

Universidad del
Rosario

**Phylogenetic relationships and evolutionary patterns of the genus
Psammolestes (Hemiptera: Reduviidae).**

Mateo Andrés Alvarado López

Director

Juan David Ramírez González PhD.

Co-director

Camilo Andrés Salazar Clavijo PhD.

Trabajo presentado como requisito para optar por el título de Biólogo

Facultad de Ciencias Naturales

Pregrado en Biología

Universidad del Rosario

Bogotá, 2021

TITLE: Phylogenetic relationships and evolutionary patterns of the genus Psammolestes (Hemiptera: Reduviidae).

ABSTRACT	3
INTRODUCTION.....	4
MATERIALS AND METHODS	5
<i>Sampling.....</i>	<i>5</i>
<i>Extraction, amplification and alignment of DNA data.....</i>	<i>5</i>
<i>Molecular phylogenetic analysis</i>	<i>6</i>
<i>Species delimitation tests</i>	<i>6</i>
<i>Population genetics analyses</i>	<i>7</i>
<i>Population structure</i>	<i>8</i>
<i>Environmental niche modeling</i>	<i>8</i>
<i>Species distribution modelling</i>	<i>8</i>
<i>Environmental variables</i>	<i>9</i>
<i>Environmental niche of the parental species</i>	<i>9</i>
<i>Assignment of the contribution of each environmental variable</i>	<i>10</i>
RESULTS.....	10
<i>Molecular phylogenetics.....</i>	<i>10</i>
<i>Species delimitation tests</i>	<i>10</i>
<i>Population genetics analyses</i>	<i>11</i>
<i>Environmental niche modeling</i>	<i>12</i>
DISCUSSION.....	12
REFERENCES	15
TABLES	30
FIGURES	31
SUPPLEMENTARY MATERIAL.....	39

**Phylogenetic relationships and evolutionary patterns of the genus *Psammolestes*
(Hemiptera: Reduviidae)**

Mateo Alvarado¹, Carolina Hernández¹, Fabián Salgado², Camilo Salazar³, Nathalia Ballesteros¹, Nicol Rueda³, Jader Oliveira⁴, Joao Aristeu da Rosa⁴, Plutarco Urbano⁵, Juan David Ramírez^{1*}

¹ Centro de Investigaciones en Microbiología y Biotecnología-UR (CIMIBIUR), Facultad de Ciencias Naturales, Universidad del Rosario, Bogotá, Colombia.

² School of Biosciences, University of Melbourne, Melbourne, Australia.

³ Grupo de Genética Evolutiva y Filogeografía, Departamento de Biología, Facultad de Ciencias Naturales y Matemáticas, Universidad del Rosario, Bogotá, Colombia.

⁴ Universidade Estadual Paulista (UNESP), Faculdade de Ciências Farmacêuticas, Araraquara, Sao Paulo 01000, Brazil.

⁵ Grupo de Investigaciones Biológicas de la Orinoquia, Fundación Universitaria Internacional del Trópico Americano (Unitropico), Yopal, Colombia.

*Correspondence: juand.ramirez@urosario.edu.co

ABSTRACT

The family Reduviidae (Hemiptera: Heteroptera) is among the most diverse families of the true bugs. The evolution and phylogenetic relationships of Rhodniini and Triatomini tribes (Triatominae) are well studied due to their epidemiological relevance as vectors of *Trypanosoma cruzi*, the parasite that causes the Chagas disease. Rhodniini is composed by the genera *Rhodnius* and *Psammolestes*, where the genetic diversity of the second one remains to be studied in comparison with *Rhodnius*, the main vector of *T. cruzi*. Therefore, we gathered 92 samples in total, 38 for *Psammolestes arthuri* in Colombia, 24 for *Psammolestes tertius* and 30 for *Psammolestes coreodes* in Brazil. We used five novel nuclear loci: tRNA Guanine (37) -N (1) methyl transferase (TRNA), Putative juvenile hormone inducible protein (PJH), Probable cytosolic iron sulfur protein assembly protein Ciao 1 (CISP), Lipoyl synthase, mitochondrial (LSM) and Uncharacterized protein for cell adhesion (UPCA), along with two previously reported loci: 28S and CYTB, to depict the phylogenetic relationships and the evolutionary patterns of the genus *Psammolestes*. Four of the seven gene topologies were not consistent with the concatenated topology, while the other three were concordant, but the general pattern is clear: *Psammolestes* is a monophyletic group, corroborating hypotheses previously suggested for the genus. Clustering analysis along with population genetics summary statistics resulted in the delimitation of three different populations. These three clusters corresponded to each one of the *Psammolestes* species known *a priori* -defined by morphology, ecology and cytogenetic methods- which suggests that populations for each one of the species has a well-supported genetic structure. Overall, our results corroborated the existence of the three previously described *Psammolestes* species,

showing that they probably diverged in allopatry, under the influence of the Guyana shield and the Amazon basin as barriers to dispersal.

INTRODUCTION

The kissing bugs, insects of the family Reduviidae (Hemiptera: Heteroptera), are among the most diverse bug families between the true bugs (Heteroptera) (Weirauch, 2008). The subfamily Triatominae excels between the subfamilies of Reduviidae due to their hematophagous behavior, but specially for being vectors of *Trypanosoma cruzi*, which causes the Chagas disease (Hwang & Weirauch, 2012). This subfamily comprises 154 species –three of which are extinct- and all of the extant species are considered to be potential vectors of *T. cruzi* (Galvão & Carcavallo, 2003). Some of the factors that make Triatominae insects potential vectors of *T. cruzi* include their feeding source, feeding speed, insect size and abiotic factors like temperature, altitude, and domiciliation capacity (Asin & Catalá, 1995; Gorchakov *et al.*, 2016; Castillo-Neyra *et al.*, 2015; Sandoval *et al.*, 2004; Waleckx *et al.*, 2015). The disease can be caused by contact of the insect's faeces infected with the protozoan *T. cruzi*, and the blood or mucosa of vertebrates (Chagas, 1909), and is one of the main public health and economic concerns in Latin America (Bern, 2015).

The Rhodniini tribe, one of the five tribes within the Triatominae subfamily, is divided in two different genera: *Rhodnius* and *Psammolestes*, which are distributed across South and Central America (Ceccarelli *et al.*, 2018). This tribe, along the Triatomini tribe, are the most diverse and therefore, the best studied tribes of the subfamily Triatominae, especially for their role as vectors of *Trypanosoma cruzi*. However, evolutionary research has been focused on the epidemiologically relevant species of the Rhodniini tribe (Schofield & Galvão, 2009). Also, the number of investigations aiming to understand the events behind the diversification of the tribe are limited (Abad-Franch *et al.*, 2009; Justi *et al.*, 2017; Dujardin 1999). These works are needed as control strategies for *T. cruzi* vectors are the most effective mechanisms to regulate the spread of the disease, considering that there is no vaccine or effective treatment for the Chagas disease (Schofield *et al.*, 2006; Townson *et al.*, 2005; Golding *et al.*, 2015). In consequence, vector control strategies require a better overall understanding of the vector, as this information will improve the evidence-based decision making on strategies, as the IVM (Integrated vector management) or the ECLAT (European Community and Latin American network for research on Triatominae) (Wilson *et al.*, 2020; Dias *et al.*, 2002; Dujardin, 2007).

Insects of the *Psammolestes* genus are found across multiple geographic regions that correspond to South American countries like Colombia, Venezuela, Brazil, and Paraguay (Ceccarelli *et al.*, 2018). Lent & Wygodzinsky in 1979 first described three *Psammolestes* species based on morphological and ecological traits, on what has been the most relevant taxonomical revision of the genus to date. However, besides of the morphological based delimitation, there is no robust molecular evidence about the existence of three different *Psammolestes* species. Nonetheless, the three species of *Psammolestes* have been included in multiple studies that aimed to describe the evolutionary relationships of the tribe (Bargues

et al., 2002a; De Paula *et al.*, 2005, 2007; Hypša *et al.*, 2002; Justi *et al.*, 2016; Lyman *et al.*, 1999; Fernando Araujo Monteiro *et al.*, 2018; Patterson & Gaunt, 2010), but only one work has focused on describing the relationships among the three species of *Psammolestes* (Oliveira *et al.*, 2018). The lack of research focused on the *Psammolestes* genus is a gap yet to be filled, especially considering that there are existing reports of *P. arthuri* infected with *T. cruzi*, reinforcing the idea that these insects might be potential vectors of the parasite (Velásquez-Ortiz *et al.*, 2019). Also, to date there are no existing investigations focused on understanding the phylogenetic and biogeographic relations, as well as the evolutionary patterns between the three *Psammolestes* species.

This study presents a first and novel comprehensive study of *Psammolestes* including phylogenetic, genetic diversity and population structure methods, in order to test two hypotheses. The first hypothesis is that *P. tertius* and *P. coreodes* are genetically more closely related between them compared to how closely related they are to *P. arthuri*. The foundation of this hypothesis is merely based on the geographic distribution of the *Psammolestes* species, as the lack of overall information about this genus did not allow us to formulate a hypothesis based on phylogenetic information. The second hypothesis considers the existence of a geographical barrier and unique niche conditions that have shaped the demography of *Psammolestes*. In general, our results support the hypotheses previously mentioned, and propose the existence of a geographical barrier that contributes to the presence of a genetic discontinuity between the *Psammolestes* species. In addition, our results disclose the role of climatic niche traits on the overall demography of the *Psammolestes* genus. More importantly, this study is the first of its kind on the genus *Psammolestes*, aiming to fill the knowledge gap that covers the evolutionary and demography patterns of its three species, adding information that might be relevant for the establishment of vector control strategies in the future.

MATERIALS AND METHODS

Sampling

We collected, with the collaboration of international research groups, a total of 92 individuals of the three *Psammolestes*' species from 12 localities in Venezuela, Colombia and Brazil (Fig. 1). Thirty-nine *P. arthuri* samples were collected in Colombia and Venezuela, specifically in the Colombian regions of Poré, Arauca, Paz de Ariporo, Monterrey, Maní and Tamará, and in the Venezuelan region of Maracay. Twenty-five *P. tertius* samples were collected in Brazil, in the regions of Minas Gerais and Bahia. Finally, twenty-seven *P. coreodes* samples were collected from the Brazilian region of Mato Grosso do Sul. The samples obtained were preserved in absolute ethanol and stored at -20°C, until needed.

Extraction, amplification and alignment of DNA data

The extraction of DNA was made from leg tissue, using the DNeasy® Blood & Tissue kit, with some modifications made to the manufacturer's protocol. Four new nuclear loci were used to explore phylogenetic relations between *Psammolestes* insects: tRNA Guanine (37) - N (1) methyltransferase (TRNA), Putative juvenile hormone inducible protein (PJH), Probable cytosolic iron sulfur protein assembly protein Ciao 1 (CISP), Lipoyl synthase,

mitochondrial (LSM), along with the previously reported Uncharacterized protein for cell adhesion (UPCA) (Caicedo-Garzón *et al.*, 2019; Nascimento *et al.*, 2019). In addition, we also included two loci previously reported to assess phylogenetic relations on insects of the Rhodniini tribe, 28S and CYTB (Monteiro *et al.*, 1999, 2000). Primers used for the amplification of these loci are listed on table S2. Amplicons were visualized on a 1% agarose gel, and positive amplicons for each locus were bidirectionally sequenced by the Sanger method in Macrogen Korea. Contigs were assembled, checked and edited in CLC Main Workbench 20.0 (<https://digitalinsights.qiagen.com>) and SeqMan NGen®. Version 12.0. (DNASTAR. Madison, WI). To resolve ambiguities in each of the individual sequences obtained from the assemblies, we used PHASE, included on DnaSp v6.12.03 (Rozas *et al.*, 2017) to implement a haplotype inference algorithm with 1,000 iterations per simulation. Sequence alignment per locus was performed using Mesquite (Maddison & Maddison, 2018), and the results were visually inspected and manually corrected if necessary using the same program.

Molecular phylogenetic analysis

To determine the phylogenetic relationships among the individuals sampled from the three *Psammolestes* species, we reconstructed maximum likelihood (ML) trees using IQ-Tree 2 (Minh *et al.*, 2020). For each locus, and for a concatenated alignment generated with the seven loci, we used *Rhodnius prolixus* as the outgroup. Substitution models per locus were accessed by the Bayesian Information Criterion (BIC; Schwarz, 1978) obtained from ModelFinder (Kalyaanamoorthy *et al.*, 2017) analyses. Then, we performed the phylogenetic reconstruction for each locus considering the result of the ModelFinder run, calculating the node support by using 10,000 UltraFast Bootstrap pseudoreplicates (Hoang *et al.*, 2018). Also, we obtained two additional node supports: aBayes (Anisimova *et al.*, 2011) and SH-aLRT (Guindon *et al.*, 2010), with the same number of pseudoreplicates as the bootstrap. After each run was finished, the convergence of the bootstrap pseudoreplicates was visually confirmed. For the concatenate analysis, an additional partition file was generated to indicate the size (in characters) of each locus and its substitution model, but the run parameters were the same as in the previous reconstructions. However, node supports were calculated by resampling both the partitions and the sites within the resampled partitions (Gadagkar *et al.*, 2005).

Species delimitation tests

To test if each *Psammolestes* species really constitutes different species at the molecular level, we used two delimitation methods: The Bayesian Phylogenetics and Phylogeography method (BPP; Yang, 2015) and the multi-rate Poisson Tree Processes method (mPTP; Kapli *et al.*, 2017). For the BPP analysis, two different inputs were generated: the first input concatenates six nuclear loci, and the other one contains the mitochondrial CYTB data, as recommended in the program. The objective of splitting the input files is to deal with the differences in the mutation rates (nuclear vs. mitochondrial rates) and in the effective population size (Yang, 2015). Then, a species tree estimation and a joint species delimitation were performed (Yang & Rannala, 2014), assigning individuals to a “species” based on the results of the phylogenetic trees previously constructed. Four models were tested, each one with a different combination of priors achieved by switching the distribution of the divergence times (τ) and

the population size parameters (θ). This combination of parameters aims to evaluate if different evolutionary scenarios influence the delimitation of the *Psammolestes* species. The models used here include combinations between large population sizes ($\theta = G(1, 10)$), shallow population sizes ($\theta = G(2, 2000)$), deep divergence times ($\tau = G(1, 10)$) and shallow divergence times ($\tau = G(2, 2000)$). Each analysis used 100,000 iterations per run, sampling every 2 iterations, and using 10% of the iterations in the chain as burn-in.

For the mPTP method, the ML concatenate tree previously obtained was used as the input file. The first step on this method is to calculate the minimum branch length on the concatenate ML tree, in order to correct the potential error generated by the sensitivity of this test to false positives when similar sequences are present. Then, using the value generated for the minimum branch length, we ran 10 MCMC replicates of 100,000,000 iterations each, sampling every 1,000 iterations; 10% of these iterations were used as the burn-in for the test. Finally, following the considerations stipulated by Satler *et al.* (2013), congruent results for both of the species delimitation methods were considered as independent species.

Population genetics analyses

We calculated summary statistics using DNAsp v6.12.03 to characterize and analyze variability at the genetic level of the three species of *Psammolestes* for the seven loci used in this study. The summary statistics calculated were: number of haplotypes (h), number of segregating sites (S), population mutation rate plus or minus one standard deviation ($\theta \pm 1$ S.D.) and average pairwise distance plus or minus one standard deviation ($\pi \pm 1$ S.D.). In addition, a relative measure (F_{ST}) and two absolute measures (D_a , D_{xy}) were calculated to estimate the genetic differentiation between *Psammolestes* species, and represented through heat maps. To evaluate deviations from panmixia, the Hudson permutation test (Hudson, Boos, & Kaplan, 1992) was used with 1,000 replicates, included as a feature on DNAsp. TCS haplotype networks (Clement *et al.*, 2002) were reconstructed for each locus to evaluate the relationships among haplotypes using PopArt v1.7 (Leigh, Bryant, & Nakagawa, 2015). To explore possible signatures of population expansion or contraction, we computed three neutrality tests: Ramos-Onsins and Rozas R_2 (R_2 ; (Ramos-Onsins & Rozas, 2002)), Tajima's D (D ; (Tajima, 1989)) and Fu & Li's F and D statistics (FF , FD ;(Fu & Li, 1993)).

To assess the possible relationship between genetic distances and geographic distances, we used two different approaches. The first approach applies the Monmonier's algorithm (Manni, Guérard, & Heyer, 2004) in the R package *adegenet* (Jombart & Ahmed, 2011) using a Delaunay triangulation to detect possible boundaries associated with geographic barriers. These possible boundaries are marked by discontinuities on the molecular and the geographic distance matrix for each locus, and for the concatenate file. To represent the results of the Monmonier's algorithm approach in a graphical way, we reconstructed the results using QGIS v3.16.0 (QGIS.org, 2021. QGIS Geographic Information System. QGIS Association. <http://www.qgis.org>) and Adobe Illustrator (Adobe Inc., 2019). The second approach consists on the evaluation of an isolation by distance scenario constructed using the Mantel test in the R package *vegan* (Dixon, 2003). For the sake of the test, we linearized the genetic distances using an F_{ST} transformation, symbolized as $1/1-F_{ST}$, and calculated the geographical distances using the function *distm* in the R package *geosphere* (Hijmans, 2019). Together, these two approaches aid to cover the limitations shown by the Mantel test (Mantel,

1967), and make a robust result to either correlate the geographic and genetic distances or to state that there is no relation between these variables.

