

## Fine-Scale Predictions of Distributions of Chagas Disease Vectors in the State of Guanajuato, Mexico

JORGE LÓPEZ-CÁRDENAS,<sup>1</sup> FRANCISCO ERNESTO GONZALEZ BRAVO,<sup>2</sup>  
PAZ MARÍA SALAZAR SCHETTINO,<sup>3</sup> JUAN CARLOS GALLAGA SOLORZANO,<sup>1</sup>  
ECTOR RAMÍREZ BARBA,<sup>4</sup> JOEL MARTINEZ MENDEZ,<sup>4</sup> V. SÁNCHEZ-CORDERO,<sup>5</sup>  
A. TOWNSEND PETERSON,<sup>6</sup> AND J. M. RAMSEY<sup>7</sup>

J. Med. Entomol. 42(6): 1068–1081 (2005)

**ABSTRACT** One of the most daunting challenges for Chagas disease surveillance and control in Mexico is the lack of community level data on vector distributions. Although many states now have assembled representative domestic triatomine collections, only two triatomine specimens had been collected and reported previously from the state of Guanajuato. Field personnel from the state's Secretaría de Salud conducted health promotion activities in 43 of the 46 counties in the state and received donations of a total of 2,522 triatomine specimens between 1998 and 2002. All specimens were identified, and live insects examined for *Trypanosoma cruzi*. In an effort to develop fine-scale distributional data for Guanajuato, collection localities were georeferenced and ecological niches were modeled for each species by using evolutionary-computing approaches. Five species were collected: *Triatoma mexicana* (Herrick-Schaeffer), *Triatoma longipennis* (Usinger), *Triatoma pallidipennis* (Stål), *Triatoma barberi* (Usinger), and *Triatoma dimidiata* (Latreille) from 201 communities located at elevations of 870–2,200 m. Based on collection success, *T. mexicana* had the broadest dispersion, although niche mapping indicates that *T. barberi* represents the greatest risk for transmission of Chagas disease in the state. *T. dimidiata* was represented in collections by a single adult collected from one village outside the predicted area for all species. For humans, an estimated 3,755,380 individuals are at risk for vector transmission in the state, with an incidence of 3,500 new cases per year; overall seroprevalences of 2.6% indicate that 97,640 individuals are infected with *T. cruzi* at present, including 29,300 chronic cases.

**KEY WORDS** Chagas disease, Guanajuato, Mexico, triatomine, vector distribution

EIGHTEEN TRIATOMINE (Triatominae: Reduviidae) species have been documented as vectors of *Trypanosoma cruzi* in Mexico (Zárate and Zárate 1985, Vidal-Acosta et al. 2000). Nine of these species are considered as primary vector species (Ramsey et al. 2003), and most belong to two species complexes (*phyllosoma* and *protracta*) in the genus *Triatoma* (Lent and Wygod-

zinsky 1979). Multiple triatomine species are frequently found within a single region or state in Mexico. As a result, >71% of the Mexican population is at direct risk of acquiring *T. cruzi* infections from triatomines, and >96% of all *T. cruzi* transmission occurs via these vectors (Ramsey et al. 2003).

One of the most significant challenges for Chagas disease surveillance and control in Mexico is the lack of fine-scale (at the level of local communities) vector distribution data, as well as the almost complete ignorance of the disease-transmission characteristics and risk areas, both by the public health community and by at-risk populations. With improved malaria control in the country, in the last decade, “additional” operative vector control activities have focused primarily in urban environments owing to increases in classical and hemorrhagic dengue transmission. (To date, vector-borne disease control activities are budgeted only for malaria in the country.) Moreover, the recent decentralization of primary health care services to the state level has emphasized the need and the opportunity to review local and state health care priorities, and despite overburdened and shrinking

<sup>1</sup> Guanajuato State Public Health Laboratory, Medical Entomology Laboratory Secretaría de Salud del Estado de Guanajuato. Calle Beta 204 Fracc. Delta, León, Guanajuato 37320, México.

<sup>2</sup> Universidad de Guanajuato, Facultad de Enfermería y Obstetricia en León, Aquiles Serdán 914, col Obregón, León, Guanajuato 37320, México.

<sup>3</sup> Medical School, Universidad Nacional Autónoma de México. Ciudad Universitaria, México D.F. 04510, México.

<sup>4</sup> Instituto de Salud Pública del Estado de Guanajuato, Tamazúca 4 Centro, Guanajuato, Guanajuato 36000, México.

