ECOGRAPHY

Research article

Vegetation cover and biodiversity reduce parasite infection in wild hosts across ecological levels and scales

Cecilia S. Andreazzi[1,2], **Luis A. Martinez-Vaquero**[3,4], **Gisele R. Winck**[1], **Thiago S. Cardoso**[1], **Bernardo R. Teixeira**[1], **Samanta C. C. Xavier**[5], **Rosana Gentile**[1], **Ana Maria Jansen**[5] **and Paulo S. D'Andrea**[1]

[1]Laboratório de Biologia e Parasitologia de Mamíferos Silvestres Reservatórios, Inst. Oswaldo Cruz, Fiocruz, Rio de Janeiro, RJ, Brazil
[2]Depto de Biodiversidad, Ecología y Evolución, Univ. Complutense de Madrid, Madrid, Spain
[3]Grupo de Sistemas Complejos and DEFE, Escuela Técnica Superior de Arquitectura, Univ. Politécnica de Madrid, Madrid, Spain
[4]Grupo Interdisciplinar de Sistemas Complejos, Univ. Carlos III de Madrid, Leganés, Madrid, Spain
[5]Laboratório de Biologia de Tripanossomatídeos, Inst. Oswaldo Cruz, Fiocruz, Rio de Janeiro, RJ, Brazil

Correspondence: Cecilia S. Andreazzi (cecilia.andreazzi@fiocruz.br)

Land use changes and biodiversity loss critically disrupts ecosystem functioning and are major drivers of infectious disease outbreaks. *Trypanosoma cruzi*, the agent of Chagas disease, is a multi-host parasite whose epidemiology has changed due to the expansion of anthropogenic activities over natural areas. We aimed to understand the ecological processes increasing parasite prevalence at the individual, the community and the landscape levels using the largest database on small mammal infection by *T. cruzi* in Brazil. We applied machine learning techniques and structural equation models to show that allometric traits and the relative abundance of rodents in the community were important predictors of infection risk, followed by variables associated with the landscape environmental quality. Natural vegetation cover change and the taxonomic and functional dimensions of biodiversity indirectly reduced infection through its effect on the abundance distribution and composition of host communities. According to our findings, approaches to biodiversity conservation and restoration based on the integration of social inclusion and human welfare would contribute to regulate the prevalence of *T. cruzi* in wild hosts, which may reduce overall transmission risk.

Keywords: biodiversity–disease relationship, Didelphimorphia, land cover change, machine learning, neotropical forest, Rodentia, *Trypanosoma cruzi*

Introduction

Land use changes are one of the main drivers of biodiversity loss in tropical ecosystems, caused by the expansion of commodity-oriented economies over natural areas, particularly agricultural and livestock sectors (Henders et al. 2015). Local extinctions, increased dominance of a few species and biotic homogenization through species introductions are among the main components of biodiversity loss (Cardinale et al.

2012, Solar et al. 2015). A critical consequence of such environmental impacts is the reduction of the regulatory function of ecosystems and their services, which are also major drivers of infectious disease outbreaks (Allen et al. 2017). Although there is growing evidence showing that local biodiversity influences parasite prevalence across several hosts (Civitello et al. 2015), the generality of this pattern remains controversial, especially for parasites that cause infectious diseases in humans (Randolph and Dobson 2012, Wood et al. 2014).

The linkages between land use changes and biodiversity loss have established the basis for most research on the ecology of infectious diseases through the dilution effect hypothesis (Keesing et al. 2010). This hypothesis proposes that biodiversity can affect parasite infection dynamics both directly and indirectly by changing the abundance, behavior, immune response and contact rates between host species and vectors, and by altering host community composition and susceptible host regulation through competition (Keesing et al. 2006, Gottdenker et al. 2014). As landscapes are modified and biodiversity is reduced, species composition changes, the number and diversity of predator species decrease, and resilient and fast-growing species dominate. A variety of traits have been already associated with human-altered landscapes, such as small body size, habitat and diet generalism, high dispersal ability and short generation time (Flynn et al. 2009). If anthropogenic processes favor the most competent hosts for a given parasite species transmission, anthropogenic changes might enhance transmission efficiency, increasing prevalence rates among the host species and, consequently, increasing overall parasite population and the risk of outbreaks and spillover to other species (Keesing and Ostfeld 2021). On the other hand, if amplifying hosts are more vulnerable to environmental degradation, then biodiversity loss will tend to reduce transmission risk. Studies testing the dilution effect hypothesis have mostly focused on species richness and abundance (Glidden et al. 2021). However, biodiversity encompasses all forms of variability among living organisms and the ecological complexes of which they are a part, including species richness, functional groups, interaction networks and heterogeneity in habitat composition (Díaz et al. 2005). This highlights the importance of encompassing the multiple dimensions of biodiversity in the disease–biodiversity debate – including individual traits, functional diversity and landscape composition and configuration – to have a clear mechanistic understanding of the ecological processes driving the association between landscape change and zoonotic disease emergence.

In Neotropical forests, small non-volant mammals (marsupials and rodents) are not only resources for higher trophic levels in the food web, but also play important ecosystem functions, such as seed predation and dispersal, and control of invertebrate populations by predation (Bovendorp et al. 2019, Magioli et al. 2021). Parasite transmission regulation in high-diversity communities is also an important ecosystem service in which transmission is diluted by lower relative encounter rates between amplifying hosts (Frainer et al. 2018). Given their smaller size and rapid reproductive rates, rodents and marsupials tend to be more resilient to environmental disturbances compared to larger mammals (Pardini et al. 2010, Dirzo et al. 2014). Defaunated sites have small mammal communities with a higher abundance of generalist species and lower diversity (Pardini et al. 2010, Bovendorp et al. 2019), indicating that the functional loss of medium and large mammals potentially affects trophic cascades, ecological functions and evolutionary history (Magioli et al. 2021). Nevertheless, many small mammal species are still vulnerable to different drivers of disturbance such as habitat loss, fragmentation and reduced landscape connectivity (Pardini et al. 2010, Palmeirim et al. 2019). Rodents and marsupials may act as important reservoirs of prevalent zoonotic diseases, such as hantaviruses (Prist et al. 2016), Chagas disease (Jansen et al. 2018) and leishmaniasis (Carreira et al. 2017). Their high taxonomic diversity and turnover of species across the landscape make them ideal models for understanding the direct and indirect effects of land use change and biodiversity loss on infectious disease dynamics.

Trypanosoma cruzi (Kinetoplastida: Trypanosomatidae), the causative agent of Chagas disease in humans, has a complex ecology that challenges transmission modeling and disease control (Jansen et al. 2015). *T. cruzi* has already been found infecting more than 100 mammalian species and its transmission may be mediated by at least three interdependent mechanisms (Jansen et al. 2015). Transmission of *T. cruzi* among wild hosts involves trophic interactions, either by blood-feeding invertebrate vectors (Triatominae) or through a trophic route that cascades along the food web, when a susceptible predator feeds on infected vectors or preys (Jansen et al. 2015). Therefore, the functional structure of host communities and their interactions is at the core of *T. cruzi* enzootic scenarios. Host communities with equal numbers of species richness but differences in species composition might differ significantly in their functional diversity due to different levels of functional redundancy (Bovendorp et al. 2019), which may result in the loss of some ecosystem functions, while increasing the resilience of transmission cycles. It has already been suggested that the dilution effect plays a role in *T. cruzi* epidemiology since transmission is increased in degraded ecosystems due to the reduction of mammal diversity and a positive selection of generalist species with high transmissibility competence. The consequent amplification of the parasite's transmission cycle occurs due to a higher abundance of competent reservoir species, and an increased prevalence of infected vectors (Xavier et al. 2012, Jansen et al. 2015).

Here, we used data on the individual interactions between rodent and marsupial hosts and *T. cruzi* in Brazil to model 1) infection probability at the individual level and 2) the landscape and biodiversity predictors of parasite prevalence at the community level. At the individual level, we hypothesized that host functional traits underlying the transmission process would predict the infection probability. Comparative studies in mammals have already shown that body size and

diet correlate with a number of life history and ecological parameters, such as pace of life, home range size, longevity and demographic rates, which ultimately affect host exposure to parasites and the resulting infection profiles (Han et al. 2015, Becker et al. 2018). At the community level, we expected positive effects of natural vegetation cover area on the taxonomic and functional diversity of small mammals, which in turn would have an indirect effect in reducing the prevalence of *T. cruzi* infection in small mammal communities by the dilution effect. We used different accuracy measures, different datasets (balanced/unbalanced), different models and different spatial scales (landscape buffers size) to ensure the robustness of our results.