Population structure

To demonstrate the presence of population structure and identify distinct genetic populations, we ran the program STRUCTURE v2.3.4 (Pritchard *et al.*, 2000). In this case, the input file was constructed as indicated by the STRUCTURE manual, with the help of the program PGDSpider v2.1.1.5 (Lischer & Excoffier, 2012) to transform the FASTA files from each locus into a haplotype STRUCTURE file. With each one of the STRUCTURE files per locus, a final unique STRUCTURE file was constructed indicating the a priori population in the first column, either or not to use the a priori assignment on the second column, the sampling site on the third column, and the rest of the columns corresponding to the haplotype data of the seven loci used in this work. The addition of the first columns representing the a priori population assignment and the sampling location aim to assist the clustering assignment based on empirical geographical information. Missing data was labeled with -9, as indicated by the program manual. We ran the program with the ADMIXTURE model including 100,000 MCMC iterations per K, using K values from 1 to 10, and applying 5 iterations per K with the same MCMC length, along with a burn-in length of 100,000. To find the K value that best fits the data, we followed the protocol proposed by Evanno *et al.* (2005), using the STRUCTURE HARVESTER (Earl, 2012; <http://taylor0.biology.ucla.edu/structureHarvester/>). Within the results shown by the web page, we looked for the modal value of the distribution of ΔK , as this value aids the location of the real K value. After obtaining the K value that best fits the data based on Evanno's methodology, we retrieved the CLUMPP file generated with the STRUCTURE HARVESTER, and ran CLUMPP (Jakobsson & Rosenberg, 2007). CLUMPP aligns the cluster membership coefficients resulted from the replicate cluster analyses of the STRUCTURE run, facilitating the interpretation of the data. In addition, CLUMPP aids the STRUCTURE result analysis by solving two different discrepancies that might affect the clustering solutions: "label switching" and "genuine multimodality" (Jakobsson & Rosenberg, 2007). Outputs from CLUMPP include an aligned cluster membership coefficient matrix for the Q values for both the population matrix and the individual matrix used as inputs. Finally, *distrupt* (Rosenberg, 2004) was used to graphically display the results from the CLUMPP program in a uniform graphic that considers the matrix of Q values from the a priori populations defined on STRUCTURE and the individual Q values matrix.

Environmental niche modeling

Species distribution modelling

A major limitation of modelling techniques is the lack of absence data (Chefaoui & Lobo, 2008; Wisz & Guisan, 2009), and as our data only had records of presence data, we created another database that included randomly generated pseudo-absences (Graham *et al.*, 2004). We generated the database using QGIS v3.16.0, this generated double the number of pseudo-absences compared to the number of presences for the three *Psammolestes* species. To aid the model construction, we used the tool BIOMOD 2 (Thuiller *et al.*, 2009) to develop distribution models for each one of the *Psammolestes* species using four algorithms. The

algorithms used were: Artificial Neural Networks (ANN; Ripley 1996), Generalized Linear Models (GLM; (McCullagh & Nelder, 1989), Generalized Boosting Models (GBM; Friedman *et al.*, 2000), and Maximum Entropy Models (MAXENT; Phillips *et al.*, 2006). The foundation of why we chose those specific algorithms is related to the fact that they are recognized as four of the most effective correlative model species distribution algorithms (Elith *et al.*, 2006), as well as that they offer a wide variety of approaches. Calibration of the data was made using 80% of the occurrence data and taking this percentage to be evaluated on the remaining 20%, and this cross-validation was repeated three times. The main probabilities of these predictions were averaged using an assemble approach. An equal weighting for presences and pseudo-absences (prevalence weights =0.5) was applied as recommended (Barbet-Massin *et al.*, 2012).

Finally, three different assembly models were generated for the three *Psammolestes* species based on the combination of the four models produced by the previously mentioned algorithms. Two metrics were used to choose the model that best predicts the distribution of the taxa: The True Skill Statistic (TSS) and the area under the curve (AUC) of the receiver-operating characteristic (ROC) (Thuiller *et al.*, 2003). To conclude, we used R and QGIS v3.16.0 to produce the potential distribution maps using the best assembly model.

Environmental variables

Nineteen climatological variables were recovered and used from CHELSA (Climatologies at High Resolution for earth's land surface areas) at spatial resolution of 1 Km (Karger *et al.*, 2017) to describe climatic variation across the occurrence range of *Psammolestes*. The high predictive power of these layers mainly in mountain regions was the main reason why they were used (Bobrowski & Udo, 2017). Topographic variable altitude values were obtained from Reuter, Nelson & Jarvis (2007). Pearson's correlation coefficient from the R package *corrplot* (Wei *et al.*, 2017) was applied to each pairwise comparison that included the 19 climatic variables and the altitude values in order to identify and remove highly correlated variables (Elith *et al.*, 2006; Wisz & Guisan, 2009). The six variables with correlation values to fit the species distribution models <0.5 were used. In this case, the values were: Altitude, annual mean temperature, isothermality, annual range of temperature, annual precipitation and precipitation seasonality.

Environmental niche of the parental species

Phillips *et al.* (2006) concept of environmental niche was the basis of the interpretation of our results. They propose that a niche consists of the subset of conditions previously present, and where environmental conditions at the occurrence localities constitute samples from the niche. R package *humboldt* (Brown & Carnaval, 2019) was used in order to estimate the equivalency of the environmental niche between the three *Psammolestes* species. A PCA was used to visualize the environmental overlap (E - space). *Schoener's D*, which ranges between 0 and 1, meaning no overlap and full overlap respectively, was calculated (Rödder and Engler, 2011). To test the significance of this metric, we compared the realized niche overlap against a null distribution of 1,000 randomly generated overlaps from the reshuffled occurrence dataset and tested whether niche background and niche equivalency were different from the

expectations by chance at $\alpha=0.05$ (Brown & Carnaval, 2019). This was done using the entire distribution of entities under comparison (niche overlap test = NOT) and using only the distribution where they overlap (niche divergence test = NDT) (Brown & Carnaval, 2019).

Assignment of the contribution of each environmental variable

The relative importance values of the six variables provided by BIOMOD 2 were used to evaluate the influence of each variable within all the models generated (Thuiller *et al.*, 2009). A randomization process was used to estimate the importance, where BIOMOD2 calculates the correlation between a prediction using all variables and a prediction where the independent variable being tested is randomly removed; this is repeated for each of the six variables. To calculate the relative importance, we subtracted this correlation from one, and therefore, higher values in the subtractions are the best predictors for the model (Thuiller *et al.*, 2009). Finally, we used a Wilcoxon signed rank to test differences among the medians of the values resulting from the calculation of the mean relative importance for each of the six environmental variables across species and subspecies.

RESULTS

Molecular phylogenetics

The ML gene reconstructions used the most suitable substitution models resulting from the Bayesian Information Criterion included in ModelFinder. The resulting substitution models for each locus were the following: HKY+F for 28S and TRNA, F81+F+I for LSM, F81+F for CISP, HKY+F+I for PJH, K2P for UPCA, and HKY+F+G4 for CYTB. The resulting ML gene topologies were not consistent, where the CYTB and PJH topologies (Fig. S3, Fig. S5) recovered *Psammolestes coreodes* and *Psammolestes tertius* as sister monophyletic clades, while 28S, CISP, LSM, TRNA and UPCA topologies (Fig. S1, Fig. S2, Fig. S4, Fig. S6, Fig. S7) didn't recover them as reciprocally monophyletic. However, in the seven gene topologies, *Psammolestes arthuri* was a well-supported monophyletic clade. Whereas the ML reconstruction using all the data recovered *P. coreodes* and *P. tertius* as sister species and *P. arthuri* as sister of the clade formed by both of them. Overall, the three *Psammolestes* species were monophyletic groups with strong node supports (Fig. 2).

Species delimitation tests

The posterior probabilities resulting from the four tested models in BPP for nDNA data were 0.9950 for deep divergence and large population size, 1 for deep divergence and small population size, 1 for shallow divergence and large population size and finally 1 for shallow divergence and small population size. For the nDNA runs, the only model that did not delimit three *Psammolestes* species was the model of deep divergence and large population size, which instead delimited two species: *Psammolestes arthuri* as the first species, and *Psammolestes coreodes* together with *Psammolestes tertius* as the second one. On the other hand, the posterior probabilities for the same models tested on mtDNA data were the following: 1 for deep divergence and large population size, 1 for deep divergence and small

population size, 0.8043 for shallow divergence and large population size, and finally 1 for shallow divergence and small population size. In comparison to the results generated with the nDNA data, the results obtained with the mtDNA data delimited three species in each one of the four models tested.

The mPTP delimitation was strongly supported (ASV=0.87), the MCMC chains converged on the same delimitation distributions (ASDDSV < 0.0000001). and three different *Psammolestes* species were delimited. Hence, the mPTP and the BPP results were concordant and suggest the same number of species.

Population genetics analyses

Population mutation rate (θ) and average pairwise distance (π) were overall low for the three species of *Psammolestes* in all of the seven loci, indicating highly conserved genetic regions (Table 1). However, in loci like 28S and CYTB where the sampling number was low in *Psammolestes tertius*, it is possible that sample size could play a major role in the calculation of these measures. Overall, the neutrality tests, even if not all of them were statistically significant, suggest an expansion process for all the three species of *Psammolestes*. This result is supported by the pattern found in the haplotype networks (Fig. 3), where central haplotypes are coupled with multiple haplotypes with singletons, suggesting that the three species are likely under expansion. The inconsistency on the statistical significance of the neutrality tests may indicate that that sampling size may play an important role on the calculation and interpretation of the expansion signal provided by these statistics, due to the fact that Fu and Li's statistics works better with bigger population sizes compared to the R_2 .

Haplotype networks were congruent with the results of the phylogenetic reconstructions. *Psammolestes arthuri* was recovered as an independent clade in the seven loci, while *Psammolestes coreodes* and *Psammolestes tertius* were more closely related to each other in comparison to *Psammolestes arthuri* (Fig 3). In agreement with the signal from the phylogenetic reconstructions and the haplotype networks, we detected strong genetic differentiation between *Psammolestes arthuri* with both *Psammolestes coreodes* and *Psammolestes tertius* (Fig. S8-S14). However, the genetic differentiation between *Psammolestes coreodes* and *Psammolestes tertius*, was considerably less than that of these with *P. arthuri*, except in CYTB, where there is a strong signal of genetic differentiation between the three *Psammolestes* species (Fig. S10).

Furthermore, a geographic barrier was greatly supported by the results of Monmonier's algorithm in all of the seven loci (Fig. S1-S7) and in the concatenate test (Fig. 2). These results were congruent between them, and they coincided with the location of the Amazon river and the Guiana shield. In addition, regional clusters were explained by the isolation by distance pattern, where the structure degree generated by the Guiana shield and the Amazon basin yields a higher structure for *P. arthuri* compared to *P. tertius* and *P. coreodes* (Fig. S15, Table S1).

Genetic clustering algorithms suggest the existence of three independent structured populations in *Psammolestes* (Fig. 4). The results here were consistent with the results from

BPP and mPTP. STRUCTURE, along with the Evanno method suggested that three populations (K) are the most probable number of clusters (Supplementary information, Fig. S16, Fig. S17). This result is concordant with the results from the phylogenetic reconstructions and the results from the haplotype networks. The permutation resulting from the CLUMPP algorithm was well supported, with an H value of 0.95, consistent for almost all the possible permutations of K=3. Overall, the results from the population structure algorithms support the hypothesis that the three *Psammolestes* species are genetically different from each other.

Environmental niche modeling

The best selected model highlighted areas with varying levels of suitability, where the areas in which the three *Psammolestes* species are distributed, are the areas with the highest probability of occurrence in the model. This model also identified a strong break in the distribution of species in the Amazon, where the conditions seem to be non-suitable for any of the three *Psammolestes* species. This result agrees with the results yielded in the phylogenetic and species delimitation tests. In addition, there is a different environmental variable playing a role on the distribution of each one of the species, where annual precipitation seems to be the main environmental feature influencing *P. tertius*, annual range of temperature for *P. coreodes*, and isothermality for *P. arthuri* (Fig. 5).

The D values of the equivalency test for the NOT and the NDT were 0.033, 0.001 and 0.003 respectively, indicating no niche overlap between the three *Psammolestes* species (Table 2). The D values for the background test were not significant. Overall, the results suggest that the genetic divergence between *Psammolestes* species is influenced by niche divergence.

DISCUSSION

Understanding the factors involved in evolution of tropical neglected diseases vectors are relevant to implement efficient vector control strategies (Bern, 2015). Previous phylogenetic relationships reconstructed for the Rhodniini tribe have indicated that *Psammolestes* was paraphyletic respect to *Rhodnius* (Bargues *et al.*, 2002b; De Paula *et al.*, 2005, 2007; Hypša *et al.*, 2002; Justi *et al.*, 2016; Lyman *et al.*, 1999; F. A. Monteiro *et al.*, 2018; Patterson & Gaunt, 2010). However, there are no existing investigations focused on depicting the phylogenetic relationships and evolutionary patterns of the *Psammolestes* species. Description of the phylogenetic relationships and evolutionary patterns of *Psammolestes* will be useful to provide information that allows the characterization of the Triatominae subfamily. Also, it provides information that will be useful for the establishment of vector control strategies specific for the *Psammolestes* species. This is especially important considering the existing reports of *Psammolestes arthuri* naturally infected with *T. cruzi* that suggest insects of this genus might be vectors of the parasite (José *et al.*, 2018; Velásquez-Ortiz *et al.*, 2019).

In this study, we corroborated the existence of three genetically structured *Psammolestes* species. Along with these findings, we identified the Guiana shield as an absolute barrier for

the dispersal of the *Psammolestes* species along the Amazonian rainforest, and the relevance of the environmental variables that deny the dispersal of these insects across other geographical regions. The genus *Psammolestes* is one of the two genus, along *Rhodnius*, that comprises the Rhodniini tribe (Ceccarelli *et al.*, 2018; F. A. Monteiro *et al.*, 2018). Based on ecological, morphological and molecular traits, *Psammolestes* has been considered as a monophyletic clade (Lent & Wygodzinsky, 1979; F. A. Monteiro *et al.*, 2000). A previous study, and the only one to date focused exclusively on the *Psammolestes* genus, corroborated the monophyletic status of the *Psammolestes* genus through cytogenetic analyses (Oliveira *et al.*, 2018). This report is consistent with our findings that the *Psammolestes* genus seems to be monophyletic, which was based on the phylogenetic analysis performed here. However, in order to confirm the previous affirmation, it is necessary to assess its relationship with the rest of the tribe. On the other hand, multiple studies have provided different hypotheses for the geographical regions that might have played an important role on the diversification not only of the Rhodniini tribe, but for the Triatominae subfamily (Abad-Franch *et al.*, 2009, 2015; Abad-Franch & Monteiro, 2007; De Paula *et al.*, 2007; Lent & Wygodzinsky, 1979). As a general pattern, these studies suggest the presence of the cordilleras, the Amazon basin, and the Guyana shield as some of the most prolific geographical barriers that have shaped the diversity of the subfamily. Based on our results, we support that the Guyana shield acts as an absolute barrier to the dispersal of *Psammolestes* (Fig. S1-S7, Fig. 2), and its role as a vicariant barrier has been successfully tested (Abad-Franch *et al.*, 2009).

Gene tree topologies yielded in this study were conflicting as four of the seven gene topologies recovered *Psammolestes coreodes* as paraphyletic respect to *Psammolestes tertius* (Fig. S1-S7). These conflicts can arise as product of multiple factors like introgression, incomplete lineage sorting, and conflicting models of selection and inheritance (e.g. mitochondrial vs. nuclear genes) (Edwards, 2009; Maddison, 1997). Also, the fact that genes do not share the same evolutionary histories may be one of the principal reasons why gene topologies are sometimes conflicting when compared to species topologies (Funk & Omland, 2003). Going even further, the development of methods to produce high throughput sequencing data also brings easily overlooked issues that can play a role on the discordance of tree topologies, for example: ambiguous data and discordance between most likely gene tree vs. species tree, gene choice, missing data, and sampling size (De La Torre-bárcena *et al.*, 2009; Degnan & Rosenberg, 2006; Lemmon *et al.*, 2009; Sanderson *et al.*, 2010; Townsend *et al.*, 2011). Reconstruction methods could also play a potential negative role on the gene tree topology reconstruction. Also, the previous limitations affect analyses like haplotype networks, Bayesian phylogenetic methods, and summary statistics calculations in the same way that they affect tree topologies (Pollock *et al.*, 2002). In order to avoid these potential conflicts, we suggest gathering a robust sampling size of $n > 30$ per population, as multiple studies demonstrated that an increased taxon sampling greatly reduce phylogenetic errors including more loci or even genomic data (Zwickl & Hillis, 2002).

The close relationship between *Psammolestes coreodes* and *Psammolestes tertius* is consistent throughout the analyses, and suggest that these species may share ancestry, either by genetic flow or incomplete lineage sorting. Other results that support this hypothesis are the results of the calculation of relative and absolute measures (i.e. F_{ST} , D_{XY} , D_a), where genetic structure was clear between *Psammolestes arthuri* with both *Psammolestes coreodes* and *Psammolestes tertius* (Table 1). Nonetheless, the genetic structure was not clear between

Psammolestes coreodes and *Psammolestes tertius*, which might be explained by a phenomenon known as the “grey zone” of speciation where any absolute measure $< 0.5\%$ corresponds to populations of the same species (Roux *et al.*, 2016). This result is consistent for all the calculations across all loci excepting for CYTB, whose different result might be explained by the phenomenon of models of selection and inheritance of mitochondrial genes (Edwards, 2009).

