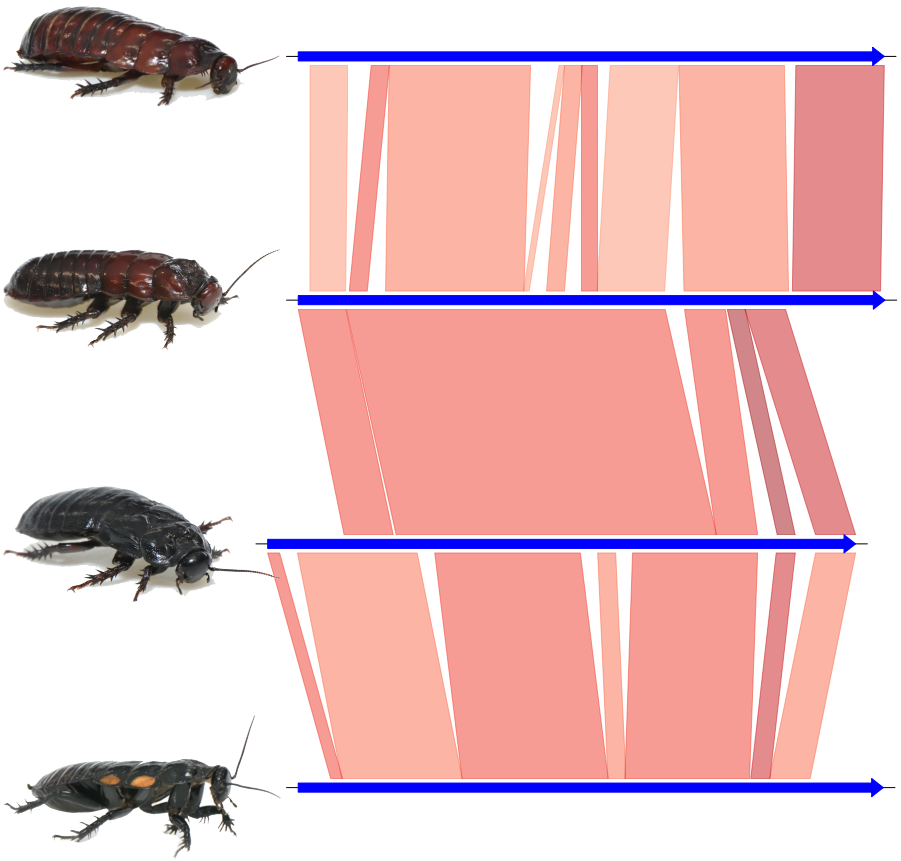
# Phylogenetic relationships and horizontal gene transfer in panesthiine and geoscapheine cockroaches

**Zhuzhi Zhang**



**The University of Sydney**

**School of Life and Environmental Sciences**

**Thesis submitted to fulfil the requirements for the degree of Doctor of Philosophy**

2026

# Authorship Attribution Statement

All chapters of this doctoral thesis comprise stand-alone manuscripts that have been prepared for submission to peer-reviewed journals. All chapters are tied to the central theme of this thesis, which explores the phylogeny and evolution of cockroaches from the subfamilies Panesthiinae and Geoscapheinae. Accordingly, there is inevitable repetition between chapters, particularly in the methods sections of Chapters 2-3.

Within this thesis, I am the sole author of Chapters 1 and 5. All other chapters are co-authored and therefore make use of the plural "we".

Chapter 2 of this thesis has been formatted for publication in a peer-reviewed journal as: Zhang Z., Walker, J. A., Rose H. A., Zwick, A., Lo N. Increased sampling of *Panesthia* improves resolution of phylogenetic relationships among Australian wood-feeding and soil-burrowing cockroaches. ZZ and NL conceived and designed the study. HAR and JAW performed taxon sampling. AZ performed DNA sequencing. ZZ analysed the molecular data and drafted the manuscript.

Chapter 3 of this thesis has been formatted for publication in a peer-reviewed journal as: Zhang Z., Fujiwara K., Maekawa K., Cheng Z., Bourguignon T., Wang Z., Lo N. Phylogenomics, Biogeography, and Morphological Evolution of Asian Panesthiinae. ZZ and NL conceived and designed the study. KF, KM, ZC, TB, YW and ZW performed sampling and laboratory experiments. ZZ analysed the molecular data, performed the relevant analyses, and drafted the manuscript.

Chapter 4 of this thesis has been formatted for publication in a peer-reviewed journal as: Zhang Z., Ewart K. M., Adams M. W. D., Rose H. A., Baker L., Jex, A., Lo N. Widespread Horizontal Gene Transfer from *Blattabacterium* Endosymbionts to Cockroach Genomes Reveals High Frequency of Non-Coding and Chimeric Inserts. ZZ, KME, MWDA and NL conceived and designed the study. HAR obtained specimens. ZZ performed laboratory

experiments. LB, MWDA, and AJ generated sequence data. ZZ and MWDA assembled and annotated the genome. KME identified HGT inserts. MWDA performed HGT insertions filtering. ZZ performed HGT verification and ancestral ancestry determination. ZZ and NL identified chimeric HGT inserts. ZZ performed the ortholog assignment and relaxed and positive selection analysis. ZZ drafted the manuscript.

Photos on the title page are by Yi Kai Tea.

Nathan Lo supervised this work and provided critical guidance on study design, analyses, and interpretation.

# Statement of Originality

I certify that, to the best of my knowledge, the content of this thesis is my own work, unless otherwise acknowledged. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual contents of this thesis are the product of my own work, and that all assistance received in preparing this thesis has been appropriately acknowledged.

# Acknowledgements

I would like to express my deepest gratitude to my principal supervisor, Professor Nathan Lo, for his invaluable guidance and support throughout my research. When I first joined the MEEP Lab at USYD, he took me to the O-Week at the School of Science tent, where he explained the ecology of *Macropanesthia rhinoceros*. Although I misunderstood his words and ended up waiting for an hour instead of ten minutes, it became a memorable and enjoyable experience. In research, Nate frequently inspired me with new ideas and encouraged me to explore my own, while in life he often showed care, asking how I was doing and warmly welcoming my girlfriend during her visits. His kindness has been both encouraging and heartwarming.

I am also sincerely grateful to my auxiliary supervisor, Professor Simon Ho. I still remember him meeting me on my first day at the university. His expertise in bioinformatics has greatly helped me in data analysis and in developing scripts for my work. Simon also gave me the opportunity to assist in the annual BEAST workshops, which enhanced my skills in teaching and presenting technical software. Beyond academic guidance, he often organized activities, such as visiting exhibitions on campus or taking us on trips, one of which allowed me to see whales with their calves and the superb lyrebird for the first time.

My heartfelt thanks also go to other members of the MEEP lab. Kyle has provided many ideas and support in our collaboration. Karen and her husband Jorge offered invaluable help in map-making and introduced us to Spanish and Mexican culture (and some great food in Sydney). Braxon generously shared his expertise in phylogenetics and lab work, always bringing energy and laughter to the group. Toby has been warm and helpful in advancing the overall project, while Maxim's work closely related to mine, led to many productive discussions.

Finally, I would like to thank my girlfriend, Sijia Yang, for her constant support and encouragement in life. Over the years, she has gifted me many cockroach-themed items and often shared them with my labmates. She also made use of her social media to recommend

events and food near the university—despite being overseas, thanks to the wonders of the internet. Her care and enthusiasm have been a great source of strength for me.

During the preparation of this thesis, I used artificial intelligence tools (OpenAI's ChatGPT) to assist with English grammar checking, language polishing, and debugging of analysis scripts. These tools were applied solely for improving clarity of expression and technical accuracy. All research ideas, data analyses, interpretations, and conclusions presented in this thesis are entirely my own.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

I certify that, if my candidature is successful, this thesis will be lodged with the University Librarian and made available for immediate use.

# Abstract

Cockroaches are an exceptionally diverse group of insects with striking ecological adaptability, occurring across tropical, subtropical, and temperate regions worldwide. Two important subfamilies of cockroaches are the Panesthiinae and Geoscapheinae, which respectively burrow into wood and soil, the latter exclusively in Australia. Members of these two subfamilies have long posed taxonomic challenges. In this thesis, I examined the phylogeny, biogeographic history, and morphological evolution of Panesthiinae and Geoscapheinae. I also undertook genome sequencing of representatives of this group and examined horizontal gene transfer (HGT) from their *Blattabacterium* symbionts. By integrating mitochondrial genomes, nuclear gene markers, and high-quality genomic data, along with extensive species sampling and bioinformatics analyses, my research has advanced our understanding of these cockroaches in multiple aspects.

Firstly, by expanding the sampling of Australian *Panesthia sloanei* and *P. tryoni*, I confirmed seven independent transitions from wood-feeding to soil-burrowing lifestyles. The early-branching position of *Geoscapheus dilatatus* + *G. robustus* within the *Panesthiinae* + *Geoscapheinae* group was resolved with high support for the first time, and the sister relationship between the *P. australis* + *P. obtusa* clade and the *P. tryoni* group was clarified. Furthermore, biogeographic reconstruction revealed that after *Panesthia* colonising Australia from Southeast Asia, this lineage became geographically isolated due to the continent-wide aridification, leading to repeated diversification and shaping its present-day distribution pattern.

Subsequently, utilising a large set of published samples along with some newly sequenced specimens, the entire Panesthiinae was firstly investigated using mitochondrial genomes. The results show that *Salganea* and *Miopanesthia* together form a well-supported monophyletic clade located at the base of Panesthiinae, while *Panesthia* is identified as a paraphyletic group. Ancestral state reconstruction of morphological characters revealed multiple independent reductions in key traits, including the degeneration or loss of wings, simplification of male genital phallomeres (L1, L2d, R2), and loss of the oothecal membrane.

These morphological changes are closely correlated with shifts in substrate use—from wood to soil—and the evolution of mating strategies.

Finally, utilising long-read sequencing technology and whole-genome sequencing data, I detected and validated HGT events in non-coding regions. I identified a substantially higher number of HGT insertions in the genomes of Panesthiinae and Geoscapheinae compared to other Blattodea species (1,543–5,037 insertions per genome versus 0–647 insertions per genome in other lineages). Most of these insertions are short (<100 bp), located in non-coding regions, and some longer insertions exhibit a chimeric structure—composed of discontinuous fragments from the *Blattabacterium* genome. Subsequent homology-based comparative analysis of these insertions across species indicates that these HGT events occurred within approximately the past 40 million years. Despite the large number of HGT events, no significant correlation was detected between HGT accumulation and the degree of relaxed selection at the genome-wide level. Furthermore, selective pressures between wood-feeding and soil-burrowing lineages did not show a significant difference.

# Table of Contents

# Chapter 1

## General introduction

### 1.1 Introduction to Blattodea

Blattodea is a widely recognised order of insects, often regarded as global pests due to their association with human habitats, where they feed on leftovers, plant material and cause damage to structures through feeding on wood. The order comprises over 7,500 extant species, including more than 2,900 termites (Krishna et al., 2013) and over 4,600 cockroaches (Beccaloni and Eggleton, 2013a). Despite their pest status, cockroaches and termites play significant ecological and economic roles as major decomposers of dead wood and may influence global climate dynamics (Beccaloni and Eggleton, 2013; Bell et al., 2007). Consequently, the systematics of both cockroaches and termites have attracted considerable research interest.

Cockroaches are a highly diverse group of insects with remarkable ecological versatility, distributed across tropical, subtropical, and temperate regions worldwide. Their distribution spans major habitat types, from deserts to tropical rainforests, and extends to specific microhabitats such as decaying wood, caves, and moss pools (Rebrina et al., 2020). The phylogenetic relationships within Blattodea are complex. Termites (Termitoidae), previously classified as a separate order (Isoptera), have been shown by molecular phylogenetic studies to be nested within Blattodea, belonging to the epifamily Termitoidae (Hellemans et al., 2024). Blattodea and Mantodea (preying mantids) are closely related, together forming the superorder Dictyoptera (Evangelista et al., 2019).

As insects that undergo incomplete metamorphosis, cockroaches have three main life stages: egg, nymph, and adult. During the reproductive season, females produce oothecae, each containing 12–25 eggs, typically arranged in two rows. In oviparous species, the ootheca develops externally, whereas in ovoviviparous species, the ootheca, or sometimes a reduced

oothecal membrane, is retained within the female's body for internal development; viviparous species give birth to live young (Rugg and Rose, 1984a). Nymphs gradually develop into adults through successive moults. The duration of each life stage varies among species; for example, the German cockroach (*Blattella germanica*) completes its life cycle in approximately 50–60 days, while other species may take several years.

## 1.2 Introduction to endosymbionts

Endosymbionts—bacteria that reside within the cells of eukaryotic hosts—are widespread in nature and have played a fundamental role in the evolution and diversification of numerous lineages (Margulis and Fester, 1991). Insects, the most speciose class within the Arthropoda, encompass over one million described species, with total diversity estimated to exceed five million (Stork, 2018). Endosymbionts are prevalent across most insect groups, conferring various advantages to their hosts, including protection against predators (Heyworth and Ferrari, 2016), resistance to insecticides (Kikuchi et al., 2012), effect on host reproduction (Serbus et al., 2008) and, most commonly, nutritional supplementation through the synthesis of essential amino acids and vitamins (Hu and Tsai, 2020; Nikoh et al., 2014; Sabree et al., 2009). These benefits link endosymbionts to host adaptive radiation (Bennett et al., 2024) and broader ecological impacts on terrestrial ecosystems (Clark et al., 2010; Degnan et al., 2025; Mitter et al., 1988; Peccoud et al., 2009).

Bacteria and fungi represent the most common insect endosymbionts, renowned for their metabolic versatility, which enables the production of compounds inaccessible to hosts. This metabolic capacity facilitates insect exploitation of novel ecological niches and is considered a major driver of extensive diversification (Kaltenpoth et al., 2025). Early microscopic examinations revealed bacterial endosymbionts residing within specialised host cells, termed bacteriocytes, in ants, sap-feeding insects, and blood-feeders (Batra and Buchner, 1968; Douglas, 1989). Antibiotic-mediated elimination of endosymbionts demonstrated that many are indispensable for host development, reproduction, and survival (Brooks and Richards, 1955; Koga et al., 2007). Such obligate, mutualistic endosymbionts—designated as primary endosymbionts—are estimated to occur in 10–15% of all insects (Wernegreen, 2002). They are strictly vertically transmitted from mothers to offspring during oogenesis (Herren et al.,

2013; Noda et al., 2025; Shan et al., 2021) and often persist for hundreds of millions of years (Jennifer J Wernegreen, 2017; Jennifer J. Wernegreen, 2017). Some insects harbour two or more co-primary endosymbionts, as observed in many Auchenorrhyncha (Brentassi and de la Fuente, 2024). In contrast, facultative (or secondary) endosymbionts are non-essential but can enhance host fitness under specific conditions (Guo et al., 2017; Kwak et al., 2025). They are not confined to specialised cells and may be transmitted both vertically and horizontally (Guo et al., 2017).

Morphological classification of bacterial endosymbionts proved challenging (Araldi-Brondolo et al., 2017), but DNA sequencing has unveiled their extensive diversity. Many primary endosymbionts belong to the Pseudomonadota, including *Buchnera aphidicola* (aphids, Douglas 1998), *Baumannia cicadellinicola* (sharpshooters, Gruwell et al., 2007), *Blochmannia floridanus* (carpenter ants, Feldhaar et al., 2007), *Carsonella ruddii* (psyllids, Spaulding and von Dohlen, 1998), and *Portiera aleyrodidarum* (whiteflies, Gruwell et al., 2007). Others, such as *Sulcia muelleri* (Auchenorrhyncha, Bennett et al., 2016) and *Blattabacterium cuenoti* (cockroaches, Tokuda et al., 2008), belong to the Bacteroidota. Sequencing-based classification has enabled direct comparisons with free-living relatives and provided insights into their evolutionary trajectories (Sabater-Muñoz et al., 2017).

For decades, insect–primary endosymbiont associations were hypothesised to stem from ancient infections of a shared ancestor (Moran et al., 2005; Jennifer J. Wernegreen, 2017). This has been corroborated in groups such as aphids, Auchenorrhyncha, cockroaches, tsetse flies, and whiteflies through molecular phylogenetics (Chen et al., 1999; Degnan et al., 2004; Lo et al., 2003; Munson et al., 1991; Thao and Baumann, 2004). Comparisons of host nuclear and mitochondrial genes with endosymbiont genomes reveal highly congruent phylogenies (Hall et al., 2016; Lo et al., 2003; Urban and Cryan, 2012), consistent with ancient co-diversification and strict vertical transmission. In contrast, facultative endosymbionts generally exhibit incongruent phylogenies, reflecting more recent associations and dynamic transmission patterns (Łukasik et al., 2015), though localised congruence has been reported (Fromont et al., 2016).

The strong co-phylogenetic signal of primary endosymbionts renders them valuable tools for investigating both bacterial and host evolution. Their strict maternal inheritance permits molecular clock calibrations using host fossil records, facilitating estimation of bacterial evolutionary timescales that are otherwise difficult to resolve. For instance, *Buchnera*–aphid associations date back over 100 million years (Nováková et al., 2013), while the partnership between *Sulcia muelleri* and Auchenorrhyncha extends beyond 260 million years, ranking among the oldest known insect–symbiont relationships (Moran et al., 2005). Furthermore, the congruence of host and endosymbiont phylogenies implies that symbiont genes can serve as proxies for reconstructing host evolutionary history, as demonstrated in whiteflies (de Moraes et al., 2018), armoured scale insects (Andersen et al., 2010), and aphids (Liu et al., 2013).

## 1.3 *Blattabacterium* endosymbionts

*Blattabacterium* is an obligate endosymbiotic bacterium belonging to the phylum Bacteroidota, which resides exclusively in cockroaches with few exceptions, such as the highly specialised cave-dwelling Nocticolidae. With the exception of *Mastotermes darwiniensis* it is also absent from termites (Bandi et al., 1995; Lo et al., 2007). It is believed that this intimate mutualistic relationship began in the common ancestor of all cockroaches more than 150 million years ago (Patiño-Navarrete et al., 2013). This endosymbiont plays an essential physiological role by recycling nitrogenous waste products into basic amino acids, thereby compensating for the nitrogen-deficient diets typical of many cockroach species (Sabree et al., 2009). *Blattabacterium* is localised within specialised cells called bacteriocytes in the host's fat body, and transmitted transovarially, leading to strict vertical inheritance (Fig. 1.1., reviewed by Mullins, 2015).

Due to its strictly vertical transmission, *Blattabacterium* has recently been employed as a complementary data source to infer cockroach phylogenetic relationships (Kinjo et al., 2018; Patiño-Navarrete et al., 2013). Given that *Blattabacterium* resides in the ovarian tissue of female cockroaches and is transmitted to offspring through the egg, it may have more opportunities to interact with the host genome and facilitate genetic exchange (Noda et al., 2025).

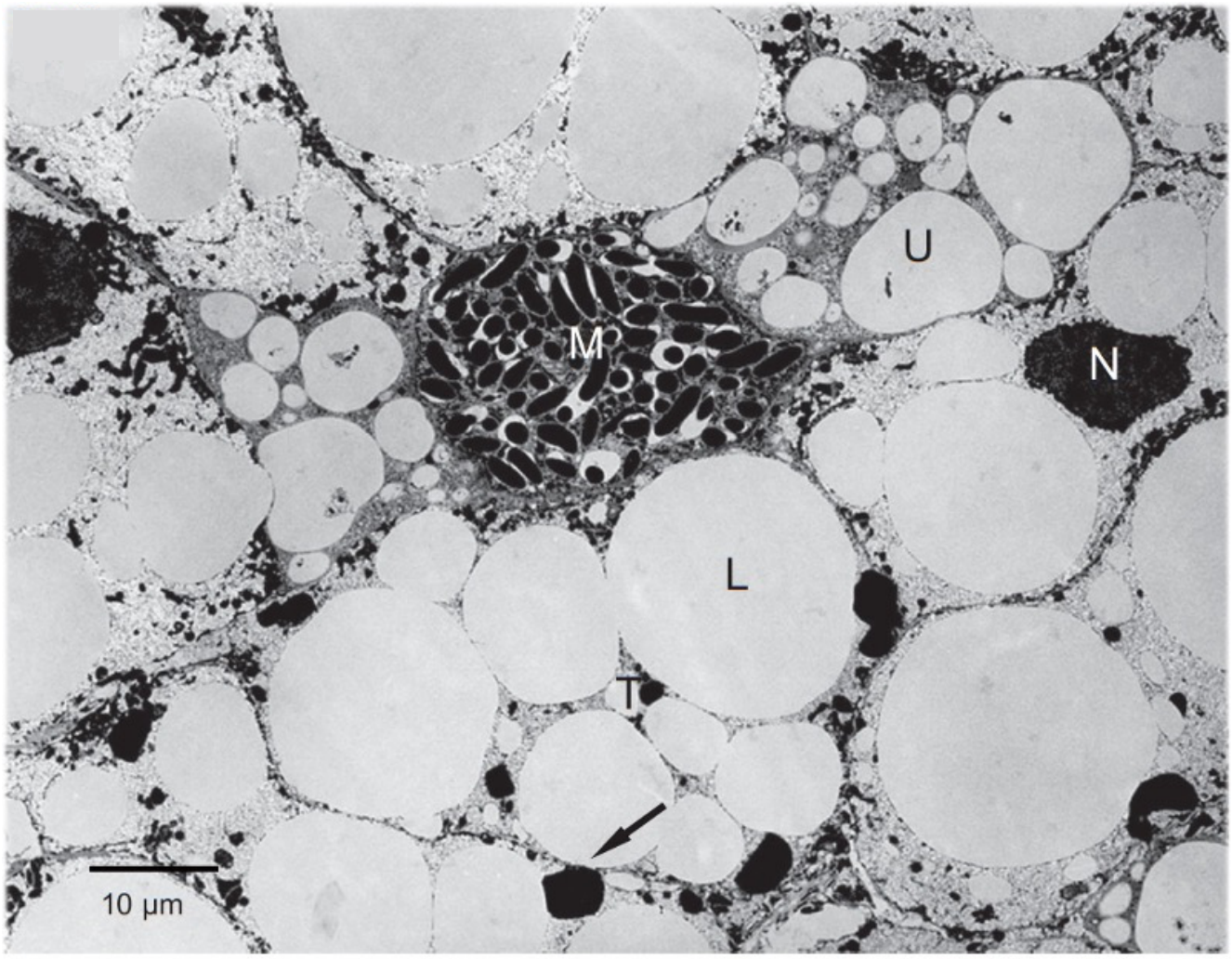



**Figure 1.1**. Ultrastructure of fat body lobes of *Periplaneta americana* with food supply. *Blattabacterium* cells are localised within mycetocytes (M). Adapted from Park et al. (2013, Fig. 3A), showing three cell types: trophocytes (T), urocytes (U) and mycetocytes (M), along with lipid droplets (L), protein components (arrow), and nucleus (N).

## 1.4 Horizontal gene transfer

Horizontal gene transfer (HGT) describes the movement of genetic material across species boundaries, in contrast to the vertical inheritance passed from parent to offspring (Good et al., 2025). In prokaryotes, like bacteria, HGT is particularly widespread. These transfers may introduce adaptive traits—such as novel metabolic capabilities, enhanced stress tolerance, detoxification mechanisms, or increased pathogenicity. If these traits offer a survival advantage, they may spread through a population via natural selection, substantially driving microbial innovation and diversity (Paquola et al., 2018). For instance, the red alga, *Galdieria sulphuraria*, which thrives in extremely acidic and high-temperature environments, appears to have acquired critical metabolic genes through HGT from archaea and bacteria, enabling its remarkable thermoacidophilic lifestyle (Schönknecht et al., 2013). While HGT from microbes occurs much less frequently in animals, ongoing research continues to uncover compelling cases across various branches of life. In insects, the whitefly, *Bemisia tabaci*, has integrated a detoxification gene from plants, allowing it to neutralise otherwise harmful plant chemical defences (Xia et al., 2021). These cases underscore HGT as an important mechanism for eukaryotic organisms to rapidly acquire advantageous traits, supplementing the slower pace of mutation-driven evolution.

*Blattabacterium* has been vertically inherited through cockroach oocytes for over 150 million years (Lo et al., 2003). During the frequent divisions of host oocytes and ootids, nuclear envelope breakdown eliminates the barrier between the nucleus and cytoplasm, creating a window in which host chromatin and cytoplasmic contents intermingle (Haraguchi and Hiraoka, 2007). Additionally, microscopic and morphometric studies have shown that *Blattabacterium* divides at surprisingly high frequencies, between 20% and 58% of cells are in the act of division at any moment, and often through asymmetric fission (Noda et al., 2024). The co-occurrence of these factors raises intriguing possibilities for molecular or genomic interactions between symbiont and host.

However, several studies investigating HGT in insects have not reported extensive transfer events in Blattodea. For instance, Li et al. (2022) surveyed most insect orders, but identified only a few HGT-derived genes in cockroaches, and none in termites. A recent preprint (Liu et

al., 2025) reported over 200 HGT events in several termite species and 2 cockroach species, predominantly derived from *Wolbachia* and not from *Blattabacterium*. Importantly, these studies relied on the alien index (AI) of annotated host genes, thereby overlooking non-coding regions that constitute most of the genome length. Further investigation of potential transfers between *Blattabacterium* and its hosts is thus warranted.

## 1.5 Phylogenetics and systematics of Blattodea

It has long been accepted that Blattodea is a member, along with mantids (Mantodea) and termites (Isoptera), of the superorder Dictyoptera, based on morphological synapomorphies present in the genitalia and wings (Beutel and Gorb, 2006; Klass, 1998; Klass and Meier, 2006; Kristensen, 1975, 1981). However, the relationships among the main clades have been debated. Recent advances in DNA sequencing and phylogenetic methodologies have strongly supported Mantodea as the sister clade to the rest of Dictyoptera (Djernæs et al., 2012; Djernæs et al., 2015; Misof et al., 2014; Terry and Whiting, 2005; Ware et al., 2008; Wipfler et al., 2019). Following the recognition of termites as eusocial cockroaches (Inward et al., 2007; Klass and Meier, 2006; Lo et al., 2000), our understanding of phylogenetic relationships within Blattodea has undergone a number of advances (Beccaloni and Eggleton, 2013b; Bläser et al., 2020; Deng et al., 2025; Djernæs and Murienne, 2022; Legendre et al., 2017; Li, 2022; Liu et al., 2023; Wang et al., 2017). Nonetheless, the systematics of the order still present numerous unresolved issues, both historical and newly identified.

Three monophyletic superfamilies, Corydioidea, Blattoidea, and Blaberoidea are widely supported by molecular data. However, consensus on the relationships among these superfamilies and within their constituent families remains elusive. Most studies recover Blaberoidea as the sister group to Corydioidea + Blattoidea (Bläser et al., 2020; Djernæs et al., 2012, 2015; Evangelista et al., 2019; Legendre et al., 2015), though some support Corydioidea (Djernæs et al., 2020; Li, 2022; Wang et al., 2017) or Blattoidea (Bourguignon et al., 2018; Liao et al., 2021) as the sister to the remaining Blattodea. Blattoidea includes the main pest cockroach genus *Periplaneta*, it has attracted the most research interest due to its inclusion of Isoptera, enabling exploration of termite evolution. The current consensus places

Cryptocercidae as the sister group to Isoptera (also known as Termitoidae), but relationships among other groups are debated.

Corydiidae + Nocticolidae form a robust clade that forms the Corydioidea. This clade includes specialist cave taxa from the genus *Nocticola*, and a number of desert specialist taxa, including *Polyphaga aegyptiaca*. Blaberoidea comprises the families Blaberidae and Ectobiidae, the latter including the pest species *Blattella germanica*. The monophyly of Blaberidae, and the paraphyly of Ectobiidae with respect to Blaberidae, are supported by recent studies (Bläser et al., 2020; Bourguignon et al., 2018; Djernæs et al., 2015; Evangelista et al., 2019; Li, 2022; Wang et al., 2017). However, which Ectobiidae subfamilies or genera are sister to Blaberidae remains contentious.

Blaberidae itself has received less focused attention, with most studies treating it as a whole. Relationships among its subfamilies remain unresolved. Beasley-Hall et al. (2021) and Lo et al. (2016) focused on Panesthiinae and Geoscapheinae, confirming their intermixing at the genus level, a finding supported by broader taxonomic analyses (Djernæs et al., 2020; Wang et al., 2017). Beyond these, no consensus exists on other subfamily relationships, as positions vary across studies with comprehensive taxon sampling (Bourguignon et al., 2018; Wang et al., 2017).

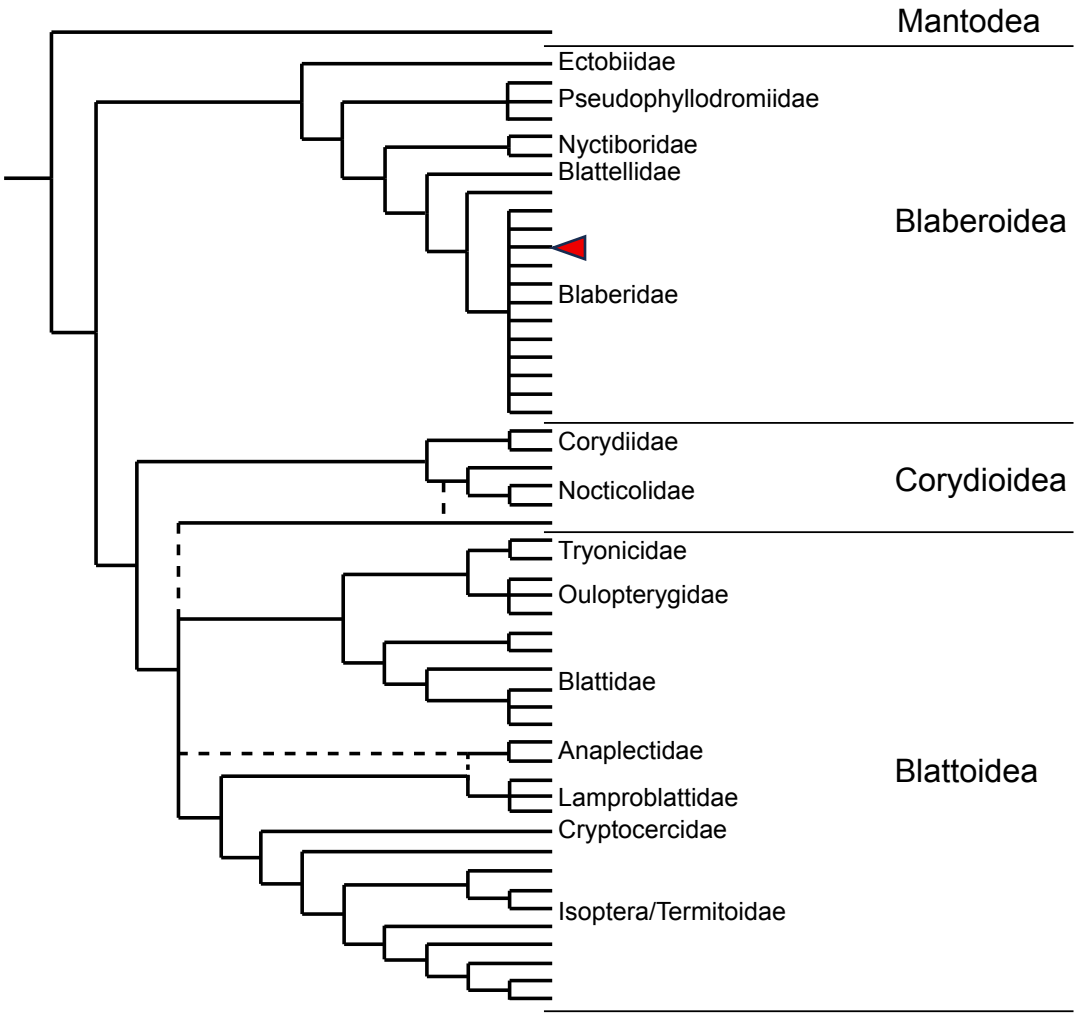



**Figure 1.2**. Current state of the family-level phylogeny of Blattodea, showing the relationships among the three major superfamilies Corydioidea, Blattoidea, and Blaberoidea. Figure adapted from Deng et al. (2025). Dashed lines indicate clades/lineages with uncertain placement. The red arrow indicate the position of Panesthiinae and Geoscapheinae, which are the main focus of this study.

## 1.6 Introduction to Panesthiinae and Geoscapheinae

Panesthiinae and Geoscapheinae are two subfamilies within Blaberidae that have long posed taxonomic challenges and remain the focus of ongoing research (The positions of these two subfamilies are shown in Fig. 1.2. Roth, 1977, 1982; Rugg and Rose, 1984). Panesthiinae are distributed across Asia and Australia, comprising 7 genera and 142 species (Beccaloni, 2024). Members of this subfamily are generally obligate wood-feeders that excavate burrows in decaying wood, which also serves as their primary food source. They exhibit subsocial behaviour, with nymphs being fed by both parents (Rugg and Rose, 1984b), although some studies have reported nocturnal movement of individuals between different colonies (Coady et al., 2025; Roth, 1977). One of the largest genera in terms of species numbers within Panesthiinae is *Panesthia*. Previous research has shown that *Panesthia* is paraphyletic, comprising several Asian and Australian lineages, and that the group colonised Australia on two separate occasions (Lo et al., 2016). However, earlier studies were limited by sampling across both Asia and Australia, and did not provide a comprehensive biogeographic or evolutionary analysis of the group.

Geoscapheinae are restricted to Australia, comprising 4 genera and 26 species, and were elevated to subfamily status on the basis of their distinctive reproductive behaviours and soil-burrowing ecology (Beccaloni, 2024). Members of this subfamily construct deep, permanent burrows in the soil and feed on dried leaf litter, which they collect and store underground. Previous studies have shown that Geoscapheinae originated from two *Panesthia* species, *P. sloanei* and *P. tryoni*, through multiple parallel lifestyle transitions during periods of aridification on the Australian continent (Perry G. Beasley-Hall et al., 2021). This phylogenetic finding fundamentally challenged earlier taxonomic classifications that were based on wing reduction, male genital morphology, and the angle of the seventh tergite. However, because of low support values at some nodes, insufficient sampling of *P. sloanei* and *P. tryoni*, no systematic taxonomic revisions were proposed.

## 1.7 Thesis outline

The primary aim of this thesis is to advance our understanding of the phylogenetic relationships within Panesthiinae and Geoscapheinae, and based on the newly inferred phylogeny, to explore the transfer of DNA from *Blattabacterium* to taxa within these groups.

Chapter 2 expands taxon sampling to investigate the evolutionary history of *P. sloanei* and *P. tryoni*, the species most closely related to Geoscapheinae. All newly sampled populations clustered with previously identified conspecific groups, confirming seven independent transitions from wood-feeding to soil-burrowing lifestyles. In addition, my results resolved the phylogenetic positions of two clades that had previously shown low support and unstable placements (*G. dilatatus* + *G. robustus* and *P. australis* + *P. obtusa*). Subsequent biogeographic analyses revealed a strong correlation between the diversification of *Panesthia* and Geoscapheinae and recognised biogeographic divisions. This study highlights the complex evolutionary history of this lineage and the repeated ecological transitions that have shaped the unique biology of these cockroaches.

Chapter 3 focuses on the phylogeny and morphological evolution of Asian Panesthiinae species. With the help of colleagues in Asia, we obtained as many Panesthiinae samples as possible and conducted phylogenetic analyses of Panesthiinae (including Geoscapheinae) using mitochondrial genomes with nuclear markers. The results divided *Panesthia* (including the nested genus *Ancaudellia*) into six clades, while *Miopanesthia* and *Salganea* were each recovered as well-supported monophyletic groups. We subsequently examined the evolution of several morphological traits in these taxa, including wings, the three male genital phallomeres (L1, L2d and R2), and the oothecal membrane.

Chapter 4 examines horizontal gene transfer events originating from the endosymbiont *Blattabacterium*. We analysed the genomes of eleven Panesthiinae + Geoscapheinae species, three other cockroaches, and four termites. The results revealed substantially more HGT insertions in Panesthiinae and Geoscapheinae (1543 - 5037) than in other cockroaches, while no *Blattabacterium*-derived HGT sequences were identified in termites. We further explored

the timing of these HGT events and identified several that occurred between ~37 million years ago, as well as several younger HGTs. In addition, we discovered cases where HGT insertions originated from multiple non-adjacent regions of the *Blattabacterium* genome but formed contiguous sequences within the host cockroach genome, and we discussed potential mechanisms underlying this phenomenon.

Finally, the General Discussion summarises the findings of this thesis by highlighting the major outcomes of each research chapter. I conclude by discussing potential future directions of research regarding the future sampling of Panesthiinae, especially for Asian genera without samples. Additionally, I also provided some suggestions for the future experimental verifications for HGT inserts.

# Chapter 2

## Increased sampling of *Panesthia* improves resolution of phylogenetic relationships among Australian wood-feeding and soil-burrowing cockroaches

**Zhuzhi Zhang**, James A. Walker, Harley A. Rose, Andreas Zwick, and Nathan Lo

## 2.1 Abstract

Soil-burrowing cockroaches (Blaberidae: Geoscapheinae) are native to Australia and have evolved an estimated seven times independently from wood-feeding ancestors, which colonised Australia ~25 Mya. The relationships among some of the major lineages of this group remain unresolved. Phylogenetically, geoscapheines are nested within two wood-feeding species *Panesthia tryoni* and *P. sloanei*, which are found in mountaintop rainforests. A number of populations of *P. tryoni* and *P. sloanei* have yet to be investigated as potentially novel sister groups of soil-burrowing taxa. We therefore collected 30 new samples of *P. tryoni* and *P. sloanei* from rainforests along the east coast of Australia and generated whole mitochondrial genomes and nuclear ribosomal markers from them. These represent the most well-known populations, including 16 locations not previously sampled. These were combined with existing genomic data from Australian *Panesthia* and Geoscapheinae for phylogenetic and biogeographic analyses. All newly sampled populations clustered with previously identified conspecific groups, confirming seven independent wood-to-soil lifestyle shifts. Notably, we recovered for the first time with strong support: 1) *Geoscapheus dilatatus* + *G. robustus* as the earliest divergence in the group; 2) the placement of *P. australis* + *P. obtusa* as sister to the *P. tryoni* group. These results suggest that the closest wood-feeding relatives of *G. dilatatus* + *G. robustus* have since gone extinct. We found that *P. sloanei* was split into two clades, with soil-burrowers evolving from only one of these clades. We examined the distribution of these lineages and found a strong correlation with recognised biogeographic divisions. This study highlights the complex evolutionary history and repeated ecological transitions within this unique group of cockroaches.

## 2.2 Introduction

The phylogenetic relationship between the main extant clades within the Blattodea is gradually being clarified, however, there are still many unresolved issues regarding the relationships both between and within individual families (Beasley-Hall et al., 2021; Beccaloni, 2014; Legendre et al., 2013, 2015, 2017; Lo et al., 2016; Maekawa et al., 2003; Wang et al., 2017). The blaberid subfamily Geoscapheinae contains 26 species of soil-burrowing cockroaches, all of which are native to Australia and found primarily in scrubland

areas of Queensland, with some species also found in western NSW, Victoria, South Australia and Western Australia (Fig. 2.1a, c). These species formed self-dug, permanent, underground burrows, typically in sandy soils (Roth, 1977; Rugg and Rose, 1984b, 1991), and feed on dried leaf litter transported to the burrow from the surface. Offspring typically remain with their mother in the burrow for several months after birth. Geoscapheinae was raised as a subfamily from within Panesthiinae, due to their unique reproductive behaviour and lifestyle (Rugg and Rose, 1984b, 1991). However, the monophyly of Geoscapheinae has been challenged in recent decades, with molecular phylogenetic studies showing that species within Geoscapheinae are the result of 6-8 multiple independent parallel evolution events from wood-feeding *Panesthia* species (Beasley-Hall et al., 2021; Legendre et al., 2014; Lo et al., 2016; Maekawa et al., 2003). Previous studies have indicated that the lineage that gave rise to Geoscapheinae arrived in Australia approximately 29 million years ago (Beasley-Hall et al., 2021; Lo et al., 2016). Because it arrived earlier than a second *Panesthia* lineage found in Australia, it is referred to as the "first wave" of Australian *Panesthia* (Adams et al., 2024) (see Fig. 2.2b, d for examples of representative first wave *Panesthia*).

The two wood-feeding species from which soil-burrowers are known to have evolved, *Panesthia sloanei* and *P. tryoni*, are found in isolated rainforested areas across the eastern seaboard of northern Australia, typically in mountainous areas (Roth, 1977). *Panesthia sloanei* is a completely apterous species, found as far south as Paluma Range National Park and as far north as Mount Finnigan (Beasley-Hall et al., 2021; Rugg and Rose, 1984, 1991). *Panesthia tryoni* is further divided into two subspecies *P. t. tryoni* and *P. t. tegminifera*, both of which are apterous. *Panesthia tryoni tryoni* is distributed between Nightcap National Park in the south and Cathu State Forest in the north. *Panesthia tryoni tegminifera* is distinguished from *P. t. tryoni* through the presence of tegmina; this subspecies is found from Bald Mountain Forest Reserve in the south, extending north to the Junuy Juluum National Park (Beasley-Hall et al., 2021; Roth, 1977; Rugg and Rose, 1984). Two other species of *Panesthia* that show affinity to Geoscapheinae are the sister taxa *P. australis* and *P. obtusa*; each of these species appears to be monophyletic (Beasley-Hall et al., 2021; Roth, 1977; Rugg and Rose, 1984).

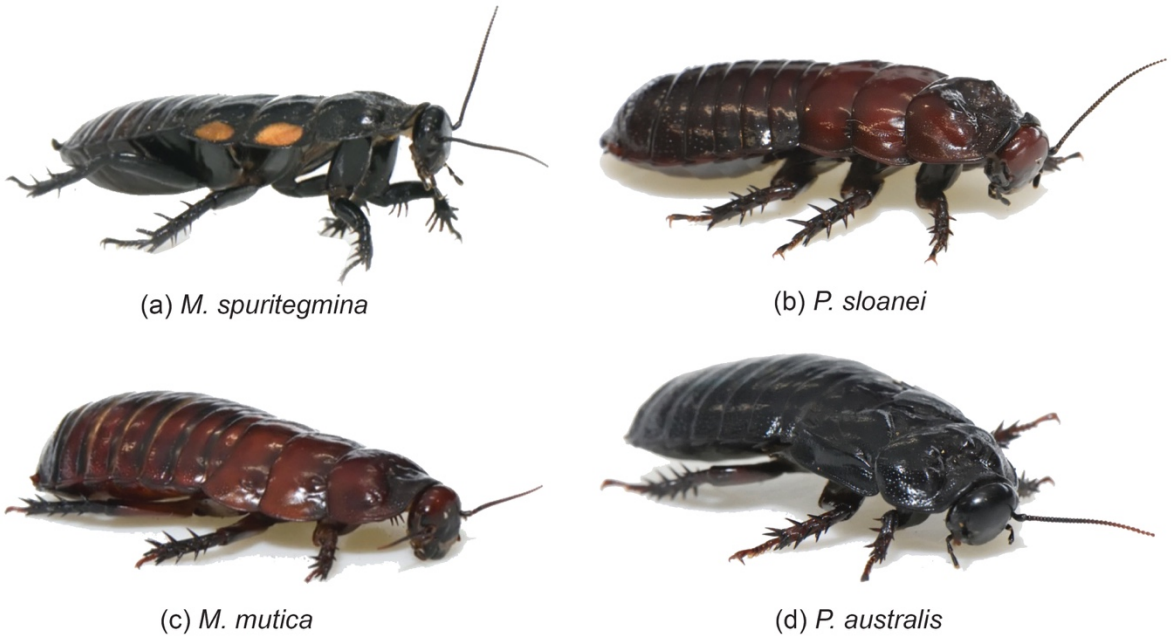



**Figure 2.1**. The appearance of four representative species of *Panesthia* and Geoscapheinae. (a) *Macropanesthia spuritegmina*, (b) *Panesthia sloanei*, (c) *Macropanesthia mutica,* (d) *Panesthia australis*. (a) and (c) are soil-burrowing species, while (b) and (d) are wood-feeding species. All photographs by Yi Kai Tea.