Material and methods

Small non-flying mammal community data

The small mammals' dataset encompasses 2054 adult individuals sampled at 250 linear transects across Brazil (Fig. 1). Study areas were chosen based on disease monitoring and investigative studies. Field methodology was standardized across all the areas (Supporting information) and sampling transects were disposed in order to cover all the different eco-epidemiological scenarios (interior forest areas, forest edges with agriculture and livestock, peri-domiciles). We produced circular landscapes (buffers) of 3, 5 and 10 km radius around each linear transect centroid (Fig. 1A–C) to define our operational communities and investigate the spatial scale of the effect of biodiversity and land cover variables on parasite infection. A minimum-size buffer of 3 km (Fig. 1B) was chosen since it is sufficient to include the home range area of the largest species, *Didelphis* spp. and *Proechimys* spp. (Rowcliffe et al. 2016). Contacting buffers were joined to produce a single one (Fig. 1C), thus varying landscape sizes. A total of 50, 45 and 36 operational communities were classified within the 3, 5 and 10 km radius buffers, respectively. We calculated the landscape and species diversity metrics for each of these operational communities. A total of 1637 individuals were tested for *T. cruzi* infection by serological and parasitological methods (Supporting information). We assessed parasite occurrence at the host individual level by considering infected individuals that had at least one diagnostic method positive, which means that it interacted with *T. cruzi* at least once during its lifetime. Parasite prevalence at the community level
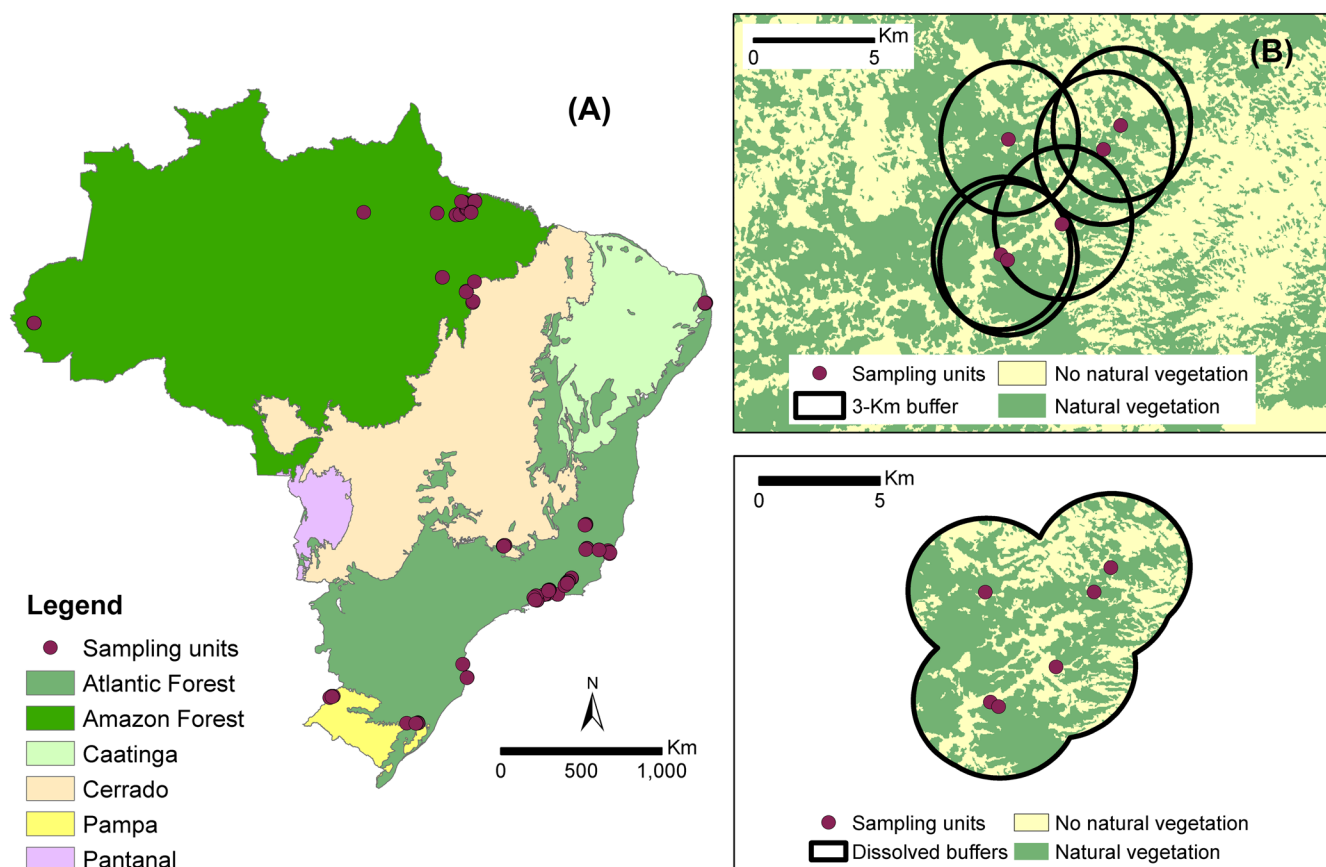

Figure 1. Small-mammal operational communities studied. Map depicting the sampling units of all communities included in the present study (A), specifying the representation of 3 km radius buffers circulating the centroid of each sampling transect (B), and the landscape built with superposed buffers as a single operational community to estimate landscape and biodiversity variables (C).

was calculated as the number of hosts that tested positive for *T. cruzi* infection (by either fresh blood smears examination, hemocultures and serological tests) divided by the number of hosts tested (Bush et al. 1997).

Land cover variables

We used MapBiomas raster layers (5th collection) for natural vegetation cover estimations to produce land cover variables, spanning from 1985 to 2019, and resolution of 30 × 30 m (Souza et al. 2020). Mapbiomas Project is a multi-institutional initiative to generate annual land cover and land use maps from automatic classification processes applied to satellite images. Mean normalized difference vegetation index (NDVI) was estimated using NASA Terra Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices Monthly (MOD13C2) ver. 6.1 layers (Didan 2021). To calculate all vegetation variables, we used raster layers corresponding to the sampling year of each study, and to calculate the loss or gain of natural vegetation area, we included the natural vegetation cover layers from five years before each study started. We computed the vegetation cover area and the perimeter of the natural vegetation fragments, both normalized by the total sampling area (VegCov and Edge, respectively). Since changes in land cover are expected to alter the composition and relative abundances of host communities with a consequent impact on infection profiles, we also estimated the proportion of natural vegetation cover that was lost or gained (VegCovChange). VegCovChange was defined as the difference in natural vegetation cover area between the year of the study and the previous five years, divided by the total sampling area, with negative values representing the gain of natural vegetation cover in the five years prior to the study beginning. A time lag of five years before the sampling period was chosen because it comprises at least one generation for all small mammal species considered in all studies. A longer time lag (10 years prior) was also analysed, but it was not significantly different from VegCovChange considering five years (Kendall's tau = 0.63, p < 0.005).

Taxonomic and functional diversity

In order to understand how taxonomic and functional diversity (hereafter species biodiversity) influences the infection probability, we applied different metrics to estimate biodiversity for each operational community considering all species sampled within. We characterized taxonomic diversity by two measures: 1) Menhinick index (Menhinick 1964), which is the total number of species sampled divided by the square root of the total number of sampled individuals, and is less affected by sampling effort than species richness (Bandeira et al. 2013); and 2) the Fisher's alpha (Fisher et al. 1943), which is the curvature parameter of the expected species–abundance relationship, and is independent of sample size. Higher values of Fisher's alpha indicate that the distribution of species abundance in the community is more equitable,

while smaller values indicate a skewed distribution with a few highly abundant species and all the others rare. As the functional diversity of small mammal communities, we calculated two different indices: 1) functional dispersion (FDis), which is the mean distance in multidimensional trait space of individual species to the centroid of all species, weighted by their abundances (Laliberté and Legendre 2010); and 2) functional evenness (FEve), which describes the distribution of abundances in a functional trait space, where low functional evenness indicates that some parts of the functional niche are underutilized (Villéger et al. 2008). For trait selection, we followed previous studies on small mammals (Bovendorp et al. 2019, Cardoso et al. 2021), including the morphological, dietary, foraging stratum and foraging activity time traits (Supporting information) which are related to host–parasite interactions (Stella et al. 2018, Cardoso et al. 2021). We obtained the morphological traits from our original data, and gathered the ecological traits from literature (Kissling et al. 2014, Wilman et al. 2014). We estimated pairwise functional distances between species using the Gower distance because the functional trait matrix includes both continuous and categorical variables (Pavoine et al. 2009). We performed the functional diversity analysis using the 'FD' package (Laliberté and Legendre 2010) in R ver. 4.0.3 (www.r-project.org).

Since rodents and marsupials respond differently to parasite infection due to their biological, ecological and epidemiological characteristics, we also characterized the species composition of the host communities as the relative proportion of rodents. We calculated the rodent abundance ratio (number of rodents captured over the total number of small mammals; RatioRodent) for each landscape to assess whether variations in this ratio influence the probability of infection at the community level.

Parasite occurrence predictors at the individual level: machine learning models

To investigate the parasite occurrence at the individual level, we considered only the individuals tested for *T. cruzi* infection, and removed those lacking information on body mass. Because infection rates across hosts tend to be highly skewed and most individuals are not infected, infection datasets are usually unbalanced. In our case, 74.2% of the individuals were negative and 25.8% positive. Hence, we created a balanced dataset for investigating how this intrinsic unbalance of infection data could affect modeling results, since machine learning algorithms are usually designed to maximize overall accuracy and thus fail in predictions by overfitting the majority class. We built the balanced dataset by adding synthetic individuals to the original unbalanced dataset, using the synthetic minority oversampling technique (SMOTE), which generates new samples of the minority class of the target variable by interpolation (Chawla et al. 2002). We applied our machine learning models to both the empirical dataset (unbalanced) and a partially synthetic dataset (balanced) since the results obtained using these two groups of datasets

are complementary. We used SMOTE from the 'unbalanced-learn 0.8.0' Python library (Lemaitre et al. 2017).

