



**HAL**  
open science

# Dynamique des génomes et évolution moléculaire des récepteurs chimiosensoriels des vecteurs de la maladie de Chagas du genre *Rhodnius*

Marie Merle

► **To cite this version:**

Marie Merle. Dynamique des génomes et évolution moléculaire des récepteurs chimiosensoriels des vecteurs de la maladie de Chagas du genre *Rhodnius*. Evolution [q-bio.PE]. Université Paris-Saclay, 2023. Français. NNT : 2023UPASL037 . tel-04100129

**HAL Id: tel-04100129**

**<https://theses.hal.science/tel-04100129>**

Submitted on 17 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Dynamique des génomes et évolution moléculaire des récepteurs chimiosensoriels des vecteurs de la maladie de Chagas du genre *Rhodnius*

Genome dynamics and molecular evolution of chemosensory  
receptors in Chagas disease vectors of the *Rhodnius* genus

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n°577 : structure et dynamique des systèmes vivants (SDSV)  
Spécialité de doctorat : Evolution  
Graduate School : Sciences de la vie et santé. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche EGCE (Université Paris-Saclay,  
CNRS, IRD), sous la direction de **Myriam HARRY**, Professeur, et le  
co-encadrement de **Florence MOUGEL**, Maître de conférences

Thèse soutenue à Paris-Saclay, le 18 avril 2023, par

**Marie MERLE**

## Composition du Jury

Membres du jury avec voix délibérative

<b>Carole SMADJA</b> Directrice de recherche, ISEM Montpellier	Présidente
<b>Carole KERDELHUÉ</b> Directrice de recherche, CBGP Montpellier	Rapporteur & Examinatrice
<b>Claudio LAZZARI</b> Professeur, IRBI Tours	Rapporteur & Examinateur
<b>Antoine BRANCA</b> Maître de conférences, EGCE Paris-Saclay	Examinateur
<b>Emmanuelle JACQUIN-JOLY</b> Directrice de recherche, iEES Paris	Examinatrice

**Titre :** Dynamique des génomes et évolution moléculaire des récepteurs chimiosensoriels des vecteurs de la maladie de Chagas du genre *Rhodnius*

**Mots clés :** phylogénomique, génomique, vecteurs de la maladie de Chagas, gènes chimiosensoriels, adaptation, Hemiptera.

**Résumé :** Les Triatominae sont des punaises hématophages vectrices de *Trypanosoma cruzi*, agent responsable de la maladie de Chagas, endémique en Amérique Latine. En milieu naturel, les punaises hématophages du genre *Rhodnius* infestent principalement les palmiers et se nourrissent sur une large gamme d'hôtes vertébrés qui trouvent refuge dans cet habitat. Cependant, certaines espèces de *Rhodnius* sont domiciliaires, c'est-à-dire qu'elles ont quitté leur habitat naturel (milieu sylvatique) et sont venues en contact étroit avec les populations humaines en colonisant leurs habitations.

Les liens de parenté entre les différentes espèces du genre *Rhodnius* ont été retracés par une approche phylogénomique reposant sur l'étude des mitogénomes et d'un set de 51 gènes nucléaires et incluant 16 espèces sur les 24 espèces décrites via l'étude de 36 échantillons. Cette approche phylogénomique a permis de statuer sur les affinités évolutives des 3 groupes d'espèces décrits (*pallescens*, *pictipes* et *prolixus*), de classer les espèces du genre *Psammolestes* au sein du genre *Rhodnius* et de montrer des événements d'introgression multiples entre espèces. Cette étude phylogénétique a été complétée par une approche de génomique comparative entre espèces adaptées à différents habitats. La taille des génomes a été estimée par deux méthodes, la cytométrie de flux et l'analyse des données moléculaires à haut débit.

Pour étudier la dynamique des génomes une phylogénie calibrée à l'aide de données fossiles a été réalisée. Onze génomes de *Rhodnius* (dont 4 domiciliaires) ont été assemblés puis annotés. L'estimation de la vitesse de gains et pertes de gènes par million d'années montre que ces événements se produisent 16 fois plus rapidement chez les espèces domiciliaires en comparaison des espèces sylvatiques. Cette étude a permis de proposer des gènes candidats dont des récepteurs chimiosensoriels impliqués dans la domiciliation des espèces du genre *Rhodnius*.

Les gènes chimiosensoriels appartiennent à des familles multigéniques dont la dynamique évolutive est rapide et à l'origine de changements dans la sensibilité sensorielle. Ils sont donc des objets d'étude de choix dans la recherche de gènes impliqués dans la reconnaissance d'un nouvel environnement. L'annotation spécifique de ces récepteurs chimiosensoriels (de 49 à 135 ORs, de 21 à 32 GRs, selon les espèces) suivie de la recherche de gènes sous sélection a permis de mettre en évidence 4 gènes candidats en lien avec la domiciliation (OR1, OR38, OR102 et ORco). Une meilleure compréhension des mécanismes d'adaptation des punaises hématophages à leur environnement pourrait être un facteur clé dans l'effort de lutte antivectorielle.

**Title :** Genome dynamics and molecular evolution of chemosensory receptors in Chagas disease vectors of the *Rhodnius* genus

**Keywords :** phylogenomics, genomics, vectors of Chagas disease, chemosensory genes, adaptation, Hemiptera.

**Abstract :** Triatominae are haematophagous bugs that can transmit *Trypanosoma cruzi*, the causative agent of Chagas disease, endemic to Latin America. In the wild, species belonging to the genus *Rhodnius* mainly infest palm trees, feeding on a wide range of vertebrate hosts that find refuge in this habitat. However, some *Rhodnius* species are domiciliary, i.e. they have left their natural habitat (sylvatic environment) and have come into close contact with human populations by colonising their dwellings, and can then transmit *T. cruzi* to humans.

The evolutionary history of species belonging to *Rhodnius* genus was assessed using a phylogenomic strategy based on whole mitogenome sequences and a set of 51 nuclear genes, characterized for 36 samples representing 16 species out of the 24 described in the genus. This approach allowed us to determine the evolutionary relationship between the 3 major groups of species (*pallescens*, *pictipes* and *prolixus*), to integrate species from *Psammolestes* genus inside the genus one, and to describe multiple introgression events between species. A comparative genomic approach was conducted between species living in different habitats. The genome size was estimated using flow cytometry and analysis of high throughput sequencing data.

Genomes from 11 *Rhodnius* species (including 4 domiciliary

species) were assembled and used to build a calibrated phylogeny including sequences from published genomes and fossil records.

The divergence times inferred from this phylogeny allowed us to study genome dynamics. It revealed that the rate of gains and losses per gene per million years is 16 times faster in domiciliary species compared to sylvatic ones. This study has made it possible to propose candidate genes including chemosensory receptors involved in the domiciliation of species of the genus *Rhodnius*.

Chemosensory genes belong to multigene families with rapid evolutionary dynamics that cause changes in sensory sensitivity. Their study is especially relevant when searching for genes involved in the recognition of a new environment. The chemosensory receptors were annotated (from 49 to 135 ORs and from 21 to 32 GRs depending on the species studied) and analyzed to detect evidence of selection pressure: 4 candidate genes (OR1, OR38, OR102 and ORco) were identified that could be involved in the domiciliation process. A better understanding of the mechanisms of adaptation of these bugs to their environment could be a key factor in vector control efforts.



## REMERCIEMENTS

---

Tout d'abord merci à [Myriam Harry](#) d'avoir cru en moi. Merci de m'avoir guidée dans les choix de Masters, de m'avoir prise en stage et de m'avoir entraînée pour le concours de thèse. Merci pour ton soutien pendant les confinements et de m'avoir donné autant de responsabilités.

Merci à [Florence Mougel](#) qui m'a beaucoup encouragée et guidée lors de ma thèse. Merci pour tes lumières et nos discussions qui m'ont motivées à continuer.

Merci aux membres de mon comité de thèse. Merci d'avoir fait en sorte que la visio ne soit pas un frein à nos échanges, et d'avoir suivi toute la thèse.

Merci aux membres de mon jury de thèse pour avoir accepté de lire et d'évaluer mon travail, d'avoir pris beaucoup de votre temps pour que je puisse passer ce diplôme.

Merci à mes frères de thèse, [Bapbap](#) que je connais depuis nos stages de M2 qui ont commencé en 2019 et qui soutient sa thèse quelques semaines avant moi. Je n'aurai pas pu mieux choisir camarade de thèse. [Mino](#) qui, avec sa bonne humeur a égayé nos journées avec ses goûts musicaux atypiques. [Mumu](#), qui malgré son humeur bien plus macabre a été notre soleil durant cette dernière année.

Mes sœurs de thèse, [Perrine](#) qui a été ma grande sœur, qu'est ce que j'aurais fait sans toi ... Tu as toujours eu les mots justes, tu es d'une grande sagesse. [Erina](#), la reine du laboratoire, d'une positivité et d'un calme effrayant mais inspirant, mon seul regret est de ne pas avoir passé plus de temps avec toi.

Merci à tous les stagiaires qui ont travaillé sur les génomes de *Rhodnius*, [Léa](#), [Cathy](#), [Gabin](#) et [Louise](#).

Merci à [Laure](#), directrice du laboratoire et à [Jean-Christophe](#), ainsi qu'à [Magalie](#) et [Valérie](#) pour la gestion et le service administratif.

Merci à tous les doctorants de l'EGCE, [Malena](#), [Ovide](#), [Camillo](#), [Carolina](#), [Simon](#), [Benjamin](#), [Mélissa](#), [Loïc](#), [Siddharth](#), [Apolline](#), [Zhen](#), [Héloïse](#). Merci aussi à [Maxime](#), à [Manuel](#), à [Sacha](#), à [Pénélope](#). Merci à mes aînées, [Clémence](#), [Elise](#) et [Julia](#). Merci à ceux avec qui j'ai partagé mon bureau au début de ma thèse, [Mélissa](#), [Sarah](#) et [Romain](#) a qui j'ai volé la place. [Joseph](#) je t'adore, je n'ai jamais vu quelqu'un réinstaller si souvent Linux sur son ordinateur. [Jérémy](#) merci pour tous ces beaux moments et désolée de t'avoir fait autant sursauté en arrivant dans ton bureau discrètement

pendant tes réflexions intenses. Merci à Héloïse, Claire, Rémi, Fabrice, Jonathan, Jean-Luc, Sylvie, Jean-Bernard et David. Merci à Hélène qui m'a donné le meilleur conseil que j'ai entendu de ma vie pour gérer les migraines (à savoir, prendre ses médicaments dès les premiers symptômes, et pas lors du pic de douleur), merci à toi et Béra de prendre soin de nous. Merci à Pierre Capy, responsable du master que j'ai réalisé et ancien directeur de l'SDSV, vous m'avez transmis votre envie d'en apprendre toujours plus et surtout de transmettre. Merci à François qui m'a donné de très bons conseils et qui m'a permis de terminer ma thèse dans les meilleures conditions possibles.

En grand merci à l'ensemble du personnel du laboratoire, discuter avec vous tous était chouette.

Merci à ma famille et tout particulièrement à ma mère. Merci à mon filleul Lucas adoré qui m'a dit que mon métier était génial. Merci à mon soutien moral, Christine, qui a été d'une pertinence nécessaire. Merci à mon compagnon Thomas, merci de m'avoir soutenu dans tous mes états, je n'aurai pas pu y arriver sans toi.

## CONTENU

---

La thèse sera présentée comme suit : une **introduction générale** sur l'état de l'art des sujets abordés, une **problématique générale** de la thèse, un **matériel et méthode général** de l'ensemble des travaux réalisés lors de la thèse puis 4 chapitres, un par article publié ou en cours de publication.

Le **chapitre 1** sera consacré à notre article sur la phylogénie des *Rhodnius* pour lequel je suis co-premier auteur avec Jonathan Filée. Cet article a été publié dans *Phylogenetics, Phylogenomics, and Systematics*, une section du journal *Frontiers in Ecology and Evolution*.

Filée, J., **Merle, M.**, Bastide, H., Mougél, F., Bérenger, J. M., Folly-Ramos, E., Almeida, C. E. & Harry, M. (2022). Phylogenomics for Chagas disease vectors of the *Rhodnius* genus (Hemiptera, Triatominae): what we learn from mito-nuclear conflicts and recommendations. *Frontiers in Ecology and Evolution*, 9, 750317.

Le **chapitre 2** porte sur les mesures des tailles de génome que nous avons réalisées à partir de deux méthodes indépendantes et publié dans *The American Journal of Tropical Medicine and Hygiene* dont je suis également co-premier auteur avec Jonathan Filée.

**Merle, M.**, Filée, J., de Oliveira, J., Almeida, C. E., Mougél, F., Bastide, H., Girondot, M., da Rosa, J.A. & Harry, M. (2022). Genome Size Variation of Chagas Disease Vectors of the Rhodniini Tribe. *The American Journal of Tropical Medicine and Hygiene*, tpm210394-tpm210394.

Le **chapitre 3** comprend l'ensemble de nos travaux sur la dynamique des génomes chez *Rhodnius* et la comparaison des génomes des espèces domiciliaires et sylvatiques. Je présente ici l'article en cours de préparation dont je serai le premier auteur.

Le **chapitre 4** porte sur l'article en cours de préparation sur l'évolution des récepteurs chimiosensoriels de 13 génomes assemblés de *Rhodnius* ainsi que l'implication potentielle de ces gènes dans le processus de domiciliation.

Chaque article est encadré par un court résumé en français ainsi qu'une conclusion en français. Ensuite le manuscrit poursuivra sur une **discussion générale** puis se terminera sur les **perspectives** de ces travaux.





## TABLE DES MATIÈRES

---

<b>Introduction</b>	<b>1</b>
1. La maladie de Chagas	1
2. La biologie des vecteurs de la maladie de Chagas et le processus de domiciliation	6
3. Biogéographie et évolution des vecteurs de la maladie de Chagas	13
4. Les apports des données moléculaires sur l'étude des insectes	20
5. Les gènes chimiosensoriels et leur lien avec l'écologie des insectes	28
<b>Objectif de la thèse</b>	<b>43</b>
<b>Matériels et Méthodes Général</b>	<b>45</b>
1. Echantillons	45
2. Extraction d'ADN et d'ARN	47
3. Phylogénies nucléaires et mitochondriales	47
4. Mesure de la taille des génomes	49
5. Assemblage des génomes	50
6. Datation	54
7. Annotation de gènes	57
8. Vitesse d'évolution des gènes	62
<b>Chapitre 1 : Phylogénomique des <i>Rhodnius</i></b>	<b>65</b>
Contexte et problématique de l'étude	65
Principaux résultats	66
Article 1 : Phylogenomics for Chagas Disease Vectors of the <i>Rhodnius</i> Genus (Hemiptera, Triatominae): What We Learn From Mito-Nuclear Conflicts and Recommendations	68
Postambule sur l'histoire évolutive des <i>Rhodnius</i>	104
<b>Chapitre 2 : Taille des génomes</b>	<b>107</b>
Contexte et problématique de l'étude	107
Principaux résultats de l'étude	108
Article 2 : Genome Size Variation of Chagas Disease Vectors of the Rhodniini Tribe	109
Postambule sur les différences de taille de génome	118
<b>Chapitre 3 : Dynamique des génomes</b>	<b>119</b>
Contexte et problématique de l'étude	119
Principaux résultats de l'étude	119
Article 3 : Genome dynamics of the <i>Rhodnius</i> genus, vector of Chagas disease	121
Postambule sur l'article en cours d'écriture	153
<b>Chapitre 4 : Gènes chimiosensoriels</b>	<b>155</b>
Contexte et problématique de l'étude	155
Principaux résultats	156
Article (en préparation) : Evolution of chemosensory receptor genes in 13 Chagas disease vectors of the genus <i>Rhodnius</i>	157

Postambule sur l'article en cours d'écriture	188
<b>Discussion Générale</b>	<b>191</b>
Histoire évolutive des espèces du genre <i>Rhodnius</i>	191
Evolution et adaptation au sein de genre <i>Rhodnius</i>	197
<b>Perspectives</b>	<b>207</b>
Etude d'autres gènes impliqués dans la perception	207
Études comportementales chez les <i>Rhodnius</i>	208
<b>Conclusion</b>	<b>211</b>
<b>Bibliographie</b>	<b>213</b>

## INTRODUCTION

---

### 1. LA MALADIE DE CHAGAS

La maladie de Chagas, endémique en Amérique Latine fait partie des 20 maladies reconnues par l'Organisation Mondiale pour la Santé (OMS) comme maladies tropicales négligées (<https://www.who.int>, 2022). Ces maladies sont définies comme affectant plus d'un million de personnes dans des régions en voie de développement avec des moyens de lutte coûteux dont l'utilisation à grande échelle est limitée. La maladie de Chagas est causée par l'organisme unicellulaire kinétoplastidé *Trypanosoma cruzi* (Kinetoplastea ; Trypanosomatidae), qui est transmis à l'Homme par les fèces infectées d'insectes hématophages appartenant à la sous-famille des Triatominae (Hemiptera ; Reduviidae). L'OMS estime entre 6 et 7 millions de personnes infectées par le parasite *T. cruzi*.

#### 1.1. Découverte de la maladie

Le plus ancien cas décrit de maladie de Chagas date d'il y a 9 000 ans chez une momie Chinchorro retrouvée entre le nord du Chili et le sud du Pérou ([Aufderheide et al. 2004](#)). Les auteurs concluent que le cycle impliquant des animaux réservoirs de la maladie de Chagas était déjà existant lors de l'établissement des premiers humains dans la zone andine. D'autres momies plus récentes ont été retrouvées dans la même région également infectées par *T. cruzi*. Certains rapports de médecins semblaient évoquer de potentiels symptômes de la maladie entre le XVI<sup>ème</sup> et le XIX<sup>ème</sup> siècle. Notamment dans un rapport, un médecin danois Theodoro J.H. Langgaard (1813-1884), vivant au Brésil, explique que certains de ses patients ont des difficultés à s'alimenter, ce qu'il appelle le "*mal de engasgo*" (difficultés à avaler en portugais), mal attribué à la maladie de Chagas par certains auteurs ([Steverding 2014](#)).

Carlos Ribeiro Justiniano Chagas, médecin hygiéniste, est appelé en 1908 sur le site de la construction d'une voie ferrée dans le nord de l'État de Minas Gerais au Brésil où les ouvriers seraient décimés par le paludisme. Il s'intéresse aux punaises qui infestent les habitations locales et il les dissèque. Il trouve des trypanosomes et décrit pour la première fois le parasite *T. cruzi* qu'il nomme d'après son mentor, le médecin Oswaldo Cruz (1872-1917) ([Steverding 2014](#)). Après avoir défini *T. cruzi* comme pathogène à l'aide d'animaux de laboratoire mis en contact avec des punaises infectées, il décrit pour la première fois la maladie et ses symptômes en 1909 chez une enfant brésilienne de 2 ans. La fillette nommée Bérénice présente une forte fièvre, une hypertrophie du foie, un œdème facial et la

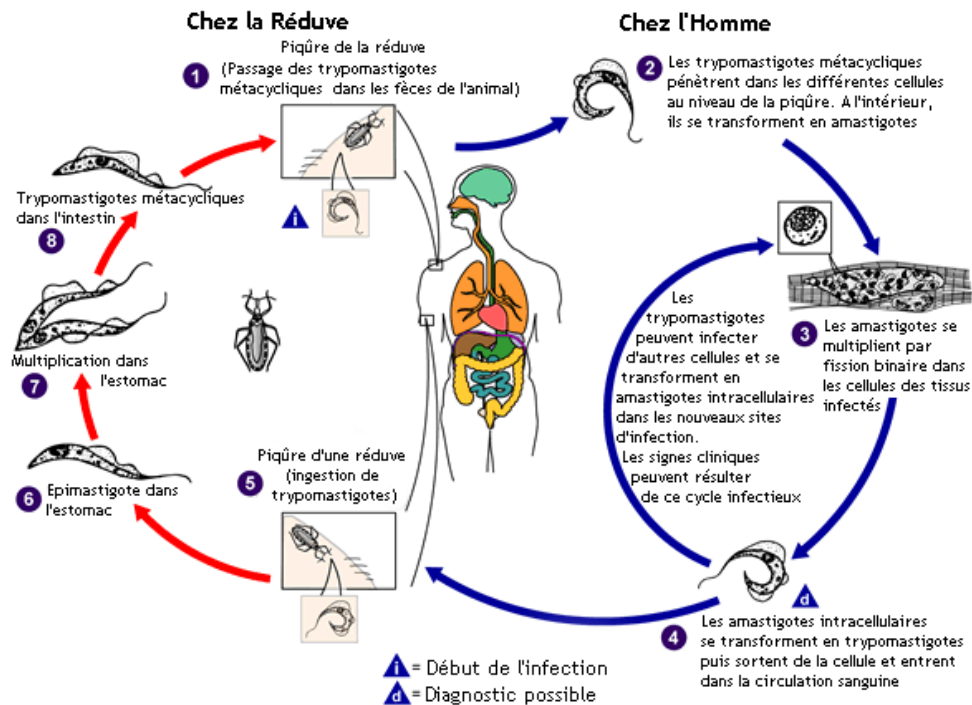
présence du même parasite dans le sang (Chagas 1909; Lidani et al. 2019). Bérénice est restée par la suite asymptomatique toute sa vie et est décédée à l'âge de 73 ans (Steверding 2014; Lidani et al. 2019). Carlos Chagas décrit le vecteur, l'agent pathogène puis ses différents stades de développement, les hôtes, les manifestations cliniques et l'épidémiologie de la maladie.

Depuis la description du premier cas, la maladie de Chagas a été reportée comme endémique dans 21 pays du continent Américain, majoritairement en Amérique Latine et touche environ 6 millions de personnes. La maladie représente une incidence annuelle moyenne de 30 000 nouveaux cas, 12 000 décès, et 8 600 contaminations de la mère à l'enfant lors de la grossesse. L'OMS et PAHO (*Pan American Health Organisation*) en 2022 estiment qu'environ 70 millions de personnes sur le continent américain vivent dans des zones d'exposition et risquent de contracter cette maladie.

## 1.2. Cycle de vie du parasite responsable de la maladie

Lorsqu'elle pique l'hôte, la punaise hématophage a un réflexe de défécation. Si la punaise est infectée par le parasite unicellulaire *T. cruzi*, celui-ci se retrouve dans les fèces. La piqûre démange l'hôte qui, en grattant, ramène les fèces contaminées au point de piqûre. Le parasite *T. cruzi* est à ce moment sous forme trypomastigote métacyclique, définie par la présence d'un flagelle. Après pénétration dans les cellules du corps de l'hôte, les parasites sous forme amastigote se multiplient par reproduction asexuée puis sortent de la cellule sous forme trypomastigote métacyclique. Les trypanosomes trypomastigotes métacycliques peuvent de nouveau pénétrer dans une autre cellule pour s'y multiplier ou rester dans le sang. Lors d'une piqûre par une punaise, celle-ci ingère du sang ainsi que des trypomastigotes métacycliques *T. cruzi* qui se retrouveront dans son système digestif pendant environ 22 jours et le cycle pourra recommencer (Neal et Miles 1977) (Figure 1).

Le vecteur est l'animal qui fait circuler le parasite *T. cruzi* alors que le l'hôte est l'animal dans lequel le parasite va pouvoir effectuer la reproduction sexuée. Un total de 180 espèces de mammifères appartenant aux ordres suivants : Didelphidomorphia, Lagomorpha, Chiroptera, Rodentia, Pilosa, Cingulata, Carnivora, Primata, Perisodactyla, ont été trouvés naturellement infectés par *T. cruzi* y compris l'Homme, qui, en plus de souffrir de la maladie, agit comme un réservoir (Herrera 2010).



**Figure 1 : Cycle de la maladie de Chagas** (<https://www.cdc.gov/>). Domaine public (traduite par Wikipedia)

### 1.3. Symptômes de la maladie

Après infection, une phase aiguë de maladie de Chagas est décrite et dure de 4 à 8 semaines. Cette phase est souvent asymptomatique ou peut engendrer, dans 5% des cas, de la fièvre, des malaises et de l'œdème unilatéral caractéristique des paupières qui se produit lorsque l'insecte pique près de l'œil, également connu sous le nom de signe de Romaña (Pino-Marín et al. 2021). Ces symptômes apparaissent 1 à 2 semaines après la piqûre des triatomes. La mort survient lors de cette phase aiguë dans 5-10% des cas symptomatiques d'une myocardite grave ou d'une méningo-encéphalite notamment chez les enfants (Rassi et al. 2010; Bruneto et al. 2021).

La maladie peut ensuite devenir chronique dans 30% des cas (Carod-Artal 2013). Lors de cette phase, la personne infectée peut être affectée de différentes pathologies : cardiaque, digestive et cardio-digestive. La forme la plus fréquente de la maladie de Chagas chronique est la forme cardiaque. Elle se développe chez 20 à 30 % des individus (Rassi et al. 2010; Machado et al. 2012), elle entraîne une défaillance cardiaque grave qui peut entraîner la mort (Pino-Marín et al. 2021). La forme digestive est observée dans 10 à 15 % des cas de phases chroniques et semble observée plus spécifiquement en Argentine, au Brésil, au Chili et en Bolivie. Cette

répartition peut être due aux différentes souches de parasites *T. cruzi* (Rassi et al. 2010). Les problèmes gastro-intestinaux se manifestent par la présence d'hypertrophie d'organes : des mégaoesophages et mégacôlons, ce qui entraîne des difficultés à se nourrir et un amaigrissement de la personne malade. La maladie peut aussi réapparaître lors de traitements immunosuppresseurs ou d'infection VIH (Rassi et al. 2010).

Le coût des soins pour les patients atteints de la maladie de Chagas a été estimé à 267 millions de dollars pour la Colombie en 2008 (Castillo-Riquelme et al. 2008).

#### 1.4. Modes de transmission de la maladie

Le mode de transmission principal est le mode vectoriel. Cette capacité vectorielle est évaluée en laboratoire à partir du temps de génération des vecteurs (plus le cycle de vie est court plus le risque d'infestation des maisons est élevé), du nombre de repas de sang effectué par le triatome et du temps entre l'alimentation et la défécation sur l'animal, puisque le parasite se trouve dans les fèces (Folly-Ramos et al. 2016). Certains triatomines n'ont jamais été décrits comme associés au parasite comme *R. neivai* (Carcavallo et Barreto 1976). La quantité de *T. cruzi* dans le système digestif et la souche de *T. cruzi* peuvent aussi jouer sur la capacité à transmettre le parasite (Rassi et al. 2010). Les piqûres de triatomines peuvent être très fréquentes, ainsi jusqu'à 58 piqûres par jour par personne ont été enregistrées au Venezuela dans des zones particulièrement infestées (Rabinovich et al. 1979).

Il existe d'autres modes de contamination. La maladie de Chagas peut être transmise par transfusion sanguine (Rassi et al. 2010). En France, le sang de donneurs ayant séjourné en Amérique Latine est systématiquement dépisté pour la maladie de Chagas (El Ghouzzi et al. 2010). La maladie peut aussi être transmise de façon congénitale, de la mère ayant la forme chronique à l'enfant, engendrant parfois des pathologies cardiaques pour l'enfant. Le taux de transmission congénitale varie selon les pays, de 5,6% (en Uruguay) à 19% (au Chili) (Carod-Artal 2013). Il a été reporté des cas de transmission par greffe d'organes. Ces modes de contamination sont dépendant de la quantité de parasites, de leur souche et de la condition médicale de l'hôte (Rassi et al. 2010).

Des cas de contaminations orales sont également reportés. Ainsi, des micro épidémies ont été observées dans la région brésilienne amazonienne en lien avec la consommation de produits à base d'açaï, (Nóbrega et al. 2009). Les baies d'açaï proviennent du palmier *Euterpe oleaca* (Arecal ; Arecaceae). L'açaï est consommé sous forme de pâte d'açaï, de boisson ou de glaces. Les fruits sont contaminés par les déjections de triatomines ou bien par des triatomines entiers broyés au moment de la fabrication des

différents produits alimentaires (Ferreira et al. 2018). La consommation de viande contaminée a aussi été reportée comme source de la maladie (Sangenis et al. 2016). L'ingestion de boissons ou de nourriture contaminées par les déjections des insectes et/ou des insectes eux-mêmes sont généralement associées avec des infections massives de parasites et provoquent une situation clinique aiguë avec des taux de mortalité allant jusqu'à 4% avec 1% en moyenne (Bruneto et al. 2021). La sévérité de la pathologie orale pourrait être expliquée par différents facteurs dont l'interaction entre certains composés de la muqueuse gastrique et certaines glycoprotéines de surface (gp35/50) des parasites, ce qui semblerait protéger les formes métacycliques du parasite de leur destruction dans l'estomac. La dégradation par la pepsine présente dans l'estomac d'autres glycoprotéines (gp90) pourrait également augmenter l'infektivité de certaines souches de trypanosomes (Cortez et al. 2006; Covarrubias et al. 2007).

### 1.5. Moyens de lutte

Il n'existe pas de vaccin préventif contre la maladie de Chagas. Les personnes infectées peuvent être traitées au benznidazole ou au nifurtimox. Cependant il existe des effets secondaires au traitement par ces molécules (cardiaques, digestifs, neurologiques) et des contre-indications (risque pour les femmes enceintes et les personnes immunodéprimées). De plus, l'efficacité de ces traitements diminue si la maladie est chronique (Rassi et al. 2010).

Le principal moyen de lutte contre la maladie est la lutte chimique contre le mode de transmission principal : les vecteurs triatomines. L'OMS recommande de pulvériser dans les habitations des insecticides à effet rémanent. Le premier insecticide utilisé pour lutter contre les vecteurs a été le DDT qui est un insecticide organochloré. Les premières campagnes de lutte anti-vectorielle ont eu lieu vers les années 1950 et le DDT, utilisé à l'époque contre les vecteurs du paludisme n'était pas efficace sur les triatomines naturellement résistantes. D'autres insecticides organochlorés ont alors été utilisés, tels que le lindane et la dieldrine. Le lindane a été utilisé au Venezuela pour lutter contre *Rhodnius prolixus* et sera majoritairement utilisé avec la dieldrine jusque dans les années 1970 avec succès (Mougabure-Cueto et Picollo 2015). Le succès de l'efficacité de ces insecticides organochlorés est néanmoins accompagné de risques potentiels pour la santé et l'environnement, c'est pour cela que les organophosphorés et carbamates moins persistants sont alors proposés comme alternatives, comme le malathion et le fénitrothion, introduits en 1975. Les pyréthrinoïdes, présentant encore moins de risques sont utilisés dès les années 1980 et ayant une activité insecticide plus élevée sont encore utilisés. Par exemple, la deltaméthrine et la cyperméthrine sont

utilisées avec succès sur le terrain pour le contrôle des triatomes en Argentine et au Brésil (Mougabure-Cueto et Picollo 2015). Cependant, en Bolivie, la plupart des populations domiciliées de *T. infestans* sont résistantes à la deltaméthrine (Lardeux et al. 2010). De plus, les campagnes d'éradication des vecteurs par pulvérisation chimique dans les habitations sont coûteuses pour les États d'Amérique Latine, la Colombie par exemple a dépensé 5 millions de dollars en 2008 (Castillo-Riquelme et al. 2008).

## 2. LA BIOLOGIE DES VECTEURS DE LA MALADIE DE CHAGAS ET LE PROCESSUS DE DOMICILIATION

### 2.1. Cycles sylvatiques et domiciliés

La plupart des espèces du genre *Rhodnius* sont inféodées aux palmiers (Abad-Franch et al. 2005; Lent et Wygodzinsky 1979; Galvão et Justi 2015; Dias et al. 2014) dans lesquels elles se nourrissent sur une large gamme d'hôtes vertébrés. Ainsi, dans la région du Tapajós en Amazonie brésilienne, 54% des palmiers échantillonnés (*Attalea maripa*, *A. speciosa*, *A. phalerata*) étaient infestés par *R. robustus* (Dias et al. 2014) et dans l'État du Ceará au Brésil, 78% des palmiers (*Attalea speciosa*, *Acrocomia* spp.) étaient infestés par *R. nasutus* (Dias et al. 2008). Certaines espèces ne sont retrouvées que sur une seule espèce de palmier comme *R. brethesi* qui est spécialisé sur le palmier *Leopoldinia piassaba* (Arecaceae ; Arecaceae) tandis que d'autres semblent généralistes comme *R. neglectus* retrouvé sur plusieurs palmiers des genres *Orbignya* et *Acrocomia*. D'autres espèces peuvent être aussi trouvées sur des broméliacées du genre *Aechmea* comme *R. pictipes* (Abad-Franch et al. 2005, Tableau 1). La spécialisation aux palmiers des espèces de *Rhodnius* n'est pas encore bien comprise (Calderón et González 2020). Différentes hypothèses ont été émises liées à la reconnaissance de l'architecture des palmiers (Romaña et al. 1999) ou à la détection de composés volatiles émis par des microorganismes (Lorenzo et al. 1999). Mais il est possible que les *Rhodnius* soient spécialisées à un hôte lui-même spécialisé à un palmier. Notamment, il a été montré que la présence de nids d'oiseaux dans les palmiers favorise l'infestation de ceux-ci par *R. pallescens* (Gurgel-Gonçalves et al. 2008).

Suite à l'anthropisation des milieux naturels (milieu sylvatique), les triatomes sont venus en contact étroit avec les populations humaines faisant alors circuler le parasite du milieu sylvatique au milieu domiciliaire. La domiciliation décrit le processus de colonisation des habitats humains par certaines espèces jusqu'alors en milieu naturel renforcée par la fragmentation des habitats naturels des vecteurs liée à la déforestation ou à l'urbanisation. Les habitations humaines deviennent donc une nouvelle



niche écologique potentielle pour les vecteurs et les humains une nouvelle source de nourriture ce qui pourrait expliquer l'augmentation importante des cas de maladie de Chagas ces dernières décennies (Lidani et al. 2019).

Ce phénomène de domiciliation est décrit pour différentes espèces de vecteurs comme les moustiques (Diptera ; Culicidae) *Aedes aegypti* en Afrique subsaharienne vecteur de la fièvre jaune et de la dengue (Powell et Tabachnick 2013), ou encore *Culex pipiens*, vecteur de la fièvre du Nil occidental (Fritz et al. 2015). Au sein du genre *Rhodnius*, on observe de multiples événements de domiciliation d'autant plus que l'interaction entre les humains et les Triatominae est assez ancienne. Le premier cas recensé d'infection humaine par *T. cruzi* date, comme mentionné plus haut, de 9 000 ans (Aufderheide et al. 2004) mais les premiers cas pourraient être plus anciens : à l'arrivée des premiers humains en Amérique Latine il y a 11 000 à 13 000 ans, les punaises étaient déjà présentes sur le territoire ainsi que *T. cruzi* (présence datée de 7 à 10 millions d'années d'après Steverding et al. 2014). On distingue différents types d'espèces selon leur statut de domiciliation : domiciliaires, en voie de domiciliation, sporadiques et sylvatiques (Tableau 1).

**Tableau 1 : Distribution géographique des espèces de *Rhodnius*, statut et inféodation aux palmiers et broméliacées** (d'après Galvão et Justi 2015 pour l'habitat, Abad-Franch et al. 2005 pour l'inféodation aux palmiers et broméliacées).

Espèces	Habitat	Spécialiste ou généraliste sur la plante hôte
<i>R. amazonicus</i> Almeida et al. 1973	Sylvatique	Données insuffisantes
<i>R. barretti</i> Abad-Franch et al. 2013	Sylvatique	Palmiers <i>Attalea butyracea</i> et <i>Oenocarpus bataua</i> et autres espèces
<i>R. brethesi</i> Matta, 1919	Sporadique dans les domiciles	Spécialisé sur <i>Leopoldinia piassaba</i>
<i>R. colombiensis</i> Mejia et al. 1999	Sporadique dans les domiciles	Spécialisé sur <i>Attalea (=Maximiliana) butyracea</i>
<i>R. dalessandroi</i> Carcavallo and Barreto, 1976	Données insuffisantes	Données insuffisantes
<i>R. domesticus</i> Neiva and Pinto, 1923	Sporadique dans les domiciles	Préférentiellement dans les broméliacées <i>Attalea sp</i>
<i>R. ecuadorensis</i> Lent and Leon, 1958	Domiciliaire en Equateur	<i>Phytelephas aequatorialis</i> , <i>Elaeis guineensis</i>
<i>R. marabensis</i> Souza et al. 2016	Sporadique dans les domiciles	données insuffisantes
<i>R. micki</i>	Données insuffisantes	données insuffisantes

Zhao et al. 2021		
<i>R. milesi</i> Carcavallo et al. 2001	Sylvatique	données insuffisantes
<i>R. montenegrensis</i> da Rosa et al. 2012	Sylvatique, mais quelques spécimens péri-domiciliaires	<i>Orbignya phalerata</i>
<i>R. nasutus</i> Stal, 1959	En voie de domiciliation au Brésil	<i>Copernicia cerifera</i> (= <i>prunifera</i> )
<i>R. neglectus</i> Lent, 1954	En voie de domiciliation au Brésil	Généraliste : <i>Orbignya maritima</i> , <i>Or. oleifera</i> , <i>Or. martiana</i> , <i>Acrocomia</i> etc
<i>R. neivai</i> Lent, 1953	Sporadique	<i>Copernicia tectorum</i> , <i>Attalea</i> spp
<i>R. pallescens</i> Lent, 1953	Domiciliaire au Panama	Généraliste : <i>A. butyracea</i> , <i>Cocos nucifera</i> , <i>E. oleifera</i> , <i>O. bataua</i> , <i>Co. Tectorum</i>
<i>R. paraensis</i> Sherlock et al. 1977	Sylvatique (nids de rongeurs arboricoles)	Données insuffisantes
<i>R. pictipes</i> Stal, 1972	Sporadique dans les domiciles	Généraliste : <i>Bromeliads A. butyracea</i> , <i>A. maripa</i> etc
<i>R. prolixus</i> Stal, 1859	Domiciliaire en Amérique Centrale	Généraliste : arbres, <i>Bromeliads A. butyracea</i> , <i>A. elegans</i> , <i>Co. tectorum</i> , <i>M. flexuosa</i> , <i>M. minor</i> , <i>Or. phalerata</i> , etc
<i>R. robustus</i> Larousse, 1927	Sporadique dans les domiciles	Généraliste : <i>Bromeliads A. butyracea</i> , <i>A. maripa</i> , <i>A. speciosa</i> , <i>O. bataua</i> , etc
<i>R. stali</i> Lent et al. 1993	En voie de domiciliation en Bolivie	<i>A. phalerata</i> (reporté une seule fois)
<i>R. zeledoni</i> Jurberg et al. 2009	Données insuffisantes	données insuffisantes
<i>Psammolestes arthuri</i> Pinto and Lent, 1935	Sylvatique	nids d'oiseaux
<i>Psammolestes coreodes</i> Bergroth, 1911	Sylvatique	nids d'oiseaux
<i>Psammolestes tertius</i> Lent and Jurberg, 1965	Sylvatique	nids d'oiseaux

Les espèces domiciliaires présentent la capacité de coloniser les environnements anthropisés et de s'établir dans cette nouvelle niche écologique, c'est-à-dire d'y réaliser leur cycle de vie. On y retrouve des œufs et tous les stades de développement de l'insecte. Ces espèces colonisent également le milieu péri-domiciliaire comme les poulaillers et tous les abris des animaux domestiques. Les espèces domiciliaires présentent des colonies établies en milieux anthropisés à partir d'habitats naturels ce qui a permis leur expansion dans des régions non endémiques depuis plusieurs décennies. Pour ces espèces, on ne trouve pas de populations en milieu naturel dans les zones non-endémiques alors qu'il peut encore y en avoir dans les zones endémiques. C'est par exemple le cas de *R. prolixus*, principal vecteur de la maladie de Chagas en Amérique Centrale où cette espèce est

largement domiciliée dans cette zone non-endémique alors qu'elle est encore présente en milieu sylvatique. En effet, après plusieurs années de controverses, il est maintenant bien établi que *R. prolixus* présente encore des populations sylvatiques au Venezuela qui peuvent ré-infester les habitations humaines après pulvérisations d'insecticides dans les maisons (Feliciangeli et al. 2007), mais également en Colombie (Guhl et al. 2009). D'autres espèces sont également considérées comme domicilières. C'est le cas de *R. ecuadoriensis* dans la zone nord du Pérou et de l'Équateur, de *R. stali* en Bolivie, et de *R. pallescens* au Panama.

Certaines autres espèces peuvent être en voie de domiciliation, c'est-à-dire que certaines populations peuvent se trouver localement en milieu domiciliaire alors que les populations restent majoritairement dans le milieu naturel. Au Brésil, des populations de *R. neglectus* et *R. nasutus* colonisent souvent les milieux humains et sont considérées en voie de domiciliation (Sarquis et al. 2004; Ramos et al. 2018).

D'autres encore peuvent être sporadiques. Les insectes visitent les habitations humaines pour se nourrir mais elles ne s'y établissent pas. C'est le cas de *R. robustus* (Lent et Wygodzinsky 1979; Galvão et Justi 2015) au Brésil, en Colombie, en Équateur, en Guyane française, au Pérou et au Venezuela.

Les espèces peuvent aussi être strictement retrouvées dans le milieu naturel et jamais reportées dans les habitations humaines, ce sont les espèces sylvatiques. C'est le cas par exemple de *R. barretti* (Abad-Franch et al. 2013) qui n'a jamais été trouvée dans les habitations humaines. C'est également le cas des espèces du genre *Psammolestes* qui ne sont retrouvées que dans les nids d'oiseaux (Lent et Wygodzinsky 1979).

Concernant les trois espèces *R. dalessandroi*, *R. micki* et *R. zeledoni*, il n'y a pas assez de données pour trancher sur leur statut écologique. *R. zeledoni*, morphologiquement proche de *R. paraensis*, n'a été décrit qu'à partir d'un seul spécimen mâle mort et séché (Jurberg et al. 2009), et *R. dalessandroi* n'est en fait connu que par sa description publiée (Carcavallo et Barreto 1976). L'espèce *R. micki* est décrite d'après deux spécimens trouvés en Bolivie et son écologie est encore mal connue (Zhao et al. 2021).

Le processus de réinfestation des habitations par des vecteurs provenant soit de l'habitat sylvatique, soit d'habitats domiciliaires ou péri-domiciliaires est à prendre en compte dans les campagnes d'éradication des vecteurs afin de coordonner les campagnes de pulvérisation d'insecticides dans les différents environnements anthropisés abritant des communautés connectées de triatomines. Ainsi, pour les populations sylvatiques de *R. prolixus*, les palmiers *Attalea* proches des habitats humains contribuent à l'infestation et à la réinfestation des maisons et des structures péri-domiciliaires (Abad-Franch et al. 2005;

de Arias et al. 2022), ou peuvent favoriser les visites sporadiques de *R. robustus* dans la région du tapajos au Brésil (Dias et al. 2014).

Il est important de bien connaître l'écologie des punaises et leur inféodation aux palmiers afin de mieux organiser les programmes d'éradication des punaises (Harry et al. 2000) d'autant plus qu'il suffirait d'un pool de quelques punaises pour recoloniser rapidement l'habitat (Schofield 2000; Gourbière et al. 2012). Des études de génétique de population ont révélé des différences entre populations de *R. prolixus* d'Amérique Centrale provenant du Honduras et de *R. prolixus* d'Amérique du Sud provenant de Bolivie (Dujardin et al. 1998a). Cette étude suggère que les populations de *R. prolixus* domiciliaires d'Amérique Centrale proviennent de populations d'Amérique du Sud qui auraient divergé génétiquement. Concernant *T. infestans*, les populations d'Uruguay sont génétiquement différentes de celles de Bolivie (Dujardin et al. 1998b). Les auteurs suggèrent que les populations uruguayennes proviennent de populations de Bolivie qui se seraient isolées à la suite de plusieurs goulets d'étranglement. Plusieurs études montrent que les habitations peuvent être ré-infestées par des individus sylvatiques comme en Argentine (Cardinal et al. 2008) et dans l'État de Rio Grande do Norte où on observe un chevauchement des cycles domiciliaires et sylvatiques de *T. cruzi* (Martins et al. 2015). C'est aussi le cas dans l'État de Rio Grande do Norte où il a été montré des flux de gènes entre des populations sylvatiques et domiciliaires de *T. brasiliensis*, ainsi que des effets de fondation multiples dans les habitations à partir d'individus sylvatiques. Ce processus menace les efforts de lutte anti-vectorielle car les populations sylvatiques sont difficilement contrôlables (Almeida et al. 2016). Il est donc crucial de prendre en compte les interconnexions entre les différents cycles domiciliaires et sylvatiques dans les programmes d'éradication des vecteurs afin d'éviter les ré-infestations rapides.

## 2.2. Les triatomines, des punaises hématophages

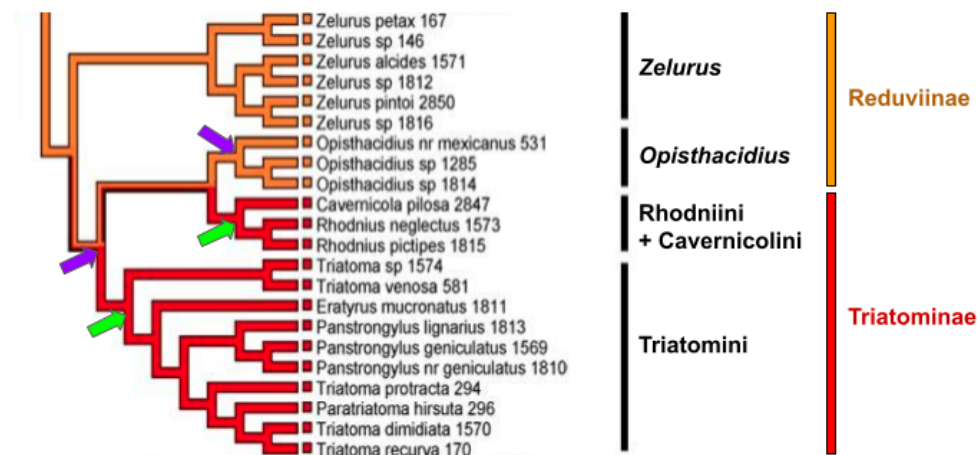
Tous les membres de la sous-famille des Triatomines sont des vecteurs potentiels de *T. cruzi*. Tous les stades et les deux sexes sont de potentiels vecteurs, du juvénile à l'adulte, mâle ou femelle. Leur cycle de vie comprend 5 stades juvéniles et leur temps de génération est de 3 à 6 mois (Gourbière et al. 2012). Les punaises sont des organismes paurométaboles hétérométaboles : il n'y a pas de larve ni de stade nymphal durant le développement mais des juvéniles qui partagent le même type d'habitat que les adultes. Contrairement aux moustiques ou aux taons par exemple, les mâles peuvent se nourrir sur des hôtes et donc être vecteurs de maladie. Outre les parasites *T. cruzi*, les triatomines peuvent faire circuler d'autres parasites comme *Trypanosoma rangeli* qui lui n'est pas responsable

de maladies chez l'Homme (Dias et al. 2010).

Les triatomines se nourrissent de sang de vertébrés. Ces insectes sont décrits comme étant obligatoirement hématophages et présentent des pièces buccales spécifiques. On peut observer une flexibilité dorsale accrue du dernier segment labial par rapport aux Reduviidae prédateurs d'arthropodes, et cet état de caractère synapomorphe a même été anciennement utilisé pour classer les triatomines par leur régime hématophage avant l'arrivée des analyses phylogénétiques (Monteiro et al. 2018). Cependant, d'autres modes d'alimentation, plus exceptionnels, ont été rapportés par certains auteurs. Certaines espèces de triatomines peuvent survivre à partir de l'hémolymphe d'insectes comme les espèces du genre *Belminus* (Lent et Wygodzinsky 1979). Pour l'espèce *Eratyrus mucronatus* vivant dans de grands arbres creux, il a été décrit que les adultes se nourrissent du sang de porcs-épics (*Coendou prehensilis*) alors que les jeunes stades peuvent se nourrir de l'hémolymphe de grands arachnides (*Amblypygi*) présents dans ce types d'arbres creux (Gaunt et Miles 2000). Cependant des essais d'alimentation en laboratoire sur des blattes n'ont pas montré ce type de comportement pour les espèces *T. infestans*, *R. stali* et *Panstrongylus rufotuberculatus* (Durán et al. 2016).

Une récente étude a montré qu'en conditions de laboratoire, l'espèce *R. prolixus* pouvait s'alimenter sur des tomates (Díaz-Albiter et al. 2016). Ces données pourraient remettre en question le dogme du régime strictement hématophage des triatomines.

Le régime hématophage chez certains réduves ancestraux prédateurs d'insectes serait apparu une ou plusieurs fois selon les auteurs (Hwang et Weirauch 2012). L'étude la plus complète contenant le plus d'espèces à ce jour propose deux scénarios plausibles différant selon la méthode utilisée (maximum de vraisemblance et méthode de parcimonie) (Figure 2). La méthode du maximum de vraisemblance suggère un scénario d'une transition unique vers l'hématophagie au nœud ancestral du clade Triatominae + *Opisthacidius* durant l'Oligocène (il y a entre 23 et 34 millions d'années) puis un retour à un comportement alimentaire généraliste chez *Opisthacidius* tandis que la méthode du maximum de parcimonie suggère que les tribus, Triatomini et Rhodniini + Cavernicolini ont acquis l'hématophagie indépendamment Hwang et Weirauch 2012. Cette étude confirme le caractère polyphylétique des triatomines proposé par Schofield 2000, qui suggérait que le caractère hématophage des punaises pouvait être apparu plusieurs fois indépendamment et que la sous-famille Triatominae décrite sur la base du régime alimentaire n'était donc peut-être pas monophylétique.



**Figure 2 : Différents scénarii évolutifs de l'acquisition de l'hématophagie proposés par Hwang et Weirauch 2012.** Les espèces en rouge sont hématophages (Triatominae) et les espèces en orange sont prédatrices d'arthropodes (Reduviinae). Scénario en vert : 2 acquisitions indépendantes de l'hématophagie, une acquisition au nœud Rhodniini + Cavernicolini et une acquisition au nœud Triatomini. Scénario en violet : une acquisition de l'hématophagie au nœud regroupant tous les Triatominae + *Opisthacidius* puis un retour à la prédation d'arthropodes au nœud *Opisthacidius*. Modifiée d'après Hwang et Weirauch 2012.

Plus de 180 espèces de mammifères sont décrites comme hôtes à partir du nombre d'espèces porteuses de *T. cruzi*. En revanche, les oiseaux, les amphibiens et les reptiles sont réfractaires au développement du parasite (Herrera 2010). Le nombre d'hôtes estimé en analysant moléculairement le repas de sang de 42 punaises des genres *Panstrongylus*, *Psammolestes*, *Rhodnius* et *Triatoma*, collectées en Colombie, va jusqu'à 67 (Arias-Giraldo et al. 2020). Deux études, l'une sur 181 et l'autre sur 60 individus de l'espèce *T. brasiliensis* collectés au nord du Brésil ont montré que les rongeurs *Galea spixii* (Rodentia ; Caviidae) et *Kerodon rupestris* (Rodentia ; Caviidae) sont des hôtes clés de ces espèces de punaises, constituant leur principale source alimentaire (et aussi des réservoirs importants de *T. cruzi*) (Lilioso et al. 2020; Almeida et al. 2016). A Ceará (nord du Brésil), 50 individus de l'espèce *T. brasiliensis*, 7 de *T. pseudomaculata*, et un de *R. nasutus* ont été collectés et l'analyse de leurs sources alimentaires a révélé que les animaux domestiques, tels que les poulets *Gallus gallus* (Galliformes ; Phasianidae) et les chèvres *Capra hircus* (Artiodactyla ; Bovidae), constituent leur principale source de nourriture (Valença-Barbosa et al. 2015). Au sein de la tribu Rhodniini, il existe des préférences d'hôtes. Comme mentionné ci-dessus, les punaises du genre *Psammolestes* se nourrissent principalement du sang des oiseaux de la famille des Furnariidae (Lent et Wygodzinsky 1979) dont elles colonisent le nid et où elles réalisent leur cycle de vie. Certains triatomines pratiquent la cleptohématophagie : ils se nourrissent sur des insectes eux-mêmes

hématophages (Sandoval et al. 2000).

Une étude sur *R. robustus* a montré que les individus pouvaient survivre à des périodes prolongées de jeûne : jusqu'à 189 jours pour les juvéniles et 137 jours pour les adultes (Pessanha et al. 2022). Le régime hématophage des punaises s'accompagne d'un microbiote intestinal particulier. Par exemple, *Rhodococcus rhodnii* (anciennement *Nocardia rhodnii*), décrit pour la première fois dans une étude du microbiote de *R. prolixus*, fournit à ses hôtes des vitamines du groupe B telles que la biotine, la nicotinamine, la thiamine, la pyridoxine ou la riboflavine (Baines 1956). Cette bactérie vit dans l'intestin moyen des punaises, et est transmise horizontalement par contamination de la surface des œufs ou par coprophagie (Wigglesworth 1936). Il a été montré des problèmes dans le développement et la survie des *Rhodnius* lorsqu'on élimine cette bactérie de leur microbiote (Salcedo-Porras et al. 2020). L'apport de vitamines par des bactéries, appelé "mutualisme nutritionnel", est très répandu chez les insectes et implique les capacités métaboliques du microbiote (Sudakaran et al. 2017).

L'étude récente du génome de *R. prolixus*, première espèce de triatomine dont le génome a été séquencé entièrement, a mis en évidence plusieurs extensions spécifiques de familles de gènes liées à l'alimentation sanguine (Mesquita et al. 2015).

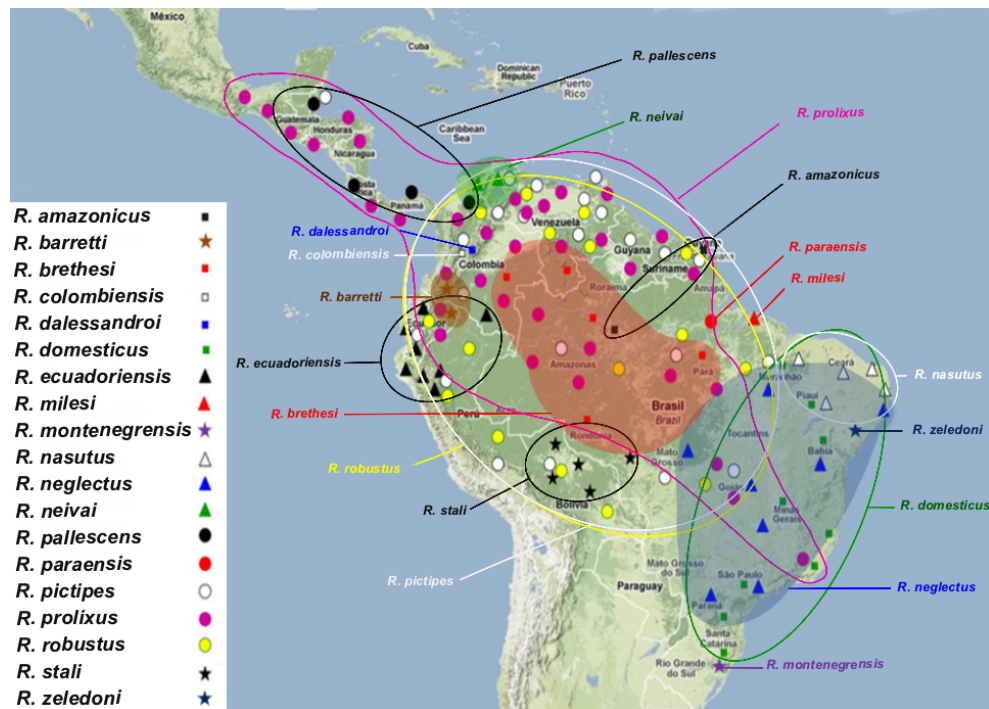
### 3. BIOGÉOGRAPHIE ET ÉVOLUTION DES VECTEURS DE LA MALADIE DE CHAGAS

#### 3.1. Répartition géographique des espèces de la tribu Rhodniini

Dans la famille des Triatominae on dénombre actuellement 157 espèces (154 espèces actuelles et 3 fossiles) (Alevi et al. 2021) réparties dans 18 genres et cinq tribus : *Alberproseniini*, *Bolboderini*, *Cavernicolini*, *Rhodniini* et *Triatomini*. Mais le nombre d'espèces décrites est susceptible d'évoluer grâce aux analyses biogéographiques et de génomique comparative. En effet, les analyses moléculaires des Triatomines invalident ou valident la pertinence d'espèces et de sous-espèces (Nascimento et al. 2019), ou peuvent permettre la découverte de nouvelles espèces.

Les Triatomines sont majoritairement en Amérique Latine et Amérique du Nord à l'exception de 8 espèces du genre *Triatoma* identifiées en Afrique, au Moyen-Orient, en Asie du Sud-Est et dans le Pacifique occidental et de 6 espèces du genre *Linshcosteus*, reportées en Inde (Gorla et al. 1997; Monteiro et al. 2018; Vieira et al. 2018). En ce qui concerne la tribu Rhodniini, qui comprend le genre *Rhodnius* (21 espèces) et le genre *Psammolestes* (3 espèces), son aire de répartition recouvre l'Amérique

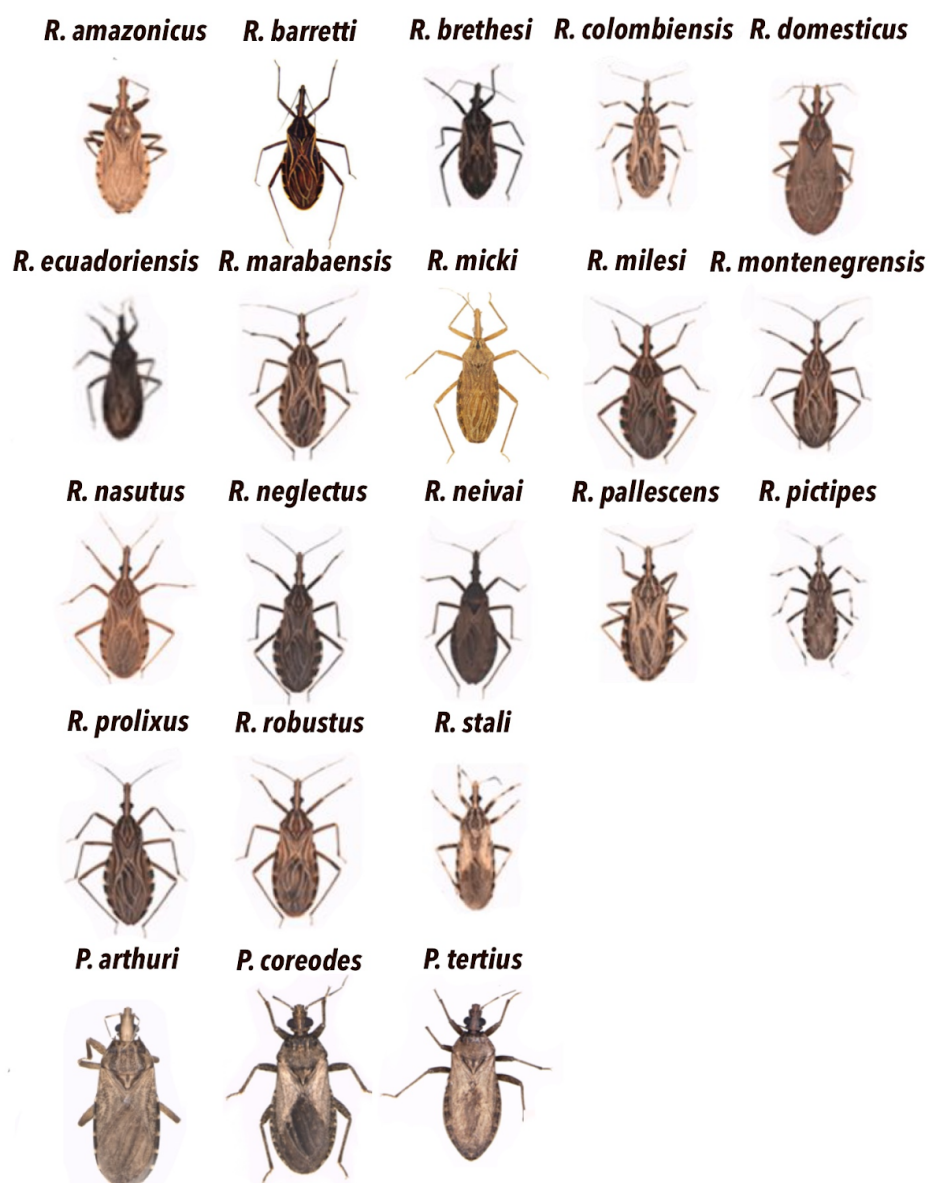
centrale et l'Amérique du Sud (Figure 3).



**Figure 3 :** Distribution géographique du genre des *Rhodnius*.

Il existe 3 groupes au sein du genre *Rhodnius* : le groupe pictipes (comprenant les espèces *R. amazonicus*, *R. brethesi*, *R. paraensis*, *R. pictipes*, *R. stali* et *R. zeldoni*), le groupe pallescens (comprenant les espèces *R. colombiensis*, *R. ecuadoriensis* et *R. pallescens*) et le groupe prolixus (comprenant les espèces *R. barretti*, *R. dalessandroi*, *R. domesticus*, *R. neivai*, *R. marabaensis*, *R. milesi*, *R. montenegrensis*, *R. neglectus*, *R. nasutus*, *R. robustus* et *R. prolixus*). *R. micki* est une espèce décrite en 2021 à partir d'un seul individu et des observations morphologiques suggèrent que cette espèce fait partie du groupe pictipes. Des données moléculaires seraient nécessaires pour confirmer cette appartenance (Zhao et al. 2021). De même pour *R. dalessandroi* et *R. zeldoni*, il n'existe pas de données moléculaires. Une espèce supplémentaire aux 21 nommées ci-dessus (Figure 4), *R. taquarussuensis* (da Rosa et al. 2017), a été proposée mais d'autres auteurs ont souligné des ressemblances avec *R. neglectus* : *R. taquarussuensis* et *R. neglectus* sont désormais considérés comme synonymes (Nascimento et al. 2019).





**Figure 4** : Les différentes espèces de *Rhodnius* et *Psammolestes* (tribu Rhodniini) d'après [Hernández et al. 2020](#) pour toutes les images sauf [Zhao et al. 2021](#) (pour *R. micki*), [Abad-Franch et al. 2013](#) (pour *R. barretti*), [Oliveira et al. 2021](#) (pour *P. tertius*) et photographies de *P. arthuri* et *P. coreodes* par Dr. J. de Oliveira. Aucune photographie n'est disponible pour *R. dalessandroi*, *R. paraensis* et *R. zeledoni*.

Ces trois groupes ont d'abord été décrits sur la base de données morphologiques mais aussi sur des données géographiques. En effet, le groupe *pallescens* a une aire de répartition limitée à l'Amérique Centrale et à l'ouest des Andes (Figure 3), d'où il tient son nom de groupe "trans-andins" aussi appelé groupe "trans-pallescens" par certains auteurs. Les sept espèces du groupe *pictipes* (*R. amazonicus*, *R. brethesi*, *R. micki*, *R. paraensis*, *R. pictipes*, *R. stali*, et *R. zeledoni*) sont cis-andines, distribuées à l'est des Andes et dans la région amazonienne. La même distribution est observée pour dix espèces du groupe *prolixus* (*R. barretti*, *R. dalessandroi*, *R. domesticus*, *R. marabaensis*, *R. milesi*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. prolixus* et *R. robustus*), la onzième *R. neivai* ayant également quelques populations trans-andines (Zhao et al. 2021).

### 3.2. Une histoire évolutive des *Rhodniini* non résolue

Il existe plusieurs conflits taxonomiques au sein de la tribu *Rhodniini*. Ici, je présente une petite synthèse des problèmes rencontrés par la communauté scientifique sur les *Rhodniini* tels que la relation entre les genres *Psammolestes* et *Rhodnius* et leur statut, le statut de certaines espèces au sein du complexe *prolixus-robustus* et les relations entre les différents groupes *pallescens*, *pictipes* et *prolixus*.

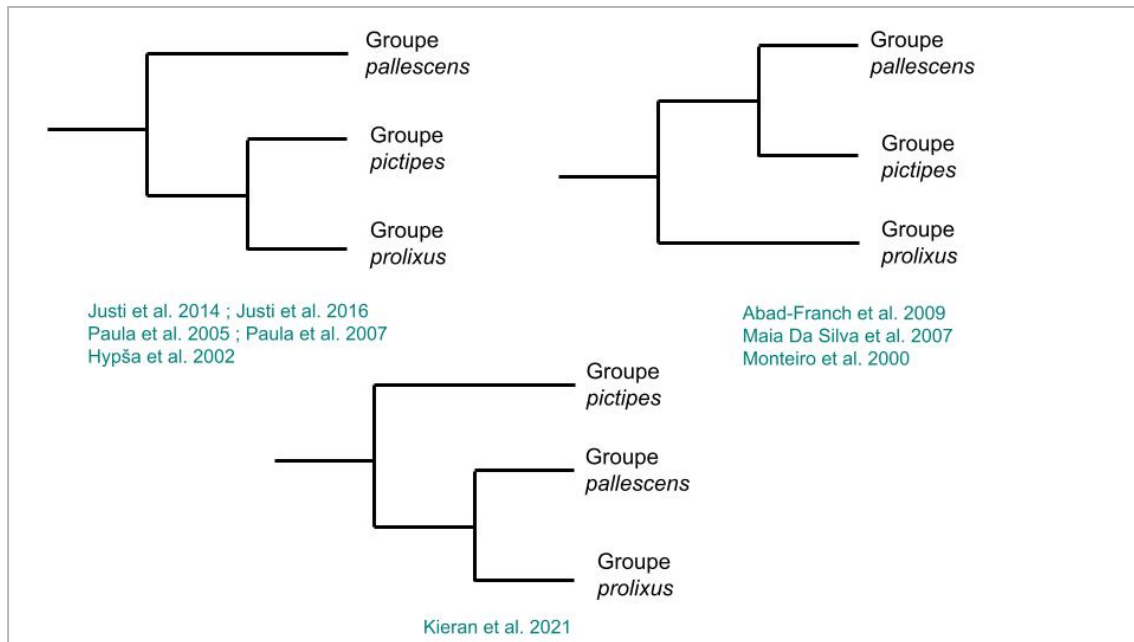
Les premières études phylogénétiques sur les punaises de la tribu *Rhodniini* ont été réalisées à partir de données d'électrophorèse de protéines, les allozymes et isozymes (Dujardin et al. 1998; Harry et al. 1992; Monteiro et al. 2002). Des espèces de *Rhodnius* ont aussi été identifiées à l'aide d'études morphologiques, morphométriques comme l'étude de la forme de l'aile de spécimens qui a permis de confirmer la présence de *R. robustus* en Bolivie (Matias et al. 2001), la taille du processus médian du pygophore (dernier segment abdominal contenant les organes génitaux) qui a permis d'identifier plusieurs espèces de *Rhodnius* (Lent et Wygodzinsky 1979; Harry 1993) et de *Triatoma* (Mendonça et al. 2014) et des études des protéines salivaires sur gel d'électrophorèse (Soares et al. 1998; Soares et al. 2001).

Les punaises du genre *Psammolestes* ont une écologie différente des punaises du genre *Rhodnius*. Elles vivent dans les nids d'oiseaux et s'alimentent sur ces hôtes tandis que les punaises du genre *Rhodnius* vivent dans les palmiers et/ou les habitations humaines et se nourrissent sur une large gamme d'hôtes vertébrés (Arias-Giraldo et al. 2020). Des différences morphologiques entre *Rhodnius* et *Psammolestes* s'ajoutent à cela (Lent et Wygodzinsky 1979). Des études sur les allozymes montrent cependant que *P. tertius* appartiendrait au genre *Rhodnius* (Monteiro et al. 2002). Des phylogénies à partir de données mitochondriales ou combinant des gènes mitochondriaux et nucléaires (16S, 18S et Wg concaténés) montrent que *P.*

*coreodes* serait dans le groupe *prolixus* (Lyman et al. 1999; Hypša et al. 2002) ainsi que *P. tertius* (Hypša et al. 2002; de Paula et al. 2021).

Certaines espèces sont difficiles à différencier les unes des autres sur des critères morphologiques générant ainsi des conflits taxonomiques. Des études pionnières utilisant des allozymes ont remis en question le statut d'espèce pour des taxons étroitement apparentés telles que les populations vénézuéliennes de *R. robustus* et *R. prolixus* (Harry et al. 1992; Harry 1993). Il a ensuite été montré, en utilisant le gène mitochondrial *cytb* pour reconstruire des phylogénies, que *R. robustus* peut être scindé en 4 clades, ROB I, ROB II, ROB III et ROB IV (Monteiro et al. 2003; Brito et al. 2019). ROB I correspond aux spécimens de *R. robustus* provenant du Venezuela et groupés avec des spécimens de *R. prolixus* du Venezuela, ce qui pourrait suggérer des événements d'introgression entre les espèces *R. prolixus* et *R. robustus* au Venezuela (Brilo et al. 2019). De plus, *R. montenegrensis* est groupe frère de spécimens de *R. robustus* appartenant au clade ROB II, ce qui suggère soit une synonymie soit des événements d'introgression. Des problèmes d'identification des espèces ont été soulevés par plusieurs auteurs notamment des problèmes d'introgression dans les élevages de *R. prolixus* et la difficulté de trouver des lignées pures de *R. prolixus* sans croisement avec des *R. robustus* (Mesquita et al. 2015; Brito et al. 2019).

Les liens phylogénétiques entre les différents groupes n'ont pas fait consensus au sein de la communauté scientifique. S'appuyant indépendamment sur des données mitochondriales ou mitochondriales et nucléaires, certains auteurs proposent le groupe *pictipes* comme groupe frère du groupe *prolixus* (Abad-Franch et al. 2009; Maia Da Silva et al. 2007; Monteiro et al. 2000) tandis que d'autres auteurs proposent le groupe *pictipes* comme groupe frère du groupe *pallescens* (Justi et al. 2014; Justi et al. 2016; de Paula et al. 2005; de Paula et al. 2007; Hypša et al. 2002) (Figure 5). Les auteurs soutenant la première hypothèse proposent que l'événement de spéciation entre le groupe *pallescens* transandin, situé à l'ouest de la cordillère des Andes et les groupes *pictipes* - *prolixus* serait de type allopatrique lié à la formation de la cordillère des Andes, notamment la formation du système Pebas entre 23 et 10 millions d'années durant la période du Miocène (Justi et al. 2016). Une troisième topologie a été proposée avec le groupe *pallescens* comme groupe frère du groupe *prolixus* à partir d'éléments ultra conservés (Kieran et al. 2021) mais avec des données fragmentaires.



**Figure 5 : Synthèse de l'état de l'art concernant les différents groupes du genre *Rhodnius*.** Les études portant sur la phylogénie des espèces ne font pas consensus en ce qui concerne les relations entre les 3 groupes *pallescens*, *pictipes* et *prolixus*.

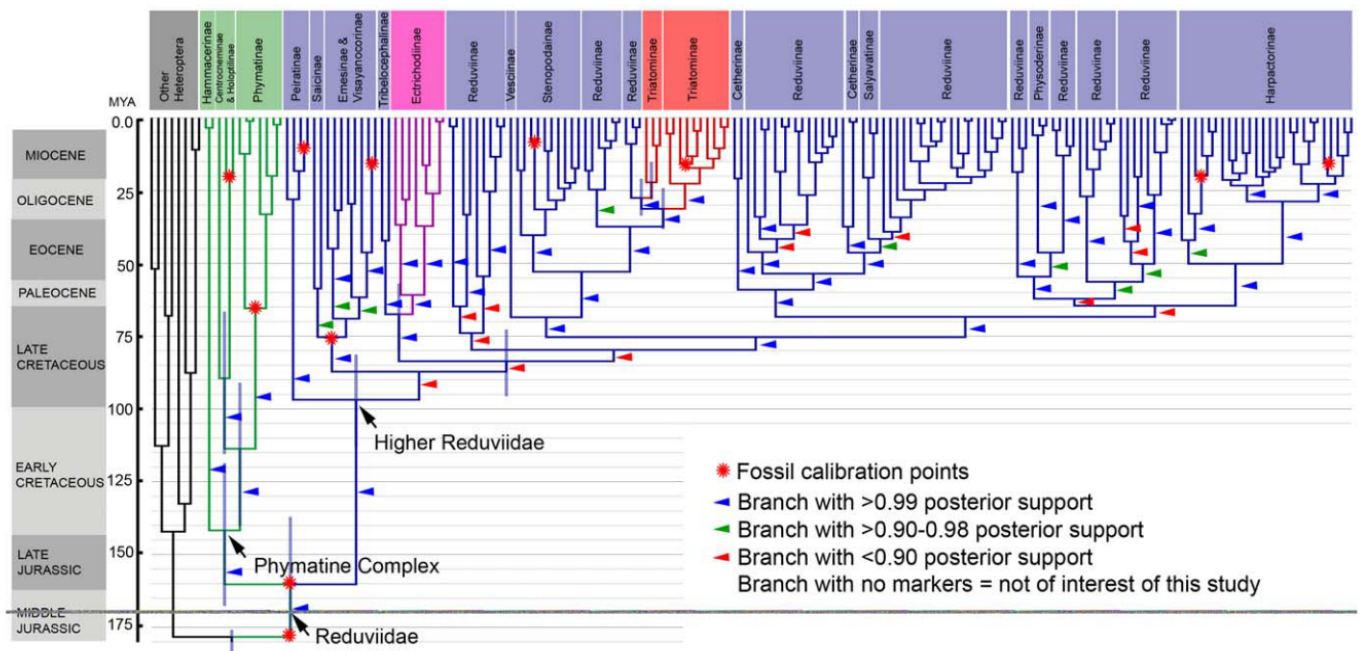
Les différences de topologie pour les trois groupes de *Rhodnius* peuvent s'expliquer par plusieurs biais. Le choix du groupe externe pour raciner la phylogénie est primordial car il peut changer la topologie de l'arbre. Ainsi, les espèces de *Psammolestes* qui ont été reportées comme appartenant au genre *Rhodnius*, ne feraient donc pas de bon candidats comme groupe externe pour raciner les phylogénies *Rhodnius* (Lyman et al. 1999; Monteiro et al. 2000; Hypša et al. 2002; Justi et al. 2014; de Paula et al. 2021). Les biais méthodologiques peuvent aussi être induits par des échantillonnages de taxons incomplets/différents ainsi que par le type et le nombre de marqueurs utilisés. Certaines discordances peuvent s'expliquer par le type de données utilisées, données mitochondriales et/ou nucléaires. Ces discordances entre les phylogénies mitochondriales et nucléaires peuvent refléter la réalité biologique. En effet, le génome mitochondrial est haploïde, hérité de façon unilatérale et a un taux d'évolution rapide contrairement au génome nucléaire. D'autres raisons peuvent être un échantillonnage incomplet des différents groupes de *Rhodnius*, des transferts horizontaux de gènes, l'introgression, ou encore l'hybridation (Toews et Brelsford 2012; Campillo et al. 2019).

### 3.3. Datation des événements de spéciation

Peu d'études portent sur l'histoire évolutive des Triatominae en

proposant des datations d'événements de spéciation. Les temps de divergence estimés entre les Triatominae et les autres réduves varient énormément selon les études, allant de 4 à 107 millions d'années (Gaunt et Miles 2002; Hwang et Weirauch 2012; Patterson et Gaunt 2010; Schofield et Galvão 2009). Pour étudier les temps de divergence entre les espèces, la méthode est de combiner des données fossiles avec des données moléculaires, afin de réaliser une phylogénie calibrée grâce aux données fossiles. Un écueil à l'estimation des temps de divergence au sein des triatomines est le faible nombre de fossiles décrits. Il existe trois fossiles de triatomines : *Triatoma dominicana* (Poinar 2005), *Paleotriatoma metaxytaxa* (Poinar 2019) et *P. hispaniolae* (Poinar 2013), tous les trois conservés dans l'ambre. Le nombre de points de calibration fossile a été démontré empiriquement comme crucial pour une datation fiable des phylogénies (Zheng et Wiens 2015).

Les datations utilisant des données moléculaires et combinant des fossiles montrent une origine des Triatomines autour de 35 à 40 millions d'années (Hwang et Weirauch 2012; Ibarra-Cerdeña et al. 2014; Justi et al. 2016). Le plus grand ensemble de données publié à ce jour est l'étude de Hwang et Weirauch 2012 comprenant 229 Reduviidae dont 70 espèces de triatomines des tribus Cavernicolini, Rhodniini et Triatomini à partir d'une concaténation de gènes mitochondriaux et nucléaires (16S rDNA, 18S rDNA, 28S D2 rDNA, 28S D3-D5 rDNA et wingless). Des données d'observation de 11 fossiles sont incluses dans cette analyse. L'âge de l'ancêtre des triatomines y est estimé entre 27 et 32 millions d'années (Figure 6).



