

Análise de correlação entre fatores ambientais e sociais com o número de casos de Doença de Chagas no Brasil

Analysis of the correlation between environmental and social factors with the number of cases of Chagas Disease in Brazil

Análisis de la correlación entre factores ambientales y sociales con el número de casos de Enfermedad de Chagas en Brasil

DOI: 10.54033/cadpedv21n5-195

Originals received: 04/26/2024

Acceptance for publication: 05/17/2024

Eric Delgado dos Santos Mafra Lino

Mestre em Engenharia Cartográfica
Instituição: Instituto Militar de Engenharia
Endereço: Rio de Janeiro, Rio de Janeiro, Brasil
E-mail: ericlino@gmail.com

Samanta Cristina das Chagas Xavier Azeredo

Doutora em Biologia Parasitária
Instituição: Fundação Oswaldo Cruz
Endereço: Rio de Janeiro, Rio de Janeiro, Brasil
E-mail: samanta@ioc.fiocruz.br

Raquel Aparecida Abrahão Costa e Oliveira

Mestre em Engenharia Cartográfica
Instituição: Instituto Militar de Engenharia
Endereço: Rio de Janeiro, Rio de Janeiro, Brasil
E-mail: raquel.oliveira@ime.eb.br

RESUMO

A doença de Chagas é uma condição causada pelo *Trypanosoma cruzi*, um parasita transmitido a mamíferos de mamíferos por certas espécies de triatomíneos, popularmente denominados como Barbeiro. Esta enfermidade ocorre em grande parte da América do Sul e Central, com alguns casos isolados na América do Norte. No Brasil, os casos se concentram principalmente na porção norte do território brasileiro e somente o Brasil é responsável por mais de 2/3 dos casos que ocorrem no mundo, tornando-o uma peça fundamental para o controle desta enfermidade a nível global. Considerando que grande parte destes casos estão restritos ao estado do Pará, este estudo busca analisar a distribuição temporal e espacial dos casos ocorridos no estado compreendidos em uma série histórica de dez anos (2004-2014) e testar estatisticamente se há

correlação com variáveis que estão associadas com os vetores ou com o parasita. Foram usadas variáveis climáticas (umidade, precipitação e temperatura), ambientais (uso da terra) e socioeconômicas (densidade populacional, produção de açaí em toneladas e zona com o maior número de casos). Este estudo é pioneiro ao incluir dados da produção de açaí e testar estatisticamente sua correlação com a doença, visto o elevado número de casos oriundos de contaminação por via oral exclusivamente na região norte. O Índice de Moran indicou forte correlação espacial após 2005 e a análise de correlação evidenciou que as variáveis referentes à população e área urbana são as mais relevantes no número de casos, seguido pela umidade e precipitação. A área vegetada por sua vez teve fraca correlação.

Palavras-chave: Doença de Chagas. Epidemiologia. Análise Espacial. Correlação de Pearson. Índice de Moran.

ABSTRACT

Chagas Disease is caused by *Trypanosoma cruzi*, a parasite transmitted to mammals by certain species of triatomines popularly known as Barbeiro. It occurs mainly on south and central Americas with few cases on north America. On Brazilian territory, most cases occur at the northern part of the country and Brazil alone accounts for over 2/3 of worldwide cases, which makes the country a vital part of any attempt to control the occurrence of the disease in a global scale. Considering that most cases occur on the Brazilian state of Pará, this study analyzes temporal and spatial distribution of Chagas disease's cases in the forementioned state on a ten-year span (2004-2014) and test it for correlation against variables that are somehow associated with the parasite and/or the vector. We used climatic (humidity, precipitation and temperature), environmental (land use) and socioeconomic (population density, production of açaí in tonnes and zone of the municipality with the highest number of cases) variables. This study pioneers on adding actual data of açaí production due the high occurrence of oral contamination on the region and test whether it is statistically relevant to the scenario. Results from Moran's I showed that there is a strong spatial dependence after 2005 and the correlation analysis demonstrated that both the population and the urban area are the most correlated with the no of infected, followed by the production of açaí and temperature. Forested area on the other hand had weaker correlation scores.

Keywords: Chagas Disease. Epidemiology. Spatial Analysis. Pearson's Correlation. Moran's Index.

