BIOLOGY, SYSTEMATICS, TAXONOMY, AND EVOLUTION OF INSECT VECTORS

EDITED BY: Jader Oliveira, Cleber Galvão, Kaio Cesar Chaboli Alevi, Mauro Marrelli, Andre Barretto Bruno Wilke and Wanzhi Cai

PUBLISHED IN: Frontiers in Ecology and Evolution and Frontiers in Tropical Diseases

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-741-2
DOI 10.3389/978-2-88974-741-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

BIOLOGY, SYSTEMATICS, TAXONOMY, AND EVOLUTION OF INSECT VECTORS

Topic Editors:
Jader Oliveira, University of São Paulo, Brazil
Cleber Galvão, Oswaldo Cruz Institute, Oswaldo Cruz Foundation (Fiocruz), Brazil
Kaio Cesar Chaboli Alevi, Sao Paulo State Universty, Brazil
Mauro Marrelli, University of São Paulo, Brazil
Andre Barretto Bruno Wilke, University of Miami, United States
Wanzhi Cai, China Agricultural University, China

Citation: Oliveira, J., Galvão, C., Alevi, K. C. C., Marrelli, M., Wilke, A. B. B., Cai, W., eds. (2022). Biology, Systematics, Taxonomy, and Evolution of Insect Vectors. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-741-2

Table of Contents

05 ***Editorial: Biology, Systematics, Taxonomy, and Evolution of Insect Vectors***
Kaio Cesar Chaboli Alevi, Cleber Galvão, Mauro Toledo Marrelli, André Barretto Bruno Wilke, Wanzhi Cai and Jader de Oliveira

08 ***Identification of* Aedes *(Diptera: Culicidae) Species and Arboviruses Circulating in Arauca, Eastern Colombia***
David Martínez, Carolina Hernández, Marina Muñoz, Yulieth Armesto, Andres Cuervo and Juan David Ramírez

18 ***Influence of Larval Habitat Environmental Characteristics on Culicidae Immature Abundance and Body Size of Adult* Aedes aegypti**
Mariana Rocha David, Edson Santos Dantas, Rafael Maciel-de-Freitas, Cláudia Torres Codeço, Alex Enrich Prast and Ricardo Lourenço-de-Oliveira

30 ***The Larval Density of Mosquitos (Diptera: Culicidae) in Jiaxiang County, Shandong Province, China: Influence of Bacterial Diversity, Richness, and Physicochemical Factors***
Haiyang Wang, Yang Wang, Peng Cheng, Huaiwei Wang, Haifang Wang, Hongmei Liu, Chongxing Zhang and Maoqing Gong

42 ***WingBank: A Wing Image Database of Mosquitoes***
Flávia Virginio, Victor Domingues, Letícia Cecília Gomes da Silva, Lucas Andrade, Kelly Rosa Braghetto and Lincoln Suesdek

57 ***Historical Biogeography and the Evolution of Hematophagy in Rhodniini (Heteroptera: Reduviidae: Triatominae)***
Alexandre Silva de Paula, Carlos Barreto, Mike Cerqueira Martins Telmo, Lileia Diotaiuti and Cleber Galvão

70 ***Evidence of Elevational Speciation in* Kerteszia cruzii *(Diptera: Culicidae) in the Ribeira Valley, São Paulo, Brazil***
Bruna Demari-Silva, Gabriel Zorello Laporta, Tatiane Marques Porangaba de Oliveira and Maria Anice Mureb Sallum

80 ***Evidence of Local Extinction and Reintroduction of* Aedes aegypti *in Exeter, California***
Erin Taylor Kelly, Lindsey K. Mack, Melina Campos, Crystal Grippin, Tse-Yu Chen, Ana L. Romero-Weaver, Kyle J. Kosinski, Katherine K. Brisco, Travis C. Collier, Eva A. Buckner, Lindsay P. Campbell, Anthony J. Cornel, Gregory C. Lanzaro, Rodrigo Rosario-Cruz, Kirk Smith, Geoffrey M. Attardo and Yoosook Lee

88 ***Egg Cannibalism Varies With Sex, Reproductive Status, and Egg and Nymph Ages in* Arma custos *(Hemiptera: Asopinae)***
Shaolong Wu, Weiai Zeng, Wan Deng, Mi Li, Wei Hu, Hailin Cai, Youzhi Li, Pengfei Xie, Lin Tan and Zhicheng Zhou

97 ***Chemosensory Gene Expression for Two Closely Relative Species* Rhodnius robustus *and* R. prolixus *(Hemiptera, Reduviidade, Triatominae) Vectors of Chagas Disease***

Axelle Marchant, Florence Mougel, Emmanuelle Jacquin-Joly, Carlos E. Almeida, Denis Blanchet, Jean-Michel Bérenger, João A. da Rosa and Myriam Harry

112 ***Phylogenomics for Chagas Disease Vectors of the* Rhodnius *Genus (Hemiptera, Triatominae): What We Learn From Mito-Nuclear Conflicts and Recommendations***

Jonathan Filée, Marie Merle, Héloïse Bastide, Florence Mougel, Jean-Michel Bérenger, Elaine Folly-Ramos, Carlos Eduardo Almeida and Myriam Harry

frontiers in Ecology and Evolution

EDITORIAL
published: 01 March 2022
doi: 10.3389/fevo.2022.858886

Editorial: Biology, Systematics, Taxonomy, and Evolution of Insect Vectors

Kaio Cesar Chaboli Alevi[1,2]*, *Cleber Galvão*[2], *Mauro Toledo Marrelli*[1], *André Barretto Bruno Wilke*[3], *Wanzhi Cai*[4] and *Jader de Oliveira*[1]

[1] *Laboratório de Entomologia em Saúde Pública, Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo (USP), São Paulo, Brazil,* [2] *Laboratório Nacional e Internacional de Referência em Taxonomia de Triatomíneos, Instituto Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil,* [3] *Laboratory for Computational Epidemiology and Public Health, Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, United States,* [4] *MOA Key Lab of Pest Monitoring and Green Management, Department of Entomology, College of Plant Protection, China Agricultural University, Beijing, China*

Keywords: biology, systematics, taxonomy, evolution, insects vectors

Editorial on the Research Topic

Biology, Systematics, Taxonomy, and Evolution of Insect Vectors

Vector insects comprise several hematophagous invertebrate species grouped in the orders Diptera, Hemiptera, Anoplura, and Siphonaptera (Kuno and Chang, 2005). These species are responsible for the transmission of many infectious diseases such as dengue, chikungunya, Zika, yellow fever, malaria, lymphatic filariasis, leishmaniasis, Chagas disease, sleeping sickness, onchocerciasis, bubonic plague, Rift Valley fever, Japanese encephalitis, West Nile fever, tungiasis, typhus, louse-borne relapsing fever, sandfly fever, Crimean-Congo hemorrhagic fever, Lyme disease, relapsing fever, spotted fever, Q fever, and hence are of great importance to public health (World Health Organization, 2022). Vector-borne diseases account for more than 17% of all infectious diseases (caused by parasites, bacteria, or viruses), causing more than 700,000 deaths annually (World Health Organization, 2022). A better understanding of insect vector biology, systematics, taxonomy, and evolution can support more effective management and control strategies to mitigate the mortality and morbidity of these diseases. This is of great importance because, for the most part, vector control is the primary means by which incidences of vector-borne diseases in humans and animals are mitigated.

The studies presented in this Research Topic focused on two important orders of vector insects, namely, Diptera and Hemiptera. The genus *Aedes* Meigen, 1818 (Diptera, Culicidae) was discussed in several works: Martínez et al.; David et al.; Kelly et al.; Wang et al.; Virgínio et al., contributing to the taxonomic (Martínez et al.; Virgínio et al.), genetic/genomic (Martínez et al.; Kelly et al.), biological (David et al.) and entomo-epidemiological knowledge (Martínez et al.; Kelly et al.) of these vectors.

Martínez et al. presented important taxonomic, genetic (demonstrating that barcodes based on the COI gene fragment are successful method for identifying *Aedes* species) and entomo-epidemiological [diagnosing natural infection for dengue virus type 1 (DENV-1) and chikungunya virus (CHIKV)] for *Ae. aegypti* (Linnaeus, 1762) and *Ae. albopictus* (Skuse, 1894) from the eastern region of Colombia.

David et al. evaluated the development of *Ae. aegypti* in response to different variables of the physicochemical of water. They observed a positive relationship between factors such as the abundance of immatures and container volume and concentration of dissolved organic carbon (DOC), average dry weight of females and wing lengths of males and females, and temperature, as

Edited and reviewed by:
Rodney L. Honeycutt, Pepperdine University, United States

***Correspondence:**
Kaio Cesar Chaboli Alevi
kaiochaboli@hotmail.com

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *20 January 2022*
Accepted: *24 January 2022*
Published: *01 March 2022*

Citation:
Alevi KCC, Galvão C, Marrelli MT, Wilke ABB, Cai W and Oliveira J (2022) Editorial: Biology, Systematics, Taxonomy, and Evolution of Insect Vectors. Front. Ecol. Evol. 10:858886. doi: 10.3389/fevo.2022.858886

well as between male mean dry weight and water conductivity. Furthermore, Wang et al. evaluated biotic (bacterial diversity) and abiotic (physicochemical factors) issues that influence the larval density of mosquitoes of the genera *Aedes*, *Culex* Linnaeus, 1758, *Anopheles* Meigen, 1818, and *Mansonia* Blanchard, 1901 in Jiaxiang County, Shandong Province, China and concluded that the density and species can be jointly affected by these evaluated factors.

Kelly et al. performed population genomics studies of *Ae. aegypti* and suggested that populations from Exeter, California (CA), were reintroduced from the southern CA cluster following the elimination of the vector in 2015 by Delta Vector Control District (DVCD) activities. Demari-Silva et al. evaluated the genetic structure of *Anopheles cruzii* Dyar and Knab, 1908 from Ribeira Valley, southeastern Brazil, and verified whether the genetic structure was associated with forest cover, elevation, slope, and vegetation physiognomy. The authors observed two distinct lineages in the studied region associated with elevation and isolation by distance and proposed that differences in observed population structure found might be associated with the distribution of bromeliad species. Virgínio et al. presented the wing image database of mosquitoes (WingBank), developed from the geometric morphometry of the wings of these vectors. This image bank contains data referring to 77 species belonging to 15 genera of Culicidae.

Studies of Chagas disease vectors are related to the Rhodniini tribe (Hemiptera, Triatominae) (Marchant et al.; Filée et al.; de Paula et al.). Marchant et al. performed a transcriptomic study to evaluate the chemosensory gene expression of *Rhodnius robustus* Larrousse, 1927 and *R. prolixus* Stål, 1859. Among the 199 transcripts annotated in this study, 93% were conserved between *R. prolixus* and the sylvatic *R. robustus*; 12 novel takeout transcripts from the *R. prolixus* genome (Mesquita et al., 2015) were annotated; a large set of transcripts were found to be differentially expressed between males and females, with a majority of them overexpressed in males; 10 transcripts were specifically expressed in one sex and absent in another; three chemosensory transcripts were found to be expressed only in the reared *R. prolixus* and only one in sylvatic *R. robustus*; and 8,596 transcripts were differentially expressed, with most overexpressed in the sylvatic *R. robustus* samples.

de Paula et al. and Filée et al. performed phylogenetic [using mitochondrial rDNA (16S), nuclear ribosomal RNA (28S) and wingless (Wg) sequences] and phylogenomic studies [using 15 mitochondrial genes (13.3 kb), partial nuclear rDNA (5.2 kb) and 51 nuclear protein-coding genes (36.3 kb)], respectively, in the Rhodniini tribe. de Paula et al. presented dating related to the origin of the tribe Rhodniini (17.91 Mya ago), of the genus *Rhodnius* (9.13 Mya ago) and observed a close relationship between Rhodniini and Salyavatinae, highlighting that these vectors diverged from this subfamily 30.43 Mya. Furthermore, the authors highlighted that the colonization of bromeliads, palms trees, and bird nests represent important events for the speciation of *Rhodnius* spp. and discussed the evolution of hematophagy in a scenario where Rhodniini's ancestor could be pre-adapted for the invasion of these ecotopes that presented favorable factors such as water availability, thermal damping, and source of food (vertebrate inquilines) for the species of Rhodniini. Filée et al. discussed taxonomic and systematic issues of the Rhoniini tribe. They confirmed the specific status of *R. montenegrensis* Rosa et al., 2012 and *R. marabaensis* Souza et al., 2016 and the synonymy of *R. taquarussuensis* Rosa et al., 2017 with *R. neglectus* Lent, 1954, questioned the specific status of *R. milesi* Carcavallo et al., 2001 (that is more likely *R. nasutus* Stål, 1859) and suggested the generic reclassification of *Psammolestes tertius* Lent and Jurberg, 1965 to *R. tertius*. The authors observed a greater relationship between *pictipes* and *pallescens* groups than with the *prolixus* group and discussed which introgression events occurred in the three vector groups.

Finally, Wu et al. carried out studies on egg cannibalism of *Arma custos* (Fabricius, 1794) (Hemiptera: Asopinae) and observed that females exhibit a higher tendency for egg cannibalism than males and that cannibalism varies not only with the developmental stage of the eggs and nymphs but also with sex and reproductive status of females.

Thus, the Research Topic Biology, Systematics, Taxonomy, and Evolution of Insects Vectors addressed an important gap in the knowledge associated with biological, systematic, taxonomic, and evolutionary processes of insect vector species and helps further our understanding of issues related to the dynamics of vector-borne disease transmission, thus providing direction for activities related to vector control programs.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

REFERENCES

Kuno, G., and Chang, G. –J. J. (2005). Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. *Clin. Microbiol. Rev.* 18, 608–637. doi: 10.1128/CMR.18.4.608-637.2005

Mesquita, R. D., Vionette-Amaral, R. J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F. A., Minx, P., et al. (2015). Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14936–14941. doi: 10.1073/pnas.1506226112

World Health Organization (2022). *Vector-Borne Diseases*. Available online at: https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases (accessed on January 19, 2022).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 26 November 2020
doi: 10.3389/fevo.2020.602190

Identification of *Aedes* (Diptera: Culicidae) Species and Arboviruses Circulating in Arauca, Eastern Colombia

*David Martínez[1], Carolina Hernández[1], Marina Muñoz[1], Yulieth Armesto[2], Andres Cuervo[2] and Juan David Ramírez[1]**

[1] Grupo de Investigaciones Microbiológicas-UR (GIMUR), Departamento de Biología, Facultad de Ciencias Naturales, Universidad del Rosario, Bogotá, Colombia, [2] Unidad Administrativa Especial De Salud De Arauca, Arauca, Colombia

The identification of vector species and their natural infection with arboviruses results in important data for the control of their transmission. However, for the eastern region of Colombia, this information is limited. Therefore, this study morphologically and molecularly identified species of the genus *Aedes* and the detection of arboviruses (Dengue, Chikungunya, Zika, and Mayaro) in female mosquitoes (individually) present in three municipalities (Saravena, Arauquita, and Tame) by amplifying the genetic material using RT-PCR (reverse transcriptase polymerase chain reaction) in the department of Arauca, eastern Colombia. Inconsistencies between morphological and molecular identification were detected in 13 individuals with *Aedes albopictus* initially determined as *Aedes aegypti* based on morphology (*n* = 13). Molecular identification showed the simultaneous presence of *A. aegypti* (n = 111) and *A. albopictus* (*n* = 58) in the urban municipalities of Saravena and Arauquita. These individuals were naturally infected with Dengue virus type 1 (DENV-1) and Chikungunya virus (CHIKV). The most frequent arbovirus was DENV-1 with an infection rate of 40.7% (11/27) for *A. aegypti* and 39.7% (23/58) for *A. albopictus*, which was followed by CHIKV with an infection rate of 1.8% for *A. aegypti* (2/111) and 6.9% for *A. albopictus* (4/58). Additionally, a mixed infection of DENV-1 and CHIKV was obtained in 4.5% of *A. aegypti* (5/111). Zika virus (ZIKV) and Mayaro virus (MAYV) infections were not detected. This study found that barcoding (fragment gene COI) is a successful method for identifying *Aedes* species. Additionally, we recommend the individual processing of insects as a more accurate strategy for arboviruses detection since the infection rate is obtained and co-infection between DENV-1 and CHIKV is also possible.

Keywords: arbovirus, dengue, Zika, Chikungunya, Mayaro, *Aedes*

OPEN ACCESS

Edited by:
Jader Oliveira,
Sao Paulo State University, Brazil

Reviewed by:
Jinbao Gu,
Southern Medical University, China
Antônio Ralph Medeiros De Sousa,
University of São Paulo, Brazil

***Correspondence:**
Juan David Ramírez
juand.ramirez@urosario.edu.co

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *02 September 2020*
Accepted: *04 November 2020*
Published: *26 November 2020*

Citation:
Martínez D, Hernández C, Muñoz M, Armesto Y, Cuervo A and Ramírez JD (2020) Identification of Aedes (Diptera: Culicidae) Species and Arboviruses Circulating in Arauca, Eastern Colombia. Front. Ecol. Evol. 8:602190. doi: 10.3389/fevo.2020.602190

INTRODUCTION

Arboviruses (arthropod-borne diseases) have a major impact on public health in Latin America (Shope and Meegan, 1997; Espinal et al., 2019). DENV, ZIKV, and CHIKV have similar epidemiological cycles and result in epidemics approximately every 3 years (Villar et al., 2015). DENV has the highest incidence due to the interaction of different factors such as socioeconomic

conditions, arbovirus control plans and the circulation of four serotypes in Latin America where Dengue virus type 1 (DENV-1) has the highest prevalence (Azevedo et al., 2009; de la Cruz et al., 2019; Liu et al., 2020).

Individuals of the *Aedes* genus are found in areas with high arbovirus rates, which is why it is associated with an infection cycle in urban areas (Beaty et al., 2016). Within the *Aedes* genus, there are approximately 190 species; however, *Aedes aegypti* and *Aedes albopictus* are important for their vector capacity of CHIKV, DENV, ZIKV, and MAYV (Freitas, 2010; Ramasamy et al., 2011; Vega-Rúa et al., 2014; Lounibos and Kramer, 2016). The morphological identification of these two species is complex where individuals are severely damaged, or their external morphological characteristics are altered such as the pattern in the scutum (mesothorax sclerite) (Patsoula et al., 2006; Sumruayphol et al., 2016). Due to these difficulties, different authors mention that the identification of insects based on deoxyribonucleic acid (DNA) can become a more efficient technique when classifying individuals (Rolo, 2020). For this reason, the concept of the DNA barcoding emerges as a new tool for species discrimination. This tool uses a small fragment of DNA which functions as a specific barcode for each species. In insects, this fragment corresponds to a sequence of ∼ 650 base pairs (bp) located at the 5′ end of the cytochrome c oxidase subunit I gene (COI) (Joyce et al., 2018). Therefore, it is important to conduct molecular identification of species using barcoding to accurately identify individuals. The precise identification of *A. aegypti* and *A. albopictus* is relevant because this information defines characteristics for these species such as their biology and vector capacity.

The positive detection of arbovirus and the frequency of infection in vectors is reported as the minimum infection rate (MIR), which is defined as the relationship between positive pools and the total number per 1,000 individuals (Gu et al., 2004). However, there are no studies with methodologies that detect natural infection for arboviruses in single insects. In Latin America, *A. aegypti* has higher MIR values for CHIKV, DENV, ZIKV, and MAYV compared to *A. albopictus*. Furthermore, no CHIKV and MAYV infection has been detected in the latter. In *A. aegypti*, DENV presents a higher range of MIR compared to other viruses and ranges from 0.2 to 26 infected individuals per 1,000 individuals. Alternatively, ZIKV has a MIR value ranging from 3 to 17 individuals and CHIKV ranges from 2 to 3 individuals. Finally, for MAYV, the only reported value is 1.75 (Diallo et al., 2014; Guerbois et al., 2016; Huerta et al., 2017; Cevallos et al., 2018; Eiras et al., 2018; Garcia-Rejon et al., 2018). These data show that *A. aegypti* presents a higher rate of natural infection by arboviruses compared to *A. albopictus* and this is attributed to the greater presence of this species in urban areas, which indicates that *A. aegypti* is the main vector of arboviruses and *A. albopictus* is the secondary vector in Latin America (Black et al., 2002; Paupy et al., 2009).

Focusing on both the frequency of the viral infection and the incriminated vector species in endemic areas is important to propose reliable prevention and control measures. Also, there is a pivotal need to understand infection rates not only on humans but also on the arbovirus vectors to reveal disease transmission dynamics. The eastern departments of Colombia present the highest incidence of DENV, CHIKV, and ZIKV. Additionally, due to its close proximity to Venezuela, attempting to identify MAYV is worthwhile (Lorenz et al., 2019). Therefore, this study morphologically and molecularly identifies *Aedes* species circulating in three municipalities of the department of Arauca (Colombia) and obtained sequences that were used to assess the precision of the molecular test used to identify *A. aegypti* and *A. albopictus*. In addition, the infection by DENV-1, CHIKV, ZIKV, and MAYV was attempted.

MATERIALS AND METHODS

Location and Capture of Insects

The insects were captured in the department of Arauca in the municipalities of Saravena (6° 57′17″ N, 71° 52′36″ W), Arauquita (7° 01′34″ N, 71° 25′38″ W), and Tame (6° 27′30″ N, 71° 44′41″ W) (**Figure 1**). These sites were chosen due to the high incidence of arboviral cases. The collection was conducted in the urban area of three municipalities; however, in the municipality of Saravena, individuals were also captured in rural areas, because in this municipality there is not a defined urban area, thus in the rural area there are different farms that have potential breeding sites. Two samplings were conducted, one in 2018 and the other in 2019, and each BG-Mosquitaire was installed for three consecutive months during the rainy season (April–May–June). Six traps were installed in the municipality of Saravena, 5 in Tame and 4 in Arauquita. In each trap, only adult females of the genus *Aedes* were chosen, and subsequently, the pools were formed using only individuals belonging to the same species and according to their morphological identification. The insects or pools were stored in Invitrogen® RNAlater (Salehi and Najafi, 2014) in jars marked with the coordinates and information regarding the house and/or the collection site and transported to the microbiology laboratory of the Universidad del Rosario. Once in the laboratory, they were individually separated in Eppendorf tubes for molecular identification of species, and in total, 169 specimens were analyzed.

Morphological and Molecular Identification of *Aedes* Species

The morphological identification of *Aedes* genus individuals was performed using a stereomicroscope following the taxonomic key proposed by Forattini (1965) to identify insects to the genus level, and the taxonomic review by Cova-Garcia et al. (1966) was used to identify species. Molecular identification of individual *Aedes* species was conducted by amplifying a fragment of the COI gene (barcoding) from the RNA of the insect. For technical reasons, only the RNA of the insect could be obtained.

RNA extraction was conducted with the Quick-RNATissue/Insect Microprep Kit (Zymo, # R2030) following the manufacturer's recommended instructions, and the RNA concentration was quantified in a NanoDrop (Thermo

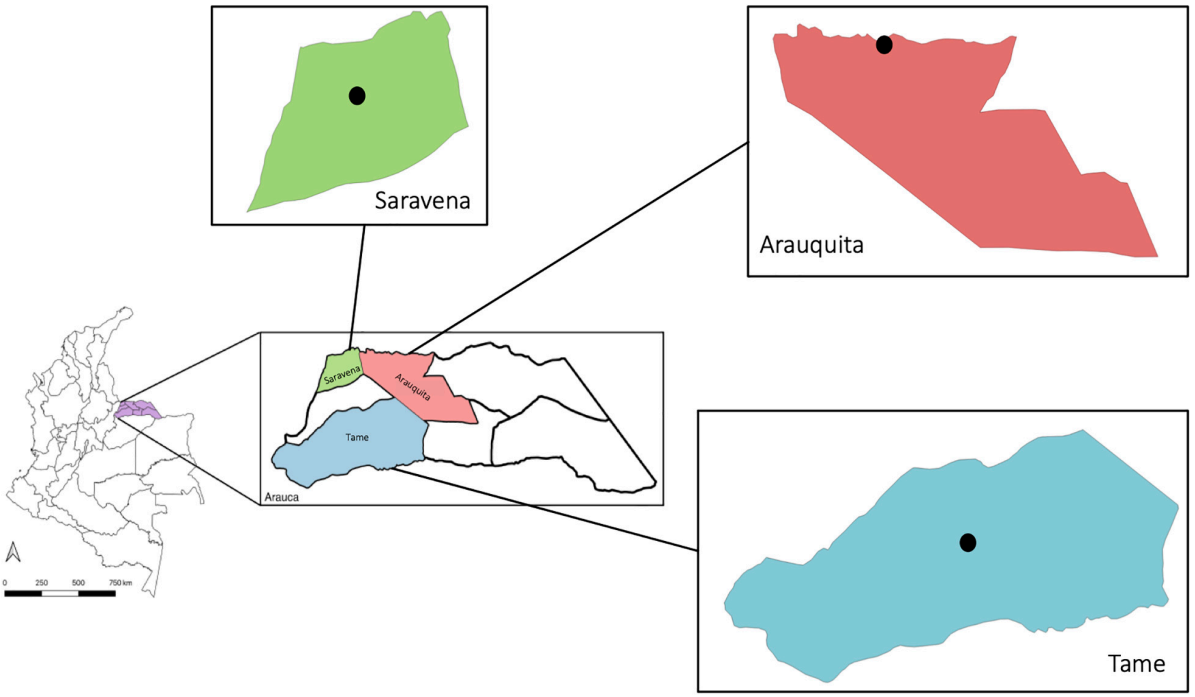

FIGURE 1 | Municipalities where captures were conducted: Saravena (green), Arauquita (pink), and Tame (blue) in the department of Arauca (purple), Colombia. The black dot indicates the urban area where the capture of individuals was made.

Scientific™). Subsequently, One-Step RT-PCR was conducted from the RNA to generate cDNA (complementary DNA) and amplify the COI gene. PCR reactions were performed in a final volume of 20 μL containing 10 μL qScript™ One-Step, Low ROX™ (Quantabio, Carlsbad, CA, United States), 10 μM of each primer and 5 μL of RNA. The primers LCO1490 (5′-GGT CAA ATC ATA AAG ATA TTG G-3) and HCO2198 (5′-TAA ACT TCA GGG TGA CCA AAA AAT CA-3′) were also used (Joyce et al., 2018). The thermal profile consisted of the following: one cycle at 50°C for reverse transcriptase activation for 10 min followed by initial denaturation at 95°C for 1 min, 45 cycles of 94°C for 10 s, 60°C for 1 min and final extension at 72°C for 10 min.

The amplification of the gene fragment was conducted using 2.5% agarose gel electrophoresis in TBE 1x buffer and SYBR® Safe (Invitrogen®, Carlsbad, CA, United States) as the intercalating agent and then visualized under UV light to observe a band of ∼650 base pairs (bp). Subsequently, the samples were sequenced by the Sanger method if the band corresponding to the fragment of the COI gene was observed.

COI (Barcoding) Gene Sequence Analysis

Sequence cleaning and alignment was performed in Seqman software (Lasergene) and the nucleotide sequence was compared to the Nucleotide Blast database[1]; the species were assigned considering the percentage of identity, the coverage value and *e*-value. The sequences were aligned using the MUSCLE algorithm in the MEGA X v10.1 software (Sohpal et al., 2010). From the alignments, a network of haplotypes was generated in PorpArt v1.7 software using the TCS algorithm (Clement et al., 2002), and haplotype diversity (h) was calculated in DNAsp v6 (Librado and Rozas, 2009). A Maximum Likelihood (ML) tree was built in the software IQtree (Nguyen et al., 2014) using the Jukes–Cantor substitution model, and the support of the nodes was established from 5,000 bootstrap replicates. The external group for the phylogenetic tree was *Culex spinosus* (KM593059.1).

Arbovirus Detection

The detection of each arbovirus (DENV-1/CHKV/ZIKV/MAYV) was conducted on the RNA collected by qRT-PCR One Step using the corresponding primers and Taqman probes (**Table 1**). Since Endogenous Flaviviral Elements have been identified in the mosquito genome (Houé et al., 2019), the chosen primers have an analytical validation to avoid cross reactions. The qRT-PCR reactions for DENV-1, CHIKV, and ZIKV (per sample) were performed in a final volume of 20 μL containing 10 μL qScript™ One-Step, Low ROX™ (Quantabio, Carlsbad, CA, United States), 10 μM of each primer, 10 μM of the probe and 5 μL of RNA. The protocol recommended by the WHO (World Health Organization [WHO], 2020) was used for the MAYV detection. The thermal profile used for all arboviruses consisted of the following: one cycle for 50°C for reverse transcriptase activation for 10 min, followed by initial denaturation at 95°C for 1 min, 40 cycles of 95°C for 15 s, 60°C for 1 min and final extension at 72°C for 10 min. The viral RNA from each

[1]https://blast.ncbi.nlm.nih.gov/Blast.cgi

TABLE 1 | Primers and probes specific for each arbovirus.

Virus	Primers and Probes	Gen (reference)
CHIKV	F: TCACTCCCTGTTGGACTTGATAGA R: TTGACGAACAGAGTTAGGAACATACC P: [FAM]-AGGTACGCGCTTCAAGTTCGGCG	NS (non-structural protein) (Lanciotti et al., 2007)
DENV	DENV-1 F: CAAAAGGAAGTCGYGCAATA R: CTGAGTGAATTCTCTCTGCTRAAC P: [FAM]-CATGTGGYTGGGAGCRCGC	NS5 (non-structural protein 5) (Santiago et al., 2013)
MAYV	F: GTGGTCGCACAGTGAATCTTTC R: CAAATGTCCACCAGGCGAAG P:[FAM]-ATGGTGGTAGGCTATCCGACAGGTC-carboxytetramethylrhodamine [TAMRA]	E2 (envelope protein 2) (Long et al., 2011)
ZIKV	F: AARTACACATACCARAACAAAGTG GT R: TCCRCTCCCYCTYTGGTCTTG′ P: [FAM]-CTYAGACCAGCTGAAR-BBQ	NS5 (non-structural protein 5) (Faye et al., 2013)

arbovirus was used as a positive control and was provided by the virology group of the National Institute of Health and the confirmation of positive tests was performed in duplicate. Finally, the infection frequency for each arbovirus species was expressed as the infection rate (infected individual/total individuals).

Statistical Analyses

Descriptive analyses of the frequency of each species and natural infection of each arbovirus were performed based on percentages. In addition, the kappa index was calculated to identify the percentage of agreement between the identification of *A. aegypti* and *A. albopictus* species by taxonomic code and molecular marker. A logistic regression, including the intercept, was performed to identify the statistical associations between arbovirus infection as a composite variable (y) and the explanatory variables: vector species, municipality and collection area. The species variable and collection area were transformed into dummies (1/0) using *A. albopictus* and the rural area as references (0). Preliminarily, a main effects model was run to identify variables potentially associated with the established outcome and a p-value < 0.2 was significant. Subsequently, a step-by-step, second-level interaction model was executed, which included interactions between vector species and the collection area or municipality.

RESULTS

Morphological and Molecular Identification of *Aedes* Species

In the two sampling periods, 169 specimens of *Aedes* were captured with 68% (n = 115) collected in the municipality of Saravena, 15.4% (n = 26) in Arauquita and 16.6% (n = 28) in Tame. Overall, differences were observed in the number of individuals identified as *A. aegypti* and *A. albopictus* between molecular and morphological identification methods (**Figure 2**). A kappa index of 0.764 and agreement of 91.1% (n = 154) were obtained between species identification methods (taxonomic key vs. molecular) (Cook, 2005). Inconsistencies were mainly evident in the identification of 13 individuals (7.7%), where they were initially identified through morphology as *A. aegypti* while the molecular marker analysis indicated that they were *A. albopictus*.

COI Gene Sequence Analysis

A 338 bp fragment of the COI gene was obtained. In the haplotype network constructed from the sequences of the individuals of the genus *Aedes*, 19 haplotypes were found (**Figure 3**). We observed two groups that corresponded to the species *A. aegypti* (Orange) and *A. albopictus* (Purple) and are differentiated by 36 bp (**Figure 3A**). Additionally, the highest number of haplotypes of the two species was found in the municipality of Saravena (Azul) with 8 haplotypes of the species *A. aegypti* and 9 haplotypes of the species *A. albopictus* (**Figure 3B**). The diversity of haplotypes found in Saravena was $h = 0.806$ and in Arauquita $h = 0.498$, and only one haplotype was found in Tame. Furthermore, the global diversity of haplotypes for all individuals captured in Arauca was $h = 0.639$.

The topology of the phylogenetic tree obtained from the COI sequence alignment revealed two well-supported clusters (bootstrap ≥ 90) that correspond to *A. aegypti* and *A. albopictus* (**Figure 3C**) with a genetic distance of 0.107 between the species. Additionally, the genus *Culex* (included as an outgroup) formed an independent clade in the species of genus *Aedes*.

Geographical Distribution of *A. aegypti* and *A. albopictus*

For the results of the geographic distribution of individuals and their natural infection by arboviruses, molecular identification of the species was considered. **Table 2** shows the information regarding the frequency of *A. aegypti* individuals (n = 111) and *A. albopictus* (n = 58) by area and the municipality of capture. The species *A. aegypti* was present in the three municipalities and only in urban areas (**Figure 4A**) and there was a greater frequency of this species in the municipality of Saravena at 64.9% (72/111). However, *A. albopictus*, which is only present in the municipalities of Saravena and Arauquita in both urban and rural areas (**Figure 4A**), showed a frequency of 74.1% (43/58) in the

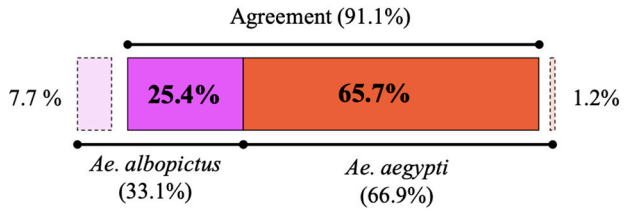

FIGURE 2 | Percentage of *A. albopictus* (purple) and *A. aegypti* (orange) individuals identified by morphological and molecular testing. The dotted lines mark the inconsistencies where individuals were initially assigned by morphology to a different species compared to the one detected by molecular methods, and the latter is the technique by which individuals were ultimately assigned for subsequent sequence analysis.

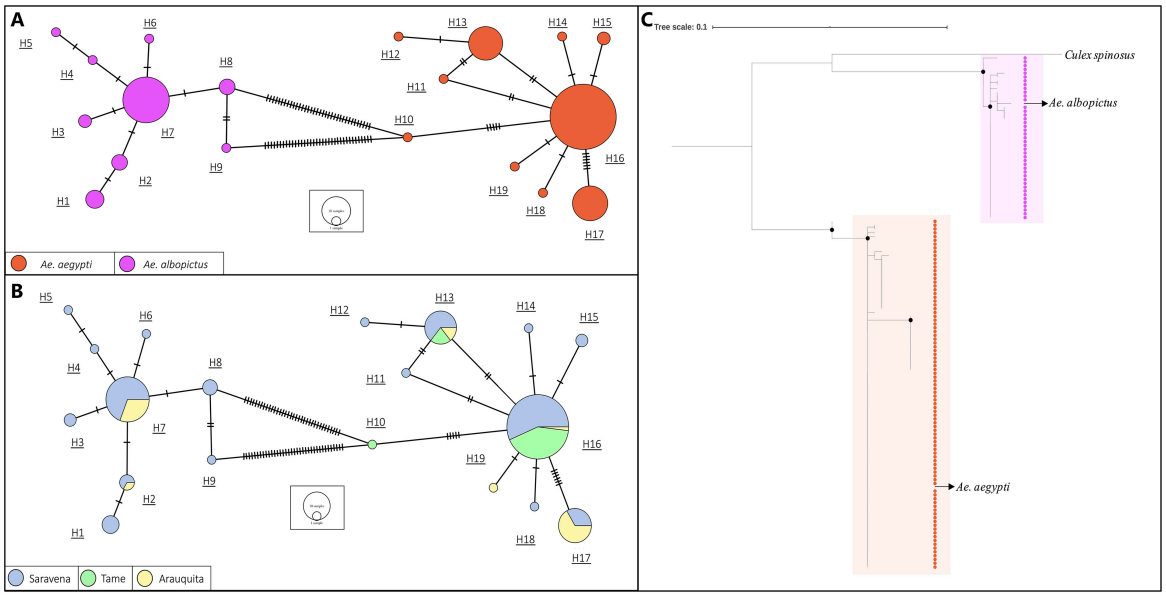

FIGURE 3 | The haplotype network based on sequences of *A. aegypti* and *A. albopictus* collected in municipalities of the department of Arauca and colored by **(A)** species and by **(B)** collection municipality. The lines in each branch indicate the number of mutational steps between haplotypes, and the size of the circle is proportional to the number of individuals. **(C)** A maximum likelihood tree based on the JC model of *A. aegypti* and *A. albopictus* indivuduals from each place where they were collected. Black dots indicate support bootstrap ≥ 90.

municipality of Arauquita. The above results show clear evidence of the presence of these two species in the urban area of the municipality of Saravena and Arauquita.

Natural Infection by Arboviruses

Of the 169 individuals collected from the genus *Aedes*, the overall rate of arboviruses (DENV-1 and CHIKV) infection was 36.1% (61/169). The arbovirus with the highest frequency was DENV-1, which was found in 82% (50/61) of infected individuals, while 9.8% (6/61) were infected by CHIKV and none of the individuals were infected with ZIKV or MAYV. In addition, a 4.5% frequency (5/111) of mixed infection by arboviruses DENV-1 and CHIKV was found in *A. aegypti*. **Table 2** shows the arbovirus infection rates in each species, and there was a higher infection rate observed in individuals of the species *A. albopictus*. In Saravena was found 18 individuals of *A. aegypti* and 17 of *A. albopictus* infected by DENV-1, while for CHIKV one individual of *A. aegypti* and two individuals of *A. albopictus* were found. Only 5 individuals of *A. aegypti* were found coinfected by DENV-1 and CHIKV. In Arauquita were found 5 individuals of *A. aegypti* and 6 individuals of *A. albopictus* infected by DENV-1, while for CHIKV one individual of *A. aegypti* and two individuals of *A. albopictus* were found in Tame, only 4 individuals of *A. aegypti* were found infected for DENV-1. The logistic regression results showed that the municipality of Arauquita is 7 times more likely to find infected individuals (OR = 7; 95% CI = 1.890–25.932; $p = 0.004$) compared to Tame. However, the municipality of Saravena was 3.6 more likely to find infected individuals compared to Tame (OR = 3.6; 95% CI = 1.165–11.025; $p = 0.026$) (**Figure 4B**) in the **Supplementary Table 1** are shown all the parameters of the models.

TABLE 2 | Number and percentage of *A. aegypti* and *A. albopictus* captured per year, area, municipality and infected with arboviruses.

	A. aegypti n (%)	*A. albopictus n* (%)	Total
Capture	111 (65, 7)	58 (34, 3)	169
Year			
2018	51 (49)	53 (51)	104
2019	60 (92.3)	5 (7.7)	65
Zone			
Rural	0	37 (100)	37
Urban	111 (84.1)	21 (15.9)	132
Municipality			
Saravena	72 (62.6)	43 (37.4)	115
Arauquita	11 (42.3)	15 (57.7)	26
Tame	28 (100)	0	28
Rate infection with arbovirus			
DENV-1	27 (54)	23 (46)	50 (29.6)
CHIKV	2 (33.3)	4 (66.7)	6 (3.5)
Mixed (DENV-1 y CHIKV)	5 (100)		5 (3)
Uninfected	77 (71.3)	31 (28.7)	108 (63,9)

DISCUSSION

The *A. aegypti* and *A. albopictus* classification of individuals is mainly determined by morphological identification (Kraemer et al., 2015), and few reports use molecular identification (Zamora-Delgado et al., 2015; Atencia et al., 2018).

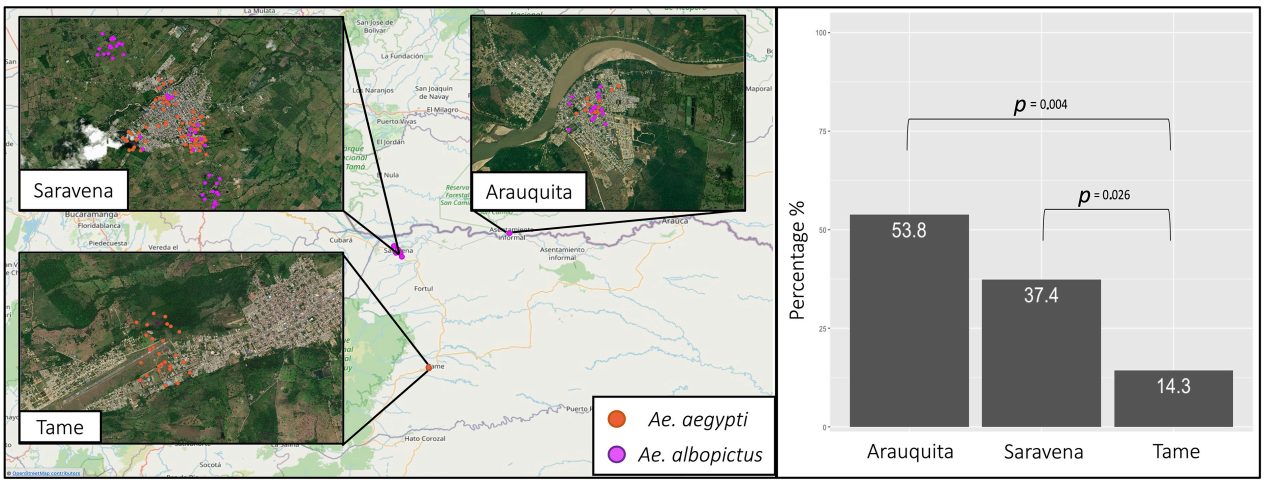

FIGURE 4 | (A) The geographic location of *A. aegypti* (orange) and *A. albopictus* (purple) collected from the municipalities of Saravena, Arauquita, and Tame. **(B)** The percentage of individuals infected by the municipalities of Arauquita, Saravena, and Tame.

Although, morphological identification is the most widely used classification technique, it has different limitations. During the transport of individuals, they are susceptible to damage and modifications in their external parts, which can generate confusion during the classification of individuals (Sumruayphol et al., 2016). Taking this into account, it is necessary to use complementary tools that allow a correct identification of individuals. Among these tools it is recommended to use molecular techniques such as DNA barcoding (Cook et al., 2005; Anoopkumar et al., 2019). In this study was described the precision of taxonomic keys and DNA barcoding to classify individuals of the genus *Aedes*.

Regarding the analysis of the COI sequences, a global diversity of haplotypes value of 0.639 was evident for species *A. aegypti* and *A. albopictus*. This value is similar to that reported by Caldera et al. (2019) for individuals captured in Colombia and Joyce et al. (2018) for El Salvador; these values are relatively high. This high haplotype diversity may be due to adaptation processes for different environmental conditions and ecosystems where the insect is exposed due to its cosmopolitan distribution. Accordingly, the high number and diversity of haplotypes present in Saravena (h = 0.806, **Figure 3B**) may be due to evolutionary pressures that generate genetic changes in these dipteran species such as the use of insecticides. Between the two species, 36 mutations (**Figure 3A**) and a genetic distance of 0.107 were found. According to Hoyos-Lopez et al. (2015), this value is found for inter-species ranges (0.022–0.2565) that were previously reported in individuals of the Culicidae family. The tree topology (**Figure 3C**) is consistent with these results since it is evident that the individuals constitute two clusters grouped by species. Considering the groups formed by species in the haplotype network (**Figure 3A**) and the clustering by species in the phylogenetic tree (**Figure 3C**), it is possible to determine the precision of molecular tests to identify *A. aegypti* and *A. albopictus*.

According to the kappa index obtained in this study between the identification methods (k = 0.7643) (Cook, 2005), a deficiency in identification through morphology can occur (**Figure 2**). This could be an indicator of an overestimate of *A. aegypti* by traditional morphology-based techniques. The main characteristic that permits differentiation between these two species is the pattern in the scutum (mesothorax sclerite). These patterns are described as susceptible to blurring during insect storage processing (Patsoula et al., 2006). Therefore, there may be a risk of incorrect morphological identification and any error in this process has a negative effect on studies that make conclusions on the biology and vector capacity of each of these insect species. In this scenario, molecular identification may become more accurate (Chan et al., 2014). However, molecular techniques are too expensive for routine use and especially in areas with high vector infestation. Therefore, we propose to use this complementary tool mainly in three situations; (i) generate a baseline on the proportions of individuals per species in areas where this information is not available, (ii) as a quality control of morphological identification, choosing a random percentage of individuals to be confirmed by DNA barcoding and (iii) when individuals are damaged and morphological identification cannot be performed.

Studies focusing on the distribution of *A. aegypti* and *A. albopictus* conclude that their presence is due to the interaction of different biotic and abiotic factors (Lozano-Fuentes et al., 2012; Liang et al., 2019). This interaction is shown in the models that show an association between land use and the presence of *A. aegypti* and *A. albopictus* and the prevalence of arboviruses. The main aspects at risk are human settlements, horticulture and livestock (Cheong et al., 2014), and this is mainly because these activities generate micro-habitats that favor the reproduction of the vector (Sarfraz et al., 2012). In accordance with the above, there is a coexistence of *A. aegypti* and *A. albopictus* in urban areas of the Saravena and Arauquita municipalities (**Figure 4A**).

Countries, such as Brazil and Panama, show evidence of this observation (Li et al., 2014; Whiteman et al., 2019). Therefore, the coexistence of these two species in the same area generates a public health risk since they are two vectors contributing to the transmission of arboviruses in the same urban area. Future studies should consider evaluating the transmission dynamics of arboviruses in this region by examining the plausible biological competition of the two species.

However, in the municipality of Arauquita the species that is found in higher frequency is *A. albopictus* (**Table 2**). This may be due to the different land uses in each municipality. For example, in Arauquita, there is a high availability of natural water due to the presence of the Arauca river in the northern part of the municipality. The above benefits the reproduction of *A. albopictus* since this species has preferences for these bodies of water compared to *A. aegypti* (Claeys et al., 2016). In this study we do not report data about immature collections to identify their most common breeding sites, however, this phenomenon has been previously reported in Vietnam and Brazil (Higa et al., 2010; Barbosa et al., 2020), and in Brazil, a spatial analysis was conducted that determined an association between the presence of vectors, depending on the type of water body (Lozano-Fuentes et al., 2012; Liang et al., 2019). Considering the aforementioned, it is possible that other factors are interacting in the municipalities that influence the geographical distribution of these species and subsequently the transmission dynamics of the arboviruses.

The logistic regression results show that in the municipality of Arauquita there is a higher risk of finding individuals infected with arboviruses (**Figure 4B**). In accordance with our results, different models (Barmak et al., 2016; Massaro et al., 2019; Zhu et al., 2019) show that human mobility is an important factor for the increase in the incidence of DENV even over short distances, i.e., between cities. Considering the aforementioned, the municipality of Arauquita is in a border area with Venezuela, and therefore, there is a high movement of humans that may maintain the virus in these areas. In accordance with the aforementioned, Lizarazo et al. (2019) showed that DENV-1 strains isolated in Venezuela were closely related to previously published sequences of this serotype in Colombia. This corroborates the influence that the movement of people between these two countries has on the prevalence of infection in insects.

DENV-1 was the arbovirus with the highest frequency in the population of infected individuals at 82% (50/61) and the overall infection rate in *A. aegypti* and *A. albopictus* was 29.6% (50/169) (**Table 2**). The DENV-1 serotype has also been detected more frequently than other arboviruses in Indonesia, Mexico, Cuba, Brazil, and Colombia (Velandia-Romero et al., 2017; Eiras et al., 2018; Garcia-Rejon et al., 2018; Gutiérrez-Bugallo et al., 2018; Rahayu et al., 2019). The high frequency of this arbovirus may be because the transmission in the urban cycle of DENV is provided by the level of viremia of the hosts (humans) since individuals at the beginning of infection have high rates of viremia (Duong et al., 2015; Martínez-Vega et al., 2015). This agrees with the results obtained in this study, given that DENV-1 is the most prevalent arbovirus in the department of Arauca (Boletín Epidemiológico Semanal (BES) Semana Epidemiológica 22, 24 al 30 de mayo, 2020). Furthermore, there is a need to conduct complementary entomo-virological surveillance to actively search for human cases and to have a complete overview of the urban transmission of arboviruses. Nevertheless, future studies should consider the detection of the other 3 DENV serotypes in vectors to complement the understanding of transmission dynamics in this endemic region and due to its proximity to Venezuelan states that are highly endemic for DENV-4 (Grillet et al., 2019).

The results in this study show a higher infection rate by DENV-1 and CHIKV in *A. albopictus* (**Table 2**) compared to *A. aegypti*. This does not agree with the values previously reported since the MIR values in *A. aegypti* are higher (Guerbois et al., 2016; Huerta et al., 2017; Garcia-Rejon et al., 2018). However, it was experimentally shown that *A. albopictus* is a vector more susceptible to infection than *A. aegypti* because the arbovirus dissemination rates were higher in nearly all strains of *A. albopictus* tested (Turell et al., 1992). The dissemination rate is associated with the CHIKV genotype, where *A. albopictus* is more susceptible to infection of the Asian genotype and this genotype is reported more frequently than the African genotype in Colombia (Camacho Burgos, 2019). The lack of natural ZIKV infection is due to the low number of cases reported in humans in the department of Arauca over the last 2 years (Boletín Epidemiológico Semanal (BES) Semana Epidemiológica 22, 24 al 30 de mayo, 2020). This confirms that this arbovirus is not circulating in the urban cycle of the three sampled municipalities.

However, the absence of individuals infected by MAYV may be because the insects processed in this study are part of the urban cycle and no studies demonstrate the efficiency of viral transmission in this cycle. However, the study by Pereira et al. (2020) shows that under laboratory conditions, *A. aegypti* and *A. albopictus* have the ability to become infected and the potential to transmit MAYV genotype D. Likewise, it is possible to infect mosquitoes with the virus from the saliva of individuals who initially tested negative. This indicates that for the infection to occur, a low viral load is necessary, which cannot be detected by PCR, and this is reflected in the MIR values of 0.4 (95% CI = 0.0–1.4) that are reported for this species (Maia et al., 2019). This is why different studies propose viral enrichment techniques that allow the viral sequences to increase, which is corroborated by next generation sequencing (NGS) (Hall et al., 2014). Therefore, it is important that future studies implement viral enrichment techniques that allow virus detection even at low concentrations.

Natural coinfection between DENV-2 and CHIKV has been previously reported in *A. albopictus*, where an association with simultaneous circulation of these two arboviruses during a dengue outbreak in Central Africa is possible (Caron et al., 2012). In this study, a mixed DENV-1 and CHIKV coinfection rate of 4.5% (5/111) was found in *A. aegypti*. This shows a simultaneous circulation of these two arboviruses within one specimen, which can lead to coinfection in humans, where an increase in coinfection cases would lead to rapid genetic evolution of the virus (Caron et al., 2012). There are two methods

to estimate the infection rate from detection by pools: MIR (Minimum Infection Rate) and MLE (Maximum Likelihood Estimate). However, to produce a correct estimate, it is necessary to obtain a sampling ≥ 1,000 individuals and sizes of small pools. Processing individually allows us to naturally observe mixed infections in vectors including the option to conduct molecular discrimination of *Aedes* species (Walter et al., 1980; Gu et al., 2004). Therefore, the importance of conducting surveillance of arboviruses in the individual vector in areas where there is no information on natural infection in the vectors is important. This generates a baseline and avoids erroneous data regarding natural arbovirus infection in each of the vectors. In places where such a baseline exists and there are large samplings, the processing of insects by pools is possible. Future studies should compare the most reliable method for expressing infection rates in arboviral diseases in terms of sensitivity and cost-effectiveness.

CONCLUSION

This study reports on the natural CHIKV infection of *A. aegypti* and *A. albopictus* in Colombia, and a higher infection rate was found in the latter species. This shows the susceptibility of this species to infection by arboviruses. In addition, DENV-1 was the arbovirus with the highest frequency in the collected individuals, which may be associated with different factors such as the transovarian transmission of the vector and human infected with high viremias that can maintain constant transmission of the arbovirus. Expression of the natural infection data using infection rates to obtain a value per individual on arbovirus infection is proposed which increased the infection rate and the finding of viral co-infections. Lastly, we demonstrated the coexistence of *A. aegypti* and *A. albopictus* in urban areas of Saravena and Arauquita. This coexistence can be explained by the different land uses in the municipalities and the availability of water bodies that allow correct reproduction and development of the vectors. Our findings could become valuable to improve national control programs for arbovirosis. However, it is important to highlight some limitations of the study, among them the lack of detection of the DENV-2 – 4 serotypes as well as the measurement of other meteorological variables that would allow us to carry out a more robust analysis. Likewise, a longer sampling time that could allow us a greater capture of individuals and a more general view of the behavior of the infection of arboviruses in *A. aegypti* and *A. albopictus*. Lastly, the non-sequencing of the positive samples for DENV1 and CHIKV and their comparison with sequences available on GenBank. In the future, studies on this region must be also conducted to depict the molecular epidemiology of this arboviral agents for high resolution epidemiological surveillance.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

YA and AC designed the protocol and carried out the capture of the individuals. DM, CH, and JR performed the processing of the samples in the laboratory. MM carried out the statistical and phylogenetic analysis of the results obtained. DM, CH, MM, and JR wrote the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

We would like to thank the Dirección de Investigación e Innovación from Universidad del Rosario for funding this study and providing the editing English service of ENAGO and cover the publication fees.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.602190/full#supplementary-material

Supplementary Table 1 | Logistic regression. The parameters of Main effects model and Second level interaction model.

REFERENCES

Anoopkumar, A. N., Puthur, S., Rebello, S., and Aneesh, E. M. (2019). Molecular characterization of Aedes, Culex, Anopheles, and Armigeres vector mosquitoes inferred by mitochondrial cytochrome oxidase I gene sequence analysis. *Biologia* 74, 1125–1138. doi: 10.2478/s11756-019-00231-0

Atencia, M. C., Pérez, M. D. J., Caldera, S. M., Jaramillo, M. C., and Bejarano, E. E. (2018). Genetic variability of *Aedes aegypti* in the department of sucre, colombia, by analysis of the nucleotide sequence of the mitochondrial ND4 gene. *Biomédica* 38, 267–276. doi: 10.7705/biomedica.v38i0.3728

Azevedo, R. S., Silva, E. V., Carvalho, V. L., Rodrigues, S. G., Neto, J. P. N., Monteiro, H. A., et al. (2009). Mayaro fever virus, Brazilian amazon. *Emerg. Infect. Dis.* 15:1830. doi: 10.3201/eid1511.090461

Barbosa, R. M. R., Melo-Santos, M. A. V. D., Silveira, Jr, J. C., Silva-Filha, M. H. N. L., Souza, W. V., Oliveira, C. M. F. D., et al. (2020). Infestation of an endemic arbovirus area by sympatric populations of Aedes aegypti and Aedes albopictus in Brazil. *Mem. Instit. Oswaldo Cruz* 115: e190437. doi: 10.1590/0074-02760190437

Barmak, D. H., Dorso, C. O., and Otero, M. (2016). Modelling dengue epidemic spreading with human mobility. *Phys. A Statist. Mech. Appl.* 447, 129–140. doi: 10.1016/j.physa.2015.12.015

Beaty, B. J., Black, W. C., Eisen, L., Flores, A. E., García-Rejón, J. E., Loroño-Pino, M., and Saavedra-Rodriguez, K. (2016). "The intensifying storm: Domestication of Aedes aegypti, urbanization of arboviruses, and emerging insecticide resistance," in *Proceedings of the Global Health Im-pacts of Vector-Borne Diseases: Workshop Summary* (Washington, DC: National Academies Press).

Black, W. C. IV., Bennett, K. E., Gorrochótegui-Escalante, N., Barillas-Mury, C. V., Fernández-Salas, I., de Lourdes Muñoz, M., et al. (2002). Flavivirus susceptibility in Aedes aegypti. *Arch. Med. Res.* 33, 379–388. doi: 10.1016/S0188-4409(02)00373-9

Boletín Epidemiológico Semanal (BES) Semana Epidemiológica 22, 24 al 30 de mayo (2020). Obtenido de. Available at : https://www.ins.gov.co/buscador-eventos/BoletinEpidemiologico/2020_Bolet%C3%ADn_epidemiológico_semana%2022.pdf (accessed July 20, 2020).

Caldera, S. M., Jaramillo, M. C., Cochero, S., Pérez Doria, A., and Bejarano, E. E. (2019). Diferencias genéticas entre poblaciones de *Aedes aegypti* de municipios del norte de Colombia, con baja y alta incidencia de dengue. *Revista* 33, doi: 10.7705/biomedica.v33i0.1573

Camacho Burgos, E. Y. (2019). *Caracterización filogenética de los virus dengue, Chikungunya y Zika circulantes en el departamento de Sucre durante el periodo 2013–2015*, Colombia: Universidad de Sucre.

Caron, M., Paupy, C., Grard, G., Becquart, P., Mombo, I., Nso, B. B. B., et al. (2012). Recent introduction and rapid dissemination of Chikungunya virus and Dengue virus serotype 2 associated with human and mosquito coinfections in Gabon, central Africa. *Clin. Infect. Dis.* 55, e45-e53. doi: 10.1093/cid/cis530

Cevallos, V., Ponce, P., Waggoner, J. J., Pinsky, B. A., Coloma, J., Quiroga, C., et al. (2018). Zika and Chikungunya virus detection in naturally infected Aedes aegypti in Ecuador. *Acta Trop.* 177, 74–80. doi: 10.1016/j.actatropica.2017.09.029

Chan, A., Chiang, L. P., Hapuarachchi, H. C., Tan, C. H., Pang, S. C., Lee, R., et al. (2014). DNA barcoding: complementing morphological identification of mosquito species in Singapore. *Paras. Vect.* 7, 1–12. doi: 10.1186/s13071-014-0569-4

Cheong, Y. L., Leitão, P. J., and Lakes, T. (2014). Assessment of land use factors associated with dengue cases in Malaysia using boosted regression trees. *Spatial Spatio Temp. Epidemiol.* 10, 75–84. doi: 10.1016/j.sste.2014.05.002

Claeys, C., Robles, C., Bertaudiere-Montes, V., Deschamps-Cottin, M., Megnifo, H. T., Pelagie-Moutenda, R., et al. (2016). Socio-ecological factors contributing to the exposure of human populations to mosquito bites that transmit dengue fever, chikungunya and zika viruses: a comparison between mainland France and the French Antilles. *Environ. Risques Santé* 15, 318–325.

Clement, M., Snell, Q., Walker, P., Posada, D., and Crandall, K. (2002). TCS: estimating gene genealogies. *Parall. Distrib. Process. Symp. Int.* 2:0184. doi: 10.1109/IPDPS.2002.1016585

Cook, R. J. (2005). Kappa. *Encyclop. Biostatist.* 4. doi: 10.1002/0470011815.b2a04023

Cook, S., Diallo, M., Sall, A. A., Cooper, A., and Holmes, E. C. (2005). Mitochondrial markers for molecular identification of Aedes mosquitoes (Diptera: Culicidae) involved in transmission of arboviral disease in West Africa. *J. Med. Entomol.* 42, 19–28. doi: 10.1603/0022-2585(2005)042[0019:mmfmio]2.0.co;2

Cova-Garcia, P., Sutil, E., and Rausseo, J. A. (1966). *Mosquitos (Culicinos) de Venezuela: Tomo I and Tomo II*. Caracas: Ministerio de Sanidad y Asistencia Social.

de la Cruz, C. H., Martinez, S. L. A., Failoc-Rojas, V. E., and Aguilar-Gamboa, F. R. (2019). Momento de considerar otras arbovirosis luego del virus mayaro. *Rev. Cubana Med. Gen. Integ.* 35.

Diallo, D., Sall, A. A., Diagne, C. T., Faye, O., Faye, O., Ba, Y., et al. (2014). Zika virus emergence in mosquitoes in southeastern Senegal, 2011. *PLoS One* 9:e109442. doi: 10.1371/journal.pone.0109442

Duong, V., Lambrechts, L., Paul, R. E., Ly, S., Lay, R. S., Long, K. C., et al. (2015). Asymptomatic humans transmit dengue virus to mosquitoes. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14688–14693. doi: 10.1073/pnas.1508114112

Eiras, A. E., Pires, S. F., Staunton, K. M., Paixão, K. S., Resende, M. C., Silva, H. A., et al. (2018). A high-risk Zika and dengue transmission hub: virus detections in mosquitoes at a Brazilian university campus. *Paras. Vect.* 11:359. doi: 10.1186/s13071-018-2883-8

Espinal, M. A., Andrus, J. K., Jauregui, B., Hull Waterman, S., Morens, D. M., Santos, J. I., et al. (2019). Arbovirosis emergentes y reemergentes transmitidas por Aedes en la Región de las Américas: implicaciones en materia de políticas de salud. *Rev. Panam. Salud. Publ.* 43.

Faye, O., Diallo, D., Diallo, M., Weidmann, M., and Sall, A. A. (2013). Quantitative real-time PCR detection of Zika virus and evaluation with field-caught mosquitoes. *Virol. J.* 10:311. doi: 10.1186/1743-422x-10-311

Forattini, O. P. (1965). *Entomologia médica: culicini: culex, Aedes e Psorophora*, Vol. 2. São Paulo: Faculdade de Higiene e Saúde Pública, Depto. de Parasitología.

Freitas, R. M. D. (2010). A review on the ecological determinants of Aedes aegypti (Diptera: Culicidae) vectorial capacity. *Oecol. Austr.* 14:726–736 doi: 10.4257/oeco.2010.1403.08

Garcia-Rejon, J. E., Ulloa-Garcia, A., Cigarroa-Toledo, N., Pech-May, A., Machain-Williams, C., Cetina-Trejo, R. C., et al. (2018). Study of Aedes aegypti population with emphasis on the gonotrophic cycle length and identification of arboviruses: implications for vector management in cemeteries. *Rev. Instit. Med. Trop. São Paul.* 60:e44 doi: 10.1590/s1678-9946201860044

Grillet, M. E., Hernández-Villena, J. V., Llewellyn, M. S., Paniz-Mondolfi, A. E., Tami, A., Vincenti-Gonzalez, M. F., et al. (2019). Venezuela's humanitarian crisis, resurgence of vector-borne diseases, and implications for spillover in the region. *Lancet Infect. Dis.* 19, e149-e161. doi: 10.1016/S1473-3099(18)30757-6

Gu, W., Lampman, R., and Novak, R. J. (2004). Assessment of arbovirus vector infection rates using variable size pooling. *Med. Vet. Entomol.* 18, 200-204 doi: 10.1111/j.0269-283X.2004.00482.x

Guerbois, M., Fernandez-Salas, I., Azar, S. R., Danis-Lozano, R., Alpuche-Aranda, C. M., Leal, G., et al. (2016). Outbreak of Zika virus infection, Chiapas State, Mexico, 2015, and first confirmed transmission by Aedes aegypti mosquitoes in the Americas. *J. Infect. Dis.* 214, 1349–1356. doi: 10.1093/infdis/jiw302

Gutiérrez-Bugallo, G., Rodríguez-Roche, R., Díaz, G., Pérez, M., Mendizábal, M. E., Peraza, I., et al. (2018). Spatio-temporal distribution of vertically transmitted dengue viruses by Aedes aegypti (Diptera: Culicidae) from Arroyo Naranjo, Havana, Cuba. *Trop. Med. Int. Health* 23, 1342–1349. doi: 10.1111/tmi.13162

Hall, R. J., Wang, J., Todd, A. K., Bissielo, A. B., Yen, S., Strydom, H., et al. (2014). Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery. *J. Virol. Methods* 195, 194–204. doi: 10.1016/j.jviromet.2013.08.035

Higa, Y., Yen, N. T., Kawada, H., Son, T. H., Hoa, N. T., and Takagi, M. (2010). Geographic distribution of Aedes aegypti and Aedes albopictus collected from used tires in Vietnam. *J. Am. Mosq. Control Assoc.* 26, 1–9. doi: 10.2987/09-5945.1

Houé, V., Bonizzoni, M., and Failloux, A. B. (2019). Endogenous non-retroviral elements in genomes of Aedes mosquitoes and vector competence. *Emerg. Microb. Infect.* 8, 542–555. doi: 10.1080/22221751.2019.1599302

Hoyos-Lopez, R., Roman Pardo, S., CASTAÑO, J. C., and GALLEGO-GÓMEZ, J. C. (2015). DNA barcode for typing of immature mosquitoes from Armenia and Circasia (Quindío, Colombia). *Rev. Colomb. Entomol.* 41, 218-227

Huerta, H., González-Roldán, J. F., Sánchez-Tejeda, G., Correa-Morales, F., Romero-Contreras, F. E., Cárdenas-Flores, R., et al. (2017). Detection of Zika virus in Aedes mosquitoes from Mexico. *Trans. Royal Soc. Trop. Med. Hygiene* 111, 328–331. doi: 10.1093/trstmh/trx056

Joyce, A. L., Torres, M. M., Torres, R., and Moreno, M. (2018). Genetic variability of the Aedes aegypti (Diptera: Culicidae) mosquito in El Salvador, vector of dengue, yellow fever, chikungunya and Zika. *Paras. Vect.* 11:637. doi: 10.1186/s13071-018-3226-5

Kraemer, M. U., Sinka, M. E., Duda, K. A., Mylne, A. Q., Shearer, F. M., Barker, C. M., et al. (2015). The global distribution of the arbovirus vectors Aedes aegypti and Ae. *albopictus*. *eLife* 4:e08347. doi: 10.7554/eLife.08347

Lanciotti, R. S., Kosoy, O. L., Laven, J. J., Panella, A. J., Velez, J. O., Lambert, A. J., et al. (2007). Chikungunya virus in US travelers returning from India, 2006. *Emerg. Infect. Dis.* 13, 764–767. doi: 10.3201/eid1305.070015

Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., et al. (2014). Urbanization increases Aedes albopictus larval habitats and accelerates mosquito development and survivorship. *PLoS Negl. Trop. Dis.* 8:e3301. doi: 10.1371/journal.pntd.0003301

Liang, L., Jia, P., Tan, X., Chen, J., and Chen, X. (2019). Potential effects of heat waves on the population dynamics of the dengue mosquito Aedes albopictus. *PLoS Neglect. Trop. Dis.* 13:e0007528. doi: 10.1371/journal.pntd.0007528

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187

Liu, Y., Lillepold, K., Semenza, J. C., Tozan, Y., Quam, M. B., and Rocklöv, J. (2020). Reviewing estimates of the basic reproduction number for dengue, Zika and chikungunya across global climate zones. *Environ. Res.* 182:109114. doi: 10.1016/j.envres.2020.109114

Lizarazo, E., Vincenti-Gonzalez, M., Grillet, M. E., Bethencourt, S., Diaz, O., Ojeda, N., et al. (2019). Spatial dynamics of Chikungunya virus, Venezuela, 2014. *Emerg. Infect. Dis.* 25:672. doi: 10.3201/eid2504.172121

Long, K. C., Ziegler, S. A., Thangamani, S., Hausser, N. L., Kochel, T. J., Higgs, S., et al. (2011). Experimental transmission of Mayaro virus by *Aedes aegypti*. *Am. J. Trop. Med. Hyg.* 85, 750–757. doi: 10.4269/ajtmh.2011.11-0359

Lorenz, C., Ribeiro, A. F., and Chiaravalloti-Neto, F. (2019). Mayaro virus distribution in South America. *Acta tropica* 198:105093. doi: 10.1016/j.actatropica.2019.105093

Lounibos, L. P., and Kramer, L. D. (2016). Invasiveness of Aedes aegypti and Aedes albopictus and vectorial capacity for chikungunya virus. *J. Infect. Dis.* 214(suppl_5), S453-S458. doi: 10.1093/infdis/jiw285

Lozano-Fuentes, S., Hayden, M. H., Welsh-Rodriguez, C., Ochoa-Martinez, C., Tapia-Santos, B., Kobylinski, K. C., et al. (2012). The dengue virus mosquito vector Aedes aegypti at high elevation in Mexico. *Am. J. Trop. Med. Hyg.* 87, 902-909. doi: 10.4269/ajtmh.2012.12-0244

Maia, L. M. S., Bezerra, M. C. F., Costa, M. C. S., Souza, E. M., Oliveira, M. E. B., Ribeiro, A. L. M., et al. (2019). Natural vertical infection by dengue virus serotype 4, Zika virus and Mayaro virus in Aedes (Stegomyia) aegypti and Aedes (Stegomyia) albopictus. *Med. Vet. Entomol.* 33, 437–442. doi: 10.1111/mve.12369

Martínez-Vega, R. A., Danis-Lozano, R., Díaz-Quijano, F. A., Velasco-Hernández, J., Santos-Luna, R., Román-Pérez, S., et al. (2015). Peridomestic infection as a determining factor of dengue transmission. *PLoS Neglect. Trop. Dis.* 9:e0004296. doi: 10.1371/journal.pntd.0004296

Massaro, E., Kondor, D., and Ratti, C. (2019). Assessing the interplay between human mobility and mosquito borne diseases in urban environments. *Sci. Rep.* 9, 1–13 doi: 10.1038/s41598-019-53127-z

Nguyen, LT, Schmidt, HA, von Haeseler, A, Minh, BQ. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Patsoula, E., Samanidou-Voyadjoglou, A., Spanakos, G., Kremastinou, J., Nasioulas, G., and Vakalis, N. C. (2006). Molecular and morphological characterization of Aedes albopictus in northwestern Greece and differentiation from Aedes cretinus and Aedes aegypti. *J. Med. Entomol.* 43, 40–54. doi: 10.1603/0022-2585(2006)043[0040:mamcoa]2.0.co;2

Paupy, C., Delatte, H., Bagny, L., Corbel, V., and Fontenille, D. (2009). Aedes albopictus, an arbovirus vector: from the darkness to the light. *Microbes Infect.* 11, 1177–1185. doi: 10.1016/j.micinf.2009.05.005

Pereira, T. N., Carvalho, F. D., De Mendonça, S. F., Rocha, M. N., and Moreira, L. A. (2020). Vector competence of *Aedes aegypti*, *Aedes albopictus*, and Culex quinquefasciatus mosquitoes for Mayaro virus. *PLoS Negl. Trop. Dis.* 14:e0007518. doi: 10.1371/journal.pntd.0007518

Rahayu, A., Saraswati, U., Supriyati, E., Kumalawati, D. A., Hermantara, R., Rovik, A., et al. (2019). Prevalence and distribution of dengue virus in *Aedes aegypti* in yogyakarta city before deployment of Wolbachia infected *Aedes aegypti*. *Int. J. Environ. Res. Public Health* 16:1742. doi: 10.3390/ijerph16101742

Ramasamy, R., Surendran, S. N., Jude, P. J., Dharshini, S., and Vinobaba, M. (2011). Larval development of Aedes aegypti and Aedes albopictus in peri-urban brackish water and its implications for transmission of arboviral diseases. *PLoS Negl. Trop. Dis.* 5:e1369. doi: 10.1371/journal.pntd.0001369

Rolo, E. S. A. (2020). DNA Barcoding and Forensic Entomology: A Molecular Approach for Diptera Species' Identification. Doctoral dissertation, University of Cape Town.

Salehi, Z, Najafi, M. (2014). RNA preservation and stabilizaton. *Biochem. Physiol. Open Access.* 3, 1–4.

Santiago, G. A., Vergne, E., Quiles, Y., Cosme, J., Vazquez, J., Medina, J. F., et al. (2013). Analytical and clinical performance of the CDC real time RT-PCR assay for detection and typing of dengue virus. *PLoS Negl. Trop. Dis.* 7:e2311. doi: 10.1371/journal.pntd.0002311

Sarfraz, M. S., Tripathi, N. K., Tipdecho, T., Thongbu, T., Kerdthong, P., and Souris, M. (2012). Analyzing the spatio-temporal relationship between dengue vector larval density and land-use using factor analysis and spatial ring mapping. *BMC Publ. Health* 12:853. doi: 10.1186/1471-2458-12-853

Shope, R. E., and Meegan, J. M. (1997). *Arboviruses in Viral Infections of Humans*. Boston, MA: Springer, 151–183. doi: 10.1007/978-1-4899-0036-4_6

Sohpal, V. K., Dey, A., and Singh, A. (2010). MEGA biocentric software for sequence and phylogenetic analysis: a review. *Int. J. Bioinform. Res. Appl.* 6, 230–240. doi: 10.1504/ijbra.2010.034072

Sumruayphol, S., Apiwathnasorn, C., Ruangsittichai, J., Sriwichai, P., Attrapadung, S., Samung, Y., et al. (2016). DNA barcoding and wing morphometrics to distinguish three Aedes vectors in Thailand. *Acta Trop.* 159, 1–10. doi: 10.1016/j.actatropica.2016.03.010

Turell, M. J., Beaman, J. R., and Tammariello, R. F. (1992). Susceptibility of selected strains of *Aedes aegypti* and *Aedes albopictus* (Diptera: Culicidae) to chikungunya virus. *J. Med. Entomol.* 29, 49–53. doi: 10.1093/jmedent/29.1.49

Vega-Rúa, A., Zouache, K., Girod, R., Failloux, A. B., and Lourenço-de-Oliveira, R. (2014). High level of vector competence of Aedes aegypti and Aedes albopictus from ten American countries as a crucial factor in the spread of Chikungunya virus. *J. Virol.* 88, 6294–6306. doi: 10.1128/jvi.00370-14

Velandia-Romero, M. L., Olano, V. A., Coronel-Ruiz, C., Cabezas, L., Calderón-Peláez, M. A., Castellanos, J. E., et al. (2017). Detección del virus del dengue en larvas y pupas de Aedes aegypti recolectadas en áreas rurales del municipio de Anapoima, Cundinamarca, Colombia. *Biomédica* 37, 193–200. doi: 10.7705/biomedica.v37i0.3584

Villar, L. A., Rojas, D. P., Besada-Lombana, S., and Sarti, E. (2015). Epidemiological trends of dengue disease in Colombia (2000-2011): a systematic review. *PLoS Negl. Trop. Dis.* 9: e0003499. doi: 10.1371/journal.pntd.0003499

Walter, SD, Hildreth, SWY and Beaty, BJ. (1980). Estimación de las tasas de infección en poblaciones de organismos utilizando grupos de tamaño variable. *Am. J. Epidemiol.* 112, 124–128. doi: 10.1093/oxfordjournals.aje.a112961

Whiteman, A., Gomez, C., Rovira, J., Chen, G., McMillan, W. O., and Loaiza, J. (2019). Aedes mosquito infestation in socioeconomically contrasting neighborhoods of panama city. *EcoHealth* 16, 210–221. doi: 10.1007/s10393-019-01417-3

World Health Organization [WHO] (2020). *Real-Time RT-PCR (TaqManTM) protocol – Mayaro virus (MAYV)*. Geneve: WHO.

Zamora-Delgado, J., Castaño, J. C., and Hoyos-López, R. (2015). DNA barcode sequences used to identify Aedes (Stegomyia) albopictus (Diptera: Culicidae) in La Tebaida (Quindío, Colombia). *Rev. Colomb. Entomol.* 41, 212-217

Zhu, G., Liu, T., Xiao, J., Zhang, B., Song, T., Zhang, Y., et al. (2019). Effects of human mobility, temperature and mosquito control on the spatiotemporal transmission of dengue. *Sci. Total Environ.* 651, 969–978. doi: 10.1016/j.scitotenv.2018.09.182

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 19 February 2021
doi: 10.3389/fevo.2021.626757

Influence of Larval Habitat Environmental Characteristics on Culicidae Immature Abundance and Body Size of Adult *Aedes aegypti*

Mariana Rocha David[1†], Edson Santos Dantas[1†], Rafael Maciel-de-Freitas[1], Cláudia Torres Codeço[2], Alex Enrich Prast[3,4] and Ricardo Lourenço-de-Oliveira[1]*

[1] Laboratório de Mosquitos Transmissores de Hematozoários, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil, [2] Programa de Computação Científica, Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil, [3] Department of Thematic Studies—Environmental Change, Linköping University, Linköping, Sweden, [4] Multiuser Unit of Environmental Analysis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

OPEN ACCESS

Edited by:
Andre Barretto Bruno Wilke, University of Miami, United States

Reviewed by:
Shannon L. LaDeau, Cary Institute of Ecosystem Studies, United States
Gabriel L. Hamer, Texas A&M University, United States

***Correspondence:**
Mariana Rocha David
maridavid@ioc.fiocruz.br

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Behavioral and Evolutionary Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *06 November 2020*
Accepted: *28 January 2021*
Published: *19 February 2021*

Citation:
David MR, Dantas ES, Maciel-de-Freitas R, Codeço CT, Prast AE and Lourenço-de-Oliveira R (2021) Influence of Larval Habitat Environmental Characteristics on Culicidae Immature Abundance and Body Size of Adult Aedes aegypti. Front. Ecol. Evol. 9:626757. doi: 10.3389/fevo.2021.626757

Aedes aegypti is adapted to live in close association with human dwellings, where it lays eggs in several man-made container types with a broad range of size, shape, and material. Biotic and abiotic conditions of larval habitats determine the abundance and body size of emerging adult mosquitoes. Here, we estimated the predictive potential of physicochemical water variables for Culicidae immature abundance and *Ae. aegypti* adult body size in four neighborhoods with distinct urban landscapes in Rio de Janeiro, Brazil. Domestic water holding containers (N = 240) were inspected for the presence of Culicidae immatures and had several physiochemical parameters measured. Larvae and pupae were counted, and pupae were reared to the adult stage for taxonomic identification. Dry weight and wing size were measured for *Ae. aegypti* adult mosquitoes (N = 981). The association between larval habitat parameters with Culicidae abundance and *Ae. aegypti* body size data was estimated through linear mixed models and generalized linear mixed models, respectively, with the neighborhood as random effect. The abundance of immature Culicidae in larval habitats (from which >90% of adults emerging from field collected pupae were *Ae. aegypti*) was positively associated with container volume and the dissolved organic carbon concentration (DOC). Female average dry weight and male and female wing lengths were positively associated with larval habitat temperature whereas male average dry weight was positively related to water conductivity. *Aedes aegypti* originating from larval habitats with *Ae. albopictus* exhibited no differences in median wing length and dry body weight when compared with specimens collected in containers exclusively colonized by *Ae. aegypti*. These results demonstrate that container water volume (characteristic easily observed in the field) and DOC (often higher in unmanaged water holding recipients) is related to higher *Ae. aegypti* immature density. Estimating the effects of physicochemical water variables on immature abundance and adult body size can provide valuable information for predicting arbovirus transmission risk in endemic settings.

Keywords: *Aedes aegypti*, larval habitat, larval density, body size, field ecology

INTRODUCTION

The incidence of arboviral diseases has increased substantially in the last few decades, with up to half of the world population currently under the risk of infection by dengue, Zika and chikungunya viruses (Bhatt et al., 2013; Messina et al., 2016; Nsoesie et al., 2016). The spread of mosquito-borne diseases to urban settlements in tropical and subtropical regions worldwide is associated with the geographic expansion of the main vectors. In the Americas, the primary vector of the three aforementioned viruses is the container-breeding mosquito *Aedes aegypti*.

Aedes aegypti is adapted to live in close association with human dwellings and its geographic distribution is constrained by environmental conditions, especially temperature (Kraemer et al., 2015). Female mosquitoes preferentially blood feed on human hosts (Harrington et al., 2001), rest inside premises and lay their eggs mostly in man-made containers located in peridomestic areas (Clements, 1992). Several container types with a broad range of size, shape and material usually receive *Ae. aegypti* eggs, including tires, plant pots, plastic pots, drains, swimming pools and water tanks (Christophers, 1960; Lourenço-de-Oliveira et al., 2004; Maciel-de-Freitas et al., 2007b). The frequency in which *Ae. aegypti* colonizes each container type in a given area is influenced by manifold socio-economic and behavioral factors.

The quality of larval habitats is determined by several factors including the nutritional resources input (e.g., organic matter, microorganisms and algae), larval density (i.e., competition), presence of predators and water physicochemical properties (Clements, 1992; Juliano, 2009). *Aedes aegypti* exhibits a decrease in larval development time and body size as temperature increases but there is a reduction in survival when the larvae are kept at extreme temperatures, as 15°C or 35°C (Rueda et al., 1990; Tun-Lin et al., 2000). Other abiotic factors, such as salinity and pH, may also impact larval survival and growth rate, hindering the presence of this mosquito species in salted water or habitats with extreme pH values (Peryassú, 1908; Clark et al., 2004a,b; Medeiros-Sousa et al., 2020). Similarly, food availability or its limitation by intraspecific or interspecific competition has significant effects on larval development time and adult emergence (Yee et al., 2006).

Food availability and competition in domestic water holding containers are considered major determinants of adult *Ae. aegypti* development and body size. Physical and chemical characteristics of larval habitats have also been correlated with the occurrence and abundance of *Ae. aegypti*, including water volume, temperature, pH and salinity (Strickman and Kittayapong, 2003; Schneider et al., 2004; Barrera et al., 2006; Medeiros-Sousa et al., 2020). Among the traits affected by the water quality, some are directly related with mosquito vectorial capacity such as survival, body size and biting behavior, and thus have a direct effect on disease transmission intensity (Garrett-Jones, 1964; Nasci, 1991; Tun-Lin et al., 2000; Cohuet et al., 2010; Sasmita et al., 2019). Increasing temperatures of larval habitats are associated with shortening larval development time, which reflects on a faster adult emergence (Tun-Lin et al., 2000). The likelihood of dengue infection in mosquitoes increased with size, i.e., *Ae. aegypti* that experienced a less competitive larval habitat (Juliano et al., 2014). On the other hand, poor nutrition and/or competition during larval development resulted in reduced adult body size, longevity and fecundity (Briegel, 1990; Tun-Lin et al., 2000; Maciel-De-Freitas et al., 2007a; Reiskin and Lounibos, 2009), but were associated with higher susceptibility to dengue virus under laboratory conditions (Alto et al., 2008). In addition, smaller females usually blood-fed more frequently than larger females (Scott et al., 2000; Farjana and Tuno, 2013) as they might emerge with insufficient energy reserves to complete oogenesis from a single bloodmeal (Macdonald, 1956). Such behavior increases the contact rate with hosts and as a consequence the risk of pathogen transmission. Identifying container variables correlated with better habitat quality for *Ae. aegypti* in urban settings has the potential to direct control approaches toward the elimination of larval habitats that potentially produce high numbers of large sized adults with presumably higher survival and fecundity rates (Chadee and Focks, 1997).

In this context, we conducted an exploratory survey to estimate the predictive potential of physicochemical water variables for immature Culicidae abundance and *Ae. aegypti* adult body size in four neighborhoods with distinct urban landscapes in Rio de Janeiro. We also addressed the effects of the co-occurrence with *Ae. albopictus* on the weight and wing size of *Ae. aegypti*.

MATERIALS AND METHODS

Study Areas

Rio de Janeiro city, Brazil, is a dengue and chikungunya endemic city with complex, heterogeneous and disordered urban structure, in which high-income areas and slums are in vicinity to each other. In this scenario, *Ae. aegypti* larval habitats have been shown to vary according to urban characteristics such as the availability of piped water, human density and housing type (Maciel-de-Freitas et al., 2007b; David et al., 2009). Field collections were conducted in four neighborhoods: Curicica (CUR), Prainha (PRA), Tubiacanga (TUB) and Vila Valqueire (VILA), representing the range of socio-demographic and economic conditions across Rio de Janeiro (**Table 1**). Therefore, it is expected that the containers sampled are representative of *Ae. aegypti* larval habitats in Rio. At each site, agents from the Rio de Janeiro Department of Health randomly visited and inspected the houses until they found approximately 10 negatives and 50 positives water-holding containers for *Ae. aegypti* immature forms (larvae and/or pupae). Dengue and chikungunya epidemic season in Rio usually goes from December-April. Data were collected April–August 2010.

Physicochemical Characterization of Mosquito Larval Habitats and Water/Sediment Analysis

Containers were classified according to type, location (indoors or outdoors) and the following parameters were registered *in loco*: dissolved O_2, water temperature (oxymeter SevenGo

TABLE 1 | Field site characteristics.

Characteristic	Field site			
	CUR	PRA	TUB	VILA
Geographic location	22°56'33″S, 43°22'52″O	22°48'00″S, 43°19'31″O	22°47'08″S, 43°13'36″O	22°53'21″'S, 43°22'21″O
Surrounding environment	Urban area	Highway and urban area	Guanabara Bay	Secondary forest
Socio-demographic description	Low-middle class urban	Slum (favela)	Low-middle class urban	Middle class urban
Paved streets	Yes	No	Partially	Yes
Piped water	Regular	Irregular	Irregular	Regular
Housing type	Standard (2 bedrooms)	Small (1 bedroom)	Standard (2 bedrooms)	Large (3–4 bedrooms)
Garbage collection	Yes	Yes	Yes	Yes
Human density (inhabitants/ha)	126	132	178	72
Housing density (houses/ha)	658.8	1,578.9	578	638.5
Field site area (ha)	17.3	1.9	1.5	19.2
N visited houses*	100–120	80	35	120–150
Sampling period	Apr/2010	Jun–Jul/2010	Jul–Aug/2010	Aug/2010

*CUR, Curicica; PRA, Prainha; TUB, Tubiacanga; VILA, Vila Valqueire. *Estimated total number of visited houses necessary to find 60 larval habitats.*

Pro, Mettler Toledo, Ohio, EUA), pH (pHmeter model 826 pH Mobile, Metrohm, São Paulo, Brazil), conductivity (codutivimeter SevenGo, Mettler Toledo, Ohio, EUA), air temperature, wind speed, light incidence (CA810 Luxmeter, Chauvin Arnoux, Paris, France) and the presence of insect predators (e.g., *Toxorhynchites* and Odonata immature stages). Water volume was calculated from the length, width and height of the water column in large containers (>10 L) or directly measured by transferring it to a graduated beaker. The surface of larval habitats was calculated using container size measurements.

A sample of the water and sediment from each container was collected using an algae sampler, which consisted of a kitchen sink plunger with a brush inside connected to an plastic syringe for sample collection (Loeb, 1981). The sediment was sampled by pressing the plunger firmly to the substrate, which was scrubbed with the brush. The contents of the plunger cup were then collected with a syringe. Two sizes of algae samplers were used according to containers surface: one using an 20ml syringe (7 cm^2 of sampling area) for containers with <50cm^2 and one with an 80ml syringe (44.2 cm^2 of sampling area) for containers >50cm^2 of surface area. The water/sediment samples from containers were transferred to plastic bags, properly identified and transported in a cool box to the laboratory, where they remained frozen (–20°C) for about 2 months.

Unfiltered frozen samples (10 ml) of water/sediment were processed for Total Phosphorus via orthophosphate determination in a spectrophotometer according to Valderrama (1981) and Total Nitrogen through nitrate detection with ion chromatography with UV detection (Metrohm, São Paulo, Brazil).

Water samples were filtered with GF/F 47 mm Whatman filters, preserved with 15 μL of phosphoric acid and kept in the fridge at 4°C for no longer than 10 days. DOC was determined by an elemental carbon analyzer (model HiperToc, Thermo Scientific, Waltham, EUA) for oxidation with sodium persulfate and determination with UV light. Samples were analyzed in analytical triplicate.

The sediment was analyzed for total organic carbon (sediment organic carbon, SOC) levels after drying in porcelain crucibles in an oven with forced air circulation for 30 min at 60°C. The remaining dry material was weighed, and the organic matter was dosed by calcination in a muffle furnace at 400°C for 12 h.

Mosquito Collection, Taxonomic Identification, and Body Measurement

All larvae and pupae were collected with a strainer and plastic pipette, transferred to a plastic bags containing water from the

TABLE 2 | Water container categories.

Category name	Description	Container types
Domestic water storage	Containers used for water storage for domestic use (e.g., cooking, dishwashing, cleaning and laundry)	Water tanks, metal drums, tanks, cisterns
Non-fixed	Decorative/religious items, drinking fountains	Vases, plant dishes, ice containers, wooden and fiberglass boxes, drinking fountains, small ornamental fountains, boats, religious/ritual objects, trash cans
Fixed	Fixed containers used for water storage for non-domestic use, fixed decorative items, building structures that eventually accumulate rainwater	Tanks in tire shops and gardens, gutters, drains, disused toilets, bathtubs, sewer, untreated pools, awnings, pipes, puddles
Removable	Abandoned containers, disposable materials and garbage	Plastic containers, buckets, bowls, bottles, cans, tires, pans, debris, car parts, plastic bags, plastic covers
Natural	Containers associated to plants, rocks and trees	Bromeliad, rock and tree holes, animal remains (hooves and shells)

larval site and stored in a cool box. For ethical issues, all the remaining water was discarded from containers. When larval site elimination was not possible (e.g., water for domestic use), the water holding container was covered or treated with larvicide by local health agents.

Samples were transported to the laboratory where immature mosquitoes were counted, and the pupae reared to adults in dechlorinated tap water at 28 ± 2°C and 70 ± 10% relative humidity. After taxonomic identification (Consoli and de Oliveira, 1994), adult *Ae. aegypti* mosquitoes were killed with ethyl acetate and dried in a heat oven at 50°C by 15 h. Subsequently, specimens were weighed in a precision scale (Denver Instrument, New York, EUA) and the wing length measured by taking the distance from the axillary incision to the apical margin excluding the fringe (Harbach and Knight, 1980).