Eight different machine learning algorithms with their default parameters applied to datasets (model building, evaluation and validation are described in Supporting information). We computed the area under the receiver operating characteristic curve (hereafter AUROC) and the area under the precision–recall curve (AUPRC) to evaluate model accuracies. We chose the gradient boosting classification (GBM) model since it is largely used in ecology (Elith et al. 2008) and builds a robust final model through the ensemble of weaker decision trees models fitting residuals from previous trees, which incrementally reduce prediction errors. To set the hyperparameters values of our GBM model, we carried out a grid search through the GBM hyperparameters (Supporting information). To estimate variable importance, we employed two different methods, the model variable importance and the permutation variable importance. The former refers to the relative influence of each variable during the splitting of the tree building processes. The latter is defined as the decrease of the model accuracy when a single value is randomly shuffled. We repeated this shuffling 10 times for each variable and calculated its standard deviation. We also computed partial dependence plots to describe the functional relationship between the infection probability and the predictor variables, which can be linear, quadratic, monotonic or more complex. To perform all machine learning analyses, we used the 'scikit-learn 0.24.1' Python library (Pedregosa et al. 2011).

Land cover and biodiversity predictors of parasite prevalence at the community level: structural equation models

To assess the direct and indirect effects of land cover and biodiversity measures on *T. cruzi* prevalence at the host community level, we used path analyses, within the structural equation models (SEMs). We hypothesized that landscape metrics could affect parasite prevalence both directly and indirectly through its effect on biodiversity measures, which in turn would directly affect parasite prevalence. We also assumed the existence of correlations within landscape cover variables and among biodiversity metrics and thus allowed paths to vary freely. We fitted a general SEM using the entire dataset but also separated SEMs for the forested biomes (Atlantic Forest and Amazon), since most of our samples were taken within those areas. To estimate partial regression coefficients and correlations and to fit SEM we used maximum likelihood (Shipley 2016) through 'lavaan' package (Rosseel 2012) in R ver. 4.0.3. In Results, we describe the proportion of the variance explained by biodiversity and land cover variables by the marginal and conditional coefficients of determination (R^2). Data used in this work can be found in the Supporting information and code is available under request.

Results

Overall, 69 different small mammal species, 49 rodents and 20 marsupials, were identified. *T. cruzi* is able to infect all those species and its role as a reservoir will depend on environmental variables, parasite subpopulation, vector ecology and host parasitemia (Jansen et al. 2018). Individual traits (morphological, dietary, foraging stratum and foraging activity time), biodiversity measures, land cover metrics and infection rates varied among the landscapes (Supporting information).

Parasite occurence predictors at the individual level

All GBM models performed better than null predictions (all models AUROC > 0.5 and AUPRC > 0.258, Supporting information). Unbalanced- and balanced-data trained models predicted the occurrence of *T. cruzi* infection in host individuals with similar level of accuracy for the 3 km buffer (Fig. 2, Supporting information), 5 km buffer (Supporting information) and 10 km buffer (Supporting information).

The relative importance of covariates was consistent across the three landscape buffers analyzed (Fig. 2, Supporting information). The variables RatioRodent and BodyMass had the highest relative importance values in all models. In unbalanced-data models, morphological traits had the highest relative importance in predicting *T. cruzi* occurrence, followed by biodiversity and land cover measures (Fig. 2). In the balanced-data models, the relative importance of biodiversity variables was higher than in unbalanced-data models, and represented the most important group of variables in predicting parasite occurrence (Fig. 2). In addition, the Ground foraging stratum also became important when considering balanced-data models. Furthermore, Menhinick, Fisher, FDis, FEve, Tail/Body, VegCov and VegCovChange displayed relative importance when considering unbalanced and/or balanced data (Fig. 2).

By evaluating the marginal effect of each variable on the predicted outcome of the model through the partial dependence plots, we observed that unbalanced and balanced-data models had similar fitted functions for the probability of *T. cruzi* infection at the individual level (Fig. 3, Supporting information). The probability of *T. cruzi* infection increased with BodyMass and decreased with RatioRodent and Edge (Fig. 3, Supporting information). The fitted function for the infection probability followed the distribution of Tail/Body ratio, with a decreased probability in individuals of species with either a lower or higher Tail/Body ratio: those that tend to be fossorial or those that have a better locomotion and balance in the arboreal strata. The probability of parasite infection peaked at lower-intermediate levels of Menhinick, but decreased at higher Fisher and FDis (Fig. 3, Supporting information), which indicates that there is a minimum host diversity required to amplify *T. cruzi* transmission; but after a certain threshold of increased host biodiversity, the infection is greatly reduced. The probability of *T. cruzi* infection in
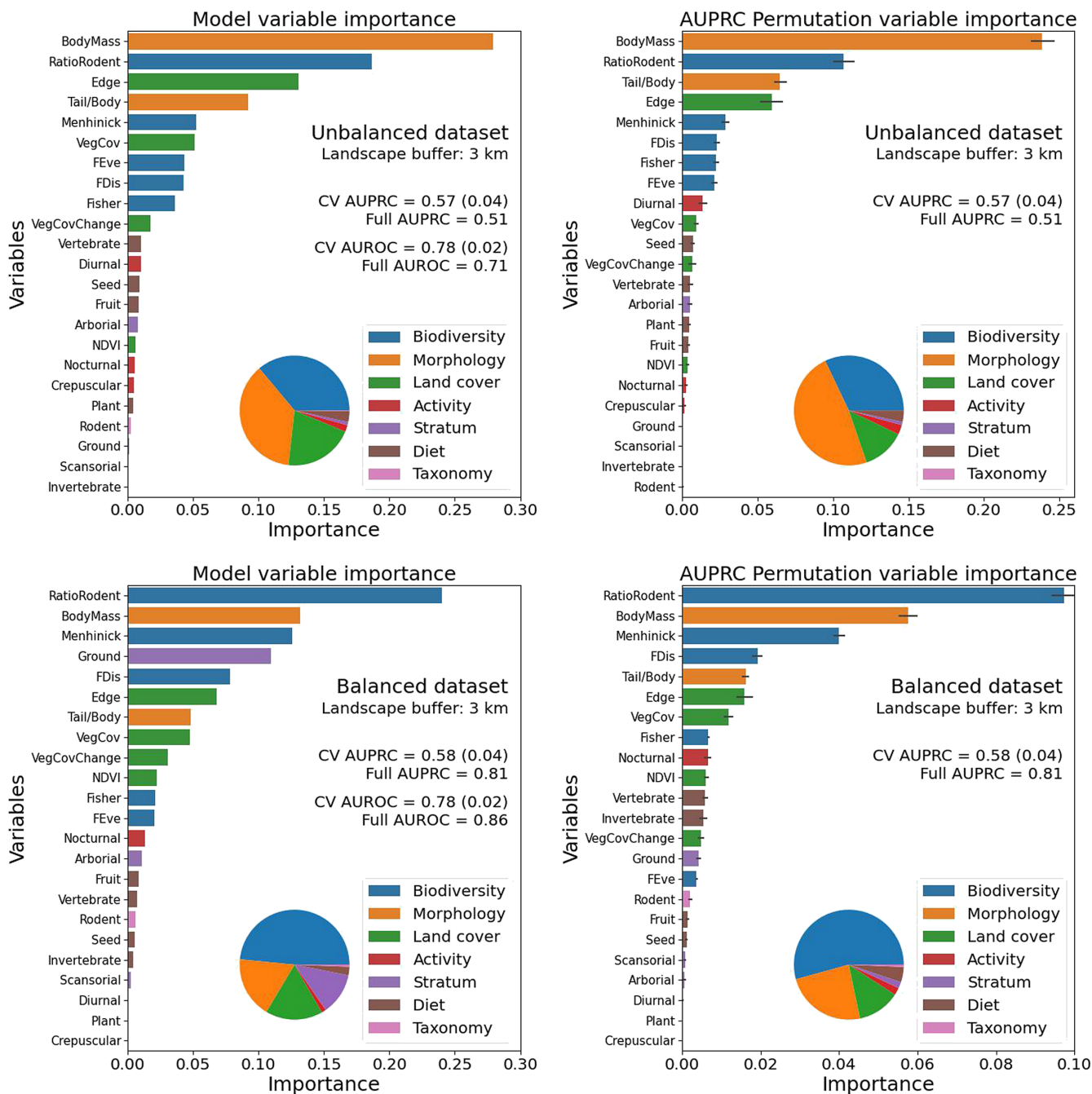
Figure 2. Variable importance and model performances considering 3 km radius buffer for unbalanced (first row) and balanced (second row) datasets. Variable importances are computed both directly from the model (first column) and using the permutation variable method via AUPRC (second column) cross-validation accuracy. Cross-validation (CV) AUROC and AUPRC accuracy values are shown with their standard deviation between parenthesis. Full AUROC and AUPRC accuracies are calculated fitting the whole input dataset to the model.

individual hosts varied along the VegCov and VegCovChange gradient. Intermediate to high levels of VegCov and landscapes that experienced negative or positive VegCovChange (those that gained or lost natural vegetation cover when compared to the previous five years) had higher probability of infection. Individuals in landscapes that remained with constant vegetation cover (VegCovChange=0) were those with the lowest probability of infection (Fig. 3, Supporting information).

