



Simulating Population Genetics of Pathogen Vectors in Changing Landscapes: Guidelines and Application with *Triatoma brasiliensis*

Francois Rebaudo^{1,2*}, Jane Costa³, Carlos E. Almeida⁴, Jean-Francois Silvain^{1,2}, Myriam Harry², Olivier Dangles^{1,2,5}

1 BEI-UR072, IRD, Gif-sur-Yvette, France, **2** LEGS-UPR9034, CNRS-UPSud11, Gif-sur-Yvette, France, **3** Laboratório de Biodiversidade Entomológica, Instituto Oswaldo Cruz - Fiocruz, Rio de Janeiro, Rio de Janeiro, Brasil, **4** Departamento de Ciências Biológicas, Faculdade de Ciências Farmacêuticas, UNESP, Araraquara, Sao Paulo, Brasil, **5** Instituto de Ecología, Campus Cotacota, Universidad Mayor San Andrés, La Paz, Bolivia

Abstract

Background: Understanding the mechanisms that influence the population dynamics and spatial genetic structure of the vectors of pathogens infecting humans is a central issue in tropical epidemiology. In view of the rapid changes in the features of landscape pathogen vectors live in, this issue requires new methods that consider both natural and human systems and their interactions. In this context, individual-based model (IBM) simulations represent powerful yet poorly developed approaches to explore the response of pathogen vectors in heterogeneous social-ecological systems, especially when field experiments cannot be performed.

Methodology/Principal Findings: We first present guidelines for the use of a spatially explicit IBM, to simulate population genetics of pathogen vectors in changing landscapes. We then applied our model with *Triatoma brasiliensis*, originally restricted to sylvatic habitats and now found in peridomestic and domestic habitats, posing as the most important *Trypanosoma cruzi* vector in Northeastern Brazil. We focused on the effects of vector migration rate, maximum dispersal distance and attraction by domestic habitat on *T. brasiliensis* population dynamics and spatial genetic structure. Optimized for *T. brasiliensis* using field data pairwise fixation index (*FST*) from microsatellite loci, our simulations confirmed the importance of these three variables to understand vector genetic structure at the landscape level. We then ran prospective scenarios accounting for land-use change (deforestation and urbanization), which revealed that human-induced land-use change favored higher genetic diversity among sampling points.

Conclusions/Significance: Our work shows that mechanistic models may be useful tools to link observed patterns with processes involved in the population genetics of tropical pathogen vectors in heterogeneous social-ecological landscapes. Our hope is that our study may provide a testable and applicable modeling framework to a broad community of epidemiologists for formulating scenarios of landscape change consequences on vector dynamics, with potential implications for their surveillance and control.

Citation: Rebaudo F, Costa J, Almeida CE, Silvain J-F, Harry M, et al. (2014) Simulating Population Genetics of Pathogen Vectors in Changing Landscapes: Guidelines and Application with *Triatoma brasiliensis*. PLoS Negl Trop Dis 8(8): e3068. doi:10.1371/journal.pntd.0003068

Editor: Philippe Solano, IRD/CIRDES, Burkina Faso

Received: April 11, 2014; **Accepted:** June 21, 2014; **Published:** August 7, 2014

Copyright: © 2014 Rebaudo et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This research was supported by Fapesp (process numbers 2010/17027-0, and 2011/22378), and the French Agence Nationale de la Recherche (project Adaptanthrop ANR-097-PEXT-009). The funders had no role in study design, data collection and analysis, decision to submit for publication, or preparation of the manuscript.

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

* Email: francois.rebaudo@ird.fr

Introduction

Human-induced landscape changes are increasingly recognized as important drivers of infectious disease outbreak and emergence events, resulting in significant threats to public health [1–3]. Worldwide the rapid modification of natural habitats has triggered intense research on the landscape epidemiology of vector diseases to describe how the temporal dynamics of host, vector, and pathogen populations interact spatially within heterogeneous and changing environments to enable transmission (see [4] for a review). Landscape changes not only affect the transmission of

endemic infections by modifying contact patterns between hosts and vectors [5], but also have an effect on selection pressure, leading to the dominance of pathogen strains and vector populations adapted to new environmental conditions [4]. While evolutionary ecologists have increasingly recognized the importance of evolutionary processes (e.g., local adaptive genetic variation in pathogen vectors) to predict population response to changing landscape conditions [6], this issue has received relatively little attention among landscape epidemiologists. As a result, we are lacking spatially and temporally explicit quantitative approaches required to understand the key causal mechanisms (e.g.,

Author Summary

Worldwide, humans are modifying landscapes at an unprecedented rate. These modifications have an influence on the ecology of pathogen vectors, yet this issue has received relatively little input from modeling research. The current study presents guidelines for the use of a modeling framework for the representation of the dynamics and spatial genetic structure of pathogen vectors. It allows considering spatiotemporal landscape modifications explicitly, to represent human-altered modifications and consequences. We applied this modeling framework to *Triatoma brasiliensis*, vector of the pathogen *Trypanosoma cruzi* responsible for the Chagas disease, in the semi-arid Northeastern Brazil. Using field data of pairwise fixation index (*F_{ST}*) from microsatellite loci, we found that migration rate, maximum dispersal distance and attraction by domestic habitat were all key parameters to understand vector spatial genetic structure at the landscape level. At the interface across disciplines, this study provides to the community of epidemiologists a testable and applicable framework to foresee landscape modification consequences on vector dynamics and genetic structure, with potential implications for their surveillance and control.

habitat selection and adaptation, migration and resulting gene flow) involved in pathogen vector response to landscape changes [7]. This is particularly true for pathogen vectors in tropical regions where landscape changes (deforestation, urbanization) occur at an accelerated rate [8] putting at risk human populations with limited resources to face disease-related challenges.

The main objective of this study is to propose a methodological framework to simulate spatial population genetics of pathogen vectors in heterogeneous and changing landscapes, in order to link observed patterns with processes. We adapted the SimAdapt simulation software [9] to simulate the evolution of both neutral and adaptive genotypes of diploid, sexually reproducing pathogen vectors introduced into a landscape. This genetic model accounts for vector dispersal and adaptation to local conditions and is coupled to a cellular automaton allowing the representation of land-use and land-cover changes. It includes landscape features known to influence vectors' genetic structure (e.g., roads, domestic, peridomestic and sylvatic habitats linked to possible loci under selection). Model simulations can be compared to field data of vector's spatial genetic structure. As a step further, the model can be used to simulate the evolution of vector spatial genetic structure in changing landscapes (e.g., prospective scenario of deforestation or urbanization).