Data Analysis

Statistical analysis was conducted with R environment (R Development Core Team, 2011). Container types were classified into five categories according to the criteria used by the Brazilian Ministry of Health (Ministério da Saúde, 2009; **Table 1**). A list of all positive and negative containers types sampled per neighborhood is available in **Supplementary Tables S1, S2**. Exploratory analyses were first performed to identify environmental differences between neighborhoods, as well as variations in physicochemical characteristics according to larval habitat categories.

Principal Component Analysis (PCA) and Random Forest (RF) classification (Liaw and Wiener, 2002) were employed to test whether larval habitats exhibited specific physicochemical profiles according to container category or geographical origin. Differences between positive and negative larval habitats for *Ae. aegypti* were also graphically evaluated through PCA. RF classifying models were trained with a random sample of ∼70% of the data (N = 139 larval habitats) and then validated with the remaining 30% (N = 60 larval habitats) using the "caret" R package. The overall model accuracy and the Kappa index, as well as the identification of most influencing parameters for classification, were obtained after applying the RF models to the validation dataset. PCA was applied on the correlation matrix after scaling variables to have unit variance since they were on different scales and therefore exhibit different variances. In this case, variables were standardized to mean equal to zero and variance equal to one. The environmental parameters "air temperature," "container surface," and "location (indoors or outdoors)" were removed from data analysis due to strong correlation ($r > 0.7$) with "water temperature," "water volume," and "light incidence," respectively. The presence of predators was excluded from the analyses since only a single predatory larva of *Toxorhynchites* sp. was observed in one bromeliad from VILA. The environmental variables considered in RF and PCA analyses are listed in **Table 3**.

In order to elucidate possible associations between larval habitats characteristics and mosquito abundances, the total count of immature per recipient was included as dependent variable in generalized linear mixed models (GLMMs) using the "lme4" R package (Bates et al., 2015). We considered the total number of Culicidae immature forms collected, since only the pupae were reared to the adult stage for taxonomic identification, and also positive and negative containers for *Ae. aegypti*. The independent variables considered in the GLMMs are listed in **Table 3**. In order to improve the quality of model fit, the variables "volume," "light incidence," "conductivity," "DOC," and "SOC" were included in the models at the log scale. The count of Culicidae immature per larval habitat was analyzed with a GLMM with negative binomial distribution with the neighborhood as random effect in order to incorporate the variation between collection sites in the models. A regression with negative binomial distribution was preferred over the traditional Poisson distribution because data exhibited over-dispersion (i.e., variance was larger than the mean), confirmed by the Pearson Chi^2 test and Dispersion Statistic >1, calculated using the "msme" R package (Hilbe and Robinson, 2013). The consistency of data with the negative binomial distribution was verified using the goodness of fit test "Minimum Chi-squared" (Pearson χ^2 = 240.7, df = 233, p = 0.35) from the "goodfit" command, implemented in the "vcd" R package (Meyer et al., 2020). After adjusting independent variables to univariate models, those with significant effects (α = 0.1) on immature Culicidae per larval habitat were used to build GLMMs. The most informative and parcimonious model was selected through second-order Akaike's information criterion scores (AICc) and level of support (AICcWt), calculated using the "AICcmodavg" package (Mazerolle, 2019). Collinearity between independent variables was checked in the best model through VIFs (Variance Inflation Factors) (Zuur et al., 2010). The assumptions of the best model were examined by checking heteroscedasticity, residuals dispersion and the presence of outliers with the R package DHARMa (Hartig, 2020). The conditional Pseudo-R-squared for GLMMs (R^2c), calculated using the MuMIn R package (Bartón, 2014), was employed as goodness-of-fit metric and a measurement of variance explained by the entire model, including both fixed and random effects.

TABLE 3 | Variables tested in multivariate analyses.

Dependent/ response variable	Analyses	Independent/ explanatory variables tested
Neighborhood or larval habitat category	Random Forest Classification and Principal Component Analysis	Dissolved O_2 (ppm), water temperature (°C), pH, conductivity (μS/cm), wind speed (m/s)*, light incidence (lx), water volume (ml), Total nitrogen (mg/L), Total phosphorous (mg/L), dissolved organic carbon (DOC, mg/L), sediment organic carbon (SOC, g/kg) and number of immature Culicidae**
Culicidae immature per larval habitat	Generalized linear mixed models (negative binomial distribution)	
Female and male body size (average weight and wing length per larval habitat)	Linear mixed models	

*Removed from body size analyses.
**Only included in body size analyses.

Regarding the body size analyses, *Ae. aegypti* average wing length and dry weight were calculated for each larval habitat according to mosquito sex. Since there was a significant, but moderate correlation between these two variables (Pearson's correlation female: $t = 11.7$, $p < 0.001$, $r = 0.69$, $r^2 = 0.48$; male: $t = 10.4$, $p < 0.001$, $r = 0.60$, $r^2 = 0.36$; **Supplementary Figure S1**), both were fitted as dependent variables at the Linear Mixed Models (LLMs) with the neighborhood as random effect (**Table 3**). As data was considered normally distributed, a gaussian distribution was applied (One-sample Kolmogorov-Smirnov test $p > 0.05$). The independent variable "wind speed" was removed from these analyses once aquatic mosquito larvae and pupae are submerged, therefore would not be directly exposed to variations in this environmental condition. Model selection and validation followed the previously mentioned procedures.

The dry weight and wing size were compared between *Ae. aegypti* originating from larval habitats with and without *Ae. albopictus* (or any other mosquito species). Differences in mosquito body size were tested between groups using the Kruskal-Wallis test as male wing size was not considered normally distributed (One-sample Kolmogorov-Smirnov test $p < 0.05$).

Ethical Considerations

Larval surveys are routinely performed by health agents from the Rio de Janeiro City Health Department (Secretaria Municipal de Saúde do Rio de Janeiro, SMS-RJ). Access to containers inside dwellings were obtained after consent of householder approving the entrance of the research team and the health agent who routinely worked at that neighborhood.

RESULTS

Larval Habitats and Mosquito Collection

In total, 240 water holding containers were selected at the four field sites, of which 204 were positive for *Ae. aegypti* immature presence. The number of sampled containers according to category and study site are shown in **Table 4** and a total of 3,184 mosquito immatures were collected, of which 1,168 were pupae. Pupae survival to the adult stage allowed the taxonomic identification of 1,064 mosquitoes, 692 males and 372 females. Although most of adult mosquitoes were identified as *Ae. aegypti* (**Table 5**), *Ae. albopictus*, *Aedes (Ochlerotatus)* sp., *Culex* sp., and *Wyeomyia* sp. were also identified in the inspected containers (**Table 5**). Of the 981 adult *Ae. aegypti*, 638 were males and 343 were females which provided the measure of body weight and wing length as a proxy for body size.

TABLE 4 | Distribution of water holding containers according to category in the four field sites.

Container category	CUR (%)	PRA (%)	TUB (%)	VILA (%)	Total (%)
Domestic water storage	7 (11.7)	21 (35)	14 (23.3)	6 (10)	48 (20)
Non-fixed	4 (6.7)	6 (10.0)	5 (8.3)	8 (13.3)	23 (9.6)
Fixed	14 (23.3)	8 (13.3)	7 (11.7)	22 (36.7)	51 (21.2)
Removable	33 (55.0)	25 (41.7)	34 (56.7)	10 (16.7)	102 (42.5)
Natural	2 (3.3)	0 (0.0)	0 (0.0)	14 (23.3)	16 (6.7)
Total	60 (100)	60 (100)	60 (100)	60 (100)	240 (100)

Physicochemical Parameters of Water and Sediment According to Geographical Origin and Between Container Types

The accuracy and Kappa index of RF models fitted to classify larval habitats using physicochemical parameters of water and sediment were higher according to geographical location than container category (**Table 6**). Overall model accuracy for neighborhood was 78.3%, with Classification sensitivity (true positive rate) varying between 63 (VILA) to 92% (TUB). Water temperature, volume and light incidence were the most important predictors for sample discrimination into neighborhood of collection. Median water temperature in TUB containers was 25.7°C meanwhile it ranged between 21.4 and 21.6°C in the other field sites. Container volume tended to be higher in TUB and contrasted to the median volume of larval habitats sampled elsewhere (TUB: median of 7.2 L; other sites: median of 0.3–0.6 L). On the other hand, median light incidence was 63.5 lx in TUB larval habitats while it ranged between 113 and 154.1 lx in the other sampled neighborhoods. A less efficient classification was observed for container categories, with the container volume as the most important predictor and overall accuracy of 50%. Raw larval habitat data can be found in **Supplementary Table S3**.

According to the PCA, the first two axes explained 27.3% of total variation. The first axis accounted for 13.8% of data variance, with DOC, light incidence and water temperature as the variables that most contributed to explain the dataset variation (**Supplementary Figure S2A**). Dissolved O2, SOC and water temperature were the main variables explaining the second component (accounting for 13.5% of variance) (**Supplementary Figure S2B**). TUB larval habitat formed a more heterogeneous cluster, compared to CUR, PRA and VILA (**Figure 1A**). No discrimination was noted between container types (**Figure 1B**) nor between positive and negative *Ae. aegypti* larval habitats (**Figure 1C**), corroborating RF classification results. Considering the higher heterogeneity between TUB larval habitats in comparison to the other field sites, GLMMs and LMMs for abundance and body size investigations were adjusted with neighborhood as a random effect.

Larval Habitat Characteristics and Immature Abundance

Univariate GLMM analyses indicated a significant association of immature Culicidae abundance in larval habitats with volume, total nitrogen, total phosphorous and DOC concentrations (**Supplementary Table S4**). Subsequently, GLMMs were fitted using these four independent variables and compared to the null model. The most informative GLMM had volume, total nitrogen,

TABLE 5 | Taxonomic identification of adult mosquitoes collected during the pupal stage in larval habitat from four neighborhoods in Rio de Janeiro, Brazil.

	Aedes aegypti (%)*	*Aedes albopictus* (%)*	*Aedes (Ochlerotatus)* sp. (%)*	*Culex* sp. (%)*	*Wyeomyia* sp. (%)*	Total
CUR	170 (93.4)	7 (3.9)	0	5 (2.7)	0	182 (100)
PRA	263 (96)	6 (2.4)	0	4 (1.6)	0	273 (100)
TUB	306 (90)	13 (3.8)	8 (2.4)	13 (3.8)	0	340 (100)
VILA	243 (91.1)	9 (2.1)	0	14 (4.8)	3 (1)	269 (100)
Total	982 (92.3)	35 (3.3)	8 (0.7)	36 (3.4)	3 (0.3)	1,064 (100)

**% relative to the total number of mosquitoes found in the respective neighborhood.*

total phosphorous and DOC as independent variables, but only volume and DOC exhibited significant effects on the number of immature Culicidae. Model R^2c was 0.37. These results suggest that the number of immature Culicidae increased with larval habitat volume and DOC concentration (**Table 7**). For example, TUB, the neighborhood in which container volume tended to be higher, exhibited 31.3 immature per larval habitat in average, while it varied between 4.7 and 8.8 in the other field sites. Descriptive analysis of immature abundance data can be found in **Supplementary Table S5**.

Larval Habitat Characteristics and *Ae. aegypti* Body Size

Female *Ae. aegypti* originated from 126 larval habitats. Regarding average dry weight and wing length per container, the univariate LMMs presented only water temperature as significant independent variable, with R^2c of 0.15 and 0.06, respectively. The models output indicates that both female weight and wing length increased with water temperature (**Table 8**). Additionally, neighborhood variance estimates were near zero, suggesting there was virtually no among-neighborhood variance left to be explained by the random intercepts (**Table 8**). Indeed, in TUB, the neighborhood with higher water temperatures, female *Ae. aegypti* had slightly larger wings, with 2.56 mm in average, while it ranged between 2.39 and 2.51 in the other field sites. However, the same pattern could not be clearly seen for the dry weight, as PRA and TUB mosquitoes had in average 0.32 mg, while it was 0.38–0.39 mg in average in VILA and CUR, respectively (**Supplementary Table S5**).

Male *Ae. aegypti* originated from 160 larval habitats. The univariate LLMs indicated a significant association between average male dry weight per container with dissolved O2, conductivity, phosphorous and DOC (**Supplementary Table S6**). Therefore, GLMMs were fitted using these four independent variables and were compared to the null model. The full model was the one with the highest support, with male dry weight increasing with water conductivity (**Table 8**), which was in median 322.50 μS/cm (ranging from 2.25 to 1,374 μS/cm), with no variation between neighborhoods or container categories, as shown by RF and PCA analysis. The R^2c was 0.09 for this LLM. For male wing length, only water temperature had a significant effect, suggesting an increase of wing length with temperature. As seen for female *Ae. aegypti,* male tended to have larger wings in TUB (2.06 mm in average against 1.90–1.98 mm in average in the other neighborhoods, **Supplementary Table S5**). However, the R^2c was quite low (0.06), as shown for the other LMMs fitted for body size measurements. Neighborhood variance estimates were near zero, suggesting there was virtually no among-neighborhood variance left to be explained by the random intercepts (**Table 8**).

TABLE 6 | Random forest classification of larval habitats into container types or neighborhoods using physicochemical characteristics of water and sediment as predictors.

Response variable	Overall accuracy (%)	Accuracy 95% C.I. (%)*	Kappa index	Most important predictors
Neighborhood	78.3	65.8–87.9	0.71	Water temperature, volume and light incidence
Container category	53.3	40–66.3	0.29	Water volume, light incidence, dissolved oxygen

** 95% confidence interval (C.I.).*

Co-occurrence With *Ae. albopictus* and *Ae. aegypti* Body Size

The two species were simultaneously observed in 12 out of 204 larval habitats (5.9%), six from TUB and in two containers from each of the other field sites. A total of 30 male and 22 female *Ae. aegypti* mosquitoes were collected from shared larval habitats with *Ae. albopictus*, while 545 males and 301 *Ae. aegypti* females originated from larval habitats where this species occurred exclusively. Due to disparities higher than 10-fold between the number of *Ae. aegypti* originating from shared and non-shared containers, a random sample of 30 female and 30 male mosquitoes was taken from the second group in order to balance sample sizes for statistical comparison. Median dry weight and wing length tended to be lower for both male and female when co-occurring with *Ae. albopictus*, although with no statistical significance (**Figures 2A–D**: Female wing length: KW $\chi^2 = 1.88$, df = 1, $p = 0.17$; Female dry weight: KW $\chi^2 = 1.93$, df = 1, $p = 0.16$; Male wing length: KW $\chi^2 = 1.99$, df = 1, $p = 0.16$; Male dry weight: KW $\chi^2 = 2.76$, df = 1, $p = 0.10$).

DISCUSSION

In the present study, we addressed the association between physicochemical properties of larval habitats and Culicidae immature abundance and *Ae. aegypti* adult body size. We showed that domestic containers from four distinct urban landscapes in Rio de Janeiro were massively colonized by

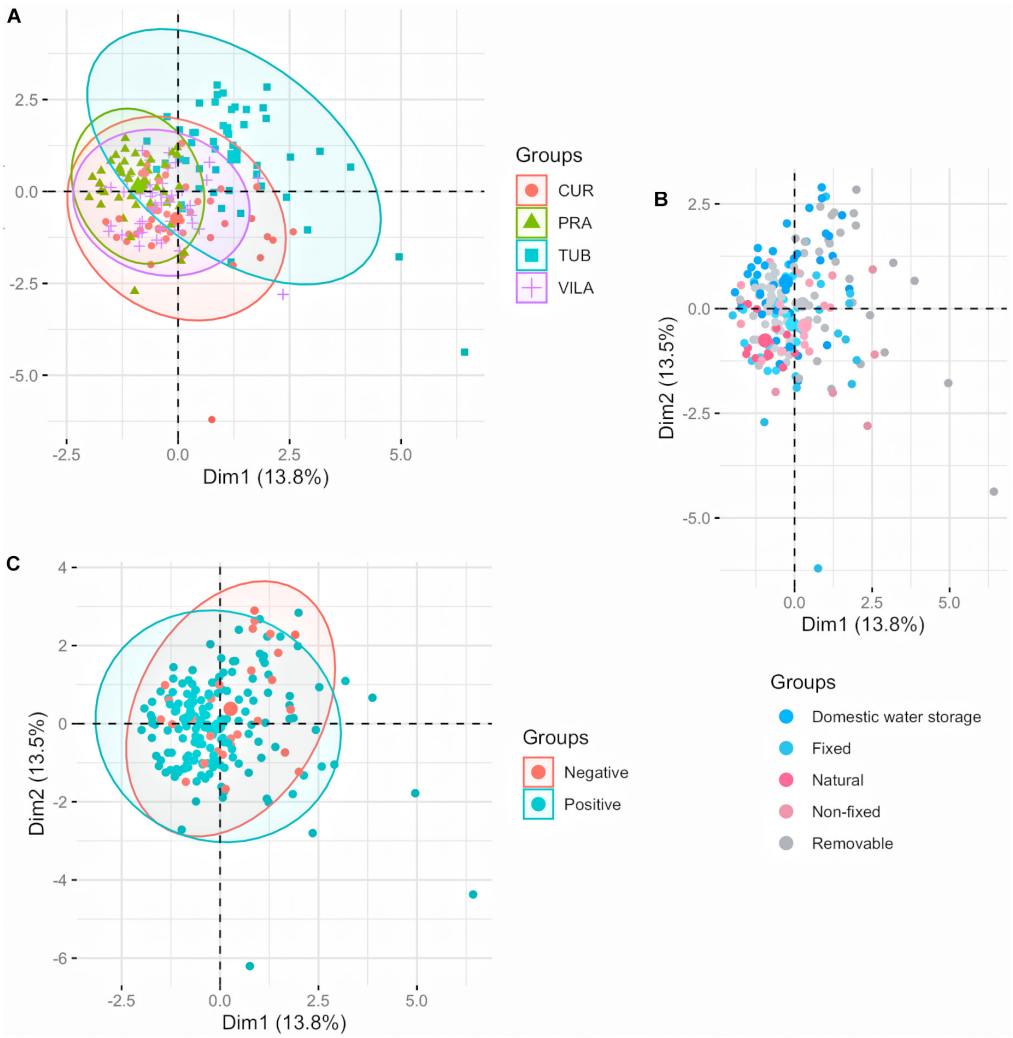

FIGURE 1 | Principal component analysis of larval habitats. Each point corresponds to a larval habitat, colored according to the field site of collection **(A)**, container category **(B)** and positive or negative for *Ae. aegypti* presence **(C)**. Overlapped points are represented by bigger circles. Container categories are described in **Table 2**.

Ae. aegypti, which accounted for >90% of adults emerging from field collected pupae. Moreover, the abundance of immature Culicidae was positively associated with container volume and the DOC concentration. Although the habitat parameters measured explained a small fraction (≤15%) of the data variation for *Ae. aegypti* body size, female average dry weight and male and female wing lengths were positively associated with larval habitat temperature whereas male average dry weight was positively related to water conductivity. Lastly, *Ae. aegypti* originating from larval habitats shared with *Ae. albopictus* exhibited no differences in median wing length and dry body weight when compared to specimens collected in containers exclusively colonized by *Ae. aegypti*.

Female *Ae. aegypti* has opportunistic egg-laying behavior, i.e., lay eggs in water-holding containers of varied characteristics. Here, we demonstrate that through RF classification and PCA that larval habitats differed in some physiochemical characteristics according to the geographical origin, but not according to container category and the presence of *Ae.*

TABLE 7 | Results of the generalized linear mixed model (negative binomial) of the number of immature Culicidae in larval habitats.

Effects	Term	Estimate	SE	90% CI	*z*-value	*p*-value
Fixed	Intercept	0.46	0.43	–0.30 to 1.23	1.05	0.29
	Log (volume)	0.14	0.03	0.10–0.19	5.34	<0.001
	Nitrogen	0.03	0.02	–0.001 to 0.07	1.58	0.11
	Phosphorous	0.07	0.04	0.002–0.14	1.62	0.10
	Log (DOC)	0.15	0.07	0.03–0.27	2.11	0.03
Random	Neighborhood (variance)	0.40	–	–	–	–

SE, Standard error; CI, confidence interval.

TABLE 8 | Results of the linear mixed models of the average wing length and body weight per larval habitat of female and male *Ae. aegypti.*

Dependent variable	Effects	Term	Estimate	SE	90% CI	df	*t*-value	*p*-value
Female average dry weight	Fixed	Intercept	0.15	0.12	–0.03 to 0.39	61.69	1.30	0.20
		Water temperature	0.008	0.005	–0.001 to 0.02	80.10	1.73	0.09
	Random	Neighborhood (variance)	0.002	–	–	–	–	–
Female average wing length	Fixed	Intercept	2.00	0.23	1.63–2.36	18.18	8.79	<0.001
		Water temperature	0.02	0.009	0.006–0.04	18.24	2.18	0.04
		Neighborhood (variance)	0.001	–	–	–	–	–
Male average dry weight	Fixed	Intercept	0.18	0.02	0.15–0.21	7.47	11.13	<0.001
		Conductivity	<0.001	< 0.001	<0.001	159.3	2.06	0.04
		Dissolved O_2	–0.002	0.001	–0.004–0.00	145.9	–1.64	0.10
		DOC	<0.001	<0.001	<0.001	159.5	1.43	0.15
		Phosphorous			–0.001–0.005			
		Neighborhood (variance)	<0.001	–	–	–	–	–
Male average wing length	Fixed	Intercept	1.62	0.15		20.29	11.06	<0.001
		Water temperature	0.01	0.006		22.27	2.43	0.02
	Random	Neighborhood (variance)	<0.001	–	–	–	–	–

SE, Standard error; CI, confidence interval.

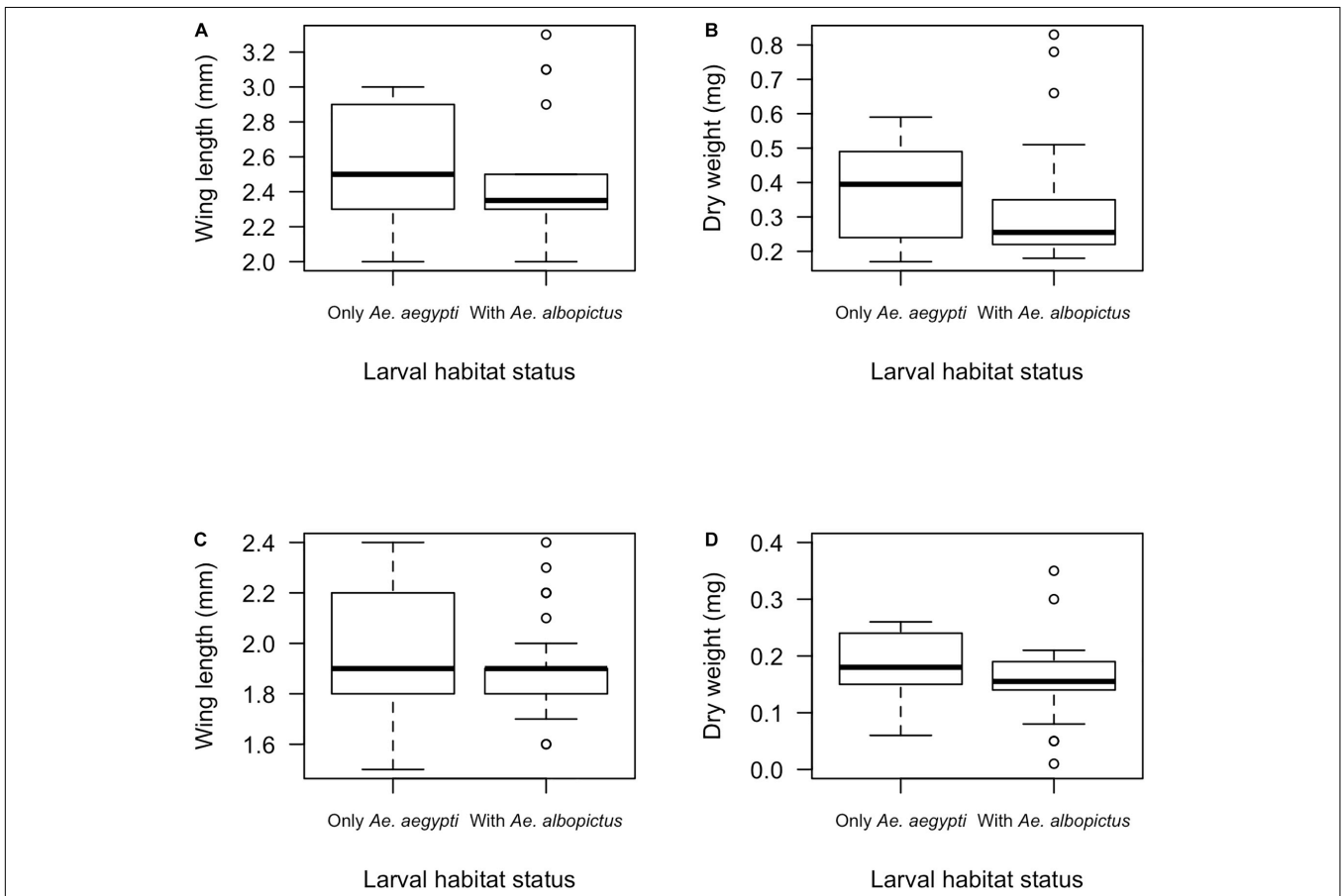

FIGURE 2 | Wing length and dry weight of female **(A,B)** and male *Ae. aegypti* **(C,D)** according the presence of no other species or the co-occurrence with *Ae. albopictus* in the larval habitat. No significant differences in body measurements were observed.

aegypti. Studies conducted in different environments showed a fuzzy association between physicochemical characteristics of larval habitats and container types. However, differences in some parameters such as total dissolved solids, water hardness, electrical conductivity, chemical oxygen demand and conductivity could be linked to container nature (i.e., natural or artificial) and water origin (Chatterjee et al., 2015; Gnanasoundari et al., 2017; Hery et al., 2020).

Potential differences in physiochemical parameters between positive and negative containers for *Ae. aegypti* could indicate oviposition preferences by female mosquitoes, egg hatching rate, immature survival and/or modifications of water following the *Ae. aegypti* colonization. Larvae development for several generations have been linked, for example, to the decrease of water turbidity, which would increase the activity of nitrifying bacteria and consequently the concentration of nitrate and nitrite ions (Darriet and Corbel, 2008). Nevertheless, any differentiation between positive and negative containers for *Ae. aegypti* immature was notice by the PCA. This result might be explained by sampling methods, since we had no control for how many generations these water containers were colonized by *Ae. aegypti,* i.e., whether it was long enough to cause any change in water and sediment parameters.

Containers exhibited differences in water physiochemical profile according to the neighborhood of location, with TUB larval habitats forming a more heterogenous cluster in relation to CUR, PRA, and VILA. Water temperature and volume, as well as the light incidence, were the most important variables distinguishing larval habitats according to its geographical location. This observation is in line with the high environmental temperatures (>35°C) usually registered in TUB in comparison to other locations in Rio de Janeiro (e.g., Dutra et al., 2015; Garcia et al., 2019). Sampling during April (autumn in Rio de Janeiro) might also have contributed to higher water temperatures in TUB comparison to the other field sites, which were sampled between June and August (autumn and winter). Median container volume was also higher in TUB, what might have been influenced by the 23.3% of domestic water storage containers sampled in TUB. Although PRA also had a high prevalence of this container category (35%), TUB houses were bigger than in PRA and as a consequence, residents usually stored water in larger containers (personal observation).

The abundance of Culicidae immature increased with container water volume and dissolved organic carbon (DOC) concentration. The container size has been recognized as an important predictor of the number of eggs laid by *Ae. aegypti* under lab and semi-field conditions, as day active mosquitoes would choose oviposition sites based on visual cues (Harrington et al., 2008; Panigrahi et al., 2014). Moreover, large containers are expected to provide more food (due to larger submerged surface and input of organic matter), as well as be more resistant to desiccation. A positive association between *Ae. aegypti* immature abundance and water volume have also been described in Puerto Rico, Dominican Republic and Bangladesh (Barrera et al., 2006; González et al., 2019; Islam et al., 2019).

Food availability is one of the main attributes that regulate *Ae. aegypti* development (Tun-Lin et al., 2000). Here, the number of immature Culicidae was positively associated to the dissolved organic carbon (DOC) concentration in the larval habitat. A higher availability of organic matter usually results in a higher flow of energy via the food web (Benavides-Gordillo et al., 2019) and could have acted as an oviposition attractant for mosquito female (together with container size) and/or food source enabling increased immature survival. In addition, it might also serve as substrate for bacterial growth (Benavides-Gordillo et al., 2019). In the urban area of Iquitos, Peru, female *Ae. aegypti* laid more eggs in larger unmanaged containers with more organic material, characteristics that were expected to maximize the progeny growth and survival. Nevertheless, those larval habitats produced mosquitoes with lower pupation probability and smaller body size when compared to specimens from smaller unmanaged containers, which received less eggs. Thus, although more attractive for mosquito oviposition, the larger containers were considered more competitive habitats for *Ae. aegypti* (Wong et al., 2012).

Considering the influence of physiochemical parameters on Culicidae immature density, we are aware that our findings should extrapolated carefully to *Ae. aegypti*, as only the pupae were reared to the adult stage for taxonomic identification. However, the sampled container types are well known to be historically colonized mainly by *Ae. aegypti* in Rio de Janeiro (Maciel-de-Freitas et al., 2007b; David et al., 2009), fact that was confirmed by the >90% of pupae identified of being of this species. Since we have no evidence of differential survival between the immature Culicidae species found, we are confident that this high proportion of *Ae. aegypti* found in emerging adults can be extrapolated to immature forms of mosquitoes present in larval habitats.

Unlike the density of immature mosquitoes, whose ∼37% of the data variance could be explained by the GLMM, a smaller fraction (≤15%) of the variation in body size measurements could be related to the addressed physiochemical characteristics of larval habitats. Such finding reinforces the complex and dynamic nature of larval habitat suitability for mosquito development (Hemme et al., 2009). Even so, the wing length of both male and female *Ae. aegypti,* as well as female dry weight, exhibited a positive significant association with water temperature of larval habitats. Water temperature is a determining factor for the development and survival of immature *Ae. aegypti* (Christophers, 1960; Couret and Benedict, 2014). In addition, water temperature also can influence the production and survival of algae, bacteria and fungi, which are part of the diet of mosquito larvae, thus altering food availability in the larval habitat (Hemme et al., 2009). In the laboratory, optimal survival of *Ae. aegypti* occurred between 20 and 30°C, with significant decrease in body size in temperatures >25°C (Tun-Lin et al., 2000). Considering the four field sites, water median temperature was 21.7°C (22.7°C in average), while it was in median 25.7°C in TUB. In all cases, the water temperature may not have been high enough to cause any development acceleration and reduction in adult body size as seen in other lab (Tun-Lin et al., 2000; Mohammed and Chadee, 2011) and field studies such as Barrera et al., 2006 (in which water temperature was 29.3°C in average).

Male average dry weight showed a positive significant association with water conductivity, a commonly used proxy for the amount of dissolved nutrients. The amount of solids dissolved in the water release ions as organic fractions of solids are consumed leading to an increase in conductivity. Thus, it could be interpreted as an indirect measure of food availability for larvae and/or substrate for microorganism growth (Hery et al., 2020). Conductivity has been linked to the

presence of *Ae. aegypti* in artificial containers and to the larval density (Chatterjee et al., 2015; Garcia-Sánchez et al., 2017), but it effects on mosquito development, nutrition and body size are still unknown.

Competition has been associated to *Ae. aegypti* body size reduction (Strickman and Kittayapong, 2003; Barrera et al., 2006; Wong et al., 2012). Surprisingly, there was no evidence for density dependent effects on adult body size in our study, as the total number of immature Culicidae did not show any association with *Ae. aegypti* wing length or dry weight. In addition, a non-significant reduction in mosquito body measurements was noticed after comparing specimens originating from containers shared with *Ae. albopictus* or exclusively colonized by *Ae. aegypti*, what might be carefully interpreted due to the low frequency of co-occurrence registered here. The strength of density-dependent effects on mosquito fitness (e.g., body size) and population dynamics are heterogenous between water containers in the field, as they vary in accordance with larval densities and food availability (Walsh et al., 2013). We hypothesized that the range of immature densities in the sampled larval habitats (0.03 immature/ml in average, with a maximum of 0.66 immature/ml) were below the larval habitats carrying capacity (i.e., larval growth was not constrained by the nutritional resources saturation), thus, density-dependent intraspecific competition effect was not strong enough to produce smaller mosquitoes, as seen by others (Barrera et al., 2006; Walsh et al., 2013).

In summary, this study expands the currently knowledge about the relation between *Ae. aegypti* immature density, adult body size and larval habitat characteristics in the field. Our results indicate that container water volume (an easily observed characteristic in the field) is a predictor of *Ae. aegypti* immature density. Although the contribution of variations in larval habitats in adult populations was not investigated here, our results reinforces the recommendation that eliminating or preventing mosquito access to larger containers have the potential to effectively reduce *Ae. aegypti* field populations. Similarly, DOC concentration (often higher in unmanaged water holding recipients) can also be a predictor of *Ae. aegypti* productivity. Larval habitat characteristics have important implications population dynamics, as well as biology parameters, such as body size, fecundity and survival, which ultimately influence vector capacity and pathogen transmission to human populations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MD analyzed the data and drafted the manuscript. ED conducted field collections and chemical analysis of larval habitat water. RM-d-F contributed to study design and field collections. CC contributed to data analysis. AP contributed to chemical analysis of larval habitat water. RL-d-O conceived the study. All authors edited the manuscript and approved the submitted version.

FUNDING

This work was supported by the US National Institutes of Health Fogarty International Center FIRCA (Grant No. R03TW007446-01A1), Conselho Nacional de Desenvolvimento Científico e Tecnológico (Grant No. 12446/2018), Preventing and Combating the Zika Virus MCTIC/FNDCT-CNPq/MEC-CAPES/MS-Decit. (Grant No. 440929/2016-4), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (Grants Nos. E-26/203.064/2016 and E-26/201.335/2016). ED received a scholarship of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—CAPES (Finance Code 001). We are grateful for the support of the Coordination for the Improvement of Higher Education Personnel—CAPES, and Fiocruz (Programa Institucional de Internacionalização da Fiocruz PrInt Fiocruz-CAPES). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank the Rio de Janeiro City Health Department (Secretaria Municipal de Saúde do Rio de Janeiro, SMS-RJ), Reginaldo Rego, Roberto Peres, Renato Carvalho and Marcelo Celestino for contributions to the fieldwork. We also thank L. Philip Lounibos, Márcio G. Pavan, and Steven A. Juliano for insightful comments and suggestions in statistical analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.626757/full#supplementary-material

Supplementary Figure 1 | Linear regressions of mean wing length and mean weight per larval habitat for *Ae. aegypti* female and male mosquitoes.

Supplementary Figure 2 | Contribution of each environmental variable to the first **(A)** and second **(B)** components of PCA. The red dashed line indicates the expected average contribution if the contribution of the variables was uniform, the expected value would be 1/11 variables, i.e., 9.09%.

Supplementary Table 1 | *Aedes aegypti* positive containers sampled per neighborhood.

Supplementary Table 2 | *Aedes aegypti* negative containers sampled per neighborhood.

Supplementary Table 3 | Raw larval habitat data.

Supplementary Table 4 | Univariate generalized linear mixed models with environmental variables to predict the number of immature Culicidae in larval habitats.

Supplementary Table 5 | Descriptive analysis of immature abundance and mosquito body size according to field sites.

Supplementary Table 6 | Univariate linear mixed models with environmental variables to predict the average male *Ae. aegypti* dry weight per larval habitat.

REFERENCES

Alto, B. W., Lounibos, L. P., Mores, C. N., and Reiskind, M. H. (2008). Larval competition alters susceptibility of adult Aedes mosquitoes to dengue infection. *Proc. R. Soc. B Biol. Sci.* 275, 463–471. doi: 10.1098/rspb.2007.1497

Barrera, R., Amador, M., and Clark, G. G. (2006). Ecological Factors Influencing Aedes aegypti (Diptera: Culicidae) Productivity in Artificial Containers in Salinas, Puerto Rico. *J. Med. Entomol.* 43, 484–492. doi: 10.1603/0022-2585200643

Bartón, K. (2014). *MuMIn: Multi-model inference. R Packag. version 1.10.0*.

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 2015:v067. doi: 10.18637/jss.v067.i01

Benavides-Gordillo, S., Farjalla, V. F., González, A. L., and Romero, G. Q. (2019). Changes in rainfall level and litter stoichiometry affect aquatic community and ecosystem processes in bromeliad phytotelmata. *Freshw. Biol.* 64, 1357–1368. doi: 10.1111/fwb.13310

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., et al. (2013). The global distribution and burden of dengue. *Nature* 496, 504–507. doi: 10.1038/nature12060

Briegel, H. (1990). Metabolic relationship between female body size, reserves, and fecundity of Aedes aegypti. *J. Insect Physiol.* 36, 165–172. doi: 10.1016/0022-1910(90)90118-Y

Chadee, D. D., and Focks, D. A. (1997). Pupal Survey: An Epidemiologically Significant Surveillance Method for Aedes aegypti: An Example Using Data from Trinidad. *Am. J. Trop. Med. Hyg.* 56, 159–167. doi: 10.4269/ajtmh.1997.56.159

Chatterjee, S., Chakraborty, A., and Sinha, S. K. (2015). Spatial distribution & physicochemical characterization of the breeding habitats of Aedes aegypti in & around Kolkata, West Bengal, India. *Indian J. Med. Res.* 142, S79–S86. doi: 10.4103/0971-5916.176631

Christophers, S. R. (1960). *A{ë}des Aegypti (L.), the Yellow Fever Mosquito: Its Life History, Bionomics, and Structure*. New York: University Press Available online at: https://books.google.com.br/books?id=rnJkjgEACAAJ.

Clark, T. M., Flis, B. J., and Remold, S. K. (2004a). Differences in the effects of salinity on larval growth and developmental programs of a freshwater and a euryhaline mosquito species (Insecta: Diptera, Culicidae). *J. Exp. Biol.* 207, 2289–2295. doi: 10.1242/jeb.01018

Clark, T. M., Flis, B. J., and Remold, S. K. (2004b). pH tolerances and regulatory abilities of freshwater and euryhaline Aedine mosquito larvae. *J. Exp. Biol.* 207, 2289–2295. doi: 10.1242/jeb.01021

Clements, A. N. (1992). *The Biology of Mosquitoes: Development, nutrition, and reproduction*. London: Chapman and Hall.

Cohuet, A., Harris, C., Robert, V., and Fontenille, D. (2010). Evolutionary forces on Anopheles: what makes a malaria vector? *Trends Parasitol.* 26, 130–136. doi: 10.1016/j.pt.2009.12.001

Consoli, R. A. G. B., and de Oliveira, R. L. (1994). *Principais mosquitos de importância sanitária no Brasil. 1st ed., ed. E. Fiocruz Rio de Janeiro: Editora FIOCRUZ* doi: 10.7476/9788575412909

Couret, J., and Benedict, M. Q. (2014). A meta-analysis of the factors influencing development rate variation in Aedes aegypti (Diptera: Culicidae). *BMC Ecol.* 14:3. doi: 10.1186/1472-6785-14-3

da Saúde, M. (2009). *Diretrizes Nacionais para a Prevenção e Controle de Epidemias de Dengue*. Brasília: Ministério da saúde

Darriet, F., and Corbel, V. (2008). Attractive properties and physicochemical modifications of water following colonization by Aedes aegypti larvae (Diptera: Culicidae). *Compt. Ren. Biol.* 331, 617–622. doi: 10.1016/j.crvi.2008.05.003

David, M. R., Lourenço-de-Oliveira, R., and Freitas, R. M. De (2009). Container productivity, daily survival rates and dispersal of Aedes aegypti mosquitoes in a high income dengue epidemic neighbourhood of Rio de Janeiro: presumed influence of differential urban structure on mosquito biology. *Mem. Inst. Oswaldo Cruz* 104, 927–932.

Dutra, H. L. C., dos Santos, L. M. B., Caragata, E. P., Silva, J. B. L., Villela, D. A. M., Maciel-de-Freitas, R., et al. (2015). From Lab to Field: The Influence of Urban Landscapes on the Invasive Potential of Wolbachia in Brazilian Aedes aegypti Mosquitoes. *PLoS Negl. Trop. Dis.* 9:e0003689. doi: 10.1371/journal.pntd.0003689

Farjana, T., and Tuno, N. (2013). Multiple Blood Feeding and Host-Seeking Behavior in Aedes aegypti and Aedes albopictus (Diptera: Culicidae). *J. Med. Entomol.* 50, 838–846. doi: 10.1603/ME12146

Garcia, G. de A., Sylvestre, G., Aguiar, R., da Costa, G. B., Martins, A. J., Lima, J. B. P., et al. (2019). Matching the genetics of released and local Aedes aegypti populations is critical to assure Wolbachia invasion. *PLoS Negl. Trop. Dis.* 13(1):e0007023. doi: 10.1371/journal.pntd.0007023

Garcia-Sánchez, D. C., Pinilla, G. A., and Quintero, J. (2017). Ecological characterization of Aedes aegypti larval habitats (Diptera: Culicidae) in artificial water containers in Girardot, Colombia. *J. Vector Ecol.* 42, 289–297. doi: 10.1111/jvec.12269

Garrett-Jones, C. (1964). the Human Blood Index of Malaria Vectors in Relation To Epidemiological Assessment. *Bull. World Health Organ.* 30, 241–261.

Gnanasoundari, A, Thangamathi, P, Ananth, S. (2017). Physico-Chemical Characteristics of Breeding Habitats and its Influence in Larval Population Density of Dengue Vector Aedes Aegypti in Thanjavur, Tamilnadu, India. *Int. J. Res. Appl. Nat. Soc. Sci.* 7, 69–76.

González, M. A., Sosa, M. A. R., Bautista, Y. E. V., Fernández, L. D., de Prada, M. B., Guerrero, K. A., et al. (2019). Micro-environmental features associated to container-dwelling mosquitoes (Diptera: Culicidae) in an urban cemetery of the Dominican republic. *Rev. Biol. Trop.* 67, 132–145. doi: 10.15517/RBT.V67I1.33158

Harbach, R., and Knight, K. (1980). *Taxonomists' Glossary of Mosquito Anatomy. Marlton*, New Jersey: Plexus Publishing.

Harrington, L. C., Edman, J. D., and Scott, T. W. (2001). Why do female Aedes aegypti (Diptera: Culicidae) feed preferentially and frequently on human blood? *J. Med. Entomol.* 38, 411–22.

Harrington, L. C., Ponlawat, A., Edman, J. D., Scott, T. W., and Vermeylen, F. (2008). Influence of Container Size, Location, and Time of Day on Oviposition Patterns of the Dengue Vector, *Aedes aegypti*, in Thailand. *Vect. Borne Zoonotic Dis.* 8, 415–424. doi: 10.1089/vbz.2007.0203

Hartig, F. (2020). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. URL: Available online at: https://cran.r-project.org/package=DHARMa.

Hemme, R. R., Tank, J. L., Chadee, D. D., and Severson, D. W. (2009). Environmental conditions in water storage drums and influences on Aedes aegypti in Trinidad, West Indies. *Acta Trop.* 112, 59–66. doi: 10.1016/j.actatropica.2009.06.008

Hery, L., Guidez, A., Durand, A. A., Delannay, C., Normandeau-Guimond, J., Reynaud, Y., et al. (2020). Natural Variation in Physicochemical Profiles and Bacterial Communities Associated with Aedes aegypti Breeding Sites and Larvae on Guadeloupe and French Guiana. *Microb. Ecol.* 81, 93–109. doi: 10.1007/s00248-020-01544-3

Hilbe, J., and Robinson, A. (2013). *Methods of Statistical Model Estimation. 1st ed*. New York: Chapman and Hall/CRC.

Islam, S., Haque, C. E., Hossain, S., and Rochon, K. (2019). Role of container type, behavioural, and ecological factors in Aedes pupal production in Dhaka, Bangladesh: An application of zero-inflated negative binomial model. *Acta Trop.* 193, 50–59. doi: 10.1016/j.actatropica.2019.02.019

Juliano, S. A. (2009). Species Interactions Among Larval Mosquitoes: Context Dependence Across Habitat Gradients. *Annu. Rev. Entomol.* 54, 37–56. doi: 10.1146/annurev.ento.54.110807.090611

Juliano, S. A., Ribeiro, G. S., Maciel-de-Freitas, R., Castro, M. G., Codeço, C., Lourenço-de-Oliveira, R., et al. (2014). She's a femme fatale: Low-density larval development produces good disease vectors. *Mem. Inst. Oswaldo Cruz.* 109, 1070–1077. doi: 10.1590/0074-02760140455

Kraemer, M. U. G., Sinka, M. E., Duda, K. A., Mylne, A. Q. N., Shearer, F. M., Barker, C. M., et al. (2015). The global distribution of the arbovirus vectors Aedes aegypti and Ae. Albopictus. *Elife*. 4:e08347. doi: 10.7554/eLife.08347

Liaw, A., and Wiener, M. (2002). Classification and Regression by randomForest. *R News*.

Loeb, S. L. (1981). An in situ method for measuring the primary productivity and standing crop of the epilithic periphyton community in lentic systems. *Limnol. Oceanogr.* 26, 394–399. doi: 10.4319/lo.1981.26.2.0394

Lourenço-de-Oliveira, R., Castro, M. G., Braks, M. A. H., and Lounibos, L. P. (2004). The invasion of urban forest by dengue vectors in Rio de Janeiro. *J. Vector Ecol.* 29, 94–100. Available online at: http://www.ncbi.nlm.nih.gov/pubmed/15266746.

Macdonald, W. W. (1956). Aedes aegypti in Malaya. II. Larval and adult biology. *Ann. Trop. Med. Parasitol.* 50, 399–414.

Maciel-De-Freitas, R., Codeço, C. T., and Lourenço-De-Oliveira, R. (2007a). Body size-associated survival and dispersal rates of Aedes aegypti in Rio de Janeiro. *Med. Vet. Entomol.* 21, 284–292. doi: 10.1111/j.1365-2915.2007.00694.x

Maciel-de-Freitas, R., Marques, W. A, Peres, R. C., Cunha, S. P., and de Oliveira, R. L. (2007b). Variation in Aedes aegypti (Diptera: Culicidae) container productivity in a slum and a suburban district of Rio de Janeiro during dry and wet seasons. *Mem. Inst. Oswaldo Cruz* 102, 489–96.

Mazerolle, M. (2019). *Model selection and multimodel inference based on (Q)AIC(c). R package version 2.2-2*. Available from https://cran.r-project.org/package=AICcmodavg. R Packag.

Medeiros-Sousa, A. R., de Oliveira-Christe, R., Camargo, A. A., Scinachi, C. A., Milani, G. M., Urbinatti, P. R., et al. (2020). Influence of water's physical and chemical parameters on mosquito (Diptera: Culicidae) assemblages in larval habitats in urban parks of São Paulo, Brazil. *Acta Trop.* 205:105394. doi: 10.1016/j.actatropica.2020.105394

Messina, J. P., Kraemer, M. U., Brady, O. J., Pigott, D. M., Shearer, F. M., Weiss, D. J., et al. (2016). Mapping global environmental suitability for Zika virus. *Elife* 5:15272. doi: 10.7554/eLife.15272

Meyer, D., Zeileis, A., and Hornik, K. (2020). *vcd: Visualizing Categorical Data*. Available online at: https://cran.r-project.org/package=vcd.

Mohammed, A., and Chadee, D. D. (2011). Effects of different temperature regimens on the development of Aedes aegypti (L.) (Diptera: Culicidae) mosquitoes. *Acta Trop.* 119, 38–43. doi: 10.1016/j.actatropica.2011.04.004

Nasci, R. S. (1991). Influence of larval and adult nutrition on biting persistence in Aedes aegypti (Diptera: Culicidae). *J. Med. Entomol.* 28, 522–526. doi: 10.1093/jmedent/28.4.522

Nsoesie, E. O., Kraemer, M. U., Golding, N., Pigott, D. M., Brady, O. J., Moyes, C. L., et al. (2016). Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Eurosurveillance* 21:30234. doi: 10.2807/1560-7917.ES.2016.21.20.30234

Panigrahi, S. K., Barik, T. K., Mohanty, S., and Tripathy, N. K. (2014). Laboratory Evaluation of Oviposition Behavior of Field Collected Aedes Mosquitoes. *J. Insects* 2014, 1–8. doi: 10.1155/2014/207489

Peryassú, A. (1908). *Os culicídeos do Brasil. Rio de Janeiro*, Brazil: Instituto de Manguinhos,.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* 1:409. doi: 10.1007/978-3-540-74686-7

Reiskin, M. H., and Lounibos, L. P. (2009). Effects of intraspecific larval competition on adult longevity in the mosquitoes Aedes aegypti and Aedes albopictus. *Med. Vet. Entomol.* 23, 62–68. doi: 10.1111/j.1365-2915.2008.00782.x

Rueda, L. M., Patel, K. J., Axtell, R. C., and Stinner, R. E. (1990). Temperature-dependent development and survival rates of Culex quinquefasciatus and Aedes aegypti (Diptera: Culicidae). *J. Med. Entomol.* 27, 892–898. doi: 10.1093/jmedent/27.5.892

Sasmita, H. I., Tu, W.-C., Bong, L.-J., and Neoh, K.-B. (2019). Effects of larval diets and temperature regimes on life history traits, energy reserves and temperature tolerance of male Aedes aegypti (Diptera: Culicidae): optimizing rearing techniques for the sterile insect programmes. *Parasit. Vectors* 12:578. doi: 10.1186/s13071-019-3830-z

Schneider, J. R., Morrison, A. C., Astete, H., Scott, T. W., and Wilson, M. L. (2004). Adult Size and Distribution of Aedes aegypti (Diptera: Culicidae) Associated with Larval Habitats in Iquitos, Peru. *J. Med. Entomol.* 41, 634–642. doi: 10.1603/0022-2585-41.4.634

Scott, T. W., Amerasinghe, P. H., Morrison, A. C., Lorenz, L. H., Clark, G. G., Strickman, D., et al. (2000). Longitudinal Studies of Aedes aegypti (Diptera: Culicidae) in Thailand and Puerto Rico: Blood Feeding Frequency. *J. Med. Entomol.* 37, 89–101. doi: 10.1603/0022-2585-37.1.89

Strickman, D., and Kittayapong, P. (2003). Dengue and its vectors in Thailand: Calculated transmission risk from total pupal counts of Aedes aegypti and association of wing-length measurements with aspects of the larval habitat. *Am. J. Trop. Med. Hyg.* 68, 209–217. doi: 10.4269/ajtmh.2003.68.209

Tun-Lin, W., Burkot, T. R., and Kay, B. H. (2000). Effects of temperature and larval diet on development rates and survival of the dengue vector Aedes aegypti in north Queensland, Australia. *Med. Vet. Entomol.* 14, 31–7.

Valderrama, J. C. (1981). The simultaneous analysis of total nitrogen and total phosphorus in natural waters. *Mar. Chem.* 10, 109–122. doi: 10.1016/0304-4203(81)90027-X

Walsh, R. K., Aguilar, C. L., Facchinelli, L., Valerio, L., Ramsey, J. M., Scott, T. W., et al. (2013). Regulation of aedes aegypti population dynamics in field systems: Quantifying direct and delayed density dependence. *Am. J. Trop. Med. Hyg.* 89, 68–77. doi: 10.4269/ajtmh.12-0378

Wong, J., Morrison, A. C., Stoddard, S. T., Astete, H., Chu, Y. Y., Baseer, I., et al. (2012). Linking oviposition site choice to offspring fitness in Aedes aegypti: Consequences for targeted larval control of dengue vectors. *PLoS Negl. Trop. Dis.* 6:e1632. doi: 10.1371/journal.pntd.0001632

Yee, D. A., Kesavaraju, B., and Juliano, S. A. (2006). Interspecific Differences in Feeding Behavior and Survival Under Food-Limited Conditions for Larval Aedes albopictus and Aedes aegypti (Diptera: Culicidae). *Ann. Entomol. Soc. Am.* 97, 720–728. doi: 10.1603/0013-87462004097

Zuur, A. F., Ieno, E. N., and Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* 1, 3–14. doi: 10.1111/j.2041-210x.2009.00001.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fevo.2021.616769

The Larval Density of Mosquitos (Diptera: Culicidae) in Jiaxiang County, Shandong Province, China: Influence of Bacterial Diversity, Richness, and Physicochemical Factors

*Haiyang Wang, Yang Wang, Peng Cheng, Huaiwei Wang, Haifang Wang, Hongmei Liu, Chongxing Zhang and Maoqing Gong**

Shandong Institute of Parasitic Diseases, Shandong First Medical University, Shandong Academy of Medical Sciences, Jining, China

OPEN ACCESS

Edited by:
Mauro Marrelli, University of São Paulo, Brazil

Reviewed by:
Hassan Vatandoost, Tehran University of Medical Sciences, Iran
Antônio Ralph Medeiros De Sousa, University of São Paulo, Brazil

***Correspondence:**
Maoqing Gong
mqgong@sdfmu.edu.cn

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *09 November 2020*
Accepted: *02 February 2021*
Published: *26 February 2021*

Citation:
Wang H, Wang Y, Cheng P, Wang H, Wang H, Liu H, Zhang C and Gong M (2021) The Larval Density of Mosquitos (Diptera: Culicidae) in Jiaxiang County, Shandong Province, China: Influence of Bacterial Diversity, Richness, and Physicochemical Factors. Front. Ecol. Evol. 9:616769. doi: 10.3389/fevo.2021.616769

As Jiaxiang County of Shandong province is an area with complex mosquito vector composition, it is necessary to investigate the relationship between bacterial diversity, physicochemical factors, and larval density. Therefore, the physicochemical properties of 46 breeding sites for six kinds of habitat types (small puddles, small water containers, paddy fields, large water containers, irrigation channels, and drainage ditches) were investigated by a multiparameter analyzer; the water's bacterial diversity was analyzed by the 16S rRNA full-length sequencing method. Spearman correlation and multiple linear regression were used to analyze the correlation between larval density and variables. The variables analyzed were dissolved oxygen, pH, hardness, turbidity, conductivity, temperature, ammonia nitrogen, water depth, and distance from the nearest house. One-Way ANOVA was used to understand whether there are differences in bacterial diversity in different habitats. Pearson linear correlation model was used to analyze the effects of bacterial diversity and richness on mosquito densities in breeding sites. A total of 3291 larvae were captured, and a total of 6 species of 4 genera were identified. The identified species were *Culex pipiens pallens*, *Aedes albopictus*, *Anopheles sinensis*, *Culex tritaeniorhynchus*, *Culex bitaeniorhynchus*, and *Mansonia uniformis*. The density and species can be jointly affected by physicochemical properties and bacterial diversity, especially Shannon index and distance from the nearest house. In general, the physicochemical parameters and bacterial diversity of different habitats were significantly different. Even for the same habitat type, the physicochemical parameters varied greatly due to different environments.

Keywords: full-length sequencing, habitat types, density, mosquito, bacterial diversity

INTRODUCTION

Mosquito is an important medical arthropod transmitting various mosquito-borne diseases such as malaria, dengue fever, and Japanese encephalitis. The larva and pupa stages of various mosquitoes are in water all the time, and the water body with positive larvae is regarded as their breeding sites (El-Ghani et al., 2012; Guo et al., 2017; Ferreira-de-Lima and Lima-Camara, 2018). Jiaxiang County is located in the Southwest Plain of Shandong Province, China. It has a temperate monsoon climate, numerous rivers and lakes, and a developed water system. All these conditions are suitable for breeding a variety of mosquitoes. *Culex pipiens pallens* and *Aedes albopictus* are the dominant mosquito species (Wang et al., 2020). The former is a potential vector of the West Nile virus (WNV) in Shandong Province (Jiang et al., 2014). The latter caused an outbreak of the dengue virus in Jiaxiang County in August 2017. There were 79 cases, which made the area a focus of attention for epidemic prevention personnel (Shi et al., 2019; Liu et al., 2020). In addition, *Anopheles sinensis* is the most important plasmodium vector in Shandong Province. In 2010, Shandong Province reported 117 malaria cases, among which 70% were imported cases. This may be related to the large rice and lotus cultivation area in southwest Shandong Province, which created a favorable breeding environment for Anopheles (Liu et al., 2015).

It is widely believed that larval source management (LSM) can effectively prevent mosquitoes' breeding and development. Researches show that LSM has an inhibitory effect on transmitting mosquito-borne diseases, whether in the countryside or a city. However, Jiaxiang County is rich in water resources, is densely populated, and its residents have a weak awareness of prevention, applying unreasonable mosquito control methods, and inadequate sanitation management. Hence, it is difficult to manage the larval breeding sites in this area (Wang et al., 2020). The Breteau Index (BI), an indicator of the density of *Aedes* mosquitoes in an area, can be used to assess the risk of dengue epidemics in that area. According to reports, the average BI of Jiaxiang County in August 2017 hit a stunning 107.27, whereas, for 2017 and 2018, it was 45.30 and 15.95, respectively (Liu et al., 2020). Areas where BI>20 will be regarded as high-risk means that any importation of external cases could cause a dengue epidemic in this area (Aryaprema and Xue, 2019). These data indirectly indicate that *Ae. albopictus* density in Jiaxiang County is high, and there is a potential risk of mosquito-borne virus transmission. In addition, in 2018, the resistance ratio of *Cx. p. pallens* to DDVP, propoxur, cypermethrin and deltamethrin was 19.76, 4.6, 438, and 351, respectively. Also, over the last decade, the resistance of *Cx. p. pallens* to different types of insecticides has risen to varying degrees (Liu et al., 2019; Wang et al., 2020). Therefore, scientific guidance is urgently needed to carry out LSM in this county. To some extent, understanding the larval density and distribution in this area, as well as the environment of breeding sites, is the first step in this plan.

While choosing the spawning site, female mosquitoes follow specific cues such as physicochemical factors including dissolved oxygen (DO), pH, hardness, turbidity, conductivity, temperature, ammonia nitrogen (AN), water depth (WD), and several biological factors (e.g., the composition of microorganism in water) (Juliano et al., 2004). They try to avoid water bodies where competitors and/or predators are already present and prefer to choose safe water bodies with specific physicochemical properties and microorganisms (Bentley and Day, 1989; Lin et al., 2008). In addition, the distance from the nearest house (DNH) is an important factor that needed to be considered to estimate the significance of the study to public health (Killeen et al., 2001). The research demonstrated that the distance between breeding sites and the nearest house determined the level of blood uptake by a female mosquito, which in turn affected the density, emergence rate of larvae, growth cycle, and malaria transmission dynamics (Killeen et al., 2001). However, with the expansion of urban sprawl to rural areas and frequent personnel exchanges, significant changes have taken place in the environment of breeding sites (Jacob et al., 2005). This affects the density and distribution of mosquito larvae, which, to a certain degree, increases the risk of transmission of mosquito-borne diseases and boosts the pressure of its prevention and control as well.

Bacteria in the breeding sites and other biological factors jointly create the external environment in which the larvae can live, exchanging materials directly with larvae intestines (Dada et al., 2018). This can directly affect the survival status and even the density of the larvae (Mukhopadhyay and Chatterjee, 2016). Hence, it is necessary to investigate the composition of bacteria of water in breeding sites. This information will enable us to estimate whether water bacterial diversity was related to larval density. To our knowledge, habitat analysis on breeding sites is very limited, and reports on bacterial diversity and richness.

of the breeding sites are short (Molineaux, 1997; Qing et al., 2020). To enrich the research findings, we analyzed the relationship between the larval density of Jiaxiang County in Shandong Province and the physicochemical properties and bacterial diversity of water in the breeding sites, and explain the correlation between them using statistics. If water's physicochemical properties and/or bacterial diversity vary, our results provide a rationale prediction of mosquito density and help develop scientific strategies for mosquito vector control.

MATERIALS AND METHODS

Selection of the Study Area

Jiaxiang County (35°11′ N, 116°06′ E) is located in the southwest of Shandong Province. Jiaxiang County is hot and rainy from June to August, with an average rainfall of 398.5 mm in summer. Zhu Zhaoxin River passes through, and the water system is developed, providing many favorable conditions for mosquito breeding and reproduction. We selected the Jintun town as the center, which had a dengue fever outbreak in August 2017, and radiated out to the surrounding towns and villages to survey the breeding sites. We invited several epidemic prevention personnel from the local health center to assist us in sampling. Water bodies with larva or pupa activity or eggs were selected as sampling sites (Nambunga et al., 2020). Then, GPS was used to locate breeding sites. The distribution of breeding sites is shown in **Figure 1**.

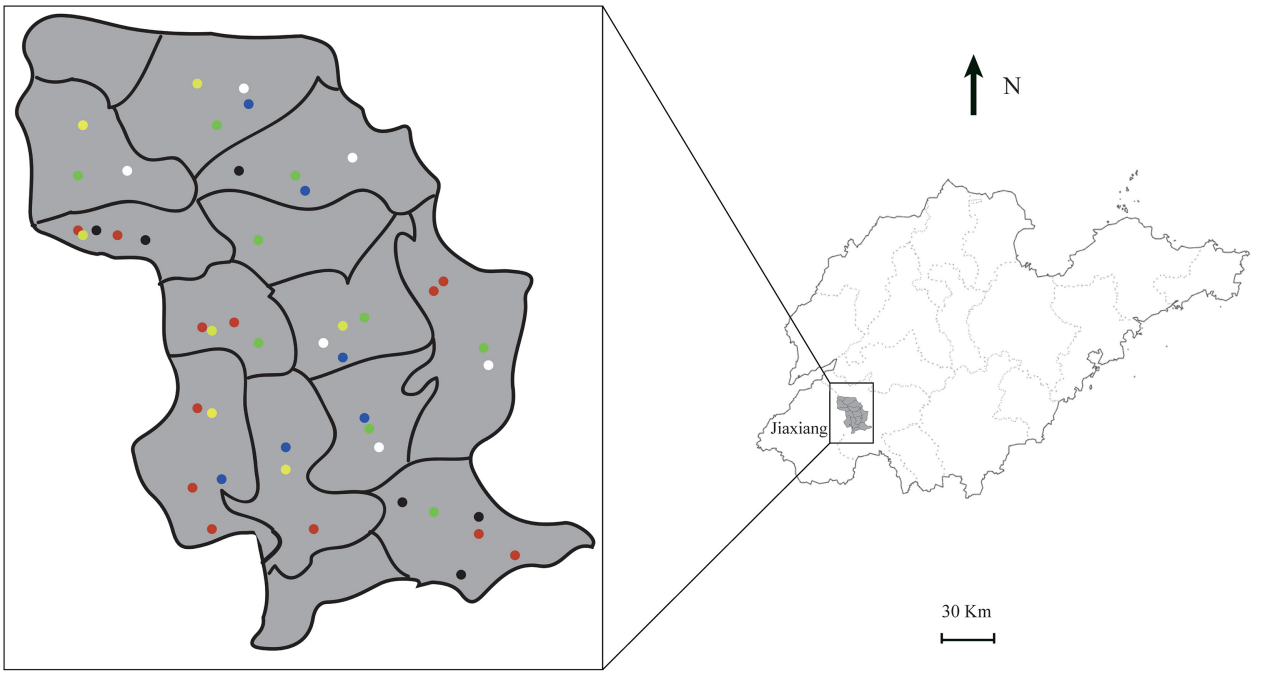

FIGURE 1 | Locations and habitats type of sampling sites in Jiaxiang County. Six black dots represent small puddles, 12 red dots representing small water containers, seven yellow dots representing paddy fields, nine green dots representing large water containers, six white dots representing irrigation channels, and six blue dots representing drainage ditches. The location of each dot displayed a sampling site.

Larva Collection and Identification

Mosquito larva was collected from each breeding site using a standard 350 mL long handle dipper. Each breeding site was dipped four times within a collection period. When the sampling site could not be immersed, siphoning or dumping was used. According to the planned date, each sampling site was sampled five times: July 7, 2019, July 14, 2019, July 23, 2019, July 31, 2019, and August 7, 2019. In the case of thunderstorms or other exceptional circumstances, sampling was postponed. If the water at the sampling site was dry, we located a same-type breeding site nearby. Larvae collected from each breeding site were brought to the Shandong Institute of Parasitic Diseases and classified into species level using standard identification keys (Wilkerson et al., 1990). Quantification of larval density (Villarreal-Treviño et al., 2015) was as follows: D (per dipper) = $\frac{\text{the total number of larvae}}{\text{total number of dips}}$, where D = mosquito density. The calculation method of mosquito species distribution was based on Kabirul's method (Bashar et al., 2016): C = $\frac{n}{N} \times 100\%$, where C = mosquito species distribution, n = number of positive habitats, and N = the total number of habitats.

Identification of the Water Physicochemical Properties

At the site, we measured the water physicochemical properties of each sampling site, including pH, turbidity (NTU: Nephelometric turbidity unit), conductivity (μS/cm), DO (mg/L), AN (mg/L), hardness (mg/L), and temperature (°C) using a portable multiparameter tester (Palintest Potacheck 2, United Kingdom). WD (cm) and DNH (m) were measured by a graduated string with a lead weight and a tapeline, respectively. After the measurement, at least 200 mL of water was taken from each sampling site and placed into a sterile conical flask. Then, the flask was placed into an icebox and brought back to the laboratory to analyze the water body's bacterial diversity and richness. The collection time, locations, environmental status, and habitat types were recorded timely and in detail.

Analysis of Bacterial Diversity and Richness

The water samples of each breeding site were centrifuged at 12000 rpm for 10 min in 4°C, and the supernatant was discarded to obtain bacterial precipitation. A Power Soil DNA Isolation kit was used to extract the total bacterial DNA, and the primer sequence 27F (AGRGTTTGATYNTGGCTCAG) and 1492R (TASGGHTACCTTGTTASGACTT) were used to amplify the full-length 16S rRNA gene. PCR reaction conditions were 95°C for 5 min, 95°C for 30 s, 50°C for 30 s, and 72°C for 1 min, with a total of 30 cycles and a reaction system of 10 μL. After the construction of magnetic beads (MagicPure Size Selection DNA Beads) purification, preparing library, and target gene sequencing, in turn, original reads were corrected to obtain Circular Consensus Sequencing (CCS) (SMRT Link, Version 8.0). The Lima software (version 1.7.0) was used to identify the CCS sequences of different samples by a barcode sequence and a remove chimera (UCHIME, version 8.1), resulting in high-quality CCS sequences. The sequences were clustered at a

similarity level of 97% (USEARCH, version 10.0) (Quast et al., 2013) and the OTU (Operational Taxonomic Units) were filtered with 0.005% of all numbers sequenced as the threshold (Kõljalg et al., 2013). Bacteria chose 16S: Silva databases (Release 132)[1] (Wang et al., 2007) and RDP Classifier (version 2.2)[2] (Larkin et al., 2007), and species were annotated with a confidence threshold of 0.8. Finally, the Alpha index analysis software Mothur (Version V.1.30)[3] was used to analyze the microbial diversity and richness of different habitat samples.

Statistical Analysis

All variables were tested by the Shapiro–Wilk test for normality and then analyzed for homogeneity of variance. We used a simple correlation coefficient to analyze the correlation between continuous variables of physicochemical factors. DNH was not considered because there was no logical relationship between variables. Results showed that the variables of physical and chemical factors were not normally distributed and had different variances, and they were not independent of each other. Therefore, it was more reasonable to use a Spearman correlation and multiple linear regression to analyze the correlation between mosquito density and physical and chemical factors (Nikookar et al., 2017). Canonical correspondence analysis (CCA) has been performed to further explore the associations between larval densities and physicochemical factors (Bashar et al., 2016). One-way ANOVA was used to explain whether there were differences in bacterial diversity and richness among different habitat types. Pearson's correlation model was used to analyze the effects of bacterial diversity and richness on larval densities in breeding sites.

All statistical analyzes were performed using Microsoft Excel 2019 for Windows and SPSS 16 (SPSS Inc., Chicago, IL, United States 2007), with a significance level of 0.05.

RESULTS

Habitat Types

A total of 62 representative potential habitats were investigated. Out of these, 46 bred larvae. We classified these 46 breeding sites into different habitat types according to the situation on the spot, such as small puddles (6), small water containers (12), paddy fields (7), large water containers (9), irrigation channels (6), and drainage ditches (6). **Figures 1**, **2** show the locations and habitat types.

Larval Density and Species Distribution

A total of 3,291 larvae were collected in all breeding sites, and a total of four genera and six species were identified, including *Cx. p. pallens* (2026/61.56%), *Ae. albopictus* (732/22.24%), *An. sinensis* (147/4.47%), *Cx. tritaeniorhynchus* (319/9.69%), *Cx. bitaeniorhynchus* (13/0.40%), and *Ma. uniformis* (54/1.64%). The larvae densities of each habitat type were: small puddles (3.38), small water containers (4.55), paddy fields (2.31), large water containers (4.42), irrigation channels (2.96), and drainage ditches (2.66), respectively (see **Table 1**). In terms of the distribution of mosquito species, *Cx. p. pallens* was the highest (C = 98.70%), which could be found in all habitat types. It was followed by *Ae. albopictus* (C = 67.83%), *Cx. tritaeniorhynchus* (C = 55.65%), *An. sinensis* (C = 25.65%), and *Ma. uniformis* (C = 10.43%). The distribution of *Cx. bitaeniorhynchus* was the smallest (C = 4.35%). *An. sinensis* was only present in paddy fields and irrigation channels, *Cx. bitaeniorhynchus* was found only in large water containers, and *Ma. uniformis* was found only in irrigation channels.

The Physical and Chemical Parameters and Bacterial Diversity of Breeding Sites

Table 2 shows the means and standard deviations of various variables for each habitat. **Figure 3** shows the mean and standard error of the OTU, bacterial diversity index (Shannon index), and the richness index (Chao 1) of breeding sites. The Shannon diversity index dilution curve **(Figure 4A**) showed each sample's microbial diversity at different sequencing quantities. The curve was flat, indicating that the sequencing quantity was large enough, and the results of OTU clustering accurately reflected the samples' real situation. The breeding sites with the highest Shannon index were drainage ditches, and the lowest was small water containers. **Figure 4B** shows the top-ten species at the richness level in each habitat. The bacteria with the highest richness in small puddles was *Hydrogenophaga*. The bacteria with the highest richness in small water containers, paddy fields, and irrigation channels were all *Curvibacter*. *Polynucleobacter* microbes were most abundant in large water containers. Results of one-way ANOVA showed that the Shannon index (df = 5, F = 5.841, p = 0.001) and Chao 1 (df = 5, F = 10.859, $p < 0.01$) were different between habitat types.

Correlations Between Variables

Table 3 shows the correlation coefficients between the variables of each physicochemical factor. Among the 36 correlation coefficients, 7 (19.44%) were statistically significant, indicating a non-random correlation between the variables with physicochemical properties. pH was negatively correlated with turbidity ($r = -0.442$, p = 0.002) and conductivity ($r = -0.502$, $p < 0.01$). A significant negative and positive correlation existed between hardness with turbidity ($r = -0.325$, p = 0.028) and with AN (r = 0.339, p = 0.021), respectively. Turbidity was positively correlated with conductivity (r = 0.357, p = 0.015) and negatively correlated with AN ($r = -0.375$, p = 0.01). Temperature was positively correlated with WD ($r = -0.407$, p = 0.005).

Relationship Between Larval Density and Variables

Spearman correlation analysis (**Table 4**) shows the correlations between larval densities and physical and chemical parameters. In the multiple linear regression model analysis, the larvae densities of various types were taken as dependent variables, and then 9 physical and chemical factors, as independent variables

[1]http://www.arb-silva.de

[2]http://sourceforge.net/projects/rdpclassifier/

[3]http://www.mothur.org/

FIGURE 2 | Representative images from the larval habitats in Jiaxiang County.

TABLE 1 | Number and density of each species in different habitat types.

Mosquito species	Habitat types						*n*	%
	Small puddles (6)	Small water containers (12)	Paddy fields (7)	Large water containers (9)	Irrigation channels (6)	Drainage ditches (6)		
Cx. p. pallens	265/2.21	677/2.82	158/1.13	523/2.91	184/1.53	219/1.82	2026/12.42	61.56
Ae. albopictus	129/1.08	358/1.49	48/0.34	197/1.09	0	0	732/4.00	22.24
An. sinensis	0	0	92/0.66	0	55/0.46	0	147/1.12	4.47
Cx. tritaeniorhynchus	12/0.10	58/0.24	25/0.18	62/0.34	62/0.52	100/0.83	319/2.21	9.69
Cx. bitaeniorhynchus	0	0	0	13/0.08	0	0	13/0.08	0.40
Ma. uniformis	0	0	0	0	54/0.45	0	54/0.45	1.64
Total	406/3.38	1093/4.55	323/2.31	795/4.42	355/2.96	319/2.66		

The numbers in front of the slash indicate the number of larvae, followed by the density.

respectively, were incorporated into the model. The variables were fitted into the regression equation according to the stepwise selection method, and the independent variables with a large contribution to the regression equation were reserved ($p < 0.05$) were retained. According to the results of multiple linear regression analyzes (**Table 5**), *Cx. p. pallens* ($F = 18.30$, $p < 0.001$) required low DNH, whereas *An. sinensis* ($F = 72.388$, $p < 0.001$) required high DNH. *Ae. albopictus* ($F = 16.239$, $p < 0.001$) required less DNH, AN, and pH, but a higher temperature. These variables explained the density of *Ae. albopictus* more strongly than other variables. *Cx. tritaeniorhynchus* ($F = 7.929$, $p = 0.001$) required high AN and WD. Compared to other variables, high WD and DNH and low pH contributed the most to *Ma. uniformis* ($F = 33.498$, $p < 0.001$). Our results showed that *Cx. bitaeniorhynchus* did not fit the regression equation as the *P* value was greater than 0.05.

Table 6 showed that the density of *Cx. p. pallens* negatively correlated with the Shannon index ($r = -0.430$, $p = 0.018$)

TABLE 2 | Means and standard deviations of physicochemical characteristics in different larval habitats.

(Mean ± SD)	Habitats types					
	Small puddles	**Small water containers**	**Paddy fields**	**Large water containers**	**Irrigation channels**	**Drainage ditches**
DO	5.56 ± 0.24	6.22 ± 0.52	6.11 ± 1.01	6.23 ± 0.65	7.30 ± 1.16	4.60 ± 0.85
pH	8.12 ± 0.56	7.08 ± 0.40	8.01 ± 0.33	7.97 ± 0.57	8.02 ± 0.47	8.40 ± 0.85
Hardness	211.67 ± 64.14	129.17 ± 54.08	167.71 ± 92.96	118.11 ± 22.66	204.33 ± 99.08	342.33 ± 182.07
Turbidity	114.18 ± 45.28	509.54 ± 70.01	98.19 ± 7.60	402.83 ± 180.71	147.13 ± 43.20	254.97 ± 191.65
Conductivity	165.65 ± 18.95	327.80 ± 113.98	152.97 ± 43.74	192.41 ± 162.93	154.63 ± 29.46	162.50 ± 11.75
Temperature	25.23 ± 1.51	26.43 ± 2.26	25.79 ± 1.98	24.33 ± 3.49	19.30 ± 2.03	24.10 ± 2.73
AN	0.54 ± 0.12	0.68 ± 0.16	2.13 ± 0.49	0.50 ± 0.12	1.64 ± 0.47	9.74 ± 5.90
WD	3.67 ± 1.63	13.00 ± 5.43	10.71 ± 1.38	51.89 ± 13.81	97.00 ± 17.65	41.67 ± 10.15
DNH	38.83 ± 16.79	5.25 ± 3.99	176.00 ± 96.08	6.56 ± 4.56	192.83 ± 88.11	34.83 ± 32.11

DO, dissolved oxygen; AN, ammonia nitrogen; WD, water depth; DNH, distance from the nearest house.

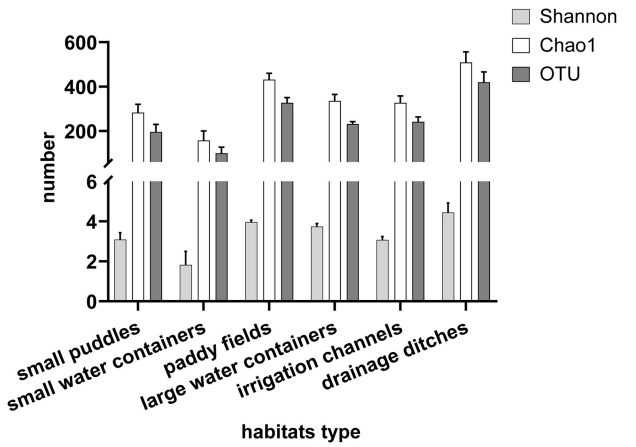

FIGURE 3 | Mean and standard error of the OTU, Shannon Index, and Chao 1 of the 6 habitat types. Shannon index is the bacterial diversity index used to measure species diversity, which was affected by species richness and community evenness in the sample. Chao 1 is the richness index, representing the number of species.

and the richness index ($r = -0.644$, $p < 0.001$). The density of *Ae. albopictus* negatively correlated with the bacterial diversity index ($r = -0.411$, $p = 0.024$) and the richness index ($r = -0.572$, $p = 0.001$). *Cx. tritaeniorhynchus* positively correlated with the Shannon index ($r = 0.493$, $p = 0.006$), while other mosquito species did not significantly correlate with bacterial diversity. **Figure 5** visually showed the correlation between larval density and variables based on the Canonical correspondence analysis (CCA).

DISCUSSION

Mosquito problems have been categorized as medical care problems and environmental issues with the progress of human society, urban sprawl, and intensification of environmental changes (Ahmed et al., 2007). Sole dependence on insecticides can no longer completely eliminate the hazards that mosquitoes had brought to us. Luckily, LSM is environment friendly and does not lead governments and inhabitants to too much economic burden and time cost. These could be rapidly realized utilizing eliminating breeding sites or using chemical or biological methods. However, developing efficient and safe strategies required a scientific understanding of the control area's larval density and environmental factors.

In recent years, due to the vigorous development in modern agriculture, the mode of agriculture in Jiaxiang County has changed. For example, paddy fields have been replaced by citrus orchards, which directly resulted in the change of some breeding sites' environment. We suspect that this may be a factor that led to the failure to obtain mosquitoes such as *An. dirus*, which prefer open water (Service, 1991; Norris, 2004). In addition, agricultural activities such as the use of pesticides and fertilizers changed the richness of microbiota and algae in the water, which are the main food sources for larvae (Noori et al., 2016). This disturbed selecting suitable sites for oviposition and incubation by specific mosquitoes.

Most of the microbiota in the breeding sites form parasitic or symbiotic relationships with the larvae, and their impact on the larvae can be classified in the following three aspects: regulating the internal and external environment and immune response of the mosquito host, promoting the developments and growth of larvae, and causing pathogenicity to larvae such as infection of the filamentous fungi *Beauveria bassiana* and *Bacillus thuringiensis* (Bti). In addition, studies have shown that bacterial diversity might be related to larval density (Dida et al., 2018). Therefore, we used molecular biology methods, namely the 16S rRNA sequencing technology, to sequence and analyze the bacterial diversity and richness of breeding sites. Results showed that different habitats have large differences in bacterial diversity and richness, even in the same habitat, because bacterial diversity and richness in water are always dynamic and affected by their environment, such as light, temperature, vegetation cover, and human activities. In general, *Cx. p. pallens* and *Ae. albopictus*, the dominant species in Jiaxiang County, both showed lower tolerance to bacteria, which also seems to explain the lower diversity of intestinal symbiotic bacteria in these two species (Coon et al., 2016). There was no statistical correlation between the density of other species and the bacterial diversity, which we hypothesized might be due to the small

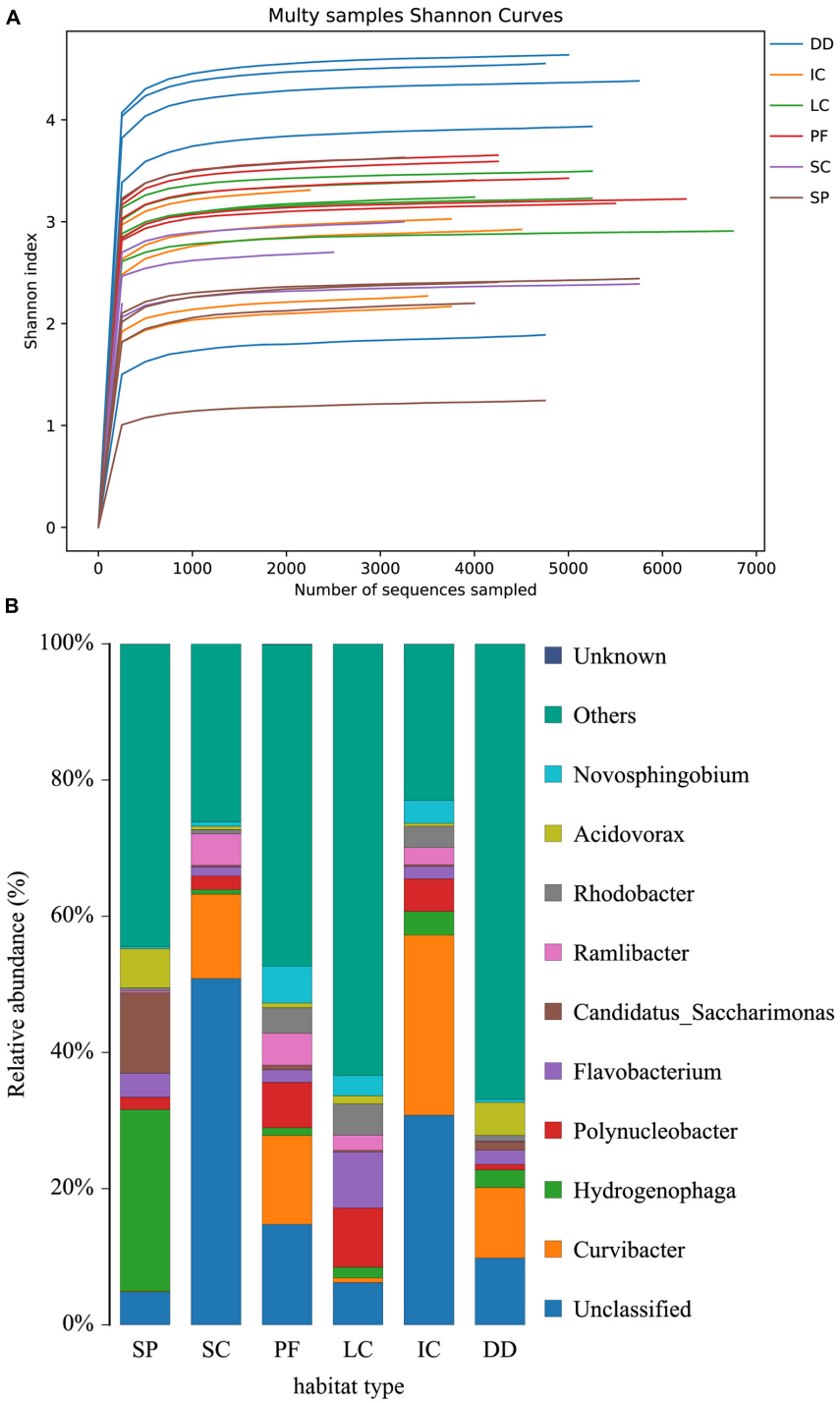

FIGURE 4 | Alpha diversity analysis: **(A)** Shannon diversity index dilution curve. It shows the microbial diversity of each sample at different sequencing quantities. The curve is flat, indicating that the sequencing quantity is large enough, and the results of Shannon index accurately reflected the real situation of the samples. **(B)** The top-ten species at the richness level in each habitat. DD, drainage ditches; IC, irrigation channels; LC, large water containers; PF, paddy fields; SC, small water containers; SP, small puddles.

number of samples we tested from water bodies or the large variation in the diversity index of different habitat types. *Cx. p. pallens* has a wide range of fitness for physical and chemical properties such as pH, hardness, turbidity, and AN, which is consistent with the results of Nikookar et al., (Nikookar et al., 2017). In other words, *Cx. p. pallens* have no strict requirements

TABLE 3 | Correlation between variables with physicochemical property.

	DO	pH	Hardness	Turbidity	Conductivity	Temperature	AN	WD
DO	1							
pH	−0.226	1						
Hardness	−0.286	0.213	1					
Turbidity	0.118	−0.442**	−0.325*	1				
Conductivity	−0.030	−0.502**	−0.212	0.357*	1			
Temperature	−0.104	−0.113	−0.198	0.179	0.088	1		
AN	−0.079	0.270	0.339*	−0.375*	−0.179	−0.188	1	
WD	0.285	0.231	0.024	0.155	−0.216	−0.407**	0.238	1

***Significant at the 0.01 level. *Significant at the 0.05 level.*

TABLE 4 | Spearman correlation coefficient between larval density and physicochemical characteristics of larval habitats.

	Cx. p. pallens		*Ae. albopictus*		*An. sinensis*		*Cx. tritaeniorhynchus*		*Cx. bitaeniorhynchus*		*Ma. uniformis*	
	r	*P*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
DO	−0.035	0.818	0.148	0.327	0.312*	0.035	−0.068	0.652	0.265	0.075	0.394**	0.007
pH	−0.311*	0.036	−0.408**	0.005	0.239	0.110	0.075	0.621	0.200	0.184	0.122	0.419
Har	−0.239	0.110	−0.442**	0.002	0.085	0.574	0.247	0.097	−0.233	0.119	0.182	0.226
Tur	0.445**	0.002	0.381**	0.009	−0.509**	<0.001	0.078	0.606	0.176	0.243	−0.189	0.209
Con	0.094	0.533	0.229	0.127	−0.293*	0.048	−0.202	0.178	−0.084	0.577	−0.124	0.413
Tem	0.312*	0.035	0.505**	<0.001	−0.259	0.082	−0.153	0.310	−0.057	0.708	−0.537**	<0.001
AN	−0.575**	<0.001	−0.672**	<0.001	−0.512**	<0.001	0.299*	0.044	−0.415**	0.004	0.252	0.091
WD	−0.078	0.607	−0.486**	0.001	0.192	0.202	0.513**	<0.001	0.311*	0.036	0.578**	<0.001
DNH	−0.678**	<0.001	−0.583**	<0.001	0.768**	<0.001	−0.133	0.377	0.187	0.213	−0.304*	0.040

*Har, hardness; Tur, turbidity; Con, conductivity; Tem, temperature. **Significant at the 0.01 level. *Significant at the 0.05 level.*

TABLE 5 | The multiple regression equation shows the contribution rate of each variable to mosquito density.

Mosquito species	Multiple regression equations	*p*	*F*
Cx. p. pallens	Y = 51.163–0.111(DNH)	<0.001	18.30
Ae. albopictus	Y = 37.511–0.056(DNH)–1.377(AN)–5.186(pH) + 1.241(Tem)	<0.001	16.239
An. sinensis	Y = −0.087 + 0.041(DNH)	<0.001	72.388
Cx. tritaeniorhynchus	Y = 3.697 + 0.627(AN) + 0.56(WD)	0.001	7.929
Cx. bitaeniorhynchus	–	0.25	1.343
Ma. uniformis	Y = 5.494 + 0.065(WD) + 0.018(DNH)– 0.979(pH)	<0.001	33.498

TABLE 6 | Pearson correlation coefficient between larval density and bacterial diversity of larval habitats.

		Cx. p. pallens	*Ae. albopictus*	*An. sinensis*	*Cx. tritaeniorhynchus*	*Cx. bitaeniorhynchus*	*Ma. uniformis*
Shannon Index	*r*	−0.430*	−0.411*	0.169	0.346	0.149	−0.096
	p	0.018	0.024	0.373	0.061	0.433	0.613
Chao 1	*r*	−0.644**	−0.572**	0.244	0.493**	0.028	−0.032
	p	< 0.001	0.001	0.194	0.006	0.884	0.867

***Significant at the 0.01 level. *Significant at the 0.05 level.*

regarding the environment they breed in, and they can even breed in less hospitable environments. This seems to explain the dominance of *Cx. p. pallens* in the less sanitary areas of northern China. In addition, Gabriel *et al.,* have shown that *Culex* (especially *Cx. p. pallens* and *Cx. tritaeniorhynchus*) are widely distributed in a variety of water bodies, providing intuitive evidence that they are highly adaptable to different habitats. Urbanization, human activities, and other factors have affected the breeding environment's physicochemical properties and water bodies (Ndaruga et al., 2004). Despite this, the density of *Culex* in Jiaxiang County remains high. Previous reports and our study give a reasonable explanation for this phenomenon.