Land cover and biodiversity predictors of parasite prevalence at the community level

Structural equation models explained nearly 55% of the variation in parasite prevalence at the community level and goodness-of-fit increased with the size of landscape buffer ($R^2 = 0.52$, 0.53 and 0.57 for 3, 5 and 10 km, respectively). RatioRodent was the only variable with significant negative direct effect on *T. cruzi* prevalence in all landscape
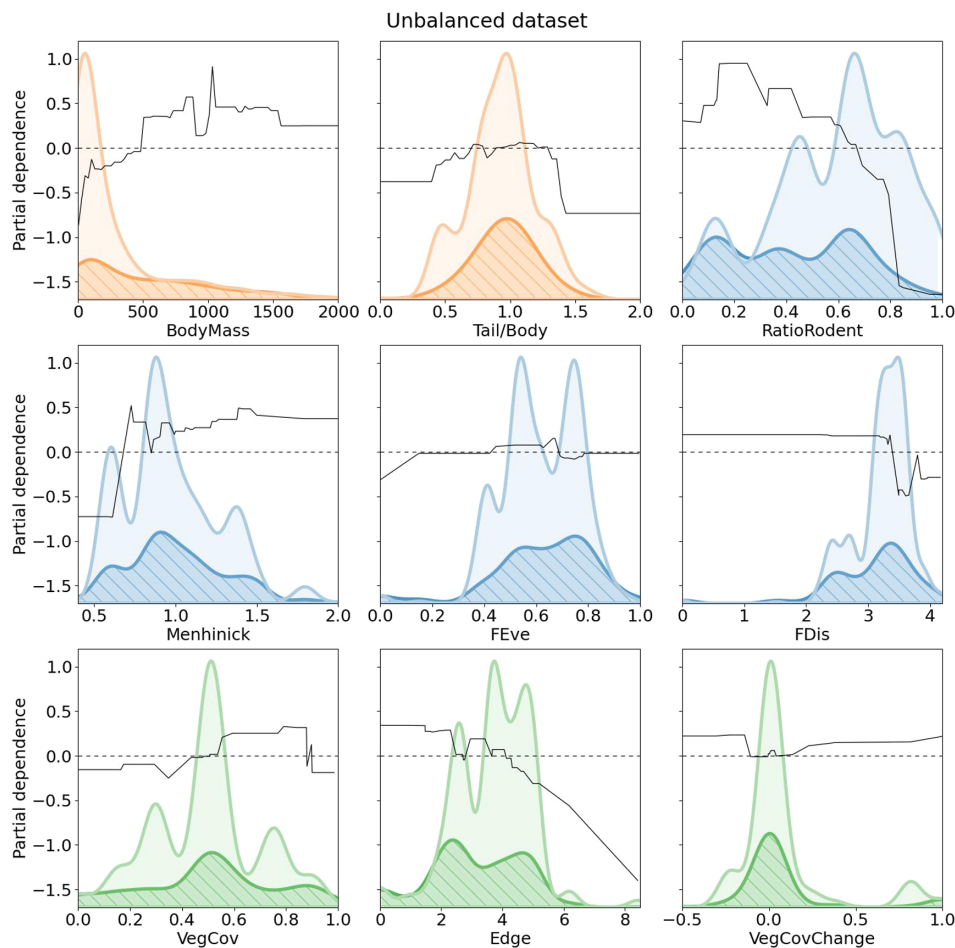
Figure 3. Partial dependence plots for the unbalanced dataset. Marginal effect of the most important variables on the predicted outcome of the GBM model considering 3 km radius buffer and unbalanced dataset. The black lines on the partial dependence plots show the functional relationship between the infection probability and the variable. Colored lines show the density distribution of the variables in the unbalanced dataset for infected individuals (darker with hatched pattern) and non-infected individuals (lighter).

buffers (β = −0.63, −0.61 and −0.80 for landscape buffers 3, 5 and 10 km, respectively; Fig. 4 show paths for 3 km landscapes, for the other buffers see Supporting information). VegCovChange had a direct positive effect on parasite prevalence, but only when considering a 10 km buffer landscape (Supporting information). All the biodiversity measures (Menhinick, Fisher, FDis and FEve) were positively correlated among them and also positively correlated with RatioRodent, having a negative indirect effect on the parasite prevalence of small mammals communities when considering the three landscape buffers (Fig. 4, Supporting information). Land cover also indirectly affected parasite prevalence through its effect on RatioRodent (Fig. 4, Supporting information), but this effect varied according to the landscape buffer considered. For the 3 and 5 km radius landscape buffer, Edge and VegCovChange positively affected RatioRodent, with an indirect negative effect on parasite prevalence (Fig. 4, Supporting information). In the 10 km-sized buffers, VegCov had a positive effect, Edge had no significant effect on RatioRodent and NDVI negatively affected it, having an indirect positive effect on *T. cruzi* prevalence (Supporting information).

When considering the Amazon and Atlantic Forest separately, we performed SEMs only for the 3 km radius landscapes since the largest landscapes had less than 20 observations per biome. SEMs explained 72 and 67% of the *T. cruzi* prevalence of small mammals communities in the Amazon and Atlantic Forest biomes, respectively. In the Amazon, parasite prevalence was negatively directly influenced by RatioRodent and indirectly affected by biodiversity variables, which were strongly correlated with each other (Fig. 5). Land cover influenced biodiversity only by the positive effect of VegCov on Fisher (Fig. 5). Higher values of VegCovChange (loss in vegetation cover) in the Amazon was negatively related to VegCov, Edge and NDVI. Unlike in the Amazon, land cover variables in the Atlantic Forest were weakly correlated with each other (with the exception of the positive correlation between NDVI and VegCov, Fig. 6) and had no effect on small mammals biodiversity measures (Fig. 6). As in the Amazon, biodiversity variables such as Menhinick, Fisher, FDis and FEve were positively correlated with each other. However, only Menhinick, Fisher and FDis were positively related to RatioRodent, and this effect was weaker than it was observed in the Amazon.
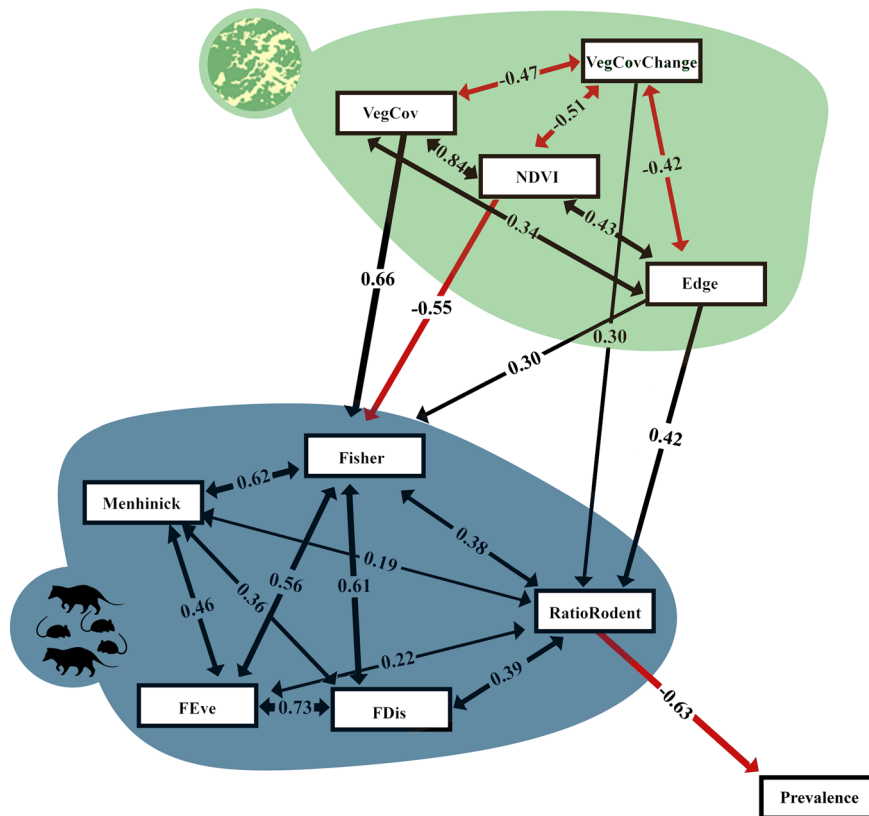
Figure 4. Structural equation model (SEM) diagram describing the relationships between biodiversity (RatioRodent, FDis, FEve, Menhinick, Fisher) and land cover (VegCov, VegCovChange, Edge, NDVI) variables and its effects on *Trypanosoma cruzi* prevalence at the community level (Infection) considering 3 km radius buffer (root mean square error of approximation < 0.05, standardized root mean square residual = 0.015). Only significant relationships (p < 0.005) are shown. Double arrow lines represent correlations, and one arrow lines represent regressions. Positive and negative pathways are indicated by black and red lines, respectively. The thickness of the arrows is scaled to illustrate the relative strength of effects and the standardized coefficients are indicated on each line.