After providing general guidelines on our model, we applied SimAdapt to explore local adaptation processes of pathogen vectors in a real-world landscape using *Triatoma brasiliensis* Neiva, 1911 (Hemiptera, Reduviidae, Triatominae) as a study model. *T. brasiliensis* is a blood sucking bug vector of the pathogen responsible for the Chagas disease (American trypanosomiasis), caused by a parasite *Trypanosoma cruzi* (Kinetoplastea, Trypanosomatidae). This disease affects approximately 10 million people in Latin America and Caribbean [10], and is recognized by the World Health Organization as one of the world's most neglected tropical disease [11]. *T. brasiliensis* represents the most important vector of *T. cruzi* in Northeastern Brazil [12,13]. As for several species of native Triatominae (e.g., *Rhodnius equatorialis*, [14]; *T. pseudomaculata*, [15,16]), *T. brasiliensis* was originally restricted to sylvatic habitats but since its description it has been increasingly found invading and establishing in peridomestic and domestic

habitats [17–20]. Among all Brazilian triatomines, *T. brasiliensis* is the one that exhibits the highest pressure for re-infestation after insecticide house spraying. As a consequence, six months (sometimes longer) after chemical treatment, human dwellings start being re-infested [21,22]. As a native vector, its eradication would require significant logistical and technical investment in the long term, and, therefore, control efforts are kept on insecticide house spraying and improvement.

Here, we used SimAdapt to explore three parameters related to the dispersal of *T. brasiliensis*: i) the migration rate; ii) the dispersal distance (i.e. the maximum distance covered by an individual during its life cycle, including passive and active dispersal); and iii) the attraction by domestic habitat (i.e. the strength that avoid individual from emigrating, as a consequence of attraction by light or availability of hosts to feed on). Model simulations were then compared to field data of *T. brasiliensis* population genetic structure and prospective scenarios of landscape change effects on *T. brasiliensis* adaptation were run.

Materials and Methods

The SimAdapt model

Digital representation of the landscape. The first step to perform simulations in SimAdapt is to build a digital representation of the study landscape. To achieve this goal, the landscape is categorized into habitat types. Optionally, as SimAdapt selection submodel can be set to consider that being adapted to a given habitat is considered advantageous while the individual is in this habitat, but disadvantageous in others habitats (see [23] for a discussion on alternative selection models), the categorization of habitat type can reflect a selective pressure resulting in a differential reproductive fitness of individuals. Once categorized into habitat types, the landscape is converted into a matrix representation. In this matrix, habitat types are converted into a discrete variable made of grid cells (raster representation, see Figure 1 in the study case below) unlike in a geographic information system (GIS), where delimited areas (here habitat types) are typically drawn continuously (vectorial representation). The tessellation defines the resolution of the landscape which should be defined according to the ecological characteristics of the vector model, in particular its dispersal capabilities. At this point, each cell of the grid corresponding to a habitat type can be parameterized with a carrying capacity and a resistance to vectors' emigration. The carrying capacity defines the maximum number of individuals that can be located into a cell, and the resistance for migration defines the permeability of the landscape (*sensu* [24]) ranging from 0 (i.e., no resistance) to 100 (i.e., impermeable barrier) – see [1,25] for reviews on the effects of

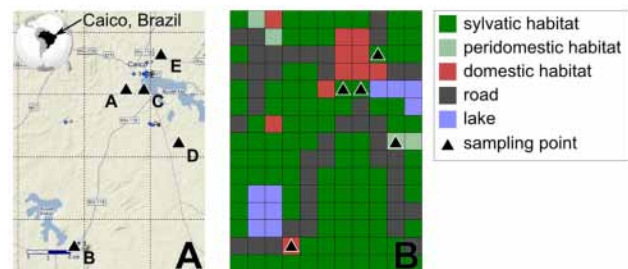


Figure 1. Study landscape (A) and matrix representation of the landscape in simAdapt (B). The different habitat types, roads and lakes are represented by different colors. The black triangles represent the location of sample points from the field data in the Caico municipality in Northeastern Brazil.

doi:10.1371/journal.pntd.0003068.g001

Table 1. Initialization of state variables in the SimAdapt model.

Description	Value	Justification/References
Population dynamics submodel		
Growth rate of the logistic model	0.3	[13]
* Migration rate (m)	[0.1:1.0]	[18]
* Maximum dispersal distance (d)	[1:5]	[29]
* Attraction by domestic habitat (l)	[1:10]	[15,51]
Genetic submodel		
Number of microsatellites loci	7	[21]
Heterozygosity at initialization (H_0)	0.5±0.2	To fit the field sample
Mutation rate for microsatellites	10e-4	Default value [46]
Number of loci under selection per habitat type	1	Default value [9]
Coefficients of selection (s and h)	$s = 0.2$ $h = 0.5$	Assumption for codominance and strong selection (prospective scenarios only)
Landscape representation		
Carrying capacity matrix for the logistic growth model (adults)	50	Based on field sample [30,31]
Resistance matrix for emigration	Heterogeneous	To fit the landscape
Habitat type matrix for natural selection	[1:3]	To fit the landscape
Location and number of individuals at initialization	everywhere (50 adult individuals)	To fit the field sample
Scenario of landscape change	Deforestation and urbanization	[8,44,45]
Sampling submodel		
Virtual sampling	Population sampled without replacement	According to field sampling design
Number of individuals sampled per sampling point	25	According to field sampling design
Number of points sampled	5	According to field sampling design
Simulation		
Number of repetitions per simulation	30	To account for stochasticity
Number of generations per simulation	100	Time to stabilize F_{ST}
Output files format	ARLEQUIN (Arlecore)	To fit the field sample analyzes

The three variables studied in this paper are highlighted with an asterisk at the beginning of the first column.
doi:10.1371/journal.pntd.0003068.t001

landscape resistance on infectious diseases and vectors. Technically, any GIS software provides the necessary features to define areas (habitat types), as well as additional layers for carrying capacity and resistance for migration. This information can be exported into SimAdapt as three different “ascii” files at a chosen resolution. Alternatively, the map can be built entirely from the graphical user interface (see Appendix S1 in [9]).