Ae. albopictus is the main vector of the dengue virus. Our study showed that *Ae. albopictus* density was negatively correlated

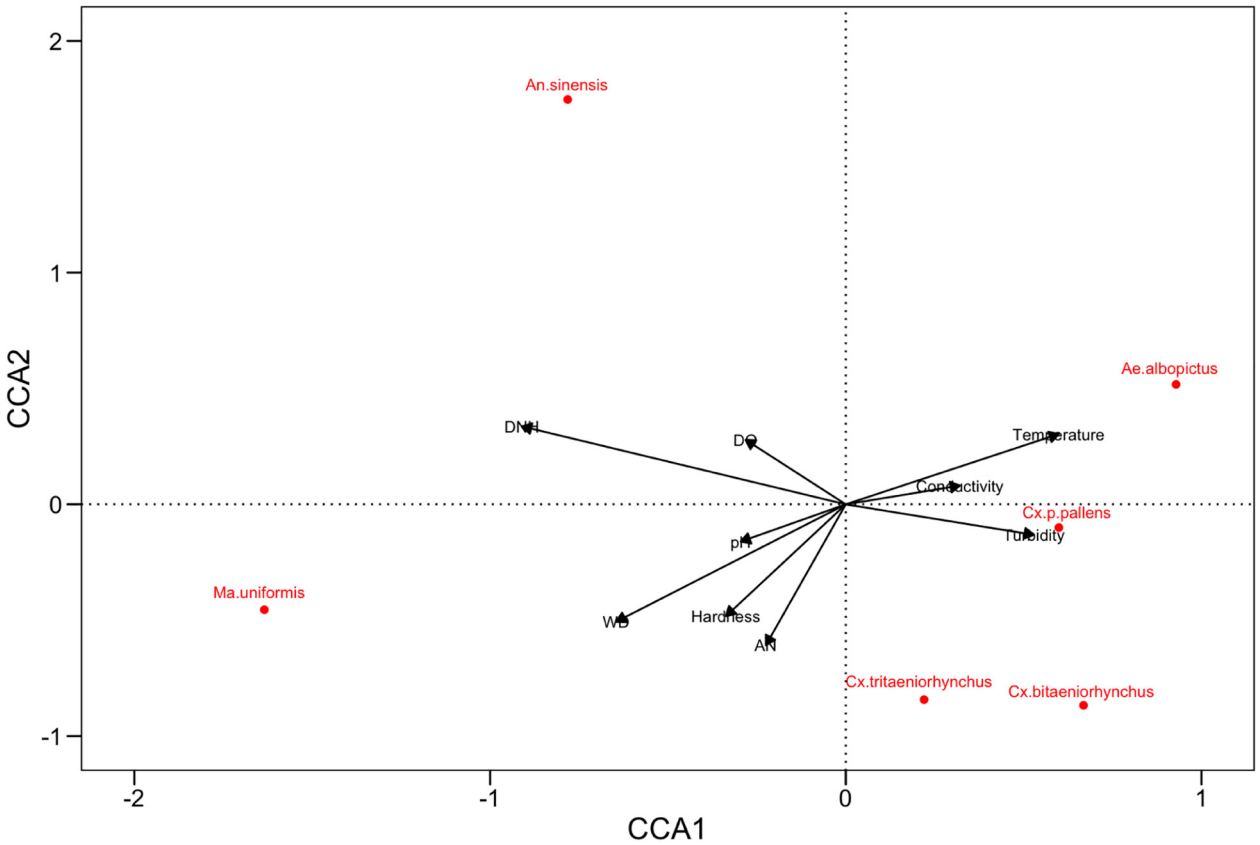

FIGURE 5 | Biplot representing the results of CCA, the relationship between the larval density (red labels) and the physicochemical variables (black arrows). The constrained inertia explained by the CCA is 31.18%, and the proportion explained by CCA1 axis and CCA2 axis was 35.35% and 10.27%, respectively.

with WD. This may be because they prefer to lay eggs in small water containers such as junked tires, plastic bottles, and surface-gathered water and have strong adaptability to low water levels (Li et al., 2014; Petric et al., 2014). There is a negative correlation between *Ae. albopictus* and DNH. In other words, the farther the breeding site is from the house, the lower the density of *Ae. albopictus,* and *vice versa*. This characteristic is observed in all species except *An. sinensis*. WHO has reported that *Ae. albopictus* prefers to lay eggs in small containers having stagnant water in people's yards or houses (World Health Organization, 1983). Based on such characteristics, dumping stagnant water to reduce the incidence of dengue fever by reducing the BI could be an effective and environment-friendly method. Moreover, *Ae. albopictus*, just like *Cx. p. pallens* showed a positive correlation with turbidity and temperature, while a negative correlation with pH and AN. They showed a wide range of adaptation to the abovementioned physical and chemical properties. As Thavara and Bashar reported, *Ae. albopictus* preferred to inhabit human-made habitats inside and outside houses, where the habitats' physical and chemical properties were more fluctuating (Thavara et al., 2004).

The increase in salt compounds and turbidity had a significant effect on the density of *An. sinensis*, and they showed low tolerance to AN levels. Kenawa found a strong positive correlation between its density and dissolved oxygen (El-Ghani et al., 2012), which is consistent with our results (r = 0.312, p = 0.035). In comparison with *Aedes* and *Culex*, *Anopheles* mosquitoes were more sensitive to other physical and chemical changes in water (El-Ghani et al., 2012). Minakawa et al. (2005) suggested that *Anopheles* preferred sunny open waters such as lotus ponds, paddy fields, and irrigation channels where breeding sites are far from houses. Our results also confirm that *Anopheles* mosquito density is positively correlated with DNH. We suspect that this may be related to the strong long-range flight ability of *An. sinensis*. The rice-growing areas and lotus ponds in Jiaxiang County are widely distributed, which provides the conditions for the breeding of *An. sinensis* in the area. Applying agricultural insecticides seem to be the most effective way to control *An. sinensis*. However, the epidemic prevention department should instruct farmers in the region to spray insecticides scientifically and rationally to slow the development of mosquito resistance. On the basis of the existing rice-planting pattern in Jiaxiang County, humid irrigation replacing the current irrigation regime can effectively eliminate the breeding sites for irrigation channels, which is also an effective method to control *An. Sinensis* (Tusting et al., 2013). Minakawa found that *Anopheles* in western Kenya were not found in sunny open water and preferred lakeside swamps. This difference suggests that even mosquitoes from

the same genus may show differences in habitat selection. The primary reason for this phenomenon was geographical isolation and ecological differences (Varela and Yadav, 2020), suggesting that we should first understand their habits and adapt to local conditions to prevent and control larvae.

Culex tritaeniorhynchus is the primary medium carrying the Japanese encephalitis virus and the dominant mosquito species in Shandong Province, especially near the South-to-North Water Diversion Project. In terms of distribution alone, they are similar to *Cx. p. pallens*, as both of them, were found in all habitat types. However, the density of *Cx. tritaeniorhynchus* that we collected was significantly lower than that of *Cx. p. pallens*. They can survive in deep water with low oxygen content, so the density and WD showed a strong positive correlation. From the results of the regression equation, it is evident that AN and WD had the highest density contribution to *Cx. tritaeniorhynchus*, which is the same as the research results of Wang et al., (Wang et al., 2020). *Cx. bitaeniorhynchus* and *Ma. uniformis* showed a low distribution, and they were only found in large water containers and irrigation channels, respectively. Combined with previous surveys, we found that the density of these two species had been low in Jiaxiang County, which did not bring trouble or danger to humans. Therefore, this is not discussed in this article.

Although the CCA results were consistent with the Spearman, it may be able to predict larval density in a much more intuitive way. We can intuitively receive such results from the biplot: the larval density of *Ae. albopictus* and *Cx. p. pallens* can be better predicted by low WD, DNH, and high temperature. Alternatively, high DNH and DO may be more useful in predicting larval density of *An. sinensis* than other variables. The larval density of *Ma. uniformis* can be affected by high WD and low temperature, while the density of *Cx. bitaeniorhynchus* and *Cx. tritaeniorhynchus* can hardly be predicted from the results of CCA. The CCA biplot indicated that *Anopheles*, *Aedes*, *Culex*, and *Mansonia* were plotted in different axis which may explain different requirement for different variables (Bashar et al., 2016).

There are several aspects that need to be improved in this study. We did not conduct a survey of algae, protists, and rotifers in breeding grounds. However, protists are the main source of nutrition for larvae (Blaustein and Chase, 2007; Addicott, 2008), and *Coelastrum*, *Scenedesmus*, *Selenastrum*, and *Tetrallantos* that are algae for *Aedes* and *Culex* are difficult to digest, which can reduce the survival rate of larvae (Howland, 1930). Microbial diversity and physicochemical parameters of water bodies change dynamically, and a large number of survey results are needed to develop a scientific and rigorous model to predict larval density. Thus, our sample number is far from enough.

CONCLUSION

This research showed that the physicochemical parameters and bacterial diversity of different habitat types were significantly different. Even for the same habitat type, the physicochemical parameters varied greatly due to different environments. The variable that contributed the most to the density of *Cx. p. pallens* was DNH. Parameters such as DNH, AN, pH, and temperature contributed significantly to the density of *Ae. albopictus*. These two kinds of mosquitoes are distributed in almost all kinds of habitats in Jiaxiang County, and their density is also the highest. This suggests that we should be vigilant against the control of mosquito-borne viruses in this area, especially the dengue virus carried by *Ae. albopictus*.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MQG and HYW designed the study. HFW and HWW participated in field surveys with support from PC. HYW extracted the data and carried out the data analysis with support from YW. HYW and YW wrote and reviewed the article. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China [81871685 (MQG)] and Shandong Provincial Natural Science Foundation [ZR2020MC048 (PC) and ZR2020KH001 (HML)].

ACKNOWLEDGMENTS

We would like to thank the population of BMKCloud for their collaboration during the analysis of bacterial diversity. We would also like to thank local villagers to help us searching for breeding sites.

REFERENCES

Addicott, J. F. (2008). Predation and prey community structure: an experimental study of the E. *Communication* 55:41. doi: 10.2307/1935141

Ahmed, T. U., Rahman, K. B. M. S., Shamsuzzaman, M., Samajpati, S., Sultana, S., Hossain, M. I., et al. (2007). Some socio-demographic factors related to dengue outbreak in Dhaka City. Bangladesh Bangladesh. *J. Zool.* 35, 213–222. doi: 10.1002/anie.200461295

Aryaprema, V. S., and Xue, R. D. (2019). Breteau index as a promising early warning signal for dengue fever outbreaks in the colombo district. Sri Lanka. *Acta Trop.* 199:105155. doi: 10.1016/j.actatropica.2019.105155

Bashar, K., Sayfur Rahman, M. D., Nodi, I. J., and Howlader, A. J. (2016). Species composition and habitat characterization of mosquito (Diptera: Culicidae) larvae in semi-urban areas of Dhaka. Bangladesh. *Pathogens Global Health* 110:48. doi: 10.1080/20477724.2016.1179862

Bentley, M. D., and Day, J. F. (1989). Chemical ecology and behavioral aspects of mosquito oviposition. *Annu. Rev. Entomol.* 34, 401–421. doi: 10.1146/annurev.en.34.010189.002153

Blaustein, L., and Chase, J. M. (2007). Interactions between mosquito larvae and species that share the same trophic level. *Annu. Rev. Entomol.* 52, 489–507. doi: 10.1146/annurev.ento.52.110405.091431

Coon, K. L., Brown, M. R., and Strand, M. R. (2016). Mosquitoes host communities of bacteria that are essential for development but vary greatly between local habitats. *Mol. Ecol.* 25, 5806–5826. doi: 10.1111/mec.13877

Dada, N., Sheth, M., Liebman, K., Pinto, J., and Lenhart, A. (2018). Whole metagenome sequencing reveals links between mosquito microbiota and insecticide resistance in malaria vectors. *Sci. Rep.* 8:2084. doi: 10.1038/s41598-018-20367-4

Dida, G. O., Anyona, D. N., Abuom, P. O., Akoko, D., Adoka, S. O., Matano, A. S., et al. (2018). Spatial distribution and habitat characterization of mosquito species during the dry season along the Mara River and its tributaries, in Kenya and Tanzania. *Infect Dis. Poverty* 7, 2–18. doi: 10.1186/s40249-017-0385-0

El-Ghani, M. A., Shehata, M. N., Mobarak, A., and Bakr, R. (2012). Factors affecting the diversity and distribution of synanthropic vegetation in urban habitats of the Nile Delta. *Egypt Rendiconti Lincei* 23, 327–337. doi: 10.1007/s12210-012-0179-z

Ferreira-de-Lima, V. H., and Lima-Camara, T. N. (2018). Natural vertical transmission of dengue virus in *Aedes aegypti* and *Aedes albopictus*: a systematic review. *Parasit Vectors* 11:77. doi: 10.1186/s13071-018-2643-9

Guo, C., Zhou, Z., Wen, Z., Liu, Y., Zeng, C., Xiao, D., et al. (2017). Global epidemiology of dengue outbreaks in 1990-2015: a systematic review and meta-analysis. *Front. Cell Infect. Microbiol.* 7:317. doi: 10.3389/fcimb.2017.00317

Howland, L. J. (1930). The nutrition of mosquito larvae, with special reference to their algal food. B. *Entomol. Res.* 21, 431–439. doi: 10.1017/S0007485300024779

Jacob, B. G., Arheart, K. L., Griffith, D. A., Mbogo, C. M., Githeko, A. K., Regens, J. L., et al. (2005). Evaluation of environmental data for identification of anopheles (Diptera: Culicidae) aquatic larval habitats in kisumu and malindi. Kenya. *J. Med. Entomol.* 22, 751–755. doi: 10.1093/jmedent/42.5.751

Jiang, S., Wang, Z., Guo, X., Zhang, Y., Li, C., Dong, Y., et al. (2014). Infection and dissemination of west nile virus in china by the potential vector. culex pipiens pallens, *J. Vector Ecol.* 39, 78–82. doi: 10.1111/j.1948-7134.2014.12073.x

Juliano, S. A., Lounibos, L. P., and O'Meara, G. F. (2004). A field test for competitive effects of *Aedes albopictus* on *A. aegypti* in South Florida: differences between sites of coexistence and exclusion?. *Oecologia* 139, 583–593. doi: 10.1007/s00442-004-1532-4

Killeen, G. F., McKenzie, F. E., Foy, B. D., Bøgh, C., and Beier, J. C. (2001). The availability of potential hosts as a determinant of feeding behaviours and malaria transmission by African mosquito populations. *Trans. R. Soc. Trop. Med. Hyg.* 95, 469–476. doi: 10.1016/s0035-9203(01)90005-7

Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F., Bahram, M., et al. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271–5277. doi: 10.1111/mec.12481

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Li, Y., Kamara, F., Zhou, G., Puthiyakunnon, S., Li, C., Liu, Y., et al. (2014). Urbanization increases *Aedes albopictus* larval habitats and accelerates mosquito development and survivorship. *PLoS Negl. TropDis.* 8:e3301. doi: 10.1371/journal.pntd.0003301

Lin, Y. S., Richard, M., Lehtinen, and Kam, Y. C. (2008). Time- and context-dependent oviposition site selection of a phytotelm-breeding frog in relation to habitat characteristics and conspecific cues. *Herpetologica* 64, 413–421. doi: 10.2307/27669731

Liu, H., Liu, L., Cheng, P., Yang, L., Chen, J., Lu, Y., et al. (2020). Bionomics and insecticide resistance of *Aedes albopictus* in Shandong, a high latitude and high-risk dengue transmission area in China. *Parasit Vectors* 13:11. doi: 10.1186/s13071-020-3880-2

Liu, H., Xie, L., Cheng, P., Xu, J., Huang, X., Wang, H., et al. (2019). Trends in insecticide resistance in Culex pipiens pallens over 20 years in Shandong, China. *Parasit Vectors* 12:167. doi: 10.1186/s13071-019-3416-9

Liu, H. M., Yang, P. P., Cheng, P., Wang, H. F., Liu, L. J., Huang, X., et al. (2015). Resistance Level of mosquito species (Diptera: Culicidae) from shandong province, China. *Int. J. Insect Sci.* 7, 47–52. doi: 10.4137/IJIS.S24232

Minakawa, N., Sonye, G., and Yan, G. (2005). Relationships between occurrence of *Anopheles gambiae* s.l. (Diptera: Culicidae) and size and stability of larval habitats'. *J. Med. Entomol.* 42, 295–300. doi: 10.1093/jmedent/42.3.295

Molineaux, L. (1997). Malaria and mortality: some epidemiological considerations. *Ann. Trop. Med. Parasitol.* 91, 811–825. doi: 10.1080/00034989760572

Mukhopadhyay, P., and Chatterjee, S. (2016). Characterization and control of symbiotic Bacillus cereus isolated from the mid gut of *Anopheles subpictus* Grassi. *J. Parasit Dis.* 40, 1414–1421. doi: 10.1007/s12639-015-0704-y

Nambunga, I. H., Ngowo, H. S., Mapua, S. A., Hape, E. E., Msugupakulya, B. J., Msaky, D. S., et al. (2020). Aquatic habitats of the malaria vector *Anopheles funestus* in rural south-eastern Tanzania. *Malar J.* 19:219. doi: 10.1186/s12936-020-03295-5

Ndaruga, A. M., George, G., Ndiritu, Nathan, N., Gichuki, and Wamicha, W. N. (2004). Impact of water quality on macroinvertebrate assemblages along a tropical stream in Kenya. *African J. Ecol.* 42, 208–216. doi: 10.1111/j.1365-2028.2004.00516.x

Nikookar, S. H., Fazeli-Dinan, M., Azari-Hamidian, S., Mousavinasab, S. N., Aarabi, M., Ziapour, S. P., et al. (2017). Correlation between mosquito larval density and their habitat physicochemical characteristics in mazandaran province, northern Iran. *PLoS Negl. Trop. Dis.* 11:e0005835. doi: 10.1371/journal.pntd.0005835

Noori, N., Lockaby, B. G., and Kalin, L. (2016). Larval development of culex quinquefasciatus in water with low to moderate. *J. Vector Ecol.* 40, 208–220. doi: 10.1111/jvec.12156

Norris, D. (2004). Mosquito-borne diseases as a consequence of land use change. *Ecohealth* 1, 19–24. doi: 10.1007/s10393-004-0008-7

Petric, D., Bellini, R., Scholte, E. J., Rakotoarivony, L. M., and Schaffner, F. (2014). Monitoring population and environmental parameters of invasive mosquito species in Europe. *Parasit Vectors* 7:187. doi: 10.1186/1756-3305-7-187

Qing, W., Zhijing, X., Guangfu, Y., Fengxia, M., Qiyong, L., Zhong, Z., et al. (2020). Variation in the microbiota across different developmental stages of *Aedes albopictus* is affected by ampicillin exposure. *Microbiologyopen* 9, 1162–1174. doi: 10.1002/mbo3.1026

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Service, M. W. (1991). Agricultural development and arthropod-borne diseases: a review. *Rev. Saú Pública* 25, 165–178. doi: 10.1590/s0034-89101991000300002

Shi, Q., Song, X., Lv, Y., Huang, X., Kou, J., Wang, H. W., et al. (2019). Potential risks associated with Japanese encephalitis prevalence in shandong province. China. *Vector Borne Zoonotic Dis.* 19, 640–645. doi: 10.1089/vbz.2018.2416

Thavara, U., Tawatsin, A., and Chompoosri, J. (2004). Evaluation of attractants and egg-laying substrate preference for oviposition by *Aedes albopictus* (Diptera: Culicidae). *J. Vector Ecol.* 29, 66–72.

Tusting, L. S., Thwing, J., Sinclair, D., Fillinger, U., Gimnig, J., Bonner, K., et al. (2013). Mosquito larval source management for controlling malaria. *Cochrane Database Syst. Rev.* 8:CD008923. doi: 10.1002/14651858.CD008923.pub2

Varela, J. N., and Yadav, V. G. (2020). A Pichia biosensor for high-throughput analyses of compounds that can influence mosquito behavior. *Microbiologyopen* 10:e1139. doi: 10.1002/mbo3.1139

Villarreal-Treviño, C. R., Penilla-Navarro, R. P., Vázquez-Martínez, M. G., Moo-Llanes, D. A., Moo-Llanes, D. A., Ríos-Delgado, J. C., et al. (2015). Larval habitat characterization of *Anopheles darlingi* from its northernmost geographical distribution in Chiapas, Mexico. *Malar J.* 14:517. doi: 10.1186/s12936-015-1037-0

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, Y., Cheng, P., Jiao, B., Song, X., Wang, H., Wang, H., et al. (2020). Investigation of mosquito larval habitats and insecticide resistance in an area with a high incidence of mosquito-borne diseases in Jining. shandong province. *PLoS One* 15:e0229764. doi: 10.1371/journal.pone.0229764

Wilkerson, R. C., Strickman, D., and Litwak, T. R. (1990). Illustrated key to the female anopheline mosquitoes of Central America and Mexico. *J. Am. Mosq. Control Assoc.* 6, 7–34.

World Health Organization (1983). Integrated vector control: seventh report of the WHO expert committee on vector biology and control [meeting held in Geneva from 7 to 13 December 1982]. *World Health Organ. Tech. Rep. Ser.* 767:72.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 16 April 2021
doi: 10.3389/fevo.2021.660941

WingBank: A Wing Image Database of Mosquitoes

Flávia Virginio[1,2†], Victor Domingues[3†], Letícia Cecília Gomes da Silva[4], Lucas Andrade[2], Kelly Rosa Braghetto[5‡] and Lincoln Suesdek[4‡]*

[1] Laboratório de Coleções Zoológicas, Instituto Butantan, São Paulo, Brazil, [2] Fundação Butantan, São Paulo, Brazil, [3] Graduando em Ciências da Computação, Faculdade Anhembi Morumbi, São Paulo, Brazil, [4] Laboratório de Parasitologia, Instituto Butantan, Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, Brazil, [5] Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil

Mosquito-borne diseases affect millions of people and cause thousands of deaths yearly. Vaccines have been hitherto insufficient to mitigate them, which makes mosquito control the most viable approach. But vector control depends on correct species identification and geographical assignment, and the taxonomic characters of mosquitoes are often inconspicuous to non-taxonomists, which are restricted to a life stage and/or even damaged. Thus, geometric morphometry, a low cost and precise technique that has proven to be efficient for identifying subtle morphological dissimilarities, may contribute to the resolution of these types of problems. We have been applying this technique for more than 10 years and have accumulated thousands of wing images with their metadata. Therefore, the aims of this work were to develop a prototype of a platform for the storage of biological data related to wing morphometry, by means of a relational database and a web system named "WingBank." In order to build the WingBank prototype, a multidisciplinary team performed a gathering of requirements, modeled and designed the relational database, and implemented a web platform. WingBank was designed to enforce data completeness, to ease data query, to leverage meta-studies, and to support applications of automatic identification of mosquitoes. Currently, the database of the WingBank contains data referring to 77 species belonging to 15 genera of Culicidae. From the 13,287 wing records currently cataloged in the database, 2,138 were already made available for use by third parties. As far as we know, this is the largest database of Culicidae wings of the world.

Keywords: relational database, open source, vector-borne disease, public health, medical entomology, geometric morphometric approach, integrative taxonomic approach

OPEN ACCESS

Edited by:
Jader Oliveira,
University of São Paulo, Brazil

Reviewed by:
Tanawat Chaiphongpachara,
Suan Sunandha Rajabhat University, Thailand
Rodrigo Gurgel-Gonçalves,
University of Brasilia, Brazil

****Correspondence:***
Flávia Virginio
flavia.virginio@butantan.gov.br

† These authors have contributed equally to this work and share first authorship

‡ These authors have contributed equally to this work and share last authorship

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *30 January 2021*
Accepted: *22 March 2021*
Published: *16 April 2021*

Citation:
Virginio F, Domingues V, da Silva LCG, Andrade L, Braghetto KR and Suesdek L (2021) WingBank: A Wing Image Database of Mosquitoes. Front. Ecol. Evol. 9:660941. doi: 10.3389/fevo.2021.660941

INTRODUCTION

Diseases whose etiological agents are dispersed by vectors, such as mosquitoes, have been a major public health problem worldwide for years. These diseases, between the 17th and 20th centuries, caused the death of more people than all other causes combined (Gubler, 1991) and interfered with the economic development of large areas around the world (Philip and Rozenboom, 1973; Calmon, 1975; Gubler, 1991). Amongst insect Families popularly known as mosquitoes, the Culicidae family represents an important group composed of several vector species of disease-causing pathogens, including viruses, worms and protozoans. All these pathogens cause important but neglected

Tropical Diseases, such as lymphatic filariasis, which is one of the leading causes of global disability. This disease affects around 120 million people in tropical and subtropical areas of the world (World Health Organization, 2020a). In 2019, it led more than 400,000 people to death, 67% of them children (World Health Organization, 2020b).

However, despite the importance of mosquitoes, their taxonomic identification based only on traditional morphology is quite complex and often difficult to conclude, i.e., Subgenus Nyssorhynchus, which is difficult to identify using only external morphology (Sallum et al., 2010). In addition, traditional taxonomy has been in crisis for decades (Godfray, 2002), with few professional specialists in the field remaining worldwide. Furthermore, frequently mosquito samples are morphologically damaged by being sampled using CDC light, BG or other mosquito traps. Even taxonomists have difficulty identifying them. Thus, the arising of new technologies has complemented the traditional technique in several studies, such as the case of geometric morphometry (GM), which is based on the fusion between geometry, biology (Bookstein, 1982) and statistics (Monteiro and Reis, 1999).

The GM allows the multivariate study of the shape of biological structures in two or three spatial dimensions. This technique allows various statistical biometric assessments, graphic representation of shape and size, preserving the physical integrity of the shape and preventing a collapse in linear measurements that do not represent the structure as a whole (Richtsmeier et al., 2002). Based on multivariate character analysis, which allows the simultaneous comparison of different characteristics in a complex body structure (Rohlf, 1993; Monteiro and Reis, 1999), this technique has been used extensively to solve different biological problems. In some cases, it is useful when applied alone, while in others its combination with any other integrative taxonomy technique can increase its accuracy (Schlick-Steiner et al., 2010; Garros and Dujardin, 2013; Lorenz et al., 2017).

In mosquitoes, the most used structure is the wings (Ruangsittichai et al., 2011; Jaramillo et al., 2015; Sumruayphol et al., 2016; Wilke et al., 2016; Chaiphongpachara and Laojun, 2019; Chaiphongpachara et al., 2019; Sauer et al., 2020; Souza et al., 2020), mainly due to its two-dimensionality, which increases precision and repeatability even when they are assembled and digitized by different operators (Lorenz and Suesdek, 2013). In general, the shape of the wing is quite informative (Klingenberg, 2010), on both hereditary, geographic, and evolutionary issues. Some studies have indicated that in Diptera, the wing shape is inheritable (Bitner-Mathé and Klaczko, 1999), has polygenic determination and is minimally influenced by epigenetic factors (Jirakanjanakit et al., 2007; Morales-Vargas et al., 2010). Although reports indicate that the shape of the wing is influenced by temperature differences and eco-geographic factors (Aytekin et al., 2009; Gómez et al., 2014), this influence is minimal (Morales Vargas et al., 2013). Therefore, the shape of the wing is presented as a more stable character than the size of the wing, which is strongly influenced by the environment (climatic and environmental factors such as larval density in the same breeding place, availability of food, temperature, etc.) (Morales Vargas et al., 2013; Gómez et al., 2014), and variations in this character should be interpreted with caution.

Any study of the Life Science field needs reliable taxonomic identification. However, every year there is a more frequent need for helpful and innovative techniques since the traditional taxonomy has been increasingly restricted to a specialized group. Thus, automated identification methods which extract and analyse informative features of species in images have emerged as a helpful tool, and consequently, the number of large databases (DBs) hosting this data has also grown. Although DBs of winged insect images are mentioned in thematic reviews in which researchers emphasize the importance of creating morphometric DBs for culicids (Dujardin, 2008; Lorenz et al., 2017) and other vectors, efforts toward building those specific to mosquitoes are still incipient. This gap has hindered classification tests or biometric identification, as well as studies of long-term biological inference in insects (Sonnenschein et al., 2015).

According to the Registry of Research Data Repositories–re3data[1], there are about 1,423 databases registered in the Life Science area, and 1,311 in Natural Sciences. These numbers together correspond to approximately six times the number of records in the Engineering area (with the lowest number of records, 482)—values updated in January 2021. This can be explained by the large growth in data production in recent years, mainly the Big Data common in the area of Molecular Biology, reflecting the massive production of gene sequences (Stephens et al., 2015). However, despite the expressive number of DBs registered in re3data, Brazil contributes only with 11 records, which certainly does not reflect the amount of DBs created and used in the country.

The myth of the Tower of Babel besides serving as a metaphor for an existing problem in the systematic of insects (Caterino et al., 2000), can also be associated with the extensive "independent production" of programs and databases related to gene annotations (Drăghici et al., 2006), for example. Thus, to avoid the same inconsistency of the aforementioned metaphor, morphometric DBs created in the Culicidology area must be able to connect to other DBs, whether they are about georeferenced, genetic annotation or bibliographic data.

The wing GM, in addition to requiring computational implementations, requires a robust DB that allows the validation of this application through tests. Furthermore, according to Garros and Dujardin (2013), for a morphometric DB to have taxonomic utility, it must share, at least, the computed cartesian coordinates and/or the images of the structures. As far as we know, there are only two repositories that share useful images for alar GM on insects, ApiClass[2] and XYOM (formerly CLIC bank;[3]). ApiClass is a specialized online system for identifying bee subspecies belonging to the species *Apis mellifera* Linnaeus, 1758, based on the wings, and which uses a Relational DB (RDB) with 5,763 images maintained

[1]https://www.re3data.org/

[2]http://apiclass.mnhn.fr

[3]http://xyom-clic.eu/clic-bank/

with the MySQL database management system (DBMS). 3 (Dujardin, 2012).

In this work, we present WingBank[4] a data platform developed to be a pioneer in the storage and sharing of wing images, cartesian coordinates of the images and other morphological information, and data of interest for integrative taxonomy, such as spatio-temporal data. In addition, it was conceived to serve as a basis for machine learning in the GM field, managing data from dozens of species and thousands of specimens of mosquitoes. WingBank architecture is composed of an RDB, managed through the DBMS Microsoft SQL Server (MS-SQL), and a prototype of a web system that fosters research by allowing users to retrieve the data they need and also contribute with new data.

This article describes: (a) the modeling and implementation of the WingBank's RDB, aimed at micro and macroevolutionary studies that use the wing geometric morphometrics technique; and (b) the development of the web system for managing biological data related to wing morphometry, which is the user interface for this RDB.

MATERIALS AND METHODS

Target Audience

The WingBank system is intended for undergraduate and graduate students, scientific researchers, and health professionals, such as entomological and epidemiological supervisors, and other interested parties, who need to download or upload images and data of wing geometric morphometry of mosquitoes around the world.

Taxonomic Classification

The taxonomic classification adopted in this work followed the one used by Walter Reed Biosystematics Unit (WRBU), Entomology Division, Walter Reed Research Institute (Walter Reed Army Institute of Research, WRAIR) in its Systematic Catalog of Culicidae[5] and presented by Ralph Harbach (Elmasri and Navathe, 2011) in his Taxonomic Mosquito Inventory[6], following the guidelines of Wiley and Liebermann (2011); Vences et al. (2013), and Wilkerson et al. (2015) for the classification of the Aedini Tribe. The informal taxonomic categories (Series, Group, Subgroup, Complex) were not adopted in the modeling and implementation of DB. Only the definitions a) *sensu lato* ("*s.l.*") with the records of species belonging to the species complex, b) "*sp*." with the records that do not have specific identification, and c) *sensu stricto* ("*s.s.*") for records identified at a specific level were adopted.

Data Organization

As a basis for the creation of the WingBank DB, around 14 thousand images of mosquito wings were used, deposited in the Entomological Collection under the curatorship of Dr Flávia Virginio, Lead Investigator of Medical Entomology Study Group, located at the Laboratório de Coleções Zoológicas (LCZ) of the Instituto Butantan (IB). Voucher specimens (wing donors) are also deposited in that collection. For this purpose, the wings were extracted from mosquitoes, mounted between a slide and coverslip as described by Virginio et al. (2015). Each wing, when deposited in the Collection, received an identification tag containing a sequential code in accordance with the policies adopted by LCZ, and with additional information, such as gender, wing side and number of the individual (e.g., FD001-IBSP-Ent 1). All wings were photographed and stored in digital format together with their morphometric and biological, taxonomic, geographical, space-time, environmental, climate and human resources information.

WingBank Creation: Step by Step

The main phases for the creation of a DB are gathering of requirements, conceptual database design, choice of a DBMS, data model mapping (also called logical database design), physical database design and finally, database system implementation and tuning (Elmasri and Navathe, 2011). Thus, for the creation of the WingBank DB, firstly the requirements specification and analysis were carried out with the researchers who own the images, based on the data collection forms and guidelines proposed by Gaffigan and Pecor (Gaffigan and Pecor, 1997) and Foley et al. (2011). This step basically consisted of collecting information regarding the interests of DB users and the program applications that will interact with it, based on occasional meetings. From this, it was possible to identify the main requirements for the elaboration of the WingBank, characterize the types of data desirable for storage in WingBank and the query and data maintenance functionalities to be offered to users through the web system.

After gathering the requirements, in the second phase, the construction of the Conceptual Database Model was elaborated. Based on the Entity-Relationship Model (ERM) (Chen, 1977), an Entity-Relationship (ER) Diagram was built. This model was chosen because it assists in the modeling of relational DBs, is very expressive and easy to understand, and supplies the need for abstraction in this phase of modeling. Subsequently, in the third phase, using the software Visual Paradigm Community 14.1[7], the ER diagram was mapped into a relational diagram. In the relational model, data is organized "in tables, nothing but tables" (Date, 2004), in other words, the types of entities are transformed into tables, the attributes in columns of the respective tables, the multivalued attributes in different tables, the relationships in foreign keys and (possibly) new tables. In the diagram of the relational scheme, the crow's foot notation (Everest, 1976) was used to indicate the cardinalities of the relationships established through foreign keys. This notation was created for conceptual data modeling, but it is often used (combined with other notations) to graphically represent cardinality and participation constraints on relationships expressed through foreign keys in a relational scheme. In the created relational scheme, the domain of each attribute was also defined.

[4]https://wingbank.butantan.gov.br/

[5]http://www.mosquitocatalog.org/

[6]http://mosquitotaxonomic-inventory.info/

[7]https://www.visual-paradigm.com/

At the fourth phase, it is delimiting important specifications related to stored databases, such as file storage structures and indexes. The scheme refinement is an optional step in the logical design, but it is commonly used. Through this, it is possible to identify potential problems in the created scheme and apply techniques to improve it. In the physical model of the WingBank DB, some final specifications were implemented, especially those related to storage and access to the DB. In this phase, some indexes were created, both to ensure the coherence of the connections between the data, and to improve the performance of future operations on the data (Elmasri and Navathe, 2011).

And finally, the last phase is about database and application programs implementation and testing, a continuous activity known as database tuning. To obtain useful knowledge from the data, it is necessary that it is clear. However, many databases are composed of redundant and inconsistent data, missing values, as well as values and fields that are not logically related to each other stored in the same table (Parsaye and Chignell, 1993; Savasere et al., 1995; Adriaans and Zantinge, 1996; Fayyad et al., 1996). To avoid mistakes before feeding the WingBank DB, part of the data cleaning was done manually, and the other part was performed together with the initial process of data entering, through SQL scripts, in a careful process of inspection, cleaning, standardization, completion and transformation of data. These steps refer to the Knowledge Discovery in Databases (KDD) methodology (Fayyad et al., 1996), except for the data mining phase.

As the raw data were not fully standardized, there was a small loss of information in the process. For example, the following were not uploaded to the database: (a) photographic records without information about gender, (b) geographic information of cases in which there was no relationship between individuals and localities, although there was information on the location of the collection. Data from other specific cases were kept, such as (c) photographic records with an unnoticed side, which did not compromise the integrity of the bank, (d) damaged wings, but with some preserved points.

There are different design patterns for the development of a system with access to the DB. In the WingBank web system, Domain-Driven Design (DDD) and Model View Controller (MVC) standards were used. The first is a development approach that focuses on understanding business rules and how they should be reflected in the system code and domain model (Evans, 2003). In addition to this technique being linked to the good practices of Object-Oriented Programming, it is a way of organizing the code so that the business rules are in delimited contexts and decoupled from the database. The DDD was used in the architecture of practically the entire application, from infrastructure to data access.

For the communication between the system model layer and the DB, the Object-Relational Mapping was used. It's a technique that consists of representing a program object in a relational way so that it can be persisted in an RDB and, when necessary, be retrieved and created as an object in main memory again, without loss of information in this process.

The MVC standard was used only for the architecture of the web interface client application. In this pattern, the model layer refers to the representation of the data on which the system operates. The vision layer is the presentation of data and processing logic to end-users, even allowing their interaction with the system. Finally, the controller responds to events, manages access to the model and view, and coordinates the flow of data between them.

The WingBank web system provides features for searching, visualizing, and downloading data that are related to a scientific publication. As this data is public, that is, it can be accessed by anyone, there is no functionality to control user access in the developed prototype. The behavior of the system was described in an Activity Diagram in the Unified Modeling Language (UML).

RESULTS

The WingBank RDB contains morphological information and their respective ecological and space-time metadata, exclusively from mosquitoes. This DB is a prototype under test and is available at https://wingbank.butantan.gov.br/. It is currently open for searching and downloading *via* the website and sending data *via* email WingBank@butantan.gov.br. This database can also be used by users as a faithful depository of images and data of geometric morphometry, as it generates unique identifiers (see Standardization of image labeling) that can be cited in scientific publications.

This DB was inspired by the actions promoted by the National Center for Biotechnology Information (NCBI), one of the Institutes that has provided the most biological databases to date (Agarwala et al., 2015). At this moment, in WingBank, 13,287 wing records are registered, of which 12,939 are already linked to images (right and/or left wings) of mosquitoes belonging to the Culicidae family. Currently, there are 13 scientific publications registered on WingBank, of which 8 are already linked to 2,138 images which, due to this relationship with one or more publications, are already available for use by third parties (**Table 1**). Of these wing images available for use by third parties, 584 are left wings, 1,476 right wings and 78 with unknown sides. It is worth mentioning, as the wing images available for other users until now represent samples of totalities located in a unique time-space, we decided that instead of linking only the files used on each study, we would link all images of those specimens located in those unambiguous times and spaces. To date, WingBank stores data collected from 1998 to 2016, in 10 different Brazilian states. Forty-three people, linked to 19 different National and International Teaching and Research Institutions, have so far contributed to WingBank with the collection of information and specimens, mounting of slides, and/or photographic and morphometric records.

The main types of data (**Figure 1**) of this DB were from Field and/or Colony, and/or Donation; Wing images, which can be right and/or left wings and can be related to some Scientific Publication and also with landmark Notes; Taxonomic Classification (Family, Subfamily, Tribe, Genus, Subgenus, Species, Subspecies), Classification Method, which is possible to report which method was used (Molecular or Morphological);

TABLE 1 | List of publications and data available on WingBank for use by third parties.

Publication Id	Publication title	Publication doi	Number of records
1	Temporal variation of wing geometry in *Aedes albopictus*	http://doi.org/10.15t90/S0074-02762012000800011	180
2	Wing diagnostic characters for *Culex quinquefasciatus* and *Culex nigripalpus* (Diptera, Culicidae)	http://doi.org/10.1590/S0085-56262011000100022	244
3	Wing sexual dimorphism of pathogen-vector culicids	http://doi.org/10.1186/s13071-015-0769-6	439
4	Comparison of wing geometry data and genetic data for assessing the population structure of *Aedes aegypti*	http://doi.org/10.1016/j.meegid.2011.11.013	378
5	High morphological and genetic variabilities of Ochlerotatus scapularis, a potential vector of filarias and arboviruses	http://doi.org/10.1186/s13071-015-0740-6	116
6	Microevolution of *Aedes aegypti*	http://doi.org/10.1371/journal.pone.0137851	219
7	Morphometrical diagnosis of the malaria vectors *Anopheles cruzii, An. homunculus* and *An. Bellator*	http://doi.org/10.1186/1756-3305-5-257	–
8	Altitudinal population structure and microevolution of the malaria vector *Anopheles cruzii* (Diptera: Culicidae)	http://doi.org/10.1186/s13071-014-0581-8	274
9	Morphogenetic characterisation, date of divergence, and evolutionary relationships of malaria vectors *Anopheles cruzii* and *Anopheles homunculus*	http://doi.org/10.1016/j.meegid.2015.08.011	–
10	Assessment of the correlation between wing size and body weight in captive *Culex quinquefasciatus*.	http://doi.org/10.1590/0037-8682-0039-2016	288
11	Artificial Neural Network Applied as a Methodology of Mosquito Species Identification.	https://doi.org/10.1016/j.actatropica.2015.09.011	–
12	Evaluation of Chemical Preparation on Insect Wing Shape for Geometric Morphometrics.	https://doi.org/10.4269/ajtmh.13-0359	–
13	Wing sexual dimorphism in *Aedes scapularis* (Diptera: Culicidae)	http://doi.org/10.1590/S1676-06032011000200016	–
			2,138

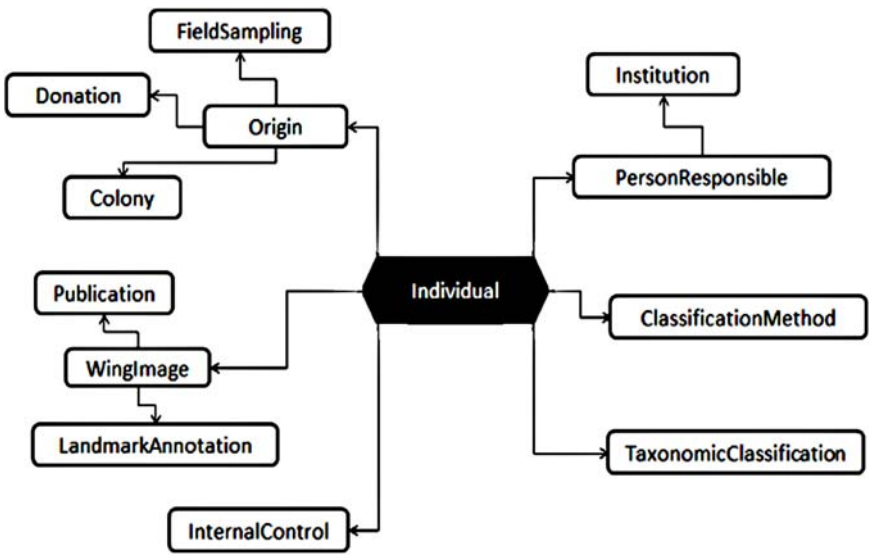

FIGURE 1 | WingBank database conceptual map.

and the Person Responsible for identifying each individual. It is important to mention that the relationship between Wing Image and a Scientific Publication is a requirement for an individual's data to be publicly available.

Conceptual Database Design

A conceptual database model must be concise, descriptive, and present the main requirements of the modeled domain. It can be expressed in different ways. The most common and that was used in this DB, is through the ER diagram. Elmasri and Navathe's graphical notation (Elmasri and Navathe, 2011) was used, which is based on the notation of Peter Chen, creator of ERM (Codd, 1970). An ER diagram is structured in (see **Figure 2**): (A) types of entities (rectangles), which represent the abstraction of a real-world object in the DB, such as, Individual, Image and Taxonomic Identification;

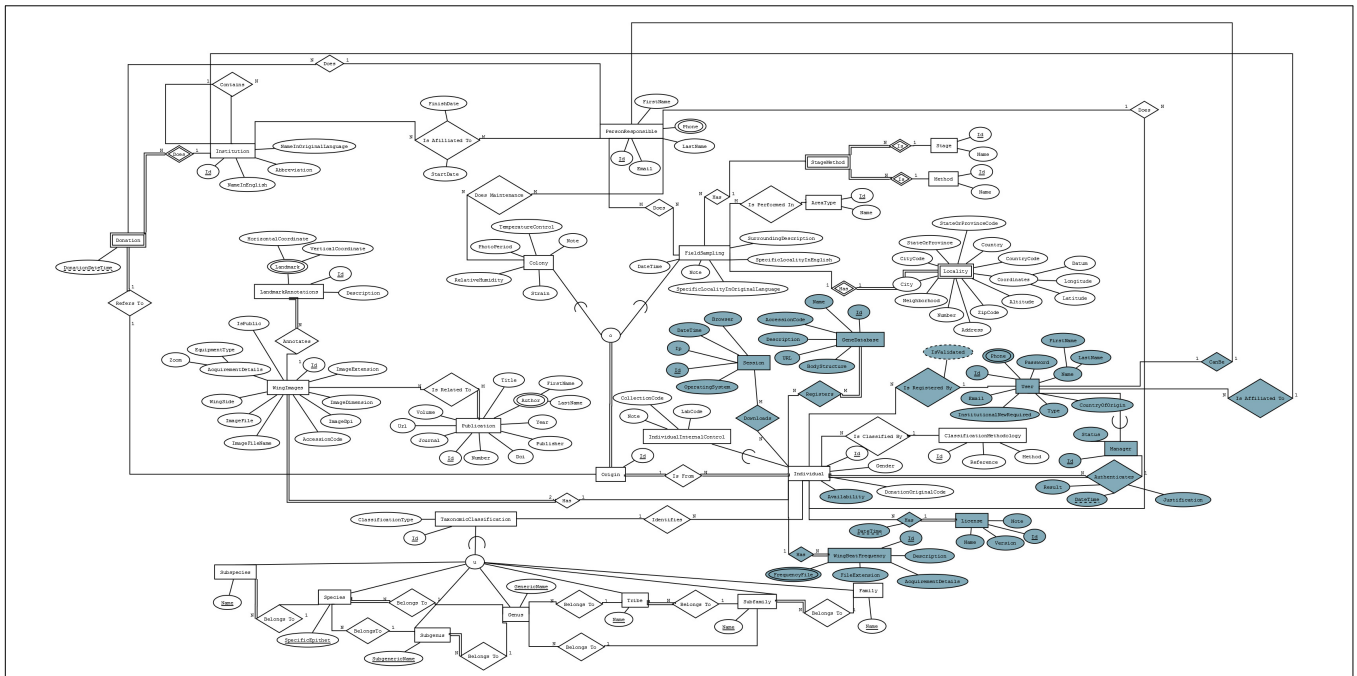

FIGURE 2 | Representation of the Conceptual Database Design through an Enhanced Entity-Relationship (EER) diagram. Rectangles: types of entities, which represent the abstraction of a real-world object. Ellipses: Attributes, which are details, properties that characterize a type of entity. Diamonds: types of relationships, which represent associations between entities, accompanied by their participation restrictions and cardinalities. Filled geometric forms: Entity, Attributes and Relationships which were planned but not yet implemented (Phase II).

(B) types of relationships (diamond-shaped boxes), which are connected by lines to the rectangular boxes representing the participating entity types, followed by the cardinality ratio and participation constraint of each relationship type, such as: [individual] 1 < has > 2 [image], [individual] 1 < is > 1 [species], that is, each individual can have up to 2 images, and each individual corresponds to only one species; (c) attributes (ellipses), which are details, properties that characterize a type of entity, such as gender of the individual and dimension of the image.

Entities are typified as strong (defined in the diagram by a simple outline, e.g., Publication) or weak (double outline, e.g., Donation). Attributes can be compound (subdivided into more than one component attribute, e.g., Coordinates); multivalued (with double contour, e.g., Phone in PersonResponsible); complex (multivalued and compound at the same time; e.g., Author in Publication); derivative (with dashed outline; e.g., Is validated). When the participation constraints of an entity in a relationship are partial, the line between them is simple; when the participation is total, the line is double (Elmasri and Navathe, 2011).

For a better understanding of the diagram that will be presented below, it is worth mentioning that this database has specializations. Specialization is the definition of subsets of entities (subclasses) of an entity type (in this case, considered a superclass), in which the subclasses have their own attributes and/or relationships and, at the same time, inherit the attributes (including, the keys) of the superclass. Therefore, an entity in a subclass is always an entity in its superclass. Specialization can be of the overlapping type (denoted by the symbol "o" inside a circle) or disjunctional type (symbol "d" inside a circle). In the disjunctional type, a superclass entity can be at most in one of the subclasses of the specialization; in the overlapping type, an entity of the superclass can be in several subclasses. A category (denoted by the symbol "u" in a circle) can be seen as a set of entities of different types that do not necessarily have attributes in common between them (Elmasri and Navathe, 2011).

In general, the symbol ∩ is added to the relationship in which you want to represent the specialization or category. The opening of this symbol is directed at the entity on which the restriction is being applied. For example, if the intention is to represent that an Individual can originate from a Field Collection, a Laboratory Colony, or that he can be collected in the Field, but then be Colonized/Maintained in the laboratory, this can be represented by directing the opening of the symbol to its Origin (which is the superclass of an overlapping specialization) as shown in **Figure 2**.

Key attributes of entity types (solely identifying their entities) are underlined (e.g., Id in most entities), while those that are partial keys are denoted by a dashed line (as in DonationDateTime). The WingBank is represented by an Enhanced Entity-Relationship (EER) diagram, as some specific restrictions have been applied, such as the total overlapping specialization of Origin (Colony and FieldSampling), the Individual specialization (IndividualInternalControl); and the Taxonomic Classification category, which is a subset of the union of entity classes: family, Subfamily, Tribe, Genus, Subgenus, Species and Subspecies. The weak entity type StageMethod has

no attributes (something unusual) because it was created to represent a binary relationship between Stage and Method and, therefore, to make it possible its association with FieldSampling. Weak entity types, like Locality, did not receive a partial key, as they participate in a 1:1 relationship with Sampling, that is, if each Sampling can have only 1 Locality, then the Locality key can be the Sampling key itself.

The Conceptual Model represents the following situation: An individual (female or male) is taxonomically classified using a method and based on a specific reference. Taxonomic information for each individual can be presented at different levels: Family, Subfamily, Tribe, Gender, Species and Subspecies. Each individual, when registered in the system, automatically receives a unique and sequential code ("WingBankCode"), which will be with it as part of the DB. The specimen may originate from a field data collection, from a Laboratory Colony, or even being collected and then colonized. In either case, it may also have come from a Donation. If the individual is the result of a collection, it must have information regarding the collected stage, collection method, date of collection, general and specific location, information about the surroundings of the collection area, type of area, etc. If the individual comes from a laboratory colony, it must have information regarding laboratory conditions, such as temperature, relative humidity, etc. If the individual originated from a donation and has a "DonationOriginalCode," this code is kept in the records. If this individual is an internal property of the Instituto Butantan, it will have a "LabCode" and if it is already deposited at the Entomological Collection of LCZ, it will receive a "CollectionCode." In addition, there are people responsible for each step individuals go through, such as collection, colonization, donation, and taxonomic classification. These people are linked to one or more institutions.

Considering the importance of sharing computed coordinates in morphometric DBs, as mentioned by Garros and Dujardin (2013), WingBank DB modeling included the types of entities LandmarkAnnotation and Landmark, which make it possible to store raw computed coordinates (Landmarks and/or Semilandmarks of the wings). Some types of entities have attributes with predefined domains (**Table 2**), as the case of the type of entity AreaType, in which the values of the Name attribute were restricted to Park, which represents Urban Park or Linear Park; Urban, which refers to areas distributed by cities; Preserved, which represents Forested Areas, Environmental Protection Areas, Conservation Units, or Permanent Preservation Areas; and Rural, which represents Agriculture and Livestock.

Other types of entities such as Stage and Method, at first, received restricted values for the "Name" attribute, such as: Immature, Egg, Larva (1st stage), Larva (2nd stage), Larva (3rd stage), Larva (4th stage), Larva, Pupa, Adult, Immature/Adult and Manual Sampling, Manual Aspiration, Metal/Plastic Scoop, Metal Ladle, Fine Mesh, Pasteur Pipette/Dropper, Manual suction pump, Entomological Manual Aspirator, Entomological Automatic Aspirator, Trap, Egg Trap, Larva Trap, Adult Trap, CDC, Shannon, New Jersey, MoquitoMagnet, Manual Sampling/Manual Aspiration, respectively. Finally, the entity type ClassificationMethodology received only two different values for the ClassificationMethod attribute: Morphological or Molecular.

TABLE 2 | Description of the domains of the Name attribute of the AreaType entity.

Domain value	Description
Park	Urban park
	Linear park
Urban	City
Preserved	Forested areas
	Environmental protection areas
	Conservation units
	Permanent preservation areas
Rural	Agriculture
	Livestock

As the data currently stored on WingBank has been collected in the past and with less homogeneity, part of the attributes may receive the special value NULL, since the values of these attributes for some of the previously collected records are unknown. It is worth mentioning that this does not mean that these fields will be permanently null. The new records are expected to be more complete. Furthermore, in order to complement the implicit and explicit restrictions expressed in the DB scheme, the semantic restrictions that apply to the data have been described in a data dictionary (**Supplementary Table 1**).

Logical Database Design

The relational data model was created by Edgar F. Codd (1970) and was considered a revolutionary idea in the 1970s (Seltzer, 2008). The main motivation was an observation made about the workflow of programmers from the company IBM (International Business Machines) where Codd worked. He needed to rewrite a large number of application programs manually whenever the content or physical organization of a database changed (IBM, 2011). Relational DBs were originally proposed to separate the physical storage of data from its conceptual representations and provide mathematical foundations for the representation and search for data (see **Figure 3**).

Standardization of Image Labeling

The images of each wing stored in the WingBank are fundamental pieces for this DB. With its implementation, they were functionally linked to other information which enables data validation and recovery, considering that MorphoJ (Klingenberg, 2011), one of the programs commonly used in GM analysis, has a feature called "Extract New Classifier from ID String." This feature allows the classification of samples to be analyzed based on each character present in the file label (input). A nomenclature standard for wing image files is proposed here, which applies to all files submitted to this DB.

Each record on the WingBank (image of the wing + other metadata), when registered in the system, automatically receives a unique and sequential access code ("WingBankCode"), which will be with it as part of the DB, and will facilitate the creation of a reference in a scientific publication, for example. This identifier

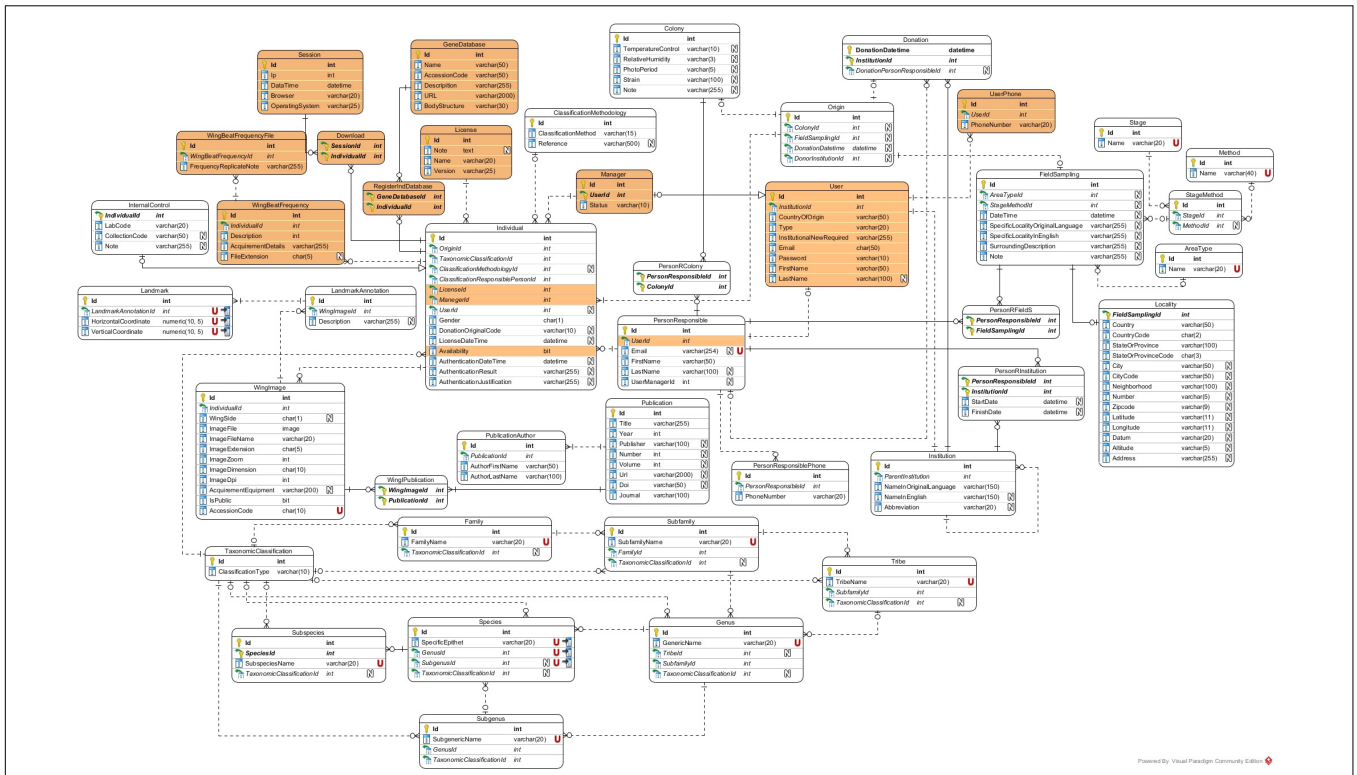

FIGURE 3 | Representation of the Logical Database Design by mapping a conceptual schema in the EER model into a relational representation. In this model, the types of entities are transformed into tables, the attributes in columns of the respective tables, the multivalued attributes in other tables, the relationships in foreign keys and (possibly) new tables. This diagram also shows the crow's foot notation which indicates the cardinalities of the relationships established by means of foreign keys. Filled boxes: Entity, Attributes and Relationships which were planned but not yet implemented (Phase II).

consists of an alphanumeric string containing a letter (prefix) representing the gender (F or M) of the individual, a sequence of 8 non-variable fixed digits, filled with leading zeros, a letter (suffix), which represents the side (R or L, right or left); and finally, when the information related to the side is non-existent, the insertion of the letter U (Unknown) is adopted, as in "F00000001R," "F00000001L," and "F00000001U," respectively.

Implementation of the Web System

Figure 4 is a diagram in UML that describes the behavior of the web search system developed for WingBank. This Activity Diagram represents the flows of the system processes, whether they are business processes or internal operations. This type of diagram also shows deviations for alternative processes, such as displaying an error message. It is then possible to observe the lanes which were used to organize the actors and components of the system at different stages of the flow, to make it more readable and show the importance of each actor in the flow. The actors of the flow arranged in the lanes are User and Google API. The components are Web and Server. As shown in the diagram, the search process always starts and ends with the user (User).

Currently, on the WingBank platform, there are two types of searches (**Figure 5**) the simple one, where the user performs the search by keywords, and (**Figure 6**) the advanced one, which allows the user a combined search in the most diverse ways. The user can filter the results by characteristics related to the collection (Sampling), that is, space-time information such as Specific Locality, which refers to specifications such as breeding type; Locality referring to the most generalized location such as an address, City, State/Province, Country or Geographic coordinates (Latitude and Longitude). It is also possible to filter searches by information related to animal biology, such as Gender and/or Wing Side.

The search for Taxonomic Classification can also be filtered by Family, Tribe, Genus and Specific Epithet. The Institution field, although it can at first be used by any user, is of greater interest to the internal users of the Instituto Butantan, as it involves searching for related Institutions (InstitutionName), Internal Code of the laboratory (LabCode) and the Collection depository (LCZCode). The results are displayed in a table, in which the user can sort the results by any of the table headings, just dragging it to the indicated field.

DISCUSSION

In recent publications referring to biological databases, Dujardin et al. (2010); Garros and Dujardin (2013) report on the importance of creating DBs related to the use of the mosquito wing geometric morphometry (WGM) technique. In addition, Lorenz et al. (2017); Jaramillo

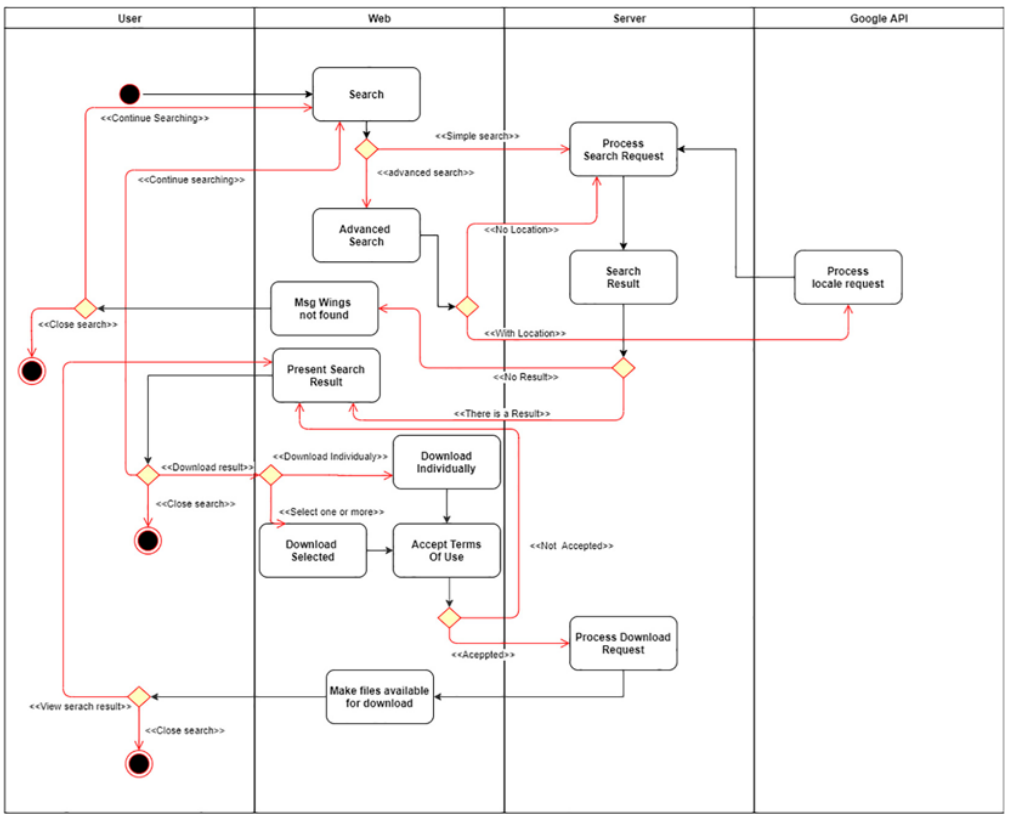

FIGURE 4 | Activity diagram for the WingBank platform search system.

TABLE 3 | List of genera and species stored in WingBank.

Genera	Number of related species
Chagasia (Cruz, 1906)	1
Aedeomyia (Theobald, 1901)	1
Runchomyia (Theobald, 1903)	1
Sabethes (Robineau-Desvoidy, 1827)	1
Trichoprosopon (Theobald, 1901)	1
Haemagogus (Willston, 1896)	2
Mansonia (Blanchard, 1901)	2
Limatus (Theobald, 1903)	2
Uranotaenia (Lynch Ambálzaga)	2
Wyeomyia (Theobald, 1901)	3
Psorophora (Robineau-Desvoidy, 1827)	5
Coquillettidia (Dyar, 1905)	5
Aedes (Meigen, 1818)	7
Anopheles (Meigen, 1818)	20
	77

et al. (2015), and Wilke et al. (2016), among other several authors (Calle et al., 2002; Jirakanjanakit and Dujardin, 2005; Dujardin, 2008; Jirakanjanakit et al., 2008; Henry et al., 2010; Vidal et al., 2011; Motoki et al., 2012; Vidal and Suesdek, 2012; Yeap et al., 2013; Börstler et al., 2014; Laurito et al., 2015; Phanitchat et al., 2019; Chaiphongpachara and Laojun, 2020; Sauer et al., 2020; Carvajal et al., 2021) who have already used to solve biological problems in Culicidae, directly and indirectly, reinforce the importance of gathering morphological, ecological and space-time data related to mosquitoes in a DB. The WingBank can enable studies of reanalysis and meta-analysis, micro and macroevolutionary and that can also contribute to integrative taxonomy, a multidisciplinary approach defended by several authors as the best approach for the diagnosis of species (Schlick-Steiner et al., 2010; Garros and Dujardin, 2013). Garros and Dujardin (2013) suggest that the need for a DB is underestimated, since the very power of morphometry to identify rates is underestimated, reinforcing that the success of the identification of rates through WGM depends on the relevance of the reference images in their level of form divergence, as well as in the classification technique. In addition, a program of automatic digitization of points, such as WINGMACHINE (Houle et al., 2003), or morphometric analysis in an agile way such as XYOM (XYOM-CLIC) can optimize the process.

However, despite all entries about the importance, as far as we know, until the creation of the WingBank in 2018 (Virginio-Fonseca, 2018), there was no relational database to raw material for analysis of alar geometric morphometry for mosquitoes with worldwide reach. The only project closest to what was expected from WingBank was MoMe-CLIC (Morphometrics in Medical Entomology—Collection of

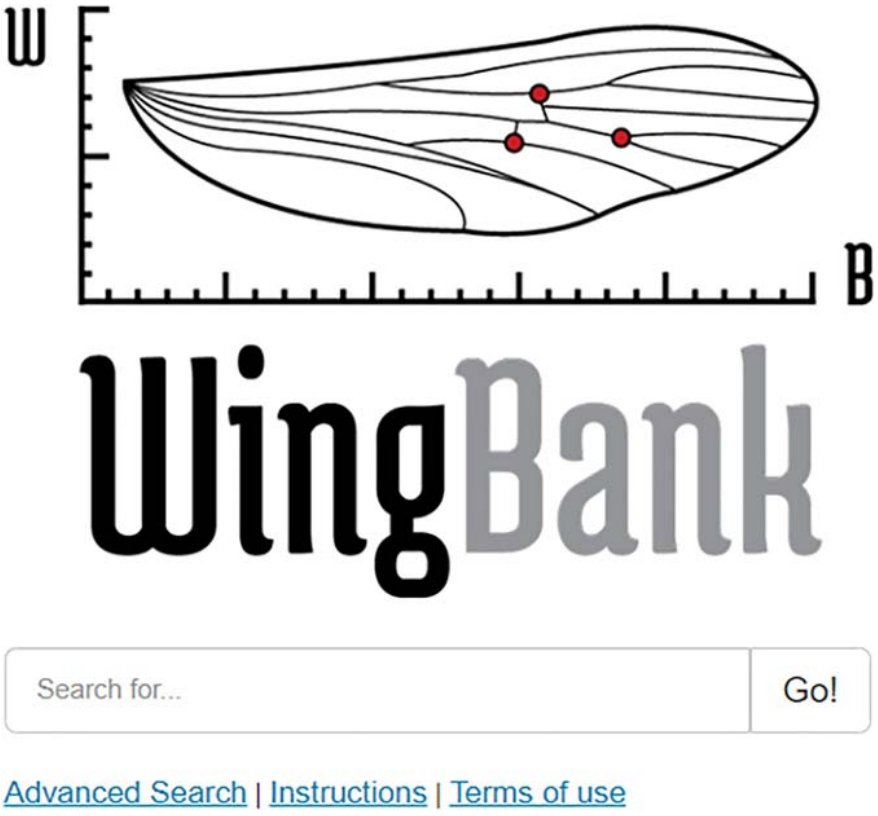

FIGURE 5 | Layout interface of home page and basic search of WingBank website.

Landmark for Identification and Characterization) which had an image repository "CLIC bank"—currently migrated to XYOM (XYOM-CLIC), created by Dujardin et al. (2010). It differs from WingBank by being a broad spectrum repository of Arthropod wing images covering different Classes and Orders, including Culicidae, Tephiritidae, Braconidae, Glossinidae, Ceratopogonidae, Psychodidae (Insecta; Diptera); Reduviidae (Insecta; Hemiptera) and Mummuciidae (Arachnida; Solifugae). With the changes that have taken place on its website over the years, XYOM became an interesting web application that dismisses downloads, installation, configuration and urdergo automatic updating. Although XYOM and WingBank have been distinctly conceived, their main goal allows them to be mutually complementary.

Currently, WingBank has data of 77 species belonging to 15 genera, as shown in **Table 3**. The WingBank shelters several mosquito species of medical and veterinary importance in Brazil and worldwide, such as *Cx. nigripalp*us, a species from which the Saint Louis virus has already been isolated (Belle et al., 1964; Chamberlain et al., 1964; Dow et al., 1964); *Cx. coronator* which has had specimens found naturally infected with several viruses, such as Saint Louis in Brazil and Trinidad and Tobago, the Venezuelan Equine Encephalitis virus in Mexico (Mackay et al., 2010) and the West Nile virus in the United States (Unlu et al., 2010); *Cx. quinquefasciatus*, species with records of participation in the transmission of West Nile fever, and lymphatic filariasis (Fernandes et al., 2016; Guedes et al., 2017).

In addition, it is possible to find in the WingBank data from *Ae. aegypti*, which participates in the transmission of urban yellow fever, Zika, chikungunya, dengue, besides being able to transmit microfilariae in urban areas (Jowett, 1986; Cirio, 2005; Lee and Rohani, 2005; Noridah et al., 2007; Chouin-Carneiro et al., 2016; Costa-Da-Silva et al., 2017a,b); *Ae. scapularis,* a species from which Ilhéus, Melao and yellow fever viruses were isolated (Spence et al., 1962; Vasconcelos et al., 2001; Pauvolid-Correa et al., 2013), and their participation in the transmission of the Rocio virus is suspected, due to their abundance in the Vale do Ribeira region at the time the epidemy occurred (Jowett, 1986). Another group present on WingBank is *Psorophora ferox*, which already was found naturally infected with Rocio virus (Lopes et al., 1981); *Haemagogus leucocelaenus*, which has specimens found naturally infected with the yellow fever virus (Shannon et al., 1938); *Coquilettidia venezuelensis,* related to the transmission of Oropouche virus (Forattini, 1965); and *Mansonia titillans*, from which Venezuelan Equine Encephalitis viruses have been isolated (Aitken, 1972; Sudia, 1972).

Furthermore, WingBank has data of several species of the *Anopheles* genus, such as *Anopheles darlingi, An. aquasalis, An. nuneztovari sl, An. oswaldoi, An. triannulatus sl, An. albitarsis sl, An. cruzii sl, An. bellator* and *An. homunculus*, which are

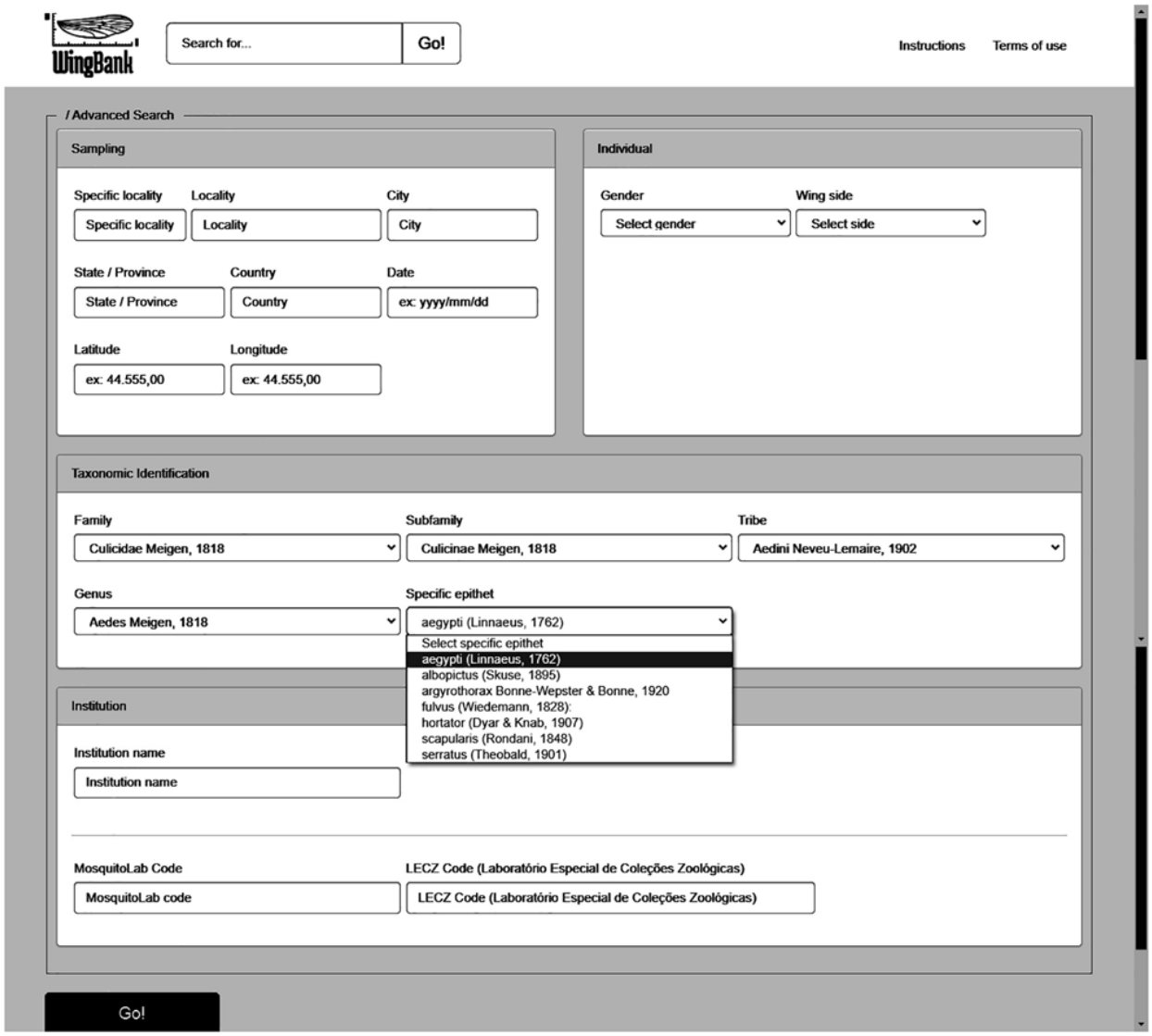

FIGURE 6 | Layout interface of advanced search of WingBank website. A detailed search for "Taxonomic Identification."

related to the transmission of malaria (Ramirez and Dessen, 1994; Tadei and Thatcher, 2000; Bourke et al., 2013). More than that, WingBank stores images and information from many other species, which may not have been studied for vectorial capacity yet, but which can be considered for future studies. Currently, data from 11 species are available for search on WingBank, because they are already related to some scientific publication: *Ae. aegypti, Ae. albopictus, Ae. scapularis, An. albitarsis s.l., An. arthuri s.l., An. cruzii s.l., An. homunculus, An. strodei s.l., An. triannulatus s.l., Cx. quinquefasciatus and Cx. nigripalpus.* This number may increase rapidly from the beginning of the use of the DB by external users, who will be able to contribute with images and information of species collected around the world.

Finally, it is known that the predictive power of this data to compose an automated identification system based on machine learning (Weeks et al., 1999; Gurgel-Gonçalves et al., 2017; Khalighifar et al., 2019; Valan et al., 2019; Motta et al., 2020) depends largely on the quality and quantity of the training sample (Kalayeh and Landgrebe, 1983; Nigam et al., 2000; Mukherjee et al., 2003; Tam et al., 2006; Dobbin et al., 2008; Kim, 2009). In this context, WingBank is a pioneer on several fronts: exclusivity regarding the alar structure and the Family Culicidae, extremely important for public health worldwide; quantitative and qualitative awareness of the data; accessibility and credibility in making the material available in a friendly interface and reliable hosting at one of the most recognized research institutions in the world: the Instituto Butantan. Therefore, it is expected that this DB will be an important ally, and possibly a watershed event in the field of creating new technologies in the area of public health.

WingBank is the largest database of Culicidae wings that we are aware of, and the first of the relational type and with the proposed intentions. The database facilitates the search for information. It already has 13,287 wing records, of which 12,939 are already linked to images, and 2,138 images are already available for use by third parties, which makes it possible to carry out future meta-analysis and reanalysis studies. Furthermore, each record received an access code (WingBankCode), which makes it more efficient and reliable to cite the records of wings into scientific publications. In addition, with WingBank it was possible to contemplate the South American culicidofauna, which stands out for its biodiversity, as well as for the great number of mosquito-related diseases present in this region.

In this context, the WingBank composed of thousands of information with important richness and density about the diversity of the Brazilian Culicid fauna can be used as a basis for programs to digitize the landmarks of the wings of mosquitoes. Making automatic identification through the use of this DB should be the next step, which will be further strengthened with each contribution made by professional colleagues from around the world. New studies with micro and macroevolutionary objectives will also be possible, and the work of employees of health fields will be facilitated, mainly if they do not know with which species they are handling. As geometric morphometrics increasingly rises, the launch of WingBank may revolutionize Medical Entomology, bringing benefits to students, professional workers and civil society. We are not waiting in the wings, this collaborative work is only beginning.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LS had the first insight into the creation of the database. FV wrote the first draft of the manuscript and materialized the idea by building the basilar structure. All authors contributed to the subsequent study conception and design. FV and LS performed the material preparation and the data collection. FV, LA, VD, and KB designed and implemented the database. LA performed the data cleaning and SQL process. All authors contributed to the versions of the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) grant 23038.005.274/2011-24, grant 032/2010—23038.001614/2016-52, and by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grants 13/05521-9, 11/18962-8, 10/15039-1, 10/14479-8, 07/01665-5, 06/05164-8 and 06/02622-5. LS has been fellow of Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grants 311805/2014-0 and 311984/2018-5, and fellow of Fundação Butantan (FB).

ACKNOWLEDGMENTS

We would like to thank the Instituto Butantan for allowing us to develop this work-study and allowing us to deposit the WingBank website on its domain, and for allowing the several mosquitoes samplings on its green area. We would also like to thank all workers and students from the Laboratório de Parasitologia and Laboratório de Coleções Zoológicas, from the same institution, for their support during sampling to provide the data. A special acknowledgment to Fernanda Almeida and Karina Zanatta (assistants from the Laboratório de Parasitologia) and Gabrielle R. de Andrade (Technologist from the Laboratório de Coleções Zoológicas) for hours of assistance on data organizing, spreadsheeting, paper diagramming and logistics, and Eliane Campos de Oliveira (Technologist from the Laboratório de Coleções Zoológicas) for years of dedication to support in curating the Entomological Collection from Instituto Butantan. Finally, we thank all taxonomists who identified the massive number of mosquitoes that we collected in the samplings.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.660941/full#supplementary-material

Supplementary Table 1 | Data dictionary of the semantic restrictions applied to WingBank.

REFERENCES

Adriaans, P., and Zantinge, D. (1996). *Data Mining*. Boston MA: Addison-Wesley.

Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., et al. (2015). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 43, D6–D17.

Aitken, T. H. G. (1972). "Habits of some mosquito hosts of VEE virus from northeastern South America, in cluding trinidad," in *Proceedings Workshop-Symposium on Venezuelan encephalitis virus. Pan American Health Organization*, 243, (Washington DC: Scientific Publ), 254–256.

Aytekin, S., Aytekin, A. M., and Alten, B. (2009). Effect of different larval rearing temperatures on the productivity (Ro) and morphology of the malaria vector anopheles superpictus grassi (Diptera: Culicidae) using geometric morphometrics. *J. Vector Ecol.* 34, 32–42. doi: 10.1111/j.1948-7134.2009.00005.x

Belle, E., Grant, L., and Page, W. (1964). The isolation of st. louis encephalitis virus from culex nigripalpus mosquitoes in jamaica. *Am. J. Trop. Med. Hygiene* 13, 452–454. doi: 10.4269/ajtmh.1964.13.452

Bitner-Mathé, B. C., and Klaczko, L. B. (1999). Heritability, phenotypic and genetic correlations of size and shape of drosophila mediopunctata

wings. *Heredity* 83, 688–696. doi: 10.1046/j.1365-2540.1999.00606.x

Bookstein, F. L. (1982). Foundations of morphometrics. *Ann. Rev. Ecol. Systemat.* 13, 451–470. doi: 10.1146/annurev.es.13.110182.002315

Börstler, J., Lühken, R., Rudolf, M., Steinke, S., Melaun, C., Becker, S., et al. (2014). The use of morphometric wing characters to discriminate female Culex pipiens and *Culex torrentium*. *J. Vector Ecol.* 39, 204–212.

Bourke, B. P., Oliveira, T. P., Suesdek, L., Bergo, E. S., and Sallum, M. A. M. (2013). A multi-locus approach to barcoding in the anopheles strodei subgroup (Diptera: Culicidae). *Parasites Vectors* 6:111.

Calle, L. D. A., Qui?ones, M. L., Erazo, H. F., and Jaramillo, O. N. (2002). Morphometric discrimination of females of five species of Anopheles of the subgenus Nyssorhynchus from Southern and Northwest Colombia. *Memórias do Instituto Oswaldo Cruz.* 97, 1191–1195. doi: 10.1590/s0074-02762002000800021

Calmon, P. (1975). *História De Dom Pedro II*. Rio De Janeiro: J. Olympio, 1–5.

Carvajal, T. M., Amalin, D. M., and Watanabe, K. (2021). Wing geometry and genetic analyses reveal contrasting spatial structures between male and female *Aedes aegypti* (L.) (Diptera: Culicidae) populations in metropolitan Manila. *Philippines. Infect. Genet. Evol.* 87:104676. doi: 10.1016/j.meegid.2020.104676

Caterino, M. S., Cho, S., and Sperling, F. A. H. (2000). The current state of insect molecular systematics: a thriving tower of babel. *Ann. Rev. Entomol.* 45, 1–54. doi: 10.1146/annurev.ento.45.1.1

Chaiphongpachara, T., and Laojun, S. (2020). Wing morphometric variability of the malaria vector *Anopheles (Cellia) epiroticus* linton et harbach (diptera: culicidae) for the duration of the rainy season in coastal areas of samut songkhram. *Thailand. Folia Parasitologica. (Praha).* 67:2020.007.

Chaiphongpachara, T., Sriwichai, P., Samung, Y., and Ruangsittichai, J. (2019). Geometric morphometrics approach towards discrimination of three member species of Maculatus group in Thailand. *Acta Tropica* 192, 66–74. doi: 10.1016/j.actatropica.2019.01.024

Chaiphongpachara, T., and Laojun, S. (2019). Comparative analysis of the internal and external outlines of wings for an outline-based geometric morphometric approach to distinguish three *Aedes mosquitoes* (Diptera: Culicidae) in Thailand. *J. Entomol. Acarol. Res.* 51, 1–10.

Chamberlain, R., Sudia, W., Coleman, P., and Beadle, L. (1964). Vector studies in the St. Louis encephalitis epidemic, tampa bay area, Florida, 1962. *Am. J. Trop. Med. Hygiene* 13, 456–461. doi: 10.4269/ajtmh.1964.13.456

Chen, P. P.-S. (1977). *The Entity-Relationship Model: Toward a Unified View of Data*. Massachusetts, MA: Library Of The Massachusetts Institute Of Technology.

Chouin-Carneiro, T., Vega-Rua, A., Vazeille, M., Yebakima, A., Girod, R., Goindin, D., et al. (2016). Differential susceptibilities of aedes aegypti and aedes albopictus from the americas to Zika Virus. *PLoS Negl. Trop. Dis.* 10:e0004543. doi: 10.1371/journal.pntd.0004543

Cirio, S. M. (2005). *Epidemiologia E Clínica De Cães Portadores De Dirofilariose Em Espaços Urbanos De Município Do Litoral Do Paraná E Aspectos Da Histologia De Culex quinquefasciatus (Say, 1823) (Diptera, Culicidae)*. Curitiba: Universidade Federal Do Paraná.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Commun. ACM* 13:377. doi: 10.1145/362384.362685

Costa-Da-Silva, A. L., Ioshino, R. S., De Araujo, H. R. C., Kojin, B. B., Zanotto, P. M. D., Oliveira, D. B. L., et al. (2017a). Laboratory strains of aedes aegypti are competent to Brazilian zika virus. *PLos One* 12:e0171951. doi: 10.1371/journal.pone.0171951

Costa-Da-Silva, A. L., Ioshino, R. S., Petersen, V., Lima, A. F., Cunha, M. D. P., Wiley, M. R., et al. (2017b). First report of naturally infected aedes aegypti with chikungunya virus genotype ECSA in the Americas. *PLoS Negl. Trop. Dis.* 11:e0005630. doi: 10.1371/journal.pntd.0005630

Date, C. J. (2004). *An Introduction To Database Systems*, 8 Edn. London: Pearson Education Inc.

Dobbin, K. K., Zhao, Y., and Simon, R. M. (2008). How large a training set is needed to develop a classifier for microarray data? *Clin. Cancer Res.* 14, 108–114. doi: 10.1158/1078-0432.ccr-07-0443

Dow, R., Coleman, P., Meadows, K., and Work, T. (1964). Isolation of St. Louis encephalitis viruses from mosquitoes in the tampa bay area of florida during the epidemic Of 1962. *Am. J. Trop. Med. Hygiene* 13, 462–468. doi: 10.4269/ajtmh.1964.13.462

Drāghici, S., Sellamuthua, S., and Khatria, P. (2006). Babel's tower revisited: a universal resource for crossreferencing across annotation databases. *Bioinformatics* 22, 2934–2939. doi: 10.1093/bioinformatics/btl372

Dujardin, J. P. (2008). Morphometrics applied to medical entomology. *Infect. Genet. Evol.* 8, 875–890. doi: 10.1016/j.meegid.2008.07.011

Dujardin, J.-P. (2012). *Morphometrics in Medical Entomology – Collection of Landmark for Identification and Characterization.* Available online at: https://xyom-clic.eu/posts/ (accessed April 5, 2021).

Dujardin, J. P., Kaba, D., and Henry, A. B. (2010). The exchangeability of shape. *BMC Res. Notes* 3:266. doi: 10.1186/1756-0500-3-266

Elmasri, R., and Navathe, S. B. (2011). *Fundamentals Of Database Systems*, 6th Edn. London: Pearson.

Evans, E. (2003). *Domain-Driven Design: Tackling Complexity In The Heart Of Software*. Boston MA: Addison Wesley.

Everest, G. (1976). *Basic Data Structure Models Explained With A Common Example*. Austin, TX: IEEE Computer Society Publications Office.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, A. P. (1996). From data mining toknowledge discovery indatabases. *AI Magazine* 17:37. Available online at: https://www.aaai.org/ojs/index.php/aimagazine/article/view/1230/1131 (accessed April 5, 2021).

Fernandes, R. S., Campos, S. S., Ferreira-de-Brito, A., Miranda, R. M., da Silva, K. A. B., de Castro, M. G., et al. (2016). Culex quinquefasciatus from Rio de Janeiro is not competent to transmit the local Zika Virus. *PLoS Negl. Trop. Dis.* 10:e0004993. doi: 10.1371/journal.pntd.0004993

Foley, D., Rueda, P., and Wilkerson, R. (2011). *Vectormap*. Available online at: http://vectormap.si.edu/ (accessed April 5, 2021).

Forattini, O. P. (1965). *Entomologia Médica*. São Paulo: Universidade De São Paulo.

Gaffigan, T., and Pecor, J. (1997). *Collecting, Rearing, Mounting and Shipping Mosquitoes.* Available online at: https://wrbu.si.edu/resources/protocols (accessed April 5, 2021).

Garros, C., and Dujardin, J.-P. (2013). "Genetic and phenetic approaches to anopheles systematics, anopheles mosquitoe – new insights into malaria vectors," in *Anopheles Mosquitoes – New insights Into Malaria Vectors* Available online at: https://www.intechopen.com/books/anopheles-mosquitoes-new-insights-into-malaria-vectors/genetic-and-phenetic-approaches-to-anopheles-systematics (accessed April 5, 2021).

Godfray, H. C. J. (2002). Challenges for taxonomy. *Nature* 417, 17–18.

Gómez, G. F., Márquez, E. J., Gutiérrez, L. A., Conn, J. E., and Correa, M. M. (2014). Geometric morphometric analysis of colombian *Anopheles albimanus* (Diptera: Culicidae) reveals significant effect of environmental factors on wing traits and presence of a metapopulation. *Acta Tropica* 135, 75–85. doi: 10.1016/j.actatropica.2014.03.020

Gubler, D. J. (1991). "Insects in disease transmission," in *Hunter Tropical Medicine*, 7th Edn, ed. G. T. Strickland (Philadelphia (PA): W. B. Saunders), 981–1000.

Guedes, D., Paiva, M., Donato, M., Barbosa, P., Krokovsky, L., Rocha, S., et al. (2017). Zika Virus replication in the mosquito culex quinquefasciatus in Brazil. *Emerg. Microbes Infect.* 6:e69.

Gurgel-Gonçalves, R., Komp, E., Campbell, L. P., Khalighifar, A., Mellenbruch, J., Mendonça, V. J., et al. (2017). Automated identification of insect vectors of chagas disease in Brazil and Mexico: the virtual vector lab. *PeerJ* 18:e3040. doi: 10.7717/peerj.3040

Henry, A., Thongsripong, P., Fonseca-Gonzalez, I., Jaramillo-Ocampo, N., and Dujardin, J. P. (2010). Wing shape of dengue vectors from around the world. *Infect. Genet. Evol.* 10, 207–214. doi: 10.1016/j.meegid.2009.12.001

Houle, D., Mezey, J., Galpern, P., and Carter, A. (2003). *BMC Evol. Biol.* 3:25. doi: 10.1186/1471-2148-3-25

IBM (2011). *IBM's 100 Icons of Progress – Relational Database.* Available online at: https://www.ibm.com/ibm/history/ibm100/us/en/icons/reldb/ (accessed April 5, 2021).

Jaramillo, N., Dujardin, J. P., Calle-Londono, D., and Fonseca-Gonzalez, I. (2015). Geometric morphometrics for the taxonomy of 11 species of anopheles (Nyssorhynchus) mosquitoes. *Med. Vet. Entomol.* 29, 26–36. doi: 10.1111/mve.12091

Jirakanjanakit, N., and Dujardin, J. P. (2005). Discrimination of *Aedes aegypti* (Diptera: Culicidae) laboratory lines based on wing geometry. *Southeast Asian J. Trop. Med. Public Health* 36, 858–861.