A positive result yielded from this study is related to the results from the species delimitation tests. Multiple studies have shown that these tests are efficient only when working with smaller data sets (Ence & Carstens, 2011; Yang & Rannala, 2014), which is the case on this investigation. It has been suggested, as multiple delimitation species methods use different approaches, to use more than one of these methods if the aim is to delimit species based on genetic data (Carstens *et al.*, 2013). The objective of this recommendation is to encourage researchers to increase the sturdiness of the results of these tests, and that is justified on the fact that if two methods yield the same result, it is more likely that the species are being correctly delimited. This study followed the recommendation previously mentioned, and we successfully delimited three species with two different methods, which corresponded to the three extant *Psammolestes* species, resulting in increased credibility for the results produced. Furthermore, these tests have been applied previously on *Rhodnius* species in order to establish if *Rhodnius taquarussuensis* and *Rhodnius neglectus* were the same species (Nascimento *et al.*, 2019), successfully concluding that they were one single *Rhodnius* species. But, only one delimitation species test was applied on that study, which, as mentioned previously, can lead to inaccurate results as a bias of the delimitation method used. Supporting the results from the species delimitation test, the STRUCTURE results recovered three clearly separated genetic clusters corresponding to the prior groups (i.e. species) (Fig. 4). These results indicate a high level of genetic structure, that could be explained by differences on ecological, behavioral or environmental features that may have shaped the insect’s genetic regions studied (Lobo, 2008). Future demographic models with genomic data, as well as characterization of reproductive isolation could eventually provide more insights about the speciation mechanisms within this genus.

There are multiple factors that can affect the significance of the results generated on a research work. In this study, the major limitation was the sample size, specifically talking about the sample size in some loci for *Psammolestes tertius*. This limitation can drastically change the results or patterns that are present on the reality, and that because of the limited sampling size, it is not going to be reflected on the analysis performed here. Sampling size has major implications on phylogenetic and phylogeographic analysis, because it has been demonstrated that a bigger sampling size reduces the error and the bias on multiple genetic analysis (Pollock *et al.*, 2002; Zwickl & Hillis, 2002). Another major limitation is the lack of data about reproductive isolation, which could allow us to provide a more confident conclusion about the existence of three different species of *Psammolestes*. However, despite these limitations, this is a pioneer study depicting the phylogenetic relationships in *Psammolestes*.

Considering the results produced here, we also evaluated potential improvement opportunities and knowledge gaps left or not covered by this work. One of the main questions generated by this work, is the question surrounding the role of the Pebas system and the

Guyana shield on the diversification events of the *Psammolestes* species. The lack of information like gene substitution rates or the absence of a fossil that would allow us to generate a coalescent approach to estimate a date for the divergence of the species, was also one of the biggest limitations of this study. The need to evaluate whether the Pebas system and the Guyana shield played a role on the diversification of the species, is not only useful for *Psammolestes*, but it can shed light on other evolutionary processes for other species of the tribe or the subfamily.

Finally, the results here presented support the existence of three genetically well-structured species of the *Psammolestes* genus, that also support the monophyly of this genus. The Guyana shield acts as an absolute barrier that blocks the dispersal of the *Psammolestes* species to the Amazonian rainforest. This study is the first of its kind that describes the phylogenetic relationships and evolutionary patterns in the *Psammolestes* genus.

REFERENCES

- Abad-Franch, F., Lima, M. M., Sarquis, O., Gurgel-Gonçalves, R., Sánchez-Martín, M., Calzada, J., Saldaña, A., Monteiro, F. A., Palomeque, F. S., Santos, W. S., Angulo, V. M., Esteban, L., Dias, F. B. S., Diotaiuti, L., Bar, M. E., & Gottdenker, N. L. (2015). On palms, bugs, and Chagas disease in the Americas. *Acta Tropica*, *151*, 126-141. <https://doi.org/10.1016/j.actatropica.2015.07.005>
- Abad-Franch, F., & Monteiro, F. A. (2007b). Biogeography and evolution of Amazonian triatomines (Heteroptera: Reduviidae): implications for Chagas disease surveillance in humid forest ecoregions. *Memórias Do Instituto Oswaldo Cruz*, *102*, 57-70. <https://doi.org/10.1590/S0074-02762007005000108>
- Abad-Franch, F., Monteiro, F. A., Jaramillo O., N., Gurgel-Gonçalves, R., Dias, F. B. S., & Diotaiuti, L. (2009b). Ecology, evolution, and the long-term surveillance of vector-borne Chagas disease: A multi-scale appraisal of the tribe Rhodniini (Triatominae). *Acta Tropica*, *110*(2), 159-177. <https://doi.org/10.1016/j.actatropica.2008.06.005>
- Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., & Gascuel, O. (2011). Survey of Branch Support Methods Demonstrates Accuracy, Power, and Robustness of Fast

- Likelihood-based Approximation Schemes. *Systematic Biology*, 60(5), 685-699.
<https://doi.org/10.1093/sysbio/syr041>
- Asin, S., & Catalá, S. (1995). Development of *Trypanosoma cruzi* in *Triatoma infestans*: Influence of Temperature and Blood Consumption. *The Journal of parasitology*, 81, 1-7. <https://doi.org/10.2307/3283997>
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, 3(2), 327-338. <https://doi.org/10.1111/j.2041-210X.2011.00172.x>
- Bargues, M. D., Marcilla, A., Dujardin, J. P., & Mas-Coma, S. (2002c). Triatomine vectors of *Trypanosoma cruzi*: A molecular perspective based on nuclear ribosomal DNA markers. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 96, S159-S164. [https://doi.org/10.1016/S0035-9203\(02\)90069-6](https://doi.org/10.1016/S0035-9203(02)90069-6)
- Bern, C. (2015). Chagas' Disease. *New England Journal of Medicine*, 373(5), 456-466.
<https://doi.org/10.1056/NEJMra1410150>
- Bobrowski, M., & Schickhoff, U. (2017). Why input matters: Selection of climate data sets for modelling the potential distribution of a treeline species in the Himalayan region. *Ecological Modelling*, 359, 92-102.
<https://doi.org/10.1016/j.ecolmodel.2017.05.021>
- Brown, J. L., & Carnaval, A. C. (2019). A tale of two niches: Methods, concepts, and evolution. *Frontiers of Biogeography*.
- Caicedo-Garzón, V., Salgado-Roa, F. C., Sánchez-Herrera, M., Hernández, C., Arias-Giraldo, L. M., García, L., Vallejo, G., Cantillo, O., Tovar, C., Rosa, J. A. da, Carrasco, H. J., Segovia, M., Salazar, C., & Ramírez, J. D. (2019). Genetic

- diversification of *Panstrongylus geniculatus* (Reduviidae: Triatominae) in northern South America. *PLOS ONE*, *14*(10), e0223963.
<https://doi.org/10.1371/journal.pone.0223963>
- Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013b). How to fail at species delimitation. *Molecular Ecology*, *22*(17), 4369-4383.
<https://doi.org/10.1111/mec.12413>
- Castillo-Neyra, R., Barbu, C. M., Salazar, R., Borrini, K., Naquira, C., & Levy, M. Z. (2015). Host-Seeking Behavior and Dispersal of *Triatoma infestans*, a Vector of Chagas Disease, under Semi-field Conditions. *PLOS Neglected Tropical Diseases*, *9*(1), e3433. <https://doi.org/10.1371/journal.pntd.0003433>
- Ceccarelli, S., Balsalobre, A., Medone, P., Cano, M. E., Gurgel Gonçalves, R., Feliciangeli, D., Vezzani, D., Wisnivesky-Colli, C., Gorla, D. E., Marti, G. A., & Rabinovich, J. E. (2018). DataTri, a database of American triatomine species occurrence. *Scientific Data*, *5*(1), 180071. <https://doi.org/10.1038/sdata.2018.71>
- Chagas, Carlos. (1909). Nova tripanozomiaze humana: Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem. *Memórias do Instituto Oswaldo Cruz*, *1*(2), 159-218.
<https://doi.org/10.1590/S0074-02761909000200008>
- Chefaoui, R. M., & Lobo, J. M. (2008b). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, *210*(4), 478-486.
<https://doi.org/10.1016/j.ecolmodel.2007.08.010>
- Clement, M., Snell, Q., Walker, P., Posada, D., & Crandall, K. (s. f.). *TCS: Estimating Gene Genealogies*. 7.

- Cruz-Guzmán, P. J., Morocoima, A., Chique, J. D., Ramonis-Quintero, J., Uzcátegui, M. T., & Carrasco, H. J. (s. f.). *Psammolestes arthuri* NATURALMENTE INFECTADO CON *Trypanosoma cruzi* ENCONTRADO EN SIMPATRÍA CON *Rhodnius prolixus* Y *Triatoma maculata* EN NIDOS DE AVES EN EL ESTADO ANZOÁTEGUI, VENEZUELA. 14.
- de Paula, A. S., Diotaiuti, L., & Galvão, C. (2007). Systematics and biogeography of *Rhodniini* (Heteroptera: Reduviidae: Triatominae) based on 16S mitochondrial rDNA sequences. *Journal of Biogeography*, 34(4), 699-712.
<https://doi.org/10.1111/j.1365-2699.2006.01628.x>
- de Paula, A. S., Diotaiuti, L., & Scho, C. J. (2005). Testing the sister-group relationship of the *Rhodniini* and *Triatomini* (Insecta: Hemiptera: Reduviidae: Triatominae). *Molecular Phylogenetics and Evolution*, 8.
- Degnan, J. H., & Rosenberg, N. A. (2006b). Discordance of Species Trees with Their Most Likely Gene Trees. *PLOS Genetics*, 2(5), e68.
<https://doi.org/10.1371/journal.pgen.0020068>
- Dias, J. C. P., Silveira, A. C., & Schofield, C. J. (2002). The impact of Chagas disease control in Latin America: A review. *Memórias do Instituto Oswaldo Cruz*, 97(5), 603-612. <https://doi.org/10.1590/S0074-02762002000500002>
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 14(6), 927-930. <https://doi.org/10.1111/j.1654-1103.2003.tb02228.x>
- Dujardin, J. P., Chavez, T., Moreno, J. M., Machane, M., Noireau, F., & Schofield, C. J. (1999). Comparison of Isoenzyme Electrophoresis and Morphometric Analysis for Phylogenetic Reconstruction of the *Rhodniini* (Hemiptera: Reduviidae:

Triatominae). *Journal of Medical Entomology*, 36(6), 653-659.

<https://doi.org/10.1093/jmedent/36.6.653>

Dujardin, J.-P., Beard, C. B., & Ryckman, R. (2007). The relevance of wing geometry in entomological surveillance of Triatominae, vectors of Chagas disease. *Infection, Genetics and Evolution*, 7(2), 161-167.

<https://doi.org/10.1016/j.meegid.2006.07.005>

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359-361.

<https://doi.org/10.1007/s12686-011-9548-7>

Edwards, S. V. (2009). Natural selection and phylogenetic analysis. *Proceedings of the National Academy of Sciences*, 106(22), 8799-8800.

<https://doi.org/10.1073/pnas.0904103106>

Elith, J., Graham*, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. M. M., Peterson, A. T., ... Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2), 129-151.

<https://doi.org/10.1111/j.2006.0906-7590.04596.x>

Ence, D. D., & Carstens, B. C. (2011b). SpedeSTEM: A rapid and accurate method for species delimitation. *Molecular Ecology Resources*, 11(3), 473-480.

<https://doi.org/10.1111/j.1755-0998.2010.02947.x>

- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611-2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Friedman, J. H. (2006). Recent advances in predictive (machine) learning. *Journal of classification*, *23*(2), 175-197.
- Fu, Y. X., & Li, W. H. (1993). Statistical tests of neutrality of mutations. *Genetics*, *133*(3), 693-709.
- Funk, D. J., & Omland, K. E. (2003b). Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 397-423. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132421>
- Gadagkar, S. R., Rosenberg, M. S., & Kumar, S. (2005). Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *304B*(1), 64-74. <https://doi.org/10.1002/jez.b.21026>
- Galvão, C., Carcavallo, R., Rocha, D. D. S., & Jurberg, J. (2003). A checklist of the current valid species of the subfamily Triatominae Jeannel, 1919 (Hemiptera, Reduviidae) and their geographical distribution, with nomenclatural and taxonomic notes. *Zootaxa*, *202*(1), 1-36. <https://doi.org/10.11646/zootaxa.202.1.1>
- Golding, N., Wilson, A. L., Moyes, C. L., Cano, J., Pigott, D. M., Velayudhan, R., Brooker, S. J., Smith, D. L., Hay, S. I., & Lindsay, S. W. (2015). Integrating vector control across diseases. *BMC Medicine*, *13*(1), 249. <https://doi.org/10.1186/s12916-015-0491-4>

- Gorchakov, R., Trosclair, L. P., Wozniak, E. J., Feria, P. T., Garcia, M. N., Gunter, S. M., & Murray, K. O. (2016). Trypanosoma cruzi Infection Prevalence and Bloodmeal Analysis in Triatomine Vectors of Chagas Disease From Rural Peridomestic Locations in Texas, 2013–2014. *Journal of Medical Entomology*, *53*(4), 911-918. <https://doi.org/10.1093/jme/tjw040>
- Graham, Catherine H., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. T. (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, *19*(9), 497-503. <https://doi.org/10.1016/j.tree.2004.07.006>
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307-321. <https://doi.org/10.1093/sysbio/syq010>
- Hijmans, R. J., Williams, E., Vennes, C., & Hijmans, M. R. J. (2017). Package ‘geosphere’. *Spherical trigonometry*, *1*, 7.
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518-522. <https://doi.org/10.1093/molbev/msx281>
- Hudson, R. R., Boos, D. D., & Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, *9*(1), 138-151. <https://doi.org/10.1093/oxfordjournals.molbev.a040703>
- Hwang, W. S., & Weirauch, C. (2012b). Evolutionary History of Assassin Bugs (Insecta: Hemiptera: Reduviidae): Insights from Divergence Dating and Ancestral State

Reconstruction. *PLOS ONE*, 7(9), e45523.

<https://doi.org/10.1371/journal.pone.0045523>

Hypša, V., Tietz, D. F., Zrzavý, J., Rego, R. O. M., Galvao, C., & Jurberg, J. (2002b).

Phylogeny and biogeography of Triatominae (Hemiptera: Reduviidae): molecular evidence of a New World origin of the Asiatic clade. *Molecular Phylogenetics and Evolution*, 23(3), 447-457. [https://doi.org/10.1016/S1055-7903\(02\)00023-4](https://doi.org/10.1016/S1055-7903(02)00023-4)

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801-1806.

<https://doi.org/10.1093/bioinformatics/btm233>

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070-3071.

<https://doi.org/10.1093/bioinformatics/btr521>

Justi, S. A., & Galvão, C. (2017). The Evolutionary Origin of Diversity in Chagas Disease Vectors. *Trends in Parasitology*, 33(1), 42-52.

<https://doi.org/10.1016/j.pt.2016.11.002>

Justi, S. A., Galvão, C., & Schrago, C. G. (2016b). Geological Changes of the Americas and their Influence on the Diversification of the Neotropical Kissing Bugs (Hemiptera: Reduviidae: Triatominae). *PLOS Neglected Tropical Diseases*, 10(4),

e0004527. <https://doi.org/10.1371/journal.pntd.0004527>

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., von Haeseler, A., & Jermin, L. S. (2017).

ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nature methods*, 14(6), 587-589. <https://doi.org/10.1038/nmeth.4285>

- Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, 33(11), 1630-1638. <https://doi.org/10.1093/bioinformatics/btx025>
- Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4(1), 170122. <https://doi.org/10.1038/sdata.2017.122>
- Leigh, J. W., & Bryant, D. (2015). popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110-1116. <https://doi.org/10.1111/2041-210X.12410>
- Lemmon, A. R., Brown, J. M., Stanger-Hall, K., & Lemmon, E. M. (2009b). The Effect of Ambiguous Data on Phylogenetic Estimates Obtained by Maximum Likelihood and Bayesian Inference. *Systematic Biology*, 58(1), 130-145. <https://doi.org/10.1093/sysbio/syp017>
- Lent, H., & Wygodzinsky, P. (1979). Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. *Bulletin of the American Museum of Natural History*, 163(3), 123-520.
- Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298-299. <https://doi.org/10.1093/bioinformatics/btr642>
- Lobo, I. (2008b). Environmental Influences on Gene Expression. *Nature Education*, 39.
- Lyman, D. F., Monteiro, F. A., Escalante, A. A., Cordon-Rosales, C., Wesson, D. M., Dujardin, J. P., & Beard, C. B. (1999). Mitochondrial DNA sequence variation

- among triatomine vectors of Chagas' disease. *The American Journal of Tropical Medicine and Hygiene*, 60(3), 377-386. <https://doi.org/10.4269/ajtmh.1999.60.377>
- Maddison, W. P., & Maddison, D. R. (2019). *Mesquite: A modular system for evolutionary analysis. Version 3.51. 2018.*
- Maddison, Wayne P. (1997). Gene Trees in Species Trees. *Systematic Biology*, 46(3), 523-536. <https://doi.org/10.1093/sysbio/46.3.523>
- Manni, F., Rard, E. G., & Heyer, E. (s. f.). *Geographic Patterns of (Genetic, Morphologic, Linguistic) Variation: How Barriers Can Be Detected by Using Monmonier's Algorithm*. 18.
- Mantel, N. (1967). The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Research*, 27(2 Part 1), 209-220.
- McCullagh, P. (81). Nelder. JA (1989), Generalized Linear Models. *CRC Monographs on Statistics & Applied Probability, Springer Verlag, New York.*
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, 37(5), 1530-1534. <https://doi.org/10.1093/molbev/msaa015>
- Monteiro, F. A., Wesson, D. M., Dotson, E. M., Schofield, C. J., & Beard, C. B. (2000b). Phylogeny and molecular taxonomy of the Rhodniini derived from mitochondrial and nuclear DNA sequences. *The American Journal of Tropical Medicine and Hygiene*, 62(4), 460-465. <https://doi.org/10.4269/ajtmh.2000.62.460>
- Monteiro, Fernando Araujo, Weirauch, C., Felix, M., Lazoski, C., & Abad-Franch, F. (2018b). Chapter Five—Evolution, Systematics, and Biogeography of the Triatominae, Vectors of Chagas Disease. En D. Rollinson & J. R. Stothard (Eds.),

Advances in Parasitology (Vol. 99, pp. 265-344). Academic Press.