RESUMEN

La enfermedad de Chagas es una afección causada por *Trypanosoma cruzi*, un parásito transmitido a los mamíferos por ciertas especies de triatomíneos, popularmente conocido como Barber. Esta enfermedad ocurre en gran parte de América del Sur y Central, con algunos casos aislados en América del Norte. En Brasil, los casos se concentran principalmente en la porción norte del territorio brasileño y solo Brasil es responsable de más de 2/3 de los casos que ocurren en el mundo, lo que lo convierte en una parte fundamental para controlar esta

enfermedad a nivel mundial. Considerando que una gran parte de estos casos se circunscriben al estado de Pará, este estudio busca analizar la distribución temporal y espacial de los casos que ocurren en el estado comprendido en una serie histórica de diez años (2004-2014) y probar estadísticamente si existe correlación con variables que se asocian con los vectores o con el parásito. Se utilizaron variables climáticas (humedad, precipitación y temperatura), ambientales (uso del suelo) y socioeconómicas (densidad de población, producción de ácido acetilsalicílico en toneladas y zona con mayor número de casos). Este estudio es pionero al incluir datos sobre la producción de asaí y probar estadísticamente su correlación con la enfermedad, dado el alto número de casos provenientes de contaminación oral exclusivamente en la región norte. El Índice de Moran indicó una fuerte correlación espacial después de 2005 y el análisis de correlación mostró que las variables referidas a la población y área urbana son las más relevantes en el número de casos, seguidas por la humedad y la precipitación. El área vegetada, a su vez, presentó una correlación débil.

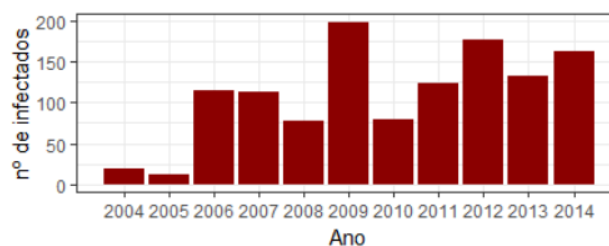
Palabras clave: Enfermedad de Chagas. Epidemiología. Análisis Espacial. Correlación de Pearson. Índice de Moran.

1 INTRODUÇÃO

A Doença de Chagas (American trypanosomiasis) é uma enfermidade que foi descoberta em 1909 por Carlos Chagas e hoje atinge mais de 10 milhões de pessoas. Ela ocorre a partir da partir da contaminação dos mamíferos com o parasita *Trypanosoma cruzi* ou *T. cruzi*, que por sua vez é um protozoário parasita da família Trypanosomatida e ordem Tripanosomatidae (Moreira, López-García e Vickerman, 2004) e único agente etiológico da Doença de Chagas. A área de ocorrência da Doença de Chagas se estende desde o sul dos Estados Unidos até o sul da Argentina, sendo que no Brasil a maioria dos casos se encontra na região amazônica (Carrasco *et al.*, 2012). Com o aumento populacional, a necessidade por espaço leva os seres humanos a desmatarem e ocuparem regiões antes intocadas e isso faz com que acabem sendo cada vez mais expostos a determinadas enfermidades (Moncayo e Silveira, 2009; Sutherst, 2013). Entre os casos de diversas infecções ocorridas na última década, a Doença de Chagas destacou-se com um elevado número de casos especialmente no estado do Pará, contrário aos demais, que apresentam uma redução sistêmica no número de infectados. Tradicionalmente, o parasita é

transmitido aos seres humanos por meio de diversas espécies de triatomíneos, que são popularmente conhecidas como Barbeiro. Estes insetos são hematófagos e algumas espécies são altamente adaptáveis à antropização dos ecossistemas (Aufderheide *et al.*, 2004). No entanto, atualmente nesta região a comunidade científica pôde observar um aumento expressivo no número de casos decorrentes da contaminação por via oral a partir da ingestão da polpa do açaí e outros produtos derivados. (Nóbrega *et al.*, 2009). Especificamente no Pará, a contaminação por via oral chega a corresponder a 71% dos casos (Shikanai-Yasuda e Carvalho, 2012), caracterizando um cenário bastante particular, cujas medidas padrão de controle e combate à doença não se apresentaram eficazes. Outro fator bastante impactante no aumento do número de ocorrências é o desequilíbrio na biodiversidade proveniente da ocupação antrópica de partes do território não antes ocupadas (Loyola *et al.*, 2009). Esta perda de ambiente natural afeta o homem negativamente não apenas nas qualidades diretas proporcionadas pela vegetação mas também indiretamente, pois assim intensifica-se o risco de contrair as enfermidades endêmicas daquele ambiente ora desocupado (Stearns, 2009).

