

RESEARCH ARTICLE

# Can You Judge a Disease Host by the Company It Keeps? Predicting Disease Hosts and Their Relative Importance: A Case Study for Leishmaniasis

Christopher R. Stephens<sup>1,2\*</sup>, Constantino González-Salazar<sup>2\*</sup>, Víctor Sánchez-Cordero<sup>3</sup>, Ingeborg Becker<sup>4</sup>, Eduardo Rebollar-Tellez<sup>5</sup>, Ángel Rodríguez-Moreno<sup>3</sup>, Miriam Berzunza-Cruz<sup>4</sup>, Cristina Domingo Balcells<sup>4</sup>, Gabriel Gutiérrez-Granados<sup>6</sup>, Mircea Hidalgo-Mihart<sup>7</sup>, Carlos N. Ibarra-Cerdeña<sup>8</sup>, Martha Pilar Ibarra López<sup>9</sup>, Luis Ignacio Iñiguez Dávalos<sup>9</sup>, María Magdalena Ramírez Martínez<sup>10</sup>



**OPEN ACCESS**

**Citation:** Stephens CR, González-Salazar C, Sánchez-Cordero V, Becker I, Rebollar-Tellez E, Rodríguez-Moreno Á, et al. (2016) Can You Judge a Disease Host by the Company It Keeps? Predicting Disease Hosts and Their Relative Importance: A Case Study for Leishmaniasis. *PLoS Negl Trop Dis* 10(10): e0005004. doi:10.1371/journal.pntd.0005004

**Editor:** Andrew Fenton, University of Liverpool, UNITED KINGDOM

**Received:** April 11, 2016

**Accepted:** August 25, 2016

**Published:** October 7, 2016

**Copyright:** © 2016 Stephens et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the Universidad Nacional Autónoma de México (PAPIIT IN215212, IN202711, IN113414). Funding was also obtained from Consejo Nacional de Ciencia y Tecnología (CONACyT) 132811 to VSC and from Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México to CRS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**1** Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico, **2** C3 - Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico, **3** Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico, **4** Unidad de Investigación en Medicina Experimental, Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de Mexico, Mexico, **5** Laboratorio de Entomología Médica, Departamento de Zoología de Invertebrados, Facultad de Ciencias Biológicas, Universidad Autónoma de Nuevo León, Nuevo León, México, **6** Facultad de Estudios Superiores-Zaragoza, Universidad Nacional Autónoma de México, México D.F., México, **7** División Académica de Ciencias Biológicas. Universidad Juárez Autónoma de Tabasco, Villahermosa, Tabasco, México, **8** Departamento de Ecología Humana, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav) Unidad Mérida, Mérida, Yucatán, México, **9** Laboratorio de Zoología, Centro Universitario de la Costa Sur, Universidad de Guadalajara, Jalisco, México, **10** Departamento de Salud y Ecología Humana, Centro Universitario de la Costa Sur, Universidad de Guadalajara, Jalisco, México

\* [stephens@nucleares.unam.mx](mailto:stephens@nucleares.unam.mx) (CRS); [cgsalazar7@gmail.com](mailto:cgsalazar7@gmail.com) (CGS)

## Abstract

Zoonoses are an important class of infectious diseases. An important element determining the impact of a zoonosis on domestic animal and human health is host range. Although for particular zoonoses some host species have been identified, until recently there have been no methods to predict those species most likely to be hosts or their relative importance. Complex inference networks infer potential biotic interactions between species using their degree of geographic co-occurrence, and have been posited as a potential tool for predicting disease hosts. Here we present the results of an interdisciplinary, empirical study to validate a model based on such networks for predicting hosts of *Leishmania (L.) mexicana* in Mexico. Using systematic sampling to validate the model predictions we identified 22 new species of host (34% of all species collected) with the probability to be a host strongly dependent on the probability of co-occurrence of vector and host. The results confirm that *Leishmania (L.) mexicana* is a generalist parasite but with a much wider host range than was previously thought. These results substantially change the geographic risk profile for Leishmaniasis and provide insights for the design of more efficient surveillance measures and a better understanding of potential dispersal scenarios.

**Competing Interests:** The authors have declared that no competing interests exist.

## Author Summary

Emerging and neglected zoonoses are an important global threat to public health. Host range, in particular, is a crucial factor in determining disease risk and the potential for adequate interventions. Here we show that *Leishmania* has a very wide host range and that Complex Inference Networks can be used to infer ecological relationships in the context of zoonoses, identifying both the potential hosts and their relative importance. These results substantially change the risk profile and potential control measures that can be used to combat the disease, allowing for the design of more efficient surveillance measures and a better understanding of potential dispersal scenarios.

## Introduction

Zoonoses are an important class of neglected [1,2] or emerging infectious diseases [3–6], accounting for more than 60% of human infectious diseases. Wildlife species that are hosts for pathogens play a fundamental role in zoonoses, threatening domestic animal and human health and global biodiversity. Although for particular zoonoses some hosts have been identified [7–14], there have been few systematic empirical studies carried out to identify the host range and the relative importance of the different hosts within that range for a given zoonosis.

Additionally, most work on disease ecology over the last 20 years has focused on single-host, single-agent systems. Recently however, there has been increasing interest in the more complex case of multi-host systems [15–20], with the realization that many zoonoses have potentially ample host ranges.

The relative importance of a species as a disease host will be highly multi-factorial, with risk factors covering many different scales, from the micro to the macro. However, there are two particularly important elements that come into play: host competence (the ability to transmit the parasite to a new host or vector) and the frequency of contact between host and pathogen or, in the case of vector borne diseases, host and vector [21].

Although it is intuitively clear that the “relative importance” of a host will depend on both its competence and its frequency of contact with the vector, it is a somewhat ambiguous concept in that it depends, for instance, on whether we are talking about human transmission or of maintaining an enzootic transmission cycle. We will take here relative importance to be associated with the probability of infection of an individual of a potential host species. Abstractly, this is a highly multi-factorial function  $P(C | X_1, X_2, \dots, X_N)$ , where, for instance,  $X_1$  could represent host competence and  $X_2$  frequency of host-vector contact. Although a host may be highly competent, if it only has infrequent contact with any disease vector, then the frequency of infected individuals will be low. Conversely, a host and a vector may have frequent contact, but the host may have low competence. All else being equal, the most important hosts will be those that have frequent contact with the disease vectors and are competent. Unfortunately, gathering information about these two aspects, especially for emerging or neglected diseases, is difficult and resource intensive [7,8,22,23]. Furthermore, how these fundamental aspects interact in multi-host systems is quite distinct from their single-host analogues. For instance, the fact that host-pathogen competencies may differ greatly among the hosts can potentially lead to a dilution effect [24–26].

Another important differentiating factor is that, multiple-host systems provide for much richer and complex scenarios for the dispersion of a disease from one geographic region to another [27]. As the characteristics of the host range play a crucial role in the emergence risk of a novel human pathogen and of the optimal interventions for combating the zoonosis [16] the

importance of predicting and identifying potential disease hosts has been widely recognized [28–30]. To do so by exhaustive, systematic search through all possible hosts would be prohibitively resource intensive. At the same time, good data often only exist for a few (presumed) focal species. As it is unknown what part of the host range has already been discovered, the undiscovered part constitutes a type of ‘epidemiological dark matter’ [31].

An early attempt at systematising the search for potential hosts [32], in the case of Ebola, considered a heuristic approach based on expert knowledge, which was used to then filter the list of potential candidates. As such, it is both subjective and subject to model bias. More recently, other methods have appeared: In [33] a small group of four suspected hosts was used as a starting point for including biotic effects indirectly by calculating the fundamental niche of these four mammal species and considering the geographical correspondence with the niche distributions of the vectors. This paper was more concerned with including information about a particular set of potential hosts into corresponding risk maps rather than identifying new hosts per se. In contrast, in [34], a classification model using a supervised learning technique was used to predict other potential rodent reservoirs based on the predictive value of a set of potentially distinguishing characteristics of already known ones. Note that this paper was concerned with the potential hosts of a large number of pathogens considered all together and therefore could not discriminate against potential hosts of one disease versus another. In this case only categorical and no spatial information was used. Moreover, as it is based on supervised learning it can be affected by bias in the data defining both the class and in the predictors. This is in evidence in that the most predictive factor found was the number of literature citations for a given species. In [35], the authors considered biotic factors as potential predictive variables for describing the geographic range of Ebola rather than trying to predict which mammals are the most likely hosts.

In contrast, in Stephens et al [36], a general framework was presented using Complex Inference Networks based on the degree of co-occurrence between different species, for inferring potential biotic interactions. The framework is also capable of including in other variable types, at distinct spatial resolutions, such as environmental layers normally associated with abiotic variables [37] allowing for a comparison of the relative importance of biotic versus abiotic factors.

The methodology differs from those of [32–35] by using as model inputs only purely spatial data, using point collection data to proxy spatial distributions of taxa and co-occurrences to infer potential biotic interactions. In particular, it uses no auxiliary information, such as expert knowledge, as in the case of [32]; fundamental niche distributions of taxa, as in the case of [33]; or specific categorical data associated with the relevant taxa, as in the case of [34].

Networks are an important tool in ecological studies [38–41]. However, their local structure—in the sense of two nodes and a link as the base element—represents an already known relation, such as in a food web [38], or in a contact network representing ticks, vertebrates and pathogens, as in [42]. In this case the local structure of the network, i.e., the individual nodes and links, only represent what is known. However, the global properties of the network can lead to new insights from an eco-systemic or community viewpoint and also to specific predictions. In contrast, Stephens et al. [36] use the local structure of networks to infer and discover previously unknown relations, such as the relation between vector and host. Although nodes are taxa, the local structure of the network is different to a traditional ecological network in that the links represent the degree of overlap between the distributions of the corresponding taxa with the idea that statistically significant degrees of co-occurrence can be an indicator of potential biotic interaction between the associated taxa, such as between a host and a vector.

In determining the host range of a zoonosis, an exhaustive empirical analysis of all potential hosts is prohibitively difficult, hence the importance of theoretical models, such as that presented in [36], for guiding observation and experiment. Although consistent with known

results, it is important to note that the theoretical predictions of [36] have not previously been tested experimentally. A theoretical model needs to be validated by experiment, as this is the only way to truly determine if the model works. In this vein, most ecological modelling of zoonoses remains untested, in that theoretical predictions are not validated using a suitable experimental validation framework. This paper presents the results of an interdisciplinary study carried out to experimentally test the predictions of [36] using, as a test, the case of Leishmaniasis in Mexico.