Initialization of the simulation model. The second step is to initialize the parameters of the model as presented in Table 1. Depending on available knowledge on vector species, unknown parameters can be estimated within a range of possible values (referred as “parameter space”), and sensitivity analyses performed to evaluate the influence of parameter variation on simulations output results.

Optimization. The third step is to optimize model parameterization using field data. It consists in the comparison of studied parameters in their range of possible values (parameter space), with observed data. Parameter optimization can take several forms depending on available data. It can vary from standard genetic algorithms [26], or random search [27], to the use of simulations previously performed as sensitivity analysis.

Prospective scenarios. Once the model has been optimized, SimAdapt can perform prospective simulations to represent the effects of landscape changes (based on scenarios of land-use

changes) on local adaptation of vector species. Technically, a user-defined scenario of land-use change is established in SimAdapt using the landscape cellular automaton submodel (see [9] for more details).

The case of *Triatoma brasiliensis* in Northeastern Brazil

As an example of application of our model in the context of vector-borne tropical diseases, we used genetic field data on *T. brasiliensis* (vector of the pathogen responsible for the Chagas disease) in the Caicó municipality in Northeastern Brazil (Rio do Grande do Norte state). Triatomine individuals were sampled during a field work performed in March 2011. Domestic habitats consisted in houses, peridomestic habitats of the area 50 meters around houses and sylvatic habitats of areas at a minimum of 200 meters from any house. In total 126 individuals were collected in five different locations (populations) chosen to cover a range of distance between sampling points and habitat types (three sylvatic, one peridomestic and one domestic populations; see Figure 1A). All individuals were genotyped according to seven microsatellites as neutral markers [21], and pairwise fixation index (F_{ST}), allelic diversity and observed heterozygosity computed using Arlequin version 3.5 [28]. A detailed description is available in supporting information S1, and F_{ST} results in Table 2.

Table 2. Pairwise fixation index (*F_{ST}*) between sampling points A, B, C, D and E (in bold italic when the associated *p*-value was below 0.05), for the observed dataset in the Caicó municipality in Northeastern Brazil.

	A	B	C	D	E
A	0				
B	<i>0.02130</i>	0			
C	<i>0.01513</i>	<i>0.02265</i>	0		
D	0.00635	0.00852	<i>0.02791</i>	0	
E	0.00000	<i>0.03253</i>	<i>0.02440</i>	<i>0.01609</i>	0

doi:10.1371/journal.pntd.0003068.t002

As the main objective of our study case was to explore the effect of vector dispersal on observed population genetic structure (using *F_{ST}* values), we focused our analysis on three specific variables: i) vector's migration rate (*m*); ii) vector's dispersal distance (*i.e.*, the maximum distance covered by an individual during its life cycle, including passive and active dispersal (*d*)); and iii) vector's attraction by a specific habitat type (*i.e.*, the strength that impedes individual from emigrating when they are located in a specific habitat type (*l*)). Note that all output files from simulations were designed as input files for Arlequin (genotypes of individuals from the 7 microsatellite loci), using the same sampling method, so that observed and simulated *F_{ST}* values could be compared.

Study landscape and its digital representation. The study landscape, in the semi-arid Northeastern Brazil, part of the municipality of Caicó, was represented as a matrix of 11 per 15 grid cells (see Figure 1). In this tessellation, one grid cell is equivalent to 2*2 km. Grid cells were categorized according to *T. brasiliensis* ecology [12,19] to include domestic habitats, peridomestic habitats, and sylvatic habitats (see [18] for a description and discussion). Physical barriers to dispersal and establishment (lakes) and potential factors of passive dispersal by human (roads), were also considered and georeferenced. This allowed us to apply a matrix of landscape resistance to our grid. Based on previous studies on *T. infestans* by [29], the maximal active dispersal distance was set to 2 km per generation over the entire landscape. To simulated passive dispersal, the dispersal distance on roads was fixed to a maximum of 10 km per generation (corresponding to *d* = 5, with *d* varying from 1 to 5 grid cells in our simulations).

Initialization of the simulation model. We set up our landscape with 50 adult individuals per grid cell, within the range of Triatominae densities documented in the literature ([30,31]). Note that using more individuals per grid cell reduced the variance in simulation results, due to a reduction in genetic drift, but did not change observed patterns. Individuals were attributed genotypes using 7 microsatellite loci with alleles randomly chosen in a normal distribution. We did not account for selection as analyses performed from *F*-statistics on field data suggested no loci under selection (using Arlequin based on [32]). The other parameters were set to their default value in SimAdapt or according to available literature (see Table 1). The three variables related to dispersal (*m*, *d*, and *l*) were defined as ranges of potential values (parameter space; see Table 1) and then further explored through a sensitivity analysis (see below).

Model validation. Validation is an important process in model development, and consist in the demonstration that the model meets performance standards under specific conditions [33,34]. To validate SimAdapt initialization for *T. brasiliensis*, we calculated *F_{ST}* using Arlequin (based on genotypes from the 7 microsatellite loci and assuming no selection for habitat type) and compared its values with *F_{ST}* from theoretical expectation in an

island model [35]. In an island model, the theoretical response of *F_{ST}* to migrant number fits a curve of the form $F_{ST} \approx 1/(4Nm+1)$, where *N* is the effective population size and *m* is the migration rate between populations. Our case slightly differed from the island model as it included a finite number of populations and dispersal mechanisms partially driven by the habitat types. We consequently expected different *F_{ST}* values than those predicted by theory [36,37]. We found a significant relationship between *F_{ST}* and *Nm* using a nonlinear least squares model for migration rate from 0.1 to 1 ($F_{ST} = 1/(5.85Nm+1.24)$, $p < 0.05$, see Figure 2). This result confirmed that, within the range of parameters' space, our simulation model behaved according to theoretical expectations thereby validating its use for *T. brasiliensis*.

