Spatial distribution of Biomphalaria spp., the intermediate host snails of Schistosoma mansoni, in Brazil

Ronaldo G.C. Scholte1,2,3, Omar S. Carvalho1, John B. Malone4, Jürg Utzinger1,2, Penelope Vounatsou1,2

1Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, P.O. Box, CH-4002 Basel, Switzerland; 2University of Basel, P.O. Box, CH-4003 Basel, Switzerland; 3Centro de Pesquisas René Rachou/Fiocruz-MG, Av Augusto de Lima 1715, Barro Preto, Belo Horizonte, CEP:30190-002, MG, Brazil; 4Department of Pathobiological Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA 70803, USA

Abstract. Schistosomiasis mansoni remains an important parasitic disease of man, endemic in large parts of sub-Saharan Africa, the Middle East, South America and the Caribbean. The aetiological agent is the trematode Schistosoma mansoni, whereas aquatic snails of the genus Biomphalaria act as intermediate hosts in the parasite life cycle. In Brazil, the distribution of Biomphalaria spp. is closely associated with the occurrence of schistosomiasis. The purpose of this study was to map and predict the spatial distribution of the intermediate host snails of S. mansoni across Brazil. We assembled snail “presence-only” data and used a maximum entropy approach, along with climatic and environmental variables to produce predictive risk maps. We identified a series of risk factors that govern the distribution of Biomphalaria snails. We find that high-risk areas for B. glabrata are concentrated in the regions of Northeast and Southeast and the northern part of the South region. B. straminea are found in the Northeast and Southeast regions, and B. tenagophila are concentrated in the Southeast and South regions. Our findings confirm that the presence of the intermediate host snails is correlated with the occurrence of schistosomiasis mansoni. The generated risk maps of intermediate host snails might assist the national control programme for spatial targeting of control interventions and to ultimately move towards schistosomiasis elimination in Brazil.

Keywords: Biomphalaria, intermediate host snail, schistosomiasis, risk mapping and prediction, ecological niche model, maximum entropy, Brazil.

Introduction

Schistosomiasis mansoni is a human parasitic disease, currently endemic in 54 countries in Africa, the Middle East, South America and the Caribbean (Chitsulo et al., 2000; Gryseels et al. 2006; Steinmann et al., 2006). In Brazil, Katz and Peixoto (2000), using data from the National Health Foundation, estimated that 6.4 million people in 18 states are affected by schistosomiasis. The causative agent is the blood-dwelling fluke Schistosoma mansoni, and the intermediate hosts are aquatic snails of the genus Biomphalaria (Mollusca: Gastropoda: Pulmonata: Planorbidae). Thus far, 11 species and one subspecies of Biomphalaria have been described in Brazil. Among these, three species (i.e. B. glabrata, B. tenagophila and B. straminea) were found naturally infected with S. mansoni. Another three species (i.e. B. amazonica, B. peregrina and B. cousini) have been considered as potential intermediate hosts (Corrêa and Paraense, 1971; Paraense and Corrêa, 1973; Caldeira et al., 2010, Teodoro et al., 2010). Clearly, B. glabrata is recognised as the best adapted intermediate host for transmitting S. mansoni due to its wide geographical distribution, high rates of infection and transmission efficiency. The distribution of this snail species is almost always associated with the occurrence of schistosomiasis mansoni, the observation of this fact was first mentioned by Lutz (1917) (Paraense, 1983; Teles and Vaz, 1987, Carvalho et al., 2008). It follows that maps of the distribution of B. glabrata and other (potential) intermediate hosts could serve as an important tool for spatial targeting of schistosomiasis control interventions (Stensgaard et al., 2012). However, only few studies assessed the distribution of Biomphalaria spp. in Brazil (Malone et al., 2005; Guimarães et al., 2009; Baboza et al., 2012). It should be noted that available data on Biomphalaria spp. in Brazil is “presence-only”, while care is indicat-
ed in absence data, since surveyed areas are large and the surveyors might have missed sites where snails occurred but were not discovered due to the limited sampling effort.

