseases, Division of Vector Borne Infectious Diseases, Fort Collins, Colorado

Abstract. Plague, caused by the bacterium *Yersinia pestis*, is established across western North America, and yet little is known of what determines the broad-scale dimensions of its overall range. We tested whether its North American distribution represents a composite of individual host–plague associations (the “Host Niche Hypothesis”), or whether mammal hosts become infected only at sites overlapping ecological conditions appropriate for plague transmission and maintenance (the “Plague Niche Hypothesis”). We took advantage of a novel data set summarizing plague records in wild mammals newly digitized from paper-based records at the Centers for Disease Control and Prevention to develop range-wide tests of ecological niche similarity between mammal host niches and plague-infected host niches. Results indicate that plague infections occur under circumstances distinct from the broader ecological distribution of hosts, and that plague-infected niches are similar among hosts; hence, evidence coincides with the predictions of the Plague Niche Hypothesis, and contrasts with those of the Host Niche Hypothesis. The “plague niche” is likely driven by ecological requirements of vector flea species.

INTRODUCTION

Spatial scale is a critical element in an integrative understanding of ecology. In disease ecology, local-scale studies have dominated, whereas synthetic, broader-scale analyses have been less frequent. Integrative, cross-scale analyses—i.e., detailed understanding of local-level processes placed in the context of regional processes of range limitation and biogeography—are rare, and have been developed for few disease systems.

Plague, caused by infections of the bacterium *Yersinia pestis*, is transmitted among susceptible hosts by bites of fleas infected previously by feeding on other hosts that were highly bacteremic. Such transmission is most effective for many flea species within the first few days after taking an infectious host blood meal, and later for a smaller number of flea species in which midgut blockage occurs. Plague bacteria also may survive outside living hosts on carcasses or in the soil, but only limited evidence exists to suggest that such mechanisms are important for long-term survival in nature. Plague likely evolved in Asia and has since spread broadly by various means, including the transport of infected hosts and fleas along overland trade routes or aboard rat-infested ships during the three historically documented pandemics. The third of these pandemics and that most relevant to this study introduced plague to North American ports in the late nineteenth century; where it afterwards “escaped” into native rodent populations, eventually spreading east through the Rocky Mountains to the western edge of the Great Plains. Wild carnivores also are frequently found to be seropositive, but probably are of relatively little importance as sources of infection for vector fleas, and thus these species likely play no direct role in the natural host-to-flea-to-host cycle. However, these animals might be important as temporary flea hosts, and may transport infected rodent fleas from one site to another. In the face of such complex host relationships, factors that maintain the geographic distribution of plague are not well understood; hence, in this contribution, we propose and test two contrasting hypotheses regarding the distribution, ecology, and geography of plague at coarse resolutions across western North America.

The “Host Niche Hypothesis” (HNH) postulates that plague distributions are mediated by host distributions, such that the distribution of plague depends on an amalgam of host ranges, and the presence of a particular host species could extend the distributional potential of the pathogen (Figure 1). Alternatively, plague may have its own distinct ecological niche, with infections occurring only in regions where hosts’ distributions overlap this ecological and geographic potential (Figure 1). This “Plague Niche Hypothesis” (PNH) suggests that plague distribution is independent of particular host distributions, but rather is mediated by other factors such as vector ecology. Under this view, plague occurs in a particular mammal taxon only if and where its distribution overlaps sites fitting the conditions of the plague niche.

Here, we test the HNH and PNH using a framework incorporating ecological niche models. The HNH predicts that plague infections in hosts will not be distinct ecologically from the overall ecology of the host; PNH, on the other hand, predicts differences between the two. More importantly, HNH predicts distinct ecological profiles for plague-infected distributions of different host species, whereas PNH would expect exposure to the bacterium. Additional potential rodent hosts are known, and plague has been proposed to be enzootic in some, although the role of deer mice (*Peromyscus* spp.) as enzootic hosts was challenged recently. Nevertheless, most mammals are clearly susceptible. Another small rodent, the grasshopper mouse, *Onychomys leucogaster*, is relatively resistant to plague in some areas, leading some to propose that it might play a role in long-term maintenance and transmission of plague over its widespread range in North America.

Wild carnivores also are frequently found to be seropositive, but probably are of relatively little importance as sources of infection for vector fleas, and thus these species likely play no direct role in the natural host-to-flea-to-host cycle. However, these animals might be important as temporary flea hosts, and may transport infected rodent fleas from one site to another. In the face of such complex host relationships, factors that maintain the geographic distribution of plague are not well understood; hence, in this contribution, we propose and test two contrasting hypotheses regarding the distribution, ecology, and geography of plague at coarse resolutions across western North America.