Sensitivity analyses. Sensitivity analyses were performed to explore the effect of three variables related to vector dispersal on triatomine genetic differentiation coefficient (*F_{ST}*): i) the effect of migration rate from *m* = 0.1 to 1 per generation; ii) the effect of

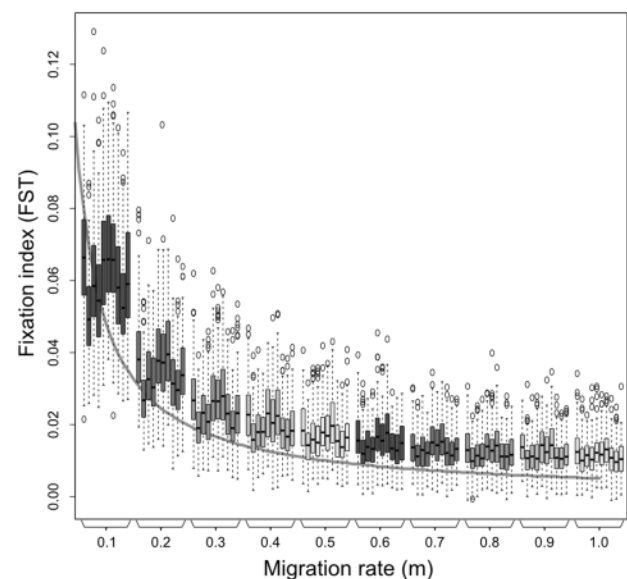


Figure 2. Validation of the model initialization using the response of *F_{ST}* to migration rate. *F_{ST}*s between sampling points (from left to right for each migration rate: AB, AC, AD, AE, BC, BD, BE, CD, CE, DE), are represented as a function of migration rate from 0.1 to 1. Simulated results are represented using boxplots of 30 repetitions, for all values of dispersal distance (*d* ranging from 1 to 5 by 1), *i.e.* 150 pairwise *F_{ST}* values per boxplot. The theoretical expectation is represented by a solid grey line ($F_{ST} \approx 1/(4Nm+1)$ with *N* = 50 and *m* ranging continuously from 0 to 1). doi:10.1371/journal.pntd.0003068.g002

maximum dispersal distance from $d=1$ to 5 grid cells per generation; and iii) the effect of attraction by domestic habitat from $l=0$ to 10 (l representing a factor that reduce emigration from domestic habitat, *i.e.* when $l=10$, $m=m/10$). The resulting sets of simulations were repeated 30 times to account for stochasticity. We performed an additional set of 30 repetitions to ensure that the number of repetitions was representative of stochastic processes (Student tests, $T_{mean} = 0.39$, $df_{mean} = 296.65$, $p_{min} > 0.23$). All analyses were performed on a random sampling of 25 individuals at the five sampling locations (*i.e.*, 125 individuals in total; see [38] for a discussion on virtual sampling). Genotypes of sampled individuals for the 7 microsatellite loci were analyzed to calculate F_{ST} after 100 generations using Arlecore from Arlequin version 3.5 [28]. This time was sufficient to observe a stabilization of the F_{ST} for all sets of simulations. Analyses of variance (ANOVA) were run to evaluate the effect of migration rate (m), dispersal distance (d) and domestic habitat attraction (l) on the F_{ST} using R software [39].

Model optimization. Model outputs from the sensitivity analysis were compared to those obtained from field data using least-square optimization [40] between simulated and observed F_{ST} . The optimization procedure generated 300 F_{ST} values (10 migration rates values \times 5 dispersal distances values \times 6 domestic habitat attraction values), repeated 30 times. Based on this data, we determined the values of migration rate, maximum dispersal distance and attraction by domestic habitat that minimized the difference between simulated and observed F_{ST} data. Additionally, we plotted a ternary plot using an Akima interpolation [41] on the mean values of the 30 repetitions to visualize the topography of the least-square optimization. To evaluate the goodness of fit of our simulation, we compared simulated and observed F_{ST} values using one-sample Student tests (see [42,43] for examples of F_{ST} comparisons). The values that minimized difference between simulated and observed data were then used to initialize the prospective scenario.

Prospective scenarios. We performed forward-time simulations of the potential effects on *T. brasiliensis* population genetics of landscape change (urbanization) within the next 50 years. Our simplified scenario of urbanization consisted in the progression of peridomestic and domestic habitats over sylvatic habitats (1 km per year to mimic worst case scenario of deforestation, see [44,45] for a discussion). The prospective scenarios included selection considering one locus under selection per habitat type (domestic, peridomestic and sylvatic), with two alleles ("generalist" and "specialized", see [9]). As selection coefficient and degree of dominance (s and h , see [46]) were unknown and because our field sampling did not allow to infer such parameters (no time series of alleles frequencies, see [47]), the degree of dominance was represented to reproduce co-dominance ($h = 0.5$), and the selection coefficient to reproduce strong selection ($s = 0.2$) (see [46]). *T. brasiliensis* dispersal variables were set up using results from the optimization procedure ($m = 0.6$; $d = 3$; $l = 2$; see Figure 3). The simulations were run over 100 triatomine generations, with one generation equivalent to approximately 6 months ([13]). Individuals were sampled every generation at all sampling points, and F_{ST} computed over time using Arlecore (see Table 1 for parameterization). Four scenarios were compared using two factors (land-use change and selection) and two states (with/without) using two-way analysis of variance of F_{ST} after 100 generations.

Results

Sensitivity analyses

As expected, we found a significant effect of migration rate (m) on the F_{ST} values (see Table 3). The effect of dispersal distance (d)

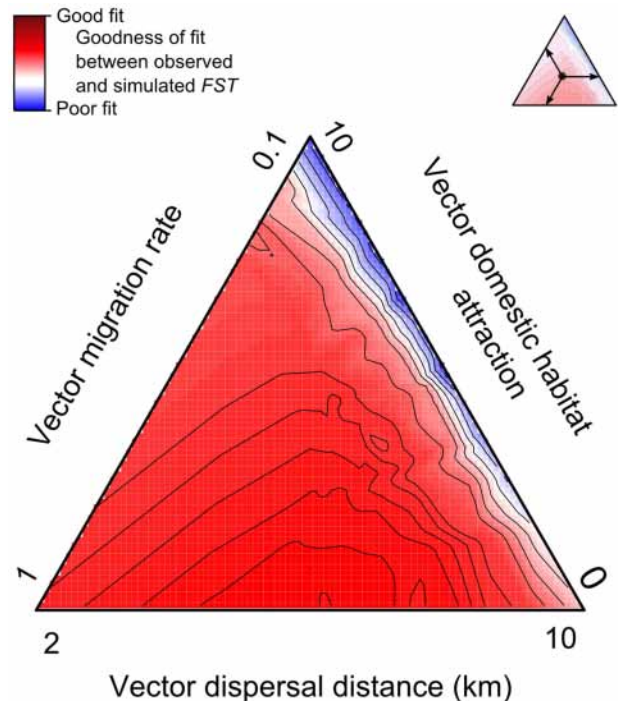


Figure 3. Goodness of fit between observed and simulated F_{ST} for *T. brasiliensis* in Northeastern Brazil. The red from blue color gradient represents an Akima interpolation of the least-square optimization between observed and simulated F_{ST} for different values of vector migration rate, dispersal distance and domestic habitat attraction. F_{ST} were computed using Arlequin over the 10 couples of sampled points. Sets of simulations were repeated 30 times for each value of migration rate (m index ranging from 0.1 to 1 by 0.1), dispersal distance (d ranging from 2 to 10 km by 2 km) and domestic habitat attraction (l index ranging from 0 to 10 by 2). The plot is represented using mean values with a gradient from blue (high value, *i.e.*, poor fit), to red (low value, *i.e.*, good fit).
doi:10.1371/journal.pntd.0003068.g003