Jirakanjanakit, N., Leemingsawat, S., and Dujardin, J. P. (2008). The geometry of the wing of *Aedes (Stegomyia) aegypti* in isofemale lines through successive generations. *Infect. Genet. Evol.* 8, 414–421. doi: 10.1016/j.meegid.2007.05.004

Jirakanjanakit, N., Leemingsawat, S., Thongrungkiat, S., Apiwathnasorn, C., Singhaniyom, S., Bellec, C., et al. (2007). Influence of larval density or food variation on the geometry of the wing of *Aedes (Stegomyia) aegypti*. *Trop. Med. Int. Health* 12, 1354–1360. doi: 10.1111/j.1365-3156.2007.01919.x

Jowett, T. (1986). "Preparation of nucleic acids," in *Drosophila A Practical Approach*, ed. D. B. Roberts (Oxford: Roberts DB Press).

Kalayeh, H. M., and Landgrebe, D. A. (1983). Predicting the required number of training samples. *IEEE Trans. Pattern Anal. Mach. Intell.* 5, 664–667. doi: 10.1109/tpami.1983.4767459

Khalighifar, A., Komp, E., Ramsey, J. M., Gurgel-Gonçalves, R., and Peterson, A. T. (2019). Deep learning algorithms improve automated identification of chagas disease vectors. *J. Med. Entomol.* 56, 1404–1410. doi: 10.1093/jme/tjz065

Kim, S. Y. (2009). Effects of sample size on robustness and prediction accuracy of a prognostic gene signature. *BMC Bioinform.* 10:147. doi: 10.1186/1471-2105-10-147

Klingenberg, C. P. (2010). Evolution and development of shape: integrating quantitative approaches. *Nat. Rev. Genet.* 11, 623–635. doi: 10.1038/nrg2829

Klingenberg, C. P. (2011). Morphoj: an integrated software package for geometric morphometrics. *Mol. Ecol. Resources* 11, 353–357. doi: 10.1111/j.1755-0998.2010.02924.x

Laurito, M., Almirón, W. R., and Ludueña-Almeida, F. F. (2015). Discrimination of four culex (Culex) species from the Neotropics based on geometric morphometrics. *Zoomorphology* 1611, 447–455. doi: 10.1007/s00435-015-0271-x

Lee, H. L., and Rohani, A. (2005). Transovarial transmission of dengue virus in aedes aegypti and aedes albopictus in relation to dengue outbreak in an urban area in Malaysia. *Dengue Bull.* 29, 106–111.

Lopes, O. D., Sacchetta, L. D., Francy, D. B., Jakob, W. L., and Calisher, C. H. (1981). Emergence of a new arbovirus disease in Brazil. 3. isolation of rocio virus from *Psorophora ferox* (Humboldt, 1819). *Am. J. Epidemiol.* 113, 122–125. doi: 10.1093/oxfordjournals.aje.a113075

Lorenz, C., Almeida, F., Almeida-Lopes, F., Louise, C., Pereira, S. N., Petersen, V., et al. (2017). Geometric morphometrics in mosquitoes: what has been measured? *Infect. Genet. Evol.* 54, 205–215. doi: 10.1016/j.meegid.2017.06.029

Lorenz, C., and Suesdek, L. (2013). Short report: evaluation of chemical preparation on insect wing shape for geometric morphometrics. *Am. J. Trop. Med. Hygiene* 89, 928–931. doi: 10.4269/ajtmh.13-0359

Mackay, A., Kramer, W., Meece, J., Brumfield, R., and Foil, L. (2010). Host feeding patterns of culex mosquitoes (Diptera: Culicidae) in east baton rouge parish, louisiana. *J. Med. Entomol.* 47, 238–248. doi: 10.1603/me09168

Monteiro, L., and Reis, S. (1999). *Princípios De Morfometria Geométrica*. Ribeirão Preto: Holos.

Morales Vargas, R., Phumala-Morales, N., Tsunoda, T., Apiwathnasorn, C., and Dujardin, J. (2013). The phenetic structure of *Aedes albopictus*. *Infect. Genet. Evol.* 13, 242–251. doi: 10.1016/j.meegid.2012.08.008

Morales-Vargas, R. E., Ya-Umphan, P., Phumala-Morales, N., Komalamisra, N., and Dujardin, J. P. (2010). Climate associated size and shape changes in *Aedes aegypti* (Diptera: Culicidae) populations from Thailand. *Infect. Genet. Evol.* 10, 580–585. doi: 10.1016/j.meegid.2010.01.004

Motoki, M. T., Suesdek, L., Bergo, E. S., and Sallum, M. A. M. (2012). Wing geometry of *Anopheles darlingi* root (Diptera:Culicidae) in five major Brazilian ecoregions. *Infect. Genet. Evol.* 12, 1246–1252. doi: 10.1016/j.meegid.2012.04.002

Motta, D., Santos, A. ÁB., Machado, B. A. S., Ribeiro-Filho, O. G. V., Camargo, L. O. A., Valdenegro-Toro, M. A., et al. (2020). Optimization of convolutional neural network hyperparameters for automatic classification of adult mosquitoes. *PLoS One* 15:e0234959. doi: 10.1371/journal.pone.0234959

Mukherjee, S., Tamayo, P., Rogers, S., Rifkin, R., Engle, A., Campbell, C., et al. (2003). Estimating dataset size requirements for classifying DNA microarray data. *J. Comput. Biol.* 10, 119–142. doi: 10.1089/106652703321825928

Nigam, K., Mccallum, A. K., Thrun, S., and Mitchell, T. (2000). Text classification from labeled and unlabeled documents using EM. *Machine Learn.* 39, 103–134.

Noridah, O., Paranthaman, V., Nayar, S., Masliza, M., Ranjit, K., Norizah, I., et al. (2007). Outbreak of chikungunya due to virus of central/east african genotype in Malaysia. *Med. J. Malaysia* 62, 323–328.

Parsaye, K., and Chignell, M. (1993). *Intelligent Database Tools And Applications: Hyperin-Formation Access, Data Quality, Visualization, Automatic Discovery, 1st Edn.* Wiley Professional Computing; John Wiley & Sons. 560.

Pauvolid-Correa, A., Kenney, J. L., Couto-Lima, D., Campos, Z. M. S., Schatzmayr, H. G., Nogueira, R. M. R., et al. (2013). Ilheus virus isolation in the pantanal, west-central Brazil. *PLoS Negl. Trop. Dis.* 7:e2318. doi: 10.1371/journal.pntd.0002318

Phanitchat, T., Apiwathnasorn, C., Sungvornyothin, S., Samung, Y., Dujardin, S., Dujardin, J. P., et al. (2019). Geometric morphometric analysis of the effect of temperature on wing size and shape in *Aedes albopictus*. *Med. Vet. Entomol.* 33, 476–484. doi: 10.1111/mve.12385

Philip, C. B., and Rozenboom, L. E. (1973). "Medico-veterinary entomology: a generation of progress. history of entomology," in *History Of Entomology*, eds R. F. Smith, T. E. Mittler, and C. N. Smith (Palo Alto (CA): Annual Reviews), 333–359.

Ramirez, C. C. L., and Dessen, E. M. B. (1994). Cytogenetic analysis of a natural-population of *Anopheles cruzii*. *Revista Brasileira De Genetica* 17, 41–46.

Richtsmeier, J. T., Deleon, V. B., and Lele, S. R. (2002). The promise of geometric morphometrics. *Yearbook Phys. Anthropol.* 45, 63–91.

Rohlf, F. J. (1993). Morphometric tools for landmark data - geometry and biology -bookstein, FL. *J. Classif.* 10, 133–136.

Ruangsittichai, J., Apiwathnasorn, C., and Dujardin, J. P. (2011). Interspecific and sexual shape variation in the filariasis vectors *Mansonia dives* and *Ma. bonneae*. *Infect. Genet. Evol.* 11, 2089–2094. doi: 10.1016/j.meegid.2011.10.002

Sallum, M. A. M., Foster, P. G., Dos Santos, C. L. S., Flores, D. C., Motoki, M. T., and Bergo, E. S. (2010). Resurrection of two species from synonymy of anopheles (Nyssorhynchus) strodei root, and characterization of a distinct morphological form from the strodei complex (Diptera: Culicidae). *J. Med. Entomol.* 47, 504–526. doi: 10.1093/jmedent/47.4.504

Sauer, F. G., Jaworski, L., Erdbeer, L., Heitmann, A., Schmidt-Chanasit, J., Kiel, E., et al. (2020). Geometric morphometric wing analysis represents a robust tool to identify female mosquitoes (Diptera: Culicidae) in Germany. *Sci. Rep.* 19:17613.

Savasere, A., Omiecinski, E., and Navathe, S. (1995). "An efficient algorithm for mining association rules in large databases," in *Proceedings of the 21st International Conference On Very Large Data Bases*, New York NY: ACM 432–444.

Schlick-Steiner, B., Steiner, F., Seifert, B., Stauffer, C., Christian, E., and Crozier, R. (2010). Integrative taxonomy: a multisource approach to exploring biodiversity. *Ann. Rev. Entomol.* 55, 421–438. doi: 10.1146/annurev-ento-112408-085432

Seltzer, M. (2008). Beyond relational databases. *Commun. ACM* 3, 50–58.

Shannon, R. C., Whitman, L., and Franca, M. (1938). Yellow fever virus in jungle mosquitoes. *Science* 88, 101–110.

Sonnenschein, A., Vanderzee, D., Pitchers, W. R., Chari, S., and Dworkin, I. (2015). An image database of drosophila melanogaster wings for phenomic and biometric analysis. *Gigascience* 4:25.

Souza, A. L., da, S., Multini, L. C., Marrelli, M. T., and Wilke, A. B. B. (2020). Wing geometric morphometrics for identification of mosquito species (Diptera: Culicidae) of neglected epidemiological importance. *Acta Tropica.* 211:105593. doi: 10.1016/j.actatropica.2020.105593

Spence, L., Anderson, C. R., Aitken, T. H., and Downs, W. G. (1962). Melao Virus, a new agent isolated from trinidadian mosquitoes. *Am. J. Trop. Med. Hygiene* 11, 687–690. doi: 10.4269/ajtmh.1962.11.687

Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., et al. (2015). Big data: astronomical or genomical? *PLoS Biol.* 13:e1002195. doi: 10.1371/journal.pbio.1002195

Sudia, W. D. (1972). *Arthropod Vectors Of Venezuelan Equine Encephalitis*. Washington, DC: Pan American Healt Organization.

Sumruayphol, S., Apiwathnasorn, C., Ruangsittichai, J., Sriwichai, P., Attrapadung, S., Samung, Y., et al. (2016). DNA barcoding and wing morphometrics to distinguish three Aedes vectors in Thailand. *Acta Tropica* 159, 1–10. doi: 10.1016/j.actatropica.2016.03.010

Tadei, W. P., and Thatcher, B. D. (2000). Malaria vectors in the brazilian amazon: anopheles of the subgenus nyssorhynchus. *Revista Do Instituto De Medicina Tropical De São Paulo* 42, 87–94. doi: 10.1590/s0036-46652000000200005

Tam, V. H., Kabbara, S., Yeh, R. F., and Leary, R. H. (2006). Impact of sample size on the performance of multiple-model pharmacokinetic simulations. *Antimicrobial Agents Chemotherapy* 50, 3950–3952. doi: 10.1128/aac.00337-06

Unlu, I., Kramer, W., Roy, A., and Foil, L. (2010). Detection of west nile virus rna in mosquitoes and identification of mosquito blood meals collected at alligator farms in Louisiana. *J. Med. Entomol.* 47, 625–633. doi: 10.1093/jmedent/47.4.625

Valan, M., Makonyi, K., Maki, A., Vondráček, D., and Ronquist, F. (2019). Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Systematic Biol.* 68, 876–895. doi: 10.1093/sysbio/syz014

Vasconcelos, P. F. C., Costa, Z. G., Travassos, Da Rosa, E. S., Luna, E., Rodrigues, S. G., et al. (2001). Epidemic of jungle yellow fever in Brazil, 2000: implications of climatic alterations in disease spread. *J. Med. Virol.* 65, 598–604. doi: 10.1002/jmv.2078

Vences, M., Guayasamin, J. M., Miralles, A., and De la Riva, I. (2013). To name or not to name: criteria to promote economy of change in linnaean classification schemes. *Zootaxa* 3636, 201–244. doi: 10.11646/zootaxa.3636.2.1

Vidal, P. O., Peruzin, M. C., and Suesdek, L. (2011). Wing diagnostic characters for *Culex quinquefasciatus* and *Culex nigripalpus* (Diptera, Culicidae). *Revista Brasileira de Entomologia*. 55, 134–137. doi: 10.1590/s0085-56262011000100022

Vidal, P. O., and Suesdek, L. (2012). Comparison of wing geometry data and genetic data for assessing the population structure of *Aedes aegypti*. *Infect. Genet. Evol.* 12, 591–596. doi: 10.1016/j.meegid.2011.11.013

Virginio-Fonseca, F. (2018). *Morfometria geométrica e banco de dados na investigação de problemas biológicos em Culicidae [tese]*. São Paulo: Instituto de Ciências Biomédicas, doi: 10.11606/T.42.2019.tde-06062018-6150917

Virginio, F., Vidal, P. O., and Suesdek, L. (2015). Wing sexual dimorphism of pathogen-vector culicids. *Parasites Vectors* 8:769.

Weeks, P. J. D., O'Neill, M. A., Gaston, K. J., and Gauld, I. D. (1999). Automating insect identification: exploring the limitations of a prototype system. *J. Appl. Entomol.* 123, 1–8. doi: 10.1046/j.1439-0418.1999.00307.x

Wiley, E. O., and Liebermann, B. S. (2011). *Phylogenetic Systematics*, 2 Edn. 432. Hoboken, NJ: Wiley online library

Wilke, A. B. B., Christe, R. D., Multini, L. C., Vidal, P. O., Wilk-Da-Silva, R., De Carvalho, G. C., et al. (2016). Morphometric wing characters as a tool for mosquito identification. *PLoS One* 11:e0161643. doi: 10.1371/journal.pone.0161643

Wilkerson, R. C., Linton, Y. M., Fonseca, D. M., Schultz, T. R., Price, D. C., and Strickman, D. A. (2015). Making mosquito taxonomy useful: a stable classification of tribe aedini that balances utility with current knowledge of evolutionary relationships. *PLos One* 10:e0133602. doi: 10.1371/journal.pone.0133602

World Health Organization (2020a). *Lymphatic Filariasis*. Geneva: World Health Organization.

World Health Organization (2020b). *World Malaria Report 2020: 20 Years of Global Progress and Challenges*. Geneva: World Health Organization.

Yeap, H. L., Endersby, N. M., Johnson, P. H., Ritchie, S. A., and Hoffmann, A. A. (2013). Body size and wing shape measurements as quality indicators of Aedes aegypti mosquitoes destined for field release. *Am. J. Trop. Med. Hygiene* 89, 78–92. doi: 10.4269/ajtmh.12-0719

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation with several of the authors, LS and KB, at time of review.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 10 May 2021
doi: 10.3389/fevo.2021.660151

Historical Biogeography and the Evolution of Hematophagy in Rhodniini (Heteroptera: Reduviidae: Triatominae)

*Alexandre Silva de Paula[1], Carlos Barreto[1,2], Mike Cerqueira Martins Telmo[1], Lileia Diotaiuti[3] and Cleber Galvão[4]**

[1] Laboratório de Sistemática e Biogeografia, Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, Ouro Preto, Brazil, [2] Department of Biology, Western University, London, ON, Canada, [3] Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Brazil, [4] Laboratório Nacional e Internacional de Referência em Taxonomia de Triatomíneos, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

The Rhodniini tribe is one of the five tribes in the subfamily Triatominae and is notorious for its domestic blood-sucking pests and vectors of *Trypanosoma cruzi* across Latin America. The human and economic costs of the Chagas disease in the American tropics are considerable, and these insects are of unquestionable importance to humans. We used mitochondrial rDNA (16S), nuclear ribosomal RNA (28S) and wingless (Wg) sequences to perform phylogenetic analysis to derive trees based on parsimony and maximum likelihood. Nucleotide sequences were used in molecular-clock analyses to estimate time divergence between species of Rhodniini. The potential distribution of each species was modeled and compared with Kappa statistic. Multivariate niches with bioclimatic variables were used to describe differences between the species using discriminant analysis. The results of this study indicate that the Rhodniini originated 17.91 Mya ago. *Rhodnius domesticus* is the oldest species having its origin at 9.13 Mya. Rhodniini are closely related to Salyavatinae that are specialist termite predators and diverged from this subfamily 30.43 Mya. Most species are clearly allopatric and have distinct bioclimatic niches. The colonization of bromeliads, palms trees and bird nests represent important events for the speciation of these taxa. The hematophagous habit can be described as a scenario where Rhodniini's ancestor could be pre-adapted for the invasion of bromeliads, palm trees, and bird nests where they would find significant water availability and thermal damping. These environments are widely used by vertebrate inquilines that would be the source of food for the species of Rhodniini. Lastly, our results show an alternative position of *Psammolestes* in the phylogenetic tree.

Keywords: MAXENT, molecular clock, Neotropical, Rhodnius, *Psammolestes*, vicariance

OPEN ACCESS

Edited by:
Anthony I. Cognato, Michigan State University, United States

Reviewed by:
João Aristeu Da Rosa, Universidade Estadual Paulista Júlio de Mesquita Filho, Brazil
Amanda Ravazi, São Paulo State University, Brazil

****Correspondence:***
Cleber Galvão
clebergalvao@gmail.com

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *28 January 2021*
Accepted: *12 April 2021*
Published: *10 May 2021*

Citation:
de Paula AS, Barreto C, Telmo MCM, Diotaiuti L and Galvão C (2021) Historical Biogeography and the Evolution of Hematophagy in Rhodniini (Heteroptera: Reduviidae: Triatominae). Front. Ecol. Evol. 9:660151. doi: 10.3389/fevo.2021.660151

INTRODUCTION

The Rhodniini tribe is one of the five tribes in the subfamily Triatominae and is notorious for its domestic blood-sucking pests and vectors of *Trypanosoma cruzi* (Chagas, 1909) (Kinetoplastida, Trypanosomatidae) across Latin America. The tribe has a wide geographical distribution ranging from Central America to the Southern Cone (Hernández et al., 2020). The human and economic

cost of the Chagas disease in the American tropics is considerable, totaling the annual burden of $627.46 million in health-care costs (Lee et al., 2013). There are about six million chronic disease carriers in the Americas and these insects are of unquestionable importance to humans (PAHO/WHO, 2020). Because of this, Rhodniini has been extensively studied due to its epidemiological importance (Schofield and Galvão, 2009). The tribe is considerably diverse, with 24 described species (21 of the genus *Rhodnius* Stål and three of *Psammolestes* Bergroth) (Zhao et al., 2021). The members of the genus *Rhodnius* have been classified into three groups—*pallescens* (trans-Andean), *pictipes* (cis-Andean), and *prolixus*—based on geographical distribution and morphology (Schofield and Galvão, 2009; Hernández et al., 2020). Although the taxonomy of *Rhodnius* has been widely studied, there are controversies regarding the number of species, the classification of these species into groups, and the phylogenetic relationships and monophyletic status of these groups that deserve further studies to elucidate *Rhodnius* taxonomic status. This is the case of *R. zeledoni* Jurberg et al., 2009, which seems to be very similar to *R. domesticus* Neiva and Pinto, 1923 (Galvão, 2014, 2020). The unique specimen was found very damaged in Sergipe state, Brazil, a region included in the distribution range of *R. domesticus*, therefore, examination of further material is essential to confirm whether or not *R. zeledoni* is a valid species. Following Monteiro et al. (2018), *R. marabaensis* dos Santos Souza et al., 2016 has been suggested to be a genetic variant very close to *R. robustus* Larrousse, 1927 (specific status recently considered valid by Castro et al., 2020) and *R. taquarussuensis* (Rosa et al., 2017), was synonymized with *R. neglectus* Lent, 1954 by Nascimento et al. (2019) based on interspecific crosses and molecular markers.

Issues in the phylogenetic relationships in *Rhodnius* have arisen from the conflicting results of morphometric analyses and molecular markers (Schofield and Galvão, 2009; Monteiro et al., 2018). Monteiro et al. (2003) and Pavan et al. (2013) proposed the paraphyly of *R. robustus* with respect to *R. prolixus* Stål, 1859, despite not including other populations of the species of the prolixus complex—*R. nasutus* Stål, 1859 and *R. neglectus*. Although rarely studied and recognized, the cross-contamination between *R. prolixus* and *R. robustus* colonies is fairly frequent. Mesquita et al. (2015) and Monteiro et al. (2018) proposed that the colony material should always be genotyped and compared with the established taxonomic standards, particularly when the research involves taxa with known cryptic variation. Misidentification of cryptic species must also be considered.

Triatominae species are almost entirely hematophagous, although there are cases of some species that feed on other invertebrates (Sandoval et al., 2004, 2010). Through evolutionary processes, Triatominae species have specialized morphological, physiological, and behavioral adaptations for feeding on blood. Among them are mouthparts, specifically saliva composition and enzymes, along with digestive symbionts that are the most noticeable (Otálora-Luna et al., 2015). The ecological, morphological and molecular changes that are involved in the transition from a predatory lifestyle to hematophagous Triatominae are still unclear due to the lack of robust and densely sampled phylogenies, preventing the development of significant hypothesis tests on these issues (Monteiro et al., 2018). The transition from predatory habits in Reduviidae to hematophagous habits in Triatominae could have occurred by colonization of vertebrate nests. Furthermore, the evolution of blood feeding in the tribes Triatomini and Rhodniini + Cavernicolini may have occurred independently (Hwang and Weirauch, 2012).

de Paula et al. (2005) presented the first molecular evidence for the polyphyletic origin of the subfamily Triatominae and showed a separation between Rhodniini and Triatomini tribes from predatory groups. In that study, it is suggested that the Reduviinae subfamily is a sister group of Triatomini, and the subfamilies Salyavatinae and Harpactorinae are sister groups of Rhodniini. According to Hwang and Weirauch (2012), Triatominae are paraphyletic with respect to the *Opisthacidius* Berg, 1879 (Reduviinae), and this contrasts prior hypotheses that found Triatominae to be monophyletic or polyphyletic. In addition, the same authors, as well as Patterson and Gaunt (2010) and Justi et al. (2016) proposed *Zelurus* Hahn, 1826 and *Opisthacidius* as sister groups to Triatominae in their phylogenetic analysis. Despite that, several issues remain underlying on the hypotheses of the origin of Rhodniini. The most common involves the period in which one group has arisen and the geologic and biogeographic events that have conditioned it. Understanding the mechanisms and processes underlying patterns of species distributions remains a central question in historical biogeography (Lomolino et al., 2016).

Molecular dating studies involving Triatominae species included Gaunt and Miles (2002); Hwang and Weirauch (2012), and Justi et al. (2016). Until the work of Justi et al. (2016), the Rhodniini tribe had been analyzed in biogeography studies by de Paula et al. (2007) and Abad-Franch et al. (2009). Other biogeographical hypotheses for Rhodniini's evolution have also been tested (Hypša et al., 2002; Hwang and Weirauch, 2012). Justi et al. (2016) presented the first comprehensive phylogenetic study aiming to understand the geological events that led to Triatominae diversification using fossils as calibration points for time estimates and correlating the geological changes in the Neotropics to Triatominae evolution. They showed that the major geological changes in the Neotropics since the Eocene have played a role in the diversification of Triatominae, including the effects of the Andean uplift in the Amazonian area, comprising the formation of the Pebas system (23–10 Mya); the formation of the Acre system (10–7 Mya); the formation of the connection between the Amazon and the Atlantic Forest; the GAARlandia (GAAR = Greater Antilles + Aves Ridge) land bridge; the period of biodiversity exchange resulting from the closure of the Isthmus of Panama (10–2.7 Mya); the Florida and Gulf Coast inundations during the Miocene; and dispersal across the Bering Land Bridge during the Eocene.

The goal of the study reported here was to expand on our analysis (de Paula et al., 2007) by providing an update on the biogeographic hypotheses to explain the modern geographic distribution of Rhodniini species. We also performed molecular dating analysis for each species and presented an evolutionary hypothesis for the evolution of hematophagy. A fundamental question that we propose in this work is to evaluate the relationship between the Rhodniini tribe and the subfamilies Reduviinae and Salyavatinae. Salyavatinae are specialist termite

predators (Cobben, 1978; van Doesburg and Forero, 2012) while Reduviinae are generalist predators (Hwang and Weirauch, 2012). Consequently, there are two possible scenarios to explain the haematophagic habit in Rhodniini that need to be elucidated. Lastly, we used bioclimatic data to show potential areas of distribution of each species.

MATERIALS AND METHODS

Sequence Alignment, Phylogenetic Analysis, and Molecular Dating

The study makes use of mitochondrial rDNA sequences (16S), nuclear ribosomal RNA (28S) and wingless (Wg) sequences currently available in GenBank. We chose these markers because they are available for the outgroup and for most of the Rhodniini species studied.

Analyses of 16S rDNA, 28S rDNA and Wg sequences were performed by aligning sequences using MUSCLE (Edgar, 2004a,b[1]), and treating the gaps as missing (?) (Swofford, 2019). *Zelurus alcides* (Stål, 1863) (Reduviinae), *Opisthacidius chinai* Lent and Wygodzinsky, 1956 (Reduviinae), and *Lisarda rhypara* (Stål, 1859) (Salyavatinae) were used as outgroups in the 16S + 28S + Wg analyses based on previous studies (de Paula et al., 2005; Patterson and Gaunt, 2010; Hwang and Weirauch, 2012; Justi et al., 2016; **Supplementary Table 1**). We used PAUP 4.0a167 (Swofford, 2019) in the phylogenetic analysis to derive trees based on parsimony and maximum likelihood (ML). Parsimony branch-and-bound search was performed on the alignment, in addition to bootstrap analysis employing a heuristic search with 1000 bootstrap replicates using 100 random stepwise addition (tree-bisection-reconnection, TBR). We defined values for bootstrap support >90% as strongly supported, 90–70% as well-supported/moderate support, and < 70% as weakly supported (Hwang and Weirauch, 2012). Characters were treated as unordered and of equal weight. The topology of the retained trees was tested with maximum likelihood (ML) using a model of estimated gamma distribution, HKY85 variant, unequal base frequencies, different substitution rates, empirical base frequencies, estimated substitution model following heuristic stepwise addition using TBR branch-swapping (de Paula et al., 2005, 2007), and molecular clock enforced. The best evolutionary model was selected under the corrected Akaike Information Criterion (AIC) (Swofford and Bell, 2017; Liu et al., 2018), likelihood estimation (−ln L), and Kishino-Hasegawa test using normal approximation (KH-test) (Swofford and Bell, 2017).

Nucleotide sequences have been the most common form of data used in molecular clock analyses. Several models using molecular information have been developed and implemented to infer divergence times, although they do not assume a strict molecular clock, and yet are commonly applied to empirical data. The divergence time estimate analysis was conducted using BEAST 2.6.2 (Bouckaert et al., 2019), which goal is to estimate the joint posterior probability of the branch rates and times given a set of sequences and calibration information. The gamma model was set to general time reversible (GTR), and base frequencies and lognormal relaxed clock (Uncorrelated) were estimated. We adjusted the scale parameter to reduce the prior density ($\alpha = 2$ and $\beta = 0.25$). The ucldMean and ucldStdev were set to exponential. Age estimates from fossil organisms are the most common form of divergence time calibration information. Fossils are part of the same diversification process (i.e., birth-death model) that generated the extant species and these data were used as age constraints on their putative ancestral nodes (Ho and Phillips, 2009; Heled and Drummond, 2010, 2012; Heath, 2012). Recent analyses using relaxed clock models calibrated with fossils converge estimates on the origin of Triatominae around 35 Mya to just over 40 Mya (Hwang and Weirauch, 2012; Ibarra-Cerdeña et al., 2014; Justi et al., 2016). We calibrated the origin of Salyavatinae to 42.0 Mya (Hwang and Weirauch, 2012) and lognormal prior distribution ($M = 1.0$ and $S = 1.25$). By default, BEAST sets the number of generations to 10,000,000 and length of chain to 1,000,000, thus these parameters were applied. The first 10% of the samples in each tree file were discarded using TREEANNOTATOR 2.6.2 (Bouckaert et al., 2019). The remaining trees were used to produce the maximum clade credibility tree that was visualized in FIGTREE 1.3.1 (Rambaut, 2018).

Ecological Niche Model

The potential distribution of each species was modeled using the MAXENT 3.4.1 (Phillips et al., 2018). The distribution data for Rhodniini was obtained from Ceccarelli et al. (2018) (see **Supplementary Table 2**). Georeferenced data were checked for inconsistencies using DIVA-GIS 7.5 (Hijmans et al., 2012), and duplicated data were removed. WorldClim bioclimatic variables version 2 (Fick and Hijmans, 2017) were used to perform the analyses. The rasters had a 2.5 min resolution, which represents the value of the 19 bioclimatic variables in the study area (see **Supplementary Table 3**). We used 10 replicate analyses based on cross validation. The distributional data for each species were separated into two sets: one for model calibration (70% of points) and one for model evaluation (30% of points) (Elith et al., 2010; Merow et al., 2013; de Paula and Barreto, 2020). The cloglog format was used to output the results in ASCII format, and the average output grids were imported into DIVA-GIS 7.5 (Hijmans et al., 2012) for assessments and analyses. To define the potential distribution for Rhodniini species, we used a 10% training presence threshold (the probability value at which 90% of the presence points fall within the potential area) to generate binary rasters with potential distribution areas for each individual species (Phillips et al., 2006, 2017; Phillips, 2017). We used the Area Under Curve (AUC) to evaluate models' performances. AUC values range from 0 (a model worse than random) to 1 (perfect discrimination) with a cut-off of 0.5, meaning the model does not have a better fit than random prediction models.

We compared potential distribution maps of related species with Kappa statistic tool in the MAP COMPARISON KIT (MCK) 3.2.1 (Visser and de Nijs, 2006; Research Institute for Knowledge Systems BV, 2010). The Kappa comparison method is based on a straightforward cell-by-cell map comparison, which considers whether each pair of cells on the two maps are equal or not. This results in a comparison map displaying the spatial distribution

[1]https://www.ebi.ac.uk/Tools/msa/muscle/

TABLE 1 | Test for significance of likelihood-score differences Kishino-Hasegawa test (KH) using normal approximation, two-tailed test.

			KH-test			
Tree	**−ln L**	**Diff−ln L**	**SD**	**T**	**P**	**AIC**
4	6101.613	(best)				12245.226
2	6104.225	2.61210	3.945	0.662	0.5080	12250.450
5	6104.380	2.76692	3.859	0.717	0.4735	12250.760
1	6110.201	8.58809	5.625	1.527	0.1270	12262.402
3	6110.204	8.59096	5.630	1.526	0.1272	12262.408

Values for KH tests is P–values for null hypothesis of no difference between trees.

of agreement. This comparison method does not require any parameters. The kappa statistic ranges from −1 to +1, where +1 indicates perfect agreement and values of zero or less indicate a performance no better than random (Cohen, 1960; van Vliet et al., 2011).

Species Discrimination

Linear discriminant function analysis (i.e., discriminant analysis) performs a multivariate test of differences between groups. This classification technique introduced by Fisher (1936) is based on the idea of discriminant function analysis to provide a set of weightings that allow the groups to be distinguished. It requires that the individuals be divided into groups and it is used when there are observations from pre-determined groups with two or more response variables recorded for each observation. It then generates a linear combination of variables that maximizes the probability of correctly assigning observations to their pre-determined groups. Linear discriminant function analysis was used to test the efficacy of the bioclimatic variables to classify the species of Rhodniini. Multivariate niches with all 19 bioclimatic variables were used to describe differences between the species (**Supplementary Table 4**). The values of bioclimatic variables were extracted from the WorldClim 2 layers (Fick and Hijmans, 2017). Twenty locations for each species were randomly selected, except for *R. brethesi* Matta, 1919 (12), *R. domesticus* (15), and *R. stali* Lent et al., 1993 (10), which selection comprised all locations. The discriminant analysis was performed using the stepwise method, Willks' Lambda (λ), and computing *F* values for pairwise distances. Wilks' λ is the ratio of within-groups sum of squares to the total sum of squares and it is designed to indicate whether a particular variable contributes significantly to explaining additional variance in the dependent variable. There are an *F* and a *p*-value associated with Wilks' λ that indicates the level of significance (Quinn and Keough, 2002; McLachlan, 2004). The analysis was performed in SPSS SOFTWARE 20 (IBM, 2011). The criteria for entry sets were *F* values of 3.84 to enter a new variable and 2.71 to remove an already entered variable. These *F* values represent significance levels of approximately 0.05 and 0.10, respectively. We defined *F* values from groups size option for prior probabilities for classification—the groups are weighted by the proportion of the number of cases in each group.

RESULTS

Sequence Alignment, Phylogenetic Analysis, and Molecular Dating

The parsimony analysis combining 16S + 28S + Wg sequences retained five trees with *CI* = 0.838, −ln L ranging from 6101.613 to 6110.204, and AIC from 12245.226 to 12262.408 (**Table 1**). The tree in **Figure 1A** shows the topology with the best scores

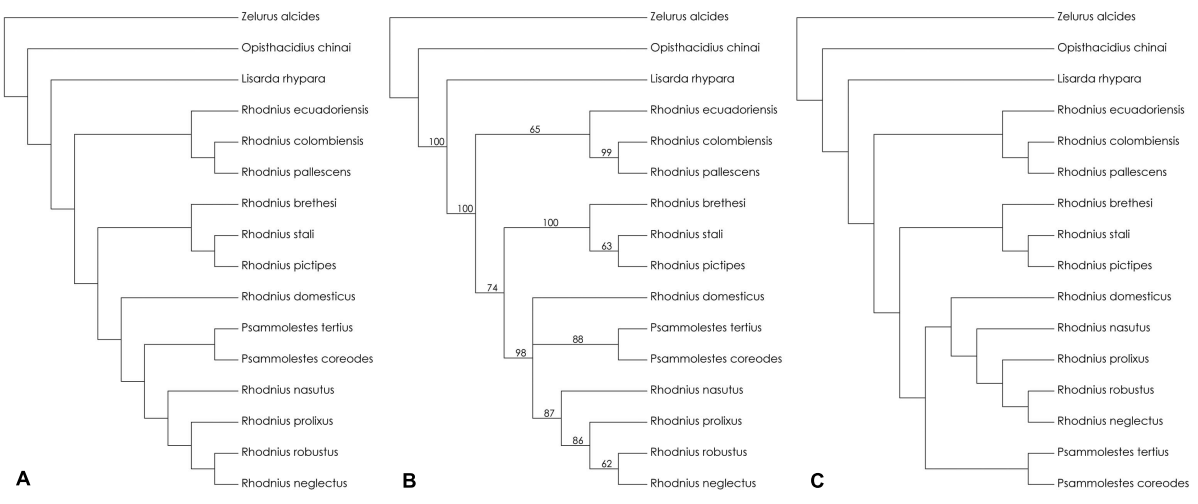

FIGURE 1 | The topology of the retained trees in the parsimony analysis combining 16S + 28S + wingless sequences retained five trees with CI = 0.838, –ln L ranging from 6101.613 to 6110.204, and AIC from 12245.226 to 12262.408 (see **Table 1**). **(A)** the topology with the best scores for the test for significance of likelihood-score differences—Kishino-Hasegawa test (KH). **(B)** bootstrap support values employing a heuristic search with 1,000 bootstrap replicates using 100 random stepwise addition (tree-bisection-reconnection, TBR). We defined values for bootstrap support > 90% as strongly supported, 90–70% as well-supported/moderate support, and <70% as weakly supported. **(C)** the topology showing the alternative position of the genus *Psammolestes*. It was used PAUP 4.0 in the phylogenetic analysis to derive trees based on parsimony.

for the test for significance of likelihood-score differences—Kishino-Hasegawa test (KH) (**Table 1**). The bootstrap values varied between 62 and 100, although the species *R. domesticus* + *Psammolestes* had no bootstrap support (**Figure 1B**). The stronger bootstrap value was seen for the *pictipes* group (100%). The *pallescens* group had weakly supported bootstrap values (65%) and the *prolixus* group was well-supported (87%). One of the five trees had a phylogenetic relationship where *Psammolestes* species—*P. tertius* and *P. coreodes*—were not retained in the clade *R. domesticus* + *prolixus* group (**Figure 1C**, −ln *L* = 6104.380, AIC = 12250.760). This topology was not showed in any publication prior to our study. The result of the maximum likelihood analysis (ML) is shown in **Figure 2** and showed better supports than the trees retained in the parsimony analysis (−ln *L* = 5895.714, AIC = 11833.428). The main difference in the topology of this tree is the split of the tribe into two evolutionary lineages—*pallescens* group + *pictipes* group, and *R. domesticus* + *Psammolestes* + *prolixus* group, contrasting the trees retained in the analysis of parsimony that showed three clades—*pallescens* group, *pictipes* group, and *R. domesticus* + *Psammolestes* + *prolixus* group. The results of our phylogenetic analyses did not show the species *Z. alcides* and *O. chinai* as closely related to Rhodniini, as previously found (Patterson and Gaunt, 2010; Hwang and Weirauch, 2012; Justi et al., 2016). The results support the hypothesis proposed in de Paula et al. (2005), where Salyavatinae was the sister clade to Rhodniini.

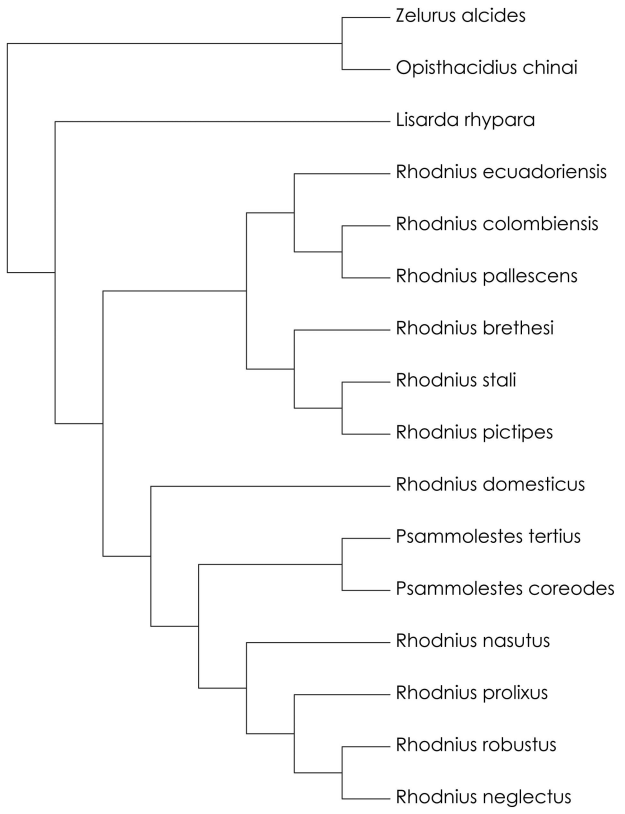

FIGURE 2 | The topology of the maximum likelihood analysis combining 16S + 28S + wingless sequences (–ln L = 5895.714, AIC = 11833.428). It was used PAUP 4.0 in the phylogenetic analysis to derive trees based maximum likelihood (ML).

The BEAST analysis (**Figure 3**) showed that *O. chinai* and *Z. alcides* are more recent than the origin of the Rhodniini and originated 12.91 Mya (95% credibility interval: 0.02; 0.67). This more recent dating for the species of the subfamily Reduviinae in our results reinforces the hypothesis that Salyavatinae is closely related to Rhodniini. Four main clades appear in this figure—*pallescens* group, *pictipes* group, *R. domesticus* + *Psammolestes*, and *prolixus* group. The posterior probability values are strong for most of the species of Rhodniini, ranging from 1 to 0.84.

Rhodniini are closely related to Salyavatinae that are specialist termite predators and diverged from this subfamily 30.43 Mya. The origin of the Rhodniini occurred 17.91 Mya (95% credibility interval: 0.02, 0.97) (**Figure 3**). Following the tribe origin, four main clades have emerged, the *pallescens* group (7.93 Mya, 95% credibility interval: 0.01, 0.47), the *pictipes* group (3.45 Mya, 0.00, 0.22), the *prolixus* group (6.92 Mya, 0.01, 0.39), and the clade *R. domesticus* + *Psammolestes* (9.13 Mya, 0.01, 0.48). In *pallescens* group, *R. ecuadoriensis* Lent and León, 1958 had its origin 7.93 Mya (0.01, 0.47), *R. colombiensis* Moreno Mejía et al., 1999, and *R. pallescens* Barber, 1932 both 1.09 Mya (0.00, 0.08). In *pictipes* group, *R. brethesi* had its origin 3.45 Mya (0.00, 0.22), and *R. pictipes* Stål, 1872 and *R. stali* originated 1.01 Mya (95% credibility interval: 0.00, 0.07). *R. domesticus* had 9.13 Mya (0.01, 0.48), *Psammolestes* species appeared 4.98 Mya (0.00, 0.27), *R. nasutus* 6.92 Mya (0.00, 0.20), *R. prolixus* 3.68 Mya (0.00, 0.13), and *R. neglectus* and *R. robustus* 2.26 Mya (0.00, 0.00). The *prolixus* group split up from *R. domesticus* 11.01 Mya (0.01, 0.60).

Ecological Niche Model

All ecological niche models had Training AUC values above 0.8 and Test AUC values above 0.8, which confirms the high sensitivity of the analysis (**Supplementary Table 5**). The average values of the 10 replicate analyses based on cross validation, considering 10% training presence binomial probability, were significant, except for the species *R. brethesi* (0.443), *R. colombiensis* (0.101), *R. domesticus* (0.154), and *R. stali* (0.298). In these species, the average result was replaced by the best results among the 10 replicate analyses—*R. brethesi* (0.004), *R. colombiensis* (0.000), *R. domesticus* (0.023), *R. stali* (0.094) (**Supplementary Table 5**). These results were used to generate binary rasters with potential distribution areas for each individual species, where 90% of the presence points fall within the potential area (**Figures 4–6**, white areas).

The distribution of the species *Rhodnius ecuadoriensis*, *R. colombiensis*, and *R. pallescens* are shown in **Figures 4A–C**, and these species form the *pallescens* group (trans-Andean). All these species have combinations of Kappa values below 0 (performance no better than random). The Fraction correct was above 0.95. These results indicate speciation by vicariance, where the species are clearly allopatric. The *pictipes* group includes *Rhodnius brethesi* (**Figure 4D**), *R. stali* (**Figure 4E**), and *R. pictipes* (**Figure 4F**) (cis-Andean). *Rhodnius brethesi* and *R. stali* are clearly allopatric (Kappa = −0.083, **Supplementary Table 6**). *Rhodnius pictipes* distribution overlapped with that of *R. brethesi*

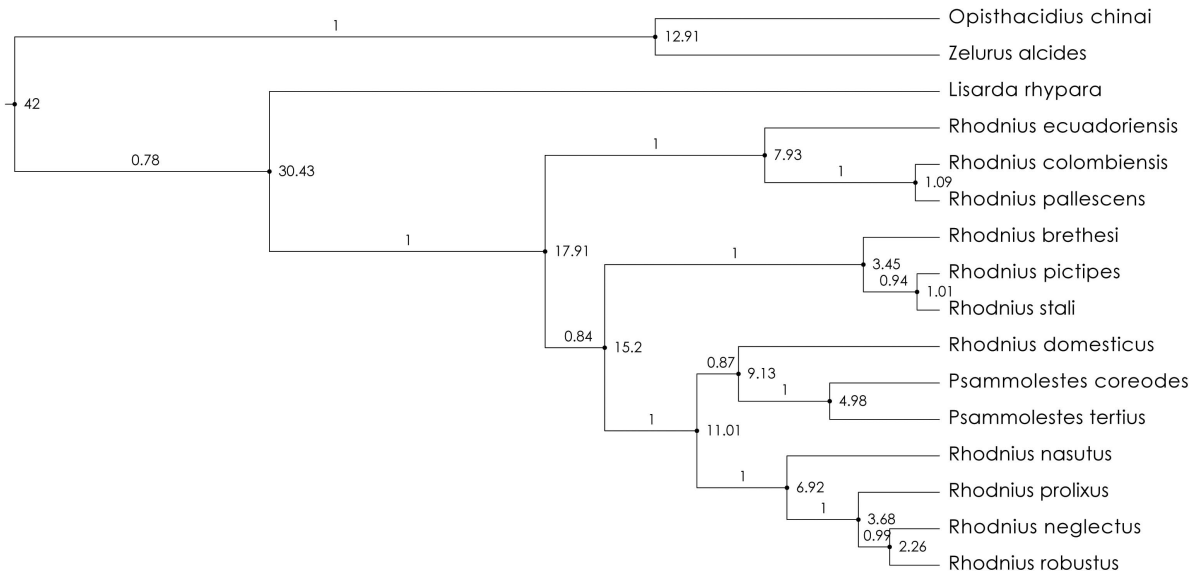

FIGURE 3 | Molecular clock analyses (BEAST 2.6.2) combining the 16S + 28S + wingless sequences, and using gamma model, generalized time reversible (GTR), and lognormal relaxed clock. The values in the nodes represents time in millions of years ago (Mya). The values above the branches are posterior probability.

(Kappa = 0.218) to the north of its distributional area. *Rhodnius pictipes* and *R. stali* are closely related but had little overlap in their areas of potential distribution (Kappa = −0.119). Based on this, we can assume that *R. stali* and *R. brethesi* originated by vicariance, and *R. pictipes* would have also originated by vicariance followed by post-speciation dispersal.

The distribution of *R. domesticus* is shown in **Figure 5A**, *P. tertius* Lent and Jurberg, 1965 in **Figure 5B** and *P. coreodes* Bergroth, 1911 in **Figure 5C**. The species *P. coreodes* and *P. tertius* are distinctively allopatric species (Kappa = −0.080). In this scenario, we can assume that *P. coreodes* and *P. tertius* originated by vicariance. *P. tertius* overlaps with *R. domesticus* (Kappa = 0.218) and *P. coreodes* and *R. domesticus* are clearly allopatric (Kappa = −0.006).

The *prolixus* group includes *R. nasutus*, *R. prolixus*, *R. robustus*, and *R. neglectus*. Their distributions are shown in **Figures 6A–D**. *R. nasutus* and *R. prolixus* are clearly allopatric (Kappa = −0.020). There is no overlap between the distribution areas of *R. nasutus* and *R. robustus* (Kappa = −0.016), and between *R. nasutus* and *R. neglectus* (Kappa = −0.006). Therefore, they are clearly allopatric. In addition, *R. prolixus* does not overlap with the areas of distribution of *R. robustus* (Kappa = −0.006) and *R. neglectus* (Kappa = −0.066). The other species in the group, *R. robustus* and *R. neglectus*, are allopatric (Kappa = −0.203). Consequently, all species of the *prolixus* group originated by vicariance.

Species Discrimination

Multivariate niches with all 19 bioclimatic variables were used to describe differences between the species. Discriminant dimensions for all Rhodniini species (**Figure 7**) were significantly different in the results of pairwise *F*-ratios for testing group differences (**Supplementary Table 7**). In total, 75.4% of original grouped cases were correctly classified, and 66.5% of cross-validated grouped cases were correctly classified. Wilks' λ showed $p < 0.05$ for all bioclimatic variables, indicating that they are indeed significant predictors for Rhodniini species (Tests of Equality of Group Means, **Supplementary Table 8**). Variables in species discrimination analysis after 12 steps (**Supplementary Table 9**) included Temperature seasonality, Precipitation seasonality, Isothermality, Mean temperature of driest quarter, Annual mean diurnal range, Annual precipitation, Precipitation of driest quarter, Precipitation of warmest quarter, Precipitation of wettest quarter, Precipitation of wettest month, Annual temperature range, and Mean temperature of wettest quarter. These results indicate that Rhodniini species clearly show distinct multivariate niches.

DISCUSSION

Historical Biogeography

The software TREEMAP was used by de Paula et al. (2007) to deduce taxon–area associations to propose a biogeographical hypothesis of the Rhodniini in the Neotropics. In this study, tanglegrams were used to show the relationship between the Neotropical areas of endemism and Rhodniini species. TREEMAP generated 12 optimal solutions with the lowest total cost. These optimal reconstructions required six vicariance events, 20 duplications (sympatry), at least three dispersals, and at least one extinction event. In our study, though, results indicate a parsimonious scenario comprising vicariance events. The origin of the Rhodniini tribe occurred 17.91 Mya (**Figure 3**), and this result is more recent than the study of Hwang and Weirauch (2012), which defined the origin of Rhodniini as 22.18 Mya. This divergence estimate is

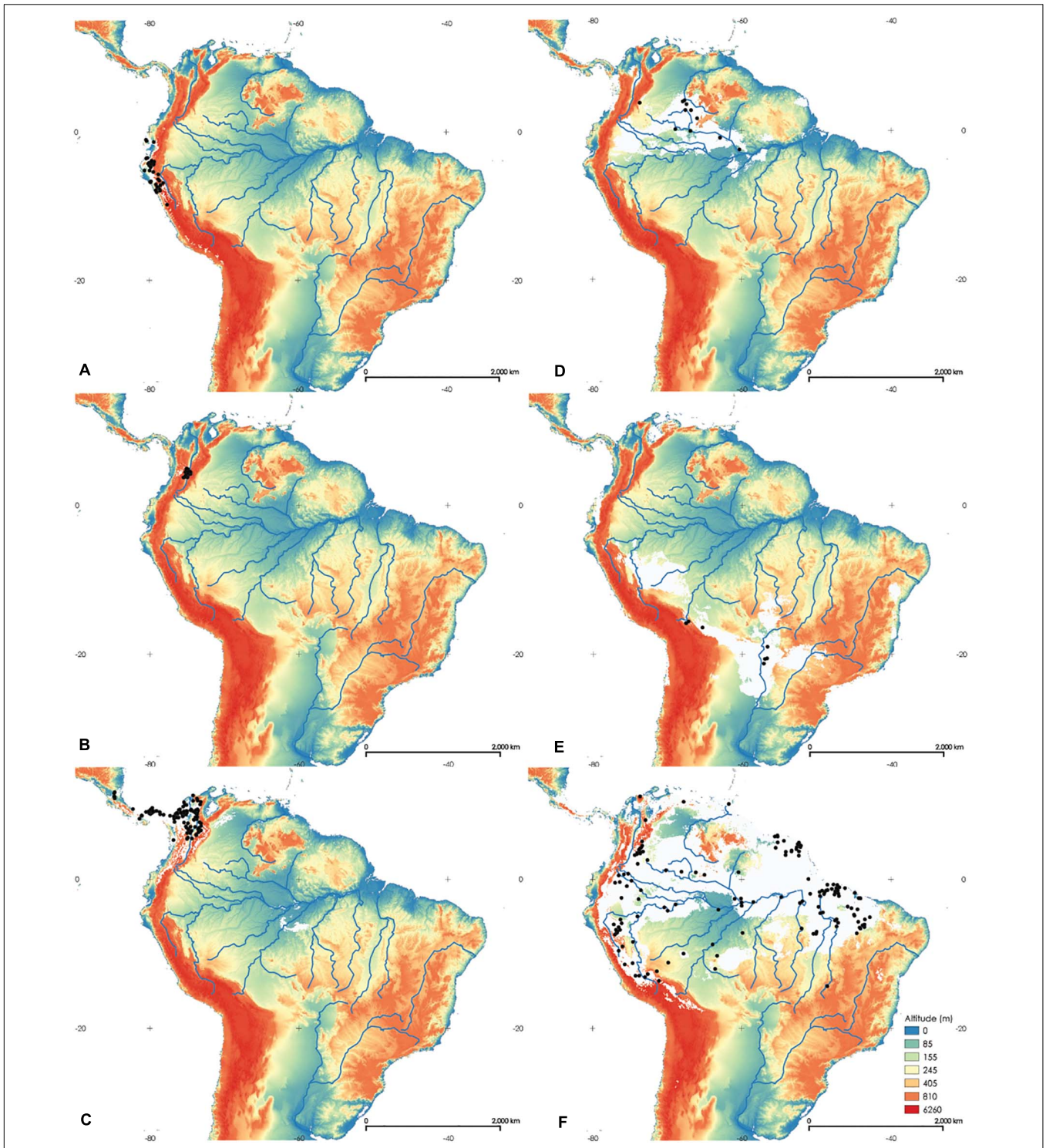

FIGURE 4 | Maps of the potential distribution of Rhodniini based on 19 bioclimatic variables in the study area. *pallescens* group—**(A)** *Rhodnius ecuadoriensis*, **(B)** *R. colombiensis*, **(C)** *R. pallescens*; *pictipes* group—**(D)** *R. brethesi*, **(E)** *R. stali*, **(F)** *R. pictipes*. The white areas comprise the probability value at which 90% of the presence points fall within the potential area. The black dots indicate records of each species.

much younger than the 107 Mya age proposed by Patterson and Gaunt (2010). Although molecular and biogeographic divergences support the diversification of the species of the Rhodniini tribe (as discussed below), karyotype and chromosomal homogeneity demonstrate that species of this tribe have not suffered structural changes in chromosomes

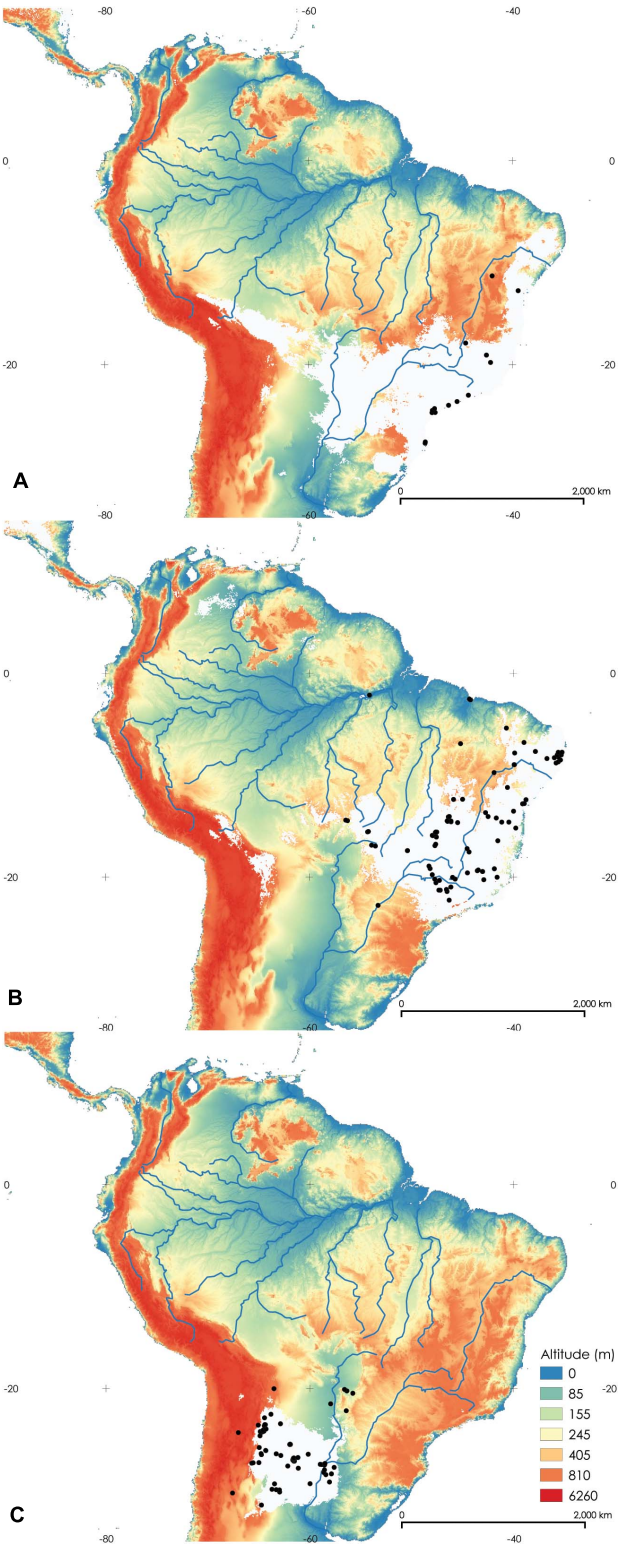

FIGURE 5 | Maps of the potential distribution of Rhodniini based on 19 bioclimatic variables in the study area. **(A)** *Rhodnius domesticus*, **(B)** *Psammolestes tertius*, and **(C)** *P. coreodes*. The white areas comprise the probability values at which 90% of the presence points fall within the potential area. The black dots indicate records of each species.

during speciation events (Alevi et al., 2018; Oliveira et al., 2018; Ravazi et al., 2021).

The formation of the Pebas system (23–10 Mya, Hoorn et al., 2010) split the clades that originate the *pallescens* (trans-Andean) and *pictipes* groups (cis-Andean) (17.91 Mya). The rapid uplift of the Eastern Cordillera between 5 and 2 Mya (Hoorn et al., 2010) may be related to the origin of the species *R. colombiensis* and *R. pallescens* (1.09 Mya). The tectonic uplifts of the Northern Andes, called Quechua 2 (9–8 Mya), may be related to the origin of *R. ecuadoriensis* (7.93 Mya). Hydrological connections between the Araguaia–Amazon basins in Early and Middle Miocene (20.4–9.0 Mya, Albert et al., 2018) would be a possible additional event that split the cis-Andean + trans-Andean clades of the *R. domesticus* + *Psammolestes* + *prolixus* group (15.2 Mya). This hydrological connection incorporated the Pantanal system in the Upper Miocene-Lower Miocene (10.0–4.5 Mya, Albert et al., 2018), resulting in the separation of Pantanal from the Atlantic Forest. This event may have contributed to the formation of *Psammolestes* species when *P. coreodes* and *P. tertius* were then originated (4.98 Mya).

New species establish themselves in new habitats under new climate regimes and can persist in a changing environment. *Rhodnius nasutus* occurs in the Borborema Province (NE Brazil), which consists of a complex mosaic of rocks composed mainly by Proterozoic metasedimentary belts that amalgamate Archean-to-Paleoproterozoic gneiss-migmatite complexes (Oliveira and Medeiros, 2018). The distribution of *R. nasutus* is bounded to the west by the Parnaíba River, which is trapped between the Amazonian craton and the Borborema orogenic belt (Daly et al., 2014). Arid biomes mainly expanded after 15 Mya ago by pronounced cooling events during the Miocene (10 Mya). Nearly all the dated shifts into the arid biome occurred after this time (Crisp et al., 2009). *Rhodnius nasutus* appeared 6.92 Mya (**Figure 3**) and occurs in Caatinga Steppe, an arid biome. We here propose that *R. nasutus* originated during cooling events, possibly when it colonized carnaúba palm trees [*Copernicia prunifera* (Miller)].

Rhodnius prolixus originated 3.68 Mya and its distribution area may be related to the hydrological connection between the Llanos-Amacuro basins that occurred between 10 and 4.5 Mya (Albert et al., 2018). *Rhodnius robustus* and *R. neglectus* were dated at 2.26 Mya and would probably be related to the hydrological connection between the Amazonas-Araguaia basins that occurred between 4.5 and 0 Mya (Albert et al., 2018). Lastly, the transcontinental Amazon Stage 2 (Albert et al., 2018) occurred since middle Pliocene to Recent (4.5–0 Mya), and may be related to the origin of *R. stali* (1.01 Mya).

Evolution of Hematophagy

Schofield (1988) proposed that different species and genera of triatomines have different ecological preferences and postulated that Triatominae may not represent a clade, but different hematophagous lineages that could have evolved independently of predatory ancestors that shared microhabitat. According to Hwang and Weirauch (2012), the evolution of blood-feeding may have occurred once or twice independently among predatory assassin bugs. All prey specialists evolved from

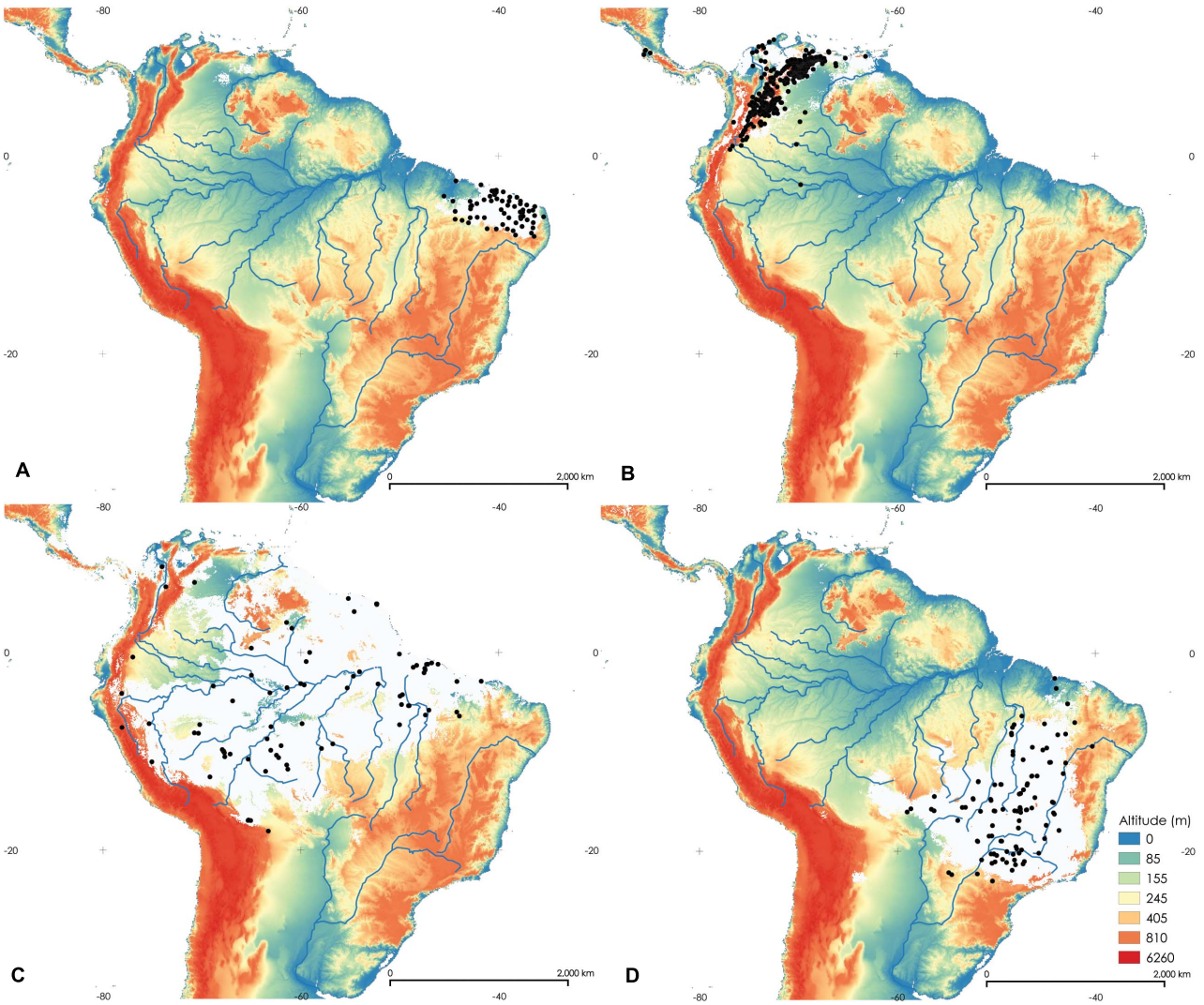

FIGURE 6 | Maps of the potential distribution of Rhodniini based on 19 bioclimatic variables in the study area. *prolixus* group—*Rhodnius nasutus* **(A)**, *R. prolixus* **(B)**, *R. robustus* **(C)**, **(D)** *R. neglectus*. The white areas comprise the probability value at which 90% of the presence points fall within the potential area. The black dots indicate records of each species.

generalist ancestors, with multiple evolutionary origins of termite and ant specializations. The hematophagous habit seen in Rhodniini is associated with major morphological, physiological and behavioral adaptations that have accumulated throughout the evolutionary history of the various lineages of blood-sucking arthropods (Lazzari et al., 2013). Hwang and Weirauch (2012) showed that the generalist predatory feeding strategy is ancestral for Reduviidae, and that all prey specialists evolved from generalist ancestors. The transition from predatory to hematophagous habits could have occurred by the colonization of vertebrate nests.

Ambrose (1999) suggested that the reduviids could be broadly divided into two groups based on whether or not they possessed tibial pads. Those without tibial pads live in tropical forest ecosystems and are known as timid predators that do not use their forelegs to capture prey, instead impaling prey items with their long rostra. Rain forest reduviids may have developed tibial pads and other features that made them more efficient predators when they migrated to deciduous scrub forest and other semi-arid habitats. Although the forelegs are more simplified in Triatominae, some species exhibit similar extendible adhesive organs (spongy pads) in both the fore and middle tibia of the adults, which allow them climbing smooth surfaces (Weirauch, 2007), a useful adaptation of hemipterans that fly and live in trees, as do species of the *Rhodnius* genus. Ambrose (1999) also considered the members of Salyavatinae ancestral to the subfamilies Triatominae and Ectrichodiinae. This hypothesis is supported by de Paula et al. (2005).

The pantropical Salyavatinae, subfamily used as outgroup in this study, are suspected specialist termite predators (Cobben, 1978; van Doesburg and Forero, 2012), and show morphological and camouflaging behavior that may represent adaptations to

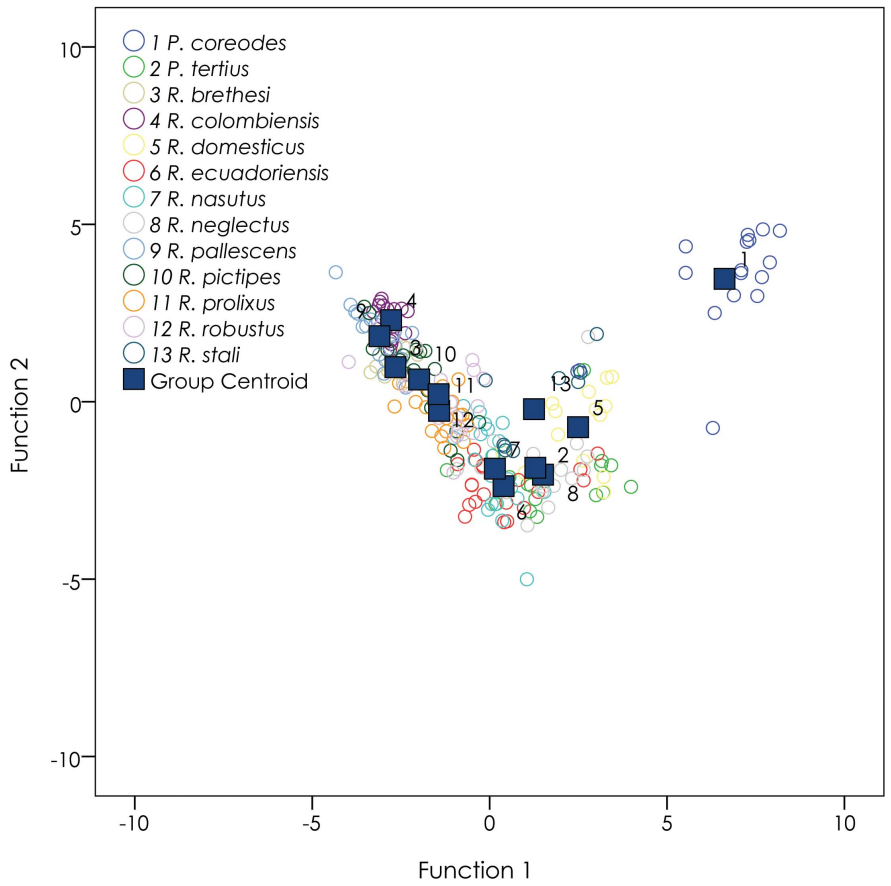

FIGURE 7 | Linear discriminant functional analysis for Rhodniini. All species were significantly different in the results of pairwise F-ratios for testing group differences (see **Supplementary Table 7**). 75.4% of original grouped cases were correctly classified. 66.5% of cross-validated grouped cases were correctly classified.

termite feeding (Weirauch, 2006). Considering Salyavatinae and Rhodniini closely related (de Paula et al., 2005, 2007), we propose a scenario where the Rhodniini's ancestor could be pre-adapted to the invasion of bromeliads and palm trees, where they would find significant water availability and thermal damping. Rhodniini are closely related to Salyavatinae that are specialist termite predators and diverged from this subfamily 30.43 Mya (**Figure 3**).

Dias et al. (2011) suggested that microclimate plays an important role in establishing a stable relationship between Rhodniini species and palm trees, further influencing the adaptation of the bugs to different palm genera. Rhodniini expresses a clear thermo and hygropreference, such as amenable microclimates in terms of temperature and relative humidity (Lazzari et al., 2013). Furthermore, Rhodniini species need to find adequate shelter for protection from daylight, which can make them vulnerable to potential predators. Besides that, food source detection in Rhodniini is achieved via the detection of air currents carrying distinctive odors, water vapors and heat (Lazzari et al., 2013). Natural habitats (burrows, nests, hollow tree-trunks) provide an abundance of shelters, although many of them may not offer suitable conditions. Shelters can be subject to high temperatures, humidity and intense illumination or may not be inhabited by colonies of conspecifics. We believe that the transition from the predatory lifestyle to hematophagy may have been a gradual process that involved specialist termite predators who sought shelter in vertebrate nests and then opportunistically fed on arthropods that inhabit nests. In laboratory conditions, *R. prolixus* nymphs showed signs of cannibalism and cleptohematophagy (Piñero et al., 1988).

The species of the tribe Rhodniini included in this study have clearly distinct bioclimatic niches (**Figure 7** and **Supplementary Table 7**). Consequently, the colonization of bromeliads, palm trees and nests represent important events for the speciation of these taxa. *Rhodnius domesticus* can be found in bromeliads, rodents and marsupial refuges, hollow trees, and under barks (Lent and Wygodzinsky, 1979; Monteiro et al., 2018). The results of our study indicate that *R. domesticus* is the oldest species of Rhodniini, having its origin at 9.13 Mya (**Figure 3**). In this scenario, the invasion of palm trees would have occurred only once, with subsequent colonization of bromeliads and bird nests in the *R. domesticus* + *Psammolestes* lineage. These clades had their origins at 15.2 Mya (*pictipes*), 7.93 Mya (*pallescens*), 11.01 Mya (*prolixus* + *R. domesticus* + *Psammolestes*), respectively (see **Figure 3**).

Psammolestes species are adapted to explore bird nest microhabitats, and seem to use palms opportunistically (Abad-Franch et al., 2015). *Psammolestes* species are wild and found in bird nests, mostly from birds in Furnariidae family (Lent and Wygodzinsky, 1979). For instance, the nest of *Furnarius rufus* (Gmelin) (Furnariidae) is a domed mud structure, with a clear separation between the breeding chamber and the outside. In general, each couple builds one nest per year, but each nest is used for one clutch or two consecutive clutches in the same breeding season. Some nests can remain in the field for 2 or 3 years, but some have a longer permanence (up to more than 8 years). Thus, the nests are widely used by other vertebrate inquilines too (Turienzo and di Iorio, 2010), which would be the source of food for the species of *Psammolestes,* in addition to birds. The species of *Psammolestes*, *P. coreodes* and *P. tertius*, have an origin dated 4.98 Mya (see **Figure 3**).

The results of this study indicate that the Rhodniini tribe originated 17.91 Mya ago, with *R. domesticus* being the oldest species, originated 9.13 Mya. The biogeographic history of the studied species comprised vicariance events. Rhodniini species clearly have distinct bioclimatic niches, and the colonization of bromeliads, palm trees and bird nests also represent important events for the speciation of these taxa. Their hematophagous habit can be described as a scenario where the ancestor of Rhodniini could have been pre-adapted to the invasion of bromeliads, palm trees, and bird nests, where it would have had available water and thermal damping. These habitats are used by vertebrates that the ancestor would have fed on. These environments are widely used by vertebrate inquilines that would be the food source for Rhodniini.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

ASDP: conceptualization, methodology, formal analysis, and writing. CB: methodology, formal analysis, and writingŮreview and editing. MT: methodology, formal analysis, and writing. LD and CG: conceptualization and writing. All authors contributed to the article and approved the submitted version.

FUNDING

Initial phases of this work were supported by the Brazilian National Council for Scientific and Technological Development (grant nos. PIBIC 129897/2009-4 and PIBIC 117649/2010-4 to CB).

ACKNOWLEDGMENTS

Special thanks to Dr. Danon Clemes Cardoso (Laboratório de Genética Evolutiva e de Populações—UFOP) for critical reading and suggestions in the first draft of the manuscript. We also thank Matthew Meehan (Western University) for final comments, and JADR and AR for their constructive criticism and time. CB and MT thank UFOP for the support of their undergraduate thesis research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.660151/full#supplementary-material

REFERENCES

Abad-Franch, F., Lima, M. M., Sarquis, O., Gurgel-Gonçalves, R., Sánchez-Martín, M., Calzada, J., et al. (2015). On palms, bugs, and Chagas disease in the Americas. *Acta Trop.* 151, 126–141. doi: 10.1016/j.actatropica.2015.07.005

Abad-Franch, F., Monteiro, F. A., Jaramillo, O. N., Gurgel-Gonçalves, R., Dias, F. B. S., and Diotaiuti, L. (2009). Ecology, evolution, and the long-term surveillance of vector-borne Chagas disease: a multi-scale appraisal of the tribe Rhodniini (Triatominae). *Acta Trop.* 110, 159–177. doi: 10.1016/j.actatropica.2008.06.005

Albert, J. S., Val, P., and Hoorn, C. (2018). The changing course of the amazon river in the Neogene: center stage for Neotropical diversification. *Neotrop. Ichthyol.* 16:e180033. doi: 10.1590/1982-0224-20180033

Alevi, K. C. C., de Oliveira, J., da Rosa, J. A., and de Azeredo-Oliveira, M. T. V. (2018). Karyotype evolution of Chagas disease vectors (Hemiptera, Triatominae). *Am. J. Trop. Med. Hyg.* 99, 87–89. doi: 10.4269/ajtmh.17-0166

Ambrose, D. P. (1999). *Assassin Bugs*. Enfield, NH: Science Publishers.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., et al. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 15:e1006650. doi: 10.1371/journal.pcbi.1006650

Castro, M. R. J., Goubert, C., Cararetto, C. M. A., Monteiro, F. A., and Vieira, C. (2020). Homology-free detection of transposable elements unveils their dynamics in three ecologically distinct *Rhodnius* species. *Genes* 11:170. doi: 10.3390/genes11020170

Ceccarelli, S., Balsalobre, A., Medone, P., Cano, M. E., Gonçalves, R. G., Feliciangeli, D., et al. (2018). Data descriptor: DataTri, a database of American triatomine species occurrence. *Sci. Data* 5:180071. doi: 10.1038/sdata.2018.71

Cobben, R. H. (1978). Evolutionary trends in heteroptera. Part II. Mouth-part structures and feeding strategies. *Meded. Landbouwhogesch. Wageningen* 78, 1–407.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* XX, 37–46.

Crisp, M. D., Arroyo, M. T. K., Cook, L. G., Gandolfo, M. A., Jordan, G. J., McGlone, M. S., et al. (2009). Phylogenetic biome conservatism on a global scale. *Nature* 458, 754–756. doi: 10.1038/nature07764

Daly, M. C., Andrade, V., Barousse, C. A., Costa, R., McDowell, K., Piggott, N., et al. (2014). Brasiliano crustal structure and the tectonic setting of the Parnaíba basin of NE Brazil: results of a deep seismic reflection profile. *Tectonics* 33, 2102–2120. doi: 10.1002/2014TC003632

de Paula, A. S., and Barreto, C. (2020). Potential distribution of *Nysius simulans* (Stål) (Hemiptera: Lygaeidae) in soybean crops in South America under current and future climate. *J. Econ. Entomol.* 113, 1702–1710. doi: 10.1093/jee/toaa089

de Paula, A. S., Diotaiuti, L., and Galvão, C. (2007). Systematics and biogeography of Rhodniini (Heteroptera: Reduviidae: Triatominae) based on

16S mitochondrial rDNA sequences. *J. Biogeogr.* 34, 699–712. doi: 10.1111/j.1365-2699.2006.01628.x

de Paula, A. S., Diotaiuti, L., and Schofield, C. J. (2005). Testing the sister-group relationship of the Rhodniini and Triatomini (Insecta: Hemiptera: Reduviidae: Triatominae). *Mol. Phylogenet. Evol.* 35, 712–718. doi: 10.1016/j.ympev.2005.03.003

Dias, F. B. S., de Paula, A. S., Belisário, C. J., Lorenzo, M. G., Bezerra, C. M., Harry, M., et al. (2011). Influence of the palm tree species on the variability of *Rhodnius nasutus* Stål, 1859 (Hemiptera, Reduviidae, Triatominae). *Infect. Genet. Evol.* 11, 869–877. doi: 10.1016/j.meegid.2011.02.008

dos Santos Souza, E., von Atzingen, N. C. B., Furtado, M. B., de Oliveira, J., Nascimento, J. D., Vendrami, D. P., et al. (2016). Description of *Rhodnius marabaensis* sp. n. (Hemiptera, Reduviidade, Triatominae) from Pará State, Brazil. *ZooKeys* 621, 45–62. doi: 10.3897/zookeys.621.9662

Edgar, R. C. (2004a). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. doi: 10.1186/1471-2105-5-113

Edgar, R. C. (2004b). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Elith, J., Kearney, M., and Phillips, S. (2010). The art of modelling range-shifting species. *Methods Ecol. Evol.* 1, 330–342. doi: 10.1111/j.2041-210x.2010.00036.x

Fick, S. E., and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi: 10.1002/joc.5086

Fisher, R. A. (1936). The use of multiple measurement in taxonomic problems. *Ann. Eugen.* 7, 179–188.

Galvão, C. (2014). *Vetores da Doença de Chagas no Brasil*. Curitiba: Sociedade Brasileira de Zoologia, 289.

Galvão, C. (2020). "Taxonomia dos vetores da doença de Chagas da forma à molécula, quase três séculos de história," in *Atualidades em Medicina Tropical no Brazil: Vetores*, eds J. de Oliveira, K. C. C. Alevi, L. M. A. Camargo, and D. U. O. Meneguetti (Rio Branco: Stricto Sensu), 9–37. doi: 10.35170/ss.ed.9786586283129.01

Gaunt, M., and Miles, M. (2002). An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Mol. Biol. Evol.* 19, 748–761.

Heath, T. A. (2012). A hierarchical Bayesian model for calibrating estimates of species divergence times. *Syst. Biol.* 61, 793–809. doi: 10.1093/sysbio/sys032

Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570–580. doi: 10.1093/molbev/msp274

Heled, J., and Drummond, A. J. (2012). Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Syst. Biol.* 61, 138–149. doi: 10.1093/sysbio/syr087

Hernández, C., da Rosa, J. A., Vallejo, G. A., Guhl, F., and Ramírez, J. D. (2020). Taxonomy, evolution, and biogeography of the Rhodniini tribe (Hemiptera: Reduviidae). *Diversity* 12:97. doi: 10.3390/d12030097

Hijmans, R. J., Guarino, L., and Mathur, P. (2012). *DIVA-GIS: A Geographic Information System for the Analysis of Species Distribution Data*. California. Patent No. 5,710,835.

Ho, S. Y. W., and Phillips, M. J. (2009). Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. *Syst. Biol.* 58, 367–380. doi: 10.1093/sysbio/syp035

Hoorn, C., Wesselingh, F. P., ter Steege, H., Bermudez, M. A., Mora, A., Sevink, J., et al. (2010). Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. *Science* 330, 927–931. doi: 10.1126/science.1194585

Hwang, W. S., and Weirauch, C. (2012). Evolutionary history of assassin bugs (Insecta: Hemiptera: Reduviidae): insights from divergence dating and ancestral state reconstruction. *PLoS One* 7:e45523. doi: 10.1371/journal.pone.0045523

Hypša, V., Tietz, D., Zrzavý, J., Rego, R., Galvão, C., and Jurberg, J. (2002). Phylogeny and biogeography of Triatominae (Hemiptera: Reduviidae): molecular evidence of a New World origin of the Asiatic clade. *Mol. Phylogenet. Evol.* 23, 447–457.

Ibarra-Cerdeña, C. N., Zaldívar-Riverón, A., Peterson, A. T., Sánchez-Cordero, V., and Ramsey, J. M. (2014). Phylogeny and niche conservatism in North and Central American Triatomine bugs (Hemiptera: Reduviidae: Triatominae), vectors of Chagas' disease. *PLoS Negl. Trop. Dis.* 8:e3266. doi: 10.1371/journal.pntd.0003266

IBM (2011). *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Cor.

Jurberg, J., Rocha, D., da, S., and Galvão, C. (2009). *Rhodnius zeledoni* sp. nov. afim de *Rhodnius paraensis* Sherlock, Guitton & Miles, 1977 (Hemiptera, Reduviidae, Triatominae). *Biota Neotrop.* 9, 123–128. doi: 10.1590/s1676-06032009000100014

Justi, S. A., Galvão, C., and Schrago, C. G. (2016). Geological changes of the Americas and their influence on the diversification of the Neotropical kissing bugs (Hemiptera: Reduviidae: Triatominae). *PLoS Negl. Trop. Dis.* 10:e0004527. doi: 10.1371/journal.pntd.0004527

Lazzari, C. R., Pereira, M. H., and Lorenzo, M. G. (2013). Behavioural biology of Chagas disease vectors. *Mem. Inst. Oswaldo Cruz* 108, 34–47. doi: 10.1590/0074-0276130409

Lee, B. Y., Bacon, K. M., Bottazzi, M. E., and Hotez, P. J. (2013). Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect. Dis.* 13, 342–348. doi: 10.1016/S1473-3099(13)70002-1

Lent, H., and Wygodzinsky, P. (1979). Revision of the Triatominae (Hemiptera, Reduviidae) and their significance as vectors of Chagas disease. *Bull. Am. Mus. Nat. Hist.* 163:3.

Liu, Y., Song, F., Jiang, P., Wilson, J. J., Cai, W., and Li, H. (2018). Compositional heterogeneity in true bug mitochondrial phylogenomics. *Mol. Phylogenet. Evol.* 118, 135–144. doi: 10.1016/j.ympev.2017.09.025

Lomolino, M. V., Riddle, B. R., and Whittaker, R. J. (2016). *Biogeography, Biological Diversity Across Space and Time*, 5th Edn. Sunderland, MA: Sinauer Associates, Inc.

McLachlan, G. J. (2004). *Discriminant Analysis and Statistical Pattern Recognition*. New York, NY: J. W. & Sons.

Merow, C., Smith, M. J., and Silander, J. A. Jr. (2013). A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography* 36, 1058–1069. doi: 10.1111/j.1600-0587.2013.07872.x

Mesquita, R. D., Vionette-Amaral, R. J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F. A., Minx, P., et al. (2015). Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14936–14941. doi: 10.1073/pnas.1506226112

Monteiro, F. A., Barrett, T. V., Fitzpatrick, S., Cordon-Rosales, C., Feliciangeli, D., and Beard, C. B. (2003). Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R. robustus*. *Mol. Ecol.* 12, 997–1006. doi: 10.1046/j.1365-294X.2003.01802.x

Monteiro, F. A., Weirauch, C., Felix, M., Lazoski, C., and Abad-Franch, F. (2018). Evolution, systematics, and biogeography of the Triatominae, vectors of Chagas disease. *Adv. Parasitol.* 99, 265–344. doi: 10.1016/bs.apar.2017.12.002

Nascimento, J. D., da Rosa, J. A., Salgado-Roa, F. C., Hernández, C., Pardo-Diaz, C., Alevi, K. C. C., et al. (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:e0211285. doi: 10.1371/journal.pone.0211285

Neiva, A., and Pinto, C. (1923). Dos hemipteros hematophagos do Norte do Brasil, com a descrição de duas novas especies. *Bras. Med.* 37, 73–76.

Oliveira, J., Alevi, K. C. C., Ravazi, A., Herrera, H. M., Santos, F. M., de Azeredo-Oliveira, M. T. V., et al. (2018). New evidence of the monophyletic relationship of the genus *Psammolestes* Bergroth, 1911 (Hemiptera, Reduviidae, Triatominae). *Am. J. Trop. Med. Hyg.* 99, 1485–1488. doi: 10.4269/ajtmh.18-0109

Oliveira, R. G., and Medeiros, W. E. (2018). Deep crustal framework of the Borborema Province, NE Brazil, derived from gravity and magnetic data. *Precambrian Res.* 315, 45–65. doi: 10.1016/j.precamres.2018.07.004

Otálora-Luna, F., Pérez-Sánchez, A. J., Sandoval, C., and Aldana, E. (2015). Evolution of hematophagous habit in Triatominae (Heteroptera: Reduviidae). *Rev. Chil. Hist. Nat.* 88:4. doi: 10.1186/s40693-014-0032-0

PAHO/WHO (2020). *Chagas Disease*. Available online at: https://www.paho.org/en/topics/chagas-disease (accessed September 6, 2020).

Patterson, J. S., and Gaunt, M. W. (2010). Phylogenetic multi-locus codon models and molecular clocks reveal the monophyly of haematophagous reduviid bugs and their evolution at the formation of South America. *Mol. Phylogenet. Evol.* 56, 608–621. doi: 10.1016/j.ympev.2010.04.038

Pavan, M. G., Mesquita, R. D., Lawrence, G. G., Lazoski, C., Dotson, E. M., Abubucker, S., et al. (2013). A nuclear single-nucleotide polymorphism (SNP) potentially useful for the separation of *Rhodnius prolixus* from

members of the *Rhodnius robustus* cryptic species complex (Hemiptera: Reduviidae). *Infect. Genet. Evol.* 14, 426–433. doi: 10.1016/j.meegid.2012.10.018

Phillips, S. J. (2017). *A Brief Tutorial on MaxEnt*. AT&T Research. Available online at: http://biodiversityinformatics.amnh.org/open_source/maxent/ (accessed March 15, 2020).

Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., and Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography* 40, 887–893. doi: 10.1111/ecog.03049

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

Phillips, S. J., Dudík, M., and Schapire, R. E. (2018). *Maxent Software for Modelling Species Niches and Distributions. (Version 3.4.1)*. Available online at: http://biodiversityinformatics.amnh.org/open_source/maxent/ (accessed December 16, 2020).

Piñero, D. F., Carcavallo, R. U., and Fernandez, E. (1988). Canibalismo y transmisión directa de *Trypanosoma cruzi* entre ninfas de *Rhodnius prolixus*. *Chagas* 5, 18–22.

Quinn, G. P., and Keough, M. J. (2002). *Experimental Design and Data Analysis for Biologists*. New York, NY: Cambridge University Press.

Rambaut, A. (2018). *FigTree*. Available online at: http://tree.bio.ed.ac.uk/software/figtree (accessed November 25, 2018).

Ravazi, A., Olaia, N., de Oliveira, J., Souza, E. D. S., da Rosa, J. A., de Azeredo-Oliveira, M. T. V., et al. (2021). Revisiting the chromosomal diversification of the genus *Rhodnius* (Stål, 1859) (Hemiptera, Triatominae). *Am. J. Trop. Med. Hyg.* 104, 656–658. doi: 10.4269/ajtmh.20-0875

Research Institute for Knowledge Systems BV (2010). *Map Comparison Kit 3 - User Manual*. Maastricht: Research Institute for Knowledge Systems BV, 53.

Rosa, J. A., Justino, H. H. G., Nascimento, J. D., Mendonça, V. J., Rocha, C. S., de Carvalho, D. B., et al. (2017). A new species of *Rhodnius* from Brazil (Hemiptera, Reduviidae, Triatominae). *ZooKeys* 675, 1–25. doi: 10.3897/zookeys.675.12024

Sandoval, C. M., Duarte, R., Gutíerrez, R., da Silva Rocha, D., Angulo, V. M., Esteban, L., et al. (2004). Feeding sources and natural infection of *Belminus herreri* (Hemiptera, Reduviidae, Triatominae) from dwellings in Cesar, Colombia. *Mem. Inst. Oswaldo Cruz* 99, 137–140. doi: 10.1590/S0074-02762004000200004

Sandoval, C. M., Ortiz, N., Jaimes, D., Lorosa, E., Galvão, C., Rodriguez, O., et al. (2010). Feeding behaviour of *Belminus ferroae* (Hemiptera: Reduviidae), a predaceous Triatominae colonizing rural houses in Norte de Santander, Colombia. *Med. Vet. Entomol.* 24, 124–131. doi: 10.1111/j.1365-2915.2010.00868.x

Schofield, C. J. (1988). "Biosystematics of the Triatominae," in *Biosystematics of Haematophagous Insects*, ed. M. W. Service (Oxford: Clarendon Press), 284–312.

Schofield, C. J., and Galvão, C. (2009). Classification, evolution, and species groups within the Triatominae. *Acta Trop.* 110, 88–100. doi: 10.1016/j.actatropica.2009.01.010

Swofford, D. L. (2019). *PAUP*. Available online at: http://phylosolutions.com/paup-test/ (accessed November 20, 2019).

Swofford, D. L., and Bell, C. D. (2017). *PAUP Manual* (accessed December 16, 2020).