The “Host Niche Hypothesis” (HNH) postulates that plague distributions are mediated by host distributions, such that the distribution of plague depends on an amalgam of host ranges, and the presence of a particular host species could extend the distributional potential of the pathogen (Figure 1). Alternatively, plague may have its own distinct ecological niche, with infections occurring only in regions where hosts’ distributions overlap this ecological and geographic potential (Figure 1). This “Plague Niche Hypothesis” (PNH) suggests that plague distribution is independent of particular host distributions, but rather is mediated by other factors such as vector ecology. Under this view, plague occurs in a particular mammal taxon only if and where its distribution overlaps sites fitting the conditions of the plague niche.

Here, we test the HNH and PNH using a framework incorporating ecological niche models. The HNH predicts that plague infections in hosts will not be distinct ecologically from the overall ecology of the host; PNH, on the other hand, predicts differences between the two. More importantly, HNH predicts distinct ecological profiles for plague-infected distributions of different host species, whereas PNH would expect exposure to the bacterium. Additional potential rodent hosts are known, and plague has been proposed to be enzootic in some, although the role of deer mice (*Peromyscus* spp.) as enzootic hosts was challenged recently. Nevertheless, most mammals are clearly susceptible. Another small rodent, the grasshopper mouse, *Onychomys leucogaster*, is relatively resistant to plague in some areas, leading some to propose that it might play a role in long-term maintenance and transmission of plague over its widespread range in North America.
Taxidea taxus

spp., latrans taxa showing some of the greatest densities of records: for 75 mammal taxa. We focused analyses on the wild mammal points that included generic or specific reference to host species. From the resulting data set, we georeferenced 3,777 occurrence results. Animals were considered plague-positive when samples was not possible to summarize information about contents of a massive paper-based archive of records of plague- distinguishing between these two views of plague ecology.

those profiles to be similar. Previous efforts have shown that the geography of Y. pestis occurrence in humans (who are incidental hosts of this bacterium) is highly predictable using the niche modeling approach—in other words, coarse-scale spatial correlates exist that provide a predictive view of human plague geography. Here, we use new niche-based tools to distinguish between these two views of plague ecology.

MATERIALS AND METHODS

Input data. We captured digitally on spreadsheets the contents of a massive paper-based archive of records of plague-positive samples in North America maintained at the Centers for Disease Control and Prevention, Ft. Collins, CO; however, given the large-scale retrospective capture of these data, it was not possible to summarize information about negative test results. Animals were considered plague-positive when samples of their tissues tested positive by one of the following assays: serology, direct immunofluorescence, or bacterial culture. From the resulting data set, we georeferenced 3,777 occurrence points that included generic or specific reference to host species for 75 mammal taxa. We focused analyses on the wild mammal taxa showing some of the greatest densities of records: Canis latrans (2,516 points), ground squirrels (including Ictonomys spp., Callospermophilus spp., and Urocitellus spp.; 150), Tamias taxus (106), tree squirrels and chipmunks (including Tamias spp., Tamiasciurus hudsonicus, and Sciurus spp.; 100), Cynomys (including Cynomys gunnisoni, C. leucurus, C. ludovicianus, and C. parvidens; 69), Peromyscus spp. (69), Ursus americanus (43), and Neotoma spp. (29; see Figure 2). Geographic coordinates were derived for all of these records following standard point and error radius protocols by means of referring to Terrain Navigator (http://www.maptech.com), Topozone (http://www.topozone.com), and Graphical Locator (http://www.esg.montana.edu). Records showing uncertainty radii of > 10 km were removed from analyses, as were redundant localities associated with the same host.

Host mammal occurrence data (i.e., general occurrences not associated with plague infections) were collated from records provided online from 29 natural history museums by the MaNIS (http://manisnet.org/) and Arctos (http://arctos.database.museum/) biodiversity data portals (see Figure 2). In addition to the taxa listed above, we gathered O. leucogaster occurrence data for inclusion in analyses, in light of its possible role in plague maintenance. Geographic coordinates had been derived previously for these records as part of the MaNIS project, following standard point and error radius protocols; again, all records showing uncertainty radii of > 10 km were removed from analyses, and all redundant localities were removed. We pooled occurrence data across host taxa in groups, as described previously for plague occurrence data. Host occurrence points were abundant (541–5,614 per species), and were distributed reasonably evenly across the ranges of each host species.