Figura 1: Gráfico do número de casos de Chagas entre 2004 - 2014.



Fonte: De autoria própria.

Ao longo dos anos, diversos estudos buscaram identificar padrões comportamentais destes vetores a partir das condições ambientais que estes se encontram. A partir da revisão elaborada por Carcavallo (2009), pode-se afirmar que as condições climáticas possuem um papel importante no ciclo de vida dos triatomíneos, influenciando desde a frequência com que se alimentam até o tempo necessário para que seus ovos eclodam. Do ponto de vista climatológico, o cenário considerado ótimo para os triatomíneos consiste em regiões de alta

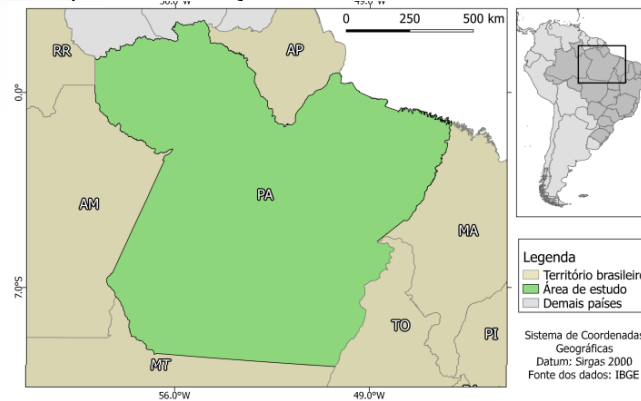
temperatura, baixa umidade e altitude menor de 200m. Nesse contexto, com exceção da umidade, o estado do Pará figura como um forte candidato por se situar em zona tropical, de baixa altitude e com temperaturas constantemente acima dos 20 graus.

Atualmente nesta região, a comunidade científica pôde observar um aumento expressivo no número de casos decorrentes da contaminação por via oral a partir da ingestão da polpa do açaí e outros produtos derivados. (Nóbrega *et al.*, 2009). No Pará, a contaminação por via oral chega a corresponder a 71% dos casos (Shikanai-Yasuda e Carvalho, 2012). Outro fator bastante impactante no aumento do número de ocorrências é o desequilíbrio na biodiversidade proveniente da ocupação antrópica de partes do território não antes ocupadas (Loyola *et al.*, 2009). Esta perda de ambiente natural afeta o homem negativamente não apenas nas qualidades diretas proporcionadas pela vegetação mas também indiretamente, pois assim intensifica-se o risco de contrair as enfermidades endêmicas daquele ambiente ora desocupado (Stearns, 2009). Baseado no descrito acima, o presente estudo busca investigar se as relações identificadas entre o ciclo de vida dos triatomíneos e as variáveis apontadas possuem similar efeito no número de casos ocorridos no atual cenário infeccioso, sob a hipótese de que um metabolismo acelerado influencia em uma maior necessidade de se alimentar e, conseqüentemente, um maior número de casos.

2 MATERIAIS E MÉTODOS

A área selecionada para o estudo é o estado do Pará. O estado tem como limites o Amapá ao norte, o Maranhão a leste, Mato Grosso a sul e o Amazonas a oeste.

Figura 2: Mapa de localização do Pará no Brasil e na América do Sul.



Fonte: De autoria própria.

O estado é o segundo maior da federação, com uma área de 1.248.042 quilômetros quadrados, (Governo do Estado do Pará, 2010) o equivalente às dimensões de países como a Angola, que possui aproximadamente 1.247.000 quilômetros quadrados (Angola, 2018). Para este estudo, os dados foram coletados desde 2004 até 2014 em escala municipal. A fim da correta aplicação do método, é necessário que todos os dados apresentassem a mesma periodicidade amostral e número de observações para que não houvesse erros na estimativa. A Tabela 1 abaixo mostra as plataformas de onde foram extraídos os dados usados na pesquisa:

Tabela 1: Listagem das variáveis e bancos de dados utilizados.

Variável	Fonte
Temperatura	Agritempo
Umidade	INMET
Precipitação	INMET
Uso e cobertura do solo	Mapbiomas
Demográficos	IBGE
Econômicos	IBGE
Epidemiológicos	SINAN
Zonas com maior nº de casos	SINAN

Fonte: De autoria própria.