<https://doi.org/10.1016/bs.apar.2017.12.002>

Nascimento, J. D., Rosa, J. A. da, Salgado-Roa, F. C., Hernández, C., Pardo-Díaz, C., Alevi, K. C. C., Ravazi, A., Oliveira, J. de, Oliveira, M. T. V. de A., Salazar, C., & Ramírez, J. D. (2019). Taxonomical over splitting in the *Rhodnius prolixus* (Insecta: Hemiptera: Reduviidae) clade: Are *R. taquarussuensis* (da Rosa et al., 2017) and *R. neglectus* (Lent, 1954) the same species? *PLOS ONE*, *14*(2), e0211285.

<https://doi.org/10.1371/journal.pone.0211285>

Oliveira, J., Alevi, K. C. C., Ravazi, A., Herrera, H. M., Santos, F. M., Azeredo-Oliveira, M. T. V. de, & Rosa, J. A. da. (2018). New Evidence of the Monophyletic Relationship of the Genus *Psammolestes* Bergroth, 1911 (Hemiptera, Reduviidae, Triatominae). *The American Journal of Tropical Medicine and Hygiene*, *99*(6), 1485-1488. <https://doi.org/10.4269/ajtmh.18-0109>

Patterson, J. S., & Gaunt, M. W. (2010b). Phylogenetic multi-locus codon models and molecular clocks reveal the monophyly of haematophagous reduviid bugs and their evolution at the formation of South America. *Molecular Phylogenetics and Evolution*, *56*(2), 608-621. <https://doi.org/10.1016/j.ympev.2010.04.038>

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3), 231-259.

<https://doi.org/10.1016/j.ecolmodel.2005.03.026>

Pollock, D. D., Zwickl, D. J., McGuire, J. A., & Hillis, D. M. (2002b). Increased Taxon Sampling Is Advantageous for Phylogenetic Inference. *Systematic biology*, *51*(4), 664-671. <https://doi.org/10.1080/10635150290102357>

- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945-959.
- Ramos-Onsins, S. E., & Rozas, J. (2002). Statistical Properties of New Neutrality Tests Against Population Growth. *Molecular Biology and Evolution*, *19*(12), 2092-2100. <https://doi.org/10.1093/oxfordjournals.molbev.a004034>
- Reuter, H. I., Nelson, A., & Jarvis, A. (2007). An evaluation of void-filling interpolation methods for SRTM data. *International Journal of Geographical Information Science*, *21*(9), 983-1008. <https://doi.org/10.1080/13658810601169899>
- Ripley, B. D., & Naylor, P. (1996). Pattern Recognition and Neural Networks. *Nature*, *381*(6579), 206-206.
- Rödger, D., & Engler, J. O. (2011). Quantitative metrics of overlaps in Grinnellian niches: Advances and possible drawbacks. *Global Ecology and Biogeography*, *20*(6), 915-927. <https://doi.org/10.1111/j.1466-8238.2011.00659.x>
- Rosenberg, N. A. (2004). distruct: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*(1), 137-138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLOS Biology*, *14*(12), e2000234. <https://doi.org/10.1371/journal.pbio.2000234>
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution*, *34*(12), 3299-3302. <https://doi.org/10.1093/molbev/msx248>

- Sanderson, M. J., McMahon, M. M., & Steel, M. (2010b). Phylogenomics with incomplete taxon coverage: The limits to inference. *BMC Evolutionary Biology*, *10*(1), 155.
<https://doi.org/10.1186/1471-2148-10-155>
- Sandoval, C. M., Duarte, R., Gutiérrez, R., Rocha, D. da S., Angulo, V. M., Esteban, L., Reyes, M., Jurberg, J., & Galvão, C. (2004). Feeding sources and natural infection of *Belminus herreri* (Hemiptera, Reduviidae, Triatominae) from dwellings in Cesar, Colombia. *Memórias do Instituto Oswaldo Cruz*, *99*(2), 137-140.
<https://doi.org/10.1590/S0074-02762004000200004>
- Schofield, C. J., & Galvão, C. (2009). Classification, evolution, and species groups within the Triatominae. *Acta Tropica*, *110*(2), 88-100.
<https://doi.org/10.1016/j.actatropica.2009.01.010>
- Schofield, Chris J., Jannin, J., & Salvatella, R. (2006). The future of Chagas disease control. *Trends in Parasitology*, *22*(12), 583-588.
<https://doi.org/10.1016/j.pt.2006.09.011>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, *6*(2), 461-464. <https://doi.org/10.1214/aos/1176344136>
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585-595.
- Thuiller, Wilfried, Araújo, M. B., & Lavorel, S. (2003). Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, *14*(5), 669-680.
<https://doi.org/10.1111/j.1654-1103.2003.tb02199.x>

- Thuiller, Wilfried, Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, 32(3), 369-373. <https://doi.org/10.1111/j.1600-0587.2008.05742.x>
- Torre-Bárcena, J. E. de la, Kolokotronis, S.-O., Lee, E. K., Stevenson, D. W., Brenner, E. D., Katari, M. S., Coruzzi, G. M., & DeSalle, R. (2009). The Impact of Outgroup Choice and Missing Data on Major Seed Plant Phylogenetics Using Genome-Wide EST Data. *PLOS ONE*, 4(6), e5764. <https://doi.org/10.1371/journal.pone.0005764>
- Townsend, T. M., Mulcahy, D. G., Noonan, B. P., Sites, J. W., Kuczynski, C. A., Wiens, J. J., & Reeder, T. W. (2011). Phylogeny of iguanian lizards inferred from 29 nuclear loci, and a comparison of concatenated and species-tree approaches for an ancient, rapid radiation. *Molecular Phylogenetics and Evolution*, 61(2), 363-380. <https://doi.org/10.1016/j.ympev.2011.07.008>
- Townson, H., Nathan, M. B., Zaim, M., Guillet, P., Manga, L., Bos, R., & Kindhauser, M. (2005). Exploiting the potential of vector control for disease prevention. *Bulletin of the World Health Organization*, 83, 942-947. <https://doi.org/10.1590/S0042-96862005001200017>
- Velásquez-Ortiz, N., Hernández, C., Herrera, G., Cruz-Saavedra, L., Higuera, A., Arias-Giraldo, L. M., Urbano, P., Cuervo, A., Teherán, A., & Ramírez, J. D. (2019b). Trypanosoma cruzi infection, discrete typing units and feeding sources among Psammolestes arthuri (Reduviidae: Triatominae) collected in eastern Colombia. *Parasites & Vectors*, 12(1), 157. <https://doi.org/10.1186/s13071-019-3422-y>
- Waleckx, E., Gourbière, S., Dumonteil, E., Waleckx, E., Gourbière, S., & Dumonteil, E. (2015). Intrusive versus domiciliated triatomines and the challenge of adapting

- vector control practices against Chagas disease. *Memórias do Instituto Oswaldo Cruz*, 110(3), 324-338. <https://doi.org/10.1590/0074-02760140409>
- Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., & Zemla, J. (2017). Package ‘corrplot’. *Statistician*, 56(316), e24.
- Weirauch, C. (2008b). Cladistic analysis of Reduviidae (Heteroptera: Cimicomorpha) based on morphological characters. *Systematic Entomology*, 33(2), 229-274. <https://doi.org/10.1111/j.1365-3113.2007.00417.x>
- Wilson, A. L., Courtenay, O., Kelly-Hope, L. A., Scott, T. W., Takken, W., Torr, S. J., & Lindsay, S. W. (2020). The importance of vector control for the control and elimination of vector-borne diseases. *PLOS Neglected Tropical Diseases*, 14(1), e0007831. <https://doi.org/10.1371/journal.pntd.0007831>
- Wisz, M. S., & Guisan, A. (2009b). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, 9(1), 8. <https://doi.org/10.1186/1472-6785-9-8>
- Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, 61(5), 854-865. <https://doi.org/10.1093/czoolo/61.5.854>
- Yang, Z., & Rannala, B. (2014b). Unguided Species Delimitation Using DNA Sequence Data from Multiple Loci. *Molecular Biology and Evolution*, 31(12), 3125-3135. <https://doi.org/10.1093/molbev/msu279>
- Zwickl, D. J., & Hillis, D. M. (2002b). Increased Taxon Sampling Greatly Reduces Phylogenetic Error. *Systematic Biology*, 51(4), 588-598. <https://doi.org/10.1080/10635150290102339>

TABLES

Table 1. Population genetics summary statistics for each species per locus. n: number of samples, h: number of haplotypes, S: number of segregating sites, θ : population mutation rate, π : average pairwise distance, D_T : Tajima's D, R_2 : Ramos-Onsins & Rozas' R_2 , Fu & Li's F, and Fu & Li's D. “*” symbolizes $p < 0.05$ and “^” symbolizes $p < 0.02$.

Statistic	28S			CISP			CYTB			LSM			PJH			TRNA			UPCA		
	P.art	P.cor	P.ter																		
n	35	16	4	20	28	25	13	21	7	36	26	23	37	28	30	50	39	22	40	29	16
h	7	6	3	7	19	9	6	6	7	8	9	7	5	8	8	15	10	7	14	11	3
S	8	6	5	6	19	11	11	27	27	29	6	7	14	17	5	38	9	7	43	12	2
θ	0.0035	0.0032	0.0049	0.0027	0.0080	0.0048	0.0070	0.0128	0.0222	0.0099	0.0022	0.0027	0.0051	0.0067	0.0019	0.0132	0.0033	0.0030	0.0167	0.0050	0.0010
π	0.0012	0.0020	0.0045	0.0012	0.0058	0.0045	0.0057	0.0082	0.0206	0.0051	0.0020	0.0018	0.0013	0.0034	0.0018	0.0051	0.0024	0.0020	0.0052	0.0056	0.0005
D_T	-1.92*	-1.27*	-0.79*	-1.71*	-0.97*	-0.21	-0.76	-1.39	-0.40	-1.68*	-0.31	-0.96	-2.38*	-1.71*	-0.15	-2.06*	-0.74	-0.97	-2.42*	0.36	-1.03
R_2	0.07	0.10	0.25	0.07*	0.08*	0.12	0.12	0.08*	0.17	0.05*	0.11	0.09	0.12	0.12	0.11	0.04*	0.08	0.09	0.13	0.13	0.13
Fu and Li's F	-2.82*	-1.53	-0.75	-1.85	-1.10	0.71	-0.59	-0.67	-0.66	-0.05	-0.33	-0.86	-3.93^	-3.00*	0.16	-1.78	-0.68	-1.38	-4.74^	0.19	-0.73
Fu and Li's D	-2.63*	-1.36	-0.79	-1.55	-0.93	0.96	-0.43	-0.23	-0.65	0.78	-0.27	-0.66	-3.79^	-2.94*	0.26	-1.17	-0.50	-1.27	-4.82^	0.07	-0.50

Table 2. Results of the niche overlap test (NOT) and of the Niche Divergence Test (NDT) for each one of the combinations between *Psammolestes* species.

Species 1	Species 2	Niche overlap test (NOT)				Niche Divergence Test (NDT)				Interpretation
		Equivalency test		Background test		Equivalency test		Background test		
		D	p value for D	p value for D (2-1)	p value for D (1-2)	D	p value for D	p value for D (2-1)	p value for D (1-2)	
<i>P. arthuri</i>	<i>P. tertius</i>	0.03309	0.00099	0.16666	0.375	0.03309	0.00099	0.14285	0.1666	Strong evidence niches have diverged
<i>P. arthuri</i>	<i>P. coreodes</i>	0.00189	0.00099	0.1578	0.768	0.00189	0.00099	0.456	0.9809	Strong evidence niches have diverged
<i>P. coreodes</i>	<i>P. tertius</i>	0.00351	0.00099	0.2	0.05882	0.00351	0.00099	0.125	0.0625	Strong evidence niches have diverged

FIGURES

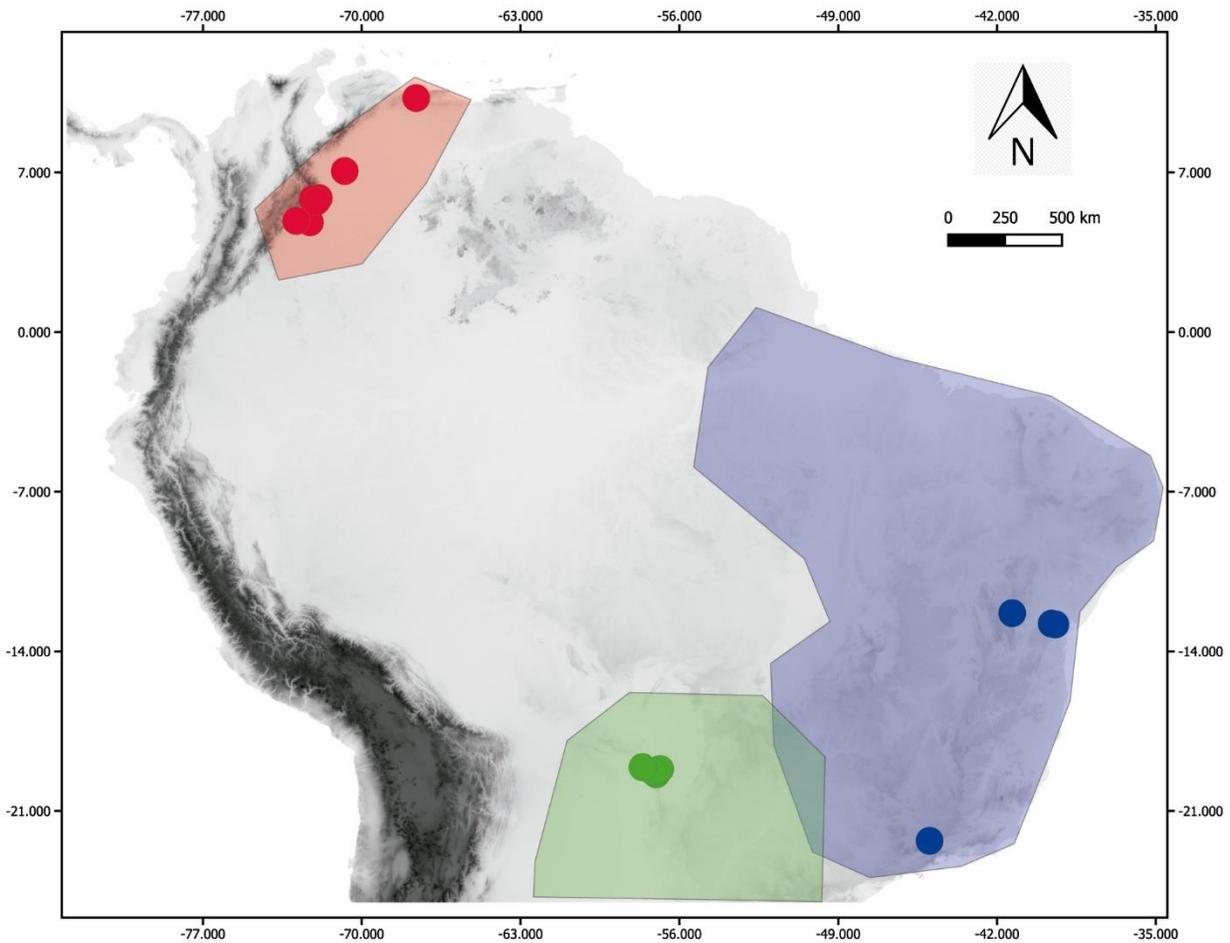


Figure 1. Distribution of the three *Psammolestes* species sampled in this study. Dots represent the sampling sites of this work while polygons symbolize previously reported sampling sites by Ceccarelli *et al.* in 2018, where the *Psammolestes* species have been found. *Psammolestes arthuri* is represented by the red color, *Psammolestes coreodes* with green, and *Psammolestes tertius* with blue.