Discussion

By using a combination of different modeling approaches applied at the individual and community levels, we showed the complex ways that host traits, multiple biodiversity components and landscape structure affect infection risk using the multi-host parasite *T. cruzi* as a model system. When considering the unbalanced dataset, morphological traits (mainly species body mass) were important predictors of the individual infection probability, followed by biodiversity measures and landscape structure. For the balanced dataset, biodiversity measures, particularly the relative frequency of rodents in the communities, and foraging stratum (ground foragers) were the most important variables to predict parasite occurrence at the individual level. At the community level, parasite prevalence was directly affected by the relative frequency of rodents in the communities and indirectly by biodiversity measures and landscape attributes.

Parasite distributions tend to be highly aggregated, with most host individuals showing no infection (Poulin 2007). Therefore, the nature of infection data is intrinsically unbalanced and reflects the heterogeneity among host individuals in their exposure or susceptibility to acquire

parasites, with parasite traits themselves explaining some of the variability in aggregation levels (Poulin 2013). This type of skewed unbalanced data is often reported as a serious obstacle to the classification performance of machine learning algorithms because of its tendency to bias the results of the models (Wang et al. 2021). We have explored the potential impact of this issue by generating a balanced dataset, which included synthetic data obtained oversampling from the minority class (infected individuals) of the original unbalanced data, and comparing the results from both datasets. Unbalanced-data models usually look better predictors, yet this effect may be spurius. They tend to produce biased predictions towards the majority class (non-infected individuals) since they are better predictors of it. Since infected individuals appear rarely, false negatives may not have a strong impact in global accuracy. In our case, the accuracy of balanced- and unbalanced-data models were similar but variable importance differed according to the approach used. Thus, unbalanced-data models inflated the importance of morphological traits because small individuals had in general a higher probability of not being infected. If, instead of on the correctness of the prediction, our interest also rested on knowing why infected individuals are so, then
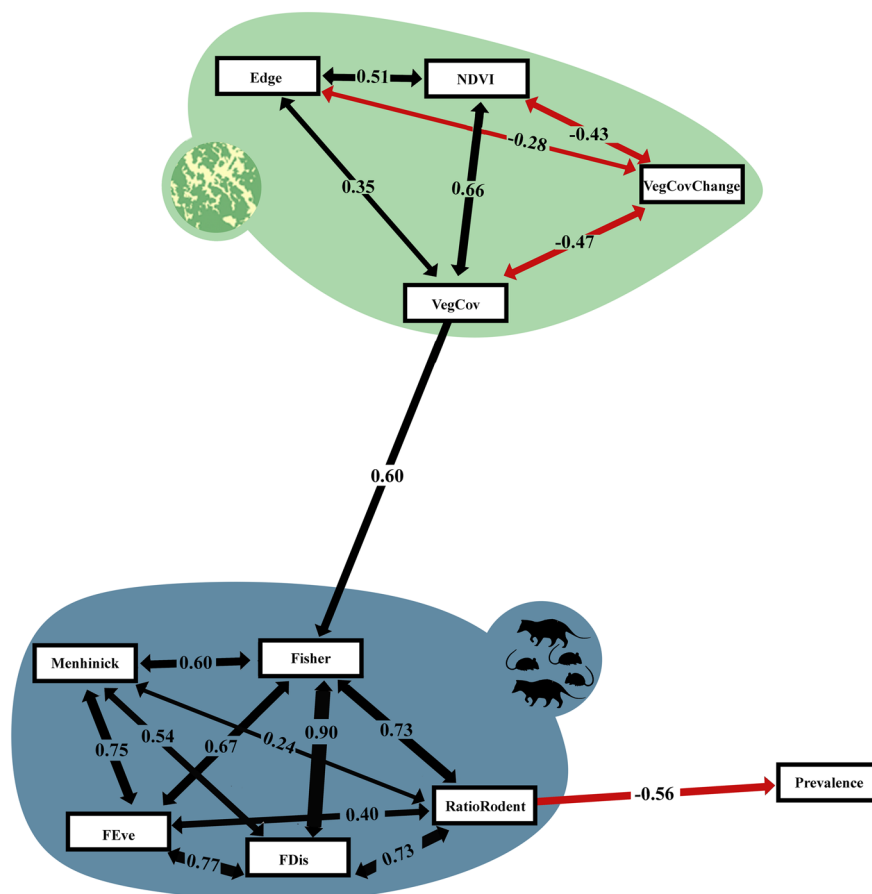
Figure 5. Structural equation model (SEM) diagram describing the relationships between biodiversity (RatioRodent, FDis, FEve, Menhinick, Fisher) and land cover (VegCov, VegCovChange, Edge, NDVI) variables and its effects on *Trypanosoma cruzi* prevalence at the community level (Infection) considering 3 km radius buffer in the Amazon biome (root mean square error of approximation < 0.05, standardized root mean square residual = 0.01). Only significant relationships (p < 0.005) are shown. Double arrow lines represent correlations, and one arrow lines represent regressions. Positive and negative pathways are indicated by black and red lines, respectively. The thickness of the arrows is scaled to illustrate the relative strength of effects and the standardized coefficients are indicated on each line.

balancing the dataset may become a solution. However, this strategy may also bias data and results. Therefore the best strategy to avoid the effect of bias is using both approaches. Our results show that biodiversity measures and landscape structure were important predictors of individual infection in all the modeling approaches used, which ratifies the importance of these variables and confirms previous results showing that host species composition and habitat quality affects zoonotic infection risk (Gottdenker et al. 2012, Xavier et al. 2012, Estrada-Peña et al. 2014, Keesing and Ostfeld 2021).

Allometric scaling has already been suggested as a general rule in parasite interactions (Kuris et al. 1980, Hechinger 2015). Body size is expected to be related with physiological, ecological and life-history traits in mammals, such as the metabolic rate, the home range size, the longevity and the reproductive strategy (Brown et al. 2004, Jetz et al. 2004), which affect host interactions with parasites. Our results indicate that allometric relationships are important drivers of host vulnerability to *T. cruzi* infection, with individuals of larger-sized species having higher *T. cruzi* occurrence than individuals of smaller-sized species. Mammals are constantly exposed to *T. cruzi* infection, which can occur once or multiple times during their lifetime (Jansen et al. 2015). Rodents and marsupials with larger body mass have larger home ranges and live longer (Brown et al. 2004), increasing their chances of exposure to infection due to a possible higher number of contact events through infected vectors and other transmission routes (e.g. oral route) (Jansen et al. 2015). Few studies have investigated the effect of body mass on *T. cruzi* occurrence (Orozco et al. 2016, Ghersi et al. 2020), but among-species comparisons suggest that larger-bodied marsupials and rodents tend to have higher prevalence and interact with a higher diversity of *T. cruzi* genotypes (Jansen et al. 2015, 2018).

Infection probability tended to follow the distribution of Tail/Body ratio values of the sampled individuals, but with a lower probability in individuals with lower or higher Tail/Body ratio. This indicates that species with adaptations to terrestrial or arboreal locomotor habits tended to show a
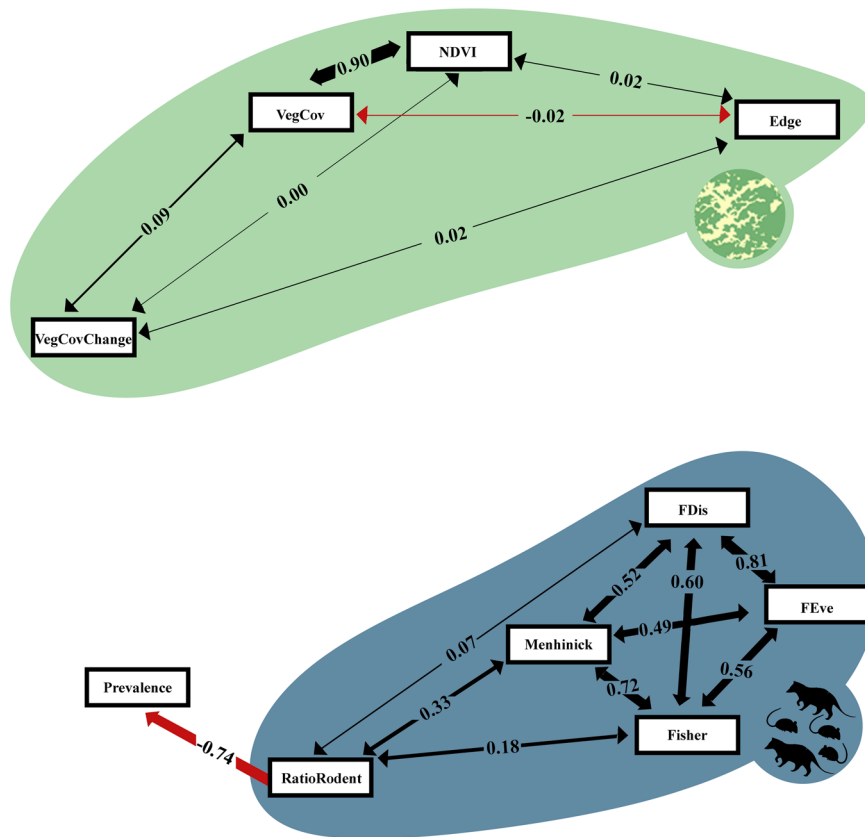
Figure 6. Structural equation model (SEM) diagram describing the relationships between biodiversity (RatioRodent, FDis, FEve, Menhinick, Fisher) and land cover (VegCov, VegCovChange, Edge, NDVI) variables and its effects on *Trypanosoma cruzi* prevalence at the community level (Infection) considering 3 km radius buffer in the Atlantic Forest biome (root mean square error of approximation < 0.05, standardized root mean square residual = 0.013). Only significant relationships (p < 0.005) are shown. Double arrow lines represent correlations, and one arrow lines represent regressions. Positive and negative pathways are indicated by black and red lines, respectively. The thickness of the arrows is scaled to illustrate the relative strength of effects and the standardized coefficients are indicated on each line.