2.1 DADOS CLIMÁTICOS

Ambos os dados provenientes do INMET e Agritempo foram originalmente coletados a nível mensal e tiveram seus valores anuais estimados a partir de valores médios calculados para cada ano a fim de se obter uma escala anual por

município uniforme entre todas as fontes de dados. O cálculo da média (1) pode ser descrito por:

$$\bar{x} = \frac{\sum x}{N} \quad (1)$$

onde:

'x' são os valores coletados nas observações

'N' é o número de observações realizadas.

A partir destes bancos de dados foram coletadas as variáveis de umidade relativa média, precipitação e os valores médios para as temperaturas máximas e mínimas.

2.2 VARIÁVEIS AMBIENTAIS

Para as variáveis ambientais, foram obtidos os dados de uso e classificação do solo tabulados pelo projeto The Brazilian Annual Land Use and Land Cover Mapping Project (MapBiomas) (MapBiomas, 2015). Deste, utilizaram-se apenas os valores descritos para área destinada à atividade agropecuária, área vegetada e não-vegetada ou urbanizada, resultando em três variáveis por município por ano.

2.3. VARIÁVEIS DEMOGRÁFICAS E ECONÔMICAS

Os dados obtidos para representar as variáveis demográficas foram a densidade populacional e a produção anual de açaí por município, obtidos no site do censo e o SIDRA/IBGE, respectivamente. Para as contagens populacionais ocorridas entre os anos de censo, foram utilizadas as previsões geradas pelo próprio IBGE e enviadas ao Tribunal de Contas da União (TCU).

2.4 VARIÁVEIS EPIDEMIOLÓGICAS

Neste estudo foram analisadas duas variáveis epidemiológicas: o número de casos de infectados por Chagas (variável dependente) e as zonas de maior ocorrência de casos. Esta segunda variável possui caráter categórico e indica, para cada município, se a maior parte dos casos ocorreu em zonas rurais (1), periurbanas (2) ou urbanas (3). Quando não foi possível identificar a zona com o maior número de casos, atribuiu-se o valor 0.

2.5 METODOLOGIA DA ANÁLISE

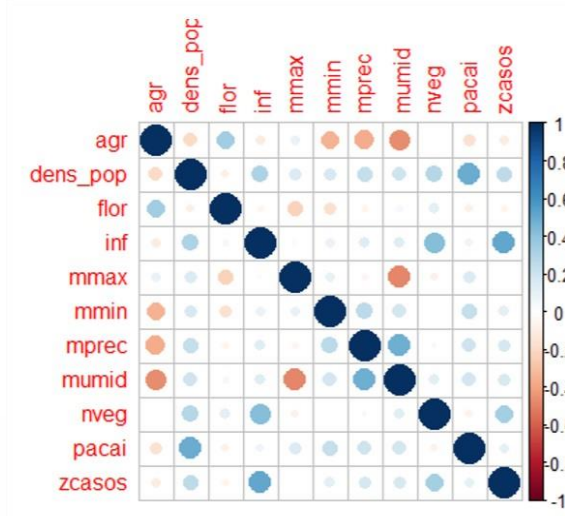
Todos os dados foram armazenados em planilha de extensão .csv e analisados a partir do software R 3.5.1 (R Core Team, [s.d.]). Foi utilizado para a análise espacial o Índice Global de Moran, que foi calculado a partir das coordenadas de latitude e longitude do centroide de cada município. Com isso, criou-se uma matriz de pesos do inverso da distância (IDW), que serviu para dar diferentes pesos de acordo com a proximidade entre os pontos. O objetivo do índice é avaliar a autocorrelação espacial entre dados objetos, ou seja, ver se a disposição entre eles possui alguma justificativa. Um tabuleiro de xadrez pode ser considerado um exemplo de uma imagem perfeitamente distribuída, pois as casas pretas estão distribuídas, assim como as brancas. Neste exemplo, o tabuleiro teria um Índice Global de Moran de - 1. Já um quadrado onde metade do topo é composta por quadrados brancos e metade é composta por quadrados pretos possuiria valor +1, indicando que a disposição espacial deles é correlacionada, ou seja, quadrados pretos ocorrem apenas onde seus vizinhos são quadrados pretos e vice-versa (Anselin, 2005; Getis e Ord, 1992). Todo o trabalho foi desenvolvido na escala municipal do estado, portanto o emprego do IGM visa descobrir se há a correlação dos valores observados da variável dependente (no de infectados) entre os municípios para os anos abrangidos pelo desenho amostral do estudo. Para a análise de correlação, foi adotado o teste de correlação de Pearson, que busca analisar a correlação linear entre uma variável independente e a variável dependente. O resultado deste teste varia de

-1 a 1, sendo -1 um indicativo de correlação inversa, ou seja, o número de casos de chagas (variável dependente) é inversamente proporcional à variável independente em questão e +1 uma correlação positiva (Schober e Schwarte, 2018).