Leishmaniasis is a significant, yet neglected tropical disease, with 350 million people in 98 countries worldwide living at risk of developing one of the many forms of the disease [43]. It is caused by infection with one of several different species of protozoan parasites of the genus *Leishmania*, which maintain their life cycle through transmission between an insect (sandflies—genus *Lutzomyia*) and a mammalian host. In Mexico, the most epidemiologically important species is *Leishmania (L.) mexicana*, though the presence of other species has been confirmed. Eleven species of *Lutzomyia* are considered to have potential medical importance. Of these, three are known vectors of either cutaneous or visceral Leishmaniasis, while four others have been found infected with *L. (L.) mexicana* [44].

Besides the clinical and social importance of Leishmaniasis [45] and the acknowledgment of its zoonotic nature [46], the identification of wildlife hosts for these parasites is sparse and non-systematic. Prior to the present study, only 8 mammalian species had been identified as hosts in Mexico [47–49]. This potential lack of knowledge of parasite hosts greatly increases the difficulty of formulating theoretical approaches to explaining and predicting disease spread or for planning better and sustainable control measures [50].

## Methods

### Ethics statement

The collection of specimens was performed according to the guidelines of the American Society for the Use of Mammalogists of Wild Mammals in Research and under a collecting permit issued by the General Direction of Wildlife of Mexico (permission number SGPA/DGVS/04631/14). The infections in mice were carried out following the National Ethical Guidelines for laboratory animals NOM-062-ZOO-1999. The project was approved by the Institutional Ethics Committee of the Medical Faculty of the National Autonomous University of Mexico (UNAM) with the registration number FMED/CI/RGG/013/01/2008.

### Rationale for the modelling methodology

The general modelling methodology of [36], is based on the idea that biotic interactions can be inferred from the locations of taxa as a function of space and time. Although biotic and ecological interactions in general are very complex, it is reasonable to state that the spatio-temporal distributions of taxa, or other ecological variables, reflect all of the factors and their causal interactions that determine them. In [36], the degree of co-occurrence between taxa was taken as an observable measure with which potential interactions could be inferred. Although co-occurrence is not equal to biological interaction, a significantly non-random co-occurrence distribution is a *necessary* condition for a biotic interaction between taxa, and as such it can be used to formulate hypotheses that can be checked experimentally. However, it is clearly not a *sufficient* condition. In the spirit of niche modelling, a biotic variable that co-occurs with a target taxon can be understood as being a niche component in the same sense as any abiotic variable, such as temperature. In fact, one would generally expect a closer causal relation between biotic variables than with abiotic variables. For example, the distribution of prey species for a

predator, such as a carnivore, should clearly influence the latter's distribution more significantly than temperature or precipitation.

In the case of many zoonoses, the predominant interaction between vector and host is due to the former feeding on the latter. This obviously requires a coincidence in space and time. Species that offer blood meals can maintain the presence of vector populations independently of the capacity to harbour a given pathogen. In other words, the interaction between host and vector is a necessary but not sufficient condition for the transmission of the pathogen. The total number of encounters between vector and potential host depends on many factors, including the abundance of both species. However, a key factor is the geographical overlap between them, as the greater the overlap the greater the probability of an encounter. Thus, for two host species, identical in all respects except their relative geographical overlap with the vectors, the host species with the larger overlap will be epidemiologically more important. Thus, vector-mammal geographical overlap is a necessary but not sufficient condition for both a feeding interaction and a pathogen transmission interaction.

Of course, there may be geographical overlap between species due to other reasons than a direct biotic interaction. Even if species distributions were random there would be overlap. It is therefore necessary to measure overlap relative to a null hypothesis, such as that associated with random distributions. Additionally, it may occur that there is a non-random overlap due to the existence of one or more confounding factors; for example, an abiotic variable, such as temperature. This can only be quantified by controlling for the presence of such a factor. As it is obviously infeasible to control systematically for every potential factor, a logic must be presented for considering a particular candidate. In summary: although geographical overlap is not a sufficient condition for biological interaction, it is necessary, and as such can be used to construct models that can then be checked explicitly by experiment to see to what extent it is predictive.

The explicit example considered in [36] was the identification of potential hosts for Leishmaniasis by studying co-occurrences between the vector species and the potential host species. A Complex Inference Network summarising the co-occurrence distributions was deduced that showed the most important potential mammal hosts for each sandfly species. Although the full network contains a great deal of structure and information, in terms of experimental validation each network observable requires an experimental protocol to be able to measure it. In particular, to work at the species level for the vectors, and associate and confirm hosts for a given vector species, would require collecting sandflies and genotyping their blood meals, as well as collecting potential host species and confirming the presence of the pathogen. In the present experimental study, we restricted attention to only potential host species and tested them for the presence of the pathogen considering the vector at the genus level only.

## Inferring vector-hosts interactions

The explicit model for predicting potential hosts was created using a database of point collections for one Class, Mammalia, and one genus, *Lutzomyia*, of the class Insecta. The mammal data set contains 37,297 unique point collections from geo-referenced localities for 419 terrestrial mammals occurring in Mexico—the full potential host range (GBIF; [www.gbif.org](http://www.gbif.org), and CONABIO; [www.conabio.gob.mx](http://www.conabio.gob.mx)). For *Lutzomyia*, there were 270 collections points taken from published literature and from national collections: Instituto de Diagnóstico y Referencia Epidemiológica (InDRE, Mexico City), the Colección Entomológica Regional, Universidad Autónoma de Yucatán (UADY, Mérida) and the Laboratorio de Medicina Tropical at the Universidad Nacional Autónoma de México (UNAM, Mexico City).

First, we divide up a geographic region of interest into spatial cells,  $x_{cs}$ —in the present case Mexico—here we used a uniform grid of 3,337 rectangular cells of size 25km x 25km. The choice

of an appropriate cell size is known in geography as the “modifiable areal unit problem”. In terms of forming a spatial grid, there are at least two important considerations: The sizes of the statistical samples of the variables and their degree of correlation. Too fine a grid and there will be no co-occurrences, too rough and there will be little to no discrimination. It was checked explicitly in [36] that the relative ranking of mammals by the model was quite insensitive to the cell size over the range 5km to 100km. See also [51]. One then counts co-occurrences in each spatial cell between different taxa, or other variables. In the present case, the co-occurrences are between the presence of *Lutzomyia*,  $B_i$ , and the presence of each distinct mammal species,  $I_k$ .

We take the taxon distribution,  $B_i$  (*Lutzomyia*), and a subset of potential niche variables. We are interested in the probability  $P(B_i | I') = N_{B_i \text{ AND } I'} / N_{I'}$ , where  $N_{B_i \text{ AND } I'}$  is the number of spatial cells where there is a co-occurrence of the taxon  $B_i$  and the niche variables  $I'$ , which we take here to be biotic variables, and  $N_{I'}$  is the number of cells where the niche variables take their stated values. The niche profile  $I'(x_\alpha)$  associated with a spatial cell  $x_\alpha$  then determines the probability of the distribution variable,  $B_i(x_\alpha)$ , in that cell, and one now has a predictive model. The problem of calculating  $P(B_i | I')$  directly is that both  $N_{B_i \text{ AND } I'}$  and  $N_{I'}$  are likely to be zero when the number of taxa or niche variables considered simultaneously is large, as there will tend to be no co-occurrences of so many variables. This can be ameliorated by considering a reduced number of both class and feature variables. For instance,  $P(B_i | I_k)$  is determined by the number of co-occurrences of the taxon  $B_i$  and the particular niche variable  $I_k$  and, in principle, allows us to find the most important statistical associations between the niche variables and the taxa distributions. However,  $P(B_i | I_k)$  being a probability does not account for sample size. For example, if  $P(B_i | I_k) = 1$ , this may be as a result of there being a coincidence of  $B_i$  and  $I_k$  in one spatial cell or 1,000. Obviously, the latter is more statistically significant. To remedy this we consider the following test statistic

$$\epsilon(B_i | I_k) = \frac{N_{I_k} (P(B_i | I_k) - P(B_i))}{(N_{I_k} P(B_i) (1 - P(B_i)))^{1/2}} \tag{1}$$

a binomial test which measures the statistical dependence of  $B_i$  on  $I_k$  relative to the null hypothesis that the distribution of  $B_i$  is independent of  $I_k$  and randomly distributed over the grid, i.e.,  $P(B_i) = N_{B_i} / N$ , where  $N_{B_i}$  is the number of grid cells with point collections of species  $B_i$  and  $N$  is the total number of cells in the grid. The sampling distribution of the null hypothesis is a binomial distribution where, in this case, every cell is given a probability  $P(B_i)$  of having a point collection of  $B_i$ . The numerator of eq (1) then, is the difference between the actual number of co-occurrences of  $B_i$  and  $I_k$  relative to the expected number if the distribution of point collections were obtained from a binomial with sampling probability  $P(B_i)$ . As we are talking about a stochastic sampling the numerator must be measured in appropriate “units”. As the underlying null hypothesis is that of a binomial distribution, it is natural to measure the numerator in standard deviations of this distribution and that forms the denominator of eq (1). In general, the null hypothesis will always be associated with a binomial distribution as in each cell we are carrying out a Bernoulli trial (“coin flip”). However, the sampling probability can certainly change.

The quantitative values of  $\epsilon(B_i | I_k)$  can be interpreted in the standard sense of hypothesis testing by considering the associated p-value as the probability that  $|\epsilon(B_i | I_k)|$  is at least as large as the observed one and then comparing this p-value with a required significance level. In the case where  $N_{I_k} > 5-10$  then a normal approximation for the binomial distribution should be adequate, in which case  $\epsilon(B_i | I_k) = 1.96$  would represent the standard 95% confidence interval. When a normal approximation is not accurate then other approximations to the cumulative probability distribution of the binomial must be used.

As  $\epsilon$  increases monotonically with the frequency of co-occurrence, we interpret a statistically significant positive correlation as inferring a potential biotic interaction. Here, between sandflies and the corresponding mammal, which in this ecological setting one would naturally interpret as the mammal being a blood source for the sandfly, and therefore a potential host. The higher the value of  $(P(B_i | I_k) - P(B_i))$  the greater the degree of spatial overlap between the species distributions and therefore the greater the risk posed by the corresponding mammal. Negative values of  $\epsilon$  correspond to spatial overlaps that are less than one would expect from the null hypothesis.

The 419 mammal species were ranked according to  $\epsilon$ . The resultant list serves as a predictive risk model, with the hypothesis that the highest ranked mammals correspond to the most important hosts, where, in the absence of other information, we assume that host competence is equal and importance is associated with the degree of spatial overlap between sandflies and mammal. All else being equal more overlap means more vector-host encounters. It should be noted that the method is not determining the physiological capacity of a mammal species to be a host but, rather, its potential epidemiological importance given that presence of mammal hosts is a necessary condition for the presence of the pathogen.