Turienzo, P., and di Iorio, O. (2010). Insects found in birds' nests from Argentina. *Furnarius rufus* (Gmelin, 1788) (Aves: Furnariidae) and their inquiline birds, the true hosts of *Acanthocrios furnarii* (Cordero & Vogelsang, 1928) (Hemiptera: Heteroptera: Cimicidae). *Zootaxa* 2700, 1–112.

van Doesburg, P. H., and Forero, D. (2012). Revision of the genus *Salyavata* Amyot & Audinet-Serville, 1843 (Heteroptera: Reduviidae: Salyavatinae). *Zool. Meded.* 86, 515–568.

van Vliet, J., Bregt, A. K., and Hagen-Zanker, A. (2011). Revisiting Kappa to account for change in the accuracy assessment of land-use change models. *Ecol. Model.* 222, 1367–1375. doi: 10.1016/j.ecolmodel.2011.01.017

Visser, H., and de Nijs, T. (2006). The map comparison kit. *Environ. Model. Softw.* 21, 346–358. doi: 10.1016/j.envsoft.2004.11.013

Weirauch, C. (2006). Anatomy of disguise: camouflaging structures in nymphs of some Reduviidae (Heteroptera). *Am. Mus. Novit.* 3542, 1–18.

Weirauch, C. (2007). Hairy attachment structures in Reduviidae (Cimicomorpha, Heteroptera), with observations on the fossula spongiosa in some other Cimicomorpha. *Zool. Anz.* 246, 155–175. doi: 10.1016/j.jcz.2007.03.003

Zhao, Y., Galvão, C., and Cai, W. (2021). *Rhodnius micki*, a new species of Triatominae (Hemiptera, Reduviidae) from Bolivia. *ZooKeys* 1012, 71–93. doi: 10.3897/zookeys.1012.547791

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 06 July 2021
doi: 10.3389/fevo.2021.707642

Evidence of Elevational Speciation in *Kerteszia cruzii* (Diptera: Culicidae) in the Ribeira Valley, São Paulo, Brazil

Bruna Demari-Silva[1], Gabriel Zorello Laporta[2], Tatiane Marques Porangaba de Oliveira[1] and Maria Anice Mureb Sallum[1]*

[1] Faculdade de Saúde Pública, Departamento de Epidemiologia, Universidade de São Paulo, São Paulo, Brazil, [2] Centro Universitário FMABC da Fundação ABC, Setor de Pós-graduação, Pesquisa e Inovação, Santo André, Brazil

Kerteszia cruzii [former *Anopheles* (*Kerteszia*) *cruzii*] is a bromeliad mosquito widespread in the Brazilian Atlantic rainforest. In South-eastern Brazil, it plays an important role in malaria transmission because it was infected with at least four *Plasmodium* species. There is robust evidence that *Ke. cruzii* is a species complex. We used single nucleotide polymorphisms (SNPs) from a nextRAD sequence (nextera-tagmented, reductively amplified DNA) to investigate the genetic structure of *Ke. cruzii* in the Ribeira Valley, South-eastern Brazil. Furthermore, we verified whether the genetic structure was associated with forest cover, elevation, slope, and vegetation physiognomy. Our results showed two distinct lineages in the studied region associated with elevation and isolation by distance. The first lineage included samples from coastal localities and the second comprised specimens from inland or mountain sites. At one sampling locality (Esteiro do Morro in Cananéia municipality), both lineages are sympatric. These results are in accordance with previously published data that showed elevated stratification in *Ke. cruzii*. However, *Fst* values did not indicate the existence of cryptic or sister species in *Ke. cruzii* in this region, we concluded that elevational speciation probably occurs, and we hypothesized that differences in population structure found might be associated with the distribution of bromeliad species.

Keywords: ***Kerteszia cruzii*, population structure, landscape, isolation by distance, elevational speciation**

OPEN ACCESS

Edited by:
Kaio Cesar Chaboli Alevi, São Paulo State Universty, Brazil

Reviewed by:
Hélcio Reinaldo Gil-Santana, Oswaldo Cruz Foundation (Fiocruz), Brazil
Carlos José Carvalho Pinto, Federal University of Santa Catarina, Brazil

****Correspondence:***
Bruna Demari-Silva
bdemari.silva@gmail.com

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Ecology and Evolution

Received: *10 May 2021*
Accepted: *01 June 2021*
Published: *06 July 2021*

Citation:
Demari-Silva B, Laporta GZ, de Oliveira TMP and Sallum MAM (2021) Evidence of Elevational Speciation in Kerteszia cruzii (Diptera: Culicidae) in the Ribeira Valley, São Paulo, Brazil. Front. Ecol. Evol. 9:707642. doi: 10.3389/fevo.2021.707642

INTRODUCTION

Kerteszia cruzii occurs in areas of the Brazilian Atlantic rainforest and is abundant, depending on the abundance of the bromeliad phytotelmata. Formerly, *Kerteszia* was classified as a subgenus of the genus *Anopheles* until 2017, when the results of a phylogenetic analysis of the complete mitochondrial genome consistently showed *Kerteszia* as a monophyletic taxon placed outside the genus *Anopheles* (Foster et al., 2017). Consequently, Foster et al. (2017) elevated *Kerteszia*, *Lophopodomyia*, *Stethomyia*, and *Nyssorhynchus* to the genus level.

Females of *Ke. cruzii* are primarily sylvatic, presenting a low degree of sinanthropy (Forattini, 2002), can feed on humans, and are more abundant at the edges of forests where they usually perform their blood repast (Forattini et al., 1986; Medeiros-Sousa et al., 2019). Although the females of some populations are capable of blood-feeding at any time of the day, they usually present two activity peaks, one at twilight and the other at dawn (Forattini et al., 1986). Females

that feed at dusk were found to live longer and, therefore, are more likely to survive and exceed the extrinsic incubation period of *Plasmodium* (Dalla Bona and Navarro-Silva, 2010).

The epidemiological importance of this species in the Atlantic Forest is unquestionable. In the seventies, this species was naturally infected with *Plasmodium brasilianum* and *Plasmodium simium* (Deane et al., 1970). Ever since, *Ke. cruzii* was found naturally infected with *Plasmodium vivax* and *P. vivax* VK247, *Plasmodium falciparum*, and *Plasmodium malariae* in São Paulo and Espírito Santo states, Brazil (Branquinho et al., 1997; Duarte et al., 2013; Kirchgatter et al., 2014; Laporta et al., 2015; Buery et al., 2018; Demari-Silva et al., 2020). Recently, studies have demonstrated that *P. vivax* and *P. simium* are almost indistinguishable and can infect humans (Brasil et al., 2017; de Alvarenga et al., 2018).

Several studies have suggested that *Ke. cruzii* is a complex of species. Analyses of the banding pattern of the ovarian polytene chromosome showed that this species encompasses three sibling species in the Boracéia and Juquitiba municipalities in São Paulo state, Brazil (Ramírez and Dessen, 2000a; Ramírez and Dessen, 2000b). A broader study using samples from Rio de Janeiro (RJ), São Paulo (SP), Santa Catarina (SC), and Bahia (BA) states in Brazil demonstrated that the specimens from Bahia state are genetically distant from the remaining populations analyzed from the South and Southeast of Brazil (Carvalho-Pinto and Lourenço-de-Oliveira, 2004).

Subsequently, studies employing the *timeless* nuclear gene reaffirmed the findings of Carvalho-Pinto and Lourenço-de-Oliveira (2004) and suggested that in Itatiaia municipality, RJ state, Brazil, there are two putative species under the name of *Ke. cruzii* (Rona et al., 2010, 2013). A study in Cananéia municipality in Southeastern Brazil employing wing geometric morphometrics and *COI* gene analyses showed that samples from hilltops differ from those found in the lowlands (Lorenz et al., 2014). Analyses employing the complete mitochondrial genome of four *Kerteszia* specimens, including four distinct populations of *K. cruzii,* showed differences in the codon composition from three localities in the South-eastern region (São Paulo and Cananéia both in São Paulo state; Itatiaia in RJ) of Brazil and one from South Brazil (Maquiné in the Rio Grande do Sul state) (Oliveira et al., 2016). Analyses using the *cpr* and *clock* nuclear genes revealed two lineages in the Serra do Mar region: one corresponding to samples from the low land coast and another corresponding to that from mountain specimens (de Rezende Dias et al., 2018). Recently, Kirchgatter et al. (2020), employing sequences from GenBank from *NADH4* and *COI*, identified three genetic lineages of this species in Brazil corresponding to the Serra do Mar, Serra da Mantiqueira, and Serra da Cantareira.

Herein, we employed *single nucleotide polymorphisms* (SNPs) from a nextRAD sequencing (nextera-tagmented, reductively amplified DNA) to verify whether landscapes presenting heterogeneous vegetation physiognomies, with distinct vegetal coverage, can influence the genetic structure of *Ke. cruzii.* The main objectives of this study were to (1) test the genetic structure of *Ke. cruzii* in different gradients of vegetation cover in six locations described in Laporta et al. (2015); (2) to search for genetic signatures associated with distinct environmental variables.

MATERIALS AND METHODS

Mosquito Collection and DNA Extraction

Mosquitoes were collected between July 2016 and December 2018 (**Table 1**). *Ke. cruzii* specimens were collected from seven field collection sites as follows: (1) Tapiraí, (2) Sete Barras, (3) Eldorado, (4) Eldorado – Toca da Onça, (5) Cananéia, Esteiro do Morro, (6) Cananéia, Sítio Itapuan, and (7) Ilha Comprida, Pedrinhas. The longitude and latitude (GPS, datum WGS84) data were obtained from each collection site and georeferenced in the Geographic Information System (ArcGIS v. 10.3.1, and QGIS v. 2.18.9) in EPSG: 4326 (WGS84, world geographic projection). The elevation map was obtained from the Shuttle Radar Topography Mission (SRTM) 1.4, interpolated to a 30-m spatial resolution. An elevation map was used to construct the terrain slope topographic layers of the study region. These topographic layers were overlapped and projected in GIS together with vegetation cover layers (i.e., ombrophilous dense forest, restinga, and mangrove) obtained from SOS Mata Atlântica/INPE and Landsat images (30-m spatial resolution).

The collection sites in Esteiro do Morro, Eldorado, Sete Barras, Toca da Onça, and Tapiraí are in transition landscapes between dense ombrophilous forests and rural environments, with forest cover gradients of 74.99% (Tapiraí), 65.37% (Esteiro do Morro and Sete Barras), and 44.66% (Eldorado). Sítio Itapuan and Pedrinhas presented heterogeneous landscapes. For example, in Sítio Itapuan, besides dense ombrophilous forest and the rural environment, restinga, water, and mangroves (97.49% of natural Atlantic Forest) are still present. The landscape in Pedrinhas is characterized as a transition between the urban environment, restinga vegetation, and mangroves (with 94.48% of the preserved natural vegetation) (Laporta et al., 2015) (**Figure 1**).

Field collections were performed using Shannon traps from 6:00 p.m. to 10:00 p.m. The mosquitoes were preserved in silica gel until morphological identification and DNA extraction. Specimens were morphologically identified using the dichotomous keys of Forattini (2002) and Consoli and Lourenço-de-Oliveira (1994) and confirmed by barcode *COI* amplification and sequencing (Bourke et al., 2018). Genomic DNA was extracted using the salt method described by Miller et al. (2007) and modified by Laporta et al. (2015) from 59 mosquito specimens (**Table 1**).

Next Generation Sequencing (NGS) and SNPs Detection

The nextRAD libraries were assessed by the SNPsaurus LLC Company, as in Russello et al. (2015). Briefly, the genomic DNA was first fragmented using Nextera reagent (Illumina, Inc.), which also ligated short adapters to the ends of the fragments; the reaction was scaled to fragment 3 ng of genomic DNA. Each fragment was then amplified using 25 cycles at 75°C, with one of the primers matching the adapter and extending eight nucleotides into the genomic DNA with the selective sequence

TABLE 1 | Field collection information.

Locality	Samples code	Municipality	Coordinate	Number of samples
Sítio Itapuan	SIT	Cananéia	−24.888583, −47.851667	10
Esteiro do Morro	EM	Cananéia	−24.809933, −47.860367	10
Pedrinhas	PE	Ilha Comprida	−24.886117, −47.782233	10
Tapiraí	TAP	Tapiraí	−24.006220, −47.500060	9
Eldorado	ELD-1	Eldorado	−24.495983, −48.131250	5
Eldorado Toca da Onça	ELD-2	Eldorado	−24.45507, −48.22217	10
Sete Barras	SB	Sete Barras	−24.302783, −47.891100	5

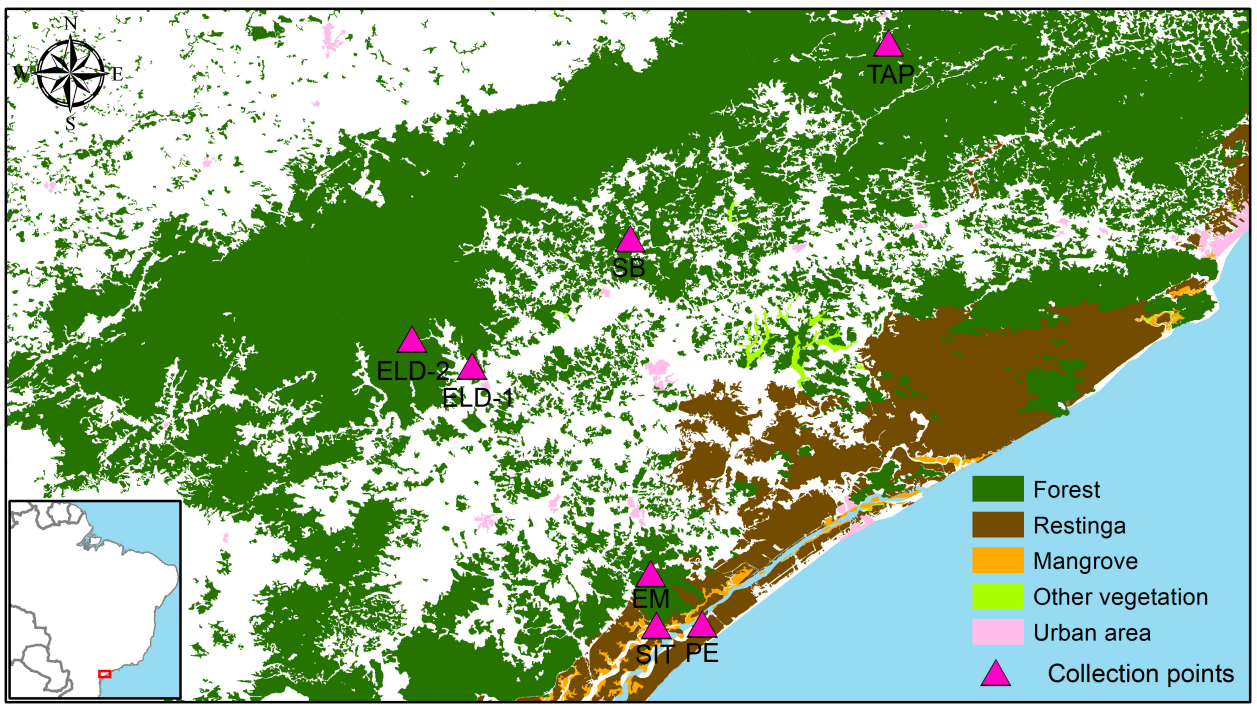

FIGURE 1 | Study region: Ribeira Valley, Southeastern São Paulo State, Brazil. Field collection points were utilized for sampling from mosquito fauna in remnants of the Brazilian Atlantic rainforest conserved in the study region. Source: SOS Mata Atlântica/INPE.

TGCAGGAG. Thus, only fragments starting with a sequence that can be hybridized by the selective sequence of the primer can be efficiently amplified. Thereafter, the amplified fragments were sequenced on a HiSeq 4000 with one lane of 150 bp reads (University of Oregon).

Genotyping analysis was performed with custom scripts (SNPsaurus, LLC) using bbduk (BBMap tools)[1] to trim the reads (**Supplementary Material S1**). A *de novo* reference was created after collecting 10 million reads equally from the samples; reads with counts less than 7 or greater than 700 were excluded from the analysis. The remaining *loci* were aligned to each other to identify allelic *loci* and collapsed allelic haplotypes to a single representative. Then, the reads were mapped to match the reference with a threshold of 95% using bbmap (BBmap tools), and genotype calling was performed with call variants in BBmap tools. Finally, the vcf archive was filtered using vcf tools (Danecek et al., 2011) to exclude: (1) alleles with frequencies less than 5%; (2) genotype calls under 50%; (3) samples with more than 50% missing data; (4) *loci* with more than 10% missing data; and (5) *loci* with a deviation of the Hardy-Weinberg equilibrium ($P < 0.01$).

Population Analyses

The vcf file was converted into the necessary formats to perform the remaining analyses in PGDspider v 2.0 (Lischer and Excoffier, 2012). To assess the genetic distances within individuals and groups of individuals, two matrices were generated using Nei's distances in R package v. 3.5.2, using StAMPP (Pembleton et al., 2013). The matrices were visualized in Splitstree4 v. 4.14.2 (Huson and Bryant, 2006). StAMPP was also employed to generate pairwise Weir and Cockerham's (1984) *Fst* matrixes. The statistical significance (*P*) of each value was determined using 100 permutation tests. To evaluate the genetic structure of the studied

[1]http://sourceforge.net/projects/bbmap/

samples, Bayesian analysis in STRUCTURE software v. 2.3.4 (Pritchard et al., 2000) and principal component analysis (PCA) in R v. 3.5.1, using the adegenet package (Jombart, 2008; Jombart and Ahmed, 2011) was performed. STRUCTURE analysis was performed using Strauto (Chhatre and Emerson, 2017) in seven runs ($K = 1–7$) with ten replicates for each run. The Markov chain Monte Carlo (MCMC) was carried out for 1 million generations and a burn-in period of 100,000 for each run. The Evanno method (Evanno et al., 2005) was implemented in STRUCTURE Harvester (Earl and vonHoldt, 2012) to determine a suitable number of clusters. The ten replicates for the best K value were combined in CLUMPP v1.1.2 (Jakobsson and Rosenberg, 2007) under the algorithm *Large K Greedy* with 2,000 permutations, and the results were visualized in Distruct v 1.1 (Rosenberg, 2004). Both CLUMPP and distruct were used through the pipeline CLUMPAK (Kopelman et al., 2015) on the webserver http://clumpak.tau.ac.il. The K-means clustering of the PCA analysis was determined for the Bayesian inference criterion (BIC), and then, we performed a discriminant PCA (dPCA) analysis.

Arlequin v. 3.5 software (Excoffier and Lischer, 2010) was employed to evaluate isolation by distance with 1,000 permutation tests and to perform a molecular variance analysis (AMOVA) (Weir and Cockerham, 1984) with 1,000 permutations. The latter was implemented in two ways: one considering the seven localities belonging to one group. At the same time, in the second, the populations were divided into two groups: one containing the lowland samples and the other containing the samples collected in the interior and mountains.

Genomic Signatures

We employed two distinct methodologies to detect the candidate *loci*. The first is based on differences in allelic frequencies in samples implemented by BayScan v 2.01 (Foll and Gaggiotti, 2008). This software uses the multinomial model Dirichlet and the reversible jump Markov chain Monte Carlo (RJ-MCMC) algorithm to obtain the posterior probability distributions. We used the default to perform this analysis, which uses 20 pilot runs with 5,000 interactions to adjust the distribution of the RJ-MCMC algorithm and a false discovery rate (FDR) value of 0.05.

The second approach was implemented in LFMM v.1.2 (Frichot et al., 2013). This methodology associates the allelic frequencies with environmental variables, using latent factor mixed models based on a Bayesian distribution, which can decrease FDR because it can estimate aleatory effects, which may be associated with genetic population events and isolation by distance. The number of latent factors was based on the STRUCTURE, PCA, and dPCA results. To decrease the FDR rates, we estimated the inflation factor according to the authors' suggestions. Based on the hypotheses of the study, the following variables were quantified in the landscape (100-km^2) surrounding the field collection points: (1) the mean elevation, (2) the mean terrain slope, and (3) the proportion of each vegetation cover (i.e., ombrophilous dense forest, restinga, and mangrove) (**Supplementary Material S2**). In addition to these variables, the distance (km) from the field collection point to the coast shore was also calculated (**Table 2**).

TABLE 2 | Field collection points and environmental variables.

FCP[1]	Elevation	Slope	Forest	Restinga	Mangrove	Distance
TAP	682	9.1	90.1	0	0	64.7
SB	43	3.2	75.2	0	0	60.2
ELD-1	107	6.4	44.7	0	0	58.6
ELD-2	203	9.8	95.5	0	0	68
EM	149	9.6	77.9	2.9	0	14.1
SIT	78	4.8	29.5	37.2	14.3	7.2
PE	17	2	9	50.1	8	2.8

[1]*Field Collection Points: TAP, Tapiraí; SB, Sete Barras; ELD-1, Eldorado; ELD-2, Eldorado Toca da Onça; EM, Esteiro do Morro; SIT, Sitio Itapuan; PE, Pedrinhas.*

RESULTS

After filtering, 4,523 *loci* per individual remained out of the original 19,906 loci from the genotype call. In addition, one sample from Sete Barras showed more than 50% missing data; thus, it was discarded from further analyses. The unrooted phylogenetic tree (**Figure 2A**), produced using the Nei's distance matrix with the individuals, showed two distinct groups: one with only lowland samples (from Pedrinhas – Ilha Comprida; Esteiro do Morro and Sítio Itapuan, both in Cananéia municipality), and another with all individuals from the inland (Tapiraí, Sete Barras, Eldorado, and Eldorado – Toca da Onça), and two individuals from Esteiro do Morro. The second network recovered using the same analytical approach as in the first round, but defining the populations of all seven sites sampled (**Figure 2B**), showed three distinct genetic groups. The first group included all specimens from inland sites, the second group included the Esteiro do Morro population only, and the third group consisted of the remaining lowland samples (Pedrinhas and Sítio Itapuan).

The pairwise divergence (*FST*) results were relatively low. However, they were statistically significant, except between Eldorado and Sete Barras (**Table 3**). The distances ranged from 0.001 (between Eldorado and Sete Barras populations) and 0.279 (between the Sítio Itapuan and Tapiraí populations). Among lowland specimens, the *Fst* values varied from 0.01 to 0.3, and the countryside samples ranged between 0.001 and 0.038. The best-fit K chosen by the Evanno method was $K = 2$. The Bayesian multilocus analysis from STRUCTURE showed Esteiro do Morro as the most heterogeneous population, while Sítio Itapuan was the most homogenous (**Figure 3**).

The first two principal components represent only 25% of the variability. However, the analysis showed the same tendency of Nei's distance and the Bayesian analysis of STRUCTURE software. The Y-axis clearly showed two groups: Pedrinhas, Esteiro do Morro, and Sítio Itapuan are in the negative quadrant, whereas Tapiraí, Eldorado, and Eldorado – Toca da Onça and Sete Barras, and three samples from Esteiro do Morro are in the positive quadrant (**Supplementary Material S3**). The most suitable K value chosen by the BIC is $K = 2$. As such, the results of the dPCA analyses clearly show two distinct groups

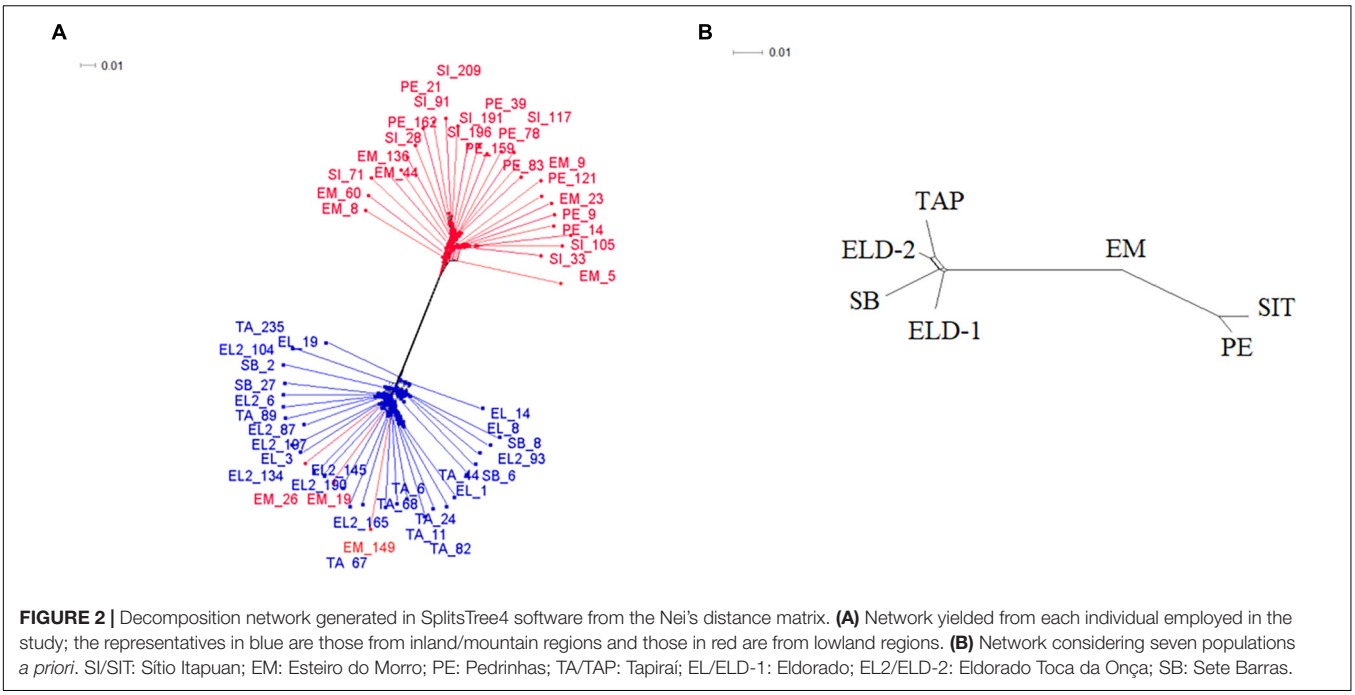

FIGURE 2 | Decomposition network generated in SplitsTree4 software from the Nei's distance matrix. **(A)** Network yielded from each individual employed in the study; the representatives in blue are those from inland/mountain regions and those in red are from lowland regions. **(B)** Network considering seven populations *a priori*. SI/SIT: Sítio Itapuan; EM: Esteiro do Morro; PE: Pedrinhas; TA/TAP: Tapiraí; EL/ELD-1: Eldorado; EL2/ELD-2: Eldorado Toca da Onça; SB: Sete Barras.

in the *X*-axis of the first discriminant variable (**Figure 4**). The AMOVA analyses showed that the variation among individuals (80.70%) was greater than among populations (19.27%) or when considering two groups (lowland × inland/mountain, 17.38%). Mantel's test suggests isolation by distance (regression coefficient = 0.57, $P = 0.003$).

The BayeScan analysis identified 18 outliers. Conversely, the LFMM analysis, which considered the environmental variables, showed that most putative *loci* under selection were associated with elevation (38 *loci*) and distance (27 *loci*). In comparison, 11 and 6 *loci* were associated with slope and mangrove, and eight were associated with forest and restinga, respectively (**Table 4**). In contrast, some *loci* were shared among environmental variables (**Table 4**). Fourteen *loci* overlapped in both analyses all associated with distance in the LFFM analysis.

TABLE 3 | Matrix of estimate of the genetic heterogeneity (*Fst*).

	SI	EM	PE	TA	EL2	EL	SB
SIT	0						
EM	**0.03**	0					
PED	**0.01**	**0.03**	0				
TAP	**0.28**	**0.12**	**0.26**	0			
ELD-2	**0.27**	**0.11**	**0.25**	**0.011**	0		
ELD-1	**0.25**	**0.1**	**0.23**	**0.029**	**0.01**	0	
SB	**0.26**	**0.1**	**0.25**	**0.038**	**0.01**	0.001	0

Values in bold were statistically significant.

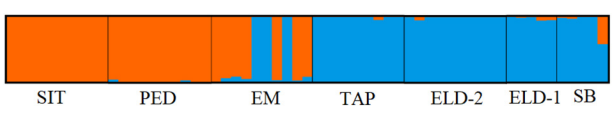

FIGURE 3 | STRUCTURE software result showing the two genetic clusters inferred by the Evanno method.

DISCUSSION

Multiple studies have indicated that the *Ke. cruzii* may represent a species complex (Ramírez and Dessen, 2000a; Ramírez and Dessen, 2000b; Carvalho-Pinto and Lourenço-de-Oliveira, 2004; Rona et al., 2010; Rona et al., 2013; de Rezende Dias et al., 2018; Kirchgatter et al., 2020). In this study, we used the nextRAD generation sequence approach to investigate the patterns of the genetic structure of *Ke. cruzii* in the Ribera Valley, Southeastern São Paulo state, Brazil. In addition, we verified the association between the genetic structure and landscape variables. The results of our analyses revealed the presence of two distinct lineages of this species in the studied regions and that they are associated with elevation and isolation by distance. The first lineage corresponds to lowland samples (from Pedrinhas, Sítio Itapuan, and Esteiro do Morro), and the second lineage is composed of specimens from inland and mountain sites (from Sete Barras, Eldorado, and Eldorado Toca da Onça, and Tapiraí). The results of PCA, Structure, and SplitTree analyses showed that in Esteiro do Morro, both lineages coexist in sympatry.

Similar results were found by de Rezende Dias et al. (2018) using sequences from the *cpr* and *clock* nuclear genes. Although using a broader geographic area – encompassing three Brazilian states, RJ, SP, and SC – de Rezende Dias et al. (2018) used specimens collected in lowland and inland/mountain localities (Serra do Mar). Their results showed one lineage that corresponds to samples from lowland sites and a second lineage

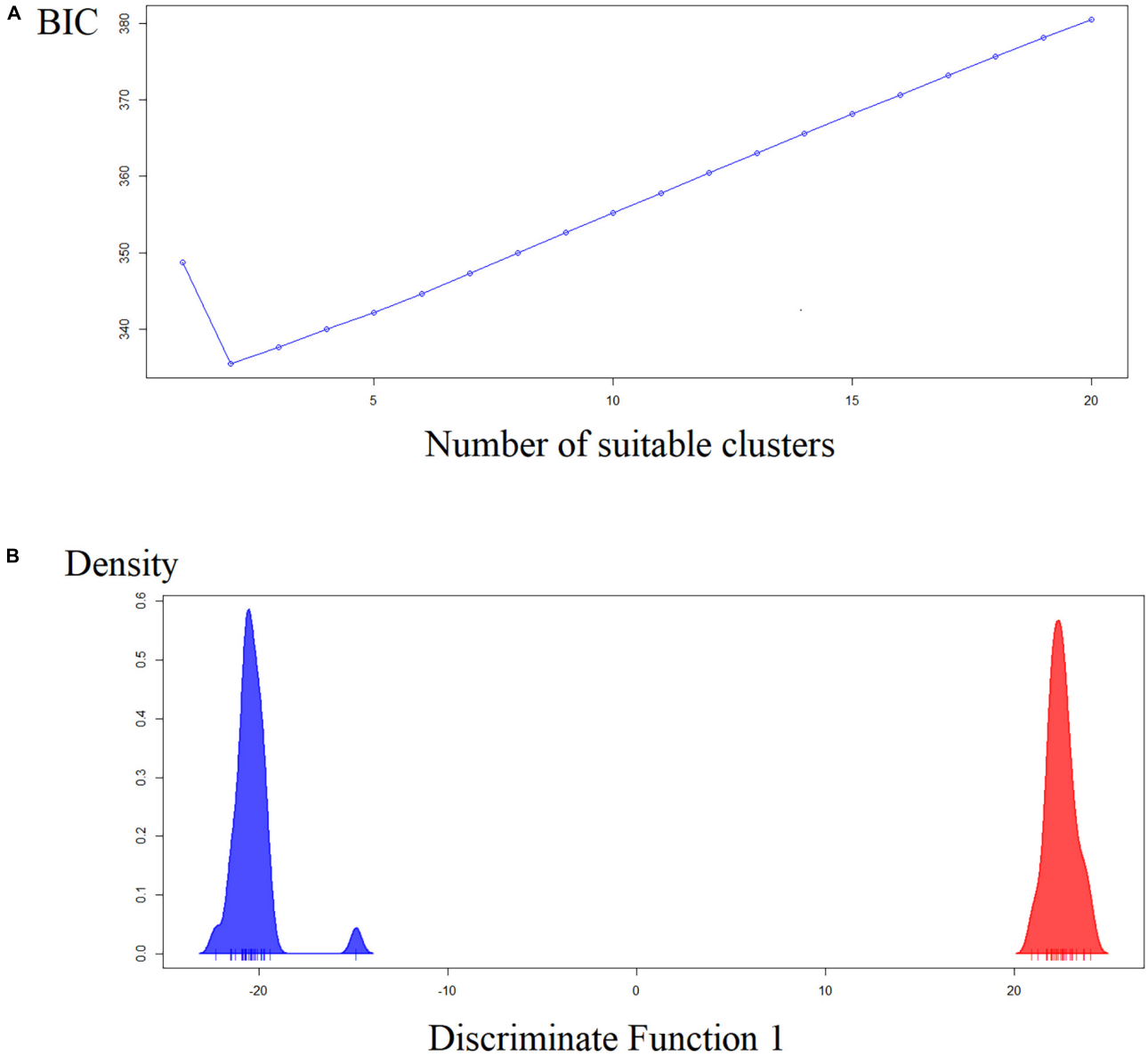

FIGURE 4 | Discriminant analysis of principal components (dPCA). **(A)** Likely number of clusters estimated by Bayesian Criterion (BIC). **(B)** Ordination showing the two clusters along the *X*-axis.

formed of specimens from the Serra do Mar mountain range. Likewise, specimens collected in Bocaina were clustered in both lineages and were found in sympatry. Unlike our results, however, their findings were not associated with isolation by distance, and the *Fst* values found between the lineages were higher (0.57) than the value found in our study (∼0.25, between lineages). Consequently, de Rezende Dias et al. (2018) strongly suggests the existence of two cryptic species under the name of *Ke. cruzii* in the study region.

Although the *Fst* values calculated using NGS datasets generated for samples employed in this study were relatively low, they were statistically significant, except for the values comparing populations from Eldorado and Sete Barras localities. The *Fst* values obtained from comparisons within lowland and inland and mountain samples were lower (ranging from 0.01 to 0.03 among lowland and 0.001 to 0.038 among inland and mountain samples) than between these regions (in which *Fst* varied from 0.1 to 0.28), evidence of restricted gene flow between these lineages. Taken together with LFMM analyses (**Table 4**), genetic differentiation was associated with elevation and distance, despite differences in vegetation type, forest cover, or slope.

Altitudinal stratification in *Ke. cruzii* was also verified by Lorenz et al. (2014) in the Cananéia municipality when using *COI* barcode sequences and wing geometric morphometrics. The *COI* haplotypes were very polymorphic; however, only two of the 60 haplotypes were shared by lowland and hilltop samples.

TABLE 4 | Putative *loci* under adaptive selection.

A						B		
Elevation	**Slope**	**Forest**	**Restinga**	**Mangrove**	**Distance**	***Locus***	**Prob**	**log10(PO)**
SNP_74	SNP_366	**SNP_472**	SNP_472	SNP_350	SNP_56	56	0.99400	2,2191
SNP_283	**SNP_648**	**SNP_648**	SNP_833	**SNP_472**	SNP_211	177	0.99360	2
SNP_401	**SNP_833**	**SNP_833**	**SNP_842**	SNP_1646	SNP_231	211	1	1.000
SNP_422	**SNP_834**	**SNP_834**	SNP_1013	SNP_1841	SNP_393	1,121	1	1,000
SNP_550	**SNP_842**	SNP_981	SNP_1114	SNP_2830	SNP_981	1,795	0.99	2.74
SNP_593	SNP_843	**SNP_3825**	SNP_3300	**SNP_4164**	SNP_1121	2,421	0.99	2.493
SNP_641	**SNP_1630**	SNP_3890	SNP_3334		SNP_1337	2,756	1	1,000
SNP_663	SNP_2609	SNP_4164	**SNP_4164**		SNP_1530	2,878	1	1,000
SNP_751	SNP_2697				SNP_1531	2,986	1	1,000
SNP_833	SNP_2699				SNP_2421	2,995	0.99	2.3
SNP_834	**SNP_3825**				SNP_2423	3,206	0.99	3.69
SNP_858					SNP_2878	3,357	1	1,000
SNP_957					SNP_2986	3,376	1	1,000
SNP_1011					SNP_2995	3,378	1	1,000
SNP_1304					SNP_3206	3,724	1	1,000
SNP_1357					SNP_3357	3,725	1	1,000
SNP_1548					SNP_3376	3,857	0.99	3.69
SNP_1630					SNP_3378	3,858	0.99	3.70
SNP_2200					SNP_3398			
SNP_2237					SNP_3424			
SNP_2345					SNP_3725			
SNP_2498					SNP_3857			
SNP_2499					SNP_3858			
SNP_2725					SNP_4016			
SNP_2746					SNP_4054			
SNP_2773					SNP_4219			
SNP_2805					SNP_4295			
SNP_3191								
SNP_3498								
SNP_3565								
SNP_3726								
SNP_3749								
SNP_3825								
SNP_3872								
SNP_4029								
SNP_4229								
SNP_4283								
SNP_4316								

A: Loci associated with environmental variables resulted from LFFM analysis; loci in bold were shared among environmental variables. B: Outliers from BayeScan analysis.

Therefore, considering the results of Lorenz et al. (2014), de Rezende Dias et al. (2018), and our results, we can consider elevational speciation, an ecological speciation in which adaptive divergence leads to dichotomous low/high altitude distribution. Elevational speciation is well known and has been observed in birds, frogs, and plants (Badyaev and Ghalambor, 2001; Caro et al., 2013; Chapman et al., 2013; Funk et al., 2016).

Usually, elevational speciation is studied and observed at high elevations (>1,000 m). However, we observed that small differences among altitudes (∼50 m) showed local adaptation in *Ke. cruzii*. Lowlands were also associated with high species richness and abundance in Culicidae, which uses bromeliads as larval habitats in Cananéia, southeast Brazil (Marques et al., 2012). Accordingly, even small variations, such as 200 m in elevation, can imply differences in the structure of Culicidae bromeliad communities. Similarly, bacterial and eukaryotic communities of phytotelma, on which the larval stages develop, also vary at low and high elevations (Gilbert et al., 2020; Malfatti et al., 2020). Additionally, Culicidae species distribution varies according to bromeliad species; for example, *Culex* (*Microculex*) spp. are more commonly found in *Vriesea friburgensis* than in *Aechmea lindenii*, whereas *Wyeomyia incaudata* and *Wy. pilicauda* are primarily associated with *A. lindenii* (Müller and Marcondes, 2006).

Furthermore, altitude is an important factor for species distribution in the bromeliads of the Atlantic Rain Forest (Brandão et al., 2009; Fontoura et al., 2012). For example, *Aechmea catendensis* and *Aechmea serragrandensis* were found clustered in lowlands, while *Echinocactus sessiliflorus* and *Aechmea guainumbiorum* were found mainly in the Submontana region in South and Southeast Brazil. Conversely, *Aechmea cephaloides* is typical of the highlands (Fontoura et al., 2012). A similar pattern occurs within *Vriesea*, the most common genus of Brazilian bromeliad, with species distribution differing at distinct altitudinal ranges (Malfatti et al., 2020). Considering (1) the evolutionary relationship between *Ke. cruzii* and bromeliads, (2) the evidence that bromeliad species follow a pattern of elevational distribution, and (3) that culicids are distributed according to bromeliad species, we can hypothesize that the lineages found herein may be associated with different bromeliad species and their elevational distribution pattern.

Although our results showed two distinct lineages in the Ribeira Valley, São Paulo, Brazil, *Fst* values did not corroborate the existence of cryptic or sister species under *Ke. cruzii* in this region. The differences between these lineages were mainly associated with isolation by distance and elevation, more than the vegetation mosaic, slope, or forest cover. Therefore, we conclude that this species may be under elevational speciation in this region and hypothesize that it is likely associated with the distribution of bromeliad species. In order to confirm our hypothesis, further investigations need to be performed in the region, to verify potential association between bromeliads species and population structure in *K. cruzii.*

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI SRA BioProject, accession no: PRJNA729988, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA729988.

AUTHOR CONTRIBUTIONS

BD-S and MAMS conceived the study. BD-S conducted the field collection work, performed the NGS and population analyses, and wrote the manuscript. GL contributed with the environmental metrics and analyses. TO helped with laboratory and NGS analyses. All authors read and agreed with the final version of this manuscript.

FUNDING

This study was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP grants no. 2016/08551-4 to BD-S; no. 2014/26229-7 to MAMS; and no. 2014/09774-1 to GL), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq grant no. 302375/2020-1 to MAMS and 307432/2019-0 to GL).

ACKNOWLEDGMENTS

The authors are grateful to the Fundação de Amparo à Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico for financial support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.707642/full#supplementary-material

REFERENCES

Badyaev, A. V., and Ghalambor, C. K. (2001). Evolution of life histories along elevational gradients: trade-off between parental care and fecundity. *Ecology* 82, 2948–2960. doi: 10.1890/0012-9658(2001)082[2948:eolhae]2.0.co;2

Bourke, B. P., Conn, J. E., Oliveira, T. M. P., Chaves, L. S. M., Bergo, E. S., Laporta, G. Z., et al. (2018). Exploring malaria vector diversity on the Amazon frontier. *Malar. J.* 17:342. doi: 10.1186/s12936-018-2483-2

Brandão, S. N., Silva, W. N., Silva, L. A. E., Fagundes, V., De Mello, C. E. R., Zimbrão, G., et al. (2009). "Analysis and visualization of the geographical distribution of atlantic forest bromeliads species," in *Proceedings of the IEEE, Symposium on Computational Intelligence and Data Mining*, (Nashville: IEEE), 375–380. doi: 10.1109/CIDM.2009.4938674

Branquinho, M. S., Marrelli, M. T., Curado, I., Natal, D., Barata, J. M. S., Tubaki, R., et al. (1997). Infecção do *Anopheles* (*Kerteszia*) cruzii por *Plasmodium vivax* e *Plasmodium vivax* variante VK247 nos municípios de São Vicente e Juquitiba, São Paulo. *Rev. Panam. Salud. Publica* 2, 122–134.

Brasil, P., Zalis, M. G., de Pina-Costa, A., Siqueira, A. M., Júnior, C. B., Silva, S., et al. (2017). Outbreak of human malaria caused by *Plasmodium simium* in the Atlantic Forest in Rio de Janeiro: a molecular epidemiological investigation. *Lancet Glob. Health* 5, e1038–e1046.

Buery, J. C., Rezende, H. R., Natal, L., Silva, L. S. D., Menezes, R. M. T., Fux, B., et al. (2018). Ecological characterization and infection of Anophelines (*Diptera*: *Culicidae*) of the Atlantic Forest in the southeast of Brazil over a 10 year period: has the behavior of the autochthonous malaria vector changed? *Mem. Inst. Oswaldo Cruz.* 113, 111–118. doi: 10.1590/0074-02760170225

Caro, L. M., Caycedo-Rosales, P. C., Bowie, R. C. K., Slabbekoorn, H., and Cadena, C. D. (2013). Ecological speciation along an elevational gradient in a tropical passerine bird? *J. Evol. Biol.* 26, 357–374. doi: 10.1111/jeb.12055

Carvalho-Pinto, C. J., and Lourenço-de-Oliveira, R. (2004). Isoenzymatic analysis of four *Anopheles* (*Kerteszia*) cruzii (*Díptera*: *Culicidae*) populations of Brazil. *Mem. Inst. Oswaldo Cruz.* 99, 471–475. doi: 10.1590/S0074-02762003000800012

Chapman, M. A., Hiscock, S. J., and Filatov, D. A. (2013). Genomic divergence during speciation driven by adaptation to altitude. *Mol. Biol. Evol.* 30, 2553–2567. doi: 10.1093/molbev/mst168

Chhatre, V. E., and Emerson, K. J. (2017). StrAuto: automation and parallelization of structure analysis. *BMC Bioinformatics* 18:192. doi: 10.1186/s12859-017-1593-0

Consoli, R. A. G. B., and Lourenço-de-Oliveira, R. (1994). *R: Principais Mosquitos de Importância Sanitária no Brasil*. Rio de Janeiro: Brasil Press, 97–114.

Dalla Bona, A. C., and Navarro-Silva, M. A. (2010). Physiological age and longevity of *Anopheles* (*Kerteszia*) cruzii Dyar & Knab (*Diptera*: *Culicidae*) in the Atlantic

Forest of Southern Brazil. *Neotrop. Entomol.* 39, 282–288. doi: 10.1590/S1519-566X2010000200021

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

de Alvarenga, D. A. M., Culleton, R., de Pina-Costa, A., Rodrigues, D. F., Bianco, C. Jr., Silva, S., et al. (2018). An assay for the identification of *Plasmodium simium* infection for diagnosis of zoonotic malaria in the Brazilian Atlantic Forest. *Sci. Rep.* 8:86. doi: 10.1038/s41598-017-18216-x

de Rezende Dias, G., Fujii, T. T. S., Fogel, B. F., Lourenço-de-Oliveira, R., Silva-do-Nascimento, T. F., Pitaluga, A. N., et al. (2018). Cryptic diversity in an Atlantic Forest malaria vector from the mountains of South-East Brazil. *Parasit. Vectors* 15:36. doi: 10.1186/s13071-018-2615-0

Deane, L. M., Ferreira Neto, J. A., Deane, M. P., and Silveira, I. P. S. (1970). *Anopheles* (*Kerteszia*) cruzii, a natural vector of the monkey malaria parasites, *Plasmodium simium* and *Plasmodium brasilianum*. *Trans. R. Soc. Med. Hyg.* 64:647. doi: 10.1016/0035-9203(70)90088-x

Demari-Silva, B., Laporta, G. Z., Oliveira, T., and Sallum, M. (2020). *Plasmodium* infection in *Kerteszia* cruzii (*Diptera*: *Culicidae*) in the Atlantic tropical rain forest, southeastern Brazil. *Infect. Genet. Evol.* 78:104061. doi: 10.1016/j.meegid.2019.104061

Duarte, A. M. R., Pereira, D. M., de Paula, M. B., Fernandes, A., Urbinatti, P. R., Ribeiro, A. F., et al. (2013). Natural infection in anopheline species and its implications for autochthonous malaria in the Atlantic forest in Brazil. *Parasit. Vectors* 6:58. doi: 10.1186/1756-3305-6-58

Earl, D. A., and vonHoldt, B. M. (2012). Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons. Gen. Resour.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847

Foll, M., and Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221

Fontoura, T., Scudeller, V. V., and da Costa, A. F. (2012). Floristics and environmental factors determining the geographic distribution of epiphytic bromeliads in the Brazilian Atlantic rain forest. *Flora* 207, 662–672. doi: 10.1016/j.flora.2012.05.003

Forattini, O. P. (2002). *Culicidologia Médica*. São Paulo: Editora da Universidade de São Paulo Press.

Forattini, O. P., Gomes, A. C., Natal, D., and Santos, J. L. F. (1986). Observações sobre a atividade de mosquitos *Culicidae* em matas primitivas da planície e perfis epidemiológicos de vários ambientes do Vale do Ribeira, SãoPaulo, Brasil. *Rev. Saude Publ.* 20, 178–203. doi: 10.1590/S0034-89101986000300002

Foster, P. G., de Oliveira, T. M. P., Bergo, E. S., Conn, J. E., Sant'Ana, D. C., Nagaki, S. S., et al. (2017). Phylogeny of Anophelinae using mitochondrial protein coding genes. *R. Soc. Open. Sci.* 4:170758. doi: 10.1098/rsos.170758

Frichot, E., Schoville, S. D., Bouchard, G., and François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol. Biol. Evol.* 30, 1687–1699. doi: 10.1093/molbev/mst063

Funk, W. C., Murphy, M. A., Hoke, K. L., Muths, E., Amburgey, S. M., Lemmon, E. M., et al. (2016). Elevational speciation in action? Restricted gene flow associated with adaptive divergence across an altitudinal gradient. *J. Evol. Biol.* 29, 241–252. doi: 10.1111/jeb.12760

Gilbert, K. J., Bittleston, L. S., Naïve, M. A. K., Kiszewski, A. E., Buenavente, P. A. C., Lohman, D. J., et al. (2020). Investigation of an elevational gradient reveals strong differences between bacterial and eukaryotic communities coinhabiting nepenthes phytotelmata. *Microb. Ecol.* 80, 334–349. doi: 10.1007/s00248-020-01503-y

Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 3, 254–267. doi: 10.1093/molbev/msj030

Jakobsson, M., and 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, 1801–1806. doi: 10.1093/bioinformatics/btm233

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart, T., and Ahmed, I. (2011). Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi: 10.1093/bioinformatics/btr521

Kirchgatter, K., Guimarães, L. O., Trujillano, H. H. Y., Arias, R. F., Cáceres, A. G., Duarte, A. M. C., et al. (2020). Phylogeny of *Anopheles* (*Kerteszia*) (*Diptera*: *Culicidae*) using mitochondrial genes. *Insects* 24:324. doi: 10.3390/insects11050324

Kirchgatter, K., Tubaki, R. M., dos Santos Malafronte, R., Alves, I. C., Lima, G. F., Guimarães, L. D. O., et al. (2014). *Anopheles* (*Kerteszia*) cruzii (*Diptera*: *Culicidae*) in peridomiciliary area during asymptomatic malaria transmission in the Atlantic Forest: molecular identification of blood-meal sources indicates humans as primary intermediate hosts. *Rev. Inst. Med. Trop. Sao Paulo.* 56, 403–409. doi: 10.1590/S0036-46652014000500006

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., and Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15, 1179–1191. doi: 10.1111/1755-0998.12387

Laporta, G. Z., Burattini, M. N., Levy, D., Fukuya, L. A., de Oliveira, T. M., Maselli, L. M., et al. (2015). *Plasmodium falciparum* in the southeastern Atlantic forest: a challenge to the bromeliad-malaria paradigm? *Malar. J.* 25:181. doi: 10.1186/s12936-015-0680-9

Lischer, H. E. L., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28, 298–299. doi: 10.1093/bioinformatics/btr642

Lorenz, C., Marques, T. C., Sallum, M. A. M., and Suesdek, L. (2014). Altitudinal population STRUCTURE and microevolution of the malaria vector *Anopheles* cruzii (*Diptera*: *Culicidae*). *Parasit Vectors* 16:581. doi: 10.1186/s13071-014-0581-8

Malfatti, E., Ferreira, P., and Utz, L. R. P. (2020). Eukaryotic communities in bromeliad phytotelmata: how do they respond to altitudinal diferences? *Diversity* 12:326. doi: 10.3390/d12090326

Marques, T. C., Bourke, B. P., Laporta, G. Z., and Sallum, M. A. M. (2012). Mosquito (*Diptera*: *Culicidae*) assemblages associated with *Nidularium* and *Vriesea* bromeliads in Serra do Mar, Atlantic Forest, Brazil. *Parasit. Vectors* 5:41. doi: 10.1186/1756-3305-5-41

Medeiros-Sousa, A. R., de Oliveira, R. C., Duarte, A. M. R. C., Mucci, L. F., Ceretti-Junior, W., and Marrelli, M. T. (2019). Effects of anthropogenic landscape changes on the abundance and acrodendrophily of *Anopheles* (*Kerteszia*) cruzii, the main vector of malaria parasites in the Atlantic Forest in Brazil. *Malar. J.* 18:110. doi: 10.1186/s12936-019-2744-8

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., and Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17, 240–248. doi: 10.1101/gr.5681207

Müller, G. A., and Marcondes, C. B. (2006). Bromeliad-associated mosquitoes from Atlantic forest in Santa Catarina Island, southern Brazil (*Diptera*, *Culicidae*), with new records for the State of Santa Catarina. *Iheringia Ser. Zool.* 96, 315–319. doi: 10.1590/S0073-47212006000300070007

Oliveira, T. M., Foster, P. G., Bergo, E. S., Nagaki, S. S., Sanabani, S. S., Marinotti, O., et al. (2016). Mitochondrial genomes of *Anopheles* (*Kerteszia*) (*Diptera*: *Culicidae*) from the Atlantic Forest, Brazil. *J. Med. Entomol.* 53, 790–797.

Pembleton, L., Cogan, N., and Forster, J. (2013). StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Resour.* 13, 946–952. doi: 10.1111/1755-0998.12129

Pritchard, J. K., Stephens, P., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959. doi: 10.1093/genetics/155.2.945

Ramírez, C. C., and Dessen, E. M. (2000a). Chromosomal evidence for sibling species of the malaria vector *Anopheles* cruzii. *Genome* 43, 143–151. doi: 10.1139/g99-103

Ramírez, C. C., and Dessen, E. M. (2000b). Chromosome differentiated populations of *Anopheles* cruzii: evidence for a third sibling species. *Genetica* 108, 73–80. doi: 10.1023/a:1004020904877

Rona, L. D., Carvalho-Pinto, C. J., and Peixoto, A. A. (2010). Molecular evidence for the occurrence of a new sibling species within the *Anopheles* (*Kerteszia*) cruzii complex in south-east Brazil. *Malar. J.* 9:33. doi: 10.1186/1475-2875-9-33

Rona, L. D., Carvalho-Pinto, C. J., and Peixoto, A. A. (2013). Evidence for the occurrence of two sympatric sibling species within the *Anopheles* (*Kerteszia*) cruzii complex in southeast Brazil and the detection of asymmetric introgression between them using a multilocus analysis. *BMC Evol. Biol.* 13:207. doi: 10.1186/1471-2148-13-207

Rosenberg, N. A. (2004). Distruct: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Russello, M. A., Waterhouse, M. D., Etter, P. D., and Johnson, E. A. (2015). From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ.* 3:e1106. doi: 10.7717/peerj.1106

Weir, B., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi: 10.2307/2408641

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Tropical Diseases

ORIGINAL RESEARCH
published: 08 July 2021
doi: 10.3389/fitd.2021.703873

Evidence of Local Extinction and Reintroduction of *Aedes aegypti* in Exeter, California

Erin Taylor Kelly[1†], *Lindsey K. Mack*[1†], *Melina Campos*[2], *Crystal Grippin*[3], *Tse-Yu Chen*[4], *Ana L. Romero-Weaver*[4], *Kyle J. Kosinski*[4], *Katherine K. Brisco*[1], *Travis C. Collier*[4], *Eva A. Buckner*[4], *Lindsay P. Campbell*[4], *Anthony J. Cornel*[1], *Gregory C. Lanzaro*[2], *Rodrigo Rosario-Cruz*[6], *Kirk Smith*[5], *Geoffrey M. Attardo*[1‡] *and Yoosook Lee*[4*‡]

[1] *Department of Entomology and Nematology, University of California – Davis, Davis, CA, United States,* [2] *Vector Genetics Laboratory, Department of Pathology, Microbiology, and Immunology, School of Veterinary Medicine, University of California – Davis, Davis, CA, United States,* [3] *Laboratory Division, Delta Vector Control District, Visalia, CA, United States,* [4] *Florida Medical Entomology Laboratory, Department of Entomology and Nematology, Institute of Food and Agricultural Sciences, University of Florida, Vero Beach, FL, United States,* [5] *Maricopa County Environmental Services Department, Phoenix, AZ, United States,* [6] *BioSA Research Lab, Natural Sciences College, Autonomous Guerrero State University, Campus "El Shalako" Las Petaquillas, Guerrero, Mexico*

Established populations of *Aedes aegypti*, a mosquito vector of multiple major arthropod-borne viruses, were first found in three California (CA) cities in 2013. From 2013 to April 2021, *Ae. aegypti* thwarted almost all control efforts to stop its spread and expanded its range to 308 cities, including Exeter, in 22 counties in CA. Population genomic analyses have suggested that multiple genetically distinct *Ae. aegypti* populations were introduced into CA. However *Ae. aegypti* collected for the first time in 2014 in Exeter, appeared to be different from three major genetic clusters found elsewhere in CA. Due to intense control efforts by the Delta Vector Control District (DVCD), *Ae. aegypti* was thought to have been eliminated from Exeter in 2015. Unfortunately, it was recollected in 2018. It was not clear if the reemergence of *Ae. aegypti* in Exeter was derived from the bottlenecked remnants of the original 2014 Exeter population or from an independent invasion from a different population derived from surrounding areas. The goal of this work was to determine which of these scenarios occurred (recovery after bottleneck or reintroduction after elimination) and if elimination and reintroduction occurred to identify the origin of the invading population using a population genomic approach. Our results support the reintroduction after elimination hypothesis. The source of reintroduction, however, was unexpectedly from the southern CA cluster rather than from other two geographically closer central CA genetic clusters. We also conducted a knockdown resistance mutation profile, which showed Exeter 2014 had the lowest level of resistant alleles compared to the other populations, could have contributed towards DVCD's ability to locally eliminate *Ae. aegypti* in 2014.

Keywords: mosquito, *Aedes aegypti*, California, population genomics, mosquito control, reintroduction assessment

OPEN ACCESS

Edited by:
Andre Barretto Bruno Wilke, University of Miami, United States

Reviewed by:
JohnPaul Mutebi, Centers for Disease Control and Prevention (CDC), United States
Tom Schmidt, The University of Melbourne, Australia

***Correspondence:**
Yoosook Lee
yoosook.lee@ufl.edu

[†]These authors have contributed equally to this work and share first authorship

[‡]These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Vector Biology, a section of the journal Frontiers in Tropical Diseases

Received: *30 April 2021*
Accepted: *15 June 2021*
Published: *08 July 2021*

Citation:
Kelly ET, Mack LK, Campos M, Grippin C, Chen T-Y, Romero-Weaver AL, Kosinski KJ, Brisco KK, Collier TC, Buckner EA, Campbell LP, Cornel AJ, Lanzaro GC, Rosario-Cruz R, Smith K, Attardo GM and Lee Y (2021) Evidence of Local Extinction and Reintroduction of Aedes aegypti in Exeter, California. Front. Trop. Dis. 2:703873. doi: 10.3389/fitd.2021.703873

INTRODUCTION

Aedes aegypti serves as a major vector of four human disease-causing viruses, including yellow fever, dengue, chikungunya, and Zika viruses, posing a major threat to public health. Records indicate this species became established in the southeastern United States of America between the 15-18th centuries (1) and then spread throughout the east coast and southern states (2). California (CA) had remained free from *Ae. aegypti* until the summer of 2013 (3, 4) when the species were first detected in Menlo Park, Clovis, and Madera (4). Since then, this species has expanded its range to 308 cities in 22 counties, as of April, 2021 (5).

Attempts to locally eliminate and even control this highly invasive species has proven to be extremely challenging. For example, the Consolidated Mosquito Abatement District (CMAD) implemented labor intensive integrated vector control management in the city of Clovis where *Ae. aegypti* were first detected in 2013. Their efforts involved extensive public education, thorough property inspections, sanitation, insecticide treatment at larval sources, and residual barrier spraying with pyrethroids. Despite these efforts, *Ae. aegypti* successfully overwintered and continued to persist in Clovis (6).

Delta Vector Control District (DVCD), which covers northern Tulare County, just south of the area covered by CMAD, first detected *Ae. aegypti* in Exeter in August of 2014. Following detection, the DVCD initiated an intensive control campaign involving thorough, routine property inspections and barrier applications, public education, breeding site treatment or removal, as well as hand and truck-mounted adulticide fogging through October. DVCD encountered significant pushback from the public due to the regular adulticide applications, mandatory property inspections, and confiscation of container habitats. However, they reported that residents followed instructions and were able to control mosquito breeding. Exeter remained free of *Ae. aegypti* from the beginning of 2015 through the summer of 2017. DVCD detected *Ae. aegypti* again in 2017 in the neighboring cities of Visalia and Farmersville and, in 2018, at multiple sites in Exeter. In response to this detection, the district attempted to mount a response similar to 2014. Despite frequent, mandatory property inspections, breeding site elimination and sanitation, and barrier spraying, the infestation persisted.

Population genetic/genomic analyses suggest that multiple genetically distinct *Ae. aegypti* populations were introduced in CA (7, 8). The *Ae. aegypti* populations in California could be largely grouped into three major genetic clusters (8). One cluster includes samples from Fresno, Madera, and Menlo Park in Central CA (8). Another cluster includes samples from Clovis in central CA adjacent to Fresno. The third cluster includes all southern CA samples, as well as 2014 Exeter samples (8). Further clustering revealed that 2014 Exeter samples are similar to an *Ae. aegypti* population from Florida rather than the populations from southern CA. However, it is important to note that the data from Lee et al. (8) does not provide definitive evidence that Exeter *Ae. aegypti* were introduced from Florida because the cluster analysis included only Florida samples for comparison. *Aedes aegypti* from southeastern states such as Louisiana and Texas are genetically similar to Florida populations based on microsatellite analysis (3).

The genetic dissimilarity of 2014 Exeter to other California *Ae. aegypti* provided an opportunity to investigate the hypothesis that (1) the Exeter mosquito population went through a severe bottleneck due to intensive insecticide spraying and control followed by a recovery after a couple of years (bottleneck and recovery) or (2) DVCD successfully eliminated its initial introduction but later new mosquitoes migrated or were reintroduced to Exeter from other cities (local extinction and reintroduction). If bottleneck and recovery occurred, then we expected 2018 Exeter samples to have similar genetic profile as 2014 Exeter samples. If localized extinction and reintroduction occurred, then we expected 2018 Exeter samples to have a similar genetic profile to any of the other three CA genetic clusters. To investigate this hypothesis, we report the population genomic analysis of 243 *Ae. aegypti* from California, Arizona, Florida, and Mexico in this paper.

MATERIALS AND METHODS

Sample Collection

Genome data was obtained from deposited sequences available for specimens originating from earlier collections in Clovis, Fresno, Madera, Menlo Park, 2014 Exeter, East Los Angeles, San Diego (CA), Vero Beach, and Key West, Florida (FL) - NCBI BioProject PRJNA385349 (8, 9). Further samples from St. Augustine (FL), Naples (FL), St. Lucie (FL), Arizona, and Mexico were collected as adults using BG Sentinel traps and the remaining samples as eggs in ovicups, which were then reared to the adult stage in the laboratory and stored in >70% alcohol prior to DNA extraction. The global positioning system (GPS) coordinates from where each site specimens originated from as well as the number of samples sequenced and genotyped are provided in **Table 1**. The sample locations map is provided in **Figure 1**.

Genome Sequencing

Protocols for sequenced specimens available in the NCBI Bioproject database are described in Lee et al. (8) and Schmidt et al. (9). For the other specimens, DNA was extracted from individual mosquitoes using a magnetic-bead based DNA extraction protocol described in Chen et al. (10). DNA concentrations for each sample were measured using the Qubit dsDNA HS Assay Kit (Life Technologies) on a Qubit instrument (Life Technologies). A genomic DNA library was constructed for each individual mosquito with the QIAseq FX DNA Library UDI kit (Qiagen, Valencia, CA) using 20 ng DNA. Enzymatic fragmentation was conducted at 32°C for 11 minutes followed by 65°C for 30 minutes. Adapter ligation was conducted at 20°C for 2 hours. PCR amplification of the constructed library was carried out for 8 cycles [(98°C for 20 seconds, 60°C for 30 seconds, 72°C for 30 second) x 8]s. Library cleanup was done using PCRClean DX (Aline Biosciences, Woburn, MA). Library concentrations were measured using Qubit (Life Technologies).

TABLE 1 | Sample collection data. Number of *Ae. aegypti* sequenced are provided in N_{WGS} column. N_{KDR} denotes the number of samples genotyped for kdr and population specific SNPs using iPLEX assay.

Location	State	District	latitude	longitude	N_{WGS}	N_{KDR}
Clovis	California	Consolidated	36.81342	-119.66665	10	60
Fresno	California	Fresno	36.83998	-119.90485	3	
Madera	California	Madera	36.92671	-120.05016	3	
Menlo Park	California	San Mateo	37.43305	-122.19881	3	
Exeter (2014)	California	Delta	36.30385	-119.15797	3	24
Exeter (2018)	California	Delta	36.30385	-119.15797	4	24
East Los Angeles	California	Greater Los Angeles	34.03515	-118.15410	3	67
San Diego	California	San Diego	32.55557	-117.05128	4	
Phoenix	Arizona	Maricopa	33.51389	-112.47583	4	
St. Augustine	Florida	Anastasia	29.90119	-81.31262	1	
Miramar	Florida	Broward	25.98629	-80.24622	2	
Naples	Florida	Collier	26.15504	-81.75737	1	
Tampa	Florida	Hillsborough	27.96606	-82.49508	2	
Fort Myers	Florida	Lee	26.65284	-81.81183	1	
Miami	Florida	Miami-Dade	25.75458	-80.22354	2	
Key Largo	Florida	Monroe	25.08723	-80.44773	2	
Key West	Florida	Monroe	24.55684	-81.78290	1	
Haverhill	Florida	Palm Beach	26.68861	-80.11346	2	
Holiday	Florida	Pasco	28.18634	-82.74527	2	
Auburndale	Florida	Polk	28.04973	-81.77675	2	
Sarasota	Florida	Sarasota	27.31105	-82.46285	3	
Sanford	Florida	Seminole	28.82482	-81.33626	2	
St. Lucie	Florida	St. Lucie	27.52948	-80.31699	1	
Vero Beach	Florida	Indian River	27.58721	-80.37340	3	
Cuernavaca	Mexico	Morelos	17.59358	-100.84823	3	
				TOTAL	68	175

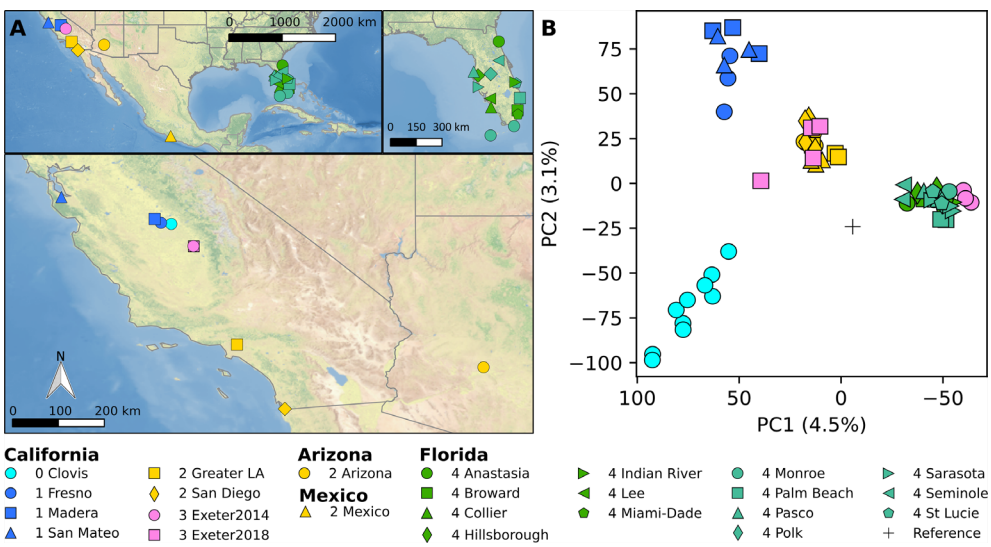

FIGURE 1 | **(A)** Map of sample location. ClearnTOPO2, a public domain dataset, was used as a basemap. **(B)** PCA analysis based on genome-wide biallelic SNPs.

Libraries were sequenced as 150 bp paired-end reads using a NovaSeq instrument (Illumina) at the University of Florida Interdisciplinary Center for Biotechnology Research (ICBR) Nextgen DNA Sequencing Core.

SNP Genotyping

Eggs were collected from the field in Clovis and Greater LA and reared in the lab under existing protocols (11). Adult collections from Exeter were obtained from Delta Vector Control District. Sixty individuals from Clovis, 67 individuals from Greater LA, 24 individuals from Exeter 2014 and 24 individuals from Exeter 2018 were collected (**Table 1**, N_{KDR}). DNA was extracted using the Zymo Quick-DNA/RNA Mini Prep Kit (#D7001) using the protocol for Solid Tissue. DNA quality and quantity were determined using a Qubit instrument (Life Technologies) and approximately 4 ng/μL was extracted from each individual.

DNA was submitted to the UC Davis Veterinary Genetics Laboratory for SNP genotyping using the iPLEX MassARRAY analysis following the protocol described in Lee et al. (12). Thirty-seven SNPs were selected from the published whole genome sequences (8), 29 of which were chosen due to their association with specific genetic clusters and eight within the Voltage Gated Sodium Channel gene. Data for 5 knockdown resistance (*kdr*) SNPs (F1534C, V410L, S723T, I915K, and V1016I) were included from (13). A full list of SNPs and primers can be found in **Table S1**. Populations were clustered using a Principal Component Analysis (PCA) performed in R v4.0.5, removing individuals with no calls for any SNPs.

Genome Sequence Data Analysis

Raw reads were trimmed using fastp (14) version 0.20.1. Trimmed reads were mapped to the Ae13CLOV028MT (Genbank ID: MH348176) first using BWA-MEM (15) version 0.7.15 following recommendation from Schmidt et al. (16) to minimize the impact of mitochondrial reads mapping to the nuclear genome due to presence of pseudogenes (17). Unmapped and mate-is-unmapped reads from mitogenome mapping were filtered using sambamba, converted to fastq files using samtools version 1.12, and mapped to AaegL5 reference genome (18) using BWA-MEM (15) version 0.7.15. Mapping statistics were calculated using Qualimap version 2.2 (19) (**Table S2**). Joint variant calling using all samples was done using Freebayes (20) version 1.0.1 with standard filters and population priors disabled.

The repeat regions were soft-masked in the AaegL5 reference genome and SNPs in these regions were excluded from analysis. Only biallelic SNPs with a minimum of 6X coverager were used for further analysis. A missing data threshold of 10% was used to filter SNPs. Hudson FST (21) and Principal Component Analysis (PCA) analyses was done in Python version 3.6.6 using the scikit-allel module version 1.2.0 (22). A phylogenetic tree based on the SNPs was constructed using the neighbor-joining algorithm as implemented in PHYLIP (23) version 3.696. Bayesian clustering method implemented in ADMIXTURE v1.3.0 (24) was used to estimate ancestry components for each individual. For this analysis, a total of 10 iterations were performed for values of *K* clusters from 1 to 10 with no prior population assignment. The results for each *K* were compiled using the online version of CLUMPAK and plotted in R. Windowed population genetic statistics such as F_{ST} and nucleotide diversity (π) were calculated using scikit-allel library with a 1Mbp window size and half-step overlapping windows.

RESULTS

Thirty-eight *Ae. aegypti* genomes were sequenced, originating from Florida (N=24), Arizona (N=4), Exeter 2018 (N=4), and Clovis (N=6). These genomes were analyzed together with 30 sequenced genomes used in other publications (N=68) (**Table 1**, N_{WGS}) (8, 9). Each sample was sequenced with a mean nuclear genome coverage of 11.9X per sample (**Table S2**; range = 8.3-37.9X; std = 5.0).

Principal component analysis (PCA) of genome-wide biallelic SNP genotypes reveal four major genetic clusters: (1) Clovis 2013 (cyan in **Figure 1**), (2) Fresno, Madera and Menlo Park (blue in **Figure 1**), (3) southern CA, Arizona, Mexico (yellow in **Figure 1**), and (4) Florida (green and sage in **Figure 1**). The *Ae. aegypti* population from Exeter 2014 cluster with Florida populations while 2018 Exeter samples mostly cluster together with southern CA samples. One sample from the 2018 Exeter group occupying the intermediate genetic space between southern CA and two other central populations appears to have genotypes intermediate between southern CA and Clovis populations (**Figure 2**).

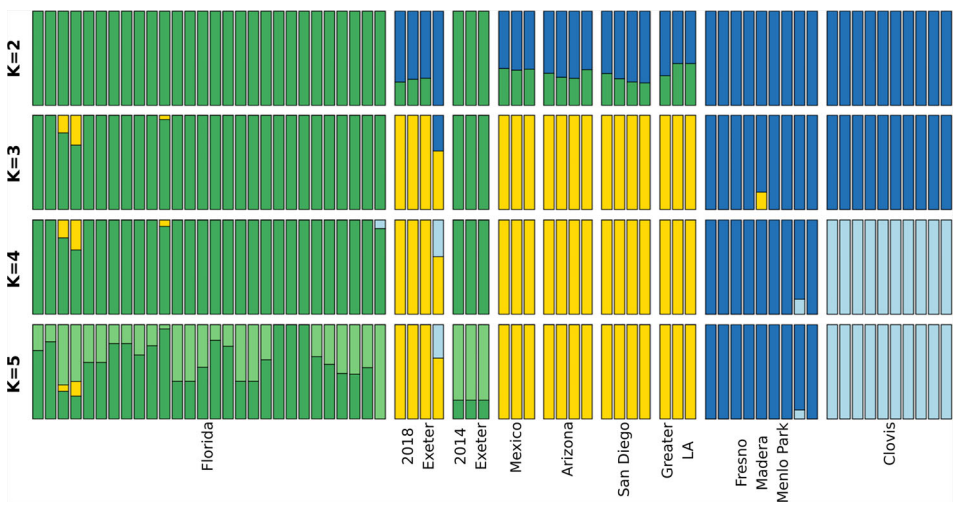

FIGURE 2 | Admixture plot of individual *Ae. aegypti* mosquitoes. Each column represents an individual with probability of each cluster represented by different colors. The total probabilities add up to 1. Individuals with high likelihood of belonging to one cluster would appear as a single color column while individuals with mixed ancestry would have multiple colors in each column.

Consistent with PCA results, Admixture analysis of genome-wide SNPs clusters 2014 Exeter with Florida mosquitoes while 2018 Exeter clusters together with southern CA, Mexico and Arizona mosquitoes (**Figure 2**). One Exeter 2018 mosquito appears to have a signature of an admixture with the Clovis cluster. This particular sample corresponds to the outlier individual in the PCA that occupied a space between the southern CA and central CA clusters (**Figure 1**). The likelihood values increase continually with higher K (**Figure S1**) indicating that further substructure may be present beyond the K=5 clusters presented in **Figure 2**.

Exeter 2014 and 2018 were the most distant population pairs (F_{ST} = 0.152; **Figure 3** and **Table S3**). Consistent with Admixture results, East Los Angeles, Arizona, Mexico, and Florida form a closely related group with F_{ST} < 0.05 (**Table S3**). Fresno, Madera, and Menlo Park also showed no differentiation with F_{ST} = 0 (**Figures 2**, **3** and **Table S3**). The closest population pair with Exeter 2014 was Florida (F_{ST} = 0.055) while the closest population pair with Exeter 2018 was East Los Angeles (F_{ST} = 0.038).

Based on published genome sequences (8), we selected a set of biallelic SNPs most informative for separating individuals into each of the major three genetic clusters found in CA. We designed a multiplex SNP genotyping assay using the iPLEX MassARRAY system to screen 37 SNPs simultaneously for their genetic background (**Table S1**). Thirty-one of these SNPs differentiate *Ae. aegypti* into each of the four genetic clusters present in California and 6 of these SNPs are associated with permethrin resistance. We genotyped 2018 Clovis, Greater Los Angeles, Exeter 2014, and Exeter 2018 samples using this new multiplex SNP assay. Our PCA analysis of SNP genotypes (**Figure 4**) shows Clovis and Greater Los Angeles forming separate genetic clusters, although the boundary is not as clear as genome-wide SNP data shown in **Figure 1**. Consistent with genome-wide SNP data, Exeter 2018 has large overlap with southern CA samples with two samples potentially having mixed ancestry from the southern CA and Clovis clusters. Exeter 2014 *Ae. aegypti* clearly clustered separately from the Clovis 2018, Exeter 2018, and Greater Los Angeles *Ae. aegypti* mosquitoes.

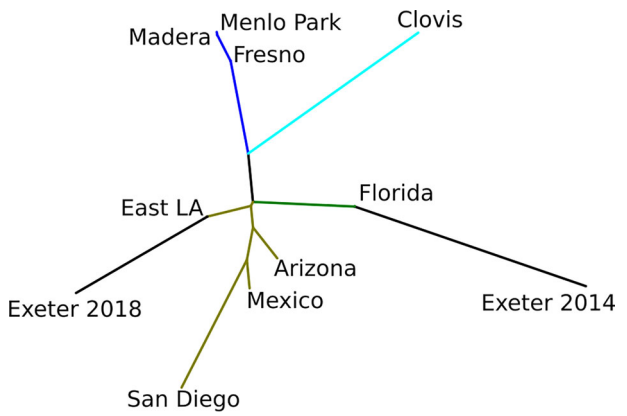

FIGURE 3 | Neighbor-joining tree based on pairwise F_{ST} distance.

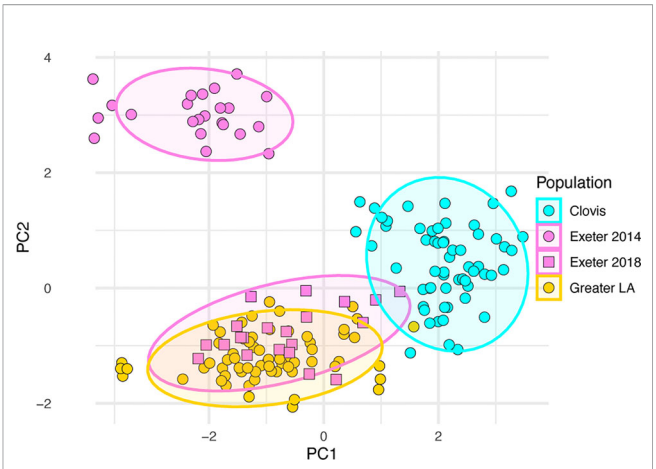

FIGURE 4 | Principal component analysis based on 37 SNPs in *Ae. aegypti* from Central and Southern California using the iPLEX MASSarray genotyping.

We calculated the windowed F_{ST}, nucleotide diversity (π), and change in π between 2018 and 2014 Exeter samples (**Figure 5**) to identify any hotspots of divergence. Any F_{ST} greater than 0.1 (red line in **Figure 5A**) indicates high genetic distance in the corresponding genomic region between two years. The genomic divergence between the two years appears to be genome-wide except in limited locations such as in the immediate vicinity of the centromeres. Overall elevated π was observed near the centromeres of Chromosome 1 and 3 (**Figure 5B**). While 2018 Exeter samples had higher π in Chromosome 1, 2014 Exeter samples had higher π in Chromosome 3 (**Figure 5C**).

We also analyzed 6 SNPs in the voltage-gated sodium channel gene, typically known as knockdown resistance gene (*kdr*) in mosquito literature (AAEL023266) located on Chromosome 3 using the method described in Mack et al. (13). Five of these 6 SNPs were previously used in Mack et al. (13). The SNP Q1853R (3:315931672) is a new addition to this study as it was only detected in the mosquitoes derived from the Exeter 2014 population. None of the other Californian populations possess this SNP and it appears to have disappeared with the eradication of that population. More than 60% of Exeter 2014 and 2018 *Ae. aegypti* had a V1016 resistant allele, while >90% of Clovis mosquitoes had this resistant allele present. Other nonsynonymous mutations in the *kdr* gene also occurred at the highest frequencies in Clovis. Except for allele F1534C, which was almost fixed in Clovis, alleles V410L, S723T and I915K were at much lower frequencies in Exeter 2014, 2018 and Greater LA *Ae. aegypti* (**Table 2**).

DISCUSSION

Our results strongly suggest that control actions taken by Delta Vector Control District personnel eliminated the local *Ae. aegypti* population in Exeter in 2014-2015. *Aedes aegypti* reinvaded Exeter in 2018 originating from different location/s than what originally invaded in 2014 and became established and are still present. We were able to confirm elimination and reinvasion in Exeter by comparative genetic analyses on 2014 and 2018 Exeter individuals.

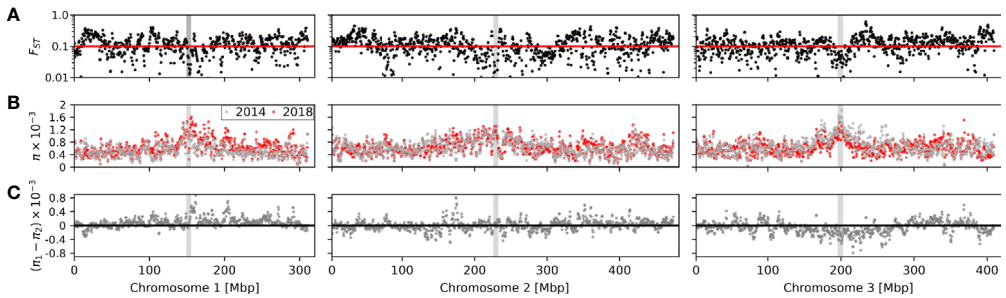

FIGURE 5 | **(A)** Windowed F_{ST}, **(B)** nucleotide diversity (π), and **(C)** change in nucleotide diversity (π_{2018} - π_{2014}) between 2014 and 2018 Exeter samples. The gray bar in the middle of each chromosome indicates location of the centromere.

TABLE 2 | Knockdown resistance (kdr) mutation genotypes per population.

Coord.	V410L 3:316080722				S723T 3:316014588				I915K 3:315999297				V1016I 3:315983763				F1534C 3:315939224				Q1853R 3:315931672			
	SS	SR	RR	%R	SS	SR	RR	%R	SS	SR	RR	%R	SS	SR	RR	%R	SS	SR	RR	%R	SS	SR	RR	%R
Clovis 2013	0	2	8	90%	0	2	8	90%	0	2	8	90%	0	2	8	90%	0	0	19	100%	10	0	0	0%
Clovis 2018	0	9	51	92.5%	0	9	51	92.5%	0	0	60	100%	0	1	59	99%	0	1	60	99%	60	0	0	0%
Exeter 2014	6	14	4	46%	6	14	4	46%	3	12	9	63%	3	12	9	63%	0	0	24	100%	17	7	0	15%
Exeter 2018	3	7	12	70%	3	7	12	70%	0	0	22	98%	3	7	12	70%	0	0	22	100%	22	0	0	0%
Greater LA 2018	11	37	19	59%	11	37	19	59%	0	15	52	88.8%	11	37	19	59%	0	1	66	99%	67	0	0	0%

Clovis 2013 data is genome sequencing data. Clovis 2018 data and Greater Los Angeles (LA) 2018 data are from Mack et al. (13). S stands for susceptible (reference) allele and R for resistant (alternate) allele. Genomic coordinates of kdr SNPs are provided on the 2nd row of this table.

Had the Exeter population experienced a severe reduction due to DVCD control efforts, followed by 3-4 years of recovery before detection again, then the 2018 Exeter individuals would be expected to be genetically similar to those from Exeter 2014. However, the 2018 Exeter samples form a genetic cluster separate from that of the 2014 Exeter cluster. Genetic differentiation between 2014 and 2018 Exeter samples appears genome-wide with the majority of genomic regions exceeding $F_{ST} > 0.1$ (**Figure 5**). This pattern is in contrast to the level of genetic differentiation we observed between Clovis 2013 and 2016 (8). Such a drastic change in genome structure would be extremely unlikely to develop within a few years simply by genetic drift or natural selection, adding additional evidence toward the local elimination and reintroduction hypothesis for *Ae. aegypti* population in Exeter, CA.

The Exeter 2018 cluster more closely aligns with the samples from the Southwestern USA and Mexico indicating that 2018 samples likely resulted from a reintroduction to Exeter. The most likely source of reinvading individuals would intuitively come from neighboring cities like Clovis and Fresno, which experience high abundances of *Ae. aegypti* in the middle to late summer. However, the genetic similarity of the Exeter 2018 to southern CA *Ae. aegypti* suggests that the founding Exeter 2018 mosquito/es were likely escapees traveling in a vehicle traveling from southern CA, demonstrating the propensity of *Ae. aegypti* for human-mediated dispersal across long distances. Our study is the first report of *Ae. aegypti* from a Central CA location sharing the majority of its genetic profile with *Ae. aegypti* from southern CA. Importantly, when *Ae. aegypti* were detected in 2018 in Exeter, they were also detected in several Visalia suburbs (unpublished data), and it remains to be learned whether the *Ae. aegypti* that are now present in Visalia and other towns neighboring Exeter are genetically similar or different to those from Exeter and whether they are migrants from the second Exeter invasion or from other neighboring central California locations. The success in eliminating the population detected in 2014 is likely due to the limited geographical scope of the population and the intensive control efforts that were implemented. However, the possibility of eliminating the new populations of *Ae. aegypti* in Exeter and suburbs of Visalia is unlikely as resources are stretched too thin to mount such an intense control effort across such an extensive area.

Southern CA *Ae. aegypti* had different frequencies of non-synonymous SNPs or mutations in the VGSC gene than in Central CA populations, including the ubiquitous V410L mutation that is known to confer resistance to pyrethroids (**Table 2**). If all six of the mutations in the VGSC gene included in this study confer some levels of resistance to pyrethroids (*kdr* resistant alleles) then Southern CA populations have a higher proportion of susceptible alleles relative to central CA populations which is consistent with a previous report (13). Although Exeter 2014 and 2018 samples had different overall genetic profiles, they both contained a greater number of pyrethroid susceptible *kdr* alleles than the Clovis population. Presence of higher frequencies of *kdr* susceptible alleles may have been a factor leading to elimination of *Ae. aegypti* in Exeter in 2014, while other neighboring central CA districts struggled to manage dispersal of their *Ae. aegypti* populations.

These findings confirm that implementation of intensive control efforts can eliminate *Ae. aegypti* locally. However, the

replacement of the Exeter 2014 population with individuals related to the Southern CA population that had a higher proportion of pyrethroid resistance alleles adds to the difficulty of adulticide control that could be further exacerbated when they hybridize to CA Central Valley populations that have an even greater number of *kdr* alleles. Elimination strategies require intensive efforts and resources, including public consent and establishment of a strong public relations program. Anticipating challenges resulting from the dispersal of resistant populations throughout California will be critical to comprehensive and sustainable control strategies.

This is the first record of southern CA-like *Ae. aegypti* genetic cluster found north of Kern county [latitude 35.21N; (25)]. The results from our study suggest that the population of southern CA *Ae. aegypti* continues to expand its range northwards and will hybridize with the existing Clovis population, corroborating previous indications of this possibility presented in Lee et al. (25). Additional statewide surveillance on the geographic distribution of *Ae. aegypti* genetic clusters combined with socio-environmental parameters may inform the nature and potential mechanisms of *Ae. aegypti* dispersal pathways within CA. Our relatively low-cost SNP genotyping assay or a similar approach could be a cost-effective way to screen a larger number of samples than current whole genome sequencing approaches and would be a useful tool to elucidate genetic mixing, origin of introductions, and pyrethroid resistance status of local populations.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NCBI BioProject repository with accession number PRJNA725510.

AUTHOR CONTRIBUTIONS

YL, GA, GL, and AC conceived experimental design. EK, LM, and GA conducted SNP genotyping analysis. AR-W, T-YC, and KK conducted genomic DNA library preparations for whole genome sequencing. YL, MC, and TC conducted genome data analysis. EK, LM, KB, CG, RR-C, KS, LC, and EB conducted sample collection or arranged sample acquisition and provided the sample metadata. All authors contributed to the article and approved the submitted version.

FUNDING

We acknowledge funding support from the Pacific Southwest Regional Center of Excellence for Vector-Borne Diseases funded by the U.S. Centers for Disease Control and Prevention (Cooperative Agreement 1U01CK000516), CDC grant NU50CK000420-04-04, the Southern IPM Center (Project S21-002) as part of USDA National Institute of Food and Agriculture Crop Protection and Pest Management Regional coordination Program (Agreement No. 2018-70006-28884), the USDA National Institute of Food and Agriculture (Hatch project 1025565), UF/IFAS Florida Medical Entomology Laboratory fellowship to T-YC, and Florida Department of Health (Contract CODQJ). The findings and conclusions in this article are those of the author(s) and do not necessarily represent the views of the funding agencies.

ACKNOWLEDGMENTS

We thank Ms. Mir Bear-Johnson at the Delta Vector Control District for providing materials for genetic analysis and sharing a detailed account of the mosquito control activities in Exeter, CA. We also thank Drs. Heinig and Lucas and their team at Collier Mosquito Control District, Dr. Sarah Shafer-Graham at Hillsborough County Mosquito Management Services, Dr. Carl Boohene at Polk County Mosquito Control District, Drs. Rui-de Xue, Steven Peper, and their team from Anastasia Mosquito Control District, Dr. Roger Jacobsen and his team at St. Lucie Mosquito Control & Coastal Management Services, Drs. Wade Brennan and Katie Cassell from Sarasota County Mosquito Management Services, for providing mosquito specimens for genetic analysis. We thank Rishi Kondapaneni, Ashley Malcolm, Brian Vasquez, and Eric Zeng at University of Florida to facilitate *Ae. aegypti* sample acquisition. We also thank Dr. David Moraga Amador at University of Florida NextGen DNA Sequencing Core and his team for their outstanding sequencing service.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fitd.2021.703873/full#supplementary-material

Supplementary Table 1 | iPLEX SNP primers and assay information

Supplementary Table 2 | Whole genome sequence metadata.

Supplementary Table 3 | Pairwise F_{ST} values.

Supplementary Figure 1 | Log likelihood values of Admixture runs.

Supplementary Figure 2 | Admixture plots for higher K values.