Several phylogenetic studies have investigated the evolution of soil-burrowing cockroaches (Beasley-Hall et al., 2021; Lo et al., 2016; Maekawa et al., 2003). To date, sampling of *P. sloanei* and *P. tryoni* in these studies has been somewhat limited. A number of populations of these species across their respective ranges have yet to be included. The possibility, therefore, remains that some of these populations may represent the sister group to some soil-burrowing clades, and thus represent additional cases of parallel evolution. For example, the clades comprising *G. dilatatus*/*G. robustus* and *Macropanesthia*. *kinkuna*/*G. crenulatus*/*M. mackerrasae* have no known extant wood-feeding sister groups.

Furthermore, the causes of the disjunct distribution of *P. sloanei* and *P. tryoni* in rainforested areas are yet to be investigated in detail. Previous studies indicate that divergent populations within each of these species diverged up to ~20 million years ago (Beasley-Hall et al., 2021; Lo et al., 2016; Maekawa et al., 2003). This coincides with the onset of arid conditions on the Australian continent (Martin, 1982, 1998, 2006), which are the likely causes of the differentiation and isolation of these rainforested areas. Further studies including additional taxa for these species, in addition to comparisons with other animal species with similar distributions, are required to resolve the biogeography of these species.

Finally, the taxonomy of *P. sloanei* and *P. tryoni* requires revision (Beasley-Hall et al., 2021; Lo et al., 2016; Maekawa et al., 2003). Molecular investigations into additional populations of *P. sloanei* and *P. tryoni* are likely to provide relevant information regarding the number of cryptic lineages present among each species. Such investigations could lead to a more accurate understanding of their species boundaries, allowing us to determine the number of transitions from wood-feeding to soil-burrowing lifestyles and contribute to subsequent genomic analyses.

We sequenced mitochondrial genomes and nuclear ribosomal gene complements from a number of populations of *P. tryoni* and *P. sloanei* that have yet to be investigated and compared these data with those previously obtained from other members of Panesthiinae and Geoscapheinae. We aimed to: (1) explore whether there are any undetected sister groups of Geoscapheinae genera; (2) estimate the divergence times of each wood-feeding and soil-burrowing lineage, to better understand the phylogeography of these groups.

## 2.3 Methods

### 2.3.1 Taxon sampling

Our study aimed to build on previous phylogenetic work by extending the sampling of the *P. tryoni* and *P. sloanei* species complexes. Samples were collected between 1987 to 2021, and stored in 70–100% ethanol. We included 30 samples of these two species, from 13 locations that have not previously been included in molecular analyses (6 for *P. sloanei* and 7 *for P. tryoni*), plus one *Panesthia* sp. from the 'second wave' clade. We combined data from these species with 41 samples previously sequenced in Beasley-Hall et al. (2021) and Lo et al. (2016), covering 4 (out of 4) genera and 24 (out of 26) species.

### 2.3.2 DNA sequencing and assembly

DNA was extracted from leg muscle tissue to avoid contamination from endosymbiotic *Blattabacterium cuenoti* bacteria in the abdominal fat bodies (Kinjo et al., 2015). DNA sequencing was outsourced to the Australian National Insect Collection, Canberra, utilising an approach suitable for highly fragmented historical DNA (see Jin et al., 2020; Zwick and Zwick, 2023). In summary, genomic DNA was extracted from proteinase K digested tissue, using a silica filter-based approach in a 384-well format. Ligation-based whole-genome shotgun DNA sequencing libraries were built using up to 5 ng of extracted DNA, by an acoustic liquid handler (Echo 525; Beckman Coulter, California, USA) to miniaturise reaction volumes for increased reaction efficiency. DNA libraries of different samples were pooled equimolar and sequenced at the Australian National University's BRF sequencing centre on an Illumina NovaSeq 6000 platform, using an S1 flow cell and a 300-cycle sequencing kit.

DNA raw reads were assembled *de novo* using SPAdes (Bankevich et al., 2012) with default settings and sampling k values of 33, 55, 77, 91, and 121. Contigs were assembled using

GENEIOUS (v.11.0.18, http://geneious.com), by mapping to reference mitogenomes of closely related species with default settings and medium sensitivity. Reference sequences were chosen to represent the closest known sister taxon to each sample, based on the phylogenetic framework of Beasley-Hall et al. (2021), sequences were sourced from GenBank. We extracted the consensus sequences when a single near-complete mitogenome contig was available; when it resulted in multiple contigs, the consensus sequence of multiple contigs was extracted, selecting the bases with the highest representation. The nuclear ribosomal sequences 18S rRNA gene, Internal Transcribed Spacer 1 (ITS1) and 28S rRNA gene were assembled following the same process. Previously sequenced read libraries from Beasley-Hall et al. (2021) were also assembled for the nuclear gene 28S following the same process. In total, 30 ingroup samples were sequenced successfully.

### 2.3.3 Phylogenetic analyses

We combined data generated in this study with mitochondrial genome and nuclear data (18S + ITS1) from Beasley-Hall et al. (2021) to create per-gene alignments using MUSCLE algorithm (Edgar, 2004) with default settings and excluded overlapping sections of genes. Mitochondrial and nuclear datasets consisted of whole mitochondrial genomes (13,571 bp) and nuclear data (1,533 bp), respectively. Protein-encoding alignments were checked visually to ensure they conformed to the correct reading frame, using invertebrate mitochondrial and standard genetic codes for each dataset. Substitutional saturation was not detected previously (Beasley-Hall et al., 2021), and the third codon position was retained in the alignment. We conducted the partitioned Maximum-likelihood (ML) analyses using IQ-TREE 2.2.2.4 (Minh et al., 2020) with the default setting and the default ModelFinder (Kalyaanamoorthy et al., 2017), and the branch support was completed by running 10,000 replicates with ultrafast bootstrap approximation (UFBoot; Hoang et al., 2018) and the SH-like approximate likelihood-ratio test (SH-aLRT) with 1000 iterations.

We estimated evolutionary timescales using BEAST2 (v. 2.7.4, Bouckaert et al., 2014), and included one fossil calibration, '*Gyna' obesa* Piton, to infer the divergence times, using an exponential distribution and soft maximum bounds, as the maximum age was not determined

very accurately (Ho and Phillips, 2009). This fossil was described as the oldest Blaberidae (Evangelista et al., 2017), and widely used as stem Blaberidae to calibrate molecular dating analyses (Beasley-Hall et al., 2021; Bourguignon et al., 2018; Evangelista et al., 2019). We followed Beasley-Hall et al. (2021) and Evangelista et al. (2019), and included '*Gyna' obesa* as the stem-Blaberidae with a minimum age of 57.7 Mya hard maximum bound and 145 Mya soft maximum bound to represent the first Blaberidae (Lin, 1980). Analyses were conducted using the optimised relaxed clock and the birth-death tree, with the built-in BEAST model test, performed for 100 million generations and convergence of the stationary distribution was checked using ESS values of >200 in Tracer (v. 1.7.2, Rambaut et al., 2018). The maximum clade credibility tree from the combined runs was produced using Treeannotator using a 10% burn-in (Bouckaert et al., 2014).

We used BioGeoBEARS (Matzke, 2013) to reconstruct the ancestral ranges of Australian Panesthiinae and Geoscapheinae. We pruned the estimated phylogram only to include one representative of each monophyletic species or lineage. Since *G. dilatatus* and *G. robustus* formed a widely distributed soil-burrowing lineage with no extant wood-feeding relatives, we excluded them from the biogeographical analyses to avoid any misleading conclusions. As *P. tryoni* and *P. sloanei* are widely distributed and include multiple lineages with significant genetic differences, the collected locations were used to differentiate the distinct lineages. Depending on the rainforest endemism and bioregions (Bryant and Krosch, 2016), *Panesthia* and Geoscapheinae were coded as belonging to the 'North QLD', 'Central QLD', 'South QLD', 'North NSW', 'Central and South NSW + Victoria', or 'Dry Inland' regions. We implemented three models with default parameters and settings of BiogeoBEARS: Dispersal-Extinction-Cladogenesis (DEC, Ree and Smith 2008), ML implementations of Dispersal-Vicariance Analysis (DIVALIKE, Ronquist, 1997) and Bayesian Analysis of Biogeography (BAYAREALIKE or BayArea, Matzke, 2014). We assessed the relative probabilities of models with the Akaike information criterion (AIC) and corrected Akaike (AICc). Because of ongoing debate surrounding the use of the jump dispersal parameter (+J), we did not include models containing a jump dispersal parameter (Ree and Sanmartín, 2018).

## 2.4 Results and discussion

Our findings indicate that, following their invasion from Asia (Maekawa et al., 2003), first-wave Australian *Panesthia* + Geoscapheinae likely became widely distributed across eastern Queensland to northern New South Wales, and also distributed in the current arid areas (Fig. 2.3). Subsequently, and consistent with previous results (Beasley-Hall et al., 2021; Djernæs et al., 2020; Legendre et al., 2013, 2015, 2017; Lo et al., 2016; Maekawa et al., 2003), this group diversified into three primary lineages: (i) an early diverging *G. dilatatus* + *G. robustus* lineage, whose closest wood-feeding relatives have likely gone extinct; (ii) a clade comprising two divergent groups of *P. sloanei*, one of which was the sister group to 9 Geoscapheinae species, including *M. rhinoceros* (Fig. 2.2), and (iii) a clade comprising representatives of various *P. tryoni* populations, their closely related Geoscapheinae taxa, and *P. australis* + *P. obtusa*. Furthermore, our results indicate that these three major clades are subdivided into seven minor clades: an early branching *Geoscapheus* group (*G. dilatatus* group in Fig. 2.2, clade F in Beasley-Hall et al., 2021); a *P. sloanei* 'southern' group (*P. sloanei* + sister soil burrowers, *P. sloanei* 'southern' group in Fig. 2.2, a part of clade G in Beasley-Hall et al., 2021); and a *P. sloanei* 'northern' group (*P. sloanei*, *P. sloanei* 'northern' group in Fig. 2.2, a part of clade G in (Beasley-Hall et al., 2021); an early branching *Panesthia* group (*P. australis* + *P. obtusa* group in Fig. 2.2, clade E in Beasley-Hall et al., 2021); a *P. tryoni* 'southern' group (*P. t. tryoni* + *P. t. tegminifera* + sister soil burrowers, *P. tryoni* 'southern' group in Fig. 2.2, clade B in Beasley-Hall et al., 2021); a *P. tryoni* 'central' group (*P. t. tryoni* + sister soil burrowers, *P. tryoni* 'central' group in Fig. 2.2, clade A in Beasley-Hall et al., 2021); a *P. tryoni* 'northern' group (*P. t. tryoni* + sister soil burrowers, *P. tryoni* 'northern' group in Fig. 2.2, clade C + D in Beasley-Hall et al., 2021). By substantially increasing sampling across the known distributions of *P. sloanei* and *P. tryoni*, we confirmed there are seven individual origins of the soil-burrowing lifestyle (Beasley-Hall et al., 2021; Legendre et al., 2015, 2017; Lo et al., 2016). However, we did not identify any novel *P. sloanei* or *P. tryoni* lineages that represent additional sister groups to soil-burrowing lineages. The distribution patterns of *P. sloanei* and *P. tryoni* are similar to those of classic wingless and closed-rainforest species, for example, the flat bug subfamily Mezirinae (Family: Aradidae), being heavily restricted by the extent of rainforest habitat (Monteith, 1997). Its northernmost range does not extend into the Cape York Peninsula Zone, while in the south, it reaches a limit near Barrington Tops (Monteith, 1997). This is discussed further below.

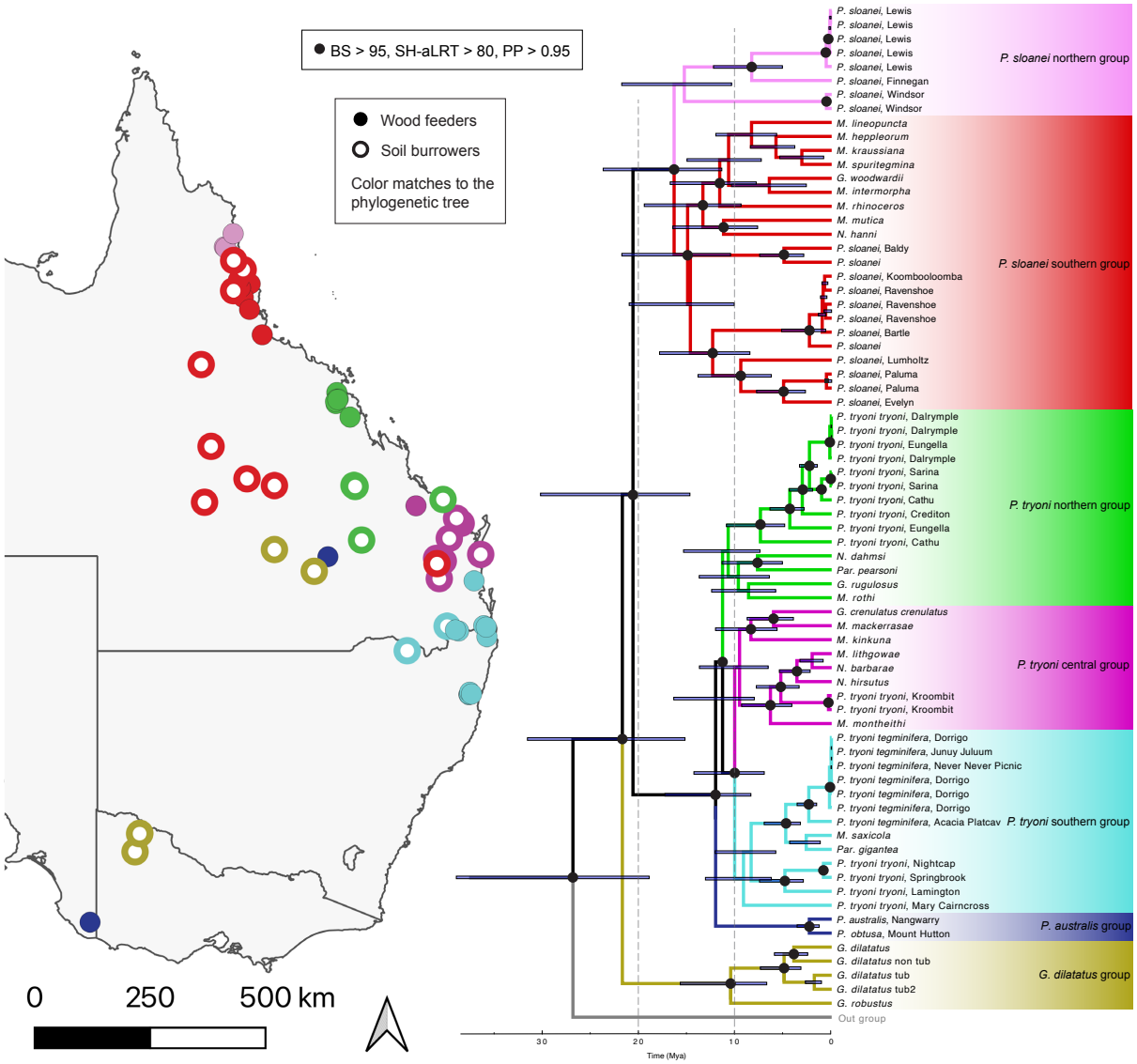



**Figure 2.2**. Dated phylogeny of *Panesthia* and Geoscapheinae inferred from complete mitochondrial genomes and nuclear ribosomal markers in BEAST and IQTREE. The evolutionary timescale (millions of years ago) was inferred using one fossil calibration, '*Gyna' obesa* Piton (Piton, 1940). Note that the locations shown on the tree are not necessarily reflective of the total distribution of some species. For example, *P. australis* has a distribution in temperate forests from Southern Queensland to Western Victoria/eastern South Australia. All specimens of the genus *Panesthia* were wood-feeders, while all specimens from other genera were soil-burrowers. BS, ultrafast bootstrap; PP, posterior probability; SH-aLRT, SH-like approximate likelihood-ratio test. Sampling locality colours correspond to groupings based on the phylogeny. The distribution of sampling localities in eastern Australia are shown on the left of the figure. Grey bars at each node represent 95% highest posterior density (HPD) divergence times.

### 2.4.1 Biogeography and evolution of *G. dilatatus* + *G. robustus*

The *G. dilatatus* + *G. robustus* clade separated from other lineages at ~25 Mya, which was reconstructed as the earliest diverged monophyletic group of the first-wave Australian *Panesthia* with high support for the first time (Fig. 2.2). We infer that the *G. dilatatus* + *G. robustus* clade may have originated from an extinct wood-feeding lineage, presumably prior to their divergence ~14 Ma (95% credible interval [CI] 5.63 – 24.03 Ma, Fig. 2.2). An alternative scenario is that the wood-feeding lineage that gave rise to this soil burrowing clade lives on in their descendents *P. sloanei* and *P. tryoni*. *Geoscapheus dilatatus* and *G. robustus* exhibit the ability to adapt to a broader range of temperature and humidity conditions compared to other Geoscapheinae species (Beasley-Hall et al., 2018, 2021a), which has allowed them to colonise areas with more variable environmental conditions than other soil-burrowing clades. This adaptability has enabled them to disperse across regions including NSW, Victoria, South Australia, and Western Australia (Fig. 2.3).

The wood-feeding ancestors of *G. dilatatus* + *G. robustus* may have been distributed in areas that no longer contain habitats suitable for their existence (Fig. 2.3). This contrasts with the case of *P. sloanei* and *P. tryoni*, which continue to exist in rainforested areas. Based on the results of our analyses, we also cannot exclude the possibility that the ancestors of *G. dilatatus* and *G. robustus* arrived in Australia independently from their sister group, but the most parsimonious explanation is that there was a single invasion.

In contrast to the widely distributed *G. dilatatus* and *G. robustus*, the distribution of soil-burrowing species that form close relationships with *P. sloanei* and *P. tryoni* is more limited, perhaps with the exception of *M. kraussiana*, which is found as far south as Western NSW. We hypothesise that the warm and humid but fragmented rainforests in the north during the early stages of aridification likely limited the dispersal of most species closely related to *P. tryoni* and *P. sloanei*, as discussed further below.

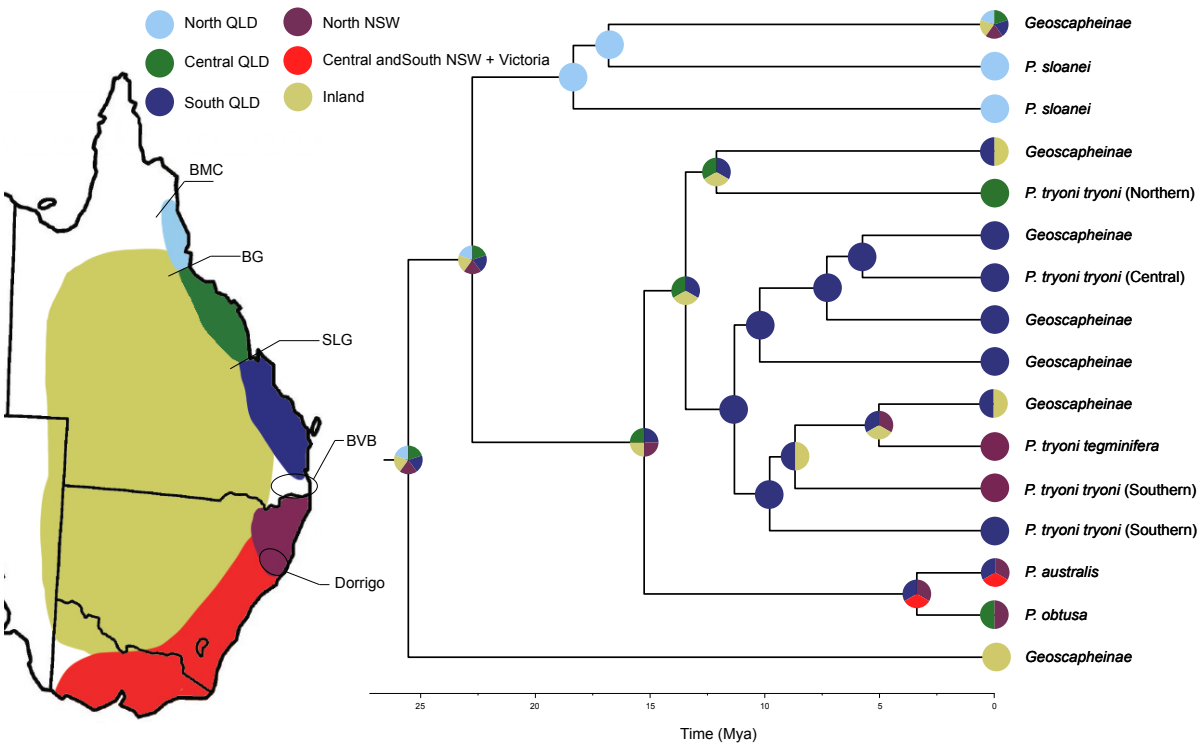



**Figure 2.3**. Ancestral geographic ranges of *Panesthia* and Geoscapheinae, estimated over the Bayesian chronogram inferred from whole mitogenomes and nuclear ribosomal markers in BEAST. The topology of this tree is consistent with that presented in Fig. 2.2, with each Geoscapheinae node comprising several Geoscapheinae species from the corresponding position in Fig. 2.2. Circles at nodes represent the most likely ancestral range, estimated using a dispersal-vicariance model in BioGeoBEARS. Multicoloured circles denote a range spanning multiple bioregions, tips are labelled with present-day distributions. Sections of pie charts do not represent the proportion of species distributions. BMC, Black Mountain Corridor; BG, Burdekin Gap; SLG, Saint Lawrence Gap; BVB, Brisbane Valley Barrier (Bryant and Krosch, 2016); Dorrigo, Dorrigo National Park.

### 2.4.2 Biogeography and evolution of *P. sloanei*

The common ancestor of all extant *P. sloanei* separated with the common ancestor of *P. tryoni* and *P. australis* + *P. obtusa* at ~23 Mya (Fig. 2.2). Although the distribution boundaries of *P. sloanei* and the *P. tryoni* groups coincide with the Burdekin Gap (BG) (Fig. 2.3), their divergence time is much earlier than those estimated by most studies of other organisms, including the yellow-bellied glider (*Petaurus australis*, Brown et al., 2006, 0.6–1.4 Mya), Fawn-footed mosaic-tailed rat (*Melomys cervinipes*, Bryant and Fuller, 2014, 0.5–1.6) and Brush-tailed rock-wallaby (*Petrogale penicillate*, Potter et al., 2012, 0.05–2.9). These shallower divergence times are likely due to the relatively high mobility of these vertebrate taxa, compared with the dispersal ability of the cockroach group examined in this study. On the other hand, a study focusing on assassin spiders (Archaeidae) showed that two groups found on either side of the St. Lawrence Gap diverged at 34-51 Mya. The authors proposed that the rainforests in central eastern Queensland had already begun to be isolated by the St. Lawrence Gap before the formation of the Burdekin Gap (Rix and Harvey, 2012), which is consistent with our results.

All extant *P. sloanei* were found in north Queensland and formed two monophyletic clades. These two clades were separated by the Black Mountain Corridor (BMC, Bryant and Krosch, 2016) around 18 Mya ago (Fig. 2.2). Other invertebrate taxa, such as the earthworm *Terrisswalkerius* sp., found from the Paluma Range to Cape York Peninsula, display evidence of an even earlier separation by the BMC (31–84 Mya, Moreau et al., 2015). However, most examined species have substantially younger divergences across the BMC, including woodland agamid lizards (*Diporiphora australis*, 4.7–9.2 Mya, Edwards and Melville, 2010), rainforest frogs (*Litoria*, 2–12 Mya, Bell et al., 2012) and the midge (*Echinocladius martini*, 6-7 Mya, Krosch, 2011). Unlike the multiple origins of soil-burrowing in *P. tryoni* groups, only one extant soil-burrowing clade originated from the *P. sloanei* clade distributed to the south of the BMC. *Panesthia sloanei* samples from newly surveyed distributions did not form new sister group relationships with any soil-burrowing taxa.

The phylogenetic relationships within the *P. sloanei* group are correlated with their geographic distributions, exhibiting a distinct north-to-south progressive separation. After the northernmost clade was isolated from other clades by the Black Mountain Corridor around 18 Mya, the southern clade diverged around 17 Mya into the soil-burrowing clade and other *P. sloanei* lineages that have retained the original wood-feeding lifestyle. Although these wood-feeding clades have lived in isolation for several million years, their behaviours and morphologies remain very similar. Even though the distribution ranges of different clades of the *P. sloanei* overlap with those of the hypothesized Pleistocene rainforest refugia (e.g., Windsor Uplands, Carbine Uplands, and Paluma Ranges; Hilbert et al., 2007; Nix and Switzer, 1991; Schneider and Moritz, 1999), the periods of isolation among them are much older than the Pleistocene and the Last Glacial Maximum. Therefore, we speculate that the formation of multiple *P. sloanei* lineages and the transition to the soil-burrowing lifestyle were driven by aridification resulting from the northward movement of the Australia continent. Only in regions that maintained rainforest refugia until the Last Glacial Maximum did wood-feeding lineages survive to the present. The stable conditions in these refugia helped preserve the *P. sloanei* clades' consistent behavioural and morphological traits. A similar pattern has also been found in some other rainforest species in the same region (Cunningham and Moritz, 1998; Oberski et al., 2018; Schneider et al., 1998; Schneider and Moritz, 1999). For example, the ancestors of midge species *Echinocladius martini* (Diptera: Chironomidae) were separated by the BMC about 8 million years ago, and gradually formed separate clades in many scattered rainforest refuges. The notion that these clades represent the same species is challenged by molecular evidence (Krosch, 2011). Krosch (2011) also provides evidence of a lineage of midge re-entering the Lewis region after divergence. This may be a phenomenon unique to flying insects with stronger long-distance dispersal capabilities than *P. sloanei*.

### 2.4.3 Biogeography and evolution of *P. tryoni* and *P. australis* groups

All newly sampled *P. tryoni* taxa have clustered with their conspecifics (Beasley-Hall et al., 2021). The common ancestor of *P. tryoni* clades and the *P. australis* group was reconstructed to be distributed in the region between the Saint Lawrence Gap and the Dorrigo National Park region. Given that the Australian continent was still relatively mesic at the time (15

Mya), the ancestor of this clade likely also lived in areas that are now part of the arid inland region (Fig. 2.3, Carpenter et al., 2011). The 'northern' group diverged from the other groups at ~13 Mya (Fig. 2). Finally, the 'central' and 'southern' groups separated at ~11 Mya (Fig. 2.2). The *P. tryoni* 'northern' group includes wood-feeding species distributed in rainforests around central Queensland, around the west of Mackay, QLD or in the Central Mackay Coast region (CMC, Australian Government Department of Climate Change, 2023), or between the Burdekin Gap and the St. Lawrence Gap. Related soil burrowers were identified as *N. dahmsi*, *Par. pearsoni*, *M. rothi* and *G. rugulosus*. These soil burrowers were found to the south of their *P. tryoni* relatives in both dry inland and coastal regions.

In the 'central' group, there are two groups of soil burrowers, *G. crenulatus*, *M. mackerrasae*, *M. kinkuna* forming the first group, and *M. lithgowae*, *N. barbarae*, *N. hirsutus* and *M. montheithi* forming the second, in which was nested *P. tryoni* from Kroombit Tops National Park (Fig. 2.2). This nested position is not, in our opinion, the result of wood-feeding re-evolving from a soil-burrowing ancestor. Rather, the wood-feeding sister groups of *M. monteithi* and (*G. crenulatus* + *M. mackerrasae* + *M. kinkuna*) are yet to be identified, or have become extinct. Soil burrowers of this group are found in the coastal regions or inland regions, but do not cross the Great Dividing Range.

The *P. tryoni* 'southern' group was found south of Brisbane, and includes the soil-burrowers *M. saxicola* and *Par. gigantea* (Fig. 2.2). These soil burrowing lineages do not appear to have dispersed to the same degree as some other soil-burrowers (e.g. *M. kraussiana*), being distributed near closely related *P. tryoni* lineages. Most representatives of the subspecies *P. t. tegminifera*, defined by their possession of vestigial forewings, were found within this group. The ancestor of the clade comprising *M. saxicola*, *Par. gigantea*, and *P. t. tegminifera* either possessed similar vestigial forewings, in which case the former lost this character, or possessed more extensive wings, which were lost completely in the former but reduced in the latter.

The *P. tryoni* 'southern' group is primarily distributed in the tropical rainforests of the Great Dividing Range, spanning from central-eastern to southeastern Queensland and northern New South Wales. This region is distinctly divided into two distribution ranges: *P. t. tryoni* occurs

in the McPherson Range along the southern boundary of the BVB, while *P. t. tegminifera* is predominantly found further south in the Dorrigo Plateau. *Panesthia tryoni tryoni* is not monophyletic. One *P. t. tryoni* sample (Mary Cairncross), which is distributed just north of the BVB, forms a sister group with *P. t. tryoni* + *P. t. tegminifera* + soil-burrowers, while the remaining individuals are found around the McPherson Range and Tweed Volcano. This distribution pattern may reflect a north-to-south fragmentation process of the *P. tryoni* 'southern' group as a result of increasing aridity, leading to isolation within rainforest remnants. Specifically, after being initially separated by the BVB, the ancestor of the *P. tryoni* 'southern' group survived in closed rainforest refugia preserved in a volcanic uplifted highland area (Sutherland, 2011). A similar pattern has been observed in the southern leaf-tailed gecko (*Saltuarius*), although species in this genus found near the McPherson Range and Dorrigo Plateau are not closely related (Couper et al., 2008). This may be because the ancestors of *Saltuarius* were not widely distributed or underwent only a single transition from rainforest-dwelling to rock-dwelling habitats (Couper et al., 2008). Additionally, a recent study on the Australian eusocial beetle lineage, *Austroplatypus*, has revealed a similar pattern of differentiation (Bickerstaff et al., 2025). This divergence is driven by the combined effects of habitat fragmentation and the limited dispersal capability of *Austroplatypus*. The study also documented instances where morphology remains conserved, yet molecular data unveil significant lineage divergence.

Notably, the earliest branching lineages of *P. t. tegminifera* are located in the north, towards the inland side of the McPherson Range, geographically isolated from other *P. t. tegminifera* populations. This distribution pattern may reflect the preservation of different populations in separate refugia during the formation of the Dorrigo Plateau and the McPherson Range. Subsequently, during rainforest range fluctuations, these populations likely experienced range expansions, although they did not reach the point of hybridisation. Although the region comprising the McPherson Range (or Macleay–McPherson Overlap) to the Dorrigo Plateau exhibits relatively high biodiversity (Colgan et al., 2009; Crisp et al., 2001), the absence of extensive arid zones between rainforests may have prevented the formation of stable geographic barriers. Consequently, it has not received adequate attention, and only a few studies have documented distribution trends similar to ours. Studies of the lizards *Ctenotus taeniolatus* and *Oedura lesueurii* also found that different species occupy distinct distribution ranges in contiguous rainforests (Colgan et al., 2009); and beech

*Nothofagus moorei*, which is distributed in several distinct areas due to their altitude requirements (Bale and Williams, 1993).

The *P. tryoni* 'northern' group was found on mountains that are located within a drier region, between the Burdekin Gap (BG) and the St. Lawrence Gap (SLG; Bryant and Krosch, 2016) (Fig. 2.3). The *P. tryoni* northern group separated from the other two groups at around 13 Mya (Fig. 2.2). Compared to other species isolated by these geographical barriers, our divergence estimates are all substantially older. For example, much later divergences were found in mammals, including wallabies (*Lampropholis delicata*, BG: 0.05–2.9, StLG: 0.1–4.4, Potter et al., 2012) and melomys (*Melomys cervinipes*, BG: 0.5–1.6, Bryant and Fuller, 2014). The larger body size of these mammals may have allowed them to move more readily between regions, allowing maintenance of gene flow during expansions and contractions, compared with small or slow invertebrates like *Panesthia* (Bryant and Krosch, 2016).

*Panesthia australis* and *P. obtusa* formed a monophyletic lineage, which was the sister group of the entire *P. tryoni* clade (Fig. 2.2). The distribution of *P. australis* and *P. obtusa* overlaps considerably with the *P. tryoni* southern group, as well as that of *P. cribrata* (Adams et al., 2024) (Fig. 2.3). More detailed habitat information for these species is necessary to explain the underlying causes of this overlap. *Panesthia australis* and *P. obtusa* were found to have different tolerances to humidity and temperature than other first-wave Panesthiinae (Beasley-Hall et al., 2018), but the causes of these differences are not understood.

### 2.4.4 Taxonomic implications

Our results support the monophyly of the early branching *G. dilatatus* + *G. robustus* group, and the early branching *P. australis* + *P. obtusa* group. However, clades comprising *P. tryoni* and *P. sloanei* clearly require taxonomic revisions. *Panesthia tryoni* is polyphyletic with respect to several species and should be divided into several species based on their phylogenetic relationships. In particular, *P. tryoni* of the northern and central groups could each be independently revised as two new species, while the southern group *P. t. tegminifera* and *P. t. tryoni* could be raised as two distinct species.

*Panesthia sloanei* should be divided into at least two species. The BMC is the main barrier of *P. sloanei*, those found in the north and the south of the BMC could be divided into two species. Meanwhile, samples found in Mount Windsor and Baldy Mountain Forest Reserve have been shown to have been isolated from other *P. sloanei* for ~16 million years, and therefore might best be classified as new species. Morphological work that is beyond the scope of this study will be required prior to any new species designations.

## 2.5 Conclusion

This study, which incorporates more extensive sampling of *P. sloanei* and *P. tryoni* than studies undertaken previously, offers new insights into the evolutionary history of the first-wave Australian *Panesthia*. We identified seven monophyletic clades, and clarified that *G. dilatatus* and *G. robustus* form an early-diverging branch distinct from the first-wave Australian *Panesthia*. We divided *P. tryoni* and its closely related soil burrowers into three groups, while *P. australis* and *P. obtusa* together form a separate group. Meanwhile, *P. sloanei* and its related soil burrowers are split into two groups. We also analysed the geographic barriers involved in the divergence of *P. tryoni* and *P. sloanei*, and compared them with other species distributed in the same regions. Our findings suggest that the current distribution pattern of the first-wave Australian *Panesthia* is the result of both geographic isolation and aridification.

# Chapter 3

## Phylogenetics, Biogeography and Morphological Evolution of Asian Panesthiinae

**Zhuzhi Zhang**, Kokuto Fujiwara, Kiyoto Maekawa, Zhuli Cheng, Thomas Bourguignon, Zongqing Wang, Nathan Lo

## 3.1 Abstract

Panesthiinae and Geoscapheinae are two sister subfamilies of Blaberidae (Blattodea), with Panesthiinae primarily adapted to a wood-burrowing lifestyle and Geoscapheinae specialised for soil-burrowing. Panesthiinae are distributed across Asia and Australia, and Geoscapheinae are only found in Australia. Although the two subfamilies exhibit marked differences in morphology, behaviour, and reproduction, molecular phylogenetic studies have shown that Geoscapheinae is nested within Panesthiinae, and its soil-burrowing lifestyle evolved multiple times independently. Previous studies of Asian Panesthiinae, however, were limited by sparse taxon sampling and few molecular markers, leaving their relationships and the evolutionary history of key morphological traits unresolved. To address this gap, we sequenced mitochondrial genomes and nuclear ribosomal DNA from 18 Asian Panesthiinae species representing five of seven genera, and combined these with published Australian data to construct a phylogenetic framework of 104 species. Divergence time estimation calibrated with fossils and molecular clocks revealed *Salganea* and *Miopanesthia* as a well-supported basal sister group, while the Asian genus *Panesthia* were recovered as paraphyletic. Biogeographic analyses supported an Asian origin of the subfamily, with two dispersal events into Australia at approximately 29 and 19 million years ago. Ancestral state reconstructions indicated repeated, independent reductions of wings, simplification of male genitalia, and loss of the oothecal membrane across multiple lineages, traits likely associated with ecological adaptation and mating strategies. This study provides a phylogenetic basis for taxonomic revision of Panesthiinae and Geoscapheinae, while also revealing the complex history of their morphological evolution.

## 3.2 Introduction

Panesthiinae and Geoscapheinae are two closely related blaberid subfamilies (Roth, 1977, 1982; Rugg and Rose, 1984), both adapted to burrowing—Panesthiinae in wood and Geoscapheinae in soil. Panesthiinae includes 142 species in seven genera restricted to Asia and Australia, and, like the distantly related genus *Cryptocercus*, all members feed exclusively on the wood they tunnel through—a relatively rare trait among animal groups. Geoscapheinae consists of 26 species in four genera, found only in Australia, and all feed on

dried leaf litter, which they transport to deep, permanent burrows up to 90 cm underground. While other cockroaches may burrow shallowly into soil or sand for temporary refuge, Panesthiinae and Geoscapheinae construct long-lasting shelters.

Although Panesthiinae and Geoscapheinae differ morphologically, behaviourally, and reproductively, phylogenetic studies have revealed that geoscapheines are derived from within Panesthiinae (Maekawa et al., 2003). The Panesthiinae evolved in Asia, and the burrowing habits of Geoscapheinae evolved multiple times independently from wood-feeding ancestors that dispersed from Asia to Australia during a 'first wave' invasion ~25 million years ago, making both subfamilies nonmonophyletic (Lo et al., 2016; Beasley-Hall et al., 2021). Australia's increasing aridity likely played a key role in this transition, as drying climates reduced access to the moist, shaded habitats these cockroaches were originally adapted to. This environmental shift imposed strong selection for behaviours that preserved similar cool and humid microclimates, driving the repeated evolution of deep soil-burrowing (Lo et al., 2016; Beasley-Hall et al., 2018; Chapter 2 of this thesis). The molecular evidence that Geoscapheinae are nested within Panesthiinae is consistent with the primary taxonomic distinction between the subfamilies being a minor difference in the angle of the seventh tergite.

The earliest branching lineage of the Panesthiinae has not yet been determined with confidence, but the lineages comprising *Salganea* and *Miopanesthia* are two candidates. One reason for this uncertainty is that molecular phylogenetic studies of Asian Panesthiinae thus far have been based on relatively few molecular markers (partial regions of ribosomal 18S rRNA, plus mitochondrial 12S and COII), and relatively low numbers of taxa (15 species in total; Maekawa et al., 2003). In contrast, recent studies of Australian Panesthiinae and Geoscapheinae have generated full mitochondrial genomes, and in some cases nuclear ribosomal complements and *Blattabacterium* genomes, to infer phylogenetic relationships among almost all known species (Adams et al., 2024; Beasley-Hall et al., 2024; Beasley-Hall et al., 2021; Beasley-Hall et al., 2021). There is a need to obtain larger molecular datasets from Asian Panesthiinae, from increased numbers of taxa, to obtain a more accurate understanding of the evolution of this group.

Within Panesthiinae, species exhibit a wide range of wing development. Some have fully developed tegmina and wings (macropterous) that extend to or beyond the end of the abdomen. Other species display varying degrees of wing reduction: from wings that cover only part of the abdomen or reach just the first abdominal segment (brachypterous)—but remain broad and meet along the midline—to small, widely separated lateral pads (micropterous), and ultimately to completely wingless (apterous) forms with no visible tegmina or wings. In some genera, such as *Miopanesthia*, males typically have wings while females do not. Additionally, several species (e.g. *P. cribrata* and *P. australis*) are dimorphic in both sexes, with individuals showing either fully developed or variously reduced wings. A study of wing evolution across Asian and Australian Panesthiinae and Geoscapheinae has yet to be undertaken.

In insects, male genitalia have long been employed to infer phylogenetic relationships, owing to their rapid diversification driven by sexual selection and their composite structure consisting of multiple elements. The effectiveness of male genitalia as a source of phylogenetic signal has also received some support (Song and Bucheli, 2010). Within Blattodea and particularly Panesthiinae, male genitalia represent a classical set of morphological characters for taxonomy. In Panesthiinae, the genitalia are typically divided into four principal phallomeres, one or more of which may be reduced or absent. The most widely used phallomeres include: L1 (the first sclerite of the left phallomere), which usually bears two lobes separated by a cleft with sclerotized margins; L2vm (the second ventromedial sclerite of the left phallomere), which is consistently present and therefore often excluded from comparative analyses; L2d (the second dorsal sclerite of the left phallomere), a sclerotized derivative of L2vm but separated from it, which may be reduced or absent; and R2 (the second sclerite of the right phallomere), which is hook-shaped when fully developed, but lacks the hook portion when reduced. Here, we follow the terminology of McKittrick (1964) and the descriptions of Roth (1977).

Both Panesthiinae and Geoscapheinae are generally ovoviviparous, and their eggs are typically enclosed within an oothecal membrane. During oviposition, the oothecal membrane is extruded, rotated, and subsequently retracted into the body. The oothecal membrane is generally thought to facilitate the arrangement of eggs into a bilaterally symmetrical double

row, while also protecting them from desiccation and external threats during the period of extrusion (Rugg and Rose, 1984a). From an evolutionary perspective, members of Blaberidae are generally considered to exhibit a trend toward reduction of the ootheca and oothecal membrane. Remarkably, within Geoscapheinae, two species have been reported to have completely lost the oothecal membrane, depositing eggs directly into the brood sac (Rugg and Rose, 1984a). This variation in reproductive traits highlights the evolutionary plasticity of oothecal structures within cockroaches and provides an important context for understanding trait evolution in Panesthiinae.

We performed sequencing of mitochondrial genomes and nuclear ribosomal complements for a total of 14 Asian Panesthiinae from 5 of the 7 genera of the group. These data were combined with those previously obtained from both Asian and Australian Panesthiinae and Geoscapheinae species, and phylogenetic and divergence dating analyses were performed. We aimed to improve the understanding of the biogeography of the Panesthiinae, as well as the evolution of key characters in the group, including wings, male genitalia (L1, L2d and R2 phallomeres), and the oothecal membrane.

## 3.3 Methods

### 3.3.1 Taxon sampling

For this study, mitogenomes and 18S rRNA gene, Internal Transcribed Spacer 1 (ITS1) and 28S rRNA gene were sequenced and assembled from 11 Asian samples previously examined in Maekawa et al. (2003) (*Ancaudellia shawi*, 2 samples of *Miopanesthia deplanata*, *Panesthia angustipennis spadica*, *P. angustipennis yayeyamensis*, *P. saussurii*, *P. transversa*, *Salganea esakii*, *S. gressiti*, *S. raggei*, *S. taiwanensis ryukyuanus*), plus 3 samples from Thomas Bourguignon and his colleagues (*A. sp. nymph*, *A. serratisima serratisima*, *S. ternatensis hirsute*). The data from these species were combined with 30 samples described in Chapter 2 of this thesis, plus 98 samples obtained from Genbank, covering 9 (out of 11) genera and 104 species, and an additional 9 outgroups.