was also significant except between two sampling sites (A and C) that were very close from each other (Euclidian distance; $F_{AC(1, 8991)} = 1.4$, $p = 0.24$). Dispersal distance explained much less variance in the ANOVA than migration rate (see differences in F -values in Table 3). The same result was found when testing for the effect of domestic habitat attraction on F_{ST} . On average, for all 10 couples of sampled locations, $58\% \pm 7$ of the variance of F_{ST} in the ANOVA models was explained by migration rate, dispersal distance, and domestic type attraction.

Optimization

The optimization procedure revealed that the difference between observed and simulated F_{ST} was lowest at high migration rates, average dispersal distances and low domestic habitat attractions (see the redder zone of the ternary plot, Figure 3). The best set of parameter values from this zone explained 50% of the F_{ST} values observed in the field (see Table 4, using $m = 0.6$; $d = 3$; and $l = 2$). Note that overall, the combinations of migration rate, dispersal distance and domestic habitat attraction could significantly explain 70% of F_{ST} values.

Prospective scenarios of land-use change

The dynamics of the spatial genetic structure of *T. brasiliensis* populations strongly differed between the urbanization and the "no land-use change" (control) scenarios (see Figure 4 for F_{ST}

Table 3. Analysis of variance (ANOVA) of the effect of migration rate, dispersal distance, and domestic habitat attraction on *FST*.

	Df	Sum Sq	Mean Sq	F value	p-value
migration rate (m)	1	22.725	22.725	67252	<0.05
dispersal distance (d)	1	0.097	0.097	288	<0.05
domestic habitat attraction (l)	1	2.960	2.960	8759	<0.05
interaction m:d	1	0.012	0.012	35	<0.05
interaction m:l	1	0.053	0.053	157	<0.05
Residuals	89982	30.406	0.0003		

FST were computed between the 10 couples of locations for the 300 parameters combinations, with 30 repetitions per combination.
doi:10.1371/journal.pntd.0003068.t003

between individuals located at sampling points A and B). In the control scenarios, *FST* reached a threshold and stabilized within 100 generations for all couples of sampled locations (*e.g.*, mean *FST* threshold value of 0.021 between A and B with and without selection, see Figure 4). Contrastingly, *FST* values did not reach such threshold in the urbanization scenarios (*e.g.*, mean *FST* threshold value of 0.030 and 0.027 between A and B with and without selection, respectively), except between individuals at sampling locations A, C and E. We found a significant effect of land-use change on *FST* after 100 generations (*e.g.*, between A and B: $F = 24.6$, $df = 1$, $p < 0.05$). Contrastingly, the effect of selection was not significant, even if *FST* values were generally higher in the scenarios including selection.

Discussion

Understanding the dispersal behavior of vector arthropods is a central issue in the control and surveillance of vector-borne tropical diseases [48,49]. Although dispersal studies on vectors have been conducted since the eighties (*e.g.*, [29]), accurate descriptions of the spatio-temporal distribution of most tropical pathogen vectors are still lacking. Along with recent advances in spatially explicit models [48], our study proposes a contribution to better characterize the dispersal behavior of the vectors of *T. cruzi*, responsible for Chagas disease. This characterization is exemplified in this study with a generic platform providing a natural description of the dispersal mechanisms [50], adapted to the social-ecological system inhabited by *T. brasiliensis*. Our study

Table 4. Comparison of *FST* values between simulated and observed data using One-sample Student tests.

Couples of sampled locations	m	d	l	P_{mean} (% >0.05)	T_{mean}	df
AB*	[0.1:1]	[1:5]	[0:10]	0.113 (29%)	5.67	29
	0.6	3	2	0.867	-0.17	29
AC*	[0.1:1]	[1:5]	[0:10]	0.077 (20%)	4.50	29
	0.6	3	2	0.06	-1.96	29
AD	[0.1:1]	[1:5]	[0:10]	<0.01 (0%)	14.18	29
	0.6	3	2	<0.01	13.03	29
AE	[0.1:1]	[1:5]	[0:10]	<0.01 (0%)	25.14	29
	0.6	3	2	<0.01	23.06	29
BC*	[0.1:1]	[1:5]	[0:10]	0.097 (28%)	5.38	29
	0.6	3	2	0.325	-1.00	29
BD	[0.1:1]	[1:5]	[0:10]	<0.01 (0%)	14.98	29
	0.6	3	2	<0.01	12.66	29
BE*	[0.1:1]	[1:5]	[0:10]	0.030 (9%)	8.12	29
	0.6	3	2	0.624	0.49	29
CD	[0.1:1]	[1:5]	[0:10]	0.038 (10%)	-4.26	29
	0.6	3	2	<0.01	-12.03	29
CE*	[0.1:1]	[1:5]	[0:10]	0.020 (5%)	9.00	29
	0.6	3	2	0.864	0.17	29
DE	[0.1:1]	[1:5]	[0:10]	0.008 (3%)	16.37	29
	0.6	3	2	<0.01	10.08	29

For each couple of sampled locations, the first line represents average values for all combinations of parameters (*m*, *d* and *l*) with the percentage of significant combinations between brackets, and the second line for *m* = 0.6, *d* = 3 and *l* = 2. Couples of sampled locations significantly explained by this parameterization are shown in bold with an asterisk.