When both presence and absence data are available, classical statistical approaches can be used to analyse the data (Corsi et al., 2000; Guisan and Zimmerman, 2000; Elith, 2002; Scott et al., 2002). However, for presence-only data, modelling techniques are required that take into account the lack of zeros in the data in order to obtain accurate estimates of the outcome distribution. Different methods have been used for presence-only data when modelling species distribution, such as ecological niche models (ENMs). Among others, prominent ENMs include generalized linear models (GLMs), generalized additive models (GAMs), generalized linear mixed models (GLMMs) (Fenton et al., 2010), BIOCLIM (Busby, 1986; Nix, 1986), DOMAIN (Carpenter et al., 1993), environmental-niche factor analysis (ENFA) (Hirzel et al., 2002), genetic algorithm for rule-set prediction (GARP) (Stockwell and Noble, 1992) and maximum entropy (MaxEnt) models (Phillips et al. 2004). MaxEnt models proved particularly useful, as they pursue a general purpose method for making predictions from presence-only data (Ponder et al., 2001; Anderson et al., 2002, 2003; Graham et al., 2004; Phillips et al., 2006; Stensgaard et al., 2012).

In this study, we employed compiled Biomphalaria spp. presence-only data from Brazil to produce predictive risk maps of the spatial distribution of the intermediate host snails of *S. mansoni*. We used a MaxEnt modelling approach using as predictors both environmental and climatic proxies readily obtained from remote sensing and climate databases. The maps presented here will be useful for the spatial targeting of schistosomiasis control interventions and future efforts emphasising schistosomiasis elimination.

**Materials and methods**

**Malacological data**

We obtained Biomphalaria spp. occurrence data (i.e. presence-only) for Brazil from two main sources. First, data were readily available from the Laboratory of Helminthiasis and Medical Malacology (LHMM) of the René Rachou Research Center (CPqRR/Fiocruz-MG). Of note, the LHMM is a reference laboratory for schistosomiasis in Brazil, performing the examination and identification of the snails of the genus *Biomphalaria* according to the demands of health services. Second, data were retrieved through a systematic review of the peer-reviewed literature and from grey literature (Carvalho et al., 2008). Our database comprises 1,977 municipalities, hence 35.5% from all the municipalities in Brazil (Fig. 1).

**Environmental and climatic data**

Environmental and climatic proxies were considered in our analyses, since these are the main predictors for the distribution of *Biomphalaria* spp. Environmental data were extracted from different freely accessible remote sensing data sources, as summarised in Table 1. In brief, land surface temperature (LST) data were used as a proxy for day and night temperature. The normalized difference vegetation index (NDVI) was employed as a proxy for moisture. A digital elevation model (DEM) was used to extract altitude data.

Climatic data were obtained from Worldclim–Global Climate Data, which provides interpolated monthly climatic information from weather stations averaged over a 50-year period (from 1950 to 2000) at a spatial resolution of 1 km (Table 1). The Worldclim data have been used successfully in ENMs and by the global climate change community (Waltari et al., 2007).

![Fig. 1. Observed spatial distribution of the intermediate host snails of *S. mansoni* in Brazil; (A) *B. glabrata*; (B) *B. straminea*; (C) *B. tenagophila* and (D) *Biomphalaria* spp. (Carvalho et al., 2008).](image-url)
MaxEnt modelling approach

We used a MaxEnt modelling approach proposed by Phillips et al. (2006) for prediction of the intermediate host snail distribution. The MaxEnt is an iterative method used to obtain predictions or make inferences from incomplete information (e.g. presence-only data). The main purpose is to estimate the unknown probability distribution of a species based on the principle of maximum entropy (Phillips et al., 2006). Entropy, as defined by Shannon (1948), is “a measure of how much “choice” is involved in the selection of an event”. A distribution with higher entropy, involves more choices. Given a set of samples (e.g. presence of species) and a set of features (e.g. environmental/climatic variables), the MaxEnt model estimates niches by finding the distribution of probabilities closest to uniform (maximum entropy), constrained by the fact that feature values match their empirical average. The final result is a model that predicts areas where the snails are most likely to occur based on environmental and/or climatic correlations.