lower probability of being infected. Scansorial species that have the ability to use a wider range of strata (intermediate values of Tail/Body ratio) may have an increased opportunity for interacting with vector species, since triatomine species are associated with a wide range of habitats (Péneau et al. 2016, Abad-Franch and Gurgel-Gonçalves 2021). In addition, the foraging stratum had a relatively high importance in the balanced-data models, with ground-forager species showing a lower probability of infection, a pattern already observed in local studies (Rademaker et al. 2009, Correa et al. 2015).

Host communities with a higher proportion of rodents had reduced *T. cruzi* infection, both at the individual and at the community levels. Previous studies have already reported higher infection rates in marsupial species, which are frequently described as reservoirs of *T. cruzi* in natural environments (Jansen et al. 2018). Despite that, rodent species may be considered as reservoirs under certain circumstances due to the presence of traits that favor the gradual accumulation of *T. cruzi* in the host populations (e.g. high longevity and potential for vertical transmission of *T. cruzi*) (Correa et al. 2015). In addition, several rodents share their microhabitat with *T. cruzi* vectors, many are

favored by disturbed environments close to human dwellings, eventually contributing to cycles of domestic or peridomestic transmission of *T. cruzi* (Orozco et al. 2014, Dario et al. 2022). We propose the use of species abundances distributions, here synthesized as the rodents/small mammals' ratio, as an important host biodiversity composition metric to describe *T. cruzi* infection profile in small mammal communities.

The amount of forest edges in the landscape decreased infection rates, both at the individual and at the community levels, which may be explained by the positive correlation between forest edges and vegetation cover, and the positive effect of these landscape variables on small mammal taxonomic diversity and relative frequency of rodents in the community. Although forest edges had already been related to increased transmission of several infectious diseases such as yellow fever (Prist et al. 2022), malaria (Medeiros-Sousa et al. 2019) and spotted fever (Scinachi et al. 2017) through its effects on vector abundance, we show that when considering the host communities, these relationships are not straightforward. In our case, forest edges were not related to vegetation cover in the Atlantic Forest and positively related to forest cover in the Amazon, which indicate that those edges may be increasing

habitat heterogeneity and favoring a higher diversification of small mammal communities and reducing the dominance of superabundant species. For instance, edge effects tend to decrease the richness and abundance of marsupial species, while rodent species were not affected by it in Amazonian forest fragments (Santos-Filho et al. 2012). The qualification of the type of edges, if it is between natural areas and agriculture, pasture, housings or degraded areas would help to better understand the effects of land use in shaping host assemblages and consequent parasite prevalence.

Although our results support the hypothesis that landscape attributes directly affect the taxonomic and functional diversity of small mammals, which in turn affects *T. cruzi* prevalence, the interplay among land cover variables and their effect on host diversity differed between biomes. Despite both being highly diverse rainforest biomes, the Amazon and the Atlantic Forest have different environmental histories. The latter is highly fragmented and holds nearly 75% of the Brazilian human population (Ribeiro et al. 2009), whereas the former currently comprises large areas of continuous primary forest due to a period of regeneration after the pre-Columbian period of higher human density (de Souza et al. 2018). Along with the evolution of environments, the different history of human occupation contributes to the marked differences in the structure of biodiversity. In the Amazon, areas with higher vegetation cover, NDVI and forest edges also had more diverse small mammal communities with a higher relative frequency of rodents, which decreased parasite prevalence. On the other side, in the Atlantic Forest the landscape attributes had no significant effect on the small mammal biodiversity and were weakly correlated with each other. Nevertheless, more diverse host communities also tended to have a higher relative frequency of rodents, which reduced *T. cruzi* prevalence.

The lowest infection risk at the individual level was found in areas with higher vegetation cover and that underwent less changes in vegetation cover. At the community level, vegetation cover increased the taxonomic diversity of host communities and changes in vegetation cover were positively related to a higher frequency of rodents in the communities. However, this relationship was not found in the Atlantic Forest and it is related to differences in the deforestation processes in these two biomes. The deforestation process in the Atlantic Forest is much older than in the Amazon and larger forest remnants are usually protected areas. On the contrary, the Amazon is currently experiencing a major deforestation front (Silva Junior et al. 2021) and largest deforestation rates are found in forested areas that are relatively accessible through edges such as roads, villages or agricultural lands (Carrero et al. 2022). Therefore, highest infection rates were mostly related to changes in host community composition driven by changes in vegetation cover. At a regional scale (10 km buffer), there was a direct effect of vegetation cover loss on increasing infection rates, considering both biomes. Thus, the effects of land use changes on infection risk is scale-dependent and this is one of the main reasons explaining contradictory results regarding the dilution effect (Johnson et al. 2015).

We recommend scale-dependency to be always taken into account when evaluating the ecosystemic function of disease regulation.

In summary, our results show that there is a minimum level of vegetation cover and of species diversity in small mammal communities that is required to maintain the transmission of *T. cruzi* among wild hosts because vegetation cover slightly increases individual infection probability at intermediate levels. However, at a community level, land cover change and deforestation increased host communities infection rates. Although the effect of functional diversity on infection rates was indirect at the community level, our results show that the functional structure of host communities affect the probability of infection at the individual level. In this sense, control measures that combine epidemiological surveillance with broader policies mitigating deforestation as well as promoting new approaches to biodiversity conservation and restoration founded on social inclusion and human welfare can be efficient to reduce the prevalence of *T. cruzi* in wild hosts.

Ethics statement – All field procedures followed the standards of capture, handling and care recommended by the Ethics Committee on the Use of Animals of the Oswaldo Cruz Foundation (license CEUA LW-39/14) and a federal license for the capture of wild mammals (13373 MMA/ICMBio/SISBIO).

Author contributions

Cecilia S. Andreazzi: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (lead); Methodology (lead); Project administration (lead); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **Luis A. Martinez-Vaquero**: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Methodology (equal); Software (lead); Validation (lead); Visualization (lead); Writing – original draft (equal); Writing – review and editing (equal). **Gisele**

R. Winck: Data curation (equal); Formal analysis (lead); Investigation (equal); Methodology (equal); Software (equal); Visualization (equal); Writing – original draft (equal); Writing – review and editing (equal). **Thiago S. Cardoso**: Data curation (equal); Investigation (equal); Visualization (equal); Writing – review and editing (equal). **Bernardo R. Teixeira**: Data curation (equal); Investigation (equal); Writing – review and editing (equal). **Samanta C. C. Xavier**: Data curation (equal); Funding acquisition (supporting); Investigation (equal); Resources (equal); Writing – review and editing (equal). **Rosana Gentile**: Data curation (supporting); Funding acquisition (supporting); Investigation (equal); Project administration (supporting); Resources (supporting); Writing – review and editing (equal). **Ana Maria Jansen**: Data curation (supporting); Funding acquisition (supporting); Investigation (equal); Project administration (supporting); Resources (equal); Writing – review and editing (equal). **Paulo S. D'Andrea**: Data curation (supporting); Funding acquisition (supporting); Investigation (equal); Project administration (supporting); Resources (equal); Writing – review and editing (equal).

Transparent peer review

The peer review history for this article is available at https://publons.com/publon/10.1111/ecog.06579.

Data availability statement

Data are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.n2z34tn1f (Andreazzi et al. 2023).

Supporting information

The Supporting information associated with this article is available with the online version.

References

Abad-Franch, F. and Gurgel-Gonçalves, R. 2021. The ecology and natural history of wild triatominae in the Americas. – In: Guarneri, A. and Lorenzo, M. (eds), Entomology in focus. Springer, pp. 387–445.

Allen, T., Murray, K. A., Zambrana-Torrelio, C., Morse, S. S., Rondinini, C. Di Marco, M., Breit, N., Olival, K. J. and Daszak, P. 2017. Global hotspots and correlates of emerging zoonotic diseases. – Nat. Commun. 8: 1124.