To characterize environmental variation across North America, we used climatic layers from the WorldClim data archive (http://www.worldclim.org/) at a spatial resolution of 5 km, matching the approximate precision of occurrence data. To avoid fitting models in overly dimensional environmental spaces, we assessed patterns of correlation among the 19 bioclimatic variables from WorldClim, and chose seven that were relatively uncorrelated: annual mean temperature, mean diurnal temperature range, maximum temperature of warmest month, minimum temperature of coldest month, annual precipitation, precipitation of wettest month, and precipitation of driest month. All of our analyses are based on fitting models in this seven-dimensional space, in a study area extending across southern Canada and the continental United States. The geographic ranges of some host taxa extend quite broadly, so this region represents the union of areas of potential occurrence for these hosts; hence, our comparisons were drawn from the same region regardless of host identity. As detailed below, the niche comparisons that form the basis for our inferences involve specification of more restricted arenas on the basis of occurrence of plague-infected hosts.

Niche modeling. Ecological niche models were generated for each host assemblage (i.e., species or set of related species) and for plague-infected individuals of each host assemblage using a maximum-entropy-based algorithm called Maxent. This algorithm calculates a probability distribution from the overall environmental data and occurrence data on the basis of the principle of maximum entropy: i.e., that the best explanation for a phenomenon is that which shows the broadest and most-spread-out probability distribution. Maxent fits this distribution subject to particular constraints, in this case, environmental values associated with known presences.

The output from Maxent is thus a continuous grid of probabilities ranging between 0 and unity, although overfitting has
been a concern in some previous studies, where output values are high only in certain heavily sampled regions. To reduce this bias, we used a threshold to convert continuous output maps into binary maps that present more conservative estimates of presence and absence. This step carries a series of assumptions, which we sought to simplify as much as possible. As a simplest option, we used the minimum probability value (defined from the logistic output from Maxent) assigned to any training occurrence point for a given assemblage as a minimum value for a prediction of suitability (i.e., potential for presence). Grid cells with values equal to or greater than this threshold were considered as indicating suitability, whereas the remaining cells were considered as indicative of unsuitability.

Measuring niche similarity. Warren and others presented tools for assessing identity and similarity of ecological niches. Given that many of the ranges of the various species examined in this study differ in their overall extents, we focused on Warren and others background similarity measures, to avoid the false-positive errors that are common in niche identity tests. We calculated the Hellinger’s-based $I$ and Schoener’s $D$ indices for all pairwise combinations of plague-infected records of the eight host assemblages and 2) comparing overall versus plague-infected records within each host assemblage.

For the background similarity tests, a sampling region must be designated, preferably one that represents the accessible geographic area for the taxon, in essence the “M” or mobility constraint in the BAM diagram presented by Soberón and Peterson. Accordingly, under a simple assumption of uniform dispersal distance, we generated minimum convex polygons around the set of plague-related points for all taxa, which we buffered by 500 km to represent an area hypothesized as accessible to the assemblage (i.e., within a reasonable distance of its present distributional area). These polygons were used as areas from which to generate random localities in Warren and others ENMTools (http://enmtools.com/).

In ENMTools, for each pairwise comparison, we drew 100 random points from the background 100 times, and generated niche models based on those points in Maxent; outputs were converted to binary format using the minimum probability value (see above) as a threshold; and $I$ and $D$ values determined (Figure 4). In comparing niche models, we asked the statistical question, “are they similar?” As such, our null hypothesis was that the two species were similar, which we
RESULTS

The Centers for Disease Control and Prevention (CDC) plague dataset included 3,777 plague-positive records for wild mammals. Different species ranged from 1 to 2,516 in numbers of plague-positive records, representing a total of ~72 species (actual numbers of species are debatable as voucher specimens are not available for examination, and taxonomic changes have made several records identified to genus only). Species-wise, most records corresponded to rodents (50 of 72 species), but most records were from carnivores (2,840 of 3,777). Among rodents, most were sciurids (28 species, 311 records) or cricetids (17 species, 114 records), in addition to castorids (1 species, 1 record), heteromyids (1 species, 1 record), and murids (3 species, 8 records); carnivores included canids (6 species, 2,545), felids (3 species, 113 records), mephitids (2 species, 11 records), mustelids (3 species, 111 records), procyonids (2 species, 17 records), and ursids (1 species, 43 records). We note that 2 records each of 2 artiodactyls (wild boars, Sus scrofa, and mule deer, Odocoileus hemionus), and 3 species (14 records) of lagomorphs (all leporids) were also included.