**Figure 6 : Estimation du temps de divergence des Reduviidae par une analyse avec le programme BEAST.** Cette analyse se base sur une phylogénie réalisée à partir de gènes mitochondriaux et nucléaires (16S rDNA, 18S rDNA, 28S D2 rDNA, 28S D3-D5 rDNA, wingless) concaténés, calibrée avec des données fossiles. Les 11 points de calibration fossiles sont notés avec des étoiles rouges. Les lignes sont colorées sur le chronogramme : taxons du groupe externe (noir), complexe Phymatine (vert), Ectrichodiinae (rose), Triatominae (rouge), autres sous-familles de réduviidés (bleu). Cette datation permet de dater l'ancêtre des Triatominae à 27–32 millions d'années. [Hwang et Weirauch 2012.](#)

## 4. LES APPORTS DES DONNÉES MOLÉCULAIRES SUR L'ÉTUDE DES INSECTES

### 4.1. Les omiques

L'ajout du préfixe « omique » à un terme moléculaire signifie une évaluation complète ou globale d'un ensemble de molécules ([Misra et al. 2019](#)). Les omiques regroupent plusieurs disciplines telles que la génomique, l'épigénomique, la transcriptomique, la protéomique, la microbiomique et la métabolomique ([Misra et al. 2019](#); [Hasin et al. 2017](#)). Le développement des différentes techniques haut débit (notamment le séquençage, la spectrométrie de masse et la résonance magnétique nucléaire) et le développement des disciplines bioinformatiques ont contribué à l'augmentation des approches omiques pour répondre aux questions scientifiques. L'approche omique a pour but par exemple de relier le génotype au phénotype, quantifier le protéome et le métabolome

mais aussi d'étudier l'organisation et l'évolution des génomes, analyser leur expression selon les stades de développement, les tissus, les variétés ou populations et les conditions environnementales.

Ces approches sont très utilisées notamment dans le domaine médical. La réduction des coûts a permis ces dernières années d'étendre l'usage des données omiques à de très nombreuses espèces, et particulièrement aux espèces non modèles. Dans le cadre des interactions entre les insectes et leur environnement, ce type d'approche a par exemple permis d'étudier les relations entre insectes et plantes hôtes ainsi que les défenses mises en place par celles-ci. Les variations mises en évidence peuvent être génétiques (génomique des populations), ou phénotypiques : changement de profil d'expression des gènes (études transcriptomiques) et/ou changement métabolique (études métabolomiques) (Barah et Bones 2015).

L'approche omique doit cependant relever plusieurs défis tels que la prise en charge d'énormes jeux de données, leur archivage, une puissance de calcul importante et des problèmes de puissance statistique. Pour répondre à ces problématiques, des plateformes de bioinformatiques fleurissent telles que Genouest à Rennes, Genotoul à Toulouse qui proposent des connexions à distance sur des serveurs afin d'effectuer des assemblages génomiques ou des annotations de génome par exemple.

Le séquençage à haut débit permet d'assembler le génome de différentes espèces et ainsi d'éclaircir leur histoire évolutive, par exemple en étudiant les contractions et expansions de gènes en lien avec l'environnement. Les transferts horizontaux (éléments transposables ou éléments transférés par les virus) peuvent aussi être mis en lumière grâce à la génomique comparative. Ces études dépendent du séquençage du génome des espèces ce qui explique la corrélation entre le développement des nouvelles techniques de séquençage et la quantité de publications sur le sujet.

#### 4.2. Les défis du séquençage et de l'annotation des génomes

L'assemblage *de novo* des génomes dépend entièrement du chevauchement des informations entre les lectures. Il en résulte que la profondeur de séquençage, c'est-à-dire le nombre de fois en moyenne qu'un locus dans le génome est lu par le séquenceur, doit être assez élevée pour pouvoir assembler les génomes. La méthodologie d'assemblage *de novo* dépendante des chevauchements entre lectures rend difficile l'assemblage des séquences répétées. La seconde méthode appelée assemblage de cartographie (ou assemblage par référence) détermine d'abord la position des lectures par rapport au génome de référence, puis

assemble les lectures en *contigs* (Li et al. 2019).

Un des défis du séquençage du génome des insectes est que parfois la quantité d'ADN contenu dans un spécimen n'est pas assez élevée pour avoir une profondeur de séquençage suffisante pour réaliser un assemblage de qualité du génome. Il faut dans ce cas regrouper plusieurs spécimens. Cependant, ces spécimens doivent être génétiquement identiques pour que la couverture augmente. Si les spécimens à combiner avant l'extraction d'ADN sont trop différents génétiquement, cela revient à séquencer plusieurs génomes différents. C'est la raison pour laquelle la plupart des insectes dont le génome a été séquencé proviennent d'élevages, ce qui assure la consanguinité et une plus grande homozygotie des insectes.

La qualité du génome assemblé est souvent évaluée par le N50, sa complétude, et la taille totale de l'assemblage. Le N50 désigne la taille du contig pour laquelle la moitié des bases de l'assemblage est comprise dans les *contigs* plus grands que cette valeur. Plus le N50 est grand, plus les contigs sont de grande taille, plus l'assemblage est de bonne qualité. Le logiciel BUSCO (Simão et al. 2015) pour *Benchmarking Universal Single-Copy Orthologs* permet de vérifier la complétude du génome. Pour ce faire, le logiciel cherche dans l'assemblage certains gènes d'un set de gènes présents en copie unique chez tous les insectes. Les gènes de référence doivent être retrouvés entiers et en un seul exemplaire dans le génome assemblé. La taille totale de l'assemblage peut être croisée avec d'autres mesures indépendantes de la taille du génome, comme la mesure de la taille des noyaux de cellule estimée par cytométrie de flux.

L'annotation des génomes dépend à la fois de leur complétude et de la taille des *contigs* mais aussi de la manière d'annoter. Il y a plusieurs méthodes pour cela. La première, la prédiction génique *ab initio* qui fait référence à l'identification de gènes codant pour des protéines à l'aide des signaux qui définissent et caractérisent leurs structures génétiques le long du génome. Elle permet d'identifier les signaux correspondant à une séquence codante. Cette technique utilisée seule fonctionne bien mieux pour l'annotation des génomes de bactéries que les génomes d'eucaryotes à cause de la longueur des gènes et de la structure intron-exon des gènes des eucaryotes (Song et al. 2012). La seconde méthode consiste à utiliser les données séquençage du transcriptome (de la même espèce ou d'une espèce proche) et de les aligner sur le génome afin de trouver les séquences transcrites et donc trouver plus facilement les séquences codantes. Cette technique est plus rapide et précise pour les eucaryotes mais nécessite un séquençage exhaustif du transcriptome pour l'espèce donnée ou pour une espèce proche (Song et al. 2012). Une dernière manière d'annoter les génomes est de transférer l'annotation déjà réalisée d'une espèce proche sur le génome de l'espèce à annoter. Cette méthode fait l'hypothèse que des espèces proches ont les mêmes gènes. Cette méthode ne permet donc pas de trouver des gènes spécifiques à l'espèce à



étudier.

#### 4.3. Les contributions du séquençage

La variabilité du contenu génique du nombre de gènes est un phénomène connu et bien documenté chez les bactéries. La part de gènes partagés par toutes les bactéries d'un même taxon est appelée core-genome. La totalité des gènes de toutes les bactéries d'un taxon est appelée pan-génome et le génome accessoire est la part de gènes qui n'est pas retrouvée chez tous les spécimens d'un taxon (Brockhurst et al. 2019). Certains gènes ou de plus longs fragments d'ADN contenant plusieurs gènes peuvent être transférés par transfert horizontal par des bactériophages ou par conjugaison permettant des gains de fonction chez les bactéries réceptrices, leur permettant d'améliorer leur pathogénicité, d'acquérir des résistances ou de s'adapter à une nouvelle niche écologique en infectant un autre type d'hôte.

Chez les insectes, ce phénomène est moins courant mais il peut y avoir des transferts horizontaux et des contractions du nombre de gènes. Les études de génomique comparative, qui nécessitent le séquençage des génomes, permettent d'identifier les séquences acquises par transferts horizontaux comme des éléments transposables ou des virus qui peuvent s'intégrer dans le génome et devenir des virus endogènes (Feschotte et Gilbert 2012). Les processus de duplication de génomes et de gènes sont bien illustrés et aboutissent à des variations importantes du nombre de gènes. Grâce au séquençage du génome de 12 espèces de drosophiles, une étude pionnière a permis de mettre en évidence que le nombre de pertes et gains par gène par million d'années n'était pas homogène dans le genre *Drosophila*. Ce taux varie entre 0,0006 et 0,0193 pertes et gains par gène et par million d'années, soit un facteur 32 entre les branches présentant des vitesses d'évolution différentes (Hahn et al. 2007).

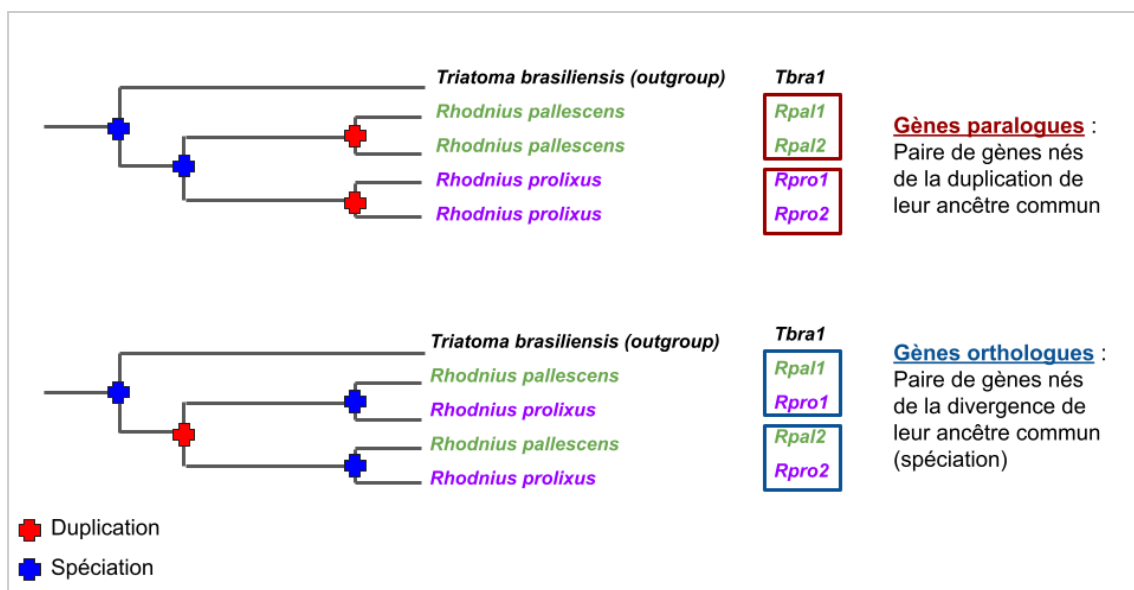
Le séquençage du génome de *Ceratosolen solmsi* (Hymenoptera ; Agaonidae) a révélé que la spécialisation à une plante hôte, ici le figuier sur lequel l'insecte est spécialisé depuis 75 millions d'années, entraîne une réduction du nombre de gènes dans certaines familles multigéniques impliquées dans la détection et la détoxification de l'environnement (Xiao et al. 2013).

#### 4.4. Évolution des familles de gènes

Une famille multigénique est composée de gènes ayant une origine commune. Ils sont issus de gènes ancestraux qui se sont dupliqués. Les gènes appartenant à une même famille multigénique peuvent être

distribués en tandem ou être très éloignés (sur des chromosomes différents par exemple), du fait des événements de recombinaison qui ont suivi la duplication ce qui complique parfois l'annotation des génomes. En effet, les séquences codantes répétées sont mal gérées par les programmes de recherche de gènes et mal annotées (Karpe et al. 2021). La copie du gène nouvellement formé peut rester active et donc être redondante ou acquérir une nouvelle fonction (synthèse d'une protéine fonctionnellement différente). Mais souvent les nouvelles copies vont accumuler des mutations qui peuvent entraîner des décalages du cadre de lecture et provoquer l'apparition de codons stop. Dans ce dernier cas, les gènes dupliqués deviennent des pseudogènes non-fonctionnels (et ultérieurement non traduits) (Kahn 1985).

Les gènes issus d'une duplication au sein d'un génome sont appelés paralogues. Les gènes orthologues sont quant-à-eux des gènes qui ont pour origine un gène ancestral transmis à deux (ou plusieurs) espèces lors d'événements de spéciation (Figure 7). Les concepts d'orthologie et de paralogie sont des éléments clés de l'étude de l'évolution des gènes (Koonin 2005). Les gènes paralogues peuvent constituer des expansions géniques lorsque les duplications sont spécifiques à une lignée. La pression de sélection positive tend alors à conserver les duplications et conduit à une néofonctionnalisation de la copie.



**Figure 7 : Arbres phylogénétiques hypothétiques illustrant les relations d'orthologie et de paralogie entre deux gènes et leurs descendants chez deux espèces (*R. pallescens* et *R. prolixus*).** Les duplications sont indiquées par des croix rouges et les événements de spéciation par les croix bleues. (Figure inspirée de la figure de Jean-Phillipe Doyon).

#### 4.5. Cytométrie de flux, études génomiques et transcriptomiques des triatomines

Avant le développement des méthodes de séquençage à haut débit, différentes études de cytologie ont été réalisées chez les Triatominae. Les premières études de caryotype des espèces du genre *Rhodnius*, *R. pictipes*, *R. nasutus* et *R. robustus* ont montré que ces espèces possèdent 20 autosomes plus les deux chromosomes sexuels XY. Parmi les 102 espèces de triatomines étudiées à ce jour, 99 d'entre elles présentent 20 autosomes. Les trois autres espèces sont *T. rubrofasciata* (avec 22 autosomes), *Panstrongylus megistus* et *T. nitida* (avec 18 autosomes) (Panzera et al. 2021). Trois systèmes sexuels chez les mâles sont observés chez les Triatominae (XY, X<sub>1</sub>X<sub>2</sub>Y, et X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>Y) (Panzera et al. 2021). Un mécanisme sexuel particulier, X<sub>1</sub>X<sub>2</sub>Y<sub>1</sub>Y<sub>2</sub> a été décrit chez *Mepraia spinolai* (Frias et Atria 1998).

Des études de cytométrie de flux ont permis d'estimer la taille du génome en picogramme pour certaines espèces de triatomines (Panzera et al. 1995). Pour les espèces du genre *Rhodnius*, la taille du génome a été estimée pour quatre espèces : *R. colombiensis* a un génome de 0,59 pg (Díaz et al. 2014), *R. ecuadoriensis* de 0,72 pg (Panzera et al. 2007), *R. pallescens* de 0,73 pg (Gómez-Palacio et al. 2012) et *R. prolixus* de 0,75 pg (Panzera et al. 2007). La taille du génome de deux espèces de *Panstrongylus* et de 14 espèces de *Triatoma* a été estimée et varie de 0,91 pg pour *T. longipennis* à 2,9 pg pour *T. delpontei* (Panzera et al. 2010). Ces données sont accessibles via le site genomesize.com (Tableau 2).

**Tableau 2 : Synthèse des données bibliographiques sur les analyses de taille de génome par cytométrie de flux** (Gregory, 2022 genomesize.com). Les méthodes utilisées sont indiquées par FCM = Cytométrie en Flux, NS = non-spécifié, FIA = Feulgen Image Analysis Densitometry, FD = Feulgen Densitometry. Les types de cellules sont indiquées par TS = testicules, S = sperm, NS = non-spécifié. Les espèces de références sont notées HS = *Homo sapiens*, NS = non-spécifié, DM = *Drosophila melanogaster*, AC = l'oignon *Allium cepa* (Liliales ; Liliaceae).

Espèces	C-value (pg)	Méthode	Type de cellules	Espèce de référence	Références
<i>Dipetalogaster maximus</i>	1,29	FCM	TS	HS	Panzera et al. (2007)
<i>Eratyrus cuspidatus</i>	1,26	FCM	TS	HS	Panzera et al. (2007)
<i>Panstrongylus geniculatus</i>	1,42	FCM	TS	HS	Panzera et al. (2007)

<i>Panstrongylus rufotuberculatus</i>	1,44	FCM	TS	HS	Panzer et al. (2007)
<i>Rhodnius colombiensis</i>	0,58	FCM	TS	HS	Diaz et al. (2014)
<i>Rhodnius ecuadoriensis</i>	0,72	FCM	TS	HS	Panzer et al. (2007)
<i>Rhodnius pallescens</i>	0,73	FCM	TS	HS	Gomez-Palacio et al. (2012)
<i>Rhodnius prolixus</i>	0,26	NS	NS	NS	Petitpierre (1996)
<i>Rhodnius prolixus</i>	0,59	FIA	S	DM	Gregory. Unpublished Data
<i>Rhodnius prolixus</i>	0,75	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma barberi</i>	1,23	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma brasiliensis</i>	1,09	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma delpontei</i>	2,90	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma delpontei</i>	1,80	FD	NS	AC	Panzer et al (1995)
<i>Triatoma dimidiata</i>	1,07	FCM	TS	HS	Panzer et al (2006)
<i>Triatoma dimidiata</i>	0,92	FCM	TS	HS	Panzer et al (2006)
<i>Triatoma dimidiata</i>	0,98	FCM	TS	HS	Panzer, et al. (2006)
<i>Triatoma infestans</i>	1,03	FD	NS	AC	Panzer et al. (1995)
<i>Triatoma infestans</i>	1,70	FCM	TS	HS	Panzer et al. (2004)
<i>Triatoma infestans</i>	1,98	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma infestans</i>	1,44	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma infestans</i>	1,50	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma longipennis</i>	0,91	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma maculata</i>	1,09	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma melanosoma</i>	1,53	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma nitida</i>	1,35	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma platensis</i>	1,33	FCM	TS	HS	Bargues et al (2006)
<i>Triatoma platensis</i>	0,87	FD	NS	AC	Panzer et al. (1995)
<i>Triatoma pseudomaculata</i>	1,13	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma rubrovaria</i>	1,17	FCM	TS	HS	Bargues et al. (2006)
<i>Triatoma rubrovaria</i>	0,54	FD	NS	AC	Panzer et al. (1995)
<i>Triatoma ryckmani</i>	1,10	FCM	TS	HS	Panzer et al. (2007)
<i>Triatoma sordida</i>	1,12	FCM	TS	HS	Bargues et al. (2006)

En ce qui concerne les données mitochondriales, les premières données sur les triatomés ont concerné l'espèce *T. dimidiata* (Dotson et Beard 2001) puis *T. infestans* (Pita et al. 2017). La démocratisation des

nouvelles techniques de séquençage a permis d'accroître à 27 le nombre d'espèces de triatomines dont le mitogénome est disponible sur NCBI (consulté en novembre 2022) dont 4 *Rhodnius* (*R. neglectus*, *R. pallescens*, *R. pictipes* et *R. prolixus*).

Peu de génomes nucléaires complets sont disponibles pour les espèces de la sous-famille Triatominae, seulement 3 en novembre 2022. Le premier génome nucléaire complet séquencé et annoté d'une espèce de triatomine a été celui de *R. prolixus* (Mesquita et al. 2015), résultat du travail d'un consortium de scientifiques de plus de 10 pays. Le génome de *R. prolixus* fait 702 Mb. Son annotation a aussi été réalisée et a mis en évidence 15 456 gènes putatifs. Un accent sur les voies de l'immunodéficiences, de la détoxification, des gènes chimiosensoriels et des gènes impliqués dans la digestion du sang a été mis dans cette étude. Le génome de *R. prolixus* contient également des gènes de la bactérie *Wolbachia* acquis horizontalement et intégrés dans le génome de l'insecte. Le second génome de triatomine assemblé est celui de *Triatoma rubrofasciata* (Liu et al. 2019), qui fait 680 Mb. Cette espèce est la seule espèce de Triatominae ayant une distribution pantropicale et est essentiellement trouvée en Amérique, en Asie, en Océanie et en Afrique (Lent et Wygodzinsky 1979). Cette espèce a colonisé l'Asie il y a 300 ans à la suite d'une importation involontaire (Monteiro et al. 2018). Ce génome est assemblé au niveau chromosomique. Le dernier génome de triatomines accessible à ce jour est celui de *Triatoma infestans* déposé sur NCBI en 2020 par l'Université du Vermont. Ce génome fait 949 Mb.

Le transcriptome de certaines espèces a également commencé à être étudié. Différentes études de sialotranscriptomes, qui permettent de connaître la composition en protéines de la salive ont été réalisées, comme chez *T. brasiliensis* (Santos et al. 2007) et *T. dimidiata* (Kato et al. 2010). Pour cette dernière espèce, l'étude du sialo-transcriptome en permettant de connaître notamment les protéines allergènes présentes dans la salive de ces espèces hématophages fournit des marqueurs d'exposition à *T. dimidiata* pour les études épidémiologiques. Le transcriptome de *R. prolixus* a été séquencé et analysé (Ribeiro et al. 2014) et un an plus tard, un transcriptome chimiosensoriel de *T. brasiliensis* a été publié (Marchant et al. 2015). Le séquençage et l'assemblage des transcrits permet de faciliter l'annotation des génomes en alignant les transcrits sur le génome ce qui permet de trouver les régions transcrites et donc de potentielles séquences codantes. Ces assemblages permettent aussi de faire des analyses transcriptomiques, telles que des comparaisons de répertoires de transcrits.

Une étude chez *T. brasiliensis*, entre populations sylvatiques et domiciliaires, montre que certains gènes chimiosensoriels ont une expression différentielle selon l'habitat (Marchant et al. 2016). Plus récemment, différents transcriptomes de triatominae ont été séquencés : *R.*

*montenegrensis* et *R. robustus* dans une analyse comparative des transcriptomes (Carvalho et al. 2017), *R. prolixus*, dans beaucoup d'études d'expressions lors du développement ou des repas de sang (ex : Latorre-Estivalis et al. 2020; Leyria et al. 2020; Ribeiro et al. 2014), *T. infestans* (Gonçalves et al. 2017; Traverso et al. 2016), *Panstrongylus lignarius*, vecteur principal de la maladie de Chagas au Pérou (Nevoa et al. 2018) et *T. rubrofasciata* (Mizushima et al. 2020). Une analyse comparative du transcriptome de *R. prolixus* (espèce domiciliaire) et *R. robustus* (espèce qui visite sporadiquement des habitations humaines) montre des différences d'expression des gènes chimiosensoriels. Ces gènes pourraient donc être impliqués dans la sensibilité des *Rhodnius* aux différentes molécules environnantes, ce qui pourrait être lié aux différences d'habitats entre *R. prolixus* et *R. robustus* (Marchant et al. 2021).

## 5. LES GÈNES CHIMIOSENSORIELS ET LEUR LIEN AVEC L'ÉCOLOGIE DES INSECTES

L'étude des systèmes de communication entre les punaises ainsi que des signaux attractifs ou répulsifs peut permettre une meilleure compréhension de leur comportement et donc leur adaptation à un nouvel environnement. Ces signaux peuvent prendre plusieurs formes. Les Triatomés sont capables de produire différents signaux vibratoires en fonction du contexte. Par exemple, lorsque les femelles strident pour refuser l'accouplement, les signaux sont propagés par un contact direct corps à corps avec le mâle (Lazzari et al. 2006). Outre ces signaux mécaniques, des systèmes de communication chimiques ont été étudiés. L'initiation du vol chez le mâle *R. prolixus* est favorisée par les phéromones sexuelles émises par les femelles (Zacharias et al. 2010). Il a été montré chez plusieurs espèces de *Triatoma* (*T. infestans*, *T. sordida* et *T. guasayana*, *T. pseudomaculata*) que les individus s'agrègent sur des papiers imprégnés de leurs propres fèces (Lorenzo et Lazzari 1998; Vitta et al. 2002). Ces fèces, ainsi que des composés cuticulaires, seraient une marque chimique pour guider les triatomés vers leurs abris (Lorenzo et Lazzari 1996). Une fois à l'intérieur du refuge, l'arrêt d'activité des punaises pourrait être provoquée par la stimulation mécanique (thigmotaxie) et les empreintes olfactives (Pires et al. 2002). Les triatomés adultes peuvent aussi produire une phéromone d'alerte répulsive, principalement de l'acide isobutyrique (Manrique et al. 2006). Cependant, à faible dose, ce composé pourrait être attractif (Guerenstein et Guerin 2001).

Certaines molécules exogènes peuvent aussi impacter le comportement des triatomés. Parmi les composés testés l'acide L-lactique, 1-octène-3-ol (Barrozo et Lazzari 2004), l'acide isobutyrique (Guerenstein et Guerin 2001) et l'ammoniac (Taneja et Guerin 1997) se sont avérés des chimiostimulants. De plus, les triatomés sont attirées vers une source de

CO<sub>2</sub> (Barrozo et Lazzari 2006; Milne et al. 2009). Des études multimodales ont montré notamment un effet synergique du CO<sub>2</sub> et du NH<sub>3</sub> (Otálora-Luna et al. 2004), ou encore du CO<sub>2</sub> et l'acide L-lactique (Barrozo et Lazzari 2004). Par ailleurs, il a été démontré que des composés volatiles de culture de levure sont attractifs pour *R. prolixus* (Lorenzo et al. 1999). Des expérimentations de conditionnement olfactif ont été réalisées chez *R. prolixus*. Ainsi, les punaises présentent la capacité à associer des odeurs simples (L-acide lactique) à des renforcements positifs (récompense sanguine) ou négatifs (choc mécanique) (Vinauger 2011a; Vinauger 2011b). De plus, le comportement de recherche d'hôte chez *R. prolixus* peut être manipulé expérimentalement, les punaises biaisant leur choix d'hôte en fonction de leur expérience négative antérieure en associant l'odeur d'un hôte à la perspective d'être blessé (Vinauger et al. 2011b). Des expériences comportementales ont montré l'importance des composés volatiles libérés par le microbiote cutané des visages et des pieds humains dans l'attraction de *R. prolixus* (Ortiz et al. 2011). Cependant, la réponse comportementale est complexe, la plupart des espèces de *Staphylococcus* testées sont des médiateurs de l'attraction, tandis que les espèces non *Staphylococcus* ne sont pas attractives (Tabares et al. 2018).

Concernant les répellents, il a été montré que le DEET très efficace contre les vecteurs hématophages a cependant un moindre effet répulsif sur les triatominae (Ramírez et al. 2020). Des répulsifs alternatifs ont été proposés dont le 4-méthylcyclohexanol (Franco et al. 2018). Récemment, de nouveaux répellents le 2-phényléthanol et le 2-mercaptoéthanol pour *R. prolixus* ainsi que le sulfure de diméthyle et le 2-mercaptoéthanol pour *T. infestans* ont été proposés (Ramírez et al. 2020). Ces composés sont structurellement similaires à des composés identifiés à partir de cultures de la bactérie opportuniste *Citrobacter koseri* de patients malades, suggérant que ces volatiles pourraient signaler un individu malsain aux insectes.

La détection des phéromones et des molécules dégagées par les hôtes impliquent le système chimiosensoriel.

## 5.1. Découverte des gènes chimiosensoriels

Les protéines chimiosensorielles permettent la communication entre les insectes, la perception de leur environnement, la recherche de nourriture, du site de ponte et de leur niche écologique (inféodation à une plante hôte par exemple). La compréhension de l'écologie des insectes fait appel à plusieurs disciplines comme l'éthologie, la neurophysiologie mais aussi la génétique par l'étude du répertoire de gènes chimiosensoriels des espèces. Dans le cas des Triatominae, les protéines chimiosensorielles pourraient jouer un rôle dans la détection des partenaires sexuels, dans la reconnaissance de l'hôte sur lequel la punaise va se nourrir, ou encore dans

la reconnaissance des palmiers et broméliacées et donc dans la spécialisation à ceux-ci.

Dans les années 2000, les gènes chimiosensoriels ne sont annotés que chez quelques espèces de vertébrés, chez le nématode *Caenorhabditis elegans* (Rhabditida ; Rhabditidae) (Troemel et al. 1995), et chez la drosophile, *Drosophila melanogaster* (Diptera ; Drosophilidae) (Clyne et al. 1999; Clyne et al. 2000). Chez cette dernière, les premiers gènes impliqués dans le système chimiosensoriel sont annotés grâce au séquençage de son génome en 1998. Deux familles de gènes sont identifiées les années suivantes : les récepteurs gustatifs (GRs) (Clyne et al. 2000) et les récepteurs olfactifs (ORs) (Clyne et al. 1999; Gao et Chess 1999; Vosshall et al. 1999).

Ces récepteurs ont été d'abord immuno-localisés, ce qui a permis de confirmer que ces gènes étaient bien exprimés dans les dendrites des neurones olfactifs présents dans les palpes maxillaires et dans les antennes, organes olfactifs de la drosophile (Elmore et Smith 2001; Dobritsa et al. 2003). La surexpression des gènes de récepteurs olfactifs dans les organes sensoriels de drosophiles génétiquement modifiées a été mise en lien avec une sensibilité plus forte à certains composants volatiles, ce qui a permis de confirmer leur statut de récepteur et de leur attribuer une fonction putative (Störtkuhl et Kettler 2001).

La fonction de ces récepteurs de drosophile a ensuite été étudiée grâce à des systèmes hétérologues dans des œufs de xénope (Wetzel et al. 2001). C'est à dire que l'on a exprimé le gène à tester (ici le gène du récepteur olfactif) chez un autre animal n'exprimant aucun autre récepteur du même type, ici les œufs de xénopes, une cellule unique assez grosse pour pouvoir la manipuler.

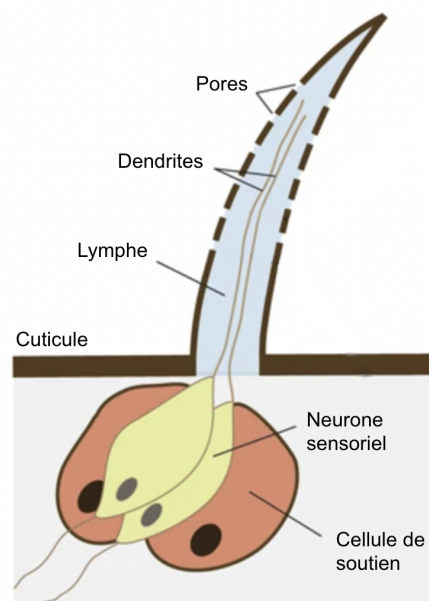
Ces études et le développement des nouvelles technologies de séquençage ainsi que la mise en place de systèmes hétérologues chez des drosophiles mutantes n'exprimant aucun récepteurs (permettant ainsi de "désorphaniser" des récepteurs olfactifs d'autres insectes, c'est à dire trouver le ligand du récepteur) ont permis une grande avancée dans le domaine de la chimio-perception des insectes. En effet, les cinq années suivantes, des récepteurs olfactifs et gustatifs ont également été annotés chez le moustique vecteur du paludisme, *Anopheles gambiae* (Diptera ; Culicidae), chez la noctuelle des bourgeons du tabac, *Heliothis virescens* (Lepidoptera ; Noctuidae) et chez le Bombyx du mûrier, *Bombyx mori* (Lepidoptera ; Bombycidae) (Hallem et al. 2006). Depuis, les gènes chimiosensoriels ont été annotés chez un très grand nombre d'insectes.

## 5.2. Fonction des récepteurs et des transporteurs chimiosensoriels

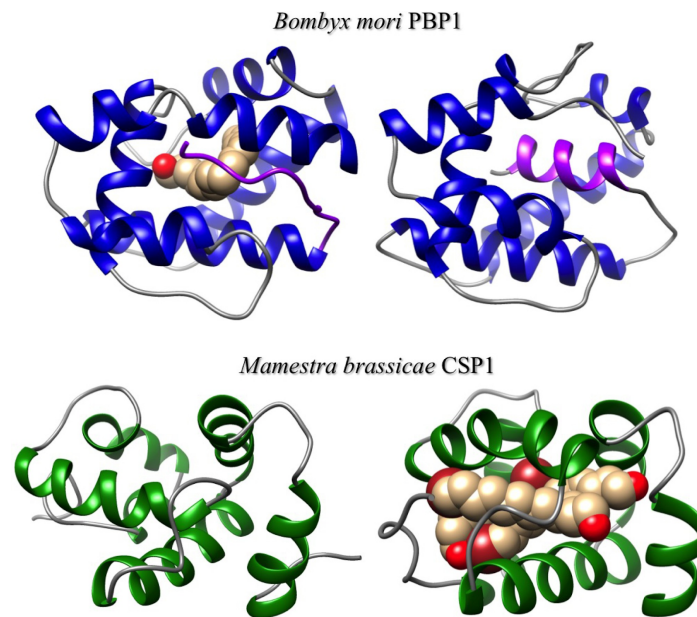
Il existe des protéines associées à ces récepteurs notamment les



protéines liant les molécules olfactives. Les OBPs (*Odorant Binding Proteins*) tout comme les CSPs (*Chemosensory Proteins*) permettent le transport des molécules volatiles souvent hydrophobes dans la lymphe de la sensille et leur transfert jusqu'à la membrane du neurone olfactif où se trouvent les récepteurs olfactifs (Figure 8). Ces protéines présentent une cavité hydrophobe formée par deux hélices  $\alpha$  et une surface hydrophile. La structure des OBPs est stabilisée par trois ponts disulfures entre des cystéines conservées, tandis que pour les CSPs, il y a seulement deux liaisons disulfures entre les cystéines (Figure 9). La fixation du composé volatile entraîne un changement de conformation de la protéine. Les OBPs sont constitués de 6 hélices  $\alpha$ . La partie C-terminale des OBPs est de longueur variable et semble agir comme un couvercle protégeant le cœur fixateur comme cela a été démontré pour une OBP chez l'abeille *Apis mellifera*. Il existe moins de conformations différentes entre les CSPs qui sont toujours composés de 5 hélices  $\alpha$  sans variation de taille rapportée pour la partie C-terminale (Pelosi et al. 2018).



**Figure 8** : Représentation schématique de la structure générale d'une sensille olfactive d'insecte d'après Sánchez-Gracia et al. 2009. Les sensilles gustatives ont une structure similaire, avec un seul pore au sommet du poil sensoriel.



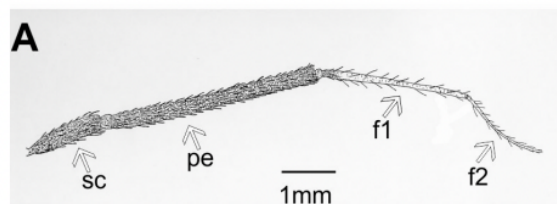
**Figure 9 : Structures tridimensionnelles de la protéine de liaison à la phéromone de *Bombyx mori* (PBP1) et de la protéine chimiosensorielle 1 de *Mamestra brassicae* (CSP1), exemples représentatifs des deux classes de protéines de liaison au ligand impliquées dans la communication chimique des insectes. A pH faible, la partie C-terminale de PBP1 en rose se plie en hélice  $\alpha$  qui remplit la poche interne et éjecte le ligand. CSP1 quant à lui peut aussi subir des changements de conformation et s'étend, ce qui permet la fixation de plusieurs ligands, ici 3 molécules de 12-bromo-dodécanol.**

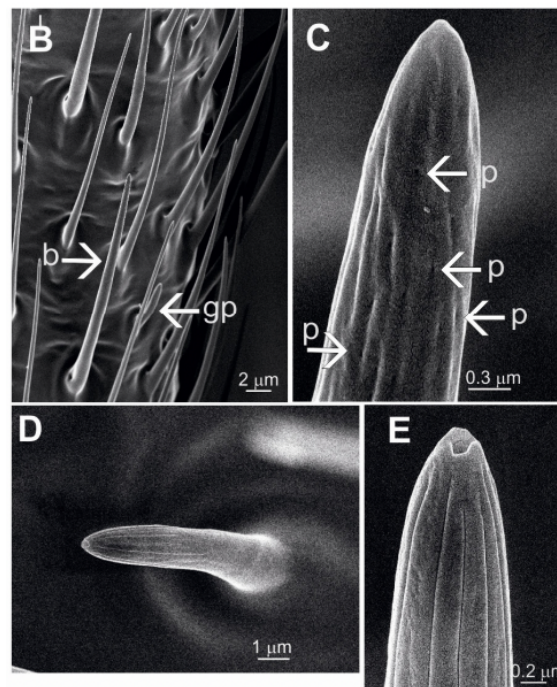
Les OBPs n'ont pas la même expression dans toutes les sensilles et peuvent modifier les interactions entre les ligands et les récepteurs olfactifs et ainsi moduler la sensibilité du neurone olfactif aux différents stimuli. Certaines OBPs peuvent être spécifiques aux phéromones (composés volatiles qui permettent la communication entre individus de la même espèce) et sont appelés plus précisément les PBP (*Pheromone Binding Proteins*). Chez le papillon *Bombyx* du mûrier, *Bombyx mori* (Lepidoptera ; Bombycidae), l'expression de PBPs spécifiques, très conservées, est localisée dans les neurones (ORNs) répondant à différents composants de phéromones, tandis que d'autres OBPs sont associées à la détection de plantes (Bruyne et Warr 2006; Pelosi et al. 2018).

Les CSPs et OBPs ont aussi d'autres fonctions comme le stockage des phéromones dans des glandes pour les relâcher plus tard dans l'environnement chez les Lépidoptères (Zhang et al. 2015), des rôles dans la réponse inflammatoire chez les moustiques hématophages Culicidae (Calvo et al. 2006), dans la formation de l'œuf chez *Apis mellifera* (Maleszka et al. 2007) et chez la fourmi *Solenopsis invicta* (Hymenoptera ; Formicidae) (Cheng et al. 2015) et dans la résistance aux insecticides chez *Bombyx mori* (Xuan et

al. 2015), l'aleurode du tabac *Bemisia tabaci* (Hemiptera ; Aleyrodidae) (Liu et al. 2016) et la chenille de la teigne des choux *Plutella xylostella* (Lepidoptera ; Plutellidae) (Bautista et al. 2015). Enfin, un rôle potentiel serait l'acheminement de pigment dans les yeux de l'insecte. Cette fonction n'a pas été démontrée fonctionnellement, mais a été proposée par homologie avec les rôles identifiés chez les vertébrés (Zhu et al. 2016).

Les récepteurs olfactifs (ORs pour *Olfactory Receptor* aussi appelé *Odorant Receptor* par certains auteurs) permettent la détection de composés volatiles dans l'environnement de l'insecte et sont présents à la surface des neurones olfactifs contenus dans les sensilles des palpes et des antennes. Les ORs sont retrouvés dans la membrane des neurones olfactifs (OSN pour *Olfactory Sensory Neurons*) qui se trouvent dans les sensilles olfactives. Les sensilles olfactives des triatomines sont présentes sur la surface de l'antenne. L'antenne est composée de trois segments, le scape (base de l'antenne), le pédicelle et deux flagellomères (à extrémité de l'antenne). Les sensilles olfactives se trouvent à la fois sur les flagellomères distaux et proximaux (Barrozo et al. 2017) (Figure 10A). Chez les triatomines de la tribu Triatomini il a été montré la présence de sensilles olfactives sur le pédicelle mais pas chez la tribu Rhodniini. Trois types de sensilles olfactives ont été décrites sur le flagellomère de triatomines : basiconique (sensille à paroi mince), le trichoïde (sensille à paroi épaisse), rainurée, en anglais *grooved-peg* (sensille à double paroi) (Barrozo et al. 2017) (Figure 10B-E).





**Figure 10 : Localisation et images en microscope électronique de différentes sensilles olfactives chez les triatomines.** D'après Barrozo et al. 2017

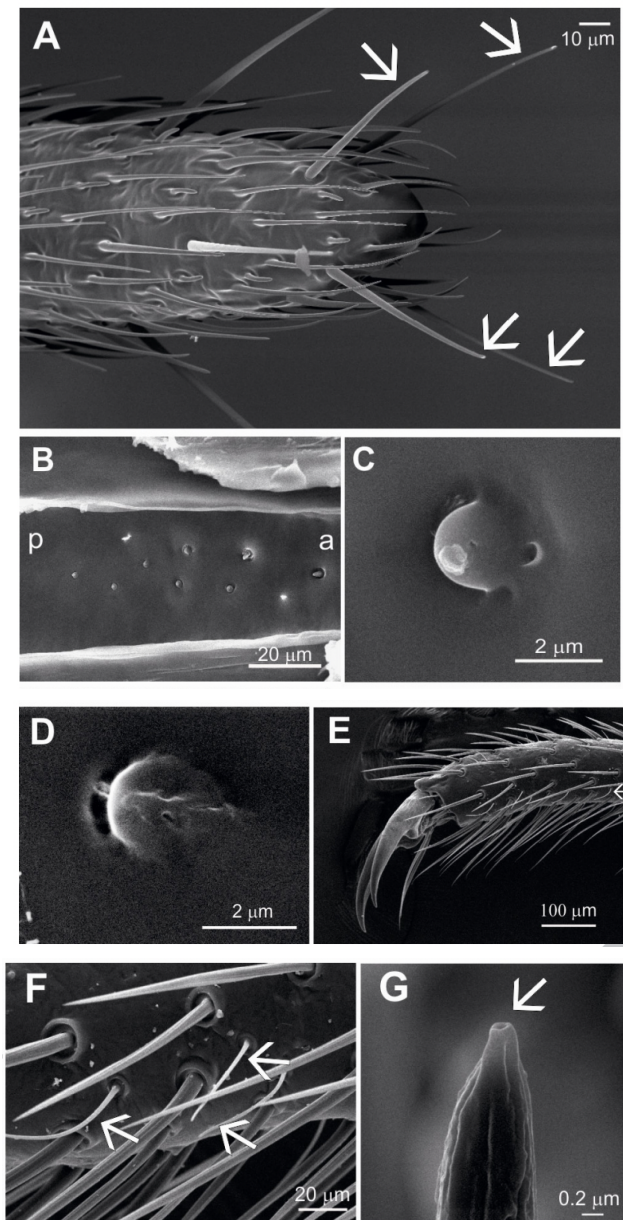
A : Représentation schématique de l'antenne de *T. infestans* adulte montrant ses quatre segments : le scape (sc), le pedicelle (pe) et les flagellomères 1 et 2 (f1, f2).

B : Région médiane du premier flagellomère de *R. prolixus* couverte d'une forte densité de sensilles, où l'on peut reconnaître deux types de sensilles olfactives : les sensilles à paroi minces basiconiques (b) et les sensilles rainurées *grooved-peg* (gp).

C : Détail d'une sensille basiconique montrant la présence de multiples pores (p) dans la cuticule.

D-E : Grossissement d'une sensille rainurée *grooved-peg* (sensille à double paroi)

Les récepteurs gustatifs (GRs pour *Gustatory Receptor*) sont impliqués dans la reconnaissance des molécules sapides. Ils sont retrouvés dans la membrane des neurones des sensilles gustatives (GSN pour *Gustatory Sensory Neurons*) qui, chez la drosophile se situent sur les ailes, les pattes, la bouche et sur l'ovipositeur (Bruyne et Warr 2006; Liman et al. 2014). Les sensilles gustatives n'ont qu'un seul pore par lequel entrent les molécules environnantes contrairement aux sensilles olfactives qui sont multi-pores (Figure 11). Les tarse et les tibias de *R. prolixus* possèdent également des sensilles gustatives (Barrozo et al. 2017). Des sensilles internes ont aussi été décrites dans l'épipharynx de *R. prolixus* (Pontes et al. 2014).



**Figure 11 : Sensilles gustatives des triatomés en microscope électronique.** D'après Barrozo et al. 2017.

A : Extrémité distale du deuxième flagellomère antennaire de *R. prolixus* montrant quatre sensilles gustatives appelées poils chaetiques (flèches blanches).

B : Région antérieure de l'épipharynx de *R. prolixus* montrant huit sensilles épipharyngéales. a : antérieur, p : postérieur. Adapté de Pontes et al. 2014.

C : Détail d'une seule sensille épipharyngéale de *R. prolixus*. Adaptés de Pontes et al. 2014.

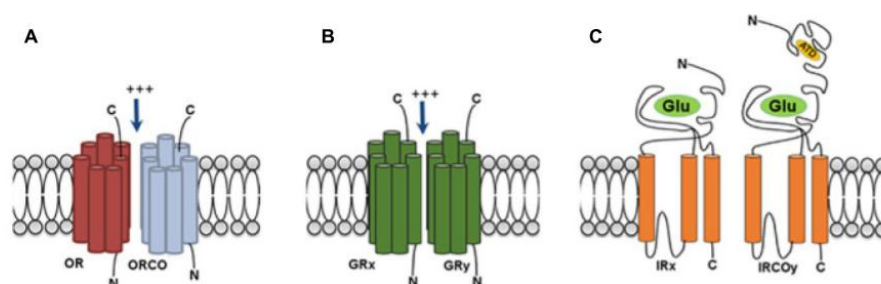
D : Détail d'une seule sensille épipharyngéale de *T. infestans*.

E : Sensilles sensorielles multiples sur un tarse de *R. prolixus*.

F-G : Détail d'une sensille unipore (flèches blanches) sur un tarse de *R. prolixus*.

Les GRs et les ORs sont constitués de 7 domaines transmembranaires

(Figure 12). Ces 7 domaines sont dans la membrane plasmique de la dendrite du neurone chimiosensoriel (OSNs ou GSNs). Lorsque les récepteurs sont activés par les ligands, ils changent de conformation, laissant entrer les cations (ions chargés positivement) ce qui dépolarise le neurone et crée un potentiel d'action (Liu et al. 2021). Sauf exceptions, la partie N-ter de la protéine est intracellulaire à l'intérieur du neurone tandis que la partie C-ter est extracellulaire dans la lymphe. Chez les vertébrés, les GRs et les ORs sont aussi des protéines GPCRs (*G Protein-Coupled Receptors*) car se sont des récepteurs couplés aux protéines G, c'est-à-dire qu'une fois activées par leur ligand, elles ont la capacité de se coupler à des protéines dépendantes du GTP, activant une cascade de réactions à l'intérieur du neurone sensoriel. Ce n'est pas le cas pour les GRs et ORs d'insectes qui ne montrent pas d'homologie évidente avec les GRs et ORs de vertébrés (Ha et Smith 2022). Il y a aussi une autre différence de conformations entre les GRs et ORs d'insectes et ceux des vertébrés : la partie C-terminale est extracellulaire chez les insectes et est intracellulaire (donc dans le neurone) chez les vertébrés (Benton et al. 2006). Chez les insectes, les GRs fonctionnent en combinaison de 2 GRs qui forment un complexe multi-protéique ce qui complexifie la désorption de ceux-ci, tandis que chaque OR est couplé à ORCO, un corécepteur commun à tous les ORs. Les GRs et les ORs transforment une information chimique contenue dans l'environnement en un signal électrique dans le neurone sensoriel, ce qui déclenche une réponse comportementale. Il existe aussi un autre type de récepteur, les récepteurs ionotropiques (IR pour *Ionotropic Receptor*) qui ont 3 domaines transmembranaires et fonctionnent en hétérodimères. Trois co-récepteurs appelés IRCO (Figure 12C) sont décrits chez *R. prolixus* comme étant exprimés à tous les stades de développement, dans les deux sexes et dans différentes conditions physiologiques (Latorre-Estivalis et al. 2016).



**Figure 12 : Schémas des structures de trois récepteurs chimiosensoriels (OR/ORCO, GR, IR/IRCO) exprimés dans les membranes des dendrites des neurones.** A : Quel que soit le récepteur olfactif (ORs), il est toujours couplé au co-récepteur ORCO. B : Les récepteurs gustatifs (GRs) fonctionnent sous forme de complexe multiprotéique. C'est une combinaison de GRs qui permet de détecter une molécule sapide et de générer un signal. C : Les récepteurs ionotropiques fonctionnent également avec des co-récepteurs IRCO (Liu et al. 2021).

Chaque type de protéines est codé par des gènes d'une même famille multigénique : celle des GRs, des ORs, des IRs, des OBPs et des CSPs.

### 5.3. Spécialisation et gènes chimiosensoriels

Plusieurs études chez les insectes suggèrent une évolution rapide des gènes codant les protéines du système chimiosensoriel en lien avec une capacité d'adaptation à un nouveau milieu ou à un nouvel hôte (Smadja et Butlin 2009). Cette spécialisation peut s'accompagner d'une spéciation si les différences du répertoire de gènes jouent sur la reproduction (reconnaissance du partenaire sexuels en modifiant la sensibilité aux phéromones sexuelles par exemple) ou sur la gamme d'hôte (habitat différent qui peut conduire à une spéciation allopatrique ou sympatrique).

Chez plusieurs espèces d'insectes, un lien a pu être établi entre la spécialisation, et donc l'étendue de la gamme d'hôte, et le nombre de gènes du répertoire chimiosensoriel. Il semblerait que les répertoires de gènes chez les espèces sténophages comme le dendroctone du pin *Dendroctonus ponderosae* (Coleoptera ; Curculionidae) et *Agrilus planipennis* (Coleoptera ; Buprestidae) sont réduits par rapport à ceux des coléoptères polyphages (Andersson et al. 2019). Chez le papillon *Vanessa cardui* (Lepidoptera ; Nymphalidae), une duplication de GRs expliquerait l'extension de la gamme d'hôtes de ces papillons (Suzuki et al. 2018). Une perte des gènes codants pour les ORs et les GRs dix fois plus importante chez *Drosophila sechellia* par rapport à *Drosophila simulans* pourrait être mise en relation avec la spécialisation récente de *D. sechellia* sur le fruit toxique *Morinda citrifolia* (McBride 2007).

La modulation du répertoire chimiosensoriel peut aussi s'exercer par la régulation de l'expression de ces gènes. Chez le moustique *Aedes aegypti*, il a été mis en évidence une surexpression d'un OR chez les individus de populations domiciliaires par rapport aux individus de populations sylvatiques, liée à la perception de l'odeur humaine (McBride et al. 2014). De même, chez deux espèces soeurs du complexe d'espèces *Anopheles*, *A. gambiae* montre une préférence de prise de repas de sang sur l'espèce humaine tandis que *A. quadriannulatus* montre une préférence pour les hôtes vertébrés/mammifères non-humains. Cette préférence est à mettre

en relation avec la surexpression de différents ORs dans les antennes de *A. gambiae*.

En ce qui concerne les *Triatoma* : chez *T. brasiliensis*, il a été montré une sous-expression des CSPs et des OBPs chez les individus de populations domiciliaires par rapport aux individus de populations sylvatiques (Marchant et al. 2016). Des expressions différentielles entre *R. prolixus* et *R. robustus* ont été observées pour plusieurs ORs et pourraient être mises en relation avec l'habitat domiciliaire ou sylvatique des espèces (Marchant et al. 2021).

Outre les pertes complètes ou gains de gènes, les gènes chimiosensoriels peuvent aussi accumuler des mutations en cas de relâchement des pressions de sélection. Ce phénomène conduit à la pseudogénéisation du gène et donc à sa perte de fonction. C'est notamment le cas chez *Drosophila sechellia*. Sa spécialisation sur le fruit toxique *M. citrifolia* semble s'accompagner de la pseudogénéisation d'un OBP (Dworkin et Gibson 2006).

#### 5.4. Des traces de sélection observées dans les gènes chimiosensoriels

Les changements de pression de sélection, en lien avec l'adaptation à une nouvelle niche écologique, peuvent aussi être liés à des gains de fonctions en accumulant des mutations non synonymes. Dans les cas des protéines chimiosensorielles, des changements d'acides aminés d'une OBP, d'une CSP ou d'un récepteur pourraient changer leur affinité avec certains ligands et donc permettre de détecter une nouvelle molécule dans l'environnement modifiant ainsi la gamme d'hôtes.

La recherche de gènes codant pour des protéines chimiosensorielles sous sélection positive a été mise en œuvre notamment chez la drosophile (Almeida et al. 2014; Gardiner et al. 2008) où les auteurs montrent qu'après une duplication de gènes chimiosensoriels la pression de sélection purificatrice est moins forte sur le gène dupliqué. Chez le puceron du pois *Acyrtosiphon pisum*, une sélection positive exercée sur les gènes d'ORs et de GRs, suite à une expansion de ces familles multigéniques, serait impliquée dans un processus de spéciation (Duvaux et al. 2015). Certains GRs et ORs de la guêpe *Ceratosolen solmsi* (Hymenoptera ; Agaonidae) dont le cycle de vie dépend de la figue, ont subi une expansion propre à cette espèce et présentent des traces de sélection positive comparativement aux orthologues de ces gènes dans les espèces proches non associées à des figuiers.



## 5.5. Études des gènes chimiosensoriels chez les *Rhodnius*

L'annotation des ORs et des GRs de *R. prolixus* a été réalisée par le consortium à l'origine du séquençage du génome (Mesquita et al. 2015). Les gènes chimiosensoriels ont été trouvés par recherche de modèles de gène et par homologie de séquence avec ceux déjà annotés chez le pou et le puceron. Cette annotation a conduit à l'annotation de 116 gènes (dont 5 pseudogènes) codants pour 111 ORs dont le co-récepteur olfactif ORco (annotation complétée dans Latorre-Estivalis et al. 2016) et 28 gènes codants pour 30 GRs par épissage alternatif. Les OBPs et CSPs ont aussi été annotés et sont respectivement au nombre de 27 et 19 (Mesquita et al. 2015). L'étude comparative des transcriptomes entre *R. prolixus* (domiciliaire) et *R. robustus* (sylvatique qui visite sporadiquement des habitations humaines) 17 transcrits chimiosensoriels (4 CSPs, 1 OBPs, 5 ORs, 1 GR et 4 IRs) sont surexprimés dans le transcriptome de *R. robustus*, mais 4 transcrits (OBP19, OR13, OR40 et OR79) étaient surexprimés dans les transcriptomes de *R. prolixus* (Marchant et al. 2021). On remarque des expansions spécifiques à *Rhodnius* comparé aux autres organismes incluant *Acyrtosiphon pisum* (Hemiptera ; Aphididae), *Pediculus humanus* (Phthiraptera ; Pediculidae), *Anopheles gambiae* (Diptera ; Culicidae), *Apis mellifera* (Hymenoptera ; Apidae), *Bombyx mori* (Lepidoptera ; Bombycidae), *Nasonia vitripennis* (Hymenoptera ; Pteromalidae), *Tribolium castaneum* (Coleoptera ; Tenebrionidae), et *Drosophila melanogaster* (Diptera ; Drosophilidae) (Mesquita et al. 2015). Dans cette étude, les récepteurs ORco de différentes espèces forment un groupe monophylétique dans la phylogénie de gènes ce qui veut dire qu'ils sont bien conservés et qu'ils n'ont pas été dupliqués. D'ailleurs, les séquences des gènes codants pour ORco sont souvent utilisées en groupe externe pour raciner les phylogénies de gènes codants pour les ORs. On dénombre une expansion probablement récente d'après les auteurs, basée sur les longueurs de branches relativement courtes, de gènes codants pour des ORs chez *R. prolixus*, et trois expansions anciennes (longueurs de branches plus importantes).

On remarque que le nombre d'ORs est très variable au sein de l'ordre des Hémiptères (Tableau 3). Il s'agit surtout d'espèces phytophages, ravageurs de cultures, d'où l'intérêt de la communauté scientifique pour ces espèces. L'espèce *Cimex lectularius* (Hemiptera ; Cimicidae), la punaise des lits qui se nourrit de sang possède 50 ORs ce qui est bien plus faible que le nombre d'ORs décrit chez *R. prolixus* et à mettre en relation avec sa spécialisation sur un nombre restreint d'hôtes.

**Tableau 3 : synthèse de nombre d'ORs et de GRs décrits chez d'autres hémiptères**

Espèce	Ordre ; Famille	nb ORs	nb GRs	Tiré de
<i>Aphis gossypii</i>	Hemiptera ; Aphididae	45	-	<a href="#">Cao et al. 2014</a>
<i>Acyrtosiphon pisum</i>	Hemiptera ; Aphididae	79	77	<a href="#">Smadja et al. 2009</a>
<i>Yemma signatus</i>	Hemiptera ; Berytidae	66	14	<a href="#">Song et al. 2020</a>
<i>Subpsaltria yangi</i>	Hemiptera ; Cicadidae	29	7	<a href="#">Qi et al. 2018</a>
<i>Cimex lectularius</i>	Hemiptera ; Cimicidae	50	-	<a href="#">Hansen et al. 2014</a>
<i>Sogatella furcifera</i>	Hemiptera ; Delphacidae	63	-	<a href="#">He et al. 2015</a>
<i>Gerris buenoi</i>	Hemiptera ; Gerridae	155	60	<a href="#">Armisen et al. 2018</a>
<i>Nysius ericae</i>	Hemiptera ; Lygaeidae	83	7	<a href="#">Zhang et al. 2016</a>
<i>Oncopeltus fasciatus</i>	Hemiptera ; Lygaeidae	122	-	<a href="#">Panfilio et al. 2019</a>
<i>Cyrtorhinus lividipennis</i>	Hemiptera ; Miridae	15	3	<a href="#">Wang et al. 2018</a>
<i>Adelphocoris lineolatus</i>	Hemiptera ; Miridae	88	4	<a href="#">Xiao et al. 2017</a>
<i>Apolygus lucorum</i>	Hemiptera ; Miridae	110	-	<a href="#">An et al. 2016</a>
<i>Halyomorpha halys</i>	Hemiptera ; Pentatomidae	138	15	<a href="#">Sun et al. 2020</a>
<i>Phenacoccus solenopsis</i>	Hemiptera ; Pseudococcidae	4	1	<a href="#">Nie et al. 2018</a>
<i>Cacopsylla chinensis</i>	Hemiptera ; Psyllidae	7	2	<a href="#">Xu et al. 2019</a>
<i>Diaphorina citri</i>	Hemiptera ; Psyllidae	46	20	<a href="#">Wu et al. 2016</a>
<i>Rhodnius prolixus</i>	Hemiptera ; Reduviidae	111	30	<a href="#">Mesquita et al. 2015</a>

En ce qui concerne les GRs, on remarque une expansion spécifique à *R. prolixus* ([Mesquita et al. 2015](#)). On remarque aussi que le GR1 forme un groupe monophylétique avec deux récepteurs déjà identifiés chez d'autres espèces comme étant des récepteurs permettant de détecter le fructose, AmGR3 chez *Apis mellifera* ([Robertson et Wanner 2006](#)) et DmGr43a chez *Drosophila melanogaster* ([Sato et al. 2011](#)). Ceci est étonnant si on considère que les triatomines sont strictement hématophages mais vient confirmer les observations de ([Díaz-Albiter et al. 2016](#)) montrant que *Rhodnius prolixus* est capable de s'alimenter sur des tomates en laboratoire. Cependant, la fonction de ce gène n'a pas été confirmée expérimentalement chez *R. prolixus*.

L'annotation de ces récepteurs permet de réaliser plus facilement des systèmes hétérologues pour "désorphaniser" les récepteurs olfactifs. Un système hétérologue dans des ovocytes de xénope a permis de montrer que l'OR80 de *Rhodnius prolixus* est capable de détecter le 2-heptanone, le  $\gamma$ -octalactone, l'acetophenone, et le 4-méthylcyclohexanol ([Franco et al. 2018](#)). De plus, une étude par qPCR, a aussi permis de mettre en évidence 21 récepteurs candidats à la détection de phéromones dont, 4 gènes

surexprimés dans les antennes des individus mâles ([Franco et al. 2018](#)).



## OBJECTIF DE LA THÈSE

---

L'objectif de mes travaux de thèse était d'étudier l'évolution des génomes avec un focus sur les gènes chimiosensoriels. Afin de faire la part signal phylogénétique et adaptatif, cette étude nécessitait au préalable une identification fiable des espèces et du contexte phylogénétique.

Les liens de parenté entre les différentes espèces du genre *Rhodnius* ont été retracés par une approche phylogénomique reposant sur l'étude des mitogénomes et d'un set de 51 gènes nucléaires et incluant 17 espèces sur les 24 espèces décrites. Ces travaux sont présentés dans le **chapitre 1**. Cette approche phylogénomique a permis de statuer sur les affinités évolutives des 3 groupes d'espèces décrits (*pallescens*, *pictipes* et *prolixus*), de proposer d'abroger le genre *Psammolestes*, de confirmer le statut taxonomique de certaines espèces et de mettre en évidence de fréquents événements d'introgession au sein de ce genre. La taille des génomes de différentes espèces a ensuite été estimée par deux méthodes, la cytométrie de flux et l'analyse des données moléculaires à haut débit présentée dans le **chapitre 2**.

Cette étude phylogénétique a été complétée par une approche de génomique comparative entre espèces. Pour étudier la dynamique des génomes, une phylogénie calibrée à l'aide de données fossiles a été réalisée. Onze génomes de *Rhodnius* (dont 4 domiciliaires) ont été assemblés puis annotés. Cette étude a permis de monter un taux différentiel de perte et gain de gènes selon l'habitat et d'identifier parmi les groupes de gènes orthologues ceux ayant le plus d'événements évolutifs ou enrichis selon l'habitat. Cette étude permet de proposer des gènes candidats dont des récepteurs chimiosensoriels impliqués dans l'adaptation à l'habitat des espèces du genre *Rhodnius* (**chapitre 3**).

Les gènes chimiosensoriels appartiennent à des familles multigéniques dont la dynamique évolutive est rapide et à l'origine de changements dans la sensibilité sensorielle. Ils sont donc des objets d'étude de choix dans la recherche de gènes impliqués dans la reconnaissance d'un nouvel environnement. L'annotation spécifique de ces récepteurs chimiosensoriels est présentée dans le **chapitre 4**, leur dynamique en termes de perte et de gain de gènes ainsi que les traces de pression de sélection sur ces gènes sont discutées.



## MATÉRIELS ET MÉTHODES GÉNÉRAL

### 1. ECHANTILLONS

Tout le matériel biologique utilisé dans ces travaux de thèse est présenté ci-dessous (Tableau 4). Les échantillons ont été soit collectés sur le terrain lors de précédentes missions réalisées par Pr. Myriam Harry (échantillons de terrains notés F dans le Tableau 4), soit fournis par des laboratoires avec lesquels nous sommes en collaboration. Le répertoire des espèces de *Rhodnius* que nous avons est assez représentatif du genre. Sur les 24 espèces de *Rhodnius*, deux espèces sont connues uniquement par leur description publiée, il s'agit de *R. dalessandroi* et *R. zeledoni* (Jurberg et al. 2009). Trois espèces sont nouvellement décrites, *R. barretti* (Abad-Franch et al. 2013) et *R. mickii* (Zhao et al. 2021) décrite à partir de 2 individus en Bolivie et *R. marabaensis* (Souza et al. 2016).

Tous les échantillonnages et envois d'échantillons en provenance du Brésil ont été réalisés en respectant les lois sur le partage de données biologiques et génétiques (voir références dans les publications respectives).

**Tableau 4 : Liste des échantillons séquencés et utilisés dans mes travaux de thèse.** En jaune les espèces déterminées moléculairement comme faisant partie du groupe *pallescens*, en vert les espèces du groupe *pictipes* et en violet les espèces du groupe *prolixus*. Les espèces en blanc sont les espèces pour lesquelles nous n'avons pas d'échantillon et qui ne sont pas déterminées moléculairement comme faisant partie de l'un de ces 3 groupes (ces espèces ont été ajoutées au tableau à titre indicatif). La dernière ligne en gris correspond à une espèce de *Triatoma*. Les spécimens sont soit de terrain, F, soit de souche de laboratoire dont le laboratoire d'origine est indiqué. Pour les souches, la référence et l'année de fondation de la souche sont indiquées si possible avec CTA : Colônias de Triatominae, Araraquara, Brésil. Pour les données de séquençage nous avons attribué un identifiant à chaque séquençage. La dernière colonne nommée FC correspond aux nombres de tête que nous avons utilisés dans l'analyse de cytométrie en flux.

Espèces	Répartition géographique	Terrain (F) ou réf. labo	Origine de l'échantillon	Transcriptomique	Illumina	PacBio	GridION	FC
<i>R. amazonicus</i> Almeida et al. 1973	Brazil, French Guiana	F	Belizon, French Guyana		Ama1A			
<i>R. barretti</i> Abad-Franch et al. 2013	Colombia, Ecuador							
<i>R. brethesi</i> Matta, 1919	Brazil, Colombia, Venezuela	CTA 222 (2009)	Igarapé Tucunaré, Brasil	18	R9WX	R9		
		F	Amazonia, Brasil		BRE25WB			
<i>R. colombiensis</i> Mejia et al. 1999	Colombia	CTA 050 (2001)	Tolima, Colombia	19	R10B	R10		
					RCOLM			
<i>R. dalessandroi</i> Carcavallo and								

Barreto, 1976								
<b>R. domesticus</b> Neiva and Pinto, 1923	Brazil	F SC strain	Santa Catalina, Brasil		DomC			
					RDOMFM		RDOMFM	
<b>R. ecuadorensis</b> Lent and Leon, 1958	South Colombia, Ecuador, North Peru	F	PD, Colombia		ECUD			
		F	Peru		ECUE			
					REM			
<b>R. marabensis</b> Souza et al. 2016		F	Benefica, Brasil		ROBB			
<b>R. micki</b> Zhao et al. 2021								
<b>R. milesi</b> Carcavallo et al. 2001	Brazil	F	Brasil		MILE			
		CTA 283	Para, Brazil		RMM			5
<b>R. montenegrensis</b> da Rosa et al. 2012	Brazil	CTA 087 (2003)	Montenegro, RO, Brasil	17	R8E	R8		
<b>R. nasutus</b> Stal, 1959	Brazil	CTA 054 (1999)	Brasil	16	R7WU	R7		8
		F	Piaui, Para, Brasil		INCP			
		F	Piaui, Para, Brasil		NASP			
		F	Piaui, Brasil		NASG			
<b>R. neglectus</b> Lent, 1954	Brazil	CTA 061 (1983)	Frutal, Minas Gerais, Brasil	14	R5WU	R5		
		CTA (2011)	Taquaruçu, Mato-Grosso do Sul, Brasil	12	R3WT			
		F	Pitangueiras, San Paulo, (CTA 58)	13	R4H			
<b>R. neivai</b> Lent, 1953	Colombia, Venezuela	F	Maracay, Venezuela		NEII			
		CTA 278	Carabobo Venezuela		NEIF		NEIF	14
<b>R. pallescens</b> Lent, 1953	Belize, Colombia, Costa Rica, Nicaragua, Panama, Venezuela	CTA s/R	Colombia	11	R1J	R1		
					RPALM			
<b>R. paraensis</b> Sherlock et al. 1977	Brazil, French Guiana							
<b>R. pictipes</b> Stal, 1972	Colombia, Ecuador, French	CTA 072 (1998)	Belem, PA, Brasil	15	R6WV	R6		
		CTA 072 (1998)	Belem, Para, Brasil		R63K			
		F	Belem, Brasil		PIC34WC			
		F	Para, Brasil		PIC3L			
		F	Guyane		PIC2WA			
		CTA 072			RPICM			
<b>R. prolixus</b> Stal, 1859	Central America, Colombia, Ecuador, Suriname, Venezuela	labo IRBI	Honduras		ProH	ProH		
		CTA 077 (1982)	Colombia	PMEF/ PFEF	ProM			10
		CTA 080 (1976)	Colombia		ProN			
		F	Estado Guarico, Venezuela		Pro100			
		F	Cojedes, Venezuela		ProYRP			
		LABO	Univ Neûchatel	PMES PFES				
<b>R. robustus</b> Larousse, 1927	Brazil, Colombia, Ecuador, French Guiana, Peru, Venezuela	F	Piaui, Para, Brasil		NEGP			
		F	Peru		RobR			
		F	Guyane	RMG RFG	Rob5s			
		F	Para, Brasil		MP			
		F	Para, Brasil		RobQ			
		CTA 272 F	Rondônia, Brazil		RRTF			
		CTA 272 M	Rondônia, Brazil		RRTM		RRTM	
<b>R. stali</b> Lent et al. 1993	Bolivia, Brazil	F	Bolivia		V-StaWZ			
		F	Alto Beni, Bolivia		U-StaWY			
<b>R. zeledoni</b>								



Jurberg et al. 2009								
<b><i>R. (Psammolestes) arthuri</i></b> Pinto and Lent, 1935								
<b><i>R. (Psammolestes) coreodes</i></b> Bergroth, 1911		F PC strain	Mato Grosso do Sul, Brazil		PSAMc PC		PSAMc PC	7
<b><i>R. (Psammolestes) tertius</i></b> Lent and Jurberg, 1965		F	Bahia, Curaça, Brazil		PSAM			
<b><i>Triatoma longipennis</i></b> Usinger, 1939		CTA 286	Mexico					10

## 2. EXTRACTION D'ADN ET D'ARN

Les différents protocoles sont donnés dans les publications respectives et je n'ai pas contribué à cette partie.

## 3. PHYLOGÉNIES NUCLÉAIRES ET MITOCHONDRIALES

Nous avons réalisé des phylogénies nucléaires et mitochondriales sur 36 des échantillons Illumina soulignés dans le [Tableau 4](#) ci-dessus, ce qui correspondait à l'intégralité des échantillons séquencés que nous avons à ce moment-là. A noter que nous n'avons pas d'échantillon pour *R. coreodes* lors de cette analyse. Cette analyse nous a permis de déterminer moléculairement l'appartenance des échantillons aux espèces. La correspondance entre les échantillons et les espèces, et l'appartenance aux groupes de *Rhodnius* (couleur jaune pour le groupe *pallescens*, vert pour *pictipes* et violet pour *prolixus*) dans le [Tableau 4](#) ci-dessus a donc été déterminée à *posteriori*.

### 3.1. Echantillons

Les phylogénies ont été réalisées à partir des données Illumina (séquencés en 2016 à la plateforme Imagif à Gif-sur-Yvette) pour avoir un échantillon d'espèces le plus large possible et des données homogènes entre échantillons.

### 3.2. Séquences d'ADN mitochondrial

Afin d'assembler les mitogénomes, les *reads* Illumina ont été sous-échantillonnées (500 Mb de lectures pour chaque échantillon) et assemblées en utilisant le logiciel Trinity avec les paramètres par défaut (Haas et al. 2013). Cette approche permet d'assembler des séquences très abondantes dans le pool de séquences comme les gènes mitochondriaux ou ribosomiaux. Chez les insectes, l'ADNmt représente en moyenne 0,42% de l'ADN total dans les projets de séquençage de génomes (Meng et al. 2019). Comme nous avons utilisé 500 000 kb de données/échantillon pour l'assemblage mitochondrial qui contient en moyenne  $(500\ 000 \times 0,42)/100 = 2\ 100$  kb d'ADNmt et étant donné que le génome mitochondrial fait environ 15 kb, nous obtenons une couverture de 140x pour ce génome. Cette stratégie d'assemblage évite également de récupérer certains NUMTS (*nuclear mitochondrial DNA*) qui sont des pseudogènes mitochondriaux intégrés dans les chromosomes nucléaires. En effet, si ces éléments étaient présents dans les données, comme ils proviennent du génome nucléaire, leur couverture serait très faible (de 5 à 30x selon les échantillons) par rapport à celle du (vrai) ADNmt et avec une faible probabilité d'être récupérés dans les données utilisées pour assembler l'ADNmt. Les contigs correspondant aux mitogénomes et à l'opéron ADNr nucléaire ont été identifiés par BLAST en utilisant comme modèles le mitogénome de *Triatoma infestans* et les séquences du *R. prolixus* de référence (Mesquita et al. 2015). Deux ensembles de données concaténées ont été générés pour les données mitochondriales : le premier comprenait les 15 gènes mitochondriaux (15 mtG), à savoir les 13 gènes codant pour des protéines mtPCG (Nad2, COI, COII, Atp8, Atp6, COIII, Nad3, Nad5, Nad4, Nad4L, Nad6, Cytb et Nad1) et les 2 gènes d'ADNr (16s rDNA et 12s rDNA). Le deuxième ensemble de données, comprenant uniquement les 13 gènes codant pour des protéines mitochondriales (mtPCG), a été utilisé pour des analyses de signature de sélection.

### 3.3 Séquence d'ADN nucléaire

Un deuxième assemblage a été réalisé pour chacun des 36 génomes avec le logiciel SOAPdenovo2 (Luo et al. 2012), la taille de k-mers optimum étant estimée à l'aide du programme Kmergenie (Chikhi et Medvedev 2014). Les gènes du set *Benchmarking Universal Single-Copy Orthologs* (BUSCO) (Waterhouse et al. 2019) ont été recherchés dans ces assemblages en utilisant le jeu de données *insecta\_odb10* comprenant 1 367 gènes communs à tous les insectes et présents en copie unique.

Dans un premier temps, nous avons conservé les gènes codant pour des protéines nucléaires (nu-PCG) pour lesquels une copie entière ou fragmentée a été trouvée dans chacun des 36 assemblages de génomes de

Rhodniini et dans le transcriptome de *T. brasiliensis* (Marchant et al. 2015), car cette espèce a été utilisée comme *outgroup*. Nous retrouvons 88 gènes entiers ou fragmentés dans tous les génomes sur les 1 367 gènes. Pour chaque gène, un alignement a été réalisé avec MAFFT (option d'homologie globale ; Katoh et al. 2019) et les alignements ont ensuite été concaténés. L'alignement concaténé a été visualisé à l'aide de Bioedit (version 7.2.6.1 ; Hall, 1999). Dans un second temps, toutes les régions ne comportant pas au moins 29 génomes alignés ont été retirées manuellement, ainsi que les lacunes ou les séquences mal alignées. Cet élagage final drastique a donné 51 nu-PCGs avec au moins 80% des séquences présentes pour chaque espèce excepté *R. brethesi*.

### 3.4. Phylogénies en maximum de vraisemblance et méthode bayésienne

Les modèles de substitution nucléotidique ont été analysés en utilisant le critère d'information bayésien (BIC) pour tous les ensembles de données. Deux types d'arbres ont été générés. MrBayes (Ronquist et al. 2012) a été utilisé pour générer des topologies d'inférence bayésienne (BI) en utilisant le schéma de partitionnement et les modèles d'évolution suggérés par PartitionFinder (Lanfear et al. 2012) et en effectuant deux exécutions indépendantes de 10 millions de générations. Des phylogénies de maximum de vraisemblance (ML) ont été réalisées en utilisant IQ-TREE (Minh et al. 2020) là aussi avec le schéma de partitionnement et les modèles d'évolution suggérés par PartitionFinder. Le soutien de chaque nœud a été estimé à partir de 1 000 répliqués et du rapport de vraisemblance approximatif avec la correction non paramétrique de Shimodaira-Hasegawa (SH-aLRT) (Guindon et al. 2010). Les arbres phylogénétiques ont été affichés et modifiés à l'aide de iTol (Letunic et Bork 2019).

## 4. MESURE DE LA TAILLE DES GÉNOMES

La mesure de la taille des génomes a été réalisée par deux méthodes indépendantes : cytométrie de flux et estimation basée sur la distribution des k-mers.

### 4.1. Echantillons

Pour réaliser une analyse en cytométrie de flux afin de mesurer les tailles de génomes nous avons besoin de plusieurs individus. Les essais préliminaires ont révélé que les punaises stockées séchées ne donnaient pas de bons résultats. Nous avons donc mesuré les noyaux des punaises

que nous avons en assez grande quantité et qui étaient stockées à -80 degrés. Nous avons donc pu mesurer le génome de 5 espèces de *Rhodnius* (*R. milesi*, *R. nasutus*, *R. neivai*, *R. prolixus*, et *R. robustus*). En ce qui concerne *Psammolestes* nous avons pu faire l'analyse en cytométrie de flux pour *P. coreodes* et l'analyse de la distribution de la taille des k-mers pour *P. tertius* car nous n'avions pas encore de données de séquençage pour *R. coreodes* et pas assez de matériel biologique pour *R. tertius*. La mesure en cytométrie de flux a été réalisée pour des mâles et pour des femelles pour les *Rhodnius* et seulement pour des individus mâles pour *R. coreodes*. En ce qui concerne le séquençage Illumina, nous avons utilisé les mêmes données de séquençage que pour la partie phylogénie ci-dessus, nous n'avions que des données pour les femelles *Rhodnius* et pour un mâle de *P. tertius*.

#### 4.2. Cytométrie de flux et estimation de la taille des génomes basée sur la distribution des k-mers

Pour la mesure en cytométrie en flux, les échantillons ont été préparés à partir de la partie inférieure de la tête d'un seul adulte, mâle ou femelle. La tête de l'insecte est riche en cellules en cours de division ce qui est idéal pour réaliser l'analyse en cytométrie de flux. Prendre la partie inférieure de la tête permet de ne pas prendre les yeux qui sont riches en pigments. Nous aurions pu prendre les gonades, souvent utilisées en cytométrie de flux cependant il était plus pratique de s'affranchir de l'étape de dissection des punaises et d'extraction des gonades et risquer de contaminer le matériel biologique à analyser avec les symbiote et parfois même le repas de sang contenu dans le tube digestif de la punaise. Chaque échantillon a été broyé à froid dans un tampon Galbraith modifié et nous avons utilisé un protocole décrit par [Siljak-Yakovlev et al. 2020](#). Les noyaux de chaque échantillon ont été mesurés en présence du standard (tomate) *Solanum lycopersicum* (Solanales ; Solanaceae) de taille de génome connue. Cela nous permettait d'estimer la taille des génomes de nos échantillons.