### Constructing risk maps from inference networks

A corresponding biotic geographic risk model can be computed by calculating the probabilities  $P(B_i | I')$ , or proxies thereof, for each spatial cell. When  $I'$  is of high dimension, this can be done using different classification models, such as neural networks, discriminant analysis, etc. A particularly transparent, simple and effective approximation is the Naive Bayes approximation:

$$P(B_i | I) = \frac{P(I|B_i)P(B_i)}{P(I)} = \frac{\prod_{k=1}^N P(I_k|B_i)P(B_i)}{P(I)} \tag{2}$$

where, in the first equality, Bayes rule has been used, and in the second it has been assumed that the niche variables  $I_k$  are independent. The product here is over the  $N$  niche variables under consideration as conditioning factors for  $B_i$ . In the case of the relationship between *Lutzomyias* and mammals,  $N$  represents the number of mammal species. A score function that can be used as a proxy for  $P(B_i | I')$  is

$$S(B_i | I') = \sum_{k=1}^N S(B_i | I_k) = \sum_{k=1}^N \ln \left( \frac{P(I_k | B_i)}{P(I_k | \bar{B}_i)} \right) \tag{3}$$

where  $\bar{B}_i$  is the complement of the set  $B_i$ . For example, if  $B_i$  is the set of cells with presence of taxon  $B_i$  then represents the set of cells without presence.  $S(B_i | I')$  is a measure of the probability to find the distribution variable  $B_i$  when the niche profile is  $I'$ . It can be applied to a spatial cell  $x_\alpha$  by determining the niche profile of the cell,  $I'(x_\alpha)$ . As an example, for two biotic niche variables,  $B_2$  and  $B_3$ , that take values 1 (corresponding to the fact that there is a point collection associated with that cell) and 0 (there is no point collection associated with the cell), the four possible biotic niche profiles of any cell are  $(B_2, B_3) = (0,0), (0,1), (1,0)$  and  $(1,1)$ . The score contributions of each biotic variable are  $S(B_i | B_2)$  and  $S(B_i | B_3)$ , calculated using the above formula. Hence,  $S(B_i | I') = S(B_i | B_2, B_3) = S(B_i | B_2) + S(B_i | B_3)$ . Thus, for any given spatial cell  $x_\alpha$  one can assign a niche profile, i.e. values of  $B_2$  and  $B_3$ , from whence it is possible to assign a corresponding score. If there is no statistical association between  $B_i$  and  $B_2$  or  $B_3$  then the corresponding score contributions are zero.

An overall zero score signifies that the probability to find  $B_i$  is the same as would be found if  $B_i$  were distributed randomly. If the score is positive then there is a higher than random

probability to find  $B_i$  present and on the contrary if the score is negative. As each niche factor is treated separately in  $\epsilon(B_i | I_k)$  or  $S(B_i | I_k)$  we can thus evaluate the relative contribution of any given niche factor and compare it to the contribution of any other. By determining the set of presence/no presence attributes in a spatial cell, [eq \(3\)](#) can be applied to each cell thereby determining the relative risk of that cell for the presence of *Lutzomyias*. As taking the ranked list as a predictive model involves several important assumptions; it is essential to test the model with experimental data. Obtaining the relevant data required sampling the spatial grid by collecting mammals and sandflies from different geographic points and then testing them for the presence of the parasite.

### Sampling design for model testing

The sampling sites were selected as follows: a 25 x 25 km grid (as discussed above) was superimposed on a map of Mexico and this was used as template to determine the sampling sites. The sampling was stratified according to altitude so that only grid squares at < 2000 masl were used. Also, we excluded any grid square with >50% of water or urban cover. Sometimes the 25 x 25 km grid square selected was sub-divided in order to have more than one sampling site per square. Once the grid was established we selected 52 localities at random in 10 Mexican states and covering many different eco-regions and associated with a large selection of vegetation types. A random sample of spatial cells was necessary in order to fairly validate the model. If we had targeted the sampling to only those cells predicted to be highest risk then we would not be able to discriminate between high and low risk in the validation. Of course, once the model is validated it can be used with confidence in the future to preferentially identify those regions of highest risk where, for example, surveillance efforts should be targeted. The sampling was carried out by four field groups who, over a period of two years, collected 922 taxonomically identified specimens of 70 distinct species (for more details of animal sampling see [S1 Text](#)).

### Laboratory methodology

Tissue samples were taken to the Laboratorio de Inmunoparasitología of the Unidad de Investigación en Medicina Experimental of the National Autonomous University of Mexico, UNAM, where PCR tests were carried out to identify the presence of the pathogen *Leishmania (L.) mexicana*.

**DNA purification.** DNA from animal tissues was purified from approximately 25 mg of starting material. DNA extractions were done with a DNeasy Blood and Tissue kit (QIAGEN, Germany), following the manufacturer's instructions. Genomic DNA was used for PCR-based amplification.

**Oligonucleotides.** To determine the presence of *Leishmania* we used oligonucleotides based on the *Leishmania* mini-circle kinetoplast DNA sequences that are conserved among species. The primer sequences used for amplification were 5'-CTRGGGGTTGGTGTA AAA TAG-3' (L.MC-1S) and 5'-TWTGAACGGGRTTCTG-3' (L.MC-1R) [52]. For the species *Leishmania (L.) mexicana* we took primer IR1, designed by [53], which corresponds to the 32 final nucleotides of the conserved sequences from the 3' region of the small subunit of the 18S ribosomal gene as our forward primer: IR1 (5'- GCT GTA GGT GAA CCT GCA GCA GCT GGA TCA TT-3'), the reverse primer was LM17 (5'-CCC CTC TCC TCC TCC CC-3') [54].

**Polymerase chain reaction amplification.** The PCR was performed using 50  $\mu$ l of the following reaction mixture: Taq PCR Master Mix (QIAGEN) (provides a premixed solution containing Taq DNA Polymerase, PCR buffer, MgCl<sub>2</sub> and dNTPs), 100 ng of the corresponding oligonucleotides, and DNA from tissues, we used 1  $\mu$ l of tissue extract corresponding to 100 ng of DNA. The amplification was performed in a Perkin Elmer 2720 thermocycler using different



conditions, which depended on the oligonucleotides used. For L.MC-1S/ L.MC-1R *Leishmania* genus we used the PCR amplification with *Leishmania* mini-circle kinetoplast DNA specific primers was performed with 30 cycles of denaturation (95°C for one minute), annealing (55°C for one minute), and polymerization (72°C for one minute). For IR1/LM17 *L. (L.) mexicana* we used 35 cycles of 1 min at 94°C, 1 min at 65°C and 1 min at 72°C. In all cases the cycles were preceded by another cycle at 94°C for 5 min and a final extension cycle of 72°C for 7 min.

An analysis of the sensitivity of the primers L.MC-1S/ L.MC-1R was made with DNA from cultured promastigotes of *L. (L.) mexicana* using: 10 ng, 1 ng, 100 pg, 10 pg, 1pg, 100 fg, 10 fg and 1 fg DNA. PCR products were analyzed using electrophoresis in 1.5% agarose gels in TAE 1 × at 80 V. Gels were stained with 0.5 µg/ml ethidium bromide and photographed under a UV light source.

The PCR amplification products from two positively identified animals to genus and species were sequenced. The sequences were then compared and aligned using the National Center for Biotechnology Information, U.S. National Library of Medicine, Basic Local Alignment Search Tool (BLAST).

## Results

In [Table 1](#), we show the list of collected species with the number of individuals that tested positive for the presence of the parasite *Leishmania (L.) mexicana* and the number that tested negative. Of the 70 mammal species collected, approximately 1/6 of all species present in Mexico, 24 (34%) had one or more samples that tested positive for the presence of the *Leishmania (L.) mexicana* parasite. Thirteen species of bats, and one of squirrel, were identified for the first time as *Leishmania* hosts in Mexico. Of the total number of collected individuals (N = 922), 62 tested positive, yielding an average infection rate across all species of 6.7%, although infection rates varied greatly, both temporally and spatially, exhibiting considerable heterogeneity, from 0% to 60%, across distinct collection sites and season of the year. In addition to the percentage that tested positive we also include the 95% confidence interval limits using the Wilson score interval [55,56] for that percentage relative to the null hypothesis that all mammal species had the same baseline infection rate of 6.7%. However, for those species where the number of positives is zero we also calculate the exact probability to obtain this result,  $(1-p)^N$ , where p is the baseline infection rate and N is the number of negative collections.

[Fig 1](#) shows the ranked values of  $\epsilon$ , the statistical measure of degree of co-occurrence used to infer potential biotic interactions (see [Eq \(1\)](#) of the [Methods](#) section) for all mammal species as determined from the complex network exhibited in [36] ([S1 Table](#)) and considering *Lutzomyia* as a genus. The horizontal axis represents the null hypothesis that sandflies are distributed randomly with respect to mammal distributions, in other words  $P(B_i | I_k) = P(B_i)$  and hence  $\epsilon = 0$ . To determine the extent to which collection biases can influence the overall distribution we randomly redistributed all collections over those spatial cells that had at least one collection. This has the effect of removing correlations between one species and another while at the same time preserving any bias associated with under sampling of certain geographical areas. This random re-assortment was repeated 50 times and average values for  $\epsilon$  determined for each species.

To test the predictive power of the risk model, and to better visualize the relationship between  $\epsilon$  and the percentage of species that tested positive, we group the list of mammal species ranked by  $\epsilon$  into deciles [57] each decile corresponding to 10% of the list, and compute the average value of  $\epsilon$  for each decile. The result can be seen in [Fig 2](#). The relative proportion of positives, P, is a strongly increasing function of  $\epsilon$ . Note that this regression is only to demonstrate the predictive power of the underlying classification model based on  $\epsilon$ , i.e., the statistical

**Table 1. List of collected species, ranked by  $\epsilon$  (see Eq 1 of Methods), with the number of individuals that tested positive for the presence of the parasite *Leishmania (L.) mexicana* and the number that tested negative.**