### 3.3.2 DNA sequencing and assembly

To reduce contamination from endosymbiotic *Blattabacterium cuenoti* bacteria in the abdominal fat body (Kinjo et al., 2015), DNA was extracted from leg muscle tissue by Maekawa's and Bourguignon's groups, and tissues from some older specimens were frozen with liquid nitrogen to facilitate pulverization. For our own Australian samples, DNA extraction and sequencing were outsourced to the Australian National Insect Collection, Canberra, utilizing an approach suitable for highly fragmented historical DNA (see Jin et al., 2020; Zwick and Zwick, 2023). DNA of samples provided by Maekawa were extracted using NucleoSpin Tissue XS (TaKaRa), and then outsourced to Beijing Genomics Institute (BGI), Hong Kong, utilising DNBSEQ low-input DNA library-KAPA for low-input historical DNA library. Samples provided by Thomas Bourguignon were processed for DNA extraction from dissected fat bodies using QIAGEN DNeasy Blood & Tissue kit, DNA libraries were prepared using NEBNext Ultra IWE FS with a modified protocol and dual unique barcodes, and sequenced using Novaseq 6000 platform at OIST specifying 2 × 151 bp paired-end reads.

DNA Raw reads were assembled *de novo* using SPAdes v.3.12.0 (Bankevich et al., 2012) with default settings and sampling k values of 33, 55, 77, 91, and 121, and contigs were imported into GENEIOUS Prime (v.11.0.18, http://geneious.com), mapped to reference sequences of closely relative species with default settings and medium sensitivity. Reference sequences were sourced from GenBank to represent the closest known sister taxa to each sample, as determined by the phylogenetic framework established by Beasley-Hall et al. (2021). Either a single near-complete mitochondrial genome contig, or multiple contigs that covered the majority of the mitochondrial genome, were then extracted. Nuclear ribosomal operons (which encode 18S and the internal transcribed spacer ITS1) were assembled following the same process.

### 3.3.3 Phylogenetic analyses

The data generated in this study and obtained from Genbank were combined to create per-gene alignments using MUSCLE algorithm (Edgar, 2004) with default settings, and overlapping gene sections were excluded. The final alignments of the whole mitochondrial genome and nuclear sequences comprised a total of 14776 bp and 1530 bp, respectively. Protein-coding alignments were visually inspected to ensure they adhered to the correct reading frame, using the invertebrate mitochondrial genetic code and the standard genetic code for the respective datasets. As substitutional saturation was not detected in a previous study (Beasley-Hall et al., 2021), the third codon position was retained in the alignment. Partitioned Maximum Likelihood (ML) analyses were performed using IQ-TREE v2.2.2.4 (Minh et al. 2020) with default settings and ModelFinder (Kalyaanamoorthy et al., 2017) for model selection. Branch support was assessed using 10,000 ultrafast bootstrap replicates (UFBoot; Hoang et al., 2018) and the SH-like approximate likelihood ratio test (SH-aLRT) with 1,000 iterations.

We estimated evolutionary timescales using BEAST2 (v. 2.7.4, Bouckaert et al., 2014). Because there are no fossils within Panesthiinae suitable for calibration, some studies infer the evolutionary timescale by specifying prior substitution rates (Adams et al., 2024). To validate this result, we inferred the evolutionary timescale using both prior substitution rates and fossil calibrations. For the substitution rates, as its conserved substitution rate across insect orders (Gaunt and Miles, 2002; Papadopoulou et al., 2010), we specified a substitution rate prior for CO1 ($1.69 \times 10^{-2} \pm 1.9 \times 10^{-3}$ substitutions/site/Myr) as the normal priors with uncertainty corresponding to the standard deviation, and the rates of remaining partitions were estimated during analysis. For the fossil calibrations, we used '*Gyna*' *obesa* Piton, using exponential distribution and soft maximum bounds, as the maximum age was not determined very accurately (Ho and Phillips, 2009). Described as the earliest-known Blaberidae (Evangelista et al., 2017), '*Gyna*' *obesa* was widely used as a stem Blaberidae calibration in molecular dating analyses (Beasley-Hall et al., 2021; Bourguignon et al., 2018; Evangelista et al., 2019), with a minimum age of 57.7 Mya hard maximum bound and 145 Mya soft maximum bound (Lin, 1980), following Beasley-Hall et al. (2021) and Evangelista et al. (2019). Analyses were performed under an optimized relaxed-clock model with a birth–death

tree prior and BEAST's built-in model test. We ran the MCMC for 100 million generations, confirmed stationarity by checking that all ESS values exceeded 200 in Tracer v1.7.2 Rambaut et al. 2018), and then used TreeAnnotator with a 10% burn-in to generate the maximum clade credibility tree (Bouckaert et al., 2014). The distribution map was development using QGIS.

### 3.3.4 Ancestral character state reconstructions (ASR)

The evolution of morphological traits, including the wing, male genitalia, and the oothecal membrane, was inferred using ancestral state reconstruction. The different states of wing morphology were classified as full wing (macropterous), reduced wing (brachypterous or micropterous) and wingless (apterous). Three traits in the male genitalia were selected as follows: L1 (well developed, reduced, absent); L2d (well developed, reduced, absent) and R2 (well developed, reduced, absent). Character states of extant species were mapped to the species tree, and the R package *phytools* v.2.4.4 (Revell, 2024) was used to perform the ancestral state estimations using the continuous-time Markov chain model. The asymmetric (ARD) model was implemented, as some studies argue that the possibility of wings being regained after loss is not as low as traditionally believed (Bank and Bradler, 2022). The most likely state of each node was used as the ancestral state. Some highly polymorphic species are represented by their assumed ancestral states, as our samples do not cover the entire distributions of several species, and a number of species in the Panesthiinae were not available for analysis.

## 3.4 Results and discussion

### 3.4.1 Phylogenetic relationships

We conducted the most comprehensive phylogenetic analysis of the cockroach subfamily Panesthiinae to date. Our study is the first to include mitogenome data for several Asian representatives of the group. Previously, data from only 3 molecular markers were available, leading to a lack of resolution of some relationships in this group. Our results support the

monophyly of several genera: *Salganea*, *Caeparia*, *Miopanesthia*, and *Ancaudellia* (Fig. 3.1). However, Asian representatives of *Panesthia*, which constitutes the most speciose genus in Panesthiinae, were found to be a paraphyletic group, mirroring previous studies demonstrating that Australian members of the genus are also a paraphyletic group with respect to Geoscapheinae (e.g. Beasley-Hall et al., 2021).

We found that *Miopanesthia* and *Salganea* together form the sister group to the remaining Panesthiinae. Although each of these genera has been found to be monophyletic in previous studies, their relative positions among the earliest branching Panesthiinae lineages have remained uncertain (Beasley-Hall et al., 2021; Lo et al., 2016; Maekawa et al., 2003). The most recent studies have placed *Salganea* as the earliest branching lineage in the group, followed by *Miopanesthia*. However, the support values for these groupings have not been consistent (Beasley-Hall et al., 2021; Lo et al., 2016). Our study provides the first evidence that these two genera form a well-supported monophyletic group at the base of the Panesthiinae tree. However, since we only examined representatives of 1 of the 8 recognised *Miopanesthia* species, and 8 of the 50 recognised *Salganea* species (including an unknown species), further work is required to further investigate the grouping of these two genera. Previous morphological comparisons of Panesthiinae did not group the two genera together, instead placing *Miopanesthia* as the sister group of *Panesthia*, and *Salganea* as the sister group of *Microdina* (Roth, 1982), although it is now clear that the characters used to propose these groupings (male genital phallomeres, wings) are unreliable for inferring relationships (Lo et al., 2016). Further studies include *Microdina* are also required to examine its position among the Panesthiinae.

The clade *Panesthia transversa* + *Caeparia* spp. was found as the sister group to all *Panesthia* and Geoscapheinae, consistent with previous findings (Maekawa et al., 2003). Our study is the first to include more than a single representative of the genus *Caeparia*, and thus is the first to support the monophyly of species within this genus. Our results are consistent with those of Lo et al. (2016) and Beasley-Hall et al. (2021), confirming the phylogenetic relationship between these two lineages.

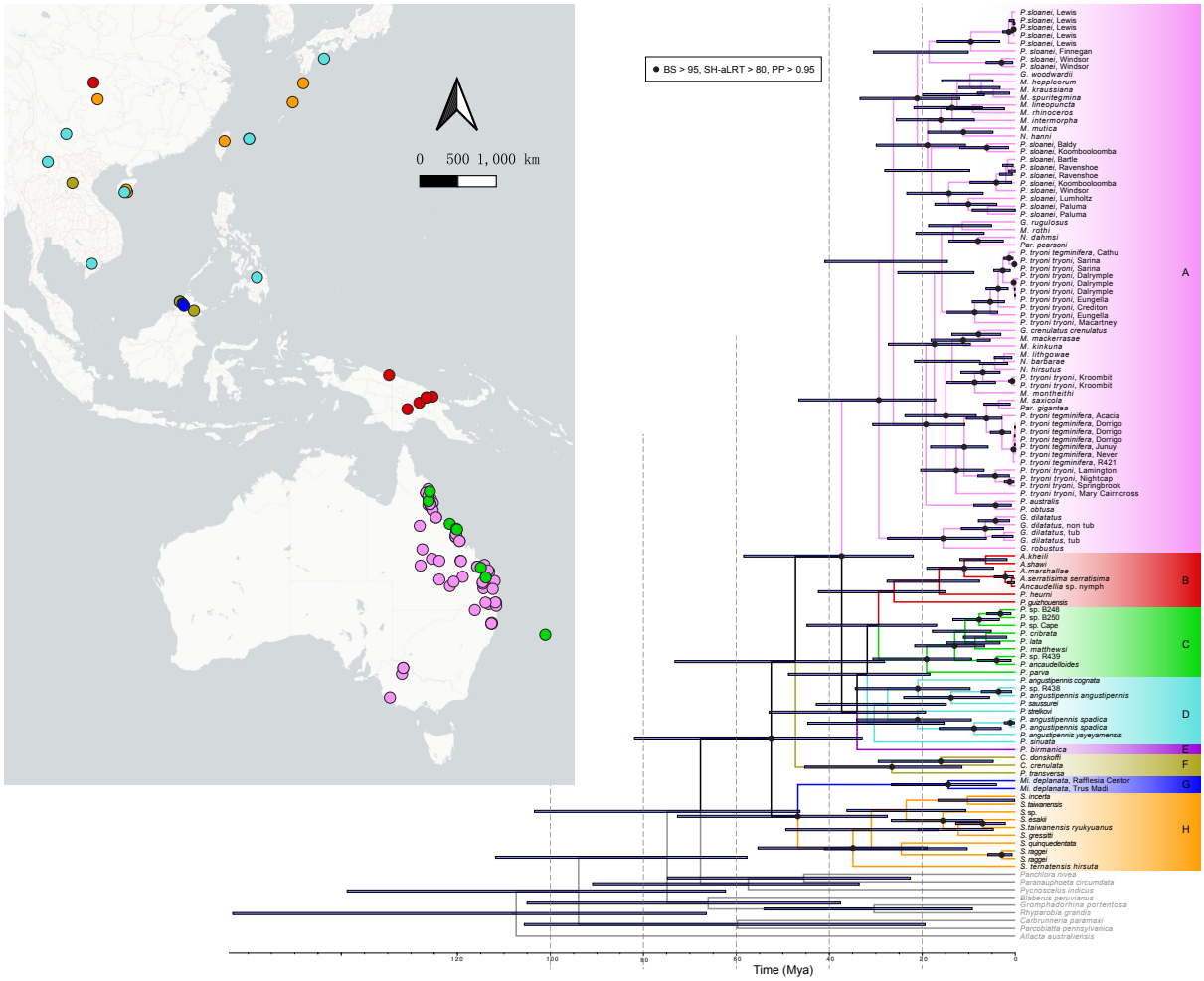



**Figure 3.1**. Dated phylogeny of Panesthiinae and Geoscapheinae inferred from complete mitochondrial genomes and nuclear ribosomal markers in BEAST and IQTREE. The evolutionary timescale was inferred using one fossil calibration, '*Gyna*' *obesa* Piton (Piton, 1940). The locations shown on the tree are not necessarily reflective of the total distribution of some species. Letters A-H on the right represent eight main clades of Panesthiinae and Geoscapheinae. BS, ultrafast bootstrap; PP, posterior probability; SH-aLRT, SH-like approximate likelihood-ratio test. Outgroups were shown in grey. Sampling locality colours correspond to groupings based on the phylogeny. Grey bars at each node represent 95% highest posterior density (HPD) divergence times.

The clade comprising Geoscapheinae soil burrowing taxa + *P. sloanei* + *P. tryoni* appears to be the first to have invaded Australia. Within this clade, we recovered *G. dilatatus* + *G. robustus* as the sister clade to all others, although with variable support (BS = 43.1, SH-aLRT = 84, PP = 1). The remaining taxa were split into 2 major groups, one comprising a paraphyletic *P. sloanei* with respect to a clade of soil-burrowing taxa including *M. rhinoceros*, and the other comprising an early-branching lineage of wood-feeding species (*P. australis* and *P. obtusa*), *P. tryoni*, and various soil-burrowing lineages. As the relationships within this clade were essentially identical to those found in a recent study (Chapter 2 of this thesis), the details are not repeated here.

In the clade comprising the remaining Australian *Panesthia* spp. (clade C in Fig. 3.1) + Asian *Panesthia* + *Ancaudellia* spp. (clades/lineages B, D, E, Fig. 3.1), the earliest branching lineage was *P. birmanica*. Roth (1979b) emphasised the behavioural similarities between *P. birmanica* and *Miopanesthia*, suggesting that the former might be a different lineage of *Panesthia.* While we did not find any close relationship between these two taxa, we did find that *P. birmanica* represents a divergent lineage. The remaining *Panesthia* within this group were split into three main groups: the 'second wave' Australian *Panesthia*; a clade comprising *P. angustipennis* and other Asian species; and finally, a clade comprising *P. guizhouensis, P. heurni,* and representatives of *Ancaudellia*. Representatives of the widespread species *P. angustipennis* were not recovered as monophyletic, with *P. saussurii* and *P. strelkovi* nested among them. *Ancaudellia* was found to form a strongly supported monophyletic group whose sister group was *P. heurni*.

### 3.4.2 Biogeography

Miopanesthia *and* Salganea

*Miopanesthia* + *Salganea* were estimated to have diverged from other Panesthiinae lineages ~53 Mya (95% CI 31.2 – 79.1), and further split into two distinct clades, *Miopanesthia* and *Salganea* ~47 Mya (95% CI 27.6 – 71.6). *Miopanesthia* and *Salganea* coexist in the same broad region, including South Asia, China, Indochina, and Southeast Asia. Notably, in areas

such as Myanmar, Thailand, Japan, and New Guinea, only *Salganea* species are present (Beccaloni, 2024). Due to the very limited sampling of *Miopanesthia* in this study (the two specimens represent only one species from two locations in Sarawak, Malaysia), it is currently not possible to reconstruct the biogeographic history of this genus. *Salganea* can be roughly divided into three major clades: a New Guinea clade (*S. ternatensis hirsuta*), a clade comprising *S. quinquedentata* and *S. raggei* (Hainan Island), and an East Asian clade (*S. incerta*, *S. taiwanensis*, *S. taiwanesis ryukyuanus*, *S. esakii* and *S. gressitti*). The New Guinea lineage was the first to diverge from the other examined *Salganea* lineages, around 35 Mya (95% CI 19.1 – 54.2). At this time, the Australian/Papua New Guinean and Southeast Asian plates were still geographically distant. Given the evidence that Panesthiinae evolved in Asia (Maekawa et al, 2003), and the fact that almost all *Salganea* species diversity is found in Asia, it appears likely that *S. ternatensis* dispersed to Papua New Guinea from Asia. When this occurred remains unclear, but it might be expected to have occurred following the collision of the Australian/Papua New Guinean plate with the Asian plate, as this would have brought the two faunas in closer proximity. Since *S. ternatensis* is also found in Sarawak, Borneo, future studies including representatives of this population, as well as others, may shed light on the origins of the Papua New Guinean *S. ternatensis*. The other two clades (the East Asian clade and *S. quinquedentata* + *S. raggei* clade) diverged around 31 Mya, yet they still exhibit significant overlap in their distribution ranges. The reasons for their divergence remain unclear.

Caeparia *and* Panesthia transversa

*Caeparia* and *P. transversa* diverged from other *Panesthia* + *Ancaudellia* approximately 47 Mya, and subsequently from each other around 27 Mya. Geographically, *Caeparia* is distributed across South Asia, Indochina, and Sumatra Island in Indonesia, whereas *P. transversa* occurs in China, Myanmar, Malaysia, and regions of Indonesia west of the Wallace Line. The current distribution of *Caeparia* exhibits notable gaps, which may reflect insufficient sampling. Similarly, the known range of *P. transversa* may also be incomplete. At present, the two lineages appear to occupy complementary geographic regions. However, more comprehensive sampling and phylogeographic studies are needed to better resolve their actual distributions and understand their evolutionary history.

*Asian* Panesthia

Among Asian *Panesthia* lineages, the earliest divergence occurred around 34 Mya, when *P. birmanica* split from the remaining three clades (Clades B–D). *Panesthia birmanica* is distributed in India, Myanmar, Thailand and Vietnam. The early divergence of this taxon suggests the origin of the genus may lie in the western side of its distribution. Subsequent divergence events followed at approximately 32 Mya (between Clade D and Clades B + C) and 29 Mya (between Clades B and C). Clade C represents the second wave of *Panesthia* that colonized Australia, while early branches of Clade B *(P. heurni* and *P. guizhouensis*) are distributed in southern China and Malaysia—regions geographically distant from *Ancaudellia*, which is restricted to New Guinea. It is generally assumed that the ancestor of Clades B + C was distributed near New Guinea, suggesting that some members of this lineage were among the earlier species to cross the present-day Wallace and Lydekker lines. Subsequently, some lineages may have dispersed westward into Southeast Asia, especially after inter-island distances were reduced due to geological changes.

Repeated crossings of biogeographic boundaries such as the Wallace and Lydekker lines are not uncommon, particularly following the narrowing of the sea between the Australian and Sunda plates. For example, in murine rodents (Muridae, Rodentia; Rowe et al., 2019), multiple eastward dispersals across the Wallace Line have been documented, along with two westward crossings of the Wallace Line and one westward crossing of Lydekker's Line. Similarly, flightless weevils (*Trigonopterus*), which have low dispersal capabilities, show evidence of multiple crossings of the Wallace Line in both directions (Tänzler et al., 2014).

*Australian groups*

Australian Panesthiinae and Geoscapheinae species have received considerable attention, and the phylogenetic relationships and evolutionary histories within these groups are relatively well resolved (Chapter 2 of this thesis; Adams et al., 2024; Beasley-Hall et al., 2021; Lo et al., 2016). Based on our results, the first invasion of this group into Australia led to the start of its diversification ~29 Mya. Diversification of the 'second wave' of *Panesthia* (Adams et al., 2024) in Australia began ~19 million years ago. The wood-feeding species arising from each of these invasions has a somewhat similar overall distribution along the

east coast of Australia, from Far North Queensland down to Victoria. However, the distribution of the soil-burrowing representatives from the first invasion includes arid inland scrubland areas, including habitats in South and Western Australia.

By the time the second-wave Australian *Panesthia* began diversifying in Australia, descendants of the first invasion had already diverged into three main lineages, possibly as a result of the reduction in rainforest habitat and the consequent formation of barriers that are believed to have commenced around that time (Byrne et al., 2008). These wood-feeding taxa evolved into soil-burrowing taxa on multiple occasions as a result of aridification from 15-20 Mya onwards, and as late as ~ 6 Mya in the case of the ancestors of *Macropanesthia saxicola* + *Parapanesthia pearsoni*. On the other hand, no second-wave *Panesthia* evolved the trait of excavating the semi-permanent and deep burrows (up to ~1 m) found in Geoscapheinae species. Nevertheless, some second-wave *Panesthia,* including *P. parva* evolved the ability to live in relatively dry habitats, while *P. matthewsi* and *P. lata* can burrow in soil under rocks. These burrows are only superficial (directly below rocks) and transient, compared with those formed by Geoscapheinae (Coady et al., 2025).

An examination of the relatives of soil-burrowing taxa reveals that their sister taxa are strictly those found in rainforests. Taxa from each invasion that are found in temperate forests, such as *P. cribrata* and *P. australis*, did not evolve soil-burrowing relatives in response to aridification. Assuming that large areas of both rainforests and temperate forests in Eastern Australia experienced conversion to arid scrubland during the last 20 million years, a potential explanation for the evolution of soil burrowing is that only transitions between rainforest and scrubland provided strong enough selective pressure to allow the evolution of this trait. An alternative explanation is that only the genomic background of lineages of the 'first wave' permitted the transition from wood-feeding to soil-burrowing.

We performed time calibrations of Panesthiinae in BEAST using both fossil and substitution rate. The resulting divergence times were similar and consistent with previous studies (Beasley-Hall et al., 2021; Lo et al., 2016). However, compared to a recent study focusing only on the second wave Australian *Panesthia*, all the divergence times obtained in our study are significantly older than their results (Adams et al., 2024). We hypothesise that this discrepancy may be due to recent rapid diversification events leading to limited

phylogenetic signal, which causes divergence time estimates to shift towards more recent times. Some studies on Bayesian molecular dating have highlighted the impact of taxon sampling on divergence time estimates (Linder et al., 2005; Soares and Schrago, 2015). For instance, some researchers have found that increasing sampling around or below a node tends to result in older estimated ages for that node (Poux et al., 2008). This suggests that when using substitution rates alone for divergence time estimation, careful consideration of taxon sampling is necessary. Including representative samples from higher-level taxa might be worth considering.

### 3.4.3 Morphological evolution in Panesthiinae and Geoscapheinae

*Wing reduction*

We examined the evolution of wings, male genitalia, and the oothecal membrane across Panesthiinae and Geoscapheinae for the first time in a formal ancestral state reconstruction analysis. Many species within these groups exhibit wing reduction to varying degrees (Fig. 3.2). The current consensus concerning wings in wood-dwelling insects is that fully developed wings hinder movement within rotten logs and that flight muscle is costly to develop and maintain (Cao and Jin, 2020; Korb, 2008; Zera and Mole, 1994). Nevertheless, a number of Panesthiinae species that inhabit dry rotten wood, *P. parva*, *P. obtusa* and *P. grayi*, have fully developed wings. This may be because it is easier to move and turn around in dry wood, or because these species need to migrate more frequently than those living in decayed wood, especially after the wood falls or gradually decomposes. Among the 'first wave' clade of Australian *Panesthia*, most species have completely lost their wings, with the exception of *P. australis* and *P. obtusa*. In the ASR analysis, these two taxa were inferred to have re-evolved wings. Although the reactivation of silenced developmental pathways to regain lost traits is thought to be possible (Collin and Miglietta, 2008; Whiting et al., 2003), it appears unlikely in this case. We believe a more reasonable explanation is that the ancestor of this entire clade was actually winged, and that wing loss occurred multiple times independently (Lo et al., 2016). *Panesthia tryoni tegminifera* retains highly reduced, lobe-like forewings, which in the ASR analysis appear as 're-evolved' wing parts. Again, we believe that these likely evolved from an ancestor with more developed wing parts, rather than re-evolving from a completely

apterous ancestor. Overall, this question cannot be resolved by our current results and may require a more detailed genomic analysis of wing developmental pathways to be answered.

Wing reduction is also very common in the second wave of Australian *Panesthia* (Fig. 3.2), though most species still retain lobe-like forewings (e.g. *P. matthewsi*, *P. lata*). In contrast, only scattered cases of wing reduction are observed in Asian Panesthiinae species (Fig. 3.2). Interestingly, *S. taiwanensis* has been recorded exhibiting a behaviour where mates chew off each other's wings before copulation (Osaki and Kasuya, 2021). If this behaviour plays a key role in establishing monogamy, it could act as a constraint against wing reduction. This may partly explain why *S. taiwanensis* retains fully developed wings, even within a lineage where reduced wings have otherwise become the norm.

*Male genitalic reductions*

In Panesthiinae, male genitalia exhibit extensive reductions. Among the first wave of Australian Panesthiinae/Geoscapheinae, the reduction of male genitalia generally aligns with phylogenetic relationships (Fig. 3.3). For example, all *P. tryoni* and their soil-burrowing relatives have lost their L2d phallomere; *G. dilatatus* and *G. robustus* exhibit a reduction in all the male genitalia phallomeres that were examined in this study. Most Asian *Panesthia* have fully developed male genitalia (fully developed L1, L2d and R2 phallomeres), but reductions were found in *Caeparia* (L1 and R2) and *Miopanesthia* (L2d). The clade exhibiting the highest male external genitalic variation is that comprising *Salganea*. Multiple independent reductions have occurred in this clade in both the L1 and R2 phallomeres. However, the mere presence or absence of such reduced characters alone is insufficient to justify taxonomic reclassification, for example, the variety of the three focused traits is not sufficient to distinguish the first wave of Australian Panesthiinae/Geoscapheinae and Asian Panesthiinae clades.

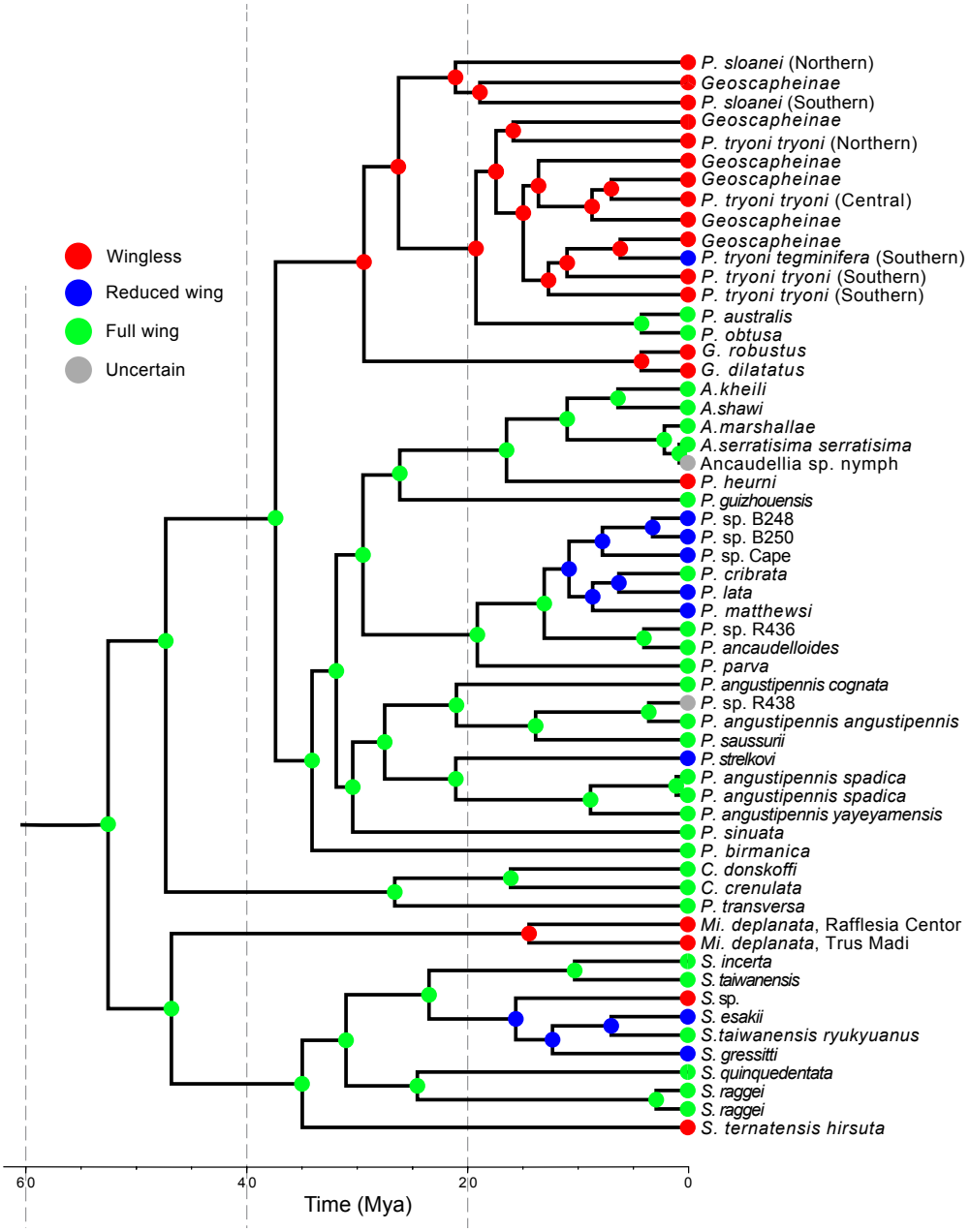



**Figure 3.2**. Evolution of wing morphology, estimated in phytools over the Bayesian chronogram inferred from whole mitogenomes and nuclear ribosomal markers in BEAST. The estimation was identical with wing re-evolution permitted or prohibited. *P. tryoni*, *P. sloanei* and Geoscapheinae samples were represented by operational taxonomic units (OTUs) in the analysis. The colour of circles at nodes indicates the most probable ancestral state (probability >50%). Red indicates wingless (apterous), blue reduced wing (brachypterous or micropterous), and green full wing (macropterous), while a grey circle indicates no data for this sample.

The ASR analysis indicates that the reduction of male genitalia may have occurred multiple times within first-wave Australian taxa, particularly in the L1 and R2 phallomeres. It is generally believed that simplification of male genitalia is associated with the reduction of female selection and repeated mating (Brennan and Prum, 2015; Cayetano et al., 2011; Kuntner et al., 2016; Rowe and Arnqvist, 2012; Simmons, 2014). However, significant simplification of male genitalia has not occurred in all soil-burrowing lineages, and the simplification of L2d is also observed in many wood-feeding species. Therefore, although shifts in family composition and possibly changes in mating frequency and sexual selection intensity have occurred during multiple transitions to a soil-burrowing lifestyle in the first-wave clade, whether these changes are the cause of genital simplification requires further in-depth and specialised research.

Simplification of the male genitalia is also known to occur in *Caeparia* and *Salganea*. Due to limited sampling, it is unclear whether simplification in *Caeparia* occurred in the ancestor or multiple times independently. In *Salganea*, reductions in L1 and R2 have clearly occurred multiple times. Although *Salganea* shares a wood-feeding lifestyle similar to Australian wood-feeding *Panesthia*, mutual wing-eating behaviour during mating is unique for *Salganea*, which may serve as a nuptial gift, and could incline *Salganea* towards monogamy (Osaki and Kasuya, 2021). This may explain the varying degrees of male genital simplification observed in this genus. However, broader descriptions of *Salganea* mating behaviour and more comprehensive statistical analyses are needed to confirm this hypothesis.

*Oothecal membrane reduction*

Oothecal membrane reduction has also occurred in a few soil-burrowing species and *Salganea*. The loss of an oothecal membrane in two lineages of soil-burrowing taxa (*G. dilatatus* and *G. robustus*, plus *M. rhinoceros*) was originally used as evidence for the creation of the Geoscapheinae subfamily (Rugg and Rose, 1984a); however, later studies showed that this reduction was not consistent among soil-burrowing species. Thus, the oothecal membrane is not reliable as a phylogenetic character.

In ovoviviparous cockroaches, the oothecal membrane is believed to assist in arranging the two rows of eggs, and to protect them before they are drawn into the internal brood sac (Rugg and Rose, 1984a). In *Panesthia*, due to a reduced number of eggs in the ootheca, there is space

for a reduction of the oothecal membrane. Earlier studies proposed that the loss of the oothecal membrane in *M. rhinoceros* and *G. dilatatus* represents the culminating stage in the trend toward oothecal membrane reduction within the family Blaberidae (Rugg and Rose, 1984a). There is also a species of *Salganea* (*S. raggei*) where the oothecal membrane has entirely disappeared, suggesting that the reduction of the oothecal membrane may not be directly tied to a soil-burrowing lifestyle.

Although the complete disappearance of the ootheca in most termites indicates that a stable nesting environment can compensate for the oothecal membrane's protective function (Nalepa and Lenz, 2000), the retention of an oothecal membrane in the majority of Panesthiinae and Geoscapheinae suggests that the nesting environment in these groups may not offer the same level of stability. The loss of the oothecal membrane in *S. raggei* suggests that there may be additional Asian lineages where the membrane has entirely disappeared. Although an earlier hypothesis assumed that the disappearance of the oothecal membrane was a prerequisite for the evolution of viviparity (Roth and Willis, 1955, 1957), the disorganised arrangement of eggs in *M. rhinoceros* and *G. dilatatus* also indicates a significant gap between membrane loss and viviparity (Rugg and Rose, 1984a). Thus, it remains premature to infer that *Salganea* and Asian Panesthiinae have evolved closer to viviparity than membrane loss alone would suggest.

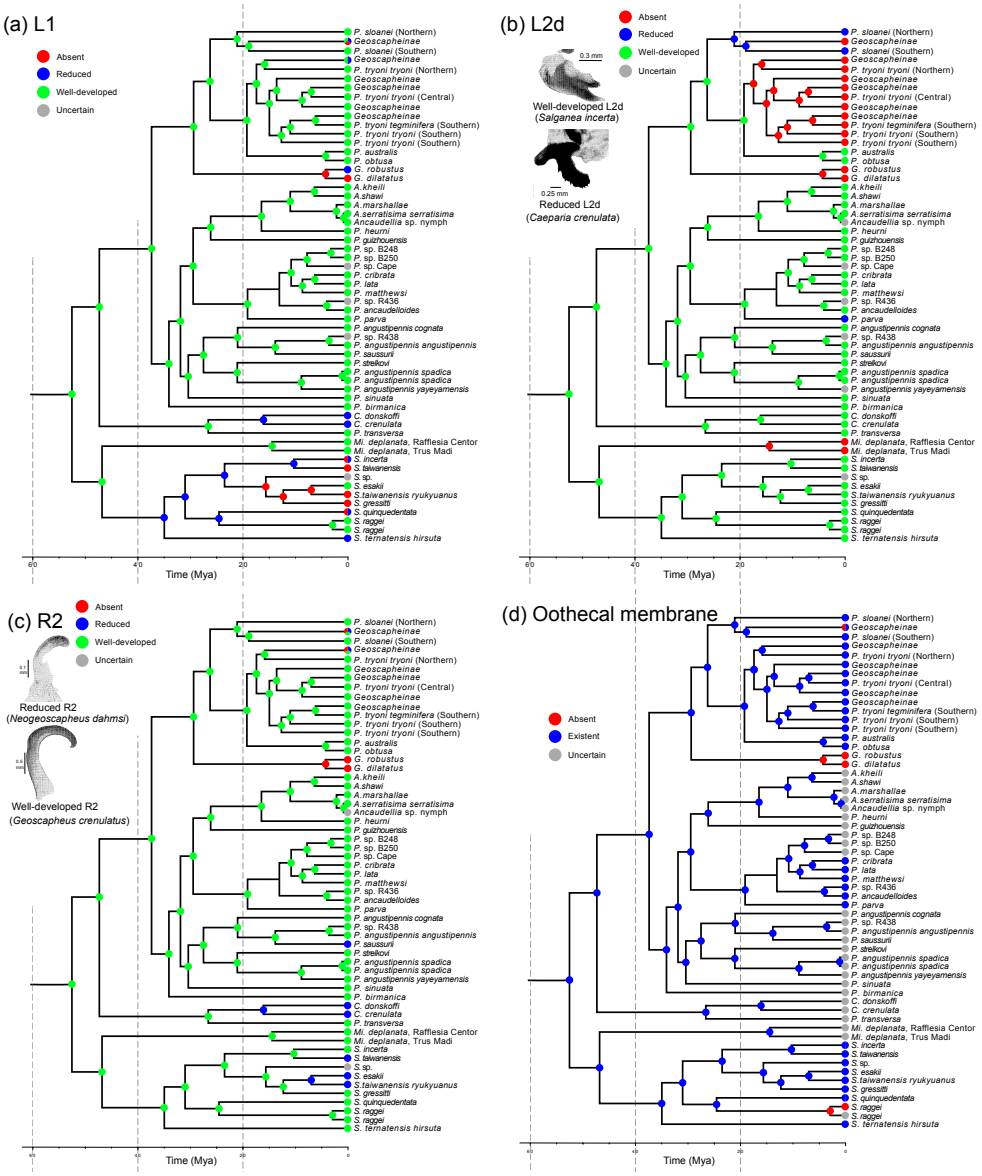



**Figure 3.3**. Evolution of male genitalia phallomeres and the oothecal membrane morphology, estimated in phytools over the Bayesian chronogram inferred from whole mitogenomes and nuclear ribosomal markers in BEAST. Traits re-evolution permitted or prohibited were identical in the estimations. *P. tryoni*, *P. sloanei* and Geoscapheinae samples were represented by operational taxonomic units (OTUs) in the analysis. The colour of circles at nodes indicates the most probable ancestral state (probability >50%). In male genitalia phallomeres (a, b and c), red indicates absent, blue for reduced phallomere, and green for well-developed phallomere, while a grey circle indicates no data for this sample; in oothecal membrane (d), red indicates absent and blue for existent oothecal membrane, grey for missing data. (a) estimation of L1 phallomere; (b) estimation of L2d phallomere; (c) estimation of R2 phallomere; (d) estimation of the oothecal membrane. Photographs adapted from Roth (1977, 1979)

### 3.4.4 Taxonomy

The taxonomy within Panesthiinae is a long-standing problem. The genus *Panesthia* was originally established by Illiger (1801) with *P. angustipennis* as the type species. Brunner (1865) later elevated *Panesthia* to the family level, and additional genera *Salganea*, *Miopanesthia*, *Caeparia*, *Geoscapheus*, and *Macropanesthia* were gradually established and included within Panesthiinae. Subsequently, Kirby (1903, 1904) redefined the family Blattidae and confirmed the subfamily status of Panesthiinae, a classification that has been maintained to the present. Roth (1977, 1979a, 1979b, 1982) incorporated Ancaudelliini, Salvagnini, and Caepariini into Panesthiinae based on the morphology of the seventh tergite, but proposed Geoscapheini as a distinct tribe. Later, Rugg and Rose (1984) elevated Geoscapheini to the subfamily Geoscapheinae based on morphology and oothecal structure. However, subsequent studies using mitochondrial markers (Lo et al., 2016) and both mitochondrial and nuclear markers (Beasley-Hall et al., 2021; Chapter 2 of this thesis) revealed that Geoscapheinae is nested within Australian *Panesthia*, challenging the subfamily status of Geoscapheinae. Taxonomic, phylogenetic, and biogeographic studies of other genera within Panesthiinae have largely focused on *Salganea* (e.g., Maekawa et al., 1999; Maekawa and Matsumoto, 2003); other genera have been poorly studied. Some studies described some new species and even new genera (e.g., Sergeev, 1984; Wang, Wang, and Che, 2014), but no phylogenetic studies were conducted. The work of Beasley-Hall et al. (2021) included some samples other than the first-wave Australian clade and *Salganea*, but only included them as outgroups.

The result of this study confirms the paraphyly of Panesthiinae. Besides the previously recognised issues of the Australian *Panesthia* and the polyphyly of Geoscapheinae, *Ancaudellia* and *Caeparia* each form monophyletic groups but are nested within different positions among lineages of *Panesthia*. Meanwhile, *Salganea* and *Miopanesthia* each form monophyletic groups and are sister groups to each other. Therefore, in revising the taxonomy of Panesthiinae and Geoscapheinae at this stage, it is clear that Geoscapheinae cannot be an independent subfamily. Additionally, none of the four Geoscapheinae genera correspond with phylogenetic relationships. Thus, the first wave of Australian *Panesthia* + Geoscapheinae might be best revised as a new genus (*Geoscapheus* in Fig. 3.4). Alternatively, they could be divided into multiple genera.

For the Asian taxa, our sampling is still incomplete, as *Annamoblatta* and *Microdina* were not included, and there are also several *Panesthia, Caeparia, Miopanesthia* and *Salganea* species that were not included. Nevertheless, based on our phylogenetic results, Asian *Panesthia* and the second-wave Australian *Panesthia* could be divided into four genera: the second-wave Australian *Panesthia* (also confirmed as monophyletic by Adams et al., 2024; Beasley-Hall et al., 2021, Gen. nov. 1 in Fig. 3.4); *P. angustipennis* + *P. saussurii* + *P. strelkovi* + *P. sinuata* (*Panesthia* in Fig. 3.4); *Ancaudellia* + *P. guizhouensis* + *P. heurni* (*Ancaudellia* in Fig. 3.4); and *P. birmanica* (Gen. nov. 2 in Fig. 3.4).

*Caeparia* formed a monophyletic group with *P. transversa* (*Caeparia* in Fig. 3.4); thus, these two lineages could together be considered a single genus. However, these two lineages have some morphological differences, especially the reduction and sclerotisation of male genital phallomeres L1 and L2 in *P. transversa*, which differ from those of *Caeparia* (Roth, 1979a, 1979b). Therefore, an alternative is that *P. transversa* is treated as an independent genus rather than merged into *Caeparia*. Since our sampling only included two *Caeparia* species, a more comprehensive sampling is necessary to confirm how the group should be redefined. A similar issue exists for *Miopanesthia*, although it does not disrupt the monophyly of *Salganea*, which has been well studied (Maekawa et al., 2001; Roth, 1979a). We would therefore propose retaining the two genera as they are, with further sampling required to validate this proposal (*Miopahesthia* and *Salganea* in Fig. 3.4).