Pour les estimations de la taille du génome basées sur les distributions de k-mer, les données génomiques obtenues par séquençage Illumina *paired-end* (15 à 32 Gb de lectures correspondant à une couverture de 30 à 65x) ont été utilisées. Après le comptage des k-mer à l'aide du logiciel Jellyfish avec une taille de k-mer fixée à 21, la taille du génome a été estimée à l'aide du programme GenomeScope ([Vurture et al. 2017](#)).

## 5. ASSEMBLAGE DES GÉNOMES

## 5.1. Données de séquençage

Les différentes données utilisées pour les assemblages sont synthétisées dans le [Tableau 5](#). Du fait de la disponibilité des échantillons, l'acquisition des données s'est échelonnée dans le temps. Par ailleurs, le séquençage Nanopore demandant l'extraction d'ADN à partir d'échantillons frais. Nous avons procédé à ce séquençage en 2021 à partir d'échantillons congelés à -80°C après leur acheminement en carboglace.

Les espèces *R. brethesi*, *R. colombiensis*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. pallescens*, *R. pictipes* et *R. prolixus* ont été séquencées en utilisant la technologie PacBio en 2016 avec une couverture 7x (Plateforme Genotoul). Ces contigs PacBio contiennent des erreurs de séquençage (environ 10%) répartis aléatoirement le long du génome, c'est pour cela qu'une étape d'autocorrection des *reads* a été réalisé à l'aide du mode correct de canu v.1.7 ([Koren et al. 2017](#)). En alignant les *reads* PacBio, le programme corrige les erreurs de séquençage.

En février 2021, nous avons séquencé en long-read *P. coreodes*, *R. domesticus*, *R. milesi*, *R. neivai* et *R. robustus*. Ce séquençage à été réalisé à Genotoul en utilisant l'appareil GridION qui fonctionne avec la technologie nanopore. Cette technique présentant un taux d'erreur plus faible que le PacBio, nous n'avons pas appliqué de correction préalable à l'assemblage. Un séquençage Illumina (40x à 80x) a été réalisé en parallèle sur ces échantillons.

Par ailleurs, un effort de re-séquençage Illumina a été réalisé pour les espèces pour lesquelles nous avons encore du matériel biologique, à savoir *R. brethesis*, *R. colombiensis*, *R. pallescens*, *R. pictipes* et *R. ecuadoriensis* avec une profondeur de séquençage importante allant de 40x à 82x selon la qualité du matériel biologique. Pour les espèces *R. montenegrensis*, *R. nasutus* et *R. neglectus* nous avons utilisé les données de séquençage Illumina de 2016 utilisées pour la phylogénie. Les adaptateurs des *reads* Illumina du séquençage de 2016 ont été retirés à l'aide de cutadapt v.1.18 ([Martin 2011](#)). Pour *R. ecuadoriensis* nous n'avions qu'un séquençage Illumina de bonne couverture (60x), nous avons donc essayé d'assembler le génome en utilisant ces données même si nous n'avions pas de séquençage *long reads* du à la trop petite quantité de matériel frais. Cet assemblage était de trop mauvaise qualité, je n'ai pas pu l'annoter.

**Tableau 5 : Récapitulatif des données utilisées pour assembler les génomes des 13 espèces** (en vert : espèces sylvatiques, en jaune : espèces sporadiques et en rouge : espèces domiciliaires).

Espèce	Sexe	Nom de l'échantillon	long read: technologie et profondeur (x)	short read: Illumina profondeur (x)
<i>P. coreodes</i>	Mâle	PSAMcor	Nanopore (24 x)	74x
<i>R. brethesi</i>	Mâle	BAM	PacBio (7x)	53x
<i>R. colombiensis</i>	Mâle	COLM	PacBio (7x)	49x
<i>R. domesticus</i>	Mâle et femelle	DOMFM	Nanopore (18x)	40x
<i>R. milesi</i>	Mâle	MM	Nanopore (15x)	56x
<i>R. montenegrensis</i>	Femelle	monR8	PacBio (7x)	18x
<i>R. nasutus</i>	Femelle	nasR7	PacBio (7x)	16x
<i>R. neglectus</i>	Femelle	negR5	PacBio (7x)	23x
<i>R. neivai</i>	Femelle	NEIF	Nanopore (19x)	62x
<i>R. pallescens</i>	Mâle	PALM	PacBio (7x)	64x
<i>R. pictipes</i>	Mâle	PICM	PacBio (7x)	79x
<i>R. prolixus</i>	Femelle	ProH2	PacBio (7x)	82x
<i>R. robustus</i>	Mâle	RTM	Nanopore (17x)	46x

## 5.2. Assemblage des génomes

Nous avons réalisé un assemblage *de novo* pour 13 espèces de *Rhodnius* en utilisant le séquençage de données de *short reads* et de *long reads*. Nous avons réalisé l'assemblage avec le programme MaSuRCA v.5 (Zimin et al. 2017) qui combine les bénéfices de deux méthodes d'assemblages : l'approche des graphes de Bruijn (découpage en k-mer de toutes les données de séquençage et détection des chevauchements possibles) et l'approche d'Overlap-Layout-Consensus (assemblage par identification des chevauchements des *reads* sans découpage préalable). En option de MaSuRCA, nous avons utilisé la version modifiée de l'assembleur CABOG (Miller et al. 2008). Le programme MaSuRCA possède une étape finale de *polishing* des données, c'est-à-dire que les *reads* Illumina sont

alignés sur le génome pour corriger les éventuelles erreurs des *long reads*. Si une région n'est couverte que par des *reads* PacBio par exemple, la séquence sera remplacée par des "N" pour noter que l'information est manquante.

Certains assemblages n'étant pas de qualité satisfaisante, nous avons tenté de les améliorer par différents tests. Nous avons testé d'autres logiciels d'assemblages tels que DBG2OLC (Ye et al. 2016) (non présentés ici), Redbean (Ruan et Li 2020) et Flye (Kolmogorov et al. 2020) (Tableau 6). Nous avons gardé la méthode qui nous donnait à la fois de bons N50 et une grande taille d'assemblage.

**Tableau 6** : Essais d'amélioration des assemblages initiaux MaSuRCA 5 (CABOG) en utilisant d'autres assembleurs (Flye et redbean) pour les espèces *R. coreodes* et *R. domesticus*.

	PSAM			RDOMFM		
	initial	readbean	flye	initial	redbean	flye
Number of contigs	5018	<b>3881</b>	5931	<b>3959</b>	6260	8989
Number of contigs in scaffolds	36	0	0	96	0	0
Number of contigs not in scaffolds	4982	3881	5931	3863	6260	8989
Total size of contigs	<b>579395283</b>	534408629	555311360	<b>538748738</b>	488576460	<b>539010022</b>
Longest contig	3647237	2681027	2045246	2735295	1213203	1097976
Shortest contig	1340	1507	500	1644	1983	500
Number of contigs > 1K nt	5018	3881	5388	3959	6260	7876
Number of contigs > 10K nt	4820	2978	3213	3484	5395	4678
Number of contigs > 100K nt	1860	1563	1555	1534	1717	1731
Number of contigs > 1M nt	19	29	36	32	1	5
Number of contigs > 10M nt	0	0	0	0	0	0
Mean contig size	115463	137699	93629	136082	78047	59963
Median contig size	70634	65535	13767	63851	49621	11124
N50 contig length	194317	<b>318079</b>	<b>333222</b>	<b>290478</b>	137764	209826
L50 contig count	812	<b>468</b>	<b>466</b>	<b>510</b>	1066	740
contig %A	33.08	33.19	33.17	32.96	33.10	33.05
contig %C	16.92	16.81	16.83	17.04	16.90	16.99
contig %G	16.93	16.82	16.85	17.06	16.88	16.96
contig %T	33.07	33.18	33.15	32.95	33.12	33.00
contig %N	0.00	0.00	0.00	0.00	0.00	0.00
contig %non-ACGTN	0.00	0.00	0.00	0.00	0.00	0.00
Number of contig non-ACGTN nt	0	0	0	0	0	0

Nous avons aussi essayé de combiner les données Illumina du séquençage de 2016 avec les données récentes afin de tester si l'augmentation de la quantité de données permettait d'améliorer la qualité de l'assemblage, ce qui ne fut pas le cas. Mélanger plusieurs jeux de données de séquençage Illumina crée du bruit et ne permet pas d'améliorer les assemblages. Un exemple de résultat est présenté ci-dessous pour *R. brethesi* est présenté en Tableau 7.

**Tableau 7 : Essai d'amélioration des assemblages MaSuRCA en utilisant les données Illumina de 2016 seules, 2020 seules ou en les combinant.** La complétude a été estimée avec programme BUSCO en utilisant le set de gènes insecta\_odb10 (Simão et al. 2015).

Assemblage <i>R. brethesi</i> en utilisant les Illumina de 2016	illumina 25x (femelle) + pacbio 7x (femelle)	Complétude : 69,8% N50: 38 671 pb taille totale : 342 263 560 pb
Assemblage <i>R. brethesi</i> en utilisant les Illumina de 2020	illumina 53x (male) + pacbio 7x (femelle)	Complétude : 89,4% N50 : 74 676 pb taille totale : 468 068 069 pb
Assemblage <i>R. brethesi</i> en utilisant les Illumina de 2016 en plus de ceux de 2020	illumina 25x (femelle) + illumina 53x (male) + pacbio 7x (femelle)	Complétude : 86,9% N50: 68 172 pb taille totale : 452 895 036 pb

## 6. DATATION

Une estimation des temps de divergence a été réalisée sur des phylogénies mitochondriales et nucléaires en utilisant des points de calibration de données fossiles.

### 6.1. Données fossiles utilisées

La calibration d'une phylogénie à l'aide de données fossiles permet d'estimer une vitesse d'évolution pour les gènes d'intérêt et ainsi d'estimer les temps de divergence entre espèces. Il n'y a pas de fossiles de *Rhodnius* décrits à ce jour, c'est pour cela que nous avons décidé d'inclure d'autres hémiptères (et des Blattodea en groupe externe) afin de pouvoir placer plusieurs points de calibration de données fossiles. Sept points de calibration ont été utilisés, dont 4 points soutenus par 2 fossiles (Tableau 8). Plusieurs bases de données de fossiles existent et ici, une partie de la bibliographie a été réalisée en utilisant <https://fossilinsectdatabase.co.uk/>.



**Tableau 8 : données fossiles utilisées pour calibrer la phylogénie**

Placement dans l'arbre	Espèce des fossiles	Auteur	Âge (millions d'années)	Nom du site
Cimicoidea	<i>Pumilanthocoris gracilis</i>	Hou et al. 2012	166.1 – 163.5	Daohugou village (Chine)
Cimicoidea	<i>Pumilanthocoris obesus</i>	Hou et al. 2012	166.1 – 163.5	Daohugou village (Chine)
Aleyrodoidea+ Aphididae	<i>Lutevanaphis permiana</i>	Szwedo et al. 2014	256 – 256	Lodeve (France)
Miridae	<i>Mirivena robusta</i>	Yao et al. 2007	166.1 – 163.5	Daohugou village (Chine)
Pentatomidae	<i>Asopus puncticollis</i>	Piton, 1940	59.2 – 56	Menat (France)
Pentatomidae	<i>Doryderes laticollis</i>	Piton 1940	59.2 – 56	Menat (France)
Reduviidae	<i>Ceresopsis ornata</i>	Becker-Migdisova 1962	201.3 – 199.3	Sogyuty (Kyrgyzstan)
Reduviidae	<i>Ceresopsis vitrea</i>	Becker-Migdisova 1962	201.3 – 199.3	Ak-Bulak-Say, Sogjuta (Kyrgyzstan)
Termitidae	<i>Nanotermes isaacae</i>	Engel et al. 2011	56 – 47.8	Tadkeshwar lignite mine (Inde)
Triatomini	<i>Panstrongylus hispaniolae</i>	Poinar 2013	20.44 – 15.97	La Toca mine (République Dominicaine)
Triatomini	<i>Triatoma dominicana</i>	Poinar 2005	20.44 – 15.97	République Dominicaine

## 6.2. Données nucléaires et mitochondriales

Deux datations ont été réalisées, une à partir des données nucléaires et l'autre à partir des séquences mitochondriales entières. Pour l'analyse faite à partir des séquences d'ADN nucléaire, nous avons utilisé les mêmes 51 gènes BUSCO que ceux décrits dans la partie 1 du matériel et méthodes général. Afin d'inclure des espèces d'hémiptères et de Blattodea dans l'analyse, les génomes entiers ont été récupérés sur NCBI (Tableau 9) et les gènes extraits avec BUSCO (Waterhouse et al. 2019) puis concaténés. Chaque gène a été aligné avec MAFFT (option d'homologie globale ; Katoh et al. 2019) et les alignements ont été ensuite concaténés. L'alignement concaténé a été visualisé à l'aide de Bioedit (version 7.2.6.1 ; Hall, 1999). Le meilleur modèle d'évolution a été estimé par PartitionFinder (Lanfear et al. 2012). Pour les séquences mitochondriales nous avons récupéré les données pour les mêmes espèces si possible, ou des espèces proches si les

données étaient indisponibles.

**Tableau 9 : Espèces et identifiants des séquences incluses dans la datation.** En raison de l'indisponibilité du génome de *P. lignarius* nous avons utilisé les données transcriptomiques pour cette espèce.

Taxon	Espèce	identifiant ADN mitochondrial	identifiant ADN génomique
Aleyrodoidea	<i>Bemisia tabaci</i>	MH714535.1	GCF_001854935.1
Aleyrodoidea	<i>Trialeurodes vaporariorum</i>	AY521265.2	GCA_011764245.1
Aphididae	<i>Acyrtosiphon pisum</i>	NC_011594.1	GCF_005508785.1
Aphididae	<i>Aphis gossypii</i>	MK994521.1	GCF_004010815.1
Aphididae	<i>Cinara cedri</i>	non disponible	GCA_902439185.1
Aphididae	<i>Therioaphis tenera</i>	MH643885.1	non disponible
Blattodea	<i>Blattella germanica</i>	EU854321.1	GCA_003018175.1
Blattodea	<i>Coptotermes formosanus</i>	NC_015800.1	GCA_013340265.1
Blattodea	<i>Cryptotermes secundus</i>	KP026283.1	GCF_002891405.2
Blattodea	<i>Periplaneta americana</i>	GU947663.1	GCA_002939525.1
Blattodea	<i>Zootermopsis nevadensis</i>	KJ958410.1	GCF_000696155.1
Cimicoidea	<i>Cimex lectularius</i>	JQ739180.1	GCF_000648675.2
Cimicoidea	<i>Orius laevigatus</i>	non disponible	GCA_018703685.1
Cimicoidea	<i>Orius sauteri</i>	KJ671626.1	non disponible
Cimicoidea	<i>Gorpis humeralis</i>	JF927830.1	non disponible
Miridae	<i>Apolygus lucorum</i>	KU234539.1	GCA_009739505.2
Miridae	<i>Cyrtorhinus lividipennis</i>	non disponible	GCA_019603395.1
Miridae	<i>Nesidiocoris tenuis</i>	JQ806057.1	GCA_902806785.1
Miridae	<i>Pachypeltis micranthus</i>	non disponible	GCA_020466155.1
Miridae	<i>Trigonotylus caelestialium</i>	KJ170899.1	non disponible
Miridae	<i>Adelphocoris lineolatus</i>	KJ020286.1	non disponible
Pentatomidae	<i>Aelia acuminata</i>	OU426986.1	GCA_911387785.1
Pentatomidae	<i>Euschistus heros</i>	MG253270.1	GCA_003667255.1
Pentatomidae	<i>Halyomorpha halys</i>	FJ685650.1	GCF_000696795.2
Pentatomidae	<i>Stiretrus anchorago</i>	non disponible	GCA_010014745.1
Pentatomidae	<i>Picromerus lewisi</i>	MW355499.1	non disponible
Triatomini	<i>Panstrongylus lignarius</i>	non disponible	PRJNA395235.fa
Triatomini	<i>Panstrongylus rufotuberculatus</i>	MF614953.1	non disponible

Triatomi	<i>Triatoma brasilliensis</i>	non disponible	non disponible
Triatomi	<i>Triatoma infestans</i>	KY640305.1	GCA_011037195.1
Triatomi	<i>Triatoma rubrofasciata</i>	MH934953.1	Triatoma_chr_assembly

### 6.3. Principe BEAST

La suite de programmes BEAST (v 2.6.3 ; [Bouckaert et al. 2014](#)) permet d'estimer les temps de divergence entre différents taxa en utilisant des points de calibration. Le programme BEAUTi de la suite BEAST a été utilisé pour créer le fichier contenant toutes les données qui permettront d'estimer les temps de divergence. Ce fichier a été ensuite mis en entrée dans BEAST.

Le modèle de Yule calibré a été utilisé. Ce modèle de spéciation est décrit comme généralement plus approprié lorsque l'on considère des séquences provenant de différentes espèces. Le modèle par défaut dans BEAUTi est l'horloge stricte avec un taux de substitution fixe égal à 1 substitution/site/million d'années. Nous avons utilisé un modèle de substitution moins strict. L'option log-normale de l'Horloge Relaxée suppose que les taux de substitution associés à chaque branche sont tirés indépendamment dans une seule distribution log-normale discrétisée ([Drummond et al. 2006](#)).

Cinq simulations ont été réalisées avec le programme BEAST en parallèle pour chaque type de données (mitochondriale et nucléaire) ce qui a permis d'accélérer l'analyse en concaténant les données ensuite. Chaque simulation a été réalisée avec une longueur de chaîne de 100 millions de générations. L'échantillonnage a été réalisé tous les 1 000 arbres. Les 25 000 premiers arbres générés dans chaque simulation n'ont pas été gardés pour créer les 2 arbres consensus avec treeannotator ([Bouckaert et al. 2014](#)). Les fichiers de sortie ont été analysés avec Tracer ([Rambaut et al. 2018](#)) et les arbres ont été édités avec Figtree ([Rambaut 2010](#)).

## 7. ANNOTATION DE GÈNES

### 7.1. Jeu de données

Nous avons annoté les 14 assemblages des espèces de *Rhodnius* (voir 5. 2. du matériel et méthode général) assemblés avec MaSuRCA. Les

transcriptomes des têtes et rostrs de 7 espèces de *Rhodnius* ont été séquencés en 2016 avec la technologie Illumina, *R. brethesi*, *R. colombiensis*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. pallescens* et *R. pictipes*. Les données transcriptomiques de *R. prolixus* et *R. robustus* de [Marchant et al. 2015](#) ont aussi été utilisés dans cette annotation.

## 7.2. Annotation des génomes à l'aide du transcriptome

Pour prendre en compte les séquences répétées, nous avons utilisé RepeatModeler. Ce programme trouve et compte les séquences répétées et les classe par famille. Ensuite, nous avons effectué un masquage de ces séquences avec le programme RepeatMasker ([Smit & Hubley, RepeatModeler Open-1.0. 2008-2015](#) <http://www.repeatmasker.org>). Ce programme masque les séquences répétées dans un génome en remplaçant les répétitions identifiées par des lettres minuscules et facilite ainsi l'annotation.

Le programme BRAKER2 ([Brůna et al. 2021](#)) combinant les méthodologies GeneMark-ES/ET et AUGUSTUS ([Brůna et al. 2021](#)) a été utilisé pour déterminer les séquences codantes. Tout d'abord, nous avons aligné les lectures brutes RNAseq dans le génome avec minimap2 ([Li 2018](#)). Si les données RNAseq n'étaient pas disponibles, nous avons utilisé les données RNAseq d'espèces proches. Pour *R. brethesi*, *R. colombiensis*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. pallescens* et *R. pictipes* nous avons un séquençage de l'ARN pour chaque espèce, pour *R. domesticus* et *P. coreodes* nous avons utilisé le séquençage transcriptomique des espèces les plus proches disponibles qui sont *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. prolixus* (de [Marchant et al. 2015](#)) et *R. robustus* (de [Marchant et al. 2015](#)). Ces mêmes transcriptomes ont été utilisés pour annoter les génomes de *R. neivai* et *R. milesi* mais aucun essai n'a donné de résultats satisfaisants. Pour *R. prolixus* et *R. robustus* nous avons utilisé *R. montenegrensis*, *R. prolixus* (de [Marchant et al. 2015](#)) et *R. robustus* (de [Marchant et al. 2015](#)) car les données de séquençage Illumina RNAseq de l'étude [Marchant et al. 2015](#) étaient trop partielles.

## 7.3. Annotation fine des gènes chimiosensoriels

Nous avons utilisé les ORs et GRs de plusieurs espèces d'hémiptères déjà décrits dans la littérature, y compris les ORs et GRs déjà annotés dans le génome de *R. prolixus* ([Mesquita et al. 2015](#)). Nous avons effectué une recherche avec exonerate dans les génomes assemblés avec MaSuRCA. Pour analyser la sortie d'exonerate et identifier des gènes chimiosensoriels, nous avons utilisé InsectOR de ([Karpe et al. 2021](#) ; [http://caps.ncbs.res.in/gws\\_ors/](http://caps.ncbs.res.in/gws_ors/)) avec les paramètres par défaut (Alignment

Cluster Cutoff 1, Completion Cutoff 300). Ce pipeline permet de prendre en compte la forte dynamique de perte et de gain des gènes codant pour les ORs et tient compte de leur distribution étroite en tandem. Il aboutit à une annotation précise et reproductible des ORs.

Le pipeline InsectOR ayant été développé pour l'extraction des ORs, nous avons voulu vérifier que nous pouvions l'utiliser pour extraire les GRs. Nous avons extrait les GRs manuellement à l'aide des fichiers de sortie *exonerate* pour quelques espèces et nous avons comparé cette annotation avec l'annotation d'InsectOR en réalisant une phylogénie (Figure 13). Cette phylogénie (contenant aussi les séquences du *R. prolixus* de référence à titre indicatif) permet de voir que les deux annotations donnent des résultats semblables. Les séquences retrouvées manuellement qui ne sont pas retrouvées dans l'annotation InsectOR sont après vérification plus fine, des séquences très partielles ou des séquences mal annotées. L'annotation par InsectOR permet même de trouver des séquences que l'annotation manuelle ne permettait pas. Les coordonnées des gènes annotés diffèrent de 3 nucléotides entre l'annotation manuelle et l'annotation d'InsectOR car je n'ai pas annoté le codon stop alors que InsectOR le prend en compte dans ses coordonnées.

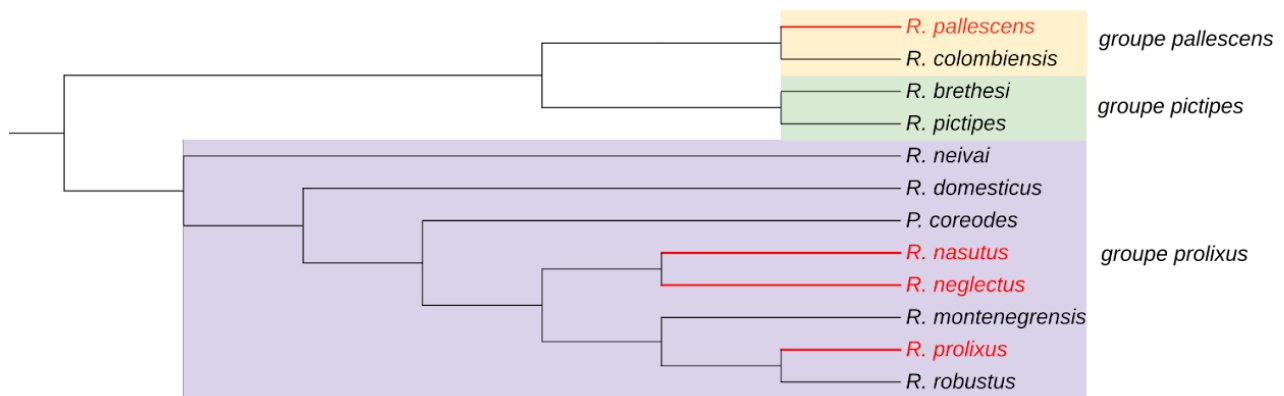


**Figure 13 : Comparaison des séquences trouvées avec InsectOR et des séquences extraites manuellement pour l'espèce *R. colombiensis*.** Les GRs annotés chez *R. prolixus* ont été ajoutés à titre indicatif. L'arbre a été réalisé avec IQTREE (1000 ultrafast bootstraps) à partir des séquences peptidiques alignées par MAFFT (option E-INS-i).

## 8. VITESSE D'ÉVOLUTION DES GÈNES

### 8.1. Jeux de données

Pour cette étude nous avons utilisé les séquences codantes prédites par BRAKER annotées sur nos assemblages MaSuRCA. Nous avons donc 4 assemblages d'espèces considérés comme domiciliaires (comprenant *R. nasutus*, *R. neglectus*, *R. pallescens* et *R. prolixus*). Nous avons 6 espèces qui visitent sporadiquement les domiciles *R. brethesi*, *R. colombiensis*, *R. domesticus*, *R. neivai*, *R. pictipes* et *R. robustus* et 2 espèces sylvatiques qui sont *R. coreodes* et *R. montenegrensis* (Figure 14).



**Figure 14 : cladogramme des espèces incluses dans l'analyse.** Les branches rouges sont les branches terminales des espèces domiciliaires notées elles aussi en rouge. En jaune sont encadrées les espèces du groupe *pallescens*, en vert les espèces du groupe *pictipes* et en violet les espèces du groupe *prolixus*. La topologie du cladogramme et l'appartenance aux groupes a été déterminée grâce à l'analyse phylogénétique décrite en partie 2 du matériel et méthodes général.

### 8.2. Taux de gains et pertes de gènes au sein du genre *Rhodnius*

Nous avons effectué une analyse d'orthologie des séquences protéiques putatives de la littérature et des séquences protéiques putatives annotées dans notre étude en utilisant OrthoMCL (Li et al. 2003) avec des paramètres par défaut (indice d'inflation 1,5). En utilisant le programme CAFE5 (Mendes et al. 2020), nous pouvons calculer le taux de gains et de pertes par gène et par million d'années (lambda). Cette analyse prend en compte les différences d'annotations dues à la complétude des génomes et à la méthode d'annotation. Nous avons analysé si nous avons des lambda



différents entre les espèces domicilières et les espèces des autres branches en calculant soit un lambda global pour tout l'arbre, soit en calculant 2 lambda, un lambda pour les branches terminales des espèces domicilières et un lambda pour les autres branches de l'arbres. En comparant les maximum de vraisemblances de ces analyses, on peut déterminer quel est le modèle le plus vraisemblable.

### 8.3. Traces de sélection dans les gènes chimiosensoriels

Les groupes d'orthologie sont définis par analyse des arbres réalisés à l'aide de IQTREE (1 000 fast-bootstraps). L'arbre des ORs est raciné grâce à ORCO et l'arbre des GRs est raciné avec la méthode du point d'inertie, moins fiable mais nous n'avons pas de groupe externe connu pour les GRs (certains auteurs utilisent les récepteurs au CO<sub>2</sub> mais qui ne sont pas retrouvés chez *Rhodnius*). Un test de recherche de trace de sélection est réalisé sur chaque groupe d'orthologie

Le package CodeML implémenté dans le logiciel PAML (Yang 2007) permet d'étudier le ratio du taux de mutations synonymes sur le taux de mutations non-synonymes (dN/dS) dans un jeu de données. Le logiciel estime les taux normalisés de sites synonymes et de sites non-synonymes et calcule la vitesse d'évolution. Il existe 3 modèles dans PAML : i) le modèle « site » ou le dN/dS peut varier le long des sites pour chaque séquence, ii) le modèle « branche » ou le dN/dS varie selon les différentes branches dans l'arbre, et iii) le modèle « branche-site » ou le ratio varie selon les deux paramètres (non utilisé ici car contesté par certains auteurs e.g. Suzuki 2008). Les deux derniers modèles prennent en compte la phylogénie des espèces et dans notre cas le caractère domicilié ou non des espèces, ce qui permet de calculer un dN/dS global pour les branches terminales des espèces domicilières et de le comparer au dN/dS global du reste de l'arbre pour chaque GR/OR. Cette méthode permet de voir des phénomènes de convergence évolutive, si un gène est soumis à une sélection différentielle selon l'habitat. Le modèle prend en compte que les espèces proches montre de l'homologie qui n'est pas due à de la convergence mais à un ancêtre commun plus proche.

Une interface graphique de ce logiciel, appelée EasyCodeML (Gao et al. 2019) a été utilisée pour réaliser les tests statistiques. EasyCodeML permet de calculer les ratios de maximum de vraisemblance et de déterminer si un modèle est significativement plus probable qu'un autre.

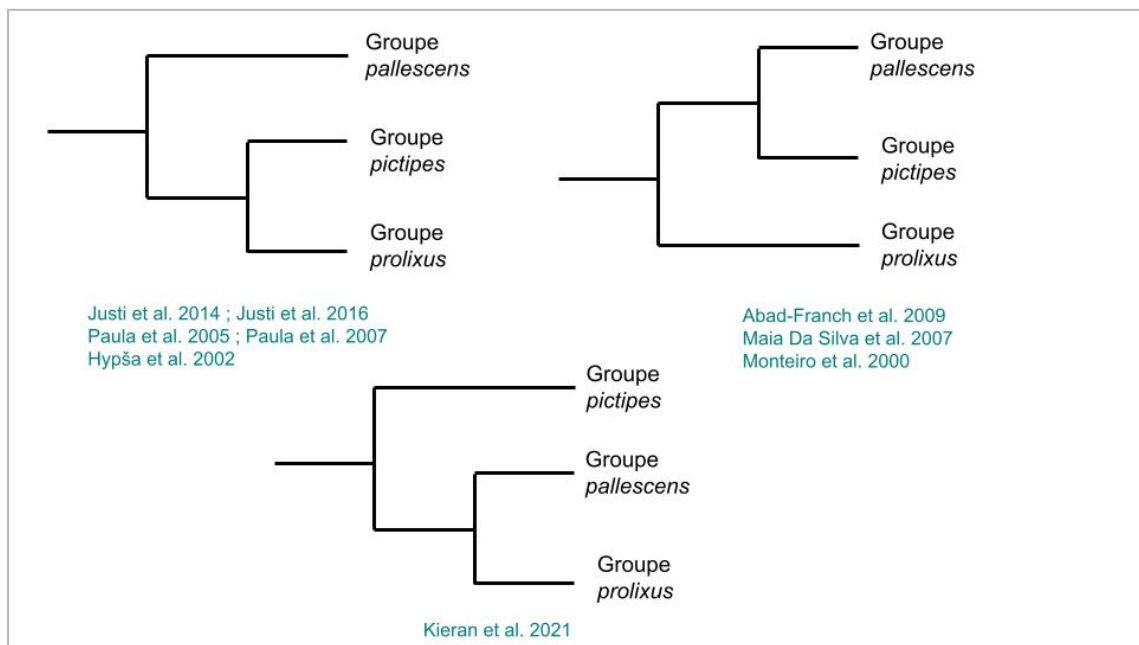


## CHAPITRE I : PHYLOGÉNOMIQUE DES *RHODNIUS*

### CONTEXTE ET PROBLÉMATIQUE DE L'ÉTUDE

#### Contexte d'étude

La phylogénie des espèces du genre *Rhodnius* était mal comprise, plusieurs topologies étaient envisagées et reconstruites par les auteurs (Figure 15) notamment en ce qui concerne les relations entre les différents groupes de *Rhodnius*. Il y a trois groupes au sein des *Rhodnius* : le groupe *pictipes* (comprenant les espèces *R. amazonicus*, *R. brethesi*, *R. paraensis*, *R. pictipes*, *R. stali* et *R. zeldoni*), le groupe *pallescens* (comprenant les espèces *R. colombiensis*, *R. ecuadoriensis* et *R. pallescens*) et le groupe *prolixus* (comprenant les espèces *R. barretti*, *R. dalessandroi*, *R. domesticus*, *R. neivai*, *R. marabaensis*, *R. milesi*, *R. montenegrensis*, *R. neglectus*, *R. nasutus*, *R. robustus* et *R. prolixus*). Selon les auteurs, les relations entre les groupes n'étaient pas congruentes et les branches mal soutenues.



**Figure 15 (Rappel Introduction) : Synthèse de l'état de l'art concernant les différents groupes du genre *Rhodnius*.** Les études portant sur la phylogénie des espèces ne font pas consensus en ce qui concerne les relations entre les 3 groupes : le groupe *pallescens*, le groupe *pictipes* et le groupe *prolixus*.

## Problématique

Afin de statuer sur le statut taxonomique de certaines espèces, de résoudre l'histoire évolutive entre les 3 groupes de *Rhodnius*, et à l'intérieur de ces différents groupes, nous avons réalisé indépendamment des phylogénies mitochondriales et nucléaires.

## Jeu de données

Cette étude repose sur un grand ensemble de données ADN sur 36 échantillons de *Rhodnius*. Les échantillons ont été séquencés à faible profondeur avec un séquençage shotgun du génome entier (technologie Illumina). Les phylogénies ont été réalisées à partir des 15 gènes mitochondriaux (13,3 kb), une partie de l'ADNr nucléaire (5,2 kb) et 51 gènes nucléaires codant pour des protéines (36,3 kb).

## PRINCIPAUX RÉSULTATS

Au niveau des espèces, nous avons confirmé le statut spécifique de *R. montenegrensis* et *R. marabaensis* et nous sommes d'accord avec la synonymie de *R. taquarussuensis* avec *R. neglectus*. Nous invitons également à réexaminer le statut spécifique de *R. milesi* qui est plus probablement incluse dans *R. nasutus* en ce qui concerne l'échantillon que nous avons analysé. Nous avons proposé de définir un complexe d'espèces *robustus* qui comprend les quatre espèces proches : *R. marabaensis*, *R. montenegrensis*, *R. prolixus* et *R. robustus*. Comme *Psammolestes tertius* était inclus dans le clade *Rhodnius*, nous recommandons fortement de reclasser cette espèce au sein du genre *Rhodnius* et la renommer comme *R. tertius*.

Au niveau des différents groupes de *Rhodnius*, les données moléculaires soutiennent systématiquement le clade des groupes *pictipes* et *pallescens*, plus apparentés entre eux qu'au groupe *prolixus*.

De plus, en comparant les topologies des arbres mitochondriaux et nucléaires, nos résultats ont démontré que divers événements d'introggression se sont produits dans les trois groupes de *Rhodnius*, dans les souches de laboratoire mais aussi dans les spécimens sauvages.

Nous avons démontré que des introgressions se sont produites fréquemment dans le groupe *prolixus*, impliquant les espèces apparentées du complexe *robustus* mais aussi les espèces apparentées *R. nasutus* et *R.*

*neglectus*. Une analyse du génome entier a mis en évidence un événement d'introgression dans le groupe pictipes entre *R. stali* et *R. brethesi* et a suggéré un flux génétique complexe entre les trois espèces du groupe *pallescens* : *R. colombiensis*, *R. pallescens* et *R. ecuadoriensis*.

Les données moléculaires soutiennent également une distribution sylvatique de *R. prolixus* au Brésil (état de Pará) et la monophylie de *R. robustus*.

## ARTICLE 1 : PHYLOGENOMICS FOR CHAGAS DISEASE VECTORS OF THE *RHODNIUS* GENUS (HEMIPTERA, TRIATOMINAE): WHAT WE LEARN FROM MITO-NUCLEAR CONFLICTS AND RECOMMENDATIONS

Publié en janvier 2022 dans *Frontiers in Ecology and Evolution* 107(1)

DOI: 10.3389/fevo.2021.750317

Jonathan Filée<sup>1†</sup>, Marie Merle<sup>1†</sup>, Héloïse Bastide<sup>1</sup>, Florence Mougel<sup>1</sup>, Jean-Michel Bérenger<sup>2,3</sup>, Elaine Folly-Ramos<sup>4</sup>, Carlos Eduardo Almeida<sup>5</sup> and Myriam Harry<sup>1\*</sup>

- (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

**Keywords:** phylogenomics, mito-nuclear conflict, introgression, selective pressure, *Rhodnius*, Chagas disease, vectors

### Abstract

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.

## 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 et 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 et Justi 2015). Whereas only three *Psammolestes* species were described, there are twenty-two *Rhodnius* species but one species, *R. taquarussuensis* (Rosa, Justino, et al. 2017), 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).

**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
<i>R. amazonicus</i> (Almeida et al., 1973)	Brazil, French Guiana	Sylvatic	Valid
<i>R. barretti</i> (Abad-Franch et al., 2013)	Colombia, Ecuador	Sylvatic	Not studied
<i>R. brethesi</i> (Matta, 1919)	Brazil (Pará, Amazonas), Colombia, Venezuela	Sylvatic Sporadically in domiciles Amazon basin	Valid
<i>R. colombiensis</i> (Mejia et al., 1999)	Colombia	Sylvatic, Sporadically in domiciles	Valid
<i>R. dalessandroi</i> (Carcavallo and Barreto, 1976)	Colombia	Sylvatic (1 site)	Dubious, no data
<i>R. domesticus</i> (Neiva and Pinto, 1923)	Brazil (atlantic forest)	Sylvatic Sporadically in domicile	Valid
<i>R. ecuadoriensis</i> (Lent and Leon, 1958)	South Colombia, Ecuador, North Peru	Sylvatic, Domiciliar Principal vector in Ecuador	Valid
<i>R. marabaensis</i> (Souza et al., 2016)	Maraba, Para, Brazil Pará	Sylvatic	Valid
<i>R. micki</i> (Zhao et al., 2021)	Bolivia (Santa Cruz, Saavedra) 2 males	No data	Not studied
<i>R. milesi</i> (Carcavallo et al., 2001)	Brazil Pará	Sylvatic	<i>R. nasutus</i> (?)
<i>R. montenegrensis</i> (Da Rosa et al., 2012)	Brazil (Acre, Rondonia)	Sylvatic	Valid
<i>R. nasutus</i> (Stal, 1959)	Brazil (caatinga)	Sylvatic In domiciliation process	Valid
<i>R. neglectus</i> (Lent, 1954)	Brazil (Cerrado São Paulo)	Sylvatic In domiciliation process	Valid
<i>R. neivai</i> (Lent, 1953)	Colombia, Venezuela	Sylvatic (some domiciliary focus)	Valid
<i>R. pallescens</i> (Lent, 1953)	Belize, Colombia, Costa Rica, Nicaragua, Panama, Venezuela	Sylvatic, Domiciliar Principal vector in Panama	Valid
<i>R. paraensis</i> (Sherlock et al., 1977)	Brazil, French Guiana	Sylvatic	Not studied
<i>R. pictipes</i> (Stal, 1972)	Belize, Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Peru, Suriname, Trinidad, Venezuela	Sylvatic Sporadically in domiciles	Valid
<i>R. prolixus</i> (Stal, 1859)	Central America, Colombia, (Ecuador) (Suriname), Venezuela	Domiciliar Principal vector in Central America	Valid
<i>R. robustus</i> (Larousse, 1927)	Brazil, Colombia, Ecuador, French Guiana, Peru, Venezuela	Sylvatic, Sporadically in domiciles (Colombia, Guyana)	Valid
<i>R. stali</i> (Lent et al., 1993)	Bolivia, Brazil (Matto Grosso, Acre)	Sylvatic, In domiciliation process in Bolivia	Valid
<i>R. taquarussuensis</i> (syn <i>R. neglectus</i> ) (Da Rosa et al., 2017)	Mato Grosso do Sul, Brazil	In rural dwelling in the city of Taquarussu	<i>R. neglectus</i>
<i>R. zeledoni</i> (Jurberg et al., 2009)	Brazil	Aracajú-Sergipe, 1 specimen	Dubious, no data
<i>Psammolestes arthuri</i> (Pinto and Lent, 1935)	Colombia, Venezuela	Sylvatic	Not studied
<i>Psammolestes coreodes</i> (Bergroth, 1911)	Argentina, Bolivia, Brazil, Paraguay	Sylvatic	Not studied
<i>Psammolestes tertius</i> (Lent and Jurberg, 1965)	Brazil	Sylvatic	<i>R. tertius</i>

In blue: *prolixus* group species, in pale green: *pictipes* group species, in darker green: *pallescens* group species.

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* (da 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; Fernando 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, Guitton, et Miles 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 et 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 et 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 due to incomplete

lineage sorting, deep coalescence, horizontal gene transfer, introgression, hybridization, hidden paralogy or lack of phylogenetic information (Toews et 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.

**TABLE 2. Specimens used in the study.**

Specimens	Morphological determination	Molecular determination	Origin	Field/Strain	Storage in French Lab
PSAM	<i>P. tertius</i>	<i>R. tertius</i>	Bahia, Curaça, Brazil	F (sylvatic)	EtOH (2003)
Ama1A	<i>R. amazonicus</i>		Belizon, French Guyana	F (sylvatic)	air dried (2014)
BRE25WB	<i>R. brethesi</i>		Amazonia, Brazil	F (sylvatic)	EtOH (2003)
R9WX	<i>R. brethesi</i>		Igarapé Tucunaré, Brazil	S (CTA 222, 2009)	EtOH (2013)
R10B	<i>R. colombiensis</i>		Tolima, Colombia	S (CTA 050, 2001)	EtOH (2013)
DomC	<i>R. domesticus</i>		Santa Catarina, Brazil	F (sylvatic)	EtOH (2006)
ECUD	<i>R. ecuadorensis</i>		PD, Colombia	F (domiciliar)	EtOH (2003)
ECUE	<i>R. ecuadorensis</i>		Peru	S	EtOH (2006)
MILE	<i>R. milesi</i>	<i>R. nasutus</i>	Brazil	S (sylvatic)	EtOH (2006)
MILEP	<i>R. milesi</i>	<i>R. prolixus</i>	Pará, Brazil	F (sylvatic)	EtOH (2003)
R8F	<i>R. montenegrensis</i>		Monte Negro, Rondônia, Brazil	S (CTA 087, 2003)	EtOH (2013)
INCP	<i>R. sp.</i>	<i>R. neglectus</i>	Coronel José Dias, Piauí, Brazil	F (sylvatic)	EtOH (2003)
NASP	<i>R. nasutus</i>		Coronel José Dias, Piauí, Brazil	F (sylvatic)	EtOH (2003)
R7WU	<i>R. nasutus</i>		Brazil	S (CTA 054, 1999)	EtOH (2013)
NasG	<i>R. nasutus</i>		Piauí, Brazil	F (sylvatic)	EtOH (2003)
R5WU	<i>R. neglectus</i>		Frutal, Minas Gerais, Brazil	S (CTA 061, 1983)	EtOH (2013)
NEII	<i>R. neivai</i>		Maracay, Venezuela	S (sylvatic)	EtOH (2006)
R1J	<i>R. pallescens</i>		Colombia	S (CTA s/R)	EtOH (2013)
PIC34WC	<i>R. pictipes</i>		Belém, Brazil	F (sylvatic)	EtOH (2003)
PIC3L	<i>R. pictipes</i>		Pará, Brazil	F (sylvatic)	EtOH (2003)
PIC2WA	<i>R. pictipes</i>		Guyane	F (sylvatic)	air dried (2014)
R6WV	<i>R. pictipes</i>		Belém, PA, Brazil	S (CTA 072, 1998)	EtOH (2003)
R63K	<i>R. pictipes</i>		Belém, Pará, Brazil	S (CTA 072, 1998)	EtOH (2003)
Pro10O	<i>R. prolixus</i>		Guarico, Venezuela	S (domiciliar)	air dried (1990)
ProYRP	<i>R. prolixus</i>		Cojedes, Venezuela	S (domiciliar)	air dried (1990)
ProN	<i>R. prolixus</i>		Colombia	S (CTA 080, 1976)	EtOH (2013)
ProM	<i>R. prolixus</i>		Colombia	S (CTA 077, 1982)	EtOH (2013)
NEGP	<i>R. robustus</i>	<i>R. prolixus</i>	Coronel José Dias, Pará, Brazil	F (sylvatique)	EtOH (2003)
RobR	<i>R. robustus</i>	<i>R. prolixus</i>	Lima, Peru	S (CTA 085, 1982)	EtOH (2006)
Rob5s	<i>R. robustus</i>		Guyane	F (sylvatic)	EtOH (2014)
ROBB	<i>R. robustus</i>	<i>R. marabaensis</i>	Maraba, Brazil	F (sylvatic)	EtOH (2003)
RobQ	<i>R. robustus</i>	<i>R. prolixus</i>	Pará, Brazil	F (sylvatic)	EtOH (2003)
R4H	<i>R. robustus</i>	<i>R. prolixus</i>	Brazil	S (sylvatic)	EtOH (2013)
V-StaWZ	<i>R. stali</i>		Bolivia	S (sylvatic)	EtOH (2006)
U-StaWY	<i>R. stali</i>		Alto Beni, Bolivia	S (sylvatic)	EtOH (2006)
R3WT	<i>R. taquarussuensis</i>	<i>R. neglectus</i>	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.

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).

**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).

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 (<https://mafft.cbrc.jp/alignment/server/>) (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 et 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 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 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 concatenated 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 et 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 et 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 ([www.vectorbase.org](http://www.vectorbase.org)) 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. 1.4.23. 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, 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.

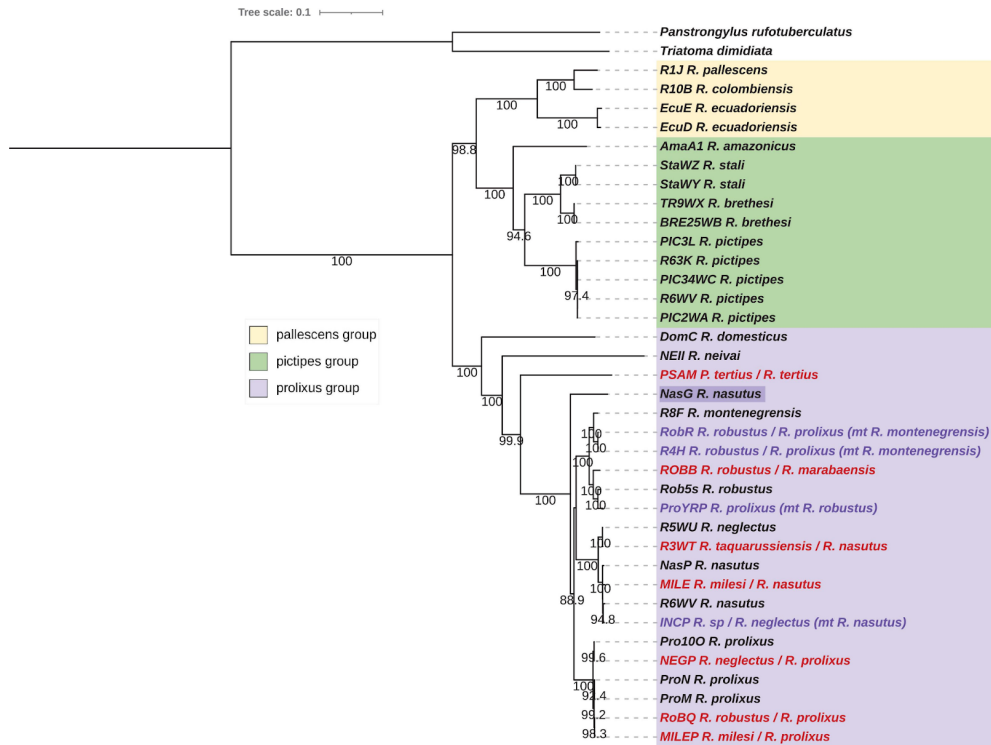
## Results

### Mitochondrial Phylogenetic Inferences

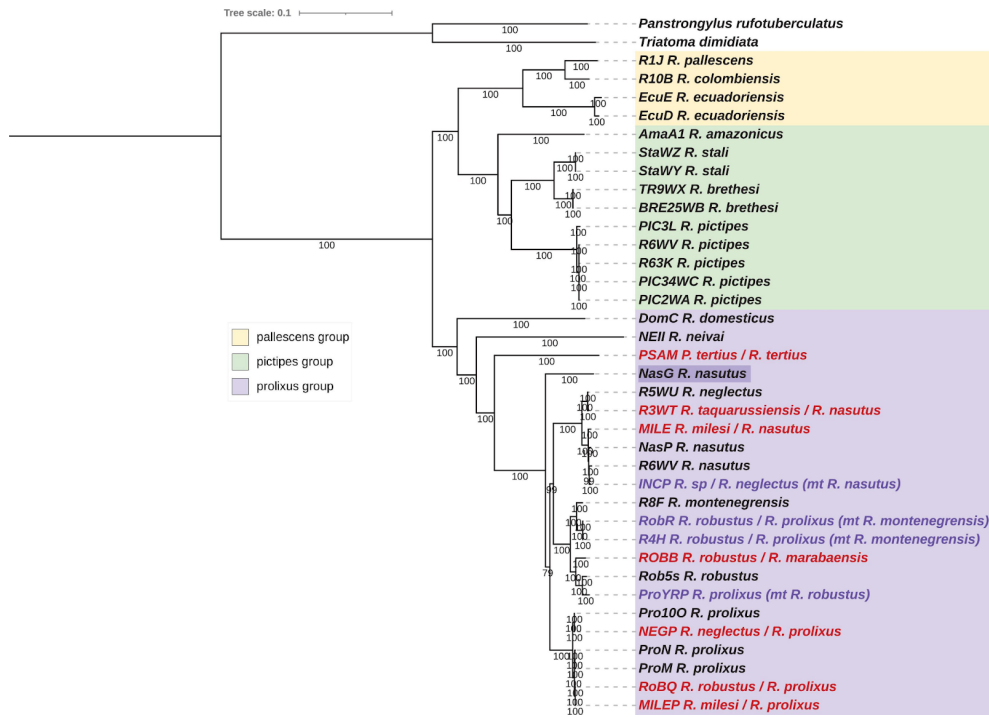
The mitochondrial dataset is composed of 36 samples from the Rhodniini tribe and the 15 mitochondrial genes. The concatenated alignment totalized 13.3 kb with a nucleotide diversity (Pi) of 0.12389. No pseudogene was retrieved from our dataset as attesting by the functional proteins obtained after the 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



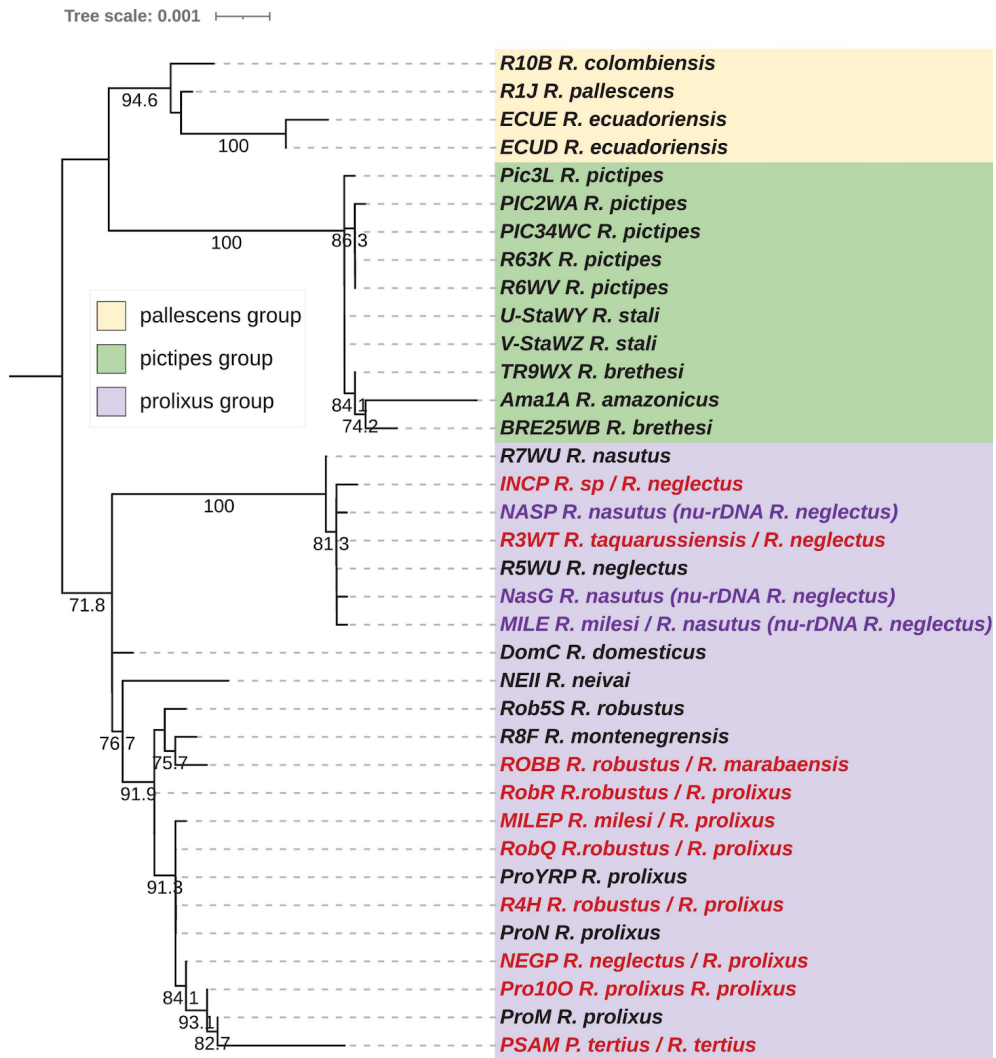
position of *R. neivai* and *R. domesticus* were weak supported.



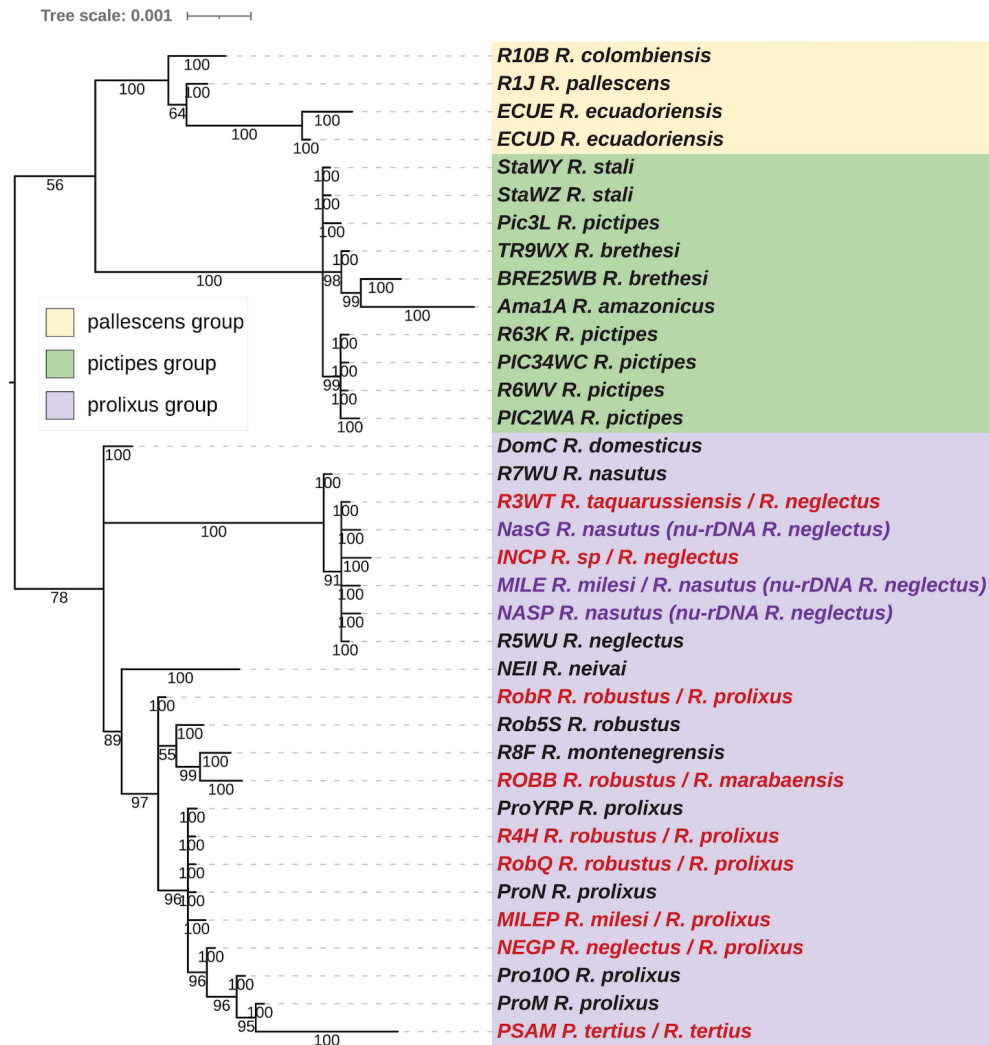
**FIGURE 1. Phylogenetic inferences for the concatenated mitochondrial (mtG) dataset for *Rhodnius* species with the Maximum Likelihood (ML) method.** Statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *T. dimidiata* and *P. rufotuberculatus* are used as outgroups. Introgressed 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.



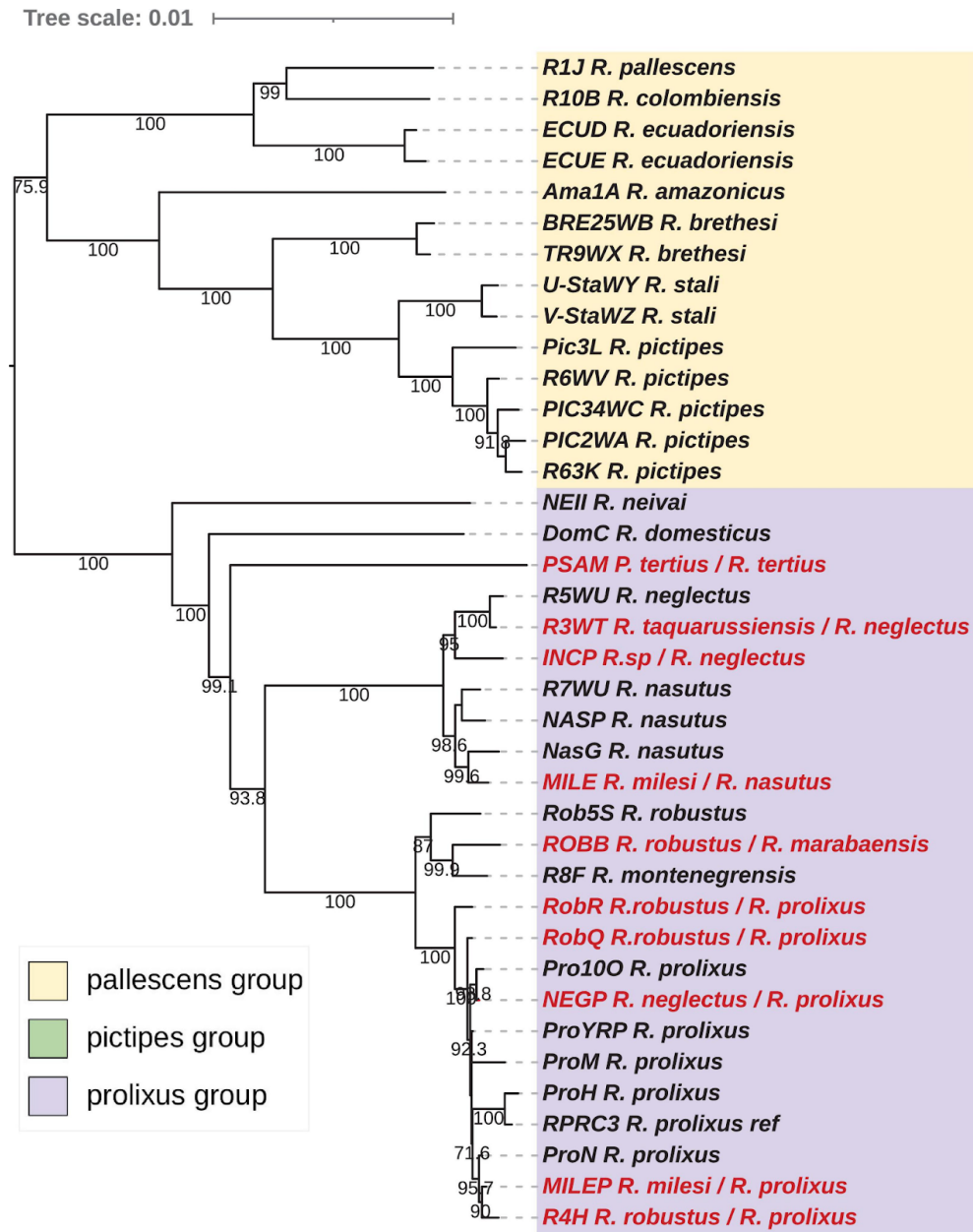
**FIGURE 2. Phylogenetic inferences for the concatenated mitochondrial (mtG) dataset for *Rhodnius* species using Bayesian Inference (BI) method.** Statistical supports (specify Bayesian posterior probabilities in percent) are indicated on the branches. *T. dimidiata* and *P. rufotuberculatus* are used 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.



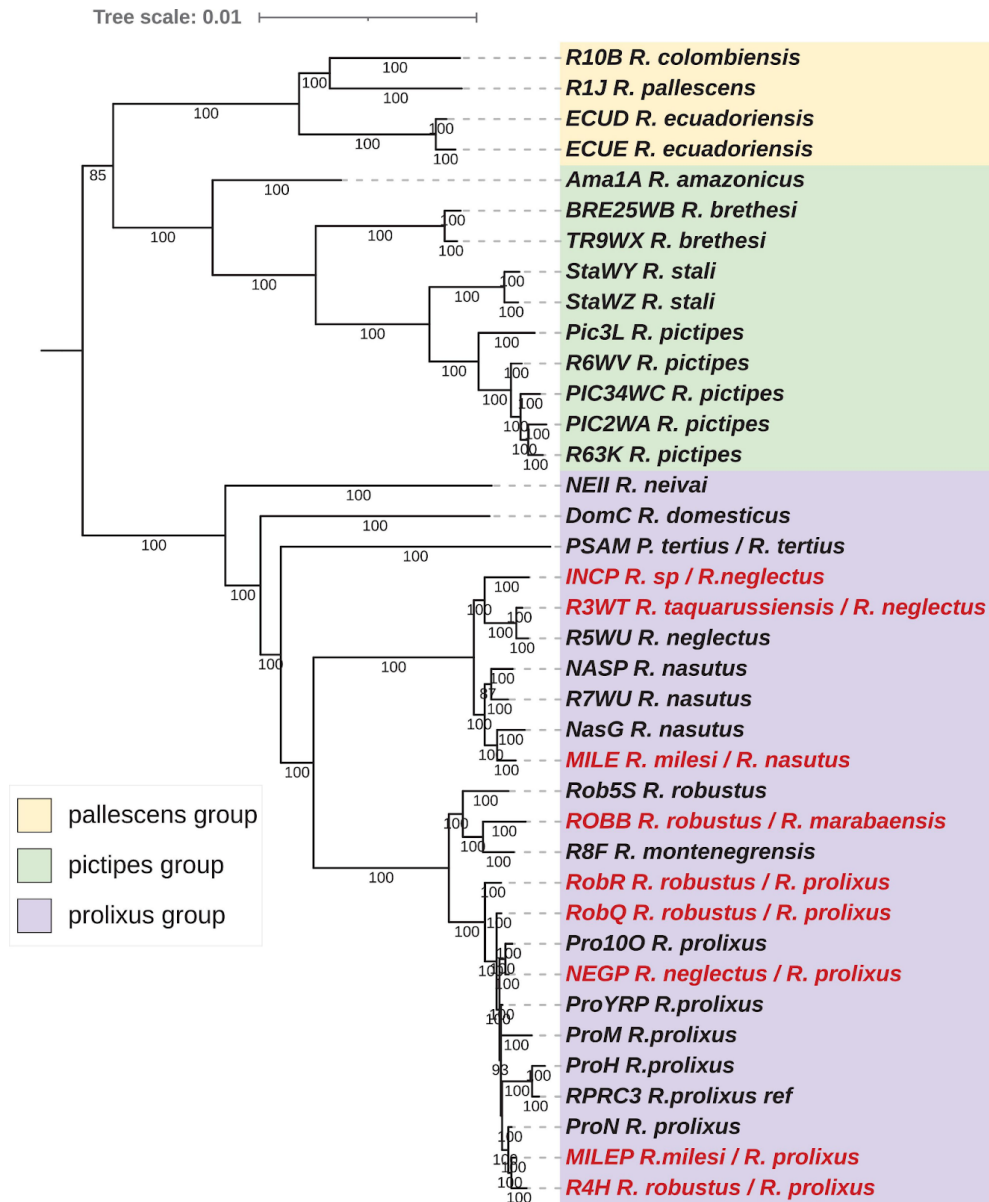
**FIGURE 3. Phylogenetic inferences for the concatenated nuclear rDNA (nu-rDNA) dataset for *Rhodnius* species with the Maximum Likelihood (ML) method.** Statistical supports (1,000 replicates SH-aLRT) are indicated on the nodes. *B. colossus* was used as outgroup. Introgressed species inferred from the nuclear phylogeny (see Figure 5) are indicated in purple, and misidentified species in red.



**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.



**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.



**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.

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* 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. montenegrensis*-*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 clade as the laboratory reference R5WU *R. nasutus* sample (MILE R7WU: 0.0069; 98.29%). More distant pairwise values were obtained between *R. montenegrensis* 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. taquarussensis* (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 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 (ML 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. taquarussensis* (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 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 =  $2 \times 10^{-9}$  for *pallescens* vs. *pictipes* and *prolixus* groups; 0.0044 for *pictipes* vs. *pallescens* and *prolixus* groups;  $9.6 \times 10^{-7}$  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 =  $13 \times 10^{-6}$ ;  $13 \times 10^{-6}$ ; 0.0009; 0.02 respectively) and 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).

**TABLE 4. Test of selection results using the PAML package.**

Datasets	<i>pallescens</i> group	<i>pictipes</i> group	<i>prolixus</i> group
mt-PCGs	<b>0.000000002</b> $\omega_0 = 0.06533$ $\omega_{\text{pall}} = 0.04302$	<b>0.004452367</b> $\omega_0 = 0.06222$ $\omega_{\text{pic}} = 0.04958$	<b>0.000000958</b> $\omega_0 = 0.05644$ $\omega_{\text{pro}} = 0.07712$
<i>Cytb</i>	0.141282571 $\omega = 0.03490$	0.248087413 $\omega = 0.03490$	<b>0.000013280</b> $\omega_0 = 0.000013280$ $\omega_{\text{pro}} = 0.06698$
<i>COI</i>	<b>0.046034763</b> $\omega_0 = 0.01252$ , $\omega_{\text{pall}} = 0.00639$	0.521251673 $\omega = 0.01134$	<b>0.000013280</b> $\omega_0 = 0.02806$ $\omega_{\text{pro}} = 0.06698$
<i>COII</i>	<b>0.040155883</b> $\omega_0 = 0.02126$ , $\omega_{\text{pall}} = 0.01023$	0.139838995 $\omega = 0.01887$	<b>0.000910664</b> $\omega_0 = 0.01453$ $\omega_{\text{pro}} = 0.04180$
<i>Nad1</i>	0.202681777 $\omega = 0.01952$	0.440764212 $\omega = 0.01952$	<b>0.020486354</b> $\omega_0 = 0.01366$ $\omega_{\text{pro}} = 0.03990$
nu-PCGs	0.696800235 $\omega = 0.04575$	0.373073414 $\omega = 0.04575$	<b>0.001175575</b> $\omega_0 = 0.05701$ $\omega_{\text{pro}} = 0.02452$

For each gene or the concatenated datasets, the null model ( $\omega$  for the whole tree) is compared by an LRT test to the model with 2 ratios:  $\omega_0$ , the ratio for branches of the two *Rhodnius* groups not tested and  $\omega_{\text{pall}}$  or  $\omega_{\text{pic}}$  or  $\omega_{\text{pro}}$ , 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.

### 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 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, *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 Triatominae 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* (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 (De Paula et al. 2007; de Paula et al. 2005) 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 + (Justi et al. 2016) 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 et 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 et 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, Havird, et Sharbrough 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 [\(da Rosa et al. 2017\)](#) 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. 2017](#) 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.*

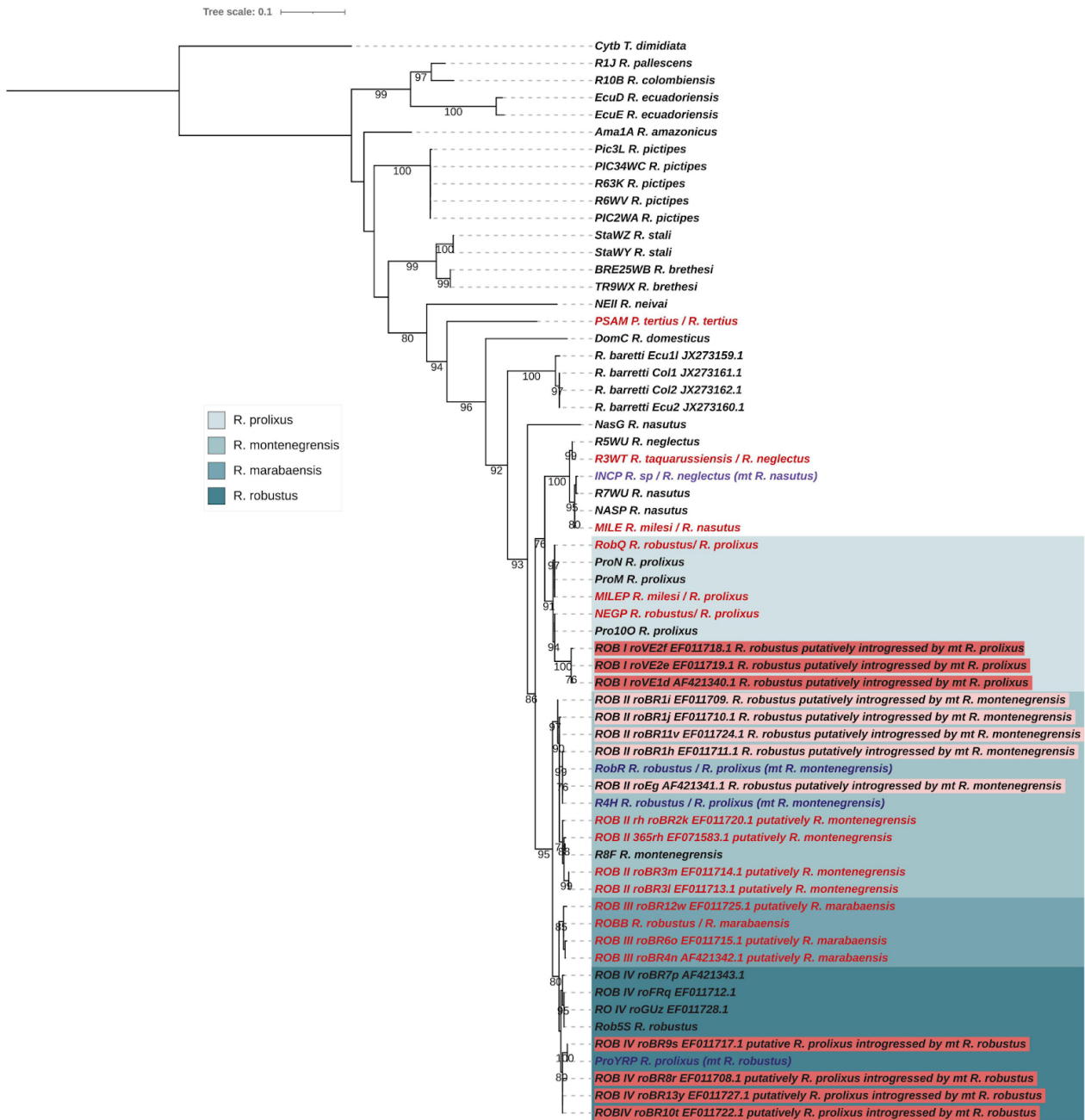
*milesi* should be considered as synonym to *R. nasutus* and we recommend reconsidering the species-specific status of this species with a larger sampling.

A crucial point concerns two species closely related to *R. robustus*, namely the two newly described *R. montenegrensis* (da 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 (da 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. nasutus* 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.





**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 synonymized 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*.

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 et 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 et Holland 2005; Matute et al. 2020). A least plausible explanation is hybrid speciation, though more common in nature than previously thought (Mavárez et Linares 2008; Schumer et al. 2014; Payseur et 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*, 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 expanding 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 *Rhodnius*. 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 et Wygodzinsky 1979](#)) 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 synonymized 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 Oliveira (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>

## POSTAMBULE SUR L'HISTOIRE ÉVOLUTIVE DES *RHODNIUS*

### Utilisation de gènes mitochondriaux et nucléaires pour comprendre l'histoire évolutive des *Rhodnius*

Nous avons démontré que les gènes mitochondriaux alors que largement utilisés pour les phylogénies, à savoir *cytb* et *COI*, sont sous sélection pour les trois groupes, *pallescens*, *pictipes* et *prolixus*. Notre étude indique également que l'utilisation des séquences marqueurs d'ADNr nucléaires (y compris 18S, 5.8S, 28S et ITS) conduit à un arbre mal résolu. Nous avons donc proposé d'utiliser des gènes nucléaires pour établir les phylogénies.

Nous recommandons fortement de réaliser des phylogénies mitochondriales et nucléaires séparées et de tirer parti des conflits mito-nucléaires afin d'avoir une vision évolutive complète de ce genre.

### Remarques sur La tribu Rhodniini : *Psammolestes* et *Rhodnius*,

Après la publication de nos travaux, une étude (Ravazi et al. 2022) va à l'encontre de notre proposition de renommer le genre *Psammolestes* en *Rhodnius* afin que le genre *Rhodnius* soit monophylétique et que la taxonomie reflète l'histoire évolutive des espèces. Cette équipe a réalisé des croisements expérimentaux entre des *Psammolestes* (*P. coreodes* et *P. tertius*) et *R. neglectus* pour évaluer le degré de compatibilité reproductive entre *Psammolestes* et *Rhodnius*. Des hybrides n'ont pas été obtenus pour les croisements entre espèces. L'incompatibilité reproductive entre ces espèces (pourtant éloignées dans l'arbre phylogénétique proposé plus haut et cité dans leur article) ainsi que la biologie différente des *Psammolestes* qui vivent dans les nids d'oiseaux et se nourrissent sur ceux-ci seraient des arguments suffisant pour les auteurs pour maintenir les deux noms de genre même si cette dénomination ne tient pas compte de l'histoire évolutive.

Nous avons démontré qu'au moins une des trois espèces de *Psammolestes* partage une histoire commune avec les *Rhodnius* (il s'agit dans cette étude de *Psammolestes tertius*, dans le chapitre 3 nous verrons que *P. coreodes* aussi se place dans le groupe *prolixus* dans la phylogénie des *Rhodnius*). L'incompatibilité reproductive et une différence d'hôte ne sont pas suffisantes à notre avis pour justifier l'existence de deux genres.





## CHAPITRE 2 : TAILLE DES GÉNOMES

---

### CONTEXTE ET PROBLÉMATIQUE DE L'ÉTUDE

#### Contexte d'étude

Lors de nos premiers essais d'assemblages génomiques nous avons des tailles d'assemblages très variables entre les espèces. La taille de l'assemblage du génome de *R. prolixus* (Mesquita et al. 2015) est de 706,8 Mb. Cependant lorsqu'on regarde le contenu de cet assemblage, 142 Mb soit 20% sont des données manquantes, remplacées par des "NNNN". Des analyses en cytométrie en flux étaient disponibles et montraient des tailles très différentes de génome chez *Rhodnius* (Tableau 15).

**Tableau 15 : Synthèse des données bibliographiques sur les analyses de taille de génome par cytométrie de flux** (Gregory 2022 genomesize.com). Les cellules de testicules ont été utilisées dans tous les cas et l'espèce de référence était *Homo sapiens*. Pour toutes les données, nous avons utilisé la conversion 1 pg = 978 Mbp.