Species	epsilon	Negative	Positive	Total	Prevalence (%)	95% confidence interval	5% confidence interval	Probability true negative
<i>Carollia sowelli</i>	8.83	42	2	44	4.55	18.07	2.28	0.00%
<i>Heteromys gaumeri</i> *	8.80	5	0	5	0.00	46.62	0.59	29.30%
<i>Peromyscus mexicanus</i>	8.79	115	6	121	4.96	12.62	3.45	0.00%
<i>Heteromys desmarestianus</i> *	8.72	30	0	30	0.00	50.52	0.50	87.51%
<i>Molossus rufus</i>	8.63	1	0	1	0.00	82.00	0.11	6.70%
<i>Glossophaga soricina</i>	8.57	9	7	16	43.75	28.92	1.25	0.00%
<i>Carollia perspicillata</i>	8.50	8	0	8	0.00	38.49	0.82	42.58%
<i>Pteronotus parnellii</i>	8.16	4	0	4	0.00	55.41	0.41	24.22%
<i>Desmodus rotundus</i>	8.15	13	1	14	7.14	30.91	1.14	0.00%
<i>Sturnira lilium</i>	8.03	58	6	64	9.38	15.59	2.72	0.00%
<i>Dermanura phaeotis</i>	8.01	35	1	36	2.78	19.69	2.06	0.00%
<i>Oryzomys couesi</i>	7.73	2	0	2	0.00	70.13	0.22	12.95%
<i>Ototylomys phyllotis</i> *	7.56	9	1	10	10.00	36.55	0.89	0.00%
<i>Sigmodon hispidus</i> *	7.28	36	4	40	10.00	18.81	2.18	0.00%
<i>Peromyscus yucatanicus</i> *	7.25	3	0	3	0.00	61.71	0.32	18.78%
<i>Didelphis virginiana</i>	7.12	3	0	3	0.00	61.71	0.32	18.78%
<i>Didelphis marsupialis</i>	6.44	11	0	11	0.00	28.92	1.25	53.37%
<i>Philander opossum</i>	6.25	6	1	7	14.29	43.42	0.67	0.00%
<i>Centurio senex</i>	6.01	1	0	1	0.00	82.00	0.11	6.70%
<i>Artibeus jamaicensis</i>	5.98	81	5	86	5.81	14.04	3.06	0.00%
<i>Artibeus lituratus</i>	5.84	36	3	39	7.69	19.02	2.15	0.00%
<i>Myotis keaysi</i>	5.61	2	0	2	0.00	70.13	0.22	12.95%
<i>Chiroderma villosum</i>	5.56	5	0	5	0.00	50.52	0.50	29.30%
<i>Saccopteryx bilineata</i>	5.30	1	0	1	0.00	82.00	0.11	6.70%
<i>Sciurus aureogaster</i>	5.23	71	8	79	10.13	40.75	0.74	42.58%
<i>Baiomys musculus</i>	5.21	2	0	2	0.00	70.13	0.22	12.95%
<i>Artibeus watsoni</i>	5.13	2	0	2	0.00	70.13	0.22	12.95%
<i>Choeroniscus godmani</i>	5.05	10	3	13	23.08	32.07	1.08	0.00%
<i>Pteronotus personatus</i>	5.03	3	1	4	25.00	55.41	0.41	0.00%
<i>Reithrodontomys mexicanus</i>	4.91	1	0	1	0.00	82.00	0.11	6.70%
<i>Oryzomys rostratus</i>	4.87	22	1	23	4.35	24.21	1.59	0.00%
<i>Micronycteris microtis</i>	4.23	1	0	1	0.00	82.00	0.11	6.70%
<i>Oligoryzomys fulvescens</i>	4.20	6	0	6	0.00	40.75	0.74	34.04%
<i>Peromyscus leucopus</i>	3.80	22	4	26	15.38	22.84	1.71	0.00%
<i>Sturnira ludovici</i>	3.79	24	1	25	4.00	23.26	1.67	0.00%
<i>Vampyrodes caraccioli</i>	3.69	1	0	1	0.00	82.00	0.11	6.70%
<i>Liomys pictus</i>	3.61	47	1	48	2.08	17.43	2.38	0.00%
<i>Glossophaga commissarisi</i>	3.49	2	6	8	75.00	40.75	0.74	0.00%
<i>Phyllostomus discolor</i>	3.48	0	1	1	100.00	82.00	0.11	0.00%
<i>Lonchorhina aurita</i>	3.48	1	0	1	0.00	82.00	0.11	6.70%
<i>Platyrrhinus helleri</i>	3.36	3	0	3	0.00	61.71	0.32	18.78%
<i>Uroderma bilobatum</i>	3.34	4	0	4	0.00	55.41	0.41	24.22%

(Continued)

Table 1. (Continued)

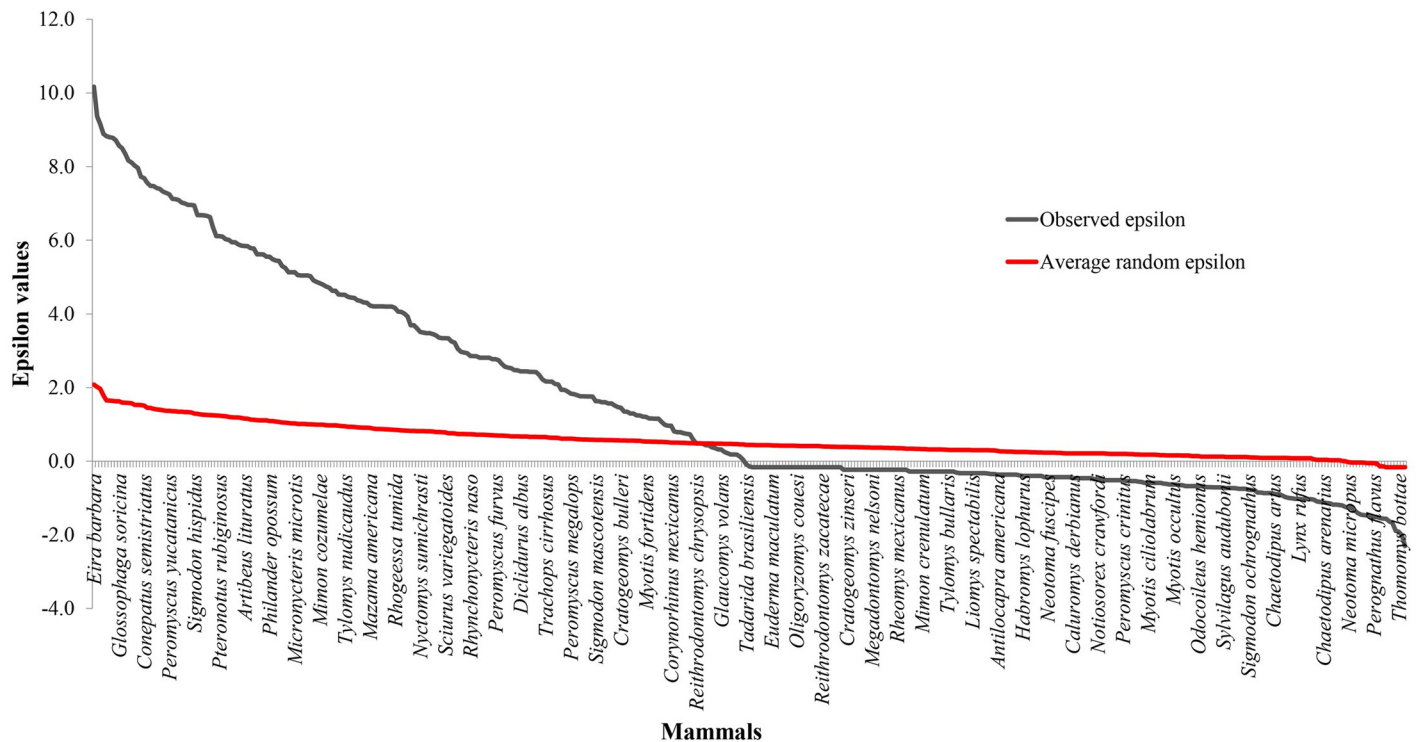
Species	epsilon	Negative	Positive	Total	Prevalence (%)	95% confidence interval	5% confidence interval	Probability true negative
<i>Urocyon cinereoargenteus</i>	2.97	1	0	1	0.00	82.00	0.11	6.70%
<i>Procyon lotor</i>	2.95	1	0	1	0.00	82.00	0.11	6.70%
<i>Myotis velifer</i>	2.58	4	0	4	0.00	55.41	0.41	24.22%
<i>Microtus mexicanus</i>	2.53	16	0	16	0.00	28.06	1.31	67.03%
<i>Myotis nigricans</i>	2.47	2	0	2	0.00	70.13	0.22	12.95%
<i>Leptonycteris curasoae</i>	2.43	1	1	2	50.00	70.13	0.22	0.00%
<i>Reithrodontomys fulvescens</i>	2.08	20	0	20	0.00	21.37	1.86	75.02%
<i>Neotoma Mexicana</i>	1.99	5	0	5	0.00	50.52	0.50	29.30%
<i>Eptesicus fuscus</i>	1.82	1	0	1	0.00	82.00	0.11	6.70%
<i>Peromyscus levipes</i>	1.34	1	0	1	0.00	82.00	0.11	6.70%
<i>Sorex saussurei</i>	1.29	3	0	3	0.00	61.71	0.32	18.78%
<i>Osgoodomys banderanus</i>	1.21	9	0	9	0.00	34.86	0.95	46.43%
<i>Liomys irroratus</i>	1.16	8	0	8	0.00	40.75	0.74	42.58%
<i>Myotis auriculus</i>	0.22	2	0	2	0.00	70.13	0.22	12.95%
<i>Tadarida brasiliensis</i>	-0.09	1	0	1	0.00	82.00	0.11	6.70%
<i>Peromyscus hylocetes</i>	-0.28	2	0	2	0.00	70.13	0.22	12.95%
<i>Antrozous pallidus</i>	-0.34	1	0	1	0.00	82.00	0.11	6.70%
<i>Peromyscus zarhynchus</i>	-0.46	2	0	2	0.00	70.13	0.22	12.95%
<i>Chaetodipus hispidus</i>	-0.71	4	0	4	0.00	55.41	0.41	24.22%
<i>Peromyscus pectoralis</i>	-0.73	2	0	2	0.00	70.13	0.22	12.95%
<i>Neotomodon alstoni</i>	-0.90	17	0	17	0.00	25.90	1.45	69.24%
<i>Baiomys taylori</i>	-1.16	10	3	13	23.08	32.07	1.08	0.00%
<i>Chaetodipus nelsoni</i>	-1.24	3	0	3	0.00	61.71	0.32	18.78%
<i>Neotoma micropus</i>	-1.27	16	0	16	0.00	28.92	1.25	67.03%
<i>Peromyscus maniculatus</i>	-1.37	58	2	60	3.33	15.97	2.64	0.00%
<i>Peromyscus eremicus</i>	-1.41	0	1	1	100.00	82.00	0.11	0.00%
<i>Perognathus flavus</i>	-1.52	1	0	1	0.00	82.00	0.11	6.70%
<i>Dipodomys merriami</i>	-2.01	1	0	1	0.00	82.00	0.11	6.70%

\*previously confirmed

doi:10.1371/journal.pntd.0005004.t001

significance of co-occurrence. Similarly, in Fig 3 we see the relative correlation between the average value of  $\epsilon$  and the percentage of individuals identified as positive. Once again, the relative proportion of positives is a strongly increasing function of  $\epsilon$ . Of course, P is a multi-factorial function,  $P(X_1, X_2, \dots, X_N)$  that depends on many factors, such as host competencies. Essentially, here we are considering a regression model for  $P(X_1, X_2, \dots, X_N)$  with respect to the variable  $X_1 = \epsilon$  and ignoring the rest as we do not have the relevant information to include them. Seen as a logistic regression at the species level, the associated relation is:  $\text{Logit } P = -1.415 + 0.186 * \epsilon$ , with a p value of 0.03 on the regression coefficient. This confirms the statistically significant relation between  $\epsilon$  as a statistical measure of geographical overlap and the probability to be a host of *Leishmania (L.) Mexicana*.