REFERENCES

1. Tabachnick WJ. The Evolutionary Relationships Among Arboviruses and the Evolutionary Relationships of Their Vectors Provides a Method for Understanding Vector-Host Interactions. *J Med Entomol* (1991) 28:297–8.
2. Crosby MC. *The American Plague : The Untold Story of Yellow Fever, the Epidemic That Shaped Our History*. New York: Berkley Books (2006).
3. Gloria-Soria A, Brown JE, Kramer V, Hardstone Yoshimizu M, Powell JR. Origin of the Dengue Fever Mosquito, Aedes Aegypti, in California. *PLoS Negl Trop Dis* (2014) 8:e3029. doi: 10.1371/journal.pntd.0003029

4. Cornel AJ, Holeman J, Nieman CC, Lee Y, Smith C, Amorino M, et al. Surveillance, Insecticide Resistance and Control of an Invasive *Aedes Aegypti* (Diptera: Culicidae) Population in California. *F1000Res* (2016) 5:194. doi: 10.12688/f1000research.8107.2
5. CDPH. *Aedes aegyptiAedes albopictusAedes Aegypti and Aedes albopictusMosquitoes in California by County/City* (2021). California Department of Health. Available at: https://www.cdph.ca.gov/Programs/CID/DCDC/CDPH%20Document%20Library/AedesDistributionMap.pdf (Accessed Apri, 2021).
6. Metzger ME, Hardstone Yoshimizu M, Padgett KA, Hu R, Kramer VL. Detection and Establishment of Aedes Aegypti and *Aedes Albopictus* (Diptera: Culicidae) Mosquitoes in Californi-2015. *J Med Entomol* (2017) 54:533–43. doi: 10.1093/jme/tjw237
7. Pless E, Gloria-Soria A, Evans BR, Kramer V, Bolling BG, Tabachnick WJ, et al. Multiple Introductions of the Dengue Vector, Aedes Aegypti, Into California. *PLoS Negl Trop Dis* (2017) 11:e0005718. doi: 10.1371/journal.pntd.0005718
8. Lee Y, Schmidt H, Collier TC, Conner WR, Hanemaaijer MJ, Slatkin M, et al. Genome-wide Divergence Among Invasive Populations of Aedes Aegypti in California. *BMC Genomics* (2019a) 20:204. doi: 10.1186/s12864-019-5586-4
9. Schmidt H, Collier TC, Hanemaaijer MJ, Houston PD, Lee Y, Lanzaro GC. Abundance of Conserved CRISPR-Cas9 Target Sites Within the Highly Polymorphic Genomes of Anopheles and *Aedes* Mosquitoes. *Nat Commun* (2020) 11:1425. doi: 10.1038/s41467-020-15204-0
10. Chen T, Vorsino AE, Kosinski KJ, Romero-Weaver AL, Buckner EA, Chiu JC, et al. A Magnetic-Bead-Based Mosquito DNA Extraction Protocol for Next-Generation Sequencing. *J Vis Exp* (2021) 170:e62354. doi: 10.3791/62354
11. Clemons A, Mori A, Haugen M, Severson DW, Duman-Scheel M. Culturing and Egg Collection of Aedes Aegypti. *Cold Spring Harb Protoc* (2010) 2010: pdb prot5507. doi: 10.1101/pdb.prot5507
12. Lee Y, Weakley AM, Nieman CC, Malvick J, Lanzaro GC. A Multi-Detection Assay for Malaria Transmitting Mosquitoes. *J Vis Exp* (2015) 96:e52385. doi: 10.3791/52385
13. Mack LK, Kelly ET, Lee Y, Brisco KK, Shen KV, Zahid A, et al. Frequency of Sodium Channel Genotypes and Association With Pyrethrum Knockdown Time in Populations of Californian Aedes Aegypti. *Parasit Vectors* (2021) 14:141. doi: 10.1186/s13071-021-04627-3
14. Chen S, Zhou Y, Chen Y, Gu J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* (2018) 34:i884–90. doi: 10.1093/bioinformatics/bty560
15. Li H. Aligning Sequence Reads, Clone Sequences and Assembly Contigs With BWA-MEM. *arXiv* (2013). doi: arXiv:1303.3997v2
16. Schmidt H, Hanemaaijer MJ, Cornel AJ, Lanzaro GC, Braack L, Lee Y. Complete Mitogenome Sequence of Aedes (Stegomyia) Aegypti Derived From Field Isolates From California and South Africa. *Mitochondrial DNA Part B Resources* (2018) 3:994–5. doi: 10.1080/23802359.2018.1495117
17. Hlaing T, Tun-Lin W, Somboon P, Socheat D, Setha T, Min S, et al. Mitochondrial Pseudogenes in the Nuclear Genome of Aedes Aegypti Mosquitoes: Implications for Past and Future Population Genetic Studies. *BMC Genet* (2009) 10:11. doi: 10.1186/1471-2156-10-11
18. Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, et al. Improved Reference Genome of Aedes Aegypti Informs Arbovirus Vector Control. *Nature* (2018) 563:501–7. doi: 10.1038/s41586-018-0692-z
19. Okonechnikov K, Conesa A, Garcia-Alcalde F. Qualimap 2: Advanced Multi-Sample Quality Control for High-Throughput Sequencing Data. *Bioinformatics* (2016) 32:292–4. doi: 10.1093/bioinformatics/btv566
20. Garrison E, Marth G. Haplotype-Based Variant Detection From Short-Read Sequencing. *arXiv* (2012).
21. Hudson RR, Slatkin M, Maddison WP. Estimation of Levels of Gene Flow From DNA Sequence Data. *Genetics* (1992) 132:583–9. doi: 10.1093/genetics/132.2.583
22. Miles A, Harding N. *Scikit-Allel - Explore and Analyse Genetic Variation* (2018). Available at: https://github.com/cggh/scikit-allel.
23. Felsenstein J. "Phylip (Phylogeny Inference Package)". In: *Seattle, 3.6 ed.* Seattle, WA: Department of Genome Sciences, University of Washington (2005). Available at: http://evolution.genetics.washington.edu/phylip.html.
24. Alexander DH, Novembre J, Lange K. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res* (2009) 19:1655–64. doi: 10.1101/gr.094052.109
25. Lee Y, Schmidt H, Collier TC, Weakley AM, Gripkey H, Person K, et al. "Population Genomics of *Aedes Aegypti* in California". In: WK Reisen, editor. *Eighty-Senventh Annual Conference of the Mosuqito and Vector Control Association of California*. Burlingame, CA: Mosquito and Vector Control Association of California (2019b). p. 23–9.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 08 September 2021
doi: 10.3389/fevo.2021.705318

Egg Cannibalism Varies With Sex, Reproductive Status, and Egg and Nymph Ages in *Arma custos* (Hemiptera: Asopinae)

Shaolong Wu[1,2,3†], Weiai Zeng[1†], Wan Deng[4], Mi Li[4], Wei Hu[2], Hailin Cai[1], Youzhi Li[3], Pengfei Xie[1], Lin Tan[3] and Zhicheng Zhou[2]**

[1] Hunan Changsha Municipal Tobacco Company, Changsha, China, [2] Hunan Province Tobacco Company, Changsha, China, [3] College of Plant Protection, Hunan Agricultural University, Changsha, China, [4] Hunan Forestry Academy, Changsha, China

Background: Egg cannibalism is common in nature. In China, *Arma custos* (Hemiptera: Asopinae) has been widely used as a natural enemy to control agricultural and forestry pests. A previous study showed that adult *A. custos* devour their eggs. However, no research has investigated the interaction between *A. custos* cannibalism and egg development. Clarifying the mechanisms involved in egg cannibalism by *A. custos* improves our understanding of the evolutionary relationships to enable more efficient mass rearing and biological control systems.

Results: Virgin females showed a lower egg cannibalism inclination than gravid females. Both virgin and mated females showed a higher egg cannibalism inclination than virgin and mated males. The first and second instar nymphs did not devour eggs. The third, fourth, and fifth instar nymphs devoured eggs. Younger eggs were more readily eaten than older eggs. Neither *A. custos* nymphs nor female adults consumed all the available eggs, allowing an emergence ratio of >70%.

Conclusion: *Arma custos* females exhibit a higher tendency for egg cannibalism than males. Egg cannibalism varies not only with the developmental stage of the eggs and nymphs but also with sex and reproductive status of *A. custos* females. These findings help us to better understand the evolutionary relationships in egg cannibalism by *A. custos* and contribute to the efficient mass rearing and realization of *A. custos* in biological control systems.

Keywords: *Arma custos*, egg cannibalism, gravid female, natural enemy, biocontrol

OPEN ACCESS

Edited by:
Cleber Galvão, Oswaldo Cruz Institute, Oswaldo Cruz Foundation (Fiocruz), Brazil

Reviewed by:
Talita Roell, University of São Paulo, Brazil
Luiz Campos, Federal University of Rio Grande do Sul, Brazil

***Correspondence:**
Lin Tan
hqltanlin@163.com
Zhicheng Zhou
zhouzchnyc@126.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Behavioral and Evolutionary Ecology, a section of the journal Frontiers in Ecology and Evolution

Received: *05 May 2021*
Accepted: *06 August 2021*
Published: *08 September 2021*

Citation:
Wu S, Zeng W, Deng W, Li M, Hu W, Cai H, Li Y, Xie P, Tan L and Zhou Z (2021) Egg Cannibalism Varies With Sex, Reproductive Status, and Egg and Nymph Ages in Arma custos (Hemiptera: Asopinae). Front. Ecol. Evol. 9:705318. doi: 10.3389/fevo.2021.705318

INTRODUCTION

In China, *Arma custos* Fallou (Hemiptera: Asopinae) (Zhao et al., 2018) is a well-known predaceous pentatomid that prefers to prey on insects of coleopterans, hymenopterans, hemipterans, and lepidopterans (Zou et al., 2012; Wu et al., 2020). Since the 1970s, it has been used to control many agricultural and forestry pests in China (Chai et al., 2000; David and Zheng, 2002; Gao et al., 2011; Zou et al., 2012). Target species include *Leptinotarsa decemlineata* (Say), *Spodoptera litura* (fabricius), *Beet armyworm* (Hübner), and *Hyphantria cunea* (Drury) (Chai et al., 2000; David and Zheng, 2002; Gao et al., 2011; Zou et al., 2012), with an efficiency of over 60%

(Zheng and Chen, 1992; Gao et al., 2011). In recent years, the application area of *A. custos* has been increasing annually, exceeding 100,000 ha in China in 2018 (Zhang and Jia, 2018).

Egg cannibalism is a very common behavior in nature (Fox, 1975; Polis, 1981), but it is unfavorable in large-scale breeding as it increases breeding costs and reduces biocontrol efficiency (Kuriwada et al., 2013). However, egg cannibalism occurs because the acquisition of nutrients directly benefits the consumer through increased growth and development rate, and the elimination of conspecific competition is also an indirect benefit (Block and Stoks, 2004; Richardson et al., 2010). Some insects cannibalize eggs to avoid starvation (Pizzatto and Shine, 2008; Dobler and Koelliker, 2010), such as flour beetles (*Tribolium confusum* Duval) (Parsons et al., 2013), European earwigs (*Forficula Auricularia* Linnaeus) (Dobler and Koelliker, 2010, 2011), and *Adalia bipunctata* Linnaeus (Agarwala and Dixon, 1992). In addition, many factors can influence egg cannibalism behavior, including predator gender (Revynthi et al., 2018a,b), predator or prey density (Kakimoto et al., 2003; Erica et al., 2011), the kinship between predator and prey (Samu et al., 1999; Hoffman, 2012; Parsons et al., 2013), and the reproductive status of the predator (Momen and AbdelKhalek, 2009; Bayoumy and Michaud, 2015; Maleknia et al., 2016).

Many Hemiptera insects are known for their cannibalistic behavior, such as *Arizona backswimmers Frank* (Zalom, 1978), *Triatoma infestans* Klug and *Triatoma sordida* Stal (Ryckman, 1951), *Arctocorisa carinata* and *Callicorixa producta* (Pajunen and Pajunen, 1991), *Xylocoris flavipes* Reuter (Richard, 1979), *Notonecta hoffmanni* (Orr et al., 1990), *Dicyphus cerastii* Wagner, and *Macrolophus pygmaeu*s Ranmbur (Goncalo et al., 2020) which would prey on its immature offspring. Our previous research showed that female *A. custos* were more active predators (ate more eggs than males) (Wu et al., 2020), while males did not participate in egg cannibalism. However, it remains unclear whether reproductive status affects egg cannibalism by *A. custos.* There may be differences in both male and female *A. custos* between mated and virgin stages. For example, Bayoumy and Michaud (2015) found that virgin female *Hippodamia convergens* Guerin-Meneville (Coleoptera) are less cannibalistic than mated adults, while males are more cannibalistic than mated adults. Meanwhile, we identified the cannibalistic behavior by adults mainly occurred 48 h after initiating the experiments. We also found that neither *A. custos* males nor females consumed all the available eggs, regardless of the predator-to-prey ratio. This may be due to egg age or because adults did not want to consume all the eggs. In addition, although egg cannibalism by larvae has been widely studied, few studies have investigated egg cannibalism by *A. custos* larvae. Pajunen and Pajunen (1991) found that female rock-pool corixids cannibalized new eggs more frequently than 1-day-old eggs. It is difficult to construct experiments in the laboratory that interact nymphs with eggs, but in the natural environment where nymphs may change and encounter other *A. custos* eggs. Significant egg cannibalism has been found in many insects, such as *Formica aquilonia* Yarrow (Hymenoptera) (Schultner et al., 2013) and *Triboliunm castaneum* Herbst (Coleoptera) (Frank and Peter, 1966). Moreover, older nymphs disperse further, requiring more energy than earlier-stage nymphs; thus, older nymphs may consume more eggs in natural environments. Therefore, it is necessary to investigate egg cannibalism by *A. custos* nymphs. This research not only improves our understanding of cannibalism evolutionary selectivity but also increases the efficiency of larger-scale breeding and release biological programs.

To identify the relationship between egg cannibalism and reproductive status, egg age, sex, and life stage of *A. custos*, the following three hypotheses were tested in the laboratory. (1) Mated female *A. custos* devour more eggs than virgin females, and reproductive status has no effect on male egg cannibalism. (2) Younger eggs are more likely to be eaten than older ones. (3) Older nymphs consume more eggs than younger nymphs.

MATERIALS AND METHODS

Experimental Insects and Conditions

Arma custos adults were collected in Langfang, Hebei Province, China, in 2018 and taken to the laboratory. Specimens were fed Chinese oak silk moth pupae (*Antherea pernyi* Guer.) purchased from a supermarket in Liaoning, China. The insects were reared in artificial climate boxes as previously described by Pan et al. (2019). Briefly, the first instar nymphs of *A. custos* from a single egg mass were placed in transparent plastic dishes and fed only water via a piece of moist absorbent cotton. Chinese oak silk moth pupae were provided food from the second instar nymph stage, with the supply replenished every 4 to 5 days (Pan et al., 2019).

Recently emerged adults (<6 h old) were paired and placed in a 6 × 10 cm Petri dish, and they would mate after 5–6 days. The Petri dish contained one Chinese oak silk moth pupae lined with a piece of paper and formed into a tube for food (diameter: 1 cm, height: 6 cm). The female had laid eggs after 4–5 d and the insects (and the remains of the moth pupa) were removed. Eggs laid on the paper were used in the experiment within 24 h.

We assessed the cannibalistic behavior of adult *A. custos* under laboratory conditions by placing recently laid eggs (<24 h old) in small plastic dishes (10 × 1.5 cm) covered with an insect-proof screen (80-μm mesh) for ventilation. All adults were 8 days old and had been starved for 24 h before beginning the experiment. The specimens used included females that had previously laid eggs and males that had previously mated. All treatments were performed at 25°C ± 2°C, 60% ± 10% relative humidity, and a photoperiod of 16:8 (L:D).

Egg Cannibalism in Different Female and Male *A. custos* Developmental Stages

We performed four treatments to verify whether the female and male *A. custos* cannibalistic behavior was affected by the age of the female. The treatments were the following: virgin and mated. The adults were placed in plastic cups containing 30 eggs (<24 h) that were not their own offspring. Each treatment had 30

replicates with 30 different adults, and each cup was considered an experimental unit.

Egg Cannibalism by Female and Male *A. custos* of Different Aged Eggs

We performed two tests to verify whether female and male *A. custos* cannibalistic behavior was related to egg developmental stage: (1) with different aged eggs and (2) supplying new eggs (<24 h) after 48 h.

(1) *Different aged egg tests*: We collected and dated eggs of a female *A. custos* and stored them in different boxes according to date so that we could distinguish the eggs' age. Five treatments were compared to verify whether the female *A. custos* cannibalistic behavior was affected by the age of the eggs in free-choice and no-choice conditions. Each treatment involved 30 replicates with 30 different adults, and each cup was considered an experimental unit.

No-choice test: one female or male *A. custos* was placed in a plastic cup containing 24, 48, 72, 96, and 120-h-old eggs. Every plastic cup contained 30 eggs that were not offspring of the female.

Free-choice test: at the beginning of the free-choice experiments, we drew five flabellate grids on the Petri dishes, which divided the dishes into five equal parts. Then six eggs of 24, 48, 72, 96, and 120 h old were put into the five equal parts of Petri dishes at random. Then, one female or male *A. custos* was placed in a plastic cup containing 30 eggs.

(2) *Supplying new eggs test*: This experiment involved two treatments, as follows: supplying new eggs (<24 h) after 48 h and not supplying new eggs (<24 h) after 48 h. The adults were placed in plastic cups containing 30 eggs that were not their offspring. Each treatment had 30 replicates with 30 different adults, and each cup was considered an experimental unit.

Egg Cannibalism in Different Nymph Developmental Stages

We performed five treatments to verify whether the *A. custos* cannibalistic behavior was affected by the age of the nymph. Treatments included: first, second, third, fourth, and fifth instar nymphs. All first, second, third, fourth and fifth instar nymphs were 1 day old and had been starved for 24 h before beginning the experiment. The nymphs were placed in plastic cups containing 30 non-related eggs (<24 h). Each treatment had 30 replicates with 30 different nymphs, and each cup was considered an experimental unit.

Data Collection

In all treatments, *A. custos* adults were monitored for 5 min post-release, and further observations were carried out every 24 h until the death of all individuals. The number of unconsumed eggs and consumed eggs (broken eggshells) was counted. We determined (1) the cannibalistic behavior of *A. custos* adults and (2) the number of consumed and unconsumed eggs. From previous research, eggs with broken and diaphanous eggshells were considered cannibalized eggs. We did not provide additional food to males and females after initiating the cannibalism experiments.

Statistical Analysis

A chi-squared test was used to estimate whether the development progress of the eggs, nymphs, and females was related to their cannibalism. An analysis of variance was performed to verify the differences between female *A. custos* egg consumption and egg emergence ratio. Bartlett's test was used to test the homogeneity of variances, and sqrt was used to analyze datasets with $p < 0.05$. Multiple comparisons were performed using Tukey's HSD test. All analyses was performed using R v.3.3.3 (R Development Core Team, 2017).

RESULTS

Egg Cannibalism by Female and Male *A. custos* at Different Developmental Stages

Virgin females showed a lower egg cannibalism inclination than gravid females. Both virgin and mated females showed a higher egg cannibalism inclination than virgin and mated males (**Figure 1**). The number of virgin females showing egg cannibalism behavior was significantly lower than that of mated females (**Figure 1A**: chi-squared test, $\chi^2 = 6.87$, $df = 1$, $p < 0.05$). The number of virgin females showing egg cannibalism behavior was four times higher than virgin males (**Figure 1A**: chi-squared test, $\chi^2 = 5.69$, $df = 1$, $p < 0.05$). The number of mated females showing egg cannibalism behavior was eight times higher than

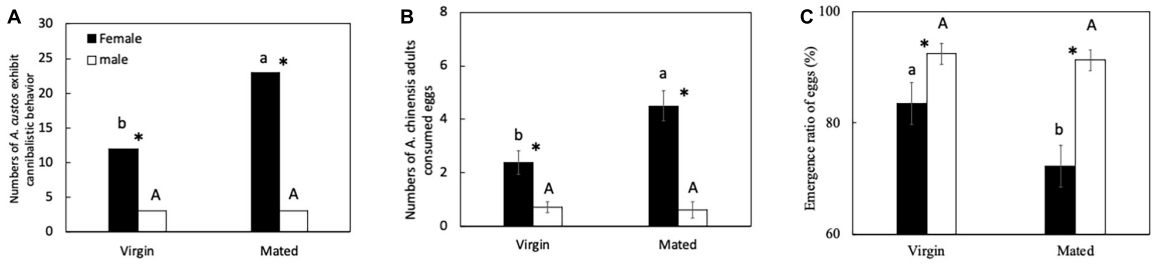

FIGURE 1 | (A–C) Predatory behavior by *A. custos* on different-aged eggs. Different lowercase letters indicate significant differences among females at $p < 0.05$ and different uppercase letters indicate significant differences among males at $p < 0.05$. The investigation was repeated 30 times. * indicates a significant difference between the black bar and white bar at $p < 0.05$ (female and male virgins or female and male mated individuals).

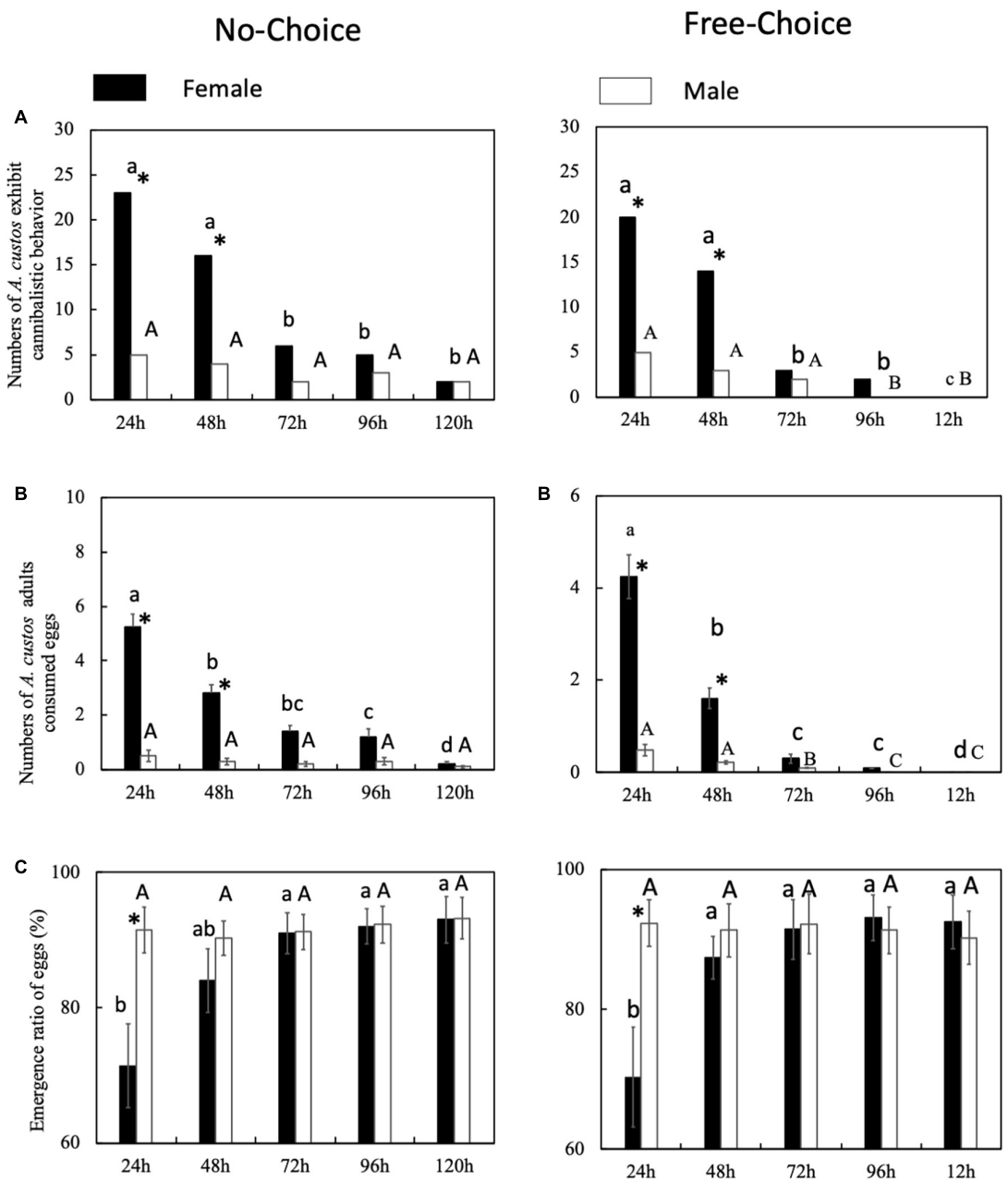

FIGURE 2 | (A–C) Predatory behavior by *A. custos* with different aged eggs under no choice and free choice conditions. Different lowercase letters indicate significant differences among females at $P < 0.05$ and different uppercase letters indicate significant differences among males at $p < 0.05$. The investigation was repeated 30 times. * indicates a significant difference between the black bar (females) and white bar (males) at $P < 0.05$.

mated males (**Figure 1A**: chi-squared test, $\chi^2 = 24.50$, $df = 1$, $p < 0.05$). Meanwhile, the number of eggs consumed by virgin female *A. custos* was lower than that of mated females (**Figure 1B**: Tukey HSD test, $p < 0.05$). Both virgin and mated females consumed more eggs than virgin and mated males (**Figure 1B**: All, Tukey HSD test, $p < 0.05$). The egg emergence ratio when cannibalized by virgin females was significantly higher than that of mated females. The egg emergence ratio when cannibalized by both virgin and mated females was significantly lower than that of virgin and mated males (**Figure 1C**: Tukey HSD test, $p < 0.05$).

Egg Cannibalism of *A. custos* to Different Ages of Eggs

The predatory behavior by female *A. custos* is affected by the egg development stage. After supplying new eggs (at 48 h), female *A. custos* exhibited a more active predatory behavior than

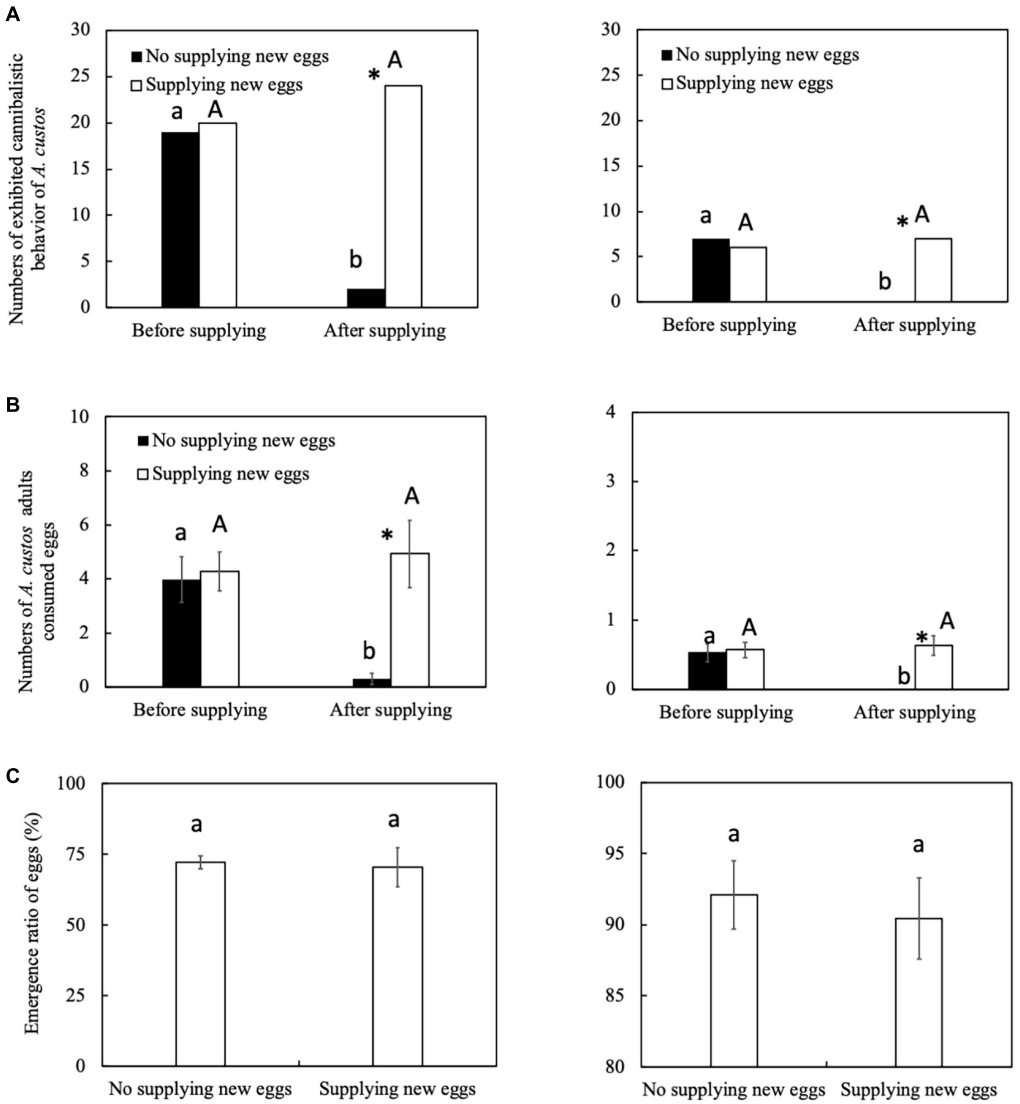

FIGURE 3 | (A–C) Predatory behavior by *A. custos* when supplied new eggs. Different lowercase or uppercase letters indicate a significant difference of different treatments at $P < 0.05$. The investigation was repeated 30 times. * indicates a significant difference between the black bar (not supplied new eggs) and white bar (supplied new eggs) at $p < 0.05$.

female *A. custos* who were not supplied with new eggs after 48 h (**Figure 2A**: before supplying new eggs, chi-squared test, $\chi^2 = 0$, $df = 1$, $p = 1$; after supplying new eggs, chi-squared test, $\chi^2 = 29.93$, $df = 1$, $p < 0.05$). Female *A. custos* who were supplied with new eggs consumed more eggs than the female *A. custos* that were not supplied with new eggs after 48 h (**Figure 2B**: before supplying new eggs, Tukey HSD test, $p = 0.78$; after supplying new eggs, Tukey HSD test, $p < 0.05$). However, supplying new eggs did not affect the emergence ratio of the eggs (**Figure 2C**: Tukey HSD test, $p < 0.05$) and more than 70% of eggs emerged (**Figure 2C**).

Female *A. custos* cannibalized old eggs less than young eggs (**Figure 3**). The cannibalism on eggs that had been laid over 72 h earlier was significantly lower than the cannibalism on younger eggs (**Figure 3A**: 48–72 h, No choice, chi-squared test, $\chi^2 = 5.81$, $df = 1$, $p < 0.05$; Free choice, $\chi^2 = 8.21$, $df = 1$, $p < 0.05$). However, eggs that had been laid over 48 h before were also consumed significantly less often than younger eggs (**Figure 3B**: 24–48 h, No choice, Tukey HSD test, $p < 0.05$; Free choice, Tukey HSD test, $p < 0.05$). In addition, in the no-choice conditions, the number of eggs consumed by female *A. custos* on 72-h-old eggs was not significantly lower than the consumption of eggs that were 48 h old (**Figure 3B**: 48–72 h, No choice, Tukey HSD test, $p > 0.05$). The number of 72-h-old eggs consumed by female *A. custos* was significantly lower than the consumption of 48-h-old eggs in the free choice conditions (**Figure 3B**: 48–72 h, Free

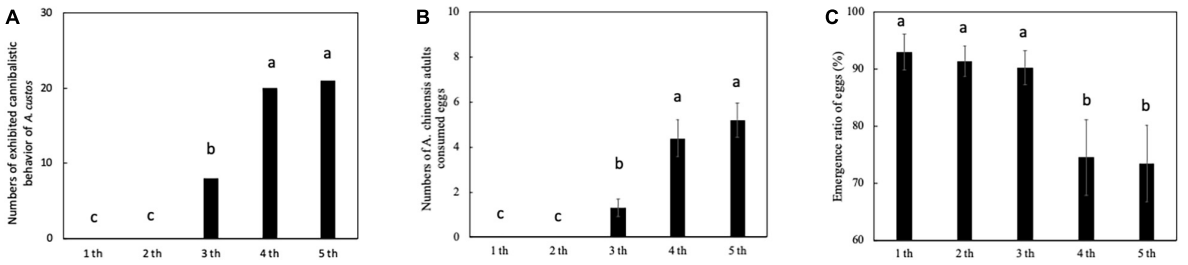

FIGURE 4 | (A–C) Predatory behavior by *A. custos* in different nymph developmental stages. Different lowercase letters indicate a significant difference in different nymph developmental stages at $p < 0.05$. The investigation was repeated 30 times.

choice, Tukey HSD test, $p < 0.05$). Meanwhile, the number of 24-h-old eggs consumed by female *A. custos* was significantly higher than the 48 h eggs regardless of the conditions (choice or free choice) (**Figure 3B**: 48–72 h, No choice, Tukey HSD test, $p < 0.05$; Free choice, Tukey HSD test, $p < 0.05$). In addition, the emergence ratio of the 24-h-old eggs was significantly lower than the other aged eggs in both the no-choice and free-choice conditions (**Figure 3C**: No choice, Tukey HSD test, $p < 0.05$; Free choice, Tukey HSD test, $p < 0.05$).

Egg Cannibalism by Nymphs at Different Developmental Stages

The developmental stage of the nymph significantly influenced the cannibalistic behavior by nymph *A. custos* (**Figure 4**). First and second instar nymphs did not cannibalize eggs. The third, fourth, and fifth instar nymphs cannibalized eggs (**Figure 4**). The number of third instar *A. custos* nymphs exhibiting cannibalistic behavior was significantly lower than that of fourth instar nymphs (**Figure 4A**: chi-squared test, $\chi^2 = 8.10$, $df = 1$, $p < 0.05$) and fifth instar nymphs (**Figure 4A**: chi-squared test, $\chi^2 = 9.61$, $df = 1$, $p < 0.05$). In addition, third instar nymph females consumed fewer eggs than fourth or fifth instar nymph females (**Figure 4B**: Tukey HSD test, $p < 0.05$). In addition, the egg emergence ratios in the first, second, and third instar nymph experiments were significantly higher than the fourth and fifth instar nymph experiments (**Figure 4C**: Tukey HSD test, All, $p < 0.05$).

DISCUSSION

Egg cannibalism is influenced by many factors (Samu et al., 1999; Kakimoto et al., 2003; Momen and AbdelKhalek, 2009; Erica et al., 2011; Hoffman, 2012; Parsons et al., 2013; Maleknia et al., 2016). Sex, reproductive status, and egg age are major factors in egg cannibalism (Momen and AbdelKhalek, 2009; Maleknia et al., 2016).

Our results indicate that virgin females showed a lower egg cannibalism inclination than gravid females, both virgin and mated females showed a higher egg cannibalism inclination than virgin and mated males, the number of eggs consumed by virgin female *A. custos* was lower than that of mated females, both virgin and mated females consumed more eggs than virgin and mated males, and both virgin and mated females showed a lower egg cannibalism inclination, supporting our first hypothesis: mated female *A. custos* devour more eggs than virgin females, and reproductive status has no effect on male egg cannibalism. We found that virgin females showed a lower egg cannibalism inclination than gravid females. This result indicates that egg cannibalism behavior by *A. custos* is influenced by the developmental stage of the female. This may be because gravid females need more energy for breeding (Neff, 2003; Miller and Zink, 2012). Therefore, we believe that oviposition may be one of the reasons for *A. custos* egg cannibalism behavior. Our results were similar to those of female *Hippodamia convergens* Guerin-Meneville (Coleoptera) (Bayoumy and Michaud, 2015) and some mite species (Schausberger, 2003). Conversely, our results were different from those of *Tribolium confusum* Duval (Parsons et al., 2013), *Coccinella undecimpunctata* L. (Bayoumy et al., 2016), and male *Hippodamia convergens* Guerin-Meneville (Coleoptera) (Bayoumy and Michaud, 2015), whose virgin adults are more cannibalistic than mated adults. However, whether female spawning behavior is related to predation remains unresolved. Does female spawning behavior stimulate egg cannibalism behavior? In future studies we will attempt to answer this question by inhibiting the expression of genes related with oviposition.

Our results also indicate that eggs aged <48 h were more likely to be cannibalized under no-choice and free-choice conditions. When we supplied new eggs after 48 h, the control group exhibited higher egg cannibalism, consuming more eggs. Therefore, we believe that egg development stage can inhibit egg cannibalism behavior by female *A. custos*. These results support our second hypothesis: younger eggs are more likely to be eaten than older ones. Furthermore, egg cannibalism behavior may be influenced by egg development as female *A. custos* no longer preyed on the eggs when the test eggs were older than 48 h. This may be because eggs have start to develop 48 h after being laid. Our results are similar to female rock-pool corixids, who cannibalized new eggs more frequently than 1-d-old eggs (Pajunen and Pajunen, 1991).

Nymphs are an important stage in the life cycle of *A. custos*. In our experiments, we supplied suitable food for nymphs from the second instar stage in the laboratory. Previous studies have shown that there are many younger cannibalistic insect that prey on eggs, such as *Formica aquilonia* Yarrow (Schultner et al., 2013) and *Triboliunm castaneum* Herbst (Frank and Peter, 1966).

Our results showed that the developmental stage of the nymphs significantly influences the cannibalistic behavior by nymph *A. custos*. We found that older nymphs cannibalize more eggs than younger nymphs, supporting our third hypothesis: older nymphs consume more eggs than younger nymphs. The first and second instar nymphs did not cannibalize eggs, which is unlike the third, fourth and fifth instar nymphs who cannibalized eggs. This result is similar to *Hippodamia convergens* Guerin-Meneville (Coleoptera) (Bayoumy and Michaud, 2015). However, the reason why both adults and nymphs avoid eating +48-h old eggs remains unresolved. It may be that adults and nymphs can distinguish between developed and undeveloped eggs? Alternatively, there are other potential explanations. For example, eggs may exude a repellent that restricts cannibalistic behavior (Narasimha et al., 2019) or it may be because adults refrain from eating more eggs (Polis, 1981; Smith and Reay, 1991; Manica, 2002). Further research is required to understand the causes of this behavior.

Previous studies have also demonstrated that selective cannibalism provides a food source for adult insects while also ensuring the survival of most of their offspring, thus maintaining the population levels (Polis, 1981; Smith and Reay, 1991; Manica, 2002). In this instance, *A. custos* females would not consume all available eggs, regardless of the predator-to-prey ratio. Similar results have also been obtained for *Anisolabis maritima* Bon. (Miller and Zink, 2012) and *Euborellia annulipes* Lucas (Jacobs and Stigall, 2019). Our results also demonstrate that neither *A. custos* nymphs nor female adults consumed all of the available eggs, and the emergence ratio was > 70%. This result indicates that most eggs were left alive. Obviously, there was a balance and a significant decision by female *A. custos* to consume less than one-third of the offspring, like *Anisolabis maritima* Bon. (Miller and Zink, 2012), and allowed the majority of the offspring to survive. This selective cannibalism would not only extend the lifespan of the adult insects but also ensure the survival of most offspring, which is more helpful in preventing populations from dwindling (Polis, 1981; Smith and Reay, 1991; Manica, 2002). Our results were similar to those for *Euborellia annulipes* Lucas (Dermaptera: Anisolabididae) (Jacobs and Stigall, 2019). This process is not only beneficial for the health of offspring, but also beneficial for females (Okada et al., 2015).

The evolution of cannibalism is driven by the balance between its benefits and costs (Hamilton, 1964). One evident benefit of egg cannibalism is starvation avoidance (Pizzatto and Shine, 2008; Dobler and Koelliker, 2010). Moreover, selective cannibalism provides an alternative food source for adult insects, while ensuring the survival of most of their offspring and maintaining the population size (Polis, 1981; Smith and Reay, 1991; Manica, 2002). Our study revealed that hungry *A. custos* do not prey on all the available eggs (>70% of the eggs were unconsumed, and males exhibit minimal egg cannibalism). Moreover, virgin females showed a lower egg cannibalism inclination than gravid females; first and second instar nymphs did not cannibalize the eggs. Egg cannibalism offers insect species a means to avoid starvation and prolong lifespan by providing an alternative source of nutrition and energy (Polis, 1981; Smith and Reay, 1991; Pizzatto and Shine, 2008; Okada et al., 2015). However, under field conditions, whether *A. custos* would exhibit the same behavior, possibly avoiding egg cannibalism by protecting at least some of its eggs and instead searching for other prey, is still unknown (Revynthi et al., 2018b). Furthermore, egg cannibalistic behavior has also been reported in the nymphs of *F. aquilonia* Yarrow (Schultner et al., 2013) and *T. castaneum* Herbst (Frank and Peter, 1966). Given that, compared with laboratory conditions, nymphs generally have little difficulty in locating eggs in the wild, it will be instructive to investigate egg cannibalism among *A. custos* nymphs, as well as egg cannibalism as a whole, under natural conditions. However, our observations were conducted under confined conditions, and it remains to be determined whether *A. custos* adult would exhibit the same behavior at more extensive spatial scales.

CONCLUSION

Overall, *A. custos* females exhibit a higher tendency for egg cannibalism than males. Egg cannibalism varies not only with the developmental stage of the eggs and nymphs but also with the sex and reproductive status of *A. custos* females. Virgin females showed a lower egg cannibalism inclination than gravid females. Both virgin and mated females showed a higher egg cannibalism inclination than virgin and mated males. First and second instar nymphs did not cannibalize the eggs. The third, fourth, and fifth instar nymphs cannibalized eggs, and younger eggs were more often eaten than older eggs. However, neither *A. custos* nymphs nor female adults consumed all of the available eggs, and the emergence ratio of the remaining eggs was >70%. Our findings help us to better understand the evolutionary relationships in egg cannibalism by *A. custos* and contribute to the efficient mass rearing and realization of *A. custos* in biological control systems.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

SW, WZ, ML, ZZ, and WD performed the experiments. SW, WZ, WH, ZZ, and WD conceived and designed the experiments, analyzed the data, and wrote the manuscript. All authors have read and approved the final manuscript.

FUNDING

Funding for this study was provided by the Major Program for Science and Technology of Hunan Province (No. 2020NK2034), Natural Science Foundation of Hunan Province (No. 2020JJ5290), and Double First-Class Construction Project of Hunan Agricultural University (No. SYL 2019029). The Natural Enemies Breeding project was supported by the Science Foundation of Hunan (No. 18-21 Aa06).

ACKNOWLEDGMENTS

We thank the following individuals: Changhua Zhang and Fangzhao Jia for supplying the insects; Anfang Zhu for feeding insects.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.705318/full#supplementary-material

REFERENCES

Agarwala, B. K., and Dixon, A. F. G. (1992). Laboratory study of cannibalism and interspecific predation in ladybirds. *Ecol. Entomol.* 17, 303–309. doi: 10.1111/j.1365-2311.1992.tb01062.x

Bayoumy, M. H., and Michaud, J. P. (2015). Egg cannibalism and its life history consequences vary with life stage, sex, and reproductive status in Hippodamia convergens (Coleoptera: Coccinellidae). *J. Econ. Entomol.* 108, 1665–1674. doi: 10.1093/jee/tov148

Bayoumy, M. H., Abou-Elnaga, A. M., Ghanim, A. A., and Mashhoot, G. A. (2016). Egg cannibalism potential benefits for adult reproductive performance and offspring fitness of coccinella undecimpunctata L. (Coleoptera: Coccinellidae). *Egypt. J. Biol. Pest. Control.* 26, 35–42.

Block, M. D., and Stoks, R. (2004). Cannibalism-mediated life history plasticity to combined time and food stress. *Oikos* 106, 587–597. doi: 10.2307/3548381

Chai, X. M., He, Z. H., Jiang, P., Wu, Z., Pan, C. R., and Hu, R. D. (2000). Studies on natural enemies of dendrolimus punctatus in zhejiang province. *Forestry Technol. Zhejiang.* 61, 1–56.

David, A., and Zheng, L. Y. (2002). Checklist and nomenclatural notes on the Chinese Pentatomidae (Heteroptera) I. asopinae. *Entomotaxonomia* 24, 107–115. doi: 10.1520/STP11985S

Dobler, R., and Koelliker, M. (2010). Kin-selected siblicide and cannibalism in the European earwig. *Behav. Ecol.* 21, 257–263. doi: 10.1093/beheco/arp184

Dobler, R., and Koelliker, M. (2011). Influence of weight asymmetry and kinship on siblicidal and cannibalistic behaviour in earwigs. *Anim Behav.* 82, 667–672. doi: 10.1016/j.anbehav.2011.06.017

Erica, L. W., Douglas, P. C., Joseph, M. K., and Andrew, R. B. (2011). The effects of food level and conspecific density on biting and cannibalism in larval long-toed salamanders, *Ambystoma macrodactylum*. *Oecol* 128, 202–209. doi: 10.1007/s004420100641

Fox, L. R. (1975). Factors influencing cannibalism, a mechanism of population limitation in the predator Notonecta hoffmanni. *Ecol.* 56, 933–941. doi: 10.2307/1936303

Frank, K. H., and Peter, S. D. (1966). Egg cannibalism by tribolium larvae. *Ecol. Soc. Am.* 47, 318–322. doi: 10.2307/1933784

Gao, Z., Wang, X., Zhang, L., Sun, Y., Fan, J., and Wang, G. (2011). Biological characteristic of Arma chinensis. *J. Eng. Heilongjiang Univ.* 2, 72–77.

Goncalo, A. D., Filipa, C., Ariadna, P., Elsa, B. S., and Elisabete, F. (2020). Intraguild predation and cannibalism among dicyphini: dicyphus cerastii vs. two commercialized species. *Entomol. Exp. Appl.* 169, 90–96. doi: 10.1111/eea.12943

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *J. Theor. Biol.* 7, 17–52.

Hoffman, C. R. (2012). *Kinship and Familiarity Affect Recognition and Foraging in the Wolf Spider, Pardosa Milvina (Araneae: Lycosidae). Ph. D, Thesis*. Miyami University.

Jacobs, A. C., and Stigall, T. (2019). Paternity and egg cannibalism in the ringlegged earwig Euborellia annulipes (Dermaptera: Anisolabididae). *Entomol. Sci.* 22, 250–257. doi: 10.1111/ens.12363

Kakimoto, T., Fujisaki, K., and Miyatake, T. (2003). Egg laying preference, larval dispersion, and cannibalism in Helicoverpa armigera (Lepidoptera: Noctuidae). *Ann. Entomol. Soc. Am.* 96, 793–798.

Kuriwada, T., Kumano, N., and Haraguchi, S. D. (2013). High population density and egg cannibalism reduces the efficiency of mass-rearing in euscepes postfasciatus (Coleoptera: Curculionidae). *Fla Entomol.* 92, 221–228. doi: 10.1653/024.092.0205

Maleknia, B., Fathipour, Y., and Soufbaf, M. (2016). Intraguild predation among three phytoseiid species, Neoseiulus barkeri, Phytoseiulus persimilis and Amblyseius swirskii. *Syst. Appl. Acarol.* 21, 417–426. doi: 10.11158/saa.21.4.4

Manica, A. (2002). Alternative strategies for a father with a small brood: mate, cannibalize or care. *Behav. Ecol. Sociobiol.* 51, 319–323. doi: 10.1007/s00265-001-0444-0

Miller, J. S., and Zink, A. G. (2012). Parental care trade-offs and the role of filial cannibalism in the maritime earwig. Anisolabis Maritima. *Anim. Behav.* 83, 1387–1394. doi: 10.1016/j.anbehav

Momen, F. M., and AbdelKhalek, A. (2009). Cannibalism and intraguild predation in the phytoseiid mites *Typhlodromips swirskii*, *Euseius scutalis* and *Typhlodromus athiasae* (Acari: Phytoseiidae). *Acari* 17, 223–229.

Narasimha, S., Nagornov, K. O., Menin, L., Mucciolo, A., Rohwedder, A., Humbel, B. M., et al. (2019). Drosophila melanogaster cloak their eggs with pheromones, which prevents cannibalism. *PLoS Biol.* 17:e2006012. doi: 10.1371/journal.pbio.2006012

Neff, B. D. (2003). Paternity and condition affect cannibalistic behavior in nest-tending bluegill sunfish. *Behav. Ecol. Sociobiol.* 54, 377–384. doi: 10.1016/j.yhbeh.2009.05.002

Okada, S., Fukuda, Y., and Takahashi, M. K. (2015). Paternal care behaviors of Japanese giant salamander *Andrias japonicus* in natural populations. *J. Ethol.* 33, 1–7. doi: 10.1007/s10164-014-0413-5

Orr, B. K., Murdoch, W. W., and Bence, J. R. (1990). Population regulation, convergence, and cannibalism in notonecta (Hemiptera). *Ecology* 71, 68–82. doi: 10.2307/1940248

Pajunen, V. I., and Pajunen, I. (1991). Oviposition and egg cannibalism in rock-pool corixids (Hemiptera: Corixidae). *Oikos* 60, 83–90.

Pan, M. Z., Zhang, H., Zhan, L., and Chen, H. (2019). Effects of starvation and prey availability on predation and dispersal of an omnivorous predator Arma chinensis Fallou. *J. Insect. Behav.* 32, 134–144. doi: 10.1007/s10905-019-09718-9

Parsons, W., Zhong, W., and Rudolf, V. H. W. (2013). Mating status and kin recognition influence the strength of cannibalism. *Anim. Behav.* 85, 365–369. doi: 10.1016/j.anbehav.2012.11.006

Pizzatto, L., and Shine, R. (2008). The behavioral ecology of cannibalism in cane toads (Bufo marinus). *Behav. Ecol. Sociobiol.* 63, 123–133. doi: 10.1007/s00265-008-0642-0

Polis, G. (1981). The evolution and dynamics of intraspecific predation. *Annu. Rev. Ecol. Syst.* 12, 251–255. doi: 10.1146/annurev.es.12.110181.001301

R Development Core Team. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. Available at: https://www.R-project.org/

Revynthi, A. M., Egas, M., Janssen, A., and Sabelis, M. W. (2018a). Prey exploitation and dispersal strategies vary among natural populations of a predatory mite. *Ecol. Evol.* 8, 10384–10394. doi: 10.1002/ece3.4446

Revynthi, A. M., Janssen, A., Egas, M. E., and Acarology, A. (2018b). Gender-specific differences in cannibalism between a laboratory strain and a field strain of a predatory mite. *Exp. Appl. Acarol.* 74, 239–247. doi: 10.1007/s10493-018-0232-4

Richard, T. A. (1979). Cannibalism in *Xylocoris flavipes* (*Hemiptera: anthocoridae*), a predator of stored-product insects. *Ent. Exp. Appl.* 25, 128–135.

Richardson, M. L., Mitchell, R. F., Reagel, P. F., and Hanks, L. M. (2010). Causes and consequences of cannibalism in non-carnivorous insects. *Annu. Rev. Entomol.* 55, 39–53. doi: 10.1146/annurev-ento-112408-085314

Ryckman, R. E. (1951). Section 1 | | recent observations of cannibalism in triatoma (Hemiptera: Reduviidae). *J. Parasitol.* 37, 433–434.

Samu, F., Toft, S., and Kiss, B. (1999). Factors influencing cannibalism in the wolf spider *Pardosa agrestis* (Araneae: Lycosidae). *Behav. Ecol. Sociobiol.* 45, 349–354. doi: 10.1007/s002650050570

Schausberger, P. (2003). Cannibalism among phytoseiid mites: a review. *Exp. Appl. Acarol.* 29, 173–191. doi: 10.1023/A:1025839206394

Schultner, E., d'Ettorre, P., and Helantera, H. (2013). Social conflict in ant larvae: egg cannibalism occurs mainly in males and larvae prefer alien eggs. *Behav. Ecol.* 24, 1306–1311. doi: 10.1093/beheco/art067

Smith, C., and Reay, P. (1991). Cannibalism in teleost fish. *Rev. Fish. Biol. Fish.* 1, 41–64. doi: 10.1007/BF00042661

Wu, S. L., Deng, W., Cai, H. L., Yang, J. S., Zeng, W. A., Zhou, Z. C., et al. (2020). The occurrence period and effect of intraspecific cannibalism behavior of arma chinensis under starvation. *Chin. J. Biol.* 36, 169–174. doi: 10.16409/j.cnki.2095-039x.2020.03.001

Zalom, F. G. (1978). Backswimmer prey selection with observations on cannibalism (Hemiptera: Notonectidae). *Southwestern Naturalist* 23, 617–622.

Zhang, C. H., and Jia, F. Z. (2018). Predatory insects Arma chinensis. *Chin. Tobacco Sci.* 24:135.

Zhao, Q., Wei, J., Wenjun, B. U., Liu, G., and Zhang, H. (2018). Synonymize Arma chinensis as Arma custos based on morphological, molecular and geographical data. *Zootaxa* 4455:161.

Zheng, Z. Y., and Chen, Y. W. (1992). Experiments on the use of Arma Custos (Fabricius) to control forist pesrs. *Chin. J. Biol. Control* 8, 155–156.

Zou, D., Wang, M., Zhang, L., Zhang, Y., Zhang, X., and Chen, H. (2012). Taxonomic and bionomic notes on Arma chinensis (Fallou) (Hemiptera: Pentatomidae: Asopinae). *Zootaxa* 3382, 41–52. doi: 10.16409/10.11646/zootaxa.3382.1.4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 06 October 2021
doi: 10.3389/fevo.2021.725504

Chemosensory Gene Expression for Two Closely Relative Species *Rhodnius robustus* and *R. prolixus* (Hemiptera, Reduviidade, Triatominae) Vectors of Chagas Disease

Axelle Marchant[1†], Florence Mougel[1†], Emmanuelle Jacquin-Joly[2], Carlos E. Almeida[3], Denis Blanchet[4], Jean-Michel Bérenger[5,6], João A. da Rosa[7] and Myriam Harry[1]*

[1] Université Paris-Saclay, CNRS, IRD, UMR EGCE, Evolution, Génomes, Comportement et Ecologie, Gif-sur-Yvette, France, [2] INRAE, Sorbonne Université, CNRS, IRD, UPEC, Université de Paris, Institute of Ecology and Environmental Sciences of Paris, Versailles, France, [3] Instituto de Biologia, Universidade Federal da Bahia (UFBA), Salvador, Brazil, [4] Laboratoire Hospitalier Universitaire de Parasitologie et Mycologie (LHUPM), UFR de Médecine, Université des Antilles et de la Guyane, Equipe EA 3593, Cayenne, French Guiana, [5] Muséum National d'Histoire Naturelle, Département Systématique & Evolution (Entomologie), Paris, France, [6] Aix Marseille Université, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France, [7] Faculdade de Ciências Farmacêuticas, Universidade Estadual Paulista (Unesp), Araraquara, Brazil

OPEN ACCESS

Edited by:
Andre Barretto Bruno Wilke, University of Miami, United States

Reviewed by:
Daniele Silva De Oliveira, Federal University of Rio de Janeiro, Brazil
Peng He, Guizhou University, China

***Correspondence:**
Myriam Harry
myriam.harry@egce.cnrs-gif.fr

[†]These authors share first authorship

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *15 June 2021*
Accepted: *02 September 2021*
Published: *06 October 2021*

Citation:
Marchant A, Mougel F, Jacquin-Joly E, Almeida CE, Blanchet D, Bérenger J-M, da Rosa JA and Harry M (2021) Chemosensory Gene Expression for Two Closely Relative Species Rhodnius robustus and R. prolixus (Hemiptera, Reduviidade, Triatominae) Vectors of Chagas Disease. Front. Ecol. Evol. 9:725504. doi: 10.3389/fevo.2021.725504

Two closely related species, *Rhodnius prolixus* and *Rhodnius robustus*, are the vectors of *Trypanosoma cruzi*, which is the causative agent of Chagas disease, but clearly exhibit clear-cut differences in their ecological behavior. *R. prolixus* is considered as a domiciliated species, whereas *R. robustus* only sporadically visits human houses in Amazonia. We performed a chemosensory gene expression study *via* RNA-sequencing (RNA-seq) for the two species and also included a laboratory introgressed *R. robustus* strain. We built an assembled transcriptome for each sample and for both sexes and compiled all in a reference transcriptome for a differential gene expression study. Because the genes specifically expressed in one condition and not expressed in another may also reflect differences in the adaptation of organisms, a comparative study of the presence/absence of transcripts was also performed for the chemosensory transcripts, namely chemosensory proteins (CSPs), odorant-binding proteins (OBPs), odorant receptors (ORs), gustatory receptors (GRs), and ionotropic receptors (IRs), as well as takeout (TO) transcripts because TO proteins have been proposed to be associated with chemosensory perception in both olfactory and taste systems. In this study, 12 novel TO transcripts from the *R. prolixus* genome were annotated. Among the 199 transcripts, out of interest, annotated in this study, 93% were conserved between *R. prolixus* and the sylvatic *R. robustus*. Moreover, 10 transcripts out of interest were specifically expressed in one sex and absent in another. Three chemosensory transcripts were found to be expressed only in the reared *R. prolixus* (CSP19, OBP9, and OR89) and only one in sylvatic *R. robustus* (OR22). A large set of transcripts were found to be differentially expressed (DE) between males and females (1,630), with a majority of them (83%) overexpressed in males. Between environmental conditions,

8,596 transcripts were DE, with most (67%) overexpressed in the sylvatic *R. robustus* samples, including 17 chemosensory transcripts (4 CSPs, 1 OBP, 5 ORs, 1 GR, 4 IR, and 2 TO), but 4 genes (OBP19, OR13, OR40, and OR79) were overexpressed in the reared samples.

Keywords: differential expression, chemosensory proteins (CSP), odorant binding protein (OBP), *Rhodnius prolixus*, *Rhodnius robustus*, Chagas disease, odorant and gustatory receptors (OR, GR)

INTRODUCTION

In Latin America, bloodsucking bugs (Hemiptera, Reduviidae, and Triatominae) are the vectors of the parasite *Trypanosoma cruzi* (Kinetoplastida and Trypanosomatidae), which is the cause of the Chagas disease affecting about 6 million people (PAHO/WHO, 2020). At present, 151 extant and 3 fossil species of Triatominae are described (Zhao et al., 2021), including 136 species restricted to Latin America. Most of the species living in natural ecotopes feed on a large variety of vertebrates, but some of them found in anthroposystems can feed on humans and transmit them the parasite *via* the infected feces.

Among the Chagas disease vectors, *Rhodnius prolixus* is one of the most famous vectors because of its epidemiological importance and its broad distribution in both Central and Northeastern Latin America. Chemical control campaigns have considerably reduced the prevalence of the disease due to this species vector. Therefore, in 2011, all previously endemic countries of Central America have been certified as free of Chagas disease transmission occurring due to *R. prolixus* (Hashimoto and Schofield, 2012), but it has been reported again in Mexico in 2019 (Antonio-Campos et al., 2019). *R. prolixus* is considered as fully domiciliated, but some researchers described the sylvatic populations of *R. prolixus* in Venezuela (Sanchez-Martin et al., 2006; Feliciangeli et al., 2007; Fitzpatrick et al., 2008) and Colombia (Guhl et al., 2009; Rendón et al., 2015).

Rhodnius robustus is closely related to *R. prolixus* and is widely distributed in the Amazon region. This species only sporadically visit human dwellings, which is being attracted by artificial light sources, as described in Peru (Cabrera et al., 2013), Guyana (Bérenger et al., 2009), Venezuela (Feliciangeli et al., 2002), and Brazil (Aguilar et al., 2007). Its sporadic visits of domiciles have favored the transmission of *T. cruzi* to humans (Castro et al., 2010) especially because this species exhibits a high prevalence of natural *T. cruzi* (Miles et al., 1983). Moreover, the mode of transmission involving an oral route through the contamination of fruit storages or palm juice (*açaí juice*) by the infected sylvatic insects, including *R. robustus,* is progressively increasing in the Amazon region (Valente et al., 1999; Coura et al., 2002; Monteiro et al., 2010).

According to Lent and Wygodzinsky (1979), *R. prolixus* and *R. robustus*—morphologically very closer to each other—are distinguished only by the leg color of the oldest larvae and some differences in the basal plates of the male genitalia. In addition, the two species are differentiated by wing morphometrics (Matias et al., 2001; Márquez and Saldamando-Benjumea, 2013) and external female genitalia (da Rosa et al., 2014). The lack of diagnostic allozyme markers for Venezuelan populations of *R. prolixus* and *R. robustus* (Harry et al., 1992; Harry, 1993a) and the absence of differences for the male median process of pygophore made Harry (1993b) to suggest that Venezuelan populations of *R. prolixus* and *R. robustus* were not distinct taxonomic units. Using *cytochrome b* sequences, the paraphyly of *R. robustus* was highlighted as the clade composed of *R. robustus* from the Orinoco region (Venezuela) was more closely related to *R. prolixus* clade in comparison to other *R. robustus* clades from the Amazon region. Then, *R. robustus* was described as a complex of five operational taxonomic units (Monteiro et al., 2003; Abad-Franch and Monteiro, 2007). In laboratory strains, the evidence of the introgression between *R. prolixus* and *R. robustus* has been reported as 11 of the 15 presuming *R. prolixus* colonies that are tested for the whole *R. prolixus* genome sequencing project, exhibited past introgression events with *R. robustus* (Mesquita et al., 2015).

Irrespective of their way of life (domiciliary or sylvatic), the survival, reproduction, and social communication of individuals rely on the chemosensory system. The detection and integration of environmental chemical signals, including smell and taste, allow organisms to detect food, hosts, and predators. This system plays a key role in ecological adaptation to new or changing hosts or habitats (Benton, 2015). The recognition of chemical molecules as signals depends on the activation of specific proteins, including odorant-binding proteins (OBPs), chemosensory proteins (CSPs), odorant receptors (ORs), and gustatory receptors (GRs). A functional OR comprises a heterodimer formed by a specific OR and an ubiquitous co-receptor, named Orco. The soluble and secreted OBP and CSP proteins are proposed to bind and transport chemical molecules to membrane receptors (Pelosi et al., 2014). ORs and GRs recognize specific ligands and send a signal to the brain, translating the chemical signal into an electrical signal. This chain reaction leads to the insect response (Jacquin-Joly and Merlin, 2004; Sánchez-Gracia et al., 2009; Bohbot and Pitts, 2015). In *R. prolixus*, the silencing of Orco using the RNA interference (RNAi) technique affects the physiology and behavior of insects by decreasing their ability to engorge on the blood, the ecdysis, and oviposition rates and by increasing the mortality rate (Franco et al., 2016).

Changes in olfactory specificity and/or sensitivity could give insects an opportunity to recognize new environmental signals. This adaptation may be driven by fast evolution *via* gene duplication/loss events of chemosensory genes (e.g., Guo and Kim, 2007; Vieira et al., 2007; Xiao et al., 2013; Engsontia et al., 2014; Suzuki et al., 2018) but also by polymorphism (e.g., Leary

TABLE 1 | *Rhodnius* samples.

Samples	Putative species	Sex	Origin	Tissues	No individuals per sample
PMEF	*R. prolixus*	M	CTA 077, Colombia, 1982	Antennae + Rostrum	18
PFEF	*R. prolixus*	F	CTA 077, Colombia, 1982	Antennae + Rostrum	18
PMES	*R. prolixus*	M	Swiss laboratory	Head	5
RFEF	*R. robustus*	F	CTA 085, Peru, 1982	Antennae + Rostrum	21
RMEF	*R. robustus*	M	CTA 085, Peru, 1982	Antennae + Rostrum	19
RFG1	*R. robustus*	F	Sylvatic, French Guyana	Head	2
RFG2	*R. robustus*	F	Sylvatic, French Guyana	Antennae + Rostrum	2
RMG	*R. robustus*	M	Sylvatic, French Guyana	Antennae + Rostrum	4

The first letter of each sample name indicates the species (P for prolixus and R for robustus); the second refers to sex (M for male and F for female), and the last two letters indicate insect origin: EF for the strains from Brazilian rearing (CTA, Colônias de Triatominae, Araraquara), ES for the strains from Swiss rearing, and G for sylvatic samples from French Guyana. All samples are from the adult stage. Raw data are available at European Nucleotide Archive (https://www.ebi.ac.uk/ena), Accession: SAMEA3900098 to SAMEA3900105.

et al., 2012; Brand et al., 2015; Wada-Katsumata et al., 2018; Ferreira et al., 2021) or a variation in gene expression (e.g., McBride et al., 2014; Purandare and Brisson, 2020). In a previous study on *Triatoma brasiliensis* (Marchant et al., 2016a), another Chagas disease vector, we showed that chemosensory genes, including *OBPs* and *CSPs*, were differentially expressed (DE) according to the populations. We also evidenced variations in the expression of *takeout* (*TO*) genes. TO proteins have been proposed to be associated with chemosensory perception in both olfactory and taste systems (Fujikawa et al., 2006) and have been involved in the circadian cycle linking the temporal cycle and feeding status information (Sarov-Blat et al., 2000; So et al., 2000; Meunier et al., 2007).

This study focuses on the chemosensory transcriptome of the two closely relative species *R. prolixus* and *R. robustus*. Wild *R. robustus* was collected in the field in French Guyana to ensure the integrity of its taxonomic status as this country seems to be outside the geographic distributional area of *R. prolixus* (Bérenger et al., 2009). Moreover, *R. robustus* Amazon lineage seems to be a monophyletic clade (Abad-Franch and Monteiro, 2007). Because there were a few sylvatic *R. prolixus* populations with uncertain status, only the rearing strains of *R. prolixus* obtained from domiciliary colonies were used. We also used the strain of *R. robustus* and tested it for its potential introgression with *R. prolixus*. To perform expression analysis, we built a reference transcriptome containing transcripts from all samples. The reads from each sample were then mapped against this reference to quantify gene expression. A comparative study was then carried out based on the differential expression study and on the presence/absence of chemosensory transcripts [OBP, CSP, OR, GR, and an ionotropic receptor (IR)] and also TO.

MATERIALS AND METHODS

Sampling, Sequencing, and RNA Extraction

Samples are given in **Table 1**. The first letter of each sample name indicates the species (P for *prolixus* and R for *robustus*); the second refers to sex (M for male and F for female), and the last two letters indicate the insect origin: EF for the strains from Brazilian rearing, ES for the strains from Swiss rearing, and G for the sylvatic samples from French Guyana. For *R. prolixus*, we used two reared strains, one from the Insetário de Triatominae da Faculdade de Ciências Farmacêuticas/Unesp/Araraquara, Brazil, including males and females (PMEF and PFEF) and another with males only (PMES) from the Swiss Biology Institute, University of Neuchâtel, tested by Marchant et al. (2016b) for the transcriptome assembly strategy. In this study, the reared samples of *R. robustus* (RFEF and RMEF) from the same Araraquara Brazilian laboratory in the same conditions like *R. prolixus* were included. The strain samples were fed on ducks once a month. Once insects reached the adult phase, they were blood-fed on ducks two times in the 1st week. The strain samples (males and females) were sacrificed 7 days after their last blood-meal and their head stored in RNA later® (Thermo Fisher Scientific, Waltham, MA, USA) before RNA extraction. Sylvatic *R. robustus* from French Guyana (RFG1, RFG2, and RMG) was collected using light traps and stored in RNA later® before RNA extraction.

The taxonomic status of all samples was checked using *cytb* sequences (data not shown) as the *cytb* of the reared *R. robustus* samples was not differentiated from that of the reared *R. prolixus*, we considered that the reared *R. robustus* was introgressed by *R. prolixus*. The sylvatic *R. robustus* samples (RMG1 and RFG3) were clustered in the *robustus* clade IV according to the definition of Monteiro et al. (2003). For each species, we first extracted RNA from entire heads to obtain a large panel of genes in the transcriptomes (PMES and RFG1 samples). The eyes were removed from the specimens prior to RNA extraction to discard the pigments that could interfere with the use of chemical reagents during the estimations of extraction and the RNA amount. Secondly, we extracted RNA from the isolated rostrums and antennae (PMEF, PFEF, RFEF, RMEF, RFG2, and RMG samples) to enrich the samples in chemosensory transcripts as some are known to be expressed at very low levels, especially chemosensory receptors. For these tissues, we grouped for each sex 2–4 individuals for the sylvatic *R. robustus* and to have some variabilities for the inbreeding strain samples, 18–21 individuals.

The extraction was carried out by TRIzol®. The samples were then sequenced using the Illumina paired-end technology by the IBENS genomics platform (Institute of Biology of the Ecole Normale Supérieure, Paris, France) according to the protocol described in Marchant et al. (2016b). Raw data are available at

the European Nucleotide Archive (https://www.ebi.ac.uk/ena), Accession: SAMEA3900098 to SAMEA3900105.

Transcriptome Assemblies

Transcriptomes from each sample were first assembled using the procedure developed by Marchant et al. (2016b) for the laboratory *R. prolixus* strains, which is clearly the most effective one compared to *de novo* strategies alone (i.e., Trinity software). The quality of individual assemblies was evaluated using several criteria, including contig lengths, N50, the total number of bases assembled, and contig number. To estimate the completeness of the transcriptome in terms of the expected gene content, we used the Benchmarking Universal Single-Copy Ortholog (BUSCO V3) assessment tool (Waterhouse et al., 2018).

A common reference transcriptome that was needed to compare the expression between samples was built using the Cuffmerge script in the Cufflinks package and contigs from all samples irrespective of the species/sex/condition. This transcriptome (RefT-PR-PA) was used for the presence-absence analysis. For differential expression analysis, a non-redundant transcriptome (RefT-PR-DE) was derived from the RefT-PR-PA transcriptome, selecting only the most expressed isoform per gene based on the indices of the fragments per kilobase per million mapped reads (FPKM).

Annotation of Transcripts of Interest

Transcripts of interest, namely chemosensory (OBPs, CSPs, ORs, and GRs) and TO transcripts, were annotated. To identify chemosensory transcripts, we searched the contigs assembled from the reads that are mapped onto the annotated genes in the *R. prolixus* genome (accessible at www.vectorbase.org). We completed the OBP annotation using OBP28 identified in Marchant et al. (2016b). For *TO* genes that were not yet annotated in the *R. prolixus* genome at the beginning of this study, transcriptome annotation was performed by running tblastn with an *e*-value threshold of 10^{-6} and using sets of queries of insect proteins from the family of interest retrieved from the GenBank database. The selected contigs were then blasted in a third validation step to the non-redundant protein database (Blastx with an *e*-value threshold of 10^{-6}). Only contigs for which the best blast hit was a protein from the searched family were kept.

Gene Phylogenies

For the CSP family, open reading frames in the identified transcripts were inspected for all the samples. We used *R. prolixus* (Mesquita et al., 2015) and *T. brasiliensis* (Marchant et al., 2016a) as reference sequences. Amino acid sequence alignments were performed using MAFFT (Katoh et al., 2019), and the peptide signals were deleted before tree reconstruction. In the final tree, only the isoforms detected in at least three sequences from the data set were retained to discard misassembled sequences.

The phylogenetic tree of Hemiptera was built using CSP sequences from phytophagous bugs: (1) three Aphidae species: *Myzus persicae* (Wang et al., 2019), *Aphis gossypii* (Gu et al., 2013), and *Acyrthosiphon pisum* (see Wang et al., 2019), (2) three Miridae plant bug species: *Adelphocoris suturalis* (Cui et al., 2017), *Apolygus lucorum* (in Wang et al., 2019), and *Adelphocoris lineolatus* (Sun et al., 2015), (3) three Delphacidae plant hopper species *Nilaparvata lugens* (in Wang et al., 2019), *Laodelphax striatellus* (Zhou et al., 2015), and *Sogatella furcifera* (Zhou et al., 2015), and (4) one Pentatomidae, *Nezara viridula* (Wu et al., 2019). For the Triatominae family, we used *T. brasiliensis* (Marchant et al., 2016a), *R. prolixus* (Mesquita et al., 2015), and only one sequence from *R. robustus* (RFG3). The Phthiraptera *Pediculus humanus* was also included because of its hematophagous diet (Mesquita et al., 2015). All CSP sequences are given in **Supplementary Table 1**.

Maximum likelihood trees were built using IQ-TREE (Nguyen et al., 2015), with the search of the best-fit model according to the Bayesian information criteria (BIC) in ModelFinder (Kalyaanamoorthy et al., 2017). Node support was assessed using SH-aLRT parallelized over the bootstrap samples to maximize load balance and efficiency. Phylogenetic trees were displayed using iTOL (Letunic and Bork, 2006).

Presence/Absence Analysis of Transcripts of Interest

We used FPKM information provided by Cuffdiff from Cufflink. Positive FPKM indicated that reads are mapped to the gene revealing gene expression while null FPKM reflected a lack of read mapping on the given gene revealing the absence of expression. We used two positive FPKM thresholds, 20 and 100, that defined three intervals ($0<FPKM<20$, $20<FPKM<100$, and $FPKM>100$) to evaluate gene expression.

Differential Expression Analysis

We performed a chemosensory gene expression study *via* RNA-seq (Trapnell et al., 2010). Read pairs from each sample were mapped against the RefT-PR-DE transcriptome using BWA with default options (Li and Durbin, 2009). Reads with multiple hits were excluded to avoid a counting bias. Count data were therefore produced using SAMtool facilities (Li et al., 2009). PMES and RFG1 samples were not used in this analysis because their RNAs were from whole heads and thus not comparable to the samples obtained from antennae and rostrum only. Furthermore, transcripts showing an expression limited to one environment were discarded from the data set before the differential expression study. Contigs with a low coverage were filtered using HTSfilter (Rau et al., 2013) using the DESeq normalization method and a threshold of five mapped reads on average. The sex status (male vs. female irrespective of their environment) and the environmental conditions (sylvatic vs. reared irrespective of their sex) were investigated as two independent factors in the model. The analyses were carried out by using DESeq2, turning off the independent filtering (Love et al., 2014). The Benjamini–Hochberg correction (Benjamini and Hochberg, 1995) was applied, and contigs were considered as DE when the *p*-adjusted value was <0.05. The contigs identified as DE were annotated by Blast on the non-redundant protein database (NR, version May 2015).

RESULTS

Transcriptome Assembly

Transcriptome assemblies showed N50, which comprises between 907 and 2,415 pb according to the recovery of samples

TABLE 2 | Summary of individual and merged transcriptome assemblies.

Assemblies	Number of paired reads	Number of contigs	Total length (bp)	Size distribution (bp)		BUSCO (%)		References
				Mean	N50	C	P	
R. prolixus lab PMES	31,125,782	81,459	88,125,904	1,082	1,656	92.25	95.98	Marchant et al., 2016b
R. robustus syl RFG1	20,372,071	93,238	124,846,519	1,339	1,847	92.03	95.76	This study
R. prolixus lab PFEF	42,792,624	148,600	137,733,087	927	1,170	86.39	92.39	Marchant et al., 2016b
R. prolixus lab PMEF	30,786,251	209,914	373,294,695	1,778	2,415	88.15	93.86	Marchant et al., 2016b
R. robustus lab RFEF	26,748,452	131,760	125,817,722	955	1,246	86.17	91.95	This study
R. robustus lab RMEF	34,455,341	205,727	328,748,578	1,598	2,091	87.42	93.64	This study
R. robustus syl RFG2	24,417,527	105,493	69,951,034	663	907	86.54	92.10	This study
R. robustus syl RMG	32,908,126	45,620	36,842,101	808	1,451	86.17	91.73	This study
RefT-PR-PA *prolixus-robustus*		184,271	483,000,356	2,621	4,318	93.55	98.79	This study
RefT-PR-DE *prolixus-robustus* one isoform per locus		136,988	338,277,756	2,469	3,885	71.77	89.92	This study

Complete BUSCO (%) (C) refers to the percentage of protein set for which complete sequence was found. Partial BUSCO (P) is the percentage of protein set for which at least 70% of the length of the sequence has been identified in the transcriptome.

with completeness from 84.68 to 91.53% of complete BUSCO sequences, going up to 96.77–98.39% when considering also partial sequences (**Table 2**).

The combined assembly of *R. prolixus* and *R. robustus* (RefT-RP-PA) comprised 184,271 transcripts and was used for annotation. Then, for each locus, only the most representative isoform (the sum of FPKM from all samples) was selected, reducing the transcriptome to 136,988 transcripts (RefT-PR-DE).

Annotation of Transcripts of Interest

As discussed in the previous publication of *R. prolixus* genome, Mesquita et al. (2015) annotated 19 *CSPs,* 27 *OBPs,* 106 *Ors*, 33 *Irs*, and 28 *Grs*, but some gene models were incomplete and/or annotated as pseudogenes. In the RefT-PR-PA transcriptome, all the 19 CSP transcripts were retrieved, but only 26 OBPs (lacking OBP2), 76 ORs, 27 IRs, and 26 GRs. The OBP28 annotated in Marchant et al. (2016b) was also retrieved in this transcriptome. Overall, 39 of the 213 genes annotated as chemosensory in the *R. prolixus* genome were not identified in our transcriptome. Their expression level may be lower than our detection threshold. They may also be expressed in other chemosensory organs such as legs or at different developmental stages or physiological conditions. However, their genomic annotation may also be erroneous as it remains speculative and no functional data are available for chemosensory genes in Triatominae. In addition to chemosensory genes, we annotated 23 TO transcripts (**Supplementary Table 2**). As 15 TOs were annotated from the antennal transcriptome of *R. prolixus* by Latorre-Estivalis et al. (2020), we used their nomenclature. From our data sets, we retrieved 12 TOs annotated by Latorre-Estivalis et al. (2020). TO1, TO4, and TO6 were absent from our data sets, but we described 12 new TOs. In total, both of the studies identified 27 TO transcripts.

In total, 199 transcripts were annotated and studied. All these transcripts were retrieved from the chemosensory samples (antennas + rostrum) but only 184 from the head samples, some ORs being not detected in this tissue. These results validated our strategy targeting chemosensory organs to enrich the transcriptome in chemosensory genes, especially chemosensory receptors, known to be expressed at very low levels.

CSP Phylogenies

Because all 19 CSPs annotated from the *R. prolixus* genome were found to be expressed in the transcriptomic samples, a phylogeny was performed for this protein family (**Figure 1**). The best model test for the reconstruction according to BIC was LG+G4. The CSP sequences were strongly conserved between the samples irrespective of the species/origin as no sequence variation in amino acids was observed for the most expressed CSPs. However, we found isoforms for the three CSPs (CSP4, CSP15, and CSP16). The CSP phylogeny confirms the cluster of the three CSPs, CSP2–4, found in the same *R. prolixus* genomic contig. Other clades were identified as CSP10, CSP11, CSP18, CSP19, and CSP12–15.

The phylogenetic CSP tree was reconstructed using 13 Hemiptera species, namely 3 Triatominae (*T. brasiliensis, R. prolixus*, and *R. robustus*), 3 Aphidae (*M. persicae*, *A. gossypii*, and *A. pisum*), 3 Miridae (*A. suturalis*, *A. lucorum*, and *A. lineolatus*), 3 Delphacidae (*N. lugens*, *L. striatellus*, and *S. furcifera*), and 1 Pentatomidae (*N. viridula*). The hematophagous Phthiraptera *P. humanus* was also included (**Figure 2**). It is noteworthy that four generalist CSP clades comprised at least 10 species attesting for the conserved orthologous sequences probably derived from an ancestral copy and suggesting that these CSPs may have a common function. By contrast, four other clades comprised three to five CSPs sequences for all species belonging to only one family. These clades represent CSP family extension, two for the Triatominae, and one for both the Miridae and the Delphacidae. No CSP extension for Aphididae was found with the used data. Moreover, one clade clustered only some CSPs from the hematophagous species, suggesting that these CSPs had probably a common function related to hematophagy.

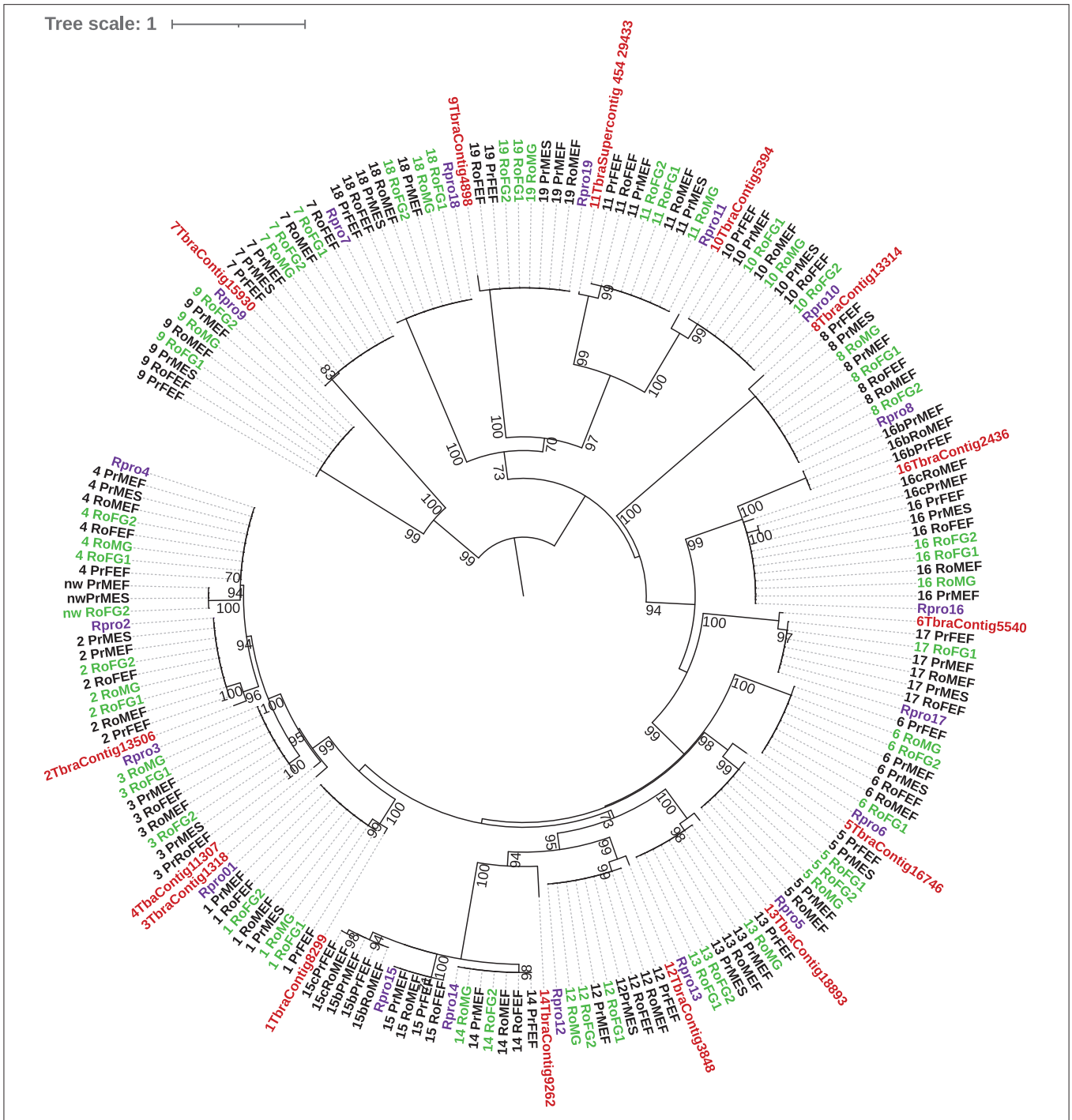

FIGURE 1 | Phylogenetic tree of chemosensory proteins (CSPs). Maximum likelihood tree in amino acids from the transcriptomic *Rhodnius robustus* and *Rhodnius prolixus* data sets (see **Table 1** for sample labeling). Sylvatic *R. robustus* is indicated in green. The predicted *R. prolixus* protein genes (Mesquita et al., 2015) are indicated in purple and *Triatoma brasiliensis* transcripts (Marchant et al., 2016a) are indicated in red. A bootstrap > 70% is indicated on the node.

Presence/Absence of Transcripts of Interest

We used the count data (FPKM) to estimate the presence/absence of transcripts in our data set (**Supplementary Table 3**). Among the 199 transcripts of interest, 16 were not expressed in the head samples (3 CSP, 1 OBP, 9 OR, 1 IR, 1GR, and 1 TO) compared to the antennas and rostrum samples. It was noted that 24 ORs were not expressed in the *R. prolixus* head samples (PMES) neither 18 in the sylvatic *R. robustus* 1 (RFG1), and 15 were only expressed in the former and 9 in the latter.

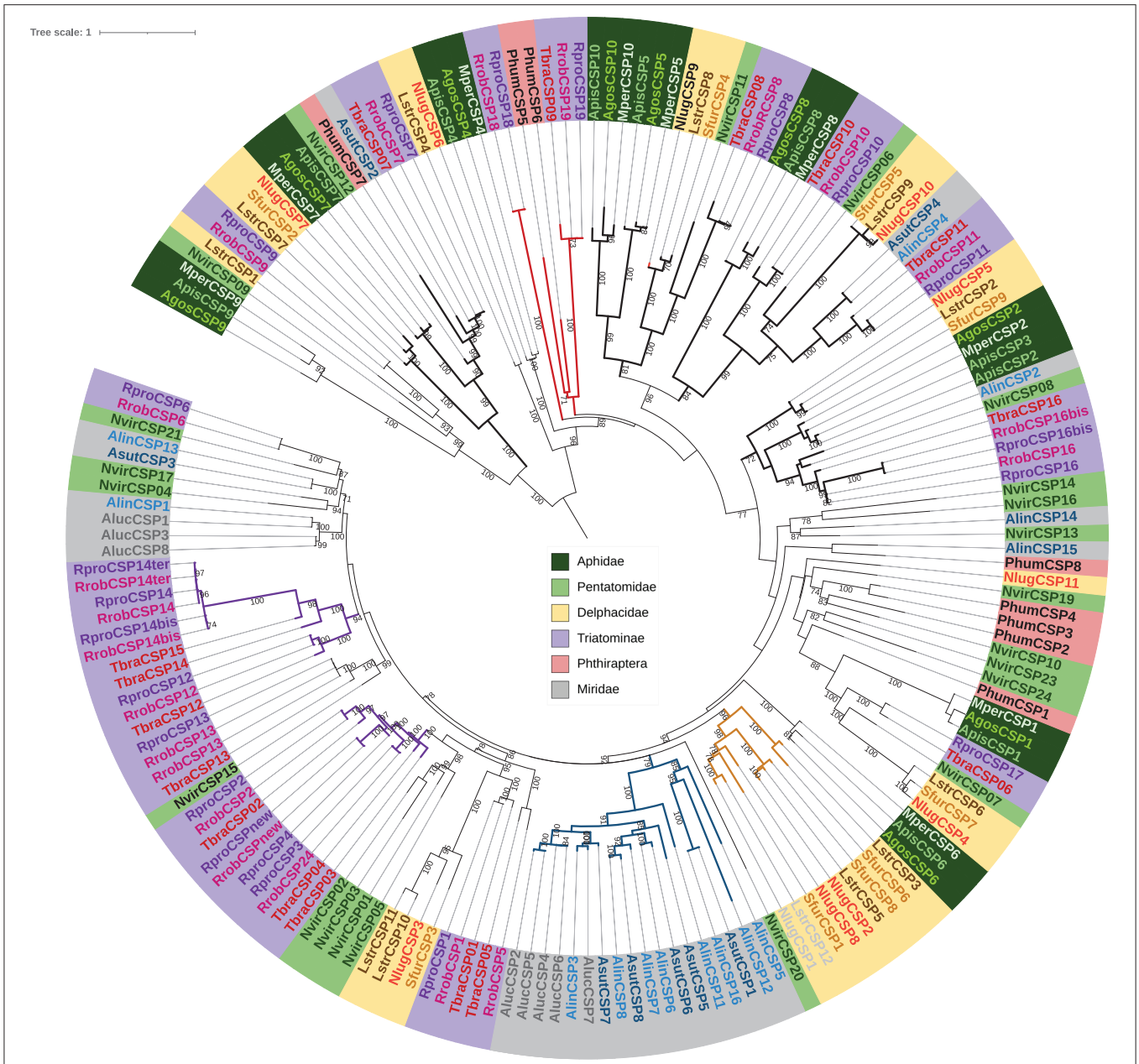

FIGURE 2 | Maximum likelihood phylogenetic tree of CSPs. Hemiptera species: Triatominae: *Rhodnius prolixus* (Rpro), *Rodnius robustus* (Rrob), *Triatoma brasiliensis* (Tbra), Aphidae *Myzus persicae* (Mper), *Aphis gossypii* (Agos), and *Acyrthosiphon pisum* (Apis); Miridae: *Adelphocoris suturalis* (Asut), *Apolygus lucorum* (Aluc), and *Adelphocoris lineolatus* (Alin); Delphacidae: *Nilaparvata lugens* (Nlug), *Laodelphax striatellus* (Lstr), and *Sogatella furcifera* (Sfur), and one Pentatomidae: *Nezara viridula* (Nvir); and Phthiraptera: *Pediculus humanus* (Phum). The four dark clades in bold indicate generalist CSP clades while the red clade indicate CSPs from the hematophagous species. The other colorated clades indicate CSP expansions, in purple for Triatominae, in blue for Miridae, in brown for Delphacidae. A bootstrap > 70% is indicated on the node. CSP sequences are given in **Supplementary Table 1**.