Andreazzi, C. S., Martinez-Vaquero, L. A., Winck, G. R., Cardoso, T. S., Teixeira, B. R., Xavier, S. C. C., Gentile, R., Jansen, A. M. and D'Andrea, P. S. 2023. Data from: Vegetation cover and biodiversity reduce parasite infection in wild hosts across ecological levels and scales. – Dryad Digital Repository, https://doi.org/10.5061/dryad.n2z34tn1f.

Bandeira, B., Jamet, J.-L., Jamet, D. and Ginoux, J.-M. 2013. Mathematical convergences of biodiversity indices. – Ecol. Indic. 29: 522–528.

Becker, D. J., Streicker, D. G. and Altizer, S. 2018. Using host species traits to understand the consequences of resource provisioning for host–parasite interactions. – J. Anim. Ecol. 87: 511–525.

Bovendorp, R. S., Brum, F. T., McCleery, R. A., Baiser, B., Loyola, R., Cianciaruso, M. V. and Galetti, M. 2019. Defaunation and fragmentation erode small mammal diversity dimensions in tropical forests. – Ecography 42: 23–35.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. and West, G. B. 2004. Toward a metabolic theory of ecology. – Ecology 85: 1771–1789.

Bush, A. O., Lafferty, K. D., Lotz, J. M. and Shostak, A. W. 1997. Parasitology meets ecology on its own terms: Margolis et al. revisited. – J. Parasitol. 83: 575–83.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S. and Naeem, S. 2012. Biodiversity loss and its impact on humanity. – Nature 486: 59–67.

Cardoso, T. S., Andreazzi, C. S., Maldonado Junior, A. and Gentile, R. 2021. Functional traits shape small mammal–helminth network: patterns and processes in species interactions. – Parasitology 148: 947–955.

Carreira, J. C. A., Magalhães, M. de A. F. M., Brazil, R. P. and da Silva, A. V. M. 2017. Leishmania in marsupials – an overview of infection records in the Americas and Australia. – Open J. Anim. Sci. 07: 315–343.

Carrero, G. C., Walker, R. T., Simmons, C. S. and Fearnside, P. M. 2022. Land grabbing in the Brazilian Amazon: stealing public land with government approval. – Land Use Policy 120: 106133.

Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P. 2002. SMOTE: synthetic minority over-sampling technique. – J. Artif. Intell. Res. 16: 321–357.

Civitello, D. J., Cohen, J., Fatima, H., Halstead, N. T., Liriano, J., McMahon, T. A., Ortega, C. N., Sauer, E. L., Sehgal, T., Young, S. and Rohr, J. R. 2015. Biodiversity inhibits parasites: broad evidence for the dilution effect. – Ecology 112: 8667–8671.

Correa, J. P., Bacigalupo, A., Fontúrbel, F. E., Oda, E., Cattan, P. E., Solari, A. and Botto-Mahan, C. 2015. Spatial distribution of an infectious disease in a small mammal community. – Sci. Nat. 102: 51.

Dario, M. A., Furtado, C., Lisboa, C. V., de Oliveira, F., Santos, F. M., D'Andrea, P. S., Roque, A. L. R., Xavier, S. C. das C. and Jansen, A. M. 2022. Trypanosomatid richness among rats, opossums and dogs in the Caatinga Biome, Northeast Brazil, a former endemic area of Chagas disease. – Front. Cell. Infect. Microbiol. 12: 851903.

de Souza, J. G., Schaan, D. P., Robinson, M., Barbosa, A. D., Aragão, L. E. O. C., Marimon, B. H., Marimon, B. S., da Silva, I. B., Khan, S. S., Nakahara, F. R. and Iriarte, J. 2018. Pre-Columbian earth-builders settled along the entire southern rim of the Amazon. – Nat. Commun. 9: 1125.

Díaz, S., Tilman, D., Fargione, J., Chapin III, F. S., Dirzo, R., Kitzberger, T., Gemmill, B., Zobel, M., Vilà, M., Mitchell, C., Wilby, A., Daily, G. C., Galetti, M., Laurance, W. F., Pretty, J., Naylor, R., Power, A., Harvell, D., Potts, S., Kremen, C., Griswold, T. and Eardley, C. 2005. Biodiversity regulation of ecosystem services. – In: Ceballos, G., Lavorel, S., Orians, G., Pacala, S. and Supriatna, J. (eds), Ecosystems and human well-being: current state and trends. Island Press, pp. 297–329.

Didan, K. 2021. MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V061. – https://data.nasa.gov/dataset/MODIS-Terra-Vegetation-Indices-Monthly-L3-Global-0/7bj2-mddy.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B. and Collen, B. 2014. Defaunation in the Anthropocene. – Science 345: 401–406.

Elith, J., Leathwick, J. R. and Hastie, T. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.

Estrada-Peña, A., Ostfeld, R. S., Peterson, A. T., Poulin, R. and de la Fuente, J. 2014. Effects of environmental change on zoonotic disease risk: an ecological primer. – Trends Parasitol. 30: 205–214.

Fisher, R. A., Corbet, A. S. and Williams, C. B. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. – J. Anim. Ecol. 12: 42.

Flynn, D. F. B., Gogol-Prokurat, M., Nogeire, T., Molinari, N., Richers, B. T., Lin, B. B., Simpson, N., Mayfield, M. M. and DeClerck, F. 2009. Loss of functional diversity under land use intensification across multiple taxa. – Ecol. Lett. 12: 22–33.

Frainer, A., McKie, B. G., Amundsen, P.-A., Knudsen, R. and Lafferty, K. D. 2018. Parasitism and the biodiversity–functioning relationship. – Trends Ecol. Evol. 33: 260–268.

Ghersi, B. M., Peterson, A. C., Gibson, N. L., Dash, A., Elmayan, A., Schwartzenburg, H., Tu, W., Riegel, C., Herrera, C. and Blum, M. J. 2020. In the heart of the city: *Trypanosoma cruzi* infection prevalence in rodents across New Orleans. – Parasit. Vectors 13: 577.

Glidden, C. K., Nova, N. K., Morgan, P., Lagerstrom, K. M., Skinner, E. B., Mandle, L., Sokolow, S. H., Plowright, R. K., Dirzo, R., De Leo, G. A. and Mordecai, E. A. 2021. Human-mediated impacts on biodiversity and the consequences for zoonotic disease spillover. – Curr. Biol. 31: R1342–R1361.

Gottdenker, N. L., Chaves, L. F., Calzada, J. E., Saldaña, A. and Carroll, C. R. 2012. Host life history strategy, species diversity and habitat influence *Trypanosoma cruzi* vector infection in changing landscapes. – PLoS Negl. Trop. Dis. 6: e1884.

Gottdenker, N. L., Streicker, D. G., Faust, C. L. and Carroll, C. R. 2014. Anthropogenic land use change and infectious diseases: a review of the evidence. – Ecohealth 11: 619–632.

Han, B. A., Park, A. W., Jolles, A. E. and Altizer, S. 2015. Infectious disease transmission and behavioural allometry in wild mammals. – J. Anim. Ecol. 84: 637–646.

Hechinger, R. F. 2015. Parasites help find universal ecological rules. – Proc. Natl Acad. Sci. USA 112: 1656–1657.

Henders, S., Persson, U. M. and Kastner, T. 2015. Trading forests: land-use change and carbon emissions embodied in production and exports of forest-risk commodities. – Environ. Res. Lett. 10: 125012.

Jansen, A. M., Xavier, S. C. C. and Roque, A. L. R. 2015. The multiple and complex and changeable scenarios of the *Trypanosoma cruzi* transmission cycle in the sylvatic environment. – Acta Trop. 151: 1–15.

Jansen, A. M., Xavier, S. C. C. and Roque, A. L. R. 2018. *Trypanosoma cruzi* transmission in the wild and its most important reservoir hosts in Brazil. – Parasit. Vectors 11: 502.

Jetz, W., Carbone, C., Fulford, J. and Brown, J. H. 2004. The scaling of animal space use. – Science 306: 266–268.

Johnson, P. T. J., Ostfeld, R. S. and Keesing, F. 2015. Frontiers in research on biodiversity and disease. – Ecol. Lett. 18: 1119–1133.

Keesing, F. and Ostfeld, R. S. 2021. Impacts of biodiversity and biodiversity loss on zoonotic diseases. – Proc. Natl Acad. Sci. USA 118: e2023540118.

Keesing, F., Holt, R. D. and Ostfeld, R. S. 2006. Effects of species diversity on disease risk. – Ecol. Lett. 9: 485–498.

Keesing, F., Belden, L. K., Daszak, P., Dobson, A., Harvell, C. D., Holt, R. D., Hudson, P., Jolles, A., Jones, K. E., Mitchell, C. E., Myers, S. S., Bogich, T. and Ostfeld, R. S. 2010. Impacts of biodiversity on the emergence and transmission of infectious diseases. – Nature 468: 647–52.