doi:10.1371/journal.pntd.0003068.t004

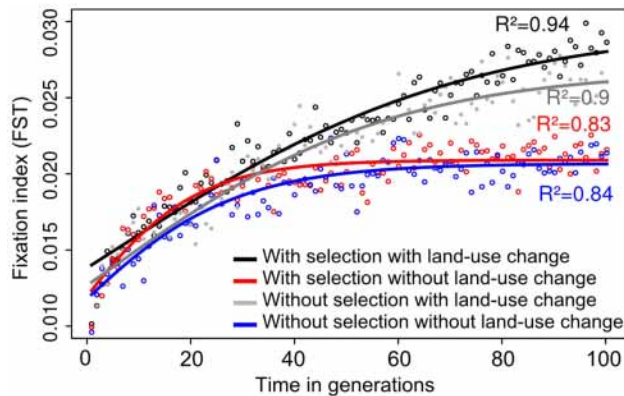


Figure 4. Evolution of *FST* through time between two sampled locations. Each point represents the mean of 30 repetitions. Curves were fitted to the general form of a sigmoid function using *nls* function in R. The different colors corresponds to scenarios with selection and with land-use change (in black); with selection and without land-use change (in red); without selection and without land-use change (in blue); and without selection with land-use change (in grey). doi:10.1371/journal.pntd.0003068.g004

successfully ranked parameters of *T. brasiliensis* dispersal and participated in explaining observed patterns (spatial population differentiation using *FST*) by linking them to processes (vectors ecology and behavior). It thus represents a contribution in understanding the underlying mechanisms of *T. brasiliensis* spatial genetic structure and population dynamics.

Inevitably, our study made a series of assumptions and simplifications inherent to the modeling process. If some were attributable to the methodological framework, others pointed gaps in the knowledge of *T. brasiliensis* ecology and genetics. The selection for habitat type submodel, for instance, relied on the theoretical basis of one locus under selection per habitat type. Further studies of *T. brasiliensis* are needed and should allow the identification candidate genes responsible for adaptation for habitat type which would considerably refine the theoretical submodel, and possibly lead to inference on selection coefficients. Moreover, additional field studies, such as those conducted for other Triatomine species (*e.g.*, studies describing and quantifying the influence on vectors of public street lights; [51,52]; or describing the vectors active and passive dispersal [53]) would help refining our results for *T. brasiliensis*. Additionally, the use of a more complex population dynamic model might have permitted to analyze the impact of other demographic parameters on population structuring, like the density of the population or the lifespan. Consequently, prospective scenarios, grounded on actual

References

- Biek R, Real LA (2010) The landscape genetics of infectious disease emergence and spread. *Molecular Ecology* 19: 3515–3531. doi:10.1111/j.1365-294X.2010.04679.x.
- Meentemeyer RK, Haas SE, Václavík T (2011) Landscape Epidemiology of Emerging Infectious Diseases in Natural and Human-Altered Ecosystems. *Annu Rev Phytopathol* 50: 379–402. doi:10.1146/annurev-phyto-081211-172938.
- Reid RS, Kruska RL, Muthui N, Taye A, Wotton S, et al. (2000) Land-use and land-cover dynamics in response to changes in climatic, biological and socio-political forces: the case of southwestern Ethiopia. *Landscape Ecology* 15: 339–355. doi:10.1023/A:1008177712995.
- Reisen WK (2010) Landscape Epidemiology of Vector-Borne Diseases. *Annu Rev Entomol* 55: 461–483. doi:10.1146/annurev-ento-112408-085419.
- Patz JA, Daszak P, Tabor GM, Aguirre AA, Pearl M, et al. (2004) Unhealthy Landscapes: Policy Recommendations on Land Use Change and Infectious Disease Emergence. *Environ Health Perspect* 112: 1092–1098. doi:10.1289/ehp.6877.
- Manel S, Segelbacher G (2009) Perspectives and challenges in landscape genetics. *Molecular Ecology* 18: 1821–1822. doi:10.1111/j.1365-294X.2009.04151.x.
- Brearely G, Rhodes J, Bradley A, Baxter G, Seabrook L, et al. (2013) Wildlife disease prevalence in human-modified landscapes. *Biol Rev* 88: 427–442. doi:10.1111/brv.12009.
- Barretto AGOP, Berndes G, Sparovek G, Wirsenius S (2013) Agricultural intensification in Brazil and its effects on land use patterns: An analysis of the 1975–2006 period. *Glob Change Biol* 19: 1804–15. doi:10.1111/gcb.12174.
- Rebaudo F, Le Rouzic A, Dupas S, Silvain J-F, Harry M, et al. (2013) SimAdapt: an individual-based genetic model for simulating landscape management impacts on populations. *Methods Ecol Evol* 4: 595–600. doi:10.1111/2041-210X.12041.
- Hotez PJ, Dumonteil E, Woc-Colburn L, Serpa JA, Bezak S, et al. (2012) Chagas Disease: “The New HIV/AIDS of the Americas.” *PLoS Negl Trop Dis* 6: e1498. doi:10.1371/journal.pntd.0001498.

knowledge and limited dataset, should be regarded as theoretical insights. They remain pertinent tools with implications in term of vectors surveillance and control. For example, our *T. brasiliensis* prospective scenarios in Northeastern Brazil revealed a significant influence of land-use on vector spatial genetic structure. It suggested that in urbanized areas, where hosts are abundant, vector population gene flow would be reduced (higher *FST* between sampling points). While anthropogenic landscape disturbance proved to increase vector infection by *T. cruzi* [54], it is more important than ever to anticipate for the effect of future landscapes on vector dynamics and spatial genetic structure, in order to establish efficient management strategies. This study on *T. brasiliensis* and associated prospective scenarios suggest that control techniques should be examined in their social-ecological context, accounting for anthropogenic features to come.

Beyond our empirical study case, the objective of the paper was to present a methodological framework for studying vectors population genetics of pathogens which can integrate its biological, ecological and sociological components. Although various studies have assessed and described the emergence of zoonoses and vector-borne diseases as the result of social-ecological interactions (*e.g.*, [55,56]), this approach has been disregarded in population and landscape genetic simulations ([2,49], but see [57]). Integrating social-ecological interactions in landscape genetics remains a key challenge, especially when considering that vectors of pathogens are localized in areas subject to anthropogenic disturbance across scales [58,59]. At the interface across disciplines, this methodological framework allows the consideration of different types of knowledge and takes into account the causes of vector spatial genetic structure at multiple levels. Our hope is that our study may provide a testable and applicable modeling framework to a broad community of epidemiologists for formulating scenarios of landscape change and foresee their consequences on vector dynamics and genetic structure, with potential implications for their surveillance and control.