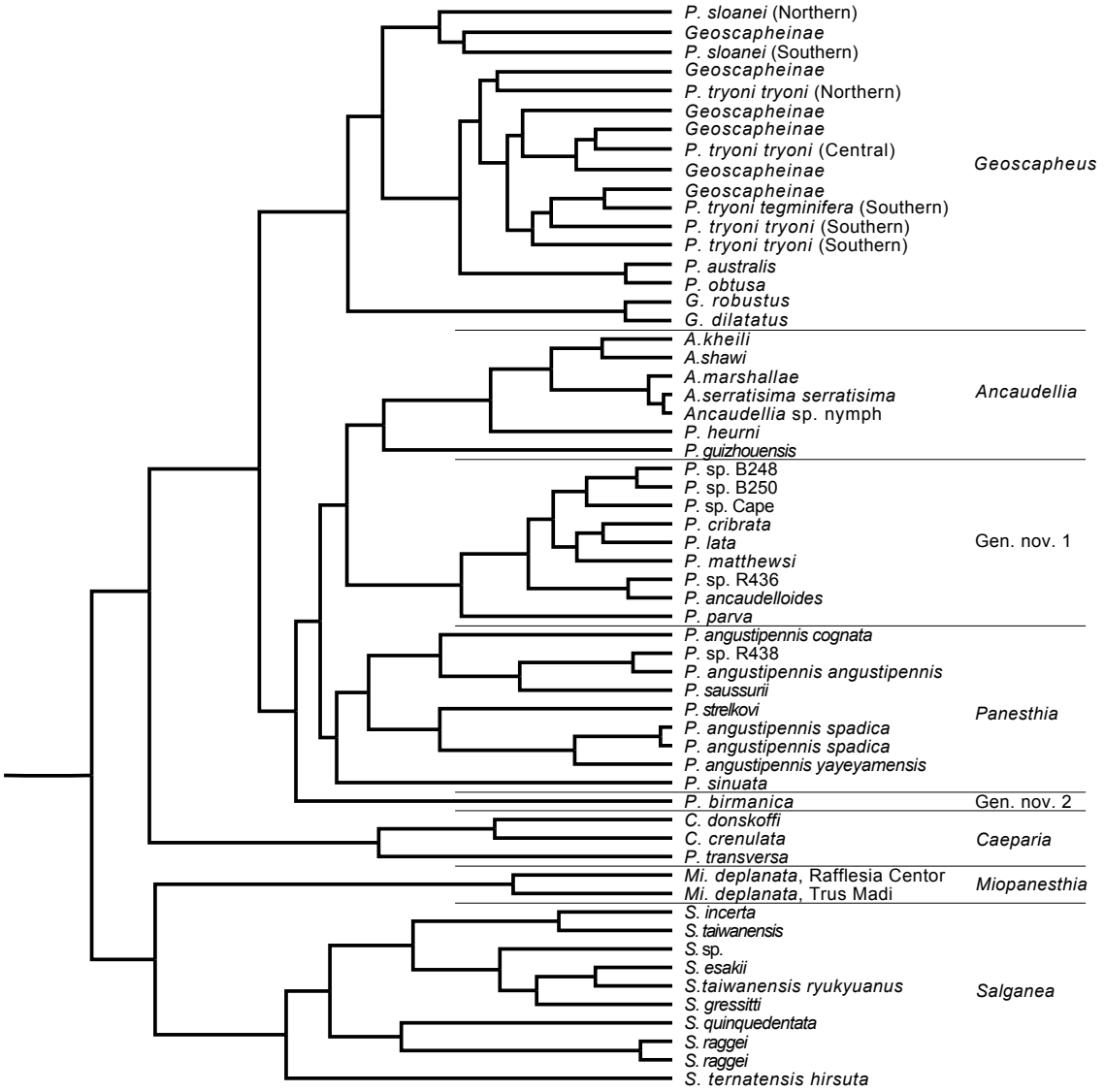



**Figure 3.4**. Taxonomic revision proposed based on our phylogenetic results. The phylogenetic tree is simplified from Fig. 3.1. Geoscapheinae represents species formerly assigned to this subfamily, which form a monophyletic clade in this position.

# Chapter 4

## Widespread Horizontal DNA Transfer from *Blattabacterium* Endosymbionts to Cockroach Genomes

**Zhuzhi Zhang**, Kyle M. Ewart, Maxim W. D. Adams, Harley Rose, Louise Baker, Aaron Jex, Nathan Lo

## 4.1 Abstract

Horizontal gene transfer (HGT) is a key mechanism driving evolutionary innovation, particularly in prokaryotes. While HGT from bacteria to eukaryotes is less common, recent studies suggest it may be underreported, especially in insects with intimate symbiotic relationships. This study investigates HGT from the obligate endosymbiont *Blattabacterium* to the genomes of 18 cockroach and termite species, leveraging long-read Oxford Nanopore sequencing for eight Panesthiinae and Geoscapheinae species. We developed a robust bioinformatic pipeline to detect both coding and non-coding HGT inserts, revealing an unprecedented number of transfers: up to 5,037 inserts in a single species, far exceeding previous estimates in insects. Most inserts were short (<100 bp), non-coding, and many were chimeric. Phylogenetic analysis dated these events to within the last 40 million years. Despite the high frequency of HGT, we found no correlation between HGT load and relaxed selection levels in wood-feeding versus soil-burrowing lineages. Our results highlight the prevalence and evolutionary dynamics of HGT in cockroaches, underscoring the importance of including non-coding regions in HGT studies and suggesting that symbiotic relationships facilitate continuous genetic exchange.

## 4.2 Introduction

Horizontal gene transfer (HGT) is a mode of genetic exchange distinct from the classical pattern of vertical inheritance from parents to offspring. It involves the transfer of genetic material across reproductive barriers between distantly related species, and its frequency varies across different domains of life. It is generally considered to be most common in prokaryotes, where DNA is transferred between bacteria through transformation, transduction, and conjugation (Arnold et al., 2022; Thomas and Nielsen, 2005). Once transferred, the foreign DNA can spread within a population through genetic drift or positive selection (Arnold et al., 2022) if it happens to be functional and confers an adaptive advantage to the new host. Numerous studies have demonstrated that HGT can endow bacteria with new adaptive traits. For example, it has been shown to introduce novel metabolic pathways (e.g., Goyal, 2022), enhance resistance to environmental stress (e.g., Kamal et al., 2021), facilitate detoxification (e.g., Bhat et al., 2024; Shoeb et al., 2012), and

increase pathogenicity (e.g., Deng et al., 2019). These cases highlight the evolutionary significance of HGT in shaping microbial function and diversity.

Although traditionally thought to be rare, instances of animals acquiring new genes through HGT are receiving increasing attention in recent years (Gilbert and Maumus, 2023; Yadav et al., 2024). For example, studies suggest that citrullinating enzymes known as peptidylarginine deiminases were acquired by vertebrates via HGT from cyanobacteria and are involved in various physiological and metabolic processes (Cummings et al., 2022). HGT appears to be more common in insects compared to vertebrates, potentially due to the widespread presence of symbionts, particularly endosymbionts, in the former. For instance, Kondo et al., (2002) found *Wolbachia pipientis* genes were transferred into the genomes of several host insects, including the fruit fly (*Drosophila ananassae*) and the bean beetle (*Callosobruchus chinensis*). Additionally, carotenoid biosynthesis genes of fungal origin have been identified in aphids (Insecta, Hemiptera) and gall midges (Insecta, Cecidomyiidae) (Cobbs et al., 2013; Moran and Jarvik, 2010).

To date, HGT studies have primarily focused on coding sequences, which only detect a small subset of fixed and expressed HGT events (Van Etten and Bhattacharya, 2020). However, sections of non-coding DNA, which constitute a significant portion of the genome of eukaryotic, may also been transferred through HGT. If insertions of foreign DNA into the genome are not removed by purifying selection, or alternatively confer greater adaptability to the species, they may persist within the genome (Husnik and McCutcheon, 2018).

Cockroaches (Insecta: Blattodea) harbour an obligate intracellular endosymbiont, *Blattabacterium*, within specialised cells of their abdominal fat body tissue. This symbiotic relationship has been maintained for over 150 million years, with phylogenetic trees of host and symbionts showing high levels of congruence (Arab et al., 2020; Lo et al., 2003). *Blattabacterium* is present in nearly all cockroach species, with the notable exception of most termites, which evolved from cockroach ancestors (Inward et al., 2007; Lo et al., 2000), and lost *Blattabacterium* early in their evolutionary history, likely due to shifts in diet and the acquisition of specialised gut microbiota. The only extant termite species retaining

*Blattabacterium* is *Mastotermes darwiniensis*, which is considered the most basal termite lineage (Kinjo et al., 2018; Sabree et al., 2009, 2012).

The genomes of *Blattabacterium* are reduced in size due to long-term mutualistic symbiosis (Sabree et al., 2009); they also contain an extremely low percentage of non-coding DNA (3.6%, corresponding to 23.4 kb in a 637 kb genome, López-Sánchez et al., 2009). They comprise genes that encode the enzymes responsible for the production of essential amino acids for their hosts, as well as vitamins including tetrahydrofolate and riboflavin (Sabree et al., 2009). The transovarial transmission of *Blattabacterium* in cockroaches facilitates close genomic interactions between the symbiont and host, which may increase the opportunities for HGT. In contrast, termites, having lost *Blattabacterium* over ~100 million years ago, are less likely to retain HGT-derived sequences unless these genes conferred a selective advantage and were maintained.

In this study, we performed detailed searches for HGTs from *Blattabacterium* to the genomes of 18 cockroach and termite genomes. Among these genomes, 8 were from the subfamilies Panesthiinae + Geoscapheinae that we sequenced for the first time, using long-read ONT technology. Following the detection of large numbers of HGTs in representatives of Panesthiinae and Geoscapheinae species, we tested the hypothesis that levels of relaxed selection were correlated with the number of HGT inserts in these Panesthiinae + Geoscapheinae species.

## 4.3 Methods

### 4.3.1 DNA extraction, genome sequencing, assembly, and annotation

We sequenced the genomes of ten Panesthiinae + Geoscapheinae species: *Macropanesthia lithgowae*, *M.kraussiana*, *Neogeoscapheus dahmsi*, *Parapanesthia gigantea*, *Panesthia tryoni tryoni* (Kroombit), *P. tryoni tegminifera*, *P. tryoni tryoni* (Eungella), *P. sloanei*, *P. lata*, and *P.*

*australis*. HMW DNA were extracted using the Phenol/Chloroform DNA isolation method (Chomczynski and Sacchi, 1987; Gautam, 2022).

Cockroaches were chilled at –30 °C for 4–5 min to induce torpor, and the femoral muscle (~3 mm) was dissected from one or two hind legs, depending on body size. Tissue was incubated overnight at 56 °C in 600 µl SDS buffer with 20 µl Proteinase K, followed by RNase A treatment (5 µl, 37 °C, 1 h). DNA was purified by phenol–chloroform extraction and isopropanol precipitation with sodium acetate (3 M, pH 5.5), washed twice with 70% ethanol, air-dried, and dissolved in 40 µl 10 mM Tris (pH 8.0). DNA quality and concentration were assessed with NanoDrop and Qubit, and samples were stored at –80 °C. The DNA of *M. lithgowae*, *M. kraussiana*, *P. tryoni tryoni* (Kroombit), *P. sloanei*, *P. lata* and *P. australis* was extracted using leg muscle only, while the DNA of *Par. gigantea*, *P. tryoni tegminifera*, *P. tryoni tryoni* (Eungella) and *N. dahmsi* were extracted using leg muscle + fat body tissue.

Following DNA extraction, samples were sent to Prof. Aaron Jex's laboratory at the Walter and Eliza Hall Institute of Medical Research (WEHI) for Oxford Nanopore Technology (long-read) sequencing, performed primarily by Dr Louise Baker, as follows. The quality and integrity of the gDNA were assessed by a 4200 TapeStation System (Agilent Technologies) using Genomic DNA ScreenTape (Cat No. 5067-5365) and reagents (Cat No. 5067-5366). The absorbance ratios were assessed by the Nanodrop Spectrophotometer (DeNovix DS-11). Quantification was determined using a Qubit 2.0 Fluorometer (Invitrogen by Life Technologies) using a dsDNA BR assay kit (Cat No. Q32850). Libraries were prepared from gDNA using the Ligation Sequencing Kit V14 (SQK-LSK114) protocol according to the manufacturer's instructions with the following modifications and options. The DNA Repair and End-Prep reaction continued for 15-20 minutes. The Adapter ligation reaction proceeded for 20 minutes, and the kit LFB wash buffer was used to enrich for longer fragments. Final library elution occurred at 37 °C to improve the recovery of long fragments. The integrity and concentration of final libraries were assessed using a 4200 TapeStation System (Agilent Technologies) and Qubit 2.0 Fluorometer (Invitrogen by Life Technologies) as previously described.

Libraries were sequenced on a PromethION 24 sequencer (Oxford Nanopore Technologies) using R10.4.1 flow cells (FLO-PRO114M) using the manufacturer's instructions. Libraries were ideally stored at 4ºC for less than a week if not sequenced immediately. Sequencing output was maximised by waiting for 30-60 minutes after loading the flow cells before starting the sequencing run. The output was additionally increased by routinely recovering the libraries after approximately 24 and 48 hours, washing the flow cells (EXP-WSH004) and reloading the libraries for sequencing for a total of 72 hours. Sequencing was performed in super-accurate base-calling mode and parameters set to also capture methylation 5mC + 5hmC modified bases.

For ONT sequencing runs that implemented duplex base calling, we used samtools v1.2 (Danecek et al., 2021) to remove redundant simplex reads. Next, adapters were trimmed from all ONT data, and reads were converted to FASTQ format using Dorado v0.6 (https://github.com/nanoporetech/dorado). A *de novo* assembly was carried out using the trimmed ONT data using Flye v 2.9.3. For *de novo* assembly, we used Flye v2.9.3 (Kolmogorov et al., 2019) on the trimmed ONT reads. After initial optimization experiments, ONT datasets with duplex base calling were assembled in "ONT corrected" mode (--nano-corr) with a minimum overlap of 10 kb (--min-overlap 10000), whereas datasets without duplex calling were assembled in "ONT raw" mode (--nano-raw) using default settings. Following assembly, we performed two additional polishing iterations. To remove putative haplotypic duplications and heterozygous overlaps, the purge_dups v1.2.5 pipeline (Guan et al., 2020) was applied. For this step, raw ONT reads were mapped back to the draft assembly using minimap2 v2.18 (Li, 2018). The complete genome assembly workflow, along with relevant scripts, is available on GitHub (https://github.com/MEEP-projects/HGT_pipeline).

The genome assemblies were annotated using FGENESH++ v7.2.2 (Salamov and Solovyev, 2000; Solovyev et al., 2006) on the Nimbus cloud platform, provided by the Pawsey Supercomputing Centre. Before annotation, we constructed a repeat library with RepeatModeler v2.0.1 (Flynn et al., 2020) to identify transposable elements. Using this custom library, repetitive regions in the assemblies were masked with RepeatMasker v4.0.6 (Smit et al., 2015), applying the -nolow option to preserve low-complexity and simple repeats. Both the masked and unmasked genome assemblies served as inputs for FGENESH++. The software was executed with the non-mammalian general pipeline

configuration and gene-finding parameters specifically optimized for *Acyrthosiphon pisum* (pea aphid). Additionally, homology-based predictions were enabled using Softberry's curated insect protein database (2020 release) via the prot_map. To ensure annotation quality, any ab initio gene predictions lacking BLAST support were discarded.

### 4.3.2 Characterising HGT inserts

The newly assembled genomes (with the exception of *P. australis* and *M. kraussiana* as the low quality of them) were combined with additional cockroach and termite genomes sourced from InsectBase 2.0 (Mei et al., 2022), NCBI and the dataset of Ewart et al. (2024) for downstream analyses. To evaluate the impact of data source and assembly approach on analytical outcomes, we included multiple genome assemblies for *Periplaneta americana*, *Blattella germanica*, and *Zootermopsis nevadensis*.

The identification steps of *Blattabacterium* insertions within cockroach and termite genomes were carried out with a multi-step analytical pipeline, a complete workflow detailing these steps is presented in Fig. S4.1. We first constructed a reference library comprising 77 *Blattabacterium cuenoti* genomes. Each genome was fragmented into 150 bp segments (alternative threshold: 100 bp) using a custom AWK script. These *Blattabacterium* fragments were aligned to the host genome using bwa mem v0.7.17 with default parameters (Li and Durbin, 2009). The resulting alignments were converted to BAM format using SAMtools, and subsequently to BED format using BEDTools v2.3 (Quinlan and Hall, 2010). Overlapping aligned fragments, as well as those within 150 bp of each other, were merged to define single putative HGT insertions. Putative inserts shorter than 50 bp were excluded using AWK.

Several filtering steps were applied to eliminate the potentially artefactual HGT insertions. First, to account for possible mis-assemblies in which authentic *Blattabacterium* sequences were mistakenly merged during assembly with cockroach sequences and mischaracterised as HGT inserts, we conservatively excluded any contigs in which putative bacterial sequence content—identified using the method described above—comprised ≥ 20% of the total contig length. In addition, we removed any putative HGT inserts exceeding 10,000 bp in length, as

these represented significant outliers in the insert length distribution. Second, we filtered out insertions that likely arose from false assembly duplications. For this filtering step, all putative HGT insertions were extracted along with 300 bp of upstream and downstream flanking sequence using BEDTools, and subjected to pairwise comparison within each genome using BLASTn v2.2.3. Inserts were considered duplicates and removed if their pairwise sequence identity exceeded 90% and their query coverage was also >90%. Third, we removed highly repetitive, low-complexity sequences. These low-entropy regions were identified using the symmetrical DUST algorithm (Morgulis et al., 2006) as implemented in Minimap2 v2.28 (Li, 2018, 2021). Any insertion with contiguous low-entropy segments comprising ≥ 50% of its total length was excluded from downstream analyses.

To determine the number of unique HGT insertions within each genome, i.e. excluding those likely resulting from duplication or transposition events, we employed a BLAST-based clustering approach. Each putative HGT insert, along with 300 base pairs of upstream and downstream flanking genomic sequence, was extracted and subjected to pairwise comparison against all other such sequences within the same genome using BLASTn v2.2.3. Insert pairs exhibiting a sequence identity greater than 70% and query coverage exceeding 70% were considered duplicates. In such cases, only one representative insert from each duplicate group was retained for further analysis. This approach ensures that the dataset reflects unique HGT events, minimising the influence of post-integration duplications or assembly artifacts.

After completing all filtering steps, we conducted an additional verification to ensure that identified putative HGT insertions originated from host cockroach genomes, rather than from *Blattabacterium* contamination during DNA extraction or assembly. This involved aligning raw (long-read) ONT sequencing reads from five species (*M. lithgowae*, *P. lata*, *N. dahmsi*, *P. tryoni tryoni* (Eungella), *P. tryoni tegminifera*) to their assembled genome using Minimap2 (Li, 2018, 2021), with default parameters and the "-a" command for the "sam" format result. For each putative HGT insertion, we extracted all reads that could align to the insert region with 500 base pairs of both upstream and downstream flanking sequences using samtools (Danecek et al., 2021). Subsequently, we analysed the Concise Idiosyncratic Gapped Alignment Report (CIGAR) of the extracted reads to calculate the combined length of matches and deletions, which was taken as the length of the genomic region to which the read

can align. This length was then compared with the length of the insertion. If the aligned region length was less than or equal to the insertion length, it indicated that the read could only be aligned to the insertion itself, but not to the flanking regions that are unequivocally part of the host cockroach genome. In such cases, the read was considered to originate from *Blattabacterium*. The analyses of the verification process are presented in Fig. 4.1. For each insertion, the number of reads classified as originating from *Blattabacterium* was counted and compared to the total number of reads corresponding to that insertion. If an insertion was composed entirely of reads from *Blattabacterium*, it was considered to represent either contamination from *Blattabacterium* or an assembly artifact.

### 4.3.3 Determining the ancestry of HGTs

We were able to estimate the minimum age of ancestral inserts across the phylogeny by comparing the sequence similarity of inserts and their flanking regions in other species to the sequence similarity of the *Blattabacterium* they harbour. Putative HGT inserts identified by the same *Blattabacterium* genome fragment in different host cockroach genomes were picked out and formed groups for further analysis. Each group represents a set of "nuclear-*Blattabacterium*" genes that may have originated from a single HGT event and have been vertically inherited into different cockroach species. These insertions with 5000 bp each of upstream and downstream flanking genomic sequence were extracted using Minimap2 (Li, 2018, 2021). After confirming that the sequences are oriented in the same direction through alignment of the insert sequences, the sequences from different species within each group were aligned in pairs to assess the homology of their flanking regions in Geneious Prime (v. 21.0.4, http://geneious.com) using Geneious aligning with default settings. If a clear homologous region could be identified in the alignment of two species (typically defined as a continuous region longer than 1,000 bp with a pairwise identity greater than 80%), the HGT inserts in these species were inferred to have originated from the same HGT event, i.e. the latest possible insertion time of the HGT occurred prior to the divergence of these two species. The divergence time between the most distantly related species pair within each group was taken as the latest possible time at which the HGT insertion could have occurred. This analysis was only conducted in all putative HGT inserts found in *D. punctata* (12 inserts) and 30 putative HGT inserts found in *P. t. tegminifera*. Two examples from *P. t. tegminifera* are shown in Fig. 4.3.

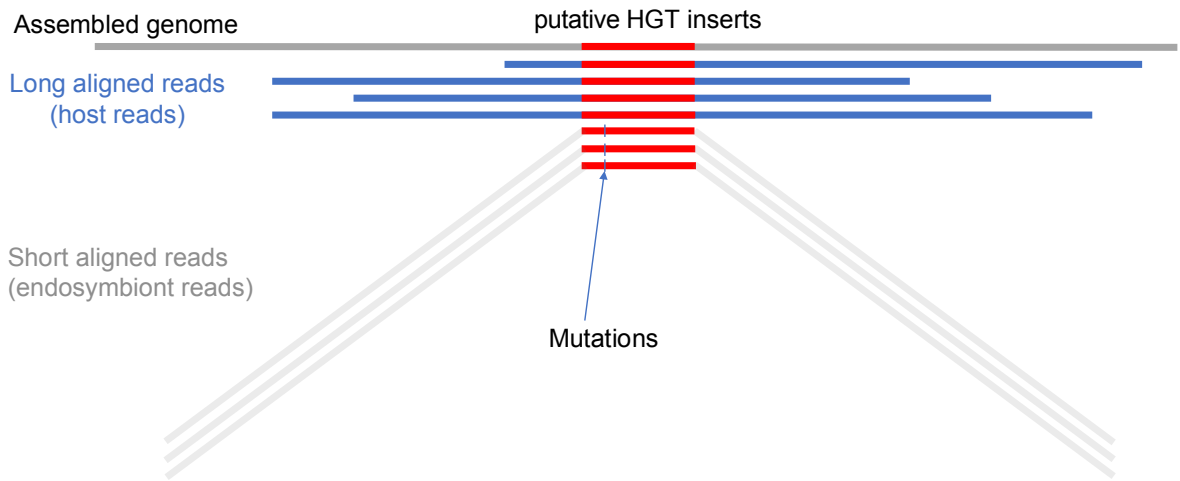



**Figure 4.1**. Schematic illustration showing methods for HGT insert detection and confirmation. The dark grey bar at the top represents the assembled genomic sequence; blue bars indicate reads broadly aligned to the genome and considered host-derived; light grey bars indicate reads aligned only at the insertion region and considered endosymbiont-derived; red marks represent the putative HGT insert sequence consistently aligned across all reads; "Mutations" denote base differences detected in short aligned reads compared with long aligned reads. Note that the "short" in "short aligned reads" refers to the length of aligned sequence between endosymbiont and host genome, rather than the reads themselves, which were of a similar length to other long reads derived from host DNA.

### 4.3.4 Orthology assignment for tests of relaxed and positive selection in Geoscapheinae

To test the hypothesis that HGT insert number is positively correlated with the level of relaxed selection observed in the genomes of Panesthiinae and Geoscapheinae taxa, we undertook selection analyses on single copy orthologues from 13 species. Proteomes of the 10 newly generated genomes were combined with those from three other Panesthiinae + Geoscapheinae genomes (*G. dilatatus*, *N. hanni* and *P. cribrata*) generated by Ewart et al. (2024). Orthologous protein sequences across all species were identified using OrthoFinder v2.5.5 (Emms and Kelly, 2015, 2019) with default parameters. Orthogroups were annotated based on the most frequently occurring gene annotation within each group. Single-copy orthologues were then extracted and aligned using OrthoFinder, employing the MAFFT algorithm for protein sequence alignment (Katoh and Standley, 2013). The corresponding coding sequences (CDS) for these single-copy orthologous proteins were retrieved from each genome and used in downstream selection analyses.

### 4.3.5 Positive and relaxed selection

To test for genes with evidence of relaxed selection, we applied the RELAX method (Wertheim et al., 2015), implemented in HyPhy (Kosakovsky Pond et al., 2020; Pond et al., 2005), to each alignment separately. RELAX fits three $d_N/d_S$ classes to the phylogeny, then tests for relaxed/intensified selection on a user-specified test branch. We performed analyses using RELAX for 13 different test branches in separate analyses. We also ran RELAX for all wood-feeding branches together (i.e. multiple test branches were considered in a single analysis) and for all soil-burrowing branches together, to assess whether there was a significant difference in the number of genes under relaxed selection between them. To assess significance, we performed separate chi-squared tests in R for relaxed selection and positive selection genes.

To elucidate potential functional links between genes under relaxed selection and the wood-feeding to soil-burrowing lifestyle transition, we annotated protein sequences of candidate genes (exclusively relaxed-selected genes in either soil-burrowing or wood-feeding lineages)

using InterProScan v5.57-90.0 (Blum et al., 2025) with default parameters. The highest-scoring InterPro entry (E-value threshold < 1e-5) for each gene was selected as the primary functional descriptor. These annotations were systematically mapped to several physiological categories implicated in substrate adaptation, for example: (1) locomotion (e.g., muscle contraction complexes, burrowing-associated cytoskeletal remodelling), (2) detoxification (e.g., cytochrome P450 systems, lignin degradation enzymes), (3) immune response (e.g., antimicrobial peptide synthesis, Toll pathway regulators), and (4) water conservation (e.g., cellulose digestion, nutrient absorption machinery). To functionally annotate the target genes, we utilised the eggNOG-mapper v2 tool, which leverages precomputed homologous gene clusters and phylogenetic trees to assign functional annotations based on single-copy Orthogroups (Cantalapiedra et al., 2021; Huerta-Cepas et al., 2019). Subsequently, Gene Ontology (GO) enrichment analysis was conducted using the clusterProfiler R package (Yu et al., 2012) through an over-representation test (ORA) to identify significantly enriched GO terms among genes under relaxed selection across multiple species.

## 4.4 Results

We *de novo* assembled and annotated high-quality genomes for eight Panesthiinae and Geoscapheinae species (genome size = 1.716 – 1.899 Gb, number of contigs = 2171 – 13000, N50 = 0.635 – 5.223 Mb, BUSCO completeness = 95.2 – 99.6%, Table 4.1). Using our novel approach, we characterised 40722 HGT inserts derived from *Blattabacterium* across the genomes of 14 cockroach species, ranging in size from 50–5037 bp, with a median size of 198 bp (Fig. 4.2). No *Blattabacterium*-derived inserts were found in termite genomes. Cockroaches were clearly divided into two groups based on the number of insertions, with Panesthiinae + Geoscapheinae genomes containing significantly more insertions than the other three cockroach species, reaching thousands (1543 - 5037) compared to 106 to 647 (depending on different versions of the genome for *B. germanica* and *P. americana* were used).

We performed a quality check of each of the thousands of insertions in five species (*M. lithgowae*, *P. lata, N. dahmsi, P. t. tryoni* Eungella, *P. t. tegminifera*), each of whose genomes had been generated using long read data. In all cases, we found ONT long reads that contained HGT insertions that were surrounded by host cockroach DNA. In those species for which no

fat-body tissue was included during extraction, all matching reads were identified as host cockroach DNA. On the other hand, for those species for which fat body had been included during RNA extraction (*N. dahmsi*, *P. t. tryoni* Eungella, and *P. t. tegminifera*), both host-derived and *Blattabacterium*-derived reads were found to align with the insertions, often with the number of reads from the latter being dominant. Nevertheless, the presence of raw ONT reads with both HGT insertion sequence plus host cockroach DNA upstream and downstream from the insert confirmed that they were not derived from symbiont DNA. For example, three species (*N. dahmsi*, *P. t. tryoni* Kroombit, and *P. t. tegminifera*) have thousands of insertions (1393–2057) where the number of *Blattabacterium* reads aligned exceeded 95% of the total. Nevertheless, a number of reads were always found that comprised the insert sequence and the host cockroach flanking sequence.

We investigated the ancestry of a subset of 30 random HGT inserts from *P. t. tegminifera* through comparisons of the HGT insert, plus 5000 bp upstream and downstream. Among these, 25 were found to be present in at least one other cockroach species, based on the presence of homologous sequence either upstream, downstream, or both. Two examples of these are shown in Fig. 4.3. Based on the presence of homologous sequence in flanking regions of these inserts, we were able to infer their minimum age through mapping onto a phylogenetic tree and using previously obtained estimates for node ages estimated in Chapter 3 of this thesis, Adams et al. (2024) and Evangelista et al. (2019). Based on the earliest divergence time of their common ancestor, the earliest could be traced back to 37 Mya (the common ancestor of the 11 examined species or the common ancestor of Panesthiinae and Geoscapheinae), with an average age of 20.67 Mya across the 30 examined inserts. We also analysed 12 insertions in *Diploptera punctata*. The results showed that the flanking regions of all insertions in *D. punctata* and similar insertions in other cockroaches did not exhibit homology, indicating that these 12 HGT insertions did not occur in the common ancestor of *D. punctata* and Panesthiinae.

**Table 4.1.** Genome assembly statistics and quality metrics for ten Panesthiinae and Geoscapheinae species, including the tissue sources used for DNA extraction.

| Species | ID | DNA extract tissue | Genome Size (Gb) | No. of contigs | N50 (Mb) | GC(%) | BUSCO complete (%) | Fragmented BUSCOs (%) | Missing BUSCOs (%) | Duplicated BUSCOs (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| *Panesthia lata* | M01 | Leg muscle | 1.765 | 11181 | 1.142 | 35.83 | 97.7 | 1.6 | 0.7 | 0.6 |
| *Panesthia sloanei* | Z01 | Leg muscle | 1.822 | 14705 | 1.667 | 35.85 | 98.6 | 0.6 | 0.8 | 1.5 |
| *Panesthia australis* | Z02 | Leg muscle | 1.795 | 11533 | 1.344 | 35.90 | 97.9 | 1.2 | 0.9 | 1.2 |
| *Macropanesthia kraussiana* | Z03 | Leg muscle | 1.795 | 9868 | 1.927 | 35.88 | 99.3 | 0.3 | 0.4 | 7.8 |
| *Macropanesthia lithgowae* | Z04 | Leg muscle | 1.824 | 6815 | 1.438 | 35.88 | 98.2 | 1.0 | 0.8 | 2.0 |
| *Panesthia tryoni tryoni* (Kroombit) | Z05 | Leg muscle | 1.716 | 13000 | 0.635 | 35.92 | 95.2 | 1.9 | 2.9 | 1.3 |
| *Parapanesthia gigantea* | Z10 | Leg muscle + fat body | 1.845 | 4559 | 1.043 | 35.67 | 98.4 | 1.6 | 0.0 | 0.4 |
| *Panesthia tryoni tegminifera* | Z11 | Leg muscle + fat body | 1.843 | 4277 | 1.188 | 35.76 | 98.8 | 0.4 | 0.8 | 0.8 |
| *Panesthia tryoni tryoni* (Eungella) | Z12 | Leg muscle + fat body | 1.862 | 5156 | 0.862 | 35.77 | 96.9 | 2.4 | 0.8 | 1.2 |
| *Neogeoscapheus dahmsi* | Z14 | Leg muscle + fat body | 1.899 | 2171 | 5.223 | 35.47 | 99.6 | 0.0 | 0.4 | 0.8 |

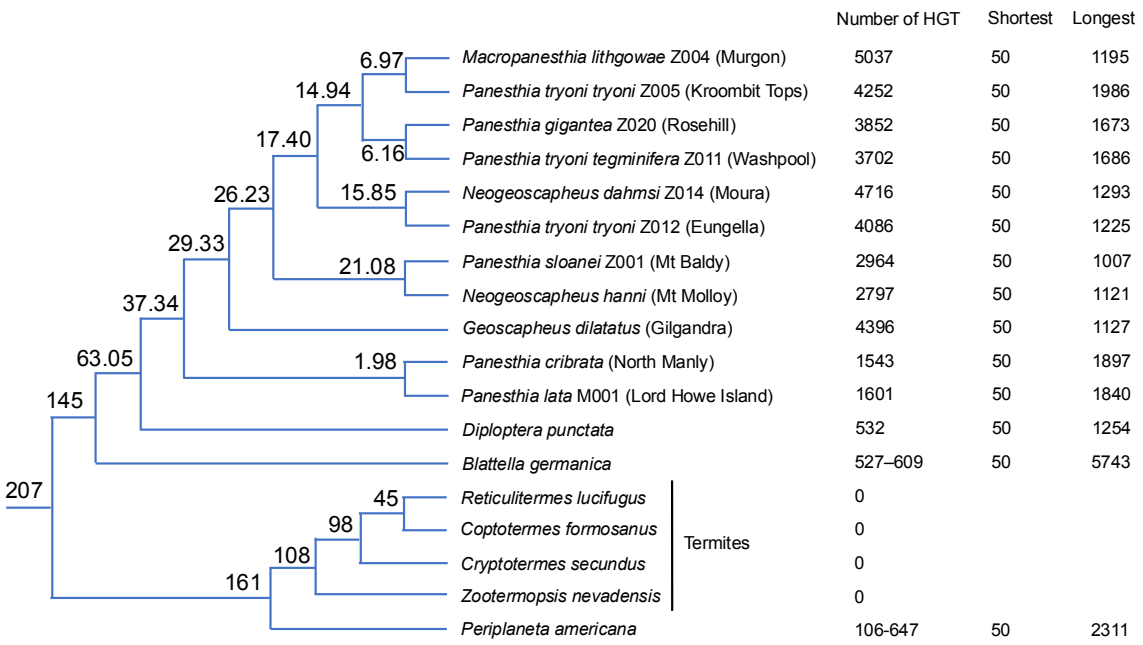



**Figure 4.2**. Nodes show divergence times (Ma) from Chapter 3 (for nodes internal to the top 12 taxa) and Evangelista et al (2019, for nodes internal to the remaining 6 taxa). One exception is the value between *P. lata* and *P. cribrata*, which is from Adams et al (2024). The tree shows only the topology, branch lengths are not scaled and therefore do not carry any evolutionary meaning. Due to differences in genome assembly quality of online genomes of *Periplaneta americana* and *Blattella germanica*, varying numbers of HGT inserts were recovered.

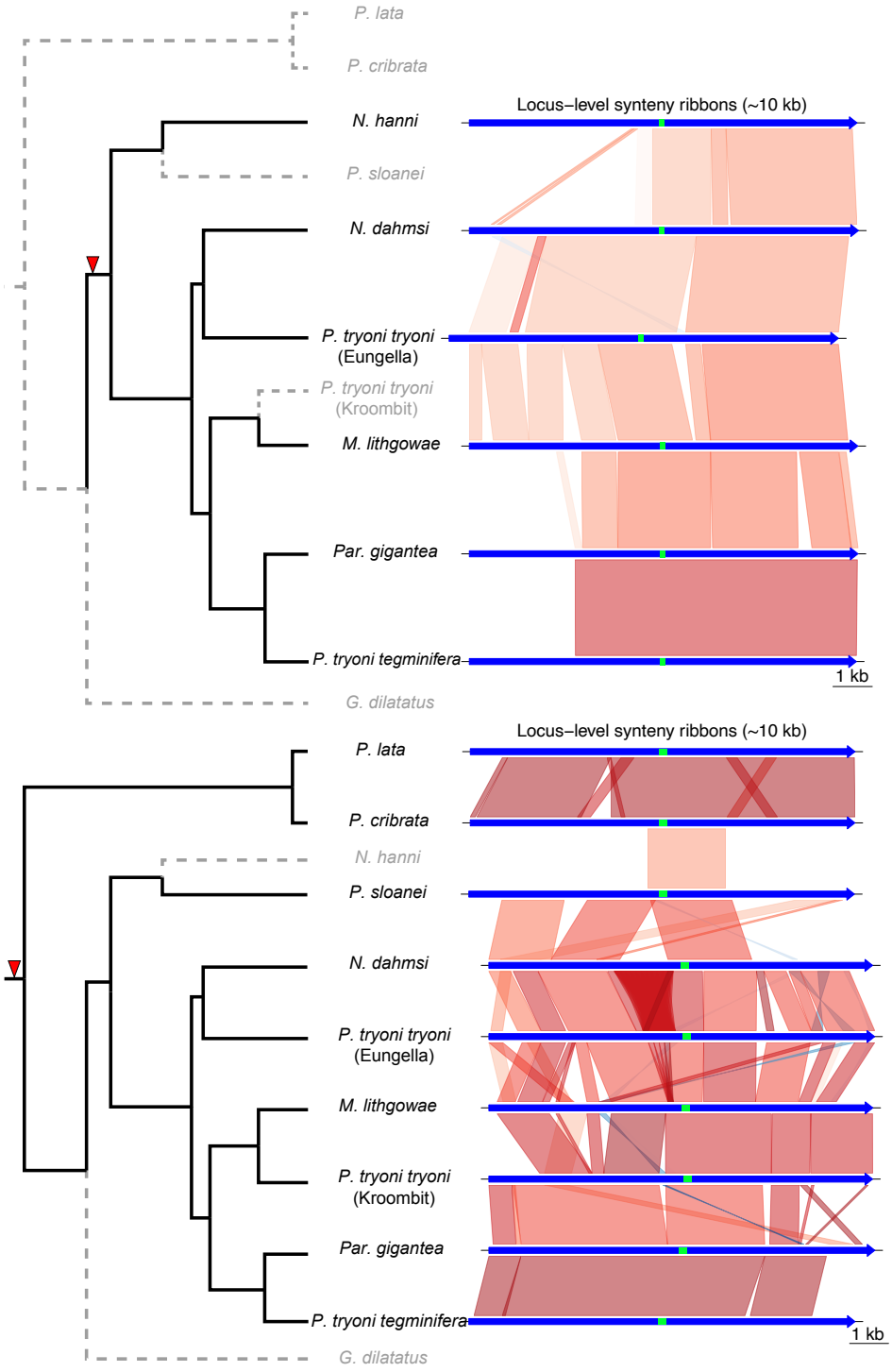



**Figure 4.3**. Examples of two HGT insertions used for 'aging' analysis. The right panel shows homology of each insertion and its 5 kb flanking regions across species: green denotes HGT inserts, blue flanking regions, with red and blue bands indicating forward and inverted homology (darker shades = higher similarity). The left phylogeny is based on Chapter 3; species lacking the insertion are shown as a dashed line. Red arrows indicate the inferred timing of HGT insertions.

During manual checking of the top 50 longest HGT inserts using BLAST, we observed the presence of short sequences (57 - 2940 bp) from distant parts of single *Blattabacterium* genomes that were joined together in a chimeric fashion, without any intervening sequence between them (Fig. 4.4). Of the 50 long HGT insertions we examined, we found evidence of such chimeras in 21 of them. These included 1 from *B. germanica*, 1 from *Per. americana*, 1 from *D. punctata*, 1 from *M. lithgowae*, 4 from *N. dahmsi*, 2 from *P. cribrata*, 3 from *Par. gigantea*, 2 from *P. lata*, 3 from *P. t. tegminifera*, 1 from *P. t. tryoni* (Kroombit), and 2 from *P. t. tryoni* (Eungella). Some of these 9 consisted of two non-contiguous sequences, while the majority of 12 consisted of more than two non-contiguous sequences, with a maximum of 9 short sections from different areas of one *Blattabacterium* genome joined together. We provide examples of an additional 5 of these chimeras in the Appendix (Fig. S4.2). In the other 29 sequences, the entire length of the sequence was derived from a single part of the *Blattabacterium* genome, and thus not chimeric.

We tested the hypothesis that the number of HGT insertions correlates with levels of relaxed selection in Panesthiinae and Geoscapheinae. Orthologous gene analyses identified widespread signals of both relaxed and positive selection across all lineages. However, these analyses did not detect a statistically significant correlation between the number of HGT events and the number of genes that had experienced relaxed selection in each species (Pearson's $r = 0.413$, two-tailed $p = 0.206$; Spearman's $\rho = 0.424$, two-tailed $p = 0.053$), nor between the wood-feeding and soil-borrowing sister groups (Pearson's $r = 0.402$, two-tailed $p = 0.221$).

We also tested whether or not the number of genes under relaxed or intensified selection differed between wood-feeding and soil-burrowing cockroaches; however, no significant difference was observed (relaxed selection: $t = 1.2793$, $p = 0.2287$, $df = 10.346$; intensified selection: $t = 0.5667$, $p = 0.5843$, $df = 9.3186$). Through a preliminary analysis of the functions of genes under relaxed selection in the wood-feeding (Panesthiinae) and soil-burrowing (Geoscapheinae) lineages, genes potentially associated with lifestyle shifts were identified in both lineages (101 in 206 in wood feeders, 30 in 124 in soil burrowers). However, gene ontology enrichment analysis did not reveal any significant enrichment between these groups (data not shown).

*P. tryoni tegminifera*: contig 33539:90566-92252

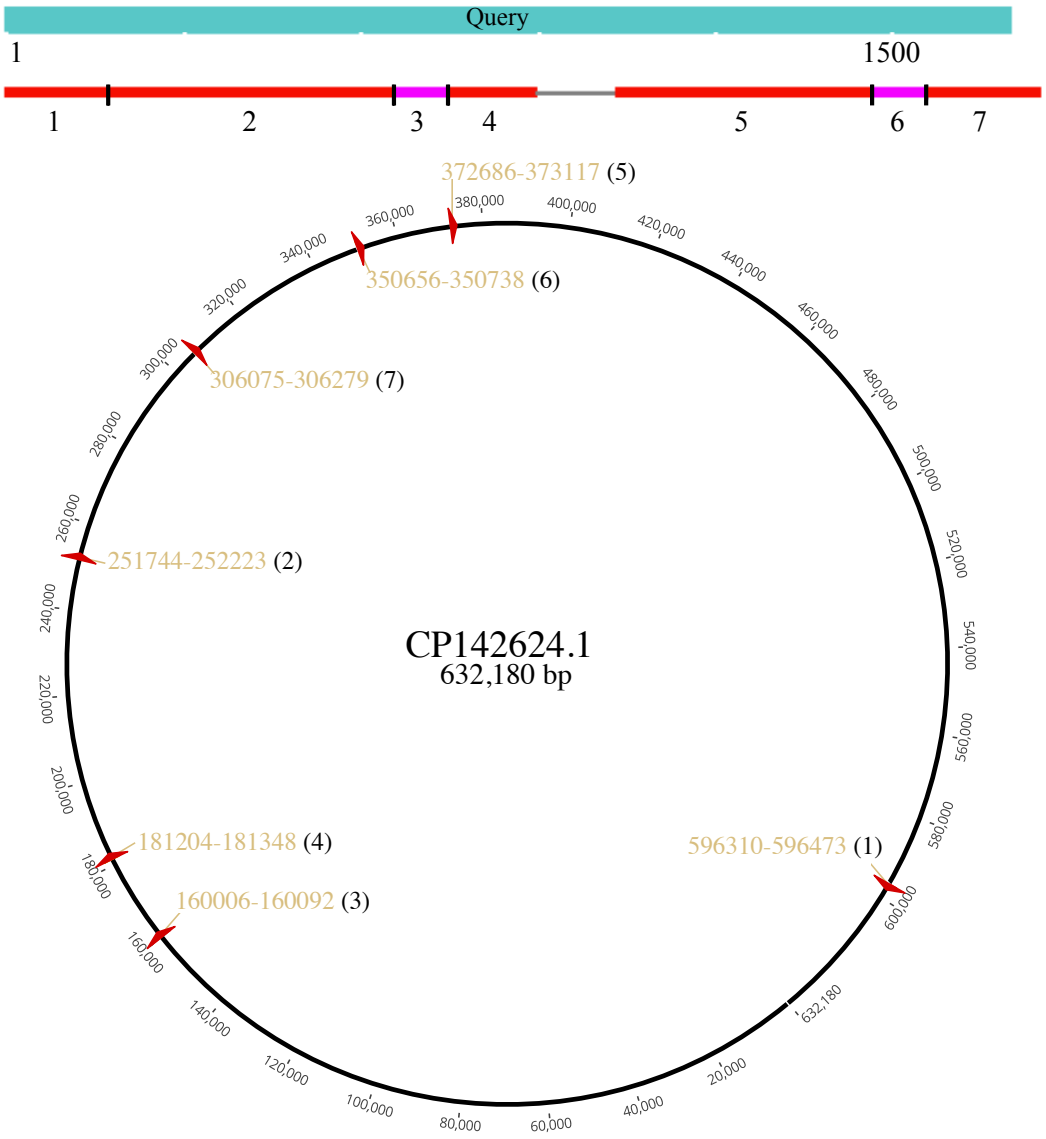



**Figure 4.4**. Example of a chimeric HGT insertion from *P. tryoni tegminifera*. The top sequence is divided into seven segments, while the circular map below represents the source genome of *Blattabacterium cuenoti* (CP142624.1). Numbers below the insertion indicate the corresponding genomic positions, with yellow labels marking specific coordinates.