Espèces	C-value	Conversion en Mb	Références
<i>Rhodnius colombiensis</i>	0,58	577	Diaz et al. (2014)
<i>Rhodnius ecuadoriensis</i>	0,72	704	Panzera et al. (2007)
<i>Rhodnius pallescens</i>	0,73	714	Gomez-Palacio et al. (2012)
<i>Rhodnius prolixus</i>	0,75	733	Panzera et al. (2007)

#### Objectif

Les génomes de *Rhodnius* semblent avoir des tailles très variables ce qui pourrait refléter les contractions importantes du contenu génomique qui pourrait être en lien avec des mécanismes d'adaptation mais cette variation de taille n'a pas été testée. Une estimation de la taille des génomes est également souvent requise pour le séquençage à haut débit afin de quantifier la couverture du séquençage. Avant de réaliser notre étude sur la dynamique des génomes chez *Rhodnius* nous avons réalisé des estimations de la taille des génomes par deux méthodes : la cytométrie de flux et l'analyse de la distribution des k-mers.

## PRINCIPAUX RÉSULTATS DE L'ÉTUDE

La cytométrie de flux et les distributions de k-mer dans les séquences génomiques ont été utilisées pour estimer la taille du génome de cinq espèces de *Rhodnius* (*R. milesi*, *R. nasutus*, *R. neivai*, *R. prolixus*, *R. robustus*) et d'une espèce de *Psammolestes* (*P. coroedes* pour FC / *P. tertius* pour les k-mers). Une autre espèce appartenant au genre *Triatoma* a été incluse dans cette étude pour comparaison.

En utilisant des modèles phylogénétiques linéaires généralisés mixtes (PGLMMs), des variations significatives de la taille du génome ont été trouvées entre les espèces et entre les sexes.

*R. prolixus* a montré la plus grande taille de génome. Dans cette étude, nous avons fourni les premières données sur la taille des génomes des femelles chez les Triatominae et avons mis en évidence que chez *R. robustus*, la taille du génome était significativement plus petite chez les femelles que chez les mâles. Bien que plus petites, les estimations de la taille des génomes basées sur la méthode de distribution k-mer présentaient le même schéma que celles de la cytométrie en flux.

## ARTICLE 2 : GENOME SIZE VARIATION OF CHAGAS DISEASE VECTORS OF THE RHODNIINI TRIBE

Publié en mai 2022 dans *The American journal of tropical medicine and hygiene* 9

DOI: 10.4269/ajtmh.21-0394

Marie Merle<sup>1†</sup>, Jonathan Filée<sup>1†</sup>, Jader de Oliveira<sup>2,3</sup>, Carlos Eduardo Almeida<sup>4</sup>, Florence Mougel<sup>1</sup>, Heloïse Bastide<sup>1</sup>, Marc Girondot<sup>5</sup>, João Aristeu da Rosa<sup>3</sup>, Myriam Harry<sup>1\*</sup>

- (1) Université Paris-Saclay, CNRS, IRD, UMR EGCE, Evolution, Génomes, Comportement et Ecologie, Gif-sur-Yvette, France
- (2) Universidade de São Paulo, Faculdade de Saúde Pública, São Paulo, Brazil
- (3) São Paulo State University, School of Pharmaceutical Sciences, Araraquara, São Paulo, Brazil;
- (4) University of Campinas–UNICAMP, São Paulo, Brazil
- (5) Université Paris-Saclay, CNRS, AgroParisTech, UMR ESE, Ecologie Systématique et Evolution, Orsay, France

### Abstract

The genome size of five *Rhodnius* species (*R. milesi*, *R. nasutus*, *R. neivai*, *R. prolixus*, and *R. robustus*) and two *Psammolestes* species (*P. coroedes* and *P. tertius*) were estimated using flow cytometry and/or k-mer distributions in genome sequences. Phylogenetic generalized linear mixed models highlighted significant genome size variations among species and between sexes, with *R. prolixus* showing the largest genome. In this study we provide the first data on female genome size in Triatominae. For five species, female genome size did not differ from males, except for *R. robustus*, where females had smaller genomes. Genome size estimations based on the k-mer distribution method were less than those estimated from flow cytometry, but both methods exhibited the same pattern of sexual differences. Further genomic studies are needed to infer whether genome size variation could be an adaptive trait in *Rhodnius*.

## Introduction

The blood-sucking bugs (Hemiptera, Reduviidae, Triatominae) are vectors of the parasite *Trypanosoma cruzi* (Kinetoplastea, Trypanosomatida), the causal agent of Chagas disease (PAHO, 2022). The Triatominae are divided into 18 genera distributed in five tribes (Hernández et al. 2020). The Rhodniini tribe is composed of two genera, *Rhodnius* and *Psammolestes*, that differ in their ecology, with *Rhodnius* species being associated with palm trees and *Psammolestes* species with bird nests (Galvão et Justí 2015). There are currently 21 *Rhodnius* species divided in three major *Rhodnius* groups-namely, *Rhodnius pictipes*, *Rhodnius prolixus*, and *Rhodnius pallescens* (Filée et al. 2022). Molecular studies showed that *Psammolestes*, which contains three described species, was nested within the *Rhodnius* clade (see references in Oliveira et al. 2018). This was confirmed using whole-genome shotgun sequencing with a very large DNA data set, including 15 mitochondrial genes, partial nuclear recombinant DNA, and 51 nuclear protein-coding genes that support the synonymy of the two genera (Filée et al. 2022). The Rhodniini species are distributed from Central America to northern Argentina, with the greatest species richness in Amazonia. The tribe encompasses a variety of species, including domiciliary vectors (e.g., *R. prolixus*), species sporadically invading anthropogenic systems (e.g., *Rhodnius nasutus*, *Rhodnius neivai*, and *Rhodnius robustus*), and strictly sylvatic species (e.g., *Rhodnius milesi*, *Psammolestes coroedes*, and *Psammolestes tertius*) (Galvão et Justí 2015).

For genome sequencing projects, genome size data are required. Performed on disease vectors such as *R. prolixus* (Mesquita et al. 2015), the outcomes on the genome structure and evolution can boost our understanding of parasite transmission and may lead to the development of innovative control methods. Indeed, the completeness of genome assemblies is assessed from genome size estimates, and independent measures based, for example, on flow cytometry (FC) help to quantify genomic library coverage. For the Rhodniini tribe, few FC data are available, except for males of the reference species *R. prolixus* (Panzera et al. 2007; Panzera et al. 2010) and for three species of the trans-Andean *pallescens* lineage namely, *Rhodnius colombiensis* (Díaz et al. 2014), *Rhodnius ecuadoriensis*, (Panzera et al. 2010) and *Rhodnius pallescens* (Gómez-Palacio et al. 2012).

## Materials and Methods

In this work, we used both FC and k-mer distributions in genome sequences to estimate the genome size of five *Rhodnius* (*R. milesi*, *R. nasutus*, *R. neivai*, *R. prolixus*, and *R. robustus*), and two *Psammolestes* (*P. coroedes* with



FC and *P. tertius* with k-mer distributions). The species *Triatoma longipennis*, one of the most important vectors of Chagas disease in Mexico, was also studied in FC to enlarge data for the Triatomini tribe. We analyzed 65 samples from both sexes that were bred under the same conditions at the Parasitology Laboratory of the FCFAR/São Paulo State University, Araraquara, Brazil (Table 1). The samples were stored at  $-80^{\circ}\text{C}$  before the experiment. These samples were prepared from the inferior part of the head from a single adult female or male. Each sample was ground in ice-cold modified Galbraith buffer, and we used a previously described protocol (Siljak-Yakovlev et al. 2020). Both nuclei from the sample and the standard (*Solanum lycopersicum* L. cv Montfavet 63-5,  $2C = 1.99$  pg) were extracted simultaneously and stained with propidium iodide for 10 min on ice in the dark. Genome sizes were measured by FC in picograms for at least 2,000 scanned nuclei per sample (Cytoflex, Beckman-Coulter) and converted from picograms to mega-base pairs ( $1$  pg =  $\sim 978$  Mbp).

**Table 1 : Genome size in Triatominae**

Species, sex	Samples	2C pg		1C pg		Mb		n	Ref.
		Mean	SE	Mean	SE	Mean	SE		
<b>Rhodniini tribe</b>									
<i>Rhodnius colombiensis</i> , M	Colombia	–	–	0.59	0.01	577	10	13	Díaz et al. <sup>9</sup>
<i>Psammolestes coroades</i> , M	Mato Grosso do Sul, Brazil	1.51	0.07	0.755	0.035	738	34	7	Our study
<i>R. milesi</i> , M	CTA 283 Pará, Brazil	1.46	0.04	0.73	0.02	714	20	5	Our study
<i>R. nasustus</i> , F	CTA 054 Brazil	1.46	0.06	0.73	0.03	714	29	4	Our study
<i>R. nasustus</i> , M	CTA 054 Brazil	1.43	0.03	0.715	0.015	700	15	4	Our study
<i>R. neivai</i> , F	CTA 278, Carabobo Venezuela	1.43	0.07	0.715	0.035	700	34	8	Our study
<i>R. neivai</i> , M	CTA 278, Carabobo Venezuela	1.44	0.01	0.72	0.005	704	5	6	Our study
<i>R. prolixus</i> , F	CTA 077, Colombia	1.54	0.04	0.77	0.02	753	20	6	Our study
<i>R. prolixus</i> , M	CTA 077, Colombia	1.54	0.02	0.77	0.01	753	10	4	Our study
<i>R. prolixus</i> , M	Santander, Colombia	–	–	0.75	0.08*	733	78	6	Panzer et al. <sup>8</sup>
<i>R. robustus</i> , F	CTA 272 Rondônia, Brazil	1.41	0.01	0.705	0.005	690	5	4	Our study
<i>R. robustus</i> , M	CTA 272 Rondônia, Brazil	1.48	0.03	0.74	0.015	724	15	7	Our study
<i>R. ecuadoriensis</i> , M	Peru	–	–	0.72	0.09*	704	88	3	Panzer et al. <sup>8</sup>
<i>R. pallescens</i> , M	Colombia, Panama	–	–	0.73	0.05	714	49	51	Gómez-Palacio et al. <sup>10</sup>
Four spp., F	–	–	–	0.73†	0.02†	714†	22†	22	–
Ten spp., M	–	–	–	0.72†	0.03†	706†	32†	106	–
<b>Triatomini tribe</b>									
<i>Dipetalogaster maxima</i>	Mexico	–	–	1.29	0.10*	–	–	–	Panzer et al. <sup>8</sup>
<i>Eratyrus cuspidatus</i> , M	Colombia	–	–	1.26	0.08	–	–	–	Panzer et al. <sup>8</sup>
<b><i>Triatoma longipennis</i>, F</b>	CTA 286 Mexico	2.49	0.09	1.24	0.04	1,218	43	5	Our study
<b><i>T. longipennis</i>, M</b>	CTA 286 Mexico	2.41	0.04	1.21	0.02	1,180	21	5	Our study
Two <i>Panstrongylus</i> , M‡	–	–	–	1.43	0.01	–	–	–	Panzer et al. <sup>8</sup>
Thirteen <i>Triatoma</i> ssp., M§	–	–	–	1.36	0.48	1,330	469	–	Panzer et al. <sup>8</sup>

CTA = Colônias de Triatominae, Araraquara; F = females; M = males. Genome size estimates in picograms (2C) are reported for Rhodniini and Triatomini tribes using flow cytometry method.

1C was estimated from 2C values. For all data, we used the conversion 1 pg = 978 Mbp. The CTA number refers to the strain number. All the species have 2n = 22 chromosomes (20A + X<sub>1</sub>X<sub>2</sub>Y), except species in bold type, with 22 of 23 chromosomes (20A + X<sub>1</sub>X<sub>2</sub>Y).

\* Estimated SE from [Neville et al. 2019](#) with a corrected coefficient of 0.09 to take into account the correction for the 1C values reported in [Oliveira et al. 2018](#).

† Mean values.

‡ *Panstrongylus geniculatus* (1.31 ± 0.06 pg) and *P. rufrotuberculatus* (1.32 ± 0.07).

§ *Triatoma barberi* (1.13 ± 0.03 pg), *T. brasiliensis* (1 ± 0.07 pg), *T. delponteii* (2.67 ± 0.17 pg), *T. dimidiata* (0.99 ± 0.06 pg), *T. infestans* (1.4 ± 0.11, 1.41 ± 0.08 pg, 1.82 ± 0.15 pg), *T. longipennis* (0.84 ± 0.06 pg), *T. maculata* (1 ± 0.07 pg), *T. nitida* (1.24 ± 0.07 pg), *T. platensis* (1.22 ± 0.03 pg), *T. pseudomaculata* (1.04 ± 0.07 pg), *T. rubrovaria* (1.08 ± 0.03 pg), *T. ryckmani* (1.01 ± 0.07 pg), *T. sordida* (1.03 ± 0.06 pg).

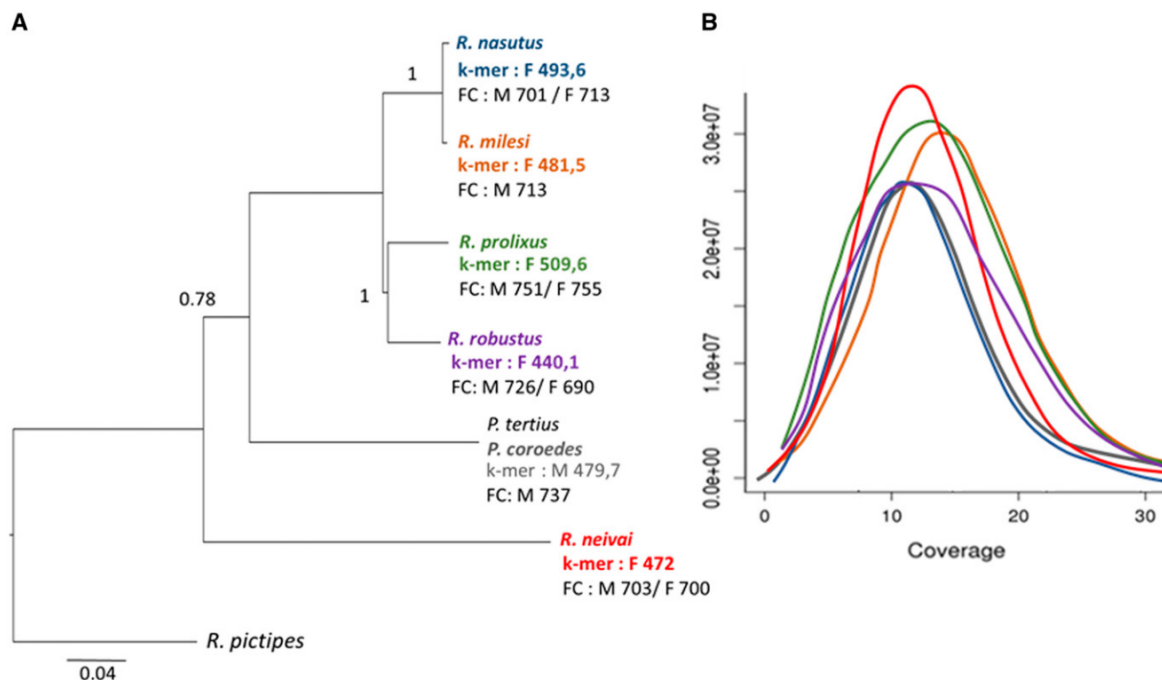
For genome size estimations based on k-mer distributions, genomic data obtained by paired-end Illumina sequencing (15–32 Gb of reads corresponding to 30–653 coverage) were used. After k-mer counting using Jellyfish software ([Marçais et Kingsford 2011](#)) with k-mer size set to 21, the genome size was estimated using GenomeScope ([Vurture et al. 2017](#)).

## Results and Discussion

For *R. prolixus*, the 1C value for males (0.77 ± 0.02 pg, 753 ± 20 Mbp) was close to that (0.75 pg) reported by [Panzer et al. 2010](#) and in the range

of the genome size assembly (702.64 Mbp) obtained after whole genome sequencing (Mesquita et al. 2015). This result validates the methodology used in our study concerning the biological material—namely, to target head tissue instead of male gonad cells as performed for Triatominae species. (Panzera et al. 2007; Panzera et al. 2010; Díaz et al. 2014; Gómez-Palacio et al. 2012). Furthermore, it simplifies the experiment and allows for the use of females regardless of their physiological status.

Among the nine species of the Rhodniini tribe, for which genome size is now available, *R. colombiensis* showed the smallest size (0.59 pg) (Díaz et al. 2014). Overall, the mean genome size (1C) was  $0.73 \pm 0.02$  pg and  $0.72 \pm 0.03$  pg for females (four species) and males (10 species), respectively, for species for which FC values are available, including our study (Table 1). *Triatoma longipennis* showed a larger genome size (1C) than Rhodniini species, with  $1.24 \pm 0.04$  pg for females and  $1.21 \pm 0.02$  pg for males, which approximates to nearly 1.2 Gb (Table 1). These values, albeit higher than those reported by Panzera et al. 2007 for this species ( $0.84 \pm 0.06$  pg), are in the range of those estimated for the genera of the Triatomini tribe (*Eratyrus*, *Dipetalogaster*, *Panstrongylus*, and *Triatoma*). In Triatominae, the genome size of *Triatoma* species (mean value,  $1.36 \pm 0.48$  pg; Table 1) is more variable than in *Rhodnius* species, with values with maximal genome size for *Triatoma delponte* ( $2n = 22$ , 2.67 pg) (Panzera et al. 2010).

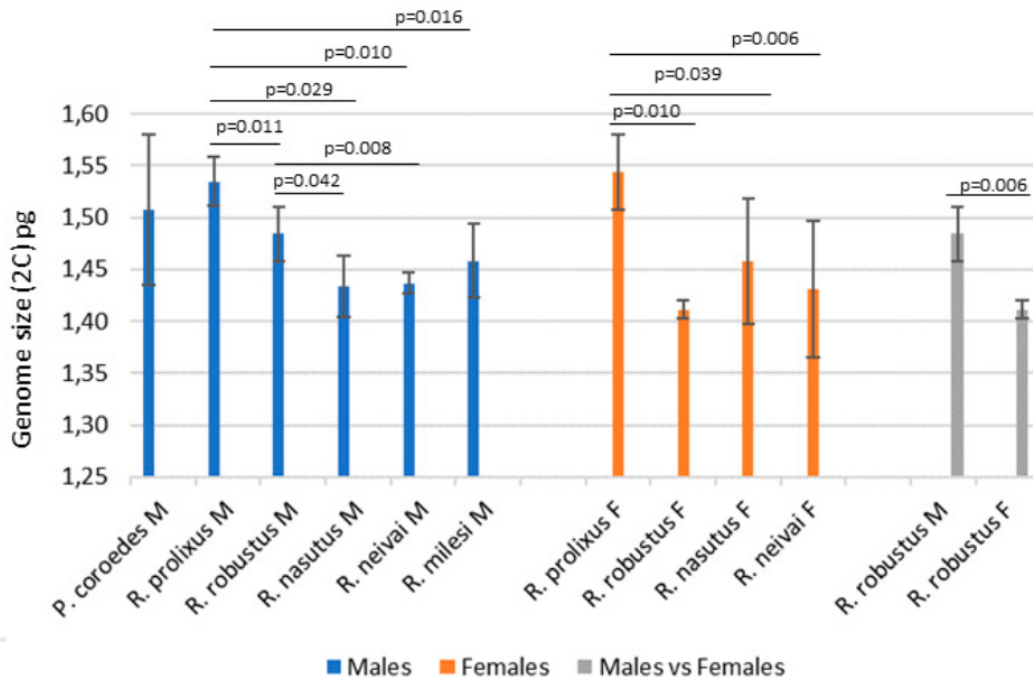


**FIGURE 1. Genome size estimations based on the k-mer distribution method for Rhodniini tribes.** (A) The k-mer values for  $k = 21$  are reported in megabases on the phylogeny performed using mitochondrial data (13 protein-coding genes: 11,094 pb) and the maximum likelihood method (MEGA software). *Rhodnius pictipes* was used as the out-group and bootstrap values are reported on the nodes ( $N = 400$ ). Genome sizes obtained with flow cytometry (FC) are given for comparison (M = males, F = females). Note that *Psammolestes tertius* was used for phylogenetic reconstruction, but *Psammolestes coroedes* for FC. (B) Distribution of the k-mer values according to the read coverage (GenomeScope) for the species studied using the same code as that used for the phylogeny.

The phylogenetic signal of genome size FC estimates was tested using Moran's I (Gittleman et al. 1990). The mitochondrial data (13 protein-coding genes: 11,094 pb; <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA4297614>) for the six studied species of the Rhodniini tribe were used to build a phylogeny (Figure 1A). Moran's I calculated with male FC genome size estimates was highly significant ( $0.354$ ;  $P < 10^{-4}$ ; 9,999 replicates), indicating that related species have a more similar genome size than random pairs of species from the phylogenetic tree. In consequence, phylogenetic generalized linear mixed models (PGLMMs) were used to analyze FC genome size variation, with species and sex as fixed effects using the R package phyr (Ives et al. 2020). PGLMMs are model-based statistics that fit observed data to underlying hypotheses about the distributions of species in a phylogeny.

Among the four models tested (a, species 1 sex; b, species; c, sex; d, none) the lowest Akaike information criterion (AIC) was for PGLMMs with species and sex as fixed effect (a: AIC = 453.70, Akaike weight [AW] = 0.973; b: AIC = 460.93, AW = 0.026; c: AIC = 513.56, AW = , 10 213 ; d: AIC = 520.29, AW =  $< 10^{-14}$ ), with a 97.3% chance for this model to be the best among the four tested ones. This analysis shows that a significant difference occurs between species when variation is freed from phylogenetic relationships.

To hierarchize the variations, we used the Mann-Whitney test (XLSTAT) recommended for small sample size (Figure 2). The greatest variation between species was found for male samples, with *R. prolixus* males having genomes significantly larger than all other species except *P. coroedes*, and with the genome size of *R. robustus* males being significantly larger than the ones of *R. nasutus* and *R. neivai*. This ranking could reflect the phylogenetic proximity between *R. prolixus* and *R. robustus* (Figure 1A), but the situation is more complex. A significant genome size difference is evidenced between *R. robustus* males and females ( $P = 0.006$ ), with females having the smallest genome (Figure 2). Indeed, for females, only *R. prolixus* showed a larger genome than the three species *R. robustus*, *R. nasutus*, and *R. neivai*.



**FIGURE 2. Genome size comparison among Rhodniini tribes and between sexes using the Mann-Whitney test.** Genome sizes (2C) are expressed in picograms. Significant pairwise differences in mean size are indicated by horizontal lines with adjusted P values. For the comparison between males (M) and females (F), only the significant pair is shown. This figure appears in color at [www.ajtmh.org](http://www.ajtmh.org).

Using the k-mer distribution values according to the read coverage on female genomic data (Figure 1B), a similar pattern with FC genome size estimates was found. *Rhodnius prolixus* displays the largest genome size (510 Mbp), whereas *R. neivai*, *R. nasutus*, *R. milesi*, and *P. tertius* have smaller genomes (472–493 Mbp), but *R. robustus* shows the smallest size (428 kb) (Figure 1A). Nevertheless, k-mer estimations are smaller than those based on FC. This is relatively common and may be explained by the usage of moderate coverage sequencing data sets that fail to represent correctly the complexity of these genomes. Indeed, repeats as well as high heterozygosity, but also sequencing errors, can decrease the accuracy of this method (Pflug et al. 2020).

Differences in genome size at inter- and intraspecific levels can result from differences in chromosome number, chromosome size, and heterochromatin/euchromatin content. Changes in genome size must involve mutational mechanisms of gain and loss of DNA, including

transposable element activity, spontaneous deletions and insertions, or duplications (Petrov 2001; Naville et al. 2019). Variations in genome size could confer a selective advantage and eventually be fixed in the genome (Pellicer et al. 2018; Bourgeois et Boissinot 2019). In our data, the chromosome number explanation can be discarded because all the species of the Rhodniini tribe share the same karyotype  $2n = 22$  ( $20A + XY$ ) (Perez et al. 1992; Alevi et al. 2015). At the Triatominae scale, the chromosome number could be considered as basically constant because only three species studied (*Triatoma barberi*, *T. longipennis*, *Triatoma nitida*) have an additional X chromosome ( $2n = 20A + X_1 X_2 Y$ ) (Panzer et al. 2010). However, there is no correlation between the chromosome number and the genome size because the genome sizes of these three species are not the largest (Table 1). But, there seems to be a positive correlation between genome size and chromosome size. The Rhodniini species have low DNA values and reduced chromosome size; the Triatomini species with large chromosomes have higher DNA values (Panzer et al. 2010). At the sex level, the significant genome size differences between males and females of *R. robustus* could result from a difference in sex chromosome size. The comparison of the karyotype of *Rhodnius* species showed that in *R. robustus*, the Y chromosome was the largest sex chromosome, whereas in *R. pictipes*, *R. nasutus*, and *R. prolixus* it was the smallest (Oliveira et al. 2018; Ravazi et al. 2018). Panzer et al. 2007 suggested that changes in heterochromatin are particularly important when considering populations or closely related species to explain genome size variation. In *Rhodnius* species, with the exception of *R. colombiensis*, *R. domesticus*, *R. nasutus*, *R. pallescens*, and *R. pictipes*, all *Rhodnius* species also have constitutive heterochromatin restricted to the sex chromosome Y (Ravazi et al. 2021). Because the species with the most heterochromatin do not have similar genome sizes and the largest genomes as one would expect if there was some ET enrichment (e.g., *R. colombiensis* exhibits the smallest *Rhodnius* genome size), this criterion does not seem to explain the genome size variation within the Rhodniini tribe.

## Conclusion

Regardless of the method used, we demonstrate that *R. prolixus* has the largest genome whereas its closest relative species, *R. robustus*, has the lowest for the females. Moreover, compared with *Triatoma* species, *Rhodnius* species have smaller genome sizes with smaller heterogeneity. Further genomic studies are needed to determine the nature of the DNA responsible for the variation observed among *Rhodnius* species and to explain the global genome size reduction in *Rhodnius* with respect to the other genus of the Triatomini tribe. Such studies could explore differential transposable element contents along with the evolutionary history of sex

chromosomes. It would also be of interest to investigate the link between genome size and environmental conditions to infer whether the genome size variation is adaptive in this genus.

### Acknowledgments

Our work has benefited from the I2BC Cytometry platform (Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell [I2BC]), Gif-sur-Yvette, France). We thank Mickaël Bourges for our training in FC and his availability.

## POSTAMBULE SUR LES DIFFÉRENCES DE TAILLE DE GÉNOME

Les résultats de cette étude ont révélé des différences de tailles de génomes. D'autres études génomiques sont nécessaires afin de déterminer la nature de l'ADN responsable de la variation observée entre les espèces de *Rhodnius* et pour expliquer la réduction globale de la taille du génome de *Rhodnius* par rapport à d'autres genres de la tribu des Triatomini. De telles études pourraient explorer le contenu différentiel en éléments transposables ainsi que l'histoire évolutive des chromosomes sexuels.

Il serait également intéressant d'étudier le lien entre la taille du génome et les conditions environnementales pour déduire si la variation de la taille du génome est adaptative dans ce genre.



## CHAPITRE 3 : DYNAMIQUE DES GÉNOMES

---

### CONTEXTE ET PROBLÉMATIQUE DE L'ÉTUDE

#### Contexte

L'essor de la génomique aide à l'identification des événements génomiques à grande échelle impliqués dans l'évolution des insectes. Des études pionnières chez la drosophile, ont montré une évolution rapide en termes de gains et de pertes de gènes associée à la diversification écologique des espèces (Emerson et Gillespie 2008; Hahn et al. 2007).

#### Objectif de l'étude

Notre étude vise à comparer la dynamique des génomes des espèces du genre *Rhodnius* de différents habitats. Une question majeure était de savoir si la domiciliation de certaines espèces s'accompagnait d'une évolution rapide en termes de gains et de pertes de gènes et impliquait des familles de gènes clés dont les expansions ou les contractions pouvaient permettre une diversification fonctionnelle. Nous avons séquencé, assemblé et annoté le génome de onze espèces de *Rhodnius* dont quatre espèces domiciliées et sept espèces sylvatiques et généré un arbre phylogénétique ultramétrique.

### PRINCIPAUX RÉSULTATS DE L'ÉTUDE

Dans cette étude nous avons assemblé et annoté 11 génomes d'espèces de *Rhodnius* vecteurs de la maladie de Chagas. L'assemblage de ces génomes à l'aide de données de séquençage *long reads* et *short reads* est de très bonne qualité. Ces assemblages nous ont permis d'explorer les différences entre les 5 espèces domiciliées (*R. domesticus*, *R. nasutus*, *R. neglectus*, *R. pallens* et *R. prolixus*) et les 6 espèces sylvatiques (*R. brethesi*, *R. colombiensis*, *R. coreodes*, *R. montenegrensis*, *R. pictipes* et *R. robustus*). Grâce à une datation estimant l'âge du dernier ancêtre commun aux *Rhodnius* à 17 millions d'années, nous avons pu calculer le taux de gains et pertes de gènes par millions d'années. Nous avons démontré que ce taux était 16 fois plus élevé pour les branches terminales des espèces domiciliées de Triatominae que pour les autres branches de l'arbre (où il ne s'est pas produit d'événement de domiciliation). Nous montrons que la domiciliation de certaines espèces s'est accompagnée d'une évolution

rapide en termes de gains et de pertes de gènes associés à la diversification écologique. Nous avons identifié des familles de gènes candidats impliqués l'adaptation aux anthroposystèmes comprenant des gènes liés à l'alimentation des insectes, à la perception sensorielle (récepteurs odorants), mais aussi à la résistance aux insecticides.

## ARTICLE 3 : GENOME DYNAMICS OF THE *RHODNIUS* GENUS, VECTOR OF CHAGAS DISEASE

Marie Merle<sup>1</sup>, Florence Mougél<sup>1</sup>, Jonathan Filée<sup>1</sup>, Claire Capdevielle Dulac<sup>1</sup>, Elaine Folly-Ramos<sup>2</sup>, Carlos Eduardo Almeida<sup>3</sup> and Myriam Harry<sup>1</sup>

(1) Université Paris-Saclay, CNRS, IRD, UMR EGCE, 91190, Gif-sur-Yvette, France.

(2) Departamento de Engenharia e Meio Ambiente – DEMA, Universidade Federal da Paraíba, João Pessoa, Brazil

(3) University of Campinas–UNICAMP, São Paulo, Brazil

### Abstract

The rise of genomics allows to identify the large-scale genomic events involved in insect evolution. Our study aimed to compare the genomes and their dynamics in species of the genus *Rhodnius* living in different habitats. A major question was whether the domiciliation of certain species was accompanied by rapid evolution in terms of gene gains and losses and which key gene families were involved in functional diversification through expansions or contractions. In this study we assembled and annotated 11 genomes of *Rhodnius* species vectors of Chagas disease (5 domiciliary species: *R. domesticus*, *R. nasutus*, *R. neglectus*, *R. pallescens* and *R. prolixus* and the 6 sylvatic species: *R. brethesi*, *R. colombiensis*, *R. coreodes*, *R. montenegrensis*, *R. pictipes* and *R. robustus*). Using fossil record information to estimate evolutionary rate, we dated the age of the last common ancestor of *Rhodnius* to be 17 million years. We then were able to calculate the rate of gene gain and loss per million years. We showed that this rate was 16 times higher for the terminal branches of the Triatominae domiciliary species than for the other branches of the tree. We show that the domiciliation of some species was accompanied by a rapid evolution in terms of gains and losses of genes associated with ecological diversification. We have identified candidate gene families involved in the adaptation to the anthroposystems including genes devoted to the feeding of bugs, the chemosensory reception (odorant receptors), but also their resistance to insecticides.

### Introduction

Arthropods are among the species that evolve rapidly and have a high due their ability to adapt to novel environments and pressures (Pélissié et al. 2018). The rise of genomics can help identify large-scale genomic events involved in insect evolution. Moreover, comparative phylogenomics allows to explore key gene families having undergone significant expansion that

favours functional diversification or regulatory evolution. Pioneer studies in *Drosophila*, showed a rapid evolution in terms of gene gains and losses associated with ecological species diversification (Emerson et Gillespie 2008; Hahn et al. 2007). Accelerated evolutionary rates are also observed in phenotypical adaptation such as body size reduction in some parasitoid wasps. Xu et al. (2021) demonstrated that a set of genes might be responsible for such convergent adaptation to miniaturization. Adaptation, and the correlated enhanced evolutionary rate could also be associated with a specific diet. In Hemiptera, mutualistic obligate endosymbiosis is observed for insects of the sub-order Sternorrhyncha and comparative analyses support the expansion of amino acid transporters involved in the transport of essential amino acids between symbionts and insect hosts to supply low amino acids abundances in their phloem-based diet (Dahan et al. 2015). For blood-sucking insects including seven diptera and two bugs, it has been shown that while the rates of gene gain and loss are slow compared to non-blood-sucking insects with a small number of small rapidly evolving gene families, hematophagous insects suffer a significantly higher number of rapid contractions (Freitas et Nery 2020). Within blood-sucking bugs (Reduviidae, Triatominae), species of the genus *Rhodnius* which includes sylvatic and recent domiciliary species constitute a good working model for testing the genome dynamics according to their habitat.

All members of Triatominae can transmit *Trypanosoma cruzi* (Trypanosomatida; Trypanosomatidae), the causative agent of Chagas disease in Latin America where about 6 million people are infected (PAHO, 2020). In the Triatominae subfamily there are currently more than 157 species including 3 fossils (Alevi et al. 2021). The genus *Rhodnius* includes 24 species since the three species of *Psammolestes* genus could be considered as *Rhodnius* species based on phylogenomics (Filée et al. 2022). The genus is divided into three groups namely *pictipes*, *pallescens* and *prolixus*, first described on the basis of morphological data but also on geographical distribution. Indeed, the *pallescens* group spread over Central America and the western Andes, from which it derives its name of "trans-Andean" group also called "trans-*pallescens*" group by some authors (Hernández et al. 2022). The seven species of the *pictipes* group are cis-Andean, distributed in the eastern part of the Andes and in the Amazon region (Justi and Galvao, 2017). The same distribution is observed for ten species of the *prolixus* group, while the eleventh species *R. neivai* also has some trans-Andean populations (Zhao et al. 2021). The phylogenetic relationships between the groups was debated (Hernandez et al. 2020) but phylogenomic studies using large molecular dataset showed that *pallescens* and *pictipes* groups are sister groups (Filée et al. 2021). The *Rhodnius* species live in palms trees (e. g. *Attalea* spp) or bromeliads where they feed on a wide variety of vertebrate hosts (Arias-Giraldo et al. 2020). Due to anthropogenic pressures, some species have left their sylvatic habitat and colonised human dwellings. These domiciliary species that transmit to humans the

Chagas disease, are able to complete their entire life cycle in the anthroposystems. While a transition to a new lifestyle would favor gene expansion and/or contraction, it is of interest to explore such events at the genomic scale. In addition, considering that the domiciliation process involves species of different evolutionary lineages (e. g. *Rhodnius* groups), analysis of such evolutionary transition represents an opportunity to study patterns of convergent molecular evolution.

Because to date only the genome of the domiciliary species *R. prolixus* is described (Mesquita et al. 2015), the first step of the study was to provide reliable genome assembly for both domiciliary and sylvatic species. In this context, we have sequenced, assembled and annotated the genome of eleven *Rhodnius* species (4 domiciliary species namely *R. nasutus*, *R. neglectus*, *R. pallescens*, *R. prolixus*; and 7 sylvatic species namely *R. brethesi*, *R. colombiensis*, *R. coroedes*, *R. domesticus*, *R. montenegrensis*, *R. pictipes*, *R. robustus*). In order to explore expansion and contraction events, we established an ultrametric phylogenetic tree describing relationships among *Rhodnius* species and more generally in Paraneoptera including Hemiptera. This phylogeny was calibrated based on seven fossil records and allowed us to estimate evolutionary rates. We showed that *Rhodnius* genomes evolve rapidly and that the domiciliation of *Rhodnius* species can even accelerate genome content change.

## Materials and Methods

### *Sampling and sequencing*

We used laboratory CTA strains from the “Insetário de Triatominae da Faculdade de Ciências Farmacêuticas”/Unesp/Araraquara, Brazil (A. Da Rosa & J. Oliveira), but *R. prolixus* samples were from IRBI lab, France (C. Lazzari). Samples were stored in absolute EtOH except for samples sequenced with the Gridlon technology freshly stored at -80°C (Table 1).

DNA was extracted from legs and alary muscles from adult bugs, using the Nucleobond AXG kit (Macherey Nagel). In order to obtain short read sequences (paired-end, 2 x 150 bp), whole-genome shotgun sequencing was performed using Illumina technology (Genotoul platform, Toulouse). For three species *R. montenegrensis*, *R. nasutus* and *R. neglectus* Illumina data is described in Filée et al. (2022).

Long read sequences were obtained over a long period of time where techniques improved rapidly. As a consequence, the dataset was obtained with different technologies and coverage. We used Pacbio technology (7x), at the I2BC (Gif-sur-Yvette, France) for 8 species and GridION technology

(18- 24x) at the Genotoul platform from (Toulouse, France).

RNA was extracted from the antennae and rostrum using the TRIzol® Reagent kit according to [Marchant et al. \(2016\)](#). The RNA was used to prepare cDNA libraries using the ScriptSeq™ Complete Kit Low Input (Epicentre). Libraries were sequenced in paired-end (2 x 100 bp) on Illumina® HiSeq 1000 platform (IMAGIF, CNRS, France).

**Table 1: Samples and Sequencing data.** In green: sylvatic and sporadic species, in yellow: sporadic species, in red: domiciliary species) In green: sylvatic species, in red: domiciliary species. Most samples are from laboratory strains (referenced as CTA or IRBI). 50 M: 50 millions of reads.

\* Obtained from [Filée et al. \(2022\)](#)

Sample description							Sequencing technology and depth (x)		
Species	<i>Rhodnius</i> group	origin	Strain	habitat	Sex	sample name	RNA	long read:	short read
<i>R. brethesi</i> Matta, 1919	<i>pictipes</i>	Igarapé Tucunaré, Brazil	CTA 222 (2009)	Sylvatic / Sporadically in dwellings (Amazon basin)	M	BAM	Illumina 50M	PacBio (7x)	53x
<i>R. colombiensis</i> Mejia et al. 1999	<i>pallescentes</i>	Tolima, Colombia	CTA 050 (2001)	Sylvatic / Sporadically in dwellings	M	COLM	Illumina 50M	PacBio (7x)	49x
<i>R. coreodes</i> Bergroth, 1911	<i>prolixus</i>	Mato Grosso do Sul, Brazil	PC (2018)	Sylvatic	M	PC		GridION (24 x)	74x
<i>R. domesticus</i> Neiva and Pinto, 1923	<i>prolixus</i>	Santa Catarina, Brazil	SC (1999)	Sylvatic / Sporadically in dwellings	M+F	DOMFM	Illumina 50M	GridION (18x)	40x
<i>R. montenegrensis</i> da Rosa et al. 2012	<i>prolixus</i>	Monte Negro, Rondônia, Brazil	CTA 087 (2003)	Sylvatic	F	monR8	Illumina 50M	PacBio (7x)	18x *
<i>R. nasutus</i> Stal, 1959	<i>prolixus</i>	Brazil	CTA 054 (1999)	In domiciliation process in Brazil	F	nasR7	Illumina 50M	PacBio (7x)	16x *
<i>R. neglectus</i> Lent, 1954	<i>prolixus</i>	Frutal, Minas	CTA 061	In	F	negR5	Illum	PacBio	23x *

		Gerais, Brazil	(1983)	domiciliation process in Brazil			ina 50M	(7x)	
<b><i>R. pallescens</i></b> Lent, 1953	<i>pallescens</i>	Colombia	CTA s/R	Domiciliar/ Principal vector in Panama	M	PALM	Illumina 50M	PacBio (7x)	64x
<b><i>R. pictipes</i></b> Stal, 1972	<i>pictipes</i>	Belém, PA, Brazil	CTA 072 (1998)	Sylvatic/Sporadically in dwellings in Amazon region	M	PICM	Illumina 50M	PacBio (7x)	79x
<b><i>R. prolixus</i></b> Stal, 1859	<i>prolixus</i>	Honduras	IRBI laboratory	Domiciliar Principal vector in Central America	F	ProH2	Illumina 50M	PacBio (7x)	82x
<b><i>R. robustus</i></b> Larousse, 1927	<i>prolixus</i>	Rondônia, Brazil	CTA 272	Sylvatic, Sporadically in dwellings in Amazon region	M	RTM	Illumina 50M	GridION (17x)	46x

### Genome assemblies

We performed a de novo hybrid assembly for each *Rhodnius* species using both long reads and shorts reads data using the MaSuRCA v.4 program (Zimin et al. 2013) with the modified version of the CABOG assembler (Miller et al. 2008). To assemble *R. montenegrensis*, *R. nasutus* and *R. neglectus* genomes, Illumina's adaptors were removed using Cutadapt v.1.18 (Martin, 2011), and PacBio contigs were auto-corrected using the mode correct of Canu v.1.7 (Koren et al. 2017). To assemble the other species, 150 pb illumina reads and Gridlon reads were used without any upstream processing.

Quality of the genome and transcriptome assemblies was assessed using QUAST (Gurevich et al. 2013) to count the contig numbers, the N50 and the size of the largest contig, and using BUSCO to estimate the completeness based on insects unique-copy orthologs (Simão et al. 2015) using the insecta\_obd10 dataset.

For genome size estimations based on k-mer distributions, genomic data obtained by paired-end Illumina sequencing were used. After k-mer counting using Jellyfish software (Marçais et Kingsford 2011) with k-mer size set to 21, the genome size was estimated using GenomeScope (Vurture et al. 2017).

### Transposable Elements and coding gene annotation

In order to annotate repeated elements, we used RepeatModeler on the contig sequences of the different species. Transposable elements were classified as non-LTR retrotransposons including short interspersed repetitive elements (SINEs), Long Interspersed Nuclear Elements (LINEs), long terminal repeat (LTR)-retrotransposons and DNA transposons.

Before the coding gene annotation, we performed a soft masking of the repeated sequences, using the RepeatMasker program (Smit, AFA, Hubley, R. RepeatModeler Open-1.0. 2008-2015 <http://www.repeatmasker.org>). Then, the RNAseq raw reads were aligned against the soft masked genome with minimap2 (Li, 2018) and coding sequences were determined using these alignments with the Braker2 pipeline combining GeneMark-ES/ET and AUGUSTUS tools (Bruna et al. 2021). For three species, as RNAseq data were not available, we used RNAseq data from the closest species. For *R. domesticus* and *R. coreodes* we used RNAseq from *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. prolixus* (Marchant et al. 2015) and *R. robustus* (Marchant et al. 2015). For *R. prolixus* and *R. robustus* we used *R. montenegrensis*, *R. prolixus* (from Marchant et al. 2015) and *R. robustus* (from Marchant et al. 2015) because the Illumina RNAseq sequencing data from Marchant et al. 2015 were too incomplete to be used alone.

### Time trees from nuclear and mitochondrial data

The ultrametric tree built from nuclear data was based on the alignment from Filée et al. 2022 including 51 BUSCO coding-genes (36.3 kb), completed with newly sequenced samples. The *R. prolixus* mitogenome available on NCBI was used as a template to find mitochondrial contigs in our MaSuRCA assemblies. The sequences retrieved were then used to build a second ultrametric tree.

Both analyses include outgroups from Blattodea and other Hemiptera. We used 7 calibration points to calibrate both trees (Supp Table 1) using fossils data (from <https://fossilinsectdatabase.co.uk/> database). Best-fit nucleotide substitution models for the mitochondrial and nuclear datasets were selected by jModelTest 2.1.4. under the Akaike Information Criterion (Posada 2008). Then, ultrametric trees were constructed with concatenated datasets using BEAST 2.6.3 (Bouckaert et al. 2014). BEAUti was used to configure the xml-file used as input for the BEAST analysis, applying following settings; a lognormal relaxed clock model, enforcement of the “estimate” option in the clock model, a Birth-Death Incomplete Sampling process Tree Prior. All other settings were default. We performed 5 runs for



10,000,000 generations, sampled every 1,000 generation, using a burnin at 25% for each run. For each run and the concatenated run, the convergence of parameters was visualized using Tracer 1.7.1 (Rambaut et al. 2018). The consensus tree was determined using TreeAnnotator (Bouckaert et al. 2014). All topologies were visualized using FigTree ver. 1.4.4 (Rambaut 2010 <http://tree.bio.ed.ac.uk/software/figtree/>).

#### *Estimating gene gain and loss along the phylogeny*

The orthologous gene families were identified by OrthoMCL (Li et al. 2003) with default parameters (inflation index 1.5). These gene families, combined with the ultrametric phylogenetic trees, were submitted to Computational Analysis of gene Family Evolution (CAFE5) software package (Mendes et al. 2020) using default parameters. CAFE estimates the gene gain and loss rate parameter, lambda ( $\lambda$ ), and the size of each gene family at ancestral nodes of the tree based on maximum likelihood approach. These MLEs (maximum likelihood estimates) can then be used to infer expansions (gain) and contractions (loss) of individual gene families on any lineage, estimated per gene per million years. The software also allows for separate lambda values i) on different branches of the tree (e. g. domiciliary species vs sylvatic ones) and ii) for loss and gain. The software takes into account errors in the observed gene family sizes among genomes even with low assembly and annotation quality thanks to the application of an error model. The maximum likelihood value of each rate parameter was estimated to retain the best model (e. g. with the higher ML value). We also calculated the evolutionary rate for each terminal branch species by using the ratio of the length of the branch and the number of gene gains or losses.

#### *Gene Ontology enrichment analysis*

GO enrichment was tested using the R package topGO (Alexa et al. 2006). GO annotations were derived from those available on vectorbase for *R. prolixus* Rpro CDC set (<https://vectorbase.org>). For each cluster of orthology (COG), the protein sequence encoded by the first gene belonging to this COG was blasted against Rpro CDC protein dataset. Each set of COG identified by the comparative approach between domiciliary and sylvatic species was thus transformed in a set of unique Rpro CDC genes. These sets were compared for their GO content to the content of the whole Rpro CDC. We compared the results of 4 algorithms implemented in topGO, namely "classic", "elim", "weight" and "weight01" that integrate hierarchical structure of Gene Ontology in different ways. Keeping nodes larger than 5, we computed Fisher tests for the three categories of "Biological Process", "Molecular Function", and "Cellular Compartment". Only the GO terms

detected as significantly enriched at the 5% level with the 4 algorithms were kept.

### Gene functional annotation

We searched for functional information for gene families of specific interest. We first focused on the 20 families with the highest number of events (gain+loss) whatever the habitat of the vector. We then looked at orthologous families showing a minimum of 3 gain and loss events in at least 2 species with contrasted changes between domiciliary and sylvatic species. Gene families were thus defined as “domiciliary +” when no change or contraction events were observed in domiciliary species while no change or expansion events occurred in sylvatic ones. Gene families with opposite distribution of changes were qualified as “sylvatic +”. For each gene family of interest, we identified the orthologs in the reference genome of *R. prolixus* in VectorBase to transfer gene annotation. Blast searches against non-redundant protein sequences set available at NCBI were conducted to refine annotation. Genes were categorized into functional classes according to their Gene Ontology (GO) using GO-slim terms for biological processes, cellular components, and molecular functions.

## Results

### Genome quality

Using whole-genome sequencing, we generated *de novo* genome assemblies for eleven *Rhodnius* species using hybrid strategy by combining Illumina short-read and PacBio/GridION long-read sequencing data. The different metrics are summarized in [Table 2](#).

The total length of the assemblies varies from 269.8 Mb (*R. nasutus*) to 579.4 Mb (*R. coreodes*). Considering the eleven species, the mean *Rhodnius* genome size estimated from ungapped assembly length is  $457 \pm 92$  Mb. The low Illumina coverage (under 20x) may explain the smaller total length of the *R. nasutus* and *R. montenegrensis* assemblies. We therefore recommend illumina sequencing depth above 30x for further *Rhodnius* genome sequencing. Assemblies performed with GridION long reads (*R. coreodes*, *R. domesticus* and *R. robustus*) showed higher N50 values (from 160,096 bp for *R. robustus* to 293,352 bp for *R. domesticus*) than those obtained from Pacbio long reads (from 30,461 bp for *R. nasutus* to 71,903 for *R. neglectus*). A notable exception is observed for *R. colombiensis* that exhibited a high N50 value (152,759 bp) despite PacBio 7x sequencing approach. The genome completeness ([Table 2](#), [supp Table 3](#)) varies from 86% (*R. pictipes*) to 95.1% (*R. colombiensis*) for well-assembled genomes. Unsurprisingly, these values

were lower for the two genomes with lower assembly length namely : 61.4% for *R. nasutus* and 79.6% for *R. montenegrensis*. The number of coding DNA sequences (CDS) varies from 17,884 (*R. colombiensis*) to 28,088 (*R. coreodes*) with an average of  $20,130 \pm 6,940$ . Even if the total length of the assembly is correlated with different parameters such as the N50 ( $r=0.873$ , padj Bonferroni=0.01) and the TE content ( $r=0.855$ , padj Bonferroni=0.02), no significant correlation is detected with the number of CDS ( $r=0.4$ , padj Bonferroni=1) (Supp Table 4).

**Table 2: Metrics of the genome assemblies of the 11 *Rhodnius* species.** CDS=number of annotated coding sequences, %CDS is the ratio between CDS length and total assembly length, BUSCO S is the percentage of singleton of the insecta\_odb10 gene set retrieved in the assemblies calculated with BUSCO v5. RE : repetitive sequences. Genome size estimations based on the k-mer distribution method.

Genome assemblies	Total ungapped length (b)	number of contigs	N50	Busco S	% RS	CDS	% CDS	Genome size estimation from k-mer method (b)
<i>R. brethesi</i> (BAM)	468,070,282	10,466	74,676	88.4	32.07	17,950	4.96	482,663,708
<i>R. colombiensis</i> (COLM)	479,825,141	7,465	152,759	95.1	33.21	17,884	4.43	476,492,503
<i>R. coreodes</i> (PC)	579,397,083	5,000	194,642	95	37.7	28,088	4.36	441,154,399
<i>R. domesticus</i> (DOMFM)	538,753,638	3,910	293,352	94.7	35.5	22,363	4.64	465,291,504
<i>R. montenegrensis</i> (R8F)	353,109,601	10,354	50,332	79.6	32.65	20,956	5.33	544,305,556
<i>R. nasutus</i> (R7WU)	269,814,734	11,363	30,461	61.4	29.45	19,024	5.96	553,651,790
<i>R. neglectus</i> (negR5)	420,276,007	9,263	71,903	86.6	32.75	23,668	5.36	517,700,729
<i>R. pallescens</i> (PALM)	410,473,381	11,565	68,364	91.2	29.41	18,919	4.76	217,461,324
<i>R. pictipes</i> (PICM)	462,908,123	11,624	69,704	86	34.75	18,632	5.00	472,466,211
<i>R. prolixus</i> (ProH)	474,000,233	13,347	57,843	88	34.55	26,724	5.25	505,302,096
<i>R. robustus</i> (RTM)	570,799,179	6,344	160,096	89.8	37.67	23,377	5.06	561,512,576

### Genome size estimation

Using k-mer estimations, the mean genome size for the *Rhodnius* species is about  $468 \pm 41$  Mb. Genome size varies from 441 Mb (*R. coreodes*) to 561 Mb (*R. robustus*, RRTM) that is a variation of 27% among the two extremes values. All the three species that belonged to another group than

the *prolixus* one exhibited small genome size (< 490 Mb, mean = 477 ± 5 Mb). In the *prolixus* group, the genome size is more variable (mean = 513 ± 46 Mb), two species exhibiting small genomes (*R. coreodes*, *R. domesticus*, <465 Mb) while the others have higher values (>505 Mb).

### Repetitive elements

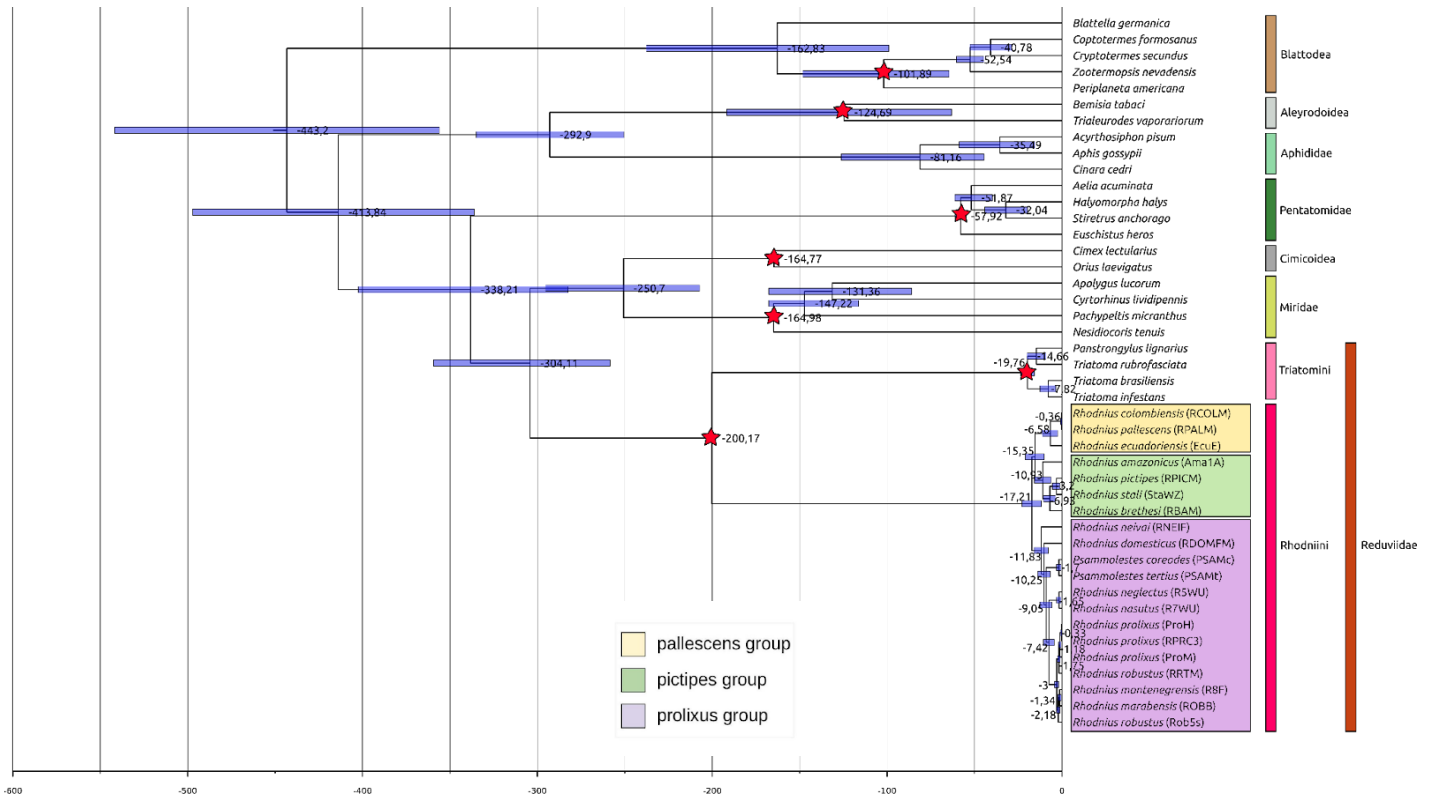
RepeatModeler was used for the annotation of the repetitive elements (Table 3). In the genomes, the content of repetitive sequences including satellite repeats varies from about 29% (*R. pallescens*) to 38% (*R. coreodes*, *R. robustus*) with an average of 33.6 ± 2.8 % (Table 2). For all the species the genome proportions of TE differs between classes with LINES non-LTR retrotransposons (6-8%) always in higher amounts than LTR retrotransposons (0.1-0.6%). Note that SINEs are only retrieved from genomes assembled with long reads sequenced with GridION technology and are present in low amounts (0.4-0.7%). Genome content in LTR retrotransposons is lower in the domiciliary species (0.14-0.25%) than in the sylvatic species (0.21-0.65%). The genome proportion of DNA transposons is the most variable between the species. For the domiciliary species this amount is low (2-3%) except for *R. prolixus* with 8%. For the sylvatic species, except for *R. colombiensis* (2%) the amount is higher (3-7%) with a remarkable value for *R. robustus* (10.9%). It results that the total of TE elements varies between the species from about 8% (*R. nasutus*) to 20% (*R. robustus*) with an average of 12.46 ± 3.74 %. Overall domiciliary species exhibit lower TE amounts (8-10%), than the sylvatic species (11-20%) except for the domiciliary *R. prolixus* (16%) and for the sylvatic *R. colombiensis* (<10%). No significant correlation is detected between the genome size estimated by the k-mer method and the TE content ( $r = -0.22$ ,  $p$ -value = 0.5367).

**Table 3: Summary table of the TE annotated in the 11 genome assemblies with RepeatMasker.** Total length (bp) and percentage reported to total assembly length of the different classes of repeats are provided.

Species	SINEs (bp)	SINEs (%)	LINEs (bp)	LINEs (%)	Total Non-LTR (%)	LTR elements (bp)	LTR elements (%)	DNA elements (bp)	DNA elements (%)	Total ET (%)	total (pb)	total %
<i>R. brethesi</i>	0	0.00	36,042,587	7.70	7.7	984,132	0.21	15,399,333	3.29	11.2	150,095,135	32.07
<i>R. colombiensis</i>	0	0.00	34,498,702	7.19	7.19	1,375,445	0.29	10,193,605	2.12	9.6	159,327,645	33.21
<i>R. coreodes</i>	3,963,263	0.68	43,205,618	7.46	8.14	2,998,221	0.52	42,843,973	7.39	16.05	218,415,46	37.

											8	70
<i>R. domesticus</i>	3,890,371	0.72	37,675,285	6.99	7.71	2,338,061	0.43	31,925,446	5.93	14.07	189,931,908	35.25
<i>R. montenegrensis</i>	0	0.00	24,825,646	7.03	7.03	1,034,675	0.29	13,431,284	3.80	11.12	115,291,763	32.65
<i>R. nasutus</i>	0	0.00	15,820,235	5.86	5.86	670,181	0.25	6,236,845	2.31	8.42	79,464,748	29.45
<i>R. neglectus</i>	0	0.00	27,258,173	6.49	6.49	1,050,014	0.25	12,389,267	2.95	9.69	137,633,382	32.75
<i>R. pallescens</i>	0	0.00	27,992,984	6.82	6.82	787,277	0.19	7,366,073	1.79	8.8	120,701,956	29.41
<i>R. pictipes</i>	0	0.00	37,410,158	8.08	8.08	2,996,732	0.65	14,234,233	3.07	11.8	160,853,630	34.75
<i>R. prolixus</i>	0	0.00	36,234,438	7.64	7.64	683,637	0.14	38,526,062	8.13	15.91	163,745,230	34.55
<i>R. robustus</i>	2,476,286	0.43	48,254,107	8.45	8.88	3,594,077	0.63	62,243,030	10.90	20.41	215,029,499	37.67
Mean (%)		0.17 ±0.29		7.25 ±0.73	7.41 ±0.85		0.35 ±0.18		4.70 ±2.97	12.46 ±3.74		33.59 ±2.78

Divergence Time



**Figure 1: Divergence time estimates based on BEAST analysis using relaxed-clock model and 7 fossil calibration points.** Placement of fossils as calibration points of clades are indicated by red stars. Blues node bars indicate 95% highest posterior density (HPD) credibility intervals.

The dating of divergence times takes into account the Hemiptera species for which a complete genome was available and allowed us to set fossil calibration points. We used the Calibrated Yule model that is generally more appropriate when considering sequences from different species according to the BEAST manual. For the ultrametric tree from nuclear data (Figure 1), we used the alignment from Filée et al. 2022 including 51 BUSCOs coding-genes (36.3 kb). The last common ancestor of the *Rhodnius* species was estimated to 17 (11.8-22.9) million years ago (MYA), the divergence of the *prolixus* group around 11.8 (8.0-15.9) MYA and the split between the *pallescens* and *pictipes* group around 15.3 (10.1-20.8) MYA. Some species of the genus *Rhodnius* are quite young since they differentiated no more than 2 MYA (Table 4, e. g. *R. pallescens* - *R. colombiensis* 0.4 (0.1-0.7) MYA). Based on the ultrametric tree from mitochondrial data (Supp Fig 1), the datation are multiplied by a factor of about four with respect to nuclear data, which is 70.9 (41.4-105.9) MYA for the last common ancestor of the *Rhodnius*

species, 54.8 (30.9-80.2) MYA for the divergence of *prolixus* group and 53.5 (30.9-81.9) MYA for the division of *pallescens* and *pictipes* groups.

**Table 4: Table of divergence times between different taxa determined with BEAST analysis for nuclear data (51 PCG, 36.6 kb).** The mean age estimate for each node (CA) is shown in the table as well as the confidence interval corresponding to the 95% highest posterior density credibility intervals (HPD).

Cladogenetic event	age (mean)	Confidence interval (95% HPD)	
		lower bound	upper bound
<b>CA <i>Rhodnius</i></b>	17.2112	11.7835	22.8589
CA <i>pallescens</i> group / <i>pictipes</i> group	15.3544	10.1228	20.8198
<b>CA <i>pallescens</i> group</b>	6.5799	2.4893	11.1934
CA <i>R. colombiens</i> / <i>R. pallescens</i>	0.3556	0.0958	0.6831
<b>CA <i>pictipes</i> group</b>	10.9294	6.4154	15.8249
CA <i>R. brethesi</i> / ( <i>R. pictipes</i> / <i>R. stali</i> )	6.9303	3.7017	10.4579
CA <i>R. stali</i> / <i>R. pictipes</i>	3.1971	1.2822	5.417
<b>CA <i>prolixus</i> group</b>	11.8286	7.9592	15.896
CA <i>R. domesticus</i> / ( <i>Psam spp, nas, neg, pro, mon, rob</i> )	10.2542	5.7642	12.5681
CA ( <i>R. nasustus, R. neglectus</i> ) / ( <i>pro, mon, rob</i> )	7.4176	4.4433	10.5407
CA <i>R. nasutus</i> / <i>R. neglectus</i>	1.6484	0.5707	2.992
CA ( <i>Rcor + Rter</i> ) / ( <i>neg, nas, pro, rob, mon</i> )	9.0456	7.9592	15.896
CA <i>P. tertius</i> / <i>R. coreodes</i>	1.6974	0.5688	3.046
CA ( <i>pro + RRTM</i> ) / ( <i>mon, rob</i> )	3.0032	1.8662	4.1733
CA RRTM / ( <i>pro</i> )	1.7508	5.7642	12.5681
CA ProM (ProH / RPRC3)	1.1806	0.9577	2.5592
CA ProH / RPRC3	0.3306	0.1044	0.6053
CA Rob5s / ( <i>R8, ROBB</i> )	2.1807	1.1641	3.2396
CA R8 / ROBB	1.3368	0.527	2.236
<b>CA Triatomini</b>	19.7615	15.7146	23.5986
CA <i>T. brasiliensis</i> / <i>T. infestans</i>	7.8192	3.3942	12.5089
CA <i>T. lignarius</i> / <i>T. rubrofasciata</i>	14.6556	9.2429	19.8952

CA Reduviidae	200.1663	198.2057	202.1271
---------------	----------	----------	----------

### Gains and losses

Two independent analyses were conducted to identify gene families with rapid expansions and/or contractions in copy numbers. A first one was performed with the complete tree including Hemiptera+Blattodea lineages. A second one was restricted to Triatominae species. A first step consisted in estimating the total number of ortholog groups (OGs) with OrthoMCL. We identified 51,029 OGs with Hemiptera+Blattodea dataset and 32,718 with Triatominae one. We then studied the variation in copy number for each OG and the gain and loss rate along the trees with CAFE5.

CAFE5 calculates an error rate (and the error rate distribution) for the estimate of gene copy number. The mean error rate was estimated at less than 10% of gene families in our datasets ( $e=0.086$ ,  $e=0.005$ , for Hemiptera+Blattodea and Triatominae respectively). Using this error distribution, we tested several models using the timetree inferred from nuclear data because the portion of the mitochondrial genome is negligible compared to the nuclear genome. The birth and death rate ( $\lambda$ ), which represents the average rate of genomic turnover (gains and losses) per gene per million years was inferred. For the Hemiptera+Blattodea tree (Table 5, Supp Fig 2), we tested three models namely i) the global model, where the whole tree has the same  $\lambda$ , ii) the Triatominae model where the blood-sucking bug terminal branches have a distinct  $\lambda$  compared to the non-blood-sucking bug branches, and iii) the *Rhodnius* model where the *Rhodnius* terminal branches have a distinct  $\lambda$  compared to the other branches. While the  $\lambda$  was similar for the global model ( $2.67 \times 10^{-3}$ ), and the Triatominae model ( $2.15 \times 10^{-3}$  vs  $2.65 \times 10^{-3}$  for the Triatominae vs non-Triatominae branches), an accelerated rate was detected in the *Rhodnius* branches with the *Rhodnius* model ( $5.90 \times 10^{-3}$  vs  $2.49 \times 10^{-3}$  for the *Rhodnius* versus non-*Rhodnius* branches). Moreover, the *Rhodnius* model exhibited the lower maximum likelihood value, meaning that it fits the data better compared to the other models.

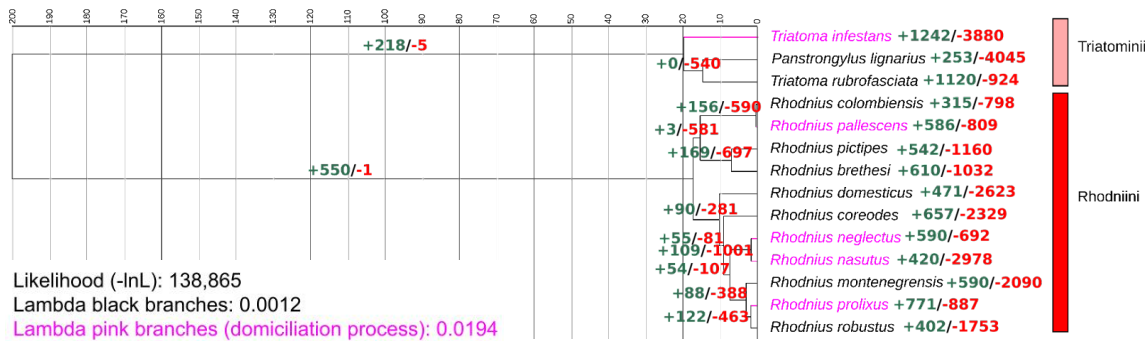
For the Triatominae tree (Table 5, Figure 2), we also tested three models namely i) the global model, where the whole tree has the same  $\lambda$ , ii) the domiciliary Triatominae model where the domiciliary *Triatoma* and *Rhodnius* terminal branches have a distinct rate compared to the other branches, and iii) the domiciliary *Rhodnius* model where the domiciliary *Rhodnius* terminal branches have a distinct  $\lambda$  compared to the other branches. The domiciliary Triatominae model exhibited the lower maximum likelihood value with a rate about 16-fold higher for the domiciliary



Triatominae terminal branches than for the other branches. The  $\lambda$  rate was *also* highly different between domiciliary and other branches in the *Rhodnius* model ( $42.66 \times 10^{-3}$  vs  $1.42 \times 10^{-3}$ ) with a maximum likelihood still low (see Table 5). This suggests that the domiciliation event in the genus *Rhodnius* is accompanied by an acceleration of the rate of gene gains/losses. Note that *R. colombiensis* showed the lower value of gene gain on the tree.

**Table 5: Results of the different models tested with the CAFE program.** The maximum likelihoods and the estimated lambda are shown for each model. For the two rate models, lambda terminal refers to the first group listed in the model. The grey lines are the estimates obtained from the Hemiptera+Blattodea tree and the white lines are the estimates from the Triatominae tree. For each tree, the most likely model is shown in bold.

Model	-lnL	lambda terminal	lambda background
Null model	375,302	-	0.0026
<i>Rhodnius</i> + background branches	374,257	0.0059	0.0025
Triatominae + background branches	375,211	0.0021	0.0026
<b>Domiciliary <i>Rhodnius</i> + background branches</b>	<b>373,896</b>	<b>0.0441</b>	<b>0.0026</b>
Domiciliary Triatominae + background branches	374,124	0.0147	0.0025
Null model	141,295	-	0.0017
<i>Rhodnius</i> + background branches	141,107	0.0020	0.0012
Domiciliary <i>Rhodnius</i> + background branches	139,181	0.0427	0.0014
<b>Domiciliary Triatominae + background branches</b>	<b>138,865</b>	<b>0.0194</b>	<b>0.0012</b>



**Figure 2: Gains and losses of genes within ortholog groups.** Data presented is from the most likely model (Domiciliary Triatominae + background branches). Gene family expansion is indicated in red, and gene family contraction is indicated in green. The two estimated gains and losses rates are indicated in events per gene per million years.

Considering the terminal branches (Supp Table 5) and excluding *R. nasutus* whose genome is incomplete, the total number of gene gains is slightly higher for domiciliary than for sylvatic species (domiciliary/sylvatic = 1.27) while total number of gene losses for domiciliary species is half as high as for sylvatic species (domiciliary/sylvatic = 0.47). Losses are more frequent compared to gains in the sylvatic species (losses/gains = 3.29) than in the domiciliary species (losses/gains = 1.22). Considering the length of the terminal branches to calculate the evolutionary rate of gain or loss, the highest rate is for domiciliary species for both gene gains (domiciliary/sylvatic = 3.57) and losses (domiciliary/sylvatic = 1.56) with about 2.29 times greater rate for gains.

Among 32,718 ortholog groups, 8,711 showed at least one event (loss or gain) with a maximal of 98 events. Among the 20 ortholog groups with the highest number of events, five are related to TE (RT\_RNaseH domain-containing protein, Integrase\_H2C2 domain-containing protein, DUF1758 domain-containing protein) and contain from 96 to 569 genes. Some species, namely the sylvatic *R. robustus*, *R. domesticus* and *R. coreodes*, are particularly enriched in these genes (Table 7, Figure 3). The other genes are related to proteolytic activity (ZnMc domain-containing protein), catalytic activity (exonuclease/endonuclease/phosphatase domain-containing protein), zinc/metal ion binding (Zinc finger protein, CCHC-type domain, PHD-type domain-containing protein), but also kinase (FAT domain-containing protein), helicase (ATP-dependent DNA helicase) and detoxification (cytochrome P450) functions. We also searched for ortholog groups with gene expansion specific to sylvatic or domiciliary *Rhodnius* habitat. We detected 16 ortholog groups expanded in sylvatic species (sylvatic +) and 26 expanded in domiciliary species (domiciliary +)

and attempted to annotate them (Table 6). Among the “sylvatic +” ortholog groups, we annotated OGs related to TE (Dimer\_Tnp\_hAT domain-containing protein, retrovirus-related Pol polyprotein from transposon), muscular physiology (neuropilin and tolloid-like protein, Pyruvate carboxylase), immunity (translation initiation inhibitor), peptidases (CUB domain-containing protein), and methyltransferase (methyltransferase NSUN7). Among the “domiciliary +”, we annotated OGs related to insecticide resistance (sodium channel protein, gamma-glutamyl kinase, ABC transporter domain-containing protein), TE (retrovirus-related Pol polyprotein from transposon, DDE-1 domain-containing protein), immunity (thioredoxin domain-containing protein), sensory reception (Odorant Receptor), microtubule-based movement (kinesin-like protein, centrosomal protein) and intracellular traffic (ADP-ribosylation factor), immunoglobulin superfamily (turtle protein), gene-silencing (maelstrom domain-containing protein), protein binding (F-box domain-containing protein), and putative proteins involved in blood-feeding (GPI mannosyltransferase, Nitrophorin domain-containing protein, CUB domain protein).

**Table 7: Gain and losses events and annotation for ortholog groups of interest.** The ID of each OG and the number of included genes are mentioned. For each species, the inferred number of changes on terminal branches is indicated (positive values correspond to gain and negative ones correspond to loss). The total number of changes cumulates gains and losses.

FamilyID	Nb of genes	pro	neg	nas	pal	rob	mon	cor	pic	bre	col	dom	Total changes	predicted function
10	212	-6	-7	20	5	12	0	0	-2	38	-1	-7	98	ZnMc domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HU38]
7	285	23	-2	-9	-1	0	-21	28	5	-1	0	7	97	RT_RNaseH domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HT13]
9	221	-11	-1	-6	4	28	-4	3	-2	9	-9	19	96	-
12	195	-2	2	-7	5	1	-10	1	-2	4	-8	44	86	Endo/exonuclease/phosphatase domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HJ12]
2	569	-10	-10	-4	5	17	0	-2	0	3	-6	20	77	RT_RNaseH domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HT13]
11	204	0	-1	-6	6	11	-7	20	7	-9	-3	2	72	Endo/exonuclease/phosphatase domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1I2R8]
6	342	1	-4	0	4	15	-2	19	12	-2	-3	-8	70	Integrase_H2C2 domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HXS5]
37	96	-1	-1	0	-3	14	0	-3	-1	0	2	39	64	DUF1758 domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HY80]
26	119	-5	-5	0	0	6	-1	26	0	0	0	12	55	RT_RNaseH domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HT13]
82	57	12	0	14	-1	-1	-1	-1	-1	1	0	23	55	-
16	152	0	-6	2	3	10	-7	-2	0	3	-5	12	50	cytochrome P450 (CYP6J)
61	66	-5	0	-2	-4	25	0	-4	-4	3	0	1	48	DUF1758 domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HY80]
38	90	-7	3	-2	0	7	-7	0	-5	14	0	2	47	-
48	77	-2	0	0	-9	11	-4	3	0	-3	3	12	47	-
39	88	8	-1	-1	0	0	-7	9	-1	0	0	18	45	Endo/exonuclease/phosphatase domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HKU3]

44	81	-3	-5	0	0	2	0	-1	-2	29	-1	1	44	FAT domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HN14]
58	68	5	0	-1	-1	-1	17	-2	-2	1	0	14	44	ATP-dependent DNA helicase (Fragment) (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1I3A7]
22	128	13	-2	0	5	0	0	7	7	-1	-5	0	40	-
14	163	10	0	4	5	0	-4	-4	-5	6	0	0	38	-
217	35	8	9	-1	-2	-1	-1	-1	-2	9	2	-1	37	-
137	45	-3	0	-3	0	1	0	0	0	0	0	0	syl +	-
528	25	0	0	0	0	0	0	0	0	0	0	4	syl +	-
932	21	0	0	0	0	0	0	0	0	4	0	2	syl +	Pyruvate carboxylase (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HWN2]
1846	16	0	0	-1	0	0	1	1	1	0	1	0	syl +	CUB domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HWE2]
2050	15	0	0	0	-1	0	0	0	1	0	0	4	syl +	Dimer_Tnp_hAT domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1IBG8]
2177	15	0	0	0	0	0	0	0	0	0	0	8	syl +	-
150	42	2	1	0	1	0	0	0	0	-1	0	0	dom +	Sodium channel prot
215	35	1	0	1	2	0	0	0	0	0	0	0	dom +	Kinesin-like protein [Source:UniProtKB/TrEMBL;Acc:T1I731]
370	28	18	0	0	0	-2	0	0	0	0	-1	-1	dom +	DDE-1 domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HWD4]
628	24	1	0	0	0	0	-1	0	-1	0	-1	-1	dom +	Gamma-glutamyl kinase (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1I111]
659	24	1	0	0	0	-1	0	-1	0	-1	-1	0	dom +	Uncharacterized protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HFR6]
777	22	2	1	0	1	0	0	0	-1	0	0	0	dom +	unspecified product
1084	20	0	1	0	0	-1	0	-1	-1	0	0	-2	dom +	ABC transporter domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HF18]
7599	11	0	0	1	0	-1	-1	-1	-1	0	0	0	dom +	Maelstrom domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1H9A4]
7750	11	1	0	0	0	-1	-1	-1	0	-1	0	0	dom +	Uncharacterized protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HWF4]
8121	10	1	0	0	0	-1	-1	-1	-1	0	0	-1	dom +	GPI mannosyltransferase 1 (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HNQ1]
8731	10	0	1	0	0	-1	0	-1	0	-1	0	-1	dom +	Cysteine-rich PDZ-binding protein [Source:UniProtKB/TrEMBL;Acc:T1HZF2]
8819	10	8	0	0	0	-1	-1	-1	0	0	0	-1	dom +	Nitrophorin domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1H8M2]
9533	9	1	0	0	0	0	0	-1	-1	0	-1	-1	dom +	Thioredoxin domain-containing protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HEY8]
9932	8	0	0	1	0	0	0	-1	-1	0	-1	-1	dom +	Odorant receptor Or74 / Or75 / Uncharacterized protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1H9N0] / Or77 / Uncharacterized protein (Fragment) [Source:UniProtKB/TrEMBL;Acc:T1HD13]
10604	8	0	1	0	0	-1	-1	-1	0	0	-1	0	dom +	CUB domain-containing protein [Source:UniProtKB/TrEMBL;Acc:T1HQ54]

### GO enrichment analysis

Among the 32,718 ortholog groups, we identified 24,390 OGs that contain genes from sylvatic and domiciliary species, 2,576 OGs that contain genes only from sylvatic species (specific to the sylvatic “pangenome”) and 5,942 OGs specific to the domiciliary “pangenome” (Figure 4). Despite the under representation of domiciliary species, we thus detected 2.3 times more OGs exclusive to this pangenome. Using these sets of exclusive OGs, we identified GO terms significantly enriched (p-value < 0.05) in domiciliary or sylvatic species compared to the whole set of genes (Table 6). Eight GO terms were found enriched in domiciliary species including biological processes, cellular components and molecular functions with highly

significant results for three GO terms namely organonitrogen compound biosynthetic processes, electron transport chain, and external encapsulating structure. Only three GO terms were detected as enriched in sylvatic species.