Supporting Information

Supporting Information S1 Material and methods for genotyping studies. (PDF)

Author Contributions

Conceived and designed the experiments: FR OD. Performed the experiments: FR. Analyzed the data: FR JC CEA JFS MH OD. Contributed reagents/materials/analysis tools: FR JC CEA JFS MH OD. Contributed to the writing of the manuscript: FR JC CEA JFS MH OD.

11. Hotez PJ, Molyneux DH, Fenwick A, Kumaresan J, Sachs SE, et al. (2007) Control of Neglected Tropical Diseases. *New England Journal of Medicine* 357: 1018–1027. doi:10.1056/NEJMra064142.
12. Sarquis O, Carvalho-Costa FA, Oliveira LS, Duarte R, D'Andrea PS, et al. (2010) Ecology of *Triatoma brasiliensis* in northeastern Brazil: seasonal distribution, feeding resources, and *Trypanosoma cruzi* infection in a sylvatic population. *Journal of Vector Ecology* 35: 385–394. doi:10.1111/j.1948-7134.2010.00097.x.
13. Soares RPP, Evangelista L das G, Laranja LS, Diotaiuti L (2000) Population dynamics and feeding behavior of *Triatoma brasiliensis* and *Triatoma pseudomaculata*, main vectors of Chagas disease in Northeastern Brazil. *Memórias do Instituto Oswaldo Cruz* 95: 151–155. doi:10.1590/S0074-0276200000200003.
14. Suarez-Davalos V, Dangles O, Villacis AG, Grijalva MJ (2010) Microdistribution of Sylvatic Triatomine Populations in Central-Coastal Ecuador. *Journal of Medical Entomology* 47: 80–88. doi:10.1603/033.047.0111.
15. Carbajal de la Fuente AL, Minoli SA, Lopes CM, Noireau F, Lazzari CR, et al. (2007) Flight dispersal of the Chagas disease vectors *Triatoma brasiliensis* and *Triatoma pseudomaculata* in northeastern Brazil. *Acta Tropica* 101: 115–119. doi:10.1016/j.actatropica.2006.12.007.
16. Harry M, Dupont L, Romana C, Demanche C, Mercier A, et al. (2008) Microsatellite markers in *Triatoma pseudomaculata* (Hemiptera, Reduviidae, Triatominae), Chagas' disease vector in Brazil. *Infect Genet Evol* 8: 672–675. doi:10.1016/j.meegid.2008.03.005.
17. Borges EC, Dujardin JP, Schofield CJ, Romanha AJ, Diotaiuti L (2000) Genetic Variability of *Triatoma brasiliensis* (Hemiptera: Reduviidae) Populations. *Journal of Medical Entomology* 37: 872–877. doi:10.1603/0022-2585-37.6.872.
18. Borges EC, Dujardin J-P, Schofield CJ, Romanha AJ, Diotaiuti L (2005) Dynamics between sylvatic, peridomestic and domestic populations of *Triatoma brasiliensis* (Hemiptera: Reduviidae) in Ceará State, Northeastern Brazil. *Acta Tropica* 93: 119–126. doi:10.1016/j.actatropica.2004.10.002.
19. Costa J, Almeida JR de, Brito C, Duarte R, Marchon-Silva V, et al. (1998) Ecotopes, Natural Infection and Trophic Resources of *Triatoma brasiliensis* (Hemiptera, Reduviidae, Triatominae). *Memórias do Instituto Oswaldo Cruz* 93: 7–13. doi:10.1590/S0074-02761998000100002.
20. Costa J, Almeida CE, Dotson EM, Lins A, Vinhaes M, et al. (2003) The epidemiologic importance of *Triatoma brasiliensis* as a Chagas disease vector in Brazil: a revision of domiciliary captures during 1993–1999. *Memórias do Instituto Oswaldo Cruz* 98: 443–449. doi:10.1590/S0074-02762003000400002.
21. Harry M, Dupont L, Quartier M, Diotaiuti L, Walter A, et al. (2009) New perspectives for population genetics of Chagas' disease vectors in the Northeastern Brazil: Isolation of polymorphic microsatellite markers in *Triatoma brasiliensis*. *Infection, Genetics and Evolution* 9: 633–637. doi:10.1016/j.meegid.2009.03.008.
22. Almeida CE, Pacheco RS, Haag K, Dupas S, Dotson EM, et al. (2008) Inferring from the Cyt B Gene the *Triatoma brasiliensis* Neiva, 1911 (Hemiptera: Reduviidae: Triatominae) Genetic Structure and Domiciliary Infestation in the State of Paraíba, Brazil. *Am J Trop Med Hyg* 78: 791–802.
23. Epperson BK, Mcrae BH, Scribner K, Cushman SA, Rosenberg MS, et al. (2010) Utility of computer simulations in landscape genetics. *Molecular Ecology* 19: 3549–3564. doi:10.1111/j.1365-294X.2010.04678.x.
24. Landguth EL, Cushman SA, Murphy M a., Luikart G (2010) Relationships between migration rates and landscape resistance assessed using individual-based simulations. *Molecular Ecology Resources* 10: 854–862. doi:10.1111/j.1755-0998.2010.02867.x.
25. Archie EA, Luikart G, Ezenwa VO (2009) Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology & Evolution* 24: 21–30. doi:10.1016/j.tree.2008.08.008.
26. Goldberg D (1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional. Available: <http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0201157675>. Accessed 20 February 2013.
27. Price WL (1983) Global optimization by controlled random search. *J Optim Theory Appl* 40: 333–348. doi:10.1007/BF00933504.
28. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564–567. doi:10.1111/j.1755-0998.2010.02847.x.
29. Schweigmann N, Vallvé S, Muscio O, Ghillini M, Alberti A, et al. (1988) Dispersal flight by *Triatoma infestans* in an arid area of Argentina. *Medical and Veterinary Entomology* 2: 401–404. doi:10.1111/j.1365-2915.1988.tb00215.x.
30. Castañera MB, Aparicio JP, Gürtler RE (2003) A stage-structured stochastic model of the population dynamics of *Triatoma infestans*, the main vector of Chagas disease. *Ecological Modelling* 162: 33–53. doi:10.1016/S0304-3800(02)00388-5.
31. Rabinovich JE, Himschoot P (1990) A population-dynamics simulation model of the main vectors of Chagas' Disease transmission, *Rhodnius prolixus* and *Triatoma infestans*. *Ecological Modelling* 52: 249–266. doi:10.1016/0304-3800(90)90019-D.