## 4.5 Discussion

In this study, using a new method, we identified a significantly higher number of HGT events in cockroach genomes than previously reported (average of 596 in pest cockroaches and average of 3,541 in Panesthiinae + Geoscapheinae, but none in termites), and showed that these HGT events occurred up to 37 Mya (range 1.98 – 37.34, average of 20.67 Mya). Li et al. (2022) examined representatives of 11 insect orders, using a method based on protein sequences and the Alien Index (this index estimates the probability of horizontal gene transfer for insect genes by measuring their sequence conservation across close and distant relatives) of genes to identify HGT insertions in most of these orders. The number of HGT insertions they found in these insects ranged from 0 to 170, with 79.0% originating from bacteria. However, they did not find any HGT in the three species of the order Blattodea. Other studies that have examined HGT have also primarily relied on coding sequences for their searches (Acuña et al., 2012; Li et al., 2021; Liu et al., 2025; Nakabachi, 2015). The highest number of insertions for any species identified in our results (5,037 in *M. lithgowae*) is several dozen times higher than the highest number previously found in any organism (170 in *Bemisia tabaci*; Li et al., 2022), with the exception of rotifers, which are known to harbour very large numbers of HGTs (up to 8% of their entire genome, as a result of anhydrobiosis, (Flot et al., 2013). The primary reason for this phenomenon is that our HGT detection method includes both genic and non-genic regions, whereas non-coding regions were largely ignored in most previous studies.

Although studies on horizontal transfer to non-coding regions are limited, some long non-coding RNAs (lncRNAs) potentially acquired via HGT have been identified in humans (Pierneef et al., 2018), and *Wolbachia* DNA, which is widely present in insects, has been found to contain a large number of potentially non-functional repeats in *Drosophila* (Klasson et al., 2014; Woolfit et al., 2015). Our method for locating HGT insertions does not distinguish between coding and non-coding regions, and therefore likely detects a large number of HGTs in non-coding regions that traditional methods did not. Although HGT insertions in non-coding regions may not confer obvious functions, non-coding sequences constitute the majority of eukaryotic genomes by length, and they may serve as sources for the origin of new genes

(Levine et al., 2006; McLysaght and Guerzoni, 2015). Accordingly, a large pool of HGT insertions in non-coding regions could still be of macro-evolutionary significance.

Secondly, it is worth noting that previous studies on HGT have primarily focused on longer transfer fragments—generally at least several hundred base pairs (≥ ~300 bp)—to effectively identify and validate their functions (Li et al., 2022; Xing et al., 2023). For example, nearly all of 1,410 HGT genes in 218 high-quality insect genomes identified by Li et al. (2022) were within the range of longer gene lengths (hundreds to thousands of bp). However, our study identified a large number of short-length HGT insertion events (less than 100 bp, average insertion count per species in Panesthiinae + Geoscapheinae = 847). Although these short sequences invariably lack complete open reading frames and are unlikely to encode functional proteins, they may still influence host genome function in multiple ways: firstly, their insertion may disrupt the structure of the original gene or its regulatory regions, leading to gene inactivation or abnormal expression, like exonization (Schmitz and Brosius, 2011; Sela et al., 2007); secondly, these short exogenous fragments may form new regulatory elements (e.g., enhancers or silencers), reshaping gene expression patterns (Ullastres et al., 2021); and finally, if inserted within the coding region and integrated into the original reading frame, they may introduce new functional domains or modification sites, expanding the molecular properties of the protein. These potential mechanisms suggest that although short HGT insertion fragments may not directly encode proteins, they still hold profound biological and evolutionary significance in genome structural modification, expression regulation, and potential functional innovation.

*Blattabacterium* migrates from the fat body cells surrounding the ovaries to the ovaries of nymphs during the formation of cockroach oocytes and infects the oocytes, allowing their vertical transmission with the host (Noda et al., 2025). We propose two hypotheses to explain the mechanism behind the extensive HGT insertions during this period: (1) RNA produced during the bacterial life cycle may integrate into the host cockroach genome via endogenous retrotransposons, which may similar to the R2 retrotransposons discovered in *Drosophila melanogaste*r (Eickbush and Eickbush, 2015; Luan et al., 1993; Nelson et al., 2023); (2) DNA fragments produced during the life cycle of *Blattabacterium* or after its death may be integrated into the cockroach genome via mechanisms similar to non-homologous end joining (NHEJ,

Doré et al., 2004) or transposon mechanisms. The identification of chimeric *Blattabacterium* sequences within HGTs is suggestive of the second mechanism occurring, probably by NHEJ. A similar case has also been reported in *Arabidopsis thaliana*, where the mitochondrial genome has been inserted into chromosome 2 in a scrambled and partially duplicated form (Fields et al., 2022). However, unlike our findings, nearly the entire mitochondrial genome was incorporated into the nuclear genome of *Arabidopsis*, rather than only a few small fragments(Fields et al., 2022). This difference may reflect distinct insertion mechanisms, although further investigation is needed to clarify this process.

Chimeric inserts may play a role in the evolution of Panesthiinae and Geoscapheinae. Chimeric genes can both generate new functional proteins through domain rearrangement or regulatory element capture (Frenkel-Morgenstern and Valencia, 2012; Fu et al., 2010), and avoid harmful mutations, leveraging the streamlined characteristics of endosymbiotic bacterial genomes (Sela et al., 2016), to efficiently produce new adaptive genes. For example, among the 28 newly formed genes (later than 1 Mya) identified on the short arm of chromosome 3 (Chr3) in rice (*Oryza sativa*), 14 are believed to originate from chimeras (Zhang et al., 2013). Related to endosymbionts, the well-known symbiogenetic genes (S-genes) originating from endosymbionts trace back to the formation of eukaryotes, arising from sequence chimeras between bacterial and archaeal sequences (Méheust et al., 2018). These genes have made significant contributions to the increase in cellular complexity. However, how chimeric genes found in cockroaches are formed, and whether they are transcribed and expressed, requires further investigation.

Ewart et al. (2024) found that, compared to their close relatives, Panesthiinae/Geoscapheinae exhibit significantly higher levels of genomic-level relaxed selection than other cockroach species, which may be related to the presence of much higher numbers (~10 times more) of HGT insertions in some of the species we examined (e.g. *M. lithgowae*). In the present study, we tested whether the variation in HGT numbers seen among the Panesthiinae/Geoscapheinae we examined (from 1543 in *P. cribrata* up to 5037 in *M. lithgowae*) was correlated with the level of relaxed selection seen in their genomes. However, we did not find any such correlation. The reasons why some taxa have higher numbers of HGTs than others, therefore, remain unclear.

Ewart et al. (2024) reported significantly more genes under relaxed selection in termites (24–31%) compared to cockroaches (2–4%), attributing this pattern to reduced effective population sizes (Ne), which dampen selection on reproductive individuals. They also mentioned that Panesthiinae + Geoscapheinae have a higher level of relaxed selection than other free-living cockroaches. As Geoscapheinae experienced multiple parallel lifestyle and family-structure shifts relative to their wood-feeding Panesthiinae ancestors, we performed similar analyses to compare the positive and relaxed selection levels in Geoscapheinae and Panesthiinae, and also explored gene changes that may have caused their parallel evolution.

We found that despite multiple independent niche shifts (wood-feeding to soil-burrowing) in the Geoscapheinae, there is no significant difference in the number of genes under relaxed selection and positive selection between the Geoscapheinae and Panesthiinae subfamilies (t-test, positive selection: $p = 0.574$, $df = 11$; relaxed selection: $p = 0.224$, $df = 11$). This suggests that: (1) Positive selection on key adaptive traits (e.g., nitrogen metabolism genes, mouthpart structure) may mask relaxed selection signals at the genome-wide level; (2) Ecological niche shifts have not led to a systematic reduction in the intensity of purifying selection—these contrasts sharply with the genomic relaxation patterns observed in termites (Ewart et al., 2024).

During the transition from a wood-feeding to a soil-burrowing lifestyle, numerous instances of parallel evolution in burrowing cockroaches occurred. For example, body shape changed from an elongate oval to more compact or nearly spherical forms, drought tolerance increased, and females formed stable family units with their offspring. These adaptive traits emerged in multiple independent burrowing lineages, despite their origins in different wood-feeding ancestors. However, these transitions did not translate into a detectable number of genes under relaxed selection. Meanwhile, an increase in intensified selection, which may result from the selection of alleles that enhance adaptation to the new environment (Beasley-Hall et al., 2024; Bolnick et al., 2010; Lahti et al., 2009), was also not observed. In soil burrowers, compared to wood feeders, genes that we found under relaxed selection are primarily concentrated in water conservation, detoxification, immunity, and feeding-related functions. In wood feeders, some genes related to these functions are also under relaxed selection, but additional genes related to motion and oxygen acquisition are present. However, the genes most commonly under relaxed

selection in our samples do not appear to be associated with key functions in these ecological niche transitions.

During arthropod adaptation to novel ecological niches, positive selection on key genes often drives early evolutionary change. This is exemplified by grasshoppers such as the migratory *Chondracris rosea*, where mitochondrial genes critical for energy metabolism (ATP8, ND4, ND5) underwent intense positive selection ($\omega > 1$), presumably to meet the high energetic demands of flight, and thus allowing driving their rapid adaptation to new habitats (Li et al., 2018). In gentoo penguins (*Pygoscelis papua*), despite background purifying selection (genome-wide $dN/dS < 1$), lineage-specific positive selection on mitochondrial ND1 (in eastern lineages) and ND4 (in southern lineages)—key for thermogenesis and ATP efficiency in local Antarctic environments—overshadowed signals of relaxed selection, demonstrating how niche-specialised adaptation dominates genomic evolution (Noll et al., 2022). This phenomenon illustrates how intense positive selection linked to essential ecological adaptations ("niche-transition genes") can eclipse background signals of relaxed selection. However, at the genome-wide level, the lifestyle transition in Geoscapheinae did not result in a detectable change in overall selection intensity. The absence of genomic relaxation in Geoscapheinae may be explained by ongoing intraspecific competition. Their burrowing habits limit dispersal capacity, resulting in persistently high local population densities. Moreover, all lineages depend on the same resource—dried leaves—intensifying resource competition. Their current distributions show substantial overlap, potentially leading to rapid post-divergence competition among lineages. This overlap and competition likely maintain strong selective pressures across the genome, preventing the emergence of a genome-wide relaxed selection signal. Consistent with habitat fragmentation theory, selective pressure due to competition may increase in isolated populations even as effective population size (Ne) decreases. Although limited dispersal ability is common across many organisms, few studies have examined its impact on genome-wide selection patterns, and no comparable cases have yet been reported.

# Chapter 5

## General discussion

### 5.1 Thesis summary and future directions

The studies in this thesis focused on the evolution of Geoscapheinae and Panesthiinae from the aspects of phylogeny, biogeography, molecular evolution, and comparative genomics. The first chapter introduces recent progress on phylogenetic research on Blattodea, and the biology of Geoscapheinae and Panesthiinae, their endosymbiont *Blattabacterium*, and horizontal gene transfer (HGT), which was found abundantly in the Geoscapheinae and Panesthiinae genomes in Chapter 4.

Chapter 2 focuses on the evolutionary history of *Panesthia sloanei* and *Panesthia tryoni*. New samples of these two species were collected across most of their known distribution ranges, and phylogenetic and biogeographic analyses were conducted using mitochondrial genomes and several nuclear markers. All samples clustered with conspecific samples from previous studies, and all previously identified wood-feeding and soil-burrowing sister groups were confirmed. Notably, two clades, *P. australis* + *P. obtusa* and *Geoscapheus dilatatus* + *G. robustus*, received strong support for the first time. In particular, the *G. dilatatus* + *G. robustus* lineage was reconstructed as the earliest diverged clade within the first-wave Australian *Panesthia* + Geoscapheinae clade, suggesting the existence of a previously unrecognised lineage of wood-feeders, and from which the ancestors of *G. dilatatus* and *G. robustus* shifted to a soil-burrowing lifestyle. Biogeographic analyses revealed that the distribution of *P. sloanei* and *P. tryoni* lineages showed a strong correlation with recognised biogeographic divisions. Overall, this study underscores the intricate evolutionary history and recurrent ecological shifts that have shaped this distinctive group of cockroaches.

Chapter 3 focuses on the evolutionary history of Panesthiinae (including Geoscapheinae). We obtained as many Asian Panesthiinae samples as possible and conducted phylogenetic

analyses using mitochondrial genomes with several nuclear markers. In addition, we analysed a set of morphological traits, including the wing, three male genital phallomeres (L1, L2d, and R2) and the oothecal membrane. By integrating molecular and morphological results, we proposed a revised classification of Panesthiinae (including Geoscapheinae). Specifically, our analyses recovered eight major lineages, with *Salganea* and *Miopanesthia* forming sister groups, *Caeparia* as sister to *Panesthia transversa*, and *Ancaudellia* nested within *Panesthia*. The placement of the Geoscapheinae clade was consistent with results from the chapter focusing on Australian taxa. Morphological analyses further revealed repeated reductions of wings, male genitalia, and the oothecal membrane in both Panesthiinae and Geoscapheinae, reflecting multiple independent evolutionary events. These patterns underscore the role of ecological shifts and mating strategies in driving convergent morphological simplification within this lineage of cockroaches.

Chapter 4 investigated horizontal gene transfer (HGT) from the obligate endosymbiont *Blattabacterium* to the genomes of 18 cockroach and termite species, with a focus on the subfamilies Panesthiinae and Geoscapheinae. Using long-read Oxford Nanopore sequencing, we assembled and annotated high-quality genomes for ten species. Using eight of which, together with ten previously published Blattodea genomes, we revealed an unprecedented number of HGT inserts—ranging from 106 to 5037 per genome—far exceeding previous estimates. Notably, Panesthiinae and Geoscapheinae species harboured thousands of inserts, whereas none were found in the termites examined. Through rigorous filtering and validation, we confirmed that these inserts are genuine integrations within host genomes, not artifacts of assembly or contamination. Age estimates based on flanking sequence homology indicated that most transfers occurred within the past 40 million years, with a median age of approximately 20.67 million years. We also identified chimeric inserts formed from non-contiguous *Blattabacterium* fragments. Contrary to expectations, we found no significant correlation between HGT abundance and levels of relaxed selection, nor did we detect a difference in selection regimes between wood-feeding and soil-burrowing lineages. This suggests that purifying selection on key adaptive traits may offset genome-wide relaxation signals. Our results underscore the role of HGT in genomic evolution and highlight the need to consider non-coding regions in future studies of horizontal transfer.

With the rapid development of sequencing technologies, phylogenetic studies have advanced considerably in recent years, and some previously unresolved relationships have gradually reached more stable conclusions. Nevertheless, many species described during the morphological taxonomy era have yet to be subject to DNA sequencing and analysis. This gap is particularly evident in insects, where specimens were often preserved in dried form, making their genetic sequences difficult to determine. At the same time, continuous improvements in genome sequencing technologies, especially in long-read platforms, have greatly increased both the efficiency and accuracy of genome assembly. These advances now enable the exploration of genomic regions that were previously difficult to analyse.

Within Blattodea, the accumulation of molecular data and the application of new analytical approaches have gradually clarified many previously unresolved phylogenetic relationships. However, due to incomplete sampling, the placement of several less-studied lineages remains uncertain. This issue is particularly evident in Panesthiinae, the main focus of this thesis. Early studies, which included relatively few species of Panesthiinae and Geoscapheinae, suggested that Geoscapheinae formed a monophyletic group (Maekawa et al., 2003). With increased taxon sampling, however, it became apparent that Geoscapheinae and Panesthiinae were intermingled, resulting in both subfamilies being rendered paraphyletic (Lo et al., 2016). Subsequent study that further expanded sampling revealed that Geoscapheinae largely originated from two key *Panesthia* species, *P. sloanei* and *P. tryoni*, although some weakly supported clades remained problematic (Beasley-Hall et al., 2021). In Chapter 2 of this thesis, I expanded sampling for *P. sloanei* and *P. tryoni*. Although no new lineages of these two species were detected, several previously weakly supported clades received strong support, but in different positions compared to previous studies. In particular, I found that the *G. dilatatus* + *G. robustus* lineage did not diverge from either *P. sloanei* or *P. tryoni*, but formed the first clade of the first-wave Australian Panesthia + Geoscapheinae. Thus, the phylogenetic relationships among the first wave of Australian *Panesthia* and Geoscapheinae have been largely clarified, although significant sampling gaps remain for the Asian representatives.

Within Asian Panesthiinae, currently available samples allow the recognition of several well-supported clades, including the two Australian *Panesthia* lineages, *Caeparia*, *Miopanesthia*, and *Salganea*. Nevertheless, several issues remain unresolved. First, the phylogenetic

relationships among the remaining Asian *Panesthia* and *Ancaudellia* are still weakly supported, and many *Panesthia* species have not yet been sampled—most notably those from India, which represents the known distributional boundary of the genus. This absence of Indian material severely limits our ability to reconstruct the early dispersal history of *Panesthia*. In addition, insufficient sampling across the numerous islands of Southeast Asia hampers our understanding of the timing and routes of insular colonisation. Although *Miopanesthia* and *Salganea* consistently form a strongly supported monophyletic group, all molecular studies of *Miopanesthia* to date have relied on just two samples of a single species (Beasley-Hall et al., 2021; Lo et al., 2016; Maekawa et al., 2003); thus, broader taxon coverage is required to robustly test the monophyly of the genus. Finally, of the seven recognised genera of Panesthiinae, two monotypic genera, *Annamoblatta* Sergeev, 1984 and *Microdina* Kirby, 1903, have never been sequenced or included in phylogenetic analyses. Incorporating these lineages represents a key priority for future research on Panesthiinae systematics.

With regard to horizontal gene transfers (HGTs) in Panesthiinae and Geoscapheinae, the analyses presented in this thesis remain somewhat preliminary due to time limitations. Future work could employ PCR amplification using primers designed from the identified insertions, followed by sequencing of flanking regions to verify their genomic integration within cockroaches. Nevertheless, the use of long-read sequencing technology provides strong evidence for the presence of such HGTs. Another important question is whether these insertions have functional significance. As a first step, the recovered sequences could be compared against the annotated coding regions to identify potentially functional inserts, which could then be subjected to more detailed characterisation. Analyses involving RNA transcripts, and examination of whether any HGT inserts are found within protein-coding genes, would allow testing of whether any inserts have a function. Additional analyses could look at regulatory regions, small RNAs, to see whether the inserts we found might be involved in these functional sequences.

In this thesis, we also reported the presence of chimeric inserts, but did not investigate them further. If functional elements can be identified within these chimeric sequences, the evolutionary origins and possible emergent functions of such mosaics would represent a

particularly intriguing avenue for future research. Moreover, the mechanisms underlying HGT insert formation remain unresolved. This may require cellular-level approaches, such as localising free DNA fragments within cockroach oocytes, and determining whether discontinuous segments from the *Blattabacterium* genome can become joined prior to integration, an observation that would provide strong evidence for the genesis of chimeric inserts. Finally, while endosymbionts are widespread across cockroaches and other insects (Eleftherianos et al., 2013; Gibson and Hunter, 2010; McCutcheon et al., 2019), large-scale transfers from endosymbionts into host genomes have not been widely reported outside of Blattodea. Whether this reflects a true biological rarity or instead highlights limitations of current detection methods remains an open and important question.

# References


Acuña, R., Padilla, B. E., Flórez-Ramos, C. P., Rubio, J. D., Herrera, J. C., Benavides, P., Lee, S.-J., Yeats, T. H., Egan, A. N., Doyle, J. J., & Rose, J. K. C. (2012). Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. *Proceedings of the National Academy of Sciences*, *109*(11), 4197–4202. https://doi.org/10.1073/pnas.1121190109

Adams, M. W. D., Walker, J. A., Rose, H. A., Jones, B. R., Zwick, A., Yang, H., Nicholls, J., Hartley, D., Bent, S., Carlile, N., Hutton, I., Ho, S. Y. W., & Lo, N. (2024). Plio-Pleistocene decline of mesic forest underpins diversification in a clade of Australian *Panesthia* cockroaches. *Systematic Entomology*. https://doi.org/10.1111/syen.12654

Andersen, J. C., Wu, J., Gruwell, M. E., Gwiazdowski, R., Santana, S. E., Feliciano, N. M., Morse, G. E., & Normark, B. B. (2010). A phylogenetic analysis of armored scale insects (Hemiptera: Diaspididae), based upon nuclear, mitochondrial, and endosymbiont gene sequences. *Molecular Phylogenetics and Evolution*, *57*(3), 992–1003.

Arab, D. A., Bourguignon, T., Wang, Z., Ho, S. Y. W., & Lo, N. (2020). Evolutionary rates are correlated between cockroach symbionts and mitochondrial genomes. *Biology Letters*, *16*(1). https://doi.org/10.1098/rsbl.2019.0702

Araldi-Brondolo, S. J., Spraker, J., Shaffer, J. P., Woytenko, E. H., Baltrus, D. A., Gallery, R. E., & Arnold, A. E. (2017). Bacterial Endosymbionts: Master Modulators of Fungal Phenotypes. *Microbiology Spectrum*, *5*(5). https://doi.org/10.1128/microbiolspec.FUNK-0056-2016

Arnold, B. J., Huang, I.-T., & Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nature Reviews Microbiology*, *20*(4), 206–218. https://doi.org/10.1038/s41579-021-00650-4

Australian Government Department of Climate Change, E. the E. and W. (2023, July 14). *Australia's bioregions (IBRA)*. https://www.dcceew.gov.au/environment/land/nrs/science/ibra#daff-page-main

Bale, C., & Williams, J. (1993). Composition and Classification of Nothofagus moorei Communities in Northern New South Wales. *Australian Systematic Botany*, *6*(5), 429. https://doi.org/10.1071/SB9930429

Bandi, C., Sironi, M., Damiani, G., Magrassi, L., Nalepa, C. A., Laudani, U., & Sacchi, L. (1995). The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *259*(1356), 293–299.

Bank, S., & Bradler, S. (2022). A second view on the evolution of flight in stick and leaf insects (Phasmatodea). *BMC Ecology and Evolution*, *22*(1), 62. https://doi.org/10.1186/s12862-022-02018-5

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Batra, L. R., & Buchner, P. (1968). Endosymbiosis of Animals with Plant Microorganisms. *Mycologia*, *60*(2), 466. https://doi.org/10.2307/3757184

Beasley-Hall, P. G., Kinjo, Y., Rose, H. A., Walker, J., Foster, C. S. P., Kovacs, T. G. L., Bourguignon, T., Ho, S. Y. W., & Lo, N. (2024). Shrinking in the dark: Parallel endosymbiont genome erosions are associated with repeated host transitions to an underground life. *Insect Science*, *31*(6), 1810–1821. https://doi.org/10.1111/1744-7917.13339

Beasley-Hall, P. G., Lee, T. R. C., Rose, H. A., & Lo, N. (2018). Multiple abiotic factors correlate with parallel evolution in Australian soil burrowing cockroaches. *Journal of Biogeography*, *45*(7), 1515–1528. https://doi.org/10.1111/jbi.13233

Beasley-Hall, Perry G, Rose, H. A., Bourguignon, T., & Lo, N. (2021). Molecular systematics and biogeography of an Australian soil-burrowing cockroach with polymorphic males, *Geoscapheus dilatatus* (Blattodea: Blaberidae). *Austral Entomology*, *60*(2), 317–329. https://doi.org/10.1111/aen.12527

Beasley-Hall, Perry G., Rose, H. A., Walker, J., Kinjo, Y., Bourguignon, T., & Lo, N. (2021). Digging deep: a revised phylogeny of Australian burrowing cockroaches (Blaberidae: Panesthiinae, Geoscapheinae) confirms extensive nonmonophyly and provides insights into biogeography and evolution of burrowing. *Systematic Entomology*, *46*(4), 767–783. https://doi.org/10.1111/syen.12487

Beccaloni, G. (2024). *Cockroach Species File Online. Version 5.0/5.0 World Wide Web electronic publication (2014)*. https://cockroach.speciesfile.org/

Beccaloni, G., & Eggleton, P. (2013a). Order Blattodea. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013). *Zootaxa*, *3703*(1), 46. https://doi.org/10.11646/zootaxa.3703.1.10

Beccaloni, G., & Eggleton, P. (2013b). Order Blattodea. In: Zhang, Z.-Q. (Ed.) Animal Biodiversity: An Outline of Higher-level Classification and Survey of Taxonomic Richness (Addenda 2013). *Zootaxa*, *3703*(1), 46. https://doi.org/10.11646/zootaxa.3703.1.10

Bell, R. C., MacKenzie, J. B., Hickerson, M. J., Chavarría, K. L., Cunningham, M., Williams, S., & Moritz, C. (2012). Comparative multi-locus phylogeography confirms multiple vicariance events in co-distributed rainforest frogs. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1730), 991–999. https://doi.org/10.1098/rspb.2011.1229

Bell, W. J., Louis M. Roth, & Christine A. Nalepa. (2007). *Cockroaches: ecology, behavior, and natural history*. JHU Press.

Bennett, G. M., Abbà, S., Kube, M., & Marzachì, C. (2016). Complete Genome Sequences of the Obligate Symbionts " *Candidatus* Sulcia muelleri" and " *Ca.* Nasuia deltocephalinicola" from the Pestiferous Leafhopper *Macrosteles quadripunctulatus* (Hemiptera: Cicadellidae). *Genome Announcements*, *4*(1). https://doi.org/10.1128/genomeA.01604-15

Bennett, G. M., Kwak, Y., & Maynard, R. (2024). Endosymbioses Have Shaped the Evolution of Biological Diversity and Complexity Time and Time Again. *Genome Biology and Evolution*, *16*(6). https://doi.org/10.1093/gbe/evae112

Beutel, R. G., & Gorb, S. N. (2006). A revised interpretation of the evolution of attachment structures in Hexapoda with special emphasis on Mantophasmatodea. *Arthropod Systematics & Phylogeny*, *64*(1), 3–25.

Bhat, A., Sharma, R., Desigan, K., Lucas, M. M., Mishra, A., Bowers, R. M., Woyke, T., Epstein, B., Tiffin, P., Pueyo, J. J., & Paape, T. (2024). Horizontal gene transfer of the Mer operon is associated with large effects on the transcriptome and increased tolerance to mercury in nitrogen-fixing bacteria. *BMC Microbiology*, *24*(1), 247. https://doi.org/10.1186/s12866-024-03391-5

Bickerstaff, J. R. M., Jordal, B. H., & Riegler, M. (2025). Biogeography and Social Family Structure Contribute to Cryptic Genomic Divergence in the Only Obligate Eusocial Beetle Species, *Austroplatypus incompertus* (Curculionidae: Platypodinae). *Molecular Ecology*, *34*(22). https://doi.org/10.1111/mec.70076

Bläser, M., Misof, B., & Predel, R. (2020). The power of neuropeptide precursor sequences to reveal phylogenetic relationships in insects: A case study on Blattodea. *Molecular*

*Phylogenetics and Evolution*, *143*, 106686. https://doi.org/10.1016/j.ympev.2019.106686

Blum, M., Andreeva, A., Florentino, L. C., Chuguransky, S. R., Grego, T., Hobbs, E., Pinto, B. L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., Salazar, G. A., Bordin, N., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Llinares-López, F., … Bateman, A. (2025). InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Research*, *53*(D1), D444–D456. https://doi.org/10.1093/nar/gkae1082

Bolnick, D. I., Ingram, T., Stutz, W. E., Snowberg, L. K., Lau, O. L., & Paull, J. S. (2010). Ecological release from interspecific competition leads to decoupled changes in population and individual niche width. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1689), 1789–1797. https://doi.org/10.1098/rspb.2010.0018

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bourguignon, T., Tang, Q., Ho, S. Y. W., Juna, F., Wang, Z., Arab, D. A., Cameron, S. L., Walker, J., Rentz, D., Evans, T. A., & Lo, N. (2018). Transoceanic dispersal and plate tectonics shaped global cockroach distributions: evidence from mitochondrial phylogenomics. *Molecular Biology and Evolution*, *35*(4), 970–983. https://doi.org/10.1093/molbev/msy013

Brennan, P. L. R., & Prum, R. O. (2015). Mechanisms and Evidence of Genital Coevolution: The Roles of Natural Selection, Mate Choice, and Sexual Conflict. *Cold Spring Harbor Perspectives in Biology*, *7*(7), a017749. https://doi.org/10.1101/cshperspect.a017749

Brentassi, M. E., & de la Fuente, D. (2024). Obligate mutualistic heritable symbiosis in sap-feeding insects: an intricate relationship in nature. *Symbiosis*, *92*(2), 159–185. https://doi.org/10.1007/s13199-023-00965-1

Brooks, M. A., & Richards, A. G. (1955). Intracellular symbiosis in cockroaches. I. Production of aposymbiotic cockroaches. *The Biological Bulletin*, *109*(1), 22–39.

Brown, M., Cooksley, H., Carthew, S. M., & Cooper, S. J. B. (2006). Conservation units and phylogeographic structure of an arboreal marsupial, the yellow-bellied glider (Petaurus australis). *Australian Journal of Zoology*, *54*(5), 305. https://doi.org/10.1071/ZO06034

Brunner von Wattenwyl, K. (1865). *Nouveau système des blattaires*. G. Braumüller. https://doi.org/10.5962/bhl.title.5652

Bryant, L. M., & Fuller, S. J. (2014). Pleistocene climate fluctuations influence phylogeographical patterns in *Melomys cervinipes* across the mesic forests of eastern Australia. *Journal of Biogeography*, *41*(10), 1923–1935. https://doi.org/10.1111/jbi.12341

Bryant, L. M., & Krosch, M. N. (2016). Lines in the land: a review of evidence for eastern Australia's major biogeographical barriers to closed forest taxa. *Biological Journal of the Linnean Society*, *119*(2), 238–264. https://doi.org/10.1111/bij.12821

Byrne, M., Yeates, D. K., Joseph, L., Kearney, M., Bowler, J., Williams, M. A. J., Cooper, S., Donnellan, S. C., Keogh, J. S., & Leys, R. (2008). Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology*, *17*(20).

Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular Biology and Evolution*, *38*(12), 5825–5829. https://doi.org/10.1093/molbev/msab293

Cao, T., & Jin, J.-P. (2020). Evolution of Flight Muscle Contractility and Energetic Efficiency. *Frontiers in Physiology*, *11*. https://doi.org/10.3389/fphys.2020.01038

Carpenter, R. J., Goodwin, M. P., Hill, R. S., & Kanold, K. (2011). Silcrete plant fossils from Lightning Ridge, New South Wales: new evidence for climate change and monsoon elements in the Australian Cenozoic. *Australian Journal of Botany*, *59*(5), 399. https://doi.org/10.1071/BT11037

Cayetano, L., Maklakov, A. A., Brooks, R. C., & Bonduriansky, R. (2011). EVOLUTION OF MALE AND FEMALE GENITALIA FOLLOWING RELEASE FROM SEXUAL SELECTION. *Evolution*, *65*(8), 2171–2183. https://doi.org/10.1111/j.1558-5646.2011.01309.x

Chen, X., Li, S., & Aksoy, S. (1999). Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus Glossina and its bacteriome-associated endosymbiont, Wigglesworthia glossinidia. *Journal of Molecular Evolution*, *48*(1), 49–58.

Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, *162*(1), 156–159. https://doi.org/10.1016/0003-2697(87)90021-2

Clark, E. L., Karley, A. J., & Hubbard, S. F. (2010). Insect endosymbionts: manipulators of insect herbivore trophic interactions? *Protoplasma*, *244*(1–4), 25–51. https://doi.org/10.1007/s00709-010-0156-2

Coady, S. K., Adams, M. W. D., Rose, H. A., Walker, J. A., Hutton, I., Gloag, R. S., & Lo, N. (2025). Island syndrome in the critically endangered Lord Howe Island cockroach Panesthia lata. *BioRxiv*, 2025–2028.

Cobbs, C., Heath, J., Stireman, J. O., & Abbot, P. (2013). Carotenoids in unexpected places: Gall midges, lateral gene transfer, and carotenoid biosynthesis in animals. *Molecular Phylogenetics and Evolution*, *68*(2), 221–228. https://doi.org/10.1016/j.ympev.2013.03.012

Colgan, D. J., O'Meally, D., & Sadlier, R. A. (2009). Phylogeographic patterns in reptiles on the New England Tablelands at the south-western boundary of the McPherson Macleay Overlap. *Australian Journal of Zoology*, *57*(5), 317. https://doi.org/10.1071/ZO08088

Collin, R., & Miglietta, M. P. (2008). Reversing opinions on Dollo's Law. *Trends in Ecology & Evolution*, *23*(11), 602–609. https://doi.org/10.1016/j.tree.2008.06.013

Couper, P. J., Sadlier, R. A., Shea, G. M., & Wilmer, J. W. (2008). A reassessment of Saltuarius swaini (Lacertilia: Diplodactylidae) in southeastern Queensland and New South Wales; two new taxa, phylogeny, biogeography and conservation. *Records of the Australian Museum*, *60*(1), 87–118. https://doi.org/10.3853/j.0067-1975.60.2008.1492

Crisp, M. D., Laffan, S., Linder, H. P., & Monro, A. (2001). Endemism in the Australian flora. *Journal of Biogeography*, *28*(2), 183–198. https://doi.org/10.1046/j.1365-2699.2001.00524.x

Cummings, T. F. M., Gori, K., Sanchez-Pulido, L., Gavriilidis, G., Moi, D., Wilson, A. R., Murchison, E., Dessimoz, C., Ponting, C. P., & Christophorou, M. A. (2022). Citrullination Was Introduced into Animals by Horizontal Gene Transfer from Cyanobacteria. *Molecular Biology and Evolution*, *39*(2). https://doi.org/10.1093/molbev/msab317

Cunningham, M., & Moritz, C. (1998). Genetic effects of forest fragmentation on a rainforest restricted lizard (Scincidae: Gnypetoscincus queenslandiae). *Biological Conservation*, *83*(1), 19–30. https://doi.org/10.1016/S0006-3207(97)00046-3

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

de Moraes, L. A., Muller, C., Bueno, R. C. O. de F., Santos, A., Bello, V. H., De Marchi, B. R., Watanabe, L. F. M., Marubayashi, J. M., Santos, B. R., & Yuki, V. A. (2018). Distribution and phylogenetics of whiteflies and their endosymbiont relationships after the Mediterranean species invasion in Brazil. *Scientific Reports*, *8*(1), 14589.

Degnan, P. H., Lazarus, A. B., Brock, C. D., & Wernegreen, J. J. (2004). Host–symbiont stability and fast evolutionary rates in an ant–bacterium association: cospeciation of Camponotus species and their endosymbionts, Candidatus Blochmannia. *Systematic Biology*, *53*(1), 95–110.

Degnan, P. H., Percy, D. M., & Hansen, A. K. (2025). Coupled evolutionary rates shape a Hawaiian insect-symbiont system. *BMC Genomics*, *26*(1), 336. https://doi.org/10.1186/s12864-025-11514-z

Deng, W., Drager, K., Li, X.-R., Han, W., Djernæs, M., Wang, Y., Martin, J., Schwartz, J., Zhang, Z., Kovacs, T., Lo, N., Hellemans, S., Che, Y.-L., Bourguignon, T., Wang, Z., & Evangelista, D. (2025). Blattodea systematics in the genomic age. *Insect Systematics and Diversity*, (in press).

Deng, Y., Xu, H., Su, Y., Liu, S., Xu, L., Guo, Z., Wu, J., Cheng, C., & Feng, J. (2019). Horizontal gene transfer contributes to virulence and antibiotic resistance of Vibrio harveyi 345 based on complete genome sequence analysis. *BMC Genomics*, *20*(1), 761. https://doi.org/10.1186/s12864-019-6137-8

Djernæs, M., Klass, K.-D., & Eggleton, P. (2015). Identifying possible sister groups of Cryptocercidae+Isoptera: A combined molecular and morphological phylogeny of Dictyoptera. *Molecular Phylogenetics and Evolution*, *84*, 284–303. https://doi.org/10.1016/j.ympev.2014.08.019

Djernæs, M., Klass, K.-D., Picker, M. D., & Damgaard, J. (2012). Phylogeny of cockroaches (Insecta, Dictyoptera, Blattodea), with placement of aberrant taxa and exploration of out-group sampling. *Systematic Entomology*, *37*(1), 65–83. https://doi.org/10.1111/j.1365-3113.2011.00598.x

Djernæs, M., & Murienne, J. (2022). Phylogeny of Blattoidea (Dictyoptera: Blattodea) with a revised classification of Blattidae. *Arthropod Systematics & Phylogeny*, *80*, 209–228. https://doi.org/10.3897/asp.80.e75819

Djernæs, M., Varadínová, Z. K., Kotyk, M., Eulitz, U., & Klass, K. D. (2020). Phylogeny and life history evolution of Blaberoidea (Blattodea). *Arthropod Systematics and Phylogeny*, *78*(1), 29–67. https://doi.org/10.26049/ASP78-1-2020-03

Doré, A. S., Drake, A. C. B., Brewerton, S. C., & Blundell, T. L. (2004). Identification of DNA-PK in the arthropods. *DNA Repair*, *3*(1), 33–41. https://doi.org/10.1016/j.dnarep.2003.09.003

Douglas, A. (1989). Mycetocyte symbiosis in insects. *Biological Reviews*, *64*(4), 409–434.

Douglas, A. E. (1998). Nutritional Interactions in Insect-Microbial Symbioses: Aphids and Their Symbiotic Bacteria *Buchnera*. *Annual Review of Entomology*, *43*(1), 17–37. https://doi.org/10.1146/annurev.ento.43.1.17

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Edwards, D. L., & Melville, J. (2010). Phylogeographic analysis detects congruent biogeographic patterns between a woodland agamid and Australian wet tropics taxa despite disparate evolutionary trajectories. *Journal of Biogeography*, *37*(8), 1543–1556.

Eickbush, T. H., & Eickbush, D. G. (2015). Integration, Regulation, and Long-Term Stability of R2 Retrotransposons. In *Mobile DNA III* (pp. 1125–1146). ASM Press. https://doi.org/10.1128/9781555819217.ch49

Eleftherianos, I., Atri, J., Accetta, J., & Castillo, J. C. (2013). Endosymbiotic bacteria in insects: guardians of the immune system? *Frontiers in Physiology*, *4*. https://doi.org/10.3389/fphys.2013.00046

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157. https://doi.org/10.1186/s13059-015-0721-2

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238. https://doi.org/10.1186/s13059-019-1832-y

Evangelista, D. A., Wipfler, B., Béthoux, O., Donath, A., Fujita, M., Kohli, M. K., Legendre, F., Liu, S., Machida, R., Misof, B., Peters, R. S., Podsiadlowski, L., Rust, J., Schuette, K., Tollenaar, W., Ware, J. L., Wappler, T., Zhou, X., Meusemann, K., & Simon, S. (2019). An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). *Proceedings of the Royal Society B: Biological Sciences*, *286*(1895), 20182076. https://doi.org/10.1098/rspb.2018.2076

Evangelista, D., Djernæs, M., & Kohli, M. (2017). Fossil calibrations for the cockroach phylogeny (Insecta, Dictyoptera, Blattodea), comments on the use of wings for their identification, and a redescription of the oldest Blaberidae. *Palaeontologia Electronica*. https://doi.org/10.26879/711

Ewart, K. M., Ho, S. Y. W., Chowdhury, A.-A., Jaya, F. R., Kinjo, Y., Bennett, J., Bourguignon, T., Rose, H. A., & Lo, N. (2024). Pervasive relaxed selection in termite genomes. *Proceedings of the Royal Society B*, *291*(2023), 20232439.

Feldhaar, H., Straka, J., Krischke, M., Berthold, K., Stoll, S., Mueller, M. J., & Gross, R. (2007). Nutritional upgrading for omnivorous carpenter ants by the endosymbiont Blochmannia. *BMC Biology*, *5*(1), 48. https://doi.org/10.1186/1741-7007-5-48

Fields, P. D., Waneka, G., Naish, M., Schatz, M. C., Henderson, I. R., & Sloan, D. B. (2022). Complete Sequence of a 641-kb Insertion of Mitochondrial DNA in the *Arabidopsis thaliana* Nuclear Genome. *Genome Biology and Evolution*, *14*(5). https://doi.org/10.1093/gbe/evac059

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G. J., Hejnol, A., Henrissat, B., Koszul, R., Aury, J.-M., Barbe, V., Barthélémy, R.-M., Bast, J., Bazykin, G. A., Chabrol, O., Couloux, A., Da Rocha, M., Da Silva, C., Gladyshev, E., … Van Doninck, K. (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. *Nature*, *500*(7463), 453–457. https://doi.org/10.1038/nature12326

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, *117*(17), 9451–9457.

Frenkel-Morgenstern, M., & Valencia, A. (2012). Novel domain combinations in proteins encoded by chimeric transcripts. *Bioinformatics*, *28*(12), i67–i74. https://doi.org/10.1093/bioinformatics/bts216

Fromont, C., Riegler, M., & Cook, J. M. (2016). Phylogeographic analyses of bacterial endosymbionts in fig homotomids (Hemiptera: Psylloidea) reveal codiversification of both primary and secondary endosymbionts. *FEMS Microbiology Ecology*, *92*(12), fiw205.