We employed the MaxEnt method in the present study to generate probability surfaces of the geographical distributions of the (potential) intermediate host snails of *S. mansoni*. The method’s principle is to estimate the spatial distribution of the snails of the genus *Biomphalaria* that are most probable to be present in a certain area, constrained to known observations (i.e. presence of intermediate hosts of *S. mansoni*). MaxEnt uses entropy as the means to generalize specific observations pertaining to the presence of a particular species.

The area under the curve (AUC) test statistics was used to measure the accuracy of the predictive distribution models. It estimates the probability of a test pixel to be correctly predicted as suitable for species presence as opposed to the probability of a randomly selected pixel of the map. The AUC ranges from zero to one, where values near one indicate high predictive performance of the model, while values smaller than 0.5 indicate low model predictive ability (Wiley et al., 2003; Allouche et al., 2006; Elith et al., 2006).

To determine the spatial distribution of the intermediate host snails of *S. mansoni* in Brazil, four models were constructed. The models were: (i) *B. glabrata*, (ii) *B. straminea*, (iii) *B. tenagophila*, and (iv) *Biomphalaria* spp., the latter including all three intermediate host species.
Statistical analysis

We used the MaxEnt software version 3.3.3e (Princeton University; Princeton, NJ, USA) to perform our analysis. Visualization and geostatistical display was carried out in ArcGIS version 9.3 (ESRI; Redlands, CA, USA). The final models were plotted with a 1 km spatial resolution.

Results

Fig. 2A shows the most suitable areas for the presence of *B. glabrata* across Brazil. The following environmental and climatic variables showed the highest contribution to the predictive model: temperature seasonality, maximum temperature of warmest month, annual precipitation and monthly mean diurnal range temperature. Our model shows a high probability of the occurrence of *B. glabrata* in the regions of Northeast and Southeast and the northern part of the South regions within Brazil.

The most suitable areas for the presence of *B. straminea* are shown in Fig. 2B. The map indicates high probability of occurrence of *B. straminea* in the Northeast and Southeast regions of Brazil. Annual precipitation, temperature seasonality and temperature annual range were the three variables with the highest contribution to the predictive ability of the model.

The prediction map in Fig. 2C shows that, the areas which are likely for *B. tenagophila* to be present, are concentrated in the Southeast and South regions of Brazil. Temperature seasonality, isothermality and monthly mean temperature of driest quarter were the most important predictors of the species’ geographical distribution.

![Fig. 2. Probability of the spatial distribution of the intermediate host snails of *S. mansoni* in Brazil; (A) *B. glabrata*; (B) *B. straminea*; (C) *B. tenagophila* and (D) *Biomphalaria* spp.](image-url)
Model-based predictions suggest that the areas most suitable for the presence of *Biomphalaria* spp. are located in the Northeast, Southeast and South regions of Brazil (Fig. 2D). The climatic parameters related to the species distribution are temperature seasonality, annual precipitation and monthly mean diurnal range.

The predictive ability of the above models were high with an AUC equal to 0.836, 0.859, 0.863 and 0.901 for *Biomphalaria* spp., *B. straminea*, *B. tenagophila* and *B. glabrata*, respectively.

**Discussion**

To our knowledge, we present the first predictive maps, using MaxEnt, of the spatial distribution of *B. glabrata*, the key intermediate host snail of *S. mansoni*, in Brazil. This snail species shows a wide distribution with a high probability of presence in the Northeast and Southeast regions and the northern part of the South region. Moreover, we present spatially explicit probability maps of the presence of *B. tenagophila* and *B. straminea*, which have also been identified as intermediate host snails of *S. mansoni* in Brazil. While *B. straminea* is concentrated in the Northeast region, *B. tenagophila* primarily occurs in the Southeast and South regions. A climatic suitability map showed that the region extending from Northeast to South has high probability of snail occurrence irrespective of their species with the highest one over 50% estimated in the Northeast, Southeast and part of South regions of Brazil. Our results agree with the literature showing the same pattern for each snail species (Carvalho et al., 2008; Barboza et al., 2012).