Modern North American plague occurrences are exclusively in the western half of the continent, mostly in the Rocky Mountains (Figure 2) and east only to the western portion of the Great Plains. Our assemblages of plague-infected hosts appear—visually, at least—to have different distributional patterns from each other: plague-infected C. latrans and Taxidea taxus occur over the broadest areas, whereas other taxa have more restricted areas of occurrence or possible incidental infections. Plague detections are unevenly distributed across the region, with concentrations in California, along the eastern edge of the Rocky Mountains, and on the Colorado Plateau.

We noted considerable variation in numbers of known plague-positive occurrence points within states (see Supporting Information), with Nevada (748) and New Mexico (670) showing highest levels and Oklahoma (7) and Nebraska (13) the lowest. Only two plague occurrence points came from Canada. As previously noted, numbers of plague-infected occurrence points were highly skewed between host assemblages, such that coyotes represented 81% of the data.

Niche models based on plague-infected hosts differed markedly from those based on overall host ranges (Figure 2): plague infections do not cover the entirety of any of the host group ranges geographically. Ecologically, in all host taxa, plague-infections do not cover the entirety of any of the host groupedly from those based on overall host ranges (Figure 2): plague infections do not cover the entirety of any of the host groupedly from those based on overall host ranges (Figure 2). Moreover, of the 17 comparisons that rejected ecological similarity, most (64%) were equivocal, in that similarity was rejected in only one metric (Figure 3). Most exceptions to similarity were associated with plague-infected C. latrans occurrences, which tended to be more distinct ecologically from the environmental “background” of other plague-infected taxa (P < 0.05) unequivalently in 5 of 7 comparisons, but equivocal in 2 of 7; Figure 3). Alternatively, comparisons of plague-infected occurrences of the other taxa and the environmental background of

Although niche models from plague occurrences tended to differ from overall ranges, niche models based on plague occurrences in different host taxa tended to be similar to one another, as comparisons rarely rejected the null hypothesis of similarity. Indeed, 70% of such comparisons between plague-infected host taxa could not reject ecological similarity in at least one of the overlap metrics (P > 0.05). Moreover, of the 17 comparisons that rejected ecological similarity, most (64%) were equivocal, in that similarity was rejected in only one metric (Figure 3). Most exceptions to similarity were associated with plague-infected C. latrans occurrences, which tended to be more distinct ecologically from the environmental “background” of other plague-infected taxa (P < 0.05) unequivalently in 5 of 7 comparisons, but equivocal in 2 of 7; Figure 3). Alternatively, comparisons of plague-infected occurrences of the other taxa and the environmental background of
Figure 4. An example of the background similarity test in ENMTools. The map shows two niche-model-based predictions 1) plague-infected-ground squirrels and 2) and a random draw of points from the M of plague-infected *Taxidea taxus*, shown as the dark gray polygon. Black cells represent predictions only from the random draw, white cells represent predictions only from the plague-infected-ground squirrels, and gray cells represent overlap between the two. The histogram shows the entire range of Schoener’s *D* values from the 100 random-draw models; the observed Schoener’s *D* is shown as the black arrow. In this example, similarity between the two plague-infected host taxa is greater than that expected given the environments surrounding the distributional area of *Taxidea*.

plague-infected *C. latrans* failed to reject similarity in either metric in all cases (*P* > 0.05). Principal components analyses (PCA) of the environmental data suggest that the plague-infected *C. latrans* niche model represents a broader suite of environmental conditions and that other taxa generally are distributed within a subset of these conditions (see Supporting Information). The remaining comparisons that rejected similarity were plague-infected ground squirrel points compared with the *Cynomys* and *Ursus americanus* plague-infected backgrounds, and plague-infected tree squirrel points compared with the plague-infected *Cynomys* background.

**DISCUSSION**

In general, our range-wide analyses showed that plague infections occur in ecologically non-random subsets of the distributions of each host assemblage, but that plague infections in different host assemblages occur under similar ecological circumstances. Hence, we found ample support for PNH expectations, but none of the predictions of HNH were fulfilled. In particular, the significant non-similarity between all plague niche models and associated host niche models is suggested by the PNH, because host species get infected with plague only where they overlap the plague “niche.” The HNH predictions of non-similarity between plague-infected niche models were not supported, given that most of these comparisons were similar, which is consistent with PNH predictions.