<sup>5</sup> Instituto de Biología, Universidad Nacional Autónoma de México, Apartado Postal 70-153, México, D.F. 04510 México.

<sup>6</sup> Natural History Museum, The University of Kansas, Lawrence, KS 66045.

<sup>7</sup> Centro de Investigaciones sobre Enfermedades Infectuosas (CISEI), Instituto Nacional de Salud Pública (INSP), Cuernavaca, Morelos, 62508 México.











(number of positive villages per total villages examined  $\times 100$ ), were calculated for communities and counties.

**Ecological Niche Modeling.** Our use of ecological niche modeling, which followed on several previous tests and prototypes (Peterson 2001, Beard et al. 2002, Costa et al. 2002, Peterson et al. 2002), introduces a new feature to the technique: the use of multitemporal, remotely sensed environmental data sets as a surrogate for climatic data (Rogers 2000, Egbert et al. 2002, Hay et al. 2002, Rogers et al. 2002). This environmental data set permits fine-scale predictions across landscapes for which climatic monitoring stations are too sparse to permit development of fine-scale climate maps.

Preliminary ecological niche models were based on unique occurrence records for all species encountered in the state, by using all known existing collection records (i.e., not part of the Guanajuato state sampling program described above) from 1948 to 2002 for the area within 250 km of Guanajuato. Three members of the *phylllosoma* species group and one from the *protracta* complex had sufficient unique occurrence localities ( $>30$  records) for analysis. All occurrence points were georeferenced to the nearest 0.1' of latitude and longitude via reference to Mexican census information ([www.conabio.gob.mx](http://www.conabio.gob.mx), [www.inegi.gob.mx](http://www.inegi.gob.mx)). The Guanajuato state health survey was reserved as an independent test and validation of model predictivity.

Ecological niches were modeled using the Genetic Algorithm for Rule-set Prediction (GARP) (Stockwell and Noble 1992, Stockwell 1999, Stockwell and Peters 1999), an evolutionary-computing software package available for public download. In general, the procedure focuses on modeling ecological niches (the conjunction of ecological conditions within which a species is able to maintain populations without immigration) (Grinnell 1917). Specifically, GARP relates ecological characteristics of known occurrence points to those of points randomly sampled from the rest of the study region, seeking to develop a series of decision rules that best summarize those factors associated with the species' presence (Peterson et al. 2002a). In GARP, input occurrence points are divided randomly and evenly into training and (intrinsic) testing data sets. GARP works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection: a method is chosen from a set of possibilities (e.g., logistic regression, bioclimatic rules), applied to the training data, and a rule is developed or evolved. Predictive accuracy is then evaluated based on 1,250 points resampled from the intrinsic test data and 1,250 points sampled randomly from the study region as a whole. Rules may evolve by a number of means that mimic DNA evolution: point mutations, deletions, and crossing over. The change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the model, and the algorithm runs either 1000 iterations or until convergence.

We used four data layers summarizing aspects of topography (elevation, slope, aspect, and topographic index [tendency to pool water]) from the U.S. Geological Survey's Hydro-1K data set, and 12 data layers summarizing monthly composite normalized difference vegetation index values ("greenness index") drawn from the Advanced Very High Resolution Radiometer satellite. Both data sets have a native spatial resolution of  $\approx 1$  km and were resolved to the same 1-km grid across the region of study. GARP's predictive abilities have been tested and proven under diverse circumstances (Peterson et al. 1999, 2001, 2002c; Peterson 2001; Anderson et al. 2002, 2003).

To take advantage of the random-walk nature of the GARP algorithm, we developed 100 replicate models of each species' ecological niche. We used the procedures described in Anderson et al. (2003) for choosing a "best subset" of the 100 models, based on optimal combinations of error statistics. The procedure is based on the observations that 1) models vary in quality, 2) variation among models involves an inverse relationship between error of omission (leaving out true distributional area) and commission (including areas not actually inhabited), and 3) best models (as judged by experts blind to error statistics) are clustered in a region of minimum omission of independent test points and moderate area predicted (an axis related directly to commission error). The relative position of the cloud of points relative to the two error axes provides an assessment of the relative accuracy of each model. To choose best subsets of models, we 1) eliminated all models that had omission error  $>5\%$  based on independent intrinsic test points, 2) calculated the median area predicted present among these zero-omission points, 3) identified the 10 models closest to the overall median area predicted, and 4) summed these best subsets models.