**Table 6: Bilan analysis enrichissement GO**

		GO.ID	Term	Annotated	Significant	Expected	Pval
over-represented in domiciliary species	Biological processes	GO:1901566	organonitrogen compound biosynthetic process	77	9	4.89	0.0029
		GO:0022900	electron transport chain	11	4	0.7	0.0034
		GO:0009060	aerobic respiration	5	2	0.32	0.035
		GO:0070887	cellular response to chemical stimulus	5	2	0.32	0.035
	Cellular components	GO:0030312	external encapsulating structure	6	3	0.33	0.0029
		GO:0031975	envelope	24	4	1.32	0.0392
	Molecular Function	GO:0016779	nucleotidyltransferase activity	11	3	0.79	0.04
		GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	5	2	0.36	0.045
over-represented in sylvatic species	Biological processes	GO:0006813	potassium ion transport	9	2	0.2	0.015
	Cellular component	GO:0016592	mediator complex	10	2	0.27	0.028
	Molecular Function	GO:0003712	transcription coregulator activity	12	2	0,36	0,049

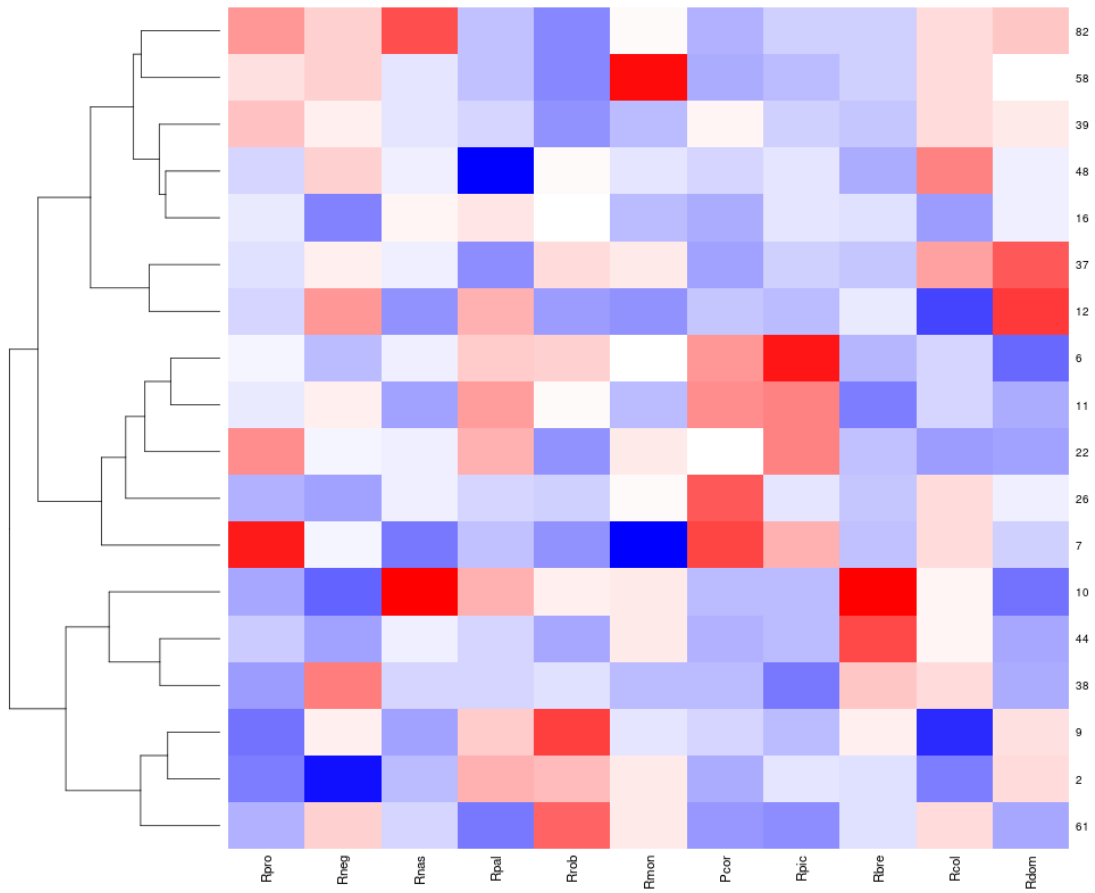
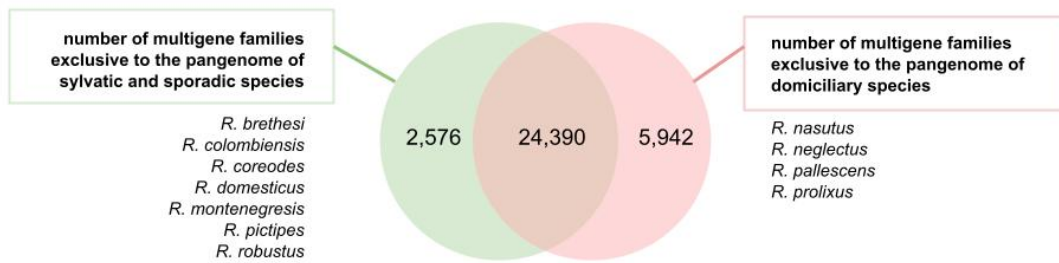


Figure 3: Heatmap



**Figure 4: Venn diagram**

## Discussion

Our study aimed to compare the dynamics of the genomes of species of the genus *Rhodnius* dependent on habitats. We questioned the link between the domiciliation of certain species and the evolution rate in terms of gene gains and losses. We then focused on key gene families specifically subjected to expansions or contractions that could allow functional diversification sustaining ecological diversification of the species. Since to date only the genome of the domiciliary species *R. prolixus* is described (Mesquita et al. 2015), we first sequenced, assembled and annotated the genome of eleven *Rhodnius* species including four domiciliary and seven sylvatic species and generate an ultrametric phylogenetic tree.

### *The domiciliation of the vectors, a recent history*

A central question lies in the dating of the domiciliation of the vectors of Chagas disease. According to discoveries in the northeast of Brazil (area of the Serra da Capivara National Park), the first prehistoric presences in caves and rock shelters containing numerous cave paintings date from 26,000 BP (Araújo et al. 2009). The authors suggest, based on current observations, that the first prehistoric artists were the first South Americans to be exposed to the transmission of the parasite by *T. brasiliensis*. But, the first human-parasite contact is dated to around 9,000 (BP or around 7,050 BC) in a Chinchorro mummy found in the Atacama Desert located between northern Chile and southern Peru and whose DNA analysis revealed the presence of the parasite *T. cruzi* (Aufderheide et al. 2004). The authors also show by studying other mummies that *T. cruzi* infections occurred in this region until the Spanish conquest in the 16th century, with a prevalence rate of 41% regardless of the period considered, which indicates that

Chagas disease was already widespread in pre-Columbian times. It is likely that the domiciliation of vectors only truly began with the domestication of small rodents (*Cavia* sp.) for consumption and rituals, around 8,000 to 6,000 years ago, and the storage of grains attracting grain-eating mammals. These food resources would have been attractive to the vectors. The earliest infection of the Chinchorro mummy therefore dates back several thousand years before the known time when rodents and camelids were domesticated (Rothhammer et al. 1985) should be considered as an incursion of humans in the sylvatic cycle of the parasite.

Without it being possible to really date these events, the growing anthropization of sylvatic environments has amplified the process of domiciliation. Anyway, according to our dated phylogenies, the different species are already differentiated long before their contact with human populations.

Indeed, the datings combining molecular data and fossils converge towards an origin of the Triatominae around 35 to 40 MYA (Ibarra-Cerdeña et al. 2014; Justí et al. 2016) and more precisely between 27 and 32 MYA for the most comprehensive study (Hwang and Weirauch 2012). According to our data, the last common ancestor of the *Rhodnius* species is estimated at 17 MYA, the divergence of the *prolixus* group around 11 MYA, the split between the *pallescens* and *pictipes* group around 15.3 MYA and the most recent speciation events are at least dated to 1 MYA.

Since speciation and domiciliation seem to be successive rather than concomitant, it is likely that species that colonized human domiciles exhibited some pre-adaptive traits. As genomic studies have shown that certain characteristics of the genomes such as great genetic diversity, the content of transposable elements, the diversification of gene families, allow rapid evolution (Pélissié et al. 2018), in this work, we analyzed some of these genomic characteristics within domiciliary and sylvatic species of *Rhodnius*.

#### *A variable genome size within the genus Rhodnius*

Using k-mer estimations, the mean genome size for the *Rhodnius* species was estimated at about  $468 \pm 41$  Mb. The species of the *pictipes* and *pallescens* groups have smaller genome size than those of the *prolixus* group. However, our data did not confirm that the size of the *R. colombiensis* genome is out of range of the values for the other *Rhodnius* species, as estimated by Díaz et al. (2014) with flow cytometry (FC) method. The genome size estimation for *R. prolixus* strain from Honduras used in this study is in agreement with the estimation for a Colombian *R. prolixus* strain (509.6 Mb) in a previous study (Merle et al. 2022). The k-mer estimations are smaller than those based on FC (Merle et al. 2022; Panzera et al. 2010) that



may be explained by the usage of moderate coverage sequencing data sets that fail to correctly represent the complexity of these genomes (Pflug et al. 2020).

We confirmed a significant variation of the genome size in the *prolixus* group since the genome of *R. robustus* male strain is 27% larger than that of the *R. coreodes* male strain. We also confirmed significant genome size variations between sexes for *R. robustus* (561 Mb for male strain in this study vs 440 Mb for female strain in Merle et al. 2022). In the *prolixus* group, our data demonstrated that the two species *R. coreoedes* and *R. domesticus* exhibited the smallest genome size, in the range of the one estimated for *R. neivai* in (Merle et al. 2022). As these three species are basal branches in the *prolixus* group phylogeny (Filée et al. 2021), a phylogenetic signal cannot be ruled out to explain the results. Moreover, the species of the other *Rhodnius* group, *pictipes* and *pallescens*, albeit our limited sampling, show similar values than the basal *prolixus* group species, that is in a parsimony point of view in favour of a small genome size for the common ancestor of the *Rhodnius* species. Under such hypothesis, a genome expansion would have taken place with the differentiation of the other species around 9-10 MYA, and could have played a major role in the widespread of the species in various ecological niches. An alternative hypothesis would be a genome reduction for the three basal species of the *prolixus* group preceded by a genome expansion in the common ancestor of this group. Such genome reduction would be correlated with the niche specialization of these three species. *Rhodnius domesticus*, the misnamed, is a rare sylvatic triatomine endemic to the Atlantic Forest associated with bromeliads, palms, marsupial and rodent nests and it occasionally invades human domiciles possibly when attracted by artificial light sources (Nascimento et al. 2021). *Rhodnius neivai* is also a rare sylvatic species in Colombia and Venezuela occurring in trunks of dead trees and on the palm tree *Copernitia tectorum*. It is also found occasionally in houses (Lent et Wygodzinsky 1979). *Rhodnius coreodes* occurring mainly in the Argentinean Chaco (Gurgel-Goncalves et Cuba 2009) is found in association with various bird species, especially the thornbird Furnariids (Carcavallo et al. 1998, Marti et al. 2014) and seems to use palms opportunistically (Abad-Franch et al. 2015). This species exhibits also specific morphological characters that had first justified the distinction of the genus *Psammolestes* before extensive molecular data undoubtedly anchor species of *Psammolestes* in the genus *Rhodnius* (Filée et al. 2021) as previously claimed by Hypša et al. (2002).

Changes in genome size must involve mutational mechanisms including transposable elements activity (Petrov 2001; Naville et al. 2019). However, in our data no correlation is detected between genome size and TE content as illustrated by the fact that both *R. coreodes* and *R. robustus*, that are respectively characterized by the smallest and largest genome size, have the highest TE contents.

### A faster evolution of some genes in *Rhodnius*

The models of genome evolution by loss and gain of genes developed in the present study show an evolution 2.4 times faster for the genus *Rhodnius* compared to other Hemiptera. In a more general study including different orders of insects and seven independent blood-sucking lineages (5 Diptera, 2 Hemiptera), hematophagous lineages were shown to have a slower rate of gain and losses of genes compared to non-hematophagous lineages (Freitas and Nery, 2020). These authors also showed that a small number of families comprised rapidly evolving genes in hematophagous species, in particular a gene family encoding heat shock proteins and another encoding chemosensory proteins. In the present study, among the most rapidly evolving gene families (i.e. with the greatest number of loss and gain events), we have identified genes linked to metabolic activities (proteolytic, catalytic, kinasic), detoxification (cytochrome P450), genes involved in metallic bonds, as well as transposable elements.

The faster evolution of certain metabolism-related gene families can be related to the hematophagous diet of the triatominae. Metal ions play an important role in all biological systems and seem to be involved in the fertility of female insects (Cardoso-Jaime et al. 2022) as in *R. prolixus* for genes involved in iron and heme fixation (Walter-Nuno et al. 2018). Metalloproteins are also thought to be involved in immunity (Cardoso-Jaime et al. 2022). The expansion of proteases with lytic activity could be an adaptation to the blood diet specifically rich in proteins as has been suggested for tsetse flies. The genomes of six species from these hematophagous diptera vectors of African trypanosomes have been explored and were shown to be specifically enriched in proteolytic genes (Attardo et al. 2019).

It is generally accepted that expansions or contractions of detoxification-related gene families are associated with functional adaptations related to diet and other environmental characteristics (Rane et al. 2016). A positive correlation between the number of detoxification genes in species and the complexity of their food sources has been reported (Rane et al. 2019) but it does not appear to be the case for aphids (Hemiptera; Aphididae) (Lin et al. 2022). Whether insects are phytophagous or hematophagous, they are exposed to various compounds as toxins from plants for the former or to the heme or iron toxicity at each blood-meal for the latter (Volonté et al., 2022). Insect cytochromes P450 are well known to be involved in the metabolic detoxification of xenobiotics including phytochemicals, insecticides and environmental pollutants (Lu et al. 2021) as previously demonstrated for triatomine bugs (Traverso et al 2017). However, a study by Volonté et al. (2022) with both phytophagous and hematophagous (*Cimex*, *R. prolixus*) Hemiptera, showed that the most

polyphagous species *Halyomorpha halys* exhibited the largest repertoire of detoxification enzymes. Regarding triatomines feeding on a wide range of vertebrate hosts, cytochromes could have a more general role related to diet.

#### *Differentiated genome dynamics in domiciliary species and expanding orthologous gene families*

In our study, the evolutionary rate of gene gain and loss, defined from groups of orthologs, is 7.5 times higher in domiciliary species than in sylvatic species. This dynamic concerns both events for domiciliary species, but with 2.3 times more gains than losses. Ratio of gene loss over gene gain is higher in sylvatic than in domiciliary species. Our comparative genomic study has allowed us to identify families of genes whose dynamics differ between domiciliary and sylvatic species. We annotated the ortholog groups that underwent frequent duplications specifically in domiciliary or sylvatic species and we identified 26 of them in the domiciliary species and 16 in the sylvatic ones. These groups therefore include candidate genes potentially involved in the domiciliation process.

Among the families of orthologs enriched in domiciliary species, we have annotated families linked in particular to transposable elements, to immunity, to resistance to insecticides, and to putative proteins involved in blood supply (nitrophorin) and chemosensory perception (odorant receptors). Among the families of orthologs enriched in sylvatic species, we have annotated families related in particular to transposable elements, muscle physiology and immunity.

It is interesting to note that genes linked to transposable elements are represented among the enriched ortholog families for both domiciliary and sylvatic species and that the analysis of TE content in the genomes showed a lower percentage for domiciliary species (8-10%) than for sylvatic species (11-20%) except for the domiciliary species *R. prolixus* (16%). The role of TEs in adaptive mechanisms is widely documented. One of the most publicized examples is that of the peppered moth (*Biston betularia*) whose black form, the frequency of which has increased in regions of England subject to industrial pollution, is linked to the insertion of the transposon carb-ET in the first intron of the cortex gene. This insertion greatly increases the amount of cortex proteins produced in the wing imaginal discs, resulting in the black coloration of the wings. The role of TEs in adaptation to insecticides has been demonstrated in *Anopheles mosquitoes* (Weedall et al. 2020). In *Anopheles coluzzii*, one of the main vectors of human malaria in sub-Saharan Africa, ET insertions have been identified near resistance or immunity genes (Vargas-Chavez et al. 2022). The domiciliary species being very strongly submitted to chemical insecticides, the resistance genes are

expected to be highly dynamic. The first vector control campaigns took place around the 1950s with the use of DDT, followed by the application of organochlorine insecticides such as lindane and dieldrin. The first findings of dieldrin resistance were reported from Venezuela in 1971 (González Valdivieso et al. 1971). Less persistent organophosphates and carbamates were then proposed as an alternative, and introduced in 1975. Then pyrethroids such as deltamethrin and cypermethrin were notably used in Argentina and Brazil (Mougabure-Cueto and Picollo 2015). However, in Bolivia, most domiciliary populations of *T. infestans* are resistant to deltamethrin (Lardeux et al. 2010). In this context, the high TE content of *R. prolixus*, a domiciliary species par excellence and highly subject to insecticides, could reflect a strong TE dynamic. Specific studies should be conducted to link the presence of TE and resistance to insecticides.

Among the expanding gene families in domiciliary species, we have annotated nitrophorins, which is a species-specific class of salivary lipocalins in the genus *Rhodnius*. In the *R. prolixus* reference genome, 14 nitrophorins have been identified (Mesquita et al. 2015). A fine study of these genes should be undertaken.

## Conclusion

In this study, by assembling and annotating 11 genomes of *Rhodnius* species vectors of Chagas disease, we demonstrated a different evolutionary rate of gain and loss per gene per million years between domiciliary and sylvatic species. This rate is 16-fold higher for the domiciliary Triatominae terminal branches than for the other branches demonstrating that the domiciliation of certain species was accompanied by rapid evolution associated with ecological diversification. We also provided an estimate of the last common ancestor of the *Rhodnius* genus which is 17 million years old.

The identification of gene families that evolve differently between domiciliary and sylvatic species could provide a better understanding of the mechanism of disease vector domiciliation. We have identified candidate gene families for domiciliation involving the feeding of bugs, the chemosensory reception (odorant receptors), but also their resistance to insecticides.

A better understanding of disease vector domiciliation could help in the development of new strategies to deal with blood-feeding vectors.

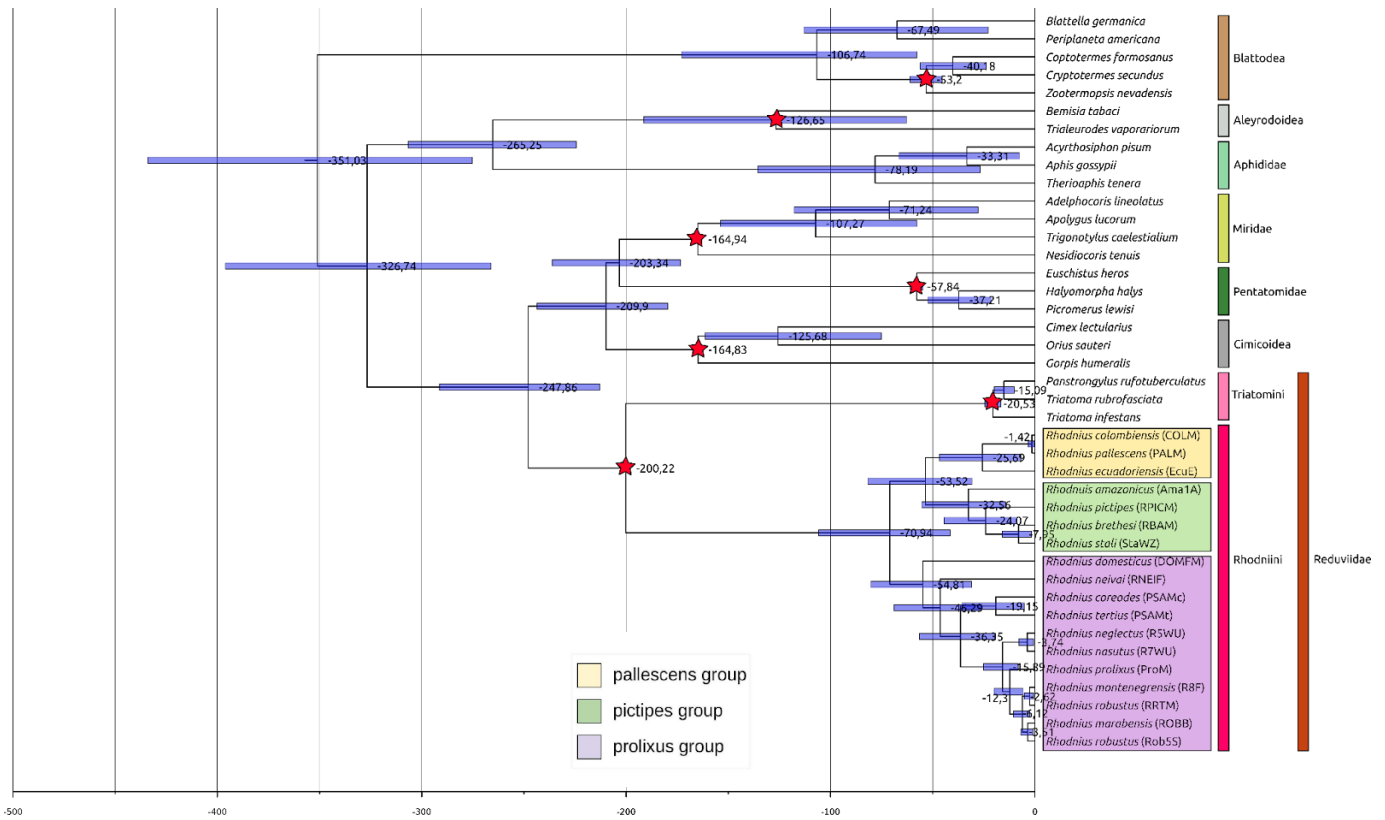
## Funding

This study was partially funded 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.

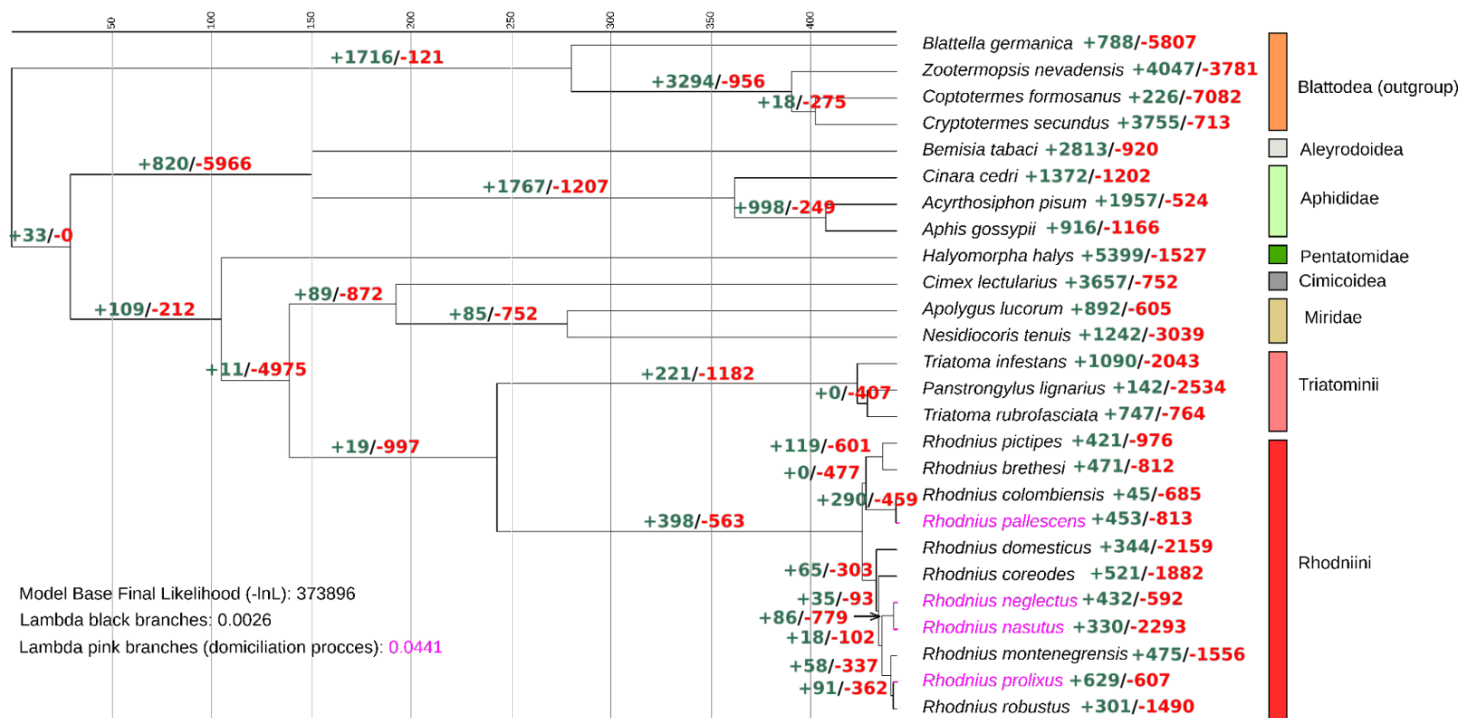
## Acknowledgements

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 Oliveira (Universidade Estadual Paulista (Unesp), Faculdade de Ciências Farmacêuticas, Araraquara, São Paulo, Brazil); and for *R. prolixus* strain from Honduras: Claudio Lazzari (Université de Tours, IRBI, France). We thank the GenOuest platform for the bioinformatics facility.

Supplementary data



**Supp Figure 1: Divergence time estimates based on BEAST analysis using relaxed-clock model and 7 fossil calibration points.** Placement of fossils as calibration points of clades indicated by red stars. Blues node bars indicate 95% highest posterior density (HPD) credibility intervals.



**Supp Figure 2: Gains and losses per gene per millions years.** Gene family expansion is indicated in red, and gene family contraction is indicated in green.

**Supp Table 1: list of all species and accession numbers for genomes and transcriptomes used.**

Taxa	Species	mitochondrial ID	genome ID
Aleyrodoidea	<i>Bemisia tabaci</i>	MH714535.1	GCF_001854935.1
Aleyrodoidea	<i>Trialeurodes vaporariorum</i>	AY521265.2	GCA_011764245.1
Aphididae	<i>Acyrtosiphon pisum</i>	NC_011594.1	GCF_005508785.1
Aphididae	<i>Aphis gossypii</i>	MK994521.1	GCF_004010815.1
Aphididae	<i>Cinara cedri</i>	-	GCA_902439185.1
Aphididae	<i>Therioaphis tenera</i>	MH643885.1	-
Blattodea	<i>Blattella germanica</i>	EU854321.1	GCA_003018175.1
Blattodea	<i>Coptotermes formosanus</i>	NC_015800.1	GCA_013340265.1
Blattodea	<i>Cryptotermes secundus</i>	KP026283.1	GCF_002891405.2
Blattodea	<i>Periplaneta americana</i>	GU947663.1	GCA_002939525.1
Blattodea	<i>Zootermopsis nevadensis</i>	KJ958410.1	GCF_000696155.1
Cimicoidea	<i>Cimex lectularius</i>	JQ739180.1	GCF_000648675.2
Cimicoidea	<i>Orius laevigatus</i>	-	GCA_018703685.1

Cimicoidea	<i>Orius sauteri</i>	KJ671626.1	-
Cimicoidea	<i>Gorpis humeralis</i>	JF927830.1	-
Miridae	<i>Apolygus lucorum</i>	KU234539.1	GCA_009739505.2
Miridae	<i>Cyrtorhinus lividipennis</i>	-	GCA_019603395.1
Miridae	<i>Nesidiocoris tenuis</i>	JQ806057.1	GCA_902806785.1
Miridae	<i>Pachypeltis micranthus</i>	-	GCA_020466155.1
Miridae	<i>Trigonotylus caelestialium</i>	KJ170899.1	-
Miridae	<i>Adelphocoris lineolatus</i>	KJ020286.1	-
Pentatomidae	<i>Aelia acuminata</i>	OU426986.1	GCA_911387785.1
Pentatomidae	<i>Euschistus heros</i>	MG253270.1	GCA_003667255.1
Pentatomidae	<i>Halyomorpha halys</i>	FJ685650.1	GCF_000696795.2
Pentatomidae	<i>Stiretrus anchorago</i>	-	GCA_010014745.1
Pentatomidae	<i>Picromerus lewisi</i>	MW355499.1	-
Triatomini	<i>Panstrongylus lignarius</i>	-	PRJNA395235.fa
Triatomini	<i>Panstrongylus rufotuberculatus</i>	MF614953.1	-
Triatomini	<i>Triatoma brasiliensis</i>	-	-
Triatomini	<i>Triatoma infestans</i>	KY640305.1	GCA_011037195.1
Triatomini	<i>Triatoma rubrofasciata</i>	MH934953.1	Triatoma_chr_assembly

**Supp Table 2: List of fossils, dates, refs**

	min	max	mean	stdev	fossils genus	author	age (MYA)	site name
Cimicoidea	166.1	163.5	164.8	1.3	<i>Pumilanthocoris gracilis</i>	Hou et al. (2012)	166.1 – 163.5	Daohugou village (China)
					<i>Pumilanthocoris obesus</i>	Hou et al. (2012)	166.1 – 163.5	Daohugou village (China)
Aleyrodoid ea+Aphidi dae	298.9	252.2	275.6	23.4	<i>Lutevanaphis permiana</i>	Szwedo et al. (2014)	256 – 256	Lodeve (France)
Miridae	166	164	165	1	<i>Mirivena robusta</i>	Yao et al. (2007)	166.1 – 163.5	Daohugou village (China)
Pentatomi dae	59.2	56	57.6	1.6	<i>Asopus puncticollis</i>	Piton (1940)	59.2 – 56	Menat (France)
					<i>Doryderes laticollis</i>	Piton (1940)	59.2 – 56	Menat (France)
Reduviidae	201.3	199.3	200.3	1	<i>Ceresopsis ornata</i>	Becker-Migdisova. (1962)	201.3 – 199.3	Sogyuty (Kyrgyzstan)
					<i>Ceresopsis vitrea</i>	Becker-Migdisova (1962)	201.3 – 199.3	Ak-Bulak-Say, Sogjuta, (Kyrgyzstan)



Termitidae	56	47.8	51.9	4.1	<i>Nanotermes isaacae</i>	Engel et al. (2011)	56 – 47.8	Tadkeshwar lignite mine (India)
Triatomini	20.4	16	18.2	2.2	<i>Panstrongylus hispaniolae</i>	Poinar et al. (2013)	20.44 – 15.97	La Toca mine (Dominican Republic)
					<i>Triatoma dominicana</i>	Poinar et al. (2005)	20.44 – 15.97	Dominican Republic

**Supp Table 3: Genome completeness estimated with BUSCOs.** The gene set used is insecta\_odb10

	Complete and single-copy BUSCOs (S)	Complete and duplicated BUSCOs (D)	Fragmented BUSCOs (F)	Missing BUSCOs (M)
<i>R. brethesi</i>	88.4	1	3.7	6.9
<i>R. colombiensis</i>	95.1	1	1.2	2.7
<i>R. coreodes</i>	93.8	2.9	0.7	2.6
<i>R. domesticus</i>	94.7	1.1	1.2	3
<i>R. montenegrensis</i>	79.6	0.8	4.3	15.3
<i>R. nasutus</i>	61.4	0.7	5.5	32.4
<i>R. neglectus</i>	86.6	1.6	2.8	9
<i>R. pallescens</i>	91.2	1.1	2.6	5.1
<i>R. pictipes</i>	86	1.4	4.2	8.4
<i>R. prolixus</i>	88	1.5	3.7	6.8
<i>R. robustus</i>	89.8	1.4	1.8	7

**Supp Table 4: Correlation metrics**

Comparaison	coeff_Spearman	Padj Bonferroni
Total.length versus contigs	-0.609	0.67
Total.length versus N50	0.873	0.01
Total.length versus Busco	0.764	0.12
Total.length versus ET	0.855	0.02
Total.length versus CDS	0.400	1.00
Total.length versus LR	0.775	0.07
Total.length versus SR	0.456	1.00

N50 versus contigs	-0.800	0.07
N50 versus Busco	0.791	0.08
N50 versus ET	0.700	0.27
N50 versus CDS	0.264	1.00
N50 versus LR	0.775	0.07
N50 versus SR	0.187	1.00

**Supp Table 5: Gene loss and gain in *Rhodnius* species in relation to habitat (yellow: domiciliary species).** The ratio is related to branch lengths (in millions of years).

	branch length (millions years)	number of gains events on the terminal branch	number of gains events on the terminal branch per million year	number of loss events on the terminal branch	number of loss events on the terminal branch per million year	
<i>R. nasutus</i>	1.64839	420	255	2978	1806.61	
<i>R. neglectus</i>	1.64839	590	358	692	419.80	
<i>R. pallescens</i>	0.35564	586	1648	809	2274.76	
<i>R. prolixus</i>	1.75082	771	440	877	500.91	
<i>R. brethesi</i>	6.93032	610	88	1032	148.91	
<i>R. colombiensis</i>	0.35564	315	886	798	2243.83	
<i>R. coreodes</i>	9.04558	657	73	2329	257.47	
<i>R. domesticus</i>	10.25418	471	46	2623	255.80	
<i>R. pictipes</i>	6.93032	542	78	1160	167.38	
<i>R. montenegrensis</i>	3.00321	590	196	2090	695.92	
<i>R. robustus</i>	1.75082	402	230	1753	1001.25	
						loss/gains
<b>sylvatic species mean</b>		512.43	228.08	1683.57	681.51	3.29
<b>domiciliary species (without <i>R. nasutus</i>) mean</b>		649.00	815	792.67	1065.16	1.22
<b>dom/syl (without <i>R. nasutus</i>)</b>		1.27	3.57	0.47	1.56	2.29

## POSTAMBULE SUR L'ARTICLE EN COURS D'ÉCRITURE

### Annotation et assemblage des génomes

L'assemblage de 11 génomes de *Rhodnius* a permis de mettre en lumière des familles de gènes impliquées dans l'adaptation de certaines de *Rhodnius* aux anthroposystèmes. Ce jeu de données ouvre la voie à de nombreuses études sur ces génomes comme des études de génomique comparative sur des familles spécifiques. Cette analyse requiert une annotation plus fine des gènes basée sur l'homologie de séquence avec des gènes déjà annotés chez des espèces proches.

L'annotation des génomes de *R. coreodes* et de *R. domesticus* avec des transcriptomes d'espèces proches a permis de les inclure dans cette analyse malgré le manque de données RNAseq sur ces espèces.

### Recherche sur les gènes candidats

Il serait intéressant d'étudier plus précisément les familles de gènes candidats identifiées dans cette étude et notamment les gènes impliqués dans l'alimentation et la chimioréception.



## CHAPITRE 4 : GÈNES CHIMIOSENSORIELS

---

### CONTEXTE ET PROBLÉMATIQUE DE L'ÉTUDE

#### Contexte

Les gènes chimiosensoriels sont annotés chez quelques hémiptères dont *R. prolixus*. Dans l'article de [Mesquita et al. 2015](#) ont été annotés 28 gènes GRs, codant 30 protéines (2 gènes ont des transcrits alternatifs), et 116 gènes ORs (ORco inclus), dont 5 pseudogènes codant 111 ORs ([Latorre-Estivalis et al. 2016](#)).

#### Objectif

L'évolution des gènes codant des protéines du système chimiosensoriel, par leur évolution rapide et leur rôle dans le comportement de l'insecte (recherche de nourriture, recherche de partenaires sexuels, recherche de lieu de ponte, acceptation ou non de l'hôte), pourrait être en lien avec la capacité de certaines punaises du genre *Rhodnius* à se domicilier.

Plus spécifiquement, l'objectif est de mettre en relation l'évolution moléculaire des gènes chimiosensoriels codant les récepteurs gustatifs et olfactifs avec l'adaptation des vecteurs à leur environnement notamment domiciliaire. Ceci pourrait permettre d'envisager des moyens alternatifs de lutte basés sur le système chimiosensoriel (brouillage sensoriel, pièges sélectifs).

L'étude a été faite au niveau interspécifique. Nous avons focalisé notre étude sur les récepteurs chimiosensoriels gustatifs et olfactifs. Les différentes séquences de GRs et ORs sont extraites à partir de données génomiques de 13 espèces de *Rhodnius* (4 domiciliaires, 9 sylvatiques). Une phylogénie des GRs et des ORs des espèces de *Rhodnius* étudiées a été réalisée comprenant les gènes déjà décrits chez *R. prolixus*. Cette phylogénie permet de formuler des hypothèses fonctionnelles pour certains de ces gènes par homologie avec des GRs et des ORs dont la fonction est connue chez d'autres espèces d'insectes. L'étude de la dynamique des gènes a été faite par la recherche de perte de gènes ou de duplication chez les espèces ayant la capacité à se domicilier par rapport aux espèces sylvatiques. Une recherche de sélection positive sur les gènes codant les GRs et ORs a également été conduite.

## PRINCIPAUX RÉSULTATS

Les récepteurs chimiosensoriels gustatifs et olfactifs ont été annotés dans le génome de 13 espèces. Deux génomes ont été séquencés et assemblés dans la présente étude. Des phylogénies ont été obtenues pour ces deux familles multigéniques et ont permis d'identifier de nombreux clusters de séquences orthologues. Les orthologues sont facilement identifiés pour les GR où un nouveau récepteur (GR29) a été décrit. Une duplication récente a été mise en évidence pour un GR (GR24). Pour les ORs certains clades montrent des relations phylogénétiques complexes résultant de multiples événements de duplication chez différentes espèces.

Sur la base d'une phylogénie calibrée, le taux de gain et de perte de gènes a pu être estimé : il est 2 fois plus élevé pour les ORs par rapport aux GRs mais un taux accéléré a été observé dans les branches conduisant aux espèces domiciliées uniquement pour les GRs. Des tests de sélection positive ont été effectués sur les grappes d'orthologie bien définies de chaque famille. Deux GRs et neuf ORs montrent des sites sous sélection positive quel que soit l'habitat des espèces. Une sélection positive a également été détectée entre les espèces domiciliées et les autres (rapport dN/dS différent) pour 4 OR supplémentaires (OR1, OR38, OR102 et ORco). Une sélection purificatrice a été détectée pour GR1, un récepteur homologue au récepteur du fructose chez la drosophile, suggérant que le sucre est également essentiel pour les espèces hématophages du genre *Rhodnius*.

## ARTICLE (EN PRÉPARATION) : EVOLUTION OF CHEMOSENSORY RECEPTOR GENES IN 13 CHAGAS DISEASE VECTORS OF THE GENUS *RHODNIUS*

Marie Merle<sup>1</sup>, Florence Mougel<sup>1</sup>, Jonathan Filée<sup>1</sup>, Louise Maille<sup>1</sup>, Elaine Folly-Ramos<sup>2</sup>, Carlos Eduardo Almeida<sup>3</sup> and Myriam Harry<sup>1\*</sup>

(1) CNRS, IRD, UMR EGCE, Evolution, Génomes, Comportement et Ecologie, Université Paris-Saclay, Gif-sur-Yvette, France

(2) Departamento de Engenharia e Meio Ambiente – DEMA, Universidade Federal da Paraíba, João Pessoa, Brazil

(3) Universidade Federal do Rio de Janeiro (UFRJ), Centro de Ciências da Saúde, Instituto de Biologia, Departamento de Zoologia, Rio de Janeiro, Brazil

### Abstract

Domiciliation process has been described among *Rhodnius* species, hematophagous bugs vector of Chagas disease. For 13 species, gustatory and odorant chemosensory receptors were annotated in the genomes, 2 of which being sequenced and assembled in the present study. Phylogenies were obtained for these two multigene families and allowed to identify numerous clusters of orthologous sequences. Orthologs are easily identified for GR where a new receptor (GR29) was described. A recent duplication was evidenced for one GR (GR24). The picture is more complicated for the larger OR family where some clades show complex phylogenetic relationships resulting from multiple duplication events in different species. Based on a calibrated phylogeny, rate of gain and loss genes could be estimated for the two families: it is 2-fold higher for OR genes compared to GR ones but only for the GRs an accelerated rate was observed in the branches leading to domiciliary species. Tests for positive selection were performed on the well-defined clusters of orthology from each family. Whatever the habitat, two GR and nine OR have few sites under positive selection. Positive selection was also detected between domiciliary species and others (different dN/dS ratio) for 4 supplementary OR (OR1, OR38, OR102 and ORco). Purifying selection was also detected for GR1, a receptor homologous to fructose receptor in *Drosophila*, suggesting that sugar is also essential for hematophagous *Rhodnius*.

### Introduction

The species of the *Rhodnius* genus are blood-sucking bugs (Hemiptera, Reduviidae, Triatominae) and vectors of the parasite *Trypanosoma cruzi* (Kinetoplastea, Trypanosomatidae), the causative agent of Chagas disease.

The *Rhodnius* genus includes 24 species since the three species of *Psammolestes* genus could be considered as *Rhodnius* species based on phylogenomics (Filée et al. 2022). In the wild, most of the species of the genus are dependent on palm trees but a few rare species are palm opportunists or are also dependent on bromeliads (Abad-Franch et al. 2005) and feed on a wide range of hosts (Ribeiro et al. 2014). Due to anthropogenic pressure disturbing or destroying natural habitats, some species considered as domiciliary species versus sylvatic species have colonized human dwellings and can feed on humans and transmit them the parasite via the infected feces. Whatever the mode of life of organisms (domiciliar or sylvatic), their survival, reproduction and social communication rely on their chemosensory system. Indeed, the detection and the 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 changing hosts or habitats (Benton et al. 2006). Three main families of receptors: olfactory receptors (OR), taste receptors (GR), and ionotropic receptors (IR) are part of the chemoreceptor proteins encoding complex olfactory information. OR recognize volatile compounds, GRs detect non-volatile chemicals while olfactory antennal IRs are associated with the detection of mainly acids and amines (Mitchell et al. 2020). Insect chemoreceptors are seven transmembrane proteins expressed on the surface of chemosensory neurons housed in hair-like structures called sensilla (Hansson et Stensmyr 2011). The genes encoding these proteins belong to multigene families, each family deriving from numerous duplication events of an ancestral gene. Depending on the rate of gain (duplication and functional specialization) and loss (pseudogenization and functional depletion), the size of the repertoires of multigene families is highly variable among species. Moreover, it is supposed to vary according to specialization, generalist species being submitted to a wide diversity of odorant signals showing larger chemosensory gene families than specialist ones. Considering GRs and ORs this was observed between the two sister species *Drosophila sechellia* recently specialized on *Morinda citrifolia*, and *D. simulans*, a generalist species: the loss rate of ORs and GRs is 10-fold higher *D. sechellia* compared to *D. simulans* (McBride 2007). In *Tessarotoma papillosa* a phytophagous Hemiptera consuming few host plants, 59 ORs were identified (Wu et al. 2019), a repertoire 2-fold smaller than the one of the generalist species *Apolygus lucorum*. A GR duplication was also proposed to be at the origin of an host range extension in the butterfly *Vanessa cardui* (Suzuki et al. 2018). Within moth species, an expansion of a specific receptor class devoted to bitter detection allows the perception of secondary metabolites emitted by plants (Wanner et Robertson 2008). By contrast, the specialist wasp *Ceratosolen solmsi*, obligate mutualist of the fig *Ficus hispida*, has restricted chemosensory repertoires (Xiao et al. 2013).

Considering hematophagous bugs, the OR and GR chemosensory gene families are annotated from genome for the domiciliary species, *R.*



*prolixus* with 116 ORs and 28 GRs (Mesquita et al. 2015) and *Cimex lectularius* highly specialised to human host but with 35 GRs but only 50 ORs described from the transcriptome (Hansen et al. 2014). By comparative protein analyses *Rhodnius prolixus*-specific expansions of chemosensory genes were identified, probably linked to the hematophagous diet (Mesquita et al. 2015).

Beside environmental adaptation via gene duplication/loss events of chemosensory genes, variation in gene expression could drive fast evolution. For Chagas disease vectors, chemosensory genes were under-expressed in the chemosensory organs of domiciliary populations compared to sylvatic populations of *Triatoma brasiliensis* (Marchant et al. 2016). Moreover, some chemosensory genes were specifically expressed in the sylvatic species *R. robustus* compared to the domiciliary *R. prolixus* (Marchant et al. 2021).

The acquisition of a novel function could be linked to specific changes in the active sites of proteins and then maintained by selection. Odorant and gustatory receptor genes show high levels of differentiation among host races in the pea aphid *Acyrtosiphon pisum* (Smadja et al. 2012). Such selection was also highlighted in chemosensory receptors of *Drosophila* species as *Drosophila suzukii* that feed on fresh fruits, while the related species *D. biarmipes* and *D. takahashii* develop on fermenting fruits (Hickner et al. 2016), or of *D. yakuba mayottensis* that feeds on the toxic fruit *Morinda citrifolia* unlike its related species *D. yakuba yakuba* (Ferreira et al. 2021).

The aim of this study was to infer the evolution of the repertoire of olfactory and gustatory receptors in the *Rhodnius* genus using 13 *Rhodnius* species including sylvatic and domiciliary species in order to document the domiciliation process.

Ortholog groups were identified with a phylogenetic approach using receptors described in the genome of *Rhodnius prolixus* as a reference (Mesquita et al. 2015). Phylogeny at the gene level was complemented by the estimation of expansion and contraction rate within the OR and GR multigene families and search for selection at the level of ortholog groups, according to the habitat lifestyle.

## Materials and Methods

### Sampling

Thirteen species were used in this study (Table 1). For 10 species, the assembled genomes were already available (Merle et al. in prep). Two species, *R. milesi* and *R. neivai* were sequenced in long reads (15x and 17 x

respectively with MinION technology, performed with GridION GXB02039) and short reads (56x and 72x respectively with Illumina technology , performed with NovaSeq6000-A00318, 150 bp paired-end). All sequencing analyses have been carried out by the Genotoul platform (Toulouse, France). Genome assembly was performed with the MaSuRCA v.4 program (Zimin et al. 2013) using the modified version of the CABOG assembler (Miller et al. 2008).

**Table 1: Sample used** (in green : non-domiciliary species, in red domiciliary species). For strains, the reference and the year of the strain foundation are given, with CTA: Colônias de Triatominae, Araraquara, Brazil.

species	origin	Field/Strain	M/F	sample name	Genome assembly
<i>R. brethesi</i> Matta, 1919	Igarapé Tucunaré, Brazil	S (CTA 222, 2009)	M	BAM	Merle et al. in prep
<i>R. colombiensis</i> Mejia et al. 1999	Tolima, Colombia	S (CTA 050, 2001)	M	COLM	Merle et al. in prep
<i>R. coreodes</i> Bergroth, 1911	Mato Grosso do Sul, Brazil	F	M	PSAMcor	Merle et al. in prep
<i>R. domesticus</i> Neiva and Pinto, 1923	Santa Catarina, Brazil	F	M and F	DOMFM	Merle et al. in prep
<i>R. milesi</i> Carcavallo et al. 2001	Brazil	S	M	MM	This study
<i>R. montenegrensis</i> da Rosa et al. 2012	Monte Negro, Rondônia, Brazil	S (CTA 087, 2003)	F	monR8	Merle et al. in prep
<i>R. nasutus</i> Stal, 1959	Brazil	S (CTA 054, 1999)	F	nasR7	Merle et al. in prep
<i>R. neglectus</i> Lent, 1954	Frutal, Minas Gerais, Brazil	S (CTA 061, 1983)	F	negR5	Merle et al. in prep
<i>R. neivai</i> Lent, 1953	Maracay, Venezuela	S	F	NEIF	This study
<i>R. pallescens</i> Lent, 1953	Colombia	S (CTA s/R)	M	PALM	Merle et al. in prep
<i>R. pictipes</i> Stal, 1972	Belém, PA, Brazil	S (CTA 072, 1998)	M	PICM	Merle et al. in prep
<i>R. prolixus</i> Stal, 1859	Honduras	S (IRBI laboratory)	F	ProH2	Merle et al. in prep
<i>R. robustus</i> Larousse, 1927	Rondônia, Brazil	S (CTA 272)	M	RTM	Merle et al. in prep

*OR and GR gene annotation*

The annotation was based on ORs from 14 hemipteran species and GR from 5 insect species already described including OR and GR already annotated in *R. prolixus* genome (see references [Table 2](#)). We performed an exonerate search in genomes assembled with MaSuRCA with the maximum intron length option set to 3,000.

Due to the dynamic of loss and gain of chemosensory receptor genes, especially the ORs with repeated tandem distribution, exonerate outputs need to be reviewed cautiously. In this way, to accurately and reproducibly process both the OR and GR results, we used the InsectOR webserver ([Karpe et al. 2021](#)) with default parameters (Alignment Cluster Cutoff 1, Completion Cutoff 300).

**Table 2: Odorant and gustatory receptors repertoires from the 14 hemipteran species used to annotate the genomic assemblies of *Rhodnius* species** (transcriptomic data \*)

Order: Family	species	# ORs	# GRs	Authors
Hemiptera: Aphididae	<i>Acyrtosiphon pisum</i>	79	77	<a href="#">Smadja et al. 2009</a>
Hemiptera: Berytidae	<i>Yemma signatus</i> *	66	-	<a href="#">Song et al. 2020</a>
Hemiptera: Cimicidae	<i>Cimex lectularius</i> *	50	35	<a href="#">Hansen et al. 2014</a>
Hemiptera: Gerridae	<i>Gerris buenoi</i>	155	-	<a href="#">Armisen et al. 2018</a>
Hemiptera: Lygaeidae	<i>Nysius ericae</i> *	83	-	<a href="#">Zhang et al. 2016</a>
Hemiptera: Lygaeidae	<i>Oncopeltus fasciatus</i> *	122	115/169	<a href="#">Panfilio et al. 2019</a>
Hemiptera: Miridae	<i>Cyrtorhinus lividipennis</i>	15	-	<a href="#">Wang et al. 2018</a>
Hemiptera: Miridae	<i>Adelphocoris lineolatus</i> *	88	-	<a href="#">Xiao et al. 2016</a>
Hemiptera: Miridae	<i>Apolygus lucorum</i> *	110	-	<a href="#">An et al. 2016</a>
Hemiptera: Pentatomidae	<i>Halyomorpha halys</i> *	138	-	<a href="#">Sun et al. 2020</a>
Hemiptera: Pseudococcidae	<i>Phenacoccus solenopsis</i> *	4	-	<a href="#">Nie et al. 2017</a>
Hemiptera: Psyllidae	<i>Cacopsylla chinensis</i> *	7	-	<a href="#">Xu et al. 2019</a>
Hemiptera: Psyllidae	<i>Diaphorina citri</i>	46	-	<a href="#">Wu et al. 2016</a>
Hemiptera: Reduviidae	<i>Rhodnius prolixus</i>	111	28/30	<a href="#">Mesquita et al. 2015</a>
Phthiraptera: Pediculidae	<i>Pediculus humanus</i>	-	5	NCBI

### Gene phylogenies

Amino acid sequence alignments were performed using MAFFT (Kato et al. 2019) for both GR and OR datasets. We used the IQTREE software (Nguyen et al. 2015) to build trees with 1,000 ultrafast bootstraps (Minh et al. 2013). The GR tree was rooted at midpoint and the ORco sequences were used to root the OR tree. Trees were viewed and edited using iTOL (Letunic et Bork 2019).

### Estimating gene gain and loss along the species phylogeny

Using the CAFE5 program (Mendes et al. 2020), and based on a calibrated tree (Merle et al. in prep), we calculated the gain and loss per gene per million years, called lambda. This analysis takes into account differences in annotations due to genome completeness and the chemosensory gene annotation thanks to the error model. The estimation relies on the assumption that gains are as frequent as losses.

We tested two models: the null model where the gain and loss rate is the same for the whole tree, and a two-rates model assuming that lambda varies between domiciliary *Rhodnius* (*R. nasutus*, *R. neglectus*, *R. pallescens* and *R. prolixus*) and the sylvatic *Rhodnius* species.

### Molecular evolution analysis

The CodeML package implemented in the PAML software (Yang 2007) was used to compare the ratio of synonymous mutations to non-synonymous mutations in the datasets. The software normalizes rates of synonymous and non-synonymous mutations with the respective number of synonymous and non-synonymous sites and calculates the rate of evolution. The model takes into account the phylogenetic signal, i.e. homology not due to convergence but inherited from a close common ancestor. The graphical interface of this software, EasyCodeML (Gao et al. 2019), was used to perform the statistical tests. EasyCodeML allows to perform likelihood ratio tests and to determine whether the model under selection is significantly more likely than the null model, i. e. the same ratio for the whole tree. The “branch” model was tested labeling the branches according to the habitat trait, either “domiciliary” or “sylvatic” (Table 1).

Ortholog groups were defined from the phylogenetic tree. Unambiguous groups (only one copy for each species or exceptionally two copies in a single species) were obtained for 42 ORs and 29 GRs allowing the search for positive selection. The other ortholog groups included too few sequences to be analyzed, paralogs derived from expansion species specific or sequences too divergent to obtain useful alignments. When a gene was duplicated in a given species, two positive selection tests were performed for each copy of the species.

One comparison test branch-model (variation between “domiciliary” and “sylvatic” branches). Four site-model comparisons were tested. The first comparison (M0 vs M3) compares a model with a single ratio (M0) with a model with three different ratios along sites (M3). The three other comparisons (M1a vs M2a, M7 vs M8, M8a vs M8) allow to detect positive selection that is taken into account in the second model (M2a, M8) but not in the first one (M1a, M7 or M8a). Beside positive selection, the compared models simulate evolution similarly: M7, M8 and M8a assume different rates of evolution between transition and transversion while M1a and M2a do not. Finally, M7 assumes a single unfixed ratio while M8a assumes this ratio is fixed to 1. Among the three comparisons devoted to the detection of positive selection, M1a vs M2a is the most stringent one. The Benjamini-Hochberg correction (Benjamini et Hochberg 1995) was applied because of the multiplicity of tests and adjusted P-values were considered as significant below 5%.

## Results

### *ORs repertory and Evolution in Rhodnius*

#### *ORs Annotation*

The number of ORs found in our assemblies varies considerably (Table 3). The total number of ORs ranges from 94 for *R. pallescens* to 182 for *R. robustus* and the number of normal complete with start codon genes varies from 84 for *R. neivai* to 24 for *R. pallescens*. The proportion of genes found as partial is also very variable, ranging from 62 for *R. robustus* to 27 for *R. colombiensis*, *R. coreodes* and *R. neivai*. We found between 3 pseudogenes in *R. pallescens* and 22 in *R. coreodes*. A significant number of the OR genes detected include the start codon: from 25 for *R. pallescens* to 99 for *R. robustus* up to 63% in *R. milesi*. This result is expected as many of the queries used, for example *R. prolixus* ORs sequences described by Mesquita et al. (2015) are not complete and lack methionine start codon. Indeed, of the 116 ORs described for *R. prolixus* only 84 are complete with a start

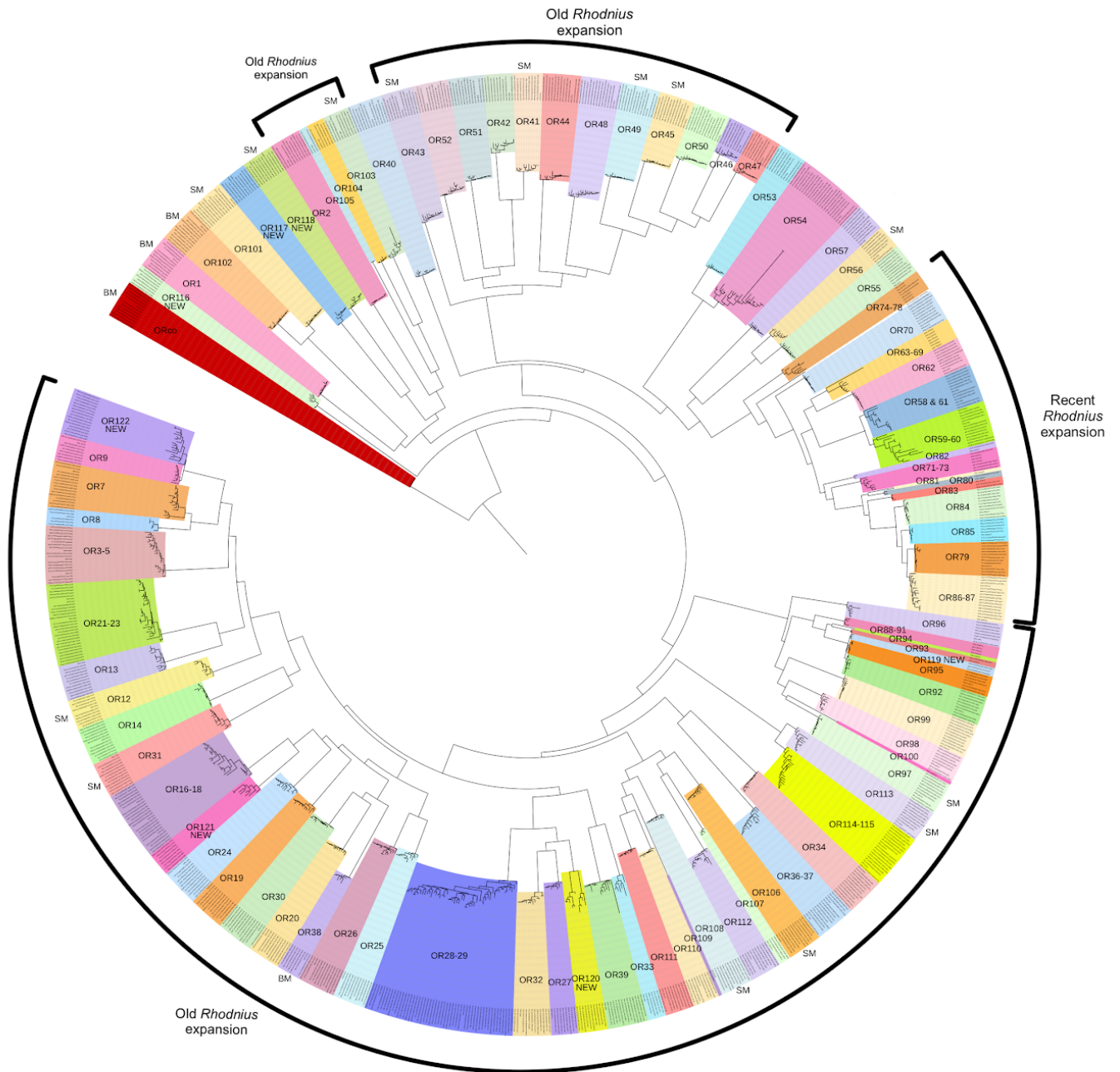
codon. Considering the total number of normal (non-pseudogenes) genes, the two domiciliary species *R. nasutus* and *R. pallelescens* exhibit very low values (<35 ORs), and among sylvatic species, *R. coreodes* and *R. brethesi* have the lower values (115, 121 ORs respectively) and *R. neivai* and *R. robustus*, the higher (>150 ORs). Duplication events are observed for 13 ORs and for 9 ORs concerned only sylvatic species (Supp Table 1).

**Table 3: Results of ORs annotation obtained with InsectOR.** Domiciliary species are colored in red

	Total number of Complete genes (including pseudogene)	Total number of Partial genes (including pseudogene)	Total number of Normal genes	Total number of Pseudo genes	Total number of genes with start codon present	Total number of genes	Complete and not pseudogene and start codon present
<i>R. nasutus</i>	59	39	92	6	44	98	34
<i>R. neglectus</i>	101	41	136	6	87	142	74
<i>R. pallelescens</i>	49	45	91	3	25	94	24
<i>R. prolixus</i>	99	47	131	15	80	146	69
<i>R. brethesi</i>	98	29	121	6	79	127	66
<i>R. colombiensis</i>	114	27	132	9	85	141	78
<i>R. coreodes</i>	110	27	115	22	82	137	70
<i>R. domesticus</i>	105	42	132	15	87	147	72
<i>R. milesi</i>	121	31	138	14	95	152	82
<i>R. montenegrensis</i>	102	48	135	15	80	150	70
<i>R. neivai</i>	135	27	150	12	91	162	84
<i>R. pictipes</i>	110	44	143	11	84	154	70
<i>R. robustus</i>	120	62	166	16	99	182	83
<i>R. prolixus</i> from Mesquita et al. (2015)	116	0	111	5	84	116	84

Orthologs groups and tree

Tree scale: 1



**Figure 1: Consensus tree built from amino acid sequences of ORs from 13 *Rhodnius* species**

The phylogeny was built with IQTREE using maximum likelihood approach and SH-aLRT support. The model used is JTT+F+G4. The bootstrap procedure was applied with 1000 ultrafast bootstraps. ORs are numbered according to Mesquita et al. (2015). Old and recent *Rhodnius* expansions described by Mesquita et al. (2015) are indicated. The species included are: reference *R. prolixus* (Mesquita et al. 2015), *R. brethesi*, *R. colombiensis*, *R. coreodes*, *R. domesticus*, *R. milesi*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. neivai*, *R. pallescens*, *R. pictipes*, *R. prolixus* and *R. robustus*.

The phylogeny of the ORs genes from the 13 *Rhodnius* species was built with the complete normal genes with the start codon, including all the ORs already annotated in *R. prolixus* (Figure 1). It allowed to define ortholog groups and to number ORs genes according to the nomenclature of (Mesquita et al. 2015). Forty-two groups contain at least one sequence for each group species and could then be identified as unambiguous ortholog groups. Three major groups show evidence of expansion events namely OR21-23, OR28-29 and OR54. For the OR28-29 group, 49 genes were identified among the 13 assemblies with up to 7 genes in the two sylvatic species *R. colombiensis* and *R. robustus*. Furthermore, six of the *R. colombiensis* genes and all the *R. robustus* ones are organized in tandem on the same contig, a structure characteristic of recent duplication. The group OR21-23 comprises 25 genes over the 13 assemblies with up to 3 genes in *R. neglectus* assembly and a tandem organization of genes in *R. neivai* et *R. milesi*. Several copies of GR24 were found in the assemblies of 11 species. We tried to identify the sequence of duplication events among the *Rhodnius* species rooting the GR24 tree with GR23 sequence (data not shown) but the evolutionary history remains unclear probably due to multiple and overlapping gains and losses events. Similarly, tandem duplications are observed in this group OR54 notably for the 3 genes detected in *R. neglectus* et *R. milesi*.

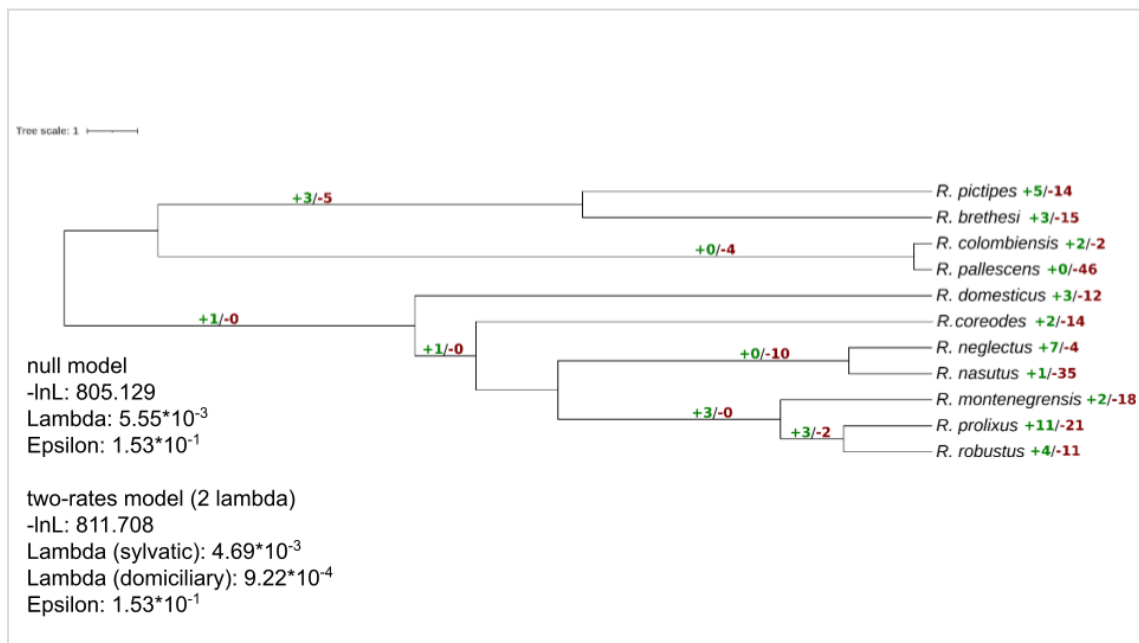
Compared to the published repertoire of *R. prolixus*, seven new ortholog groups were identified on the phylogeny (OR116 to OR122) and 18 ORs (OR63 to 69, OR81, OR75 to 78, OR88 to 91, OR94, OR100 and OR109) were not detected in any of the assemblies. Their loss could reflect a dynamic part of the ORs repertoire with shared ancient duplications at the Rhodniini tribe level and more recent ones species specific. Their absence may also be the consequence of technical bias: they may have been only partially retrieved by insectOR and consequently excluded from the phylogeny. This explanation is conceivable as the reference sequences in *R. prolixus* lack the N-terminal section of the protein for these genes which includes the start codon. Fortunately, orthologs of partial *R. prolixus* genes were identified on their entire length, thanks to InsectOR procedure which can take advantage of homology with different orthologs. For example OR71-73 are completely annotated in *R. neivai* while the sequence is only partial in the reference *R. prolixus* (Latorre-Estivalis et al. 2016). Similarly,



complete orthologs of OR74 were identified in *R. milesi* and *R. nasutus* and of OR85 in 7 assemblies despite partial sequences in *R. prolixus*.

### Gains and losses

The null model with a single evolutionary rate is more likely for OR genes than the two-rates model which assumes rate variation between domiciliary and sylvatic branches (-lnL of 805.129 and 811.708 respectively). The overall rate (including gain and loss events) is  $5.55 \times 10^{-3}$  per gene per million years (Figure 2). Most of the variations inferred are on terminal branches and evolution mostly reflects loss of genes.



**Figure 2: Gains and losses of ORs genes.** The phylogenetic tree was constructed from a concatenated alignment of 51 single-copy nuclear genes with time-measured analyses inferred using relaxed molecular clock models (Merle et al. in preparation). Gene family expansion and contraction were calculated with CAFE5 and are indicated in red and green respectively. Results are those from the null model with a unique evolutionary rate over all the tree.

## Selection analysis

Selection analysis was performed on the ortholog groups that could be found with the gene phylogeny including the ORco (co-receptor) and one receptor newly described in the present study (Table 4). Branch-model was tested that assumes different non synonymous/synonymous ratios among lineages (precisely one rate for domiciliary species and one rate for the sylvatic species). Significant results were found for ORCO, OR1, OR38 and OR102 indicating that differential selection occurred on these OR between branches specific to domiciliary species compared to that of the other branches. Site-model comparison where the ratio is assumed to vary in different ways among sites along sequences revealed clear signal of positive selection, i. e. for the three models M1a vs M2a, M7 vs M8, M8a vs M8, for 9 ORs (OR12, OR31, OR41, OR45, OR56, OR103, OR106, OR113 and the newly characterized OR118) and weaker signal i.e. for one or two of the models aforementioned, for 6 supplementary genes namely OR001, OR43, OR49, OR97, OR101 and OR108 (Table 4). All the data results including the model-site, M0 vs M3 that reveals a variation among sites rather than a positive selection are given in the Supp Table 2.

**Table 4: CodeML results for ORs.** For each OR and each test, the adjusted P.value (according to Benjamini and Hochberg) is indicated when lower 5%.

	Branch-model	Site-Model comparison		
		M1a_vs._M2a	M7_vs._M8	M8a_vs._M8
OR001	4.87E-02 $\omega_{syl}=0.14548$ $\omega_{dom}=0.34146$		4.87E-02	
OR012		4.15E-02	2.35E-02	1.14E-02
OR031		1.14E-04	1.09E-05	2.21E-05
OR038	4.73E-02 $\omega_{syl}=0.26777$ $\omega_{dom}=1.11510$			
OR041		1.49E-03	1.79E-03	8.16E-04
OR043				4.73E-02
OR045		2.46E-04	8.84E-08	6.16E-04
OR049				4.78E-02
OR056		3.10E-03	1.40E-03	6.43E-04
OR097			1.73E-02	
OR101			4.12E-03	
OR102	6.84E-04 $\omega_{syl}=0.19955$ $\omega_{dom}=0.78138$			

OR103		1.84E-05	3.81E-02	2.50E-02
OR106		1.04E-02	5.71E-03	2.33E-03
OR108			1.64E-02	1.14E-02
OR113		4.99E-05	1.60E-06	1.40E-05
OR118 (NEW)		2.00E-07	3.10E-07	7.41E-08
ORco	8.61E-03 $\omega_{syl}=0.02208$ $\omega_{dom}=0.53860$			

### GR repertory and evolution in *Rhodnius*

#### GR Annotation

According to the species, the total number of GR genes retrieved from genomes varies from 26 for *R. nasutus* to 44 for *R. pictipes* and the number of normal (non pseudogenes) complete genes and with a start codon GRs varies from 20 for *R. nasutus* to 30 for *R. neivai* (Table 4). From 0 (*R. nasutus* and *R. neglectus*) to 5 (*R. prolixus* and *R. coreodes*) pseudogenes were found in the assemblies. No pseudogenes were described in *R. prolixus* from Mesquita et al. (2015). The number of partial genes varies from 1 for *R. neglectus* to 13 for *R. pictipes*.

One disadvantage of using InsectOR to annotate genes is that it does not allow alternative transcripts identification. For example in *R. prolixus* there are 2 GRs that have two alternative transcripts (Mesquita et al. 2015): GR14-RA and GR14-RB, GR23-RA and GR23-RB. These alternative transcripts were not found, not even in our *R. prolixus* assembly. Focusing on the number of normal complete genes, it varies for the domiciliary species from 26 (*R. nasutus*) to 32 (*R. pallescens*) and for sylvatic species from 32 (*R. colombiensis*, *R. domesticus*) to 36 (*R. neivai*, *R. robustus*), except for *R. montenegrensis* (28), and *R. coreodes* (31). Considering the complete normal genes with start codon, the number varies from 20 to 25 for the domiciliary species and from 25 (*R. brethesi*) to 30 (*R. neivai*) except for *R. coreodes* and *R. montenegrensis* (23). For the domiciliary *R. nasutus*, the lower completeness of the genome (Merle et al. in prep) could be invoked for the lower GR number but for the three sylvatic species.

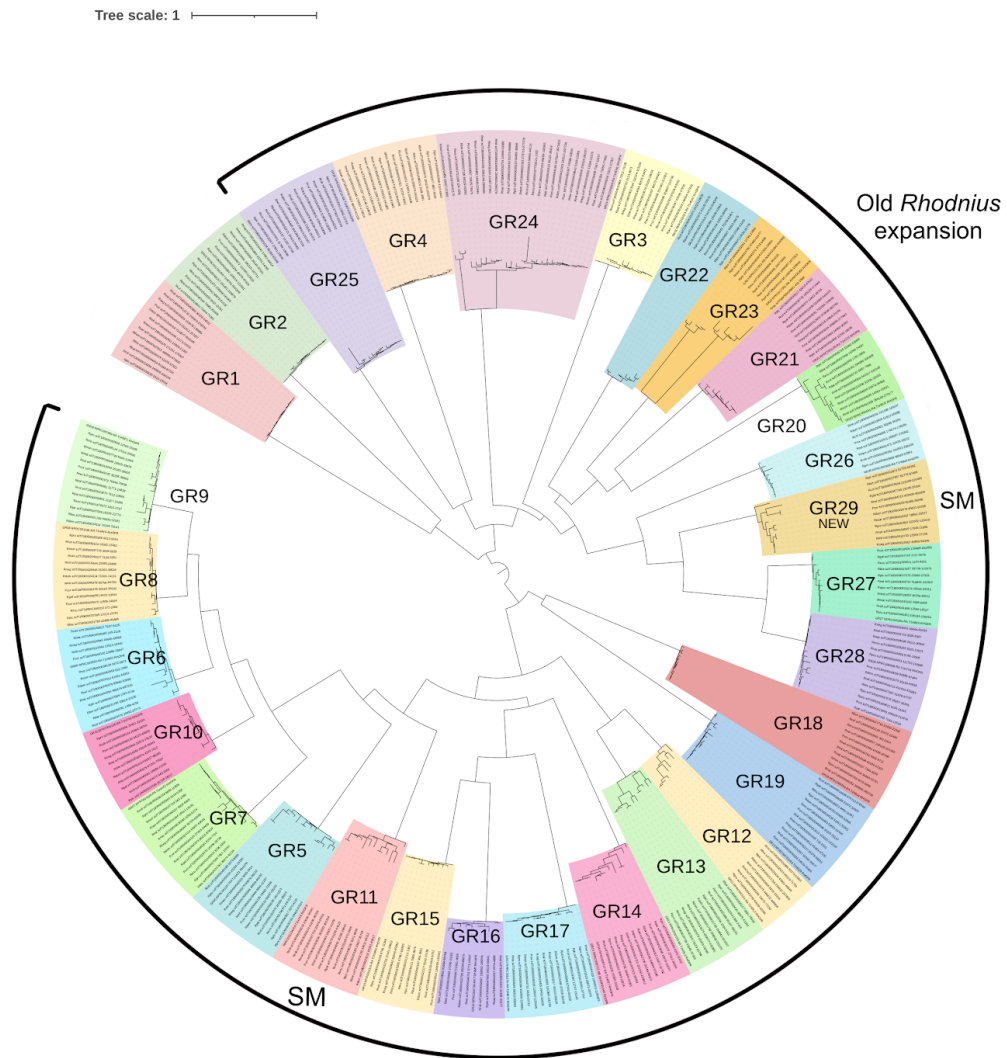
**Table 4: Results of GRs annotation obtained with InsectOR.** Domiciliary species are colored in red.

	Total number of Complete genes (including pseudogene)	Total number of Partial genes (including pseudogene)	Total number of Normal genes	Total number of Pseudo genes	Total number of genes with start codon present	Total number of genes	complete and not pseudogenes and start codon present
<i>R. nasutus</i>	21	7	28	0	22	28	20
<i>R. neglectus</i>	25	1	26	0	23	26	23
<i>R. pallescens</i>	24	11	32	3	25	35	24
<i>R. prolixus</i>	30	6	31	5	31	36	25
<i>R. brethesi</i>	29	10	35	4	26	39	25
<i>R. colombiensis</i>	28	6	32	2	28	34	28
<i>R. coreodes</i>	28	8	31	5	29	36	23
<i>R. domesticus</i>	31	3	32	2	29	34	28
<i>R. milesi</i>	29	6	34	1	29	35	28
<i>R. montenegrensis</i>	24	6	28	2	26	30	23
<i>R. neivai</i>	32	5	36	1	32	37	30
<i>R. pictipes</i>	31	13	41	3	33	44	28
<i>R. robustus</i>	31	6	36	1	31	37	29
<i>R. prolixus</i> from Mesquita et al. (2015)	28	0	28	0	28	28	28

### Ortholog groups and tree

Ortholog groups of GRs were determined on the phylogeny using normal complete genes and with a start codon, including all the GRs already annotated in *R. prolixus* (Figure 1). For instance, sequences that were found with GR1 in *R. prolixus* are considered to be orthologs of GR1. Most of the groups include one gene from each species with short terminal branches on the phylogeny. Such results illustrate the high level of conservation of GRs genes that facilitated their annotation. Missing orthologs in some species have to be interpreted cautiously as their lack is not necessarily the result of

their loss. They may be too divergent from the query to be detected. Their absence may result from the incompleteness of the assembly. Or the gene may be split on two different contigs, both parts being at best considered as partial genes.