Fu, B., Chen, M., Zou, M., Long, M., & He, S. (2010). The rapid generation of chimerical genes expanding protein diversity in zebrafish. *BMC Genomics*, *11*(1), 657. https://doi.org/10.1186/1471-2164-11-657

Gaunt, M. W., & Miles, M. A. (2002). An Insect Molecular Clock Dates the Origin of the Insects and Accords with Palaeontological and Biogeographic Landmarks. *Molecular*

*Biology and Evolution*, *19*(5), 748–761. https://doi.org/10.1093/oxfordjournals.molbev.a004133
Gautam, A. (2022). *Phenol-Chloroform DNA Isolation Method* (pp. 33–39). https://doi.org/10.1007/978-3-030-94230-4_3
Gibson, C. M., & Hunter, M. S. (2010). Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecology Letters*, *13*(2), 223–234. https://doi.org/10.1111/j.1461-0248.2009.01416.x
Gilbert, C., & Maumus, F. (2023). Sidestepping Darwin: horizontal gene transfer from plants to insects. *Current Opinion in Insect Science*, *57*, 101035. https://doi.org/10.1016/j.cois.2023.101035
Good, B. H., Bhatt, A. S., & McDonald, M. J. (2025). Unraveling the tempo and mode of horizontal gene transfer in bacteria. *Trends in Microbiology*, *33*(8), 853–865. https://doi.org/10.1016/j.tim.2025.03.009
Goyal, A. (2022). Horizontal gene transfer drives the evolution of dependencies in bacteria. *IScience*, *25*(5), 104312. https://doi.org/10.1016/j.isci.2022.104312
Gruwell, M. E., Morse, G. E., & Normark, B. B. (2007). Phylogenetic congruence of armored scale insects (Hemiptera: Diaspididae) and their primary endosymbionts from the phylum Bacteroidetes. *Molecular Phylogenetics and Evolution*, *44*(1), 267–280. https://doi.org/10.1016/j.ympev.2007.01.014
Guan, D., McCarthy, S. A., Wood, J., Howe, K., Wang, Y., & Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics*, *36*(9), 2896–2898. https://doi.org/10.1093/bioinformatics/btaa025
Guo, J., Hatt, S., He, K., Chen, J., Francis, F., & Wang, Z. (2017). Nine facultative endosymbionts in aphids. A review. *Journal of Asia-Pacific Entomology*, *20*(3), 794–801. https://doi.org/10.1016/j.aspen.2017.03.025
Hall, A. A. G., Morrow, J. L., Fromont, C., Steinbauer, M. J., Taylor, G. S., Johnson, S. N., Cook, J. M., & Riegler, M. (2016). Codivergence of the primary bacterial endosymbiont of psyllids versus host switches and replacement of their secondary bacterial endosymbionts. *Environmental Microbiology*, *18*(8), 2591–2603. https://doi.org/10.1111/1462-2920.13351
Haraguchi, T., & Hiraoka, Y. (2007). Breakdown and Reformation of the Nuclear Envelope. In K. Nagata & K. Takeyasu (Eds.), *Nuclear Dynamics: Molecular Biology and Visualization of the Nucleus* (pp. 89–106). Springer Japan. https://doi.org/10.1007/978-4-431-30130-1_4
Hellemans, S., Rocha, M. M., Wang, M., Romero Arias, J., Aanen, D. K., Bagnères, A.-G., Buček, A., Carrijo, T. F., Chouvenc, T., Cuezzo, C., Constantini, J. P., Constantino, R., Dedeine, F., Deligne, J., Eggleton, P., Evans, T. A., Hanus, R., Harrison, M. C., Harry, M., … Bourguignon, T. (2024). Genomic data provide insights into the classification of extant termites. *Nature Communications*, *15*(1), 6724. https://doi.org/10.1038/s41467-024-51028-y
Herren, J. K., Paredes, J. C., Schüpfer, F., & Lemaitre, B. (2013). Vertical Transmission of a *Drosophila* Endosymbiont Via Cooption of the Yolk Transport and Internalization Machinery. *MBio*, *4*(2). https://doi.org/10.1128/mBio.00532-12
Heyworth, E. R., & Ferrari, J. (2016). Heat stress affects facultative symbiont-mediated protection from a parasitoid wasp. *PLoS One*, *11*(11), e0167180.
Hilbert, D. W., Graham, A., & Hopkins, M. S. (2007). Glacial and interglacial refugia within a long-term rainforest refugium: The Wet Tropics Bioregion of NE Queensland, Australia. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *251*(1), 104–118. https://doi.org/10.1016/j.palaeo.2007.02.020

Ho, S. Y. W., & Phillips, M. J. (2009). Accounting for Calibration Uncertainty in Phylogenetic Estimation of Evolutionary Divergence Times. *Systematic Biology*, *58*(3), 367–380. https://doi.org/10.1093/sysbio/syp035

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Hu, F.-Y., & Tsai, C.-W. (2020). Nutritional Relationship between Bemisia tabaci and Its Primary Endosymbiont, Portiera aleyrodidarum, during Host Plant Acclimation. *Insects*, *11*(8), 498. https://doi.org/10.3390/insects11080498

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., von Mering, C., & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, *47*(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

Husnik, F., & McCutcheon, J. P. (2018). Functional horizontal gene transfer from bacteria to eukaryotes. *Nature Reviews Microbiology*, *16*(2), 67–79. https://doi.org/10.1038/nrmicro.2017.137

Illiger, J. C. W. (1801). *Magazin für insektenkunde* (Vol. 1). K. Reichard.

Inward, D., Beccaloni, G., & Eggleton, P. (2007). Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches. *Biology Letters*, *3*(3), 331–335. https://doi.org/10.1098/rsbl.2007.0102

Jin, M., Zwick, A., Ślipiński, A., de Keyzer, R., & Pang, H. (2020). Museomics reveals extensive cryptic diversity of Australian prionine longhorn beetles with implications for their classification and conservation. *Systematic Entomology*, *45*(4), 745–770. https://doi.org/10.1111/syen.12424

Kaltenpoth, M., Flórez, L. V, Vigneron, A., Dirksen, P., & Engl, T. (2025). Origin and function of beneficial bacterial symbioses in insects. *Nature Reviews Microbiology*, 1–17.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kamal, S. M., Simpson, D. J., Wang, Z., Gänzle, M., & Römling, U. (2021). Horizontal Transmission of Stress Resistance Genes Shape the Ecology of Beta- and Gamma-Proteobacteria. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.696522

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kikuchi, Y., Hayatsu, M., Hosokawa, T., Nagayama, A., Tago, K., & Fukatsu, T. (2012). Symbiont-mediated insecticide resistance. *Proceedings of the National Academy of Sciences*, *109*(22), 8618–8622.

Kinjo, Y., Bourguignon, T., Tong, K. J., Kuwahara, H., Lim, S. J., Yoon, K. B., Shigenobu, S., Park, Y. C., Nalepa, C. A., Hongoh, Y., Ohkuma, M., Lo, N., & Tokuda, G. (2018). Parallel and Gradual Genome Erosion in the Blattabacterium Endosymbionts of Mastotermes darwiniensis and Cryptocercus Wood Roaches. *Genome Biology and Evolution*, *10*(6), 1622–1630. https://doi.org/10.1093/gbe/evy110

Kinjo, Y., Saitoh, S., & Tokuda, G. (2015). An Efficient Strategy Developed for Next-Generation Sequencing of Endosymbiont Genomes Performed Using Crude DNA Isolated from Host Tissues: A Case Study of <i>Blattabacterium cuenoti</i>

Inhabiting the Fat Bodies of Cockroaches. *Microbes and Environments*, *30*(3), 208–220. https://doi.org/10.1264/jsme2.ME14153
Kirby, W. F. (1903). LVII.—Notes on Blattidæ &c., with descriptions of new genera and species in the collection of the British Museum, South Kensington—No. I. *Journal of Natural History*, *11*(64), 404–415.
Kirby, W. F. (1904). *A synonymic catalogue of Orthoptera*. Printed by order of the Trustees [by Taylor and Francis]. https://doi.org/10.5962/bhl.title.6745
Klass, K.-D. (1998). The proventriculus of the Dicondylia, with comments on evolution and phylogeny in Dictyoptera and Odonata (Insecta). *Zoologischer Anzeiger*, *237*(1), 15–42.
Klass, K.-D., & Meier, R. (2006). *A phylogenetic analysis of Dictyoptera (Insecta) based on morphological characters*. https://www.researchgate.net/publication/263429967
Klasson, L., Kumar, N., Bromley, R., Sieber, K., Flowers, M., Ott, S. H., Tallon, L. J., Andersson, S. G. E., & Dunning Hotopp, J. C. (2014). Extensive duplication of the Wolbachia DNA in chromosome four of Drosophila ananassae. *BMC Genomics*, *15*(1), 1097. https://doi.org/10.1186/1471-2164-15-1097
Koga, R., Tsuchida, T., Sakurai, M., & Fukatsu, T. (2007). Selective elimination of aphid endosymbionts: effects of antibiotic dose and host genotype, and fitness consequences. *FEMS Microbiology Ecology*, *60*(2), 229–239. https://doi.org/10.1111/j.1574-6941.2007.00284.x
Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8
Kondo, N., Ijichi, N., Shimada, M., & Fukatsu, T. (2002). Prevailing triple infection with *Wolbachia* in *Callosobruchus chinensis* (Coleoptera: Bruchidae). *Molecular Ecology*, *11*(2), 167–180. https://doi.org/10.1046/j.0962-1083.2001.01432.x
Korb, J. (2008). Termites, hemimetabolous diploid white ants? *Frontiers in Zoology*, *5*(1), 15. https://doi.org/10.1186/1742-9994-5-15
Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2020). HyPhy 2.5—A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, *37*(1), 295–299. https://doi.org/10.1093/molbev/msz197
Krishna, K., Grimaldi, D. A., Krishna, V., & Engel, M. S. (2013). Treatise on the Isoptera of the World. *Bulletin of the American Museum of Natural History*, *377*(7), 1–200. https://doi.org/10.1206/377.1
Kristensen, N. P. (1975). The phylogeny of hexapod "orders". A critical review of recent accounts. *Journal of Zoological Systematics and Evolutionary Research*, *13*(1), 1–44.
Kristensen, N. P. (1981). Phylogeny of insect orders. *Annual Review of Entomology*, *26*(1), 135–157.
Krosch, M. N. (2011). Phylogeography of Echinocladius martini Cranston (Diptera: Chironomidae) in closed forest streams of eastern Australia. *Australian Journal of Entomology*, no-no. https://doi.org/10.1111/j.1440-6055.2011.00822.x
Kuntner, M., Cheng, R.-C., Kralj-Fišer, S., Liao, C.-P., Schneider, J. M., & Elgar, M. A. (2016). The evolution of genital complexity and mating rates in sexually size dimorphic spiders. *BMC Evolutionary Biology*, *16*(1), 242. https://doi.org/10.1186/s12862-016-0821-y
Kwak, Y., Argandona, J. A., Miao, S., Son, T. J., & Hansen, A. K. (2025). A dual insect symbiont and plant pathogen improves insect host fitness under arginine limitation. *MBio*, *16*(4). https://doi.org/10.1128/mbio.03588-24

Lahti, D. C., Johnson, N. A., Ajie, B. C., Otto, S. P., Hendry, A. P., Blumstein, D. T., Coss, R. G., Donohue, K., & Foster, S. A. (2009). Relaxed selection in the wild. *Trends in Ecology & Evolution*, *24*(9), 487–496. https://doi.org/10.1016/j.tree.2009.03.010

Legendre, F., D'haese, C. A., Deleporte, P., Pellens, R., Whiting, M. F., Schliep, K., & Grandcolas, P. (2013). *The evolution of social behaviour in Blaberid cockroaches with diverse habitats and social systems: phylogenetic analysis of behavioural sequences*. https://academic.oup.com/biolinnean/article/111/1/58/2415826

Legendre, F., D'Haese, C. A., Deleporte, P., Pellens, R., Whiting, M. F., Schliep, K., & Grandcolas, P. (2014). The evolution of social behaviour in Blaberid cockroaches with diverse habitats and social systems: phylogenetic analysis of behavioural sequences. *Biological Journal of the Linnean Society*, *111*(1), 58–77. https://doi.org/10.1111/bij.12199

Legendre, F., Grandcolas, P., & Thouzé, F. (2017). Molecular phylogeny of Blaberidae (Dictyoptera, Blattodea) with implications for taxonomy and evolutionary studies. *European Journal of Taxonomy*, (291). https://doi.org/10.5852/ejt.2017.291

Legendre, F., Nel, A., Svenson, G. J., Robillard, T., Pellens, R., & Grandcolas, P. (2015). Phylogeny of Dictyoptera: Dating the Origin of Cockroaches, Praying Mantises and Termites with Molecular Data and Controlled Fossil Evidence. *PLOS ONE*, *10*(7), e0130127. https://doi.org/10.1371/journal.pone.0130127

Levine, M. T., Jones, C. D., Kern, A. D., Lindfors, H. A., & Begun, D. J. (2006). Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proceedings of the National Academy of Sciences*, *103*(26), 9935–9939. https://doi.org/10.1073/pnas.0509809103

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H. (2021). New strategies to improve minimap2 alignment accuracy. *Bioinformatics*, *37*(23), 4572–4574. https://doi.org/10.1093/bioinformatics/btab705

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Li, X. (2022). Phylogeny and age of cockroaches: a reanalysis of mitogenomes with selective fossil calibrations. *Deutsche Entomologische Zeitschrift*, *69*(1), 1–18. https://doi.org/10.3897/dez.69.68373

Li, X.-D., Jiang, G.-F., Yan, L.-Y., Li, R., Mu, Y., & Deng, W.-A. (2018). Positive Selection Drove the Adaptation of Mitochondrial Genes to the Demands of Flight and High-Altitude Environments in Grasshoppers. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00605

Li, Y., Liu, Z., Liu, C., Shi, Z., Pang, L., Chen, C., Chen, Y., Pan, R., Zhou, W., Chen, X., Rokas, A., Huang, J., & Shen, X.-X. (2022). HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell*, *185*(16), 2975-2987.e10. https://doi.org/10.1016/j.cell.2022.06.014

Li, Y., Zhou, Y., Jing, W., Xu, S., Jin, Y., Xu, Y., & Wang, H. (2021). Horizontally acquired cysteine synthase genes undergo functional divergence in lepidopteran herbivores. *Heredity*, *127*(1), 21–34. https://doi.org/10.1038/s41437-021-00430-z

Liao, S., Wang, Y., Jin, D., Chen, R., Wang, Z., & Che, Y. (2021). Exploring the relationship of Homalosilpha and Mimosilpha (Blattodea, Blattidae, Blattinae) from a morphological and molecular perspective, including a description of four new species. *PeerJ*, *9*, e10618. https://doi.org/10.7717/peerj.10618

Lin, Q. B. (1980). Fossil insects from the Mesozoic of Zhejiang and Anhui provinces. *Division and Correlation of the Mesozoic Volcano-Sedimentary Strata in Zhejiang and Anhui Provinces*, 211–234.

Linder, H. P., Hardy, C. R., & Rutschmann, F. (2005). Taxon sampling effects in molecular clock dating: An example from the African Restionaceae. *Molecular Phylogenetics and Evolution*, *35*(3), 569–582. https://doi.org/10.1016/j.ympev.2004.12.006

Liu, C., Hellemans, S., Kinjo, Y., Mikhailova, A. A., Aumont, C., Weng, Y. M., Bucek, A., Husnik, F., Sobotnik, J., Harrison, M., mcmahon, D., & bourguignon, thomas. (2025). *Horizontal gene transfers are widespread across termite genomes but do not confer metabolic innovations*. https://doi.org/10.1101/2025.03.16.643567

Liu, J., Zhang, J., Han, W., Wang, Y., He, S., & Wang, Z. (2023). Advances in the understanding of Blattodea evolution: Insights from phylotranscriptomics and spermathecae. *Molecular Phylogenetics and Evolution*, *182*, 107753. https://doi.org/10.1016/j.ympev.2023.107753

Liu, L., Huang, X., Zhang, R., Jiang, L., & Qiao, G. (2013). Phylogenetic congruence between Mollitrichosiphum (Aphididae: Greenideinae) and Buchnera indicates insect–bacteria parallel evolution. *Systematic Entomology*, *38*(1), 81–92.

Lo, N., Bandi, C., Watanabe, H., Nalepa, C., & Beninati, T. (2003). Evidence for Cocladogenesis Between Diverse Dictyopteran Lineages and Their Intracellular Endosymbionts. *Molecular Biology and Evolution*, *20*(6), 907–913. https://doi.org/10.1093/molbev/msg097

Lo, N., Beninati, T., Stone, F., Walker, J., & Sacchi, L. (2007). Cockroaches that lack Blattabacterium endosymbionts: the phylogenetically divergent genus Nocticola. *Biology Letters*, *3*(3), 327–330.

Lo, N., Tokuda, G., Watanabe, H., Rose, H., Slaytor, M., Maekawa, K., Bandi, C., & Noda, H. (2000). Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches. *Current Biology*, *10*(13), 801–804.

Lo, N., Tong, K. J., Rose, H. A., Ho, S. Y. W., Beninati, T., Low, D. L. T., Matsumoto, T., & Maekawa, K. (2016). Multiple evolutionary origins of Australian soil-burrowing cockroaches driven by climate change in the Neogene. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1825), 20152869. https://doi.org/10.1098/rspb.2015.2869

López-Sánchez, M. J., Neef, A., Peretó, J., Patiño-Navarrete, R., Pignatelli, M., Latorre, A., & Moya, A. (2009). Evolutionary Convergence and Nitrogen Metabolism in Blattabacterium strain Bge, Primary Endosymbiont of the Cockroach Blattella germanica. *PLoS Genetics*, *5*(11), e1000721. https://doi.org/10.1371/journal.pgen.1000721

Luan, D. D., Korman, M. H., Jakubczak, J. L., & Eickbush, T. H. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. *Cell*, *72*(4), 595–605. https://doi.org/10.1016/0092-8674(93)90078-5

Łukasik, P., Guo, H., van Asch, M., Henry, L. M., Godfray, H. C. J., & Ferrari, J. (2015). Horizontal transfer of facultative endosymbionts is limited by host relatedness. *Evolution*, *69*(10), 2757–2766. https://doi.org/10.1111/evo.12767

Maekawa, K., Kon, M., Araya, K., & Matsumoto, T. (2001). Phylogeny and biogeography of wood-feeding cockroaches, Genus Salganea Stål (Blaberidae: Panesthiinae), in southeast Asia based on mitochondrial DNA sequences. *Journal of Molecular Evolution*, *53*(6), 651–659. https://doi.org/10.1007/s002390010252

Maekawa, K., Lo, N., Kitade, O., Miura, T., & Matsumoto, T. (1999). Molecular Phylogeny and Geographic Distribution of Wood-Feeding Cockroaches in East Asian Islands. *Molecular Phylogenetics and Evolution*, *13*(2), 360–376. https://doi.org/10.1006/mpev.1999.0647

Maekawa, K., Lo, N., Rose, H. A., & Matsumoto, T. (2003). The evolution of soil-burrowing cockroaches (Blattaria: Blaberidae) from wood-burrowing ancestors following an

invasion of the latter from Asia into Australia. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1521), 1301–1307. https://doi.org/10.1098/rspb.2003.2359
Maekawa, K., & Matsumoto, T. (2003). Biogeography of Japanese wood-feeding cockroaches genus Salganea and Panesthia (Blaberidae: Panesthiinae). *Molecular Phylogenetics and Evolution*, *27*(1), 156–159. https://doi.org/10.1016/S1055-7903(02)00399-8
Margulis, L., & Fester, R. (1991). *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. MIT press.
Martin, H. A. (1982). Changing Cenozoic barriers and the Australian paleobotanical record. *Annals of the Missouri Botanical Garden*, 625–667.
Martin, H. A. (1998). Tertiary climatic evolution and the development of aridity in Australia. *Proceedings-Linnean Society Of New South Wales*, *119*, 115–136.
Martin, H. A. (2006). Cenozoic climatic change and the development of the arid vegetation in Australia. *Journal of Arid Environments*, *66*(3), 533–563.
Matzke, N. J. (2013). *Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing*. University of California, Berkeley.
Matzke, N. J. (2014). Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. *Systematic Biology*, *63*(6), 951–970. https://doi.org/10.1093/sysbio/syu056
McCutcheon, J. P., Boyd, B. M., & Dale, C. (2019). The Life of an Insect Endosymbiont from the Cradle to the Grave. *Current Biology*, *29*(11), R485–R495. https://doi.org/10.1016/j.cub.2019.03.032
McKittrick, F. A. (1964). Evolutionary studies of cockroaches. *Memoirs of the Cornell University Agricultural Experiment Station*, *389*, 1.
McLysaght, A., & Guerzoni, D. (2015). New genes from non-coding sequence: the role of de novo protein-coding genes in eukaryotic evolutionary innovation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1678), 20140332. https://doi.org/10.1098/rstb.2014.0332
Méheust, R., Bhattacharya, D., Pathmanathan, J. S., McInerney, J. O., Lopez, P., & Bapteste, E. (2018). Formation of chimeric genes with essential functions at the origin of eukaryotes. *BMC Biology*, *16*(1), 30. https://doi.org/10.1186/s12915-018-0500-0
Mei, Y., Jing, D., Tang, S., Chen, X., Chen, H., Duanmu, H., Cong, Y., Chen, M., Ye, X., & Zhou, H. (2022). InsectBase 2.0: a comprehensive gene resource for insects. *Nucleic Acids Research*, *50*(D1), D1040–D1045.
Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015
Misof, B., Liu, S., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., Frandsen, P. B., Ware, J., Flouri, T., Beutel, R. G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A. J., Aspöck, U., Aspöck, H., Bartel, D., … Zhou, X. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science*, *346*(6210), 763–767. https://doi.org/doi:10.1126/science.1257570
Mitter, C., Farrell, B., & Wiegmann, B. (1988). The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *The American Naturalist*, *132*(1), 107–128.
Monteith, G. B. (1997). *Revision of the Australian flat bugs of the subfamily Mezirinae (Insecta: Hemiptera: Aradidae).*

Moran, N. A., & Jarvik, T. (2010). Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids. *Science*, *328*(5978), 624–627. https://doi.org/10.1126/science.1187113

Moran, N. A., Tran, P., & Gerardo, N. M. (2005). Symbiosis and Insect Diversification: an Ancient Symbiont of Sap-Feeding Insects from the Bacterial Phylum *Bacteroidetes*. *Applied and Environmental Microbiology*, *71*(12), 8802–8810. https://doi.org/10.1128/AEM.71.12.8802-8810.2005

Moreau, C. S., Hugall, A. F., McDonald, K. R., Jamieson, B. G. M., & Moritz, C. (2015). An Ancient Divide in a Contiguous Rainforest: Endemic Earthworms in the Australian Wet Tropics. *PLOS ONE*, *10*(9), e0136943. https://doi.org/10.1371/journal.pone.0136943

Morgulis, A., Gertz, E. M., Schäffer, A. A., & Agarwala, R. (2006). A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *Journal of Computational Biology*, *13*(5), 1028–1040.

Mullins, D. E. (2015). Physiology of environmental adaptations and resource acquisition in cockroaches. *Annual Review of Entomology*, *60*(1), 473–492.

Munson, M. A., Baumann, P., Clark, M. A., Baumann, L., Moran, N. A., Voegtlin, D. J., & Campbell, B. C. (1991). Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *Journal of Bacteriology*, *173*(20), 6321–6324.

Nakabachi, A. (2015). Horizontal gene transfers in insects. *Current Opinion in Insect Science*, *7*, 24–29. https://doi.org/10.1016/j.cois.2015.03.006

Nalepa, C. A., & Lenz, M. (2000). The ootheca of *Mastotermes darwiniensis* Froggatt (Isoptera: Mastotermitidae): homology with cockroach oothecae. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267*(1454), 1809–1813. https://doi.org/10.1098/rspb.2000.1214

Nelson, J. O., Slicko, A., & Yamashita, Y. M. (2023). The retrotransposon R2 maintains *Drosophila* ribosomal DNA repeats. *Proceedings of the National Academy of Sciences*, *120*(23). https://doi.org/10.1073/pnas.2221613120

Nikoh, N., Hosokawa, T., Moriyama, M., Oshima, K., Hattori, M., & Fukatsu, T. (2014). Evolutionary origin of insect– *Wolbachia* nutritional mutualism. *Proceedings of the National Academy of Sciences*, *111*(28), 10257–10262. https://doi.org/10.1073/pnas.1409284111

Nix, H. A., & Switzer, M. A. (1991). *Rainforest animals: atlas of vertebrates endemic to Australia's wet tropics*. Australian National Parks and Wildlife Service.

Noda, T., Harumoto, T., Katsuno, T., Moriyama, M., & Fukatsu, T. (2025). Cockroach bacteriocytes migrate into the ovaries for vertical transmission of the bacterial endosymbiont Blattabacterium. *BioRxiv*, 2025–2026.

Noda, T., Mizutani, M., Harumoto, T., Katsuno, T., Koga, R., & Fukatsu, T. (2024). Frequent and asymmetric cell division in endosymbiotic bacteria of cockroaches. *Applied and Environmental Microbiology*, *90*(10). https://doi.org/10.1128/aem.01466-24

Noll, D., Leon, F., Brandt, D., Pistorius, P., Le Bohec, C., Bonadonna, F., Trathan, P. N., Barbosa, A., Rey, A. R., Dantas, G. P. M., Bowie, R. C. K., Poulin, E., & Vianna, J. A. (2022). Positive selection over the mitochondrial genome and its role in the diversification of gentoo penguins in response to adaptation in isolation. *Scientific Reports*, *12*(1), 3767. https://doi.org/10.1038/s41598-022-07562-0

Nováková, E., Hypša, V., Klein, J., Foottit, R. G., von Dohlen, C. D., & Moran, N. A. (2013). Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont Buchnera aphidicola. *Molecular Phylogenetics and Evolution*, *68*(1), 42–54.

Oberski, J. T., Sharma, P. P., Jay, K. R., Coblens, M. J., Lemon, K. A., Johnson, J. E., & Boyer, S. L. (2018). A dated molecular phylogeny of mite harvestmen (Arachnida: Opiliones: Cyphophthalmi) elucidates ancient diversification dynamics in the Australian Wet Tropics. *Molecular Phylogenetics and Evolution*, *127*, 813–822. https://doi.org/10.1016/j.ympev.2018.06.029

Osaki, H., & Kasuya, E. (2021). Mutual wing-eating between female and male within mating pairs in wood-feeding cockroach. *Ethology*, *127*(5), 433–437. https://doi.org/10.1111/eth.13133

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the Insect Mitochondrial Molecular Clock: The Mid-Aegean Trench Calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. https://doi.org/10.1093/molbev/msq051

Paquola, A. C. M., Asif, H., Pereira, C. A. de B., Feltes, B. C., Bonatto, D., Lima, W. C., & Menck, C. F. M. (2018). Horizontal Gene Transfer Building Prokaryote Genomes: Genes Related to Exchange Between Cell and Environment are Frequently Transferred. *Journal of Molecular Evolution*, *86*(3–4), 190–203. https://doi.org/10.1007/s00239-018-9836-x

Park, M. S., Park, P., & Takeda, M. (2013). Roles of fat body trophocytes, mycetocytes and urocytes in the American cockroach, Periplaneta americana under starvation conditions: An ultrastructural study. *Arthropod Structure & Development*, *42*(4), 287–295. https://doi.org/10.1016/j.asd.2013.03.004

Patiño-Navarrete, R., Moya, A., Latorre, A., & Peretó, J. (2013). Comparative Genomics of Blattabacterium cuenoti: The Frozen Legacy of an Ancient Endosymbiont Genome. *Genome Biology and Evolution*, *5*(2), 351–361. https://doi.org/10.1093/gbe/evt011

Peccoud, J., Simon, J.-C., McLaughlin, H. J., & Moran, N. A. (2009). Post-Pleistocene radiation of the pea aphid complex revealed by rapidly evolving endosymbionts. *Proceedings of the National Academy of Sciences*, *106*(38), 16315–16320. https://doi.org/10.1073/pnas.0905129106

Pierneef, R., Clasen, F. J., Bezuidt, O. K. I., & Reva, O. N. (2018). Long Non-coding RNAs in the Human Genome Acquired by Horizontal Gene Transfer. *The Open Bioinformatics Journal*, *11*(1), 182–196. https://doi.org/10.2174/1875036201811010182

Piton, L. E. (1940). Paléontologie du Gisement Éocène de Menat (Puy-de-Dôme) (Flore et Faune). *Imprimeries P. Vallier France*.

Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, *21*(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Potter, S., Cooper, S. J. B., Metcalfe, C. J., Taggart, D. A., & Eldridge, M. D. B. (2012). Phylogenetic relationships of rock-wallabies, Petrogale (Marsupialia: Macropodidae) and their biogeographic history within Australia. *Molecular Phylogenetics and Evolution*, *62*(2), 640–652. https://doi.org/10.1016/j.ympev.2011.11.005

Poux, C., Madsen, O., Glos, J., de Jong, W. W., & Vences, M. (2008). Molecular phylogeny and divergence times of Malagasy tenrecs: Influence of data partitioning and taxon sampling on dating analyses. *BMC Evolutionary Biology*, *8*(1), 102. https://doi.org/10.1186/1471-2148-8-102

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904.

Rebrina, F., Alegro, A., Hristov, G., Ternjej, I., & Brigić, A. (2020). Open karst habitats promote the diversity of ground-dwelling orthopterans and cockroaches (Insecta:

Orthoptera, Blattodea) along a temporary river. *Journal of Insect Conservation*, *24*(6), 1017–1030. https://doi.org/10.1007/s10841-020-00271-x
Ree, R. H., & Sanmartín, I. (2018). Conceptual and statistical problems with the <scp>DEC</scp> +J model of founder-event speciation and its comparison with <scp>DEC</scp> via model selection. *Journal of Biogeography*, *45*(4), 741–749. https://doi.org/10.1111/jbi.13173
Ree, R. H., & Smith, S. A. (2008). Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. *Systematic Biology*, *57*(1), 4–14. https://doi.org/10.1080/10635150701883881
Revell, L. J. (2024). phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, *12*, e16505. https://doi.org/10.7717/peerj.16505
Rix, M. G., & Harvey, M. S. (2012). Phylogeny and historical biogeography of ancient assassin spiders (Araneae: Archaeidae) in the Australian mesic zone: Evidence for Miocene speciation within Tertiary refugia. *Molecular Phylogenetics and Evolution*, *62*(1), 375–396. https://doi.org/10.1016/j.ympev.2011.10.009
Ronquist, F. (1997). Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology*, *46*(1), 195–203. https://doi.org/10.1093/sysbio/46.1.195
Roth, L. M. (1977). A taxonomic revision of the Panesthiinae of the world. I. The Panesthiinae of Australia (Dictyoptera : Blattaria : Blaberidae). *Australian Journal of Zoology Supplementary Series*, *25*(48), 1. https://doi.org/10.1071/AJZS048
Roth, L. M. (1979a). A taxonomic revision of the Panesthiinae of the world. II.* The genera Salgana Stål Microdina Kirby and Caeparia Stål (Dictyoptera: Blattaria: Blaberidae). *Australian Journal of Zoology Supplementary Series*, *27*(69), 1–201.
Roth, L. M. (1979b). A Taxonomic revision of the Panesthiinae of the world. III.* The genera Panesthia Serville and Miopanesthia Saussure (Dictyoptera: Blattaria: Blaberidae). *Australian Journal of Zoology Supplementary Series*, *27*(74), 1–276.
Roth, L. M. (1982). A taxonomic revision of the panesthiinae of the world. IV.* The genus *Ancaudellia* Shaw, with additions to parts 1–3, and a general discussion of distribution and relationships of the components of the subfamily (Dictyopter : Blattaria : Blaberidae). *Australian Journal of Zoology Supplementary Series*, *30*(82), 1. https://doi.org/10.1071/AJZS082
Roth, L. M., & Willis, E. R. (1955). Intra-uterine nutrition of the “beetle-roach” Diploptera dytiscoides (Serv.) during embryogenesis, with notes on its biology in the laboratory (Blattaria: Diplopteridae). *Psyche: A Journal of Entomology*, *62*(2), 55–68.
Roth, L. M., & Willis, E. R. (1957). An analysis of oviparity and viviparity in the Blattaria. *Transactions of the American Entomological Society (1890-)*, *83*(4), 221–238.
Rowe, K. C., Achmadi, A. S., Fabre, P., Schenk, J. J., Steppan, S. J., & Esselstyn, J. A. (2019). Oceanic islands of Wallacea as a source for dispersal and diversification of murine rodents. *Journal of Biogeography*, *46*(12), 2752–2768. https://doi.org/10.1111/jbi.13720
Rowe, L., & Arnqvist, G. (2012). SEXUAL SELECTION AND THE EVOLUTION OF GENITAL SHAPE AND COMPLEXITY IN WATER STRIDERS. *Evolution*, *66*(1), 40–54. https://doi.org/10.1111/j.1558-5646.2011.01411.x
Rugg, D., & Rose, H. A. (1984a). Reproductive biology of some Australian cockroaches (Blattodea: Blaberidae). *Australian Journal of Entomology*, *23*(2), 113–117. https://doi.org/10.1111/j.1440-6055.1984.tb01922.x
Rugg, D., & Rose, H. A. (1984b). THE TAXONOMIC SIGNIFICANCE OF REPRODUCTIVE BEHAVIOUR IN SOME AUSTRALIAN COCKROACHES

(BLATTODEA: BLABERIDAE). *Australian Journal of Entomology*, *23*(2), 118–118. https://doi.org/10.1111/j.1440-6055.1984.tb01923.x
Rugg, D., & Rose, H. A. (1991). Biology of Macropanesthia rhinoceros Saussure (Dictyoptera: Blaberidae). *Annals of the Entomological Society of America*, *84*(6), 575–582. https://doi.org/10.1093/aesa/84.6.575
Sabater-Muñoz, B., Toft, C., Alvarez-Ponce, D., & Fares, M. A. (2017). Chance and necessity in the genome evolution of endosymbiotic bacteria of insects. *The ISME Journal*, *11*(6), 1291–1304.
Sabree, Z. L., Huang, C. Y., Arakawa, G., Tokuda, G., Lo, N., Watanabe, H., & Moran, N. A. (2012). Genome Shrinkage and Loss of Nutrient-Providing Potential in the Obligate Symbiont of the Primitive Termite Mastotermes darwiniensis. *Applied and Environmental Microbiology*, *78*(1), 204–210. https://doi.org/10.1128/AEM.06540-11
Sabree, Z. L., Kambhampati, S., & Moran, N. A. (2009). Nitrogen recycling and nutritional provisioning by *Blattabacterium* , the cockroach endosymbiont. *Proceedings of the National Academy of Sciences*, *106*(46), 19521–19526. https://doi.org/10.1073/pnas.0907504106
Salamov, A. A., & Solovyev, V. V. (2000). Ab initio Gene Finding in *Drosophila* Genomic DNA. *Genome Research*, *10*(4), 516–522. https://doi.org/10.1101/gr.10.4.516
Schmitz, J., & Brosius, J. (2011). Exonization of transposed elements: A challenge and opportunity for evolution. *Biochimie*, *93*(11), 1928–1934. https://doi.org/10.1016/j.biochi.2011.07.014
Schneider, C. J., Cunningham, M., & Moritz, C. (1998). Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Molecular Ecology*, *7*(4), 487–498. https://doi.org/10.1046/j.1365-294x.1998.00334.x
Schneider, C., & Moritz, C. (1999). Rainforest refugia and Australia's Wet Tropics. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1415), 191–196. https://doi.org/10.1098/rspb.1999.0621
Schönknecht, G., Chen, W.-H., Ternes, C. M., Barbier, G. G., Shrestha, R. P., Stanke, M., Bräutigam, A., Baker, B. J., Banfield, J. F., Garavito, R. M., Carr, K., Wilkerson, C., Rensing, S. A., Gagneul, D., Dickenson, N. E., Oesterhelt, C., Lercher, M. J., & Weber, A. P. M. (2013). Gene Transfer from Bacteria and Archaea Facilitated Evolution of an Extremophilic Eukaryote. *Science*, *339*(6124), 1207–1210. https://doi.org/10.1126/science.1231707
Sela, I., Wolf, Y. I., & Koonin, E. V. (2016). Theory of prokaryotic genome evolution. *Proceedings of the National Academy of Sciences*, *113*(41), 11399–11407. https://doi.org/10.1073/pnas.1614083113
Sela, N., Mersch, B., Gal-Mark, N., Lev-Maor, G., Hotz-Wagenblatt, A., & Ast, G. (2007). Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome. *Genome Biology*, *8*(6), R127. https://doi.org/10.1186/gb-2007-8-6-r127
Serbus, L. R., Casper-Lindley, C., Landmann, F., & Sullivan, W. (2008). The Genetics and Cell Biology of *Wolbachia* -Host Interactions. *Annual Review of Genetics*, *42*(1), 683–707. https://doi.org/10.1146/annurev.genet.41.110306.130354
Sergeev, M. G. (1984). A new genus of cockroaches of the subfamily Panesthiinae (Blattariae, Blaberidae) from Vietnam. *Zoologicheskii Zhurnal*, *63*, 1895–1899.
Shan, H., Liu, Y., Luan, J., & Liu, S. (2021). New insights into the transovarial transmission of the symbiont Rickettsia in whiteflies. *Science China Life Sciences*, *64*(7), 1174–1186. https://doi.org/10.1007/s11427-020-1801-7
Shoeb, E., Badar, U., Akhter, J., Shams, H., Sultana, M., & Ansari, M. A. (2012). Horizontal gene transfer of stress resistance genes through plasmid transport. *World Journal of*

*Microbiology and Biotechnology*, *28*(3), 1021–1025. https://doi.org/10.1007/s11274-011-0900-6
Simmons, L. W. (2014). Sexual selection and genital evolution. *Austral Entomology*, *53*(1), 1–17. https://doi.org/10.1111/aen.12053
Smit, A., Hubley, R., & Green P. (2015). *RepeatMasker Open-4.0*. Http://Www.Repeatmasker.Org.
Soares, A. E. R., & Schrago, C. G. (2015). The influence of taxon sampling on Bayesian divergence time inference under scenarios of rate heterogeneity among lineages. *Journal of Theoretical Biology*, *364*, 31–39. https://doi.org/10.1016/j.jtbi.2014.09.004
Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology*, *7*(S1), S10. https://doi.org/10.1186/gb-2006-7-s1-s10
Song, H., & Bucheli, S. R. (2010). Comparison of phylogenetic signal between male genitalia and non-genital characters in insect systematics. *Cladistics*, *26*(1), 23–35. https://doi.org/10.1111/j.1096-0031.2009.00273.x
Spaulding, A. W., & von Dohlen, C. D. (1998). Phylogenetic Characterization and Molecular Evolution of Bacterial Endosymbionts in Psyllids (Hemiptera: Sternorrhyncha). *Molecular Biology and Evolution*, *15*(11), 1506–1513. https://doi.org/10.1093/oxfordjournals.molbev.a025878
Stork, N. E. (2018). How many species of insects and other terrestrial arthropods are there on Earth? *Annual Review of Entomology*, *63*(2018), 31–45.
Sutherland, F. L. (2011). Diversity within geodiversity, underpinning habitats in New South Wales volcanic areas. *Proceedings of the Linnean Society of New South Wales*, *132*, 37–54.
Tänzler, R., Toussaint, E. F. A., Suhardjono, Y. R., Balke, M., & Riedel, A. (2014). Multiple transgressions of Wallace's Line explain diversity of flightless *Trigonopterus* weevils on Bali. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782), 20132528. https://doi.org/10.1098/rspb.2013.2528
Terry, M. D., & Whiting, M. F. (2005). Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics*, *21*(3), 240–257. https://doi.org/10.1111/j.1096-0031.2005.00062.x
Thao, M. L., & Baumann, P. (2004). Evolutionary relationships of primary prokaryotic endosymbionts of whiteflies and their hosts. *Applied and Environmental Microbiology*, *70*(6), 3401–3406.
Thomas, C. M., & Nielsen, K. M. (2005). Mechanisms of, and Barriers to, Horizontal Gene Transfer between Bacteria. *Nature Reviews Microbiology*, *3*(9), 711–721. https://doi.org/10.1038/nrmicro1234
Tokuda, G., Lo, N., Takase, A., Yamada, A., Hayashi, Y., & Watanabe, H. (2008). Purification and partial genome characterization of the bacterial endosymbiont Blattabacterium cuenoti from the fat bodies of cockroaches. *BMC Research Notes*, *1*(1), 118. https://doi.org/10.1186/1756-0500-1-118
Ullastres, A., Merenciano, M., & González, J. (2021). Regulatory regions in natural transposable element insertions drive interindividual differences in response to immune challenges in Drosophila. *Genome Biology*, *22*(1), 265. https://doi.org/10.1186/s13059-021-02471-3
Urban, J. M., & Cryan, J. R. (2012). Two ancient bacterial endosymbionts have coevolved with the planthoppers (Insecta: Hemiptera: Fulgoroidea). *BMC Evolutionary Biology*, *12*(1), 87.

Van Etten, J., & Bhattacharya, D. (2020). Horizontal Gene Transfer in Eukaryotes: Not if, but How Much? *Trends in Genetics*, *36*(12), 915–925. https://doi.org/10.1016/j.tig.2020.08.006

Wang, X., Wang, Z., & Che, Y. (2014). A taxonomic study of the genus Panesthia (Blattodea, Blaberidae, Panesthiinae) from China with descriptions of one new species, one new subspecies and the male of Panesthia antennata. *ZooKeys*, *466*(466), 53–75. https://doi.org/10.3897/zookeys.466.8111

Wang, Z., Shi, Y., Qiu, Z., Che, Y., & Lo, N. (2017). Reconstructing the phylogeny of Blattodea: robust support for interfamilial relationships and major clades. *Scientific Reports*, *7*(1), 3903. https://doi.org/10.1038/s41598-017-04243-1

Ware, J. L., Litman, J., Klass, K. D., & Spearman, L. A. (2008). Relationships among the major lineages of Dictyoptera: The effect of outgroup selection on dictyopteran tree topology. *Systematic Entomology*, *33*(3), 429–450. https://doi.org/10.1111/j.1365-3113.2008.00424.x

Wernegreen, J. J. (2002). Genome evolution in bacterial endosymbionts of insects. *Nature Reviews Genetics*, *3*(11), 850–861.

Wernegreen, Jennifer J. (2017). Ancient bacterial endosymbionts of insects: Genomes as sources of insight and springboards for inquiry. *Experimental Cell Research*, *358*(2), 427–432. https://doi.org/10.1016/j.yexcr.2017.04.028

Wernegreen, Jennifer J. (2017). In it for the long haul: evolutionary consequences of persistent endosymbiosis. *Current Opinion in Genetics & Development*, *47*, 83–90. https://doi.org/10.1016/j.gde.2017.08.006

Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L., & Scheffler, K. (2015). RELAX: Detecting Relaxed Selection in a Phylogenetic Framework. *Molecular Biology and Evolution*, *32*(3), 820–832. https://doi.org/10.1093/molbev/msu400

Whiting, M. F., Bradler, S., & Maxwell, T. (2003). Loss and recovery of wings in stick insects. *Nature*, *421*(6920), 264–267. https://doi.org/10.1038/nature01313

Wipfler, B., Letsch, H., Frandsen, P. B., Kapli, P., Mayer, C., Bartel, D., Buckley, T. R., Donath, A., Edgerly-Rooks, J. S., Fujita, M., Liu, S., Machida, R., Mashimo, Y., Misof, B., Niehuis, O., Peters, R. S., Petersen, M., Podsiadlowski, L., Schütte, K., … Simon, S. (2019). Evolutionary history of Polyneoptera and its implications for our understanding of early winged insects. *Proceedings of the National Academy of Sciences*, *116*(8), 3024–3029. https://doi.org/10.1073/pnas.1817794116

Woolfit, M., Algama, M., Keith, J. M., McGraw, E. A., & Popovici, J. (2015). Discovery of Putative Small Non-Coding RNAs from the Obligate Intracellular Bacterium Wolbachia pipientis. *PLOS ONE*, *10*(3), e0118595. https://doi.org/10.1371/journal.pone.0118595

Xia, J., Guo, Z., Yang, Z., Han, H., Wang, S., Xu, H., Yang, X., Yang, F., Wu, Q., Xie, W., Zhou, X., Dermauw, W., Turlings, T. C. J., & Zhang, Y. (2021). Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell*, *184*(7), 1693-1705.e17. https://doi.org/10.1016/j.cell.2021.02.014

Xing, B., Yang, L., Gulinuer, A., & Ye, G. (2023). Research progress on horizontal gene transfer and its functions in insects. *Tropical Plants*, *2*(1), 1–12. https://doi.org/10.48130/TP-2023-0003

Yadav, A., Suhag, A., Jaiwal, R., Chaudhary, D., & Jaiwal, P. K. (2024). Current progress and challenges of horizontal gene transfers in whiteflies (Bemisia tabaci) for their sustainable management. *Journal of Asia-Pacific Entomology*, *27*(1), 102216. https://doi.org/10.1016/j.aspen.2024.102216

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: A Journal of Integrative Biology*, *16*(5), 284–287.

Zera, A. J., & Mole, S. (1994). The physiological costs of flight capability in wing-dimorphic crickets. *Population Ecology*, *36*(2), 151–156. https://doi.org/10.1007/BF02514930

Zhang, C., Wang, J., Marowsky, N. C., Long, M., Wing, R. A., & Fan, C. (2013). High Occurrence of Functional New Chimeric Genes in Survey of Rice Chromosome 3 Short Arm Genome Sequences. *Genome Biology and Evolution*, *5*(5), 1038–1048. https://doi.org/10.1093/gbe/evt071

Zwick, P., & Zwick, A. (2023). Revision of the African Neoperla Needham, 1905 (Plecoptera: Perlidae: Perlinae) based on morphological and molecular data. *Zootaxa*, *5316*(1), 1–194.

# Appendix

## Supplementary material for Chapter 2

**Supplementary Table S2.1**. List of newly collected and samples from previous studies used in this research. Sequences generated presently will be uploaded to GenBank, and the table entries will be updated, upon acceptance of the manuscript. Dashes (-) indicate missing sequences or geographic data, or 12S and COII sequences already included in the Mitogenome. Australian regional abbreviations: Queensland (QLD), West Australia (WA), South Australia, (SA), Victoria (VIC). Location abbreviations: Mountain (Mt), National Park (NP). All ingroup samples not placed in *Panesthia* belong to Geoscapheinae. Samples used as outgroups were shown in grey. Identifications of specimens to species level were performed by H.A. Rose and J.A. Walker.

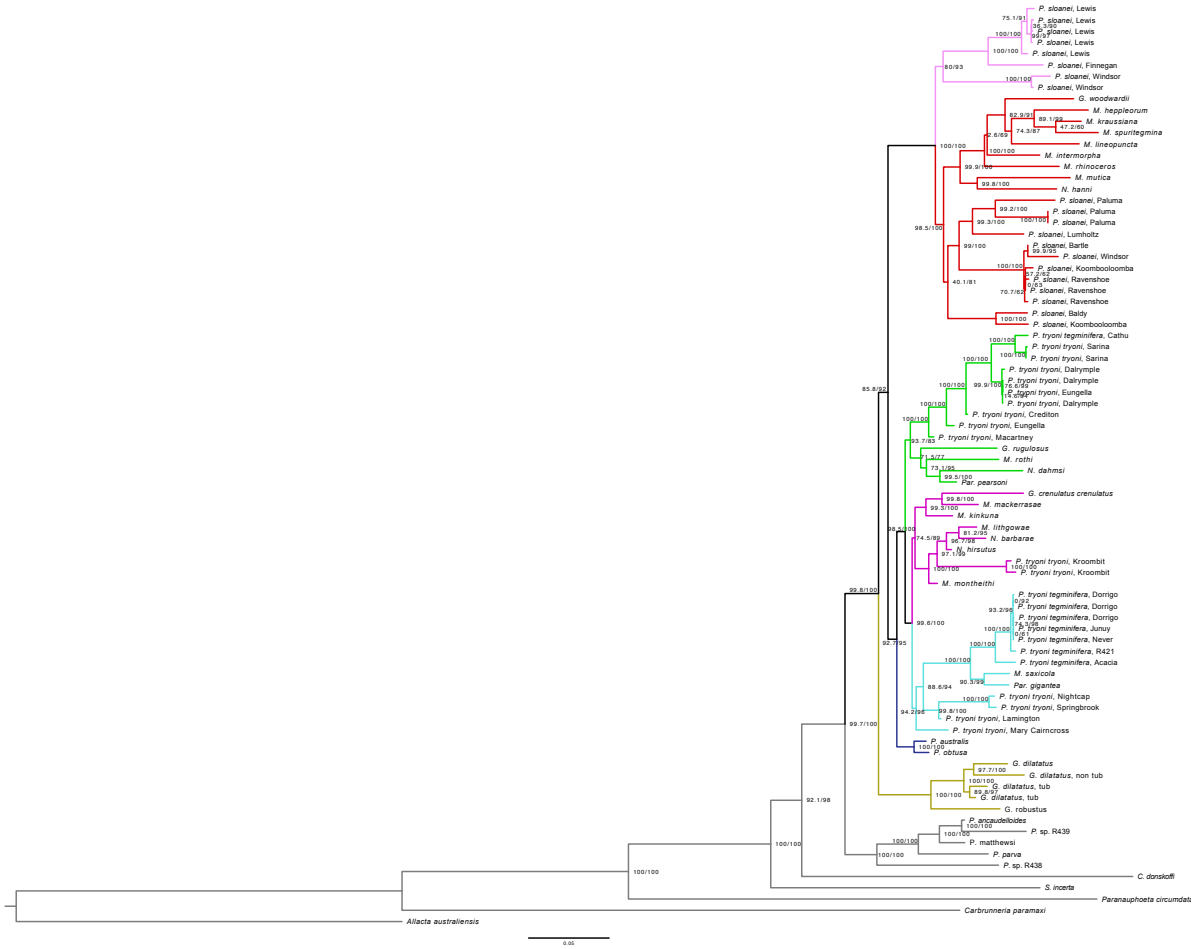
| Genus | Species | Sequence Source | Latitude (° S) | Longitude (° E) | Sample ID | Mitogenome | *12S* | *COII* | *18S* | *ITS1* | *28S* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Geoscapheus* | *crenulatus crenulatus* | Beasley-Hall et al. (2021) | 25.945 | 153.091 | Geocren | MW996579 | - | - | MW365869 | MW365805 | Assembly |
| *Geoscapheus* | *dilatatus* | Beasley-Hall et al. (2021) | ~25.794 | ~146.585 | Gdila | MW600987 | - | - | KU577825 | MW365806 | - |
| *Geoscapheus* | *dilatatus* non tub | Beasley-Hall et al. (2021) | 26.480 | 147.84 | Gdilantu | MW354074 | - | - | MW365870 | MW365806 | Assembly |
| *Geoscapheus* | *dilatatus* tub | Beasley-Hall et al. (2021) | 34.756 | 142.339 | Gdilatub | MW600999 | - | - | KU577824 | KU577871 | - |
| *Geoscapheus* | *dilatatus* tub | Beasley-Hall et al. (2021) | 35.339 | 142.188 | Gdilatub2 | MW354075 | - | - | MW365871 | MW365807 | Assembly |
| *Geoscapheus* | *robustus* | Beasley-Hall et al. (2021) | 34.767 | 142.332 | Georobu | MW996606 | - | - | KU577821 | KU577868 | - |
| *Geoscapheus* | *rugulosus* | Beasley-Hall et al. (2021) | 23.768 | 149.094 | Grugulo | MW996580 | - | - | MW365873 | MW365809 | - |
| *Geoscapheus* | *woodwardii* | Beasley-Hall et al. (2021) | 22.535 | 144.586 | Gwood | MW996581 | - | - | MW365874 | MW365810 | - |
| *Macropanesthia* | *heppleorum* | Beasley-Hall et al. (2021) | 17.620 | 145.297 | Mhepp | MW996582 | - | - | MW365875 | KU577857 | - |
| *Macropanesthia* | *intermorpha* | Lo et al. (2016) | 19.948 | 144.275 | Minter | - | KU577783 | KU577689 | KU577836 | - | - |
| *Macropanesthia* | *kinkuna* | Beasley-Hall et al. (2021) | 24.966 | 152.475 | Mkink | MW996583 | - | - | MW365876 | MW365812 | Assembly |
| *Macropanesthia* | *kraussiana* | Beasley-Hall et al. (2021) | 24.299 | 144.38 | Mkrau | MW996584 | - | - | MW365877 | MW365813 | - |
| *Macropanesthia* | *lineopunctata* | Beasley-Hall et al. (2021) | 23.557 | 145.719 | Mline | - | - | KU577686 | KU577837 | - | - |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Macropanesthia* | *lithgowae* | Beasley-Hall et al. (2021) | 26.155 | 151.911 | Mlith | MW354066 | - | - | KU577829 | KU577875 | - |
| *Macropanesthia* | *mackerrasae* | Beasley-Hall et al. (2021) | 25.439 | 152.100 | Mmack | MW996585 | - | - | MW365879 | MW365815 | Assembly |
| *Macropanesthia* | *montheithi* | Beasley-Hall et al. (2021) | 26.711 | 151.772 | Mmont | MW996586 | - | - | MW365880 | MW365816 | - |
| *Macropanesthia* | *mutica* | Beasley-Hall et al. (2021) | 16.952 | 145.598 | Mcpamuti | MW354067 | - | - | MW365881 | MW365817 | Assembly |
| *Macropanesthia* | *rhinoceros* | Lo *et al*. (2016) | 23.722 | 146.585 | Mrhin | MG882202 | - | - | KU577805 | KU577858 | DQ874230 |
| *Macropanesthia* | *rothi* | Beasley-Hall et al. (2021) | 24.261 | 151.944 | Mroth | MW354068 | - | - | MW365882 | MW365818 | - |
| *Macropanesthia* | *saxicola* | Lo et al. (2016) | 28.988 | 150.773 | Msaxi | - | KU577779 | KU577683 | KU577834 | KU577882 | - |
| *Macropanesthia* | *spuritegmina* | Lo et al. (2016) | 26.238 | 151.713 | Mspur | - | KU577744 | - | - | - | - |
| *Neogeoscapheus* | *barbarae* | Lo et al. (2016) | 26.070 | 151.700 | Nbarb | - | KU577772 | KU577676 | - | KU577878 | - |
| *Neogeoscapheus* | *dahmsi* | Beasley-Hall et al. (2021) | 25.493 | 149.253 | Ndahm | MW996587 | - | - | MW365883 | MW365819 | - |
| *Neogeoscapheus* | *hanni* | Beasley-Hall et al. (2021) | 16.665 | 145.299 | Nhann | MW996588 | - | - | MW365884 | MW365820 | Assembly |
| *Neogeoscapheus* | *hirsutus* | Beasley-Hall et al. (2021) | 24.810 | 152.330 | Nhirs | MW996589 | - | - | MW365885 | MW365821 | - |
| *Panesthia* | *australis* | Beasley-Hall et al. (2021) | 37.563 | 140.769 | Paust | MW996591 | - | - | MW365888 | MW365824 | - |
| *Panesthia* | *obtusa* | Beasley-Hall et al. (2021) | 26.020 | 148.271 | Pobtus | MW996592 | - | - | MW365890 | MW365826 | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | ~15.818 | ~145.284 | PsloFinn | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | Beasley-Hall et al. (2021) | 16.233 | 145.007 | PsloaWinB22 | MW996594 | - | - | MW365892 | - | - |
| *Panesthia* | *sloanei* | newly sequenced | 16.261 | 145.042 | PsloWinR1319 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 16.590 | 145.290 | PsloLewB26 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 16.590 | 145.290 | PsloLewR1318 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 16.590 | 145.290 | PsloLewR418 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 16.590 | 145.290 | PsloLewR4210 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | Beasley-Hall et al. (2021) | 16.590 | 145.290 | PanesloB25 | MW996596 | - | - | MW365894 | MW365829 | - |
| *Panesthia* | *sloanei* | newly sequenced | ~17.277 | ~145.450 | PsloBal | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 17.402 | 145.82 | PsloBar | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 17.525 | 145.511 | PsloRavR4212 | Assembly | - | - | Assembly | Assembly | Assembly |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panesthia* | *sloanei* | newly sequenced | 17.525 | 145.511 | PsloRavB23 | Assembly | - | - | - | - | - |
| *Panesthia* | *sloanei* | Beasley-Hall et al. (2021) | 17.525 | 145.511 | PsloaPalB24 | MW996597 | - | - | MW365895 | - | - |
| *Panesthia* | *sloanei* | newly sequenced | 17.830 | 145.604 | PsloKoo | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 17.830 | 145.604 | PsloJ122 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 18.212 | 145.799 | PsloLum | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *sloanei* | newly sequenced | 19.010 | 146.200 | PsloPalR417 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 28.364 | 152.394 | PtrteAca | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 30.296 | 152.746 | PtrteJun | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 30.355 | 152.802 | PtrteNev | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 30.358 | 152.731 | PtrteDoJ103 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 30.373 | 152.72 | PtrteDoB27 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | newly sequenced | 30.373 | 152.72 | PtrteDoB28 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tegminifera* | Beasley-Hall et al. (2021) | 30.373 | 152.725 | PtrteR421 | MW996599 | - | - | MW365897 | MW365831 | - |
| *Panesthia* | *tryoni tryoni* | newly sequenced | ~20.812 | ~148.567 | PtrteCat | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 20.834 | 148.553 | PtrtrB215 | MW996602 | - | - | MW365900 | MW365834 | - |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.035 | 148.597 | PtrtrDalJ100 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.05 | 148.55 | PtrtrDalR1315 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.05 | 148.55 | PtrtrDalR434 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.057 | 148.581 | PtrtrEunJ101 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 21.118 | 148.511 | PtrtrEunB214 | MW996601 | - | - | MW365899 | MW365833 | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.185 | 148.525 | PtrtrCre | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 21.605 | 148.972 | PtrtrSarR1314 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 21.605 | 148.972 | PtrtrSarR420 | MW354072 | - | - | MW365901 | MW365835 | - |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 24.413 | 151.039 | PtrtrKroR419 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 24.413 | 151.039 | PtrtrKroR1316 | MW996600 | - | - | MW365898 | MW365832 | Assembly |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 26.781 | 152.882 | PtrtrMar | MW996604 | - | - | MW365903 | MW365837 | - |
| *Panesthia* | *tryoni tryoni* | Beasley-Hall et al. (2021) | 28.199 | 153.187 | PtrtrLam | MW996603 | - | - | MW365902 | MW365836 | - |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 28.241 | 153.266 | PtrtrSpr | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *tryoni tryoni* | newly sequenced | 28.546 | 153.287 | PtrtrNig | Assembly | - | - | Assembly | Assembly | Assembly |
| *Parapanesthia* | *gigantea* | Beasley-Hall et al. (2021) | ~28.213 | ~152.031 | Prpagiga | MW354073 | - | - | MW365904 | MW365838 | - |
| *Parapanesthia* | *pearsoni* | Beasley-Hall et al. (2021) | 23.801 | 149.134 | Pppear | MW996605 | - | - | KU577816 | KU577863 | - |
| *Panesthia* | sp. | newly sequenced | | | PR438 | Assembly | - | - | Assembly | Assembly | Assembly |
| *Panesthia* | *ancaudellioides* | Beasley-Hall et al. (2021) | | | Panca | MW354069 | - | - | MW365886 | MW365822 | - |
| *Panesthia* | *angusti angusti* | Beasley-Hall et al. (2021) | | | Pangu | MW996590 | - | - | KU577852 | - | - |
| *Panesthia* | *matthewsi* | Beasley-Hall et al. (2021) | | | Pmatt | MW354071 | - | - | MW365889 | MW365825 | - |
| *Panesthia* | *parva* | Beasley-Hall et al. (2021) | | | Pparv | MW996593 | - | - | MW365891 | MW365827 | Assembly |
| *Salganea* | *incerta* | Wang et al. (2023) | | | SAIN | OQ736973 | - | - | Assembly | Assembly | Assembly |
| *Caeparia* | *donskoffi* | Wang et al. (2023) | | | CADO | OQ736919 | - | - | Assembly | Assembly | Assembly |
| *Paranauphoeta* | *circumdata* | Bourguignon et al. (2018) | | | Paracirc | MG882225 | - | - | - | - | - |
| *Carbrunneria* | *paramaxi* | Bourguignon et al. (2018) | | | Carbpara | MG882214 | - | - | - | - | - |
| *Allacta* | *australiensis* | Bourguignon et al. (2018) | | | Allaaust | MG882127 | - | - | - | - | - |

**Figure S2.1**. Maximum-likelihood phylogeny of the Australian Panesthiinae and Geoscapheinae, inferred from whole mitogenomes and nuclear markers in IQTREE. Node support: 10,000 ultrafast bootstrap replicates/1000 SH-like approximate likelihood ratio test replicates. Scale bar denotes substitutions/site.



## Supplementary material for Chapter 3

**Supplementary Table S3.1**. List of samples included in this study. Sequences for samples marked "ZZ" were assembled by Zhuzhi Zhang and will be uploaded to GenBank in due course. Dashes (-) indicate missing sequences or geographic data. Question marks (?) indicate two samples provided by Wang and colleagues, which have been included in previous studies but have not yet been assigned GenBank accession numbers. For some taxa, only 12S, COII, 16S, 18S or ITS1 sequences from previous studies were available for analysis. Samples used as outgroups were shown in grey.

| Genus | Species | Name | Lat | Long | mito genome | 12S | COII | 16S | 18S | ITS1 | 28S |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Ancaudellia* | *kheili* | Akhei | ~5.85717S | ~144.23176 | - | AB036135 | AB036095 | - | - | - | - |
| *Ancaudellia* | *marshallae* | Amars | ~2.591S | ~140.667 | - | AB036136 | AB036096 | - | AB036188 | - | - |
| *Ancaudellia* | sp. | ZU006 | 5.139S | 145.773 | ZZ | - | - | - | - | - | - |
| *Ancaudellia* | *serratisima serratisima* | Aserser | 5.228S | 145.08 | ZZ | - | - | - | - | - | - |
| *Ancaudellia* | shawi | AshawJ | ~6.616S | ~142.827 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Caeparia* | crenulata | Cacren | ~4.962N | ~117.689 | - | AB036145 | AB036103 | - | AB036197 | - | - |
| *Caeparia* | *donskoffi* | CADO | ~19.638N | ~103.359 | OQ736919 | - | - | - | ZZ | ZZ | ZZ |
| *Geoscapheus* | *crenulatus crenulatus* | Geocren | 25.945S | 153.091 | MW996579 | - | - | - | MW365869 | MW365805 | ZZ |
| *Geoscapheus* | *dilatatus* | Gdila | ~25.794S | ~146.585 | MW600987 | - | - | - | KU577825 | MW365806 | - |
| *Geoscapheus* | *dilatatus non tub* | Gdilantu | 26.48S | 147.84 | MW354074 | - | - | - | MW365870 | MW365806 | ZZ |
| *Geoscapheus* | *dilatatus tub* | Gdilatub2 | 35.339S | 142.188 | MW354075 | - | - | - | MW365871 | MW365807 | ZZ |
| *Geoscapheus* | *dilatatus tub* | Gdilatub | 34.756S | 142.339 | MW600999 | - | - | - | KU577824 | KU577871 | - |
| *Geoscapheus* | *robustus* | Georobu | 34.767S | 142.332 | MW996606 | - | - | - | KU577821 | KU577868 | - |
| *Geoscapheus* | *rugulosus* | Grugulo | 23.768S | 149.094 | MW996580 | - | - | - | MW365873 | MW365809 | - |
| *Geoscapheus* | *woodwardii* | Gwood | 22.535S | 144.586 | MW996581 | - | - | - | MW365874 | MW365810 | - |
| *Macropanesthia* | *heppleorum* | Mhepp | 17.62S | 145.297 | MW996582 | - | - | - | MW365875 | KU577857 | - |
| *Macropanesthia* | *intermorpha* | Minter | 19.948S | 144.275 | - | KU577783 | KU577689 | - | KU577836 | - | - |
| *Macropanesthia* | *kinkuna* | Mkink | 24.966S | 152.475 | MW996583 | - | - | - | MW365876 | MW365812 | ZZ |
| *Macropanesthia* | *kraussiana* | Mkrau | 24.299S | 144.38 | MW996584 | - | - | - | MW365877 | MW365813 | - |
| *Macropanesthia* | *lineopuncta* | Mline | 23.557S | 145.719 | - | - | KU577686 | - | KU577837 | - | - |
| *Macropanesthia* | *lithgowae* | Mlith | 26.155S | 151.911 | MW354066 | - | - | - | KU577829 | KU577875 | - |
| *Macropanesthia* | *mackerrasae* | Mmack | 25.439S | 152.1 | MW996585 | - | - | - | MW365879 | MW365815 | ZZ |
| *Macropanesthia* | *montheithi* | Mmont | 26.711S | 151.785 | MW996586 | - | - | - | MW365880 | MW365816 | - |
| *Macropanesthia* | *mutica* | Mcpamuti | 16.952S | 145.598 | MW354067 | - | - | - | MW365881 | MW365817 | ZZ |
| *Macropanesthia* | *rhinoceros* | Mrhin | 23.772S | 146.585 | MG882202 | - | - | - | KU577805 | KU577858 | DQ874230 |
| *Macropanesthia* | *rothi* | Mroth | 24.213S | 151.903 | MW354068 | - | - | - | MW365882 | MW365818 | - |
| *Macropanesthia* | *saxicola* | Msaxi | 28.988S | 150.773E | - | KU577779 | KU577683 | - | KU577834 | KU577882 | - |
| *Macropanesthia* | *spuritegmina* | Mspur | 26.238S | 151.713 | - | KU577744 | - | - | - | - | - |
| *Miopanesthia* | deplanata | MideplRC | 5.777N | 116.343 | ZZ | - | - | - | ZZ | ZZ | ZZ |

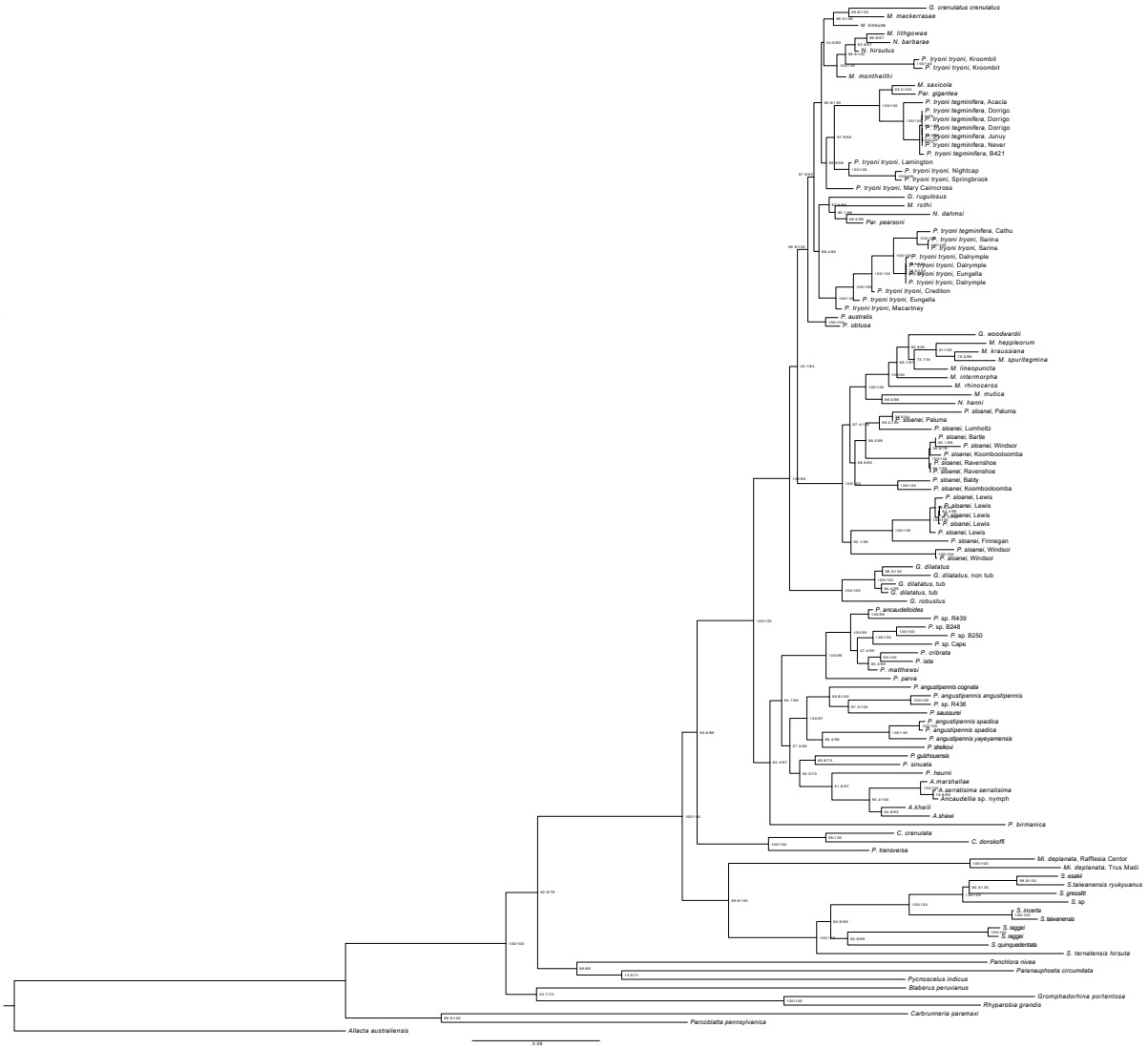
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Miopanesthia* | deplanata | MideplTM | 5.55N | 116.517 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Neogeoscapheus* | *barbarae* | Nbarb | 26.07S | 151.7 | - | KU577772 | KU577676 | - | - | KU577878 | - |
| *Neogeoscapheus* | *dahmsi* | Ndahm | 25.493S | 149.335 | MW996587 | - | - | - | MW365883 | MW365819 | - |
| *Neogeoscapheus* | *hanni* | Nhann | 16.665S | 145.299 | MW996588 | - | - | - | MW365884 | MW365820 | ZZ |
| *Neogeoscapheus* | *hirsutus* | Nhirs | 24.81S | 152.33 | MW996589 | - | - | - | MW365885 | MW365821 | - |
| *Panesthia* | sp. | PCape | 19.732S | 147.814 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | sp. | PR439 | - | - | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | sp. | PB248 | 20.271S | 148.582 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | sp. | PB250 | 20.341S | 148.678 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | sp. | PR438 | - | - | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *ancaudellioides* | Panca | 16.091S | 145.461 | MW354069 | - | - | - | MW365886 | MW365822 | - |
| *Panesthia* | *angusti angusti* | Pangu | 10.429N | 105.664 | MW996590 | - | - | - | KU577852 | - | - |
| *Panesthia* | *angustipennis cognata* | PAAU | ~24.95N | ~102.64 | OQ736943 | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | angustipennis spadica | Pangspa | - | - | OL685387 | - | - | - | AB036190 | - | - |
| *Panesthia* | angustipennis spadica | PangspaJ | ~32.734N | ~133.005 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | angustipennis yayeyamensis | PangyayJ | ~24.427N | ~124.183 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *australis* | Paust | 37.563S | 140.769 | MW996591 | - | - | - | MW365888 | MW365824 | - |
| *Panesthia* | birmanica | Pbirm | ~24.427N | ~124.183 | - | MF286795 | MF287046 | MF286867 | - | - | MF286926 |
| *Panesthia* | *cribrata* | Pcrib | 24.527S | 151.471 | - | KU577792 | KU577714 | - | KU577845 | KU577891 | DQ874240 |
| *Panesthia* | *guizhouensis* | PAGU | ~30.321N | ~105.855 | OQ736944 | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | heurni | Pheur | ~6.018N | ~116.030 | - | AB036141 | AB036099 | - | AB036193 | - | - |
| *Panesthia* | *lata* | Plata | 31.499S | 159.068 | - | AB036171 | AB036125 | - | KU577840 | KU577887 | |
| *Panesthia* | *matthewsi* | Pmatt | 25.57S | 152.053 | MW354071 | - | - | - | MW365889 | MW365825 | |
| *Panesthia* | *obtusa* | Pobtus | 26.02S | 148.271 | MW996592 | - | - | - | MW365890 | MW365826 | ZZ |
| *Panesthia* | *parva* | Pparv | 17.185S | 145.31864 | MW996593 | - | - | - | MW365891 | MW365827 | ZZ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panesthia* | saussurii | PsausJ | ~8.814N | ~125.104 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | sinuata | Psinu | ~21.947N | ~100.474 | - | MF286797 | MF287043 | MF286869 | - | - | MF286928 |
| *Panesthia* | *sloanei* | PsloLewB26 | 16.59S | 145.29 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloBar | 17.402S | 145.82 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloFinn | ~15.818S | ~145.284 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloKoo | 17.83S | 145.6037 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloJ122 | 17.83S | 145.6037 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloLewR1318 | 16.59S | 145.29 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloWinR1319 | 16.261S | 145.042 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloLewR418 | 16.59S | 145.29 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloLewR4210 | 16.59S | 145.29 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloRavR4212 | 17.525S | 145.511 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloWinB21 | 16.261S | 145.042 | - | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloRavB23 | 17.525S | 145.511 | ZZ | - | - | - | X | X | X |
| *Panesthia* | *sloanei* | PsloPalB29 | 19.01S | 146.2 | - | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloLum | 18.212S | 145.799 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloPalR417 | 19.01S | 146.2 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloRavR4211 | 17.525S | 145.511 | - | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloBal | ~17.277S | ~145.450 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *sloanei* | PsloaWinB22 | 16.233S | 145.007 | MW996594 | - | - | - | MW365892 | - | - |
| *Panesthia* | *sloanei* | PsloaPalB24 | 17.525S | 145.511 | MW996597 | - | - | - | MW365895 | - | - |
| *Panesthia* | *sloanei* | PanesloB25 | 16.59S | 145.29 | MW996596 | - | - | - | MW365894 | MW365829 | - |
| *Panesthia* | *strelkovi* | PAST | ~18.613N | ~109.508 | OQ736946 | - | - | - | - | - | ZZ |
| *Panesthia* | *transversa* | PtransvJ | ~6.018N | ~116.030 | - | ZZ | ZZ | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteDoB27 | 30.373S | 152.72 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteDoB28 | 30.373S | 152.72 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteDoJ103 | 30.358S | 152.731 | ZZ | - | - | - | ZZ | ZZ | ZZ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Panesthia* | *tryoni tegminifera* | PtrteAca | 28.364S | 152.394 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteNev | 30.355S | 152.802 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteJun | 30.296S | 152.746 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tegminifera* | PtrteR421 | 30.373S | 152.725 | MW996599 | - | - | - | MW365897 | MW365831 | - |
| *Panesthia* | *tryoni tryoni* | PtrtrCre | 21.185S | 148.525 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrSpr | 28.241S | 153.266 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrDalJ100 | 21.035S | 148.597 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrEunJ101 | 21.057S | 148.581 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrNig | 28.546S | 153.287 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrteCat | ~20.812S | ~148.567 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrSarR1314 | 21.61S | 148.97 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrDalR1315 | 21.05S | 148.55 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrKroR419 | 24.41S | 151.05 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrDalR434 | 21.05S | 148.55 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrLam | 28.199S | 153.187 | MW996603 | - | - | - | MW365902 | MW365836 | - |
| *Panesthia* | *tryoni tryoni* | PtrtrMar | 26.781S | 152.882 | MW996604 | - | - | - | MW365903 | MW365837 | - |
| *Panesthia* | *tryoni tryoni* | PtrtrEunB214 | 21.118S | 148.511 | MW996601 | - | - | - | MW365899 | MW365833 | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrB215 | 20.834S | 148.553 | MW996602 | - | - | - | MW365900 | MW365834 | - |
| *Panesthia* | *tryoni tryoni* | PtrtrKroR1316 | 24.41S | 151.05 | MW996600 | - | - | - | MW365898 | MW365832 | ZZ |
| *Panesthia* | *tryoni tryoni* | PtrtrSarR420 | 21.61S | 148.97 | MW354072 | - | - | - | MW365901 | MW365835 | - |
| *Parapanesthia* | *gigantea* | Prpagiga | ~28.213S | ~152.031 | MW354073 | - | - | - | MW365904 | MW365838 | - |
| *Parapanesthia* | *pearsoni* | Pppear | 23.8S | 149.13 | MW996605 | - | - | - | KU577816 | KU577863 | - |
| *Salganea* | *incerta* | SAIN | ~28.600N | ~106.342 | OQ736973 | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *quinquedentata* | SAQU | ~18.613N | ~109.830 | OQ736974 | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *raggei* | SARA | ~18.898N | ~109.704 | OQ736975 | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | sp. | SD3 | - | - | ? | - | - | - | ZZ | ZZ | ZZ |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Salganea* | *taiwanensis* | WN | - | - | ? | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *ternatensis hirsuta* | Sterhir | 5.228S | 145.08 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *esakii* | SaesakJ | ~30.257N | ~130.564 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *gressiti* | SagreJ | ~24.181N | ~121.28 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *raggei* | SaragJ | - | - | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Salganea* | *taiwanensis ryukyuanus* | Satairyu | ~28.292N | ~129.319 | ZZ | - | - | - | ZZ | ZZ | ZZ |
| *Blaberus* | *peruvianus* | Bperu | - | - | MG882135 | - | - | - | - | - | - |
| *Panchlora* | *nivea* | Pancnive | - | - | MG882155 | - | - | - | EF363244 | - | KF372451 |
| *Pycnoscelus* | *indicus* | Pycnindi | - | - | MG882158 | - | - | - | - | - | - |
| *Paranauphoeta* | *circumdata* | Paracirc | - | - | MG882225 | - | - | OQ738143 | - | - | - |
| *Rhyparobia* | *grandis* | Rhypgran | - | - | MG882231 | - | - | - | - | - | - |
| *Gromphadorhina* | *portentosa* | Gromport | - | - | MG882150 | - | - | - | EF383466 | - | EF383626 |
| *Parcoblatta* | *pennsylvanica* | Parcpenn | - | - | MG882171 | - | - | - | - | - | - |
| *Carbrunneria* | *paramaxi* | Carbpara | - | - | MG882214 | - | - | - | - | - | - |
| *Allacta* | *australiensis* | Allaaust | - | - | MG882127 | - | - | - | - | - | - |

**Figure S3.1**. Maximum-likelihood phylogeny of the Panesthiinae and Geoscapheinae, inferred from whole mitogenomes and nuclear markers in IQTREE. Node support: 10,000 ultrafast bootstrap replicates/1000 SH-like approximate likelihood ratio test replicates. Scale bar denotes substitutions/site.



## Supplementary material for Chapter 4

**Figure S4.1**. Flow diagram illustrates the procedures for identify, filter, and age *Blattabacterium cuenoti* inserts in the cockroach and termite genomes.

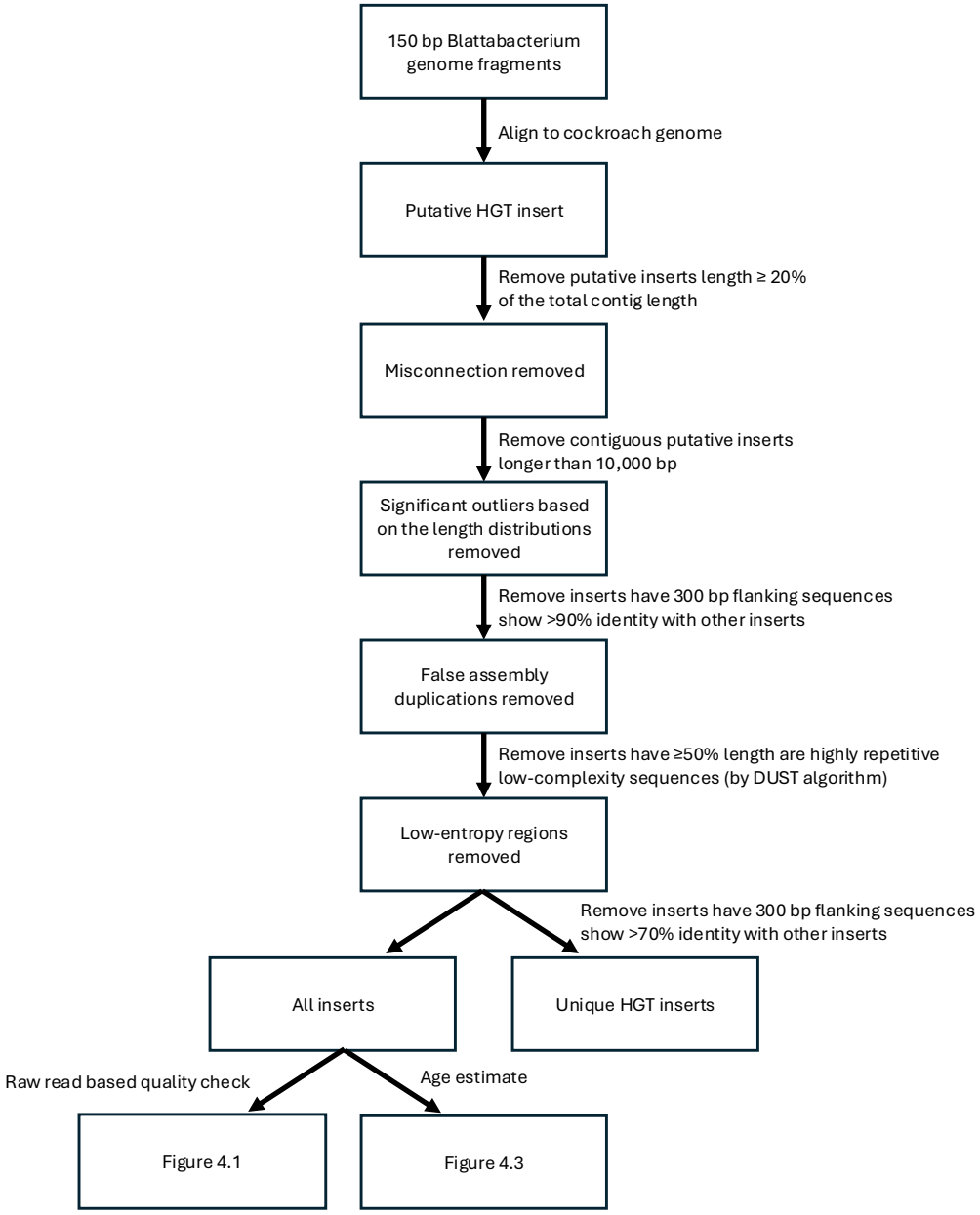
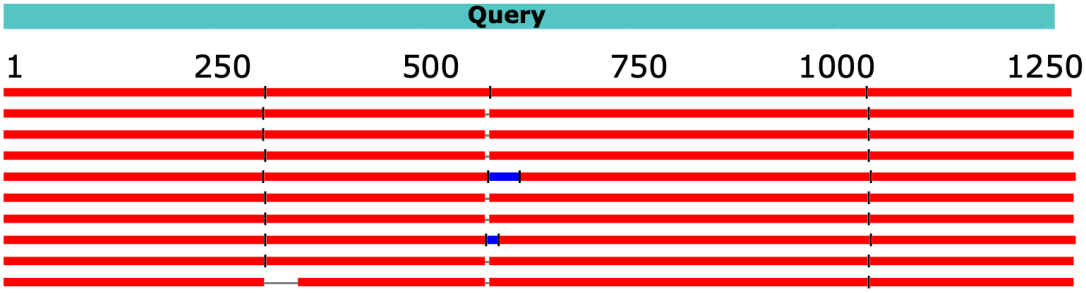


**Figure S4.2**. Five examples of chimeric inserts, illustrated using BLAST results. For each insert, the ten best BLAST hits are depicted, with the alignments of each fragment in the top hit.