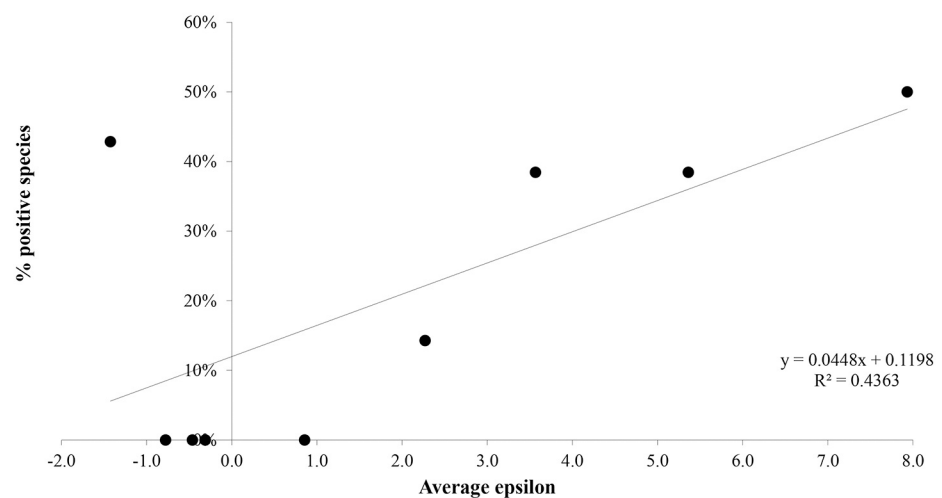
In Fig 4 we show a graph of prevalence at the species level versus  $\epsilon$  for the 24 positive species, while in Fig 5, we compare the disease risk maps determined by using only the eight



**Fig 1. Graph of ranked epsilon values for all mammal species compared to that of a random distribution.** The Average random epsilon line represents the distribution of overlaps found by randomly redistributing all collections over those spatial cells that had at least one collection.

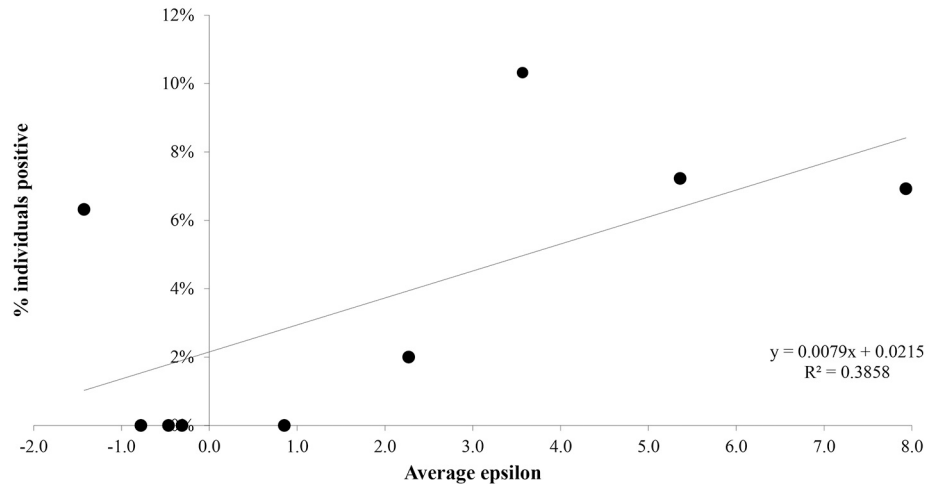
doi:10.1371/journal.pntd.0005004.g001

previously confirmed hosts of *Leishmania* versus one using the set resulting from our analysis, where by risk we mean probability of presence of the vector. A clear distinction can be seen between the areas of higher risk between the two models with the present model indicating a much higher degree of risk of presence of *Lutzomyia* in other than the southeast of Mexico.



**Fig 2. Percentage of species identified as positive for presence of *L. (L.) mexicana* in relation to mean epsilon value.**

doi:10.1371/journal.pntd.0005004.g002

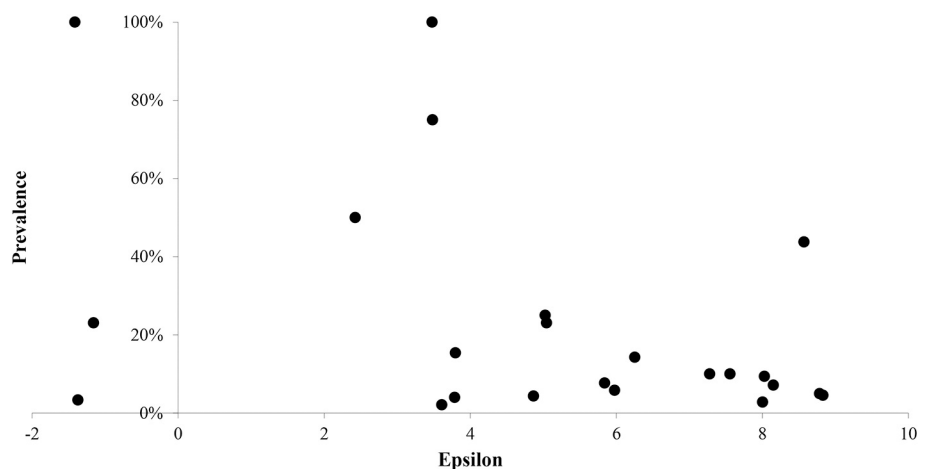


**Fig 3. Graph of percentage of individuals identified as positive for presence of *L. (L.) mexicana* versus mean epsilon value.**

doi:10.1371/journal.pntd.0005004.g003

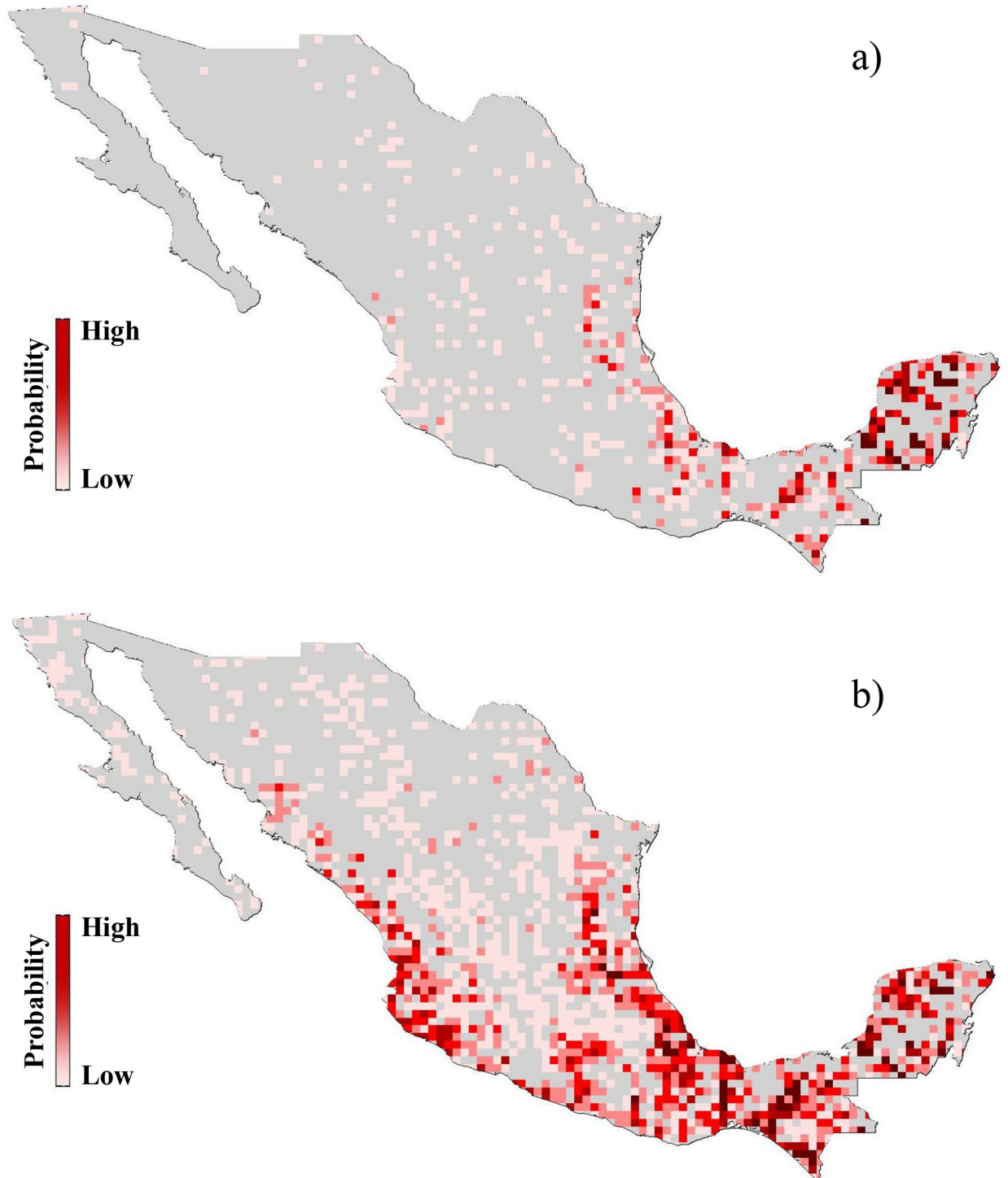
### Discussion

Host range is an important factor in the dynamics of a zoonosis, both as a variable that affects the overall risk of presence of the disease as well as in terms of determining optimal interventions. An essential factor in determining the importance of a host is its co-distribution with the disease vector. In this paper, we reported the results of an extensive, interdisciplinary, empirical investigation carried out to test the predictions of a model to predict the relative importance of mammal hosts for the pathogen *Leishmania (L.) mexicana* associated with the emerging disease Leishmaniasis in Mexico. Once again, we emphasise that *importance* here is an “all else being equal” notion, i.e., that the greater the overlap between species the greater the probability of a vector-host interaction and therefore a greater number of infected individuals and a greater probability of transmission. Of course, many other factors—host competence, host abundance etc. will influence the epidemiological importance of a given host.