An additional possibility for consideration is the idea that a single host may drive the geographic distribution of plague, such that its ecological characteristics would be similar to those of plague infections in other taxa. However, of the 64 pairwise comparisons of overall host niche models to plague-infected “backgrounds,” 63 rejected similarity. These results suggest strongly that host distributions differ from plague-infected distributions, and that these taxa do not drive plague distributions at geographic scales. Furthermore, the eight comparisons of the potential enzootic host *O. leucogaster* to plague-infected “backgrounds” rejected similarity in all cases, suggesting that its distribution does not reflect plague occurrence in other taxa.

The sole occurrence of ecological similarity between a host assemblage and plague-infected individuals of another host assemblage was between overall occurrences of *Cynomys* and plague-infected *C. latrans*. This similarity presents an intriguing scenario: *Cynomys* exhibits epizootic cycles with extremely high mortality, and coyotes are known to feed on carcasses and move among *Cynomys* colonies, such that they have been suggested as a dispersal vector for passing plague between prairie-dog towns. Inferring that *C. latrans* movements may be driving *Cynomys* outbreaks is intriguing and tempting, but premature, because the bacterium may be maintained by a variety of factors. Possibly, seropositive *C. latrans* individuals dispersed into novel areas not representative of plague within our dataset, which would explain the general non-similarity of other plague-infected backgrounds to plague-infected *C. latrans* (Figure 3). More conservatively, if we interpret *C. latrans*-related plague incidence as the ecological extent of plague distribution, as suggested by PCA, then *Cynomys* distribution is simply coincident with plague distribution.

We found widespread similarity among plague-infected host distributions, which suggests strongly that plague occurs consistently within a distinct ecological subset of North America. This study is developed at a scale distinct from that of the individual enzootic and epizootic cycles, which are manifested at local scales. Community interactions and host and vector population fluctuations no doubt influence host organisms, and therefore likely mediate plague cycles at local scales. Studies to date linking plague occurrences to environmental correlates focused in large part on host dynamics during epizootics, limiting the possibility of characterization of areas of enzootic plague transmission.

Plague-infected *C. latrans* points cover the broadest manifestation of the distribution of plague, both geographically and ecologically. The large size and transient behavior of individuals of this species lends well to easy detection, and their ubiquitous presence across many habitats provides an easy point of reference. We do not interpret these results as suggesting *C. latrans* as a driver for plague, but they do suggest that emerging local outbreaks may be signaled by increased rates of seropositivity for *C. latrans*. In a recent study, niche models of ground squirrel taxa in California were associated ecologically with plague-positive *C. latrans* occurrence sites, providing another example of the general association of plague presence and coyotes.
So what factors drive the “plague niche”? The broad extent of the Third Pandemic suggests that the pathogen itself responds little to the outside environment, and can occur where appropriate hosts and vectors are present, although factors such as temperature and humidity could affect the survival of flea vectors and the development of Y. pestis within the flea. The analyses presented herein argue against significant effects of particular host lineages. Rather, we focus on the distribution and ecology of plague vector species (Siphonaptera) as likely key determinants of plague distributional characteristics. The variety of flea species able to transmit plague (albeit with different efficiencies) is impressive (13, 53); a previous study modeled distributions of plague-associated flea species, but did not test for associations with plague occurrence patterns. Clearly, the task of incorporating flea ecology and distributions represents a significant challenge as a next step for understanding the geography and ecology of plague transmission patterns.

Received January 21, 2010. Accepted for publication May 27, 2010.

Acknowledgments: Sarah Anderson and Charles Hoxmeier provided invaluable assistance in capturing plague-host occurrence data. Financial support: ATP, SPM, and CE were supported by NSF EpSCOR funds.

Note: Supplementary materials are available at www.ajtmh.org.

Authors’ addresses: Sean P. Maher and A. Townsend Peterson, Department of Ecology and Evolutionary Biology and Biodiversity Institute, University of Kansas, Lawrence, KS, E-mails: maher@ku.edu and town@ku.edu. Christine Ellis, Genesis Laboratories, Inc., Wellington, CO, E-mail: dmbnnyl1@yahoo.com. Kenneth L. Gage and Russell E. Enscore, Centers for Disease Control and Prevention, National Center for Infectious Diseases, Division of Vector Borne Infectious Diseases, Fort Collins, CO, E-mails: klg@cdc.gov and rce0@cdc.gov.

REFERENCES