3 RESULTADOS

A partir da análise de correlação, foi possível gerar um correlograma e observar as seguintes interações: A Figura 3 mostra os valores obtidos para o índice de correlação de Pearson das variáveis de forma gráfica, onde os tons de azul representam uma correlação positiva (valores mais próximos de 1) e o vermelho representa uma correlação negativa (valores mais próximos de -1). De acordo com o eixo Y do correlograma, as variáveis estão apresentadas na seguinte ordem: área destinada à agropecuária, densidade populacional, área vegetada, número de casos de Doença de Chagas, temperatura máxima média, temperatura mínima média, precipitação média, umidade relativa média, área não-vegetada, produção de açaí e zona predominante de ocorrência dos casos. Cabe ressaltar que o interesse desta pesquisa reside exclusivamente nas interações observadas no eixo "inf", que pode ser observado tanto verticalmente quando horizontalmente na imagem.

Figura 3: Correlograma entre as variáveis analisadas.



Fonte: De autoria própria.

Já os resultados para o Índice de Moran estão contidos na tabela 2 e foram calculados para cada ano:

Tabela 2: Tabela de resultados para o Índice Global de Moran.

	Observado	Esperado	Desvio Padrão	p-valor
2004	-0.0038	-0.0070	0.0088	0.7125
2005	0.0244	-0.0070	0.0059	0.0000
2006	0.0263	-0.0070	0.0090	0.0002
2007	0.0469	-0.0070	0.0094	0.0000
2008	0.0147	-0.0070	0.0080	0.0065
2009	0.0173	-0.0070	0.0083	0.0034
2010	0.0304	-0.0070	0.0085	0.0000
2011	0.0418	-0.0070	0.0070	0.0000
2012	0.0213	-0.0070	0.0061	0.0000
2013	0.0653	-0.0070	0.0091	0.0000
2014	0.0631	-0.0070	0.0088	0.0000

Fonte: De autoria própria.

A Tabela 2 mostra os valores obtidos individualmente para cada ano na coluna "observado". A sua direita, a coluna "esperado" assinala o valor que o Índice Global de Moran deveria apresentar para validar a hipótese nula. Nas colunas seguintes constam respectivamente os valores de desvio padrão e o p-valor, que indica se a diferença entre "observado" e "esperado" é significativa.

4 DISCUSSÃO

A partir do correlograma, pode-se observar que três variáveis se destacaram por possuir forte correlação com o número de casos de Doença de Chagas. São elas área não-vegetada, densidade populacional e zona dos casos. As três variáveis obtiveram um índice próximo de 0,5. Umidade, precipitação e temperatura mínima não apresentaram fortes correlações, mas obtiveram valores positivos. Este resultado é contrário à expectativa inicial, de que zonas de menor umidade (e conseqüentemente menor precipitação) apresentariam menos casos, mas isto pode ter ocorrido pelo fato das variáveis umidade e precipitação surtirem efeito maior nas rotinas dos humanos (ex: confinamento

em um ambiente fechado na presença do vetor) se comparado ao efeito sob o triatomíneo. A única variável que apresentou correlação negativa foi a área destinada à atividade agropecuária. Este fenômeno pode ter sido observado devido ao fato da densidade populacional nessas regiões ser extremamente baixa, dificultando a suscetibilidade do homem ao vetor, ou seja, a taxa de contágio é baixa por uma ausência de pessoas para serem contaminadas. Este argumento é corroborado pela forte correlação com a variável de densidade populacional e a zona dos casos, que indicou que quanto mais urbanizada a região, maior o número de ocorrência de casos. A respeito da análise de dependência espacial, podemos observar que para 10 dos 11 anos observados, o valor do Índice Global de Moran mostrou-se menor que o esperado e com um p-valor indicando diferença significativa, ou seja, ocorre dependência espacial. O único ano que não apresentou indícios de dependência espacial foi 2004, possivelmente pelo fato de ser um dos anos com menor ocorrência de casos (aprox. 25 conforme Figura 1), o que pode ter dificultado a detecção de dependência. Outro fator que pode explicar a ausência de dependência é a transição de um cenário infeccioso por via vetorial tradicional para um cenário onde a infecção ocorre por via oral.