**Fig 4. Graph of prevalence versus epsilon by species.**

doi:10.1371/journal.pntd.0005004.g004



**Fig 5. Risk maps of Leishmaniasis: a) determined by using only the 8 previously confirmed hosts, b) determined by using the 22 new confirmed hosts and previously confirmed hosts of *L. (L.) mexicana*.**

doi:10.1371/journal.pntd.0005004.g005

As the distribution of Fig 1 predicts the most important potential mammal hosts of *Leishmania (L.) mexicana*, the deviation of the set of real  $\epsilon$  values from the two random benchmarks considered, the  $\epsilon = 0$  line and the average random  $\epsilon$  curve, shows that the distribution of *Lutzomyia* is strongly, positively correlated with a large number of mammal species, and, further, that this fact is not explainable by collection biases. Furthermore, the strong asymmetry of the distribution, with 149 potential hosts associated with statistically significant positive correlations with sandflies, compared with only two species with statistically significant negative correlations, is consistent with a hypothesis that sandflies are generalists that are capable of feeding on, and potentially infecting, whatever potential mammal species are available, as has been suggested for several mosquitoes [58]. If sandflies were specialists, associated with a few focal species, one would expect to see high  $\epsilon$  values only for those particular species, due to the fact that the focal species are an important and necessary biotic element in the niche of the sandfly, while the other mammals would be incidental and therefore one would expect to see either random association or a positive association mediated by, say, abiotic variables.

The most important results of this paper are in Figs 2 and 3, where we clearly see the significant correlation between the probability for a species or individual to be a host and  $\epsilon$  as a measure of the statistical significance of the degree of overlap between vector and host distributions. This is evidence that although many other factors, such as species abundance, species competence etc., enter, the fact that a vector and host must co-occur in order to interact leaves a significant predictive imprint. In other words, the more overlap the more opportunity for host-vector interactions. The multi-factorial nature of the complex relation between host and vector is implicit in that the relation between host probability and  $\epsilon$ , although statistically significant, is not characterized by a very high value of  $R^2$ . Especially noteworthy is the decile 1 result, where 3 out of 7 species were identified as hosts—*Baiomys taylori*, *Peromyscus maniculatus* and *Peromyscus eremicus*. These were collected from sites in Jalisco and Nuevo Leon, states from where collections of *Lutzomyia* were previously scarce or non-existent. However, recently, it has been confirmed that various species of *Lutzomyias* are relatively common in these areas. Thus, we believe that a part of the low ranking of these species is due to a systematic bias in the historic collection of *Lutzomyias* towards the southern part of Mexico. The impact of this on the relation between the percentage of positive host species or individuals and  $\epsilon$  is substantial. A regression using only the first 9 deciles yields an  $R^2 = 0.92$  versus 0.44 for all 10 deciles for the species- $\epsilon$  relation, while for the individuals- $\epsilon$  relation the corresponding figures are 0.65 and 0.39. Figs 2 and 3 also illustrate what we mean by host importance. The fact that the percentage of infected individuals decreases as a function of epsilon is consistent with the fact that in the higher deciles (higher  $\epsilon$  values) there is a higher probability of a vector encountering an infected host than in the lower deciles (lower  $\epsilon$  values). This does not imply, of course, that the pathogen is transmitted. It is however, once again, a necessary if not sufficient condition.

Besides validating both the general methodology as well as the specific model for Leishmaniasis of [36], our results also provide an extensive list of new hosts for *Leishmania (L.) mexicana* in Mexico that substantially changes what we know about the transmission cycle of the pathogen and the potential efficacy of interventions and/or surveillance efforts. As 33% of collected species tested positive for the presence of *L. (L.) mexicana* our results are completely consistent with the prediction that *L. (L.) mexicana* is a very generalist pathogen and that *Lutzomyia* is a very generalist genus. Furthermore, if we consider the probability that a species is a true negative at the 95% confidence level with respect to the null hypothesis of a 6.7% infection rate across all species, then we see that there are no true negatives. The closest is *Heteromys desmarestianus* with  $N = 30$  and a probability of 88% of being a true negative. Interestingly, though, *Heteromys desmarestianus* has previously been identified as a host [49,59]. We also note that, of the 24 identified host species, only four—*Phyllostomus discolor*, *Peromyscus eremicus*,

*Glossophaga commissarisi* and *Glossophaga soricina*—are associated with a prevalence that, with 95% confidence according to the Wilson score interval, is greater than the average prevalence across all species. Moreover, the first two species were associated with only one collection so that one would not expect the Wilson score interval to be reliable. The exact probability for an observed 100% prevalence with  $N = 1$  is 6.7%.

Our principle result, as stated previously, is the validation of the methodology and the explicit model of [36] for Leishmaniasis. However, we may also further analyse our experimental results. We have noted that they are quite consistent with the hypothesis that most host species have prevalence values that are compatible with the null hypothesis of a constant prevalence of 6.7% across species. For those positive species with  $N > 20$  in fact prevalence is very stable. Of course, heterogeneities are to be expected. This can be due not only to intrinsic differences in host competence but also, potentially, to spatial heterogeneity associated with epidemiological “hotspots” where many variables together may be favourable for transmission. In Fig 4 we see that there is no noticeable relation between prevalence and  $\epsilon$ . We would argue that there should be a dependence if prevalence is averaged for a given species over several geographical locations that systematically sample the range of that species. However, the sampling intensity of the present data is not capable of showing such effects at a species by species level. The fact that the species by species infection rates are consistent with a relatively constant 6.7% prevalence is compatible with the hypothesis that the competence of the different mammal host species is relatively homogeneous. However, as can be seen in Fig 4, the degree of dispersion of the data is large. Partly this is due to the fact that several of the observed prevalences are associated with very small sample sizes. For instance, the species with prevalence of 100% are associated with only one individual. These can be considered as outliers. This heterogeneity in sample size at the species level is one reason why a coarse grained analysis, as observed in Figs 2 and 3, is more appropriate. Taken at face value, the relative homogeneity of prevalence would also argue against any potential dilution effect due to higher biodiversity as this depends on strong competence heterogeneity amongst hosts. In fact, these results would be consistent with the fact that *Lutzomyias* do not differentiate very much among different potential sources of blood meal [60].

The results of this study stand in stark contrast to our previous understanding of the hosts of different species of *Leishmania*: Of the more than 2000 mammal species on the American continent only about 50 (2.5%) have been identified as hosts of *Leishmania* [44], while in Mexico, 8 out of 419 (2.1%) have been identified as hosts [but see 61]. If our result that 33% of collected mammal species are hosts is extrapolated to other non-collected species then potentially hundreds of mammal species could be implicated as hosts of *Leishmania*. Of the collected species that tested positive, 13 were bats, identified for the first time as hosts of *Leishmania (L.) mexicana* in Mexico [61]. We also identified the grey squirrel as a peri-domestic host species with a high degree of contact with human settlements [62].

If sandflies are generalists, in that they feed off a large number of species across many genera, and pathogen competencies are relatively uniform, as is consistent with our observations, then one would also hypothesize that *Leishmania (L.) mexicana* is also a generalist in that it can infect a large number of potential host species. In the case of Chagas disease, the impressive genetic plasticity and associated adaptability of *Trypanosoma cruzi* has been amply studied [63–65], as well as its epidemiological implications. Our results suggest that *L. (L.) mexicana* should also demonstrate a high degree of genetic plasticity and adaptability to be able to infect such a wide array of mammal species. Recent work on the genetics of *Leishmania* seems to be consistent with this viewpoint [66,67].

Given that previously there were only eight confirmed host species of *Leishmania* in Mexico, the results of this paper change the risk landscape for this neglected disease in Mexico,



both geographically, and in terms of its control or elimination, with similar potential consequences for other countries. Although it is known that *Lutzomyias* have an ample geographic range and therefore are a risk element for Leishmaniasis in many areas of Mexico, what the present work demonstrates is that there are a large number, and potentially very many more, of hosts involved in the transmission cycle of Leishmaniasis. This complicates both interventions and surveillance. In terms of surveillance we would argue that our model yields a good first approximation as to which host species to survey—those that have been identified as hosts and have the highest  $\epsilon$  values. Of course, further field and laboratory work should be carried out to better understand the underlying factors influencing host competence, both at the species level and across different geographical locations for the same species. It is also necessary to test species that are high on the model list but up to now have not been checked. In terms of interventions, with such a large host range, that spans both sylvatic and peri-domestic species, it will be very difficult to eliminate any enzootic transmission cycle, with consequential difficulties in the long term elimination of the vector. Geographically, as can be seen in Fig 5, a risk map derived from the distributions of the eight host species known before the results of this study is completely different to a risk map associated with the 30 species of host that we have now confirmed indicating much higher risk in states such as Jalisco and Nuevo Leon that have until recently been considered low risk compared to the south east.

Our results show that within species collection data there is a great deal of useful information about interspecific interactions and community structure that may be deduced with innovative modelling techniques, such as complex inference networks, and applied to important problem areas such as multi-host diseases. It also shows the importance of the systematic and unbiased collection of data associated with the distributions of potential vectors and potential reservoirs. For instance, the models created showed a higher risk for presence of *Lutzomyia* in the north of Mexico, along the border with the US, than had previously been the case. In fact, recent human cases have been reported in this region, and recent field work has shown that the presence of *Lutzomyia* is more extensive than had been previously thought [44,68].

## Supporting Information

**S1 Table. Ranked list of mammals according to their epsilon values.**  
(DOCX)

**S1 Text. Supplementary Field Methodology.**  
(DOCX)

## Acknowledgments

We deeply thank Rodolfo Dirzo and three anonymous referees for their reviews and for their significant comments and suggestions to improve this manuscript.