32. Beaumont MA, Nichols RA (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proc R Soc Lond B* 263: 1619–1626. doi:10.1098/rspb.1996.0237.
33. Oreskes N, Shrader-Frechette K, Belitz K (1994) Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science* 263: 641–646. doi:10.1126/science.263.5147.641.
34. Rykiel Jr EJ (1996) Testing ecological models: the meaning of validation. *Ecological Modelling* 90: 229–244. doi:10.1016/0304-3800(95)00152-2.
35. Wright S (1931) Evolution in Mendelian Populations. *Genetics* 16: 97–159.
36. Pearse DE, Crandall KA (2004) Beyond FST: Analysis of population genetic data for conservation. *Conservation Genetics* 5: 585–602. doi:10.1007/s10592-003-1863-4.
37. Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: FST|[ne]|1/(4Nm+1). *Heredity* 82: 117–125. doi:10.1038/sj.hdy.6884960.
38. Zurell D, Berger U, Cabral JS, Jeltsch F, Meynard CN, et al. (2010) The virtual ecologist approach: simulating data and observers. *Oikos* 119: 622–635. doi:10.1111/j.1600-0706.2009.18284.x.
39. R Core Team (2014) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Available: <http://www.R-project.org>.
40. Whittle P, Whittle P, Whittle P, Whittle P, Mathematician NZ, et al. (1963) Prediction and regulation by linear least-square methods. English Universities Press London.
41. Akima H (1970) A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. *J ACM* 17: 589–602. doi:10.1145/321607.321609.
42. Banks SC, Ling SD, Johnson CR, Piggott MP, Williamson JE, et al. (2010) Genetic structure of a recent climate change-driven range extension. *Molecular Ecology* 19: 2011–2024. doi:10.1111/j.1365-294X.2010.04627.x.
43. Hoban S, Bertorelle G, Gaggiotti OE (2012) Computer simulations: tools for population and evolutionary genetics. *Nature Reviews Genetics* 13: 110–122. doi:10.1038/nrg3130.
44. Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, et al. (2008) Climate Change, Deforestation, and the Fate of the Amazon. *Science* 319: 169–172. doi:10.1126/science.1146961.
45. Sierra R (2000) Dynamics and patterns of deforestation in the western Amazon: the Napo deforestation front, 1986–1996. *Applied Geography* 20: 1–16. doi:10.1016/S0143-6228(99)00014-4.
46. Hamilton M (2011) Population Genetics. John Wiley & Sons. 407 p.
47. Mathieson I, McVean G (2013) Estimating Selection Coefficients in Spatially Structured Populations from Time Series Data of Allele Frequencies. *Genetics* 193: 973–984. doi:10.1534/genetics.112.147611.
48. Barbu C, Dumonteil E, Gourbière S (2010) Characterization of the Dispersal of Non-Domiciliated *Triatoma dimidiata* through the Selection of Spatially Explicit Models. *PLoS Negl Trop Dis* 4: e777. doi:10.1371/journal.pntd.0000777.
49. Luz PM, Struchiner CJ, Galvani AP (2010) Modeling Transmission Dynamics and Control of Vector-Borne Neglected Tropical Diseases. *PLoS Negl Trop Dis* 4: e761. doi:10.1371/journal.pntd.0000761.
50. Bonabeau E (2002) Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences* 99: 7280–7287. doi:10.1073/pnas.082080899.
51. Minoli SA, Lazzari CR (2006) Take-off activity and orientation of triatomines (Heteroptera: Reduviidae) in relation to the presence of artificial lights. *Acta Tropica* 97: 324–330. doi:10.1016/j.actatropica.2005.12.005.
52. Pacheco-Tucuch FS, Ramirez-Sierra MJ, Gourbière S, Dumonteil E (2012) Public Street Lights Increase House Infestation by the Chagas Disease Vector *Triatoma dimidiata*. *PLoS ONE* 7: e36207. doi:10.1371/journal.pone.0036207#pone.0036207-CarbajaldelaFuente1.
53. Abrahan LB, Gorla DE, Catalá SS (2011) Dispersal of *Triatoma infestans* and other Triatominae species in the arid Chaco of Argentina: Flying, walking or passive carriage? The importance of walking females. *Memórias do Instituto Oswaldo Cruz* 106: 232–239. doi:10.1590/S0074-02762011000200019.
54. Gottdenker NL, Chaves LF, Calzada JE, Saldaña A, Carroll CR (2012) Host Life History Strategy, Species Diversity, and Habitat Influence *Trypanosoma cruzi* Vector Infection in Changing Landscapes. *PLoS Negl Trop Dis* 6: e1884. doi:10.1371/journal.pntd.0001884.
55. Gazzinelli A, Correa-Oliveira R, Yang G-J, Boatin BA, Kloos H (2012) A Research Agenda for Helminth Diseases of Humans: Social Ecology, Environmental Determinants, and Health Systems. *PLoS Negl Trop Dis* 6: e1603. doi:10.1371/journal.pntd.0001603.
56. Vorou RM, Papavassiliou VG, Tsiodras S (2007) Emerging zoonoses and vector-borne infections affecting humans in Europe. *Epidemiology & Infection* 135: 1231–1247. doi:10.1017/S0950268807008527.
57. Remais JV, Xiao N, Akullian A, Qiu D, Blair D (2011) Genetic Assignment Methods for Gaining Insight into the Management of Infectious Disease by Understanding Pathogen, Vector, and Host Movement. *PLoS Pathog* 7: e1002013. doi:10.1371/journal.ppat.1002013.
58. Hemme RR, Thomas CL, Chadee DD, Severson DW (2010) Influence of Urban Landscapes on Population Dynamics in a Short-Distance Migrant Mosquito: Evidence for the Dengue Vector *Aedes aegypti*. *PLoS Negl Trop Dis* 4: e634. doi:10.1371/journal.pntd.0000634.
59. Mills JN, Gage KL, Khan AS (2010) Potential Influence of Climate Change on Vector-Borne and Zoonotic Diseases: A Review and Proposed Research Plan. *Environ Health Perspect* 118: 1507–1514. doi:10.1289/ehp.0901389.