Projection of the rule-sets for these best subsets models onto geography provided distributional predictions for species' ranges. Model quality was tested and validated via the independent Guanajuato state triatomine survey data.  $\chi^2$  tests were used to compare observed success in predicting distributions of test points with those expected under random models (proportional area predicted present provides an estimate of occurrence points correctly predicted were the prediction to be random with respect to the distribution of the test points). Predicted presence was defined as the area in which all best-subsets models agreed in prediction of presence.

**Disease Prevalence and Burden.** Based on the areas identified by the ecological niche models at different levels of model agreement, human settlements in the state were ranked from lowest (1) to highest (10) for vector transmission risk. Three strata were defined based on this ranking: 1–3, low risk; 4–7, medium risk; and 8–10, high risk. Community, demographic, and housing information from the 2000 National Census (INEGI 2000) were combined with the risk statistics, and total population at risk, seropositive population, incidence, mortality, and disease burden were calculated by county based on an overall





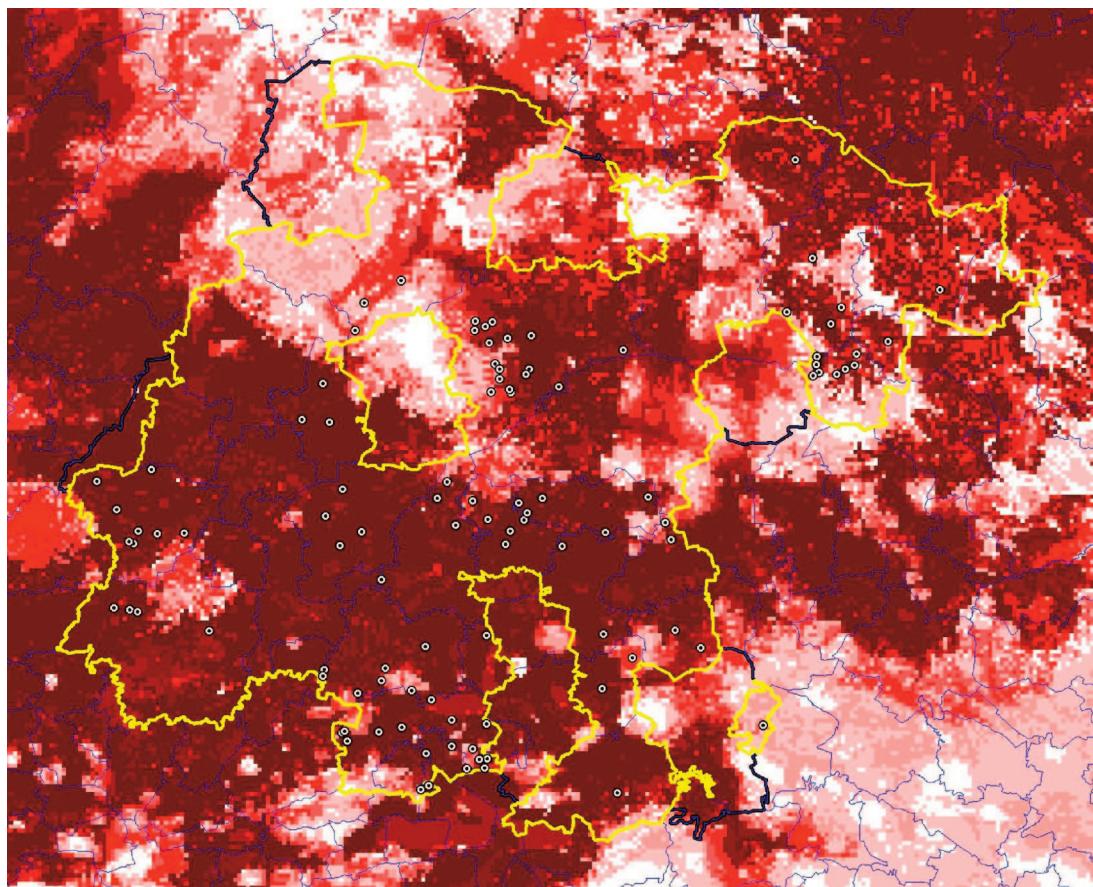


Fig. 1. Niche prediction area for *T. barberi* in the state of Guanajuato. Highest risk is indicated in dark red and lowest in white; collection sites are marked by open circles in the map.

population to calculate overall prevalence, and given a recent serological survey that found a seroprevalence of 2.6% for anti-*T. cruzi* antibody in the state, an estimated 97,640 inhabitants have Chagas disease in the state, of which 29,293 would currently be in the chronic phase (30% of all cases). Incidence for Chagas cases is estimated as 3,485 per year, in the absence of any vector control measures. With 29,293 current chronic cases, we estimate that the state and/or the population spends an average U.S.\$20,505,100 in treatment alone, independent of whether the cases are recognized as Chagas, or as "cardiac insufficiency," where most of the current cases in Mexico are classified.

Table 5. Infection of triatomines with *T. cruzi*

Species	Examined	Positive <i>T. cruzi</i>	<i>T. cruzi</i> infection (%)
<i>T. barberi</i>	163	31	19.0
<i>T. longipennis</i>	22	10	45.5
<i>T. pallidipennis</i>	13	1	7.7
<i>T. mexicana</i>	1009	29	2.9

## Discussion

Four triatomine vector species of Chagas disease have now been identified and their distributions modeled in the state of Guanajuato: *T. barberi*, *T. longipennis*, *T. mexicana*, and *T. pallidipennis*. All of these species are well adapted to domestic habitats, with colonization both inside and outside houses, and their *T. cruzi* transmission potential verified, with infection rates lowest for *T. mexicana* (6%) and highest for *T. longipennis* (46%). Despite collection bias, a result of more intensive health promotion and survey activities in *T. mexicana*-infested communities (highest collection success), this species has similar albeit slightly lower potential risk compared with *T. barberi* by using niche modeling. Both species are the most widely dispersed triatomines within the state and may infest all communities in certain counties. Neither *T. longipennis* nor *T. pallidipennis* were collected from >10% of communities within given counties. These preliminary data will require verification through entomological surveys by using standardized methods and trained vector technicians.

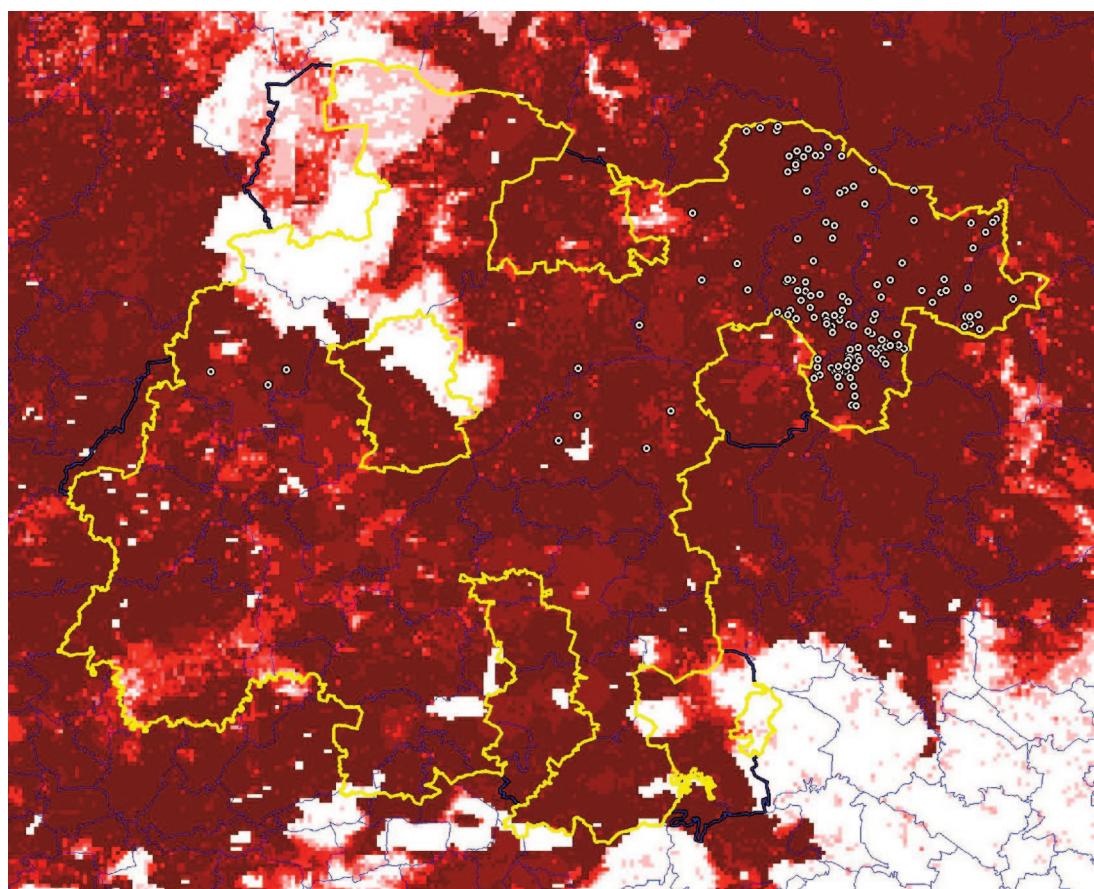


Fig. 2. Niche prediction area for *T. mexicana* in the state of Guanajuato. Highest risk is indicated in dark red and lowest in white; collection sites are marked by open circles in the map.

Most Mexican states are currently attempting to determine the distributions of their triatomine vectors to define at-risk populations for Chagas disease transmission (Ramsey et al. 2003) and to evaluate requirements for surveillance and control programs. This task is daunting, given the extent of the territory, and the lack of funds and personnel with which to conduct such surveys—indeed, it is clear that without local assistance and without more accessible methodology to measure vector distributions, diagnosis of the situation will remain obscure. This study succeeded in initiating a long-needed awareness program within the state for triatomines and Chagas disease, as well as intensive collection of vectors by local residents and vector-borne disease technicians. Nevertheless, it is clear that these collections must be supervised and supplemented with specific entomological surveys, at least in the initial stages of a surveillance program.

This study also presents a powerful example of the role that ecological niche modeling can play in this challenge (Peterson et al. 2002b). Niche-based predictions for species' distributions can be powerful tools with which health care services can outline vec-

tor species' distributional areas, establish potential disease transmission risk areas, and determine real versus potential infestation. In our analysis, triatomine collection data from states within 250 km of Guanajuato coincided precisely with the local residents' collection data, and ecological niche models developed from the two data sets were virtually identical. In other words, using just the existing data from *outside* the state, we predicted the triatomine distributions with accuracy, thereby obviating the need for preliminary surveys. Such predictions could be used to design and guide on-ground survey efforts, thereby focusing effort considerably, and reducing costs significantly. These models can then be refined via secondary surveys conducted by trained technicians in sentinel counties, so that specific entomological indicators can be evaluated.

The case of the single individual of *T. dimidiata* collected is an important example of the potential use of ecological niche models to define risk. Using data sets from states surrounding Guanajuato, we conclude that this species has virtually no potential distributional area within the state. Collection of a lone adult

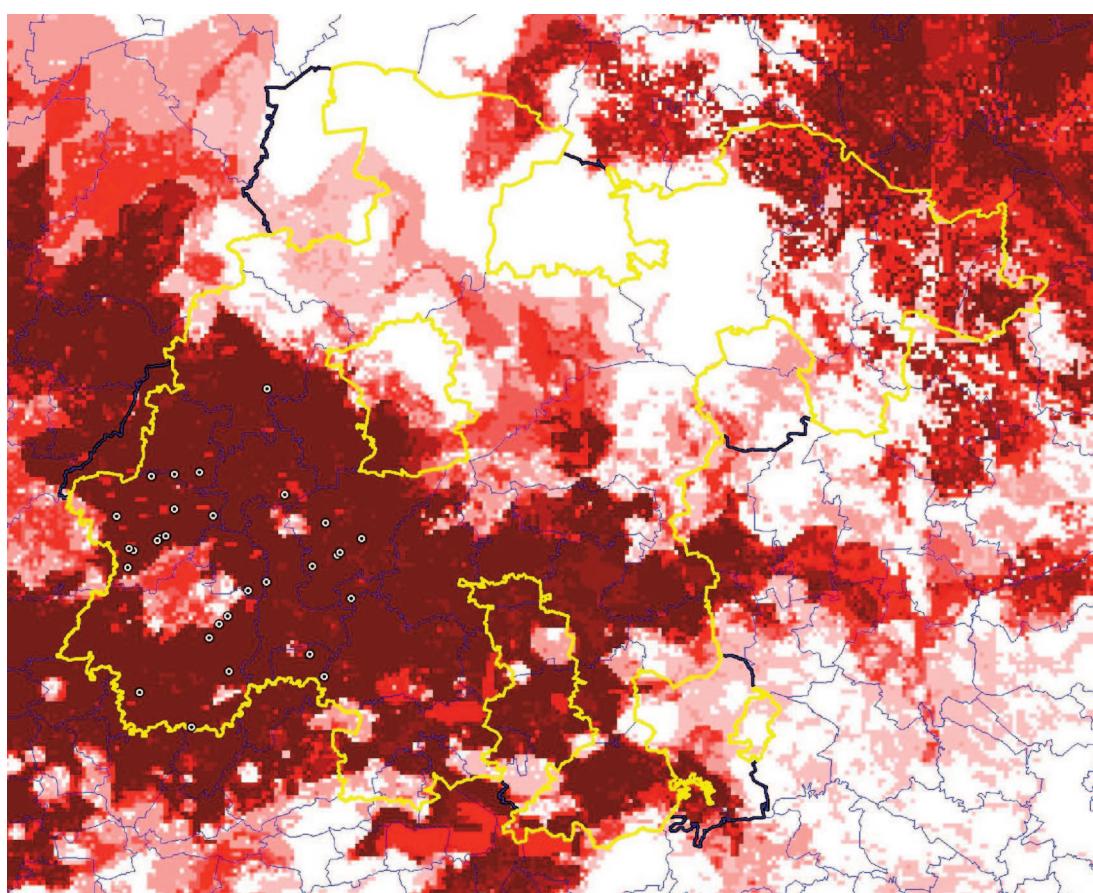


Fig. 3. Niche prediction area for *T. longipennis* in the state of Guanajuato. Highest risk is indicated in dark red and lowest in white; collection sites are marked by open circles in the map.

specimen can therefore be considered spurious, probably the result of human-assisted transport. Nevertheless, the lack of potential niche areas suggests that even though this species is transported periodically to the state, it probably cannot establish in sylvan habitats.

Chagas disease transmission risk is a function of the presence of triatomines in domestic habitats, which in turn is dependent on the presence of triatomines in surrounding sylvan areas. Niche modeling provides the only tool available to date for defining potential

dispersion of individual triatomine species and hence potential transmission risk.

The use of prediction models for Chagas surveillance programs is currently feasible at both local and statewide levels. Control programs can use these models to determine at-risk populations, to monitor control program activities and to allow control strategies to be stratified according to real or potential risk (Ramsey et al. 2003). Potential infestation areas will require specific surveillance activities, different from actual infestation areas, and information processing from population-based collections or specific entomological surveys can feed the models so that they can accurately and rapidly track program needs. Most importantly, as demonstrated in this study, they can identify preliminary risk areas so that vector control technicians can direct their surveys cost-effectively for more precise and local infestation indicators.

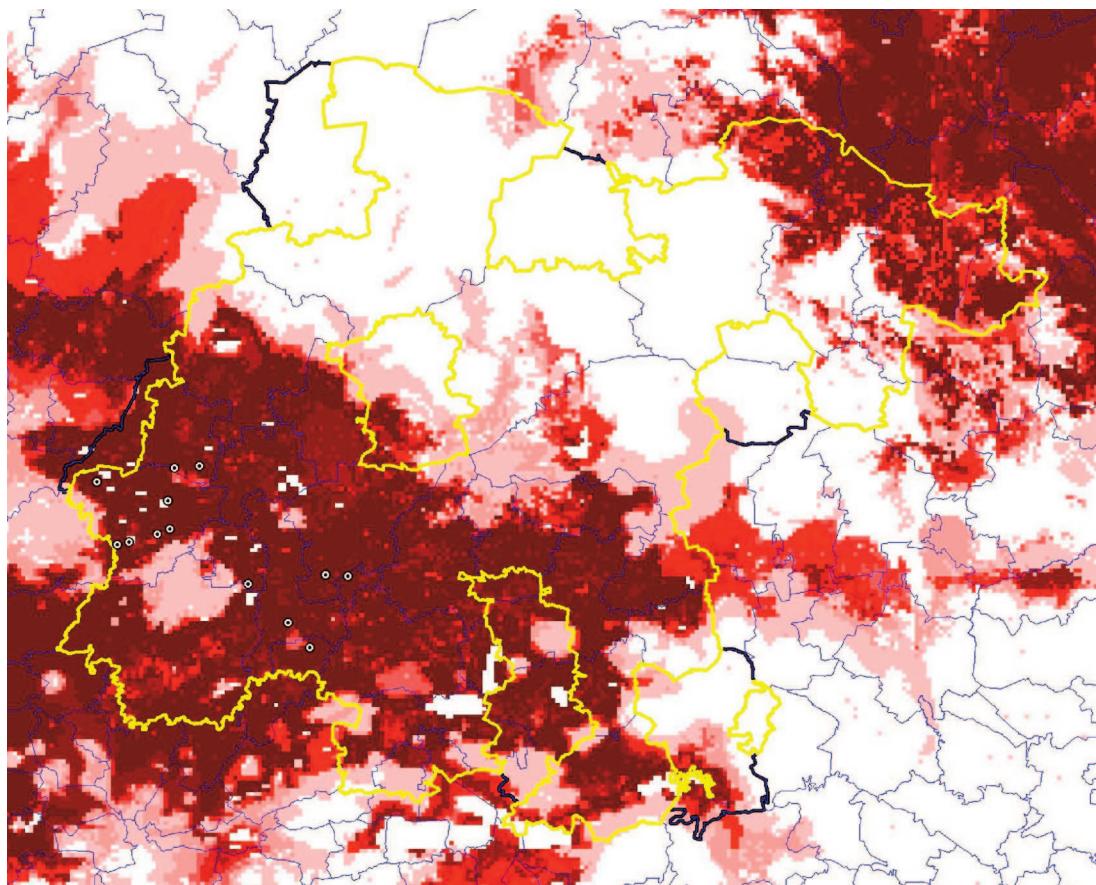
Table 6. Predicted risk (high, score 15–20; medium, score 8–14; low, score 1–7) for vector transmission for all communities in Guanajuato by using GARP, and stratified by vector species

Species	High(%)	Medium(%)	Low(%)	Total
<i>T. barberi</i>	8,299 (96.9)	221 (2.6)	41 (0.5)	8,561
<i>T. longipennis</i>	5,132 (72.9)	601 (8.5)	1,307 (18.6)	7,040
<i>T. pallidipennis</i>	5,272 (65.9)	800 (10.0)	1,922 (24.1)	7,994
<i>T. mexicana</i>	7,504 (88.9)	497 (5.9)	436 (5.2)	8,437

Note that sympatry and niche overlap among species results in repetition of at-risk communities.

#### Acknowledgments

We thank the invaluable collaboration of all personnel from the primary health care services in the state of Guana-



**Fig. 4.** Niche prediction area for *T. pallidipennis* in the state of Guanajuato. Highest risk is indicated in dark red and lowest in white; collection sites are marked by open circles in the map.

juato, especially the state and Sanitary Jurisdiction epidemiologists and technical personnel from the vector control program. We thank Vianey Vidal from the National Laboratory for Diagnosis and Epidemiological Reference for assistance with quality control of taxonomic identifications and Gely Peña Arguntia for dedication in health promotion and county collection activities. We also thank the Vector Control Coordinators from the states of Mexico, Morelos, San Luis Potosí, and Veracruz for provision of data regarding *Triatoma* distributions not reported in the scientific literature and to the Comisión Nacional para la Conservación y Uso de la Biodiversidad (CONABIO) for access to mammal distributional data. This study was financed through the Guanajuato State Health budget, federal project funds from the National Institute of Public Health Mexico to J.M.R., and by the U.S. National Science Foundation and Consejo Nacional para Ciencia y Tecnología to A.T.P.

#### References Cited

- Anderson, R. P., M. Laverde, and A. T. Peterson. 2002. Geographical distributions of spiny pocket mice in South America: insights from predictive models. *Global Ecol. Biogeogr.* 11: 131–141.
- Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Using intermodel variation in error components to select best subsets of ecological niche models. *Ecol. Model.* 162: 211–232.
- Beard, C. B., G. Pye, F. J. Steurer, Y. Salinas, R. Campman, A. T. Peterson, J. M. Ramsey, R. A. Wirtz, and L. E. Robinson. 2002. Chagas disease in a domestic transmission cycle in southern Texas, USA. *Emerg. Infect. Dis.* 9: 103–105.
- Biagi, F. 1956. Nueva Infección natural con *Schizotrypanum cruzi*. Variabilidad cromática y otras notas sobre triatomas Mexicanas. *Prensa Med. Mex.* 21: 123–126.
- Costa, J., A. T. Peterson, and C. B. Beard. 2002. Ecological niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae). *Am. J. Trop. Med. Hyg.* 67: 516–520.
- Egbert, S. L., E. Martinez-Meyer, M. A. Ortega-Herra, and A. T. Peterson. 2002. Use of datasets derived from time-series AVHRR imagery as surrogates for land cover maps in predicting species' distributions. *Proceedings IEEE 2002. Int. Geo. Rem. Sen. Symp. (IGARSS)* 4: 2337–2339.
- Grinnell, J. 1917. Field tests of theories concerning distributional control. *Am. Nat.* 51: 115–128.
- Hay, S. I., M. F. Myers, N. Maynard, and D. J. Rogers. 2002. From remote sensing to relevant sensing in human health. *Photogrammetric Eng. Remote Sensing* 68: 109–111.
- Lent, H., and P. Wygodzinsky. 1979. Revision of the Triatominae (Hemiptera: Reduviidae), and their significance as vectors of Chagas disease. *Bull. Amer. Mus. Nat. Hist.* 163: 520.