## Author Contributions

**Conceptualization:** CRS CGS VSC IB.

**Formal analysis:** CRS CGS VSC IB ERT ARM MBC CDB GGG MHM CNIC MPIL LIID MMRM.

**Methodology:** CRS CGS VSC IB.

**Resources:** IB ERT ARM MBC CDB GGG MHM CNIC MPIL LIID MMRM.

**Writing – original draft:** CRS CGS VSC IB.

## References

1. Feasey N, Wansbrough-Jones M, Mabey DCW, Solomon AW. Neglected tropical diseases. *Br Med Bull.* 2010; 93: 179–200. doi: [10.1093/bmb/ldp046](https://doi.org/10.1093/bmb/ldp046) PMID: [20007668](https://pubmed.ncbi.nlm.nih.gov/20007668/)
2. Hotez PJ, Molyneux DH, Fenwick A, Kumaresan J, Sachs SE, Sachs JD, et al. Control of neglected tropical diseases. *N Engl J Med.* 2007; 357: 1018–1027. doi: [10.1186/1741-7015-8-67](https://doi.org/10.1186/1741-7015-8-67) PMID: [17804846](https://pubmed.ncbi.nlm.nih.gov/17804846/)
3. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature.* 2008; 451: 990–993. doi: [10.1038/nature06536](https://doi.org/10.1038/nature06536) PMID: [18288193](https://pubmed.ncbi.nlm.nih.gov/18288193/)
4. Morens DM, Folkers GK, Fauci AS. The challenge of emerging and re-emerging infectious diseases. *Nature.* 2004; 430: 242–249. PMID: [15241422](https://pubmed.ncbi.nlm.nih.gov/15241422/)
5. Binder S, Levitt AM, Sacks JJ, Hughes JM. Emerging infectious diseases: public health issues for the 21st century. *Science.* 1999; 284: 1311–1313. PMID: [10334978](https://pubmed.ncbi.nlm.nih.gov/10334978/)
6. Daszak P, Cunningham AA, Hyatt AD. Emerging infectious diseases of wildlife: threats to biodiversity and human health. *Science.* 2000; 287: 43–449. doi: [10.1126/science.287.5452.443](https://doi.org/10.1126/science.287.5452.443)
7. Brunner JL, LoGiudice K, Ostfeld RS. Estimating reservoir competence of *Borrelia burgdorferi* hosts: prevalence and infectivity, sensitivity, and specificity. *J Med Entomol.* 2008; 45: 139–147. PMID: [18283955](https://pubmed.ncbi.nlm.nih.gov/18283955/)
8. Laurenson MK, Norman RA, Gilbert L, Reid HW, Hudson PJ. Identifying disease reservoirs in complex systems: Mountain hares as reservoirs of ticks and louping-ill virus, pathogens of red grouse. *J Anim Ecol.* 2003; 72: 177–185. doi: [10.1046/j.1365-2656.2003.00688.x](https://doi.org/10.1046/j.1365-2656.2003.00688.x)
9. de Thoisy B, Lacoste V, Germain A, Muñoz-Jordán J, Colón C, Mauffrey J-F, et al. Dengue infection in neotropical forest mammals. *Vector-Borne Zoonotic Dis.* 2009; 9: 157–170. doi: [10.1089/vbz.2007.0280](https://doi.org/10.1089/vbz.2007.0280) PMID: [18945183](https://pubmed.ncbi.nlm.nih.gov/18945183/)
10. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, et al. Fruit bats as reservoirs of Ebola virus. *Nature.* 2005; 438: 575–576. doi: [10.1038/438575a](https://doi.org/10.1038/438575a) PMID: [16319873](https://pubmed.ncbi.nlm.nih.gov/16319873/)
11. Coura JR, Junqueira AC V, Fernandes O, Valente SAS, Miles MA. Emerging Chagas disease in Amazonian Brazil. *Trends Parasitol.* 2002; 18: 171–176. PMID: [11998705](https://pubmed.ncbi.nlm.nih.gov/11998705/)
12. Ashford RW. The leishmaniasis as emerging and reemerging zoonoses. *Int J Parasitol.* 2000; 30: 1269–1281. doi: [10.1016/S0020-7519\(00\)00136-3](https://doi.org/10.1016/S0020-7519(00)00136-3) PMID: [11113254](https://pubmed.ncbi.nlm.nih.gov/11113254/)
13. Kilpatrick AM, Daszak P, Jones MJ, Marra PP, Kramer LD. Host heterogeneity dominates West Nile virus transmission. *Proc Biol Sci.* 2006; 273: 2327–2333. doi: [10.1098/rspb.2006.3575](https://doi.org/10.1098/rspb.2006.3575) PMID: [16928635](https://pubmed.ncbi.nlm.nih.gov/16928635/)
14. Mills JN, Childs JE. Ecologic studies of rodent reservoirs: Their relevance for human health. *Emerg Infect Dis.* 1998; 4: 529–537. doi: [10.3201/eid0404.980403](https://doi.org/10.3201/eid0404.980403) PMID: [9866729](https://pubmed.ncbi.nlm.nih.gov/9866729/)
15. Woolhouse MEJ, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis.* 2005; 11: 1842–1847. doi: [10.3201/eid1112.050997](https://doi.org/10.3201/eid1112.050997) PMID: [16485468](https://pubmed.ncbi.nlm.nih.gov/16485468/)
16. Viana M, Mancy R, Biek R, Cleaveland S, Cross PC, Lloyd-Smith JO, et al. Assembling evidence for identifying reservoirs of infection. *Trends Ecol Evol.* Elsevier Ltd; 2014; 29: 270–279. doi: [10.1016/j.tree.2014.03.002](https://doi.org/10.1016/j.tree.2014.03.002)
17. Cleaveland S, Laurenson MK, Taylor LH. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos Trans R Soc Lond B Biol Sci.* 2001; 356: 991–999. doi: [10.1098/rstb.2001.0889](https://doi.org/10.1098/rstb.2001.0889) PMID: [11516377](https://pubmed.ncbi.nlm.nih.gov/11516377/)
18. Woolhouse MEJ, Taylor LH, Haydon DT. Population biology of multihost pathogens. *Science.* 2001; 292: 1109–1112. doi: [10.1126/science.1059026](https://doi.org/10.1126/science.1059026) PMID: [11352066](https://pubmed.ncbi.nlm.nih.gov/11352066/)
19. Buhnerkempe MG, Roberts MG, Dobson AP, Heesterbeek H, Hudson PJ, Lloyd-Smith JO. Eight challenges in modelling disease ecology in multi-host, multi-agent systems. *Epidemics.* Elsevier B.V.; 2015; 10: 26–30. doi: [10.1016/j.epidem.2014.10.001](https://doi.org/10.1016/j.epidem.2014.10.001)
20. Dobson A. Population dynamics of pathogens with multiple host species. *Am Nat.* 2004; 164 Suppl: S64–S78. doi: [10.1086/424681](https://doi.org/10.1086/424681)
21. Keesing F, Holt RD, Ostfeld RS. Effects of species diversity on disease risk. *Ecol Lett.* 2006; 9: 485–498. doi: [10.1111/j.1461-0248.2006.00885.x](https://doi.org/10.1111/j.1461-0248.2006.00885.x) PMID: [16623733](https://pubmed.ncbi.nlm.nih.gov/16623733/)
22. Richter D, Spielman A, Komar N, Matuschka F-R. Competence of American Robins as Reservoir Hosts for Lyme Disease Spirochetes. *Emerg Infect Dis.* 2000; 6: 133–138. PMID: [10756146](https://pubmed.ncbi.nlm.nih.gov/10756146/)
23. Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, Hettler D, et al. Experimental infection of North American birds with the New York 1999 strain of West Nile virus. *Emerg Infect Dis.* 2003; 9: 311–322. doi: [10.3201/eid0903.020628](https://doi.org/10.3201/eid0903.020628) PMID: [12643825](https://pubmed.ncbi.nlm.nih.gov/12643825/)

24. Ostfeld RS, Keesing F. Biodiversity and Disease Risk: the Case of Lyme Disease. *Conserv Biol.* 2000; 14: 722–728.
25. Randolph SE, Dobson ADM. Pangloss revisited: a critique of the dilution effect and the biodiversity-buffers-disease paradigm. *Parasitology.* 2012; 139: 847–863. doi: [10.1017/S0031182012000200](https://doi.org/10.1017/S0031182012000200) PMID: [22336330](https://pubmed.ncbi.nlm.nih.gov/22336330/)
26. Wood CL, Lafferty KD, DeLeo G, Young HS, Hudson PJ, Kuris AM. Does biodiversity protect humans against infectious disease? *Ecology.* 2014; 95: 817–832. doi: [10.1890/13-1041.1](https://doi.org/10.1890/13-1041.1) PMID: [24933803](https://pubmed.ncbi.nlm.nih.gov/24933803/)
27. González-Salazar C, Stephens CR. Constructing ecological networks: a tool to infer risk of transmission and dispersal of leishmaniasis. *Zoonoses Public Health.* 2012; 59 Suppl 2: 179–93. doi: [10.1111/j.1863-2378.2012.01479.x](https://doi.org/10.1111/j.1863-2378.2012.01479.x) PMID: [22958263](https://pubmed.ncbi.nlm.nih.gov/22958263/)
28. Haydon DT, Cleaveland S, Taylor LH, Laurenson MK. Identifying reservoirs of infection: A conceptual and practical challenge. *Emerg Infect Dis.* 2002; 8: 1468–1473. doi: [10.3201/eid0812.010317](https://doi.org/10.3201/eid0812.010317) PMID: [12498665](https://pubmed.ncbi.nlm.nih.gov/12498665/)
29. Rhyan JC, Spraker TR. Emergence of diseases from wildlife reservoirs. *Vet Pathol.* 2010; 47: 34–39. doi: [10.1177/0300985809354466](https://doi.org/10.1177/0300985809354466) PMID: [20080482](https://pubmed.ncbi.nlm.nih.gov/20080482/)
30. Klemperer MS, Unnasch TR, Hu LT. Taking a bite out of vector-transmitted infectious diseases. *N Engl J Med.* 2007; 356: 2567–2569. doi: [10.1056/NEJMp078081](https://doi.org/10.1056/NEJMp078081) PMID: [17582066](https://pubmed.ncbi.nlm.nih.gov/17582066/)
31. Orton RJ, Wright CF, Morelli MJ, Juleff N, Thébaud G, Knowles NJ, et al. Observing micro-evolutionary processes of viral populations at multiple scales. *Philos Trans R Soc Lond B Biol Sci.* 2013; 368: 20120203. doi: [10.1098/rstb.2012.0203](https://doi.org/10.1098/rstb.2012.0203) PMID: [23382425](https://pubmed.ncbi.nlm.nih.gov/23382425/)
32. Peterson AT, Carroll DS, Mills JN, Johnson KM. Potential mammalian filovirus reservoirs. *Emerg Infect Dis.* 2004; 10: 2073–2081. PMID: [15663841](https://pubmed.ncbi.nlm.nih.gov/15663841/)
33. Samy AM, Annajar BB, Dokhan MR, Boussaa S, Peterson AT. Coarse-resolution Ecology of Etiological Agent, Vector, and Reservoirs of Zoonotic Cutaneous Leishmaniasis in Libya. *PLoS Negl Trop Dis.* 2016; 10: 1–16. doi: [10.1371/journal.pntd.0004381](https://doi.org/10.1371/journal.pntd.0004381)
34. Han BA, Schmidt JP, Bowden SE, Drake JM. Rodent reservoirs of future zoonotic diseases. *Proc Natl Acad Sci.* 2015; 112: 201501598. doi: [10.1073/pnas.1501598112](https://doi.org/10.1073/pnas.1501598112)
35. Olivero J, Fa JE, Real R, Farfán MÁ, Márquez AL, Vargas JM, et al. Mammalian biogeography and the Ebola virus in Africa. *Mamm Rev.* 2016; 1–14. doi: [10.1111/mam.12074](https://doi.org/10.1111/mam.12074)
36. Stephens CR, Heau JG, González C, Ibarra-Cerdeña CN, Sánchez-Cordero V, González-Salazar C. Using biotic interaction networks for prediction in biodiversity and emerging diseases. *PLoS One.* 2009; 4: e5725. doi: [10.1371/journal.pone.0005725](https://doi.org/10.1371/journal.pone.0005725) PMID: [19478956](https://pubmed.ncbi.nlm.nih.gov/19478956/)
37. González-Salazar C, Stephens CR, Marquet PA. Comparing the relative contributions of biotic and abiotic factors as mediators of species' distributions. *Ecol Modell.* Elsevier B.V.; 2013; 248: 57–70. doi: [10.1016/j.ecolmodel.2012.10.007](https://doi.org/10.1016/j.ecolmodel.2012.10.007)
38. Pimm SL, Lawton JH, Cohen JE. Food web patterns and their consequences. *Nature.* 1991; 350: 669–674.
39. Proulx SR, Promislow DEL, Phillips PC. Network thinking in ecology and evolution. *Trends Ecol Evol.* 2005; 20: 345–353. doi: [10.1016/j.tree.2005.04.004](https://doi.org/10.1016/j.tree.2005.04.004) PMID: [16701391](https://pubmed.ncbi.nlm.nih.gov/16701391/)
40. Bascompte J, Jordano P, Olesen JM. Asymmetric Coevolutionary Networks Facilitate Biodiversity Maintenance. *Science.* 2011; 431: 3–6. doi: [10.1126/science.1123412](https://doi.org/10.1126/science.1123412)
41. Okuyama T, Holland JN. Network structural properties mediate the stability of mutualistic communities. *Ecol Lett.* 2008; 11: 208–216. doi: [10.1111/j.1461-0248.2007.01137.x](https://doi.org/10.1111/j.1461-0248.2007.01137.x) PMID: [18070101](https://pubmed.ncbi.nlm.nih.gov/18070101/)
42. Estrada-Peña A, de la Fuente J, Ostfeld RS, Cabezas-Cruz A. Interactions between tick and transmitted pathogens evolved to minimise competition through nested and coherent networks. *Sci Rep.* Nature Publishing Group; 2015; 5: 10361. doi: [10.1038/srep10361](https://doi.org/10.1038/srep10361)
43. World Health Organization. Control of the leishmaniasis. *World Health Organ Tech Rep Ser.* 2010; 949: 22–26.
44. González C, Rebollar-Téllez EA, Ibáñez-bernal S, Becker I, Martínez-meyer E, Peterson AT, et al. Current knowledge on Leishmania vectors in Mexico: how species' geographic distributions relate to transmission areas. *Am J Trop Med Hyg.* 2011; 85: 839–846. doi: [10.4269/ajtmh.2011.10-0452](https://doi.org/10.4269/ajtmh.2011.10-0452) PMID: [22049037](https://pubmed.ncbi.nlm.nih.gov/22049037/)
45. Bern C, Maguire JH, Alvar J. Complexities of assessing the disease burden attributable to leishmaniasis. *PLoS Neglected Tropical Diseases.* 2008. p. e313. doi: [10.1371/journal.pntd.0000313](https://doi.org/10.1371/journal.pntd.0000313) PMID: [18958165](https://pubmed.ncbi.nlm.nih.gov/18958165/)
46. Gramiccia M, Gradoni L. The current status of zoonotic leishmaniasis and approaches to disease control. *Int J Parasitol.* 2005; 35: 1169–80. doi: [10.1016/j.ijpara.2005.07.001](https://doi.org/10.1016/j.ijpara.2005.07.001) PMID: [16162348](https://pubmed.ncbi.nlm.nih.gov/16162348/)