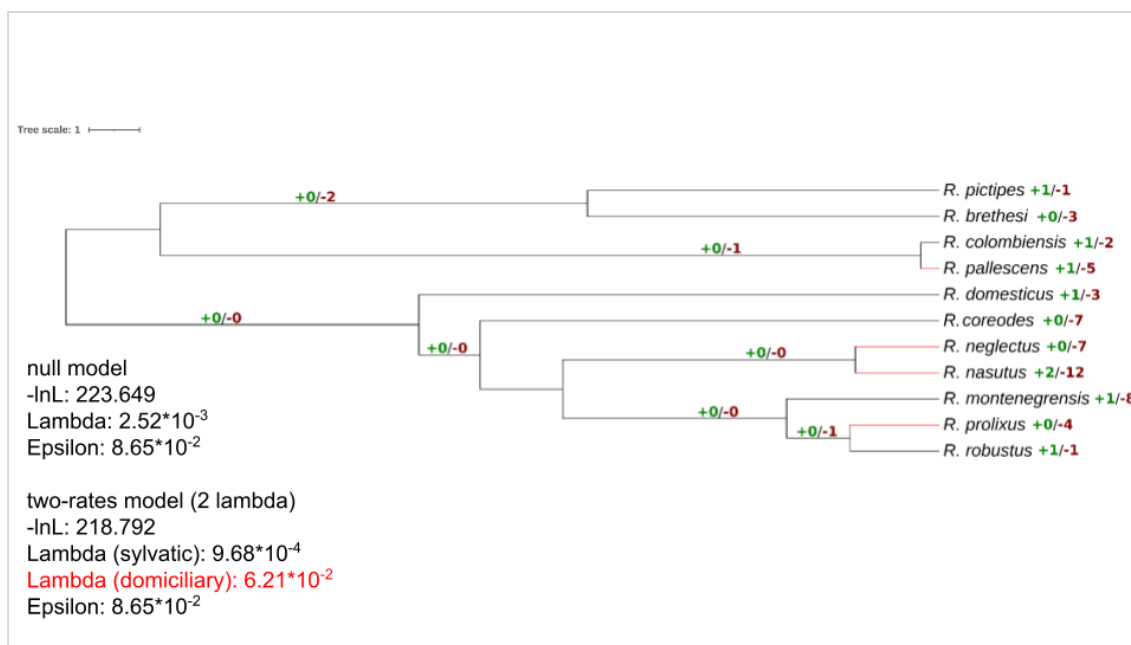
**Figure 3: Consensus tree built from amino acid sequences of GR from 13 *Rhodnius* species.** The phylogeny was built with IQTREE using maximum likelihood approach and SH-aLRT support. The model used is JTT+F+G4. The bootstrap procedure was applied with 1000 ultrafast bootstraps. GR are numbered according to Mesquita et al. (2015). Old *Rhodnius* expansion described by Mesquita et al. (2015) is indicated. The species included are: reference *R. prolixus* (Mesquita et al. 2015), *R. brethesi*, *R. colombiensis*, *R. coreodes*, *R. domesticus*, *R. milesi*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. neivai*, *R. pallescens*, *R. pictipes*, *R. prolixus* and *R. robustus*.

A novel gene, named GR29, was identified in this study but is absent in the annotation of the reference genome of *R. prolixus* (Mesquita et al. 2015). By contrast, it was detected in the assemblies of all other species. It is phylogenetically close to GR27 et GR28 (bootstrap of 100) and is also in tandem organization with these two genes on the assemblies indicating that these three genes are probably recent paralogs (Figure 3).

Several copies of GR24 were found in the assemblies of 11 species indicating that duplication events occurred in some lineages. This gene is also part of an expansion specific to *Rhodnius* in Mesquita et al. (2015). We tried to identify the sequence of duplication events among the *Rhodnius* species rooting the GR24 tree with GR23 sequence (data not shown) but the evolutionary history remains unclear probably due to multiple and overlapping gains and losses events.

#### Gains and losses

On the same GR dataset, CAFE5 was applied to estimate gains and losses of genes. The two-rates model assuming different rate between domiciliary branches (in red on Figure 4) and other branches was more likely than the one-rate model where the same gain and loss rate is applied to the whole tree (logLikelihood are respectively 218.792 and 223.649). The evolutionary rate ( $\lambda$ ) is about  $61.3 \times 10^{-3}$  gain or loss/gene/Ma for the domiciliary species versus  $0.97 \times 10^{-3}$  gain or loss/gene/MYA for the sylvatic species that is about 60-fold larger. Regarding gains and losses separately, there are for all the species considered more loss than gain events and particularly in domiciliary species but also in the two sylvatic *R. montenegrensis* and *R. coreodes* (Figure 4).



**Figure 4: Gains and losses of GRs genes.** The phylogenetic tree was built from a concatenated alignment of 51 single-copy nuclear genes with time-measured analyses inferred using relaxed molecular clock models (Merle et al. in prep). Gene family expansion and contraction were calculated with CAFE5 and are indicated in red and green respectively. Results from a two-rates model (one rate for domiciliary lineages, red branches, and one rate for other branches) are presented.

### Selection analysis

Because GR23 has two alternative transcripts published and because InsectOR fails to detect alternative transcripts, this gene was excluded from the selection analysis. GR14 was kept because all genes detected in the assemblies were unambiguously similar to the transcript GR14-RB. Two independent analyses were performed for GR11 and GR13 that comprises two duplicates in *R. robustus*.

No significant results were found for the branch-model showing that GRs are not under differential selective pressure depending on domiciliary/non-domiciliary habitat (Table 5). Significant results for model-sites comparisons, i. e. for the three models M1a vs M2a, M7 vs M8, M8a vs M8, reflecting occurrence of positive selection are observed for few GR. All model-site comparisons are significant for both duplicates of GR11 and for GR29 confirming that these genes are under positive selection in *Rhodnius* species. The significant result found for a single comparison for GR7 may be indicative of a weak signal (Table 5, supp Table 4).

**Table 5: CodeML results for GRs.** For each GR and each test, the adjusted p-value (according to Benjamini and Hochberg) is indicated when lower 5%.

	Site-Model comparison		
	M1a_vs._M2a	M7_vs._M8	M8a_vs._M8
<b>GR07</b>		3.95E-02	
<b>GR11 (<i>R. robustus</i> dup1)</b>	<b>2.30E-04</b>	<b>7.25E-05</b>	<b>6.76E-05</b>
<b>GR11 (<i>R. robustus</i> dup2)</b>	<b>6.90E-04</b>	<b>2.30E-04</b>	<b>1.49E-04</b>
<b>GR29</b>	<b>1.24E-05</b>	<b>1.15E-05</b>	<b>2.84E-06</b>

## Discussion

The present study allowed to finely describe two multigene families involved in insect chemosensory detection for 13 species belonging to the *Rhodnius* genus. The identification of orthologs was more easy for gustatory receptors for which sequences are well conserved than for odorant receptors. GRs are associated by pairs in the membrane of neurons and the sapid molecules are detected by the association of the two GRs (Liu et al. 2021). Changes in one GR may have consequences on the protein assembly and destroy the ability to detect specific chemicals. The OR family evolved from a lineage of GRs (Robertson et al. 2003). Each OR work is in combination with a co-receptor, it is the same for all ORs (ORco) and highly conserved, and forms in the membrane of olfactory sensory neurons a heterotetrameric receptor complex (Butterwick et al. 2018). It seems that each receptor includes a unique ligand binding site that varies in its specificity probably explaining the diversity of the OR family.

### *Chemosensory receptors panel in Hemiptera*

For the Chagas disease vector *R. prolixus*, 28 complete GR genes are described while the OR repertory includes 116 genes (Mesquita et al. 2015). After OR gene models were repaired on the basis of de novo transcriptome (Latorre-Estivalis et al. 2016), there are 84 complete genes, 5 pseudogenes, 22 genes with N-terminus part missing, 2 genes with C-terminus part missing. In our dataset one novel GR (GR29) was described, present in all the 13 *Rhodnius* species studied but *R. robustus*. For the OR dataset, we obtained a total of 118 complete genes with 7 new OR clusters with 8 genes,



new duplications (paralogues) for 9 already described genes and complete genes for 8 of the 22 OR genes with N-terminus part missing. The expanding repertory of the OR genes attests to a greater dynamics for this multigene family than for that of GRs. Indeed, the death and birth models showed that the gain and loss rate per gene per million years was 2-fold higher for OR genes ( $\lambda=0.006$ ) compared to GR ones ( $\lambda=0.003$ ). This rate estimated for the whole genome in triatomines is about 0.002 (Merle et al. in prep) indicating that the OR family is highly dynamic while the GR one is stable. For comparison, in the genus *Drosophila*, the GR rate is 3.7-fold higher ( $\lambda=0.011$ ) than in *Rhodnius* genus while the OR rate ( $\lambda=0.006$ ) is similar (McBride et Arguello 2007; Vieira et al. 2007; Gardiner et al. 2008). Moreover, positive selection was detected for 9 ORs and 2 GRs genes independently from sylvatic or domiciliary habitat. Such selection is indicative of diverging evolution and probably variation in the function of the encoded receptors specific to the *Rhodnius* genus.

As the insect chemosensory system plays a key role in ecological adaptation to new or changing habitat or diet, it was legitimate to ask whether the process of domiciliation of Chagas disease vectors was accompanied by a contraction of the repertory of the chemosensory genes due to the reduction of the panel of sensory stimuli in the anthropical environment, and/or specialization to this new habitat and the acquisition of specific duplications by gain of function. Even the number of chemosensory genes must be considered with caution in comparative studies due to the type of data (genomic, transcriptomic), the methodology of sequencing (short/long reads) and the depth of sequencing used, it appears for Hemiptera that *Rhodnius* species exhibited an intermediate range between generalists and specialists. For example, the polyphagous Pentatomidae pest bug *Hyalomorpha halys* that can feed on a large range of host plants including fruits, vegetables, crops, nuts, weeds, and wild plants, comprises 138 expressed OR genes (Sun et al. 2020). For the semi-aquatic bug, *Gerris buenoi*, with a diet highly variable because it feeds on various arthropods trapped by surface tension, an increase in the number of chemosensory genes was retrieved from the genome namely 155 ORs with large specific expansions and 60 GRs encoding 135 GR proteins (Armisen et al. 2018). For the milkweed bug Lygaeidae *Oncopeltus fasciatus* a moderate number OR genes (122) linked with the seed-feeding diet of this species was retrieved from genomic data, a number similar to that of *Rhodnius* species with their blood-sucking diet, but the milkweed bug exhibits a greater number of GRs (115 vs 28) (Panfilio et al. 2019). For Hemiptera specialist insects, a lower number of chemosensory receptors was reported like in aphid, specialists on one or a few plant hosts with 79 ORs and 77 GRs for *Acyrtosiphon pisum* but with specific expansions (Smadja et al. 2009). The number of chemosensory receptors is drastically reduced in the hematophagous bed-bug *Cimex lectularius* with 50 ORs and 35 GRs described from the transcriptome (Hansen et al. 2014). But whatever the lifestyle of the insects, specific events of expansions are evidenced showing

the species adaptation to the environment. In our data, at least three major ortholog groups (OR21-23, OR28-29, OR54) evidence recent expansion events with a tandem organisation of the copies.

#### *Chemosensory receptor gene repertory and specialisation in Rhodnius*

Interestingly OR duplication events concerned mainly sylvatic species. Indeed, if the habitat of the Triatominae bugs is considered, for GRs, the number of receptors is similar (*R. brethesi*, *R. colombiensis*) or higher for the sylvatic species except for *R. coreodes* and *R. montenegrensis* compared to the domiciliary species. For the ORs, two sylvatic species, *R. coreodes* and *R. brethesi* have the lower values in the range of that of domiciliary ones. These results are in favor of a loss of chemosensory receptor genes for the domiciliary species probably linked with a limitation of the diversity of the putative hosts that is corroborated for the GRs with a higher evolutionary rate for the domiciliary species than for the sylvatic species. However this similar pattern is observed for some sylvatic species namely *R. coreodes*, *R. colombiensis* or *R. montenegrensis* indicating some adaptation or specialisation to their sylvatic environment. For *R. coreodes*, this species is found in association with various bird species, especially the thornbird Furnariids (Marti et al. 2014), that constitute a restricted range of hosts compared to other *Rhodnius* species. This specialised ecology coupled with specific morphological characters had justified the recognition of a separate genus *Psammolestes* but extensive molecular data undoubtedly anchor the species of *Psammolestes* in the genus *Rhodnius* (Filée et al. 2022) as previously claimed by Hypša et al. 2002. *Rhodnius colombiensis* is restricted to the inter Andean valley of Magdalena River in central Colombia and inhabiting the *Attalea butyracea* palm tree species (Moreno Mejia et al. 1999). *Rhodnius montenegrensis* was sampled also in restricted areas namely on *Attalea phalerata* and *Attalea speciosa* in the municipality of Monte Negro, Rondônia, Brazil (da Rosa et al. 2012; Bilheiro et al. 2018) and in a rural dwelling the municipality of Rio Branco, Acre (Meneguetti et al. 2015). *Rhodnius brethesi* is only found in piassaba palm trees (*Leopoldinia piassaba*) in the microregion of Rio Negro, state of Amazonas, Brazil (Souza et al. 2017). These three species can be considered as a specialist infesting only one type of palm tree (Abad-Franch et al. 2015). By contrast, for both GRs and ORs, the two sylvatic species *R. neivai* and *R. robustus*, have the higher values. For *R. robustus*, its wide distribution in the Amazon region and the diversity of the palm tree and also bromeliads where it is found (Abad-Franch et al. 2005) could explain the large panel of its chemosensory gene repertory. However, *R. neivai* is limited to arid areas in central-western Venezuela (Machado-Allison et Pérez 1967) and northeastern Colombia (Lent et Wygodzinsky 1979) and if it was found in human dwelling (H. Lent et Jurberg 1969) or in the palm tree *Copernitia tectorum* (Carcavallo et al. 1976), only scarce data is available on its wild

biotope. Nevertheless, this species has a basal position in the *prolixus* group phylogeny (Filée et al. 2022) and seems to retain a generalist panel of chemosensory receptor genes.

Moreover, in our data, positive selection between domiciliar species and the background branches (i. e. sylvatic species and internal branches) was detected for four ORs namely ORco, OR1, OR38 and OR102. As domiciliation occurred independently on different branches on the *Rhodnius*, modifications on these genes accumulate in a convergent way. These OR and GR genes constitute candidate genes of interest to desorphanisation, in order to identify the molecules recognized by such receptors. The role of the co-receptor ORco, well conserved among insects, in the odorant perception was demonstrated with mutant lines in the hawkmoth *Manduca sexta*, the mosquito *A. aegypti* and *Anopheles gambiae* and the screwworm fly *Cochliomyia hominivorax*, (Liu et al. 2017; Fandino et al. 2019; DeGennaro et al. 2013; Paulo et al. 2021) that failed to perform successfully fundamental tasks such as food research or reproduction. In *R. prolixus*, ORco gene silencing used to infer its biological functions showed disruption in *R. prolixus* host-finding behavior including blood ingestion, molting, mortality, and egg laying (Franco et al. 2016). Positive selection on ORco for domiciliary species reinforced the major role played by this receptor for the *Rhodnius* physiology. Attempts to deorphanize *R. prolixus* ORs were performed using the *Xenopus* oocyte recording system for four ORs (RproOR1, 3, 74, and 80) significantly enriched in male antennae (Franco et al. 2018). However, even if none of these receptors responded to putative sex pheromone, RproOR80, was identified sensitive to repellents particularly 4-methylcyclohexanol. It will be of interest to see whether mutations fixed by selection have led to a change in ligand recognition especially for domiciliary species. The study of Smadja et al. (2012) shows that some chemoreceptor genes could play a key role in local adaptation and reproductive isolation in the pea aphid *A. pisum* by divergence under selection. Such polymorphism was also observed between populations of *Drosophila yakuba yakuba* and *Drosophila yakuba mayottensis* for which a signal of soft selective sweep was detected on GRs that could explain a loss of sensitivity for bitterness, probably for the hexanoic acid, allowing this species to develop on the toxic plant *Morinda citrifolia* (Ferreira et al. 2021).

#### *Putative role of chemosensory receptors in Rhodnius*

Homology studies could be performed to putatively assess some function of the chemosensory receptors but as a number of GR and OR ortholog groups are specific to *Rhodnius* species, the task is difficult. However, one GR, the GR1 is well conserved across the taxa and its homologs DmGR43a described in *Drosophila melanogaster* genome and

AmGR3 described in *Apis mellifera* genome, both detect fructose with high specificity (Sato et al. 2011; Takada et al. 2018). This GR shows evidence of purifying selection in the present study ( $dN/dS=0.09125$  see [supp data Table 4](#)) and it was found in all assemblies. The sequence conservation between *R. prolixus* sequence and its homologs in insects suggests that the bug receptor could also detect the fructose. This GR1 is furthermore highly expressed in the antenna of *R. prolixus* adults (Latorre-Estivalis et al. 2017). It is somewhat surprising for an hematophagous insect to maintain a high level of sugar detection. However, it has been shown that *R. prolixus* could be fed on saccharose solution or even on tomato (Díaz-Albiter et al. 2016). The authors have also shown that sugar consumption led the bug to ingest larger blood meals improving by the way its fitness and its lifespan. In sylvatic conditions, *Rhodnius* species live on palm trees where they can find sugar resources either in sap or fruits. Indeed, we recently evidenced that palm DNA is detected in the insect gut of *R. robustus* (Da Lage et al. [submitted](#)). They probably have conserved a sugar metabolic pathway and the ability to detect sugar from their phytophagous ancestor from which they diverged 230 million years ago (Wootton 1981). Among other gustatory receptors, it has been suggested that GR5 and GR11 could be involved in bitter recognition by homology with known receptors (Mesquita et al. 2015). Such a category of molecules is indicative of poison for most animals and used as a warning by numerous organisms (including plants and animals) to avoid predation. Insecticides, mostly derived from vegetal products, often have such bitter taste. Their massive spreading to limit Chagas vector populations may have induced strong selective pressure on bugs. Indeed, GR11 comprises sites showing evidence of positive selection that could modulate the sensitivity to bitter taste.

Functional identification of ORs is easier than for GRs functioning in combination and could be conducted in mutant drosophila or in xenopus ovocytes. It has been shown for example that OR4 in the mosquito *Aedes aegyptii* detects sulcatone, an odorant molecule produced by humans (McBride 2007). OR3 from the louse *Pediculus humanus* (PhOR3) participates to the detection of the insecticide DEET (Iwamatsu et al. 2016). *Cimex lectularius* OR9 (ClecPR9) detects decanal and OR5 (ClecOR5) detects nonanal and octanal (Liu et Liu 2015; Liu et al. 2017). Some sequence similarities have been found between such receptors and *Rhodnius* homologs namely between PhOR3 and RpOR101 and, RpOR102, but also ClecOR9 and RpOR105 as well as ClecOR5 and RpOR103. However, the sequence divergence is high enough to prevent any direct transposition of function.

## Conclusion

In this study, we annotated gustatory and olfactory receptors from 13 *Rhodnius* species and highlighted their potential link with domiciliation.

Evolution of these multigene families is based on gains and losses of genes by duplications, pseudogenisations and/or chromosomal rearrangements. Moreover, some process involving transposable elements may facilitate variation in repertoire size as it has been suggested for SINE transposons, abundant in the vicinity of GR genes in *Spodoptera littoralis* (Meslin et al. 2022). Accelerated gains and losses have been observed for GRs in domiciliary species and accelerated amino acid modifications of 4 ORs, including ORco, were detected in *Rhodnius* species adapted to human habitat. As ORco is involved in various physiological functions (Franco et al. 2016), it could be an interesting target to develop new control vector strategies. Indeed, factors that drive chemosensory genes evolution are multiples and comprise host recognition, reproduction, habitat specialization but also insecticide aversion. Identifying specific factors and the genes that respond to them is of major importance in the understanding domiciliation process and in a perspective of vector control.

### Funding

This study was partially funded by 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 the new CTA strains used in this study: Joao Aristeu da Rosa and Jader Oliveira (Universidade Estadual Paulista (Unesp), Faculdade de Ciências Farmacêuticas, Araraquara, São Paulo, Brazil);

## Supplementary Data

**Supplementary Table 1: Matrix of presence or absence of the different ORs by ortholog groups found in the 13 assemblies** (in red the domiciliary species)

	<i>R. colombiensis</i>	<i>R. pallescens</i>	<i>R. brethesi</i>	<i>R. pictipes</i>	<i>R. neivai</i>	<i>R. domesticus</i>	<i>P. coreodes</i>	<i>R. neglectus</i>	<i>R. nasutus</i>	<i>R. milesi</i>	<i>R. robus tus</i>	<i>R. montenegr ensis</i>	<i>R. prolixus</i>	TOTAL
OR001	1	1	1	1	1	1	1	1	1	1			1	11
OR002	1	1	1		1	1	1	1		1	1		1	10
OR008	1							1		1	1		1	5
OR009	1				1	1	1	1		1	1	1		8
OR012	1	1	1		1	1		1		1	1	1		9
OR013	1		1	1	1	1	1	1	1	1	1		1	11
OR014	1	1	1	1	1	1	1	1	1	1	1	1		12
OR016-018	1		1		3	1	2	2		3	3	2	3	21
OR019	1		1	1	1	1	1	1		1	1	1	1	11
OR020	1		1		1	1	1	1		1	1	1	1	10
OR021-023	2	1	1	2	3	2	1	3	1	3	2	2	2	25
OR024	1	1	1	2	1	1	1	1		1	1	1		12
OR025	1		1	1	1			1	1	1	1	1	1	10
OR026	1	1	1	1	1	1	1	1		1	1	1	1	12
OR027	1		1		1	1	1			1	1		1	8
OR028-29	7		2	4	5	4	3	3		4	7	6	4	49
OR03-5	1		1	1	2		1	1	1	2	3	1	3	17
OR030	1	1	1	1	1	1	1	1	1	1	1	1		12
OR031	1		1	1	1	1	1	1	1	1		1	1	11
OR032	1	1	1	1	1	1	1		1	1	1	1	1	12
OR033	1			1	1					1	1			5
OR034	1	1	1	1	1	1	1	1	1	1		1	1	12
OR036-37		1	1		1	1	1	2		2	2	1		12
OR038				1		1	1	1		1		1	1	7
OR039	1		1	1	1	1	1	1	1	1	1	2	1	13
OR040	1		1	1	1	1	1	1	1	1	1		1	11
OR041	1		1	1	1		1	1	1		1		1	9
OR042	1		1	1	1	1		1		1	1		1	9

OR043	1		1		1	1	1	1	1	1	1		1	10
OR044	1	1	1	1	1	1	1	1		1	1	1	1	12
OR045	1		1	1	1	1	1	1		1	1	1	1	11
OR046	1			1	1	1	1	1		1	1			8
OR047	1			1	1	1	1	1		1	1	1		9
OR048	1		1	2	2	1	1	1		1	2	1		13
OR049	1	1	1	1		1	1	1	1	1	1	1	1	12
OR050	1		1	1	1	1	1	1	1	1	1	1	1	12
OR051	1	1	1	1	1	1	1			1	1	1	1	11
OR052	1		1	1	1	1	1	1		1		1	2	11
OR053	1		1		1	1	1			1	1	1	1	9
OR054	2		1	2	3	2	2	3	1	3	1	2	1	23
OR055	1		1	1	1	1	1	1			1	1	1	10
OR056	1		1	1	1	1			1			1		7
OR057	1		1		1	1	1	1		1		1	1	9
OR058 & 061	1		2	2				1			2	2	1	11
OR059-60	2		2			2		2		1	1	1	1	12
OR062	1		1	1	1	1			1		1	1	1	9
OR063-69														0
OR07	1		1	2	1	1	2	2	1	1	2		1	15
OR070	1				1	1	1	1	1	1	1	1	1	10
OR071-73				1	3									4
OR074-78									1	1				2
OR079	1		1	1	1	1	1	1		1	1	1	1	11
OR080							1							1
OR081														0
OR082	1													1
OR083											1		1	2
OR084				1	1		1	1	2	2	1		1	10
OR085			1	2	1					1		1	1	7
OR086-87	1		3	2	1	1				1	2	1	2	14
OR088-91														0
OR092					1	2	1	1	1	1	1		1	9
OR093												1		1
OR094														0
OR095								1	1	1	1	1	1	6

OR096	1	1		1	1	1	1				1			7
OR097	1	1	1		1	1	1	1		1	1	1		10
OR098	1		1		1		1	1		1	1	1		8
OR099	1			1	1	1	1	1	1	1	1	1	1	11
OR100														0
OR101	1		1	1	1	1	1	1	1	1	1	1	1	12
OR102	1	1	1	1	1	1	1	1		1	1	1	1	12
OR103	1			1	1	1	1	1	1		1			8
OR104					1		1	1		1		1		5
OR105											1		1	2
OR106	1		1	1	1	1	1	1	1	1	1		1	11
OR107					1	1							1	3
OR108	1		1	1	1	1	1	1		1		1	1	10
OR109														0
OR110		1		1	1	1	1	1				1	1	8
OR111	1	1			1	1	1	1		1	1	1		9
OR112			2	1		2		2		1	1		1	10
OR113	1		1	1	1	1	1		1	1	1	1		10
OR114-115	1	1	1	1	2	2	2			2	2	2	1	17
OR116 (NEW)	1	1			1	1				1	1	1		7
OR117 (NEW)	1	1		1		1	1	1		1	1	1	1	10
OR118 (NEW)	1	1		1	1	1	1	1	1	1		1	1	11
OR119 (NEW)	1		1	1										3
OR120 (NEW)	1			1			1	2	1	1	1	1	1	10
OR121 (NEW)	1		1		2		2			1			1	8
OR122 (NEW)	1		4	4			1	1		1	2	2		16
ORco	1	1	1		1	1	1	1	1	1	1	1	1	12
TOTAL	78	24	66	70	84	72	70	74	34	82	83	70	69	876



**Supplementary Table 2: CodeML results for ORs.** For each OR and each test, the adjusted p-value (according to Benjamini and Hochberg) is indicated.

	Branch-model	Site-Model comparison			
		M0_vs_M3	M1a_vs_M2a	M7_vs_M8	M8a_vs_M8
<b>OR001</b>	<b>4.87E-02</b> $\omega_{syl}=0.14548$ $\omega_{dom}=0.34146$	4.20E-08	5.95E-01	4.87E-02	3.26E-01
<b>OR002</b>	$\frac{1.14E-01}{\omega=0.05975}$	9.52E-01	1.00E+00	1.00E+00	1.00E+00
<b>OR012</b>	1.09E-01 $\omega=0.42579$	2.01E-04	<b>4.15E-02</b>	<b>2.35E-02</b>	<b>1.14E-02</b>
<b>OR014</b>	1.00E+00 $\omega=0.28615$	1.57E-02	1.00E+00	1.00E+00	1.00E+00
<b>OR019</b>	5.35E-01 $\omega=0.43200$	2.81E-05	1.00E+00	1.00E+00	9.51E-01
<b>OR020</b>	1.00E+00 $\omega=0.20982$	8.17E-08	8.05E-01	2.12E-01	3.18E-01
<b>OR025</b>	9.51E-01 $\omega=0.38503$	1.62E-08	3.86E-01	3.78E-01	1.87E-01
<b>OR026</b>	5.92E-01 $\omega=0.31168$	1.62E-08	9.52E-01	4.72E-01	6.55E-01
<b>OR027</b>	5.16E-01 $\omega=0.20528$	5.25E-08	9.90E-01	2.30E-01	4.57E-01
<b>OR030</b>	1.28E-01 $\omega=0.26114$	8.57E-02	1.00E+00	1.00E+00	8.38E-01
<b>OR031</b>	5.73E-01 $\omega=0.69636$	1.62E-08	<b>1.14E-04</b>	<b>1.09E-05</b>	<b>2.21E-05</b>
<b>OR032</b>	2.12E-01 $\omega=0.24638$	1.62E-08	1.00E+00	6.19E-01	8.05E-01
<b>OR034</b>	7.49E-01 $\omega=0.24638$	5.21E-02	9.90E-01	4.94E-01	7.49E-01
<b>OR038</b>	<b>4.73E-02</b> $\omega_{syl}=0.26777$ $\omega_{dom}=1.11510$	5.52E-02	1.00E+00	1.00E+00	1.00E+00
<b>OR040</b>	7.55E-01 $\omega=0.17153$	1.46E-03	1.00E+00	1.00E+00	5.46E-01
<b>OR041</b>	7.71E-02 $\omega=0.46667$	1.62E-08	<b>1.49E-03</b>	<b>1.79E-03</b>	<b>8.16E-04</b>
<b>OR042</b>	2.64E-01 $\omega=0.38206$	1.62E-08	9.08E-01	7.55E-01	5.95E-01
<b>OR043</b>	2.66E-01 $\omega=0.49784$	4.11E-07	1.21E-01	8.44E-02	4.73E-02
<b>OR045</b>	2.43E-01	1.62E-08	<b>2.46E-04</b>	<b>8.84E-08</b>	<b>6.16E-04</b>

	$\omega=0.11020$				
<b>OR046</b>	6.02E-01 $\omega=0.46783$	7.19E-05	3.28E-01	2.66E-01	1.30E-01
<b>OR047</b>	8.05E-01 $\omega=0.18757$	8.21E-06	4.53E-01	8.57E-02	1.91E-01
<b>OR049</b>	9.52E-01 $\omega=0.23481$	1.63E-04	1.71E-01	6.60E-02	4.78E-02
<b>OR050</b>	5.92E-01 $\omega=0.23350$	2.00E-05	9.52E-01	5.73E-01	6.60E-01
<b>OR051</b>	3.28E-01 $\omega=0.17047$	2.02E-05	9.89E-01	1.30E-01	8.70E-01
<b>OR053</b>	1.00E+00 $\omega=0.12039$	2.22E-02	1.00E+00	9.82E-01	1.00E+00
<b>OR055</b>	1.46E-01 $\omega=0.29931$	4.50E-03	1.00E+00	1.00E+00	8.38E-01
<b>OR056</b>	9.90E-01 $\omega=0.55850$	1.58E-07	<b>3.10E-03</b>	<b>1.40E-03</b>	<b>6.43E-04</b>
<b>OR057</b>	1.00E+00 $\omega=0.23020$	6.07E-06	1.00E+00	5.92E-01	9.12E-01
<b>OR070</b>	7.55E-01 $\omega=0.21132$	8.44E-02	1.00E+00	1.00E+00	1.00E+00
<b>OR097</b>	5.95E-01 $\omega=0.14975$	8.21E-06	7.49E-01	1.73E-02	4.53E-01
<b>OR099</b>	5.92E-01 $\omega=0.21858$	3.92E-05	8.28E-01	3.73E-01	5.34E-01
<b>OR101</b>	2.41E-01 $\omega=0.15979$	1.62E-08	5.34E-01	4.12E-03	2.66E-01
<b>OR102</b>	<b>6.84E-04</b> <b><math>\omega_{syl}=0.19955</math></b> <b><math>\omega_{dom}=0.78138</math></b>	1.62E-08	3.61E-01	1.30E-01	1.45E-01
<b>OR103</b>	5.73E-01 $\omega=0.75792$	1.62E-08	<b>1.84E-05</b>	<b>3.81E-02</b>	<b>2.50E-02</b>
<b>OR104</b>	4.53E-01 $\omega=0.26831$	5.74E-01	1.00E+00	1.00E+00	1.00E+00
<b>OR106</b>	1.00E+00 $\omega=0.56398$	3.00E-08	<b>1.04E-02</b>	<b>5.71E-03</b>	<b>2.33E-03</b>
<b>OR108</b>	8.05E-01 $\omega=0.29322$	1.62E-08	5.07E-02	1.64E-02	1.14E-02
<b>OR110</b>	9.52E-01 $\omega=0.22089$	2.00E-07	6.17E-01	7.29E-02	3.37E-01
<b>OR111</b>	6.14E-01 $\omega=0.21793$	1.13E-01	1.00E+00	9.44E-01	8.87E-01

<b>OR113</b>	2.23E-01 $\omega=0.36315$	1.62E-08	<b>4.99E-05</b>	<b>1.60E-06</b>	<b>1.40E-05</b>
<b>OR118 (NEW)</b>	9.55E-1 $\omega=0.68395$	1.62E-08	<b>2.00E-07</b>	<b>3.10E-07</b>	<b>7.41E-08</b>
<b>ORco</b>	<b>8.61E-03</b> $\omega_{syl}=0.02208$ $\omega_{dom}=0.53860$	9.20E-02	1.00E+00	1.18E-01	1.00E+00

**Supplementary Table 3: Matrix of presence or absence of the different GRs by ortholog groups found in the 13 assemblies (in red the domiciliary species).**

	<i>R. colombiensis</i>	<i>R. pallens</i>	<i>R. ecuadoriensis</i>	<i>R. brethesi</i>	<i>R. pictipes</i>	<i>R. neivai</i>	<i>R. domesticus</i>	<i>P. coreodes</i>	<i>R. neglectus</i>	<i>R. nasutus</i>	<i>R. milesi</i>	<i>R. robustus</i>	<i>R. montenegrensis</i>	<i>R. prolixus</i>	TOTAL
GR01	1	1		1	1	1	1	1	1	1	1	1	1		10
GR02	1	1	1	1	1	1	1	1	1	1		1	1	1	10
GR03	1	1				1	1	1	1	1	1	1	1	1	9
GR04	1	1	1	1	2	1	1	1	1		1	1	1	1	11
GR05	1	1	1	1	1	1	1	1	1		1	1	1	1	10
GR06	1	1		1	1	2	1	1	1	2	1	1	1		12
GR07	1	1	1	1	1	1	1	1	1	2	1	1	1	1	12
GR08	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11
GR09	1	2	1	1	1	1	1	1	1		1	1	1	1	10
GR10	1	1			1	1	1	1	1	1	1	1		1	9
GR11	1	1	1	1	1	1	1	1			1	2	1		9
GR12	1			1	1	1	1				1	1		1	7
GR13	1	1	1	1	1		2		1		1	1		1	8
GR14	1	1		1	1	1	1		1	1	1	1		1	9
GR15		1	1		1	1	1	1	1		1	1		1	8
GR16				1	1	1	1	1		1	1	1		1	9
GR17	1	1		1	1	1	1	1	1	1	1	1	1	1	11
GR18	1		1	1		1		1	1	1	1	1	1	1	9
GR19	1	1		1	1	1	1	1	1	1	1	1	1	1	11
GR20	1	1	1	1	1	1				1	1	1	1		7
GR21	1	1			1	1	1	1	1		1	1		1	8
GR22	1			1	1	1	1				1	1		1	7

GR23		1	1	1	1	1					1	1	1	1	7
GR24	3		2	2	2	2	2	2	2	2	2	1	3	1	21
GR25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11
GR26	1					1	1	1			1	1	1	1	7
GR27	1	1	1	1	1	1	1		1			1	1	1	8
GR28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	11
GR29 (NEW)	1	1		1	1	1	1	1	1	1	1		1	1	10
TOTAL	28	24	17	25	28	30	28	23	23	20	28	29	23	25	282

**Supplementary Table 4: CodeML results for GRs.** For each GR and each test, the adjusted p-value (according to Benjamini and Hochberg) is indicated.

	Branch-Model	Site-Model comparison			
		M0_vs_M3	M1a_vs_M2a	M7_vs_M8	M8a_vs_M8
<b>GR01</b>	$\frac{1.47E-01}{\omega=0.09125}$	1.00E+00	1.00E+00	1.00E+00	8.98E-01
<b>GR02</b>	$\frac{2.78E-01}{\omega=0.09409}$	2.18E-01	1.00E+00	1.00E+00	7.43E-01
<b>GR03</b>	$\frac{1.04E-01}{\omega=0.25523}$	1.00E+00	1.00E+00	1.00E+00	1.00E+00
<b>GR04</b>	$\frac{9.57E-01}{\omega=0.14568}$	3.39E-02	1.00E+00	6.09E-01	8.26E-01
<b>GR05</b>	$\frac{4.89E-01}{\omega=0.55392}$	2.30E-04	1.88E-01	1.04E-01	8.38E-02
<b>GR06 (<i>R. neivai</i> dup1)</b>	$\frac{1.36E-01}{\omega=0.55392}$	2.53E-06	3.09E-01	2.78E-01	1.40E-01
<b>GR06 (<i>R. neivai</i> dup2)</b>	$\frac{1.67E-01}{\omega=0.46374}$	3.91E-06	2.17E-01	1.66E-01	9.92E-02
<b>GR07</b>	$\frac{1.00E+00}{\omega=0.26108}$	3.00E-08	2.90E-01	3.95E-02	1.19E-01
<b>GR08</b>	$\frac{3.06E-01}{\omega=0.17440}$	2.84E-06	1.00E+00	7.29E-01	1.00E+00
<b>GR09</b>	$\frac{3.09E-01}{\omega=0.17705}$	2.84E-06	1.00E+00	7.29E-01	1.00E+00
<b>GR10</b>	$\frac{1.00E+00}{\omega=0.31022}$	6.06E-05	2.96E-01	1.21E-01	8.38E-02

<b>GR11 (<i>R. robustus</i> dup1)</b>	6.49E-01 $\omega=0.58604$	3.00E-08	<b>2.30E-04</b>	<b>7.25E-05</b>	<b>6.76E-05</b>
<b>GR11 (<i>R. robustus</i> dup2)</b>	5.40E-01 $\omega=0.54703$	3.00E-08	<b>6.90E-04</b>	<b>2.30E-04</b>	<b>1.49E-04</b>
<b>GR12</b>	1.46E-01 $\omega=0.49059$	3.50E-02	9.97E-01	7.29E-01	7.29E-01
<b>GR13 (<i>R. domesticus</i> dup1)</b>	3.09E-01 $\omega=0.62374$	1.15E-05	3.28E-01	2.17E-01	1.60E-01
<b>GR13 (<i>R. domesticus</i> dup2)</b>	2.84E-01 $\omega=0.64962$	4.00E-07	1.39E-01	9.31E-02	5.64E-02
<b>GR14</b>	2.17E-01 $\omega=0.64962$	3.61E-06	6.61E-01	4.92E-01	3.02E-01
<b>GR15</b>	<u>1.40E-01</u> <u><math>\omega=0.06724</math></u>	3.95E-02	1.00E+00	4.92E-01	1.00E+00
<b>GR16</b>	4.17E-01 $\omega=0.18838$	3.00E-08	6.54E-01	1.12E-01	4.92E-01
<b>GR17</b>	7.31E-01 $\omega=0.10735$	7.45E-05	1.00E+00	1.00E+00	5.04E-01
<b>GR18</b>	8.61E-01 $\omega=0.14446$	2.03E-01	1.00E+00	1.00E+00	9.99E-01
<b>GR19</b>	1.00E+00 $\omega=0.37532$	8.19E-02	1.00E+00	9.99E-01	8.61E-01
<b>GR20</b>	2.60E-01 $\omega=0.39688$	8.53E-03	1.00E+00	1.00E+00	7.57E-01
<b>GR21</b>	7.33E-01 $\omega=0.29678$	1.19E-04	2.60E-01	1.22E-01	7.36E-02
<b>GR22</b>	1.00E+00 $\omega=0.10999$	1.77E-02	1.00E+00	5.84E-01	7.61E-01
<b>GR25</b>	9.31E-02 $\omega=0.26563$	2.30E-04	7.30E-01	4.51E-01	4.34E-01
<b>GR26</b>	1.00E+00 $\omega=0.30665$	7.73E-01	7.29E-01	5.89E-01	4.29E-01
<b>GR27</b>	8.76E-01 $\omega=0.21591$	2.99E-05	8.26E-01	3.73E-01	4.70E-01
<b>GR28</b>	7.35E-01 $\omega=0.30947$	3.23E-05	5.06E-01	3.80E-01	2.17E-01
<b>GR29 (NEW)</b>	9.99E-01 $\omega=0.47073$	3.00E-08	<b>1.24E-05</b>	<b>1.15E-05</b>	<b>2.84E-06</b>

## POSTAMBULE SUR L'ARTICLE EN COURS D'ÉCRITURE

### Vérifier les pertes de gènes chimiosensoriels

Les pertes de gènes sont compliquées à valider. En effet, ce n'est pas parce que le gène n'a pas été retrouvé dans notre assemblage qu'il a été perdu, il peut simplement ne pas avoir été assemblé lors de l'assemblage du génome avec le programme MaSuRCA (couverture de séquençage locale trop faible par exemple). Il est possible d'essayer de compléter l'annotation en recherchant les séquences manquantes dans les transcriptomes. Nous avons des données transcriptomiques d'antenne et rostre pour 8 espèces de *Rhodnius* (*R. pallescens*, *R. colombiensis*, *R. brethesi*, *R. pictipes*, *R. nasutus*, *R. neglectus*, *R. montenegrensis*, *R. prolixus* et *R. robustus*). Nous pourrions par exemple rechercher un homologue d'OR1, manquant dans l'assemblage de *R. robustus*, dans le transcriptome de cette même espèce.. Cette méthodologie nécessite un assemblage des transcriptomes car la famille multigénique des ORs comprend des gènes plus longs que les *reads* illumina mais surtout des gènes ayant une homologie commune et donc des domaines conservés. Elle est de plus peu fiable. En effet, un résultat de blast montrant une séquence du transcriptome de *R. robustus* ressemblant à la séquence requête de l'OR1 de *R. prolixus* ne serait pas gage de la présence de l'OR1 dans le transcriptome de *R. robustus*, cette séquence pouvant correspondre à un autre OR. Cette difficulté se pose d'autant plus lorsque l'on étudie des gènes récemment dupliqués souvent en tandem sur le génome. De plus, nous n'avons pas les transcriptomes de *R. coroedes*, *R. milesi*, *R. neivai* et *R. domesticus*. Rechercher les ORs et les GRs directement dans les séquences brutes Illumina pré-assemblage poserait le même problème dû à la similarité de séquences des protéines chimiosensorielles appartenant à une même famille multigénique. Une technique de vérification potentiellement efficace serait de créer des amorces PCR spécifiques des régions comprenant les GRs/ORs recherchés. Une autre amélioration pourrait consister à faire du *scaffolding* des génomes, notamment en utilisant la méthode Hi-C qui permet un séquençage à l'échelle chromosomique. Cette méthode est basée sur l'immobilisation de la conformation de la chromatine et le séquençage des régions d'interaction afin d'avoir une cartographie des contacts de la chromatine. Cette approche permet d'avoir accès à l'organisation structurale des chromosomes et donc potentiellement dans nos données de mettre en évidence les tandems d'ORs et de GRs et de compléter des séquences partielles ou réparties sur deux contigs différents. Cette méthode a été utilisée ce qui a été réalisé sur l'assemblage de référence de *R. prolixus* de [Mesquita et al. \(2015\)](#) par l'équipe de [Mathers et al. \(2021\)](#) mais pour lequel nous n'avons pas vérifié la pertinence de ce scaffolding par manque de temps.

## Autres protéines chimiosensorielles

Pour compléter le répertoire des gènes chimiosensoriels, les IRs sont en cours d'étude en utilisant la méthodologie mise en place pour annoter les GR et les ORs. La méthode est la même que décrite plus haut. J'ai réalisée une base de donnée à partir des IRs déjà décrits d'espèces proches, et les IRs putatifs sont recherchés par homologie de séquence avec le programme *exonerate* dans les génomes des différentes espèces de *Rhodnius* et la sortie du programme *exonerate* est analysée avec *InsectOR*. La méthode fonctionne assez bien. Les OBPs et les CSPs ont été extraits par Gabin RIGNAULT que j'ai co-encadré lors de son stage de M2 (janvier 2022 - juin 2022). Les phylogénies réalisées montrent que les gènes sont très bien conservés entre les espèces. Un article est en préparation concernant ces résultats.

## Structure des protéines

Les GRs et ORs sous sélection détectés avec PAML présentent des mutations non-synonymes, c'est-à-dire que le codon code un acide aminé qui diffère selon l'habitat. Il serait intéressant de rechercher si ces mutations ont un impact sur les propriétés des récepteurs notamment structurales et conformationnelles. Les différences d'acides aminés dans les séquences des protéines chimiosensorielles entre deux espèces peuvent se situer sur une des 7 régions transmembranaires du récepteur ou impacter la reconnaissance des ligands et permettre, par exemple, la reconnaissance d'un autre composé volatil. Une nouvelle plateforme (*iORbase*, <https://www.iorbase.com> ; Li et al. 2022) permet l'analyse structurale et fonctionnelle des iORs par la prédiction de la structure des protéines par homologie avec des ORs présents dans la base de données. Cette plateforme offre la possibilité de calculer les affinités de liaison et les résidus de liaison entre les ORs et les phéromones.





## DISCUSSION GÉNÉRALE

---

L'objectif de mes travaux de thèse était d'étudier l'évolution des génomes avec un focus sur les gènes chimiosensoriels. Afin de faire la part du signal phylogénétique et du signal adaptatif, cette étude nécessitait au préalable une identification fiable des espèces et du contexte phylogénétique.

Sur les 24 espèces de *Rhodnius* au sens large en incluant les 3 espèces du genre *Psammolestes* que nous avons invalidé (voir ci-après), nos études portent sur 18 espèces avec un nombre variable selon les études du fait de la disponibilité ou non de spécimens frais (*long reads*, FC). Ainsi ont été étudiées 6 (FC), 11 (évolution des génomes), 13 (évolution des gènes chimiosensoriels), et 17 espèces (phylogénomique). Pour les raisons invoquées ci-dessous, les espèces *R. dalessandroi* et *R. zeledoni* ne peuvent pas être étudiées. Nous n'avons pas eu accès à des échantillons ou aux données moléculaires des espèces nouvellement décrites *R. barretti* et *R. micki*. En revanche, la collection de M. Harry contenait des individus identifiés comme *R. robustus* mais qui après analyse phylogénomique se sont avérés appartenir à *R. marabaensis* (voir ci-après). Nos collections d'insectes et celles de nos partenaires (Collections M. Harry, J.-M. Béranger, J.A. Da Rosa) ne contenaient pas l'espèce *R. arthuri*. Enfin, pour *R. parensis*, un essai d'extraction d'ADN a été tenté sur un échantillon de 3 pattes d'un individu conservé à sec (J.-M. Béranger collection, Guyane 2021) mais l'ADN obtenu n'était pas en quantité suffisante pour faire une banque Illumina.

### HISTOIRE ÉVOLUTIVE DES ESPÈCES DU GENRE *RHODNIUS*

Si le genre *Rhodnius* a été décrit par Lent en 1859, avec l'espèce *R. prolixus* (Lent et Wygodzinsky 1979), de nouvelles espèces ont été décrites récemment et à l'heure de l'utilisation de séquences moléculaires pour la reconnaissance des espèces, ces données sont partielles pour *R. barretti* (Abad-Franch et al. 2013), non publiées pour *R. marabaensis* (Souza et al. 2016) ou manquantes pour *R. micki* (Zhao et al. 2021). Cependant, des données moléculaires invalident la reconnaissance de l'espèce *R. taquarussuensis* (da Rosa et al. 2017), mise en synonymie avec *R. neglectus* (Nascimento et al. 2019). En revanche, l'espèce *R. colombiensis* a été identifiée sur des caractères moléculaires. Cette espèce a été initialement confondue avec une forme sylvatique de *R. prolixus* occupant des palmiers dans la région centrale des vallées andines de Colombie mais son statut taxonomique en tant qu'espèce distincte a été confirmé par des analyses isoenzymatiques et moléculaires (Moreno Mejia et al. 1999; Jaramillo et al.

2001). Pour deux espèces, le statut spécifique n'est pas testable moléculairement : *R. zeledoni*, morphologiquement proche de *R. paraensis*, n'a été décrite qu'à partir d'un seul spécimen mâle mort et séché (Jurberg et al. 2009), et *R. dalessandroi* n'est en fait connue que par sa description publiée (Carcavallo et Barreto 1976). Enfin, pour l'espèce *R. paraensis* initialement décrite comme *R. domesticus* (Sherlock et al. 1977) présente au Brésil et en Guyane (Jean-Michel Bérenger et al. 2009), il y a très peu de spécimens et en collection.

Trois groupes ont été reconnus sur des arguments morphologiques et biogéographiques (Dujardin et al. 1999) : *pallescens*, *pictipes*, *prolixus*. Dans chacun des groupes, des questions taxonomiques se posent ainsi que la relation phylogénétique entre les trois groupes.

Les principales sources de divergence entre les phylogénies peuvent provenir d'écueils méthodologiques et/ou refléter la réalité biologique. Un biais méthodologique peut être induit par des échantillonnages de taxons incomplets ou différents selon les phylogénies mais aussi par le type et le nombre de marqueurs utilisés. Pour limiter le biais méthodologique dû à un ensemble de données limité, nous avons utilisé dans une étude phylogénomique un grand ensemble de données génomiques mitochondriales (13,3 kb) et nucléaires (5,3 kb d'ADNr ; 51 gènes codants, 36,3 kb) et un grand nombre d'échantillons (36 échantillons représentant 17 espèces). De plus, la phylogénie datée ainsi que les analyses d'orthologie réalisées dans l'étude portant sur la comparaison des génomes fournissent des éléments complémentaires.

Nous avons démontré que les gènes mitochondriaux bien que largement utilisés pour les phylogénies, à savoir *cytb* et *COI*, sont sous sélection pour les trois groupes, *pallescens*, *pictipes* et *prolixus* (Filée et al. 2022). Notre étude indique également que l'utilisation des séquences marqueurs d'ADNr nucléaires (y compris 18S, 5.8S, 28S et ITS) conduit à un arbre mal résolu. Nous avons donc proposé d'utiliser des gènes nucléaires pour établir les phylogénies. Cependant, si les relations phylogénétiques entre les trois groupes de *Rhodnius* ont été résolues, il persiste des incongruences dans le placement de certains échantillons entre nos phylogénies mitochondriales et nucléaires. Ces incongruences peuvent être dues à différents facteurs dont des transferts horizontaux de gènes, de l'introggression, ou de l'hybridation (Toews et Brelsford 2012; Campillo et al. 2019).

### Les groupes cis-andins *prolixus* et *pictipes* ne sont pas monophylétiques

Nos études ont permis de résoudre les relations entre les trois groupes de *Rhodnius*. A l'exception de quelques gènes mitochondriaux, les ensembles de données mitochondriales d'une part et les données

nucléaires d'autre part placent *pictipes* et *pallescens* comme groupes frères, indépendamment de l'outgroup utilisé.

Nos résultats remettent en question l'événement de spéciation entre le groupe *pallescens* trans-andin et les groupes cis-andins *pictipes* et *prolixus*. [Justi et al. \(2016\)](#) avait proposé que cet événement de spéciation allopatrique était dû au soulèvement des Andes du Nord qui a conduit à la formation d'une vaste zone humide il y a 10 à 23 Ma appelée système Pebas ([Hoorn et al. 2010](#)). Cet événement aurait donc, selon eux, isolé les ancêtres des groupes cis-andins et trans-andins. Notre scénario est en faveur d'un événement de spéciation entre le groupe (*pallescens* + *pictipes*) et le groupe *prolixus* puis un autre événement de spéciation entre le groupe cis-andin *pictipes* et le groupe trans-andin *pallescens*. La divergence entre le groupe (*pallescens* + *pictipes*) et le groupe *prolixus* est estimée à 17 millions d'années. En revanche, le temps de divergence entre le groupe *pallescens* et le groupe *pictipes* de 15 millions d'années est compatible avec la formation du système Pebas. Cet événement géologique aurait pu avoir une conséquence sur la spéciation entre le groupe trans-andin *pallescens* et le groupe cis-andin *pictipes*.

### Diversité des espèces et statut taxonomique

Le groupe trans-andin *pallescens*, limité aux régions côtières de l'Équateur, de la Colombie, du Panama et du Costa Rica, est composé de trois espèces : *R. pallescens*, *R. colombiensis* et *R. ecuadoriensis*. *Rhodnius ecuadoriensis* est limitée au sud de l'Équateur et au nord du Pérou. Cette espèce occupe les palmiers et est également domiciliaire. Elle est géographiquement séparée de *R. pallescens* et *R. colombiensis* par les montagnes andines et par les forêts pluviales de la côte pacifique colombienne. *Rhodnius pallescens* est largement distribué en Amérique centrale et en Colombie alors que *R. colombiensis*, sylvatique, est limitée à la vallée inter-andine de la rivière Magdalena dans le centre de la Colombie. Bien que les espèces *R. colombiensis* et *R. pallescens* aient été trouvées dans la même région en Colombie et habitent la même espèce de palmier (*A. butyracea*), aucun hybride naturel n'a été signalé ([Díaz et al. 2014](#)). Cependant, certains auteurs considèrent que *R. pallescens* peut être un complexe de deux lignées divergentes ([Díaz et al. 2014](#)). Dans notre étude phylogénomique, les espèces *R. colombiensis* et *R. pallescens* sont plus affines entre elles qu'avec *R. ecuadoriensis* mais nous mettons en évidence des flux géniques complexes entre les trois espèces de ce groupe. Ces résultats ne sont pas en accord avec ceux de [Justi et al. \(2016\)](#) proposant une plus grande affinité phylogénétique entre *R. ecuadoriensis* et *R. colombiensis* qu'avec *R. pallescens*. Cet auteur avait proposé la fermeture de l'isthme du Panama (2,76-10,1Ma) comme un événement ayant participé à la spéciation entre *R. pallescens* et l'ensemble (*R. ecuadoriensis* + *R.*

*colombiensis*) du fait de la dispersion de *R. pallescens* au Panama lors de la fermeture de l'isthme. Cette hypothèse n'est pas validée par nos données, que ce soit par les estimations de datation ou les analyses phylogénomiques. Nous avons en effet estimé que l'événement de spéciation entre *R. pallescens* et *R. colombiensis* datait de moins d'un million d'années.

Concernant le groupe cis-andin *pictipes*, distribué à l'est des Andes, il comprend sept espèces (*R. amazonicus*, *R. brethesi*, *R. micki*, *R. paraensis*, *R. pictipes*, *R. stali*, et *R. zeledoni*). Cependant, pour les raisons invoquées ci-dessus nous n'avons pu étudier que 4 de ces espèces. Le statut de *R. amazonicus* proche de *R. pictipes* a été questionné. L'espèce *R. amazonicus* a été initialement décrite par Almeida et al. (1973) puis invalidée par Lent et Wygodzinsky (1979) en raison de la description faite sur une seule femelle mal conservée collectée dans la commune de Manaus (Amazonas, Brésil). L'espèce a ensuite été revalidée par Bérenger et Pluot-Sigwalt (2002) à partir de spécimens mâles et femelles collectés en Guyane française et plus récemment par (da Rosa et al. 2017) d'après des spécimens du Brésil (Para). Dans notre analyse phylogénomique nous confirmons le statut de *R. amazonicus* et présentons les premières données moléculaires sur cette espèce. L'espèce *R. stali* a été décrite à partir de spécimens de muséum (Collection Oswaldo Cruz de Lent) identifiés comme *R. pictipes* et provenant de l'état du Mato Grosso (Lent et al. 1993). Cette espèce a été depuis échantillonnée dans l'État d'Acre au Brésil (Meneguetti et al. 2015). L'espèce *R. brethesi* est décrite au Brésil (Amazonas, Para) et au Venezuela (Lent et Wygodzinsky 1979). Nos analyses phylogénomiques confirment la reconnaissance de l'espèce *R. stali* à part entière mais mettent en évidence une introgression entre *R. stali* et *R. brethesi*.

Le groupe *prolixus*, cis-andin sauf pour quelques populations trans-andines de *R. neivai*, est le plus divers avec quatorze espèces (*R. arturi*, *R. barretti*, *R. coeodes*, *R. dalessandroi*, *R. domesticus*, *R. marabaensis*, *R. milesi*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. neivai*, *R. prolixus*, *R. robustus* et *R. tertius*). Ce groupe pose le plus de questions taxonomiques. Des études pionnières utilisant des allozymes avaient remis en question le statut d'espèces étroitement apparentées pour, *R. robustus* et *R. prolixus* après l'analyse de populations vénézuéliennes sylvatiques décrites comme étant *R. robustus* (Harry 1993; Harry et al. 1992). Récemment, des études moléculaires sur une espèce nouvellement décrite *R. montenegrensis* (da Rosa et al. 2012) ont suggéré qu'il ne s'agit pas d'une véritable espèce mais d'un sous-ensemble de *R. robustus* (Abad-Franch et al. 2013; Monteiro et al. 2018). Cependant, les deux espèces *R. montenegrensis* et *R. marabaensis* ont été reconnues par Castro et al. (2020). De plus, certains auteurs considèrent *R. milesi* comme une variante de *R. neglectus* du sud-est de l'Amazonie (Abad-Franch et al. 2013). Beaucoup de mauvaises identifications entre les espèces *R. milesi*, *R. nasutus*, *R. neglectus*, *R. prolixus* et *R. robustus* ne facilitent pas l'analyse des données bibliographiques

concernant la distribution des espèces et leur possible sympatrie. Notre étude phylogénomique a mis en évidence de mauvaises identifications pour un spécimen de *R. milesi* (*R. prolixus*), de *R. neglectus* (*R. prolixus*) et de *R. robustus* (*R. marabaensis*). Nous avons confirmé le statut taxonomique de *R. montenegrensis* et de *R. marabaensis* et la mise en synonymie de *R. taquarussuensis* avec *R. neglectus*. Nous mettons en évidence des introgressions entre différentes espèces, *R. robustus* et *R. montenegrensis*, *R. robustus* et *R. prolixus*, *R. nasutus* et *R. neglectus*. Nous avons également résolu la paraphylie de l'espèce *R. robustus* pointée par [Monteiro et al. \(2000\)](#) par la reconnaissance des espèces *R. montenegrensis* et *R. marabaensis* et l'introgression de *R. robustus* du Venezuela par *R. prolixus*.

Nous avons démontré que l'introgression pouvait être un événement biologique puisque des échantillons de terrain montrent de l'introgression. Il est possible aussi que nous ayons certains spécimens hybrides issus du terrain. En ce qui concerne les souches de laboratoire, on ne peut pas écarter le mélange d'échantillons ayant conduit à de l'introgression ou à de l'hybridation post échantillonnage de terrain. Cependant, il est impossible de dire si les souches de laboratoire présentant un tel schéma ont été accidentellement introgressées ou ont été fondées avec des spécimens naturellement introgressés ou hybridés en milieu naturel.

### La mort du genre *Psammolestes*

Concernant le genre *Psammolestes*, des différenciations morphologiques et écologiques avec le genre *Rhodnius* ont justifié le classement taxonomique dans deux genres distincts. Les différences morphologiques portent sur une tête plus courte, un rostre plus large ainsi que des pattes plus fortes avec des fémurs plus larges. Écologiquement, les espèces du genre *Psammolestes*, *P. arthuri*, *P. coreodes*, *P. tertius* sont associées aux nids d'oiseaux, en particulier les Furnariidés ([Carcavallo et al. 1998](#); [Marti et al. 2014](#)) et semblent utiliser les palmiers de manière opportuniste ([Abad-Franch et al. 2015](#)) tandis que les espèces de *Rhodnius* vivent principalement dans les palmiers où ils se nourrissent sur une très large gamme d'hôtes vertébrés ([Lent et Wygodzinsky 1979](#)).

Les caractères morphologiques des espèces de *Psammolestes*, peuvent être considérés comme adaptatifs, ce qui doit exclure leur utilisation pour établir des liens de parenté. Il existe d'ailleurs des convergences morphologiques avec une autre espèce de *Rhodnius*, *R. paraensis*.

Cette espèce, *R. paraensis*, présente également une écologie particulière car elle est trouvée dans les nids des rongeurs arboricoles. C'est une espèce de plus petite taille que les espèces du genre *Psammolestes* avec également une tête courte, des pattes robustes, des antennes courtes et

robustes. Cette proximité morphologique a conduit [Lent et Wygodzinsky \(1979\)](#) à suggérer que ces caractères sont plésiomorphes au sein du genre *Rhodnius*. Sous cette hypothèse, l'allongement progressif de la tête et des appendices que l'on trouve chez la plupart des autres espèces est considéré comme dérivé. *R. paraensis* aurait conservé l'état ancestral de ces caractères. Il n'existe pas aujourd'hui de données moléculaires permettant de placer *R. paraensis* sur la phylogénie et de polariser ces caractères morphologiques et écologiques.

Notre travail a clairement établi en utilisant à la fois de grands ensembles de données mitochondriales et nucléaires que *P. tertius* est inclus dans le genre *Rhodnius* et plus précisément dans le groupe *prolixus*. Cette position était déjà relevée par différents auteurs (e.g., [Lyman et al. 1999](#); [Monteiro et al. 2000, 2002](#); [Hypša et al. 2002](#); [Justi et al. 2014](#); [de Paula et al. 2021](#)). L'analyse du génome de *P. coreodes* et la réalisation de la phylogénie datée montre que cette espèce forme un groupe monophylétique avec *P. tertius* au sein du groupe *prolixus*. Un échantillon de *P. arthuri* nous permettrait de confirmer que toutes les espèces de *Psammolestes* sont dans le groupe *prolixus*. Au niveau génomique, nous n'avons pas montré de différences notables dans le contenu en gènes de *P. coreodes* par rapport aux *Rhodnius*. Nous avons aussi montré que *P. coreodes* partage le répertoire de gènes chimiosensoriels des autres espèces de *Rhodnius*. Les GRs et les ORs de *P. coreodes* forment toujours des groupes monophylétiques avec les GRs et ORs de *Rhodnius*. Il n'y a pas de différences notables dans ces répertoires de gènes.

Le genre *Psammolestes* étant inclus au sein du genre *Rhodnius*, les différences de morphologie et d'écologie à elles seules ne permettent pas de justifier la conservation des deux genres. Il est plus vraisemblable que les caractères de robustesse soient des innovations morphologiques liées à l'adaptation à une nouvelle niche écologique que sont les nids et pourraient s'expliquer par des modifications impliquant un nombre restreint de gènes. L'incompatibilité reproductrice entre deux espèces de *Psammolestes* (*P. coreodes* et *P. tertius*) et *R. neglectus* proposée par [Ravazi et al. \(2022\)](#) pour maintenir le genre *Psammolestes* suite à notre publication ([Filée et al. 2022](#)) est hors de propos.

Aussi, nous revendiquons l'abandon du genre *Psammolestes* et avons utilisé les nominations *R. coreodes* et *R. tertius* dans nos publications postérieures à celle de phylogénomique ([Filée et al. 2022](#)).

## EVOLUTION ET ADAPTATION AU SEIN DE GENRE *RHODNIUS*

Notre étude visait à comparer la dynamique des génomes d'espèces du genre *Rhodnius* dépendantes de différents habitats. Une question majeure était de savoir si la domiciliation de certaines espèces s'accompagnait d'une évolution rapide en termes de gains et pertes de gènes associée à la diversification écologique des espèces et impliquant des familles de gènes clés dont les expansions ou contractions pouvaient permettre la diversification fonctionnelle. Puisqu'à ce jour seul le génome de l'espèce domiciliaire *R. prolixus* est décrit (Mesquita et al. 2015), nous avons séquencé, assemblé et annoté le génome de treize espèces de *Rhodnius* dont quatre espèces domiciliaires et neuf espèces sylvatiques et généré; Nous avons également généré un arbre phylogénétique ultramétrique afin de dater les événements de divergences et d'estimer les vitesses d'évolution.

Cependant, l'acquisition des données génomiques était dépendante de la disponibilité et de la qualité du matériel frais pour le séquençage en long reads. Les assemblages ont été réalisés une première fois avec une couverture de séquençage Illumina d'environ 20x et le séquençage long read PacBio (7x) sur des échantillons conservés en alcool. Les données préliminaires sur l'étude de la dynamique des génomes ont été acquises avec 8 assemblages (*R. brethesi*, *R. colombiensis*, *R. montenegrensis*, *R. nasutus*, *R. neglectus*, *R. pallescens*, *R. pictipes*, *R. prolixus* et *R. robustus*). Au vu des résultats préliminaires (étapes d'annotation, groupes d'orthologie et calculs des vitesses de gains et pertes le long des branches) et des pourcentages de complétude peu satisfaisants pour certains assemblages, nous avons refait des extractions à partir d'individus frais congelés après transport en carboglace depuis le Brésil. Pour 5 espèces, *R. coreodes*, *R. domesticus*, *R. milesi*, *R. neivai*, et *R. robustus*, nous avons pu faire du séquençage GridION (avril 2021) étant donné que la technologie de séquençage avait évolué depuis 2016 (date du séquençage PacBio). Pour ces 5 espèces et 5 autres espèces, un séquençage Illumina a été fait ou refait. Comme expliqué dans le matériel et méthodes général, des essais de combinaisons ont été réalisés, combinant les anciennes données Illumina (de 2016) et les nouvelles données Illumina (fin 2021) et les *long reads* pour voir si cela nous permettait d'obtenir des assemblages plus complets. Un réel effort a été fait pour trouver la procédure permettant d'obtenir les meilleurs assemblages, plus de 50 assemblages (toutes espèces confondues) ont été réalisés.

## La domiciliation, une histoire récente

Une question centrale réside dans la datation de la domiciliation des vecteurs de la maladie de Chagas. Les premiers contacts avérés Homme-parasite sont datés d'il y a environ 9 000 (BP soit environ 7 050 BC) chez une momie Chinchorro retrouvée dans le désert d'Atacama situé entre le nord du Chili et le sud du Pérou et dont l'analyse d'ADN a révélé la présence du parasite *T. cruzi* (Aufderheide et al. 2004). Cette période correspond à celle à laquelle le peuple Chinchorro initie la colonisation côtière des embouchures de rivières (environ 7 500 BC) et la pratique d'une économie maritime jusqu'à 1 500 BC (Guhl 2017). Les auteurs montrent également en étudiant d'autres momies que les infections à *T. cruzi* sont présentes dans cette région jusqu'à la conquête espagnole au XVI<sup>ème</sup> siècle, avec un taux de prévalence de 41 % quelle que soit la période considérée ce qui indique que la maladie de Chagas était déjà largement répandue à l'époque précolombienne.

Se pose alors la question de savoir à partir de quel moment on peut parler de domiciliation, c'est-à-dire l'établissement de colonies de vecteurs dans les habitations humaines. Selon des découvertes dans le nord-est du Brésil (zone du parc national de la Serra da Capivara), les premières présences préhistoriques dans des grottes et des abris sous roche contenant de nombreuses peintures rupestres dateraient de 26 000 BP (Araújo et al. 2009). Les auteurs suggèrent en s'appuyant sur des observations actuelles, que les premiers artistes préhistoriques ont été les premiers sud-américains à être exposés à la transmission du parasite par *T. brasiliensis*. Dans ce cas, nous sommes dans une situation d'infection en milieu naturel, les populations humaines allant dans les lieux sylvatiques où résident les vecteurs. Il est probable que la domiciliation des vecteurs ne commence véritablement qu'avec la domestication de petits rongeurs (*Cavia* sp.) pour la consommation et pour les rituels, il y a environ 8 000 à 6 000 ans, et le stockage de grains attirant des mammifères granivores. Ces ressources alimentaires auraient été attractives pour les vecteurs. La plus ancienne infection de la momie de Chinchorro date donc de plusieurs milliers d'années avant l'époque connue où les rongeurs et les camélidés ont été domestiqués (Rothhammer et al. 1985). Les données de paléoparasitologie analysées par Araújo et al. (2009) montrent pour une momie datée de 1 150 ans, la présence d'une pathologie digestive (lésions digestive et présence du génotype TCI du parasite) et suggèrent une contamination par voie orale suite à l'ingestion de rongeurs contaminés ou de vecteurs eux-mêmes.

Sans que l'on puisse réellement dater ces événements, l'anthropisation grandissante des milieux sylvatiques a amplifié le processus de domiciliation. Quoi qu'il en soit, d'après nos phylogénies datées, les



différentes espèces étaient déjà différenciées bien avant leur contact avec les populations humaines.

En effet, les datations combinant des données moléculaires avec des données fossiles convergent vers une origine des Triatomines autour de 35 à 40 millions d'années (Ibarra-Cerdeña et al. 2014; Justi et al. 2016) et plus précisément entre 27 et 32 millions d'années pour l'étude la plus complète (Hwang et Weirauch 2012). D'après nos données, l'âge du dernier ancêtre commun au genre *Rhodnius* est estimé à 17 Ma, la divergence du groupe *prolixus* vers 11 Ma et la scission entre le groupe *pallescens* et *pictipes* vers 15,3 Ma. Les espèces différenciées le plus récemment sont au moins datées d'1 Ma.

Il est probable que la colonisation des habitations humaines a engendré des pressions évolutives sur les punaises. On peut se demander si le pas de temps est suffisant pour observer des traces de cette spécialisation dans les génomes. Des événements de spécialisation très rapides sont documentés. Par exemple, chez la drosophile, on observe deux événements de spécialisation indépendant sur le même fruit (*Morinda citrifolia*, toxique pour les autres drosophiles) par convergence évolutive. Le temps de divergence entre *D. sechellia* (spécialiste) et *D. simulans* (généraliste) est estimé à environ 4 800 ans et le temps de divergence entre *D. yakuba mayottensis* (spécialiste) et *D. yakuba yakuba* (généraliste) est estimé à 400 ans. Des différences génomiques sont observées malgré ces divergences récentes telles que des allèles différents sur l'OR22 entre les spécialistes et leur espèce sœur généralistes ou encore une inversion chromosomique entre *D. sechellia* et *D. simulans* (Ferreira 2022).

D'autres exemples de spécialisation récente peuvent être cités. Ainsi, le déplacement, de certaines populations de l'espèce de mouche *Rhagoletis pomonella* de son hôte ancestral indigène, l'aubépine (*Crataegus spp.*), vers le pommier *Malus pumila* domestique introduit en Amérique de Nord depuis l'Europe est daté de seulement 150 ans (Berlocher et Feder 2002). Ce changement de plante hôte par la mouche semble lié à la reconnaissance des composés volatiles produits par ce nouvel hôte : les *R. pomonella* du pommiers et les *R. pomonella* de l'aubépine n'ayant pas les mêmes comportement en fonction des molécules volatiles proposées (et leurs hybrides ont une réponse plus faible aux molécules volatiles des deux fruits) (Linn et al. 2004).

Des études de génomique ont montré que certaines caractéristiques du génome comme une grande diversité génétique, le contenu en éléments transposables, la diversification des familles de gènes, permettent une évolution rapide notamment liée à la sélection positive de gènes régulateurs ou structuraux (Pélissié et al. 2018).

Dans ce travail, nous avons analysé certaines de ces caractéristiques

génomiques au sein des espèces domiciliaires et sylvatiques.

### Une taille de génome variable au sein du genre *Rhodnius* et une évolution plus rapide de certains gènes

L'étude de la taille des génomes n'a pas révélé d'évidence quant au lien entre taille de génome et domiciliation. En revanche, il existe une variation inter-spécifique au sein du genre *Rhodnius* notamment pour le groupe *prolixus* et même intraspécifique dans le cas de *R. robustus*. Ainsi, les trois espèces *R. coreodes*, *R. domesticus* et *R. neivai* présentent un plus petit génome. Il est difficile de statuer sur cette différence car ces trois espèces ont une position basale dans le groupe. Le partage d'une petite taille de génome pourrait refléter un signal phylogénétique. Mais ces trois espèces sont aussi toutes trois sylvatiques et présentent des spécialisations écologiques. La réduction de la taille du génome pourrait être un caractère dérivé précédé de l'augmentation de la taille du génome pour toutes les espèces du groupe *prolixus*.

Dans nos données, l'influence du nombre de chromosomes sur la taille des génomes peut être écartée car toutes les espèces du genre *Rhodnius* partagent le même caryotype  $2n = 22$  (20A + XY).

Les modifications de la taille du génome peuvent impliquer des mécanismes mutationnels, notamment liés à l'activité des éléments transposables (Petrov 2001; Naville et al. 2019). Cependant, dans nos données, il n'y a pas de corrélation simple entre la taille du génome et le contenu en ETs (éléments transposables), comme l'illustre le fait que *R. coreodes* et *R. robustus*, qui sont respectivement caractérisées par la plus petite et la plus grande taille de génome, ont les contenus en ETs les plus élevés.

Les modèles d'évolution des génomes par perte et gain de gènes montrent pour la part codante du génome une évolution 2,4 fois plus rapide pour le genre *Rhodnius* comparé aux autres Hémiptères (Merle et al. en préparation). Dans une étude plus générale incluant différents ordres d'insectes dont sept lignées indépendantes d'insectes hématophages (5 diptères, 2 hémiptères), les lignées hématophages ont un taux de gain et de pertes de gènes plus lent que les lignées non hématophages (Freitas et Nery, 2020), hypothèse que nous n'avons pas testé dans notre jeu de données. Ces auteurs montrent également qu'il existe pour les espèces hématophages un petit nombre de familles de gènes évoluant rapidement, notamment une famille codant pour des protéines de choc thermique et une autre codant pour des protéines chimiosensorielles. Parmi les familles de gènes évoluant le plus rapidement (c'est-à-dire avec le plus grands nombre d'événements de pertes et de gains), nous avons identifié des gènes liés à des activités métaboliques (protéolytiques, catalytiques,

kinasiques), de détoxification (cytochrome P450), des gènes impliqués dans les liaisons métalliques, ainsi que des éléments transposables.

L'évolution plus rapide de certains gènes liés au métabolisme (contractions/expansions) peut être mise en relation avec le régime alimentaire hématophage des punaises. Les ions métalliques jouent un rôle important dans tous les systèmes biologiques et semblent impliqués dans la fertilité des insectes femelles (Cardoso-Jaime et al. 2022) comme chez *R. prolixus* pour les gènes impliqués dans la fixation du fer et la liaison à l'hème (Walter-Nuno et al. 2018). Les métalloprotéines seraient aussi impliquées dans l'immunité (Cardoso-Jaime et al. 2022). L'expansion des protéases à activité lytique pourrait être une adaptation au régime sanguin spécifiquement riche en protéines comme cela a été suggéré pour les glossines, diptères hématophages vecteur de trypanosomes africains. Le génome de six de ces espèces a été exploré et cette étude a révélé qu'ils étaient spécifiquement enrichis en gènes protéolytiques (Attardo et al. 2019).

Il est généralement admis que les expansions ou les contractions des familles de gènes impliquées dans la détoxification sont associées à des adaptations fonctionnelles liées au régime alimentaire et à d'autres caractéristiques de l'environnement (Rane et al. 2016). Une corrélation positive entre le nombre de gènes de détoxification et la complexité des sources de nourriture des espèces a été rapportée (Rane et al. 2019). Cela ne semble cependant pas être le cas pour les pucerons (Hemiptera ; Aphididae) (Lin et al. 2022). Que les insectes soient phytophages ou hématophages, ils sont exposés aux toxines, aux répulsifs ou aux composés anti-digestifs des plantes pour les premiers ou à des quantités toxiques d'hème, de fer et d'acides aminés à chaque repas pour les seconds (Volonté et al. 2022). Les cytochromes P450 des insectes sont bien connues pour être impliquées dans la détoxification métabolique des xénobiotiques, tels que les composés phytochimiques, les insecticides et les polluants environnementaux (Lu et al. 2021) comme cela a déjà été démontré pour les Triatomes (Traverso et al. 2017). Cependant, dans l'étude de Volonté et al. (2022) portant sur des hémiptères comportant des phytophages et des hématophages (*Cimex*, *R. prolixus*), l'espèce la plus polyphage *Halyomorpha halys* présente le plus grand répertoire d'enzymes de détoxification. Concernant les triatomes se nourrissant sur une large gamme d'hôtes vertébrés, les cytochromes pourraient avoir un rôle plus général lié à l'alimentation.

Les éléments transposables pourraient potentiellement avoir un impact à la fois sur l'architecture du génome et sur la régulation des gènes (Vargas-Chavez et al. 2022). Chez *Spodoptera littoralis*, ce rôle a été montré pour la dynamique des GRs. La présence de rétrotransposons de classe I, les transposons SINE, est significativement enrichie à proximité des clusters de gènes de récepteurs gustatifs, ce qui suggère un mécanisme médié par

les transposons pour la formation de ces groupes (Meslin et al. 2022). Dans nos données, la dynamique des transposons pourrait expliquer les forts événements de duplication identifiés dans certains groupes d'orthologie, notamment les ORs.

En effet, chez les espèces du genre *Rhodnius*, les gains et pertes se produisent 3 fois plus rapidement pour les ORs que pour les autres gènes du génome. Cette dynamique est plus importante pour les ORs que pour les GRs. En effet, les modèles de perte et de gains de gènes ont montré que ce taux par gène par million d'années était 2 fois plus élevé pour les ORs. Les GRs chez les espèces du genre *Rhodnius* semblent avoir une certaine stabilité. A titre de comparaison, dans le genre *Drosophila*, le taux d'évolution des GRs est 3,6 fois plus élevé que dans nos données alors que le taux pour les ORs est similaire (McBride et Arguello 2007; Vieira et al. 2007; Gardiner et al. 2008).

Parmi les gènes chimiosensoriels plus finement analysés, nous avons mis en évidence des gènes sous sélection quel que soit l'habitat, soulignant l'importance de ces récepteurs dans la physiologie du genre *Rhodnius*.