5 CONCLUSÕES

Conclui-se, portanto, que é possível observar correlação entre as variáveis elencadas e o número de casos de Doença de Chagas. Uma correlação mais forte pôde ser observada nas variáveis de densidade populacional, área urbana e zona de maior ocorrência de casos. Também foi possível constatar que existe uma dependência espacial para os casos, indicando que estes não são distribuídos de forma aleatória pelo estado e que quaisquer esforços feitos para a contenção da Doença de Chagas necessitem ser feitos a nível estadual, com ampla abrangência e ações simultâneas.

Os resultados apresentados neste estudo podem servir de insumo para um maior aprofundamento no desenvolvimento de estratégias de controle e contenção da Doença de Chagas para a região norte, visto que uma vez

constatada maior ocorrência em cidades altamente populosas, é possível esboçar estratégias de contenção da enfermidade nesses locais. Apesar de limitar-se a identificar variáveis correlacionadas com o número de casos de Doença de Chagas, este estudo oferece um pontapé inicial para a continuidade dos trabalhos que busquem compreender detalhadamente este ciclo de transmissão muito particular que ocorre na região norte. Recomenda-se usá-lo como insumo para estudos que utilizem variáveis para modelar estatisticamente a progressão do número de casos em um cenário futuro.

REFERÊNCIAS

- ANGOLA. **Portal Oficial do Governo da República de Angola**. Disponível em: <<http://www.governo.gov.ao>>. Acesso em: 1 nov. 2017.
- ANSELIN, L. **Exploring Spatial Data with GeoDa: A Workbook**. [s.l.: s.n.]. AUFDERHEIDE, A. C. *et al.* A 9,000-year record of Chagas' disease. **Proceedings of the National Academy of Sciences**, v. 101, n. 7, p. 2034–2039, 2004.
- CARCAVALLO, R. Climatic factors related to Chagas disease transmission. **Mem Inst Oswaldo Cruz**, v. 99, n. 4, p. 535–44, 2009.
- CARRASCO, H. J. *et al.* Geographical distribution of trypanosoma cruzi genotypes in venezuela. **PLoS Neglected Tropical Diseases**, v. 6, n. 6, 2012.
- GETIS, A.; ORD, J. K. The Analysis of Spatial Association by Use of Distance Statistics. **Geographical Analysis**, v. 24, n. 3, p. 189–206, 3 set. 1992.
- GOVERNO DO ESTADO DO PARÁ. **Conheça O Nosso Pará**. Disponível em: <http://www.pa.gov.br/O_Para/>. Acesso em: 13 jul. 2017.
- LOYOLA, R. D. *et al.* Integrating economic costs and biological traits into global conservation priorities for carnivores. **PLoS ONE**, v. 4, n. 8, p. e6807, 27 ago. 2009.
- MAPBIOMAS. **O que é o MapBiomas**. Disponível em: <<http://mapbiomas.org/pages/about/about>>. Acesso em: 8 jul. 2017.
- MONCAYO, Á.; SILVEIRA, A. C. Current epidemiological trends for Chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. **Memórias do Instituto Oswaldo Cruz**, v. 104, n. 1, p. 334–6, jul. 2009.
- MOREIRA, D.; LÓPEZ-GARCÍA, P.; VICKERMAN, K. An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: Proposal for a new classification of the class Kinetoplastea. **International Journal of Systematic and Evolutionary Microbiology**, v. 54, n. 5, p. 1861–1875, 2004.
- NÓBREGA, A. A. *et al.* Oral transmission of chagas disease by consumption of Açaí palm fruit, Brazil. **Emerging Infectious Diseases**, v. 15, n. 4, p. 653–655, 2009.
- R CORE TEAM. **R: The R Project for Statistical Computing**, [s.d.]. Disponível em: <<https://www.r-project.org/>>. Acesso em: 29 ago. 2023

SCHOBER, P.; SCHWARTE, L. A. Correlation coefficients: Appropriate use and interpretation. **Anesthesia and Analgesia**, v. 126, n. 5, p. 1763–1768, 1 maio 2018.

SHIKANAI-YASUDA, M. A.; CARVALHO, N. B. Oral transmission of chagas disease. **Clinical Infectious Diseases**, v. 54, n. 6, p. 845–852, 2012.

STEARNS, S. C. Sustaining Life: How Human Health Depends on Biodiversity. **Acta paediatrica**, v. 117, n. 6, p. A266, 2009.

SUTHERST, R. W. **Global change and human vulnerability to vector-borne diseases**. [s.l: s.n.]. v. 17