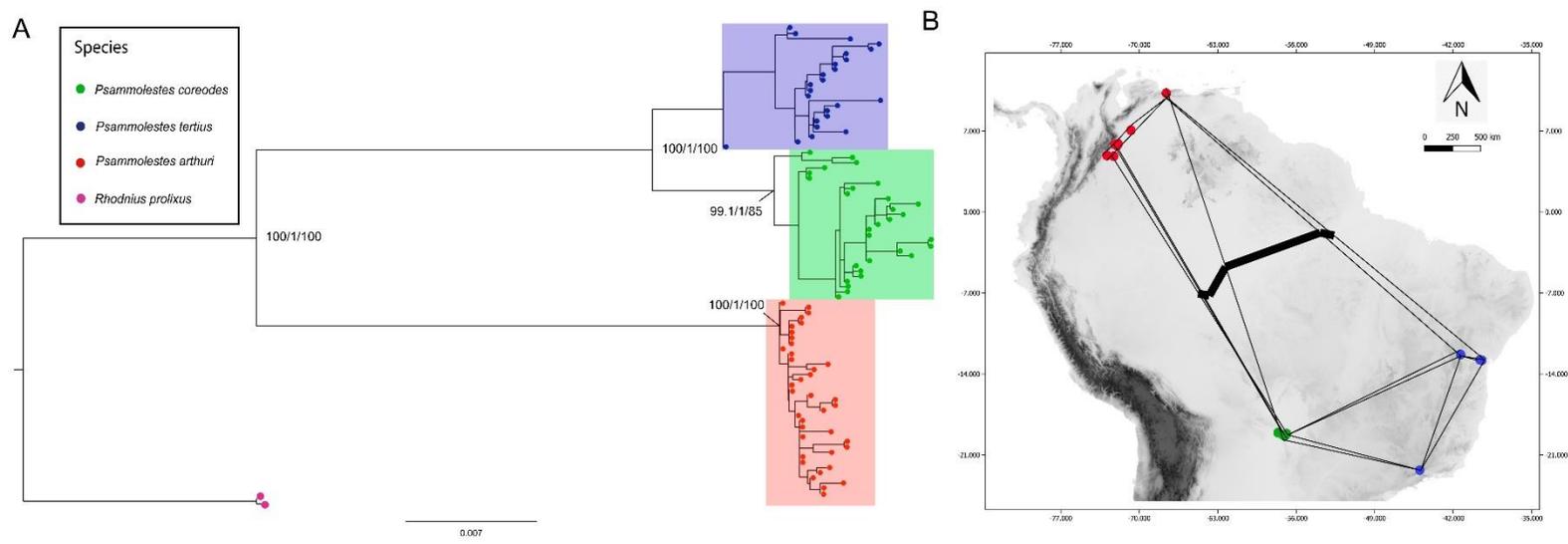
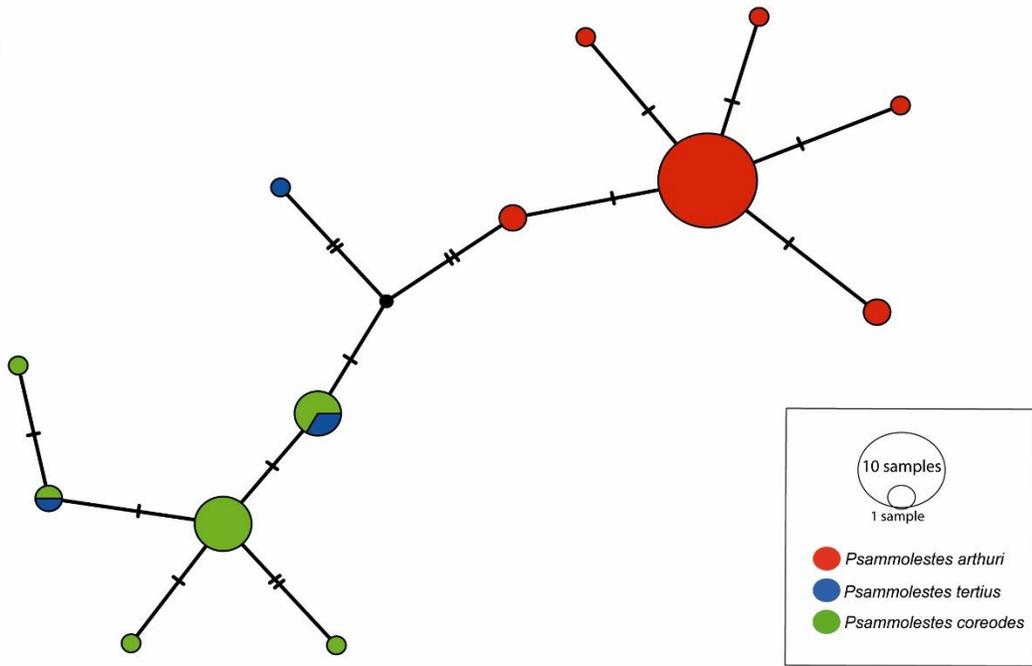
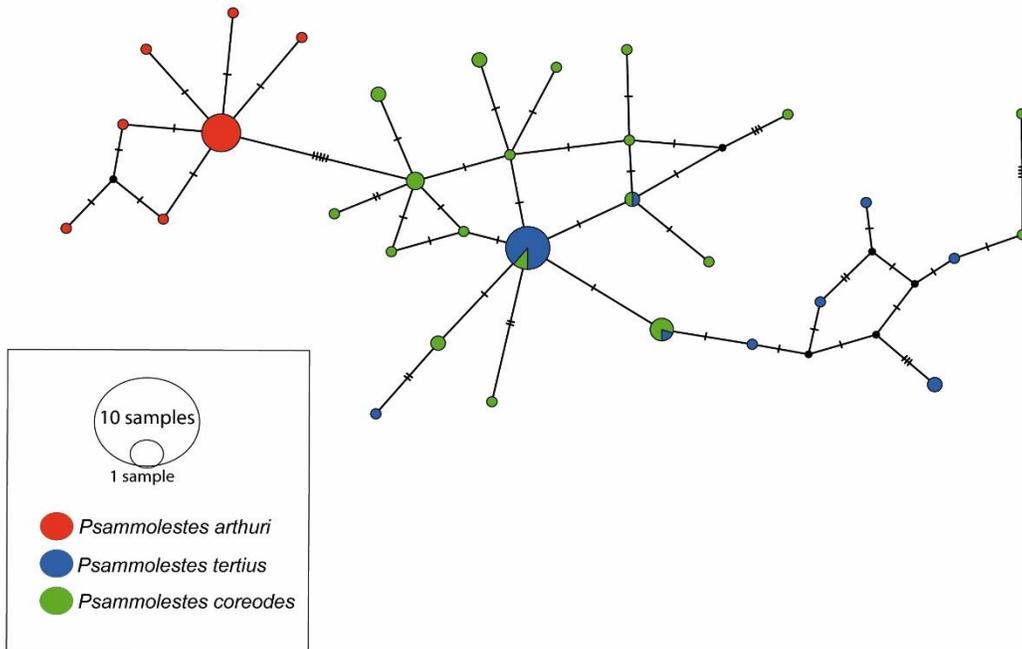


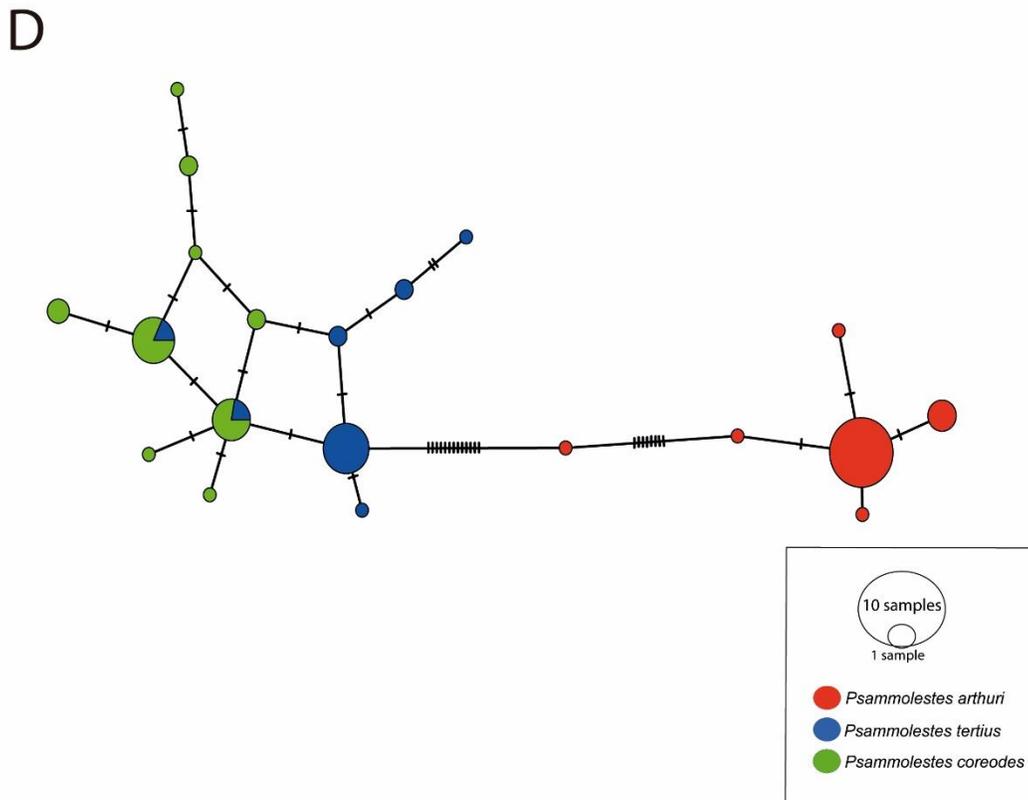
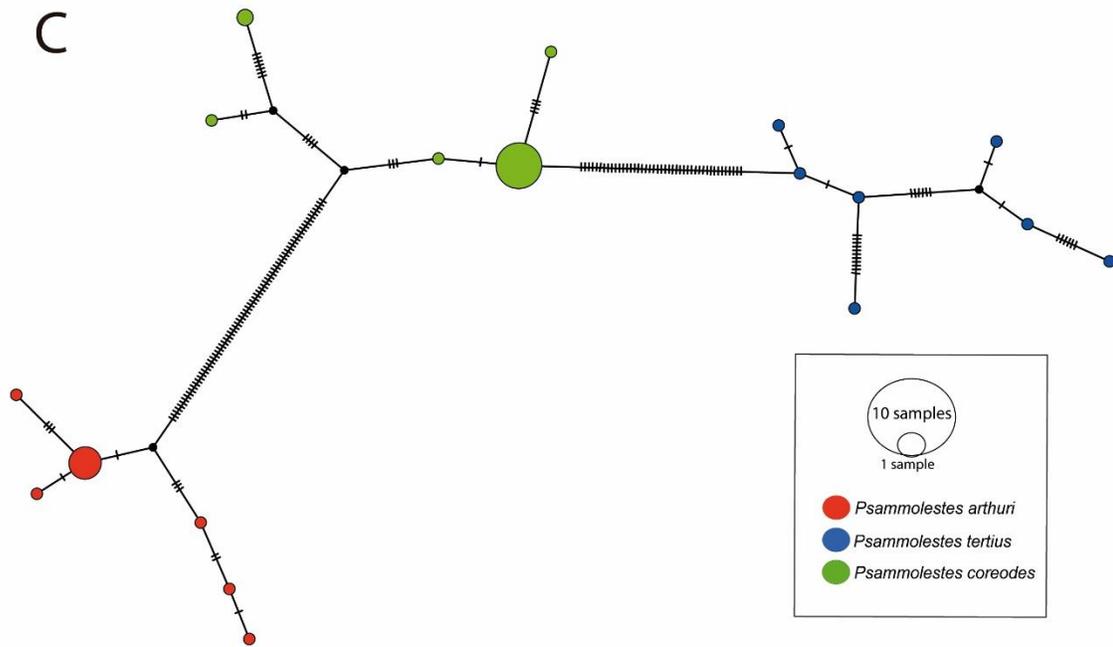
Figure 2. (A) Phylogenetic reconstruction with the ML algorithm based on the seven molecular loci used in this study and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

A

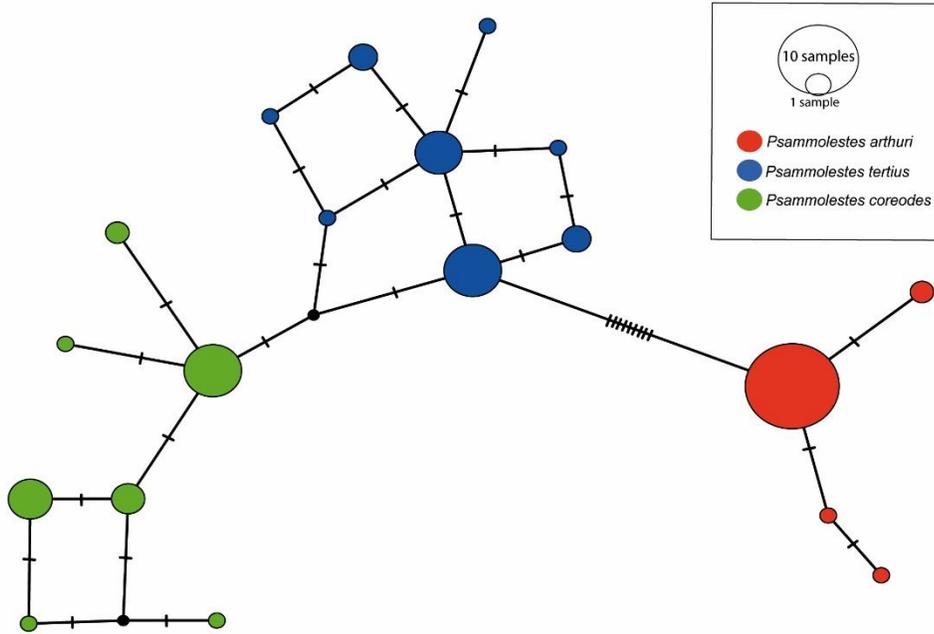


B

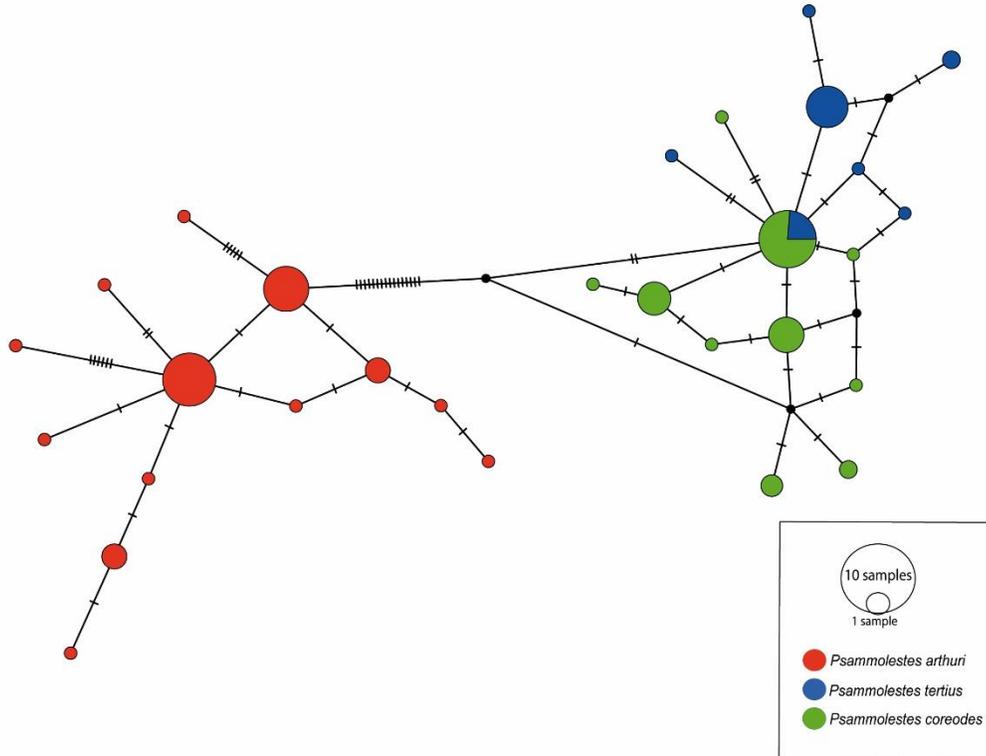




E



F



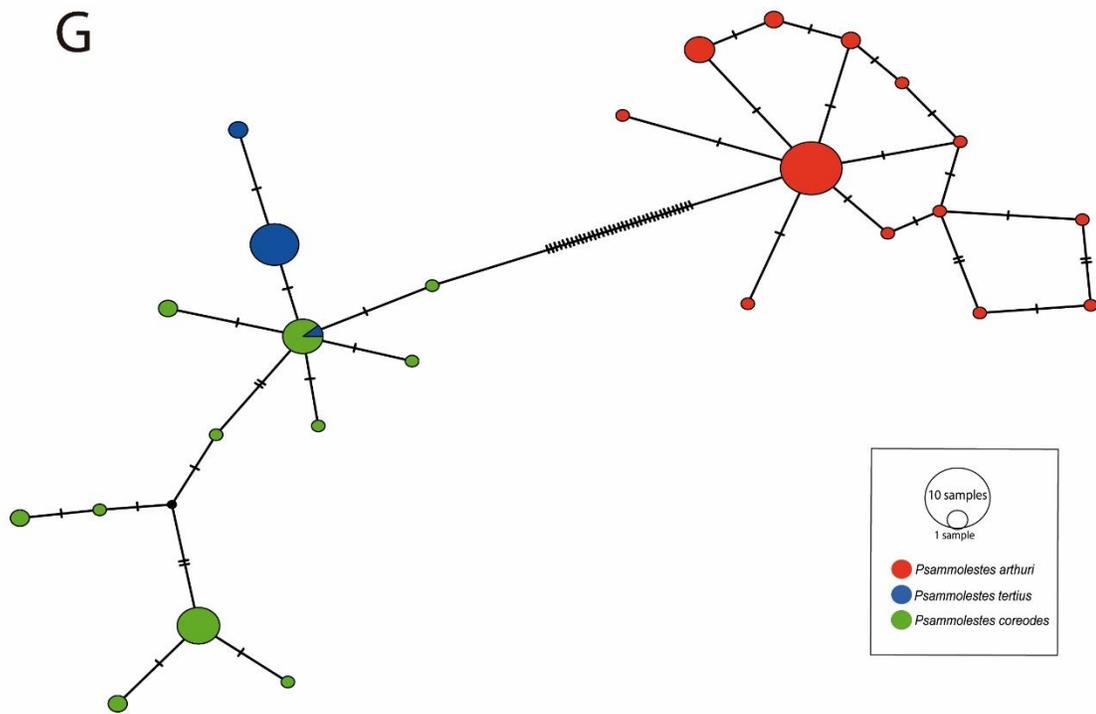


Figure 3. Haplotype networks obtained from the molecular data of the markers (A) 28S, (B) CISP, (C) CYTB, (D) LSM, (E) PJH, (F) TRNA, (G) UPCA.

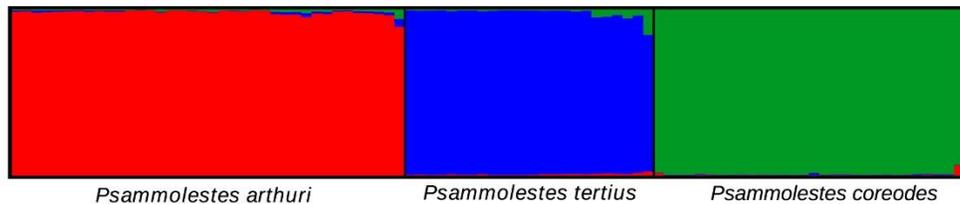


Figure 4. Graphical output from the *distruct* software plotted using $K=3$ and the matrix of aligned Q values from populations and individuals obtained from CLUMPP. Input sequences were organized from left to right in the following order: *Psammolestes arthuri*, *Psammolestes tertius*, and *Psammolestes coreodes*. Each bar represents an individual, and the color of the bar represents the likelihood of that individual of belonging to a population. Red color represents the likelihood of belonging to *Psammolestes arthuri*, blue to *Psammolestes tertius*, and green to *Psammolestes coreodes*.

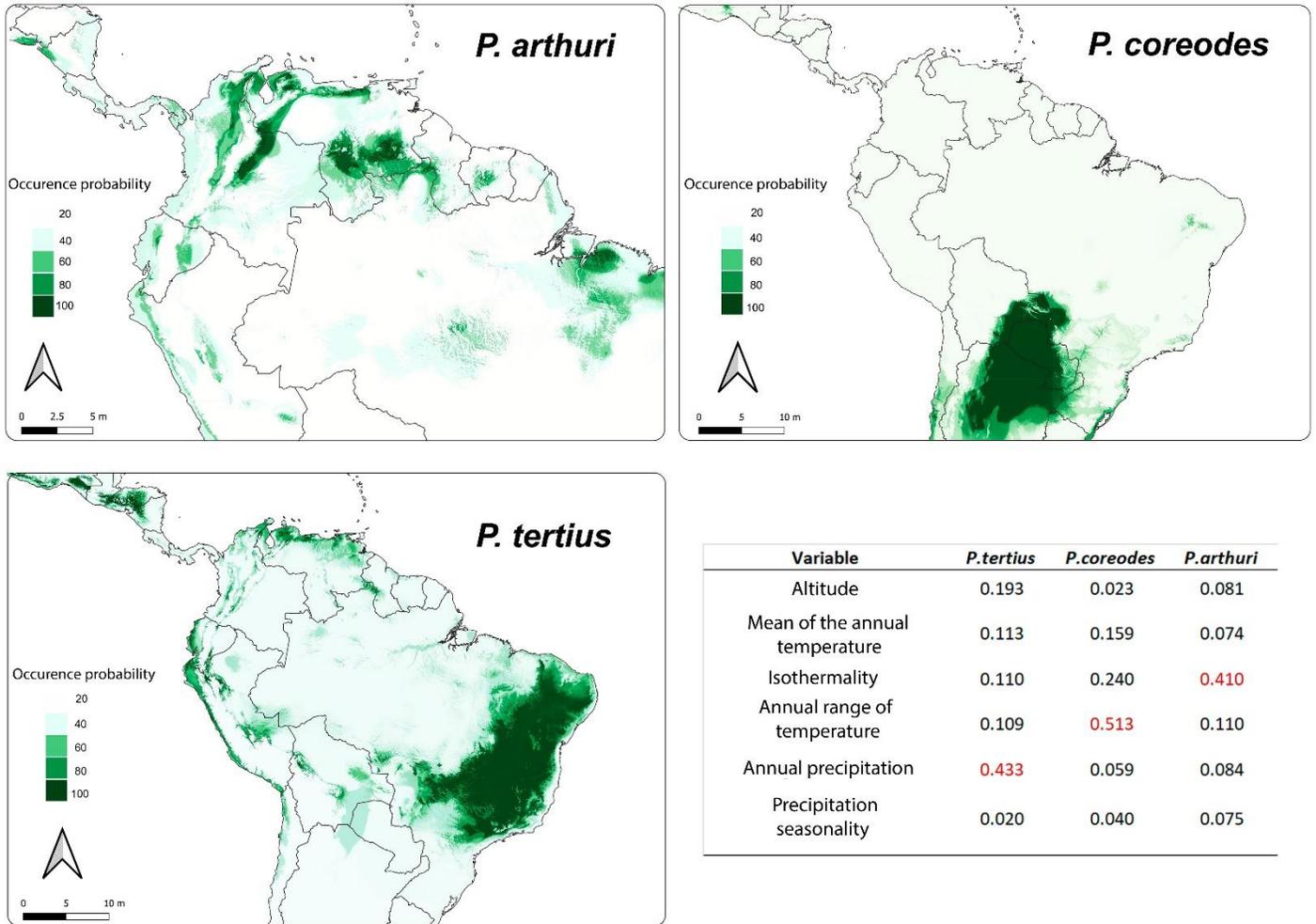


Figure 5. Results from the environmental niche modeling test. Highlighted areas in dark green correspond to the areas where it's more probable to find individuals of each one of the three *Psammolestes* species. Table shows the six variables measured on the models, and highlighted in red the variable that influences the most the distribution of each one of the species.

SUPPLEMENTARY MATERIAL

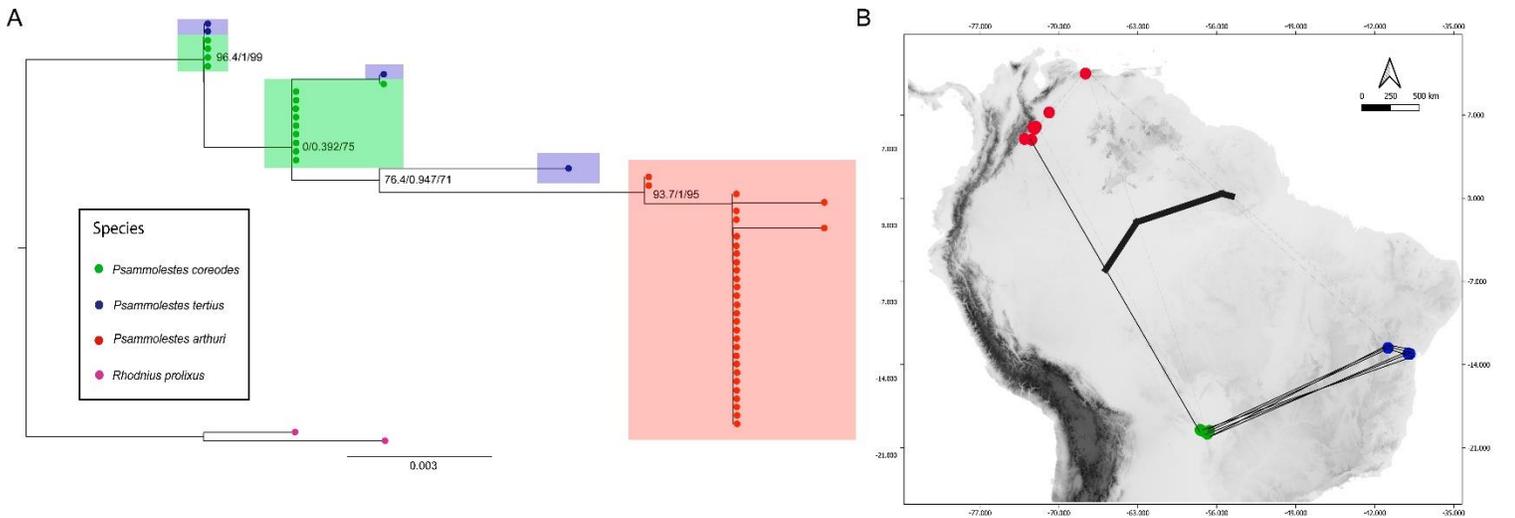


Figure S1. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus 28S and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

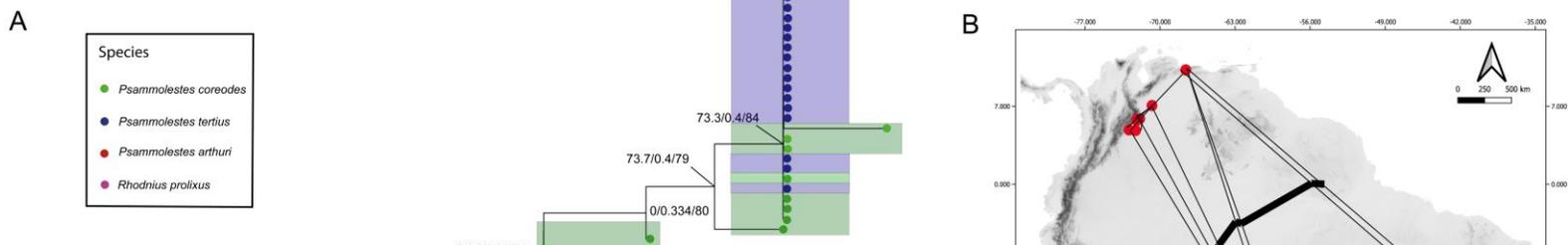


Figure S2. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus CISP and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

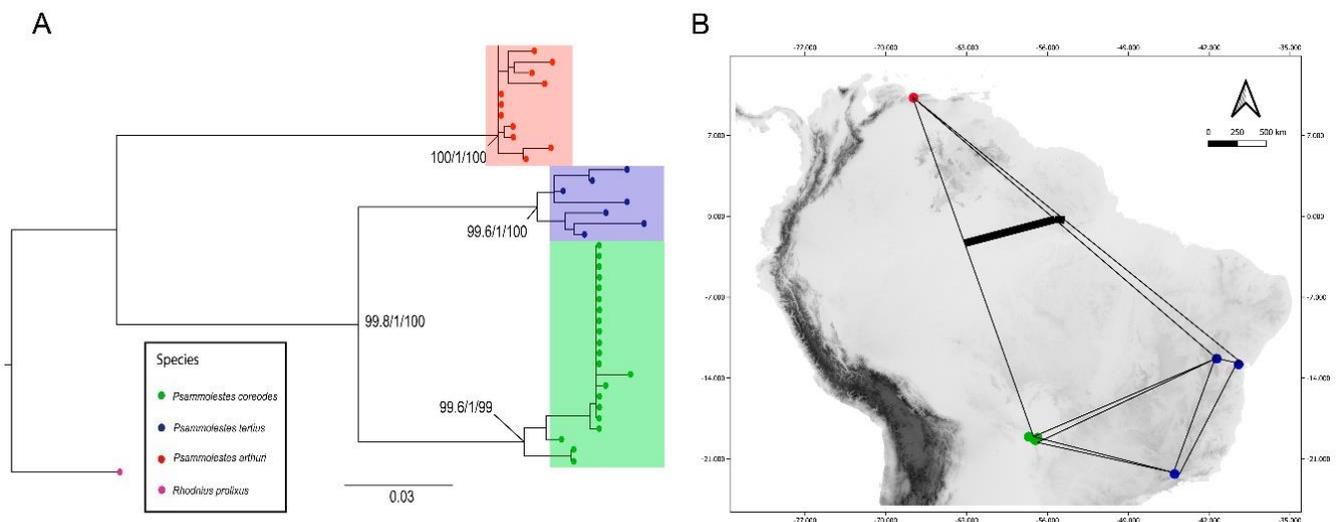


Figure S3. (A) Phylogenetic reconstruction with the ML algorithm based on the mitochondrial locus CYTB and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

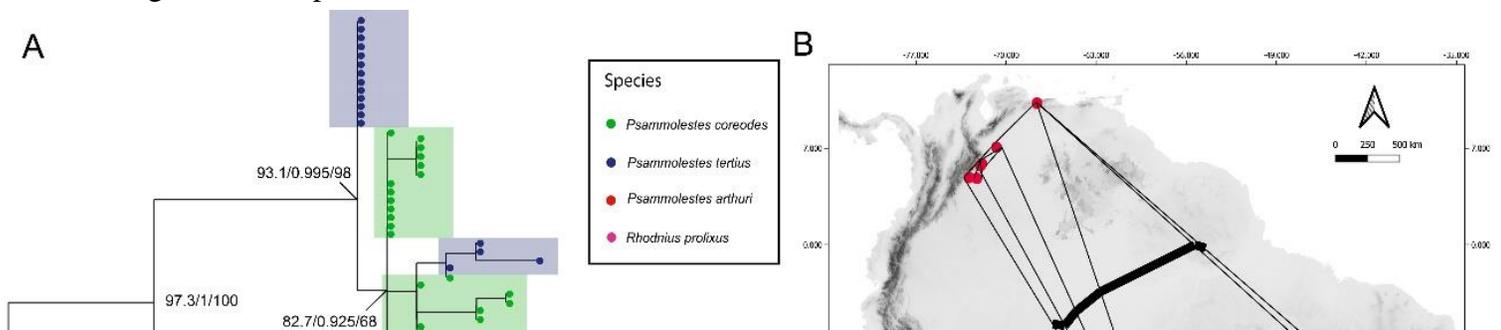


Figure S4. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus LSM and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

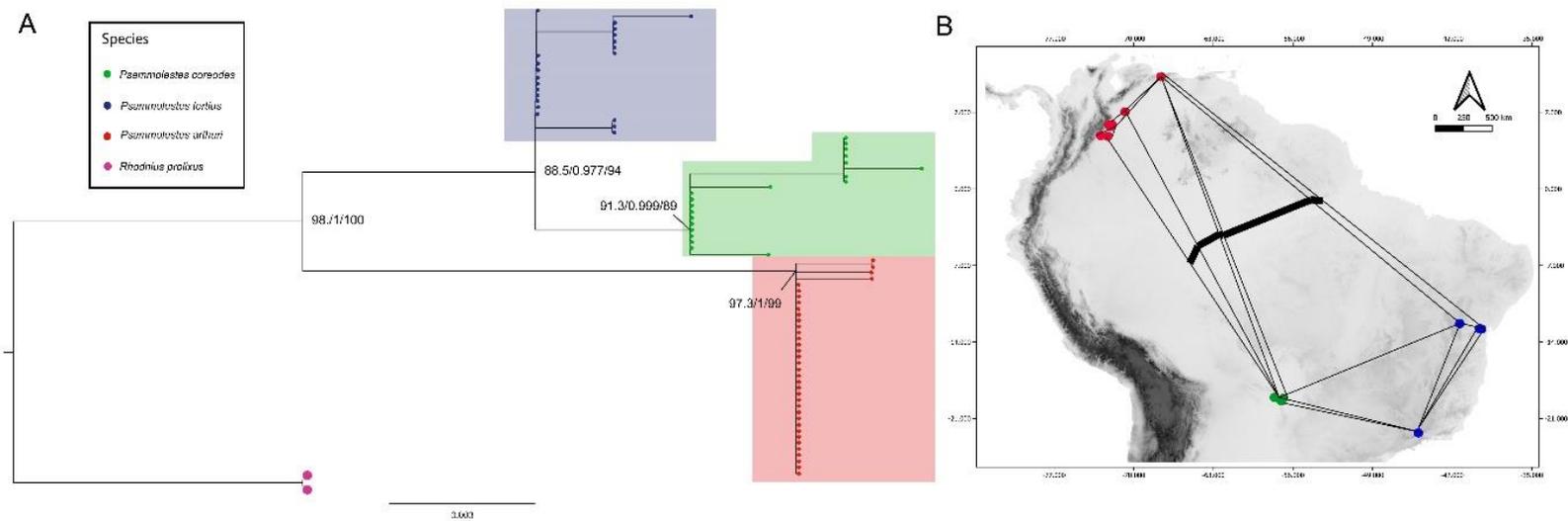


Figure S5. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus LSM and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent

Psammolestes arthuri, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

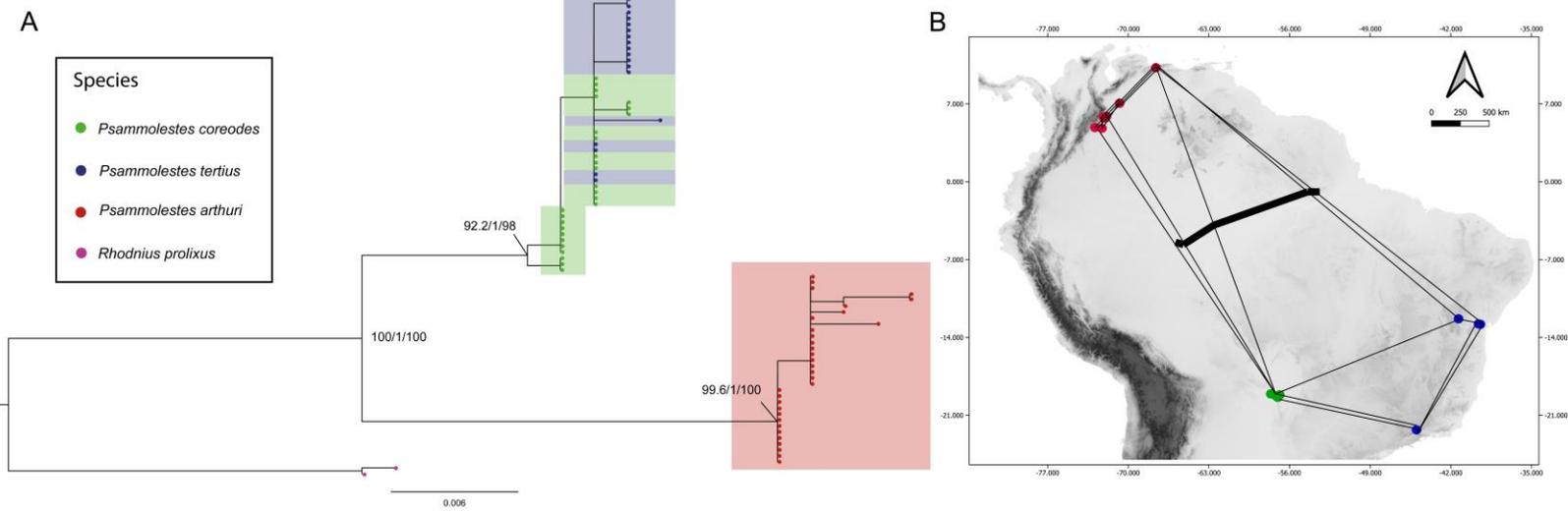


Figure S6. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus TRNA and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

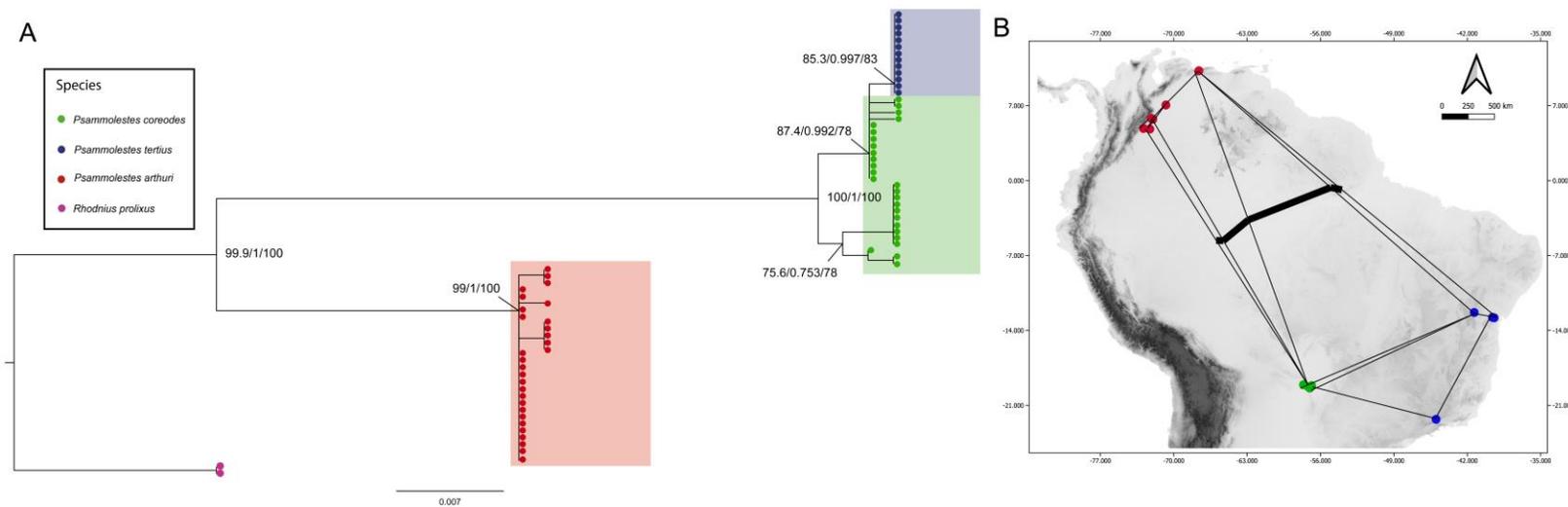


Figure S7. (A) Phylogenetic reconstruction with the ML algorithm based on the nuclear locus TRNA and (B) geographical barrier test (Monmonier's algorithm) with the thick black line representing the main geographical barrier, and thin lines being the Voronoi tessellation and Delaunay triangulation. Bootstrap values higher than 60 are shown. Red dots represent *Psammolestes arthuri*, blue dots represent *Psammolestes tertius*, and green dots represent *Psammolestes coreodes*.

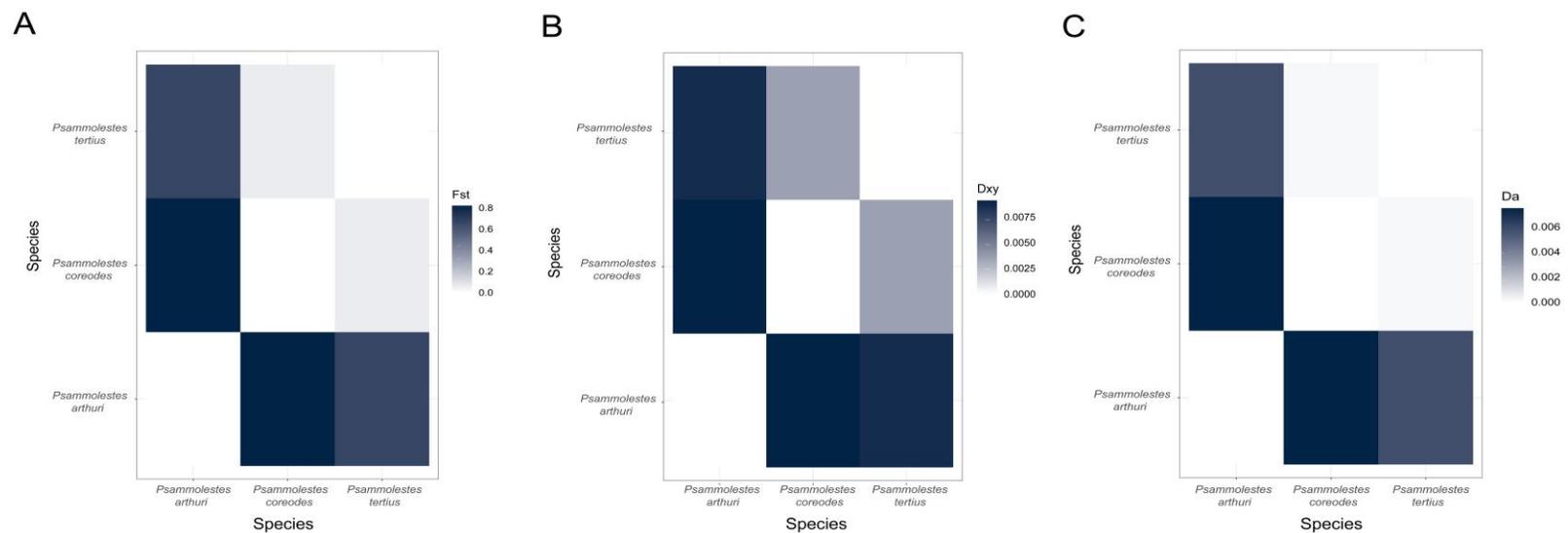


Figure S8. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus 28S.

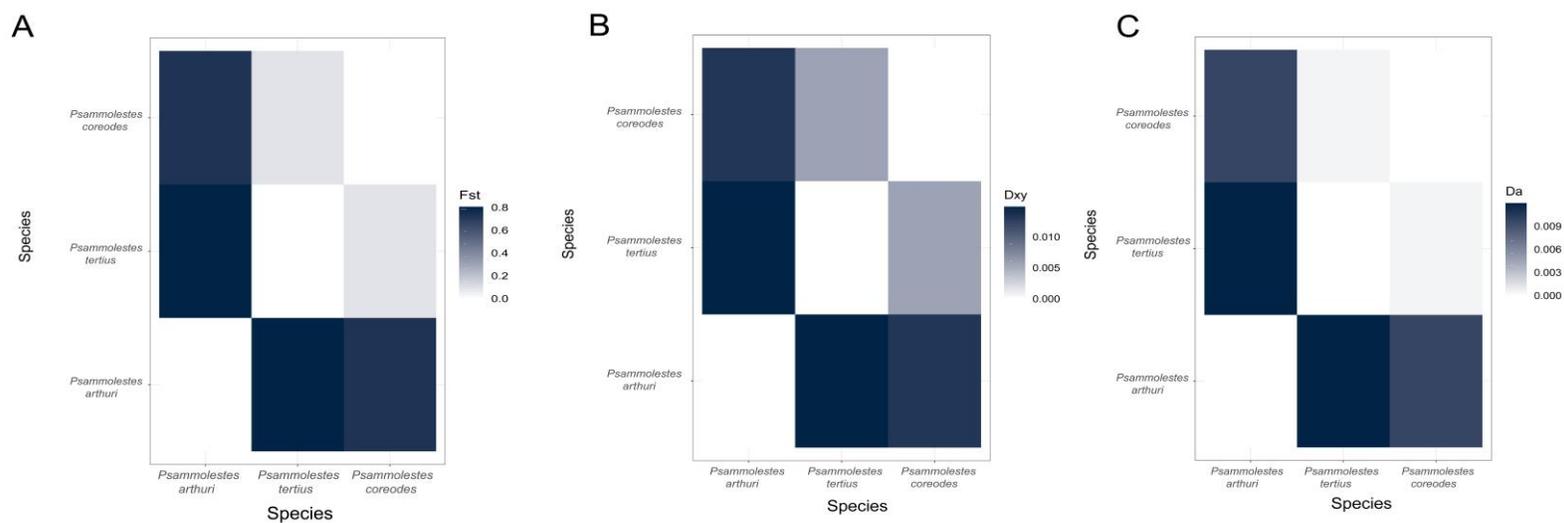


Figure S9. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus CISP.

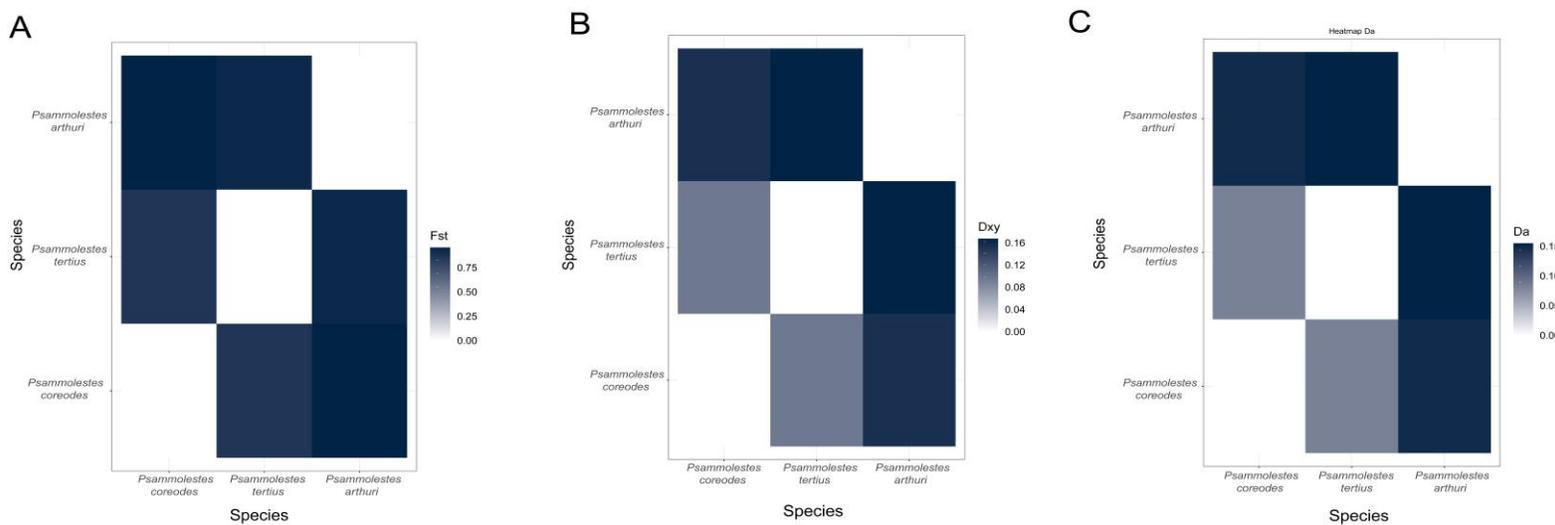


Figure S10. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the mitochondrial locus CYTB.

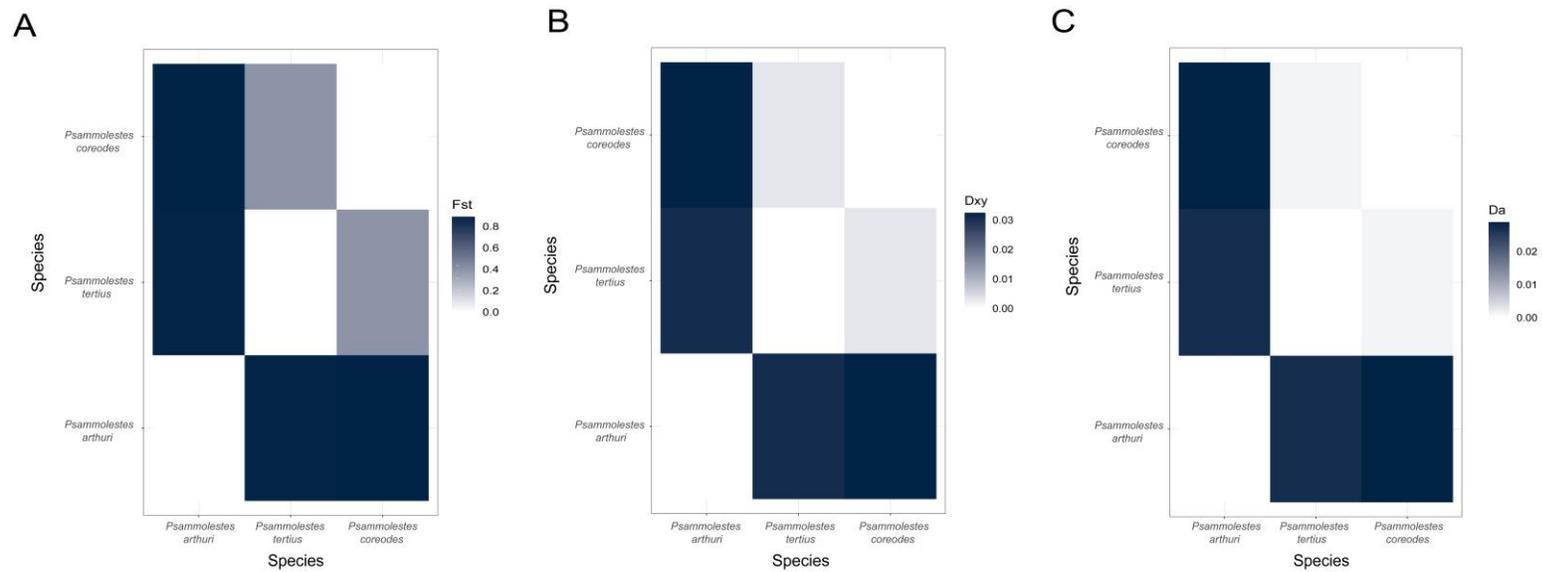


Figure S11. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus LSM.

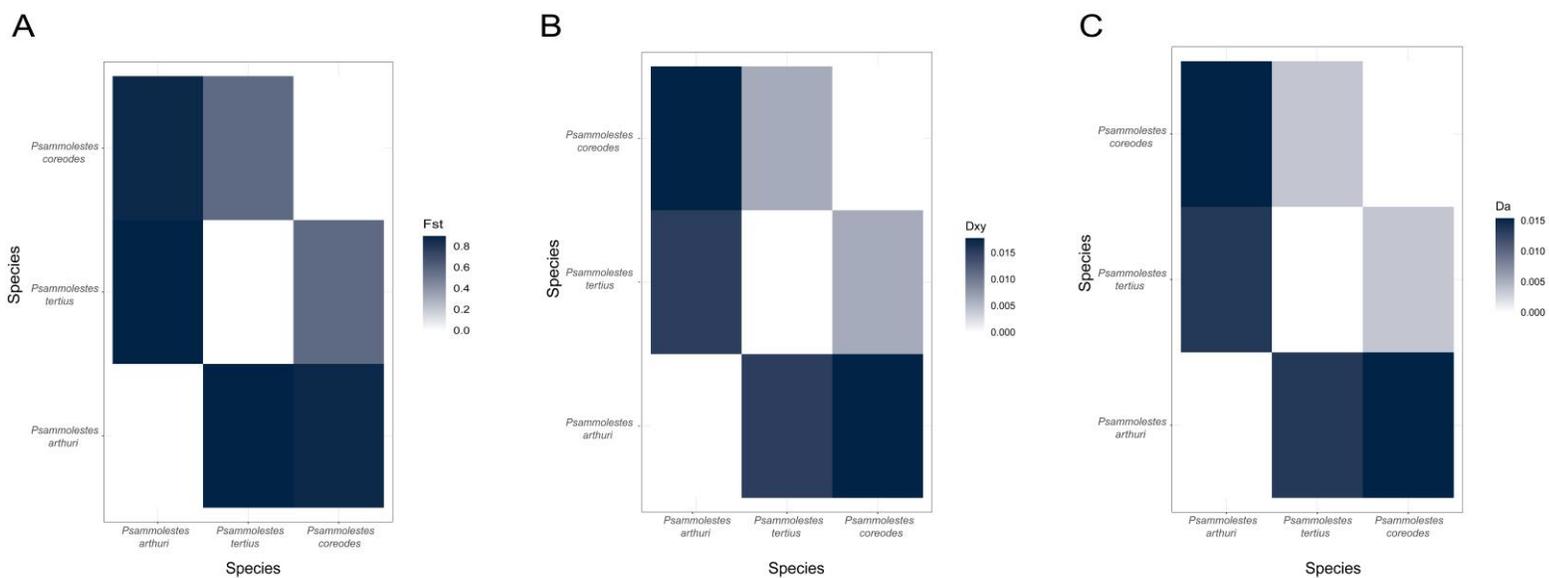


Figure S12. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus PJH.