Kissling, W. D., Dalby, L., Fløjgaard, C., Lenoir, J., Sandel, B., Sandom, C., Trøjelsgaard, K. and Svenning, J.-C. 2014. Establishing macroecological trait datasets: digitalization, extrapolation and validation of diet preferences in terrestrial mammals worldwide. – Ecol. Evol. 4: 2913–2930.

Kuris, A. M., Blaustein, A. R. and Alio, J. J. 1980. Hosts as islands. – Am. Nat. 116: 570–586.

Laliberté, E. and Legendre, P. 2010. A distance-based framework for measuring functional diversity from multiple traits. – Ecology 91: 299–305. https://cran.r-project.org/web/packages/FD/index.html.

Lemaitre, G., Nogueira, F. and Aridas, C. K. 2017. Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. – J. Mach. Learn. Res. 18: 1–5.

Magioli, M., Ferraz, K. M. P. M. B., Chiarello, A. G., Galetti, M., Setz, E. Z. F., Paglia, A. P., Abrego, N., Ribeiro, M. C. and Ovaskainen, O. 2021. Land-use changes lead to functional loss of terrestrial mammals in a Neotropical rainforest. – Perspect. Ecol. Conserv. 19: 161–170.

Medeiros-Sousa, A. R., de Oliveira Christe, R., de Castro Duarte, A. M. R., Mucci, L. F., Ceretti-Junior, W. and Marrelli, M. T. 2019. Effects of anthropogenic landscape changes on the abundance and acrodendrophily of *Anopheles* (*Kerteszia*) *cruzii*, the main vector of malaria parasites in the Atlantic Forest in Brazil. – Malar. J. 18: 110.

Menhinick, E. F. 1964. A Comparison of some species-individuals diversity indices applied to samples of field insects. – Ecology 45: 859–861.

Orozco, M. M., Enriquez, G. F., Cardinal, M. V., Piccinali, R. V. and Gürtler, R. E. 2016. A comparative study of *Trypanosoma cruzi* infection in sylvatic mammals from a protected and a disturbed area in the Argentine Chaco. – Acta Trop. 155: 34–42.

Orozco, M. M., Piccinali, R. V., Mora, M. S., Enriquez, G. F., Cardinal, M. V. and Gürtler, R. E. 2014. The role of sigmodontine rodents as sylvatic hosts of *Trypanosoma cruzi* in the Argentinean Chaco. – Infect. Genet. Evol. 22: 12–22.

Palmeirim, A. F., Figueiredo, M. S. L., Grelle, C. E. V., Carbone, C. and Vieira, M. V. 2019. When does habitat fragmentation matter? A biome-wide analysis of small mammals in the Atlantic Forest. – J. Biogeogr. 46: 2811–2825.

Pardini, R., Bueno, A. A., Gardner, T. A., Prado, P. I. and Metzger, J. P. 2010. Beyond the fragmentation threshold hypothesis: regime shifts in biodiversity across fragmented landscapes. – PLoS One 5: e13666.

Pavoine, S., Vallet, J., Dufour, A.-B., Gachet, S. and Daniel, H. 2009. On the challenge of treating various types of variables: application for improving the measurement of functional diversity. – Oikos 118: 391–402.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, É. 2011. Scikit-learn: machine learning in python. – J. Mach. Learn. Res. 12: 2825–2830.

Péneau, J., Nguyen, A., Flores-Ferrer, A., Blanchet, D. and Gourbière, S. 2016. Amazonian triatomine biodiversity and the transmission of Chagas disease in French Guiana: In Medio Stat Sanitas. – PLoS Negl. Trop. Dis. 10: e0004427.

Poulin, R. 2007. Are there general laws in parasite ecology? – Parasitology 134: 763–776.

Poulin, R. 2013. Explaining variability in parasite aggregation levels among host samples. – Parasitology 140: 541–546.

Prist, P. R., Uriarte, M., Tambosi, L. R., Prado, A., Pardini, R., D'Andrea, P. S. and Metzger, J. P. 2016. Landscape, environmental and social predictors of hantavirus risk in São Paulo, Brazil. – PLoS One 11: e0163459.

Prist, P. R., Reverberi Tambosi, L., Filipe Mucci, L., Pinter, A., Pereira de Souza, R., de Lara Muylaert, R., Roger Rhodes, J., Henrique Comin, C., da Fontoura Costa, L., Lang D'Agostini, T., Telles de Deus, J., Pavão, M., Port-Carvalho, M., Del Castillo Saad, L., Mureb Sallum, M. A., Fernandes Spinola, R. M. and Metzger, J. P. 2022. Roads and forest edges facilitate yellow fever virus dispersion. – J. Appl. Ecol. 59: 4–17.

Rademaker, V., Herrera, H. M., Raffel, T. R., D'Andrea, P. S., Freitas, T. P. T., Abreu, U. G. P., Hudson, P. J. and Jansen, A. M. 2009. What is the role of small rodents in the transmission cycle of *Trypanosoma cruzi* and *Trypanosoma evansi* (Kinetoplastida Trypanosomatidae)? A study case in the Brazilian Pantanal. – Acta Trop. 11: 102–107.

Randolph, S. E. and Dobson, A. D. M. 2012. Pangloss revisited: a critique of the dilution effect and the biodiversity–buffers–disease paradigm. – Parasitology 139: 847–863.

Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J. and Hirota, M. M. 2009. The Brazilian Atlantic Forest: how much is left, and how is the remaining forest distributed? Implications for conservation. – Biol. Conserv. 142: 1141–1153.

Rosseel, Y. 2012. lavaan: an R package for structural equation modeling. – J. Stat. Softw. 48: 1–36. https://cran.r-project.org/web/packages/lavaan/index.html.

Rowcliffe, J. M., Jansen, P. A., Kays, R., Kranstauber, B. and Carbone, C. 2016. Wildlife speed cameras: measuring animal travel speed and day range using camera traps. – Remote Sens. Ecol. Conserv. 2: 84–94.

Santos-Filho, M., Peres, C. A., da Silva, D. J. and Sanaiotti, T. M. 2012. Habitat patch and matrix effects on small-mammal persistence in Amazonian forest fragments. – Biodivers. Conserv. 21: 1127–1147.

Scinachi, C. A., Takeda, G. A. C. G., Mucci, L. F. and Pinter, A. 2017. Association of the occurrence of Brazilian spotted fever and Atlantic rain forest fragmentation in the São Paulo metropolitan region, Brazil. – Acta Trop. 166: 225–233.

Shipley, B. 2016. Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference with R. – Cambridge Univ. Press.

Silva Junior, C. H. L., Pessôa, A. C. M., Carvalho, N. S., Reis, J. B. C., Anderson, L. O. and Aragão, L. E. O. C. 2021. The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. – Nat. Ecol. Evol. 5: 144–145.

Solar, R. R. de C., Barlow, J., Ferreira, J., Berenguer, E., Lees, A. C., Thomson, J. R., Louzada, J., Maués, M., Moura, N. G., Oliveira, V. H. F., Chaul, J. C. M., Schoereder, J. H., Vieira, I. C. G., Mac Nally, R. and Gardner, T. A. 2015. How pervasive is biotic homogenization in human-modified tropical forest landscapes? – Ecol. Lett. 18: 1108–1118.

Souza, C. M., Shimbo, J. Z., Rosa, M. R., Parente, L. L., Alencar, A. A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., Ferreira, L. G., Souza-Filho, P. W. M., de Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., Weber, E. J., Lenti, F. E. B., Paternost, F. F., Pareyn, F. G. C., Siqueira, J. V., Viera, J. L., Neto, L. C. F., Saraiva, M. M., Sales, M. H., Salgado, M. P. G., Vasconcelos, R., Galano, S., Mesquita, V. V. and Azevedo, T. 2020. Reconstructing three decades of land use and land cover changes in Brazilian Biomes with Landsat Archive and Earth Engine. – Remote Sens. 12: 2735.

Stella, M., Selakovic, S., Antonioni, A. and Andreazzi, C. S. 2018. Ecological multiplex interactions determine the role of species for parasite spread amplification. – eLife 7: e32814.

Villéger, S., Mason, N. W. H. and Mouillot, D. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. – Ecology 89: 2290–2301.

Wang, S., Dai, Y., Shen, J. and Xuan, J. 2021. Research on expansion and classification of imbalanced data based on SMOTE algorithm. – Sci. Rep. 11: 24039.

Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M. and Jetz, W. 2014. EltonTraits 1.0: species-level foraging attributes of the world's birds and mammals. – Ecology 95: 2027–2027, https://figshare.com/collections/EltonTraits_1_0_Species-level_foraging_attributes_of_the_world_s_birds_and_mammals/3306933.

Wood, C. L., Lafferty, K. D., DeLeo, G., Young, H. S., Hudson, P. J. and Kuris, A. M. 2014. Does biodiversity protect humans against infectious disease? – Ecology 95: 817–832.

Xavier, S. C. das C., Roque, A. L. R., Lima, V. dos S., Monteiro, K. J. L., Otaviano, J. C. R., Ferreira da Silva, L. F. C. and Jansen, A. M. 2012. Lower richness of small wild mammal species and Chagas disease risk. – PLoS Negl. Trop. Dis. 6: e1647.