Considering the antennas and rostrum samples (**Figure 3**), 10 transcripts of interest were exclusively expressed according to sex and 19 to the environment (reared vs. sylvatic samples). Four transcripts seemed to be expressed specifically in females (OBP9, OR72, OR89, and IR75b) and six in males (OBP7, OBP18, OR20, OR21, GR24, and TO23). A comparison between environments (reared vs. sylvatic samples) highlights 1 transcript (OR22) specifically expressed in sylvatic samples and 14 specifically detected in the reared samples, two of them (OBP19 and GR16) having a FPKM higher than 20.

Considering the reared and sylvatic *R. robustus* samples, both of the samples lacked the three transcripts (CSP19, OBP9, and OR89) that are found to be only expressed in *R. prolixus* samples but shared the presence of three specific transcripts, OR60, OR75,

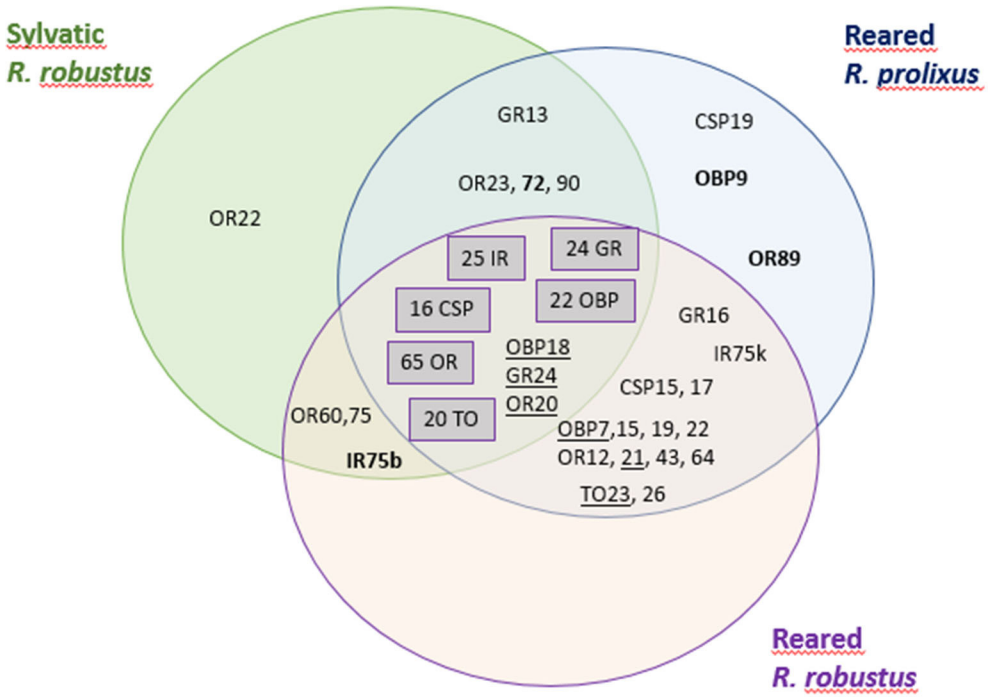

FIGURE 3 | Transcripts of interest [CSP, Odorant-binding protein (OBP), gustatory receptor (GR), odorant receptor (OR), ionotropic receptor (IR), Takeout (TO)] were detected according to the antennae and rostrum samples. Transcripts shared by all samples are indicated in gray boxes. Bold indicates the transcripts specifically expressed in females and underline indicates the transcripts specifically expressed in males.

and IR75b, the latter had a high FPKM (>100). It was noted that GR13 transcript was absent in the reared *R. robustus* while the GR16 in sylvatic *R. robustus*.

Expression data are also very useful to confirm gene identification and therefore annotation. Mesquita et al. (2015) reported that the genomic sequences of *Gr2, Or33, Or52*, and *Or106* diverged importantly from their counterparts from the same families, with evidence for an additional intron in the first exon. All four genes were expressed in our transcriptome and showed high RPKM values in our data (ranging from 22 for GR2 to 804 RPKM for OR106) validating their corresponding gene models (**Supplementary Table 4**).

DE Genes

Figure 4 shows the ACP made from count data normalized with the rlog function in DESeq2. The reared samples were closer with respect to each other irrespective of their sex and species, but the sylvatic *R. robustus* samples were much more dispatched between sexes. This high divergence of sylvatic samples could reflect a sampling bias as these samples were composed of only 2–4 individuals, while the reared samples consisted of a pool of about 20 highly inbred individuals per sample. Similarly, sylvatic samples are highly variable due to uncontrolled environmental conditions compared to the reared samples obtained from controlled rearing.

The performance of sample clustering according to the expression data of the 30 most DE contigs in the samples confirmed the observations provided by the ACP. In particular, both sylvatic *R. robustus* males and females were well-discriminated from other samples. Within the reared clade, sex was more discriminating than species condition (**Figure 5**).

When the sex condition was considered, numerous contigs (1,630) were found as DE between males and females, with a majority of them (83%) overexpressed in males (**Table 3**). When reared vs. sylvatic was considered irrespective of the sex, many more contigs (8,596) were found as DE with most (67%) overexpressed in sylvatic samples (**Table 3**).

Among the 199 transcripts of interest annotated in the reference transcriptome, 170 (85%) were sufficiently expressed to be analyzed using DESeq2 (14 CSPs, 22 OBPs, 66 ORs, 26 IRs, 20 GRs, and 22 TOs). Among them, 21 were DE (**Table 4**). When the sexes were compared, no DE transcript was identified. When the sylvatic and reared samples were compared, 17 transcripts of interest were DE and most overexpressed (81%) in the sylvatic *R. robustus* samples: 4 CSPs (CSP1–3 and CSP11), 1 OBP (OBP23), 5 ORs (OR10, OR14, OR22, OR33, and OR57), 4 IRs (IR41a, IR41c, IR75e, and IR101), 1 GR (GR27), and 2 TO (TO7 and TO12). Interestingly, in a vector control perspective, four annotated transcripts were overexpressed in the reared samples (OBP19, OR13, OR40, and OR79).

DISCUSSION

Because the insect chemosensory system plays a key role in the survival, reproduction, foraging, social communication of individuals, and ecological adaptation, we focused on

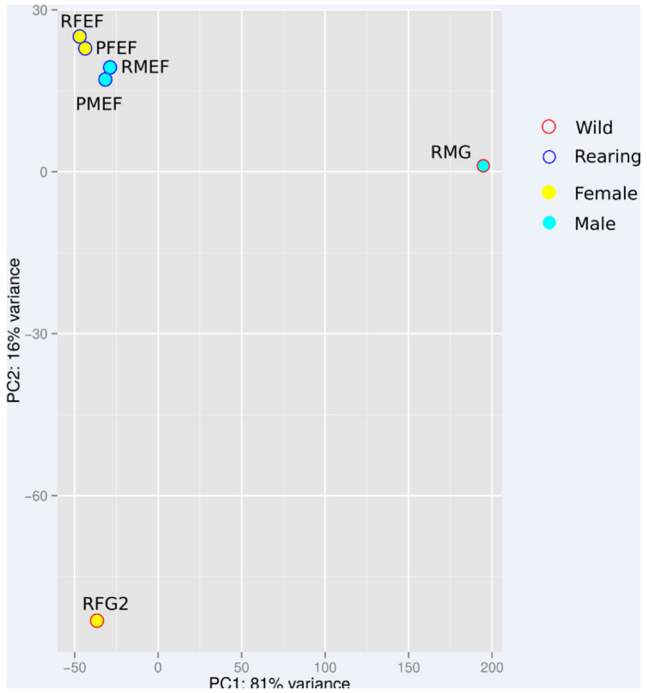

FIGURE 4 | ACP based on count data normalized by the function of rlog DESseq2. see **Table 1** for a sample description.

chemosensory transcriptomes of males and females of *R. prolixus* and *R. robustus* samples.

Divergence of Expression: Males vs. Females

In this study, a high number of transcripts (1,630) common to both *R. prolixus* and sylvatic *R. robustus* were identified as DE in the chemosensory transcriptome between the two sexes, with a majority of contigs (83%) overexpressed in males that probably reflects sexual dimorphism or is associated with specific sexual behaviors. Male chemosensory genes could be implicated in the perception of odors or pheronomes emitted by females. In *R. prolixus,* it was shown that odors from metasternal glands were more frequently emitted by females in greater quantities compared to males and may be involved in sexual communication (Pontes et al., 2008). Moreover, *R. prolixus* females release a pheromone that promotes flight initiation by males (Zacharias et al., 2010). However, most of the DE genes are poorly annotated and their precise function is unknown. Concerning the chemosensory transcripts, if none were DE, 10 were specifically expressed in one sex, OBP9, OR72, OR89, IR75b in females, OBP7, OBP18, OR20, OR21, GR24, and TO23 in males. One (IR75b) was expressed only in *R. robustus* females (reared and wild), two (OBP9 and OR89) only in *R. prolixus* females, and three (OBP7, OR21, and TO23) only in the reared male samples (*R. prolixus* and *R. robustus*). Similar expression profiles for antennal transcriptomes for males and females were notified by Latorre-Estivalis et al. (2017), whereas a significant differential expression for diverse chemosensory genes was observed between larvae and adults for *R. prolixus*. Using quantitative PCR analysis on *R. prolixus,* a significant expression in male antennae was observed for OBP26 and OBP27 (Oliveira et al., 2018) and also for OR1, OR3, OR74, and OR80 (Franco et al., 2018). Concerning TO transcripts, we evidenced that TO23 was exclusively expressed in reared males. For *R. prolixus*, TO2 was highlighted by Latorre-Estivalis et al. (2020) as significantly downregulated in male antennae are compared to those of larvae. It is, however, difficult to compare previous *R. prolixus* transcriptomic studies with our study because the data and the methods differed, namely antennae and rostrum in this study vs. antennas only in the other cited studies, and RPKM vs. quantitative PCR (qPCR).

In other insects, several studies have also shown significant differences in the expression of chemosensory genes between sexes, as for *Camponotus floridanus* and *Harpegnathos saltator* ant species (Zhou et al., 2012), the mosquito *Anopheles gambiae* (Pitts et al., 2011), or the moth *Sesamia nonagrioides* (Acín et al., 2009). For another Chagas disease vector, *T. brasiliensis* for which chemosensory genes were annotated (Marchant et al., 2015), in our previous study, we found three CSP (CSP14, CSP15, and CSP16) and one OBP (OBP4) overexpressed in males (Marchant et al., 2016a). Interestingly, the CSP16 clustered with CSPs of other Hemiptera and could suggest a common function related to sex, while the other two CSPs (CSP14 and CSP15) were in a CSP expansion specific to Triatominae. In *Drosophila, Gr32a* and *Gr68a* genes are expressed only in males and are necessary for effective courtship behavior (Bray and Amrein, 2003; Miyamoto and Amrein, 2008). *OBP99b* gene of *Drosophila* (also called *tsx*) was expressed in the adipose tissue of males but not in females while *OBP99a* gene was specifically expressed in females (Fujii, 2002). These two sex-specific *OBPs* genes are involved in the perception of pheromones and the modulation of reproductive behavior in *Drosophila*. In *Bombyx mori*, the *OR-1* was exclusively expressed in male antennae and was involved in the perception of the sex pheromone (Sakurai et al., 2004). In *Drosophila,* several TO proteins have also been identified as expressed exclusively in male heads and may be related to sex (Dauwalder et al., 2002).

Functional experiments are required to investigate if the transcripts evidenced in our study as exclusively or DE in one sex, could be involved in courtship behavior, mating control, and/or other sex-specific behaviors such as searching for oviposition sites in females. Functional studies of chemosensory genes in *R. prolixus* are still sparse. Oliveira et al. (2018) by silencing *R. prolixus OBP27* using RNAi and examining the sexual behavior of the phenotype suggested that this OBP is likely involved in the reception of sex pheromones. Another study by Franco et al. (2018) focused on four ORs using the *Xenopus* oocyte expression system, but none ORs responded to sex pheromones.

Divergence of Expression: *R. prolixus* vs. *R. robustus* and Environmental Factors

While *R. robustus* is considered sylvatic but sporadically visits houses, *R. prolixus* is widely domiciliated. As the species-specific status of *R. robustus* is debatable for populations of the Orinoco

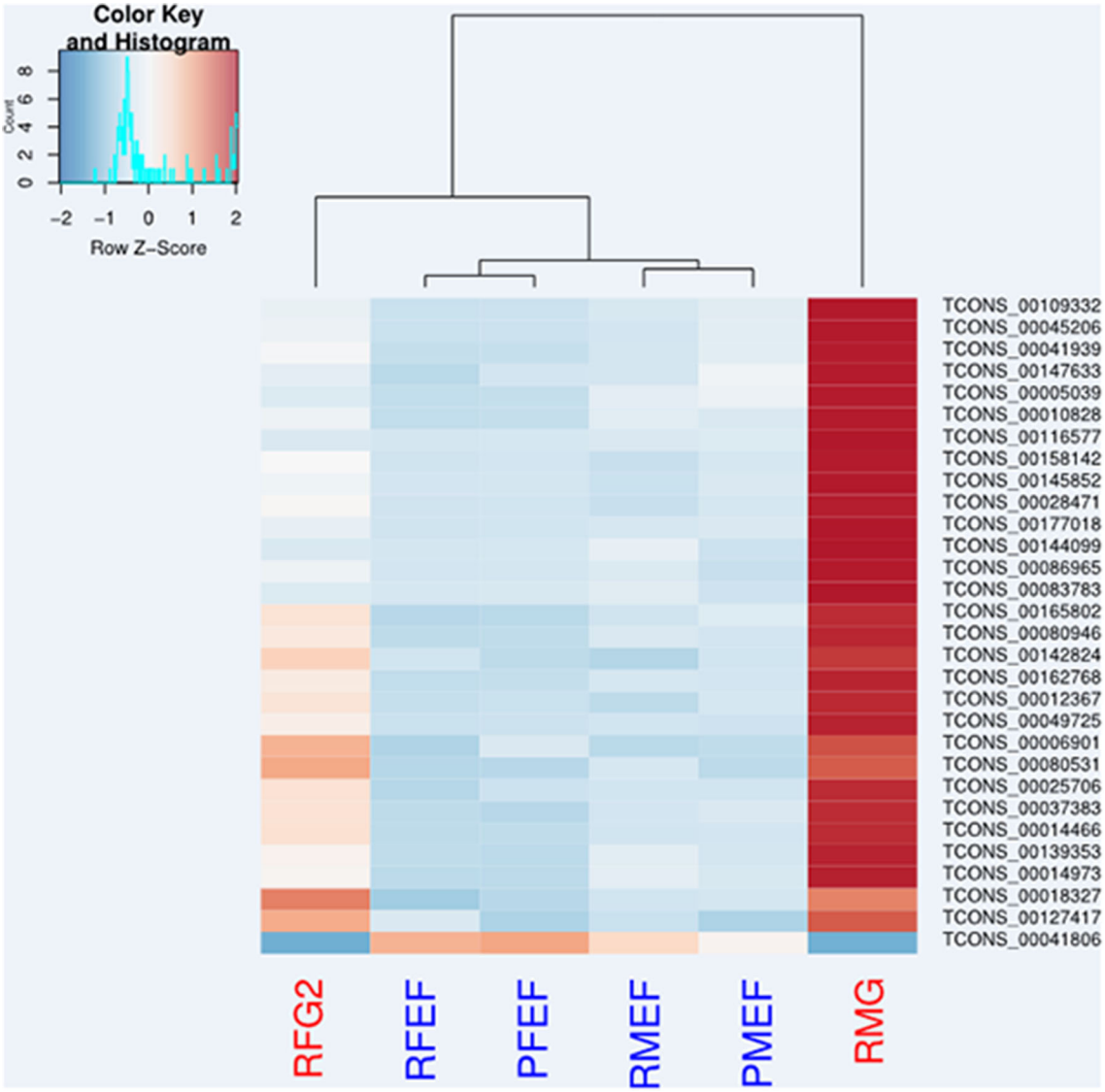

FIGURE 5 | Sample-to-sample distances. Heatmap performed using DESeq2 shows the Euclidean distances between sylvatic (red) and reared samples (blue) as calculated from the regularized log transformation expression data of the 30 most differentially expressed (DE) contigs between environment-species. see sample information (**Table 1**).

region (see Harry et al., 1992; Abad-Franch and Monteiro, 2007), we used for this study *R. robustus* from the Amazon region, i.e., French Guyana clearly differentiated from *R. prolixus* (Abad-Franch and Monteiro, 2007). For the laboratory strains used in this study, the introgression between *R. prolixus* and *R. robustus* has been evidenced at both mitochondrial and nuclear levels, a phenomenon also reported in other studies (Mesquita et al., 2015).

Among the chemosensory genes studied, 93.33% (182/199) were conserved between *R. prolixus* and the sylvatic *R. robustus* that support the close genetic relatedness between the two species. In the comparison between the sylvatic *R. robustus* and the reared *R. prolixus*, two factors overlap: the ecology of the vectors and their taxonomic status. However, the comparison with the reared *R. robustus* samples could be of interest. While introgressed with *R. prolixus*, the reared *R. robustus* samples did not express all the *R. prolixus* chemosensory genes like the wild *R. robustus*. Contrariwise, some genes were only

TABLE 3 | Summary of differential expression analysis evaluated using DESeq2 on two conditions: males (M) vs. females (F), and sylvatic/reared samples.

DE analysis	+	−	Total
M/F	1,358 (83.31%)	272	1,630
Sylvatic/reared	5,774 (67,38%)	2,822	8,596

"+" indicates contigs overexpressed in the first condition and "–" indicates contigs under expressed in the first condition.

TABLE 4 | Transcripts of interest detected as differentially expressed (DE) with DESeq2 in sylvatic vs. reared samples.

Transcripts	log2Fold change	Padj
CSP1	4,646	0,008
CSP2	5,386	0,009
CSP3	5,518	0,008
CSP11	5,788	0,006
OBP19	−4,391	0,039
OBP23	5,070	0,006
OR10	6,370	0,004
OR13	−5,165	0,011
OR14	4,164	0,029
OR22	6,243	0,001
OR33	4,730	0,003
OR40	−4,828	0,005
OR57	7,146	0,002
OR79	−3,801	0,045
IR41a	5,840	0,002
IR41c	4,319	0,018
IR75e	4,722	0,005
IR101	4,106	0,029
GR27	4,472	0,040
TO3	3,892	0,025
TO11	3,599	0,029

Negative Log2fold indicates transcripts under expressed in the sylvatic condition and positive Log2fold (gray highlight) indicates transcripts overexpressed in this condition.

expressed in *R. robustus* samples irrespective of their origin. This result could be related to species-specific expression of these transcripts in *R. prolixus* (i.e., CSP19, OBP9, and OR89) or in *R. robustus* (i.e., OR60, OR75, and IR75b). Moreover, the OR22 was only expressed in the sylvatic *R. robustus* samples, but not in the reared *R. robustus*; thus, this OR may be important for the adaptation to sylvatic environment. Similarly, the 12 chemosensory genes exclusively expressed in the reared samples (*R. prolixus* and *R. robustus*) could reflect an adaptation to the reared conditions. This ability of the reared *R. robustus* samples was potentially derived from some phenotypic plasticity or from their introgression by *R. prolixus*. In all, 8,596 contigs were DE between reared and sylvatic samples most being overexpressed (67%) in the sylvatic *R. robustus*, including 17 chemosensory genes. Only four chemosensory transcripts were overexpressed in the reared samples (OBP19, OR13, OR40, and OR79). The phenotypic plasticity of an organism can allow it to adapt to different environments without necessarily involving a genetic change but entailing gene expression modulation. Developmental and population genetic studies have shown that in some cases, up to 50% of the genome may have a variable expression depending on the environment (Snell-Rood et al., 2010). Various environmental factors could explain the differences between the two conditions. Some factors could not be controlled in our sampling plan. For instance, the feeding status of the wild insects at the time of catching may be highly variable. Furthermore, light traps capture moving individuals, a behavior not selected for the reared sample. The light trapping of *R. robustus* in the field was necessary because the sampling of *Rhodnius* in palm trees is difficult in a dense forest where trees are mostly inaccessible.

In addition to these uncontrolled environmental factors, the sylvatic/domiciliary status is likely to influence gene expression through host adaptation. Previous studies showed differences in the expression of chemosensory genes between insect species or populations adapted to different hosts. This is the case for the two moth species, one adapted to cotton (*Helicoverpa armigera*) and another to tobacco (*H. assulta*), which differ in the expression of OBP (Li et al., 2013). The differential expression analysis between the Kenyan population of *S. nonagrioides* sampled in sylvatic Poaceae plants (*Typha domingensis*) and the French population sampled in corn crops also revealed differences in the expression of an OBP transcript and one OR (Glaser et al., 2015). In a previous study on the Chagas disease vector, *T. brasiliensis* (Marchant et al., 2016a), we highlighted a significant overexpression of many chemosensory genes in the sylvatic samples of *T. brasiliensis* compared to the domiciliary ones. The phylogenetic reconstruction identified that OBP and CSP are the orthologs in *T. brasiliensis* and *R. prolixus*. Interestingly, one OBP (OBP23) and two CSPs (CSP1 and CSP11) were overexpressed in sylvatic samples from both *T. brasiliensis* and *R. robustus*. In the phylogenetic tree, the Triatominae CSP1 were in a specific clade, but the closest CSPs were from Delphacidae, including NlugCSP3. Binding characteristics studies of this *N. lugens* CSP and behavioral assays revealed that plant metabolites and flavoring agents (as non-adecane and 2-tridecanone) have high binding affinities in fluorescence competition-binding assays and displayed strong attractiveness to *N. lugens* (Waris et al., 2020). Concerning Triatominae CSP11, it is clustered with CSPs of the other Hemiptera in a generalist clade, including SfurCSP5. This *S. furcifera* CSP presented an antennae-enriched expression pattern, suggesting its potential olfactory function (Zhou et al., 2015). Using ligand-binding experiments *in vitro*, Chen et al. (2018) suggested that SfurCSP5 plays a key role in host odorant perception as SfurCSP5 bound several rice plant volatiles. Such a function in odorant perception could be therefore suggested for the Triatominae CSP1 and CSP11.

Apart from chemosensory genes, two contigs annotated as TO seemed to be over-expressed in the sylvatic samples compared to the reared/domiciliary ones. Similar results were obtained for *T. brasiliensis* (Marchant et al., 2016a). TO proteins are involved in the circadian cycle, food metabolism, and foraging behavior (Sarov-Blat et al., 2000; So et al., 2000), and in the regulation of locomotor activity during foraging activity (Meunier et al., 2007). In the case of Chagas disease vectors, the overexpression of TO transcripts in the sylvatic *R. robustus* samples compared to the reared ones may be related to dispersal behavior differences vs. aggregation. The implication of this gene family in behavioral changes in migratory *Locusta migratoria* was also noticed (Guo et al., 2011). These locusts adapt their behavior in response to population density changes, moving from a gregarious to a solitary migration phase and a *TO* gene LmigTo1 was overexpressed in solitary individuals. By contrast,

in *Danaus plexippus* (monarch butterfly of North America) an under-expression of TO protein in migrants compared to sedentary summer was shown (Zhu et al., 2008). This change in the expression of TO could be related to the sunshine duration variation used by migrants to help their navigation to their wintering areas.

Differences in expression between samples from different environments could be attributed to the adaptation of bloodsucking bugs to different hosts and thus utilizing different olfactory cues. In mosquitos, a comparison between anthropophilic *Aedes aegypti* and sylvatic *Aedes aegypti formosus* showed chemosensory changes, leading to a preference for human scent (McBride et al., 2014), and these changes were correlated with increased expression of the *AaegOr4* that binds a molecule presented at high levels in human scent.

CONCLUSIONS

Rhodnius prolixus and *R. robustus*, two closely related species vectors of Chagas disease, exhibited a large common panel of chemosensory genes. This study showed species-specific expression of some chemosensory genes and the overexpression of some other genes according to sex or environment (sylvatic/rearing). These differences in expression may induce differences in chemosensory sensitivity by a regulation at the perireceptor level, *via* the differential expression of OBPs and CSPs, supposed to transport odorants through the sensillum lymph to the chemosensory receptors, and at the receptor level, *via* the regulation of the OR and GR expression. Further electrophysiological studies comparing chemosensory sensitivity between sex and the insect origin may support this hypothesis.

Further studies are also needed at the genus scale, including different domiciliary and sylvatic *Rhodnius* species to link the differences in the chemosensory expression of chemosensory genes to environmental conditions that could reflect the adaptation of bloodsucking bugs to different environments and/or hosts, and thus utilizing different olfactory cues. Moreover, studies are required to functionally characterize the genes specifically or DE in the domiciliated species to identify the key cues involved in environmental preferences. Such types of research will open up promising perspectives in vector regulation based on the manipulation in chemosensory behavior.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Materials**.

AUTHOR CONTRIBUTIONS

AM and FM: data analysis and writing. EJ-J: study design and sample extractions. CA: study design. DB and J-MB: field collections. JR: strain collections. MH: study conception, data analysis, and writing. All authors edited the manuscript and approved the submitted version.

FUNDING

This study was funded by the French Agence Nationale de la Recherche (ADAPTANTHROP project, ANR-097-PEXT-009) and supported by the LabEx BASC (University Paris Saclay, France) and the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process number 2016/08176-9, 2017/50329-0). AM was funded by the Idex Paris Saclay, France. Financial support had a role neither in study design, collection, analysis, and interpretation of data nor in manuscript writing.

ACKNOWLEDGMENTS

We would like to thank Marie-Laure Martin-Magniette (IPS2, INRA Paris-Saclay), who helped us to develop the statistical strategy of the differential expression analysis. The authors are very grateful to the engineers of the bioinformatics platforms Genouest (Université de Rennes 1) and eBio (Université Paris Sud) for technical support. We also thank Malcolm Eden for a comprehensive review of the English text.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.725504/full#supplementary-material

Supplementary Table 1 | List of the chemosensory protein (CSP) sequences used for the phylogenetic trees of **Figures 1**, **2**. Hemiptera species: Triatominae: *Rhodnius prolixus* (Rpro), *Rhodnius robustus* (Rrob), *Triatoma brasiliensis* (Tbra), Aphidae *Myzus persicae* (Mper), *Aphis gossypii* (Agos), and *Acyrthosiphon pisum* (Apis); Miridae: *Adelphocoris suturalis* (Asut), *Apolygus lucorum* (Aluc), and *Adelphocoris lineolatus* (Alin); Delphacidae: *Nilaparvata lugens* (Nlug), *Laodelphax striatellus* (Lstr), and *Sogatella furcifera* (Sfur), and Pentatomidae: *Nezara viridula* (Nvir). Phthiraptera species: *Pediculus humanus* (Phum).

Supplementary Table 2 | List of contigs encoding putative Takeout proteins. Protein name, contig number, best BLAST hit and its ID on the non-redundant protein database, identity percentages (% ident), score and E-value are indicated.

Supplementary Table 3 | Number of transcripts of interest detected by the presence of mapped reads in the different samples. The number of transcripts is indicated in each box. The names of transcripts not detected are in parentheses.

Supplementary Table 4 | FPKM indices for each transcript of interest and each studied condition (Male/Female and Reared/Sylvatic). ID of the gene annotated in *Rhodnius prolixus* genome (https://www.vectorbase.org).

REFERENCES

Abad-Franch, F., and Monteiro, F. A. (2007). Biogeography and evolution of Amazonian triatomines (Heteroptera: Reduviidae): implications for Chagas disease surveillance in humid forest ecoregions. *Mem. Inst. Oswaldo Cruz* 102, 57–70. doi: 10.1590/S0074-02762007005000108

Acín, P., Carrascal, M., Abián, J., Guerrero, A., and Quero, C. (2009). Expression of differential antennal proteins in males and females of an important crop pest, *Sesamia nonagrioides*. *Insect Biochem. Mol. Biol.* 39, 11–19. doi: 10.1016/j.ibmb.2008.09.007

Aguilar, H. M., Abad-Franch, F., Dias, J. C. P., Junqueira, A. C. V., and Coura, J. R. (2007). Chagas disease in the Amazon Region. *Mem.*

Inst. Oswaldo Cruz 102, 47–56. doi: 10.1590/S0074-02762007005000098

Antonio-Campos, A., Cuatepotzo-Jiménez, V., Noguéz-García, J., Alejandre-Aguilar, R., and Rivas, N. (2019). Distribution of triatomine (Hemiptera: Reduviidae) vectors of Chagas disease in the state of Hidalgo, Mexico. *J. Vector Ecol.* 44, 179–186. doi: 10.1111/jvec.12342

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Benton, R. (2015). Multigene family evolution: perspectives from insect chemoreceptors. *Trends Ecol. Evol.* 30, 590–600. doi: 10.1016/j.tree.2015.07.009

Bérenger, J.-M., Pluot-Sigwalt, D., Pagès, F., Blanchet, D., and Aznar, C. (2009). The triatominae species of French Guiana (Heteroptera: Reduviidae). *Mem. Inst. Oswaldo Cruz* 104, 1111–1116. doi: 10.1590/S0074-02762009000800007

Bohbot, J. D., and Pitts, R. J. (2015). The narrowing olfactory landscape of insect odorant receptors. *Frontiers Ecol. Evol.* 3:39. doi: 10.3389/fevo.2015.00039

Brand, P., Ramírez, S. R., Leese, F., Quezada-Euan, J. J., Tollrian, R., and Eltz, T. (2015). Rapid evolution of chemosensory receptor genes in a pair of sibling species of orchid bees (Apidae: Euglossini). *BMC Evol. Biol.* 15:176. doi: 10.1186/s12862-015-0451-9

Bray, S., and Amrein, H. (2003). A putative drosophila pheromone receptor expressed in male-specific taste neurons is required for efficient courtship. *Neuron* 39, 1019–1029. doi: 10.1016/S0896-6273(03)00542-7

Cabrera, R., Vega, S., Valderrama, Y., Cabanillas, K., Fernandez, C., Rodriguez, O., et al. (2013). New focus of active transmission of Chagas disease in indigenous populations in the Peruvian Amazon basin. *Rev. Soc. Bras. Med. Trop.* 46, 367–372. doi: 10.1590/0037-8682-1195-2013

Castro, M. C., Barrett, T. V., Santos, W. S., Abad-Franch, F., and Rafael, J. A. (2010). Attraction of Chagas disease vectors (Triatominae) to artificial light sources in the canopy of primary Amazon rainforest. *Mem. Inst. Oswaldo Cruz* 105, 1061–1064. doi: 10.1590/S0074-02762010000800019

Chen, G. L., Pan, Y. F., Ma, Y. F., Wang, J., He, M., and He, P. (2018). Binding affinity characterization of an antennae-enriched chemosensory protein from the white-backed planthopper, *Sogatella furcifera* (Horvath), with host plant volatiles. *Pest. Biochem. Physiol.* 152, 1–7. doi: 10.1016/j.pestbp.2018.09.006

Coura, J. R., Junqueira, A. C. V., Fernandes, O., Valente, S. A. S., and Miles, M. A. (2002). Emerging Chagas disease in Amazonian Brazil. *Trends Parasitol.* 18, 171–176. doi: 10.1016/S1471-4922(01)02200-0

Cui, H. H., Gu, S. H., Zhu, X. Q., Wei, Y., Liu, H. W., Khalid, H. D., et al. (2017). Odorant-binding and chemosensory proteins identified in the antennal transcriptome of *Adelphocoris suturalis* Jakovlev. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 24, 139–145. doi: 10.1016/j.cbd.2016.03.001

da Rosa, J. A., Mendonça, V. J., Gardim, S., de Carvalho, D. B., de Oliveira, J., Nascimento, J. D., et al. (2014). Study of the external female genitalia of 14 *Rhodnius* species (Hemiptera, Reduviidae, Triatominae) using scanning electron microscopy. *Parasites Vectors* 7:17. doi: 10.1186/1756-3305-7-17

Dauwalder, B., Tsujimoto, S., Moss, J., and Mattox, W. (2002). The *Drosophila* takeout gene is regulated by the somatic sex-determination pathway and affects male courtship behavior. *Genes Dev.* 16, 2879–2892. doi: 10.1101/gad.1010302

Engsontia, P., Sangket, U., Chotigeat, W., and Satasook, C. (2014). Molecular evolution of the odorant and gustatory receptor genes in lepidopteran insects: implications for their adaptation and speciation. *J. Mol. Evol.* 79, 21–39. doi: 10.1007/s00239-014-9633-0

Feliciangeli, M. D., Dujardin, J.-P., Bastrenta, B., Mazzarri, M., Villegas, J., Flores, M., et al. (2002). Is *Rhodnius robustus* (Hemiptera: Reduviidae) responsible for Chagas disease transmission in western Venezuela? *Trop. Med. Int. Health* 7, 280–287. doi: 10.1046/j.1365-3156.2002.00853.x

Feliciangeli, M. D., Sanchez-Martin, M., Marrero, R., Davies, C., and Dujardin, J.-P. (2007). Morphometric evidence for a possible role of *Rhodnius prolixus* from palm trees in house re-infestation in the State of Barinas (Venezuela). *Acta Trop.* 101, 169–177. doi: 10.1016/j.actatropica.2006.12.010

Ferreira, E. A., Lambert, S., Verrier, T., Marion-Poll, F., and Yassin, A. (2021). Soft selective sweep on chemosensory genes correlates with ancestral preference for toxic noni in a specialist drosophila population. *Genes* 12:32. doi: 10.3390/genes12010032

Fitzpatrick, S., Feliciangeli, M. D., Sanchez-Martin, M. J., Monteiro, F. A., and Miles, M. A. (2008). Molecular genetics reveal that silvatic *Rhodnius prolixus* do colonise rural houses. *PLoS Negl. Trop. Dis.* 2:e210. doi: 10.1371/journal.pntd.0000210

Franco, T. A., Oliveira, D. S., Moreira, M. F., Leal, W. S., and Melo, A. C. (2016). Silencing the odorant receptor co-receptor *RproOrco* affects the physiology and behavior of the Chagas disease vector *Rhodnius prolixus*. *Insect Biochem. Mol. Biol.* 69, 82–90. doi: 10.1016/j.ibmb.2015.02.012

Franco, T. A., Xu, P., Brito, N. F., Oliveira, D. S., Wen, X., Moreira, M. F., et al. (2018). Reverse chemical ecology-based approach leading to the accidental discovery of repellents for *Rhodnius prolixus*, a vector of Chagas diseases refractory to DEET. *Insect Biochem. Mol. Biol.* 103, 46–52. doi: 10.1016/j.ibmb.2018.10.004

Fujii, S. (2002). Genes expressed in the *Drosophila* head reveal a role for fat cells in sex-specific physiology. *EMBO J.* 21, 5353–5363. doi: 10.1093/emboj/cdf556

Fujikawa, K., Seno, K., and Ozaki, M. (2006). A novel Takeout-like protein expressed in the taste and olfactory organs of the blowfly, *Phormia regina*. *FEBS J.* 273, 4311–4321. doi: 10.1111/j.1742-4658.2006.05422.x

Glaser, N., Gallot, A., Legeai, F., Harry, M., Kaiser, L., Le Ru, B., et al. (2015). Differential expression of the chemosensory transcriptome in two populations of the stemborer *Sesamia nonagrioides*. *Insect Biochem. Mol. Biol.* 65, 28–34. doi: 10.1016/j.ibmb.2015.07.008

Gu, S. H., Wu, K. M., Guo, Y. Y., Field, L. M., Pickett, J. A., Zhang, Y. J., et al. (2013). Identification and expression profiling of odorant binding proteins and chemosensory proteins between two wingless morphs and a winged morph of the cotton aphid *Aphis gossypii* Glover. *PLoS ONE* 8:e73524. doi: 10.1371/journal.pone.0073524

Guhl, F., Pinto, N., and Aguilera, G. (2009). Sylvatic triatominae: a new challenge in vector control transmission. *Mem. Inst. Oswaldo Cruz* 104, 71–75. doi: 10.1590/S0074-02762009000900012

Guo, S., and Kim, J. (2007). Molecular evolution of *Drosophila* odorant receptor genes. *Mol. Biol. Evol.* 24, 1198–1207. doi: 10.1093/molbev/msm038

Guo, W., Wang, X., Ma, Z., Xue, L., Han, J., Yu, D., et al. (2011). *CSP* and *Takeout* genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet* 7:e1001291. doi: 10.1371/journal.pgen.1001291

Harry, M. (1993a). Isozymic data question the specific status of some blood-sucking bugs of the genus *Rhodnius*, vectors of Chagas disease. *Trans. R. Soc. Trop. Med. Hyg.* 87:492. doi: 10.1016/0035-9203(93)90054-T

Harry, M. (1993b). Use of the median process of the pygophore in the identification of *Rhodnius nasutus, R. neglectus, R. prolixus* and *R. robustus* (Hemiptera: Reduviidae). *Ann. Trop. Med. Parasitol.* 87, 277–282. doi: 10.1080/00034983.1993.11812767

Harry, M., Galindez, I., and Cariou, M. L. (1992). Isozyme variability and differentiation between *Rhodnius prolixus, R.robustus* and *R.pictipes*, vectors of Chagas disease in Venezuela. *Med. Vet. Entomol.* 6, 37–43. doi: 10.1111/j.1365-2915.1992.tb00032.x

Hashimoto, K., and Schofield, C. J. (2012). Elimination of *Rhodnius prolixus* in Central America. *Parasit Vectors* 5:45. doi: 10.1186/1756-3305-5-45

Jacquin-Joly, E., and Merlin, C. (2004). Insect olfactory receptors: contributions of molecular biology to chemical ecology. *J. Chem. Ecol.* 30, 2359–2397. doi: 10.1007/s10886-004-7941-3

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings Bioinformatics* 20, 1160–1166. doi: 10.1093/bib/bbx108

Latorre-Estivalis, J. M., Robertson, H. M., Walden, K. K., Ruiz, J., Gonçalves, L. O., Guarneri, A. A., et al. (2017). The molecular sensory machinery of a Chagas disease vector: expression changes through imaginal moult and sexually dimorphic features. *Sci. Rep.* 7, 1–16. doi: 10.1038/srep40049

Latorre-Estivalis, J. M., Sterkel, M., Ons, S., and Lorenzo, M. G. (2020). Transcriptomics supports local sensory regulation in the antenna of the kissing-bug *Rhodnius prolixus*. *BMC Genomics* 21, 1–14. doi: 10.1186/s12864-020-6514-3

Leary, G. P., Allen, J. E., Bunger, P. L., Luginbill, J. B., Linn, C. E., Macallister, I. E., et al. (2012). Single mutation to a sex pheromone receptor provides adaptive specificity between closely related moth species. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14081–14086. doi: 10.1073/pnas.1204661109

Lent, H., and Wygodzinsky, P. (1979). Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. *Bull. Am. Mus. Nat. Hist*. 163, 123–520.

Letunic, I., and Bork, P. (2006). Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23, 127–128. doi: 10.1093/bioinformatics/btl529

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, H., Zhang, H., Guan, R., and Miao, X. (2013). Identification of differential expression genes associated with host selection and adaptation between two sibling insect species by transcriptional profile analysis. *BMC Genomics* 14:582. doi: 10.1186/1471-2164-14-582

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15, 1–21. doi: 10.1186/s13059-014-0550-8

Marchant, A., Mougel, F., Almeida, C., Jacquin-Joly, E., Costa, J., and Harry, M. (2015). De novo transcriptome assembly for a non-model species, the blood-sucking bug *Triatoma brasiliensis*, a vector of Chagas disease. *Genetica* 143, 225–239. doi: 10.1007/s10709-014-9790-5

Marchant, A., Mougel, F., Jacquin-Joly, E., Costa, J., Almeida, C. E., and Harry, M. (2016a). Under-expression of chemosensory genes in domiciliary bugs of the Chagas disease vector *Triatoma brasiliensis*. *PLoS Negl. Trop. Dis*. 10:e0005067. doi: 10.1371/journal.pntd.0005067

Marchant, A., Mougel, F., Mendonça, V., Quartier, M., Jacquin-Joly, E., da Rosa, A., et al. (2016b). Comparing *de novo* and reference-based transcriptome assembly strategies by applying them to the blood-sucking bug *Rhodnius prolixus*. *Insect Biochem. Mol. Biol*. 69, 25–33. doi: 10.1016/j.ibmb.2015.05.009

Márquez, E. J., and Saldamando-Benjumea, C. I. (2013). *Rhodnius prolixus* and *Rhodnius robustus*-like (Hemiptera, Reduviidae) wing asymmetry under controlled conditions of population density and feeding frequency. *J. Biosci*. 38, 549–560. doi: 10.1007/s12038-013-9332-9

Matias, A., De la Riva, J. X., Torrez, M., and Dujardin, J.-P. (2001). *Rhodnius robustus* in Bolivia identified by its wings. *Mem. Inst. Oswaldo Cruz* 96, 947–950. doi: 10.1590/S0074-02762001000700010

McBride, C. S., Baier, F., Omondi, A. B., Spitzer, S. A., Lutomiah, J., Sang, R., et al. (2014). Evolution of mosquito preference for humans linked to an odorant receptor. *Nature* 515, 222–U151. doi: 10.1038/nature13964

Mesquita, R. D., Vionette-Amaral, R. J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F. A., Minx, P., et al. (2015). Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc. Natl. Acad. Sci. U.S.A*. 112, 14936–14941. doi: 10.1073/pnas.1506226112

Meunier, N., Belgacem, Y. H., and Martin, J.-R. (2007). Regulation of feeding behaviour and locomotor activity by takeout in Drosophila. *J. Exp. Biol*. 210, 1424–1434. doi: 10.1242/jeb.02755

Miles, M. A., Arias, J. R., and Souza, A. (1983). Chagas' disease in the Amazon Basin: V. Periurban palms as habitats of *Rhodnius robustus* and *Rhodnius pictipes*-Triatomine vectors of Chagas' disease. *Mem. Inst. Oswaldo Cruz* 78, 391–398. doi: 10.1590/S0074-02761983000400002

Miyamoto, T., and Amrein, H. (2008). Suppression of male courtship by a *Drosophila* pheromone receptor. *Nat. Neurosci*. 11, 874–876. doi: 10.1038/nn.2161

Monteiro, F. A., Barrett, T. V., Fitzpatrick, S., Cordon-Rosales, C., Feliciangeli, D., and Beard, C. B. (2003). Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R. robustus*. *Mol. Ecol*. 12, 997–1006. doi: 10.1046/j.1365-294X.2003.01802.x

Monteiro, W. M., Magalhães, L. K., Santana Filho, F. S., Borborema, M., Silveira, H., Barbosa, M., et al. (2010). *Trypanosoma cruzi* TcIII/Z3 genotype as agent of an outbreak of Chagas disease in the Brazilian Western Amazonia. Trop. *Med. Int. Health* 15, 1049–1051. doi: 10.1111/j.1365-3156.2010.02577.x

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol*. 32, 268–274. doi: 10.1093/molbev/msu300

Oliveira, D. S., Brito, N. F., Franco, T. A., Moreira, M. F., Leal, W. S., and Melo, A. C. (2018). Functional characterization of odorant binding protein 27 (RproOBP27) from *Rhodnius prolixus* antennae. *Front. Physiol*. 9:1175. doi: 10.3389/fphys.2018.01175

PAHO/WHO (2020). *Chagas Disease*. Available online at: https://www.paho.org/en/topics/chagas-disease

Pelosi, P., Iovinella, I., Felicioli, A., and Dani, F. R. (2014). Soluble proteins of chemical communication: an overview across arthropods. *Front. Physiol*. 5:320. doi: 10.3389/fphys.2014.00320

Pitts, R. J., Rinker, D. C., Jones, P. L., Rokas, A., and Zwiebel, L. J. (2011). Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. *BMC Genomics* 12:271. doi: 10.1186/1471-2164-12-271

Pontes, G. B., Bohman, B., Unelius, C. R., and Lorenzo, M. G. (2008). Metasternal gland volatiles and sexual communication in the triatomine bug, *Rhodnius prolixus*. *J. Chem. Ecol*. 34, 450–457. doi: 10.1007/s10886-008-9431-5

Purandare, S. R., and Brisson, J. A. (2020). Divergent chemosensory gene expression accompanies ecological specialisation of pea aphid morphs. *Ecol. Entomol*. 45, 364–368. doi: 10.1111/een.12803

Rau, A., Gallopin, M., Celeux, G., and Jaffrézic, F. (2013). Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics* 29, 2146–2152. doi: 10.1093/bioinformatics/btt350

Rendón, L. M., Guhl, F., Cordovez, J. M., Erazo, D., Rendón, L. M., Guhl, F., et al. (2015). New scenarios of *Trypanosoma cruzi* transmission in the Orinoco region of Colombia. Mem. Inst. *Oswaldo Cruz* 110, 283–288. doi: 10.1590/0074-02760140403

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., et al. (2004). Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc. Natl. Acad. Sci. U.S.A*. 101, 16653–16658. doi: 10.1073/pnas.0407596101

Sánchez-Gracia, A., Vieira, F. G., and Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103, 208–216. doi: 10.1038/hdy.2009.55

Sanchez-Martin, M. J., Feliciangeli, M. D., Campbell-Lendrum, D., and Davies, C. R. (2006). Could the Chagas disease elimination programme in Venezuela be compromised by reinvasion of houses by sylvatic *Rhodnius prolixus* bug populations? *Trop. Med. Int. Health* 11, 1585–1593. doi: 10.1111/j.1365-3156.2006.01717.x

Sarov-Blat, L., So, W. V., Liu, L., and Rosbash, M. (2000). The Drosophila takeout gene is a novel molecular link between circadian rhythms and feeding behavior. *Cell* 101, 647–656. doi: 10.1016/S0092-8674(00)80876-4

Snell-Rood, E. C., Van Dyken, J. D., Cruickshank, T., Wade, M. J., and Moczek, A. P. (2010). Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays* 32, 71–81. doi: 10.1002/bies.200900132

So, W. V., Sarov-Blat, L., Kotarski, C. K., McDonald, M. J., Allada, R., and Rosbash, M. (2000). *Takeout*, a novel *Drosophila* gene under circadian clock transcriptional regulation. *Mol. Cell. Biol*. 20, 6935–6944. doi: 10.1128/MCB.20.18.6935-6944.2000

Sun, L., Zhou, J. J., Gu, S. H., Xiao, H. J., Guo, Y. Y., Liu, Z. W., et al. (2015). Chemosensillum immunolocalization and ligand specificity of chemosensory proteins in the alfalfa plant bug *Adelphocoris lineolatus* (Goeze). *Sci. Rep*. 5, 1–13. doi: 10.1038/srep08073

Suzuki, H. C., Ozaki, K., Makino, T., Uchiyama, H., Yajima, S., and Kawata, M. (2018). Evolution of gustatory receptor gene family provides insights into adaptation to diverse host plants in nymphalid butterflies. *Genome Biol. Evol*. 10, 1351–1362. doi: 10.1093/gbe/evy093

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol*. 28, 511–515. doi: 10.1038/nbt.1621

Valente, S. A., da, S., Valente, V. da, C., and Fraiha Neto, H. (1999). Considerations on the epidemiology and transmission of Chagas disease in the Brazilian amazon. *Mem. Inst. Oswaldo Cruz* 94, 395–398. doi: 10.1590/S0074-02761999000700077

Vieira, F. G., Sánchez-Gracia, A., and Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes:

purifying selection and birth-and-death evolution. *Genome Biol.* 8:R235. doi: 10.1186/gb-2007-8-11-r235

Wada-Katsumata, A., Robertson, H. M., Silverman, J., and Schal, C. (2018). Changes in the peripheral chemosensory system drive adaptive shifts in food preferences in insects. *Front. Cell. Neurosci.* 12:281. doi: 10.3389/fncel.2018.00281

Wang, Q., Zhou, J. J., Liu, J. T., Huang, G. Z., Xu, W. Y., Zhang, Q., et al. (2019). Integrative transcriptomic and genomic analysis of odorant binding proteins and chemosensory proteins in aphids. *Insect Mol. Biol.* 28, 1–22. doi: 10.1111/imb.12513

Waris, M. I., Younas, A., Ameen, A., Rasool, F., and Wang, M. Q. (2020). Expression profiles and biochemical analysis of chemosensory protein 3 from *Nilaparvata lugens* (Hemiptera: Delphacidae). *J. Chem. Ecol.* 46, 363–377. doi: 10.1007/s10886-020-01166-6

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., et al. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35, 543–548. doi: 10.1093/molbev/msx319

Wu, Z. Z., Cui, Y., Qu, M. Q., Lin, J. H., Chen, M. S., Bin, S. Y., et al. (2019). Candidate genes coding for odorant binding proteins and chemosensory proteins identified from dissected antennae and mouthparts of the southern green stink bug *Nezara viridula*. *Comp. Bioch. Physiol. Part D Genomics Proteomics* 31:100594. doi: 10.1016/j.cbd.2019.100594

Xiao, J. H., Yue, Z., Jia, L. Y., Yang, X. H., Niu, L. H., Wang, Z., et al. (2013). Obligate mutualism within a host drives the extreme specialization of a fig wasp genome. *Genome Biol.* 14, 1–18. doi: 10.1186/gb-2013-14-12-r141

Zacharias, C. A., Pontes, G. B., Lorenzo, M. G., and Manrique, G. (2010). Flight initiation by male *Rhodnius prolixus* is promoted by female odors. *J. Chem. Ecol.* 36, 449–451. doi: 10.1007/s10886-010-9779-1

Zhao, Y., Galvão, C., and Cai, W. (2021). *Rhodnius micki*, a new species of Triatominae (Hemiptera, Reduviidae) from Bolivia. *ZooKeys* 1012:71. doi: 10.3897/zookeys.1012.54779

Zhou, W., Yuan, X., Qian, P., Cheng, J., Zhang, C., Gurr, G., et al. (2015). Identification and expression profiling of putative chemosensory protein genes in two rice planthoppers, *Laodelphax striatellus* (Fallén) and *Sogatella furcifera* (Horváth). *J. Asia Pacific Entomol.* 18, 771–778. doi: 10.1016/j.aspen.2015.09.006

Zhou, X., Slone, J. D., Rokas, A., Berger, S. L., Liebig, J., Ray, A., et al. (2012). Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoS Genet* 8:e1002930. doi: 10.1371/journal.pgen.1002930

Zhu, H., Casselman, A., and Reppert, S. M. (2008). Chasing Migration Genes: A brain expressed sequence tag resource for summer and migratory monarch butterflies (*Danaus plexippus*). *PLoS ONE* 3:e1345. doi: 10.1371/journal.pone.0001345

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Ecology and Evolution

ORIGINAL RESEARCH
published: 12 January 2022
doi: 10.3389/fevo.2021.750317

Phylogenomics for Chagas Disease Vectors of the *Rhodnius* Genus (Hemiptera, Triatominae): What We Learn From Mito-Nuclear Conflicts and Recommendations

Jonathan Filée[1†], Marie Merle[1†], Héloïse Bastide[1], Florence Mougel[1], Jean-Michel Bérenger[2,3], Elaine Folly-Ramos[4], Carlos Eduardo Almeida[5] and Myriam Harry[1]*

[1] CNRS, IRD, UMR EGCE, Evolution, Génomes, Comportement et Ecologie, Université Paris-Saclay, Gif-sur-Yvette, France, [2] Département Systématique and Evolution, Muséum National d'Histoire Naturelle, Paris, France, [3] IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection, Aix Marseille Université, Marseille, France, [4] Departamento de Engenharia e Meio Ambiente – DEMA, Universidade Federal da Paraíba, João Pessoa, Brazil, [5] Instituto de Biologia, Universidade Federal da Bahia (UFBA), Salvador, Brazil

OPEN ACCESS

Edited by:
Kaio Cesar Chaboli Alevi, São Paulo State University, Brazil

Reviewed by:
Raquel Lima Cordon, University of Vermont, United States
Tiago Belintani, State University of Campinas, Brazil

***Correspondence:**
Myriam Harry
myriam.harry@egce.cnrs-gif.fr;
myriam.harry@u-psud.fr

[†]These authors share first authorship

Specialty section:
This article was submitted to Phylogenetics, Phylogenomics, and Systematics, a section of the journal Frontiers in Ecology and Evolution

Received: *30 July 2021*
Accepted: *29 November 2021*
Published: *12 January 2022*

Citation:
Filée J, Merle M, Bastide H, Mougel F, Bérenger J-M, Folly-Ramos E, Almeida CE and Harry M (2022) Phylogenomics for Chagas Disease Vectors of the Rhodnius Genus (Hemiptera, Triatominae): What We Learn From Mito-Nuclear Conflicts and Recommendations. Front. Ecol. Evol. 9:750317. doi: 10.3389/fevo.2021.750317

We provide in this study a very large DNA dataset on *Rhodnius* species including 36 samples representing 16 valid species of the three *Rhodnius* groups*, pictipes, prolixus* and *pallescens.* Samples were sequenced at low-depth with whole-genome shotgun sequencing (Illumina technology). Using phylogenomics including 15 mitochondrial genes (13.3 kb), partial nuclear rDNA (5.2 kb) and 51 nuclear protein-coding genes (36.3 kb), we resolve sticking points in the *Rhodnius* phylogeny. At the species level, we confirmed the species-specific status of *R. montenegrensis* and *R. marabaensis* and we agree with the synonymy of *R. taquarussuensis* with *R. neglectus.* We also invite to revisit the species-specific status of *R. milesi* that is more likely *R. nasutus*. We proposed to define a *robustus* species complex that comprises the four close relative species: *R. marabaensis, R. montenegrensis, R. prolixus* and *R. robustus.* As *Psammolestes tertius* was included in the *Rhodnius* clade, we strongly recommend reclassifying this species as *R. tertius*. At the *Rhodnius* group level, molecular data consistently supports the clustering of the *pictipes* and *pallescens* groups, more related to each other than they are to the *prolixus* group. Moreover, comparing mitochondrial and nuclear tree topologies, our results demonstrated that various introgression events occurred in all the three *Rhodnius* groups, in laboratory strains but also in wild specimens. We demonstrated that introgressions occurred frequently in the *prolixus* group, involving the related species of the *robustus* complex but also the pairwise *R. nasutus* and *R. neglectus*. A genome wide analysis highlighted an introgression event in the *pictipes* group between *R. stali* and *R. brethesi* and suggested a complex gene flow between the three species of the *pallescens* group, *R. colombiensis, R. pallescens* and *R. ecuadoriensis*. The molecular data supports also a sylvatic distribution

of *R. prolixus* in Brazil (Pará state) and the monophyly of *R. robustus*. As we detected extensive introgression events and selective pressure on mitochondrial genes, we strongly recommend performing separate mitochondrial and nuclear phylogenies and to take advantages of mito-nuclear conflicts in order to have a comprehensive evolutionary vision of this genus.

Keywords: phylogenomics, mito-nuclear conflict, introgression, selective pressure, *Rhodnius*, Chagas disease vectors

INTRODUCTION

The blood-sucking bugs (Hemiptera, Reduviidae, Triatominae) are vectors of the parasite *Trypanosoma cruzi* (Chagas, 1909) (Kinetoplastea, Trypanosomatida). Most species of live in natural habitats where they feed on a large variety of vertebrates. But some of them can be found in human dwellings responsible for the Chagas disease endemic to Latin America where about 6 million people are infected (PAHO, 2020). The Triatominae subfamily is distributed in five tribes, the most diverse among them being the Rhodniini and Triatomini, which represent more than 90% of the known species. The Rhodniini tribe is composed of two *Rhodnius* and *Psammolestes* genera, which show some morphological differentiations (Lent and Wygodzinsky, 1979) that has justified the erecting of two different genera. They also differ in their ecology, the *Rhodnius* species mainly associated with palm trees and the *Psammolestes* species with bird nests (Galvão and Justi, 2015). Whereas only three *Psammolestes* species were described, there are twenty-two *Rhodnius* species but one species, *R. taquarussuensis* (Rosa et al., 2017a), has been recently synonymized with *R. neglectus* (Nascimento et al., 2019) (see **Table 1** with authors and year of description). The *Rhodnius* genus is divided into three major groups, *pictipes, prolixus* and *pallescens* (Hernández et al., 2020). The three species of the *pallescens* group (*R. colombiensis, R. ecuadoriensis,* and *R. pallescens*) are *trans*-Andean, distributed in the west of the Andes. The seven species of the *pictipes* group (*R. amazonicus, R. brethesi, R. micki, R. paraensis, R. pictipes, R. stali,* and *R. zeledoni*) are *cis*-Andean, distributed in the east of the Andes and in the Amazon region. The same distribution is observed for ten species of the *prolixus* groups (*R. barretti, R. dalessandroi, R. domesticus, R. marabaensis, R. milesi, R. montenegrensis, R. nasutus, R. neglectus, R. prolixus,* and *R. robustus*), the eleventh *R. neivai* having also some *trans*-Andean populations (Zhao et al., 2021).

Some species are difficult to differentiate from each other on morphological criteria thus generating taxonomic conflicts. Pioneer studies using allozymes had questioned the species-specific status of closely related species such as Venezuelan populations of *R. robustus* and *R. prolixus* (Harry et al., 1992; Harry, 1993). Concerning some newly described species, molecular data are sparse for *R. barretti* (Abad-Franch et al., 2013), not released in a database for *R. marabaensis* (Souza et al., 2016) or lacking for *R. micki* (Zhao et al., 2021). Recently, molecular studies on a newly described species *R. montenegrensis* (Rosa et al., 2012) suggested that it is not a true species but a part of the *R. robustus* variability (Abad-Franch et al., 2013; Monteiro et al., 2018). However, the two species *R. montenegrensis* and *R. marabaensis* were recognized by Castro et al. (2020) based on the geographical origin of their samples. Moreover, some authors consider *R. milesi* to be a *R. neglectus* variant from south-eastern Amazonia (Abad-Franch et al., 2013). Otherwise, only a few specimens have been collected for *R. paraensis* first described as *R. domesticus* (Sherlock et al., 1977) and *R. amazonicus* close to *R. pictipes* but this latter species was revalidated by Bérenger and Pluot-Sigwalt (2002). Finally, the species-specific status for two species is questionable and difficult to deepen since *R. zeledoni* reported as morphologically close to *R. paraensis* was described from only one dead and dried male specimen (Jurberg et al., 2009), and *R. dalessandroi* is actually known only from its published description (Carcavallo and Barreto, 1976).

Some *Rhodnius* species are important vectors of *T. cruzi*, as *R. prolixus,* the main Chagas disease vector within the genus, which extends from Central America to the Andean countries and the Amazon basin. Control programs in Central America were thought to have achieved the elimination of this vector but it was reported again in Mexico in 2019 (Antonio-Campos et al., 2019). Three other species are domiciliated in some countries, namely *R. ecuadoriensis* in the northern zone of Peru and Ecuador, *R. stali* in Bolivia, and *R. pallescens* in Panama. In Brazil, *R. neglectus* and *R. nasutus* often colonize human environments (Galvão and Justi, 2015).

Phylogenetic studies performed on the Rhodniini tribe, mainly with mitochondrial genes, showed that *Psammolestes* was nested within the *Rhodnius* clade but discrepancies remained about its phylogenetic position within the *prolixus* group (reviewed in Hernández et al., 2020). Moreover, several molecular studies have also focused on the relationship between the three groups. Two tree topologies were obtained either *pictipes* and *pallescens*, or *pictipes* and *prolixus* groups as sister taxa, unrelated to the type of marker used, mitochondrial, nuclear, or both (reviewed in Hernández et al., 2020). Recently, the third configuration, *prolixus* and *pallescens* groups as sister taxa, was also reported by Kieran et al. (2021) using ultraconserved elements.

The principal sources of discrepancies for phylogenies may come from methodological pitfalls and/or reflect the biological reality. Methodological bias could therefore be induced by incomplete/different taxon samplings but also by the type and the number of markers used. However, discordance between mitochondrial and nuclear phylogenies are expected because the mitochondrial genome is haploid uniparentally inherited in most animals unlike the nuclear one and has faster rates of evolution. Incongruences between single gene phylogenies can be

TABLE 1 | Distribution and habitat of the species of the Rhodniini tribe and molecular species-specific status inferred from this study.

Species of the Rhodniini tribe	Distribution	Habitat	Molecular species-specific status
R. amazonicus (Almeida et al., 1973)	Brazil, French Guiana	Sylvatic	Valid
R. barretti (Abad-Franch et al., 2013)	Colombia, Ecuador	Sylvatic	Not studied
R. brethesi (Matta, 1919)	Brazil (Pará, Amazonas), Colombia, Venezuela	Sylvatic Sporadically in domiciles Amazon basin	Valid
R. colombiensis (Mejia et al., 1999)	Colombia	Sylvatic, Sporadically in domiciles	Valid
R. dalessandroi (Carcavallo and Barreto, 1976)	Colombia	Sylvatic (1 site)	Dubious, no data
R. domesticus (Neiva and Pinto, 1923)	Brazil (atlantic forest)	Sylvatic Sporadically in domicile	Valid
R. ecuadoriensis (Lent and Leon, 1958)	South Colombia, Ecuador, North Peru	Sylvatic, Domiciliar Principal vector in Ecuador	Valid
R. marabaensis (Souza et al., 2016)	Maraba, Para, Brazil Pará	Sylvatic	Valid
R. micki (Zhao et al., 2021)	Bolivia (Santa Cruz, Saavedra) 2 males	No data	Not studied
R. milesi (Carcavallo et al., 2001)	Brazil Pará	Sylvatic	*R. nasutus (?)*
R. montenegrensis (Da Rosa et al., 2012)	Brazil (Acre, Rondonia)	Sylvatic	Valid
R. nasutus (Stal, 1959)	Brazil (caatinga)	Sylvatic In domiciliation process	Valid
R. neglectus (Lent, 1954)	Brazil (Cerrado São Paulo)	Silvatic In domiciliation process	Valid
R. neivai (Lent, 1953)	Colombia, Venezuela	Sylvatic (some domiciliary focus)	Valid
R. pallescens (Lent, 1953)	Belize, Colombia, Costa Rica, Nicaragua, Panama, Venezuela	Silvatic, Domiciliar Principal vector in Panama	Valid
R. paraensis (Sherlock et al., 1977)	Brazil, French Guiana	Sylvatic	Not studied
R. pictipes (Stal, 1972)	Belize, Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Peru, Suriname, Trinidad, Venezuela	Sylvatic Sporadically in domiciles	Valid
R. prolixus (Stal, 1859)	Central America, Colombia, (Ecuador) (Suriname), Venezuela	Domiciliar Principal vector in Central America	Valid
R. robustus (Larousse, 1927)	Brazil, Colombia, Ecuador, French Guiana, Peru, Venezuela	Sylvatic, Ssporadically in domiciles (Colombia, Guyana)	Valid
R. stali (Lent et al., 1993)	Bolivia, Brazil (Matto Grosso, Acre)	Sylvatic, In domiciliation process in Bolivia	Valid
R. taquarussuensis (syn R. neglectus) (Da Rosa et al., 2017)	Mato Grosso do Sul, Brazil	In rural dwelling in the city of Taquarussu	*R. neglectus*
R. zeledoni (Jurberg et al., 2009)	Brazil	Aracajú-Sergipe, 1 specimen	Dubious, no data
Psammolestes arthuri (Pinto and Lent, 1935)	Colombia, Venezuela	Sylvatic	Not studied
Psammolestes coreodes (Bergroth, 1911)	Argentina, Bolivia, Brazil, Paraguay	Sylvatic	Not studied
Psammolestes tertius (Lent and Jurberg, 1965)	Brazil	Sylvatic	*R. tertius*

In blue: prolixus group species, in pale green: pictipes group species, in darker green: pallescens group species.

due to incomplete lineage sorting, deep coalescence, horizontal gene transfer, introgression, hybridization, hidden paralogy or lack of phylogenetic information (Toews and Brelsford, 2012; Campillo et al., 2019). For the Rhodniini tribe, only the complete mitogenome of *R. pictipes* (Zhao et al., 2019) and the complete nuclear genome of *R. prolixus* (Mesquita et al., 2015) are available. For all the phylogenies yet performed, the datasets were indeed limited in terms of the number of genes. So far, the most complete studies for the Rhodniini tribe were performed by Justi et al. (2014) using ten species and 1–4 genes depending on the species, Kieran et al. (2021) using from eight to sixteen species and ultraconserved elements and/or rDNA data, and Paula et al. (2021) using thirteen species and three genes.

In this study, we aimed to resolve major phylogenetic conflicts within the *Rhodnius* genus using phylogenomics from low-depth whole-genome shotgun sequencing and to explore mito-nuclear conflicts in order to have a better understanding of the evolution of this genus.

Data were obtained by Illumina technology for 36 samples of the Rhodniini tribe including 17 putative species identified using morphological characters. We performed both mitochondrial and nuclear phylogenies using a set 15 mitochondrial genes (13 protein-coding mitochondrial plus 2 rDNA genes) and two sets of nuclear markers (nu-rDNA genes and 51 protein-coding genes). The protein-coding datasets were tested for selection.

The outcome of this work led us to identified 16 valid *Rhodnius* species in our molecular dataset and to formulate recommendations for both taxonomic and phylogenetic issues.

MATERIALS AND METHODS

Material and Genomic Data

For this study, field specimens and laboratory strains were used with a special focus on the *prolixus* group for field specimens (**Table 2**). From our field collection, were selected specimens with doubts about their morphological identification (MILEP, NEGP), difficult to determine (INCP, NasG), or exhibited incoherent *cytb* sequences in a preliminary study (ROBB, ROBR, ROBQ). Some specimens were very old since they were also used by Harry (1993) and Harry (1994). We used laboratory strains from the Insetário de Triatominae da Faculdade de Ciências Farmacêuticas/Unesp/Araraquara, Brazil. Samples were either stored in absolute EtOH or air-dried.

DNA was extracted from legs and alary muscles from adult bugs, from 1 for field specimens to 6 for laboratory strains using the Qiagen DNEasy tissue kit. Genomic data was obtained from 36 samples of the Rhodniini tribe corresponding to 17 species identified using morphological characters. A whole-genome shotgun sequencing was performed using Illumina technology (100 bp paired-end, Imagif platform, Gif-sur-Yvette, France) generating from 3.7 to 21.1 Gb data per sample corresponding to a genome depth from 5 to 30x (**Table 3**).

Sequences have been deposited in the NCBI BioProject database *via* the accession number PRJNA429761.

Mitogenomes and Nuclear rDNA Operon Assembly and Annotation

In order to assemble the mitogenomes, the reads were subsampled (500 kb of reads for each sample) and assembled using the Trinity software with default parameters (Haas et al., 2013). This approach allows to assembly sequences that are highly abundant in the sequence pool like mitochondrial or ribosomal genes. In insect, mtDNA represents on average 0.42% of the total DNA in genome sequence project (Meng et al., 2019). As we used 500 kb of data/sample for the mitochondrial assembly which in average contains $(500 \times 0.42)/100 = 2{,}100$ kb of mtDNA and given that the mitochondrial genome is about 15 kb, we obtain coverage of 140x for this genome. This assembly strategy also prevent to pick up some NUMS that are mitochondrial pseudogenes in nuclear chromosomes, because if these elements were present in the data, as they are nuclear, their coverage would be very low (from 5 to 30x according to the samples) compared to that of the (true) mtDNA and with a low probability to be retrieved from the data used to assemble the mtDNA. Contigs corresponding to the mitogenomes and the nuclear rDNA operon were identified by BLAST using as templates the *Triatoma infestans* mitogenome and the *R. prolixus* reference genome rDNA. In most cases, the mitogenome and the nuclear rDNA operon of each sample appeared as a single contig, sometimes as two or three contigs.

The annotation of the *Rhodnius* mitogenomes was achieved using the MITOS pipeline (Bernt et al., 2013). Each mitochondrial and nuclear rDNA datasets was independently aligned with MAFFT[1] (Katoh et al., 2019), visualized with Bioedit (version 7.2.6.1; Hall, 1999), and ambiguously aligned regions or gapped positions were manually corrected or removed. Two concatenated datasets were generated for the mitochondrial data. The first included the 15 mitochondrial genes (15 mtG), namely the 13 mtPCG (*Nad2, COI, COII, Atp8, Atp6, COIII, Nad3, Nad5, Nad4, Nad4L, Nad6, Cytb* and *Nad1*) and the 2 rDNA genes (*16 rDNA* and *12 rDNA*). The second dataset including only the 13 mitochondrial protein-coding genes (mtPCG) was used for the selection signature analyses. This dataset was translated in order to test for the presence of pseudogenes (**Supplementary Table 1**).

For the nuclear rDNA genes (nu-rDNA), we retrieved from the data the complete *18S* and *28S* genes but the partial *ITS1, 5.8S,* and *ITS2* genes. The alignment for this dataset is given in the **Supplementary Table 1**.

Genome Assembly and Protein-Coding Gene Annotation

A second assembly was carried out for each of the 36 genomes with the SOAPdenovo2 software (Luo et al., 2012) with k-mers estimated using the Kmergenie program (Chikhi and Medvedev, 2014). Benchmarking Universal Single-Copy Orthologs (BUSCO) genes (Waterhouse et al., 2019) were searched in assemblies using the insecta_odb10 dataset comprising 1,367 genes universal to all insects. In a first step, we kept the 88 nuclear protein-coding genes (nu-PCGs) for which an entire or fragmented copy was found in each of the 36 Rhodniini genome assemblies and in the *T. brasiliensis* transcriptome (Marchant et al., 2015), as this species was used as an outgroup. Each set of genes was aligned with MAFFT (global homology option; Katoh et al., 2019) and then concatenated. The concatenated alignment was visualized using Bioedit (version 7.2.6.1; Hall, 1999). All regions without at least 29 aligned genomes were retrieved as well as gaps or misaligned sequences. This drastic final trimming resulted in 51 nu-PCGs with a sequence coverage of at least 80% except for *R. brethesi* (**Table 3**). In order to have *R. prolixus* samples validated without (or low) introgression, we also search the 51 nu-PCGs in both the *R. prolixus* used by Mesquita et al. (2015) for the reference genome (VectorBase: RproC3.4) and a Honduras strain reared in a French laboratory (**Table 1**). The

[1]https://mafft.cbrc.jp/alignment/server/

TABLE 2 | Specimens used in the study.

Specimens	Morphologicaldetermination	Moleculardetermination	Origin	Field/Strain	Storage inFrench Lab
PSAM	*P. tertius*	*R. tertius*	Bahia, Curaça, Brazil	F (sylvatic)	EtOH (2003)
Ama1A	*R. amazonicus*		Belizon, French Guyana	F (sylvatic)	air dried (2014)
BRE25WB	*R. brethesi*		Amazonia, Brazil	F (sylvatic)	EtOH (2003)
R9WX	*R. brethesi*		Igarapé Tucunaré, Brazil	S (CTA 222, 2009)	EtOH (2013)
R10B	*R. colombiensis*		Tolima, Colombia	S (CTA 050, 2001)	EtOH (2013)
DomC	*R. domesticus*		Santa Catarina, Brazil	F (sylvatic)	EtOH (2006)
ECUD	*R. ecuadorensis*		PD, Colombia	F (domiciliar)	EtOH (2003)
ECUE	*R. ecuadorensis*		Peru	S	EtOH (2006)
MILE	*R. milesi*	*R. nasutus*	Brazil	S (sylvatic)	EtOH (2006)
MILEP	*R. milesi*	*R. prolixus*	Pará, Brazil	F (sylvatic)	EtOH (2003)
R8F	*R. montenegrensis*		Monte Negro, Rondônia, Brazil	S (CTA 087, 2003)	EtOH (2013)
INCP	*R.* sp.	*R. neglectus*	Coronel José Dias, Piauí, Brazil	F (sylvatic)	EtOH (2003)
NASP	*R. nasutus*		Coronel José Dias, Piauí, Brazil	F (sylvatic)	EtOH (2003)
R7WU	*R. nasutus*		Brazil	S (CTA 054, 1999)	EtOH (2013)
NasG	*R. nasutus*		Piauí, Brazil	F (sylvatic)	EtOH (2003)
R5WU	*R. neglectus*		Frutal, Minas Gerais, Brazil	S (CTA 061, 1983)	EtOH (2013)
NEII	*R. neivai*		Maracay, Venezuela	S (sylvatic)	EtOH (2006)
R1J	*R. pallescens*		Colombia	S (CTA s/R)	EtOH (2013)
PIC34WC	*R. pictipes*		Belém, Brazil	F (sylvatic)	EtOH (2003)
PIC3L	*R. pictipes*		Pará, Brazil	F (sylvatic)	EtOH (2003)
PIC2WA	*R. pictipes*		Guyane	F (sylvatic)	air dried (2014)
R6WV	*R. pictipes*		Belém, PA, Brazil	S (CTA 072, 1998)	EtOH (2003)
R63K	*R. pictipes*		Belém, Pará, Brazil	S (CTA 072, 1998)	EtOH (2003)
Pro10O	*R. prolixus*		Guarico, Venezuela	S (domiciliar)	air dried (1990)
ProYRP	*R. prolixus*		Cojedes, Venezuela	S (domiciliar)	air dried (1990)
ProN	*R. prolixus*		Colombia	S (CTA 080, 1976)	EtOH (2013)
ProM	*R. prolixus*		Colombia	S (CTA 077, 1982)	EtOH (2013)
NEGP	*R. robustus*	*R. prolixus*	Coronel José Dias, Pará, Brazil	F (sylvatique)	EtOH (2003)
RobR	*R. robustus*	*R. prolixus*	Lima, Peru	S (CTA 085, 1982)	EtOH (2006)
Rob5s	*R. robustus*		Guyane	F (sylvatic)	EtOH (2014)
ROBB	*R. robustus*	*R. marabaensis*	Maraba, Brazil	F (sylvatic)	EtOH (2003)
RobQ	*R. robustus*	*R. prolixus*	Pará, Brazil	F (sylvatic)	EtOH (2003)
R4H	*R. robustus*	*R. prolixus*	Brazil	S (sylvatic)	EtOH (2013)
V-StaWZ	*R. stali*		Bolivia	S (sylvatic)	EtOH (2006)
U-StaWY	*R. stali*		Alto Beni, Bolivia	S (sylvatic)	EtOH (2006)
R3WT	*R. taquarussuensis*	*R. neglectus*	Taquarussu, Mato-Grosso do Sul, Brazil	S (CTA, 2011)	EtOH (2013)

Morphological determination prior the study and molecular determination posterior the study are indicated.
Specimens are either from the field, F or from laboratory strain, S, and the original habitat (sylvatic/domiciliary) is indicated in the brackets.
For strains, the reference and the year of the strain foundation are given, with CTA: Colônias de Triatominae, Araraquara, Brazil.
The storage conditions in the French laboratory is indicated, with EtOH: absolute ethanol.

corresponding BUSCO genes and the alignment are given in the **Supplementary Table 1**.

Diversity Parameters

In order to compare the variation rate between the datasets, the nucleotide diversity (Pi) was estimated for each dataset using DnaSP 6 (Rozas et al., 2017). Pairwise distances were calculated between samples using MegaX (Kumar et al., 2018) using the Maximum Composite Likelihood model and all the substitutions. The similarity between pairwise samples was calculated in percentage using the ratio of the total number of conserved sites and the total number of both conserved and variable.

Phylogenetic Analysis

All the individual mitochondrial genes and the three concatenated datasets namely the 15 mtG, the nu-rDNA, and the

TABLE 3 | Genomic data.

Samples	Data (Gb)	Genome depth (x)	mtG (bp)	mt sequence coverage (%)	nu-rDNA (pb)	nu-rDNA sequence coverage (%)	nu-PCGs (bp)	nu-PCG sequence coverage (%)
PSAM	9.3	13.29	13,300	1.00	5,189	1.00	33,793	0.93
Ama1A	5.9	8.43	13,300	1.00	5,189	1.00	34,434	0.94
BRE25WB	8.4	12.00	13,300	1.00	5,189	1.00	34,604	0.95
R9WX	7.7	11.00	13,300	1.00	5,189	1.00	28,108	0.77
R10B	12.2	17.43	12,301	0,92	5,189	1.00	34,137	0.94
DomC	16.7	23.86	13,300	1.00	5,189	1.00	35,078	0.96
ECUD	19	27.14	13,300	1.00	5,189	1.00	34,421	0.94
ECUE	6.6	9.43	13,300	1.00	5,189	1.00	30,504	0.84
MILE	11.2	16.00	13,245	1.00	5,189	1.00	35,692	0.98
MILEP	7.1	10.14	13,300	1.00	5,189	1.00	33,338	0.91
R8F	12.4	17.71	13,300	1.00	5,189	1.00	33,051	0.91
INCP	3.7	5.29	13,300	1.00	5,189	1.00	26,246	0.72
NasG	10.7	15.29	13,214	0.99	5,189	1.00	36,444	1.00
R7WU	11.4	16.29	12,203	0,92	5,189	1.00	32,565	0.89
NASP	4.1	5.86	13,300	1.00	5,189	1.00	29,385	0.81
R5WU	16.1	23.00	13,266	1.00	5,189	1.00	35,098	0.96
NEII	9.2	13.14	13,300	1.00	5,189	1.00	35,976	0.99
R1J	12.7	18.14	13,300	1.00	5,189	1.00	35,085	0.96
PIC34WC	9.5	13.57	13,300	1.00	5,189	1.00	33,932	0.93
PIC3L	19	27.14	13,300	1.00	5,189	1.00	31,363	0.86
PIC2WA	8.9	12.71	13,300	1.00	5,189	1.00	33,049	0.91
R6WV	12.6	18.00	13,300	1.00	5,189	1.00	33,95	0.93
R63K	12.6	18.00	13,300	1.00	5,189	1.00	34,032	0.93
Pro10O	8	11.43	13,300	1.00	5,189	1.00	36,135	0.99
ProYRP	7.3	10.43	13,300	1.00	5,189	1.00	33,794	0.93
ProN	7.6	10.86	13,300	1.00	5,189	1.00	35,202	0.97
ProM	12.7	18.14	13,300	1.00	5,189	1.00	35,437	0.97
NEGP	5.5	7.86	13,300	1.00	5,189	1.00	32,121	0.88
RobR	7.9	11.29	13,300	1.00	5,189	1.00	36,343	1.00
Rob5s	21.1	30.14	12,301	0.95	5,189	1.00	36,234	0.99
ROBB	6.4	9.14	13,300	1.00	5,189	1.00	29,699	0.81
RobQ	8.5	12.14	13,300	1.00	5,189	1.00	35,368	0.97
R4H	12.2	17.43	10,500	0.79	5,189	1.00	35,149	0.96
V-StaWZ	8.7	12.43	13,300	1.00	5,189	1.00	33,67	0.92
U-StaWY	10.8	15.43	13,300	1.00	5,189	1.00	30,769	0.84
R3WT	12.9	18.43	13,300	1.00	5,189	1.00	35,532	0.97
Tbra							30,347	0.83
Tdim			13,225	0.99				
Pruf			12,473	0.94				
Bcol					5,189	1.00		

For all the samples used are given: the quantity in gigabase (Gb) of the data generating by the whole-genome shotgun sequencing by Illumina technology (Data), the corresponding genome depth (x), given that 0.7 Mb (R. prolixus genome reference) was used for the calculation, the length sequences in base pair (bp) for the mtG, the nu-rDNA and the nu-PCGs datasets and the corresponding sequence coverage in percentage.
*See **Table 1** for the Rhodniini samples description.*
The Triatominae outgroups used were: T. brasiliensis (Tbra), T. dimidiata (Tdim), P. rufotuberculatus (Pruf), and B. colossus (Bcol).

nu-PCG alignments, were submitted to phylogenetic analyses. For each dataset, we used as outgroup the closest relatives identified in the sequence database. For the mitochondrial dataset, as some mitogenomes were sequenced for Triatominae, we used *T. dimidiata* (AF301594; Dotson and Beard, 2001) and *Panstrongylus rufotuberculatus* (NC_042682; Zhao et al., 2019). For the nu-rDNA dataset as only partial data is available for Triatominae, we used as outgroup the Reduviidae *Brontostoma colossus* (NC_024745; Kocher et al., 2014) and for the nu-PCGs, *T. brasiliensis* (Marchant et al., 2015). For the three concatened alignments, the sequence coverage is given for each sample in the **Table 3**. The optimal partitioning scheme and models of nucleotide substitution for each concatenated dataset were first analyzed using PartitionFinder v2.1.1 (Lanfear et al., 2012). For the concatenated mt-PCGs, alignments were also separated into single codon positions. Optimal partitioning

and nucleotide substitution models were subsequently analyzed using the Bayesian Information Criteria (BIC) for all datasets. For the concatenated datasets two types of trees were generated. MrBayes (Ronquist et al., 2012) was used, to generate Bayesian Inference (BI) topologies using the partitioning scheme and models of evolution suggested by PartitionFinder and by performing two independent 10-million generation runs. Maximum Likelihood phylogenies (ML) were inferred using IQ-TREE (Minh et al., 2020) using the partitioning scheme and models of evolution suggested by PartitionFinder and the approximate likelihood ratio test with the non-parametric Shimodaira–Hasegawa correction (SH-aLRT) (Guindon et al., 2010) was used to estimate the support for each node with 1,000 replicates. For individual mitochondrial genes, Maximum Likelihood phylogenies (ML) were inferred using IQ-TREE (Minh et al., 2020) after searching the best substitution model with ModelFinder (Kalyaanamoorthy et al., 2017) and SH-aLRT branch test with 1,000 replicates. Phylogenetic trees were displayed and modified using iTol (Letunic and Bork, 2019). Gene partitions and best substitution models for the concatenated dataset are given in the **Supplementary Table 1**.

Detection of Selection Signatures

Two approaches were used to unravel potential selective pressures during the evolution of *Rhodnius* mtDNA genomes and more specifically to identify whether the specific groups, namely *pallescens, pictipes* and *prolixus*, evolved at different rates and were under unique signatures of selection.

First, we applied the McDonald–Kreitman test (McDonald and Kreitman, 1991) using DnaSP 6 (Rozas et al., 2017) to investigate the rates of adaptive evolution in mitochondrial and nuclear genes. We used the test to compare synonymous and non-synonymous variations within and between species groups. Under neutrality (estimated with the neutrality index NI and the Fisher exact test), the ratio of replacement to synonymous fixed substitutions between species groups should be the same as the ratio of replacement to synonymous polymorphisms within species groups.

Positive selection was also investigated using branch-site models to test for differential selection on the three species groups with CodeML (PAML package v4.9; Yang, 2007) using the graphical user interface EasyCodeML (Gao et al., 2019). The mt-PCG and the nu-PCG datasets were used for this test but also the four mitochondrial genes usually used for *Rhodnius* phylogenies, namely *Cytb, COI, COII, Nad1*.

Introgression Analysis

To detect signatures of introgression in the *pallescens* and *pictipes* groups, we used the ABBA-BABA test (Green et al., 2010). The ABBA-BABA test considers that for phylogenetically-informative di-allelic sites in a four-taxon bifurcating tree, the most frequent configuration should be (A,(A,(B,B))), congruent with the species tree. Incomplete lineage sorting or homoplasy could lead to equal amounts of loci with (A,(B,(B,A))) and (B,(A,(B,A))) incongruent configurations. However, introgression could be detected if either one of the two configurations significantly exceeds the other. For the *pallescens* group, we used the three samples, *R. pallescens* R1J, *R. colombiensis* R10B and *R. ecuadoriensis,* ECUD with *R. neivai* NEII as the outgroup and for the *pictipes* outgroup, *R. stali* STAWZ, *R. brethesi* R9WX, *R. pictipes* R63K and *R. amazonicus* Ama1A as the outgroup. Raw Illumina reads of each species were mapped to the *R. prolixus* reference sequences[2] using minimap2 (v. 2.17-r941; Li, 2018) with default parameters. The mapped reads were filtered for a mapping quality of 20 with samtools (v. 1.9; Li et al., 2009). BAM files were cleaned by removing unmapped reads and sorted by coordinate using Picard v. 1.4.2[3]. The sorted BAM files were then combined in a mpileup file with samtools and finally converted to a synchronized file (one species per column) with custom scripts from PoPoolation2 (v. 1.201; Kofler et al., 2011). Using a customized perl script on the synchronized file, we counted the three configurations for all diallelic loci with a minimum coverage of 10 reads in the four species.

RESULTS

Mitochondrial Phylogenetic Inferences

The mitochondrial dataset is composed of 36 samples from the Rhodniini tribe and the 15 mitochondrial genes. The concatenated alignment totalized 13.3 kb with a nucleotide diversity (Pi) of 0.12389. No pseudogene was retrieved from our dataset as attesting by the functional proteins obtained after the traduction (**Supplementary Table 1**). For the concatenated dataset mtG, highly congruent trees were obtained from both the Maximum-Likelihood (**Figure 1**) and the Bayesian Inference (**Figure 2**) methods and showed *pallescens* and *pictipes* groups as sister clades with high statistical support (ML: 98.8, BI:100). For the individual genes, various topologies were obtained (**Supplementary Figures 1–15**). For nine genes, the same topology as for the concatenated dataset was obtained, namely for *Cytb* (**Supplementary Figure 8**), Nad1 (**Supplementary Figure 9**), *Nad5* (**Supplementary Figure 14**), *Nad6* (**Supplementary Figure 15**) and also for *Atp6* (**Supplementary Figure 3**), *Atp8* (**Supplementary Figure 4**), *COI* (**Supplementary Figure 6**), *COIII* (**Supplementary Figure 7**), and *Nad2* (**Supplementary Figure 10**), but for these five genes a paraphyly was observed for the *prolixus* group, *R. neivaii*, *R. domesticus*, and/or *P. tertius*, according to genes being outside the prolixus group. For four genes, the *pictipes* and the *prolixus* groups were sister clades, namely for *12S* (**Supplementary Figure 1**), *16S* (**Supplementary Figure 2**), *Nad3* (**Supplementary Figure 11**), *Nad4* (**Supplementary Figure 12**), and the *Nad4L* but for this gene the *pictipes* group is less supported (**Supplementary Figure 13**). The third topology, that is the *pallescens* and *prolixus* groups as sister clades, was displayed for the COII gene (**Supplementary Figure 6**) but the position of *R. neivaii* and *R. domesticus* were weak supported.

The analysis of rare insertion/deletion (InDel) patterns in the alignment indicated the presence of an InDel of 40 pb in the tRNA Cys that supported the genetic proximity between the *pallescens*

[2]www.vectorbase.org

[3]http://broadinstitute.github.io/picard/

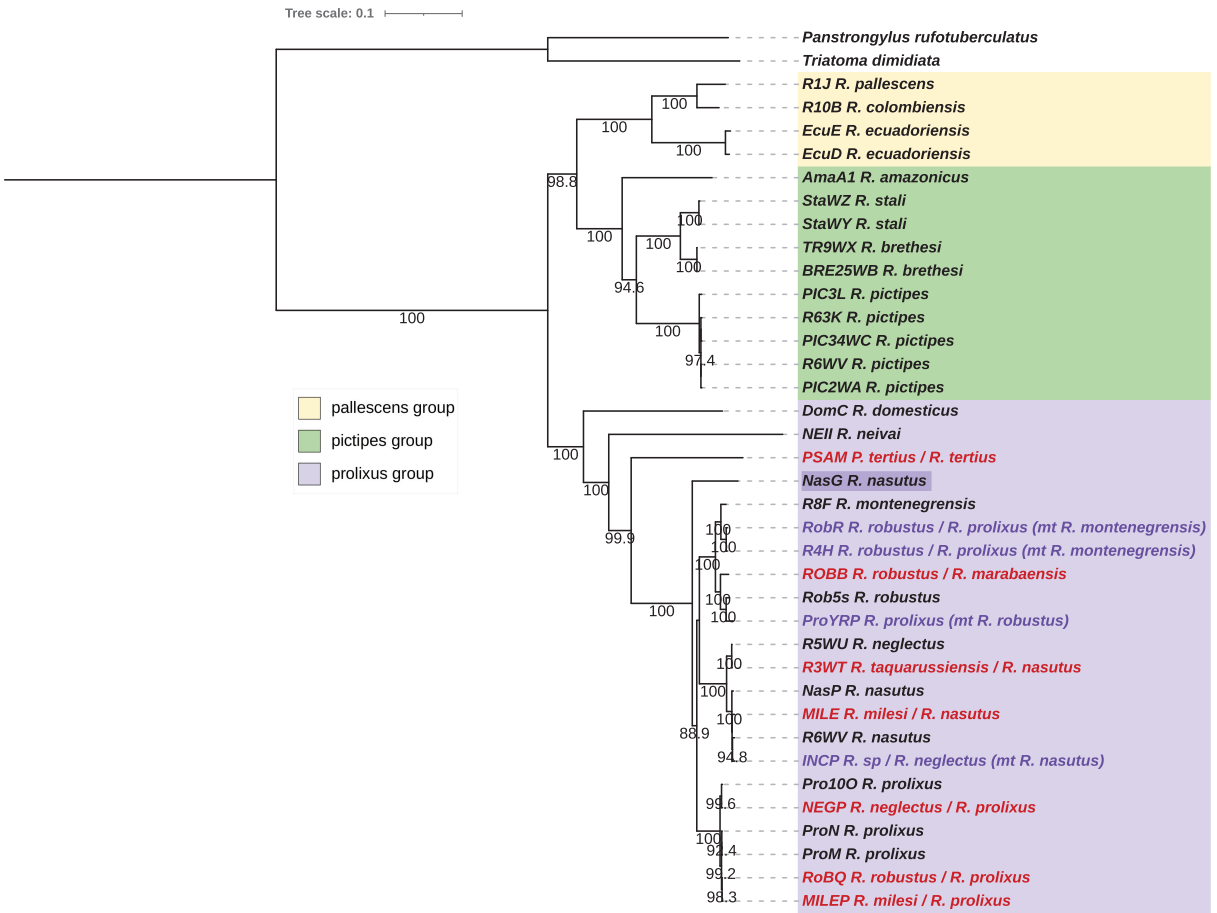

FIGURE 1 | Phylogenetic inferences for the concatenated mitochondrial (mtG) dataset for *Rhodnius* species with the Maximum Likelihood (ML) method. Statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *T. dimidiata* and *P. rufotuberculatus* are used as outgroups. Introgressed species inferred from the nuclear phylogeny (see **Figure 5**) are indicated in purple, misidentified species in red, and species with incongruent position are highlighted in dark purple.

and the *pictipes* groups and the split with the *prolixus* group (**Supplementary Figure 16**).