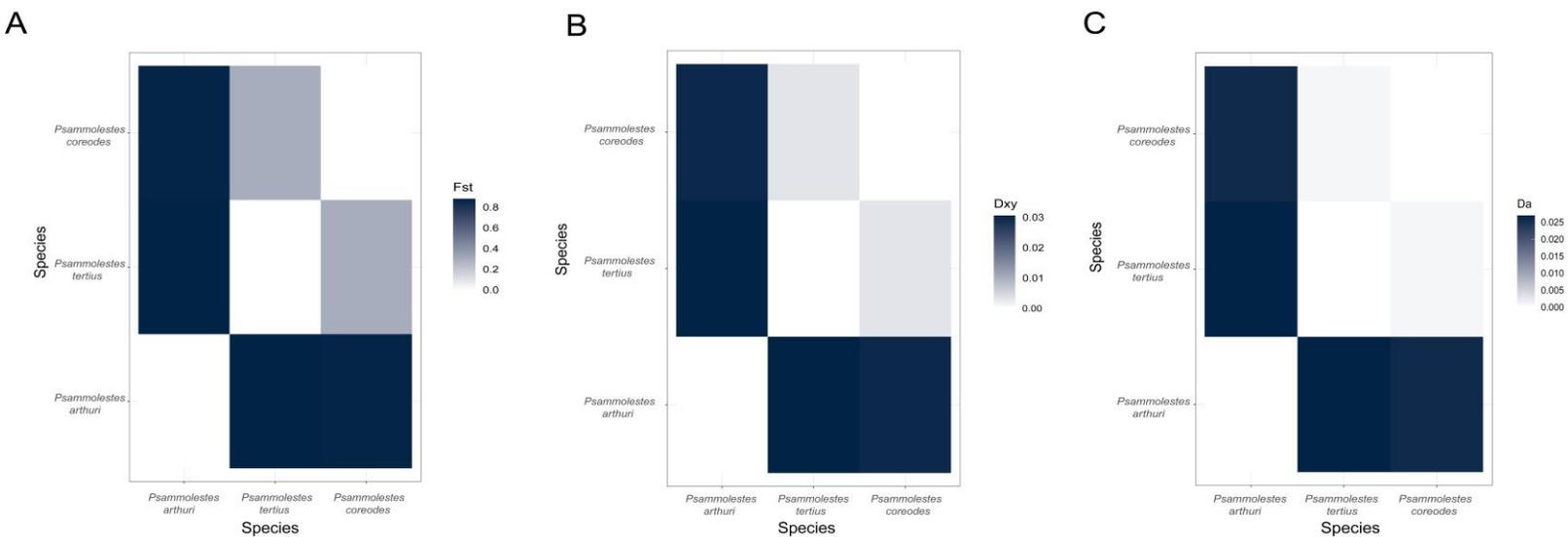


Figure S13. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus TRNA.

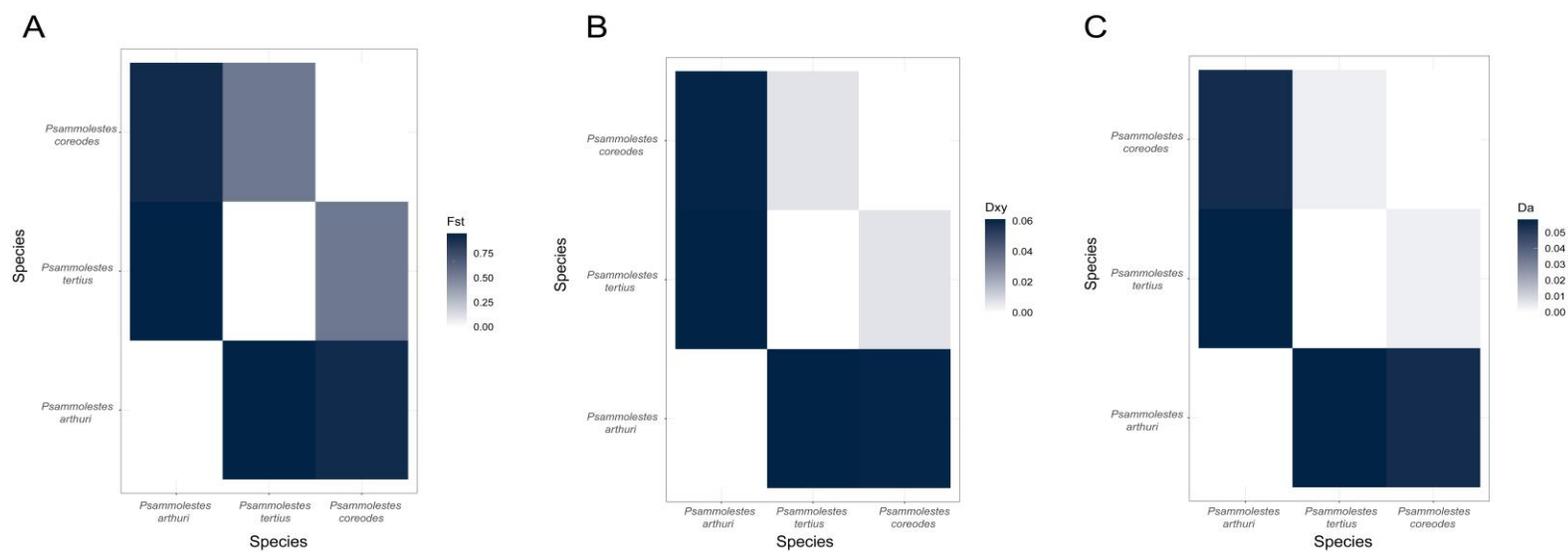


Figure S14. Heat maps calculated for three different statistics: A) F_{ST} , B) D_{XY} and C) D_a for three species based on the molecular data obtained from the nuclear locus UPCA.

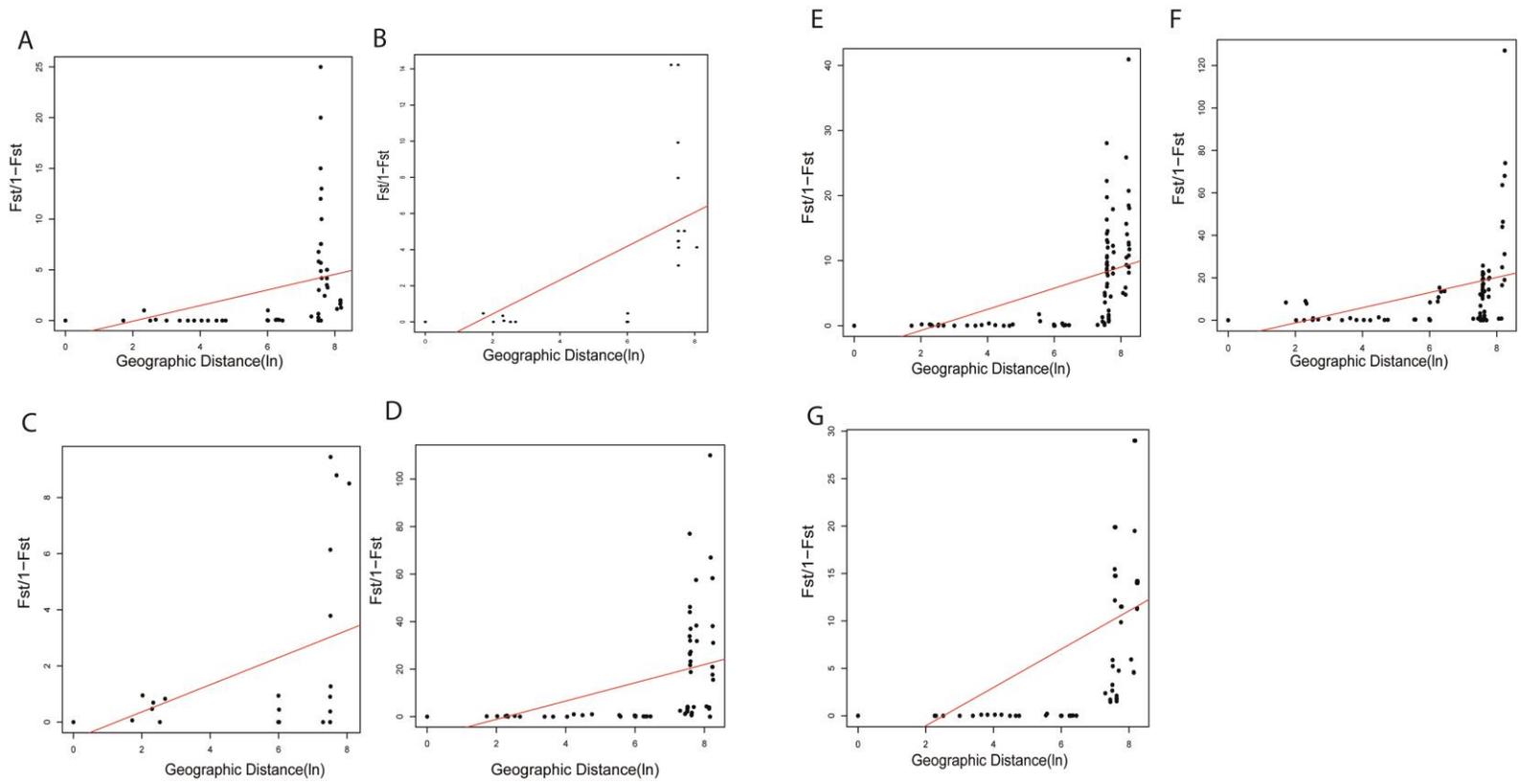


Figure S15. Linear correlations of IBD test for the loci used: (A) 28S, (B) CISP, (C) CYTB, (D) LSM, (E) TRNA, (F) UPCA, and (G) PJH. R^2 and p values are included in table S2.

Table S1. Mantel's test for isolation by distance (IBD) and linear correlation results. The result of the Mantel's test is shown in the two first columns of the table, and the results of the Pearson's correlation test correspond to the third column. The last two columns show the results of the linear correlation tested between geographical and genetic distances.

Locus	Mantel's R	Mantel's p-value	Pearson's R	R²	Corr. p-value
28S	0,3791	0,003	0,3884	0,1437	1,07e-05
CISP	0,5112	0,002	0,5158	0,2623	1,01e-14
CYTB	0,6205	0,041	0,6372	0,3934	8.51e-07
LSM	0,4150	0,001	0,4205	0,1719	1,24e-08
PJH	0,6218	0,001	0,5947	0,3492	3,83e-15
TRNA	0,5281	0,001	0,5234	0,2702	3,50e-15
UPCA	0,3727	0,001	0,3776	0,1382	4,83e-08

Table S2. Genes included in this study, the primers used to obtain their corresponding sequence, and the length of each one of them. “*” symbolizes a new locus used for the delimitation of the *Psammolestes* species.

Gen name	Primer Sequence	Size (bp)
<i>tRNA (Guanine (37) -N (1) methyltransferase *</i>	Forward: GGGCCACGTTTCTAACAAAA Reverse: CAATTGGAATGCTGCTGAAA	842
<i>Putative juvenile hormone inducible protein *</i>	Forward: CCCTTTTAGCAAAATGTTCCA Reverse: TGCCATTATTGCAAGCAGAA	720
<i>Probable cytosolic iron sulfur protein assembly protein Ciao 1*</i>	Forward: TTATCTGCGCAAGCAGTAGC Reverse: TAAGACTTTGGGGGAAGCAA	706
<i>Lipoyl synthase, mitochondrial *</i>	Forward: AAAAAGCCCATTTTCGTTTCC Reverse: AATGGGCCACATTATTCAA	768
<i>Uncharacterized protein (Cell adhesion) *</i>	Forward: TGAAAGGGATCGTACCTTGG Reverse: CCTCCAGACTGATGGCTTGT	795
<i>Cytochrome b</i>	Forward: GGACG(AT)GG(AT)ATTTATTATGG ATC Reverse: GC(AT)CCAATTCA(AG)GTTA(AG)T AA	840
28S	Iniciador Forward: GCGAGTCGTGTTGCTTGATAG TGCAG Iniciador Reverse: TTGGTCCGTGTTTCAAGACGG G	696

Table S3. Taxon sampling information for *Psammolestes* populations.

Sample ID	Species	Place	Latitude	Longitude
224	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
225	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
226	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
227	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
228	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
229	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
230	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
231	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
232	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
233	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
234	<i>Psammolestes arthuri</i>	Colombia, Casanare, Maní	-72.281577	4.817266
499	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
500	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
502	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
504			-39.428831	-12.817339

	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves		
505	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
506	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
507	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
508	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
510	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso do Sul, Corumbá	-57.01804	-19.38609
511	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso do Sul, Corumbá	-57.01804	-19.38609
512	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso do Sul, Corumbá	-57.01804	-19.38609
513	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso do Sul, Corumbá	-57.01804	-19.38609
514	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
515	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
516	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
517	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
613	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339

614	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
615	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
616	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
617	<i>Psammolestes tertius</i>	Brazil, Bahia, Castro Alves	-39.428831	-12.817339
623	<i>Psammolestes tertius</i>	Brazil, Bahia, Santa Teresinha	-39.591431	-12.783287
625	<i>Psammolestes tertius</i>	Brazil, Bahia, Santa Teresinha	-39.591431	-12.783287
630	<i>Psammolestes tertius</i>	Brazil, Bahia, Santa Teresinha	-39.591431	-12.783287
631	<i>Psammolestes tertius</i>	Brazil, Bahia, Santa Teresinha	-39.591431	-12.783287
632	<i>Psammolestes tertius</i>	Brazil, Bahia, Santa Teresinha	-39.591431	-12.783287
633	<i>Psammolestes tertius</i>	Brazil, Bahia, Seabra	-41.32896	-12.32449
634	<i>Psammolestes tertius</i>	Brazil, Bahia, Seabra	-41.32896	-12.32449
635	<i>Psammolestes tertius</i>	Brazil, Bahia, Seabra	-41.32896	-12.32449
636	<i>Psammolestes tertius</i>	Brazil, Bahia, Seabra	-41.32896	-12.32449
637	<i>Psammolestes tertius</i>	Brazil, Bahia, Seabra	-41.32896	-12.32449
647	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609

648	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
649	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
650	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
652	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
654	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
655	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.01804	-19.38609
658	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-56.849279	-19.16075
660	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-56.849279	-19.16075
662	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-56.849279	-19.16075
663	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-56.849279	-19.16075
664	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-56.849279	-19.16075
665	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
666	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467

667	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
669	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
670	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
671	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
672	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
679	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
680	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
681	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
682	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
683	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
688	<i>Psammolestes coreodes</i>	Brazil, Mato Grosso de Sul, Corumbá	-57.164048	-19.268467
689	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
690	<i>Psammolestes arthuri</i>	Venezuela,	-67.6017333	10.2633702

		Aragua, Maracay		
691	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
692	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
693	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
703	<i>Psammolestes tertius</i>	Brazil, Minas Gerais, Itanhandú	-44.973346	-22.309973
705	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
706	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
707	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
709	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
712	<i>Psammolestes arthuri</i>	Venezuela, Aragua, Maracay	-67.6017333	10.2633702
727	<i>Psammolestes arthuri</i>	Colombia, Casanare, Paz de Ariporo	-71.892955	5.881236
728	<i>Psammolestes arthuri</i>	Colombia, Casanare, Paz de Ariporo	-71.892955	5.881236
729	<i>Psammolestes arthuri</i>	Colombia, Casanare, Paz de Ariporo	-71.892955	5.881236

731	<i>Psammolestes arthuri</i>	Colombia, Casanare, Tamará	-72.16405	5.829558
732	<i>Psammolestes arthuri</i>	Colombia, Casanare, Tamará	-72.16405	5.829558
733	<i>Psammolestes arthuri</i>	Colombia, Casanare, Tamará	-72.16405	5.829558
737	<i>Psammolestes arthuri</i>	Colombia, Casanare, Poré	-71.99643	5.727289
738	<i>Psammolestes arthuri</i>	Colombia, Casanare, Poré	-71.99643	5.727289
742	<i>Psammolestes arthuri</i>	Colombia, Casanare, Poré	-71.99643	5.727289
759	<i>Psammolestes arthuri</i>	Colombia, Casanare, Monterrey	-72.892421	4.875227
760	<i>Psammolestes arthuri</i>	Colombia, Casanare, Monterrey	-72.892421	4.875227
761	<i>Psammolestes arthuri</i>	Colombia, Casanare, Monterrey	-72.892421	4.875227
765	<i>Psammolestes arthuri</i>	Colombia, Arauca, Arauca	-70.751657	7.063543