### Certains gènes chimiosensoriels sont sous sélection chez les espèces du genre *Rhodnius*, cas du GR1

Nos résultats ont montré que 9 ORs et 2 GRs présentent des sites sous sélection, cependant du fait de l'absence d'homologie avec des séquences d'autres espèces et de l'absence d'expériences d'analyse fonctionnelle pour ces récepteurs, nous ne pouvons pas donner d'éléments physiologiques expliquant une telle sélection.

Pour un autre récepteur, le GR1, notre analyse de sélection a montré que la séquence de ce récepteur est sous sélection purificatrice (dN/dS proche de zéro) donc très conservée au sein du genre *Rhodnius*. Ce GR1 chez *R. prolixus* montre une homologie avec les récepteurs associés à la reconnaissance du fructose (Mesquita et al. 2015). Les gènes codant pour les récepteurs du fructose sont décrits comme hautement conservés chez les insectes (Wu et al. 2020). Ce gène a été retrouvé dans 12 des 13 assemblages de cette étude. Étonnamment, nous ne l'avons pas retrouvé dans l'assemblage de *R. prolixus*. En recherchant plus spécifiquement ce GR dans l'assemblage, les deux premiers exons sur les 6 attendus ont pu être retrouvés sur la fin du contig scf718000040521 de la position 44 308 à la position 45 217, ce contig comprenant 46 351 bp. Il est donc vraisemblable que l'absence de ce GR dans l'assemblage de *R. prolixus* soit artefactuelle, d'autant que l'on sait que ce gène a été retrouvé dans le génome de *R. prolixus* de référence de Mesquita et al. (2015).

La présence de ce récepteur au fructose pourrait paraître étonnante

chez un insecte hématophage. Cependant, il est possible d'alimenter *R. prolixus* avec une solution de saccharose (composé de glucose et de fructose), et même avec de la tomate. [Díaz-Albiter et al. \(2016\)](#) en concluent que « tout le monde aime le sucre ». Ces auteurs montrent également que l'absorption de sucre a des conséquences sur la fitness des individus qui ensuite se gorgent d'une plus grande quantité de sang lors d'un repas sanguin et ont une durée de vie plus longue. Cette consommation de sucre végétal pourrait permettre aux individus de survivre en période de famine dû à l'absence d'hôte sur lequel se nourrir ([Díaz-Albiter et al. 2016](#)). Ce gène aurait été hérité d'un ancêtre phytophage. En effet, l'hématophagie chez les Triatominae est un caractère dérivé qui serait apparu il y a environ 85 millions d'années à partir de punaises Reduviidae prédatrices qui se seraient elles-mêmes spécialisées sur l'hémolymphe d'autres insectes il y a environ 230 millions d'années à partir d'un ancêtre phytophage ([Wootton 1981](#)). Ce scénario évolutif est entre autres documenté par le fait que l'on observe de l'hématophagie facultative chez certaines punaises phytophages (références dans [Díaz-Albiter et al. \(2016\)](#)). Dans une récente étude, nous avons démontré qu'en milieu sylvatique *R. robustus* pouvait se nourrir à partir d'éléments végétaux, de l'ADN de palmier ayant été identifié dans le tube digestif des spécimens testés. De plus, cette espèce possède une amylase fonctionnelle ([Da Lage et al. soumis](#)). La présence de gènes d'amylase avait déjà été notée chez *R. prolixus* ([Mesquita et al. 2015](#)) et chez *P. lignarius* ([Nevoa et al. 2018](#)).

L'alimentation sur des plantes pourrait fournir des nutriments ou des molécules importantes pour la physiologie des vecteurs et/ou impliquées dans la digestion du sang. Chez *R. prolixus*, la bactérie symbiotique intestinale *Rhodococcus rhodnii* fournit à ses hôtes les vitamines du groupe B manquant dans le sang ingéré et la bactérie endosymbiotique du genre *Wolbachia* pourrait assurer ou compléter le mutualisme nutritionnel apporté par les symbiotes intestinaux ([Filée et al. 2023](#)). La mutualité nutritionnelle pourrait être plus complexe impliquant les ressources végétales. Fait intéressant, une pratique ethno-botanique courante dans le traitement des maladies veineuses et inflammatoires repose sur l'ingestion de feuilles ou d'écorce de palmier. Une étude expérimentale a montré qu'un traitement par voie orale avec du palmier a un effet anti-thrombotique significatif chez la souris, lié à une augmentation de la capacité des macrophages à produire de l'oxyde nitrique et à un processus de coagulation lent. L'étude suggère une utilisation potentielle de palmier comme agent prophylactique pour prévenir les événements thrombotiques chez l'Homme ([Azevedo et al. 2007](#)). Ces mécanismes sont à rapprocher de ceux induits par différentes protéines salivaires lors de l'alimentation sanguine.

## Une dynamique de génome différenciée chez les espèces domiciliées et des familles de gènes orthologues en expansion

Dans notre étude, le taux d'évolution en termes de perte et de gain de gènes, défini à partir de groupes d'orthologues, est 16 fois plus élevé chez les espèces domiciliées que chez les espèces sylvatiques. Le ratio entre pertes et gain de gènes est plus élevé chez les espèces sylvatiques que domiciliées. Notre étude de génomique comparative a permis de définir des familles de gènes dont la dynamique n'est pas la même chez les espèces domiciliées et sylvatiques. Nous avons annoté les groupes d'orthologie qui ont subi de fortes duplications soit chez les espèces domiciliées soit chez les espèces sylvatiques et nous en avons identifié 26 dans le premier cas et 16 dans le second. Ces familles de gènes comprennent donc des gènes candidats potentiellement impliqués dans le processus de domiciliation.

Parmi les familles d'orthologues enrichies chez les espèces domiciliées, nous avons annoté des familles liées notamment aux éléments transposables, à l'immunité, à la résistance aux insecticides, et à des protéines putatives impliquées dans l'alimentation sanguine (nitrophorine) et dans la réception chimiosensorielle (ORs). Parmi les familles d'orthologues enrichies chez les espèces sylvatiques, nous avons annoté des familles liées notamment aux éléments transposables, à la physiologie musculaire et à l'immunité.

Il est intéressant de noter que les gènes liés aux éléments transposables sont représentés parmi les familles d'orthologues enrichies pour les espèces domiciliées comme pour les espèces sylvatiques. D'autre part, l'analyse du contenu des génomes en ETs a montré un pourcentage plus bas pour les espèces domiciliées (8-10%) que pour les espèces sylvatiques (11-20%) excepté pour l'espèce domiciliée *R. prolixus* (16%). Le rôle des ETs dans les mécanismes adaptatifs est largement documenté. L'un des exemples les plus médiatiques est celui de la phalène du bouleau (*Biston betularia*) pour laquelle la forme noire dont la fréquence s'est accrue dans les régions d'Angleterre soumise à la pollution industrielle, est liée à l'insertion du transposon *carb*-ET dans le premier intron du gène *cortex*. Cette insertion a pour effet de fortement augmenter la quantité de protéines *cortex* produite dans les disques imaginaires des ailes, aboutissant à la coloration noire des ailes. Le rôle des ETs dans l'adaptation aux insecticides a été aussi démontré chez les moustiques anophèles (Weedall et al. 2020). Chez *Anopheles coluzzii*, un des principaux vecteurs du paludisme humain en Afrique subsaharienne, des insertions d'ETs ont été identifiées à proximité de gènes de résistance ou d'immunité (Vargas-Chavez et al. 2022). Les espèces domiciliées étant très fortement soumises aux insecticides chimiques, il est attendu que les gènes de résistance montrent une forte dynamique pour les espèces inféodées à l'habitat humain. Les premières campagnes de lutte anti-vectorielle ont eu

lieu vers les années 1950 avec l'emploi de DDT, puis d'insecticides organochlorés tels que le lindane et la dieldrine. Les premières constatations de résistance à la dieldrine ont été reportées du Venezuela en 1971 (González Valdivieso et al. 1971). Les organophosphorés et carbamates moins persistants sont alors proposés comme alternative, et introduits en 1975. Puis les pyréthriinoïdes comme la deltaméthrine et la cyperméthrine sont notamment utilisés en Argentine et au Brésil (Mougabure-Cueto et Picollo 2015). Cependant, en Bolivie, la plupart des populations domiciliaires de *T. infestans* sont résistantes à la deltaméthrine (Lardeux et al. 2010). Des études spécifiques sont à conduire pour mettre en lien la présence d'ETs et la résistance aux insecticides.

Parmi les familles de gènes en expansion chez les espèces domiciliaires, nous avons annoté des nitrophorines, qui sont une catégorie de lipocalines salivaires spécifiques aux espèces du genre *Rhodnius*. Dans le génome de référence de *R. prolixus*, 14 nitrophorines ont été identifiées (Mesquita et al. 2015). Une étude fine serait à entreprendre sur ces gènes, comme celle conduite sur les récepteurs chimiosensoriels montrant des expansions chez les espèces domiciliaires.

### Les espèces domiciliaires montrent une sélection positive de certains récepteurs chimiosensoriels

Dans nos données, les GRs montrent une vitesse d'évolution, en terme de gain et perte de gènes très significativement différenciée entre les espèces domiciliaires et sylvatiques, puisque ces gènes évoluent 60 fois plus vite chez ces espèces. Par contre, aucun GR n'est sous sélection positive. En revanche, nous n'avons pas mis en évidence de différence de taux d'évolution en fonction de l'habitat pour les ORs. Mais les analyses de sélection montrent que 4 ORs, les OR1, OR38, OR102 et ORco, sont sous sélection positive chez les espèces domiciliaires et pourraient avoir un rôle dans l'adaptation aux anthroposystèmes. Il sera intéressant de voir si les mutations fixées par sélection ont conduit à un changement dans la reconnaissance des ligands pour ces récepteurs. Cependant des analyses structurales et/ou fonctionnelles sont nécessaires. Avec l'aide de telles études couplées à des études génomiques montrant une sélection positive sur 4 récepteurs gustatifs, il a été suggéré chez *D. yakuba mayottensis* une adaptation à un fruit toxique, le morinda *M. citrifolia*, par une diminution de la sensibilité à l'amertume de la plante hôte (Ferreira et al. 2021). Une telle modification de fonction, impliquant une sensibilité moindre à certains composés initialement répulsifs, peut être envisagée concernant par exemple la résistance aux insecticides chez les espèces domiciliaires.

Il est étonnant de trouver l'ORco parmi les gènes sous sélection chez les espèces domiciliaires car ORco est plutôt sous sélection purificatrice

chez les insectes, notamment chez les diptères (Paulo et al. 2021). Ce gène a un rôle primordial chez les insectes comme le montrent les études chez des mutants ORco (DeGennaro et al. 2013; Liu et al. 2017; Paulo et al. 2021; Fandino et al. 2019). Ces auteurs ont montré que la délétion d'ORco impacte des fonctions diverses telles que la recherche de nourriture, la reproduction et le vol dirigé par l'odeur. Il est possible que la modification de la séquence en acide aminé de la protéine ORco entraîne une altération de la perception des odeurs chez les espèces domiciliées entraînant un changement d'habitat.

La sélection positive sur ORco pour les espèces domiciliées montre le rôle majeur joué par ce récepteur pour la physiologie des espèces de *Rhodnius* dans ce contexte environnemental. Comme ORco est impliqué dans diverses fonctions physiologiques (Franco et al. 2016), il pourrait être une cible intéressante pour développer de nouvelles stratégies de contrôle des vecteurs.

En effet, les facteurs impliqués dans l'évolution des gènes chimiosensoriels sont multiples et comprennent la reconnaissance de l'hôte, la reproduction, la spécialisation de l'habitat mais aussi l'aversion aux insecticides. L'identification de ces facteurs et des gènes qui y répondent est d'une importance majeure dans la compréhension du processus de domiciliation et dans une perspective de lutte antivectorielle.



## PERSPECTIVES

---

### ÉTUDE D'AUTRES GÈNES IMPLIQUÉS DANS LA PERCEPTION

Grâce à l'avancée du séquençage à haut-débit et aux programmes pour assembler les génomes, il a été possible de séquencer et d'annoter les récepteurs de gènes chimiosensoriels ORs et GRs de 13 espèces de *Rhodnius*. De plus, l'annotation des OBPs et CSPs est en cours par notre équipe (Rignault et al, en préparation). Ces travaux ont été centrés autour des récepteurs chimiosensoriels mais d'autres gènes jouant un rôle dans la perception environnementale et pouvant être impliqués dans la spécialisation des espèces de *Rhodnius* à leur habitat, pourraient être étudiés en détail. Ainsi, l'étude de Marchant et al. (2016) a montré une sous-expression dans les organes chimiosensoriels des populations domiciliaires de *T. brasiliensis* de gènes d'OBPs et de CSPs mais également de gènes de protéines membranaires des neurones sensoriels (SNMPs), de canaux ioniques du potentiel (TRPs) et de protéines de détoxification comme le cytochrome P450. Dans notre étude, nous avons également démontré que ces gènes de détoxification connaissent une évolution rapide chez les espèces domiciliaires.

Ainsi, les protéines membranaires des neurones sensoriels (SNMPs) jouent également un rôle important dans la chimioréception des insectes. Les SNMPs des insectes sont composés de deux protéines contenant un domaine transmembranaire dans les neurones olfactifs des antennes et localisées dans les membranes des dendrites. La superfamille des canaux ioniques du potentiel (TRPs) est quant-à elle essentielle à la perception de l'environnement. Les TRPs sont impliqués dans la phototransduction, la thermosensation, la vision, l'hygrosensation, la mécanosensation et dans la chimiosensation. Un TRP appartenant à la famille des récepteurs vanilloïdes (TRPVs) pourrait être impliqué dans la haute sensibilité à la chaleur chez *R. prolixus* qui utilise ces capacités pour percevoir les hôtes à sang chaud (Zermoglio et al. 2015). Le glutathione-S-transferases et le cytochrome P450s seraient impliqués dans la détoxification et la résistance aux insecticides et sont surexprimés dans les populations d'insectes résistants aux pyréthrinoïdes (Nkya et al. 2013; Enayati et al. 2005; Marcombe et al. 2012).

Enfin, pour les gènes *takeout* exprimés dans les organes chimiosensoriels, il a été proposé qu'ils soient associés à la perception chimiosensorielle dans les systèmes gustatifs et olfactifs (Fujikawa et al. 2006) et qu'ils soient impliqués dans le cycle circadien en reliant le cycle temporel et les informations sur l'état de l'alimentation (Sarov-Blat et al. 2000; So et al. 2000; Meunier et al. 2007).

## ÉTUDES COMPORTEMENTALES CHEZ LES *RHODNIUS*

L'identification des récepteurs chimiosensoriels ORs et GRs couplée à l'écologie des *Rhodnius* permet de tester des hypothèses sur l'adaptation des vecteurs au milieu domiciliaire mais également au milieu sylvatique du fait de l'inféodation aux palmiers de certaines espèces. Cependant, l'étude de leur dynamique ou des traces de sélection qu'ils portent, n'est que la première étape dans la compréhension fine de ces mécanismes adaptatifs. L'attribution précise d'une fonction à un gène est un long chemin qui passe par la désorphanisation de ces gènes, c'est-à-dire l'identification des ligands donc des odeurs reconnus par les récepteurs chimiosensoriels.

### Désorphanisation des ORs et des GRs

Comme nous l'avons mentionné, concernant les vecteurs de la maladie de Chagas, une seule équipe a jusqu'à présent conduit une étude de désorphanisation en utilisant une méthode d'expression dans des ovocytes de xénope. Ceci a conduit à identifier pour un OR de *R. prolixus*, RproOR80, et 4 composés reconnus au sein d'un panel de 109 molécules odorantes testées et qui pourraient être utilisés comme répulsifs chez *R. prolixus* (Franco et al 2018). Une autre méthode de désorphanisation est largement utilisée chez les insectes (Montagné et al. (2012). Elle repose sur l'expression hétérologue chez la drosophile en utilisant des cellules olfactives de drosophiles "neurone vide" (c'est-à-dire qui n'expriment aucun OR endogène excepté ORco) dans lequel l'OR d'intérêt est exprimé. Les drosophiles génétiquement modifiées sont stimulées avec un panel de molécules odorantes et des enregistrements d'électrophysiologie extracellulaire sont effectués sur les sensilles olfactives antennaires des drosophiles (De Fouchier et al. 2017). Les trois gènes codant les OR1, OR38 et OR102 qui ont montré une sélection positive pour les espèces domiciliaires pourraient être testés par cette méthode. Le panel de stimulation sera composé de divers produits chimiques biologiquement pertinents pour les triatomines (revue dans Barrozo et al. 2016).

Les systèmes hétérologues avec des GRs sont plus compliqués à mettre en place chez la drosophile à cause de leur fonctionnement en combinatoire. Dans la méthode "drosophile neurone vide", il faudrait co-exprimer les GRs d'intérêt dans une sensille olfactive. Ce montage expérimental en système hétérologue a déjà été réalisé pour exprimer des GRs d'*Anopheles gambiae* dans (Freeman et al. 2014) ou des GRs de drosophile (dans ce cas, il s'agit d'une sur-expression ectopique) (Wisotsky et al. 2011). De plus, il est compliqué de connaître la ou les combinaison(s)

de GRs naturellement exprimée(s) par *R. prolixus*. Une difficulté supplémentaire est de rendre volatiles les molécules sapides en les vaporisant. Une autre approche serait de créer des mutants de délétion d'un GR candidat chez *R. prolixus* et de tester un panel de molécules. Par comparaison avec le phénotype sauvage, il serait possible de déterminer l'implication du GR dans la reconnaissance de chaque molécule sapide.

Parmi les GR d'intérêt identifiés dans notre étude, il serait intéressant de confirmer la sensibilité du GR1 au fructose. La sensibilité des GR5 et GR11 pourrait également être testée pour confirmer ou non leur reconnaissance de molécules amères, hypothèse formulée du fait de leur homologie avec des récepteurs déjà décrits (Mesquita et al. 2015).

## Tests de choix

Ces études de désorphanisation pourraient être complétées par des tests de choix. Nous pourrions utiliser un labyrinthe de Y comme cela a déjà été fait chez *R. prolixus* (e. g. Vinauger et al. 2011) pour tester si les espèces domiciliaires et sylvatiques réagissent différemment face à un panel de molécules volatiles du milieu domiciliaire comme des composés provenant d'humains et d'animaux du milieu péri-domiciliaire.

Nous pourrions aussi tester s'il y a une attirance différentielle des espèces domiciliaires et sylvatiques *Rhodnius* par des composés volatiles émis par les palmiers. Nous pourrions également tester l'attraction des espèces sylvatiques, pour lesquelles nous avons mis en évidence une ingestion de palmiers (Da Lage et al. soumis), par des odeurs de fruits ou de palmiers. Des tests de choix seraient ainsi réalisés entre odeur de fruit ou palmier et odeur de sang sur des individus soumis à différentes conditions (individus à jeun depuis longtemps ou récemment engorgés).

Par ailleurs, comme évoqué en discussion, le sucre pourrait être bénéfique pour la fitness des *Rhodnius*. Il serait possible d'utiliser le dispositif appelé MultiCAFE (pour *Multi Capillary Feeder*, alimenteur multi-capillaire) développé au laboratoire EGCE (Colombi, en préparation) à partir d'un dispositif plus simple, CAFE mis en place par (Ja et al. 2007). Le dispositif étant assez petit car développé pour des drosophiles, il faudrait l'agrandir afin de pouvoir y insérer des *Rhodnius* qui font environ 3,5 cm. Ce dispositif nous permettrait par exemple de tester la consommation de solution de sucre plus ou moins complexe comparée à la consommation d'eau afin de tester si les *Rhodnius* recherchent le sucre ou seulement à s'hydrater. D'autres sources alimentaires pourraient être testées comme du jus de baies d'acai (fruits de palmiers) ou de la sève de palmier afin de tester si les *Rhodnius* sont attirées par des composés végétaux et tenter d'apporter des éléments de réponse sur l'inféodation des *Rhodnius* aux

palmiers.

Cependant, ces études de choix portant sur différentes espèces de *Rhodnius* demandent en amont d'avoir à disposition un certain nombre d'espèces et d'individus et, si possible pour les espèces sylvatiques, des individus fraîchement échantillonnés pour s'affranchir d'une réponse conditionnée par un élevage en laboratoire.

## CONCLUSION

---

Approfondir les travaux de recherche sur la sensorialité chimiosensorielle des *Rhodnius* permettrait d'identifier les molécules reconnues par celles-ci. En connaissant les molécules reconnues et les comportements induits, il serait possible de réaliser des pièges sensoriels attractifs à disposer à distance des habitations humaines. Cependant, à ce jour, aucun piège appâté avec des substances odorantes synthétiques n'a été testé dans de tels essais sur le terrain. Les pièges attractifs doivent être pratiques, bon marché et efficaces afin d'être utilisés à grande échelle.

Par ailleurs, des répulsifs adaptés pourraient être utilisés si par exemple les récepteurs putativement sensibles à l'amer étaient fonctionnellement désorphanisés. Ainsi, l'usage de répulsifs corporels contenant des molécules amères pourrait être envisagé pour limiter les risques de piqûre. Des études seraient à conduire sur la prophylaxie d'ingestion de molécules amères comme l'ingestion de café ou la prise de quinine, car comme le soulignent [Pontes et al. \(2014\)](#), le pic de la caféine après une seule tasse de café dépasse la gamme de doses détectées par *R. prolixus*. Le développement de nouvelles stratégies ciblant le système gustatif des insectes hématophages, en utilisant des anti-appétants ou des composés amers, pourrait aider à diminuer les interactions hôte-vecteur et ainsi empêcher la transmission.

L'identification des composés volatiles ou sapides spécifiques et des gènes qui y répondent est d'une importance majeure dans la compréhension des mécanismes d'adaptation à l'habitat ainsi que dans une perspective de lutte antivectorielle.



## BIBLIOGRAPHIE

---

- Abad-Franch, F., F. S. Palomeque, H. M. Aguilar, et M. A. Miles. 2005. « Field Ecology of Sylvatic *Rhodnius* Populations (Heteroptera, Triatominae): Risk Factors for Palm Tree Infestation in Western Ecuador ». *Tropical Medicine & International Health: TM & IH* 10 (12): 1258-66. <https://doi.org/10.1111/j.1365-3156.2005.01511.x>.
- Abad-Franch, Fernando, Marli M. Lima, Otília Sarquis, Rodrigo Gurgel-Gonçalves, María Sánchez-Martín, José Calzada, Azael Saldaña, et al. 2015. « On palms, bugs, and Chagas disease in the Americas ». *Acta Tropica*, Ecology and diversity of *Trypanosoma cruzi*, 151 (novembre): 126-41. <https://doi.org/10.1016/j.actatropica.2015.07.005>.
- Abad-Franch, Fernando, Fernando A. Monteiro, Nicolás Jaramillo O., Rodrigo Gurgel-Gonçalves, Fernando Braga Stehling Dias, et Liléia Diotaiuti. 2009. « Ecology, Evolution, and the Long-Term Surveillance of Vector-Borne Chagas Disease: A Multi-Scale Appraisal of the Tribe Rhodniini (Triatominae) ». *Acta Tropica* 110 (2-3): 159-77. <https://doi.org/10.1016/j.actatropica.2008.06.005>.
- Abad-Franch, Fernando, Márcio G Pavan, Nicolás Jaramillo-O, Francisco S Palomeque, Carolina Dale, Duverney Chaverra, et Fernando A Monteiro. 2013. « *Rhodnius barretti*, a new species of Triatominae (Hemiptera: Reduviidae) from western Amazonia ». *Memórias do Instituto Oswaldo Cruz* 108 (Suppl 1): 92-99. <https://doi.org/10.1590/0074-0276130434>.
- Albert, James S., Pedro Val, et Carina Hoorn. 2018. « The changing course of the Amazon River in the Neogene: center stage for Neotropical diversification ». *Neotropical Ichthyology* 16 (3): e180033. <https://doi.org/10.1590/1982-0224-20180033>.
- Alevi, K. C. C., A. Ravazi, V. J. Mendonça, J. A. Rosa, et M. T. V. Azeredo-Oliveira. 2015. « Karyotype of *Rhodnius montenegrensis* (Hemiptera, Triatominae) ». *Genetics and Molecular Research: GMR* 14 (1): 222-26. <https://doi.org/10.4238/2015.January.16.5>.
- Alevi, Kaio Cesar Chaboli, Jader de Oliveira, Dayse da Silva Rocha, et Cleber Galvão. 2021. « Trends in Taxonomy of Chagas Disease Vectors (Hemiptera, Reduviidae, Triatominae): From Linnaean to Integrative Taxonomy ». *Pathogens* 10 (12): 1627. <https://doi.org/10.3390/pathogens10121627>.
- Almeida, Carlos Eduardo, Leslie Faucher, Morgane Lavina, Jane Costa, et Myriam Harry. 2016. « Molecular Individual-Based Approach on *Triatoma brasiliensis*: Inferences on Triatomine Foci, *Trypanosoma cruzi* Natural Infection Prevalence, Parasite Diversity and Feeding Sources ». *PLoS Neglected Tropical Diseases* 10 (2): e0004447. <https://doi.org/10.1371/journal.pntd.0004447>.
- Almeida, Flávio Barbosa de, Elcy Israel Santos, et Gigio Sposina. 1973. « Triatomíneos da Amazônia III. ». *Acta Amazonica* 3 (2): 43-46. <https://doi.org/10.1590/1809-43921973032043>.
- Almeida, Francisca C., Alejandro Sánchez-Gracia, Jose Luis Campos, et Julio Rozas. 2014. « Family Size Evolution in *Drosophila* Chemosensory Gene Families: A Comparative Analysis with a Critical Appraisal of Methods ». *Genome Biology and Evolution* 6 (7): 1669-82. <https://doi.org/10.1093/gbe/evu130>.
- Andersson, Martin N., Christopher I. Keeling, et Robert F. Mitchell. 2019. « Genomic content of chemosensory genes correlates with host range in wood-boring beetles (*Dendroctonus ponderosae*, *Agrilus planipennis*, and *Anoplophora glabripennis*) ». *BMC Genomics* 20 (1): 690. <https://doi.org/10.1186/s12864-019-6054-x>.
- Antonio-Campos, Alberto, Alejandra Nicolás-Cruz, José Ignacio Girón-Arias, Nancy Rivas, et Ricardo Alejandro-Aguilar. 2019. « Presence of *Rhodnius prolixus* Stål, 1859 (Hemiptera: Reduviidae) in Oaxaca, Mexico, Ten Years after the Certification of Its Elimination ». *Journal of Vector Ecology* 44 (2): 293-95. <https://doi.org/10.1111/jvec.12363>.
- Araújo, Adauto, Ana Maria Jansen, Karl Reinhard, et Luiz Fernando Ferreira. 2009. « Paleoparasitology of Chagas Disease: A Review ». *Memórias Do Instituto Oswaldo Cruz* 104 (suppl 1): 9-16. <https://doi.org/10.1590/S0074-02762009000900004>.
- Arias, Antonieta Rojas de, Carlota Monroy, Felipe Guhl, Sergio Sosa-Estani, Walter Souza Santos, et Fernando Abad-Franch. 2022. « Chagas Disease Control-Surveillance in the Americas: The Multinational Initiatives and the Practical Impossibility of Interrupting Vector-Borne *Trypanosoma cruzi* Transmission ». *Memórias Do Instituto Oswaldo Cruz* 117. <https://doi.org/10.1590/0074-02760210130>.

- Arias-Giraldo, Luisa M., Marina Muñoz, Carolina Hernández, Giovanni Herrera, Natalia Velásquez-Ortiz, Omar Cantillo-Barraza, Plutarco Urbano, Andrés Cuervo, et Juan David Ramírez. 2020. « Identification of Blood-Feeding Sources in *Panstrongylus*, *Psammolestes*, *Rhodnius* and *Triatoma* Using Amplicon-Based next-Generation Sequencing ». *Parasites & Vectors* 13 (1): 1-14. <https://doi.org/10.1186/s13071-020-04310-z>.
- Armisen, David, Rajendhran Rajakumar, Markus Friedrich, Joshua B. Benoit, Hugh M. Robertson, Kristen A. Panfilio, Seung-Joon Ahn, et al. 2018. « The genome of the water strider *Gerris buenoi* reveals expansions of gene repertoires associated with adaptations to life on the water ». *BMC Genomics* 19 (1): 832. <https://doi.org/10.1186/s12864-018-5163-2>.
- Attardo, Geoffrey M., Adly M. M. Abd-Alla, Alvaro Acosta-Serrano, James E. Allen, Rosemary Bateta, Joshua B. Benoit, Kostas Bourtzis, et al. 2019. « Comparative genomic analysis of six *Glossina* genomes, vectors of African trypanosomes ». *Genome Biology* 20 (1): 187. <https://doi.org/10.1186/s13059-019-1768-2>.
- Aufderheide, Arthur C., Wilmar Salo, Michael Madden, John Streitz, Jane Buikstra, Felipe Guhl, Bernardo Arriaza, et al. 2004. « A 9,000-Year Record of Chagas' Disease ». *Proceedings of the National Academy of Sciences* 101 (7): 2034-39. <https://doi.org/10.1073/pnas.0307312101>.
- Azevedo, Ana Paula S., Jardel C. Farias, Graciomar C. Costa, Susanne C. P. Ferreira, Walimir C. Aragão-Filho, Paulo R. A. Sousa, Mayara T. Pinheiro, et al. 2007. « Anti-Thrombotic Effect of Chronic Oral Treatment with *Orbiglyna phalerata* Mart ». *Journal of Ethnopharmacology* 111 (1): 155-59. <https://doi.org/10.1016/j.jep.2006.11.005>.
- Baines, S. 1956. « The Role of the Symbiotic Bacteria in the Nutrition of *Rhodnius prolixus* (Hemiptera) ». *Journal of Experimental Biology* 33 (3): 533-41. <https://doi.org/10.1242/jeb.33.3.533>.
- Barah, Pankaj, et Atle M. Bones. 2015. « Multidimensional approaches for studying plant defence against insects: from ecology to omics and synthetic biology ». *Journal of Experimental Botany* 66 (2): 479-93. <https://doi.org/10.1093/jxb/eru489>.
- Barnabé, Christian, Simone Frédérique Brenière, Jean-François Guégan, Denis Blanchet, et Christine Aznar. 2018. « Molecular Characterization of *Rhodnius robustus* Specimens, Potential Vectors for Chagas Disease in French Guiana, South America ». *Infection, Genetics and Evolution* 59 (avril): 28-31. <https://doi.org/10.1016/j.meegid.2018.01.019>.
- Barrozo, R. B., et C. R. Lazzari. 2006. « Orientation Response of Haematophagous Bugs to CO<sub>2</sub>: The Effect of the Temporal Structure of the Stimulus ». *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* 192 (8): 827-31. <https://doi.org/10.1007/s00359-006-0120-y>.
- Barrozo, Romina B., et Claudio R. Lazzari. 2004. « Orientation Behaviour of the Blood-sucking Bug *Triatoma infestans* to Short-chain Fatty Acids: Synergistic Effect of L-Lactic Acid and Carbon Dioxide ». *Chemical Senses* 29 (9): 833-41. <https://doi.org/10.1093/chemse/bjh249>.
- Barrozo, Romina B., Carolina E. Reisenman, Pablo Guerenstein, Claudio R. Lazzari, et Marcelo G. Lorenzo. 2017. « An inside Look at the Sensory Biology of Triatomines ». *Journal of Insect Physiology* 97 (février): 3-19. <https://doi.org/10.1016/j.jinsphys.2016.11.003>.
- Bautista, Ma Anita M, Binny Bhandary, Asela J Wijeratne, Andrew P Michel, Casey W Hoy, et Omprakash Mittapalli. 2015. « Evidence for Trade-Offs in Detoxification and Chemosensation Gene Signatures in *Plutella xylostella* ». *Pest Management Science* 71 (3): 423-32. <https://doi.org/10.1002/ps.3822>.
- Benjamini, Yoav, et Yosef Hochberg. 1995. « Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing ». *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1): 289-300.
- Benton, Richard, Silke Sachse, Stephen W. Michnick, et Leslie B. Vosshall. 2006. « Atypical Membrane Topology and Heteromeric Function of *Drosophila* Odorant Receptors In Vivo ». *PLOS Biology* 4 (2): e20. <https://doi.org/10.1371/journal.pbio.0040020>.
- Bérenger, Jean-Michel, Dominique Pluot-Sigwalt, Frédéric Pagès, Denis Blanchet, et Christine Aznar. 2009. « The triatominae species of French Guiana (Heteroptera: Reduviidae) ». *Memórias do Instituto Oswaldo Cruz* 104 (8): 1111-16. <https://doi.org/10.1590/S0074-02762009000800007>.
- Bérenger, J-M, et D Pluot-Sigwalt. 2002. « *Rhodnius amazonicus* Almeida, Santos & Sposina, 1973, Bona Species, Close to *R. pictipes* Stål, 1872 (Heteroptera, Reduviidae, Triatominae) ». *Memórias do Instituto Oswaldo Cruz* 97 (1): 73-77. <https://doi.org/10.1590/S0074-02762002000100011>.
- Berlocher, Stewart H., et Jeffrey L. Feder. 2002. « Sympatric Speciation in Phytophagous Insects: Moving Beyond Controversy? » *Annual Review of Entomology* 47 (1): 773-815. <https://doi.org/10.1146/annurev.ento.47.091201.145312>.



- Bernt, Matthias, Alexander Donath, Frank Jühling, Fabian Externbrink, Catherine Florentz, Guido Fritsch, Joern Pütz, Martin Middendorf, et Peter F. Stadler. 2013. « MITOS: Improved de Novo Metazoan Mitochondrial Genome Annotation ». *Molecular Phylogenetics and Evolution* 69 (2): 313-19. <https://doi.org/10.1016/j.ympev.2012.08.023>.
- Bilheiro, Adriana Benatti, João Aristeu da Rosa, Jader de Oliveira, Tiago Belintani, Gilberto Fontes, Jansen Fernandes Medeiros, Dionatas Ulises de Oliveira Meneguetti, et Luís Marcelo Aranha Camargo. 2018. « First Report of Natural Infection with *Trypanosoma cruzi* in *Rhodnius montenegrensis* (Hemiptera, Reduviidae, Triatominae) in Western Amazon, Brazil ». *Vector Borne and Zoonotic Diseases (Larchmont, N.Y.)* 18 (11): 605-10. <https://doi.org/10.1089/vbz.2018.2266>.
- Bouckaert, Remco, Joseph Heled, Denise Kühnert, Tim Vaughan, Chieh-Hsi Wu, Dong Xie, Marc A. Suchard, Andrew Rambaut, et Alexei J. Drummond. 2014. « BEAST 2: A Software Platform for Bayesian Evolutionary Analysis ». *PLOS Computational Biology* 10 (4): e1003537. <https://doi.org/10.1371/journal.pcbi.1003537>.
- Bourgeois, Yann, et Stéphane Boissinot. 2019. « On the Population Dynamics of Junk: A Review on the Population Genomics of Transposable Elements ». *Genes* 10 (mai): 419. <https://doi.org/10.3390/genes10060419>.
- Brito, Raíssa N., Juliana A. Geraldo, Fernando A. Monteiro, Cristiano Lazoski, Rita C. M. Souza, et Fernando Abad-Franch. 2019. « Transcriptome-Based Molecular Systematics: *Rhodnius Montenegrensis* (Triatominae) and Its Position within the *Rhodnius prolixus-Rhodnius robustus* Cryptic-Species Complex ». *Parasites & Vectors* 12 (1): 305. <https://doi.org/10.1186/s13071-019-3558-9>.
- Brockhurst, Michael A., Ellie Harrison, James P.J. Hall, Thomas Richards, Alan McNally, et Craig MacLean. 2019. « The Ecology and Evolution of Pangenomes ». *Current Biology* 29 (20): R1094-1103. <https://doi.org/10.1016/j.cub.2019.08.012>.
- Brůna, Tomáš, Katharina J Hoff, Alexandre Lomsadze, Mario Stanke, et Mark Borodovsky. 2021. « BRAKER2: Automatic Eukaryotic Genome Annotation with GeneMark-EP+ and AUGUSTUS Supported by a Protein Database ». *NAR Genomics and Bioinformatics* 3 (1): lqaa108. <https://doi.org/10.1093/nargab/lqaa108>.
- Bruneto, Eduardo G., Miguel M. Fernandes-Silva, Cristina Toledo-Cornell, Silvia Martins, João M. B. Ferreira, Valeria R. Corrêa, Joicyly M. da Costa, et al. 2021. « Case-Fatality From Orally-Transmitted Acute Chagas Disease: A Systematic Review and Meta-Analysis ». *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 72 (6): 1084-92. <https://doi.org/10.1093/cid/ciaa1148>.
- Bruyne, Marien de, et Coral G. Warr. 2006. « Molecular and Cellular Organization of Insect Chemosensory Neurons ». *BioEssays* 28 (1): 23-34. <https://doi.org/10.1002/bies.20338>.
- Butterwick, Joel A., Josefina Del Mármol, Kelly H. Kim, Martha A. Kahlson, Jackson A. Rogow, Thomas Walz, et Vanessa Ruta. 2018. « Cryo-EM Structure of the Insect Olfactory Receptor Orco ». *Nature* 560 (7719): 447-52. <https://doi.org/10.1038/s41586-018-0420-8>.
- Calderón, Johan M., et Camila González. 2020. « Co-occurrence or dependence? Using spatial analyses to explore the interaction between palms and *Rhodnius* triatomines ». *Parasites & Vectors* 13 (avril). <https://doi.org/10.1186/s13071-020-04088-0>.
- Calvo, Eric, Ben J. Mans, John F. Andersen, et José M. C. Ribeiro. 2006. « Function and Evolution of a Mosquito Salivary Protein Family ». *Journal of Biological Chemistry* 281 (4): 1935-42. <https://doi.org/10.1074/jbc.M510359200>.
- Campillo, Luke C., Kevin J. Burns, Robert G. Moyle, et Joseph D. Manthey. 2019. « Mitochondrial Genomes of the Bird Genus *Piranga*: Rates of Sequence Evolution, and Discordance between Mitochondrial and Nuclear Markers ». *Mitochondrial DNA Part B* 4 (2): 2566-69. <https://doi.org/10.1080/23802359.2019.1637286>.
- Carcavallo, R., et P. Barreto. 1976. « A new species of *Rhodnius* in Colombia. » *Boletín de la Dirección de Malariología y Saneamiento Ambiental* 16 (2): 176-83.
- Carcavallo, R. U., M. E. F. Rodríguez, R. Salvatella, S. I. C. Casas, I. S. Sherlock, C. Galvão, D. S. Rocha, et al. 1998. « Atlas of Chagas Disease Vectors in the Americas ». *Atlas of Chagas Disease Vectors in the Americas*, 1109-1109.
- Carcavallo, R. U., R. J. Tonn, et J. C. Jimenez. 1976. « Notes on the biology, ecology and geographical distribution of *Rhodnius neivai* (Hemiptera, Reduviidae). » *Boletín de la Dirección de Malariología y Saneamiento Ambiental* 16 (2): 169-71.
- Cardinal, Marta V., Marta A. Lauricella, Leonardo A. Ceballos, Leonardo Lanati, Paula L. Marcet, Mariano J. Levin, Uriel Kitron, Ricardo E. Gürtler, et Alejandro G. Schijman. 2008. « Molecular Epidemiology of Domestic and Sylvatic *Trypanosoma cruzi* Infection in Rural Northwestern Argentina ». *International Journal for Parasitology* 38 (13): 1533-43. <https://doi.org/10.1016/j.ijpara.2008.04.010>.

- Cardoso-Jaime, Victor, Nichole A Broderick, et Krystal Maya-Maldonado. 2022. « Metal Ions in Insect Reproduction: A Crosstalk between Reproductive Physiology and Immunity ». *Current Opinion in Insect Science* 52 (août): 100924. <https://doi.org/10.1016/j.cois.2022.100924>.
- Carod-Artal, Francisco Javier. 2013. « American Trypanosomiasis ». In *Handbook of Clinical Neurology*, 114:103-23. Elsevier. <https://doi.org/10.1016/B978-0-444-53490-3.00007-8>.
- Carvalho, Danila Blanco de, Carlos Congrains, Samira Chahad-Ehlers, Heloisa Pinotti, Reinaldo Alves de Brito, et João Aristeu da Rosa. 2017. « Differential Transcriptome Analysis Supports *Rhodnius montenegrensis* and *Rhodnius robustus* (Hemiptera, Reduviidae, Triatominae) as Distinct Species ». *PLOS ONE* 12 (4): e0174997. <https://doi.org/10.1371/journal.pone.0174997>.
- Castillo-Riquelme, Marianela, Felipe Guhl, Brenda Turriago, Nestor Pinto, Fernando Rosas, Mónica Flórez Martínez, Julia Fox-Rushby, Clive Davies, et Diarmid Campbell-Lendrum. 2008. « The Costs of Preventing and Treating Chagas Disease in Colombia ». *PLoS Neglected Tropical Diseases* 2 (11): e336. <https://doi.org/10.1371/journal.pntd.0000336>.
- Castro, Marcelo R. J., Clément Goubert, Fernando A. Monteiro, Cristina Vieira, et Claudia M. A. Carareto. 2020. « Homology-Free Detection of Transposable Elements Unveils Their Dynamics in Three Ecologically Distinct *Rhodnius* Species ». *Genes* 11 (2): 170. <https://doi.org/10.3390/genes11020170>.
- Chagas, Carlos. 1909. « Nova tripanozomíaze humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem ». *Memórias do Instituto Oswaldo Cruz* 1 (2): 159-218. <https://doi.org/10.1590/S0074-02761909000200008>.
- Cheng, Daifeng, Yongyue Lu, Ling Zeng, Guangwen Liang, et Xiaofang He. 2015. « Si-CSP9 Regulates the Integument and Moulting Process of Larvae in the Red Imported Fire Ant, *Solenopsis invicta* ». *Scientific Reports* 5 (1): 9245. <https://doi.org/10.1038/srep09245>.
- Chikhi, R., et P. Medvedev. 2014. « Informed and Automated K-Mer Size Selection for Genome Assembly ». *Bioinformatics* 30 (1): 31-37. <https://doi.org/10.1093/bioinformatics/btt310>.
- Clyne, P. J., C. G. Warr, et J. R. Carlson. 2000. « Candidate Taste Receptors in *Drosophila* ». *Science (New York, N.Y.)* 287 (5459): 1830-34.
- Clyne, P. J., C. G. Warr, M. R. Freeman, D. Lessing, J. Kim, et J. R. Carlson. 1999. « A Novel Family of Divergent Seven-Transmembrane Proteins: Candidate Odorant Receptors in *Drosophila* ». *Neuron* 22 (2): 327-38. [https://doi.org/10.1016/s0896-6273\(00\)81093-4](https://doi.org/10.1016/s0896-6273(00)81093-4).
- Cortez, Mauro, Marcelo R. Silva, Ivan Neira, Daniele Ferreira, Gisela R. S. Sasso, Alejandro O. Luquetti, Anis Rassi, et Nobuko Yoshida. 2006. « *Trypanosoma cruzi* Surface Molecule Gp90 Downregulates Invasion of Gastric Mucosal Epithelium in Orally Infected Mice ». *Microbes and Infection* 8 (1): 36-44. <https://doi.org/10.1016/j.micinf.2005.05.016>.
- Covarrubias, Charles, Mauro Cortez, Daniele Ferreira, et Nobuko Yoshida. 2007. « Interaction with Host Factors Exacerbates *Trypanosoma cruzi* Cell Invasion Capacity upon Oral Infection ». *International Journal for Parasitology* 37 (14): 1609-16. <https://doi.org/10.1016/j.ijpara.2007.05.013>.
- Dahan, Romain A., Rebecca P. Duncan, Alex CC Wilson, et Liliانا M. Dávalos. 2015. « Amino acid transporter expansions associated with the evolution of obligate endosymbiosis in sap-feeding insects (Hemiptera: Sternorrhyncha) ». *BMC Evolutionary Biology* 15 (1): 52. <https://doi.org/10.1186/s12862-015-0315-3>.
- De Paula, Alexandre Silva, Liléia Diotaiuti, et Cleber Galvão. 2007. « Systematics and Biogeography of Rhodniini (Heteroptera: Reduviidae: Triatominae) Based on 16S Mitochondrial rDNA Sequences ». *Journal of Biogeography* 34 (4): 699-712. <https://doi.org/10.1111/j.1365-2699.2006.01628.x>.
- DeGennaro, Matthew, Carolyn S. McBride, Laura Seeholzer, Takao Nakagawa, Emily J. Dennis, Chloe Goldman, Nijole Jasinskiene, Anthony A. James, et Leslie B. Vosshall. 2013. « Orco Mutant Mosquitoes Lose Strong Preference for Humans and Are Not Repelled by Volatile DEET ». *Nature* 498 (7455): 487-91. <https://doi.org/10.1038/nature12206>.
- Degnan, James H., et Noah A. Rosenberg. 2009. « Gene Tree Discordance, Phylogenetic Inference and the Multispecies Coalescent ». *Trends in Ecology & Evolution* 24 (6): 332-40. <https://doi.org/10.1016/j.tree.2009.01.009>.
- Dias, Fernando Braga Stehling, Cláudia Mendonça Bezerra, Evandro Marques de Menezes Machado, Cláudio Casanova, et Liléia Diotaiuti. 2008. « Ecological Aspects of *Rhodnius nasutus* Stål, 1859 (Hemiptera: Reduviidae: Triatominae) in Palms of the Chapada Do Araripe in Ceará, Brazil ». *Memórias Do Instituto Oswaldo Cruz* 103 (décembre): 824-30. <https://doi.org/10.1590/S0074-02762008000800014>.

- Dias, Fernando Braga Stehling, Alexandre Silva de Paula, Carlota Josefovicz Belisário, Marcelo Gustavo Lorenzo, Cláudia Mendonça Bezerra, Myriam Harry, et Liléia Diotaiuti. 2011. « Influence of the Palm Tree Species on the Variability of *Rhodnius nasutus* Stål, 1859 (Hemiptera, Reduviidae, Triatominae) ». *Infection, Genetics and Evolution* 11 (5): 869-77. <https://doi.org/10.1016/j.meegid.2011.02.008>.
- Dias, Fernando Braga Stehling, Marion Quartier, Liléia Diotaiuti, Guy Mejía, Myriam Harry, Anna Carolina Lustosa Lima, Robert Davidson, Frédéric Mertens, Marc Lucotte, et Christine A. Romaña. 2014. « Ecology of *Rhodnius robustus* Larrousse, 1927 (Hemiptera, Reduviidae, Triatominae) in *Attalea* Palm Trees of the Tapajós River Region (Pará State, Brazilian Amazon) ». *Parasites & Vectors* 7 (avril): 154. <https://doi.org/10.1186/1756-3305-7-154>.
- Dias, Fernando Braga Stehling, Marion Quartier, Christine A. Romaña, Liléia Diotaiuti, et Myriam Harry. 2010. « Tamandua *Tetradactyla* Linnaeus, 1758 (Myrmecophagidae) and *Rhodnius robustus* Larrousse, 1927 (Triatominae) Infection Focus by *Trypanosoma rangeli* Tejera, 1920 (Trypanosomatidae) in *Attalea phalerata* Mart. Ex Spreng (Arecaceae) Palm Tree in the Brazilian Amazon ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 10 (8): 1278-81. <https://doi.org/10.1016/j.meegid.2010.06.020>.
- Díaz, Sebastián, Francisco Panzera, Nicolás Jaramillo-O, Ruben Pérez, Rosina Fernández, Gustavo Vallejo, Azael Saldaña, Jose E. Calzada, Omar Triana, et Andrés Gómez-Palacio. 2014. « Genetic, Cytogenetic and Morphological Trends in the Evolution of the *Rhodnius* (Triatominae: Rhodniini) Trans-Andean Group ». Édité par Pedro Lagerblad Oliveira. *PLoS ONE* 9 (2): e87493. <https://doi.org/10.1371/journal.pone.0087493>.
- Díaz-Albiter, Hector Manuel, Tainá Neves Ferreira, Samara Graciane Costa, Gustavo Bueno Rivas, Marcia Gumiel, Danilo Rufino Cavalcante, Márcio Galvão Pavan, et al. 2016. « Everybody Loves Sugar: First Report of Plant Feeding in Triatomines ». *Parasites & Vectors* 9 (1). <https://doi.org/10.1186/s13071-016-1401-0>.
- Dobritsa, Anna A., Wynand van der Goes van Naters, Coral G. Warr, R. Alexander Steinbrecht, et John R. Carlson. 2003. « Integrating the Molecular and Cellular Basis of Odor Coding in the *Drosophila* Antenna ». *Neuron* 37 (5): 827-41. [https://doi.org/10.1016/s0896-6273\(03\)00094-1](https://doi.org/10.1016/s0896-6273(03)00094-1).
- Dotson, E. M., et C. B. Beard. 2001. « Sequence and Organization of the Mitochondrial Genome of the Chagas Disease Vector, *Triatoma dimidiata* ». *Insect Molecular Biology* 10 (3): 205-15. <https://doi.org/10.1046/j.1365-2583.2001.00258.x>.
- Drummond, Alexei J., Simon Y. W. Ho, Matthew J. Phillips, et Andrew Rambaut. 2006. « Relaxed Phylogenetics and Dating with Confidence ». Édité par David Penny. *PLoS Biology* 4 (5): e88. <https://doi.org/10.1371/journal.pbio.0040088>.
- Dujardin, J. P., M. Muñoz, T. Chavez, C. Ponce, J. Moreno, et C. J. Schofield. 1998. « The Origin of *Rhodnius prolixus* in Central America ». *Medical and Veterinary Entomology* 12 (1): 113-15. <https://doi.org/10.1046/j.1365-2915.1998.00092.x>.
- Dujardin, J. P., C. J. Schofield, et M. Tibayrenc. 1998. « Population Structure of Andean *Triatoma infestans*: Allozyme Frequencies and Their Epidemiological Relevance ». *Medical and Veterinary Entomology* 12 (1): 20-29. <https://doi.org/10.1046/j.1365-2915.1998.00076.x>.
- Durán, Pamela, Edda Siñani, et Stéphanie Depickère. 2016. « On Triatomines, Cockroaches and Haemolymphagy under Laboratory Conditions: New Discoveries ». *Memórias Do Instituto Oswaldo Cruz* 111 (octobre): 605-13. <https://doi.org/10.1590/0074-02760160027>.
- Duvaux, Ludovic, Quentin Geissmann, Karim Gharbi, Jing-Jiang Zhou, Julia Ferrari, Carole M. Smadja, et Roger K. Butlin. 2015. « Dynamics of Copy Number Variation in Host Races of the Pea Aphid ». *Molecular Biology and Evolution* 32 (1): 63-80. <https://doi.org/10.1093/molbev/msu266>.
- Dworkin, Ian, et Greg Gibson. 2006. « Epidermal Growth Factor Receptor and Transforming Growth Factor- $\beta$  Signaling Contributes to Variation for Wing Shape in *Drosophila melanogaster* ». *Genetics* 173 (3): 1417-31. <https://doi.org/10.1534/genetics.105.053868>.
- El Ghouzzi, Marie-Hélène, Elisabeth Boiret, Françoise Wind, Claudine Brochard, Sébastien Fittere, Luc Paris, Dominique Mazier, Nicole Sansonetti, et Philippe Bierling. 2010. « Testing Blood Donors for Chagas Disease in the Paris Area, France: First Results after 18 Months of Screening ». *Transfusion* 50 (3): 575-83. <https://doi.org/10.1111/j.1537-2995.2009.02476.x>.
- Elmore, T., et D. P. Smith. 2001. « Putative *Drosophila* Odor Receptor OR43b Localizes to Dendrites of Olfactory Neurons ». *Insect Biochemistry and Molecular Biology* 31 (8): 791-98. [https://doi.org/10.1016/s0965-1748\(00\)00184-3](https://doi.org/10.1016/s0965-1748(00)00184-3).
- Emerson, Brent C., et Rosemary G. Gillespie. 2008. « Phylogenetic Analysis of Community Assembly and Structure over Space and Time ». *Trends in Ecology & Evolution* 23 (11): 619-30. <https://doi.org/10.1016/j.tree.2008.07.005>.

- Enayati, A. A., H. Ranson, et J. Hemingway. 2005. « Insect Glutathione Transferases and Insecticide Resistance ». *Insect Molecular Biology* 14 (1): 3-8. <https://doi.org/10.1111/j.1365-2583.2004.00529.x>.
- Fandino, Richard A., Alexander Haverkamp, Sonja Bisch-Knaden, Jin Zhang, Sascha Bucks, Tu Anh Thi Nguyen, Katrin Schröder, et al. 2019. « Mutagenesis of odorant coreceptor Orco fully disrupts foraging but not oviposition behaviors in the hawkmoth *Manduca sexta* ». *Proceedings of the National Academy of Sciences* 116 (31): 15677-85. <https://doi.org/10.1073/pnas.1902089116>.
- Feliciangeli, M. Dora, Maria Sanchez-Martin, Rosalba Marrero, Clive Davies, et Jean-Pierre Dujardin. 2007. « Morphometric Evidence for a Possible Role of *Rhodnius prolixus* from Palm Trees in House Re-Infestation in the State of Barinas (Venezuela) ». *Acta Tropica* 101 (2): 169-77. <https://doi.org/10.1016/j.actatropica.2006.12.010>.
- Ferreira, Erina. 2022. « Bases génomiques et mécanismes évolutifs d'un cas récurrent de spéciation écologique chez la drosophile ». <http://www.theses.fr>. 9 décembre 2022. <http://www.theses.fr/s229031>.
- Ferreira, Erina A., Sophia Lambert, Thibault Verrier, Frédéric Marion-Poll, et Amir Yassin. 2021. « Soft Selective Sweep on Chemosensory Genes Correlates with Ancestral Preference for Toxic Noni in a Specialist *Drosophila* Population ». *Genes* 12 (1): 32. <https://doi.org/10.3390/genes12010032>.
- Ferreira, Renata Trotta Barroso, Maria Luiza Cabral, Ronald Sodré Martins, Paula Finamore Araujo, Sérgio Alves da Silva, Constança Britto, Maria Regina Branquinho, Paola Cardarelli-Leite, et Otacilio C. Moreira. 2018. « Detection and Genotyping of *Trypanosoma cruzi* from Açai Products Commercialized in Rio de Janeiro and Pará, Brazil ». *Parasites & Vectors* 11 (1): 233. <https://doi.org/10.1186/s13071-018-2699-6>.
- Feschotte, Cédric, et Clément Gilbert. 2012. « Endogenous viruses: insights into viral evolution and impact on host biology ». *Nature Reviews Genetics* 13 (4): 283-96.
- Filée, Jonathan, Kenny Agésilas-Lequeux, Laurie Lacquehay, Jean Michel Bérenger, Lise Dupont, Vagner Mendonça, João Aristeu da Rosa, et Myriam Harry. 2023. « *Wolbachia* Genomics Reveals a Potential for a Nutrition-Based Symbiosis in Blood-Sucking Triatomine Bugs ». *bioRxiv*. <https://doi.org/10.1101/2022.09.06.506778>.
- Filée, Jonathan, Marie Merle, Héloïse Bastide, Florence Mougel, Jean-Michel Bérenger, Elaine Folly-Ramos, Carlos Eduardo Almeida, et Myriam Harry. 2022. « Phylogenomics for Chagas Disease Vectors of the *Rhodnius* Genus (Hemiptera, Triatominae): What We Learn From Mito-Nuclear Conflicts and Recommendations ». *Frontiers in Ecology and Evolution* 9. <https://www.frontiersin.org/article/10.3389/fevo.2021.750317>.
- Fitzpatrick, Sinead, Maria Dora Feliciangeli, Maria J. Sanchez-Martin, Fernando A. Monteiro, et Michael A. Miles. 2008. « Molecular Genetics Reveal That Silvatic *Rhodnius prolixus* Do Colonise Rural Houses ». Édité par Ricardo E. Gurtler. *PLoS Neglected Tropical Diseases* 2 (4): e210. <https://doi.org/10.1371/journal.pntd.0000210>.
- Folly-Ramos, Elaine, L. Lynnette Dornak, Guilherme Orsolon, Teresa Cristina Monte Gonçalves, Mauricio Liliuso, Jane Costa, et Carlos Eduardo Almeida. 2016. « Vector Capacity of Members of *Triatoma brasiliensis* Species Complex: The Need to Extend Chagas Disease Surveillance to *Triatoma melanica* ». *Journal of Vector Ecology* 41 (1): 48-54. <https://doi.org/10.1111/jvec.12193>.
- Fouchier, Arthur de, William B. Walker, Nicolas Montagné, Claudia Steiner, Muhammad Binyameen, Fredrik Schlyter, Thomas Chertemps, et al. 2017. « Functional evolution of Lepidoptera olfactory receptors revealed by deorphanization of a moth repertoire ». *Nature Communications* 8 (juin): 15709. <https://doi.org/10.1038/ncomms15709>.
- Franco, Thiago A., Daniele S. Oliveira, Monica F. Moreira, Walter S. Leal, et Ana C. A. Melo. 2016. « Silencing the odorant receptor co-receptor RproOrco affects the physiology and behavior of the Chagas disease vector *Rhodnius prolixus* ». *Insect Biochemistry and Molecular Biology*, Special issue: *Rhodnius prolixus* genome, 69 (février): 82-90. <https://doi.org/10.1016/j.ibmb.2015.02.012>.
- Franco, Thiago A., Pingxi Xu, Nathália F. Brito, Daniele S. Oliveira, Xiaolan Wen, Monica F. Moreira, C. Rikard Unelius, Walter S. Leal, et Ana C. A. Melo. 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 biochemistry and molecular biology* 103 (décembre): 46-52. <https://doi.org/10.1016/j.ibmb.2018.10.004>.
- Freeman, Erica Gene, Zev Wisotsky, et Anupama Dahanukar. 2014. « Detection of sweet tastants by a conserved group of insect gustatory receptors ». *Proceedings of the National Academy of Sciences of the United States of America* 111 (4): 1598-1603. <https://doi.org/10.1073/pnas.1311724111>.
- Freitas, Lucas, et Mariana F. Nery. 2020. « Expansions and contractions in gene families of independently-evolved blood-feeding insects ». *BMC Evolutionary Biology* 20 (1): 87. <https://doi.org/10.1186/s12862-020-01650-3>.

- Frias, Daniel, et Juan Atria. 1998. « Chromosomal Variation, Macroevolution and Possible Parapatric Speciation in *Mepraia Spinolai* (Porter) (Hemiptera: Reduviidae) ». *Genetics and Molecular Biology* 21 (juin): 179-84. <https://doi.org/10.1590/S1415-47571998000200002>.
- Fritz, M. L., E. D. Walker, J. R. Miller, D. W. Severson, et I. Dworkin. 2015. « Divergent Host Preferences of Above- and below-Ground *Culex pipiens* Mosquitoes and Their Hybrid Offspring ». *Medical and Veterinary Entomology* 29 (2): 115-23. <https://doi.org/10.1111/mve.12096>.
- Fujikawa, Kazuyo, Keiji Seno, et Mamiko Ozaki. 2006. « A Novel Takeout-like Protein Expressed in the Taste and Olfactory Organs of the Blowfly, *Phormia regina* ». *FEBS Journal* 273 (18): 4311-21. <https://doi.org/10.1111/j.1742-4658.2006.05422.x>.
- Galvão, Cleber, et Sílvia A. Justi. 2015. « An Overview on the Ecology of Triatominae (Hemiptera:Reduviidae) ». *Acta Tropica, Ecology and diversity of *Trypanosoma cruzi**, 151 (novembre): 116-25. <https://doi.org/10.1016/j.actatropica.2015.06.006>.
- Gao, Fangluan, Chengjie Chen, Daej A. Arab, Zhenguo Du, Yehua He, et Simon Y. W. Ho. 2019. « EasyCodeML: A Visual Tool for Analysis of Selection Using CodeML ». *Ecology and Evolution* 9 (7): 3891-98. <https://doi.org/10.1002/ece3.5015>.
- Gao, Q., et A. Chess. 1999. « Identification of Candidate *Drosophila* Olfactory Receptors from Genomic DNA Sequence ». *Genomics* 60 (1): 31-39. <https://doi.org/10.1006/geno.1999.5894>.
- Gardiner, Anastasia, Daniel Barker, Roger K. Butlin, William C. Jordan, et Michael G. Ritchie. 2008. « *Drosophila* Chemoreceptor Gene Evolution: Selection, Specialization and Genome Size ». *Molecular Ecology* 17 (7): 1648-57. <https://doi.org/10.1111/j.1365-294X.2008.03713.x>.
- Garvin, Michael R., Joseph P. Bielawski, Leonid A. Sazanov, et Anthony J. Gharrett. 2015. « Review and Meta-Analysis of Natural Selection in Mitochondrial Complex I in Metazoans ». *Journal of Zoological Systematics and Evolutionary Research* 53 (1): 1-17. <https://doi.org/10.1111/jzs.12079>.
- Gaunt, Michael, et Michael Miles. 2000. « The Ecotopes and Evolution of Triatomine Bugs (Triatominae) and Their Associated Trypanosomes ». *Memórias Do Instituto Oswaldo Cruz* 95 (4): 557-65. <https://doi.org/10.1590/S0074-02762000000400019>.
- Gaunt, Michael W., et Michael A. Miles. 2002. « An Insect Molecular Clock Dates the Origin of the Insects and Accords with Palaeontological and Biogeographic Landmarks ». *Molecular Biology and Evolution* 19 (5): 748-61. <https://doi.org/10.1093/oxfordjournals.molbev.a004133>.
- Gittleman, John L., et Mark Kot. 1990. « Adaptation: Statistics and a Null Model for Estimating Phylogenetic Effects ». *Systematic Zoology* 39 (3): 227. <https://doi.org/10.2307/2992183>.
- Gómez-Palacio, Andrés, Nicolás Jaramillo-O., Harling Caro-Riaño, Sebastián Díaz, Fernando A. Monteiro, Ruben Pérez, Francisco Panzera, et Omar Triana. 2012. « Morphometric and Molecular Evidence of Intraspecific Biogeographical Differentiation of *Rhodnius pallescens* (HEMIPTERA: REDUVIIDAE: RHODNIINI) from Colombia and Panama ». *Infection, Genetics and Evolution* 12 (8): 1975-83. <https://doi.org/10.1016/j.meegid.2012.04.003>.
- Gonçalves, Leilane O, M de Oliveira Luciana, C D'Ávila Pessoa Grasielle, Aline CL Rosa, Marinely G Bustamante, Carlota J Belisário, Daniela M Resende, Lileia G Diotaiuti, et Jeronimo C Ruiz. 2017. « Insights from tissue-specific transcriptome sequencing analysis of *Triatoma infestans* ». *Memórias do Instituto Oswaldo Cruz* 112 (6): 456-57. <https://doi.org/10.1590/0074-02760160473>.
- González Valdivieso, F., B. Sánchez Díaz, et F. Nocerino. 1971. « Susceptibility of *R. prolixus* to chlorinated insecticides in Venezuela. » *Boletín de la Dirección de Malariología y Saneamiento Ambiental* 11 (1/2): 47-52.
- Good, Jeffrey M., Dan Vanderpool, Sara Keeble, et Ke Bi. 2015. « Negligible Nuclear Introgression despite Complete Mitochondrial Capture between Two Species of Chipmunks: INTROGRESSION IN CHIPMUNKS ». *Evolution* 69 (8): 1961-72. <https://doi.org/10.1111/evo.12712>.
- Gorla, D. E., J. P. Dujardin, et C. J. Schofield. 1997. « Biosystematics of Old World Triatominae ». *Acta Tropica* 63 (2): 127-40. [https://doi.org/10.1016/S0001-706X\(97\)87188-4](https://doi.org/10.1016/S0001-706X(97)87188-4).
- Gourbière, S, P Dorn, F Tripet, et E Dumonteil. 2012. « Genetics and evolution of triatomines: from phylogeny to vector control ». *Heredity* 108 (3): 190-202. <https://doi.org/10.1038/hdy.2011.71>.

- Green, Richard E., Johannes Krause, Adrian W. Briggs, Tomislav Maricic, Udo Stenzel, Martin Kircher, Nick Patterson, et al. 2010. « A Draft Sequence of the Neandertal Genome ». *Science* 328 (5979): 710-22. <https://doi.org/10.1126/science.1188021>.
- Guerenstein, P., et P. Guerin. 2001. « Olfactory and behavioural responses of the blood-sucking bug *Triatoma infestans* to odours of vertebrate hosts ». *Journal of Experimental Biology* 204 (3): 585-97. <https://doi.org/10.1242/jeb.204.3.585>.
- Guhl, F. 2017. « Chagas Disease in Pre-Colombian Civilizations ». In *American Trypanosomiasis Chagas Disease*, 23-46. Elsevier. <https://doi.org/10.1016/B978-0-12-801029-7.00002-2>.
- Guhl, Felipe, Nestor Pinto, et Germán Aguilera. 2009. « Sylvatic Triatominae: A New Challenge in Vector Control Transmission ». *Memórias Do Instituto Oswaldo Cruz* 104 (juillet): 71-75. <https://doi.org/10.1590/S0074-02762009000900012>.
- Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, et Olivier Gascuel. 2010. « New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0 ». *Systematic Biology* 59 (3): 307-21. <https://doi.org/10.1093/sysbio/syq010>.
- Gurevich, Alexey, Vladislav Saveliev, Nikolay Vyahhi, et Glenn Tesler. 2013. « QUASt: Quality Assessment Tool for Genome Assemblies ». *Bioinformatics* 29 (8): 1072-75. <https://doi.org/10.1093/bioinformatics/btt086>.
- Gurgel-Gonçalves, Rodrigo, Fernando Abad-Franch, Jonatas B. C. Ferreira, Daniella B. Santana, et César A. Cuba Cuba. 2008. « Is *Rhodnius prolixus* (Triatominae) Invading Houses in Central Brazil? ». *Acta Tropica* 107 (2): 90-98. <https://doi.org/10.1016/j.actatropica.2008.04.020>.
- Gurgel-Gonçalves, Rodrigo, et César Augusto Cuba Cuba. 2009. « Predicting the Potential Geographical Distribution of *Rhodnius neglectus* (Hemiptera, Reduviidae) Based on Ecological Niche Modeling ». *Journal of Medical Entomology* 46 (4): 952-60. <https://doi.org/10.1603/033.046.0430>.
- Ha, Tal Soo, et Dean P. Smith. 2022. « Recent Insights into Insect Olfactory Receptors and Odorant-Binding Proteins ». *Insects* 13 (10): 926. <https://doi.org/10.3390/insects13100926>.
- Haas, Brian J, Alexie Papanicolaou, Moran Yassour, Manfred Grabherr, Philip D Blood, Joshua Bowden, Matthew Brian Couger, et al. 2013. « De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis ». *Nature Protocols* 8 (8): 1494-1512. <https://doi.org/10.1038/nprot.2013.084>.
- Hahn, Matthew W., Mira V. Han, et Sang-Gook Han. 2007. « Gene Family Evolution across 12 *Drosophila* Genomes ». *PLOS Genetics* 3 (11): e197. <https://doi.org/10.1371/journal.pgen.0030197>.
- Hall, T.A. 1999. « BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. ». *Nucleic Acids Symposium Series* 41: 95-98.
- Hallem, Elissa A., Anupama Dahanukar, et John R. Carlson. 2006. « Insect Odor and Taste Receptors ». *Annual Review of Entomology* 51 (1): 113-35. <https://doi.org/10.1146/annurev.ento.51.051705.113646>.
- Hansen, Immo A., Stacy D. Rodriguez, Lisa L. Drake, David P. Price, Brittny N. Blakely, John I. Hammond, Hitoshi Tsujimoto, Erika Y. Monroy, William A. Maio, et Alvaro Romero. 2014. « The Odorant Receptor Co-Receptor from the Bed Bug, *Cimex lectularius* L ». *PLOS ONE* 9 (11): e113692. <https://doi.org/10.1371/journal.pone.0113692>.
- Hansson, Bill S., et Marcus C. Stensmyr. 2011. « Evolution of Insect Olfaction ». *Neuron* 72 (5): 698-711. <https://doi.org/10.1016/j.neuron.2011.11.003>.
- Harrison, Richard G. 1989. « Animal Mitochondrial DNA as a Genetic Marker in Population and Evolutionary Biology ». *Trends in Ecology & Evolution* 4 (1): 6-11. [https://doi.org/10.1016/0169-5347\(89\)90006-2](https://doi.org/10.1016/0169-5347(89)90006-2).
- Harry, M. 1993. « Isozymic data question the specific status of some blood-sucking bugs of the genus *Rhodnius*, vectors of Chagas disease ». *Transactions of The Royal Society of Tropical Medicine and Hygiene* 87 (4): 492. [https://doi.org/10.1016/0035-9203\(93\)90054-T](https://doi.org/10.1016/0035-9203(93)90054-T).
- Harry, M., I. Galindez, et M. L. Cariou. 1992. « Isozyme Variability and Differentiation between *Rhodnius prolixus*, *R. robustus* and *R. pictipes*, Vectors of Chagas Disease in Venezuela ». *Medical and Veterinary Entomology* 6 (1): 37-43. <https://doi.org/10.1111/j.1365-2915.1992.tb00032.x>.

- Harry, Myriam, Fernando Lema, et Christine A. Romaña. 2000. « Chagas' Disease Challenge ». *The Lancet* 355 (9199): 236. [https://doi.org/10.1016/S0140-6736\(05\)72114-0](https://doi.org/10.1016/S0140-6736(05)72114-0).
- Hasin, Yehudit, Marcus Seldin, et Aldons Lulis. 2017. « Multi-Omics Approaches to Disease ». *Genome Biology* 18 (1): 1-15. <https://doi.org/10.1186/s13059-017-1215-1>.
- Hernández, Carolina, João Aristeu da Rosa, Gustavo A. Vallejo, Felipe Guhl, et Juan David Ramírez. 2020. « Taxonomy, Evolution, and Biogeography of the Rhodniini Tribe (Hemiptera: Reduviidae) ». *Diversity* 12 (3): 97. <https://doi.org/10.3390/d12030097>.
- Herrera, Leidi. 2010. « Una revisión sobre reservorios de *Trypanosoma (Schizotrypanum) cruzi* (Chagas, 1909), agente etiológico de la Enfermedad de Chagas ». *Boletín de Malaria y Salud Ambiental* 50 (1): 3-15.
- Hickner, Paul V., Chissa L. Rivaldi, Cole M. Johnson, Madhura Siddappaji, Gregory J. Raster, et Zainulabeuddin Syed. 2016. « The making of a pest: Insights from the evolution of chemosensory receptor families in a pestiferous and invasive fly, *Drosophila suzukii* ». *BMC Genomics* 17 (1): 648. <https://doi.org/10.1186/s12864-016-2983-9>.
- Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartín, et al. 2010. « Amazonia Through Time: Andean Uplift, Climate Change, Landscape Evolution, and Biodiversity ». *Science* 330 (6006): 927-31. <https://doi.org/10.1126/science.1194585>.
- Hwang, Wei Song, et Christiane Weirauch. 2012. « Evolutionary History of Assassin Bugs (Insecta: Hemiptera: Reduviidae): Insights from Divergence Dating and Ancestral State Reconstruction ». *PLOS ONE* 7 (9): e45523. <https://doi.org/10.1371/journal.pone.0045523>.
- Hypša, Václav, David F Tietz, Jan Zrzavý, Ryan O. M Rego, Cleber Galvao, et José Jurberg. 2002. « Phylogeny and Biogeography of Triatominae (Hemiptera: Reduviidae): Molecular Evidence of a New World Origin of the Asiatic Clade ». *Molecular Phylogenetics and Evolution* 23 (3): 447-57. [https://doi.org/10.1016/S1055-7903\(02\)00023-4](https://doi.org/10.1016/S1055-7903(02)00023-4).
- Ibarra-Cerdeña, Carlos N., Alejandro Zaldívar-Riverón, A. Townsend Peterson, Víctor Sánchez-Cordero, et Janine M. Ramsey. 2014. « Phylogeny and Niche Conservatism in North and Central American Triatomine Bugs (Hemiptera: Reduviidae: Triatominae), Vectors of Chagas' Disease ». *PLOS Neglected Tropical Diseases* 8 (10): e3266. <https://doi.org/10.1371/journal.pntd.0003266>.
- Ives, A, R Dinnage, LA Nell, M Helmus, et D Li. 2020. « phyr: model based phylogenetic analysis ». <https://daijiang.github.io/phyr/>.
- Iwamatsu, Takuma, Hidefumi Mitsuno, Daisuke Miyamoto, Takeshi Sakurai, et Ryohei Kanzaki. 2016. « PhOR3 contributes to the reception of DEET in body lice, *Pediculus humanus corporis* ». In *CHEMICAL SENSES*, 41:E209-E209. OXFORD UNIV PRESS GREAT CLARENDON ST, OXFORD OX2 6DP, ENGLAND.
- Ja, William W., Gil B. Carvalho, Elizabeth M. Mak, Noelle N. de la Rosa, Annie Y. Fang, Jonathan C. Liong, Ted Brummel, et Seymour Benzer. 2007. « Prandiology of *Drosophila* and the CAFE assay ». *Proceedings of the National Academy of Sciences* 104 (20): 8253-56. <https://doi.org/10.1073/pnas.0702726104>.
- Jaramillo, C., M. F. Montaña, L. R. Castro, G. A. Vallejo, et F. Guhl. 2001. « Differentiation and Genetic Analysis of *Rhodnius prolixus* and *Rhodnius colombiensis* by RDNA and RAPD Amplification ». *Memorias Do Instituto Oswaldo Cruz* 96 (8): 1043-48. <https://doi.org/10.1590/s0074-02762001000800003>.
- Jurberg, José, Dayse Rocha, et Cleber Galvão. 2009. « *Rhodnius zeledoni* sp. nov. afim de *Rhodnius paraensis* Sherlock, Guitton & Miles, 1977 (Hemiptera, Reduviidae, Triatominae) ». *Biota Neotropica* 9 (mars): 123-28. <https://doi.org/10.1590/S1676-06032009000100014>.
- Justi, Silvia A., Cleber Galvão, et Carlos G. Schrago. 2016. « Geological Changes of the Americas and Their Influence on the Diversification of the Neotropical Kissing Bugs (Hemiptera: Reduviidae: Triatominae) ». Édité par Ananias A. Escalante. *PLOS Neglected Tropical Diseases* 10 (4): e0004527. <https://doi.org/10.1371/journal.pntd.0004527>.
- Justi, Silvia Andrade, Claudia A. M. Russo, Jacenir Reis dos Santos Mallet, Marcos Takashi Obara, et Cleber Galvão. 2014. « Molecular phylogeny of Triatomini (Hemiptera: Reduviidae: Triatominae) ». *Parasites & Vectors* 7 (1): 149. <https://doi.org/10.1186/1756-3305-7-149>.
- Kahn, A. 1985. « Gènes, pseudogènes, familles multigéniques et évolution ». *médecine/sciences* 1 (4): 214. <https://doi.org/10.4267/10608/3341>.

- Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K F Wong, Arndt von Haeseler, et Lars S Jermiin. 2017. « ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates ». *Nature Methods* 14 (6): 587-89. <https://doi.org/10.1038/nmeth.4285>.
- Karpe, Snehal Dilip, Vikas Tiwari, et Sowdhamini Ramanathan. 2021. « InsectOR—Webserver for Sensitive Identification of Insect Olfactory Receptor Genes from Non-Model Genomes ». *PLOS ONE* 16 (1): e0245324. <https://doi.org/10.1371/journal.pone.0245324>.
- Kato, Hiroto, Ryan C. Jochim, Eduardo A. Gomez, Ryo Sakoda, Hiroyuki Iwata, Jesus G. Valenzuela, et Yoshihisa Hashiguchi. 2010. « A Repertoire of the Dominant Transcripts from the Salivary Glands of the Blood-Sucking Bug, *Triatoma dimidiata*, a Vector of Chagas Disease ». *Infection, Genetics and Evolution* 10 (2): 184-91. <https://doi.org/10.1016/j.meegid.2009.10.012>.
- Katoh, Kazutaka, John Rozewicki, et Kazunori D Yamada. 2019. « MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization ». *Briefings in Bioinformatics* 20 (4): 1160-66. <https://doi.org/10.1093/bib/bbx108>.
- Kieran, Troy J., Eric R.L. Gordon, Alejandro Zaldívar-Riverón, Carlos N. Ibarra-Cerdeña, Travis C. Glenn, et Christiane Weirauch. 2021. « Ultraconserved Elements Reconstruct the Evolution of Chagas Disease-vectoring Kissing Bugs (Reduviidae: Triatominae) ». *Systematic Entomology* 46 (3): 725-40. <https://doi.org/10.1111/syen.12485>.
- Kofler, R., R. V. Pandey, et C. Schlotterer. 2011. « PoPoolation2: Identifying Differentiation between Populations Using Sequencing of Pooled DNA Samples (Pool-Seq) ». *Bioinformatics* 27 (24): 3435-36. <https://doi.org/10.1093/bioinformatics/btr589>.
- Kolmogorov, Mikhail, Derek M. Bickhart, Bahar Behsaz, Alexey Gurevich, Mikhail Rayko, Sung Bong Shin, Kristen Kuhn, et al. 2020. « MetaFlye: Scalable Long-Read Metagenome Assembly Using Repeat Graphs ». *Nature Methods* 17 (11): 1103-10. <https://doi.org/10.1038/s41592-020-00971-x>.
- Koonin, Eugene V. 2005. « Orthologs, Paralogs, and Evolutionary Genomics ». *Annual Review of Genetics* 39 (1): 309-38. <https://doi.org/10.1146/annurev.genet.39.073003.114725>.
- Koren, Sergey, Brian P. Walenz, Konstantin Berlin, Jason R. Miller, Nicholas H. Bergman, et Adam M. Phillippy. 2017. « Canu: Scalable and Accurate Long-Read Assembly via Adaptive k-Mer Weighting and Repeat Separation ». *Genome Research* 27 (5): 722-36. <https://doi.org/10.1101/gr.215087.116>.
- Kumar, Sudhir, Glen Stecher, Michael Li, Christina Knyaz, et Koichiro Tamura. 2018. « MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms ». Édité par Fabia Ursula Battistuzzi. *Molecular Biology and Evolution* 35 (6): 1547-49. <https://doi.org/10.1093/molbev/msy096>.
- Lanfear, Robert, Brett Calcott, Simon Y. W. Ho, et Stephane Guindon. 2012. « PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses ». *Molecular Biology and Evolution* 29 (6): 1695-1701. <https://doi.org/10.1093/molbev/mss020>.
- Lardeux, Frédéric, Stéphanie Depickère, Stéphane Duchon, et Tamara Chavez. 2010. « Insecticide Resistance of *Triatoma infestans* (Hemiptera, Reduviidae) Vector of Chagas Disease in Bolivia ». *Tropical Medicine & International Health* 15 (9): 1037-48. <https://doi.org/10.1111/j.1365-3156.2010.02573.x>.
- Latorre Estivalis, Jose Manuel, Lucila Traverso, Gina Pontes, et Marcelo Gustavo Lorenzo. 2022. « The Antennal Transcriptome of *Triatoma infestans* Reveals Substantial Expression Changes Triggered by a Blood Meal ». *BMC Genomics* 23 (1): 861. <https://doi.org/10.1186/s12864-022-09059-6>.
- Latorre-Estivalis, Jose Manuel, Emerson Soares de Oliveira, Barbara Beiral Esteves, Leticia Santos Guimarães, Marina Neves Ramos, et Marcelo Gustavo Lorenzo. 2016. « Patterns of Expression of Odorant Receptor Genes in a Chagas Disease Vector ». *Insect Biochemistry and Molecular Biology* 69 (février): 71-81. <https://doi.org/10.1016/j.ibmb.2015.05.002>.
- Latorre-Estivalis, Jose Manuel, Hugh M. Robertson, Kimberly K. O. Walden, Jerônimo Ruiz, Leilane Oliveira Gonçalves, Alessandra A. Guarneri, et Marcelo Gustavo Lorenzo. 2017. « The Molecular Sensory Machinery of a Chagas Disease Vector: Expression Changes through Imaginal Molt and Sexually Dimorphic Features ». *Scientific Reports* 7: 40049. <https://doi.org/10.1038/srep40049>.
- Latorre-Estivalis, Jose Manuel, Marcos Sterkel, Sheila Ons, et Marcelo Gustavo Lorenzo. 2020. « Transcriptomics supports local sensory regulation in the antenna of the kissing-bug *Rhodnius prolixus* ». *BMC Genomics* 21 (1): 101. <https://doi.org/10.1186/s12864-020-6514-3>.



- Lazzari, Claudio, Gabriel Manrique, et P. Schilman. 2006. Vibrational communication in Triatominae (Heteroptera: Reduviidae).
- Lent, H., et J. Jurberg. 1969. « The genus *Rhodnius* Stål, 1859, with a study of the genitalia of the species (Hemiptera, Reduviidae, Triatominae). » *Revista Brasileira de Biologia* 29 (4): 487-560.
- Lent, Herman, José Jurberg, et Cleber Galvão. 1993. « *Rhodnius stali* n. sp., afim de *Rhodnius pictipers* Stal, 1872 (Hemiptera, Reduviidae, Triatominae) ». *Memórias do Instituto Oswaldo Cruz* 88 (décembre): 605-14. <https://doi.org/10.1590/S0074-02761993000400019>.
- Lent, Herman, et Pedro W. Wygodzinsky. 1979. « Revision of the Triatominae (Hemiptera, Reduviidae), and Their Significance as Vectors of Chagas' Disease. Bulletin of the AMNH; v. 163, Article 3 ». *Bulletin of the American Museum of Natural History* 163 (3): 123-520.
- Letunic, Ivica, et Peer Bork. 2019. « Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments ». *Nucleic Acids Research* 47 (W1): W256-59. <https://doi.org/10.1093/nar/gkz239>.
- Leyria, Jimena, Ian Orchard, et Angela B. Lange. 2020. « Transcriptomic Analysis of Regulatory Pathways Involved in Female Reproductive Physiology of *Rhodnius prolixus* under Different Nutritional States ». *Scientific Reports* 10 (1): 11431. <https://doi.org/10.1038/s41598-020-67932-4>.
- Li, F., X. Zhao, M. Li, K. He, C. Huang, Y. Zhou, Z. Li, et J. R. Walters. 2019. « Insect Genomes: Progress and Challenges ». *Insect Molecular Biology* 28 (6): 739-58. <https://doi.org/10.1111/imb.12599>.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, et 1000 Genome Project Data Processing Subgroup. 2009. « The Sequence Alignment/Map Format and SAMtools ». *Bioinformatics* 25 (16): 2078-79. <https://doi.org/10.1093/bioinformatics/btp352>.
- Li, Heng. 2018. « Minimap2: pairwise alignment for nucleotide sequences ». *Bioinformatics* 34 (18): 3094-3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Li, Li, Christian J. Stoeckert, et David S. Roos. 2003. « OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes ». *Genome Research* 13 (9): 2178-89. <https://doi.org/10.1101/gr.1224503>.
- Lidani, Kárita Cláudia Freitas, Fabiana Antunes Andrade, Lorena Bavia, Flávia Silva Damasceno, Marcia Holsbach Beltrame, Lara J. Messias-Reason, et Thaisa Lucas Sandri. 2019. « Chagas Disease: From Discovery to a Worldwide Health Problem ». *Frontiers in Public Health* 7. <https://www.frontiersin.org/article/10.3389/fpubh.2019.00166>.
- Lilioso, Maurício, Carolina Reigada, Dayane Pires-Silva, Fernanda von H. M. Fontes, Cleanne Limeira, Jackeline Monsalve-Lara, Elaine Folly-Ramos, Myriam Harry, Jane Costa, et Carlos Eduardo Almeida. 2020. « Dynamics of Food Sources, Ecotypic Distribution and *Trypanosoma cruzi* Infection in *Triatoma brasiliensis* from the Northeast of Brazil ». *PLOS Neglected Tropical Diseases* 14 (9): e0008735. <https://doi.org/10.1371/journal.pntd.0008735>.
- Liman, Emily R., Yali V. Zhang, et Craig Montell. 2014. « Peripheral Coding of Taste ». *Neuron* 81 (5): 984-1000. <https://doi.org/10.1016/j.neuron.2014.02.022>.
- Lin, Chung-Ping, et Bryan N Danforth. 2004. « How Do Insect Nuclear and Mitochondrial Gene Substitution Patterns Differ? Insights from Bayesian Analyses of Combined Datasets ». *Molecular Phylogenetics and Evolution* 30 (3): 686-702. [https://doi.org/10.1016/S1055-7903\(03\)00241-0](https://doi.org/10.1016/S1055-7903(03)00241-0).
- Linn, Charles E., Hattie R. Dambroski, Jeffrey L. Feder, Stewart H. Berlocher, Satoshi Nojima, et Wendell L. Roelofs. 2004. « Postzygotic isolating factor in sympatric speciation in *Rhagoletis* flies: Reduced response of hybrids to parental host-fruit odors ». *Proceedings of the National Academy of Sciences* 101 (51): 17753-58. <https://doi.org/10.1073/pnas.0408255101>.
- Liu, Feng, Zhou Chen, et Nannan Liu. 2017. « Molecular Basis of Olfactory Chemoreception in the Common Bed Bug, *Cimex lectularius* ». *Scientific Reports* 7 (1): 45531. <https://doi.org/10.1038/srep45531>.
- Liu, Feng, Zhou Chen, Zi Ye, et Nannan Liu. 2021. « The Olfactory Chemosensation of Hematophagous Hemipteran Insects ». *Frontiers in Physiology* 12. <https://www.frontiersin.org/article/10.3389/fphys.2021.703768>.
- Liu, Feng, et Nannan Liu. 2015. « Human Odorant Reception in the Common Bed Bug, *Cimex lectularius* ». *Scientific Reports* 5 (1): 15558. <https://doi.org/10.1038/srep15558>.

- Liu, Guoxia, Hongmei Ma, Hongyan Xie, Ning Xuan, Xia Guo, Zhongxue Fan, Balaji Rajashekar, Philippe Arnaud, Bernard Offmann, et Jean-François Picimbon. 2016. « Biotyping Characterization, Developmental Profiling, Insecticide Response and Binding Property of *Bemisia tabaci* Chemosensory Proteins: Role of CSP in Insect Defense ». *PLOS ONE* 11 (5): e0154706. <https://doi.org/10.1371/journal.pone.0154706>.
- Liu, Qin, Yunhai Guo, Yi Zhang, Wei Hu, Yuanyuan Li, Dan Zhu, Zhengbin Zhou, Jiatong Wu, Nansheng Chen, et Xiao-Nong Zhou. 2019. « A Chromosomal-Level Genome Assembly for the Insect Vector for Chagas Disease, *Triatoma rubrofasciata* ». *GigaScience* 8 (8): giz089. <https://doi.org/10.1093/gigascience/giz089>.
- Liu, Qun, Wei Liu, Baosheng Zeng, Guirong Wang, Dejun Hao, et Yongping Huang. 2017. « Deletion of the *Bombyx mori* Odorant Receptor Co-Receptor (BmOrco) Impairs Olfactory Sensitivity in Silkworms ». *Insect Biochemistry and Molecular Biology* 86 (juillet): 58-67. <https://doi.org/10.1016/j.ibmb.2017.05.007>.
- Lorenzo, M. G., et Claudio R. Lazzari. 1998. « Aggregation Behaviour and Interspecific Responses in Three Species of Triatominae ». *Memórias Do Instituto Oswaldo Cruz* 93 (janvier): 133-37. <https://doi.org/10.1590/S0074-02761998000100025>.
- Lorenzo, M. G., G. Manrique, H. H. R. Pires, M. G. de Brito Sánchez, L. Diotaiuti, et C. R. Lazzari. 1999. « Yeast Culture Volatiles as Attractants for *Rhodnius prolixus*: Electroantennogram Responses and Captures in Yeast-Baited Traps ». *Acta Tropica* 72 (1): 119-24. [https://doi.org/10.1016/S0001-706X\(98\)00093-X](https://doi.org/10.1016/S0001-706X(98)00093-X).
- Lorenzo, Marcelo G., et Claudio R. Lazzari. 1996. « The Spatial Pattern of Defaecation in *Triatoma infestans* and the Role of Faeces as a Chemical Mark of the Refuge ». *Journal of Insect Physiology* 42 (9): 903-7. [https://doi.org/10.1016/0022-1910\(96\)00008-X](https://doi.org/10.1016/0022-1910(96)00008-X).
- Lu, Kai, Yuanyuan Song, et Rensen Zeng. 2021. « The Role of Cytochrome P450-Mediated Detoxification in Insect Adaptation to Xenobiotics ». *Current Opinion in Insect Science* 43 (février): 103-7. <https://doi.org/10.1016/j.cois.2020.11.004>.
- Luo, Ruibang, Binghang Liu, Yinlong Xie, Zhenyu Li, Weihua Huang, Jianying Yuan, Guangzhu He, et al. 2012. « SOAPdenovo2: An Empirically Improved Memory-Efficient Short-Read de Novo Assembler ». *GigaScience* 1 (1): 18. <https://doi.org/10.1186/2047-217X-1-18>.
- Lyman, D. F., F. A. Monteiro, A. A. Escalante, C. Cordon-Rosales, D. M. Wesson, J. P. Dujardin, et C. B. Beard. 1999. « Mitochondrial DNA Sequence Variation among Triatomine Vectors of Chagas' Disease ». *The American Journal of Tropical Medicine and Hygiene* 60 (3): 377-86. <https://doi.org/10.4269/ajtmh.1999.60.377>.
- Machado, Fabiana S., Linda A. Jelicks, Louis V. Kirchoff, Jamshid Shirani, Fnu Nagajyothi, Shankar Mukherjee, Randin Nelson, et al. 2012. « Chagas Heart Disease: Report on Recent Developments ». *Cardiology in Review* 20 (2): 53-65. <https://doi.org/10.1097/CRD.0b013e31823efde2>.
- Machado-Allison, Carlos E., et Jaime Ramírez Pérez. 1967. *Chupos*. Vol. 2. Dirección de Cultura, Universidad Central de Venezuela.
- Maia Da Silva, F., A. C. V. Junqueira, M. Campaner, A. C. Rodrigues, G. Crisante, L. E. Ramirez, Z. C. E. Caballero, et al. 2007. « Comparative Phylogeography of *Trypanosoma rangeli* and *Rhodnius* (Hemiptera: Reduviidae) Supports a Long Coexistence of Parasite Lineages and Their Sympatric Vectors ». *Molecular Ecology* 16 (16): 3361-73. <https://doi.org/10.1111/j.1365-294X.2007.03371.x>.
- Maleszka, J., S. Forêt, R. Saint, et R. Maleszka. 2007. « RNAi-Induced Phenotypes Suggest a Novel Role for a Chemosensory Protein CSP5 in the Development of Embryonic Integument in the Honeybee (*Apis mellifera*) ». *Development Genes and Evolution* 217 (3): 189-96. <https://doi.org/10.1007/s00427-006-0127-y>.
- Manrique, Gabriel, Ana C. R. Vitta, Raquel A. Ferreira, Carlos L. Zani, C. Rikard Unelius, Claudio R. Lazzari, Lileia Diotaiuti, et Marcelo G. Lorenzo. 2006. « Chemical Communication in Chagas Disease Vectors. Source, Identity, and Potential Function of Volatiles Released by the Metasternal and Brindley's Glands of *Triatoma infestans* Adults ». *Journal of Chemical Ecology* 32 (9): 2035-52. <https://doi.org/10.1007/s10886-006-9127-7>.
- Marçais, Guillaume, et Carl Kingsford. 2011. « A fast, lock-free approach for efficient parallel counting of occurrences of k-mers ». *Bioinformatics* 27 (6): 764-70. <https://doi.org/10.1093/bioinformatics/btr011>.
- Marchant, A., F. Mougél, C. Almeida, E. Jacquín-Joly, J. Costa, et M. Harry. 2015. « De Novo Transcriptome Assembly for a Non-Model Species, the Blood-Sucking Bug *Triatoma brasiliensis*, a Vector of Chagas Disease ». *Genetica* 143 (2): 225-39. <https://doi.org/10.1007/s10709-014-9790-5>.

- Marchant, Axelle, Florence Mougel, Emmanuelle Jacquin-Joly, Carlos E. Almeida, Denis Blanchet, Jean-Michel Bérenger, João A. da Rosa, et Myriam Harry. 2021. « Chemosensory Gene Expression for Two Closely Relative Species *Rhodnius robustus* and *R. prolixus* (Hemiptera, Reduviidae, Triatominae) Vectors of Chagas Disease ». *Frontiers in Ecology and Evolution* 9: 639. <https://doi.org/10.3389/fevo.2021.725504>.
- Marchant, Axelle, Florence Mougel, Emmanuelle Jacquin-Joly, Jane Costa, Carlos Eduardo Almeida, et Myriam Harry. 2016. « Under-Expression of Chemosensory Genes in Domiciliary Bugs of the Chagas Disease Vector *Triatoma brasiliensis* ». Édité par Ana Rodriguez. *PLOS Neglected Tropical Diseases* 10 (10): e0005067. <https://doi.org/10.1371/journal.pntd.0005067>.
- Marcombe, Sébastien, Romain Blanc Mathieu, Nicolas Pocquet, Muhammad-Asam Riaz, Rodolphe Poupardin, Serge Sélidor, Frédéric Darriet, et al. 2012. « Insecticide Resistance in the Dengue Vector *Aedes Aegypti* from Martinique: Distribution, Mechanisms and Relations with Environmental Factors ». *PLOS ONE* 7 (2): e30989. <https://doi.org/10.1371/journal.pone.0030989>.
- Marti, G. A., M. G. Echeverria, E. Waleckx, M. L. Susevich, A. Balsalobre, et D. E. Gorla. 2014. « Triatominae in Furnariid Nests of the Argentine Gran Chaco ». *Journal of Vector Ecology: Journal of the Society for Vector Ecology* 39 (1): 66-71. <https://doi.org/10.1111/j.1948-7134.2014.12071.x>.
- Martin, Marcel. 2011. « Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads ». *EMBnet.Journal* 17 (1): 10-12. <https://doi.org/10.14806/ej.17.1.200>.
- Martins, Kiev, Cléber de Mesquita Andrade, Andressa Noronha Barbosa-Silva, Gerson Barbosa do Nascimento, Egler Chiari, Lúcia Maria da Cunha Galvão, et Antonia Cláudia Jácome da Câmara. 2015. « *Trypanosoma cruzi* III Causing the Indeterminate Form of Chagas Disease in a Semi-Arid Region of Brazil ». *International Journal of Infectious Diseases* 39 (octobre): 68-75. <https://doi.org/10.1016/j.ijid.2015.08.012>.
- Mathers, Thomas C, Roland H M Wouters, Sam T Mugford, David Swarbreck, Cock van Oosterhout, et Saskia A Hogenhout. 2021. « Chromosome-Scale Genome Assemblies of Aphids Reveal Extensively Rearranged Autosomes and Long-Term Conservation of the X Chromosome ». *Molecular Biology and Evolution* 38 (3): 856-75. <https://doi.org/10.1093/molbev/msaa246>.
- Matias, A., JX de la Riva, M. Torrez, et J. P. Dujardin. 2001. « *Rhodnius robustus* in Bolivia Identified by Its Wings ». *Memórias Do Instituto Oswaldo Cruz* 96 (octobre): 947-50. <https://doi.org/10.1590/S0074-02762001000700010>.
- Matute, Daniel R, Aaron A Comeault, Eric Earley, Antonio Serrato-Capuchina, David Peede, Anaïs Monroy-Eklund, Wen Huang, Corbin D Jones, Trudy F C Mackay, et Jerry A Coyne. 2020. « Rapid and Predictable Evolution of Admixed Populations Between Two *Drosophila* Species Pairs ». *Genetics* 214 (1): 211-30. <https://doi.org/10.1534/genetics.119.302685>.
- Mavárez, Jesús, et Mauricio Linares. 2008. « Homoploid Hybrid Speciation in Animals ». *Molecular Ecology* 17 (19): 4181-85. <https://doi.org/10.1111/j.1365-294X.2008.03898.x>.
- McBride, Carolyn S. 2007. « Rapid evolution of smell and taste receptor genes during host specialization in *Drosophila sechellia* ». *Proceedings of the National Academy of Sciences* 104 (12): 4996-5001. <https://doi.org/10.1073/pnas.0608424104>.
- McBride, Carolyn S, et J Roman Arguello. 2007. « Five *Drosophila* Genomes Reveal Nonneutral Evolution and the Signature of Host Specialization in the Chemoreceptor Superfamily ». *Genetics* 177 (3): 1395-1416. <https://doi.org/10.1534/genetics.107.078683>.
- McBride, Carolyn S., Felix Baier, Aman B. Omondi, Sarabeth A. Spitzer, Joel Lutomiah, Rosemary Sang, Rickard Ignell, et Leslie B. Vosshall. 2014. « Evolution of mosquito preference for humans linked to an odorant receptor ». *Nature* 515 (7526): 222-27. <https://doi.org/10.1038/nature13964>.
- McDonald, John H., et Martin Kreitman. 1991. « Adaptive Protein Evolution at the Adh Locus in *Drosophila* ». *Nature* 351 (6328): 652-54. <https://doi.org/10.1038/351652a0>.
- Mendes, Fábio K, Dan Vanderpool, Ben Fulton, et Matthew W Hahn. 2020. « CAFE 5 models variation in evolutionary rates among gene families ». *Bioinformatics* 36 (22-23): 5516-18. <https://doi.org/10.1093/bioinformatics/btaa1022>.
- Mendonça, Vagner, Kaio Alevi, Livia Medeiros, Juliana Nascimento, Maria Azeredo-Oliveira, et João da Rosa. 2014. « Cytogenetic and morphologic approaches of hybrids from experimental crosses between *Triatoma lenti* Sherlock & Serafim, 1967 and *T. sherlocki* Papa et al., 2002 (Hemiptera: Reduviidae) ». *Infection, Genetics and Evolution* 26 (janvier): 123-31.

- Meneguetti, Dionatas Ulises de Oliveira, Simone Delgado Tojal, Pedro Raimundo Mathias de Miranda, João Aristeu da Rosa, et Luis Marcelo Aranha Camargo. 2015. « First report of *Rhodnius montenegrensis* (Hemiptera, Reduviidae, Triatominae) in the State of Acre, Brazil ». *Revista da Sociedade Brasileira de Medicina Tropical* 48 (4): 471-73. <https://doi.org/10.1590/0037-8682-0029-2015>.
- Meng, Guanliang, Yiyuan Li, Chentao Yang, et Shanlin Liu. 2019. « MitoZ: A Toolkit for Animal Mitochondrial Genome Assembly, Annotation and Visualization ». *Nucleic Acids Research* 47 (11): e63-e63. <https://doi.org/10.1093/nar/gkz173>.
- Meslin, Camille, Pauline Mainet, Nicolas Montagné, Stéphanie Robin, Fabrice Legeai, Anthony Bretaudeau, J Spencer Johnston, et al. 2022. « *Spodoptera littoralis* genome mining brings insights on the dynamic of expansion of gustatory receptors in polyphagous noctuidae ». *G3 Genes|Genomes|Genetics* 12 (8): jkac131. <https://doi.org/10.1093/g3journal/jkac131>.
- Mesquita, Rafael D., Raquel J. Vionette-Amaral, Carl Lowenberger, Rolando Rivera-Pomar, Fernando A. Monteiro, Patrick Minx, John Spieth, et al. 2015. « Genome of *Rhodnius prolixus*, an Insect Vector of Chagas Disease, Reveals Unique Adaptations to Hematophagy and Parasite Infection ». *Proceedings of the National Academy of Sciences* 112 (48): 14936-41. <https://doi.org/10.1073/pnas.1506226112>.
- Meunier, Nicolas, Yesser Hadj Belgacem, et Jean-René Martin. 2007. « Regulation of Feeding Behaviour and Locomotor Activity by Takeout in *Drosophila* ». *The Journal of Experimental Biology* 210 (Pt 8): 1424-34. <https://doi.org/10.1242/jeb.02755>.
- Miller, Jason R., Arthur L. Delcher, Sergey Koren, Eli Venter, Brian P. Walenz, Anushka Brownley, Justin Johnson, Kelvin Li, Clark Mobarry, et Granger Sutton. 2008. « Aggressive Assembly of Pyrosequencing Reads with Mates ». *Bioinformatics (Oxford, England)* 24 (24): 2818-24. <https://doi.org/10.1093/bioinformatics/btn548>.
- Milne, M. A., E. J. Ross, D. E. Sonenshine, et P. Kirsch. 2009. « Attraction of *Triatoma dimidiata* and *Rhodnius prolixus* (Hemiptera: Reduviidae) to Combinations of Host Cues Tested at Two Distances ». *Journal of Medical Entomology* 46 (5): 1062-73. <https://doi.org/10.1603/033.046.0513>.
- Minh, Bui Quang, Minh Anh Thi Nguyen, et Arndt von Haeseler. 2013. « Ultrafast Approximation for Phylogenetic Bootstrap ». *Molecular Biology and Evolution* 30 (5): 1188-95. <https://doi.org/10.1093/molbev/mst024>.
- Minh, Bui Quang, Heiko A Schmidt, Olga Chernomor, Dominik Schrempf, Michael D Woodhams, Arndt von Haeseler, et Robert Lanfear. 2020. « IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era ». Édité par Emma Teeling. *Molecular Biology and Evolution* 37 (5): 1530-34. <https://doi.org/10.1093/molbev/msaa015>.
- Misra, Biswapriya B., Carl Langefeld, Michael Olivier, et Laura A. Cox. 2019. « Integrated Omics: Tools, Advances and Future Approaches ». *Journal of Molecular Endocrinology* 62 (1): R21-45. <https://doi.org/10.1530/JME-18-0055>.
- Mitchell, R. F., T. M. Schneider, A. M. Schwartz, M. N. Andersson, et D. D. McKenna. 2020. « The Diversity and Evolution of Odorant Receptors in Beetles (Coleoptera) ». *Insect Molecular Biology* 29 (1): 77-91. <https://doi.org/10.1111/imb.12611>.
- Mizushima, Daiki, Ahmed Tabbabi, Daisuke S. Yamamoto, Le Trung Kien, et Hiroto Kato. 2020. « Salivary Gland Transcriptome of the Asiatic *Triatoma rubrofasciata* ». *Acta Tropica* 210 (octobre): 105473. <https://doi.org/10.1016/j.actatropica.2020.105473>.
- Montagné, Nicolas, Thomas Chertemps, Isabelle Brigaud, Adrien François, Marie-Christine François, Arthur de Fouchier, Philippe Lucas, Mattias C. Larsson, et Emmanuelle Jacquin-Joly. 2012. « Functional Characterization of a Sex Pheromone Receptor in the Pest Moth *Spodoptera littoralis* by Heterologous Expression in *Drosophila* ». *European Journal of Neuroscience* 36 (5): 2588-96. <https://doi.org/10.1111/j.1460-9568.2012.08183.x>.
- Monteiro, F. A., C. Lazoski, F. Noireau, et A. M. Solé-Cava. 2002. « Allozyme Relationships among Ten Species of Rhodniini, Showing Paraphyly of *Rhodnius* Including *Psammolestes* ». *Medical and Veterinary Entomology* 16 (1): 83-90. <https://doi.org/10.1046/j.0269-283x.2002.00343.x>.
- Monteiro, F. A., D. M. Wesson, E. M. Dotson, C. J. Schofield, et C. B. Beard. 2000. « Phylogeny and Molecular Taxonomy of the Rhodniini Derived from Mitochondrial and Nuclear DNA Sequences. » *The American Journal of Tropical Medicine and Hygiene* 62 (4): 460-65. <https://doi.org/10.4269/ajtmh.2000.62.460>.
- Monteiro, Fernando A., Toby V. Barrett, Sinead Fitzpatrick, Celia Cordon-Rosales, Dora Feliciangeli, et Charles B. Beard. 2003. « Molecular Phylogeography of the Amazonian Chagas Disease Vectors *Rhodnius prolixus* and *R. robustus* ». *Molecular Ecology* 12 (4): 997-1006. <https://doi.org/10.1046/j.1365-294x.2003.01802.x>.

- Monteiro, Fernando Araujo, Christiane Weirauch, Márcio Felix, Cristiano Lazoski, et Fernando Abad-Franch. 2018. « Chapter Five - Evolution, Systematics, and Biogeography of the Triatominae, Vectors of Chagas Disease ». In *Advances in Parasitology*, édité par D. Rollinson et J. R. Stothard, 99:265-344. Academic Press. <https://doi.org/10.1016/bs.apar.2017.12.002>.
- Moreno Mejia, Jaime, Cleber Galvão, et José Jurberg. 1999. « *Rhodnius colombiensis* sp. N. da Colombia com quadros comparativos entre estruturas fâlicas do gênero *Rhodnius* Stal, 1859 (Hemiptera, Reduviidae., Triatominae) ». *Entomología y Vectores* 6 (novembre): 601-17.
- Mougabure-Cueto, Gastón, et María Inés Picollo. 2015. « Insecticide Resistance in Vector Chagas Disease: Evolution, Mechanisms and Management ». *Acta Tropica* 149 (septembre): 70-85. <https://doi.org/10.1016/j.actatropica.2015.05.014>.
- Nascimento, Juliana Damieli, João Aristeu da Rosa, Fabian C. Salgado-Roa, Carolina Hernández, Carolina Pardo-Díaz, Kaio Cesar Chaboli Alevi, Amanda Ravazi, 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 (2): e0211285. <https://doi.org/10.1371/journal.pone.0211285>.
- Naville, Magali, Simon Henriët, Ian Warren, Sara Sumic, Magnus Reeve, Jean-Nicolas Volff, et Daniel Chourrout. 2019. « Massive Changes of Genome Size Driven by Expansions of Non-Autonomous Transposable Elements ». *Current Biology* 29 (7): 1161-1168.e6. <https://doi.org/10.1016/j.cub.2019.01.080>.
- Neal, R. A., et R. A. Miles. 1977. « The Sensitivity of Culture Methods to Detect Experimental Infections of *Trypanosoma cruzi* and Comparison with Xenodiagnosis ». *Revista Do Instituto De Medicina Tropical De Sao Paulo* 19 (3): 170-76.
- Nevoa, Jessica C., Maria T. Mendes, Marcos V. da Silva, Siomar C. Soares, Carlo J. F. Oliveira, et José M. C. Ribeiro. 2018. « An Insight into the Salivary Gland and Fat Body Transcriptome of *Panstrongylus lignarius* (Hemiptera: Heteroptera), the Main Vector of Chagas Disease in Peru ». *PLOS Neglected Tropical Diseases* 12 (2): e0006243. <https://doi.org/10.1371/journal.pntd.0006243>.
- Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, et Bui Quang Minh. 2015. « IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies ». *Molecular Biology and Evolution* 32 (1): 268-74. <https://doi.org/10.1093/molbev/msu300>.
- Nkya, Theresia Estomih, Idir Akhouayri, William Kisinza, et Jean-Philippe David. 2013. « Impact of Environment on Mosquito Response to Pyrethroid Insecticides: Facts, Evidences and Prospects ». *Insect Biochemistry and Molecular Biology* 43 (4): 407-16. <https://doi.org/10.1016/j.ibmb.2012.10.006>.
- Nóbrega, Aglaêr A., Marcio H. Garcia, Erica Tatto, Marcos T. Obara, Elenild Costa, Jeremy Sobel, et Wildo N. Araujo. 2009. « Oral Transmission of Chagas Disease by Consumption of Açaí Palm Fruit, Brazil ». *Emerging Infectious Diseases* 15 (4): 653-55. <https://doi.org/10.3201/eid1504.081450>.
- Oliveira, Jader, Kaio Cesar Chaboli Alevi, Amanda Ravazi, Heitor Miraglia Herrera, Filipe Martins Santos, Maria Tercília Vilela de Azeredo-Oliveira, et João Aristeu da Rosa. 2018. « New Evidence of the Monophyletic Relationship of the Genus *Psammolestes* Bergroth, 1911 (Hemiptera, Reduviidae, Triatominae) ». *The American Journal of Tropical Medicine and Hygiene* 99 (6): 1485-88. <https://doi.org/10.4269/ajtmh.18-0109>.
- Oliveira, Jader de, João Aristeu da Rosa, Felipe Mendes Fontes, David Campos Andrade, Rubens Riscala Madi, et Cláudia Moura de Melo. 2021. « *Psammolestes tertius* Lent & Jurberg, 1965 (Hemiptera, Reduviidae, Triatominae): first report in Sergipe State, Brazil ». *Revista da Sociedade Brasileira de Medicina Tropical* 54: e0708-2020. <https://doi.org/10.1590/0037-8682-0708-2020>.
- Ortiz, Mario Iván, Alejandro Suárez-Rivillas, et Jorge Molina. 2011. « Behavioural Responses to Human Skin Extracts and Antennal Phenotypes of Sylvatic First Filial Generation and Long Rearing Laboratory Colony *Rhodnius prolixus* ». *Memorias Do Instituto Oswaldo Cruz* 106 (4): 461-66. <https://doi.org/10.1590/s0074-02762011000400013>.
- Otálora-Luna, Fernando, Jean-Luc Perret, et Patrick M. Guerin. 2004. « Appetence Behaviours of the Triatomine Bug *Rhodnius prolixus* on a Servosphere in Response to the Host Metabolites Carbon Dioxide and Ammonia ». *Journal of Comparative Physiology A* 190 (10): 847-54. <https://doi.org/10.1007/s00359-004-0540-5>.
- Pan American Health Organization, World Health Organization, n.d. Chagas disease. Available at: <https://www.paho.org/en/topics/chagas-disease>. Accessed March 21, 2022.

- Panfilio, Kristen A., Iris M. Vargas Jentsch, Joshua B. Benoit, Deniz Ereyilmaz, Yuichiro Suzuki, Stefano Colella, Hugh M. Robertson, et al. 2019. « Molecular Evolutionary Trends and Feeding Ecology Diversification in the Hemiptera, Anchored by the Milkweed Bug Genome ». *Genome Biology* 20 (1): 64. <https://doi.org/10.1186/s13059-019-1660-0>.
- Panzer, F., R. Perez, Y. Panzer, F. Alvarez, E. Scvortzoff, et R. Salvatella. 1995. « Karyotype Evolution in Holocentric Chromosomes of Three Related Species of Triatomines (Hemiptera-Reduviidae) ». *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 3 (3): 143-50. <https://doi.org/10.1007/BF00710707>.
- Panzer, F., R. Pérez, Y. Panzer, I. Ferrandis, M.J. Ferreira, et L. Calleros. 2010. « Cytogenetics and Genome Evolution in the Subfamily Triatominae (Hemiptera, Reduviidae) ». *Cytogenetic and Genome Research* 128 (1-3): 77-87. <https://doi.org/10.1159/000298824>.
- Panzer, Francisco, Inés Ferrandis, Janine Ramsey, Paz M. Salazar-Schettino, Margarita Cabrera, Carlota Monroy, María D. Bargas, et al. 2007. « Genome Size Determination in Chagas Disease Transmitting Bugs (Hemiptera-Triatominae) by Flow Cytometry ». *The American Journal of Tropical Medicine and Hygiene* 76 (3): 516-21.
- Panzer, Francisco, Sebastian Pita, et Pedro Lorite. 2021. « Chromosome Structure and Evolution of Triatominae: A Review ». In *Triatominae - The Biology of Chagas Disease Vectors*, édité par Alessandra Guarneri et Marcelo Lorenzo, 65-99. *Entomology in Focus*. Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-64548-9\\_4](https://doi.org/10.1007/978-3-030-64548-9_4).
- Patterson, James S., et Michael W. Gaunt. 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 ». *Molecular Phylogenetics and Evolution* 56 (2): 608-21. <https://doi.org/10.1016/j.ympev.2010.04.038>.
- Paula, Alexandre Silva de, Carlos Barreto, Mike Cerqueira Martins Telmo, Lileia Diotaiuti, et Cleber Galvão. 2021. « Historical Biogeography and the Evolution of Hematophagy in Rhodniini (Heteroptera: Reduviidae: Triatominae) ». *Frontiers in Ecology and Evolution* 9. <https://www.frontiersin.org/article/10.3389/fevo.2021.660151>.
- Paula, Alexandre Silva de, Liléia Diotaiuti, et Christopher John Schofield. 2005. « Testing the Sister-Group Relationship of the Rhodniini and Triatomini (Insecta: Hemiptera: Reduviidae: Triatominae) ». *Molecular Phylogenetics and Evolution* 35 (3): 712-18. <https://doi.org/10.1016/j.ympev.2005.03.003>.
- Paulo, Daniel F., Ana C. M. Junqueira, Alex P. Arp, André S. Vieira, Jorge Ceballos, Steven R. Skoda, Adalberto A. Pérez-de-León, et al. 2021. « Disruption of the Odorant Coreceptor Orco Impairs Foraging and Host Finding Behaviors in the New World Screwworm Fly ». *Scientific Reports* 11 (1): 11379. <https://doi.org/10.1038/s41598-021-90649-x>.
- Payseur, Bret A., et Loren H. Rieseberg. 2016. « A Genomic Perspective on Hybridization and Speciation ». *Molecular Ecology* 25 (11): 2337-60. <https://doi.org/10.1111/mec.13557>.
- Pellicer, Jaume, Oriane Hidalgo, Steven Dodsworth, et Ilia J. Leitch. 2018. « Genome Size Diversity and Its Impact on the Evolution of Land Plants ». *Genes* 9 (2): 88. <https://doi.org/10.3390/genes9020088>.
- Pelosi, Paolo, Immacolata Iovinella, Jiao Zhu, Guirong Wang, et Francesca R. Dani. 2018. « Beyond Chemoreception: Diverse Tasks of Soluble Olfactory Proteins in Insects ». *Biological Reviews of the Cambridge Philosophical Society* 93 (1): 184-200. <https://doi.org/10.1111/brv.12339>.
- Perez, Ruben, Yanina Panzer, Sandra Scafiezzo, Maria Cristina Mazzella, Francisco Panzer, Jean Pierre Dujardin, et Ekaterina Scvortzoff. 1992. « Cytogenetics as a tool for Triatomine species distinction (Hemiptera-Reduviidae) ». *Memórias do Instituto Oswaldo Cruz* 87 (3): 353-61. <https://doi.org/10.1590/S0074-02761992000300004>.
- Pessanha, Thaila Santos, Marcelo Pires, et Alena Mayo Iñiguez. 2022. « Molecular Detection of Blood Meal Source up to Three Months since the Last Meal: Experimental Starvation Resistance in Triatomines ». *Acta Tropica* 232 (août): 106507. <https://doi.org/10.1016/j.actatropica.2022.106507>.
- Petrov, D. A. 2001. « Evolution of Genome Size: New Approaches to an Old Problem ». *Trends in Genetics: TIG* 17 (1): 23-28. [https://doi.org/10.1016/s0168-9525\(00\)02157-0](https://doi.org/10.1016/s0168-9525(00)02157-0).
- Pflug, James M., Valerie Renee Holmes, Crystal Burrus, J. Spencer Johnston, et David R. Maddison. 2020. « Measuring Genome Sizes Using Read-Depth, k-Mers, and Flow Cytometry: Methodological Comparisons in Beetles (Coleoptera) ». *G3 (Bethesda, Md.)* 10 (9): 3047-60. <https://doi.org/10.1534/g3.120.401028>.

- Pinho, Ana Paula, Teresa Cristina M Gonçalves, Regina Helena Mangia, Nédia S Nehme Russell, et Ana Maria Jansen. 1998. « The Occurrence of *Rhodnius prolixus* Stal, 1859, Naturally Infected by *Trypanosoma cruzi* in the State of Rio de Janeiro, Brazil (Hemiptera, Reduviidae, Triatominae) ». *Memórias do Instituto Oswaldo Cruz* 93 (2): 141-43. <https://doi.org/10.1590/S0074-02761998000200001>.
- Pino-Marín, Antonia, Germán José Medina-Rincón, Sebastian Gallo-Bernal, Alejandro Duran-Crane, Álvaro Ignacio Arango Duque, María Juliana Rodríguez, Ramón Medina-Mur, Frida T. Manrique, Julian F. Forero, et Hector M. Medina. 2021. « Chagas Cardiomyopathy: From Romaña Sign to Heart Failure and Sudden Cardiac Death ». *Pathogens* 10 (5): 505. <https://doi.org/10.3390/pathogens10050505>.
- Pires, H. H. R, M. G Lorenzo, L Diotaiuti, C. R Lazzari, et A. N Lorenzo Figueiras. 2002. « Aggregation Behaviour in *Panstrongylus megistus* and *Triatoma infestans*: Inter and Intraspecific Responses ». *Acta Tropica* 81 (1): 47-52. [https://doi.org/10.1016/S0001-706X\(01\)00185-1](https://doi.org/10.1016/S0001-706X(01)00185-1).
- Pita, Sebastián, Francisco Panzera, Jesús Vela, Pablo Mora, Teresa Palomeque, et Pedro Lorite. 2017. « Complete Mitochondrial Genome of *Triatoma infestans* (Hemiptera, Reduviidae, Triatominae), Main Vector of Chagas Disease ». *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 54 (octobre): 158-63. <https://doi.org/10.1016/j.meegid.2017.06.003>.
- Poinar, George. 2005. « *Triatoma dominicana* sp. n. (Hemiptera: Reduviidae: Triatominae), and *Trypanosoma antiquus* sp. n. (Stercoraria: Trypanosomatidae), the First Fossil Evidence of a Triatomine-Trypanosomatid Vector Association ». *Vector borne and zoonotic diseases (Larchmont, N.Y.)* 5 (février): 72-81. <https://doi.org/10.1089/vbz.2005.5.72>.
- Poinar, George. 2013. « *Panstrongylus hispaniolae* Sp. n. (Hemiptera: Reduviidae: Triatominae), a New Fossil Triatomine in Dominican Amber, with Evidence of Gut Flagellates ». *Palaeodiversity* 6: 8.
- Poinar, George. 2019. « A Primitive Triatomine Bug, *Paleotriatoma metaxytaxa* Gen. et Sp. Nov. (Hemiptera: Reduviidae: Triatominae), in Mid-Cretaceous Amber from Northern Myanmar ». *Cretaceous Research* 93 (janvier): 90-97. <https://doi.org/10.1016/j.cretres.2018.09.004>.
- Pontes, G., S. Minoli, I. O. Insaurralde, M. G. de Brito Sanchez, et R. B. Barrozo. 2014. « Bitter Stimuli Modulate the Feeding Decision of a Blood-Sucking Insect via Two Sensory Inputs ». *Journal of Experimental Biology* 217 (20): 3708-17. <https://doi.org/10.1242/jeb.107722>.
- Powell, Jeffrey R, et Walter J Tabachnick. 2013. « History of domestication and spread of *Aedes aegypti* - A Review ». *Memórias do Instituto Oswaldo Cruz* 108 (Suppl 1): 11-17. <https://doi.org/10.1590/0074-0276130395>.
- Rabinovich, J.E., J.A. Leal, et D.Feliciangeli de Piñero. 1979. « Domiciliary biting frequency and blood ingestion of the Chagas's disease vector *Rhodnius prolixus* Ståhl (Hemiptera: Reduviidae), in Venezuela ». *Transactions of The Royal Society of Tropical Medicine and Hygiene* 73 (3): 272-83. [https://doi.org/10.1016/0035-9203\(79\)90082-8](https://doi.org/10.1016/0035-9203(79)90082-8).
- Rambaut, A. 2010. « FigTree v1.3.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. » <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rambaut, Andrew, Alexei J Drummond, Dong Xie, Guy Baele, et Marc A Suchard. 2018. « Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7 ». *Systematic Biology* 67 (5): 901-4. <https://doi.org/10.1093/sysbio/syy032>.
- Ramírez, Melanie, Mario I. Ortiz, Pablo Guerenstein, et Jorge Molina. 2020. « Novel repellents for the blood-sucking insects *Rhodnius prolixus* and *Triatoma infestans*, vectors of Chagas disease ». *Parasites & Vectors* 13 (1): 142. <https://doi.org/10.1186/s13071-020-04013-5>.
- Ramos, Leandro, Gabriela Castro, Janis Souza, Jader Oliveira, João da Rosa, Luis Camargo, R.M. Cunha, et Dionatas Meneguetti. 2018. « First report of *Rhodnius neglectus* (Hemiptera, Reduviidae, Triatominae) from the State of Acre, Brazil, and the Brazilian Western Amazon Region ». *Revista da Sociedade Brasileira de Medicina Tropical* 51 (avril): 212-14. <https://doi.org/10.1590/0037-8682-0320-2017>.
- Rane, Rahul V., Amol B. Ghodke, Ary A. Hoffmann, Owain R. Edwards, Tom K. Walsh, et John G. Oakeshott. 2019. « Detoxifying Enzyme Complements and Host Use Phenotypes in 160 Insect Species ». *Current Opinion in Insect Science* 31 (février): 131-38. <https://doi.org/10.1016/j.cois.2018.12.008>.
- Rane, Rahul V., Tom K. Walsh, Stephen L. Pearce, Lars S. Jermiin, Karl Hj Gordon, Stephen Richards, et John G. Oakeshott. 2016. « Are Feeding Preferences and Insecticide Resistance Associated with the Size of Detoxifying Enzyme Families in Insect Herbivores? » *Current Opinion in Insect Science* 13 (février): 70-76. <https://doi.org/10.1016/j.cois.2015.12.001>.

- Rassi, Anis, Anis Rassi, et José Antonio Marin-Neto. 2010. « Chagas Disease ». *The Lancet* 375 (9723): 1388-1402. [https://doi.org/10.1016/S0140-6736\(10\)60061-X](https://doi.org/10.1016/S0140-6736(10)60061-X).
- Ravazi, A, K C C Alevi, J Oliveira, J A Rosa, et M T V Azeredo-Oliveira. 2018. « Cytogenetic Analysis in Different Populations of *Rhodnius prolixus* and *R. Nasutus* from Different Countries of South America ». *Braz. J. Biol.* 78 (1).
- Ravazi, Amanda, Nicoly Olaia, Jader de Oliveira, Eder dos Santos Souza, João Aristeu da Rosa, Maria Tercília Vilela de Azeredo-Oliveira, et Kaio Cesar Chaboli Alevi. 2021. « Revisiting the Chromosomal Diversification of the Genus *Rhodnius* (Stål, 1859) (Hemiptera, Triatominae) ». *The American Journal of Tropical Medicine and Hygiene* 104 (2): 656-58. <https://doi.org/10.4269/ajtmh.20-0875>.
- Ravazi, Amanda, Jader Oliveira, Fabricio Campos, Fernanda Madeira, Yago Visinho dos Reis, Ana Oliveira, Maria Azeredo-Oliveira, João da Rosa, Cleber Galvão, et Kaio Alevi. 2022. « Trends in evolution of the Rhodniini tribe (Hemiptera, Triatominae): experimental crosses between *Psammolestes tertius* Lent & Jurberg, 1965 and *P. coreodes* Bergroth, 1911 and analysis of the reproductive isolating mechanisms ». *Parasites & Vectors* 14 (septembre). <https://doi.org/10.1186/s13071-021-04854-8>.
- Ribeiro, José M. C., Fernando A. Genta, Marcos H. F. Sorgine, Raquel Logullo, Rafael D. Mesquita, Gabriela O. Paiva-Silva, David Majerowicz, et al. 2014. « An Insight into the Transcriptome of the Digestive Tract of the Bloodsucking Bug, *Rhodnius prolixus* ». *PLoS Neglected Tropical Diseases* 8 (1). <https://doi.org/10.1371/journal.pntd.0002594>.
- Robertson, Hugh M., et Kevin W. Wanner. 2006. « The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family ». *Genome Research* 16 (11): 1395-1403. <https://doi.org/10.1101/gr.5057506>.
- Robertson, Hugh M., Coral G. Warr, et John R. Carlson. 2003. « Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster* ». *Proceedings of the National Academy of Sciences* 100 (suppl\_2): 14537-42. <https://doi.org/10.1073/pnas.2335847100>.
- Romaña, C.A., J.C.N. Pizarro, E. Rodas, et E. Guilbert. 1999. « Palm Trees as Ecological Indicators of Risk Areas for Chagas Disease ». *Transactions of the Royal Society of Tropical Medicine and Hygiene* 93 (6): 594-95. [https://doi.org/10.1016/S0035-9203\(99\)90059-7](https://doi.org/10.1016/S0035-9203(99)90059-7).
- Ronquist, Fredrik, Maxim Teslenko, Paul van der Mark, Daniel L. Ayres, Aaron Darling, Sebastian Höhna, Bret Larget, Liang Liu, Marc A. Suchard, et John P. Huelsenbeck. 2012. « MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space ». *Systematic Biology* 61 (3): 539-42. <https://doi.org/10.1093/sysbio/sys029>.
- Rosa, João Aristeu da, Hernany Henrique Garcia Justino, Juliana Damieli Nascimento, Vagner José Mendonça, Claudia Solano Rocha, Danila Blanco de Carvalho, Rossana Falcone, Maria Tercília Vilela de Azeredo-Oliveira, Kaio Cesar Chaboli Alevi, et Jader de Oliveira. 2017. « A New Species of *Rhodnius* from Brazil (Hemiptera, Reduviidae, Triatominae) ». *ZooKeys* 675 (mai): 1-25. <https://doi.org/10.3897/zookeys.675.12024>.
- Rosa, João Aristeu da, Cláudia Solano Rocha, Sueli Gardim, Mara Cristina Pinto, Vagner José Mendonça, Júlio César Rente Ferreira Filho, Elaine Oliveira Costa De Carvalho, et al. 2012. « Description of *Rhodnius montenegrensis* n. sp. (Hemiptera: Reduviidae: Triatominae) from the state of Rondônia, Brazil ». *Zootaxa* 3478 (décembre): 62-76. <https://doi.org/10.5281/ZENODO.212770>.
- Rosa, João Aristeu, Eder Santos Souza, Adelson da Costa Teixeira, Rosenildo Ribeiro Barbosa, Aecildo José de Souza, Tiago Belintani, Juliana Damieli Nascimento, Hélcio Reinaldo Gil-Santana, et Jader de Oliveira. 2017. « Third Record of *Rhodnius amazonicus* and Comparative Study with *R. pictipes* (Hemiptera, Reduviidae, Triatominae) ». *Acta Tropica* 176 (décembre): 364-72. <https://doi.org/10.1016/j.actatropica.2017.09.003>.
- Rothhammer, Francisco, Marvin J. Allison, Lautaro Núñez, Vivien Standen, et Bernardo Arriaza. 1985. « Chagas' Disease in Pre-Columbian South America ». *American Journal of Physical Anthropology* 68 (4): 495-98. <https://doi.org/10.1002/ajpa.1330680405>.
- Rozas, Julio, Albert Ferrer-Mata, Juan Carlos Sánchez-DelBarrio, Sara Guirao-Rico, Pablo Librado, Sebastián E Ramos-Onsins, et Alejandro Sánchez-Gracia. 2017. « DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets ». *Molecular Biology and Evolution* 34 (12): 3299-3302. <https://doi.org/10.1093/molbev/msx248>.
- Ruan, Jue, et Heng Li. 2020. « Fast and Accurate Long-Read Assembly with Wtdbg2 ». *Nature Methods* 17 (2): 155-58. <https://doi.org/10.1038/s41592-019-0669-3>.



- Rubinoff, Daniel, et Brenden S. Holland. 2005. « Between Two Extremes: Mitochondrial DNA Is Neither the Panacea nor the Nemesis of Phylogenetic and Taxonomic Inference ». Édité par Vincent Savolainen. *Systematic Biology* 54 (6): 952-61. <https://doi.org/10.1080/10635150500234674>.
- Salcedo-Porras, Nicolas, Claudia Umaña-Díaz, Ricardo de Oliveira Barbosa Bitencourt, et Carl Lowenberger. 2020. « The Role of Bacterial Symbionts in Triatomines: An Evolutionary Perspective ». *Microorganisms* 8 (9): 1438. <https://doi.org/10.3390/microorganisms8091438>.
- Sánchez-Gracia, A., F. G. Vieira, et J. Rozas. 2009. « Molecular Evolution of the Major Chemosensory Gene Families in Insects ». *Heredity* 103 (3): 208-16. <https://doi.org/10.1038/hdy.2009.55>.
- Sandoval, C. M., M. I. Joya, R. Gutierrez, et V. M. Angulo. 2000. « Cleptohaematophagy of the Triatomine Bug *Belminus herrerri* ». *Medical and Veterinary Entomology* 14 (1): 100-101. <https://doi.org/10.1046/j.1365-2915.2000.00210.x>.
- Sangenis, Luiz Henrique Conde, Marco Antonio Prates Nielebock, Ceumara da Silva Santos, Mateus Curty Carriello da Silva, et Glauber Motta Ribeiro Bento. 2016. « Transmissão da doença de Chagas por consumo de carne de caça: revisão sistemática ». *Revista Brasileira de Epidemiologia* 19 (12): 803-11. <https://doi.org/10.1590/1980-5497201600040010>.
- Santos, Adriana, José Marcos C. Ribeiro, Michael J. Lehane, Nelder Figueiredo Gontijo, Artur Botelho Veloso, Mauricio R. V. Sant'Anna, Ricardo Nascimento Araujo, Edmundo C. Grisard, et Marcos Horácio Pereira. 2007. « The Sialotranscriptome of the Blood-Sucking Bug *Triatoma brasiliensis* (Hemiptera, Triatominae) ». *Insect Biochemistry and Molecular Biology* 37 (7): 702-12. <https://doi.org/10.1016/j.ibmb.2007.04.004>.
- Sarquis, Otília, José Borges-Pereira, José Roberto Mac Cord, Taís Ferreira Gomes, Pedro Hernán Cabello, et Marli Maria Lima. 2004. « Epidemiology of Chagas Disease in Jaguaruana, Ceará, Brazil. I. Presence of Triatomines and Index of *Trypanosoma cruzi* Infection in Four Localities of a Rural Area ». *Memórias Do Instituto Oswaldo Cruz* 99 (mai): 263-70. <https://doi.org/10.1590/S0074-02762004000300004>.
- Sarov-Blat, L., W. V. So, L. Liu, et M. Rosbash. 2000. « The *Drosophila* Takeout Gene Is a Novel Molecular Link between Circadian Rhythms and Feeding Behavior ». *Cell* 101 (6): 647-56. [https://doi.org/10.1016/s0092-8674\(00\)80876-4](https://doi.org/10.1016/s0092-8674(00)80876-4).
- Sato, Koji, Kana Tanaka, et Kazushige Touhara. 2011. « Sugar-Regulated Cation Channel Formed by an Insect Gustatory Receptor ». *Proceedings of the National Academy of Sciences of the United States of America* 108 (28): 11680-85. <https://doi.org/10.1073/pnas.1019622108>.
- Schofield, C. 2000. « *Trypanosoma cruzi* - the Vector-Parasite Paradox ». *Memórias Do Instituto Oswaldo Cruz* 95 (4): 535-44. <https://doi.org/10.1590/s0074-02762000000400016>.
- Schofield, C. J., et Cleber Galvão. 2009. « Classification, Evolution, and Species Groups within the Triatominae ». *Acta Tropica, The Biology of Triatominae - Vectors of Chagas Disease*, 110 (2): 88-100. <https://doi.org/10.1016/j.actatropica.2009.01.010>.
- Schumer, Molly, Gil G. Rosenthal, et Peter Andolfatto. 2014. « How Common Is Homoploid Hybrid Speciation? » *Evolution* 68 (6): 1553-60. <https://doi.org/10.1111/evo.12399>.
- Sherlock, Ítalo A., Neide Guitton, et Michael Miles. 1977. « *Rhodnius paraensis* espécie nova do Estado do Pará, Brasil (Hemiptera, Reduviidae, Triatominae) ». *Acta Amazonica* 7 (1): 71-74. <https://doi.org/10.1590/1809-43921977071071>.
- Siljak-Yakovlev, Sonja, Edina Muratović, Faruk Bogunić, Neđad Bašić, Marija Šolić, Aldin Boškailo, Mickael Bourge, et al. 2020. « Genome size of Balkan flora: a database (GeSDaBaF) and C-values for 51 taxa of which 46 are novel ». *Plant Systematics and Evolution* 306 (mars). <https://doi.org/10.1007/s00606-020-01670-y>.
- Simão, Felipe A., Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, et Evgeny M. Zdobnov. 2015. « BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs ». *Bioinformatics* 31 (19): 3210-12. <https://doi.org/10.1093/bioinformatics/btv351>.
- Sloan, Daniel B., Justin C. Havird, et Joel Sharbrough. 2017. « The On-again, Off-again Relationship between Mitochondrial Genomes and Species Boundaries ». *Molecular Ecology* 26 (8): 2212-36. <https://doi.org/10.1111/mec.13959>.
- Smadja, C., et R. K. Butlin. 2009. « On the Scent of Speciation: The Chemosensory System and Its Role in Premating Isolation ». *Heredity* 102 (1): 77-97. <https://doi.org/10.1038/hdy.2008.55>.
- Smadja, C., P. Shi, R. K. Butlin, et H. M. Robertson. 2009. « Large Gene Family Expansions and Adaptive Evolution for Odorant and Gustatory Receptors in the Pea Aphid, *Acyrtosiphon pisum* ». *Molecular Biology and Evolution* 26 (9): 2073-86. <https://doi.org/10.1093/molbev/msp116>.

- Smadja, Carole M., Björn Canbäck, Renaud Vitalis, Mathieu Gautier, Julia Ferrari, Jing-Jiang Zhou, et Roger K. Butlin. 2012. « Large-Scale Candidate Gene Scan Reveals the Role of Chemoreceptor Genes in Host Plant Specialization and Speciation in the Pea Aphid ». *Evolution; International Journal of Organic Evolution* 66 (9): 2723-38. <https://doi.org/10.1111/j.1558-5646.2012.01612.x>.
- So, W. V., L. Sarov-Blat, C. K. Kotarski, M. J. McDonald, R. Allada, et M. Rosbash. 2000. « Takeout, a Novel *Drosophila* Gene under Circadian Clock Transcriptional Regulation ». *Molecular and Cellular Biology* 20 (18): 6935-44. <https://doi.org/10.1128/MCB.20.18.6935-6944.2000>.
- Soares, R. P., N. F. Gontijo, A. J. Romanha, L. Diotaiuti, et M. H. Pereira. 1998. « Salivary Heme Proteins Distinguish *Rhodnius prolixus* from *Rhodnius robustus* (Hemiptera: Reduviidae: Triatominae) ». *Acta Tropica* 71 (3): 285-91. [https://doi.org/10.1016/s0001-706x\(98\)00065-5](https://doi.org/10.1016/s0001-706x(98)00065-5).
- Soares, R P P, S E Barbosa, E C Borges, A J Romanha, J P Dujardin, C J Schofield, et L Diotaiuti. 2001. « Genetic Studies of *Psammolestes tertius* (Hemiptera: Reduviidae: Triatominae) Using Male Genital Morphology, Morphometry, Isoenzymes, and Random Amplified Polymorphic DNA ». *Biochemical Genetics*, 13.
- Song, Can, ShaoJun Liu, Jun Xiao, WeiGuo He, Yi Zhou, QinBo Qin, Chun Zhang, et Yun Liu. 2012. « Polyploid Organisms ». *Science China Life Sciences* 55 (4): 301-11. <https://doi.org/10.1007/s11427-012-4310-2>.
- Souza, A. C., S. Catalá, A. L. Carbajal de la Fuente, et A. C. V. Junqueira. 2017. « Phenotypic Variability of the Amazonian Species *Rhodnius brethesi* (Hemiptera: Reduviidae) ». *Journal of Medical Entomology* 54 (4): 909-16. <https://doi.org/10.1093/jme/tjx038>.
- Souza, Eder dos Santos, Noé Carlos Barbosa Von Atzingen, Maria Betânia Furtado, Jader de Oliveira, Juliana Damieli Nascimento, Daniel Pagotto Vendrami, Sueli Gardim, et João Aristeu da Rosa. 2016. « Description of *Rhodnius marabaensis* sp. n. (Hemiptera, Reduviidae, Triatominae) from Pará State, Brazil ». *ZooKeys*, no 621 (octobre): 45-62. <https://doi.org/10.3897/zookeys.621.9662>.
- Steverding, Dietmar. 2014. « The history of Chagas disease ». *Parasites & Vectors* 7 (07): 317. <https://doi.org/10.1186/1756-3305-7-317>.
- Störtkuhl, Klemens F., et Raffael Kettler. 2001. « Functional analysis of an olfactory receptor in *Drosophila melanogaster* ». *Proceedings of the National Academy of Sciences of the United States of America* 98 (16): 9381-85. <https://doi.org/10.1073/pnas.151105698>.
- Sudakaran, Sailendharan, Christian Kost, et Martin Kaltenpoth. 2017. « Symbiont Acquisition and Replacement as a Source of Ecological Innovation ». *Trends in Microbiology* 25 (5): 375-90. <https://doi.org/10.1016/j.tim.2017.02.014>.
- Sun, Dongdong, Yuan Huang, Zhenjie Qin, Haixia Zhan, Jinping Zhang, Yang Liu, et Shiyong Yang. 2020. « Identification of Candidate Olfactory Genes in the Antennal Transcriptome of the Stink Bug *Halyomorpha halys* ». *Frontiers in Physiology* 11. <https://www.frontiersin.org/article/10.3389/fphys.2020.00876>.
- Sun, J.-T., P.-Y. Jin, A.A. Hoffmann, X.-Z. Duan, J. Dai, G. Hu, X.-F. Xue, et X.-Y. Hong. 2018. « Evolutionary Divergence of Mitochondrial Genomes in Two Tetranychus Species Distributed across Different Climates ». *Insect Molecular Biology* 27 (6): 698-709. <https://doi.org/10.1111/imb.12501>.
- Suzuki, Hiromu C, Katsuhisa Ozaki, Takashi Makino, Hironobu Uchiyama, Shunsuke Yajima, et Masakado Kawata. 2018. « Evolution of Gustatory Receptor Gene Family Provides Insights into Adaptation to Diverse Host Plants in Nymphalid Butterflies ». Édité par Belinda Chang. *Genome Biology and Evolution* 10 (6): 1351-62. <https://doi.org/10.1093/gbe/evy093>.
- Suzuki, Yoshiyuki. 2008. « False-Positive Results Obtained from the Branch-Site Test of Positive Selection ». *Genes & Genetic Systems* 83 (4): 331-38. <https://doi.org/10.1266/ggs.83.331>.
- Tabares, Marcela, Mario Ortiz, Mabel Gonzalez, Chiara Carazzone, Martha J. Vives Florez, et Jorge Molina. 2018. « Behavioral Responses of *Rhodnius prolixus* to Volatile Organic Compounds Released in Vitro by Bacteria Isolated from Human Facial Skin ». *PLoS Neglected Tropical Diseases* 12 (4): e0006423. <https://doi.org/10.1371/journal.pntd.0006423>.
- Takada, Tomoyuki, Taiyo Sasaki, Ryoichi Sato, Shingo Kikuta, et Maki N. Inoue. 2018. « Differential Expression of a Fructose Receptor Gene in Honey Bee Workers According to Age and Behavioral Role ». *Archives of Insect Biochemistry and Physiology* 97 (2): e21437. <https://doi.org/10.1002/arch.21437>.
- Taneja, J., et P. M. Guerin. 1997. « Ammonia Attracts the Haematophagous Bug *Triatoma infestans*: Behavioural and Neurophysiological Data on Nymphs ». *Journal of Comparative Physiology A* 181 (1): 21-34. <https://doi.org/10.1007/s003590050089>.

- Toews, David P. L., et Alan Brelsford. 2012. « The Biogeography of Mitochondrial and Nuclear Discordance in Animals: BIOGEOGRAPHY OF MITO-NUCLEAR DISCORDANCE ». *Molecular Ecology* 21 (16): 3907-30. <https://doi.org/10.1111/j.1365-294X.2012.05664.x>.
- Traverso, Lucila, Andrés Lavore, Ivana Sierra, Victorio Palacio, Jesús Martínez-Barnette, José Manuel Latorre-Estivalis, Gaston Mougabure-Cueto, et al. 2017. « Comparative and Functional Triatomine Genomics Reveals Reductions and Expansions in Insecticide Resistance-Related Gene Families ». *PLOS Neglected Tropical Diseases* 11 (2): e0005313. <https://doi.org/10.1371/journal.pntd.0005313>.
- Traverso, Lucila, Ivana Sierra, Marcos Sterkel, Flavio Francini, et Sheila Ons. 2016. « Neuropeptidomics in *Triatoma infestans*. Comparative Transcriptomic Analysis among Triatomines ». *Journal of Physiology-Paris* 110 (3): 83-98. <https://doi.org/10.1016/j.jphysparis.2016.12.005>.
- Troemel, Emily R., Joseph H. Chou, Noelle D. Dwyer, Heather A. Colbert, et Cornelia I. Bargmann. 1995. « Divergent Seven Transmembrane Receptors Are Candidate Chemosensory Receptors in *C. elegans* ». *Cell* 83 (2): 207-18. [https://doi.org/10.1016/0092-8674\(95\)90162-0](https://doi.org/10.1016/0092-8674(95)90162-0).
- Valença-Barbosa, Carolina, Fabiano Araújo Fernandes, Helena Lucia Carneiro Santos, Otília Sarquis, Myriam Harry, Carlos Eduardo Almeida, et Marli Maria Lima. 2015. « Molecular Identification of Food Sources in Triatomines in the Brazilian Northeast: Roles of Goats and Rodents in Chagas Disease Epidemiology ». *The American Journal of Tropical Medicine and Hygiene* 93 (5): 994-97. <https://doi.org/10.4269/ajtmh.15-0156>.
- Valente, Vera, Sebastião Valente, Rodolfo Carcavallo, Dayse Rocha, Cleber Galvão, et José Jurberg. 2001. « Considerações sobre uma nova espécie do gênero *Rhodnius* Stal, do Estado do Pará, Brasil (Hemiptera, Reduviidae, Triatominae) ». *Entomología y Vectores* 8 (janvier): 65-80.
- Vargas-Chavez, Carlos, Neil Michel Longo Pendy, Sandrine E. Nsango, Laura Aguilera, Diego Ayala, et Josefa González. 2022. « Transposable Element Variants and Their Potential Adaptive Impact in Urban Populations of the Malaria Vector *Anopheles coluzzii* ». *Genome Research* 32 (1): 189-202. <https://doi.org/10.1101/gr.275761.121>.
- Vieira, Caroline, Yanna Praça, Kaio Bentes, Paula Santiago, Sofia Silva, Gabriel Silva, Flávia Motta, Izabela Bastos, Jaime Santana, et Carla Araújo. 2018. « Triatomines: Trypanosomatids, Bacteria, and Viruses Potential Vectors? » *Frontiers in Cellular and Infection Microbiology* 8 (novembre): 405. <https://doi.org/10.3389/fcimb.2018.00405>.
- Vieira, Filipe G., Alejandro Sánchez-Gracia, et Julio Rozas. 2007. « Comparative Genomic Analysis of the Odorant-Binding Protein Family in 12 *Drosophila* Genomes: Purifying Selection and Birth-and-Death Evolution ». *Genome Biology* 8 (11): R235. <https://doi.org/10.1186/gb-2007-8-11-r235>.
- Vinauger, Clément. 2011a. « Apprentissage et mémoire chez les insectes vecteurs de maladies humaines ». Thèse de doctorat, Tours. <https://www.theses.fr/2011TOUR4045>.
- Vinauger, Clément, Laura Buratti, et Claudio R. Lazzari. 2011b. « Learning the Way to Blood: First Evidence of Dual Olfactory Conditioning in a Blood-Sucking Insect, *Rhodnius prolixus*. II. Aversive Learning ». *Journal of Experimental Biology* 214 (18): 3039-45. <https://doi.org/10.1242/jeb.057075>.
- Vitta, Ana Cristina R., Alicia N. Lorenzo Figueiras, Claudio R. Lazzari, Liléia Diotaiuti, et Marcelo G. Lorenzo. 2002. « Aggregation Mediated by Faeces and Footprints in Triatoma Pseudomaculata (Heteroptera: Reduviidae), a Chagas Disease Vector ». *Memórias Do Instituto Oswaldo Cruz* 97 (septembre): 865-67. <https://doi.org/10.1590/S0074-02762002000600020>.
- Volonté, Mariano, Lucila Traverso, Jose Manuel Latorre Estivalis, Francisca Cunha Almeida, et Sheila Ons. 2022. « Comparative analysis of detoxification-related gene superfamilies across five hemipteran species ». *BMC Genomics* 23 (1): 757. <https://doi.org/10.1186/s12864-022-08974-y>.
- Vosshall, L. B., H. Amrein, P. S. Morozov, A. Rzhetsky, et R. Axel. 1999. « A Spatial Map of Olfactory Receptor Expression in the *Drosophila* Antenna ». *Cell* 96 (5): 725-36. [https://doi.org/10.1016/S0092-8674\(00\)80582-6](https://doi.org/10.1016/S0092-8674(00)80582-6).
- Vurture, Gregory W, Fritz J Sedlazeck, Maria Nattestad, Charles J Underwood, Han Fang, James Gurtowski, et Michael C Schatz. 2017. « GenomeScope: fast reference-free genome profiling from short reads ». *Bioinformatics* 33 (14): 2202-4. <https://doi.org/10.1093/bioinformatics/btx153>.
- Walter-Nuno, Ana B., Mabel L. Taracena, Rafael D. Mesquita, Pedro L. Oliveira, et Gabriela O. Paiva-Silva. 2018. « Silencing of Iron and Heme-Related Genes Revealed a Paramount Role of Iron in the Physiology of the Hematophagous Vector *Rhodnius prolixus* ». *Frontiers in Genetics* 9. <https://www.frontiersin.org/articles/10.3389/fgene.2018.00019>.

- Wanner, K. W., et H. M. Robertson. 2008. « The Gustatory Receptor Family in the Silkworm Moth *Bombyx mori* Is Characterized by a Large Expansion of a Single Lineage of Putative Bitter Receptors ». *Insect Molecular Biology* 17 (6): 621-29. <https://doi.org/10.1111/j.1365-2583.2008.00836.x>.
- Waterhouse, Robert M., Mathieu Seppey, Felipe A. Simão, et Evgeny M. Zdobnov. 2019. « Using BUSCO to Assess Insect Genomic Resources ». In *Insect Genomics: Methods and Protocols*, édité par Susan J. Brown et Michael E. Pfrender, 59-74. *Methods in Molecular Biology*. New York, NY: Springer. [https://doi.org/10.1007/978-1-4939-8775-7\\_6](https://doi.org/10.1007/978-1-4939-8775-7_6).
- Weedall, Gareth D., Jacob M. Riveron, Jack Hearn, Helen Irving, Colince Kamdem, Caroline Fouet, Bradley J. White, et Charles S. Wondji. 2020. « An Africa-Wide Genomic Evolution of Insecticide Resistance in the Malaria Vector *Anopheles funestus* Involves Selective Sweeps, Copy Number Variations, Gene Conversion and Transposons ». *PLOS Genetics* 16 (6): e1008822. <https://doi.org/10.1371/journal.pgen.1008822>.
- Wetzel, C. H., H. J. Behrendt, G. Gisselmann, K. F. Störtkuhl, B. Hovemann, et H. Hatt. 2001. « Functional Expression and Characterization of a *Drosophila* Odorant Receptor in a Heterologous Cell System ». *Proceedings of the National Academy of Sciences of the United States of America* 98 (16): 9377-80. <https://doi.org/10.1073/pnas.151103998>.
- Wigglesworth, V. B. 1936. « Symbiotic Bacteria in a Blood-Sucking Insect, *Rhodnius prolixus* Stål. (Hemiptera, Triatomidae) ». *Parasitology* 28 (2): 284-89. <https://doi.org/10.1017/S0031182000022459>.
- Wisotsky, Zev, Adriana Medina, Erica Freeman, et Anupama Dahanukar. 2011. « Evolutionary Differences in Food Preference Rely on Gr64e, a Receptor for Glycerol ». *Nature Neuroscience* 14 (12): 1534-41. <https://doi.org/10.1038/nn.2944>.
- Wootton, R. J. 1981. « Palaeozoic Insects ». *Annual Review of Entomology* 26 (1): 319-44. <https://doi.org/10.1146/annurev.en.26.010181.001535>.
- Wu, Yuan-Ming, Yang-Yang Liu, et Xiang-Sheng Chen. 2020. « Genomic content of chemosensory receptors in two sister blister beetles facilitates characterization of chemosensory evolution ». *BMC Genomics* 21 (1): 589. <https://doi.org/10.1186/s12864-020-06974-4>.
- Wu, Zhong-Zhen, Yang Cui, Meng-Qiu Qu, Jia-Hao Lin, Ming-Shun Chen, Shu-Ying Bin, et Jin-Tian Lin. 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* ». *Comparative Biochemistry and Physiology*, 7.
- Xiao, Jin-Hua, Zhen Yue, Ling-Yi Jia, Xin-Hua Yang, Li-Hua Niu, Zhuo Wang, Peng Zhang, et al. 2013. « Obligate Mutualism within a Host Drives the Extreme Specialization of a Fig Wasp Genome ». *Genome Biology* 14 (12): R141. <https://doi.org/10.1186/gb-2013-14-12-r141>.
- Xuan, Ning, Xia Guo, Hong-Yan Xie, Qi-Nian Lou, Xing-Bo Lu, Guo-Xia Liu, et Jean-François Picimbon. 2015. « Increased Expression of CSP and CYP Genes in Adult Silkworm Females Exposed to Avermectins ». *Insect Science* 22 (2): 203-19. <https://doi.org/10.1111/1744-7917.12116>.
- Yang, Ziheng. 2007. « PAML 4: Phylogenetic Analysis by Maximum Likelihood ». *Molecular Biology and Evolution* 24 (8): 1586-91. <https://doi.org/10.1093/molbev/msm088>.
- Ye, Chengxi, Christopher M. Hill, Shigang Wu, Jue Ruan, et Zhanshan Ma. 2016. « DBG2OLC: Efficient Assembly of Large Genomes Using Long Erroneous Reads of the Third Generation Sequencing Technologies ». *Scientific Reports* 6 (1). <https://doi.org/10.1038/srep31900>.
- Zacharias, Claudia A., Gina B. Pontes, Marcelo G. Lorenzo, et Gabriel Manrique. 2010. « Flight Initiation by Male *Rhodnius prolixus* Is Promoted by Female Odors ». *Journal of Chemical Ecology* 36 (4): 449-51. <https://doi.org/10.1007/s10886-010-9779-1>.
- Zermoglio, Paula F., José M. Latorre-Estivalis, José E. Crespo, Marcelo G. Lorenzo, et Claudio R. Lazzari. 2015. « Thermosensation and the TRPV Channel in *Rhodnius prolixus* ». *Journal of Insect Physiology* 81 (octobre): 145-56. <https://doi.org/10.1016/j.jinsphys.2015.07.014>.
- Zhang, Ya-Nan, Xiu-Yun Zhu, Li-Ping Fang, Peng He, Zhi-Qiang Wang, Geng Chen, Liang Sun, Zhan-Feng Ye, Dao-Gui Deng, et Jin-Bu Li. 2015. « Identification and Expression Profiles of Sex Pheromone Biosynthesis and Transport Related Genes in *Spodoptera litura* ». *PLOS ONE* 10 (10): e0140019. <https://doi.org/10.1371/journal.pone.0140019>.
- Zhao, Yisheng, Cleber Galvão, et Wanzhi Cai. 2021. « *Rhodnius micki*, a New Species of Triatominae (Hemiptera, Reduviidae) from Bolivia ». *ZooKeys* 1012 (janvier): 71-93. <https://doi.org/10.3897/zookeys.1012.54779>.

Zhao, Yisheng, Manjie Jiang, Yunfei Wu, Fan Song, Wanzhi Cai, et Hu Li. 2019. « Mitochondrial Genomes of Three Kissing Bugs (Reduviidae: Triatominae) and Their Phylogenetic Implications ». *International Journal of Biological Macromolecules* 134 (août): 36-42. <https://doi.org/10.1016/j.ijbiomac.2019.05.020>.

Zheng, Yuchi, et John J. Wiens. 2015. « Do Missing Data Influence the Accuracy of Divergence-Time Estimation with BEAST? » *Molecular Phylogenetics and Evolution* 85 (avril): 41-49. <https://doi.org/10.1016/j.ympev.2015.02.002>.

Zhu, Jiao, Immacolata Iovinella, Francesca Romana Dani, Yu-Ling Liu, Ling-Qiao Huang, Yang Liu, Chen-Zhu Wang, Paolo Pelosi, et Guirong Wang. 2016. « Conserved chemosensory proteins in the proboscis and eyes of Lepidoptera ». *International Journal of Biological Sciences* 12 (11): 1394-1404. <https://doi.org/10.7150/ijbs.16517>.

Zimin, Aleksey V., Guillaume Marçais, Daniela Puiu, Michael Roberts, Steven L. Salzberg, et James A. Yorke. 2013. « The MaSuRCA genome assembler ». *Bioinformatics* 29 (21): 2669-77. <https://doi.org/10.1093/bioinformatics/btt476>.

Zimin, Aleksey V., Daniela Puiu, Ming-Cheng Luo, Tingting Zhu, Sergey Koren, Guillaume Marçais, James A. Yorke, Jan Dvořák, et Steven L. Salzberg. 2017. « Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm ». *Genome Research* 27 (5): 787-92. <https://doi.org/10.1101/gr.213405.116>.