

Preplanned Studies

Modelling the Ecological Suitability of the Chagas Disease Vector *Triatoma rubrofasciata* — China

Yunhai Guo^{1,✉}; Yunliang Shi^{2,✉}; Yi Zhang¹; Yuanyuan Li¹; Mengxia Wang¹; Qin Liu^{1,✉}

Summary

What is already known on this topic?

Triatoma rubrofasciata is a potential vector that can transmit American trypanosomiasis and was widely recorded in South of China.

What is added by this report?

Because of the low density of the triatomines, more habitats have not been discovered. This study mainly focused on predicting the geographical distribution of *T. rubrofasciata* under current and future climatic conditions in China using the MaxEnt model.

What are the implications for public health practice?

The result showed that the distribution of *T. rubrofasciata* was largely affected by annual mean temperature and possessed a high potential for expansion in southern China in the future. Our predictions are useful for targeting surveillance efforts in high-risk areas and increasing the efficiency and accuracy of public health investigations and vector control efforts in China.

Chagas disease, caused by *Trypanosoma cruzi*, is considered the fourth most transmitted disease after malaria, tuberculosis, and schistosomiasis by World Health Organization (WHO) (1). Chagas disease has become a global health issue in recent decades due to the spreading worldwide (2). Although there has not yet been any recorded cases of Chagas disease in China, one of the potential vectors, *Triatoma rubrofasciata* that can transmit *T. cruzi*, was widely recorded in southern China (3). In this study, the maximum entropy method (MaxEnt) for species distribution modeling (SDM) was used to predict the geographical distribution of *T. rubrofasciata*, under current and future climatic conditions (BCC-CSM-1 RCP26) in China. A total of 184 locations with the species occurrence were recorded; 19 bioclimatic variables derived from monthly data were used in the modeling process. The results showed that the distribution of *T. rubrofasciata* was largely affected by annual mean

temperature and minimum temperature of the coldest month. According to species response curves, this species preferred habitats with annual mean temperatures of 22.5 °C. Climate change projections indicated that by 2050 and 2070, the habitats of the *T. rubrofasciata* would be expanded.

T. rubrofasciata is one of the potential vectors of Chagas disease that is recorded to be the most widely distributed worldwide. In China, we started monitoring triatomines and imported Chagas disease in 2016. During a survey from 2016–2018, 170 habitats were recorded (3), but because of low density of the triatomines, many habitats were likely underrepresented. Therefore, predicting these habitats through appropriate environmental data is important to make up for this underrepresentation.

MaxEnt is a machine learning model that estimates a target probability distribution by calculating the probability distribution of maximum entropy (4). MaxEnt has been used extensively to model the distribution of several vector-reliant disease-causing pathogens (5–6). Many scholars believed that the MaxEnt model was better than other models in predicting the spatial distribution of species, especially in cases of incomplete distribution data (6–7). MaxEnt is also capable of projecting shifts in species distribution under various climate change scenarios (4). In this study, MaxEnt software was used to predict the potential distribution of *T. rubrofasciata*. A total of 184 records have been collected in the database. These came from two sources: 1) published records on *T. rubrofasciata* in journals, books, and reports and from Shanghai Insect Museum and Museum of Insects Institute of Guangdong Province in China (8–11); and 2) specimen field surveys by the authors from 2016 (3). After screening duplicates and poor data, 116 records were obtained for the construction of the final model.

For present environmental data (1950–2000), a total of 19 bioclimatic layers were downloaded from Worldclim database (www.worldclim.org) (Supplementary Table S1, available in <http://weekly.chinacdc.cn/>). All data used for SDM had a spatial

resolution of 1 km² (30 arc seconds). The data of distribution points and environmental variables were inputted into MaxEnt software. The maximum number of iterations was set to 2000, bootstrap repeated 10 times, and random seed was turned on. Then, the response curves and jackknife test functions were checked to analyze the importance of environmental variables that affected the distribution of *T. rubrofasciata*. The establishment of the model needs to randomly select 75% of the distribution points, and the test of the model is completed by data of the remaining 25% of the occurrence points to construct an optimal habitat model of *T. rubrofasciata*. The prediction results of MaxEnt model were tested by the area under the receiver operating characteristic curve (ROC). Maximum entropy modeling was used with the MaxEnt software (version 3.3.3k, American Museum of Natural History, NY, USA) for quantifying relative risk of invasion and mapping of the potential geographic distribution of *T. rubrofasciata* in China. A jackknife procedure was used to calculate the significance of the contribution of each variable to the model. The area under the receiver operating characteristic curve [area under curve (AUC)] was used to evaluate model performance. The potential range of changes in the suitable areas under different scenarios

of climate change in the future was analyzed by using the “distribution changes between binary SDMs” tool. When performing the quick recursive to binary conversion, the threshold value was 0.5. ArcMap was loaded for all maps..

The MaxEnt model prediction has high AUC (0.989) indicating good model performance for *T. rubrofasciata* (Supplementary Figure S1, available in <http://weekly.chinacdc.cn/>). The jackknife test showed that the distribution of *T. rubrofasciata* was largely affected by annual mean temperature (Bio1), mean temperature of the coldest quarter (Bio11), and minimum temperature of coldest month (Bio6) (Figure 1). Annual mean temperature (Bio1) explained 63.4% of the total variance and was thus identified as the main factor affecting the spatial distribution of *T. rubrofasciata*. Species distribution maps showed that Guangdong, Guangxi, Fujian, and Taiwan, China of the study area were recognized as high potential habitats of *T. rubrofasciata* in the present conditions, by 2050, and by 2070 (Figure 2). The averaged future predictions of MaxEnt for 2050 and 2070 (binary SDMs) revealed an expansion in suitable habitats of *T. rubrofasciata* of 180,433 km² and 167,495 km² respectively, which reached more northern latitudes of southern China (Figure 3).

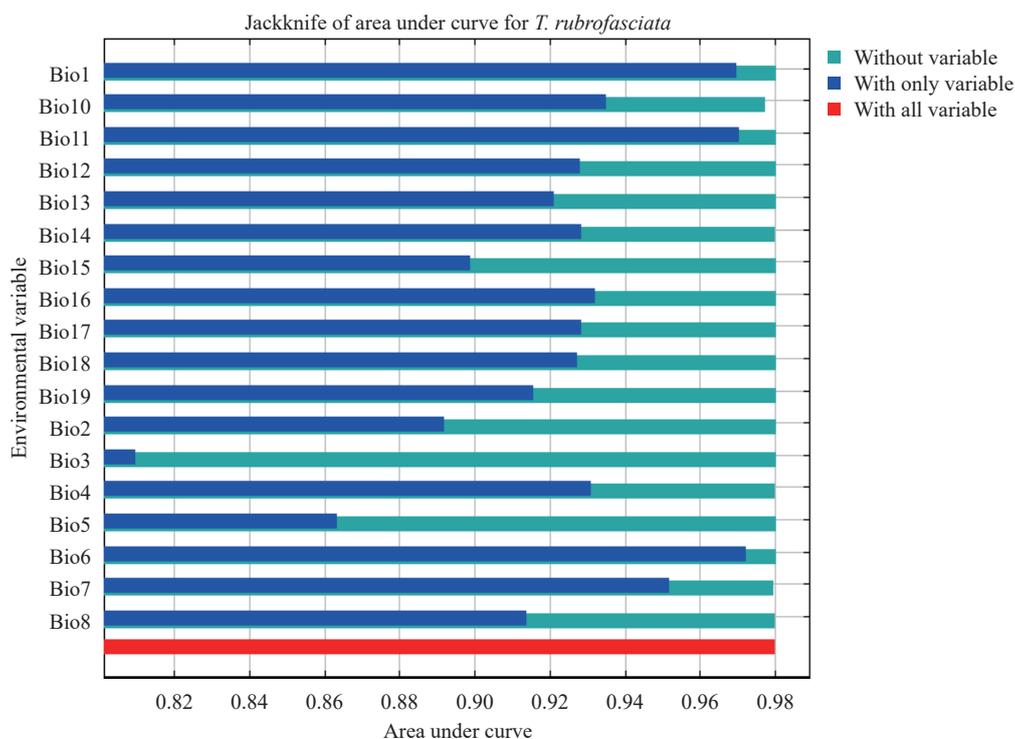


FIGURE 1. Relative predictive power of different environmental variables based on the jackknife of regularized training gain in maximum entry modeling for *T. rubrofasciata* in China.

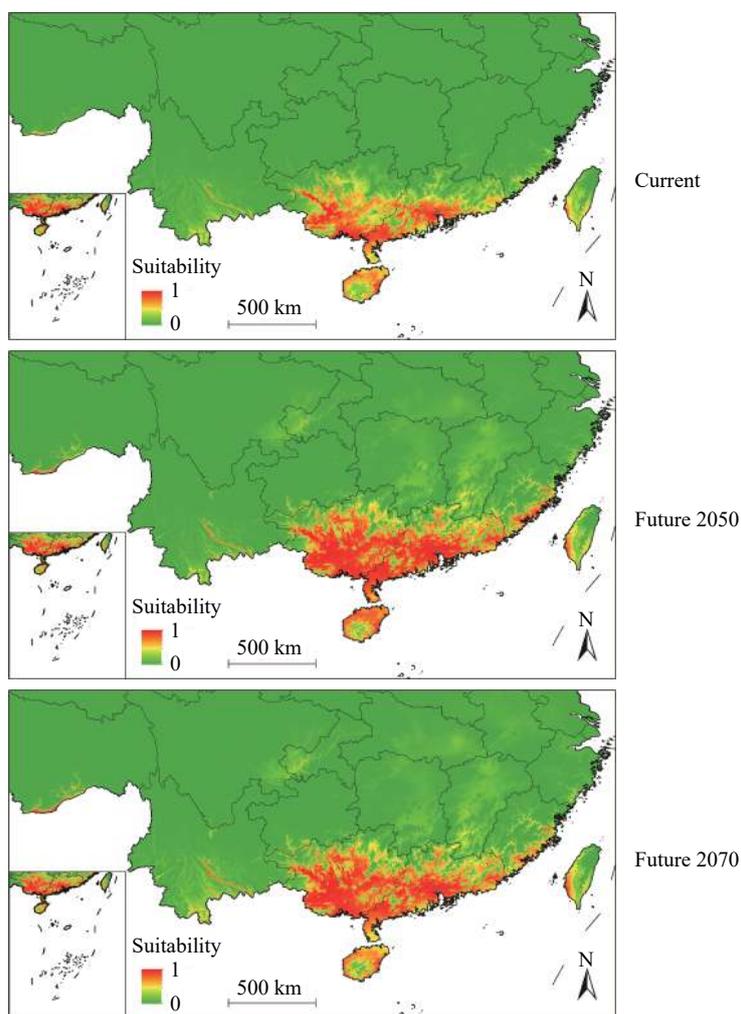


FIGURE 2. Distribution maps of *T. rubrofasciata* under current and future climatic change condition (2050 and 2070) based on A2a/HadCM3 scenario.

DISCUSSION

T. rubrofasciata, which was recorded as a potential vector that can transmit American trypanosomiasis, was the most widely distributed worldwide. During the survey of 2016–2018, the habitats of *T. rubrofasciata* was widely recorded in southern China, as well as exhibiting domestic and peri-domestic behavior invading chicken coops and human dwellings. However, due to its low density, more habitats would likely be underrepresented, including, for example, historical reports in Yunnan Province that were not corroborated in current surveys (2). For this purpose, we used MaxEnt to identify regions that offer climatically suitable conditions for this species.

The discriminatory capacity of the model displayed good predictive performance, which was also reflected in the AUC values of over 0.95. Studies have indicated

that potential hotspots of triatomine species diversity in tropical and subtropical regions between 21°N and 24°S latitude and a major limiting factor of triatomine distribution was temperature seasonality (12–13). In our study, after removing auto-correlated parameters (Figure 1), MaxEnt indicated that the current distribution of *T. rubrofasciata* was more affected by temperature variables (Bio1, Bio6, and Bio11). Among them, Bio1 (annual mean temperature) was the most affected and explained 63.4% of the total variance. The International Panel of Climate Change (IPCC) has predicted that annual mean temperatures will increase up to 5.8 °C by the end of this century (14). As the global climate warms, the risk of *T. rubrofasciata* expansion may increase.

The information obtained from the predictive maps of the current distribution of *T. rubrofasciata* was compared to those assembled from the series of future predictions in 2050 and 2070 with a medium GHG

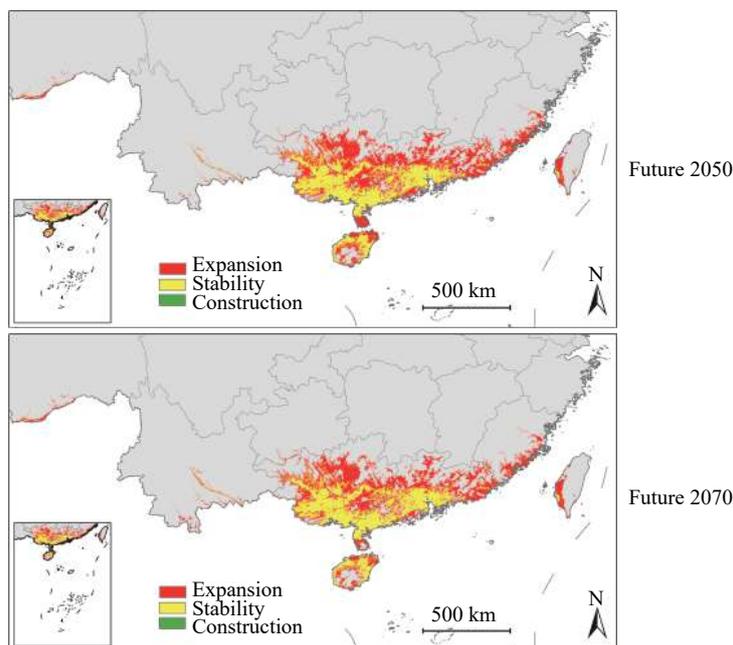


FIGURE 3. Distribution maps of *T. rubrofasciata* under future climatic change condition (2050 and 2070) based on binary species distribution modelling (SDMs).

concentration scenario (RCP26). Global climate models trained on the existing potential distribution showed a relatively stable estimate for the amount of land area that was classified as suitable for *T. rubrofasciata*. The two emissions scenarios we modeled (RCP 26) produced very similar predicted distributions, although the pathway of high concentrations of GHG predicted slightly less overall area (Figure 2). There was a predicted shift in the distribution with suitable areas moving from lower latitude and presumably warmer climates to climates at higher latitude where conditions may become more suitable (Figure 2). Regions of high predicted probability in 2050 were found in Guangxi, Guangdong and Hainan, and good suitable areas were expanded in Southern of Fujian, Jiangxi, Hunan, Guizhou, and Yunnan.

In general, using this model, it was easy to find climatically suitable habitats of *T. rubrofasciata* in China. It was useful for targeting surveillance and increasing the efficiency and accuracy of public health investigations and vector control efforts in China in the future. Background surveys should focus on southern China and set up more sampling survey points, while a small number of sampling survey points could be set up in the southern areas of Yunnan, Guizhou, Jiangxi, and Fujian.

However, the present study has some limitations. Although more than 100 records were obtained for the construction the model, the data was mainly

concentrated in Guangdong and Guangxi, so its representation has certain limitations. More data would be collected to add to the model to get more accurate prediction. The widely used “MaxEnt” software for modeling species distributions from presence-only data tends to produce models with high-predictive performance but low-ecological interpretability, which are more complex but not necessarily predictively better-than subset selection. The simpler models implement two different kinds of model fitting: maximum entropy fitting for presence-only data and logistic regression for presence-absence data would be new motivation field. With improvement of the model, some model which could cover the shortage of the data, eg, MIAMaxent R package etc, which would be involved in the habitats of *T. rubrofasciata* prediction in future study (15).

Conflicts of Interest: No conflicts of interest were reported.

Acknowledgments: All participants of this surveillance program.

Funding: Supported by National Science and Technology Project (No.2018ZX10101002), National Key Research and Development Program of China (Grant No.2016YFC1202000), and Guangxi Natural Science Foundation Program (2018GXNSFAA 294092).

doi: 10.46234/ccdcw2021.182

Corresponding author: Qin Liu, liuqin@nipd.chinacdc.cn.

¹ National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (Chinese Center for Tropical Diseases Research); NHC Key Laboratory of Parasite and Vector Biology; WHO Collaborating Centre for Tropical Diseases; National Center for International Research on Tropical Diseases, Shanghai, China; ² Institute of Parasitic Disease Prevention and Control, Guangxi Zhuang Autonomous Region Center for Disease Control and Prevention, Nanning, Guangxi Zhuang Autonomous Region, China.
[‡] Joint first authors.

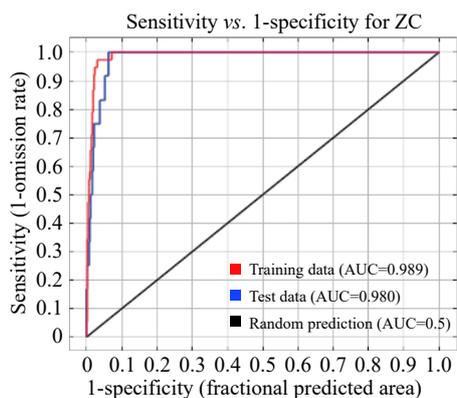
Submitted: July 23, 2021; Accepted: August 24, 2021