We employed a MaxEnt modelling approach, which has been proposed as a particularly useful ENM, to analyse a comprehensive dataset with more than 1,900 presence-only data points for *Biomphalaria* (used as outcome measure) in relation to climatic and environmental factors (used as covariates). ENMs have been widely used to address questions of species distribution and ecology in biodiversity-related studies using presence-only data (Kadmon and Heller, 1998; Gottfried et al., 1999; Peterson et al., 1999, 2001, 2002; Bakkenes et al., 2002; Peterson and Shaw, 2003; Phillips et al., 2006). More recently, ENMs, and specifically MaxEnt models, have been applied to deepen our understanding of vector and intermediate host distribution (Costa et al., 2002; Batista and Gurgel-Gonçalves, 2009; Stensgaard et al., 2012). However, only few studies used ENMs to predict the distribution of the intermediate host snails of schistosomiasis, with a notable exception of recent predictive risk maps of *Biomphalaria* spp. in Africa (Stensgaard et al., 2012).

In all our models, both temperature and precipitation were identified as important environmental features governing the distribution of *Biomphalaria* spp. in Brazil. Indeed, our models showed high predictive accuracy with AUC values above 0.8. According to Phillips and Dudík (2008), AUC values above 0.75 can be considered as to fall within the best model-fit category.

Noteworthy issues arising from our work can be summarised as follows. First, our results confirm that the distribution of aquatic snails is governed by environmental and climatic features, and hence the Worldclim database is useful for ENMs. Second, the MaxEnt model employed in this study proved useful to analyse and predict presence-only data for understanding geographical and ecological distributions of species in disease transmission cycles. Third, our maps of probability of the intermediate host snails’ distribution follow the same pattern of the predicted prevalence of schistosomiasis in Brazil. Hence, the *Biomphalaria* distribution map developed here might be used as predictor for *S. mansoni* distribution. Since the presence of the intermediate host snails of *S. mansoni* is a proxy for the disease distribution, and it is widely acknowledged which species is the most adapted intermediate host, the methodology and the generated risk maps are useful for decision makers to delineate priority areas for schistosomiasis control interventions, so that limited resources can be allocated most effectively. Our maps can guide the spatial targeting of control interventions and are of particular relevance now that efforts are getting under way to eliminate schistosomiasis in Brazil and elsewhere (Rollinson et al., 2012).

Some important issues related to the nature and precision of the *Biomphalaria* species data need to be considered when interpreting our findings. The *Biomphalaria* species data were extracted from peer-reviewed articles and grey literature. In most cases the data records were not georeferenced, and hence no precise spatial information is available. Most of the time, the data were at municipality level. Our assumption then was that the species found in a given municipality is uniformly distributed within this municipality along the drainage network. It would be interesting to conduct malacological surveys in selected municipalities, and georeference the occurrence of snails in order to determine whether data at higher spatial resolution would further improve model accuracy. Since a global neglected tropical disease (GNTD) database is now
available (Hürlimann et al., 2011) and spatially explicit parasite prevalence data (e.g. S. mansoni) can be freely obtained, it will be important to further populate the GNTD database with information on occurrence and density of intermediate host snails (and vectors for other diseases) to assist global control and elimination efforts targeting schistosomiasis and other neglected tropical diseases.

Acknowledgements

This investigation received financial support from the UBS Optimus Foundation, the Swiss-Brazilian joint research programme (BJRP 011008) and CNPq (Process 150386/2012-5).

References


Guisan A, Zimmerman NE, 2000. Predictive habitat distribu-