47. Canto-Lara SB, Van Wynsberghe NR, Vargas-González A, Ojeda-Farfán FF, Andrade-Narváez FJ. Use of Monoclonal Antibodies for the Identification of *Leishmania* spp. Isolated from Humans and Wild Rodents in the State of Campeche, Mexico. *Mem Inst Oswaldo Cruz*. 1999; 94: 305–309.
48. Van Wynsberghe NR, Canto-Lara SB, Damián-Centeno AG, Itzá-Ortiz MF, Andrade-Narváez FJ. Retention of *Leishmania (Leishmania) mexicana* in naturally infected rodents from the state of Campeche, Mexico. *Mem Inst Oswaldo Cruz*. 2000; 95: 595–600. PMID: [10998205](#)
49. Chablé-Santos JB, Van Wynsberghe NR, Canto-Lara SB, Andrade-Narváez FJ. Isolation of *Leishmania (L.) mexicana* from wild rodents and their possible role in the transmission of localized cutaneous leishmaniasis in the State of Campeche, Mexico. *Am J Trop Med Hyg*. 1995; 53: 141–145. PMID: [7677214](#)
50. Rudge JW, Webster JP, Lu D-B, Wang T-P, Fang G-R, Basáñez M-G. Identifying host species driving transmission of schistosomiasis japonica, a multihost parasite system, in China. *Proc Natl Acad Sci U S A*. 2013; 110: 11457–62. doi: [10.1073/pnas.1221509110](#) PMID: [23798418](#)
51. Sierra R, Stephens CR. Exploratory analysis of the interrelations between co-located boolean spatial features using network graphs. *Int J Geogr Inf Sci*. 2012; 26: 441–468. doi: [10.1080/13658816.2011.594799](#)
52. Kato H, Uezato H, Katakura K, Calvopiña M, Marco JD, Barroso PA, et al. Detection and identification of *Leishmania* species within naturally infected sand flies in the andean areas of Ecuador by a polymerase chain reaction. *Am J Trop Med Hyg*. 2005; 72: 87–93. doi: [10.1187/2005.07.087](#) PMID: [15728872](#)
53. Cupolillo E, Grimaldi G, Momen H, Beverley SM. Intergenic region typing (IRT): A rapid molecular approach to the characterization and evolution of *Leishmania*. *Mol Biochem Parasitol*. 1995; 73: 145–155. doi: [10.1016/0166-6851\(95\)00108-D](#) PMID: [8577322](#)
54. Berzunza-Cruz M, Bricaire G, Suazo NS, Pérez-Montfort R, Becker I. PCR for identification of species causing American cutaneous leishmaniasis. *Parasitol Res*. 2009; 104: 691–699. doi: [10.1007/s00436-008-1247-2](#) PMID: [19002715](#)
55. Newcombe RG. Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods. *Stat Med*. 1998; 17: 857–872. PMID: [9595616](#)
56. Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *J Am Stat Assoc*. 1927; 22: 209–212. Available: [http://www.med.mcgill.ca/epidemiology/hanley/tmp/proportion/wilson\\_jasa\\_1927.pdf](http://www.med.mcgill.ca/epidemiology/hanley/tmp/proportion/wilson_jasa_1927.pdf)
57. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2d ed. New York: John Wiley & Son, Inc; 2004.
58. Chaves LF, Harrington LC, Keogh CL, Nguyen AM, Kitron UD. Blood feeding patterns of mosquitoes: random or structured? *Front Zool*. 2010; 7: 3. doi: [10.1186/1742-9994-7-3](#) PMID: [20205866](#)
59. Van Wynsberghe NR, Canto-Lara SB, Sosa-Bibiano EI, Rivero-Cárdenas NA, Andrade-Narváez FJ. Comparison of small mammal prevalence of *Leishmania (Leishmania) mexicana* in five foci of cutaneous leishmaniasis in the State of Campeche, Mexico. *Rev Inst Med Trop Sao Paulo*. 2009; 51: 87–94. doi: [10.1590/S0036-46652009000200006](#) PMID: [19390737](#)
60. Ready PD. Biology of phlebotomine sand flies as vectors of disease agents. *Annu Rev Entomol*. 2013; 58: 227–250. doi: [10.1146/annurev-ento-120811-153557](#) PMID: [23317043](#)
61. Berzunza-Cruz M, Rodríguez-Moreno Á, Gutiérrez-Granados G, González-Salazar C, Stephens CR, Hidalgo-Mihart M, et al. *Leishmania (L.) mexicana* Infected Bats in Mexico: Novel Potential Reservoirs. *PLoS Negl Trop Dis*. 2015; 9: 1–15. doi: [10.1371/journal.pntd.0003438](#)
62. Ramos-Lara N, Koprowski JL. Deforestation and knowledge gaps threaten conservation of less charismatic species: status of the arboreal squirrels of Mexico. *Mammalia*. 2014; 78: 1–11. doi: [10.1515/mammalia-2013-0115](#)
63. Teixeira ARL, Gomes C, Lozzi SP, Hecht MM, Bussacos AC, Nitz N, et al. Environment, interactions between *Trypanosoma cruzi* and its host, and health. *Cad Saude Publica*. 2009; 25 Sup 1: s32–s44.
64. Macedo AM, Machado CR, Oliveira RP, Pena SDJ. *Trypanosoma cruzi*: genetic structure of populations and relevance of genetic variability to the pathogenesis of Chagas disease. *Mem Inst Oswaldo Cruz*. 2004; 99: 1–12.
65. Llewellyn M, Rivett-Carnac JB, Fitzpatrick S, Lewis MD, Yeo M, Gaunt MW, et al. Extraordinary *Trypanosoma cruzi* diversity within single mammalian reservoir hosts implies a mechanism of diversifying selection. *International Journal for Parasitology*. 2011; 41: 609–614. doi: [10.1016/j.ijpara.2010.12.004](#) PMID: [21232539](#)
66. Miles MA, Llewellyn M, Lewis MD, Yeo M, Baleela R, Fitzpatrick S, et al. The molecular epidemiology and phylogeography of *Trypanosoma cruzi* and parallel research on *Leishmania*: looking back and to the future. *Parasitology*. 2009; 136: 1509–1528. doi: [10.1017/S0031182009990977](#) PMID: [19691868](#)

67. Baidouri F El, Diancourt L, Berry V, Chevenet F, Pratlong F, Marty P, et al. Genetic structure and evolution of the *Leishmania* genus in Africa and Eurasia: what does MLSA tell us. *PLoS Negl Trop Dis*. 2013; 7: e2255. doi: [10.1371/journal.pntd.0002255](https://doi.org/10.1371/journal.pntd.0002255) PMID: [23785530](https://pubmed.ncbi.nlm.nih.gov/23785530/)
68. Roque ALR, Jansen AM. Wild and synanthropic reservoirs of *Leishmania* species in the Americas. *Int J Parasitol Parasites Wildl*. The Authors; 2014; 3: 251–262. doi: [10.1016/j.ijppaw.2014.08.004](https://doi.org/10.1016/j.ijppaw.2014.08.004) PMID: [25426421](https://pubmed.ncbi.nlm.nih.gov/25426421/)