- López-Cárdenas, J. 2001. Identificación y distribución espacial de las especies de vectores de la enfermedad de Chagas en el Estado de Guanajuato, de 1998–2000. M.S. thesis, Universidad de Guanajuato, León, Guanajuato, México.
- Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103: 599–605.
- Peterson, A. T., and D. A. Viegialis. 2001. Predicting species invasions using ecological niche modeling. BioScience 51: 363–371.
- Peterson, A. T., J. Soberon, and V. Sanchez-Cordero. 1999. Conservation of ecological niches in evolutionary time. Science (Wash. DC) 285: 1265–1267.
- Peterson, A. T., V. Sanchez-Cordero, J. Soberon, J. Bartley, R. H. Buddemeier, and A. G. Navarro-Siguenza. 2001. Effects of global climate change on geographic distributions of Mexican Cracidae. Ecol. Model. 144: 21–30.
- Peterson, A. T., M. A. Ortega-Huerta, J. Bartley, V. Sanchez-Cordero, J. Soberon, R. H. Buddemeier, and D.R.B. Stockwell. 2002a. Future projections for Mexican faunas under global climate change scenarios. Nature (Lond.) 416: 626–629.
- Peterson, A. T., V. Sanchez-Cordero, C. B. Beard, and J. M. Ramsey. 2002b. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. Emerg. Infect. Dis. 8: 662–667.
- Peterson, A. T., D.R.B. Stockwell, and D. A. Kluza. 2002c. Distributional prediction based on ecological niche modeling of primary occurrence data, pp. 617–623. In J. M. Scott [ed.], Predicting species occurrences: Issues of scale and accuracy. Island Press, Washington, DC.
- Peterson, A. T., D. A. Viegialis, and J. Andreasen. 2005. Migratory birds as critical transport vectors for West Nile Virus in North America. Emerg. Inf. Dis. (in press).
- Ramsey, J. M., R. Ordoñez, A. Tello Lopez, J. L. Pohls, V. Sanchez-Cordero, and A. T. Peterson. 2003. Actualidades sobre la epidemiología de la enfermedad de Chagas en México, pp. 85–105. In J. M. Ramsey, A. Tello Lopez, and J. L. Pohls [eds.], Iniciativa para la vigilancia y control de la enfermedad de Chagas en México. Instituto Nacional de Salud Pública, México.
- Rogers, D. J. 2000. Satellites, space, time and the African trypanosomiasis, pp. 130–171. In S. I. Hay, S. Randolph, and D. J. Rogers [eds.], Remote sensing and GIS in public health. Academic, New York.
- Rogers, D. J., S. E. Randolph, R. W. Snow, and S. I. Hay. 2002. Satellite imagery in the study and forecast of malaria. Nature (Lond.) 415: 710–715.
- Salazar-Schettino, P. M., I. de Haro, J. Jiménez, and E. García Carrera. 1983. Dos nuevas localidades de transmisores de la enfermedad de Chagas en la República Mexicana. Salud Pública Méx. 25: 77–82.
- Stockwell, D.R.B. 1999. Genetic algorithms II, pp. 123–144. In A. H. Fielding [ed.], Machine learning methods for ecological applications. Kluwer Academic Publishers, Boston, MA.
- Stockwell, D.R.B., and I. R. Noble. 1992. Induction of sets of rules from animal distribution data: a robust and informative method of analysis. Math. Comp. Simul. 33: 385–390.
- Stockwell, D.R.B., and D. P. Peters. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. Int. J. Geogr. Inf. Sys. 13: 143–158.
- Vidal-Acosta, V., S. Ibáñez-Bernal, and C. Martínez-Campos. 2000. Infección Natural de chinches Triatominae con *Trypanosoma cruzi* asociadas a la vivienda humana en México. Salud Pública de Mex. 42: 496–503.
- Zárate, L. G., and R. J. Zárate. 1985. A checklist of the Triatominae (Hemiptera: Reduviidae). Int. J. Entomol. 27: 102–127.

Received 19 March 2004; accepted 25 January 2005.