REFERENCES

- Liu Q, Zhou XN. Preventing the transmission of American trypanosomiasis and its spread into non-endemic countries. *Infect Dis Poverty* 2015;4:60. <http://dx.doi.org/10.1186/s40249-015-0092-7>.
- Pérez-Molina JA, Molina I. Chagas disease. *Lancet* 2018;391(10115):82 – 94. [http://dx.doi.org/10.1016/S0140-6736\(17\)31612-4](http://dx.doi.org/10.1016/S0140-6736(17)31612-4).
- Liu Q, Wu FW, Shi YL, Lu WC, Xie HG, Guo YH, et al. Distribution of triatomines, the vector of chagas disease — southern China, 2016–2018. *China CDC Wkly* 2020;2(33):629 – 33. <http://dx.doi.org/10.46234/ccdcw2020.174>.
- Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecol Model* 2006;190(3 – 4):231 – 59. <http://dx.doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Richman R, Diallo D, Diallo M, Sall AA, Faye O, Diagne CT, et al. Ecological niche modeling of *Aedes* mosquito vectors of chikungunya virus in southeastern Senegal. *Parasit Vectors* 2018;11(1):255. <http://dx.doi.org/10.1186/s13071-018-2832-6>.
- Sage KM, Johnson TL, Teglas MB, Nieto NC, Schwan TG. Ecological niche modeling and distribution of *Ornithodoros hermsi* associated with tick-borne relapsing fever in western North America. *PLoS Negl Trop Dis* 2017;11(10):e0006047. <http://dx.doi.org/10.1371/journal.pntd.0006047>.
- Artun O. Ecological niche modeling for the prediction of cutaneous leishmaniasis epidemiology in current and projected future in Adana, Turkey. *J Vector Borne Dis* 2019;56(2):127 – 33. <http://dx.doi.org/10.4103/0972-9062.263726>.
- Shi YL, Wei YB, Feng XY, Liu JF, Jiang ZH, Ou FQ, et al. Distribution, genetic characteristics and public health implications of *Triatoma rubrofasciata*, the vector of Chagas disease in Guangxi, China. *Parasit Vectors* 2020;13(1):33. <http://dx.doi.org/10.1186/s13071-020-3903-z>.
- Chen MX. Prevention and treatment of the dermatitis caused by *Triatoma*. *Hainan Med* 1986;1:35.
- Zhang G. Study on the medical significance of the *Triatoma*. *Hainan Med* 1991;4:8 – 9.
- Wang Y, Peng L. Four cases of anaphylactic shocks caused by triatomine bugs. *Clin Focus* 2006;21:1059 – 61.
- Eberhard FE, Cunze S, Kochmann J, Klimpel S. Modelling the climatic suitability of Chagas disease vectors on a global scale. *Elife* 2020; 9:e52072. <http://dx.doi.org/10.7554/eLife.52072>.
- Ceccarelli S, Balsalobre A, Susevich ML, Echeverria MG, Gorla DE, Marti GA. Modelling the potential geographic distribution of triatomines infected by *Triatoma virus* in the southern cone of south America. *Parasit Vectors* 2015;8:153. <http://dx.doi.org/10.1186/s13071-015-0761-1>.
- IPCC. Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 2001. <https://www.oalib.com/references/11734021>.
- Vollering J, Halvorsen R, Mazzoni S. The MIAMaxent R package: variable transformation and model selection for species distribution models. *Ecol Evol* 2019;9(21):12051 – 68. <http://dx.doi.org/10.1002/ece3.5654>.

SUPPLEMENTARY TABLE S1. Climatic variables used for predicting the potential distribution of Chagas disease vector *Triatoma rubrofasciata* in China based on MaxEnt.

Code	Variable	Source
Bio1	Annual mean temperature	http://www.worldclim.org/
Bio2	Mean diurnal range	http://www.worldclim.org/
Bio3	Isothermality	http://www.worldclim.org/
Bio4	Temperature seasonality	http://www.worldclim.org/
Bio5	Max temperature of the warmest month	http://www.worldclim.org/
Bio6	Min temperature of the coldest month	http://www.worldclim.org/
Bio7	Temperature annual range	http://www.worldclim.org/
Bio8	Mean temperature of the wettest quarter	http://www.worldclim.org/
Bio9	Mean temperature of the driest quarter	http://www.worldclim.org/
Bio10	Mean temperature of the warmest quarter	http://www.worldclim.org/
Bio11	Mean temperature of the coldest quarter	http://www.worldclim.org/
Bio12	Annual precipitation	http://www.worldclim.org/
Bio13	Precipitation of the wettest month	http://www.worldclim.org/
Bio14	Precipitation of the driest month	http://www.worldclim.org/
Bio15	Precipitation seasonality	http://www.worldclim.org/
Bio16	Precipitation of the wettest quarter	http://www.worldclim.org/
Bio17	Precipitation of the driest quarter	http://www.worldclim.org/
Bio18	Precipitation of the warmest quarter	http://www.worldclim.org/
Bio19	Precipitation of the coldest quarter	http://www.worldclim.org/



SUPPLEMENTARY FIGURE S1. Receiver operating characteristic curve [abbreviated as area under curve (AUC)] was used to evaluate model performance by MaxEnt model.