Additional chimeric sequence 1:

*Diploptera punctata* JASPKZ010002709.1: 268362-269615 (1253 bp)




Blattabacterium sp. DPU chromosome

Sequence ID: **CP049785.1** Length: 623008 Number of Matches: 6

Range 1: 519604 to 520049

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 801 bits(887) | 0.0() | 445/446(99%) | 0/446(0%) | Plus/Minus | |

```
Query  567     GAAGAATAGGATCCATAAATTCTAAGATTTTTTGTTACAGAAGGCAttttttttAATTTT  626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  520049  GAAGAATAGGATCCATAAATTCTAAGATTTTTTGTTACAGAAGGCATTTTTTTTTAATTTT  519990

Query  627     CCTATTTCGATAGCTTTATCAACAGCATCTTCTGCCATTTTTCTATATGTAGTCCATTTT  686
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519989  CCTATTTCGATAGCTTTATCAACAGCATCTTCTGCCATTTTTCTATATGTAGTCCATTTT  519930

Query  687     CCCCCTATAATACTAATAAGTCCAGAAGAACTAATCATAAGTTTATGAGATCTAGAAATA  746
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519929  CCCCCTATAATACTAATAAGTCCAGAAGAACTAATCATAAGTTTATGAGATCTAGAAATA  519870

Query  747     TCTTTAGTTTTAGTAATATCAGAACAAGAATTTTTAGGaaaaaaaGAGGCCGTAATCCA  806
               ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Sbjct  519869  TCTTTAGTTTTAGTAATATCAGAACAAGAATTTTTAGGAAAAAAAAGAGGCCGCAATCCA  519810

Query  807     GAGAATGCACTTAATATATCACtttttttGTATGAAATACAAAATATTTGTTAAAAGTT  866
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519809  GAGAATGCACTTAATATATCACTTTTTTTTGTATGAAATACAAAATATTTGTTAAAAGTT  519750

Query  867     TGTAATATAAAATCTATTTCTTTTTCTAAAGGTTTTGGTTCAAGAACACTTTTTTCCAAA  926
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519749  TGTAATATAAAATCTATTTCTTTTTCTAAAGGTTTTGGTTCAAGAACACTTTTTTCCAAA  519690

Query  927     AAAGTATCTGTAGTTCCAACTAAAACATGATCACACCATGGGACACAAAATAATACTCTT  986
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  519689  AAAGTATCTGTAGTTCCAACTAAAACATGATCACACCATGGGACACAAAATAATACTCTT  519630

Query  987     CCATCTGCAGTTTTTGGAATAACTAT  1012
               |||||||||||||||||||||||||
Sbjct  519629  CCATCTGCAGTTTTTGGAATAACTAT  519604
```

Range 2: 498063 to 498366

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 538 bits(596) | 5e-152() | 303/305(99%) | 1/305(0%) | Plus/Plus | |

```
Query  1       CTTTTATATaaaaaaaTTATTGAAAAATTTCCAAATATTGAGTTTATCGCAAGTGGAGGA  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  498063  CTTTTATATAAAAAAATTATTGAAAAATTTCCAAATATTGAGTTTATCGCAAGTGGAGGA  498122

Query  61      ATTAGTAATATAGATGATGTAGATCAATTATTTAATTTAGGTTGTAGTGGAGTCATTATT  120
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  498123  ATTAGTAATATGGATGATGTAGATCAATTATTTAATTTAGGTTGTAGTGGAGTCATTATT  498182

Query  121     GGAAAAGCTGTATATGAAAATAAAATATCATTATCTGATCTTAAAGATTGGATAAGaaaa  180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  498183  GGAAAAGCTGTATATGAAAATAAAATATCATTATCTGATCTTAAAGATTGGATAAG-AAA  498241

Query  181     aaaaaaTAATAATAAATCAATATGTTAGCTAAACGTATTATTCCTTGTTTGGACATTAAA  240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  498242  AAAAAATAATAATAAATCAATATGTTAGCTAAACGTATTATTCCTTGTTTGGACATTAAA  498301

Query  241     AATGGAAGAACCGTAAAAGGAATAAATTTTAAACATTTAAAAGATGCAGGAGATCCAATA  300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  498302  AATGGAAGAACCGTAAAAGGAATAAATTTTAAACATTTAAAAGATGCAGGAGATCCAATA  498361

Query  301     CAATT  305
              |||||
Sbjct  498362  CAATT  498366
```

Range 3: 172184 to 172451

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 480 bits(531) | 4e-134() | 267/268(99%) | 0/268(0%) | Plus/Minus | |

```
Query  302     AATTGGAAACAttttttttAATACCaaaaatacttcttgtaaaataaaaaaatttatctc  361
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Sbjct  172451  AATTGGAAACATTTTTTTTAATACCAAAAATACTTCCTGTAAAATAAAAAAATTTATCTC  172392

Query  362     aaaaaaaaatttcgttgagtggattaaaaaatattctattcaaaaaaCGGATTATCTTCT  421
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  172391  AAAAAAAAATTTCGTTGAGTGGATTAAAAAATATTCTATTCAAAAAACGGATTATCTTCT  172332

Query  422     CATTAAAGGATCAAGAAATATTGCATTAGAAAGTCTTATTTCCTTAATTTGATAGAAATT  481
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  172331  CATTAAAGGATCAAGAAATATTGCATTAGAAAGTCTTATTTCCTTAATTTGATAGAAATT  172272

Query  482     CTTTTTCTTGGTATATTGGATCTTAAAGATTAAATTTGTATACAATGATTTCATAAAATT  541
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  172271  CTTTTTCTTGGTATATTGGATCTTAAAGATTAAATTTGTATACAATGATTTCATAAAATT  172212

Query  542     TTTCTTCATGAAAGAAATTACCAAAGAA  569
              ||||||||||||||||||||||||||||
Sbjct  172211  TTTCTTCATGAAAGAAATTACCAAAGAA  172184
```

Range 4: 521046 to 521289

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 436 bits(483) | 5e-121() | 243/244(99%) | 0/244(0%) | Plus/Plus | |

```
Query  1010    TATTTAATATTCCAATTACCATGCTTCCAGAGGTAAAATCTTCTAGTGAAATTTTTGGTT  1069
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521046  TATTTAATATTCCAATTACCATGCTTCCAGAGGTAAAATCTTCTAGTGAAATTTTTGGTT  521105

Query  1070    ATACAACAGGACATATTTTATCCCATAAAATCCCTATATCTGGGATAGCTGGAGATCAAC  1129
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521106  ATACAACAGGACATATTTTATCCCATAAAATCCCTATATCTGGGATAGCTGGAGATCAAC  521165

Query  1130    AAGCCGCTCTTTTTGGTCAGATGTGTACCAAAATTGGGATGGTGAAAAATACTTATGGAA  1189
               |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521166  AAGCTGCTCTTTTTGGTCAGATGTGTACCAAAATTGGGATGGTGAAAAATACTTATGGAA  521225

Query  1190    CAGGATGTTTTATGTTAATGAATGTAGGAAATAATCCTGTTTTTTCTAGAAATAATTTAA  1249
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521226  CAGGATGTTTTATGTTAATGAATGTAGGAAATAATCCTGTTTTTTCTAGAAATAATTTAA  521285

Query  1250    TAAC  1253
               ||||
Sbjct  521286  TAAC  521289
```

Range 5: 520139 to 520168

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 44.6 bits(48) | 0.003() | 29/31(94%) | 1/31(3%) | Plus/Plus | |

```
Query  460     TTTCCTTAATTTGATAGAAATTCTTTTTCTT  490
               ||||||||||||  | ||||||||||||||||
Sbjct  520139  TTTCCTTAATTTCA-AGAAATTCTTTTTCTT  520168
```

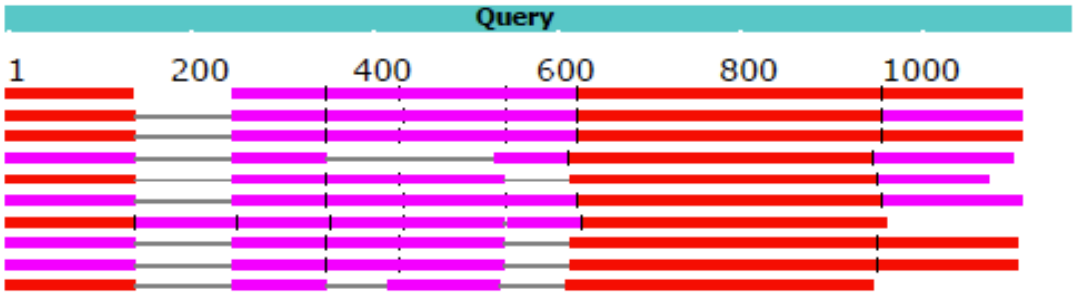
Range 6: 386042 to 386071

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 42.8 bits(46) | 0.011() | 28/30(93%) | 1/30(3%) | Plus/Plus | |

```
Query  530     TTTCATAAAATTTTTCTTCATG-AAAGAAA  558
               || |||||||||||||||||| ||||||||
Sbjct  386042  TTACATAAAATTTTTCTTCATGAAAAGAAA  386071
```

Additional chimeric sequence 2:

*Neogeoscapheus dahmsi* contig_7972: 1051443-1052600 (1157 bp)




Blattabacterium cuenoti strain BPAU chromosome
Sequence ID: **CP142618.1** Length: 632444 Number of Matches: 6
Range 1: 340984 to 341323

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 457 bits(247) | 9e-128() | 312/341(91%) | 13/341(3%) | Plus/Minus | |

```
Query  596     CTATTAATTTTAATT--aaaaaaaaa-TA-TTTACGCTATCAACAATGTTGTTCTATTTT  651
Sbjct  341323  CTATTAATTTTAATTTTAAAAAAAAATTACTTCAAGCTATCAACAACGTTGTTCTATTTT  341264
Query  652     TTGATAAAATAGAATTAACAACTAATATGTTGCAATATAATTTGATATCAAATCAAAttt  711
Sbjct  341263  TTGATAACAAAAAATTAACAACTAATATGTTGGAATATAATTTGATATCAAATCAAATTT  341204
Query  712     ttttAT-AAAAAATGA-CATTCTCCATTATGGAGATTATACCATATGTAGTAAAGAA-Ct  768
Sbjct  341203  TTT-ATAAAAAAAAAAGCATTCTCCATTATGGAGATTATACCATATGTAGTAAAGAAGGT  341145
Query  769     -ttttttACATCAATA---CAAAAATAGAGTTAAAACATGAAATAAAATTGACTGGAAAA  824
Sbjct  341144  CTTTTTTACATAAATAAAAAAAAAATAGAGTTAAAACATGAAATCAAATTGATTGGAAAA  341085
Query  825     AATTATACTGTATATGCTAATACGTTAGTATATCTATTTAAACAAGATCAAATCCATTT-  883
Sbjct  341084  AATTATACTGTATATGCCAATACGTTAGTATATCTATTTAAACAAGATCAAATCCATTTT  341025
Query  884     CATGATCCCGCCATCATAGTACAAAATACAAATTTTGATAA  924
Sbjct  341024  CATGATCCCGCTATCATAGTACAAAATACAAATTTTGATAA  340984
```

Range 2: 586822 to 586969

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 213 bits(115) | 2e-54() | 138/149(93%) | 2/149(1%) | Plus/Minus | |

```
Query  924     ACCAAAGTTTTTTCAGAATTAAAATTTGTAAGTTGATCAAAACAATTCATAATTCTGAAA  983
Sbjct  586969  ACCAAAGTTTTTTCAGAATTAAAATTTGTAAGTTGATCAAAACAATTTACAATTCTAAAA  586910
Query  984     CGTTGACCCGGTGTTATAGGCTTTAAttttttttATTGACATTAATTAAAATTCTTCTTT  1043
Sbjct  586909  CGTTGACCAGGTGTTATAGGTTTTAACttttttt-ATTGACATTAGTTAAAATTCTTTTT  586851
Query  1044    GTTTAAAAAATCAATTTTT-GATATTCAA  1071
Sbjct  586850  GTTTAAAAAATCAATTTTTTGATCTTCAA  586822
```

Range 3: 330048 to 330165

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 169 bits(91) | 5e-41() | 109/118(92%) | 0/118(0%) | Plus/Plus | |

```
Query  406     TAGATTTAAAAATTCGACAATCTATAATTAACTCATGTGTCACAATATTATTTTCTCCCG  465
               ||||||||||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||
Sbjct  330048  TAGATTTAAAAATTCGACAATCTATAATTAACTCATGTGCCACAGTATTATTTTCTCCCA  330107

Query  466     CATACAATAGATTATAAAATTTCACCAATTTTTGTTTGATGTAATTTGCATTTAACAC  523
               |||||||||  |||||||||||||| ||| ||||| |||| ||||||||||||||||||||
Sbjct  330108  CATACAATATGTTATAAAATTTCATCAACTTTTCTTTTATGTAATTTGCATTTAACAC  330165
```

Range 4: 527873 to 527969

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 158 bits(85) | 1e-37() | 93/97(96%) | 0/97(0%) | Plus/Plus | |

```
Query  241     AAAAAAGTTTTCAGTTCTCAAAACTTTATTGATTAAAAAATTGCAATCTATTTTGAATGA  300
               ||||||| ||||||||||||| ||||||||||||||||||| ||||||||||||||||||||
Sbjct  527873  AAAAAAATTTTCAGTTCTCAGAACTTTATTGATTAATAAATTGCAATCTATTTTAAATGA  527932

Query  301     AAAATTATGTCATAATTCTGTTTTGAATGaaaaaaaa  337
               |||||||||||||||||||||||||||||||||||||
Sbjct  527933  AAAATTATGTCATAATTCTGTTTTGAATGAAAAAAAA  527969
```

Range 5: 623980 to 624056

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 119 bits(64) | 5e-26() | 73/77(95%) | 2/77(2%) | Plus/Plus | |

```
Query  339     aaTCTTCG-AAAAAAATTACATTTATTTTTAGTTGTTTGGATAAATGATCAGCTAAAAATT  397
               |||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623980  AATCTTCGAAAAAAAAATACATTTATTTTTAGTTGTTTGGATAAATGATCAGCTAAAAATT  624039

Query  398     TTAAAGAATA-GATTTA  413
               |||||||||| | |||||
Sbjct  624040  TTAAAGAATAAGTTTTA  624056
```

Range 6: 48695 to 48769

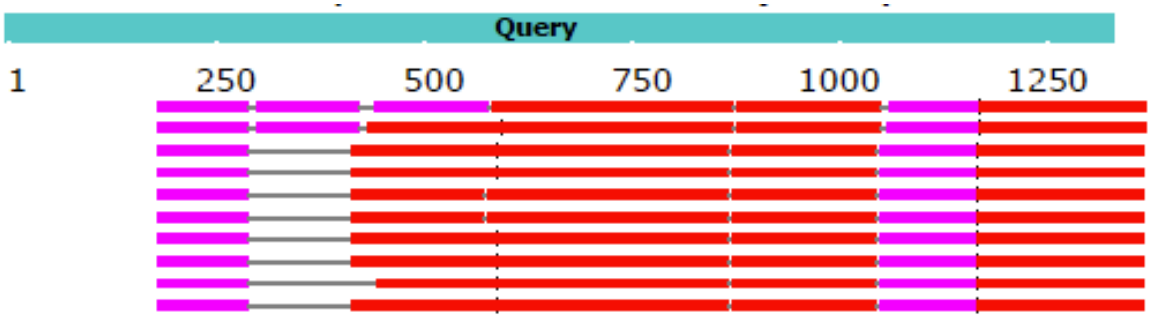
| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 115 bits(62) | 6e-25() | 71/75(95%) | 2/75(2%) | Plus/Minus | |

```
Query  524     GTTGATTACGATATCTCGAGTTTTTGTAGGTATAGAAACCACGGTGTATGACT--TAATT  581
               |||||||||||||||  |||||||||||||||||||||||||||||||||  |||||
Sbjct  48769   GTTGATTACGATATCTTAAGTTTTTGTAGGTATAGAAACCACGGTGTATGACTCCTAATT  48710

Query  582     TTCTAATTCCTTTCC  596
               |||||||||||||||
Sbjct  48709   TTCTAATTCCTTTCC  48695
```

Additional chimeric sequence 3:

*Panesthia cribrata* ctg. 000829F: 236708-238036 (1328 bp)




Blattabacterium cuenoti BPAA DNA, complete genome
Sequence ID: **AP012548.1** Length: 632490 Number of Matches: 7
Range 1: 605592 to 605875

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 462 bits(250) | 2e-129() | 273/284(96%) | 2/284(0%) | Plus/Plus | |

```
Query  570     TTCCTTTTTTAAGGAGGAAGAAAATATAGCTaaaaaaaGACGATCTAATCCTAAAGATGT  629
               ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Sbjct  605592  TTCCTTTTTAAGGAAGAAGAAAATATAGCTAAAAAAAGACGATCTAATCCTAAAGATGT  605651

Query  630     TTCTATAACATAAGGAATAGAatttttttttcattcaaaaattcgaatttttttt-cg-  687
               |||||||||||||||||| |||||||||| ||||||||||||||||||||||||| ||
Sbjct  605652  TTCTATAACATAAGGAATATAATTTTTTTTCCATTCAAAAATTCGAATTTTTTTTTCGA  605711

Query  688     aaaaaattcatgatttttAAATCAAAATCTCTACGAGAATGAATCCCTTCTATTTCTTG  747
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Sbjct  605712  AAAAAATTCATGATTCTTTAAATCAAAATCTCTACGAGAATGAATTCCTTCTATTTCTTG  605771

Query  748     AAATCCAAAAGGAAAATGAAATTCTATATCTGATTCCGCACTTGCATAATGAGACAAATG  807
               |||||||||||||||||||| |||||||||| |||||||||||||||||| |||||||
Sbjct  605772  AAATCCAAAAGGAAAATGAAATTTTATATCTGATCCCGCACTTGCATAATGAGATAAATG  605831

Query  808     ATCATGATCACATAACTGATAATATGTTTTATCTTCTAAATTTA  851
               |||||||||||||||||||||||||||||||||| |||||||||
Sbjct  605832  ATCATGATCACATAACTGATAATATGTTTTATCTCCTAAATTTA  605875
```

Range 2: 182234 to 182431

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 322 bits(174) | 4e-87() | 190/198(96%) | 0/198(0%) | Plus/Minus | |

```
Query  1131    CTGACCCATTTTCAGATAATGCATTTTCAACTAAAATTTTAATATTTTGCCATTGTAAAA  1190
               |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  182431  CTGATCCATTTTCAGATAATGCATTTTCAACTAAAAATTTAATATTTTGCCATTGTAAAA  182372

Query  1191    ATGAAATATCTTGAAATTCATCTATGAAATAATGTTTATATTGTGTGCCTATTTTTTCAT  1250
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct  182371  ATGAAATATCTTGAAATTCATCTATGAAATAATGTTTATATTGTGTGCCTATTTTTTCAT  182312

Query  1251    ATATTTTTGGAAATGTTCCTTCAACAATTCTTTCATAAAGAATTTTATTTAATTCTGCAT  1310
               |||||||||||||||||||||||||| |||||| || ||||||||||||||||||||||
Sbjct  182311  ATATTTTTGGAAATGTTCCTTCAAGAATTCTCTCGTAAAGAATTTTATTTAATTCTGCAT  182252

Query  1311    TTAAGAttttttttttt  1328
               ||||  ||  |||||||||
Sbjct  182251  TTAAAATAATTTTTTTTT  182234
```

Range 3: 103094 to 103260

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 270 bits(146) | 2e-71() | 160/167(96%) | 0/167(0%) | Plus/Minus | |

```
Query  853     AAAATAGAGATATAAGATTTTATATTTCCAATGATATTATTATGTGTAAAAGGGATGTAG  912
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  103260  AAAATAGAGATATAAGATTTTATATTTCCAATGATATTATTATGTGTAAAAGGGATGTAG  103201

Query  913     ATAGGAGATATAGGCATTCGTAGTCCATAATGAATTTTGCTGACTAATATGAAATCTCCT  972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  103200  ATAGGAGATATAGGCATTCGTAATCCATAATGAATTTTACTGACTAATATGAAATCTCCT  103141

Query  973     ACTAATAAAGTTCTTTTCATAGAAGGAGTGGGAATCACAGAAGATTG  1019
               ||||||||| ||||||| |||||||| ||||||||||||| ||| |||
Sbjct  103140  ACTAATAAGGTTCTTTCCATAGAAGAAGTGGGAATCACAAAAGGTTG  103094
```

Range 4: 24045 to 24160

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 198 bits(107) | 7e-50() | 114/117(97%) | 1/117(0%) | Plus/Minus | |

```
Query  296     TGAAGTTTGAATGCGTTCTTCTTGAAACGttttttttttaacgttttttGGAAAAAACGAG  355
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24160   TGAAGTTTGAATGCGTTCTTCTTGAAACGttttttttt-AACGttttttGGAAAAAACGAG  24102

Query  356     TAATGATATACATAAAAGGAATAGTGAAAGAAACTATAAAAGATAATTTTTAATGAA  412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24101   TAATGATATACATAAAAGGAATAGTGAAAAAAACTATAAAAGATAATTTTTGATGAA  24045
```

Range 5: 78195 to 78326

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 193 bits(104) | 3e-48() | 125/134(93%) | 5/134(3%) | Plus/Plus | |

```
Query  434    ATAAAAAGAAAATGAATCTA---TTTTTTATCTTTTATAAAATTCATGATAGATTTGTTA  490
              ||||||| |||| ||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  78195  ATAAAAA-AAAA-GAATCTATTTTTTTTTATCTTTTATAAAATTCATGATAGATTTGTTA  78252

Query  491    GAACCAAAAAGAATTAATATATCCCCCTTTTGTAAAACAGTCTCTCCTGTAACTAATCCT  550
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  78253  GAACCAAAAAGAGTTAATATATCCCCTTTTGTAAAACAGTTCTCCTGTCACTAATCCT  78312

Query  551    ATGACTTTTCTTGT  564
              ||||||||||||||
Sbjct  78313  ATGACTTTTCTTGT  78326
```

Range 6: 504481 to 504582

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 161 bits(87) | 1e-38() | 97/102(95%) | 0/102(0%) | Plus/Plus | |

```
Query  183     ctaaaattTATGAATGTTATACAAAACGTTGTTTTCAAGCAAACGCGTTAGATTTTGATG  242
               ||||||||||| ||| ||||||||||||||||||||||||||||| || ||||||||||||
Sbjct  504481  CTAAAATTTATAAATTTTATACAAAACGTTGTTTTCAAGCAAATGCATTAGATTTTGATG  504540

Query  243     ATATATTACTTCATACTAATTATTAAttttttATTTTCCAA  284
               |||||||||||||||||||||||||| |||||||||||||
Sbjct  504541  ATATATTACTTCATACTAATTATTTATTTTTTTATTTTCCAA  504582
```

Range 7: 174666 to 174765

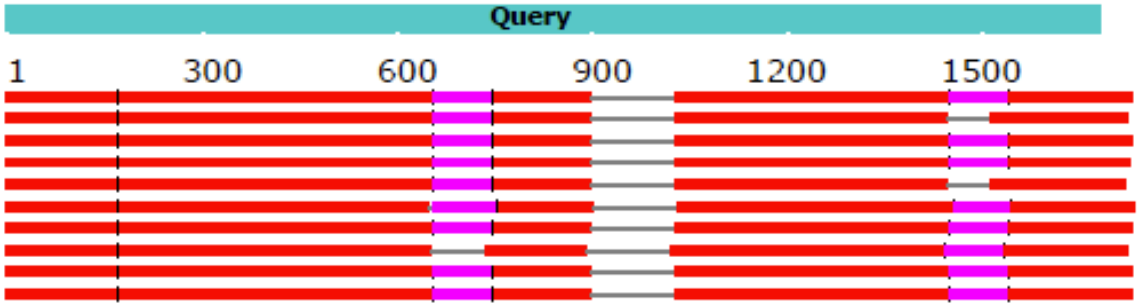
| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 158 bits(85) | 1e-37() | 95/100(95%) | 0/100(0%) | Plus/Minus | |

```
Query  1032    TAATAAATAATTATAAAAGTTTGTTAATAGATTGAATTAATATTTCTTTCTCTATTATAG  1091
              |||||||||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||
Sbjct  174765  TAATAAATAATTATAAAAGTTTGTTAATAGATTGAATTAATATTTTTTTTCTATTATAG  174706

Query  1092    AAACttttttttGATAAAGATATTGGAGTCTCCTCTGAATC  1131
              ||||||| ||||||||||||| |||||||||| |||||||
Sbjct  174705  AAACTTTCTTTGATAAAGATACTGGAGTCTCCCCTGAATC  174666
```

Additional chimeric sequence 4:

*Parapanesthia gigantea* contig_13648: 515186-516859 (1673 bp)




Blattabacterium cuenoti strain BMK chromosome

Sequence ID: **CP142624.1** Length: 632180 Number of Matches: 7

Range 1: 251744 to 252223

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 837 bits(453) | 0.0() | 472/481(98%) | 2/481(0%) | Plus/Plus | |

```
Query  157     CTCTTCGAGTATTTTCTAGATATTTTGGAATTTTCAATAATATAACATCTACTACTACAA  216
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  251744  CTTTTCGAGTATTTTCTAGATATTTTGGAATTTTCAATAATATAACATCTACTACTACAA  251803

Query  217     AAAATAGATTCCGAGTTACGGTTTCAAAAATAATGGATTTTGAACATCCAGGAATTTTGA  276
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  251804  AAAATAGATTCCGAGTTACGGTTTCAAAAATAATGGATTTTGAACATCCAGGAATTTTTA  251863

Query  277     ATCAAGCAATTATGGATTTAGGTTCTATTTTATGTATTCCAAAAAGTCCTAAATGTTTAT  336
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  251864  ATCAAGCAATTATGGATTTAGGTTCTATTTTATGTACTCCAAAAAGTCCTAAATGTTTAT  251923

Query  337     TATGTCCAGTTCAAGATTCTTATTTTTCTATTCAAAAATGGAACTGTACATAAATTACCT  396
               |||||||||||||||||||| |||||||||| | ||||||||||||||||||||||||
Sbjct  251924  TATGTCCAGTTCAAGATTCTTGTTTTTCTATCC-AAAATGGAACTGTACATAAATTACCT  251982

Query  397     GT-aaaaaaataaaaaaaTTCATAAGACATAGATttttttATTATCTTTTCATATGTGAT  455
               || |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  251983  GTAAAAAAAATAAAAAGATTCATAAGACATAGATTTTTTTATTATCTTTTCATATGTGAT  252042

Query  456     CATAACAAAAATATTTGTTTAAATAAAAGATCAACTAAAGATATATGGAAGGGTCTTTAT  515
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  252043  CATAACAAAAATATTTGTTTAAATAAAAGATCAACTAAAGATATATGGAAGGGTCTTTAT  252102

Query  516     GATTTTCCTTTAATAGAATCGAAACAAAATCTTTCAATTCATGAAATCATAGATGAAACT  575
               |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  252103  GATTTTCCTTTAATAGAATCGAAACAAAATCTTTCAATTCATGAAATCATAGATAAAACT  252162

Query  576     TGGGAAAAATTTAGAGTGAGGTTTTCTAAAAATGTGATTTATAAAGTAGAACAAAAACTG  635
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  252163  TGGGAAAAATTTAGAGTGAGGTTTTCTAAAAATGTGATTTATAAAGTAGAACAAAAACTG  252222

Query  636     A  636
               |
Sbjct  252223  A  252223
```

Range 2: 372686 to 373117

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 542 bits(293) | 4e-153() | 392/434(90%) | 30/434(6%) | Plus/Plus | |

```
Query  999     attttaaaaataatatccatttatgaaaatggtttttaaaattaacagattttgaatttt  1058
Sbjct  372686  ATTTTAAAAATAATATCCATTTATGAAAATGGTTTTTTAAAATTAACAGATTTTGAATTT  372745

Query  1059    ttaatttcagaaaaaacaaaaatAGTATCTATTAGTCATATATATCTAATGTTTTAGGGA  1118
Sbjct  372746  TTAATTTCAGAAAAAACAAAAATAGTATCTATTAGTC--ATATATCTAATGTTTTAGGGA  372803

Query  1119    TAATTAATCCTGTTCAAGATATTATTAAAAAAGCTCGTGAATATTGAGCTTTAGTTTTGA  1178
Sbjct  372804  TTATTAATCCTGTTCAAGATATTATTAAAAAAGCTCATGAATATGGAGCTTTAGTTTTGA  372863

Query  1179    TTGATGGAGTTCAAGTCCCTTCTAATTTAGACTTAG------------------------  1214
Sbjct  372864  TTGATGGAGCTCAAGTCCCTTCTAATTTAGACTTAGATGTACAAAATTTAAATGTTGATT  372923

Query  1215    ---ATGTTTTTTCTGCTCATAAAATGTATGGACCTACTGGAATTGGTATATTATATGTaa  1271
Sbjct  372924  TTTATGTTTTTTCTGCTCATAAAATGTATGGACCTACTGGAATTGGTATATTATATGGAA  372983

Query  1272    aaaaaaaaa-aaTAGAAAAATTATATCCTTATCAATTTGGAGGGGAAATGATTAAAAATG  1330
Sbjct  372984  AAAAAAAAATATTAGAAAAATTATATCCTTATCAATTTGGAGGGGAAATGATTAAAAATG  373043

Query  1331    TGAGTTTTGATAAAGCAACTTACTCAGATATATCGTTTAATTTTTAGGCAGGAACTCCAA  1390
Sbjct  373044  TGAGTTTTGATAAAACAACTTACTCAGATATACCGTTTAAATTTGAGGCAGGAACTCCGA  373103

Query  1391    ACATAGAAGGAATT  1404
Sbjct  373104  ATATAGAAGGAATT  373117
```

Range 3: 306075 to 306279

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 337 bits(182) | 2e-91() | 198/205(97%) | 3/205(1%) | Plus/Plus | |

```
Query  1472    TTGTGTTTCTACTTTTTTAAAAA---CTTCATTTTTATATACTATAGGAAAATTTTTCCA  1528
Sbjct  306075  TTGTGTTTCTACTTTTTTAAAAATTTCTTCATTTTTATATACTATAGGAAAATTTTTCCA  306134

Query  1529    TTCTTCTTCTAAAAAAGAACTCAATACGTTCCATTTAATATTACTTGCCTCATTATATTT  1588
Sbjct  306135  TTCTTCTTCTAAAAAAGAACTCAATATGTTCCATTTAATATTACTTGCCTCATTATATTT  306194

Query  1589    TACATGAAGAATATTTTAAttttttttttccatatttttacatcatcattactaataat  1648
Sbjct  306195  TACATGAAGAATATTTTCATATTTTTTTTCCATATTTTTTACATCATCATTACTAATAAT  306254

Query  1649    ttcttctttttccaaatttttttcTG  1673
Sbjct  306255  TTCTTCTTTTTCCAATTTTTTTCTG  306279
```

Range 4: 596310 to 596473

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 259 bits(140) | 4e-68() | 157/165(95%) | 2/165(1%) | Plus/Minus | |

```
Query  1       AACGGCTTTCACATGCTTGCAATAATAATTCTCCGATTATTAttttttttttCTTGAAGCC  60
Sbjct  596473  AACGACTTTCACATGCTTGCAATAATAATTCTCCGGTTATTCCTTTTTTTT-TTGAAGCC  596415
Query  61      TCAAGAAATAAACTTTCAAATTGTCATTCCAATATTCCATAAGT-TATCTTGCTTTTTGT  119
Sbjct  596414  TCAAGAAATAAACTTTCAAATTGTCGTTCCAATATTCCATAAGTATATCTTGCTTTTTGT  596355
Query  120     TTTTCTATCAATTGAATAAAATATTCTGAACGTTTTCCTCTTCGA  164
Sbjct  596354  TTTTCTATCAATTGAATAAAATATTCTGAACGTTTTCCTCTACGA  596310
```

Range 5: 181204 to 181348

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 235 bits(127) | 7e-61() | 139/145(96%) | 0/145(0%) | Plus/Minus | |

```
Query  722     TGTATTTCATTTTTAGTAGACCCTGATCCTCCATGAAAAACTAAAGACACTGGttttttt  781
Sbjct  181348  TGTATTTCATTTTTAGTAGAACCTGATCCTCCATGAAAAACTAAAGACACTGGTTTTTTT  181289
Query  782     GTCTTAGTATGAAATTTGTTTTGTATATATTCTTGTGTATTTTTTAAAATTTCGGGATGA  841
Sbjct  181288  GTCTTAGTATGAAATTTTTTTTGTATATATTCTTGTGTATTTTTTAAAATTTCTGGACGA  181229
Query  842     AGCATAACATTTCCAGGTTTataaa  866
Sbjct  181228  AGCATCACATTACCAGGTTTATAAA  181204
```

Range 6: 160006 to 160092

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 139 bits(75) | 6e-32() | 83/87(95%) | 0/87(0%) | Plus/Minus | |

```
Query  637     ATATTCAAAATCCAATTCAGGACAAGATTCTAGTTTATATAAAACAGATATGGATTTTCT  696
Sbjct  160092  ATATTGAAAATCCAATTCAGGACAAGATTCTAGTTTATATAAAACAGATATGGATTTTTT  160033
Query  697     TACTATTTTTTCTGTAGATTGAAAATG  723
Sbjct  160032  TACTATTTCTTCTATAGATTGAAAATG  160006
```

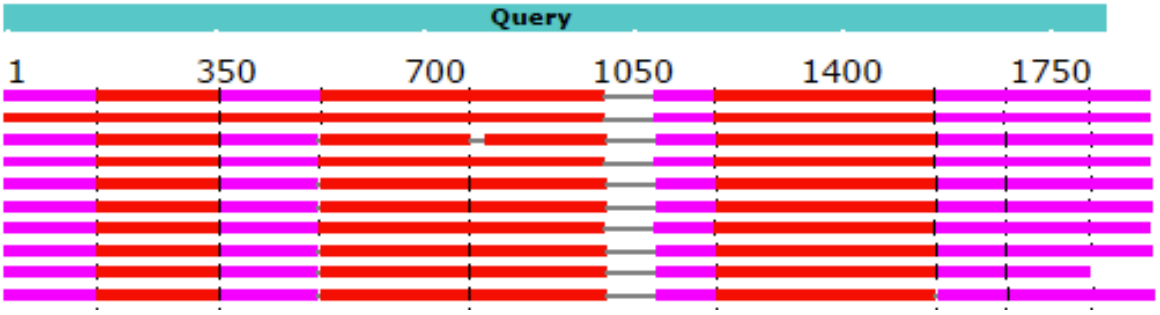
Range 7: 350656 to 350738

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 106 bits(57) | 6e-22() | 76/85(89%) | 2/85(2%) | Plus/Plus | |

```
Query  1405    GTTTTTATTCAAGATCCAAAACTTGGATaaaaaaaaTTAAAATTTTTTGGGTAAATATGT  1464
Sbjct  350656  GTTTTTATTCAAGATCCAAAACTTGGATAAAAAAAATATAAAA-TTTTTGGATAAATATGT  350714
Query  1465    AGAAATCTTGTGTTTCTACTTTTTT  1489
Sbjct  350715  AGAAATCTTTTATA-CTACTTGTTT  350738
```

Additional chimeric sequence 5:

*Panesthia lata* contig_1558: 4419266-4421106 (1840 bp)




Blattabacterium cuenoti BPAY DNA, complete genome, strain: BPAY
Sequence ID: **AP014609.1** Length: 632370 Number of Matches: 10
Range 1: 51691 to 52053

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 543 bits(601) | 9e-149() | 338/363(93%) | 0/363(0%) | Plus/Plus | |

```
Query  1141   CTCAAGTAATAGGAATTTGGAACCTTGTTTCTATAGAATTTATACGAAAATCAGATGAAA  1200
              ||||||||||||| |||||||||| |||||||| ||||||||||||||||||||||| ||
Sbjct  51691  CTCAAGTAATAGAAATTTGGAATCTTGTTTTTATAGAATTTATACGAAAATCAGATGGAA  51750

Query  1201   CATTGGAAAAACTTTATACAAAACATTTAGATACAGGAATGGGTTTGAAAAGATTATGCA  1260
              |||||||||||||| | |||||||||| |||||||||||||||| ||| ||||||||||||
Sbjct  51751  CATTGGAAAAACTTGACACAAAACATGTAGATACAGGAATGGGATTGGAAAGATTATGCA  51810

Query  1261   TGGTTTTACAAGGAAAAATATCTAGTTATGAAACTGATAATTTTTATCCTATTATTCAAT  1320
              |||| ||||||| |||||| ||||||||||||||||||||| ||||||||||||||||||
Sbjct  51811  TGGTATTACAAGAAAAAATTTCTAGTTATGAAACTGATATTTTTTATCCTATTATTCAAG  51870

Query  1321   ATATAGAAGACGATATAGGAAATGTTTATAATATAAAGGATTTTAATCAACATGTATCTA  1380
              |||||||| ||||| |||||||||||||||||| ||| |||||||||| |||||||||
Sbjct  51871  ATATAGAATACGATTTAGGAAATGTTTATAATAGAAAAGATTTTAATCAAAATGTATCTA  51930

Query  1381   TACGGATTATTGCAGATCATCTAAGAGCTATTGTTATTTCTATTTTAGATGGACAATTAC  1440
              |||||||||| |||||||||||||||||||||||||||| |||||||||||||||||||
Sbjct  51931  TACGGATTATAGCAGATCATCTAAGAGCTATTGTTTTTTCTATTTTAGATGGACAATTAC  51990

Query  1441   CATCCAATAACGGAGCTGGTTATGCGATAAGAAGGATACTTAGAAGAGCCCTTATTTACT  1500
              |||||||||||||||||||||||||| |||||||| |||||||||||||| ||||||| ||
Sbjct  51991  CATCCAATAACGGAGCTGGTTATGTGATAAGAAGAATACTTAGAAGAGCCGTTATTTTCT  52050

Query  1501   GTA  1503
              |||
Sbjct  52051  GTA  52053
```

Range 2: 407510 to 407734

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 288 bits(319) | 3e-72() | 201/225(89%) | 4/225(1%) | Plus/Plus | |

```
Query  743     aaaaatcttaacaatttaaaaaaaaaT--AGtttttggagcatttaattttaaatttttct  800
Sbjct  407510  AAAGATTTTAATAATTTTAAAAAAAATTTAGTTTTTGGAGCATTTAATTTTAAGTTTTCT  407569
Query  801     ataattaaaattgtattatatactaatttttgttctataagaaatttttctgactatattt  860
Sbjct  407570  ATAATTAAAATTTTATTACGTACTAATTTTTGTTCTATGAGAAATTTTTCTGACTATATTT  407629
Query  861     tttgtacgtttatttaattttattaaatatttttCTTGGTTAAGGACCAAAG-CTCTTCTC  919
Sbjct  407630  TTAGTACGTTTGTTTAATTTTATTAAATACTTTCTTGGTTTAGGACCAAAAACTCTTCCC  407689
Query  920     CCTCGTCTAAAAATA-GAATTCTTATGTCCCCTTTTCTGGAGCCT  963
Sbjct  407690  CCTCCTCTAAAAATAGGATTTTTTATGTTCCCTTTTCTAGAGCCT  407734
```

Range 3: 449834 to 450074

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 283 bits(313) | 1e-70() | 214/247(87%) | 14/247(5%) | Plus/Plus | |

```
Query  505     ataaaatataaaatattccctttatgaaaaatttttttGCTTTGGCAAACAGAATTCCTTC  564
Sbjct  449834  ATAAAATGTAAAATACTTCCTTTATGAAAAATTTTT-GCTTTGGCAAACAGAATTCCTTT  449892
Query  565     T-TTATGTTTATGGATCGAATATGATTTGTAGAAATTTCAGTACTAAAT-CATTAA----  618
Sbjct  449893  TCTTATTTTTATGGATCGAATATGATTTGTAGAAATTTCAATGCTAAAAACATTAAAATT  449952
Query  619     -tttttttGTTCATTTTTAATATATAACTTATaaaaaaaaTAGAAATTTCAATACTTTCA  677
Sbjct  449953  GTTTTTTTTTCATTTATAATAT-----TTATAAAAGAAATAGAACTTCCAACACTTTCA  450007
Query  678     GCTAAAATTATAATTGCTCCTCCATGtaaaaagtgaaaa-gtttaaatattttagaattt  736
Sbjct  450008  GCTAAAATTATAGTTGCTCCTCCATGTAAAAAACCAAAAGGTTGAAATATTTTAGAATTT  450067
Query  737     ataggca  743
Sbjct  450068  ATAGGCA  450074
```

Range 4: 539533 to 539742

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 270 bits(299) | 8e-67() | 185/210(88%) | 9/210(4%) | Plus/Minus | |

```
Query  143     TTGTGGTTTTATAGCTTTATGAGGTGTAATCACAATTTCCAAAAATTCATTTTCTTTGAT  202
Sbjct  539742  TTGTGTTTTTATAGCTTTATGAGCTTTAATCACAGTTTCCAAAAATTCATTTTCTTTGAT  539683
Query  203     TTCCTTAATTTCTCCTTATATCATACTAATAGAA---------GCTCCCAGTATAAAATC  253
Sbjct  539682  TTCCTTCATTTCTCCTTCTATCATAATAATAGAATGATTTGAAGCTCCGACTATTAAATC  539623
Query  254     TATATCCGCTTTCTTTAACTGGTCTAAACTAGGATTAATAATGAATTTTCCATTTAAACG  313
Sbjct  539622  TATATCCGATTTTTTTAACTGATCTAAACTAGGATTAATAATGAATTTTCCATTTAAACG  539563
Query  314     GATAATATGTATTTCTGACATAGGAGCATT  343
Sbjct  539562  GATAATACGTATTTCTGATATAGGACCATT  539533
```

Range 5: 478954 to 479095

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 194 bits(214) | 5e-44() | 128/142(90%) | 0/142(0%) | Plus/Minus | |

```
Query  1       ATATCATAAACAttttttaattttgtttttATAACAGTAAGCATTACATTATGATTACAT  60
Sbjct  479095  ATATCATAAATATTTTTTAGTTTTGTTTTTATAAAAGTACGCATTACATTATGATTGCAT  479036
Query  61      AATTCTTGTACTATTTGTTGGCGTGATTTTATGTGAAGTAAATTTTGTAAAGGAAAAGGA  120
Sbjct  479035  AATTCTTGTACTATTTGATGACGTGTTTTTATATAAAATAAATTTTTTAAAGGAAAAAGA  478976
Query  121     ATCCAATTTTTTAATAATCTTC  142
Sbjct  478975  ATCCAATTTTTTAATAACCTTC  478954
```

Range 6: 575018 to 575196

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 183 bits(202) | 1e-40() | 150/179(84%) | 14/179(7%) | Plus/Plus | |

```
Query  342     TTTACAATAGAATAATTTTTTGAAACTATTTCTGGACTAAAAATTTTAGATCACCATTC-  400
Sbjct  575018  TTTACGATAGAAAAATTTTTTGAAACTATTTCTGGACTAAAAATTTTAGATCATCATTCA  575077
Query  401     --ACTGCttta-----ttttccttttttaaaaaagatgattt-----atagaaaatttcac  449
Sbjct  575078  ATACTGCTTTACTTTTTTTATTTTTTAAAAAAAGATGATTTTATGGAAAAAAAATATCAC  575137
Query  450     ttttttttttAATT-GGGGACTAAAATATT-AATttttttcagaatatagattttattat  506
Sbjct  575138  GATTTTTTTAATTGGGGGCCTAAAATATTAAATGATTTTCAGAATATAGATTTTAATAT  575196
```

Range 7: 229715 to 229830

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 170 bits(188) | 6e-37() | 108/116(93%) | 1/116(0%) | Plus/Plus | |

```
Query  1726    TTGCATCTTTTTAAAAACTTATATTCTCCAGACTGATAAAATTGATTATAATATCCCATA  1785
Sbjct  229715  TTGCATCTTTTTAAAAATTTATCTTCTCCAAACTGATAAAATTGATTATAATATCCCATA  229774
Query  1786    AGAATAATAGGAAC-TTTATttttttCTTAATGTTTTCGATTTTATTAAATAATAA  1840
Sbjct  229775  AGAATAATAGGAACTTTTATTTTTTCCTTAATTTTTTCAATTTGATTAAATAATAA  229830
```

Range 8: 550189 to 550326

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 165 bits(182) | 3e-35() | 123/141(87%) | 5/141(3%) | Plus/Plus | |

```
Query  1605    CAGttttttccttttttttttaaaaaaaaaaTTTAAATTTAGAATCAA-ATCAATTTGTTGT  1663
Sbjct  550189  CAGTTTTTCCTTTTTTT---AAAGAAAAAATTTAAATTTTGAATCAATATAAATTTGTTGT  550245
Query  1664    AAGTAAATGAGTCATCC-GAAATAAAATTTCATTTTCAAAATGAAAAATAGATTTATATG  1722
Sbjct  550246  AAGCAAATGATACATCCCGAAATAAAATTTCATTTTCAAAATGAAAAATAGATTTATATC  550305
Query  1723    TAATTGCATCTTTTTAAAAAC  1743
Sbjct  550306  TCATTGTTTCATTTGAAACAC  550326
```

Range 9: 200638 to 200745

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 149 bits(164) | 2e-30() | 99/109(91%) | 1/109(0%) | Plus/Minus | |

```
Query  1503    AAGAAATGGGAATCAGGAGGTCCTTTTATTGGAAATACTTGGGaaaaaaaTGAATCAATG  1562
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200745  AAAAAATGGGAATTAGGAGATCCTGTTATTCGAAACACTTGG-AAAAAAATGAATCAATG  200687

Query  1563    GGTTTATAATGGGTTTGAAGAAACTTATAAAAATTTGGGAATCAGtttt  1611
               ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200686  GGTTTATAATGGATTTGAAGAAACTTATAAAAAATTGGGAATTAGTTTT  200638
```

Range 10: 165748 to 165849

| Score | Expect | Identities | Gaps | Strand | Frame |
|---|---|---|---|---|---|
| 113 bits(124) | 1e-19() | 86/102(84%) | 10/102(9%) | Plus/Minus | |

```
Query  1049    aacacattttgaaatttgaatttattagatgattttttt----AATCTTTT-------TGT  1098
               ||||||||||||||||||||||||||||||||||||||||    ||||||||       |||
Sbjct  165849  AATACATTTTTAAATTTGAATTTATTAGATAATTTTTTTTCTAAATCTTTTAATTTTTGT  165790

Query  1099    AAATCTCCGTTAAATAATTCTTTGAAACTATCTGCTGAGCCT  1140
               |||||||| ||||||||||||||| ||||||||||||||| |||
Sbjct  165789  AAATCTCCATTAAATAATTCTTTAAAACTATCTGCTGAACCT  165748
```