According to the concatenated mtG trees (**Figures 1**, **2**), for the *pictipes* group, *R. stali* and *R. brethesi* were closer to each other than to *R. pictipes*. In the *pallescens* group, *R. pallescens* and *R. colombiensis* were closer to each other than to *R. ecuadoriensis*. The *prolixus* group comprised samples of *R. domesticus, R. neivai, R. montenegrensis, R. taquarussuensis* (syn *R. neglectus*), *R. neglectus, R. nasutus, R. robustus, R. prolixus* and also *P. tertius*. Indeed, the *Psammolestes* species was robustly included in the *prolixus* group independently of the considered mitochondrial dataset. However, in the two trees, ML (**Figure 1**) and BI (**Figure 2**), the genetic proximity between the three clades (*R. montenegresis-R. robustus/R. nasutus-R. neglectus/R. prolixus*) is not resolved. For *R. amazonicus* some discrepancies were observed according to the mitochondrial dataset. For the concatenated dataset, this taxon was related to the *pictipes* group at a basal position but for the individual datasets it was either related to the *pallescens* group or the *pictipes* group.

Three Brazilian specimens seemed to have an aberrant position in the trees. Specimens NEGP and MILEP, morphologically identified as *R. neglectus* and *R. milesi* respectively, were related to *R. prolixus*. The NASG specimen, identified as *R. nasutus,* is at a basal position in the ML tree (but with a weak bootstrap) and nested within the *prolixus* group but not related to the *R. nasutus-R. neglectus* clade in the BI tree. The non-identified INCP sample was related to *R. nasutus*.

For samples identified as *R. robustus* and *R. prolixus,* two clades were inferred, each of them comprising samples from the two species. One is composed of all the *R. robustus* samples except the field specimen RobQ from Brazil, the laboratory strain *R. montenegrensis* and one *R. prolixus* sample, YRP, a Venezuelan laboratory strain from Cojedes. The other grouped all other *R. prolixus* samples with the RobQ.

Concerning species with questionable species-specific status in addition to the topologies observed in the phylogeny, the pairwise distances (**Supplementary Table 2**) and the similarity index can be considered using the mitochondrial PCG concatenated dataset. *R. neglectus* (R5WU) and *R. taquarussuensis* (syn *R. neglectus*) (R3WT) appeared to be very closely related taxa (pairwise distance = 0.0012; similarity = 99.88%). Unexpectedly, in the phylogeny, *R. milesi* (MILE) was included in the same

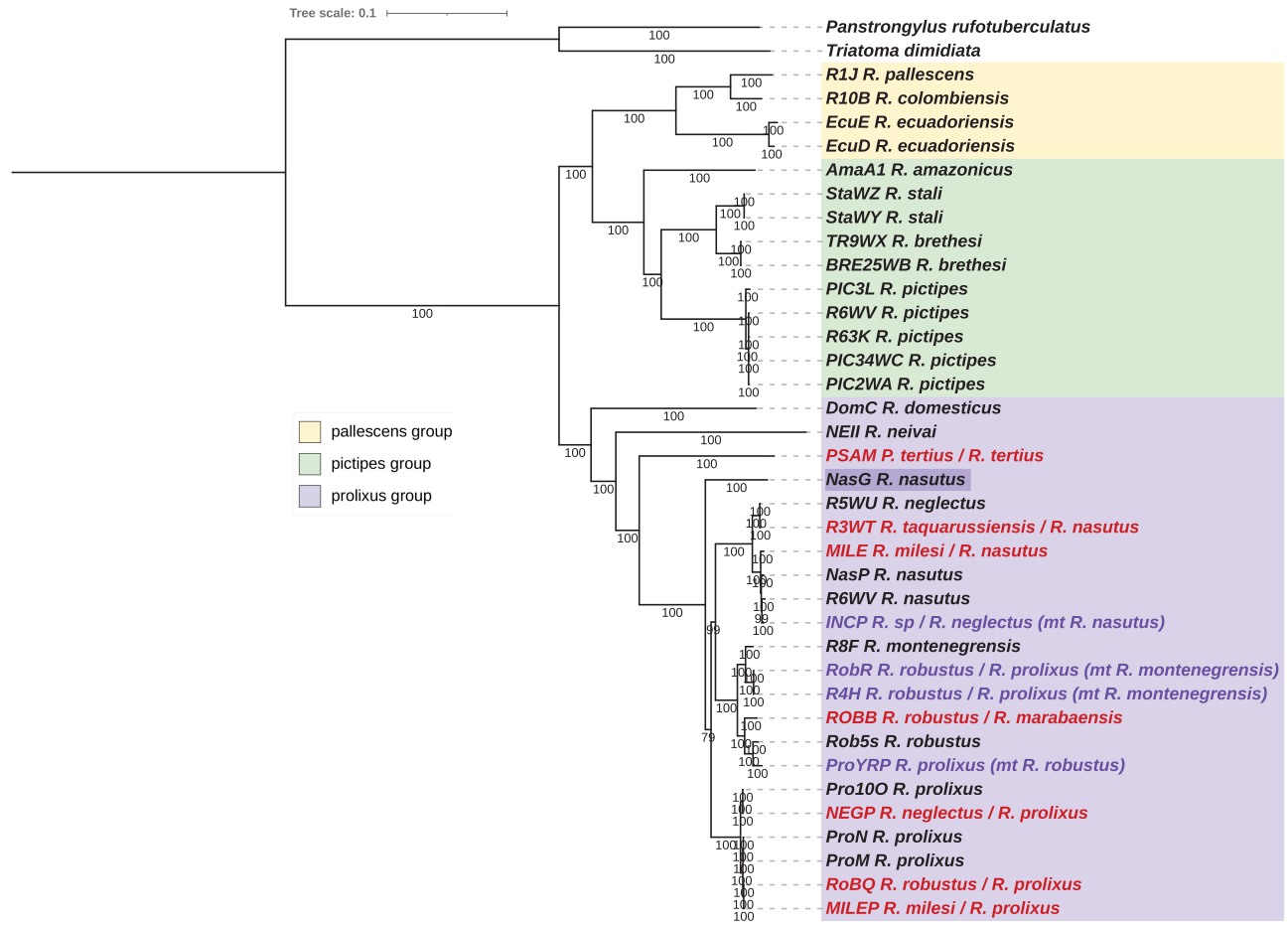

FIGURE 2 | Phylogenetic inferences for the concatenated mitochondrial (mtG) dataset for *Rhodnius* species using Bayesian Inference (BI) method. Statistical supports (specify Bayesian posterior probabilities in percent) are indicated on the branches. *T. dimidiata* and *P. rufotuberculatus* are used as outgroups. Introgressed species inferred from the nuclear phylogeny (see **Figure 6**) are indicated in purple, misidentified species in red, and species with incongruent position are highlighted in dark purple.

clade as the laboratory reference R5WU *R. nasutus* sample (MILE R7WU: 0.0069; 98.29%). More distant pairwise values were obtained between *R. montegrenensis* and *R. robustus,* the smallest being with the *R. robustus* samples from Peru (R8F RobR: 0.0149; 98.56%) and the highest with the Amazonian ones, the Brazilian ROBB (R8F ROBB: 0.0311; 97.13%) and the Guyanese Rob5s (R8F Rob5s: 0.0326; 97%). For comparison, the pairwise values between *R. nasutus* (RW7) and *R. neglectus* (R5WU) that are close but clearly distinct species, were smaller (0.0202; 98.09%) than those between *R. robustus* and *R. montenegrensis*.

Nuclear rDNA Phylogenetic Inference

The nuclear rDNA dataset included 36 taxa from the Rhodniini tribe for an alignment length of 5.2 kb and a Pi of 0.0089. ML and BI rDNA phylogenetic analysis confirmed the three major Rhodniini groups (*pallescens, pictipes*, and *prolixus*) but not supported by high bootstrap values (**Figures 3**, **4**). However, an InDel of 19 pb in the 28S rDNA genes supported the genetic proximity between the *pallescens* and the *pictipes* groups (**Supplementary Figure 16**).

For the resolution within the *Rhodnius* groups, only the BI tree gave some reliable results (**Figure 4**). The *pallescens* group exhibited a good resolution level, with the two species *R. pallescens* and *R. ecuadoriensis* closer related than to *R. colombiensis*. For the *prolixus* group, only the R7WU *R. nasutus* is differentiated from the other *R. neglectus* samples including *R. taquarenssensis* (syn. *R. neglectus*) (R3WT). Two *R. robustus* specimens (RobR, Rob5s) were clustered with *R. montenegrensis* R8F. Note that *Psammolestes* was clustered with the *R. prolixus* species. The *pictipes* group was the least resolute.

Nuclear Protein-Coding Gene Phylogenetic Inferences

The nu-PCG dataset included 36 taxa from the Rhodniini tribe. The trees were reconstructed using 51 PCGs for an alignment length of 36.3 kb with a Pi of 0.02648. The tree topologies between

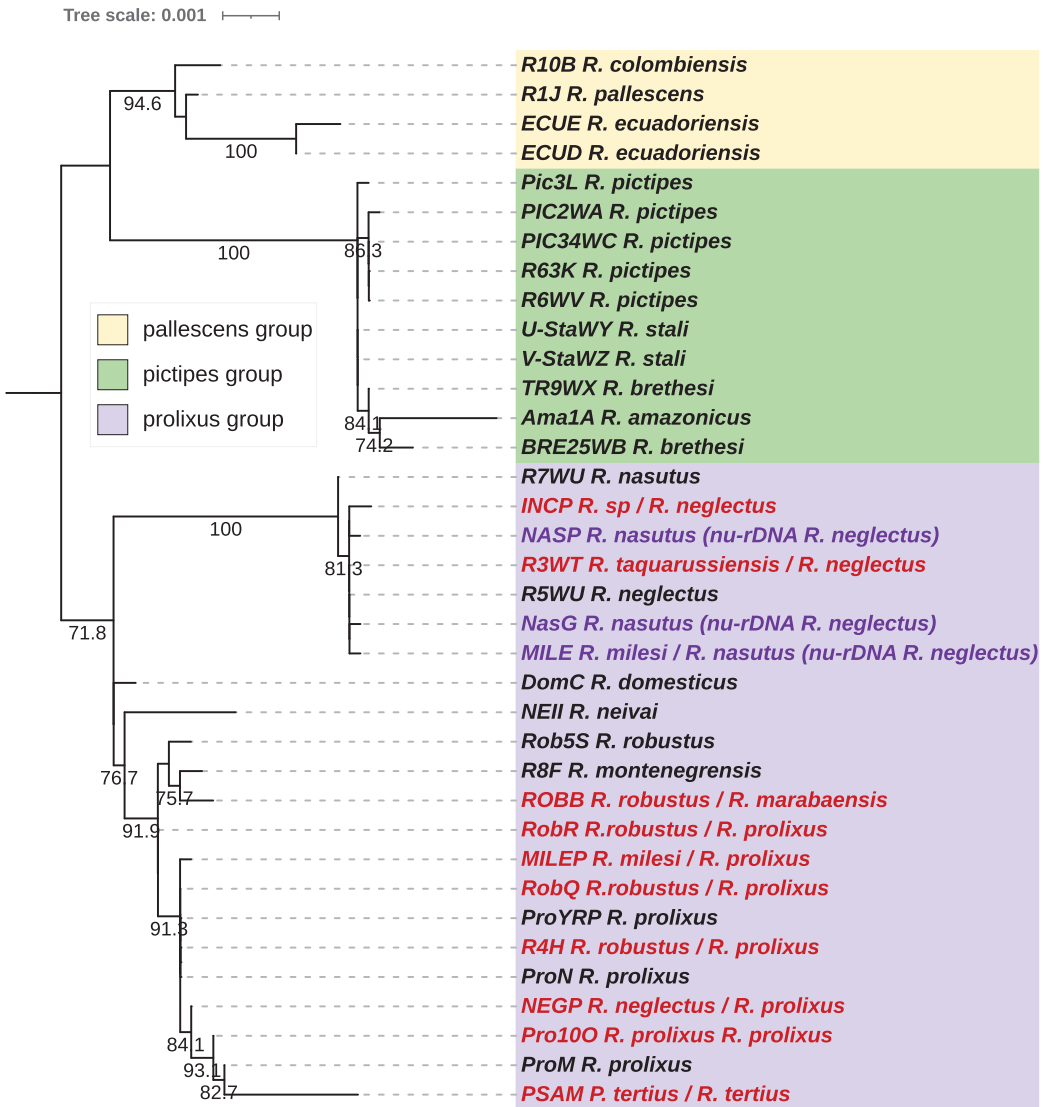

FIGURE 3 | Phylogenetic inferences for the concatenated nuclear rDNA (nu-rDNA) dataset for *Rhodnius* species with the Maximum Likelihood (ML) method. Statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *B. colossus* was used as outgroup. Introgressed species inferred from the nuclear phylogeny (see **Figure 5**) are indicated in purple, and misidentified species in red.

the reconstruction methods, ML (**Figure 5**) and BI (**Figure 6**) are fully congruent and robust.

As for the mitochondrial dataset, the three groups, *pictipes*, *pallescens*, and *robustus,* were robustly identified (MI and BI trees) with the same topology with *pallescens* and *pictipes* as sister groups. Moreover, the same topology was found for the *pallescens* group, *R. pallescens* and *R. colombiensis* being closer to each other than to *R. ecuadoriensis.* Discrepancies were found for the *pictipes* and the *prolixus* groups. Unlike the relationship revealed with mitochondrial data, *R. stali* and *R. pictipes* appeared with nuclear data closer to each other than to *R. brethesi*. Similar to what was observed with the concatenated mitochondrial dataset, *R. amazonicus* was related to the *pictipes* group at a basal position. For *R. robustus* and *R. prolixus* samples, two well supported related clades were observed. The Guyanese *R. robustus* (Rob5s), the Brazilian *R. robustus* (ROBB) and the *R. montenegrensis* (R8F) samples were grouped in the same clade. All the other *R. robustus* and *R. prolixus* samples were in another clade.

Concerning the pairwise distances (**Supplementary Table 2**) and the similarity index, *R. neglectus* (R5WU) and *R. taquarussuensis* (syn *R. neglectus*) (R3WT) were very closely related (pairwise distance = 0.00076; similarity = 99.82%), and *R. milesi* (MILE) was closest to R7WU *R. nasutus* (MILE R7WU: 0.00192; 99.81%). More distant pairwise values than for the two previous pairs of species were obtained between *R. montegrenensis* and *R. robustus,* e.g., RobB (R8F RobB: 0.00309; 99.62%) and Rob5s (R8F Rob5s: 0.00408; 99.59%) and were of the same order of magnitude as between the Guyanese *R. robustus* (Rob5s) and *R. prolixus* samples e.g., Pro100 (Rob5s Pro100: 0.00485; 99.52%). Similar distant pairwise values were

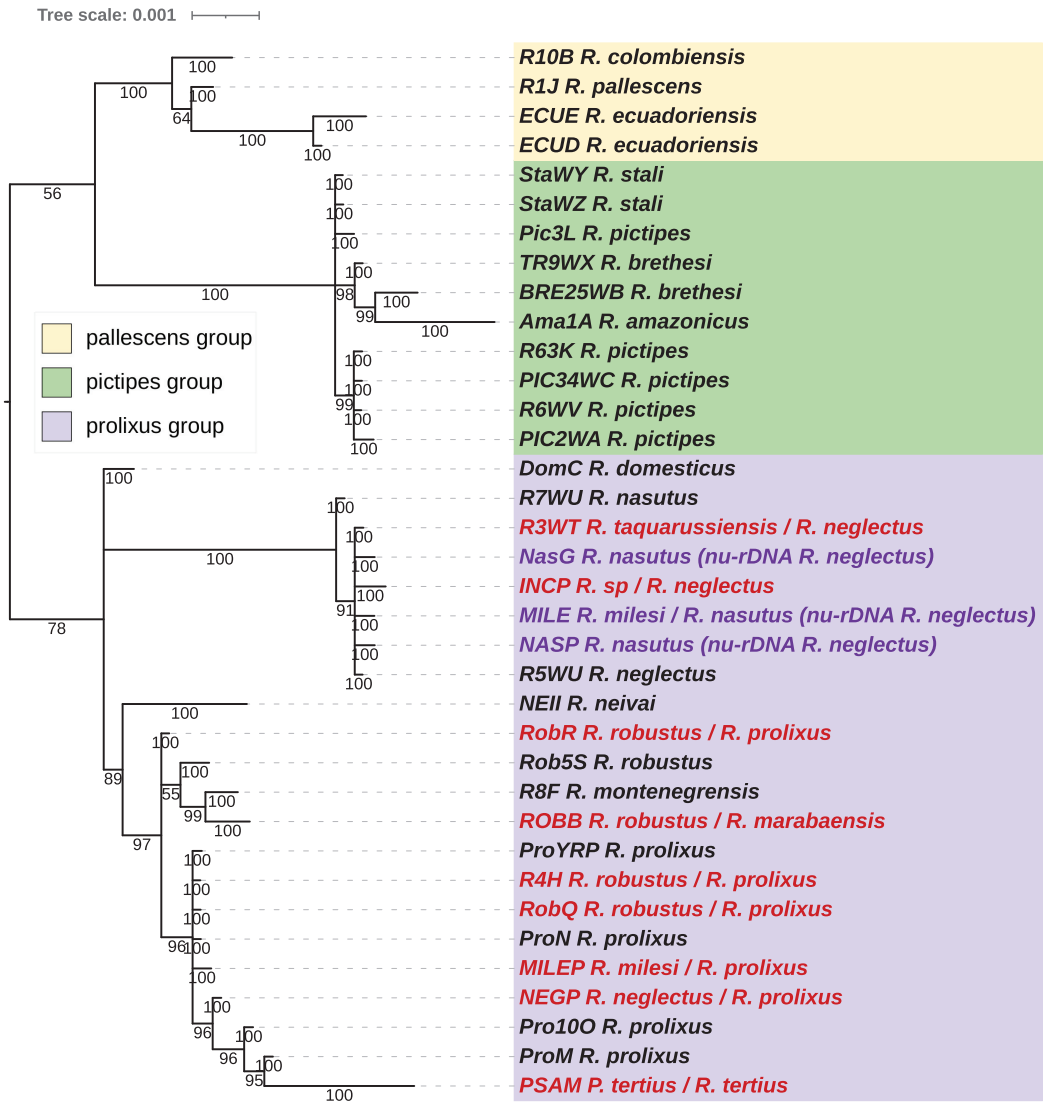

FIGURE 4 | Phylogenetic inferences for the concatenated nuclear rDNA (nu-rDNA) dataset for *Rhodnius* species using Bayesian Inference (BI) method. The statistical supports (specify Bayesian posterior probabilities in percent) are indicated on the branches. *B. colossus* was used as outgroup. Introgressed species inferred from the nuclear phylogeny (see **Figure 5**) are indicated in purple, and misidentified species in red.

obtained between *R. nasutus* and *R. neglectus* (R7WU R5WU: 0.0036; 99.64%) or between the R5WU *R. neglectus* and INCP (R5WU INCP: 0.004; 98.16%) or NASP (R5WU NASP: 0.00396; 99.65%). It should be of interest to note that NasG, placed outside the *R. nasutus-R. neglectus* clade for the concatenated mitochondrial dataset, was close related to R7WU *R. neglectus* (NASG R7WU: 0.00182; 99.82%).

Test of Selection

The McDonald–Kreitman test detected significant positive selection in the concatenated mitochondrial dataset for the three group comparisons, *prolixus-pallescens* groups (NI = 0.696; *p*-value = 0.023), *prolixus-pictipes* groups (NI = 0.609; *p*-value = 0.003), and *pallescens-pictipes* groups (NI = 0.779; *p*-value = 0.04), revealing an excess of divergent non-synonymous mutations between the group tested. For the concatenated nuclear PCG dataset no selection was detected for any pairwise, *prolixus-pallescens* groups (NI = 0.754; *p*-value = 0.41), *prolixus-pictipes* groups (NI = 0.583, *p*-value = 0.199), and *pallescens-pictipes* groups (NI = 0.931; *p*-value = 0.792).

Selective pressure was also tested using PAML method (**Table 4**). For the concatenated mitochondrial dataset, selective pressure was detected in each *Rhodnius* group test against the others (LRT *p*-value = 2×10^{-9} for *pallescens* vs. *pictipes* and *prolixus* groups; 0.0044 for *pictipes* vs. *pallescens* and *prolixus* groups; 9.6×10^{-7} for *prolixus* vs. *pictipes* and *pallescens* groups). For the single mitochondrial genes tested, namely *cytb, COI, COII, Nad1*, all were under selective pressure for the *prolixus* group tested against the *pictipes* and *pallescens* groups (LRT *p*-value = 13×10^{-6}; 13×10^{6}; 0.0009; 0.02 respectively) and

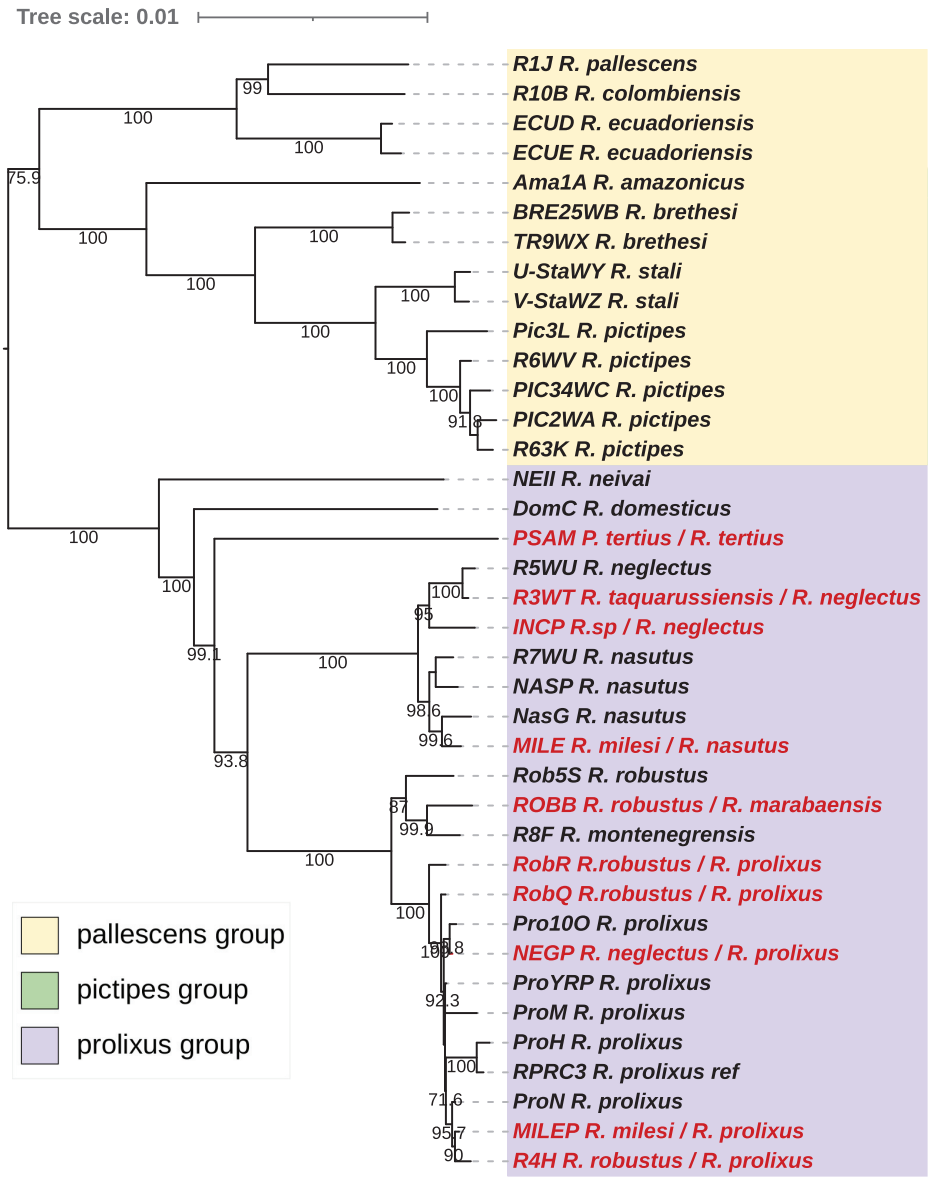

FIGURE 5 | Phylogenetic inferences for the concatenated nuclear protein-coding gene (nu-PCGs) dataset for *Rhodnius* species with the Maximum Likelihood (ML) method. The statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *T. brasiliensis* was used as outgroup. Misidentified species are indicated in red.

two, *COI* (LRT *p*-value = 0.046) and *COII*, for the *pallescens* group tested against the *pallescens* and *prolixus* groups (LRT *p*-value = 0.04). For the concatenated nuclear dataset no selective pressure was detected for any of the groups (LRT *p*-value = 0.47; 0.10; 0.14 for *pallescens, pictipes* and *prolixus* groups, respectively, tested against the others groups).

Genome-Wide Examination for *pallescens* and *pictipes* Groups

We conducted the ABBA-BABA test of introgression in the *pictipes* group for which a discordance between the mitochondrial and nuclear PCG trees was noticed. *R. stali* and *R. brethesi* were indeed clustered together in the mitochondrial tree while *R. stali* was closer to *R. pictipes* in the nuclear tree.

The three species *R. brethesi*, *R. stali* and *R. pictipes* were tested for introgression and *R. amazonicus* as the outgroup. The ABBA-BABA results were in accordance with the nuclear phylogeny of the species and provided a strong support for a mitochondrial *R. stali* introgression in *R. brethesi*. Out of 6,437 bi-allelic SNPs, we counted 5,571 AABB, namely *R. pictipes* closer to *R. stali*, 409 ABBA, namely *R. brethesi* closer to *R. pictipes* and 457 BABA, namely *R. brethesi* closer to *R. stali*. No signal of further nuclear introgression was found between *R. brethesi* and *R. stali* since

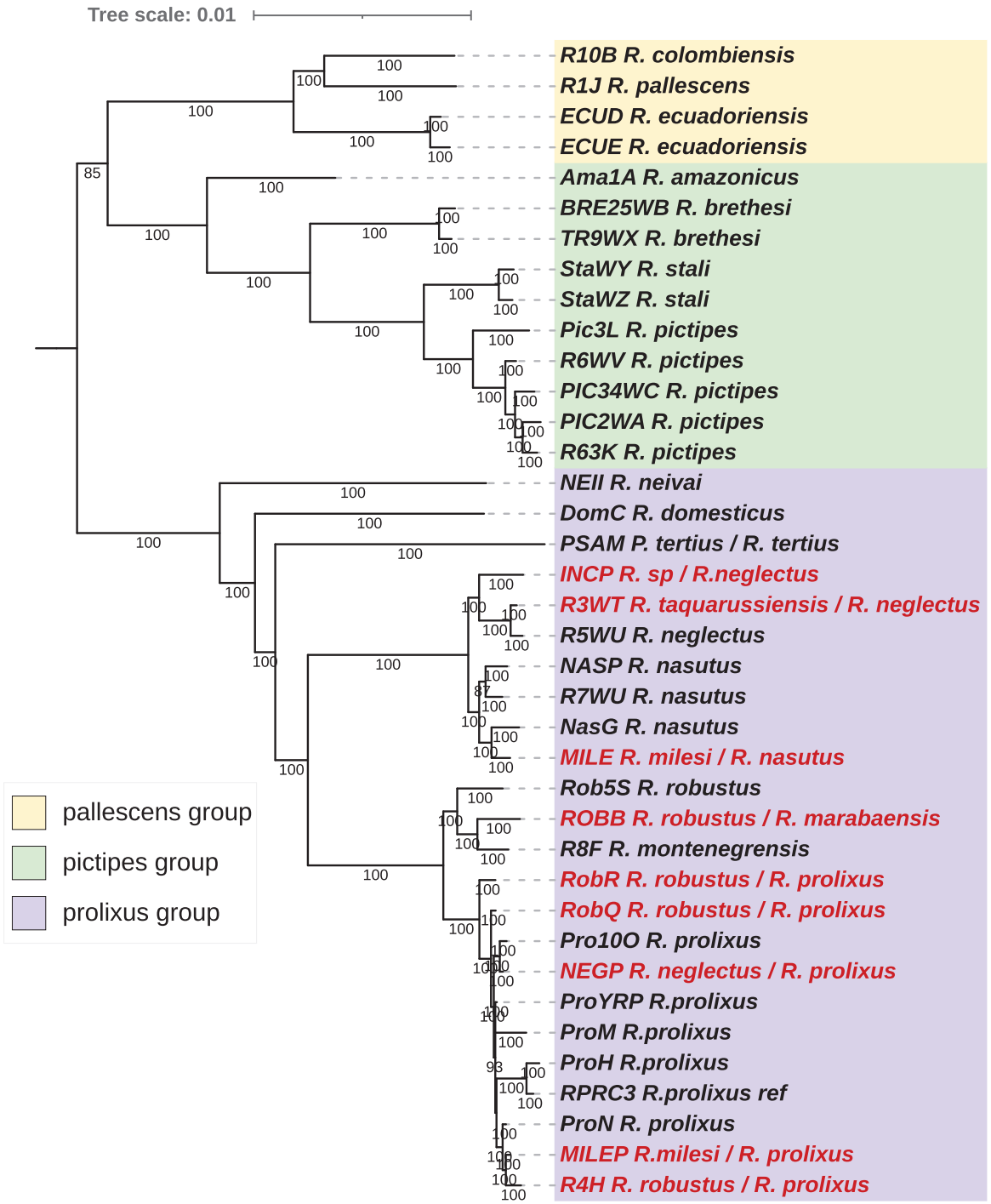

FIGURE 6 | Phylogenetic inferences for the concatenated nuclear protein-coding gene (nu-PCGs) dataset for *Rhodnius* species using Bayesian Inference (BI) method. The statistical supports (specify Bayesian posterior probabilities in percent) are indicated on the branches. *T. brasiliensis* was used as outgroup. Misidentified species are indicated in red.

both ABBA and BABA counts were almost equal and far below the AABB (i.e., the nuclear tree topology) SNP counts.

We conducted the same analysis in the *pallescens* group, for which the mitochondrial and nuclear PCG trees are congruent showing *R. colombiensis* and *R. pallescens* clustered together, but in the rDNA phylogeny, *R. pallescens* and *R. colombiensis* are closer. We found a strong signal for pervasive and widespread phylogenetic incongruence between the three species,

TABLE 4 | Test of selection results using the PAML package.

Datasets	*pallescens* group	*pictipes* group	*prolixus* group
mt-PCGs	**0.000000002** ω0 = 0.06533 ωpall = 0.04302	**0.004452367** ω0 = 0.06222 ωpic = 0.04958	**0.000000958** ω0 = 0.05644 ωpro = 0.07712
Cytb	0.141282571 ω = 0.03490	0.248087413 ω = 0.03490	**0.000013280** ω0 = 0.000013280 ωpro = 0.06698
COI	**0.046034763** ω0 = 0.01252, ωpall = 0.00639	0.521251673 ω = 0.01134	**0.000013280** ω0 = 0.02806 ωpro = 0.06698
COII	**0.040155883** ω0 = 0.02126, ωpall = 0.01023	0.139838995 ω = 0.01887	**0.000910664** ω0 = 0.01453 ωpro = 0.04180
Nad1	0.202681777 ω = 0.01952	0.440764212 ω = 0.01952	**0.020486354** ω0 = 0.01366 ωpro = 0.03990
nu-PCGs	0.696800235 ω = 0.04575	0.373073414 ω = 0.04575	**0.001175575** ω0 = 0.05701 ωpro = 0.02452

For each gene or the concatenated datasets, the null model (ω for the whole tree) is compared by an LRT test to the model with 2 ratios: ω0, the ratio for branches of the two Rhodnius groups not tested and ωpall or ωpic or ωpro, the ratio for the Rhodnius group tested.
The model that best explains the data is reported in the table with the p-value. Significant p-values are indicated in bold.

R. pallescens, R. colombiensis, and *R. ecuadoriensis*. Out of a total of 18,910 di-allelic SNPs, the ABBA-BABA test gave 7,666 sites with the configuration AABB, namely *R. pallescens* is closer to *R. colombiensis*, 4,084 with the ABBA configuration supporting a closer relationship between *R. ecuadoriensis* and *R. colombiensis*, and 7,160 with the configuration BABA supporting a closer relationship between *R. ecuadoriensis* and *R. pallescens*. Remarkably, the BABA counts even approached those of the AABB species tree counts, indicating that the exact position of *R. pallescens* could not be determined with certainty. In order to test if the ABBA or BABA configurations are restricted to specific genomic regions, the evaluated SNPs were mapped to the reference genome of *R. prolixus*. We did not observe any specific localization, the three configurations were widespread and overlapping on all scaffolds of the reference genome of *R. prolixus*.

DISCUSSION

Phylogenetic Relationship Between the Three *Rhodnius* Groups

In this study, our goal was to provide a large genomic dataset in order to elucidate key issues on phylogenetic relationships for yet the most taxonomically controversial Tritaominae tribe, the Rhodniini, using both mitochondrial and nuclear datasets analyzed separately. Previous *Rhodnius* phylogenetic studies used a limited number of genes, which could explain the divergent topologies found for the *Rhodnius* groups. Three topologies were recovered independently of the type of marker used. Considering studies with at least eight *Rhodnius* species and reliable phylogenetic reconstruction methods (Maximum Parsimony, ML, BI), the *pictipes-pallescens* clade was recovered using only mitochondrial markers, as *cytb* (Maia da Silva et al., 2007), as well as using a combination of mitochondrial and nuclear genes, as *16S* + *cytb* + *28S* (Monteiro et al., 2000). The *pictipes-prolixus* clade was obtained using also only mitochondrial markers: 16S (Paula et al., 2005, 2007) or *16S* + *cytb* (Hypša et al., 2002), as well as mitochondrial and nuclear markers together, namely *18S* + *28S* + *16s* + *Cytb* + *COI* + *COII* (Justi et al., 2014) or *18S* + *28S* + *wg* + *16S* (Justi et al., 2016). It is to highlight that Paula et al. (2021) by combining the three genes, *16S, 18S, and wg*, obtained contrasted results according to the type of analysis used, namely *pictipes-prolixus* clade with MP and *pictipes-pallescens* with ML. Using ultraconserved elements and/or rDNA data, Kieran et al. (2021) obtained the *pallescens-prolixus* clade. To limit the methodological bias due to a limited dataset, we used in this study a large genomic dataset of both mitochondrial (13.3 kb) and nuclear data (5.3 kb of rDNA; 51 PCGs, 36.3 kb). The large dataset was able to resolve the relationship between the three *Rhodnius* groups. Except for some single mitochondrial genes, both the concatenated mitochondrial and nuclear datasets showed *pictipes* and *pallescens* as sister groups, irrespective of the outgroup used.

Variation Rate Among Partitions

The incongruence of phylogenetic signals between different genomic regions can result from the rate of variation among partitions. Our study indicates that the nuclear rDNA markers (including *18S, 5.8S, 28S*, and *ITS*) led to a poorly resolved tree. Indeed, the nucleotide divergence of the rDNA marker dataset is low (max 1.99%) compared to the mitochondrial *Rhodnius* dataset (max 15.06%), and the nucleotide diversity (Pi) is 14 times slower than the mitochondrial dataset one. It results that slow-evolving rDNA genes are less suitable to infer phylogenetic relationships. Mitochondrial genes, which evolve more rapidly than the rDNA and the set of insect universal nuclear genes (BUSCO genes), seem to be candidates for more accurate phylogenetic reconstruction. In our dataset, based on Pi estimates, the mitochondrial genes evolve 4.7 times faster than the nuclear PCGs, which is in accordance with what was found in insects, the mitochondrial genes being estimated to evolve 2–9 times faster than the nuclear protein-coding genes (Lin and Danforth, 2004). Moreover, mitochondrial genes are easier to use for both technical and evolutionary reasons. Indeed, mitochondrial primers are widely available and as these mitochondrial genes are clonally inherited and non-recombining, they are not subjected to recombination and heterozygosity like the nuclear markers. Nevertheless, mitochondrial genes have some clear disadvantages (Lin and Danforth, 2004). Since all mitochondrial genes are linked to the same DNA unit, they cannot provide an independent estimate of phylogeny in the same way that unlinked single-copy, as the nuclear genes do (Harrison, 1989). Furthermore, the higher rate of substitution can be disadvantageous to resolve divergences between taxa of more than 5–10 million years, while they are useful to resolve

relationships between closely related taxa that have diverged relatively recently.

Selective Pressure on Mitochondrial Genes

In our study, we demonstrated that mitochondrial genes widely used for phylogenies, namely *cytb* and *COI,* were under selection for all the three groups, *pallescens*, *pictipes* and *prolixus.* By contrast, no selective pressure was evidenced for the nuclear PCGs. In the literature, there is increasing evidence that the genetic diversity of mtDNA may therefore be shaped not only by random genetic drift but also by natural selection given the functional importance of mitochondrial polymorphisms. Positive selection on the mitogenome has been shown for various species including vertebrates and insects (reviewed by Sun et al., 2018). This process may be related to energy production linked to migration but also aging, size, diet, neuronal function, and thermoregulation (see: Garvin et al., 2015; Sloan et al., 2017). For the *Rhodnius* species, selective pressure is likely linked to fitness-related traits that have to be investigated.

Species Diversity in the *Rhodnius* Genus

A delicate point of this work is to give molecular pertinent arguments about the species-specific status of some controversial or recently described species.

Concerning the status of the *Psammolestes* genus, some morphological differentiations with the *Rhodnius* genus has justified the erecting of two distinct genera as shorter head, stronger legs, wider femora and wider rostrum but also the lifestyle of the *Psammolestes* species in nests of birds while *Rhodnius* species live mainly in palm tree (Lent and Wygodzinsky, 1979). Our work clearly established using both large mitochondrial and nuclear datasets that *P. tertius* is included into the *Rhodnius* genus and the *prolixus* group. This position into the *Rhodnius* genus was also pointed out by previous studies (e.g., Lyman et al., 1999; Monteiro et al., 2000; Hypša et al., 2002; Justi et al., 2014; Paula et al., 2021). Thus, according to the cladistic principles and the rule of species name precedence, we strongly recommended reclassifying the *P. tertius* species as *Rhodnius tertius*. Complementary studies combining morphological and molecular approaches will be useful to know to what extent the morphological characters put forward in *Psammolestes* are part of some adaptive processes acting on *Rhodnius* species from differentiated environments. It is interesting to note that among the species of the *Rhodnius* genus, *R. paraensis* which presents a particular ecology because found in the nests of arboreal rodents, is a species of small size as *Psammolestes* with also short head, stout legs, short and stout antennae that led Lent and Wygodzinsky (1979) to suggest that these characters "make this one of the most plesiomorphic species of *Rhodnius*, if the progressive elongation of the head and appendages found in most other species is considered as derived." Phylogenetic studies including *R. paraensis* and *Psammolestes* species are required to determine the polarity (ancestral vs. derived) of the character states. Since the three *Psammolestes* species are morphologically and ecologically close (Lent and Wygodzinsky, 1979), share cytogenetic characteristics indicating a chromosomal homogeneity (Oliveira et al., 2018), and are phylogenetically closely related (rDNA, Kieran et al., 2021), we recommended to re-examine also the taxonomic status of the two other *Psammolestes* species, *P. arthuri* and *P. coreodes*.

The species *R. amazonicus* was initially described by Almeida et al. (1973). It was then invalidated by Lent and Wygodzinsky (1979) because of the description made on only one female collected in the municipality of Manaus (Amazonas, Brazil) with several *R. pictipes* specimens and was then seen as a poorly preserved specimen of *R. pictipes*. The species was then revalidated by Bérenger and Pluot-Sigwalt (2002) from both male and female specimens collected in French Guyana and more recently recorded by Rosa et al. (2017b) in Pará, Brazil. We provided the first molecular data for this species and confirmed by both our robust mitochondrial and nuclear phylogenies its species-specific status as *bona species*. Moreover, this taxon belonged to the *pictipes* group at a basal position.

In contrast, the large molecular data set of our study definitively closes the question about the species-specific status of *R. taquarussuensis* (syn *R. neglectus*), described by Rosa et al. (2017a) from a female, which invaded a rural dwelling in the city of Taquarussu, Mato Grosso do Sul, Brazil. Because of its very close genetic similarity for both mitochondrial and nuclear genes, we demonstrated that this sample belongs undoubtedly to the *R. neglectus* species. This result corroborated the revision performed by Nascimento et al. (2019) using two mitochondrial and four nuclear genes leading the authors to state that the morphological and constitutive heterochromatin pattern differences between *R. taquarussuensis* and *R. neglectus* are likely intraspecific polymorphism of *R. neglectus* and consequently to synonymize *R. taquarussuensis* with *R. neglectus*.

In our dataset, we included *R. milesi,* a less studied species. This species was described from Pará, Brazil and morphologically similar to *R. dalessandroi* (Valente et al., 2001). Yet, for this latter species, only the published description and illustration are available leading Lent and Wygodzinsky (1979) to not include it in their taxonomical work. Doubts about the species-specific status of *R. milesi* had been issued by Monteiro et al. (2018) in their review. The authors concluded to the proximity of *R. milesi* with *R. neglectus* based on molecular *cytb* and *ITS* sequences but the accession number of the *R. milesi* sequences are not provided. In our dataset, for both the nu-PCG and mitochondrial phylogenies, MILE is congruently related to the R7WU *R. nasutus* sample with low nuclear genetic pairwise values (<0.002) but is clustered with *R. neglectus* in the rDNA phylogeny. A possible explanation is that the MILE specimen used in this study is a *R. nasutus* with a probable introgression with some nuclear DNA from *R. neglectus*. This molecular finding is in accordance with the morphological observation made by Carcavallo et al. (in Valente et al., 2001) observing that *R. milesi* male genitalia have a second phallosome process, only found in *R. nasutus*. With respect to our study, we conclude that *R. milesi* should be considered as synonym to *R. nasutus* and we recommend reconsidering the species-specific status of this species with a larger sampling.

A crucial point concerns two species closely related to *R. robustus*, namely the two newly described *R. montenegrensis* (Rosa et al., 2012) and *R. marabaensis* (Souza et al., 2016). *Rhodnius montenegrensis* was described from a strain established from eight specimens collected in a palm tree in the municipality of Monte Negro, Rondonia, Brazil and morphologically identical to specimens previously caught in the same locality (Rosa et al., 2012). The authors assessed their description on morphological and *cytb* sequence divergences with *R. robustus*. But, previously, using a large set of *R. prolixus* and *R. robustus* samples, Monteiro et al. (2003) based on *cytb* phylogeny have described four clades for *R. robustus* and proposed to see *R. robustus* as a complex of species and concluded for a paraphyly for *R. robustus*. Thus, the validity of the newly described species was pointed out by Abad-Franch et al. (2013) by declaring that "*R. montenegrensis* likely represents one of the *R. robustus* lineages of Monteiro et al. (2003)." Using *cytb* and ITS2 genes from transcriptomic data, Brito et al. (2019) evidenced that *R. montenegrensis* and *R. robustus* clade II are likely the same species but this does "not invalidate the former as a separate species." Concerning *R. marabaensis,* this species was described from Marabá, Pará, Brazil (Souza et al., 2016). Unfortunately, no DNA sequence is available in public database for the described species, the authors only stated that the *cytb* sequence shown 99% of identity with that of *R. robustus*. In this study, based on the nuclear genes, we evidenced similar genetic distance values between the three pairwise involving Rob5s *R. robustus* from field French Guyana, *R. montenegrensis* (R8F), and ROBB. Similar distant pairwise values were also obtained between the two species *R. nasutus* and *R. neglectus*. Indeed, if we recognized *R. nasustus* and *R. neglectus* as valid species, the three entities must be too, especially since they displayed congruent position in all trees, mitochondrial and nuclear ones. Thus, we state that the sample Rob5s is representative of *R. robustus,* R8F of *R. montenegrensis* and the specimen ROBB has also a species-specific status. It is worth noting that ROBB was caught in a palm tree near the Benfica Field Station (5°16′S, 49°50′E) at 50 km from Marabá, Pará, Brazil. Since *R. marabaensis* was described from Marabá (Souza et al., 2016), even if molecular data for this are not publicly available, a parsimonious hypothesis could be formulated that ROBB is putatively *R. marabaensis*. The same assumption was made by Castro et al. (2020) assimilated the *R. robustus* clade III to *R. marabaensis* based on the geographical origin of the samples. For the four close related species *R. marabaensis, R. montenegrensis, R. robustus,* and *R. prolixus*, we proposed to use the terminology of "*robustus* species complex."

In order to corroborate the inference about the *R. robustus* clades defined using *cytb* sequences (Monteiro et al., 2003), despite we demonstrated that this gene was under selective pressure and that *Rhodnius* phylogeny is not well resolved using this marker, we nonetheless performed an ML phylogeny including the *R. robustus* clade sequences and also cytb sequences from *R. barretti* (**Figure 7**). In the *cytb* phylogeny: (i) *R. montenegrensis* R8F was robustly included in the ROB II clade, (ii) ROBB in the ROB III clade assimilated to *R. marabaensis* and (iii) *R. robustus* Rob5s in the ROB IV clade.

Note that the new species *R. barretti* described from both Colombia and Ecuador and morphologically close to *R. robustus* (Abad-Franch et al., 2013), showed on the *cytb* phylogeny in Monteiro et al. (2018) a basal position to *R. neglectus, R. nasutus,* and the *robustus* species complex. In the *cytb* phylogeny (**Figure 7**), the NASG has a position similar to *R. barretti*. If only mitochondrial genes were analyzed, due to the high divergence observed, the proposal of a specific status for this specimen could arise. But the nuclear data are unequivocal, NasG is relative to *R. nasutus* but having a mitochondrial introgression likely by DNA from a species of the *robustus* complex (see below). If the position of *R. barretti* in the *cytb* is irrefutable, more molecular data and especially nuclear ones are needed to discard putative introgression and definitively assess its phylogenetic position.

Discordance Between Mitochondrial and Nuclear Phylogenomic Inferences

If the mitochondrial genes used for phylogenetics suffer from disadvantages, discordance between mitochondrial and nuclear phylogenies could highlight evolutionary processes and be a potential source of additional information about the evolutionary history of a recent clade. Such mito-nuclear discordance is not rare and was reported in numerous taxa. Incomplete lineage sorting (ILS) could explain this discordance by the conservation of ancestral polymorphisms in multiple lineages after species splits (Degnan and Rosenberg, 2009). However, polymorphisms in mtDNA are expected to be lost relatively rapidly through genetic drift because of the reduced effective population size in mitochondrial genomes. This leads ILS to unlikely be the predominant cause of mito-nuclear discordance, notably when a large number of nuclear loci are congruent in phylogenies but in conflict with mitochondrial ones (Good et al., 2015). Another more likely cause of discordance could be introgression, highlighted between otherwise well-defined species by differentially affecting nuclear vs. mitochondrial DNA. In many cases, mitochondrial introgression occurs with few or no detectable movement of nuclear genes (Sloan et al., 2017). Nonetheless, evidence for mitochondrial introgression with accompanying nuclear genes usually selected against in the recipient species was described (Rubinoff and Holland, 2005; Matute et al., 2020). A least plausible explanation is hybrid speciation, though more common in nature than previously thought (Mavárez and Linares, 2008; Schumer et al., 2014; Payseur and Rieseberg, 2016). In this study, for the mitochondrial genes, misleading phylogenetic reconstructions could be due to sequence convergence as we found the occurrence of selective pressure on these genes, but also to introgression, with mitochondrial introgression being the most likely process rather than the nuclear one (Good et al., 2015).

The *pictipes* Group

In our study, we aim to test signatures of introgression using the ABBA-BABA test (Green et al., 2010) for the *pictipes* group for which a discordance between the mitochondrial and nuclear PCG trees was noticed. However, the mapping on *R. prolixus,*

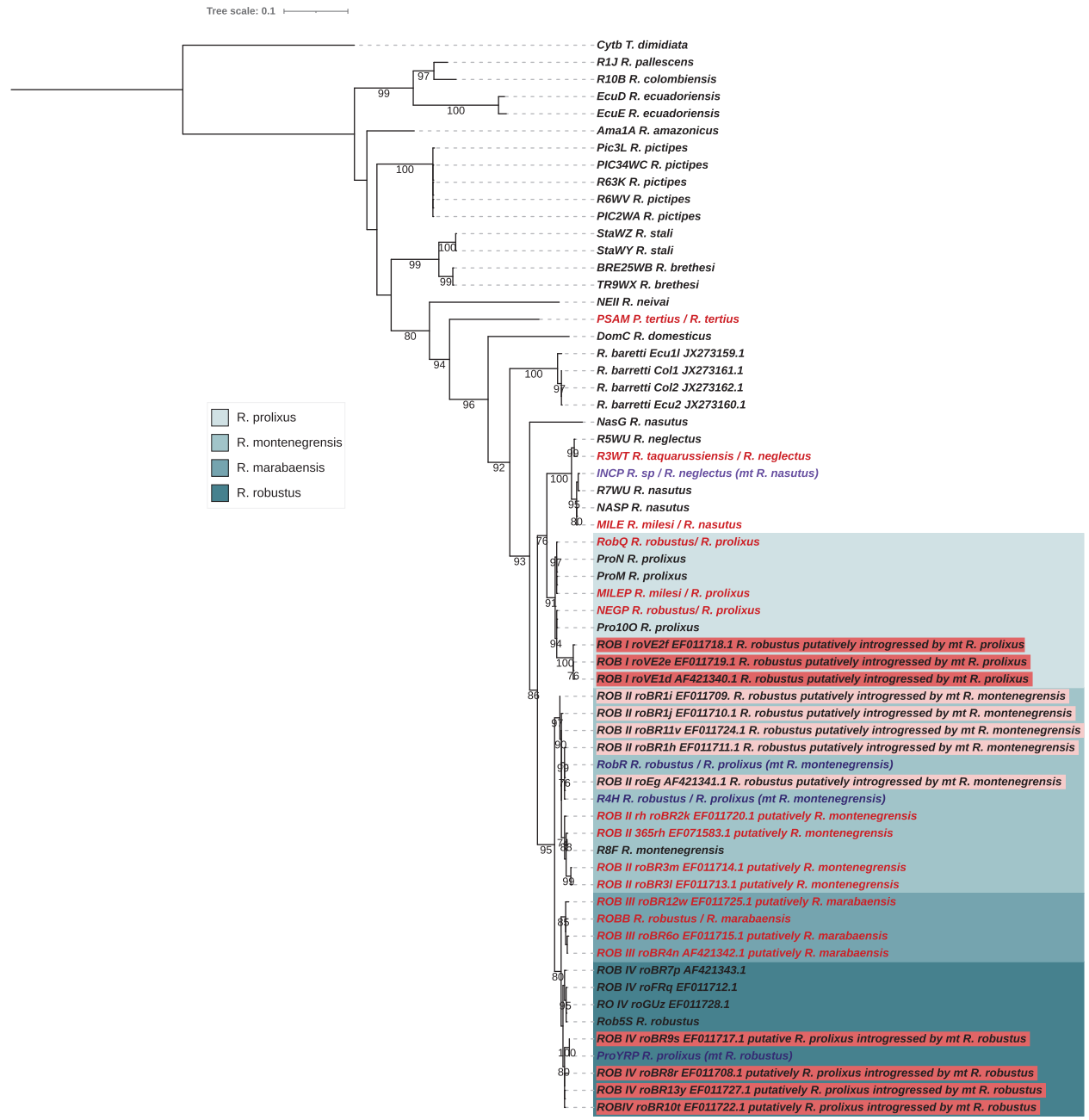

FIGURE 7 | Phylogenetic inferences and interpretation for the *Cytb* gene for *Rhodnius* species with the Maximum Likelihood (ML) method including Monteiro et al. (2003) samples. The statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *T. dimidiata* was used as outgroup. For the dataset of this study, introgressed species inferred from the nuclear phylogeny (see **Figure 5**) are indicated in purple, misidentified species in red. For Monteiro et al. (2003) samples, based on the inferences made for our samples, we proposed putative introgressed species, highlighted in the tree and putative misidentifications for some others, indicated in red. For ROB I clade clustered with *R. prolixus* samples and composed of Venezuelan *R. robustus* samples, they putatively are *R. robustus* introgressed by mitochondrial DNA from *R. prolixus*. For ROB II clade, some samples clustered with the *R. montenegrensis* could be synonymized with *R. montenegrensis* while the other could be *R. prolixus* introgressed by mitochondrial DNA from *R. montenegrensis*. For the ROB III clade, specimens clustered with *R. marabaensis* and could be synonymize as *R. marabaensis*. For the ROB IV clade, some samples are clustered with *R. robustus* while others could be *R. prolixus* introgressed by mitochondrial DNA from *R. robustus*.

the only reference genome available for the *Rhodnius* species, could be a limitation for this study. Withal, the genome-wide analysis supports in the *pictipes* group the hypothesis of a mitochondrial introgression between *R. brethesi* and *R. stali* as the source of incongruence. Based on the nuclear phylogenies, we can suggest an introgression of *R. stali* with mitochondrial DNA from *R. brethesi*. This assumption is supported by the present distribution of the two species, *R. stali* being further

south (Bolivia, Brazil: Mato Grosso and Acre) than *R. brethesi* (Colombia, Venezuela, Brazil: Pará and Amazonas) and also with the morphological similarity of the two species. However, an unsampled (or undiscovered) species of the *pictipes* group as the source of the introgressed mitogenome cannot be excluded. The recent discovery of a new species *R. micki* (Zhao et al., 2021) from Bolivia could be a putative candidate as the source of mitochondrial introgression since this species is morphologically similar to *R. pictipes* and *R. stali* but *R. brethesi* was not included in the morphometric study and molecular data from *R. micki* is still lacking. A selective sweep of mitochondrial DNA linked to an expending endosymbiont could also be considered. In this way, a study of *Wolbachia* genomic data (Filée et al., unpublished) showed that the same *Wolbachia* strain is shared by species of the *pictipes* group (*R. brethesi, R. amazonicus, R. stali,* and *R. pictipes)* but that it differs from that of the *prolixus* group (*R. robustus, R. prolixus, R. neglectus,* and *R. nasutus*). Indeed, this endosymbiont does not seem to be involved in this mechanism.

The *pallescens* Group

For the *pallescens* group, we tested signatures of introgression because if the mitochondrial and nu-PCG trees are congruent showing *R. colombiensis* and *R. pallescens* clustered together, in the rDNA phylogeny, *R. pallescens* and *R. colombiensis* are closer. Such divergent topologies were already obtained in the literature, indicating either a close relationship between *R. pallescens* and *R. colombiensis* (Hypša et al., 2002) or between *R. ecuadoriensis and R. colombiensis* (Justi et al., 2016). However, the first topology is coherent with the vector distribution since the Chocó rainforests along the Colombian Pacific coast is assumed to separate *R. ecuadoriensis* from *R. pallescens–R. colombiensis* (Abad-Franch et al., 2009). Although the tests were carried out using *R. prolixus* for the mapping, three hypotheses could explain the numerous SNPs supporting either each grouping (ABBA- or BABA distribution): Incomplete Lineage Sorting (ILS), hybrid speciation and introgression. ILS could be the result of a fast radiation of the three species that led to the sharing of ancestral polymorphisms. However, this hypothesis cannot explain alone the under-representation of the ABBA configuration, i.e., less allele sharing between *R. ecuadoriensis* and *R. colombiensis* than between *R. pallescens* and *R. colombiensis*. A least conventional yet plausible explanation is hybrid speciation. The ambiguous positioning of *R. pallescens* as closer to either *ecuadoriensis* or *colombiensis* could be the result of crosses between *R. ecuadoriensis* males and *R. colombiensis* females leading to the third species *R. pallescens*, which will reflect the mitochondrial tree topology. But hybrid speciation remains to be demonstrated in *Rhodniu*s. Introgression, which results from the occurrence of gene flow after speciation, could explain the ABBA under-representation if *R. ecuadoriensis* and *R. colombiensis* were genetically isolated but each still hybridizing with *R. pallescens*. Indeed, Díaz et al. (2014) demonstrated in the laboratory the crossbreeding between *R. pallescens* and *R. colombiensis* although the laboratory hybrids were infertile. We did not observe ABBA or BABA configurations restricted to specific genomic regions, the three configurations were widespread and overlapping on all scaffolds of the reference genome of *R. prolixus* used for the read mapping. Nonetheless, we cannot rule out the possibility that multiple rearrangements have taken place since the divergence of the three species and *R. prolixus*. At least, nuclear *R. ecuadoriensis* introgression in *R. pallescens* could explain the ABBA under-representation since the nuclear rDNA showed a closest genetic proximity between these two species. Therefore, genomic studies with large population sampling in the *R. pallescens* species group wherein different demographic models could be tested are strongly needed in the future.

The *prolixus* Group

In this study, the robust nu-PCG phylogeny splits the *prolixus* group into two major clades, the *neglectus-nasutus* clade and the *robustus* species complex clade (*R. marabaensis, R. montenegrensis, R. robustus,* and *R. prolixus*). In these two clades, we observed several discrepancies between mitochondrial and nuclear trees.

The wild Piauí-Brazil NasG specimen showed an incongruent position between the mitochondrial and nuclear PCG trees. This specimen showed a basal position for the two clades in the mitochondrial tree but it is clustered with the *R. neglectus* samples in the nu-rDNA tree and with R7WU *R. nasutus* in the nu-PCGs tree. Considering both its morphology is similar to *R. nasutus* and the substantial data for the nu-PCG dataset (36.3 kb), the mitochondrial introgression is the most likely process rather than the nuclear one. The nature of the mitochondrial introgression remains speculative. Unexpectedly, an introgression with a member of the *robustus* complex species cannot be ruled out. Moreover, some nuclear introgression for this *R. nasutus* sample could be hypothesized probably by rDNA from *R. neglectus.*

Mitochondrial introgression was also observed for two other field specimens from Piauí. The INCP sample exhibited a dark morph and an atypical phenotype which cannot be assigned to *R. nasutus* or *R. neglectus*. The molecular study evidenced that this sample is related to *R. neglectus* in the nuclear tree but to *R. nasutus* in the mitochondrial one, thus INCP could be considered as a *R. neglectus* introgressed by mitochondrial DNA from *R. nasutus*. The reverse introgression is exhibited by the NASP sample, that is a *R. nasutus* introgressed by mitochondrial DNA from *R. neglectus*. Introgressions between *R. nasutus* and *R. neglectus* was not yet reported, including from field specimens but are not surprising due to the close genetic proximity between the two species as revealed by genetic distance values but also to the geographic distribution of the species. Results obtained by Abad-Franch et al. (2009) *via* geometric morphometrics showed that *R. neglectus* and *R. nasutus* are sympatric in the Brazilian Cerrado-Caatinga transitional area. There is also some evidence of populations of *R. neglectus* occurring also in the true Caatinga biome where *R. nasutus* is distributed (Dias et al., 2011), which could favor hybridization. Moreover, chromatic variation in *R. nasutus* was already observed and adaptive phenotypic plasticity was suggested related to microhabitat features (Abad-Franch et al., 2009). It will be of interest using molecular markers to test the various *R. nasutus* chromatic forms for introgression with *R. neglectus*, that could explain the darker morph observed.

For the *robustus* species complex, three field specimens NEGP, RobQ and MILEP, all originated from Pará Brazil,

and morphologically identified as *R. neglectus, R. robustus* and *R. milesi* respectively, a misidentification is likely since all the trees (mtG, rDNA, nu-PCGs) are congruent with *R. prolixus* species. This exemplifies the difficulty to identify the species of the *prolixus* group and to admit the occurrence of sylvatic *R. prolixus*. If the occurrence of *R. prolixus* in Venezuela is now well documented (i.e., Fitzpatrick et al., 2008), in Brazil the controversy persists but it is actually accepted that *R. prolixus* occurs basically in the states of Amazonas and Pará (Pinho et al., 1998). Indeed, the discovery of *R. neglectus* in this region (Lent, 1954) but also the widely distribution of *R. robustus* and the description of *R. milesi* from Pará cast doubt on past identifications of *R. prolixus* and guided the future ones. By instance, it was suggested that *R. neglectus*, and not *R. prolixus*, was the species invading houses in central Brazil (Gurgel-Gonçalves et al., 2008). In our study, we then evidence by molecular data the presence of sylvatic *R. prolixus* in the Brazilian Pará state.

For two strain samples, R4H and RobR, a mito-nuclear conflict was evidenced. In the nuclear tree, these two samples were clustered with *R. prolixus* samples but in the mitochondrial tree, they were closely related to *R. montenegrensis* and likely have experienced mitochondrial introgression with this species. In addition, for the RobR sample, the rDNA tree suggested an introgression by *R. robustus*. The domiciliary Trujillo-Venezuelan *R. prolixus* specimen (ProYRP) was also introgressed by *R. robustus* but for mitochondrial DNA. Similar results were obtained for Venezuelan populations by Fitzpatrick et al. (2008) using both the mitochondrial *cytb* gene and the *D2* nuclear gene, a population of domiciliated *R. prolixus* identified as *R. robustus* was introgressed with mitochondrial DNA from *R. robustus* (Hap3, **Figure 7**). Introgression events of *R. prolixus* with *R. robustus* in Venezuela could explain the results obtained by Harry (1994) with morphometric data in which no clear-cut differentiation was evidenced between Venezuelan *R. prolixus* and *R. robustus* samples from Trujillo.

Considering the *cytb* phylogeny including Monteiro et al. (2003) samples, the apparent paraphyly of *R. robustus* pointed out by the authors is resolved by the consideration of introgression events (**Figure 7**):

(i) ROB I clade is clustered with *R. prolixus* samples and is composed of Venezuelan *R. robustus* samples. Due to the widespread of introgressions between *R. robustus* and *R. prolixus* especially in this region, we suggested that samples from ROB I clade are *R. robustus* introgressed by mitochondrial DNA from *R. prolixus*, but nuclear markers are needed to affirm or refute this hypothesis;

(ii) ROB II clade is split into two. One comprises samples clustered with the *R. montenegrensis* strain of reference (R8F) while the other comprises samples clustered with two of our samples (RobR, R4H) that we demonstrated the introgression of *R. prolixus* by mitochondrial DNA from *R. montenegrensis*. Then, we putatively suggest that field specimens from Carauari, Amazonas (RoBR1, haplotypes h, i, j) are introgressed, and only a part of ROB II could be synonymized with *R. montenegrensis*.

(iii) specimens of the ROB III clade clustered with ROBB from Marabá and assimilated to *R. marabaensis*. So, we conclude that this clade could be synonymize as *R. marabaensis* has used by Castro et al. (2020).

(iv) ROB IV clade is also split into two. One comprises samples including *R. robustus* from French Guyana and clustered with also our sample from this region. The same clustering was found by Barnabé et al. (2018) for an extensive survey of *R. robustus* from French Guyana. The other comprises samples clustered with our Venezuelan sample ProYRP for which we demonstrated the introgression of *R. prolixus* by mitochondrial DNA from *R. robustus* and the Hap3 also a *R. prolixus* introgressed by mitochondrial DNA from *R. robustus* and sharing the same haplotype than the roBR9 (Fitzpatrick et al., 2008). Then, we suggested that this part of samples from ROB IV clade are *R. prolixus* introgressed by mitochondrial DNA from *R. robustus*.

In this study, we demonstrated that introgression could be a biological event as field samples were introgressed. Misplacement in phylogeny for some *Rhodnius* lab strains could also result from a past biological event and not only be the fact of mixed or misidentified samples as pointed out by Brito et al. (2019). Indeed, at this time, it is impossible to state if lab strains with such pattern were accidentally introgressed or were founded with specimens naturally introgressed.

CONCLUSION

In this study, we provide a very large dataset on *Rhodnius* species including 36 samples. With low-depth whole-genome shotgun sequencing, we resolve sticking points in the *Rhodnius* phylogeny using phylogenomic approaches based on 15 mtG (13.3 kb), nu-rDNA genes (5.2 kb), and 51 nu-PCGs (36.5 kb).

Out the 17 putative species determined on phenotype, using molecular data, we identified 16 valid *Rhodnius* species (**Table 1**). We confirmed the species-specific status of *R. montenegrensis* and *R. marabaensis* and we agree with the synonymy of *R. taquarussuensis* with *R. neglectus*. We also invite to revisit the species-specific status of *R. milesi* that is more likely *R. nasutus*. We added *R. marabaensis* in our dataset, first identified as *R. robustus*. We propose the *robustus* complex species for including the four close related species *R. marabaensis, R. montenegrensis, R. prolixus,* and *R. robustus*. Moreover, we confirm the position of *P. tertius* into the *Rhodnius* genus and strongly recommended reclassifying *P. tertius* as *R. tertius*. For the newly described species *R. micki* and *R. barretti*, molecular data and especially nuclear ones are needed to assess their phylogenetic position. The same is true for *R. paraensis* rarely sampled since this species has been collected only from arboreal rodent nests in the Amazon region in Brazil (Sherlock et al., 1977) but also with light trap in French Guyana (Bérenger et al., 2009). For *R. zeledoni* and *R. dalessandroi* only known from their description, no molecular data can be obtained.

Both mitochondrial and nuclear data coherently support that the *pictipes* and *pallescens* groups are more related to each other

than they are to the *prolixus* group. We can hypothesize that the *pictipes* and *pallescens* groups more likely diverged from an ancestral form when the Northern Andean uplift before the formation of the Pebas system around 23–10 MYA and that the latter event has induced the diversification of the *prolixus* group when this system started to be drained eastward into the Atlantic Ocean. This subsequently led to hydrological changes promoting allopatric speciation or favoring dispersals (Albert et al., 2018). However, robust time-calibrated phylogenies are now needed to infer the historical and spatial origin of the *Rhodnius* lineages.

By comparing the topologies obtained for mitochondrial and nuclear phylogenies but also by performing genome-wide analysis, our results demonstrated unexpected introgression events in all the different *Rhodnius* groups, in laboratory strains but also in wild specimens. In the *pictipes* group, we formulated the hypothesis of cytoplasmic introgression between *R. brethesi* and *R. stali* and in the *pallescens* group the occurrence of complex introgression. In the *prolixus* group introgression occurred between various species: (i) from *R. robustus* or *R. montenegrensis* mitochondrial DNA to *R. prolixus*,(ii) from *R. robustus* mitochondrial DNA to *R. prolixus*, (iii) from *R. nasutus* mitochondrial DNA to *R. neglectus,* (iv) the reverse introgression from *R. neglectus* mitochondrial DNA from to *R. nasutus* and (v) putatively from one species of the *robustus* complex to *R. nasutus.* The taking into account the introgression events makes it possible to explain the paraphyly observed for *R. robustus* in some mitochondrial phylogenies but leads to abolish it by considering nuclear genes and restores the monophyly for this species. The extensive introgression especially in the *prolixus* group does not question the species boundaries for the pairwise *R. nasutus* and *R. neglectus* nor for the species of the *robustus* complex, and even that complicates the identification of the species, it is challenging from an evolutionary point of view. Further studies are needed in order to date the introgressions found in field samples which can result from old events at the time of the speciation or from recent hybridization between sympatric species.

This study exemplifies that the evolutionary inferences from only mitochondrial markers notably by using only the *cytb* gene or incomplete dataset could be misleading. Because of the persuasive introgression which confuse not only the molecular issue but surely also the phenotypic one, we strongly recommend the use of both mitochondrial and nuclear markers prior any study on *Rhodnius* species and to perform separately mitochondrial and nuclear phylogenies in order to take advantages from mito-nuclear conflicts for having a comprehensive evolutionary vision.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI (accession: PRJNA429761).

AUTHOR CONTRIBUTIONS

JF, MM, and HB: data analysis and writing. FM and EF-R: study design. CA: writing. J-MB: field collections and *Rhodnius* taxonomy. MH: study conception, DNA extractions, data analysis, and writing. All authors edited the manuscript and approved the submitted version.

FUNDING

This study was funded by the RADIANT project (Labex BASC, University Paris Saclay, France), the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process number 2016/08176-9, 2017/50329-0), and the Fundação de Apoio à Pesquisa do Estado da Paraíba – FAPESQ (process 47896.673.31653.11082021). Financial support had neither role in study design, collection, analysis and interpretation of data nor manuscript writing.

ACKNOWLEDGMENTS

We would like to thank all our collaborators who provided us with some of the samples used. For the CTA strains: Joao Aristeu da Rosa and Jader Olivieira (Universidade Estadual Paulista (Unesp), Faculdade de Ciências Farmacêuticas, Araraquara, São Paulo, Brazil); for some Guyanese *R. robustus*: Denis Blanchet (Laboratoire Hospitalier Universitaire de Parasitologie et Mycologie, Université des Antilles et de la Guyane); for Venezuelan fieldtrip and lab strains: colleagues from the Jose Vincente Scorza's team (Universidad of Los Andes, Trujillo, Venezuela); for *R. pallescens* and Brazilian Marabá and Pará specimens: Christina Romana (Université Paris Descartes); for Bolivian samples: our late colleague François Noireau; and for *R. prolixus* strain from Honduras: Claudio Lazzari (Université de Tours, IRBI, France). We would like also to thank Morgane Lavina and Claire Capdevielle-Dulac for their help in molecular biology experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2021.750317/full#supplementary-material

REFERENCES

Abad-Franch, F., Monteiro, F. A., Jaramillo, N., Gurgel-Gonçalves, R., Dias, F. B. S., and Diotaiuti, L. (2009). Ecology, evolution, and the long-term surveillance of vector-borne Chagas disease: a multi-scale appraisal of the tribe Rhodniini (Triatominae). *Acta Trop.* 110, 159–177. doi: 10.1016/j.actatropica.2008.06.005

Abad-Franch, F., Pavan, M. G., Jaramillo-O, N., Palomeque, F. S., Dale, C., Chaverra, D., et al. (2013). *Rhodnius barretti*, a new species of Triatominae (Hemiptera: Reduviidae) from western Amazonia. *Mem. Instit. Oswaldo Cruz* 108, 92–99. doi: 10.1590/0074-0276130434

Albert, J. S., Val, P., and Hoorn, C. (2018). The changing course of the amazon river in the neogene: center stage for neotropical diversification. *Neotrop. Ichthyol.* 16:e180033.

Almeida, F. B. D., Santos, E. I., and Sposina, G. (1973). Triatomíneos da Amazonia III. *Acta Amazon.* 3, 43–46.

Antonio-Campos, A., Nicolás-Cruz, A., Girón-Arias, J. I., Rivas, N., and Alejandre-Aguilar, R. (2019). Presence of *Rhodnius prolixus* Stäl, 1859 (Hemiptera: Reduviidae) in Oaxaca, Mexico, ten years after the certification of its elimination. *J. Vector Ecol.* 44, 293–295. doi: 10.1111/jvec.12363

Barnabé, C., Brenière, S. F., Guégan, J. F., Blanchet, D., and Aznar, C. (2018). Molecular characterization of *Rhodnius robustus* specimens, potential vectors for Chagas disease in French Guiana, South America. *Inf. Gen. Evol.* 59, 28–31. doi: 10.1016/j.meegid.2018.01.019

Bérenger, J. M., and Pluot-Sigwalt, D. (2002). *Rhodnius amazonicus* Almeida, Santos & Sposina, 1973, bona species, close to *R. pictipes* Stål, 1872 (Heteroptera, Reduviidae, Triatominae). *Mem. Instit. Oswaldo Cruz* 97, 73–77. doi: 10.1590/s0074-02762002000100011

Bérenger, J. M., Pluot-Sigwalt, D., Blanchet, D., and Aznar, C. (2009). The Triatominae species of French Guiana (Heteroptera: Reduviidae). *Mem. Instit. Oswaldo Cruz*. 104, 1111–1116. doi: 10.1590/S0074-02762009000800007

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., et al. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phyl. Evol.* 69, 313–319. doi: 10.1016/j.ympev.2012.08.023

Brito, R. N., Geraldo, J. A., Monteiro, F. A., Lazoski, C., Souza, R. C., and Abad-Franch, F. (2019). Transcriptome-based molecular systematics: *Rhodnius montenegrensis* (Triatominae) and its position within the *Rhodnius prolixus–Rhodnius robustus* cryptic–species complex. *Parasit. Vectors* 12, 1–16. doi: 10.1186/s13071-019-3558-9

Campillo, L. C., Burns, K. J., Moyle, R. G., and Manthey, J. D. (2019). Mitochondrial genomes of the bird genus *Piranga*: rates of sequence evolution, and discordance between mitochondrial and nuclear markers. *Mitochondrial DNA Part B* 4, 2566–2569. doi: 10.1080/23802359.2019.1637286

Carcavallo, R., and Barreto, P. (1976). A new species of *Rhodnius* in Colombia. *Bol. Dir. Malariol. San. Amb.* 16, 176–183.

Castro, M. R., Goubert, C., Monteiro, F. A., Vieira, C., and Carareto, C. (2020). Homology-free detection of transposable elements unveils their dynamics in three ecologically distinct *Rhodnius* species. *Genes* 11:170. doi: 10.3390/genes11020170

Chikhi, R., and Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics* 30, 31–37. doi: 10.1093/bioinformatics/btt310

Degnan, J. H., and Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332–340. doi: 10.1016/j.tree.2009.01.009

Dias, F. B. S., de Paula, A. S., Belisario, C. J., Lorenzo, M. G., Bezerra, C. M., Harry, M., et al. (2011). Influence of the palm tree species on the variability of *Rhodnius nasutus* Stål, 1859 (Hemiptera, Reduviidae, Triatominae). *Infect. Gen. Evol.* 11, 869–877. doi: 10.1016/j.meegid.2011.02.008

Díaz, S., Panzera, F., Jaramillo-O, N., Pérez, R., Fernández, R., Vallejo, G., et al. (2014). Genetic, cytogenetic and morphological trends in the evolution of the *Rhodnius* (Triatominae: Rhodniini) trans-Andean group. *PLoS One* 9:e87493. doi: 10.1371/journal.pone.0087493

Dotson, E. M., and Beard, C. (2001). Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. *Insect Mol. Biol.* 10, 205–215. doi: 10.1046/j.1365-2583.2001.00258.x

Fitzpatrick, S., Feliciangeli, M. D., Sanchez-Martin, M. J., Monteiro, F. A., and Miles, M. A. (2008). Molecular genetics reveal that silvatic *Rhodnius prolixus* do colonise rural houses. *PLoS Neg. Trop. Dis.* 2:e210. doi: 10.1371/journal.pntd.0000210

Galvão, C., and Justi, S. A. (2015). An overview on the ecology of Triatominae (Hemiptera: Reduviidae). *Acta Trop.* 151, 116–125. doi: 10.1016/j.actatropica.2015.06.006

Gao, F., Chen, C., Arab, D. A., Du, Z., He, Y., and Ho, S. Y. (2019). EasyCodeML: a visual tool for analysis of selection using CodeML. *Ecol. Evol.* 9, 3891–3898. doi: 10.1002/ece3.5015

Garvin, M. R., Bielawski, J. P., Sazanov, L. A., and Gharrett, A. J. (2015). Review and meta-analysis of natural selection in mitochondrial complex I in metazoans. *J. Zool. Syst. Evol. Res.* 53, 1–17. doi: 10.1111/jzs.12079

Good, J. M., Vanderpool, D., Keeble, S., and Bi, K. (2015). Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution* 69, 1961–1972. doi: 10.1111/evo.12712

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., et al. (2010). A draft sequence of the Neandertal genome. *Science* 328, 710–722. doi: 10.1126/science.1188021

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Gurgel-Gonçalves, R., Abad-Franch, F., Ferreira, J. B., Santana, D. B., and Cuba, C. A. C. (2008). Is *Rhodnius prolixus* (Triatominae) invading houses in central Brazil? *Acta Trop.* 107, 90–98. doi: 10.1016/j.actatropica.2008.04.020

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Harrison, R. G. (1989). Mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* 4, 6–11. doi: 10.1016/0169-5347(89)90006-2

Harry, M. (1993). Isozymic data question the species-specific status of some blood sucking bugs of the genus *Rhodnius*, vectors of Chagas' disease. *Trans. R. Soc. Trop. Med. Hyg.* 87:492. doi: 10.1016/0035-9203(93)90054-t

Harry, M. (1994). Morphometric variability in the Chagas' disease vector *Rhodnius prolixus*. *Jpn. J. Genet.* 69, 233–250. doi: 10.1266/jjg.69.233

Harry, M., Galindez, I., and Cariou, M. L. (1992). Isozyme variability and differentiation between *Rhodnius prolixus, R. robustus* and *R. pictipes*, vectors of Chagas disease in Venezuela. *Med. Vet. Entomol.* 6, 37–43. doi: 10.1111/j.1365-2915.1992.tb00032.x

Hernández, C., da Rosa, J. A., Vallejo, G. A., Guhl, F., and Ramírez, J. D. (2020). Taxonomy, evolution, and biogeography of the Rhodniini tribe (Hemiptera: Reduviidae). *Diversity* 12:97.

Hypša, V., Tietz, D. F., Zrzavı, J., Rego, R. O. M., Galvao, C., and Jurberg, J. (2002). Phylogeny and biogeography of Triatominae (Hemiptera: Reduviidae): molecular evidence of a New World origin of the Asiatic clade. *Mol. Phyl. Evol.* 23, 447–457. doi: 10.1016/s1055-7903(02)00023-4

Jurberg, J., Rocha, D. D. S., and Galvão, C. (2009). *Rhodnius zeledoni* sp. nov. afim de *Rhodnius paraensis* Sherlock, Guitton & Miles, 1977 (Hemiptera, Reduviidae, Triatominae). *Biota Neotrop.* 9, 123–128.

Justi, S. A., Galvao, C., and Schrago, C. G. (2016). Geological changes of the Americas and their influence on the diversification of the Neotropical kissing bugs (Hemiptera: Reduviidae: Triatominae). *PLoS Negl. Trop. Dis.* 10:e0004527. doi: 10.1371/journal.pntd.0004527

Justi, S. A., Russo, C. A., dos Santos Mallet, J. R., Obara, M. T., and Galvão, C. (2014). Molecular phylogeny of Triatomini (Hemiptera: Reduviidae: Triatominae). *Parasit. Vectors* 7, 1–12. doi: 10.1186/1756-3305-7-149

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Kieran, T. J., Gordon, E. R., Zaldívar-Riverón, A., Ibarra-Cerdeña, C. N., Glenn, T. C., and Weirauch, C. (2021). Ultraconserved elements reconstruct the evolution of Chagas disease-vectoring kissing bugs (Reduviidae: Triatominae). *Syst. Entomol.* 46, 725–740. doi: 10.1111/syen.12485

Kocher, A., Kamilari, M., Lhuillier, E., Coissac, E., Péneau, J., Chave, J., et al. (2014). Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial

genome (Heteroptera, Reduviidae). *Gene* 552, 184–194. doi: 10.1016/j.gene.2014.09.033

Kofler, R., Pandey, R. V., and Schlötterer, C. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27, 3435–3436. doi: 10.1093/bioinformatics/btr589

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35:1547. doi: 10.1093/molbev/msy096

Lanfear, R., Calcott, B., Ho, S. Y., and Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701. doi: 10.1093/molbev/mss020

Lent, H., and Wygodzinsky, P. (1979). Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas' disease. *Bull. Am. Mus. Nat. Hist.* 163, 123–520.

Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lin, C. P., and Danforth, B. N. (2004). How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. *Mol. Phyl. Evol.* 30, 686–702. doi: 10.1016/S1055-7903(03)00241-0

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18.

Lyman, D. F., Monteiro, F. A., Escalante, A. A., Cordon-Rosales, C., Wesson, D. M., Dujardin, J. P., et al. (1999). Mitochondrial DNA sequence variation among triatomine vectors of Chagas' disease. *Am. J. Trop. Med. Hyg.* 60, 377–386. doi: 10.4269/ajtmh.1999.60.377

Maia da Silva, F., Junqueira, A. C. V., Campaner, M., Rodrigues, A. C., Crisante, G., Ramirez, L. E., et al. (2007). Comparative phylogeography of *Trypanosoma rangeli* and *Rhodnius* (Hemiptera: Reduviidae) supports a long coexistence of parasite lineages and their sympatric vectors. *Mol. Ecol.* 16, 3361–3373. doi: 10.1111/j.1365-294X.2007.03371.x

Marchant, A., Mougel, F., Almeida, C., Jacquin-Joly, E., Costa, J., and Harry, M. (2015). De novo transcriptome assembly for a non-model species, the blood-sucking bug *Triatoma brasiliensis*, a vector of Chagas disease. *Genetica* 143, 225–239. doi: 10.1007/s10709-014-9790-5

Matute, D. R., Comeault, A. A., Earley, E., Serrato-Capuchina, A., Peede, D., Monroy-Eklund, A., et al. (2020). Rapid and predictable evolution of admixed populations between two *Drosophila* species pairs. *Genetics* 214, 211–230. doi: 10.1534/genetics.119.302685

Mavárez, J., and Linares, M. (2008). Homoploid hybrid speciation in animals. *Mol. Ecol.* 17, 4181–4185.

McDonald, J. H., and Kreitman, M. (1991). Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351, 652–654. doi: 10.1038/351652a0

Meng, G., Li, Y., Yang, C., and Liu, S. (2019). MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Res.* 47:e63. doi: 10.1093/nar/gkz173

Mesquita, R. D., Vionette-Amaral, R. J., Lowenberger, C., Rivera-Pomar, R., Monteiro, F. A., Minx, P., et al. (2015). Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14936–14941. doi: 10.1073/pnas.1506226112

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015

Monteiro, F. A., Barrett, T. V., Fitzpatrick, S., Cordon-Rosales, C., Feliciangeli, D., and Beard, C. B. (2003). Molecular phylogeography of the Amazonian Chagas disease vectors *Rhodnius prolixus* and *R. robustus*. *Mol. Ecol.* 12, 997–1006. doi: 10.1046/j.1365-294x.2003.01802.x

Monteiro, F. A., Weirauch, C., Felix, M., Lazoski, C., and Abad-Franch, F. (2018). Evolution, systematics, and biogeography of the Triatominae, vectors of Chagas disease. *Adv. Parasitol.* 99, 265–344. doi: 10.1016/bs.apar.2017.12.002

Monteiro, F. A., Wesson, D. M., Dotson, E. M., Schofield, C. J., and Beard, C. B. (2000). Phylogeny and molecular taxonomy of the Rhodniini derived from mitochondrial and nuclear DNA sequences. *Am. J. Trop. Med. Hyg.* 62, 460–465. doi: 10.4269/ajtmh.2000.62.460

Nascimento, J. D., Da Rosa, J. A., Salgado-Roa, F. C., Hernández, C., Pardo-Diaz, C., Alevi, K. C. C., et al. (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:e0211285. doi: 10.1371/journal.pone.0211285

Oliveira, J., Alevi, K. C. C., Ravazi, A., Herrera, H. M., Santos, F. M., de Azeredo-Oliveira, M. T. V., et al. (2018). New evidence of the monophyletic relationship of the genus *Psammolestes* Bergroth, 1911 (Hemiptera, Reduviidae, Triatominae). *Am. J. Trop. Med. Hyg.* 99:1485. doi: 10.4269/ajtmh.18-0109

PAHO (2020). *Chagas Disease*. Available online at: https://www.paho.org/en/topics/chagas-disease (accessed September 6, 2020)

Paula, A. S., Barreto, C., Telmo, M. C. M., Diotaiuti, L., and Galvão, C. (2021). Historical biogeography and the evolution of hematophagy in Rhodniini (Heteroptera: Reduviidae: Triatominae). *Front. Ecol. Evol.* 9:660151. doi: 10.3389/fevo.2021.660151

Paula, A. S., Diotaiuti, L., and Galvao, C. (2007). Systematics and biogeography of Rhodniini (Heteroptera: Reduviidae: Triatominae) based on 16S mitochondrial rDNA sequences. *J. Biogeogr.* 34, 699–712.

Paula, A. S., Diotaiuti, L., and Schofield, C. J. (2005). Testing the sister-group relationship of the Rhodniini and Triatomini (Insecta: Hemiptera: Reduviidae: Triatominae). *Mol. Phyl. Evol.* 35, 712–718. doi: 10.1016/j.ympev.2005.03.003

Payseur, B. A., and Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Mol. Ecol.* 25, 2337–2360. doi: 10.1111/mec.13557

Pinho, A. P., Gonçalves, T. C. M., Mangia, R. H., Russell, N. S. N., and Jansen, A. M. (1998). The occurrence of *Rhodnius prolixus* Stal, 1859, naturally infected by *Trypanosoma cruzi* in the State of Rio de Janeiro, Brazil (Hemiptera, Reduviidae, Triatominae). *Mem. Instit. Oswaldo Cruz* 93, 141–143. doi: 10.1590/s0074-02761998000200001

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/sys029

Rosa, J. A., Justino, H. H. G., Nascimento, J. D., Mendonça, V. J., Rocha, C. S., de Carvalho, D. B., et al. (2017a). A new species of *Rhodnius* from Brazil (Hemiptera, reduviidae, triatominae). *ZooKeys* 675, 1–25.

Rosa, J. A., Souza, E. S., da Costa Teixeira, A., Barbosa, R. R., de Souza, A. J., Belintani, T., et al. (2017b). Third record of *Rhodnius amazonicus* and comparative study with *R. pictipes* (Hemiptera, Reduviidae, Triatominae). *Acta Trop.* 176, 364–372. doi: 10.1016/j.actatropica.2017.09.003

Rosa, J. A., Rocha, C. S., Gardim, S., Pinto, M. C., Mendonca, V. J., Filho, J. C. R. F., et al. (2012). Description of *Rhodnius montenegrensis* n. sp. (Hemiptera: Reduviidae: Triatominae) from the state of Rondônia, Brazil. *Zootaxa* 3478, 62–76.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248

Rubinoff, D., and Holland, B. S. (2005). Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Syst. Biol.* 54, 952–961. doi: 10.1080/10635150500234674

Schumer, M., Rosenthal, G. G., and Andolfatto, P. (2014). How common is homoploid hybrid speciation? *Evolution* 68, 1553–1560.

Sherlock, ÍA., Guitton, N., and Miles, M. (1977). *Rhodnius paraensis* espécie nova do Estado do Pará, Brasil (Hemiptera, Reduviidae, Triatominae). *Acta Amazon.* 7, 71–74.

Sloan, D. B., Havird, J. C., and Sharbrough, J. (2017). The on-again, off-again relationship between mitochondrial genomes and species boundaries. *Mol. Ecol.* 26, 2212–2236. doi: 10.1111/mec.13959

Souza, E., Von Atzingen, N. C. B., Furtado, M. B., De Oliveira, J., Nascimento, J. D., Vendrami, D. P., et al. (2016). Description of *Rhodnius marabaensis* sp. n. (Hemiptera, Reduviidae, Triatominae) from Pará State, Brazil. *ZooKeys* 621, 45–62. doi: 10.3897/zookeys.621.9662

Sun, J. T., Jin, P. Y., Hoffmann, A. A., Duan, X. Z., Dai, J., Hu, G., et al. (2018). Evolutionary divergence of mitochondrial genomes in two *Tetranychus* species distributed across different climates. *Insect Mol. Biol.* 27, 698–709. doi: 10.1111/imb.12501

Toews, D. P., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* 21, 3907–3930. doi: 10.1111/j.1365-294X.2012.05664.x

Valente, V. D. C., Valente, S. A. D. S., Carcavallo, R. U., Rocha, D. D. S., Galvão, C., and Jurberg, J. (2001). Considerações sobre uma nova espécie do gênero *Rhodnius* stal, do Estado do Pará, Brasil (Hemiptera, Reduviidae, Triatominae. *Entomol. Vect.* 8, 65–80.

Waterhouse, R. M., Seppey, M., Simão, F. A., and Zdobnov, E. M. (2019). "Using BUSCO to assess insect genomic resources," in *Insect Genomics*, eds S. Brown and M. Pfrender (New York, NY: Humana Press), 59–74. doi: 10.1007/978-1-4939-8775-7_6

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.

Zhao, Y., Galvão, C., and Cai, W. (2021). *Rhodnius micki*, a new species of Triatominae (Hemiptera, Reduviidae) from Bolivia. *ZooKeys* 1012:71. doi: 10.3897/zookeys.1012.54779

Zhao, Y., Jiang, M., Wu, Y., Song, F., Cai, W., and Li, H. (2019). Mitochondrial genomes of three kissing bugs (Reduviidae: Triatominae) and their phylogenetic implications. *Int. J. Biol. Macromol.* 134, 36–42. doi: 10.1016/j.ijbiomac.2019.05.020

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership