

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

2022

THE EVOLUTION OF MORPHOLOGICAL DIVERSITY AND SEXUAL DIMORPHISM IN STICK AND LEAF INSECTS

Romain Pascal Gilbert Eric Boisseau

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Boisseau, Romain Pascal Gilbert Eric, "THE EVOLUTION OF MORPHOLOGICAL DIVERSITY AND SEXUAL DIMORPHISM IN STICK AND LEAF INSECTS" (2022). *Graduate Student Theses, Dissertations, & Professional Papers*. 12008.

<https://scholarworks.umt.edu/etd/12008>

This Dissertation is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

THE EVOLUTION OF MORPHOLOGICAL DIVERSITY AND SEXUAL
DIMORPHISM IN STICK AND LEAF INSECTS

By

ROMAIN PASCAL GILBERT ERIC BOISSEAU

B.S. Biological Sciences, Sorbonne Université, Paris, France, 2013
M.S. Ecology, Biodiversity and Evolution, Sorbonne Université, Paris, France, 2016
ENS graduate degree, École Normale Supérieure, Paris, France, 2016

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in Ecology and Evolution

The University of Montana
Missoula, MT

December 2022

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Dr. Douglas J. Emlen, Chair
Division of Biological Sciences

Dr. H. Arthur Woods
Division of Biological Sciences

Dr. Bret W. Tobalske
Division of Biological Sciences

Dr. Erick Greene
Division of Biological Sciences

Dr. Sven Bradler
Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute
of Zoology and Anthropology, University of Göttingen, Göttingen, Germany

© COPYRIGHT

by

Romain Pascal Gilbert Eric Boisseau

2022

All Rights Reserved

The evolution of morphological diversity and sexual dimorphism in stick and leaf insects

Chairperson: Dr. Douglas J. Emlen

Abstract

How has the diversity of life forms come to be? This question is at the core of evolutionary biology and can be addressed at different scales: by studying the processes that drive modifications within populations of organisms generation after generation (microevolution), or by investigating patterns of changes on the tree of life over long periods of time (macroevolution). Understanding the ultimate drivers of morphological diversity eventually entails connecting microevolutionary processes with macroevolutionary patterns. My dissertation investigates the diversification of body and egg form and its drivers in a relatively small but particularly diverse insect order: the stick and leaf insects (Phasmatodea).

As masters of camouflage, the 3,400 described species of phasmids are an ideal system to study morphological evolution as they vary tremendously in body morphology, going from long slender branch mimics to wide, flat animals that look exactly like leaves. This remarkable diversity of forms enables phasmids to avoid detection by visually-hunting predators. Even their remarkably diverse hard-shelled eggs resemble a wide variety of plant seeds. In addition, males and females of the same species often look very different from each other, with females in extreme cases more than ten times the size of the males.

In chapters one, two and three, I investigate the patterns of variation of female body morphology, sexual dimorphism and egg morphology respectively, and potential ecological, life history and biomechanical correlates in a phylogenetic context. I describe repeated convergence towards multiple body forms associated with habitat transitions but find substantial variation in the strength of convergence and underlying evolutionary paths. Then, I show that variation in the extent of sexual dimorphism is best explained by variation in selective pressures acting on males, namely locomotor (flight) performance and male competition (sexual selection). Finally, I show that variation in egg size and shape is driven by variation in life history strategies, mechanical constraints and oviposition strategy.

In chapters four, five and six, I investigate the microevolutionary processes behind the primary macroevolutionary forces driving variation in sexual dimorphism. In chapter four, I show in leaf insects (*Phyllium philippinicum*) that larger males are poor flyers, suggesting that selection for flight performance favors smaller male body sizes in this species, and reinforcing the broader taxonomic findings of chapter two. In chapters five and six, I describe how a change in the mating system of thorny devil stick insects (*Eurycantha calcarata*) switched the direction of sexual selection and led to the evolution of exceptionally large male body sizes and exaggerated hindleg weapons, confirming the pervasive role of sexual selection in driving variation in male size and sexual dimorphism. Collectively, my research contributes to our understanding of the forces that shape the evolution of morphology in animals and their eggs.

Résumé

Comment a été engendrée la diversité des formes du vivant? La réponse à cette question fondamentale en biologie évolutive peut être étudiée à différentes échelles: en se concentrant sur les processus à l'origine des modifications au sein des populations d'organismes génération après génération (microévolution), ou en recherchant les grandes tendances de changement dans l'arbre de la vie sur de longues périodes de temps (macroévolution). Caractériser les facteurs à l'origine de la diversité morphologique implique donc de réconcilier les processus microévolutifs et les tendances macroévolutives. Ainsi, cette thèse de doctorat s'intéresse aux origines de la diversification des formes de corps et d'œufs chez un ordre d'insectes relativement petit en nombre d'espèces mais particulièrement diversifié du point de vue morphologique: les phasmes (Phasmatodea).

Les 3400 espèces de phasmes décrites forment un modèle idéal pour étudier l'évolution de la morphologie de ces experts du camouflage. Leur apparence corporelle varie énormément, allant de formes allongées ressemblant à des brindilles à des formes larges et aplaties imitant à la perfection celles de feuilles. Cette diversité morphologique remarquable permet aux phasmes de passer inaperçus auprès de leurs prédateurs visuels. Même leurs œufs à coquille dure ressemblent à des graines diverses. Enfin, les mâles et les femelles de la même espèce ont souvent des apparences très éloignées, la femelle atteignant parfois plus de dix fois la taille du mâle.

Dans les chapitres un, deux et trois, je me focalise sur les grandes tendances de diversification de, respectivement, la morphologie des femelles, du dimorphisme sexuel et de la morphologie des œufs, ainsi qu'aux potentiels facteurs écologiques, biomécaniques et d'histoire de vie à leur origine dans un contexte phylogénétique. Concernant les femelles, je décris la convergence répétée de leur morphologie vers un nombre restreint de morphotypes associés à des changements d'habitat. En revanche, j'observe que les trajectoires évolutives vers les formes convergentes et la force de cette convergence varient considérablement entre les lignées. Ensuite, je montre que la variation du degré de différence entre mâles et femelles s'explique principalement par la variation de pressions de sélection s'appliquant spécifiquement aux mâles, en particulier la performance locomotrice en vol et la compétition entre mâles (sélection sexuelle). Enfin, je conclus que les variations de comportements de ponte, de traits d'histoire de vie, et de contraintes mécaniques sont des facteurs moteurs de la diversification de la taille et de la forme des œufs.

Dans les chapitres quatre, cinq et six, je me tourne vers les processus microévolutifs sous-jacents des facteurs principaux expliquant la variation du dimorphisme sexuel entre espèces. Dans le chapitre quatre, je démontre que, chez les phasmes feuille (*Phyllium philippinicum*), les mâles les plus lourds ont une piètre performance en vol suggérant que la sélection sur la performance en vol favorise les mâles plus petits chez cette espèce, et renforçant les résultats à grande échelle du chapitre deux. Dans les chapitres cinq et six, je caractérise comment un changement dans le système d'appariement des phasmes cuir (*Eurycantha calcarata*) a inversé le sens de la sélection sexuelle et a conduit à l'évolution de mâles exceptionnellement imposants et armés de pattes postérieures redoutables. Ceci confirme le rôle prépondérant de la sélection sexuelle comme moteur de la diversification de la taille des mâles et du dimorphisme sexuel. L'ensemble de mes résultats contribue ainsi à une meilleure compréhension des facteurs qui façonnent l'évolution de la morphologie des animaux et de leurs œufs.

Acknowledgements / Remerciements

This work would have never been possible without the love and support of the many amazing people that I have had the chance to cross path with while working on this dissertation and throughout my life.

My first thank you goes to **Doug Emlen**, my Ph.D. advisor who, throughout the years, went from my mentor, to my role model, my friend and (he will hate it) my second father. Doug, thank you for your trust since the first day, and saying “yes” to hosting a random French student in your lab eight years ago. Your flawless support has made me not only go through graduate school but enjoy every bit of my time in Missoula. I always felt like my scientific success was as important to you as my personal life. Thank you for encouraging me to spend time away from the lab with my loved ones for months while you were taking care of my animals. I feel immensely grateful to have shared these years beside you. You pushed me out of comfort zone many times, but wow was it worth it! You pushed me to start my new system, to pursue my own ideas, to go see my critters in New Guinea and, in the end, you simply made my childhood dream of “becoming a phasmatologist” a reality. Thank you for all the time editing my writing, chatting about science or just life. Thank you for your kindness and sometimes for your recklessness and finally thank you for your unconditional support of my relationship with Camille.

Thank you **Art Woods** for our friendship and for your flawless support and excitement about any harebrained (=farfelu) project that could get us to work together. I mean, who would think carrying a bomb-looking expensive device across the ocean would be a good idea? Not the French customs agents! You have been a guardian during all my time in Missoula and I always felt I could turn to you about anything. Ca va me faire chier de plus te voir et de ne plus lancer des palets pour s’exclamer que c’est de la merde de taureau. Lame !

Thank you **Bret Tobalske** for opening wide the doors of your lab and introducing me to the world of biomechanics since the very first day. Merci pour ton enthousiasme pour tous les projets que je t’ai proposés et ton soutien sans faille. Merci pour ces conversations scientifiques en Français qui m’ont permis de ne pas trop perdre la main ! Je suis très reconnaissant de t’avoir en ami et je suis sûr que notre amour des montagnes et de l’épicurisme à la française nous réuniront bientôt.

Thank you **Sven Bradler**, the master of stick insects, what I am saying? The KING of stick insects! Thank you for sharing your resources, your phasmid collection and your love of beer with me! I am so grateful for your kindness, that you took me under your wing and that you happily shared your system with me. I can’t wait for you to honor your bet and offer me a night of beer. Onto many more years of stick insect research and collaboration.

And finally, thank you **Erick Greene**, my last francophile committee member. Thank you for your constant smile, your kindness and your “bonne humeur” very Québécoise. It is always a joy to chat science or life with you. Thank you for accepting (and actually volunteering, if I remember well) to be on my committee. I hope you will visit me in Europe, we have cool birds too!

Thank you to all my other co-authors and collaborators, especially **Thies Büscher** for embarking on a remote collaborative project with me. I can’t wait to meet you in person. Thank you to all the people that made my trip to New Guinea possible and unforgettable: **Luc** and **Katie** (et Jannelle!) **Bonneau** (merci pour le coup de la panne en bateau!); **Mark Ero**, **Simon Makai**, **Richard Dikrey**, **Marnie Light**, **Banabas Sapau**, **Pal Mana**, **Seset Komda**, **Gibson Gumbira**, **Tabitha Manjobie**, **Thom Batari** and everyone at Dami and Higaturu. Thank you so much for your warm welcome.

Thank you to the Emlen lab past and present for making every day in the lab a great one whether we had seemingly meaningless or profound interactions. **Devin O'Brien** for training me how to pronounce “Doug” properly... and for the many friendship moments that culminated that night in Paris (and I guess for the sciency stuff too...). **Cerisse Allen**, for keeping this lab running and not go bankrupt and for your kindness and availability. Thank you **Jill del Sol**, **Chelsey Ponsness**, **Jesse Weber**, **Jema Rushe**, **Sophie Fitzgerald**, **Wenfei Tong**, **Nicole Lopez** for all the scientific and non-scientific moments shared. And thank you to all the undergrads that I got to “mentor” over the years, with special thanks to **Nate Barton**, **Makenna Carlsberg**, **Lexi Klawitter** and **Madeline Kleemann**. You guys rock!

Thank you to the EE Faculty, postdocs and students for providing such a supportive and thrilling scientific environment. I have learned so much by watching amazing talks, chatting and sharing beers with all of you over the years. Thank you to the DBS office staff and especially **Jill Burke**, **Ruth Johnson** and **Janean Clark** who have always done their best to make my life easier. Thank you to the international office and especially **Bodhi Murphy** and **Caroline McLean** for being so accommodating and friendly. Thank you to all the custodians and especially **Melissa** who cleaned our lab everyday and kept me company when I was working late.

I also want to thank my previous mentors and advisors who made my love for science grow. I especially want to thank **Marlène Goubault**, **Audrey Dussutour**, **Kate Barry**, **Shawn Wilder** and **Mónica Arias** for their trust and support (past and/or present).

Then I want to thank all my Missoula friends for the support, laughter and unforgettable times spent together. First my cohort (aka the “boy float” gang & close associates): **Tony** (my one and true love), **Megan** (who always makes sure no man’s left behind), **Victoria** (bad touch queen), **Kory** (bathroom selfie fan), **Luke** (HA HA HAHAAAAHA), **Valery** (game night queen), **Sam** (boxed wine lover), **Emily K** (Oula peer pressurer), **Peter** (Uuup! Dooown! Up! Down! Up! Down! Shoulder! Shoulder! Hip! Hip!), **Cole** (not here to suck ice), **Emily M** (cocktail goddess), **Joe** (do not join his sampling trips, it’s a trap!), **Julia** (how do you keep on going with Joe?). And all the others: **Gerard** (Extreme G racing pro), **Devin** (aspiring Sip 'n Dip merman), **Jill** (party queen), **Fernando** (you forgot your cane), Steve Lane (squarespace pro), **Joely** (you have interesting friends), **Rachel** (cheese butter lover), **Sophie** (Cow girl candidate), **Chelsey** (trash can lover), **Lars** (“oh yeah, I’ll get you a fish”), **Jake** (door smasher), **Nathanael** (beer can smasher), **Adam** (ka-muhm-behrt), **William** (Anjou RPZ), **Zooey** (soccer badass), **Yana** (let’s go to Bodega!), **Hyeok** (roomie!!).

Merci à mes ami(e)s Français(es) qui malgré la distance et le temps ont toujours été (et sont toujours) là pour moi. Je fais un big up aux potos de prépa (aka la Mifa) en particulier **Doumane**, **Grissot**, **Gourtay**, **Toto**, **Coco**, **Valou**, **Gloggy**. Je salue également la fine équipe **Alexis**, **Hugo**, **David** et les autres potos d’Ulm **Bertrand**, **Guillaume**, **Lucie**, **Anne**, **Adam**, **Loulou**, **Queue**. Les potes de lycée, **Matthieu**, **Guillaume**, **Célestine**, **Maeva**, **Marie Eve**, **Laura** et **Maxime**. Mon grand ami d’enfance **Pierre-Antoine**, avec qui ma passion des insectes et particulièrement des phasmes a commencé. Et je finirai par une mention spéciale pour tous les amis qui sont venus jusque dans le Montana pour me voir en bravant leur peur des ours: **Matthieu**, **Guillaume**, **Célestine**, **Thomas**, **David**, **Marianne**, **Cyril**, **Valentin**, **Lucas**, **Bertrand**, **Alexis**, **Alison**.

Je remercie également ma famille et tout particulièrement mes parents et ma sœur qui m’ont toujours soutenu dans tous mes choix, que ce soit en m’autorisant à transformer ma chambre en véritable zoo et paradis pour arthropode en tout genre, en félicitant ma réussite scolaire ou en acceptant

ma décision de partir pendant ce qui semblait être une éternité à l'autre bout du monde pour faire cette thèse. Merci **Papa** et **Maman** pour tout l'amour et le soutien que vous m'avez donné, j'ai eu la chance de grandir dans un foyer merveilleux. Merci **Mathilde** pour nos chamailleries de frère et sœur. Merci **Papi**, **Mamie Solange** et **Mamie Jacqueline** (ou Titour) pour tout votre amour, d'être mes premiers fans, et de veiller sur moi à chaque instant. Merci à ma famille plus étendue et amis de la famille, les **Boisseau**, les **Garry**, les **Cantin**, les **Lemaire**, les « **voisins** », mon parrain **Pascal** et ma marraine **Thérèse**. Enfin merci à ma belle-famille de m'avoir chaleureusement accueilli parmi eux.

Mais si il y a une personne que je dois remercier et sans qui cette thèse, si elle avait existé, n'aurait été que l'ombre d'elle-même, c'est bien entendu **Camille**. Il fallait être sacrément zinzins pour se dire que la distance, le décalage horaire, les connexions internet pourries pendant quatre ans valaient bien la peine. Ma vie est plus belle que jamais depuis que tu la partages. Merci pour notre complicité. Je me réjouis d'affronter l'avenir avec toi.

Finally, I want to thank the little dark angel that dashed into my life on a cold autumn night and held me company through my sleepless writing nights: **Spooky** the abandoned black cat and soon to be world traveler and stuck-up French doctor.

Table of contents

Abstract	iii
Résumé	iv
Acknowledgements / Remerciements	v
Table of contents	viii
Introduction	1
Chapter 1	9
Global convergence towards multiple ecomorphs in stick and leaf insects	
Chapter 2	80
Sexual selection and flight as predictors of sexual size and shape dimorphism in stick and leaf insects	
Chapter 3	121
Physiology, mechanical constraints, and ecology drive the diversification of stick and leaf insect eggs	
Chapter 4	160
Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect	
Chapter 5	206
Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects	
Chapter 6	245
The function of male exaggerated weaponry in a gregarious stick insect	

Introduction

One of the most intriguing challenges in evolutionary biology is to understand the origin and maintenance of morphological diversity (Carroll, 2005; Darwin, 1859; Gould, 1989). Take one maple, one oak and one raspberry leaf: their morphological differences are obvious: size, silhouette, texture, edges, color... But it is also evident that they share a similar layout, descended from an ancestral template that would have been over- or downsized, stretched, twisted or bent in multiple ways. This pattern is also true for all living things, with varying degrees of homologies and discrepancies (Thompson, 1917). How do we explain the similarities and differences between living organisms?

One way to address this question is to focus on broad patterns of change as they unfolded on the tree of life over extended periods of time (macroevolution). How are lineages evolving over time? How fast? Are specific morphological changes consistently associated with changes in ecology, life history or behavior? Such broad comparative approaches can identify driving factors associated with morphological diversification at large spatial and temporal scales. For instance, one could identify that independent plant lineages colonizing more arid habitats repeatedly evolve smaller leaves (Wright et al., 2017), leading to the prediction that dry environments select for smaller leaves.

Another way to approach this question is to study the processes that drive modifications within populations of organisms generation after generation, therefore at a much smaller spatial and temporal scale (microevolution). How does the distribution in phenotypic values change from one generation to the next? How does a particular phenotype affect fitness? Such studies can identify what specific selective pressures are currently acting on a phenotype and how they fluctuate over time. If we reconsider the example of the evolution of leaf size, one could test the prediction of macroevolutionary studies and show that within a species, individuals with smaller leaves are favored in dry but not in wet habitats (Dudley, 1996) and even go further into identifying the relevant performance metrics targeted by selection (e.g., water use efficiency). As both macroevolution and microevolution are fundamentally governed by the same processes, macroevolutionary patterns should largely reflect microevolutionary processes. This link is often implicitly assumed in evolutionary biology but the extent of the continuity between macro- and microevolution is still debated (Reznick and Ricklefs, 2009; Saupé and Myers, 2021; Simons, 2002). Fully understanding the ultimate drivers of morphological diversity therefore necessitates bridging macroevolutionary patterns with microevolutionary processes.

In this dissertation, I studied aspects of the morphological diversification of stick and leaf insects (order Phasmatodea) at both scales, leveraging recent efforts at reconstructing the phylogenetic

history of this poorly known clade. This group includes approximately 3,400 described species of large terrestrial herbivores (Bradler and Buckley, 2018). They are predominantly distributed in the tropics with few species living in temperate regions, which might explain why they have been largely understudied. Moreover, phasmids are usually nocturnal and show extreme forms of masquerade (i.e., imitating various inanimate objects such as twigs or leaves) or crypsis (i.e., background matching) making them especially challenging to observe and collect in the wild (Bedford, 1978). Plant mimicry is even brought to quasi-perfection through specialized behaviors such as catalepsy (i.e., stillness and death feigning) and wind swaying behaviors (Bedford, 1978; Bian et al., 2016). This impressive array of camouflage strategies is the main antipredator defense of most phasmids and has led to the evolution of an incredible and unique diversity of body sizes and morphologies ranging from elongated tubular bodies and thin legs of various sizes (twig mimicry) (Shelomi and Zeuss, 2017) to widened and flattened bodies with leaf-like extensions (Bank et al., 2021a) or bulky and robust body forms in ground-dwelling species (Bank et al., 2021b; Buckley et al., 2009).

In chapter one, I evaluate the diversity of body morphology and the extent of convergent evolution in stick and leaf insects in relation to habitat use. In many species, the advergence of phasmid morphology towards diverse plant forms led to the evolution of extreme morphological specializations such as extremely elongated body silhouettes or foliaceous lateral extensions on the abdomen and legs. What are the evolutionary implications of such camouflage specializations? Are the morphotypes exhibited by phasmids evolutionary one-offs or did they repeatedly evolve in independent lineages? How likely is it that two lineages evolve similar morphologies (Blount et al., 2018)? Recent phylogenetic studies have shown conflicting results with prior character-based taxonomic classifications suggesting a high degree of convergent evolution that might have misled taxonomists (Bradler et al., 2014; Bradler et al., 2015; Robertson et al., 2018; Simon et al., 2019). For instance, the traditionally recognized subfamily Eurycanthinae comprising flightless, robust and strongly armored species was showed to be highly polyphyletic, exemplifying a spectacular case of convergent evolution towards the so-called “tree lobster” morphotype (Buckley et al., 2009). However, no study yet has thoroughly quantified the extent of morphological convergence across phasmid species. Using morphological data on 212 species, I find repeated and widespread convergence towards 18 discrete morphotypes. Transitions between vegetation layers led lineages towards predictable regions of the morphospace, highlighting the deterministic nature of morphological evolution. Nevertheless, using detailed reconstructions of the evolutionary trajectories of convergent lineages on the phasmid morphospace, I show considerable

variation in the underlying evolutionary paths, and in the extent and strength of convergence between morphotypes, suggesting contingency and emphasizing nuances in the relative determinism of evolution.

In chapter two, I explore yet another aspect of the morphological diversity of phasmids: sexual dimorphism in body size and shape. Chapter one only considers females for simplicity as sexual dimorphism is very variable in stick insects with lineages where males closely resemble females in morphology, and lineages where those differences are extreme. So extreme that males and females of the same species were sometimes originally described as separate genera (Cumming et al., 2020). As is most common in arthropods, the mobile males are often relatively small, slender and fully winged and flight capable, while the sedentary females are usually flightless with shortened wings or no wings at all (Bank and Bradler, 2022). In this chapter, I investigate the drivers of the variation in sexual dimorphism in body size and shape in phasmids. Interspecific patterns of variation of sexual size dimorphism suggest that females are overall larger than males because of stronger historical positive selection on female size (likely related to selection for increased fecundity) rather than stronger negative selection on male size. However, variation around the average size difference between males and females is only explained by variation in selective pressures acting on male size specifically, namely the variation in flight capacity and intensity of male-male competition. This study goes against the traditional view that variation in female-biased sexual size dimorphism is primarily driven by variation in fecundity. Positive selection on female size accounts for females being, on average, 300% larger than males but it does not account for variation around this average in the extent of sexual dimorphism between species. Instead, this interspecific variation appears driven by variation in selective pressures acting on male size. Species with flight capable males are relatively more dimorphic (i.e., males smaller than expected) while species exhibiting direct male-male interactions and mate guarding are relatively less dimorphic (i.e., males larger than expected). These results suggest that selection for increased flight performance favors smaller male sizes while selection for increased fighting performance favors larger male sizes. I test these two hypotheses at the microevolutionary level in specific systems in chapters four, five and six.

The breathtaking morphological diversity of phasmids is not restricted to body morphology: even their remarkable hard-shelled eggs closely resemble plant seeds and display a stunning diversity in size and shape associated with a variety of egg-laying and dispersion strategies (Clark Sellick, 1988; Clark Sellick, 1997; Goldberg et al., 2015; O’Hanlon et al., 2020; Robertson et al., 2018). Some lineages of stick insects and plants even convergently adapted to dispersion by ants (myrmecochory). The eggs

of these lineages bear a capitulum resembling a seed elaiosome, a lipid rich appendage known to be an adaptation for dispersion by ants in plants (Stanton et al., 2015). In chapter three, I explore the physiological, ecological and biomechanical drivers of the morphological diversification of the egg capsule. Analogous to a very broad recent study on insect eggs (Church et al., 2019), I find that oviposition ecology plays a major role in the evolution of egg size and shape. Compared to species that simply drop their eggs to the ground (the ancestral state), species that glue or bury their eggs lay relatively larger eggs that in turn trade-off with a lower lifetime egg number. Females with slenderer abdomens (as is common in stick insects) lay relatively more elongated eggs, most likely allowing an easier passage of the eggs through the narrower oviduct. I also demonstrate that larger eggs display a disproportionately low metabolic rate and consequently develop more slowly, which sharply contrasts with the results of Church et al. (2019) obtained across several orders of insects. Finally, I show that flight capable females and females living in more temperate environments invest relatively less in egg production and consequently lay relatively smaller eggs. Overall, my results suggest that variation in life history, oviposition technique, mechanical constraints in the female body, flight capacity and macroclimate underlie the morphological diversification of stick insect eggs.

Chapters four, five and six focus on specific phasmid species on which I test some of the predictions derived from chapter two. Chapter two identifies flight capacity and male-male competition as the main predictors of variation in sexual dimorphism across phasmids, at the macroevolutionary level. These patterns therefore predict that flight performance should favor smaller male sizes and that intense male-male competition should favor larger male sizes at the microevolutionary level.

In chapter four, I use a species of leaf insects (*Phyllium philippinicum*) that exhibit a typical pattern of sexual dimorphism for a phasmid: males have much smaller and slenderer bodies and longer antennae than the wide leaf mimicking females (Hennemann et al., 2009). The mobile males are able to fly using their long hindwings to search for females while the sedentary females are completely flightless and can only use their long forewings to parachute if falling from high perches. I use a combination of empirical measures of flight performance and modelling of body aerodynamics to show that large body sizes impair the flight performance of males. In addition, large body sizes are detrimental to substrate adhesion and increase the risk of detaching from the substrates on which males walk or land while searching for females. Therefore, my results support the hypothesis that

smaller flying males benefit from an increased locomotor performance, and are consistent with the broad findings of chapter two.

In chapters five and six, I look into the striking exception of a lineage that displays a very reduced sexual dimorphism: the New Guinean thorny devil stick insects (*Eurycantha calcarata* and *horrida*). I set out to investigate why this lineage of tree lobsters evolved uniquely large males that, in addition, display enlarged hindlegs endowed with a sharp spine (Buckley et al., 2009; Carlberg, 1989; Hsiung, 1987). Using field observations in Papua New Guinea, I describe how, contrary to its sister species *E. insularis*, *E. calcarata* switched from solitary to communal roosting inside tree cavities. This shift in roosting behavior changed the distribution of females from scattered to clumped, increasing their monopolisability (chapter five). Male thorny devils use their hindlegs in fights against rivals for access to strategic positions on tree trunks close to cavity entrances enabling them to intercept and mate with females as they leave the cavities at dusk. Therefore, this system changed from a scramble competition to a defense-based polygyny mating system, which in turn greatly affected the evolution of male morphology.

In chapter six, I look at the details of the fighting behavior of *E. calcarata* paying special attention to the function of their exaggerated and sexually dimorphic hindlegs. I show that while males do appear to mutually assess each other during fights, they do not use their hindlegs for this assessment. The enlarged male hindlegs are not used to threaten rivals or to advertise fighting prowess, and instead function as pure tools of battle for grasping and squeezing opponents in rare escalated fights. Results from the two latter chapters therefore indicate that a change in the intensity of male-male competition and thus in the strength of sexual selection can lead to drastic changes in the pattern of sexual dimorphism of a lineage. In thorny devils, larger males are more likely to win fights against rivals and are consequently able to mate with more females. Once again, these results align with those of chapter two, which found a potent effect of changes in mating systems on male morphological evolution at the macroevolutionary scale.

Altogether, my findings further our understanding of the selective pressures that have shaped the evolution of morphology in animals and their eggs. By using diverse methods at different scales, this work demonstrates how macro- and microevolutionary studies can be integrated and used in concert to offer new insights into the evolution of the diversity of life.

References

- Bank, S. and Bradler, S.** (2022). A second view on the evolution of flight in stick and leaf insects (Phasmatodea). *BMC Ecology and Evolution* **22**, 1–17.
- Bank, S., Cumming, R. T., Li, Y., Henze, K., Le Tirant, S. and Bradler, S.** (2021a). A tree of leaves: phylogeny and historical biogeography of the leaf insects (Phasmatodea: Phylliidae). *Communications Biology* **4**, 1–12.
- Bank, S., Buckley, T. R., Büscher, T. H., Bresseel, J., Constant, J., de Haan, M., Dittmar, D., Dräger, H., Kahar, R. S., Kang, A., et al.** (2021b). Reconstructing the nonadaptive radiation of an ancient lineage of ground-dwelling stick insects (Phasmatodea: Heteropterygidae). *Systematic Entomology* **46**, 487–507.
- Bedford, G. O.** (1978). Biology and ecology of the Phasmatodea. *Annual Review of Entomology* **23**, 125–149.
- Bian, X., Elgar, M. A. and Peters, R. A.** (2016). The swaying behavior of *Extatosoma tiaratum*: Motion camouflage in a stick insect? *Behavioral Ecology* **27**, 83–92.
- Blount, Z. D., Lenski, R. E. and Losos, J. B.** (2018). Contingency and determinism in evolution: Replaying life’s tape. *Science* **362**, eaam5979.
- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.
- Bradler, S., Robertson, J. A. and Whiting, M. F.** (2014). A molecular phylogeny of Phasmatodea with emphasis on Necrosiinae, the most species-rich subfamily of stick insects. *Systematic Entomology* **39**, 205–222.
- Bradler, S., Cliquennois, N. and Buckley, T. R.** (2015). Single origin of the Mascarene stick insects: ancient radiation on sunken islands? *BMC Evolutionary Biology* **15**, 196.
- Buckley, T. R., Attanayake, D. and Bradler, S.** (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1055–1062.
- Carlberg, U.** (1989). Aspects of Defensive Behaviour of *Eurycantha calcarata* Lucas Females and the Evolution of Scorpion Mimicry in the Phasmida (Insecta). *Biologisches Zentralblatt* **108**, 257–262.
- Carroll, S. B.** (2005). *Endless forms most beautiful: the new science of evo devo and the making of the animal kingdom*. New York, New York, USA: WW Norton & Company.
- Church, S. H., Donoughe, S., de Medeiros, B. A. S. and Extavour, C. G.** (2019). Insect egg size and shape evolve with ecology but not developmental rate. *Nature* **571**, 58–62.
- Clark Sellick, J. T.** (1988). The capitula of phasmid eggs: an update with a review of the current state of phasmid ootaxonomy. *Zoological Journal of the Linnean Society* **93**, 273–282.
- Clark Sellick, J. T.** (1997). Descriptive terminology of the phasmid egg capsule, with an extended key to the phasmid genera based on egg structure. *Systematic Entomology* **22**, 97–122.
- Cumming, R. T., Tirant, S. Le, Teemsma, S. N., Hennemann, F. H., Willemse, L. and Büscher, T. H.** (2020). Lost lovers linked at long last: elusive female *Nanophyllum* mystery solved after a century of being placed in a different genus (Phasmatodea, Phylliidae). *ZooKeys* **969**, 43–84.
- Darwin, C.** (1859). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, UK: J. Murray.

- Dudley, S. A.** (1996). Differing selection on plant physiological traits in response to environmental water availability: a test of adaptive hypotheses. *Evolution* **50**, 92–102.
- Goldberg, J., Bresseel, J., Constant, J., Kneubühler, B., Leubner, F., Michalik, P. and Bradler, S.** (2015). Extreme convergence in egg-laying strategy across insect orders. *Scientific reports* **5**, 7825.
- Gould, S. J.** (1989). *Wonderful life: the Burgess Shale and the nature of history*. New York, New York, USA: W. W. Norton & Company.
- Hennemann, F. H., Conle, O. V., Gottardo, M. and Bresseel, J.** (2009). On certain species of the genus *Phyllium* Illiger, 1798, with proposals for an intra-generic systematization and the descriptions of five new species from the Philippines and Palawan (Phasmatodea: Phylliidae: Phylliinae: Phylliini). *Zootaxa* **2322**, 1–83.
- Hsiung, C.-C.** (1987). Aspects of the biology of the Melanesian stick-insect *Eurycantha calcarata* Lucas (Cheleutoptera: Phasmatidae). *Journal of Natural History* **21**, 1241–1258.
- O’Hanlon, J. C., Jones, B. R. and Bulbert, M. W.** (2020). The dynamic eggs of the Phasmatodea and their apparent convergence with plants. *The Science of Nature* **107**, 1–12.
- Reznick, D. N. and Ricklefs, R. E.** (2009). Darwin’s bridge between microevolution and macroevolution. *Nature* **457**, 837–842.
- Robertson, J. A., Bradler, S. and Whiting, M. F.** (2018). Evolution of oviposition techniques in stick and leaf insects (Phasmatodea). *Frontiers in Ecology and Evolution* **6**, 1–15.
- Saupe, E. E. and Myers, C. E.** (2021). Macroevolution. In *Evolutionary Developmental Biology* (ed. Nuño de la Rosa, L. and Müller, G. B.), pp. 149–167. Cham, Switzerland: Springer.
- Shelomi, M. and Zeuss, D.** (2017). Bergmann’s and Allen’s Rules in Native European and Mediterranean Phasmatodea. *Frontiers in Ecology and Evolution* **5**, 1–13.
- Simon, S., Letsch, H., Bank, S., Buckley, T. R., Donath, A., Liu, S., Machida, R., Meusemann, K., Misof, B., Podsiadlowski, L., et al.** (2019). Old world and new world Phasmatodea: phylogenomics resolve the evolutionary history of stick and leaf insects. *Frontiers in Ecology and Evolution* **7**, 1–14.
- Simons, A. M.** (2002). The continuity of microevolution and macroevolution. *Journal of Evolutionary Biology* **15**, 688–701.
- Stanton, A. O., Dias, D. A. and O’Hanlon, J. C.** (2015). Egg Dispersal in the Phasmatodea: Convergence in Chemical Signaling Strategies Between Plants and Animals? *Journal of Chemical Ecology* **41**, 689–695.
- Thompson, D. W.** (1917). *On growth and form*. Cambridge, United Kingdom: Cambridge University Press.
- Wright, I. J., Dong, N., Maire, V., Prentice, I. C., Westoby, M., Díaz, S., Gallagher, R. V., Jacobs, B. F., Kooyman, R., Law, E. A., et al.** (2017). Global climatic drivers of leaf size. *Science* **357**, 917–921.

CHAPTER 1

Global convergence towards multiple ecomorphs in stick and leaf insects

Global convergence towards multiple ecomorphs in stick and leaf insects

Romain P. Boisseau^{1,*}, Sven Bradler², Douglas J. Emlen¹

¹Division of Biological Sciences, University of Montana, 32 Campus Dr, Missoula, MT 59812, United States of America;

²Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute of Zoology and Anthropology, University of Göttingen, Göttingen, Germany

*Corresponding author: Romain P. Boisseau

Keywords: Phasmatodea | ecomorph | homoplasy | macroevolution | diversification

Abstract

Convergent evolution of similar forms is often used as evidence for deterministic adaptive evolution. However, variation in the extent of convergence and in the underlying evolutionary trajectories of putatively convergent lineages are often unclear. We investigated patterns of morphological convergence across the stick and leaf insects (order Phasmatodea) using 27 adult female morphological traits from more than 200 species, and show repeated and widespread convergence towards 18 discrete morphotypes. Using integrative multivariate phylogenetic comparative methods and reconstructions of the evolutionary trajectories of convergent lineages on the phasmid morphospace, we show that the extent and nature of convergence varies considerably between morphotypes. Phasmid phenotypic evolution appeared largely driven by transitions between vegetation layers, which led lineages towards predictable and partially overlapping regions of the morphospace: ground dwellers repeatedly evolved robust and wingless bodies while canopy dwellers repeatedly colonized multiple regions of the morphospace ranging from very elongated to wide and flattened body forms. The extraordinary body forms of stick and leaf insects exemplify deterministic evolution, yet the details of their evolution reveal both nuance and contingency, highlighting the importance of quantitative approaches for characterizing convergence.

Significance

In the history of life, we find many instances where unrelated lineages evolved similar body forms when submitted to similar environmental challenges. However, such events of convergence may vary considerably in their strength and underlying evolutionary patterns: How similar are the convergent lineages? How different were their ancestors? How did their morphology change over time? Evaluating these patterns is essential to shed light on how the deterministic nature of evolution interacts with its dependence upon past events. We show repeated convergence towards multiple body forms in the stick and leaf insects associated with habitat transitions. Species colonizing similar niches converged in morphology but we show substantial variation in the strength of convergence and underlying evolutionary paths, emphasizing variation in the relative determinism of evolution.

Introduction

Adaptation to environmental challenges is driven by the deterministic force of natural selection acting upon variation that has arisen through a series of mostly unpredictable and stochastic events (Blount et al., 2018). Discerning the relative contributions of contingency and selection remains a core objective of evolutionary biology (Agrawal, 2017; Conway Morris, 2003; Conway Morris, 2010; Gould, 1989; Losos, 2017; Vermeij, 2006). At the macroevolutionary scale, lineages independently invading similar environments constitute evolutionary “replays” across time and space and provide opportunities to investigate the repeatability versus contingency of evolution (Dick et al., 2009; Losos et al., 1998). Adaptive convergent evolution can occur when different lineages experience similar environmental challenges and natural selection pushes them towards the same few high-fitness solutions, despite their distinct evolutionary histories (Conway Morris, 2003; Losos, 2011; McGhee, 2011). This is particularly likely for closely related taxa that share the majority of their evolutionary past, but becomes less likely when lineages share deeper splits, illustrating a shift in the balance between contingency and determinism (Conte et al., 2012; Ord and Summers, 2015).

In the past two decades, examples of adaptive convergence have multiplied (Conway Morris, 2015; Losos, 2010; McGhee, 2011) but have mainly focused on relatively small groups of closely related taxa for which phylogenies could be reconstructed. Phylogenomic tools now enable the strong reconstruction of deep evolutionary histories permitting the investigation of convergent evolution at

larger time and spatial scales. Moreover, evidence for morphological convergence is often qualitative but methods have been recently developed to properly identify and quantify the extent of convergence (Arbuckle et al., 2014; Castiglione et al., 2019; Ingram and Mahler, 2013; Speed and Arbuckle, 2017; Stayton, 2015). Together these advances now make it possible to rigorously quantify patterns of convergent evolution across large clades of morphologically diverse and distantly related species (Borko et al., 2021; Grossnickle et al., 2020; Pigot et al., 2020; Rincon-Sandoval et al., 2020).

One difficulty in characterizing convergence is that the degree of morphological similarity can vary widely between so-called convergent taxa. “Complete convergence” refers to a significant overlap in morphospace between convergent lineages (Meachen-Samuels, 2012) while “incomplete convergence” results from lineages independently evolving some similarities but still occupying a distinct portion of the morphospace (Grossnickle et al., 2020; Herrel et al., 2004; Stayton, 2006)(**Figure 1**). Finally, “imperfect convergence” describes taxa that experienced shifts in similar directions onto a distinct region of the morphospace but may end up more disparate than their ancestors were (Collar et al., 2014)(**Figure 1**). Applying the aforementioned methods, recent studies have shown that incomplete convergence is a common phenomenon, even among “textbook” examples of convergence (Bright et al., 2016; Grossnickle et al., 2020; Law, 2022), emphasizing the need to accurately characterize morphological convergence to provide useful insights into the complexity of form-function relationships and the evolution of similar forms.

In this study, we assess the morphological diversity of stick and leaf insects (order Phasmatodea) in relation with habitat use. This mesodiverse lineage (~3,400 species) of large herbivorous insects is predominantly distributed in tropical regions around the globe, but inhabits diverse micro- and macrohabitats ranging from the mostly bare ground of deserts to the lush canopies of tropical rainforests (Bedford, 1978; Brock and Hasenpusch, 2009). Most species exhibit stunning forms of camouflage through crypsis and masquerade (Bradler and Buckley, 2018). Crypsis refers to antipredator strategies that prevent detection (Ruxton et al., 2019). For instance, the appearance of the animal (color, pattern, texture) may help it blend into the background (background matching) or may break up its outline making it difficult to perceive its shape (disruptive coloration)(Webster et al., 2013). On the other hand, masquerade does not impair detection but rather identification as a prey (Skelhorn et al., 2010). Stick insects are textbook examples of both crypsis and masquerade as most species resemble objects of the environment irrelevant to predators such as sticks, leaves, bark pieces or moss. Advergence towards such diverse objects produced a large morphological diversity ranging from

elongated tubular bodies with long slender legs to more robust habitus, sometimes with leaf-shaped expansions and/or spines. Because the morphological specializations exhibited by phasmids are quite extreme, they constitute an ideal group in which to investigate how contingent such specializations are.

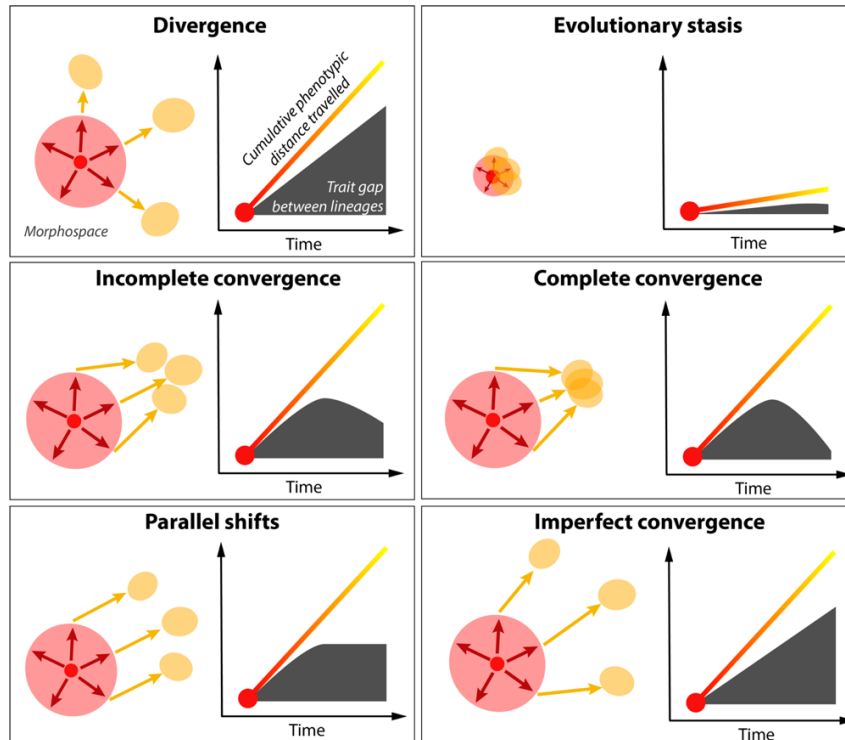


Figure 1: Schematic representation of and conceptual predictions for different patterns of phenotypic evolution. On the left of each panel, the evolutionary trajectories of lineages of interest are illustrated on a 2-D morphospace, starting with their most recent common ancestor (the big red dot). Red colors correspond to older events, yellow colors to more recent ones. On the right of each panel, the plot shows the cumulative phenotypic distances travelled by each lineage of interest combined, and the average phenotypic distance (i.e., gap) between lineages over time.

Recent phylogenetic studies on phasmids have shown conflicting results with prior taxonomic classifications based on morphological characters, suggesting a high degree of phenotypic convergence across Phasmatodea (Bradler et al., 2014; Bradler et al., 2015; Buckley et al., 2009; Glaw et al., 2019; Robertson et al., 2018; Simon et al., 2019). For example, the group of “tree lobsters” composed of flightless, robust and strongly armored species, including the famous Lord Howe island stick insect, had been traditionally recognized as the subfamily Eurycanthinae which was shown to be highly polyphyletic, illustrating a spectacular case of convergent evolution (Buckley et al., 2009). Moreover, several authors have already reported qualitatively that apterous, stockier, spinier and darker body forms tend to be found close to the ground while more elongated and winged ones tend to rest higher

up in the vegetation, suggesting a role of habitat in driving these convergent patterns (Bradler and Buckley, 2018; Cliquennois, 2008). The convergent evolution of ovipositor structures and oviposition strategies (e.g., burying eggs in the soil) has also been hypothesized to be associated with habitat transitions (Goldberg et al., 2015; Robertson et al., 2018). Here, we quantitatively assess the presence and extent of convergent evolution in body morphology associated with transitions in habitat use in Phasmatodea. We tested the hypothesis that independent colonization of similar vegetation layers are associated with the invasion of restricted and distinct portions of the morphospace. As stick insects rely on substrate appearance for the efficiency of crypsis and on the surrounding items' diversity for masquerade, we expected structurally disparate habitats to display different adaptive landscapes leading lineages to distinct and predictable regions of the morphospace.

To address this question, we reconstructed a time-calibrated multilocus phylogeny of the order Phasmatodea comprising 314 taxa and an associated morphospace of overall female body morphology for 212 species. Using a variety of comparative approaches, we characterized discrete morphotypes for which we assessed and quantified the degree and pattern of convergent evolution, following the conceptual framework laid out in **Figure 1**.

Results

Reconstruction of the phasmid phylogeny. We used genetic data (from 3 nuclear and 4 mitochondrial genes) from 314 phasmid taxa and fossil-calibrated Bayesian inferences to reconstruct the evolutionary history of Phasmatodea. The relationships between the major euphasmatodean clades that arose during an ancient radiation were constrained to match the basal topology inferred in previous phylogenomic studies (Simon et al., 2019; Tihelka et al., 2020). The inferred Maximum Clade Credibility (MCC) tree was overall strongly supported and was largely congruent with previous studies (**Figure 2**) (Bank et al., 2021b; Bradler et al., 2015; Robertson et al., 2018), providing a robust framework for all subsequent comparative analyses. 16 major clades were recovered and appeared largely defined by geographic distribution and ecozones (**Figure 2**).

Reconstruction of the phasmid morphospace and definition of morphotypes. We assembled a morphological dataset comprising more than 1300 adult female specimens from 212 species included in the phylogeny and including 24 quantitative measurements and qualitative data on color and body texture (Figure S1). From this dataset, we reconstructed a multidimensional morphospace using a mixed Principal Component Analysis (PCAmix). The PCA revealed large variation between phasmid species in relative body width (PC1, 44% of the total variation), relative body height (i.e., how flattened the body is; PC2, 11%), relative wing size (PC3, 11%), body texture (i.e., how smooth or rough the body cuticle is; PC4, 8%) and relative head size (PC5, 5%) (**Figure 3**, S2). The first five PCs together accounted for 77% of the total variation.

Morphological variation in phasmids appeared bimodal between the wide-leaf mimicking Phylliidae clade and the rest of the phasmids, which display a more stick-like, tubular morphology (**Figure 3**). Phylliids, often referred to as true leaf insects, are characterized by an exceptionally widened and flat abdomen giving them the appearance of wide angiosperm leaves (Bank et al., 2021a). The morphological distinctiveness of this clade had led to the designation of the separate “Phyllioptera” order as sister to all other phasmatodeans in the past (Crampton, 1916). All the other phasmid clades appeared centered around a single core on the morphospace, varying mostly in relative body width ranging from extremely elongated to more robust body silhouettes. Species with extreme morphologies were scattered at the periphery of this central core, often only projecting out along a single axis. For instance, the large-headed palm stick insects (subfamily Megacraniinae) mostly stand out along the PC5 axis that separates species based on relative head size (**Figure 3**). Interestingly, it appeared that only

the Phylliidae occupy their own distinct region of the morphospace, while all the other major phasmid clades occupy largely overlapping regions, highlighting the potential for widespread convergence (**Figure 3**). Most of the morphological diversity is found in the Euphasmatodea, consistent with their much greater species diversity ($n > 3400$ species), compared to Timematodea ($n = 21$ species) which is morphologically homogeneous (**Figure 3**). The explosive morphological diversification of Euphasmatodea can be visualized in Video S1.

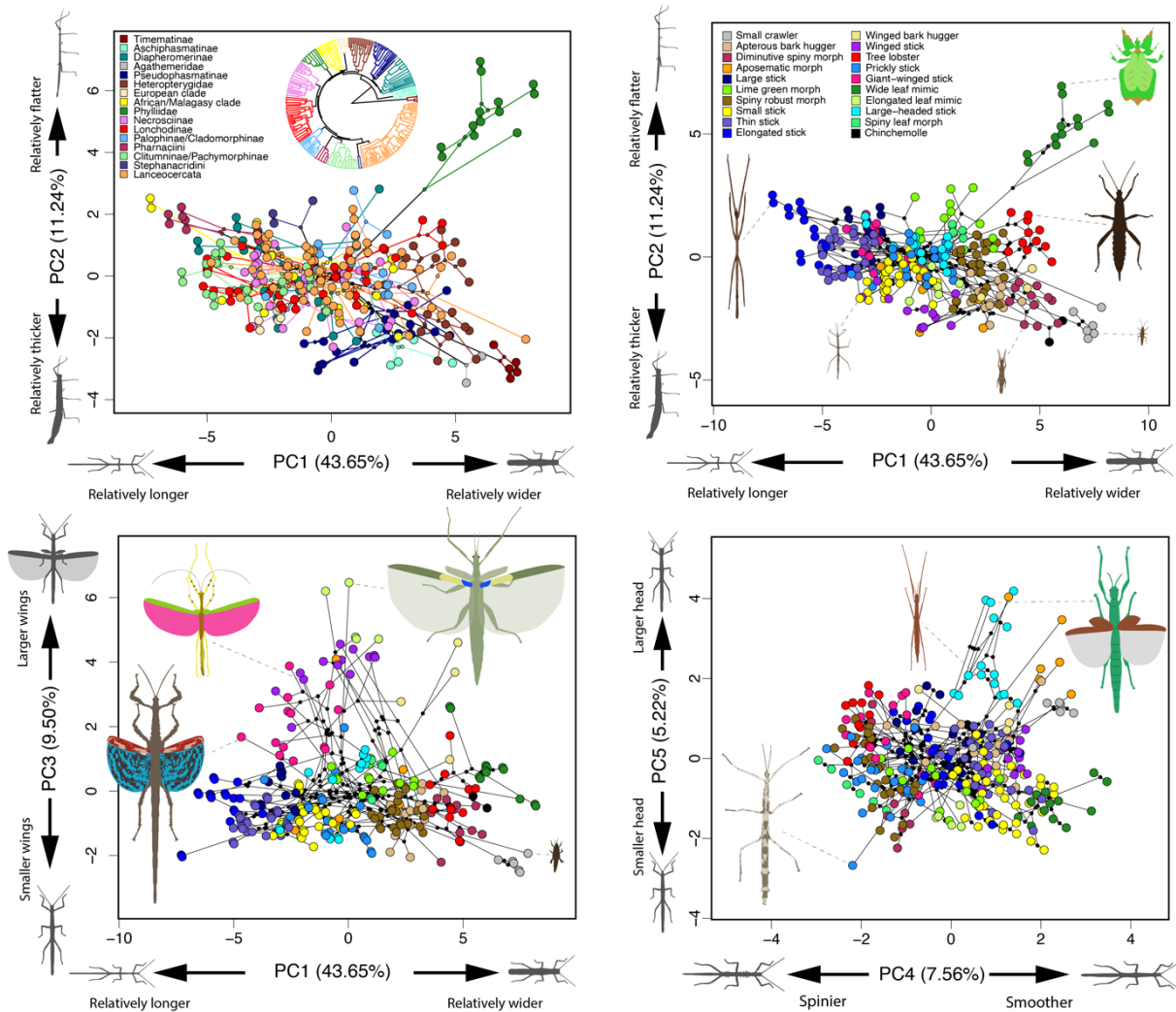


Figure 3: The phasmid phylomorphospace. The first five different dimensions are shown. Species are colored by phylogenetic clade (see Figure 1) in the top left panel, and by assigned morphotype otherwise (see Figure 4). Black lines between points represent the underlying phylogenetic relationships between species. Phasmid illustrations drawn by first author.

We then used a hierarchical clustering approach to define and assign species to morphotypes (i.e., clusters) occupying relatively distinct regions of the multidimensional morphospace (Figure 4, S3-5). As expected, among the 20 morphotypes defined, we recovered the aforementioned *wide leaf mimic* morphotype and the previously recognized *tree lobster* morphotype, which includes the thorny devil stick insects (*Eurycantha* spp.) and the Lord Howe Island stick insects (*Dryococelus australis*) (Buckley et al., 2009). Using random forest machine learning models, we identified the main morphospace axes that were most helpful for these predictive models to infer morphotype state from the morphological data and therefore the axes best distinguishing each morphotype (Table S4). This analysis revealed that morphotypes are often only standing out along a few dimensions of the morphospace. For instance, *spiny leaf mimics* (including the Australian spiny leaf insects, genus *Extatosoma*) were best distinguished by PC9 (i.e., mostly reflecting body shape solidity) due to their jagged body edges mimicking dried leaf edges, lichen or moss (Table S4, Figure S2-3).

The repeated evolution of morphotypes. Ancestral state reconstruction based on stochastic character mapping suggested that the *small stick* morphotype, characterized by its central position on the morphospace, was the ancestral morphotype of Euphasmatodea (Figure S6). It further revealed that the *chinchemolles* (clade Agathemeridae) and the *wide leaf mimic* morphotype (clade Phylliidae) each had unique origins and were therefore evolutionary one-offs. *Chinchemolles*, as originally named by native South Americans, are uniquely fat, robust and smooth insects adapted to the arid slopes of the Andes range and known for their foul defensive smell (from the Quechua chinche: “fetid”, and moyo: “udder”) (Camousseight, 1995; Cubillos and Vera, 2020) (Figure S4). All other morphotypes appeared to have originated at least twice (e.g., *small crawler* morphotype) and up to at least 10 times (e.g. *thin stick* morphotype), indicating widespread morphological convergence in the order (Figure S6).

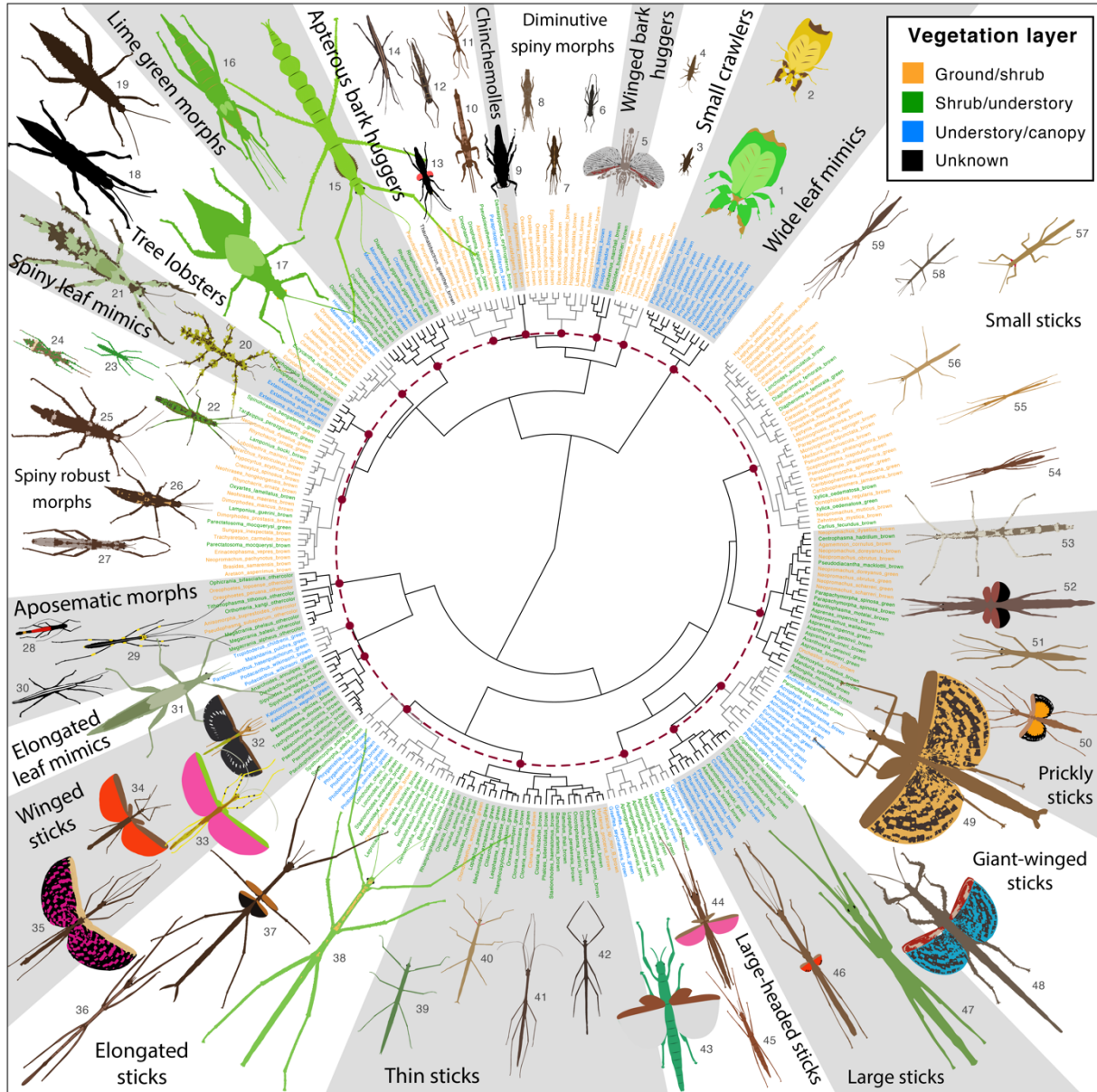


Figure 4: Hierarchical cluster dendrogram based on 24 continuous and 3 discrete functional morphological variables using the Ward’s method. Tip labels are colored according to extant habitat. The dashed maroon circle correspond to the height threshold used to delineate morphotypes. Intersection between the circle and dendrogram branches are shown as maroon dots. Scaled adult female illustrations were drawn by the first author and correspond to the species listed in Table S3.

Assessing and quantifying morphological convergence. We quantified the extent and significance of convergence for each morphotype using six new pattern-based metrics. We first calculated the C1-C4 indexes, which measure the increase in similarity between convergent taxa through time (Stayton, 2015). Specifically, C1 measures the proportion of the maximum ancestral phenotypic distance between two lineages subsequently reduced by convergence. C2 measures the absolute amount of

convergence. C3 and C4 measure the proportion of convergence in the total amount of phenotypic evolution in the smallest monophyletic group containing the two taxa or in the entire tree. We used 1,000 simulations of undirected and random character evolution to derive p-values for each metric and assess whether the observed convergence (as measured by the different metrics) was stronger than would be expected by chance. Most putatively convergent morphotypes showed statistically significant evidence of convergence in overall morphology, as measured by these metrics (**Figure 5**).

The *aposematic morphotype* did not show significant convergence for any of the metrics ($p > 0.07$, **Figure 5**). However, this is expected as these aposematic lineages are quite morphologically diverse and only are grouped together into a single morphotype because of their contrasting showy coloration, a principal component (PC 8 in Figure S2, Table S1) that was not included in the calculation of convergence metrics (PC1-PC4). The same logic applies to the *large-beaded sticks*, which show only marginal evidence of convergence, as this morphotype is mainly characterized by a relatively large head (mostly reflected by PC5; **Figure 5**, Table S1).

Interestingly, the strength of convergence as assessed by C1-4, varied widely between morphotypes (**Figure 5**). For example, *tree lobsters*, *small crawlers* and *winged sticks* showed a very high degree of convergence on average (i.e., C1=0.8, 0.75 and 0.87 respectively) while *giant-winged sticks*, *lime green morphs* and *winged bark huggers* showed a lower extent of convergence (0.44, 0.44 and 0.49). Although these latter lineages did get closer to one another on the morphospace, much of their ancestral phenotypic distance remains, suggesting largely incomplete convergence.

We next calculated the Wheatsheaf index (w), which also assesses the strength of convergence (Arbuckle et al., 2014). Here, a higher w indicates a high degree of phenotypic proximity among the convergent taxa and/or a greater distinctiveness of those taxa relative to all the others. w also showed significant convergence for all morphotypes except the *aposematic morph*. It similarly revealed considerable variation in the extent of convergence between morphs and further highlighted some strongly convergent and distinct morphotypes, most notably the *small crawlers* (characterized by their very small size and robust habitus, $w=6.15$) and the *spiny leaf mimics* (characterized by their lateral indented foliaceous abdominal extensions, $w=4.97$)(**Figure 5**).


















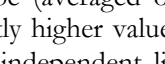
	C1	C2	C3	C4	C5	w	
Small crawlers <i>Timema californicum</i> , <i>Microcanachus matileorum</i>	0.746	7.215	0.335	0.014	2	6.153	
<i>p</i> = 0.002	<i>p</i> < 0.001	<i>p</i> = 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	
Diminutive spiny morph <i>Ommatopseudes harmani</i> , <i>Epidares nolimetangere</i> , <i>Labidophasma rouxi</i>	0.646	3.757	0.297	0.007	7	3.441	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.001	<i>p</i> = 0.01	<i>p</i> < 0.001		
Tree lobsters <i>Haaniella dehaanii</i> , <i>Eurycantha coronata</i> , <i>Dryococelus australis</i> , <i>Canachus alligator</i>	0.798	5.08	0.333	0.01	6	3.309	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.009	<i>p</i> < 0.001		
Spiny robust morph <i>Lobolibeira maineri</i> , <i>Crocoxylius spinosus</i> , <i>Sungriya inexpectata</i> , <i>Paracatantonia moccquayesi</i> , <i>Neohispana naevosa</i> , <i>Ertanacophasma vepres</i> , <i>Neopromachus dysellus</i> , <i>Cnispus rachus</i> , <i>Rhynchacris ornata</i> , <i>Dimorphodius prostratis</i> , <i>Micrarachus hystericuleus</i>	0.551	2.276	0.201	0.004	21	3.139	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	
Apterous bark huggers <i>Dinophasma saginatum</i> , <i>Paraprisopus arillarum</i> , <i>Anisocnirpa ferruginea</i> , <i>Malacomorpha jamaicana</i> , <i>Pseudophasma subapterum</i> , <i>Peruphasma schultzei</i> , <i>Pseudosiphonotus inopletis</i> , <i>Thaumobactron guentheri</i>	0.57	2.452	0.223	0.004	13	3.15	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.012	<i>p</i> < 0.001		
Winged bark huggers <i>Prisopus berosus</i> , <i>Neoclidetes buescheri</i> , <i>Epicharmus marchali</i>	0.487	3.124	0.207	0.006	3	2.321	
<i>p</i> = 0.009	<i>p</i> < 0.001	<i>p</i> = 0.012	<i>p</i> = 0.005	<i>p</i> = 0.23	<i>p</i> = 0.007		
Lime green morph <i>Diapherodes gigantea</i> , <i>Cranidium gibbosum</i> , <i>Macrophasma biroi</i> , <i>Paramenexenus laetus</i> , <i>Monandroptera acanthomera</i>	0.444	1.59	0.149	0.003	7	2.206	
<i>p</i> = 0.001	<i>p</i> = 0.006	<i>p</i> = 0.013	<i>p</i> = 0.019	<i>p</i> = 0.14	<i>p</i> = 0.004		
Elongated leaf mimics <i>Podacanthus wilkinsoni</i> , <i>Malandania pulchra</i>	0.739	2.952	0.317	0.076	4	2.815	
<i>p</i> = 0.001	<i>p</i> < 0.001	<i>p</i> = 0.005	<i>p</i> = 0.008	<i>p</i> < 0.001	<i>p</i> < 0.001		
Spiny leaf mimics <i>Extatosoma tiaratum</i> , <i>Trychopeplus laciniatus</i>	0.648	2.521	0.158	0.005	3	4.973	
<i>p</i> = 0.015	<i>p</i> = 0.038	<i>p</i> = 0.113	<i>p</i> = 0.084	<i>p</i> = 0.089	<i>p</i> < 0.001		
Winged sticks <i>Trachythorax maculicollis</i> , <i>Metriophasma dioctes</i>	0.865	4.24	0.319	0.009	11	2.342	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.006	<i>p</i> = 0.009	<i>p</i> = 0.004	<i>p</i> = 0.001		
Giant-winged sticks <i>Anchiale briareus</i> , <i>Bactrododema hecticum</i> , <i>Lopaphus sphaerulus</i> , <i>Achrioptera punctipes</i>	0.441	1.732	0.176	0.003	7	2.171	
<i>p</i> = 0.001	<i>p</i> = 0.001	<i>p</i> = 0.002	<i>p</i> = 0.014	<i>p</i> = 0.081	<i>p</i> < 0.001		
Large sticks <i>Cladomorpha phyllinus</i> , <i>Allenobostra brocki</i> , <i>Pharnacia ponderosa</i> , <i>Phasmotania lanyuhensis</i> , <i>Phaenophasos khaoyaiensis</i>	0.673	3.306	0.274	0.006	7	3.738	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.074	<i>p</i> < 0.001		
Elongated sticks <i>Bacteria pisioria</i> , <i>Spathomorpha adula</i> , <i>Baculifractum insignis</i> , <i>Leptocaulinus insularis</i> , <i>Lonchodes chani</i> , <i>Phobaeticus kirbyi</i> , <i>Cuniculina cuniculus</i> , <i>Ctenomorpha marginipennis</i>	0.688	4.41	0.326	0.008	12	2.588	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.013	<i>p</i> < 0.001		
Small sticks <i>Diapheromera femorata</i> , <i>Clonopsis gallica</i> , <i>Zehntneria mystica</i> , <i>Carausius morosus</i> , <i>Parapachymorpha spiniger</i> , <i>Monoignosis spinosa</i>	0.539	1.796	0.188	0.003	17	3.106	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.001	<i>p</i> = 0.003	<i>p</i> < 0.001		
Thin sticks <i>Oncotophasma martini</i> , <i>Phanocloidea nodulosa</i> , <i>Phalacis tuberculatus</i> , <i>Leioptasma lucidus</i> , <i>Lopaphus perakensis</i> , <i>Onites sempiterni</i> , <i>Chondrostethus woodfordi</i> , <i>Hyrtacus procerus</i> , <i>Hyrtacus sp.</i> , <i>Medauroidea extradentata</i> , <i>Ciliarchus hookeri</i>	0.668	3.046	0.255	0.006	14	3.21	
<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> = 0.006	<i>p</i> < 0.001		
Prickly sticks <i>Antongilia muricata</i> , <i>Centrophasma hadritum</i> , <i>Pseudodacantha macklotii</i> , <i>Manduria systopedon</i> , <i>Neopromachus obrutus</i> , <i>Pterinoxylus crassus</i> , <i>Agamemnon cornutus</i> , <i>Parapachymorpha spinosa</i> , <i>Acanthosylla geisovii</i>	0.359	1.288	0.119	0.002	18	2.259	
<i>p</i> < 0.001	<i>p</i> = 0.004	<i>p</i> = 0.012	<i>p</i> = 0.024	<i>p</i> = 0.031	<i>p</i> < 0.001		
Large-headed sticks <i>Graeffea leverii</i> , <i>Apterograeffea reunionensis</i>	0.414	1.35	0.205	0.019	9	1.899	
<i>p</i> = 0.07	<i>p</i> = 0.045	<i>p</i> = 0.062	<i>p</i> = 0.13	<i>p</i> = 0.11	<i>p</i> < 0.001		
Aposematic morph <i>Orthomeria kangii</i> , <i>Oncophanes peruana</i> , <i>Tithonophasma tithonus</i> , <i>Anisocnirpa bigresoides</i> , <i>Pseudophasma subapterum</i> , <i>Megacrania batossi</i> , <i>Ophicrania bifasciatus</i>	0.21	0.92	0.078	0.002	18	1.724	
<i>p</i> = 0.074	<i>p</i> = 0.094	<i>p</i> = 0.24	<i>p</i> = 0.28	<i>p</i> = 0.083	<i>p</i> = 0.23		

Figure 5: Convergence indexes calculated for each putatively convergent morphotype (averaged over all possible pairs of taxa) and associated p-values. Bolded values correspond to significantly higher values than expected under simulated undirected character evolution. The names of the specific independent lineages included in each analysis are listed under each morphotype name. Warmer colors indicate higher values for each convergence index (*i.e.*, stronger convergence). Drawings on the right illustrate one taxon belonging to each morphotype and are scaled relative to one another. Illustrations by first author.

We used the C5 metric to quantify the frequency of convergence into specific regions of the morphospace (Stayton, 2015). C5 counts the number of outside lineages that enter or cross into a hypervolume of the morphospace defined by the convergent taxa over time. C5 also varied considerably with most morphotypic niches being colonized more often than expected by chance, suggesting directed evolution towards these regions (**Figure 5**). Again, and as expected, the morphospacial region occupied by *aposematic morphs* did not show more colonizations than expected by chance ($p=0.08$). *Spiny leaf mimics*, *winged bark huggers*, *giant-winged sticks*, *large sticks* and *lime green* phasmids also showed non-significant C5s (**Figure 5**). This may be explained by the large morphospacial regions that they define, which may constitute large targets for undirected evolution and therefore make it likely that many lineages cross these regions by chance.

Evolutionary trajectories and patterns of convergence. To test the different predictions derived from each evolutionary pattern laid out in **Figure 1**, we chose for each of the 18 putatively convergent morphotypes a maximum of four independent lineages to represent their evolutionary trajectories on the morphospace. We then reconstructed their morphological evolution along the two most important principal component axes characterizing each morphotype (e.g., head size, spininess) starting at the position of the most recent common ancestor of Phasmatodea (Table S1). We estimated ancestral values every 500,000 years and, for each time step, calculated the cumulative phenotypic distance travelled by each pair of lineages combined and their trait gap (i.e., Euclidean distance between the two lineages of interest). We then calculated C1-C4 and associated p-values for each pair only considering these two principal component axes. These analyses revealed extensive variation in convergence patterns across morphotypes, and even between lineages converging towards the same morphotype (**Figure 6**, S7-23). First, most lineages assigned to a given morphotype showed extremely strong, significant and putatively complete convergence along their most characteristic axes (i.e., high C1-C4 values, lineages travelling long distances on the morphospace and ending up very close to one another, and the trait gap over time displaying a conspicuous “hump” [gray shading in **Figure 1**]). This is the case, for instance, for all the *tree lobster* lineages (**Figure 6**) and most *apterous bark huggers* (Figure S8), *diminutive spiny morphs* (Figure S9), *elongated sticks* (Figure S11), *large headed sticks* (Figure S13), *large sticks* (Figure S14), *small crawlers* (Figure S16), *spiny leaf mimics* (Figure S19), *spiny robust morphs* (Figure S20), *thin sticks* (Figure S21) and *winged sticks* (Figure S23).

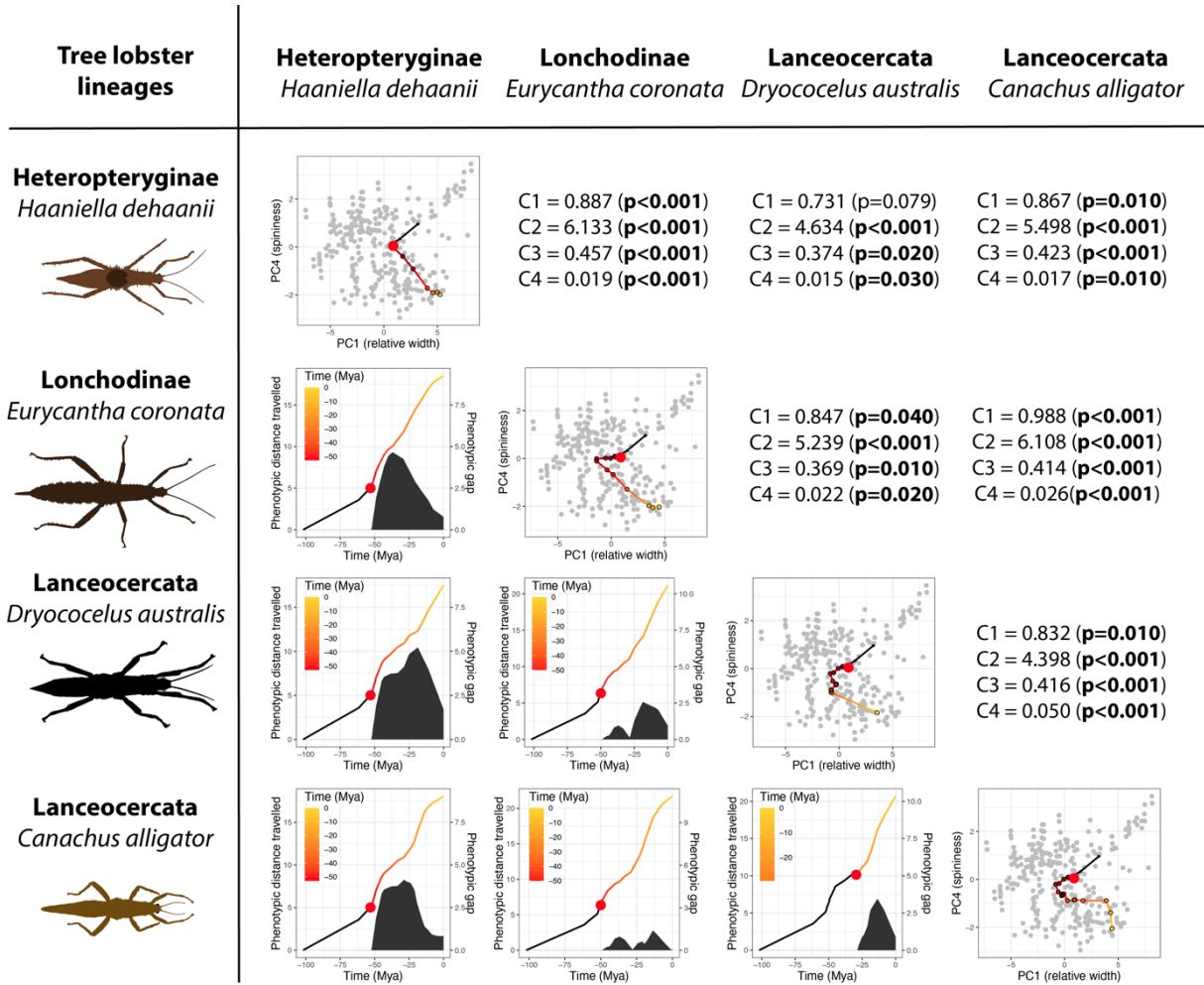


Figure 6: Convergent evolutionary trajectories of the four lineages of tree lobsters. The diagonal of the table shows the position of ancestral nodes for each lineage on the morphospace along the two axes most distinguishing tree lobsters. Colors transition from red to yellow as we go from the position of the most recent common ancestor of all four lineages (the big red dot) to present time. The lower panels in the table show the cumulative phenotypic distance travelled by each lineage pair through the 2-D morphospace and the corresponding trait gap over time. The numbers in the upper part of the table show the values and associated p-values ($n=1000$ simulations) of C1-C4 indexes calculated for each lineage pair only considering the two axes of the morphospace represented. Phasmid drawings by first author.

In contrast, other lineages showed weaker and largely incomplete convergence such as most *lime green morphs* (Figure S15), *prickly sticks* (Figure S16), *small sticks* (Figure S18) and *winged bark huggers* (Figure S22). Still other lineages invaded a similar region of the morphospace but then subsequently diverged within this region (i.e., imperfect convergence), such as the *aposematic morphs* (Figure S7), the *elongated leaf mimics* (Figure S10) and the *giant-winged sticks* (Figure S12). Remarkably, many convergent lineages shifted in parallel towards the same morphospacial region. The most striking cases can be

found in *elongated sticks* (Figure S11), *large sticks* (Figure S14), *lime green morphs* (Figure S15), *prickly sticks* (Figure S16), *spiny robust morphs* (Figure S20) and *winged sticks* (Figure S23). Finally, we could identify a few cases of evolutionary stasis where lineages did not move much on the morphospace after splitting. This is the case for the *apterous bark bugger* lineages *Dinophasma saginatum* and *Anisomorpha ferruginea* (Figure S8) or the *small stick* lineages *Neopromachus muticus* and *Monoiognosis bipunctata* (Figure S18).

Habitat transitions. The shrub/understory layer was reconstructed as ancestral in euphasmatodeans and the associated habitat transition matrix identified asymmetric transition rates between vegetation layers suggesting that transitions from the shrub/understory layer to the other two layers were the most frequent (**Figure 7A**).

Habitat hypervolume size and overlap. We calculated the size of the multidimensional hypervolumes occupied by each habitat on the morphospace and the overlap between them using dynamic range boxes (Junker et al., 2016). Overall, canopy dwelling species occupied the largest volume on the morphospace, ground dwellers the smallest (**Figure 7B-F**). Specifically, canopy dwellers spanned the greatest ranges along the PC1 (i.e., relative width), PC2 (i.e., relative height) and PC3 (i.e., relative wing size) axes (**Figure 7G**). This reflects the considerable variation in body morphology of canopy dwellers that go from extremely elongated and cylindrical stick insects (e.g., Figure S5 #48-54) to wide and flat leaf insects (e.g. Figure S5 #42) and wingless to fully winged species (e.g. Figure S5 #48-50, 46-47). In contrast, ground dwellers either included small stick-mimicking species (e.g., Figure S3 #19) or stockier and more robust taxa (Figure S4 #20-23, 29-34). Overlap, as measured by the portion of the hypervolume of habitat A also occupied by the hypervolume of habitat B, was overall strongest between adjacent vegetation layers (i.e., between shrub and canopy dwellers, and between shrub and ground dwellers) (**Figure 7H-J**) but varied along specific axes (**Figure 7K-O**). For instance, ground and canopy dwellers spanned largely dissimilar ranges specifically along PC2 (i.e., canopy dwellers are typically flatter than ground dwellers) (**Figure 7L**) and PC3 (i.e., many canopy dwellers are winged while ground dwellers typically are not) (**Figure 7M**), while they tend to be more similar along PC1 (relative body width) (**Figure 7K**) and PC4 (spininess) (**Figure 7N**).

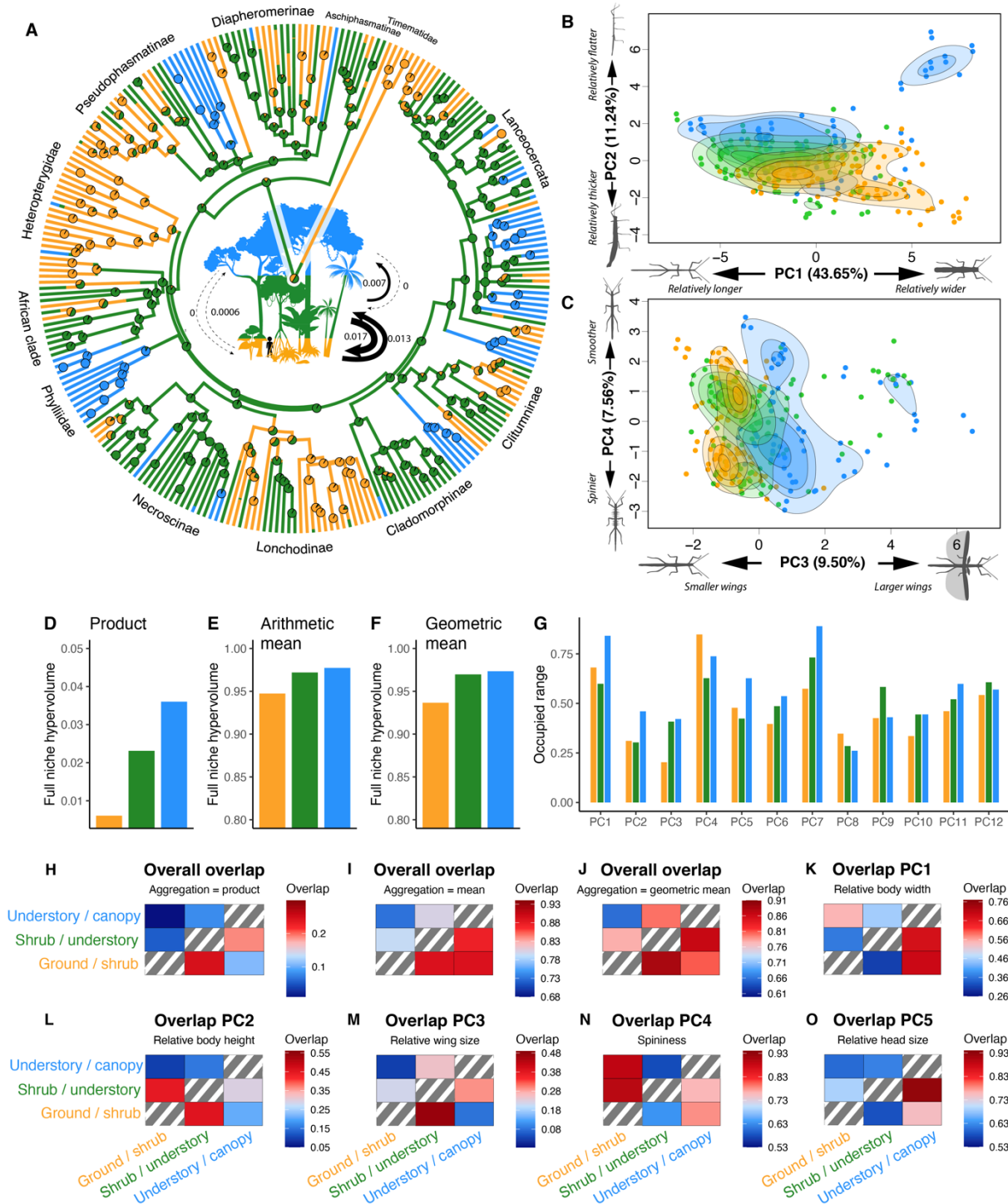


Figure 7: Habitat transitions and morphospace occupation and overlap between different habitats. **A:** Ancestral state reconstruction of habitat layer using stochastic character mapping and a transition matrix (inset) estimated by maximum likelihood. **B-C:** Two dimensional kernel density distributions of the species in each vegetation layer on the morphospace. **D-F:** 12-dimensional hypervolume size for each habitat state obtained using the dynamic range boxes approach with three different aggregation methods. **G:** Occupied range size by each habitat state along the first 12 axis of the morphospace. **H-O:** Heatmaps showing the portion of the hypervolume of habitat in y covered by the hypervolume of habitat in x. **H-J** show overlap of 12-dimensional hypervolumes aggregated as product, arithmetic and geometric mean. **K-O** show overlap along individual dimensions. Note different scales for each heatmap.

Quantifying the match between morphological traits and habitat. To further investigate whether lineages occupying different vegetation layers evolved to occupy distinct regions of the morphospace, we tested whether habitat was predictable given the position of a species on the morphospace, using random forest models. We found that the machine learning classification algorithm was able to predict habitat from morphological traits better when dimensionality was higher (Figure S24). The accuracy of predictions was limited when only based on the first morphospace axis (48%), despite PC1 accounting for almost half of the phenotypic variance (Figure S2). But, remarkably, the predictive performance of the algorithm quickly plateaued around 80% after including the first 6 axes only. Mistakes were most common between adjacent vegetation layers further indicating a morphological overlap between adjacent habitats (Figure S24).

Discussion

By integrating several recently developed comparative analyses associated with a robust phylogenetic reconstruction, we found strong evidence for widespread morphological convergent evolution towards a multitude of morphotypes in stick and leaf insects. Phasmid morphological evolution appeared closely associated with transitions in habitat use as lineages invading new vegetation layers converged in specific and predictable regions of the morphospace. Given this strong association between form and function, phasmid morphotypes are likely to constitute ecomorphs. Ground phasmids appeared consistently more robust and with relatively smaller wings and exhibited the lowest morphological diversity while canopy dwellers showed the highest morphological variability, adverting towards very elongated “stick-like” and often winged body forms all the way to wide and flattened body shapes mimicking leaves. In between, shrub dwellers, likely the ancestral habitat of phasmids, occupied a rather central place on the morphospace.

These patterns of morphological evolution indicate that the extraordinary diversity displayed by stick and leaf insects is the result of replicate adaptive radiations in different geographic regions associated with widespread morphological convergence between lineages invading similar niches (Schluter, 2000). Euphasmatodea shows a deep radiation at the base of the group (~60–55Mya) following the K-T boundary (Simon et al., 2019), which corresponds with the origin of most major clades and with dispersal across vast regions of the globe (**Figure 2**). Several of these euphasmatodean

clades subsequently radiated into multiple different morphotypes which colonized diverse habitats – namely Lanceocercata (Australasia and Mascarene islands), Cladomorphae (Caribbean islands), Lonchodinae (Indomalaya/ Australasia), Necrosiinae (Indomalaya), African/Malagasy clade (Afrotropics), Pseudophasmatinae (Nearctic and Neotropics) and Diapheromerinae (Nearctic and Neotropics).

In contrast, other morphologically homogeneous clades seem to have undergone speciation without niche differentiation (i.e., non-adaptive radiation) driven mostly by allopatry. This is the case for Phylliidae (wide leaf mimics and canopy-dwellers) and the Heteropterygidae (spiny and robust ground-dwellers), which are distributed on many islands of Indomalaya and Australasia (Bank et al., 2021b; Bank et al., 2021a).

Ecomorphological convergence is found in other arthropod groups (e.g., in the cave-dwelling amphipod genus *Niphargus* (Amphipoda, Crustacea)(Borko et al., 2021; Trontelj et al., 2012); in the Hawaiian stick spiders (Araneae, Chelicerata)(Gillespie et al., 2018); or in the fancy case caterpillar genus *Hyposmocoma* (Lepidoptera, Insecta)(Kawahara and Rubinoff, 2013)). Ecomorphological convergence has also been suggested at deeper time and spatial scales in arthropods, notably in Mantodea (Insecta)(Svenson and Whiting, 2009), Tettigoniidae (Orthoptera, Insecta)(Mugleston et al., 2018) and crabs (Decapoda, Crustacea)(Luque et al., 2019; Wolfe et al., 2021). In Phasmatodea, we identified and quantified convergence towards at least 18 ecomorphs globally. These widespread and recurrent transitions towards the same morphospacial regions suggest strong evolutionary determinism resulting from a similar use of habitat.

The details of convergence varied widely across ecomorphs, however, ranging from strong convergence of lineages displaying very different evolutionary trajectories (e.g., *tree lobster* ecomorph, **Figure 6**) to imperfect convergence, where lineages converged initially only to diverge into novel regions of the morphospace later (e.g., *giant-winged stick* ecomorph, Figure S12). Such disparity in the patterns of convergence suggest varying levels of historical constraints and/or selective regimes. For example, in some cases convergence only happened along one or a few specific axes of the 15-dimensional morphospace, such as the convergence between *large-headed stick insects*, which show very strong signals of convergence along their two most distinguishing axes, corresponding to head size and head shape (Figure S14), but low and sometimes non-significant convergence indexes on the rest of their morphology (**Figure 5**). Convergence in head morphology in these lineages may be explained

by their independent diet shift to feed exclusively on the hard leaves of palms (Arecaceae) and screwpines (Pandanaeae) (Brock and Hasenpusch, 2009; Cliquennois and Brock, 2002; Moldowan et al., 2016), which likely favored large mandibular muscles (Hennemann, 2020). However, diet may not affect the rest of the body's morphology, which may instead be influenced by host plant height or leaf morphology.

In contrast, convergence between lineages may be widespread across most axes of the morphospace, but absent along select others in revealing ways. The *giant-winged sticks* show largely incomplete but significant convergence in overall morphology (**Figure 5**) but not along one of their most distinguishing axes, which reflects variation in wing size (Figure S13). While the females of these large, elongated and flightless canopy-dwelling species all possess wings, they vary considerably in the relative sizes of these wings. Their brachypterous and often brightly colored wings are typically used in startle displays directed at predators, and likely also for parachuting if falling from the canopy (Bank and Bradler, 2022; Bedford, 1978; Robinson, 1968; Zeng et al., 2020). Since these wings are not used for powered flight, there may not be an “optimum” wing size for all giant-winged stick lineages and selection likely varies between lineages depending on predation risk, predator communities, risk of falling, host tree height and wing production cost (Maginnis, 2006). In contrast, for lineages harboring wings that enable powered flight, selection may strongly and universally favor relatively large wings to maintain a low wing loading and efficient flight performances (Boisseau et al., 2022; Zeng et al., 2020). Consistently, the flight capable canopy-dwelling *winged sticks* show strong convergence along the wing size axis (Figure S24).

Overall, phenotypic diversity, as estimated by morphospacial hypervolume, was highest in the canopy layer and lowest near the ground. These differences may be the result of different levels of visual niche diversity. Canopy dwellers include winged and wingless, elongated and wide, and spiny and smooth taxa imitating bark texture or mimicking sticks of varying sizes and leaves of varying shapes. In contrast, ground dwellers either looked like small sticks (typically when resting in the grass layer) or appeared robust, wingless and spiny, mostly matching their background coloration and texture and sometimes mimicking small pieces of bark. These results contrast with patterns seen in vipers, monitor lizards and Columbiforme birds, where arboreality is associated with lower levels of morphological variation compared to terrestriality (Collar et al., 2011; de Alencar et al., 2017; Lapiedra et al., 2013). Challenges imposed by the gappy canopy habitat may strongly constrain body form in snakes and lizards, due to the requirements of arboreal locomotion. But most female stick insects

hardly move at all, relaxing this locomotor constraint. Instead, female phasmid body forms likely result from visually hunting predators selecting for increasingly effective masquerade and camouflage, which in forest canopies appears to have yielded extreme and diverse morphological specializations. Male phasmids, in contrast, often must travel long distances in search of sedentary females (Boisseau et al., 2020; Boisseau et al., 2022), and their morphology is likely to be more constrained as a result. Future studies will be needed to test this specifically.

The extraordinary radiation of stick insect morphologies was no doubt driven by diverse selective factors including habitat structural characteristics (e.g., height, density), camouflage (i.e., habitat visual characteristics), diet, specific locomotor capacities (e.g., flight capable versus flightless lineages) and locomotor needs (e.g., females burying their eggs in the ground may need to move more than females simply flinging their eggs away (O’Hanlon et al., 2020)). Therefore, a more precise characterization of species ecological niches would be desirable to more clearly associate specific morphotypes with specific natural histories. However, such precise descriptions of habitat use are challenging in phasmatodeans for many reasons, owing to their elusive and often solitary behavior, their spectacular camouflage, and to a dearth of natural history observations. We hope that more precise habitat characterizations will be possible in the future, but this will require extensive field work and more detailed descriptions of collection sites.

The observation in stick and leaf insects of astonishing and repeated cases of convergence at the global scale epitomizes the idea that evolution can be relatively predictable even at deep phylogenetic scales, and appears to go against Gould’s views on the pervasiveness of historical contingency. However, by looking more closely at the details of evolutionary trajectories, quantifying patterns of convergence across lineages, we show evolution’s path to be more nuanced, and perhaps better viewed as a continuum (Losos, 2017). Our study highlights the power of recent approaches for rigorously characterizing the forms of convergence, shedding light on the myriad ways that determinism and contingency interact to shape the evolution of organismal form.

Materials and Methods

Taxonomic sampling and phylogenetic reconstruction. Extended and detailed materials and methods are reported in the SI Appendix, Supplementary Materials and Methods. Well-supported phylogenies for 38 phasmid lineages representing all major clades of Phasmatodea were recently constructed using next-generation sequencing (transcriptomes), yielding topologies that resolved most of the deep nodes within this group with high confidence (Simon et al., 2019; Tihelka et al., 2020). Here we reconstructed a phylogeny with 314 species representing all major phasmid lineages (9% of the known phasmid species diversity and 33% of currently recognized generic diversity), and one species of Embioptera (the sister clade of Phasmatodea) as outgroup, constraining the basal topology to match the transcriptome-based trees (Tihelka et al., 2020). Regions of 3 nuclear (*i.e.*, 18S rRNA (18S), 28S rRNA (28S) and histone subunit 3 (H3)) and 4 mitochondrial genes (*i.e.*, 12S rRNA (12S), 16S rRNA (16S), cytochrome-c oxidase subunit I (COI) and cytochrome-c oxidase subunit II (COII)) were extracted from Genbank, aligned and concatenated (6,778bp total) to reconstruct the phylogenetic history of phasmids using Bayesian inferences in BEAST 2 (v. 2.6.3)(Bouckaert et al., 2014). Divergence time was estimated using 6 unambiguous crown-group phasmid fossils as minimum calibration points (Table S1).

Habitat data. We broadly classified the habitat of stick insects as the typical vegetation layer in which they rest during the daytime (*i.e.*, when they are exposed to visually hunting predators). We defined three broad categories: ground/shrub, shrub/understory and understory/canopy. These categories included overlapping names to emphasize that boundaries between inhabited vegetation layers are loose as phasmids often navigate between them. Ground/shrub dwelling taxa were defined as often resting below 1.5 meters above the ground. These taxa can be found in the leaf litter but more typically laying on the base of trunks, mossy logs, under bark, on low vegetation on in tall grass. Understory/canopy dwelling taxa are typically found high up in the tree canopies (>5 meters above the ground) and rarely close to the ground. Finally, shrub/understory species are typically found at intermediate heights in shrubs, in the canopy of small trees, or on the trunks of large trees but do not usually climb higher than four or five meters.

Morphological data. We examined 1365 adult female specimens from species included in the phylogeny for which we had access to undamaged female specimens and/or suitable photographs (n=212 species). We gathered high quality photographs of specimens in dorsal and/or lateral view

from our own collection at the University of Göttingen (Germany), other museum collections, the published literature and other online sources (Dataset S1). Depending on material availability, we measured pictures of between 1 and 20 different individuals per species (mean= 6.4 individuals per species). We collected 23 continuous measurements (Figure S1) that together contained biologically relevant information about overall body size and shape, width and length of different body segments (notably the head), leg length, hindwing size and the length of the subgenital plate (whose function is often related to oviposition). We also qualitatively scored body texture and overall body coloration to include these critical aspects of camouflage when reconstructing the morphospace. The raw and average datasets are available in Dataset S1.

The phasmid morphospace. We built a multidimensional morphospace using a Principal Component Analysis (PCA) mixing continuous and categorical data (PCAmix)(Chavent et al., 2017). Continuous measurements were \log_{10} -transformed. We included body volume and body area as body size measurements. Then, we included the rest of the measured traits after controlling for the effect of size by substituting original trait values with the residuals calculated from a phylogenetically-corrected linear regression against body volume. Because wing length and wing area included zeros for wingless species, we divided the non-transformed measurements by body length or body length squared respectively, to obtain and include measures of relative wing length and area. In total, we included 23 continuous (priorly mean-centered on zero and scaled to unit variance) and three categorical variables (Figure S1).

Definition of morphotypes. We next used our multidimensional morphospace data to cluster species into distinct morphotypes by running a hierarchical clustering algorithm (using the Ward's method) to define morphotypes based on overall proximity on the morphospace. Using the resulting dendrogram, we defined a height threshold to delineate separate groups of morphologically similar taxa (i.e., morphotypes).

Ancestral state reconstruction of morphotypes. Morphotypes, as defined by our clustering analysis, were mapped on our Maximum Clade Credibility (MCC) tree, to establish whether they had single or multiple origins. We ran ancestral state reconstructions using stochastic character mapping as implemented in the R package “phytools” (Revell, 2012). Given the large number of morphotypes ($n=20$), only the “equal rate” transition model (assuming a single transition rate between morphotypes) could be run.

Visualizing morphological convergence. To visualize how pairs of putatively convergent lineages evolved similar trait values and test the predictions laid out in **Figure 1**, we reconstructed ancestral trait values (i.e., ancestral morphospace coordinates) at each node and at 500,000 year time intervals along each branch using a fast estimation of maximum likelihood ancestral state as implemented in “phytools” (Pigot et al., 2020).

Using these ancestral morphospace coordinates, we visualized the trajectories of convergent lineages on the morphospace. Because different morphotypes occupy regions of the morphospace that are best characterized by different axes (e.g., wide-leaf mimics differ from other morphotypes primarily in their relatively wide [PC1] and flat [PC2] bodies), we used the two trait axes that best distinguished each morphotype (Table S4, SI Appendix). We visualized these trajectories for each morphotype with multiple origins, including a representative taxon per independent origin (maximum four). We also visualized how all phasmid lineages colonized the morphospace over time, by making a dynamic 2-D phylomorphospace (Video S1).

To quantify the cumulative distance travelled across the morphospace over time by each possible pair of convergent lineages for a given morphotype, we used the position of each lineage along the two axes for each time step and summed the Euclidean distance between successive time points. In order to visualize if and how much putatively convergent taxa phenotypically diverged before coming back together, we calculated the trait gap between each possible pair of convergent lineages at each time step after they split (i.e., the phenotypic distance between the ancestral states of the two separated lineages).

Quantifying convergence. To quantify the strength of convergence towards specific morphotypes, we calculated the recently proposed C1 to C5 pattern-based metrics as implemented in the R package “convevol” (Stayton, 2015) as well as the Wheatsheaf index (w), implemented in the package “windex” (Arbuckle and Minter, 2015), for each morphotype after selecting one lineage per independent origin. The first four morphospace axes (PC1-PC4) were used. The associated p-values test the hypothesis that convergence is significantly stronger for that morphotype than would be expected by chance.

Association between morphology and habitat. To test for associations between body form and habitat, we first used a nonparametric approach (dynamic range boxes, as implemented in the R package “dynRB”) (Junker et al., 2016) to estimate hypervolumes of morphospace occupied by animals from each vegetation layer and to quantify the extent of morphological overlap of species from

different layers. To estimate volumes, this method calculates the range sizes of each habitat for each dimension and aggregates them by either calculating their product, arithmetic or geometric mean. We plotted the estimated range size for each habitat for the full niche and the first 12 dimensions individually. The overlap between hypervolumes is calculated as the portion of the hypervolume of habitat A covered by the hypervolume of habitat B ($\text{port}(A,B)$) and vice versa ($\text{port}(B,A)$). Using heatmaps, we represented pairwise range overlaps between habitats for the full hypervolumes and for the first six dimensions individually, in order to identify the phasmid morphospace axes best differentiating the three habitat zones.

Then, to quantify the match between morphology and habitat, we tested whether we could predict the habitat of a species given its position on the morphospace using random forest machine learning models. The algorithm was first trained on a randomly downsampled dataset including 75% of all species with available morphological and habitat data and then used to predict habitat for the remaining 25% of the data. We compared the predictive performance of models including one to 15 morphospace axes and repeated this step for 1,000 randomly sampled datasets for each set of the morphospace axes, and averaged prediction accuracies.

Acknowledgements

This study would not have been possible without the work of many passionate phasmid enthusiasts and breeders who, over the years, documented the biology of many species. We are therefore very grateful to the amateur and professional phasmid community for publicly sharing these notes and observations along with many high quality pictures. We are also thankful to Nicolas Cliquennois for insightful discussions regarding Malagasy and Mascarene stick insects. We finally thank Bruno Kneubühler, Nicolas Cliquennois and Paul Brock for permission to use their pictures.

Author Contributions

R.P.B., S.B. and D.J.E. designed research; R.P.B. and S.B. performed research; R.P.B. analyzed data; R.P.B., S.B. and D.J.E. wrote the paper.

Funding

This study was supported by the German Research Foundation (DFG Grants BR 2930/5-1 and BR 2930/7-1 to S.B.).

References

- Agrawal, A. A.** (2017). Toward a predictive framework for convergent evolution: Integrating natural history, genetic mechanisms, and consequences for the diversity of life. *American Naturalist* **190**, S1–S12.
- Arbuckle, K. and Minter, A.** (2015). Windex: Analyzing convergent evolution using the wheatsheaf index in R. *Evolutionary Bioinformatics* **2015**, 11–14.
- Arbuckle, K., Bennett, C. M. and Speed, M. P.** (2014). A simple measure of the strength of convergent evolution. *Methods in Ecology and Evolution* **5**, 685–693.
- Bank, S. and Bradler, S.** (2022). A second view on the evolution of flight in stick and leaf insects (Phasmatodea). *BMC Ecology and Evolution* **22**, 1–17.
- Bank, S., Cumming, R. T., Li, Y., Henze, K., Le Tirant, S. and Bradler, S.** (2021a). A tree of leaves: phylogeny and historical biogeography of the leaf insects (Phasmatodea: Phylliidae). *Communications Biology* **4**, 1–12.
- Bank, S., Buckley, T. R., Büscher, T. H., Bresseel, J., Constant, J., de Haan, M., Dittmar, D., Dräger, H., Kahar, R. S., Kang, A., et al.** (2021b). Reconstructing the nonadaptive radiation of an ancient lineage of ground-dwelling stick insects (Phasmatodea: Heteropterygidae). *Systematic Entomology* **46**, 487–507.
- Bedford, G. O.** (1978). Biology and ecology of the Phasmatodea. *Annual Review of Entomology* **23**, 125–149.
- Blount, Z. D., Lenski, R. E. and Losos, J. B.** (2018). Contingency and determinism in evolution: Replaying life's tape. *Science* **362**, eaam5979.
- Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G. and Emlen, D. J.** (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes* **181**, 104263.
- Boisseau, R. P., Büscher, T. H., Klawitter, L. J., Gorb, S. N., Emlen, D. J. and Tobalske, B. W.** (2022). Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect. *BMC Ecology and Evolution* **22**, 1–18.
- Borko, Š., Trontelj, P., Seehausen, O., Moškrič, A. and Fišer, C.** (2021). A subterranean adaptive radiation of amphipods in Europe. *Nature Communications* **12**, 1–12.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A. and Drummond, A. J.** (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology* **10**, e1003537.
- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.

- Bradler, S., Robertson, J. A. and Whiting, M. F.** (2014). A molecular phylogeny of Phasmatodea with emphasis on Necrosciinae, the most species-rich subfamily of stick insects. *Systematic Entomology* **39**, 205–222.
- Bradler, S., Cliquennois, N. and Buckley, T. R.** (2015). Single origin of the Mascarene stick insects: ancient radiation on sunken islands? *BMC Evolutionary Biology* **15**, 1–10.
- Bright, J. A., Marugán-Lobón, J., Cobb, S. N. and Rayfield, E. J.** (2016). The shapes of bird beaks are highly controlled by nondietary factors. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 5352–5357.
- Brock, P. D. and Hasenpusch, J. W.** (2009). *The complete field guide to stick and leaf insects of Australia*. Collingwood, Victoria, Australia: CSIRO publishing.
- Buckley, T. R., Attanayake, D. and Bradler, S.** (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1055–1062.
- Camousseight, A.** (1995). Revision taxonomica del genero *Agathemera* (Phasmatodea: Pseudophasmatidae) en Chile. *Revista Chilena de Entomología* **22**, 35–53.
- Castiglione, S., Serio, C., Tamagnini, D., Melchionna, M., Mondanaro, A., Di Febbraro, M., Profico, A., Piras, P., Barattolo, F. and Raia, P.** (2019). A new, fast method to search for morphological convergence with shape data. *PLoS ONE* **14**, 1–20.
- Chavent, M., Kuentz-Simonet, V., Labenne, A. and Saracco, J.** (2017). Multivariate Analysis of Mixed Data: The R Package PCAmixdata. *arXiv*.
- Cliquennois, N.** (2008). Révision des Anisacanthidae, famille endémique de phasmes de Madagascar (Phasmatodea: Bacilloidea). *Annales de la Société Entomologique de France* **44**, 59–85.
- Cliquennois, N. and Brock, P.** (2002). *ApteroGraeffea*, un nouveau genre de phasme de la Réunion et de l'île Ronde (Phasmatodea, Platycraninae). *Bulletin de la Société entomologique de France* **107**, 387–395.
- Collar, D. C., Schulte, J. A. and Losos, J. B.** (2011). Evolution of extreme body size disparity in monitor lizards (*Varanus*). *Evolution* **65**, 2664–2680.
- Collar, D. C., Reece, J. S., Alfaro, M. E., Wainwright, P. C. and Mehta, R. S.** (2014). Imperfect morphological convergence: Variable changes in cranial structures underlie transitions to durophagy in moray eels. *American Naturalist* **183**,
- Conte, G. L., Arnegard, M. E., Peichel, C. L. and Schluter, D.** (2012). The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B: Biological Sciences* **279**, 5039–5047.
- Conway Morris, S.** (2003). *Life's solution: inevitable humans in a lonely universe*. Cambridge, United Kingdom: Cambridge University press.
- Conway Morris, S.** (2010). Evolution: Like any other science it is predictable. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 133–145.
- Conway Morris, S.** (2015). *The runes of evolution: how the universe became self-aware*. West Conshohocken, Pennsylvania, USA: Templeton Press.
- Crampton, G. C.** (1916). The lines of descent of the lower Pterygotan insects, with notes on the relationships of the other forms. *Entomological news* **27**, 244–258.
- Cubillos, C. and Vera, A.** (2020). Comparative morphology of the eggs from the eight species in the genus *Agathemera* Stål (Insecta: Phasmatodea), through phylogenetic comparative method approach. *Zootaxa* **4803**, 523–543.

- de Alencar, L. R. V., Martins, M., Burin, G. and Quental, T. B.** (2017). Arboreality constrains morphological evolution but not species diversification in vipers. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171775.
- Dick, M. H., Lidgard, S., Gordon, D. P. and Mawatari, S. F.** (2009). The origin of ascophoran bryozoans was historically contingent but likely. *Proceedings of the Royal Society B: Biological Sciences* **276**, 3141–3148.
- Gillespie, R. G., Benjamin, S. P., Brewer, M. S., Rivera, M. A. J. and Roderick, G. K.** (2018). Repeated diversification of ecomorphs in hawaiian stick spiders. *Current Biology* **28**, 941–947.
- Glaw, F., Hawlitschek, O., Dunz, A., Goldberg, J. and Bradler, S.** (2019). When giant stick insects play with colors: Molecular phylogeny of the Achriopterini and description of two new splendid species (Phasmatodea: Achrioptera) from Madagascar. *Frontiers in Ecology and Evolution* **7**, 1–18.
- Goldberg, J., Bresseel, J., Constant, J., Kneubühler, B., Leubner, F., Michalik, P. and Bradler, S.** (2015). Extreme convergence in egg-laying strategy across insect orders. *Scientific reports* **5**, 1–7.
- Gould, S. J.** (1989). *Wonderful life: the Burgess Shale and the nature of history*. New York, New York, USA: W. W. Norton & Company.
- Grossnickle, D. M., Chen, M., Wauer, J. G. A., Pevsner, S. K., Weaver, L. N., Meng, Q. J., Liu, D., Zhang, Y. G. and Luo, Z. X.** (2020). Incomplete convergence of gliding mammal skeletons. *Evolution* **74**, 2662–2680.
- Hennemann, F. H.** (2020). Megacraniinae—The Palm Stick Insects: A new subfamily of Old World Phasmatodea and a redefinition of Platycraniinae Brunner v. Wattenwyl, 1893 (Phasmatodea: “Anareolatae”). *Zootaxa* **4896**, 151–179.
- Herrel, A., Vanhooydonck, B. and Van Damme, R.** (2004). Omnivory in lacertid lizards: Adaptive evolution or constraint? *Journal of Evolutionary Biology* **17**, 974–984.
- Ingram, T. and Mahler, D. L.** (2013). SURFACE: Detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods in Ecology and Evolution* **4**, 416–425.
- Junker, R. R., Kuppler, J., Bathke, A. C., Schreyer, M. L. and Trutschnig, W.** (2016). Dynamic range boxes – a robust nonparametric approach to quantify size and overlap of n-dimensional hypervolumes. *Methods in Ecology and Evolution* **7**, 1503–1513.
- Kawahara, A. Y. and Rubinoff, D.** (2013). Convergent evolution of morphology and habitat use in the explosive Hawaiian fancy case caterpillar radiation. *Journal of Evolutionary Biology* **26**, 1763–1773.
- Lapedra, O., Sol, D., Carranza, S. and Beaulieu, J. M.** (2013). Behavioural changes and the adaptive diversification of pigeons and doves. *Proceedings of the Royal Society B: Biological Sciences* **280**, 20122893.
- Law, C. J.** (2022). Different evolutionary pathways lead to incomplete convergence of elongate body shapes in carnivorous mammals. *Systematic biology* **71**, 788–796.
- Losos, J. B.** (2010). Adaptive radiation, ecological opportunity, and evolutionary determinism. *American Naturalist* **175**, 623–639.
- Losos, J. B.** (2011). Convergence, adaptation, and constraint. *Evolution* **65**, 1827–1840.
- Losos, J. B.** (2017). *Improbable destinies, fate, chance, and the future of evolution*. New York, New York, USA: Riverhead books.
- Losos, J. B., Jackman, T. R., Larson, A., De Queiroz, K. and Rodríguez-Schettino, L.** (1998). Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115–2118.
- Luque, J., Feldmann, R. M., Vernygora, O., Schweitzer, C. E., Cameron, C. B., Kerr, K. A., Vega, F. J., Duque, A., Strange, M., Palmer, A. R., et al.** (2019). Exceptional preservation of mid-Cretaceous marine arthropods and

- the evolution of novel forms via heterochrony. *Science Advances* **5**, 3875–3899.
- Maginnis, T. L.** (2006). Leg regeneration stunts wing growth and hinders flight performance in a stick insect (*Sipyloidea sipyilus*). *Proceedings of the Royal Society B* **273**, 1811–1814.
- McGhee, G. R.** (2011). *Convergent evolution: limited forms most beautiful*. Cambridge, Massachusetts, USA: MIT Press.
- Meachen-Samuels, J. A.** (2012). Morphological convergence of the prey-killing arsenal of sabertooth predators. *Paleobiology* **38**, 1–14.
- Moldowan, P. D., Copsey, J. A., Zuël, N., Tatayah, V. and Cole, N.** (2016). Sticks and stones: notes on the ecology and conservation of an endemic stick insect (*Apterograeffea marshallae*) and the restoration of an island ecosystem. *Pbelsuma* **24**, 72–79.
- Mugleston, J. D., Naegle, M., Song, H. and Whiting, M. F.** (2018). A comprehensive phylogeny of tettigoniidae (Orthoptera: Ensifera) reveals extensive ecomorph convergence and widespread taxonomic incongruence. *Insect Systematics and Diversity* **2**, 1–27.
- O’Hanlon, J. C., Jones, B. R. and Bulbert, M. W.** (2020). The dynamic eggs of the Phasmatodea and their apparent convergence with plants. *The Science of Nature* **107**, 1–12.
- Ord, T. J. and Summers, T. C.** (2015). Repeated evolution and the impact of evolutionary history on adaptation. *BMC Evolutionary Biology* **15**, 1–12.
- Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., Seddon, N., Trisos, C. H., Weeks, B. C. and Tobias, J. A.** (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. *Nature Ecology & Evolution* **4**, 230–239.
- Revell, L. J.** (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223.
- Rincon-Sandoval, M., Duarte-Ribeiro, E., Davis, A. M., Santaquiteria, A., Hughes, L. C., Baldwin, C. C., Soto-Torres, L., Acero, A. P., Walker, H. J., Carpenter, K. E., et al.** (2020). Evolutionary determinism and convergence associated with water-column transitions in marine fishes. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 33396–33403.
- Robertson, J. A., Bradler, S. and Whiting, M. F.** (2018). Evolution of oviposition techniques in stick and leaf insects (Phasmatodea). *Frontiers in Ecology and Evolution* **6**, 1–15.
- Robinson, M. H.** (1968). The defensive behavior of *Pterinoxylus spinulosus* Redtenbacher, a winged stick insect from Panama (Phasmatodea). *Psyche* **75**, 195–207.
- Ruxton, G. D., Allen, W. L., Sherratt, T. N. and Speed, M. P.** (2019). *Avoiding attack: the evolutionary ecology of crypsis, aposematism, and mimicry*. Oxford, UK: Oxford University Press.
- Schluter, D.** (2000). *The Ecology of Adaptive Radiation*. Oxford, UK: Oxford University Press.
- Simon, S., Letsch, H., Bank, S., Buckley, T. R., Donath, A., Liu, S., Machida, R., Meusemann, K., Misof, B., Podsiadlowski, L., et al.** (2019). Old world and new world Phasmatodea: phylogenomics resolve the evolutionary history of stick and leaf insects. *Frontiers in Ecology and Evolution* **7**, 1–14.
- Skelhorn, J., Rowland, H. M., Speed, M. P. and Ruxton, G. D.** (2010). Masquerade: camouflage without crypsis. *Science* **327**, 51.
- Speed, M. P. and Arbuckle, K.** (2017). Quantification provides a conceptual basis for convergent evolution. *Biological Reviews* **92**, 815–829.

-
- Stayton, C. T.** (2006). Testing Hypotheses of Convergence With Multivariate Data: Morphological and Functional Convergence Among Herbivorous Lizards. *Evolution* **60**, 824.
- Stayton, C. T.** (2015). The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* **69**, 2140–2153.
- Svenson, G. J. and Whiting, M. F.** (2009). Reconstructing the origins of praying mantises (Dictyoptera, Mantodea): the roles of Gondwanan vicariance and morphological convergence. *Cladistics* **25**, 468–514.
- Tihelka, E., Cai, C., Giacomelli, M., Pisani, D. and Donoghue, P. C. J.** (2020). Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatodea) reveal a Permian–Triassic co-origination with insectivores. *Royal Society Open Science* **7**, 201689.
- Trontelj, P., Blejec, A. and Fišer, C.** (2012). Ecomorphological convergence of cave communities. *Evolution* **66**, 3852–3865.
- Vermeij, G. J.** (2006). Historical contingency and the purported uniqueness of evolutionary innovations. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 1804–1809.
- Webster, R. J., Hassall, C., Herdman, C. M., Godin, J. G. J. and Sherratt, T. N.** (2013). Disruptive camouflage impairs object recognition. *Biology Letters* **9**, 1–5.
- Wolfe, J. M., Luque, J. and Bracken-Grissom, H. D.** (2021). How to become a crab: Phenotypic constraints on a recurring body plan. *BioEssays* **43**, 2100020.
- Zeng, Y., Malley, C. O., Singhal, S., Rahim, F., Park, S., Chen, X. and Dudley, R.** (2020). A tale of winglets: evolution of flight morphology in stick insects. *Frontiers in Ecology and Evolution* **8**, 1–15.

Supplementary Information

Supplementary Materials and Methods

We used R software (v4.1.1) (R Core Team, 2021) for most analyses and data manipulation, as described below.

DNA sequence alignment. DNA sequences were obtained for 314 species of Phasmatodea and one species of Embioptera (the sister order of Phasmatodea (Misof et al., 2014)). We sampled regions of 7 genes: nuclear 18S rRNA (18S), 28S rRNA (28S) and histone subunit 3 (H3), and mitochondrial 12S rRNA (12S), 16S rRNA (16S), cytochrome-c oxidase subunit I (COI) and cytochrome-c oxidase subunit II (COII). Our species list was largely similar to the taxon list of a previous study (Robertson et al., 2018) (Dataset S1). Sequences were extracted from Genbank and corresponded to data generated in previously published studies (Bank et al., 2021; Bradler et al., 2014; Bradler et al., 2015; Forni et al., 2021; Glaw et al., 2019; Robertson et al., 2018). They were reoriented and aligned using the R package “DECIPHER” (Wright, 2016). Protein coding sequences (i.e., H3, COI, COII) were first converted into amino acid sequences, aligned and reverse translated (*AlignTranslation*: “DECIPHER”; *R function*: “R package”). Frameshift errors occurring after the accidental insertion or deletion of one or two nucleotides during DNA sequencing were corrected after comparison of each sequence with an arbitrarily chosen unshifted sequence (*CorrectFrameshifts*: “DECIPHER”). Ribosomal genes (i.e., 18S, 28S, 12S and 16S) were first transcribed into RNA and then aligned while taking into account the secondary structure of non-coding RNA (*AlignSeqs*: “DECIPHER”). Gappy columns were then removed if they contained less than 5% nucleotide (*del.colgapsonly*: “ape”) (Paradis and Schliep, 2019). As alignment extremities varied considerably in length, we trimmed each multiple sequence alignment (MSA) from the beginning and end using MACSE 2.05 (Ranwez et al., 2018) until a 30% nucleotide coverage was achieved (`java -jar macse.jar -prog trimAlignment -align alignment.fasta -min_percent_NT_at_ends 0.3`). All alignments were visually checked for conspicuously misaligned sections using Geneious v.2021.1 (www.geneious.com) and these were manually corrected if possible or otherwise excluded (n=13 excluded), and the MSAs were concatenated (*concatenate*: “ape”) (Schliep et al., 2020).

Partition scheme. To find the best-fitting partition scheme, we used PartitionFinder 2 (Lanfear et al., 2017) which separated the data into five subsets: mitochondrial rRNA 12S+16S (GTR+I+G+X), nuclear rRNA 18S+28S (GTR+I+G+X), protein codon positions 1+2 (GTR+I+G+X), COX position 3 (GTR+G+X), and H3 position 3 (HKY+I+G+X).

Phylogenetic reconstruction. Phylogenetic reconstructions were performed using this combined and partitioned dataset of 6,778 bp on the Cipres Science Gateway (www.phylo.org/) (Miller et al., 2011). Tree and divergence time estimation through Bayesian analyses were simultaneously performed in BEAST 2 (v. 2.6.3) (Bouckaert et al., 2014). The optimal substitution model for each partition was selected using the bModelTest package in BEAST (v.1.2.1, “allreversible” (Bouckaert and Drummond, 2017)) and the partition scheme inferred by PartitionFinder.

We used 6 unambiguous crown-group phasmid fossils as minimum calibration points that were assigned to specific nodes on the tree using synapomorphy-based anatomical evidence (Table S1). We modeled all fossil constraints as a lognormal prior distribution with the minimum age as offset, a log-mean of 1.0 and a log-SD of 1.0. Tree and clock parameters were linked across all partitions and the analysis used an uncorrelated log-normal relaxed molecular clock model (UCLD) with a Yule model of speciation prior. We constrained our phylogenetic inferences to adopt the same backbone topology between major subclades as the most recent transcriptomic study (Tihelka et al., 2020). The analysis was run for 200 million generations with parameters and trees sampled every 500 generations. Convergence was assessed in Tracer v.1.7.1 (Rambaut et al., 2018) (ESS>200) and a Maximum Clade Credibility (MCC) tree was built in TreeAnnotator v.2.6.6 after the removal of the first 10% of trees as burn in.

Habitat data. Despite being charismatic and popular in museums, private collections, and as pets, most phasmids have not been studied in the wild, and the habitat use of many species is surprisingly poorly understood (Bradler and Buckley, 2018). Most phasmids are active at night, and very little is known about these behaviors. However, it is during the daytime while these insects are resting that they are exposed to visually hunting predators (including entomologists), and collection records sometimes include information about the locations of these animals at the time they are found.

Therefore, we chose to broadly classify the habitat of stick insects as the typical vegetation layer in which they rest during the daytime. We surveyed the literature, field guides and online databases for observations of where each species is typically found (see Dataset S1) and defined three broad

categories: ground/shrub, shrub/understory and understory/canopy. These habitats included overlapping names to emphasize that boundaries between inhabited vegetation layers are loose as phasmids often navigate between them. Ground/shrub dwelling taxa were defined as often resting below 1.5 meters above the ground. These taxa can be found in the leaf litter but more typically laying on the base of trunks, mossy logs, under bark, on low vegetation or in tall grass. Understory/canopy dwelling taxa are typically found high up in the tree canopies (>5 meters above the ground) and rarely close to the ground. Finally, shrub/understory species are typically found at intermediate heights in shrubs, in the canopy of small trees, or on trunk of large trees but do not climb higher than four or five meters. However, it should be acknowledged that this broad classification does not account for the full diversity of substrates and host plants that phasmids live on. Future refinements are desirable as more information on the natural history of phasmids becomes available.

Morphological data. We measured 1,359 adult female specimens from 212 species included in the phylogeny. Males are less morphologically variable than females and are not considered in this study, as sexual dimorphism is extremely strong and variable across stick insect species and is the focus of a companion study (Chapter 2). We gathered high quality photographs of live or pinned specimens in dorsal and/or lateral view from our own collection at the University of Göttingen (Germany), field guides (Brock and Hasenpusch, 2009; Seow-Choen, 2005; Seow-Choen, 2016; Seow-Choen, 2017; Seow-Choen, 2018), the published literature, and various online databases (Brock et al., 2021) (Dataset S1). Pictures were only included if the body of the specimen was relatively straight and undamaged. We measured between 1 and 18 different individuals per species (mean = 5.5 individuals per species). To standardize scale on each photograph (which were sometimes unscaled), we measured each of the 23 continuous traits described in Figure S1 relative to body length using ImageJ (v1.51)(Schneider et al., 2012). We collected data on body length (excluding ovipositor and subgenital plate) from the literature in addition to directly measuring it on properly scaled photographs (Dataset S1). We then took the median of each relative measurement across all females of a given species and we multiplied this by the median of body length to obtain absolute measurements for each species. We checked the congruence between data obtained from specimens in our collection at the University of Göttingen and specimens obtained from all other sources (Figure S26), and between data collected from pinned specimens and from photographs of live animals (Figure S27). In all cases, our measurements were highly congruent across sources.

Finally, we qualitatively scored body texture and overall body coloration (Dataset S1). We assigned a separate, binary score to the dorsal surface of the mesothorax and abdomen for texture (0 = smooth, 1 = rough and/or spiny). We qualitatively scored the dominant body coloration of each species using pictures of live specimens in captivity or in the wild (wings folded), as “green” (i.e., from light to dark green), “brown” (i.e., from light to dark brown, orange and black) or “other” (i.e., aposematic/contrasting coloration). Most phasmids exhibit colorations ranging from light green to dark brown or even black, while a few purportedly aposematic taxa have evolved conspicuous and contrasting coloration patterns including red, yellow or white.

Because body coloration is often extremely plastic in phasmids, we duplicated taxa that exhibited several color morphs that fit in separate categories, as well as taxa exhibiting a mixed brown/green color pattern or with no unambiguously dominant color. We acknowledge that such a broad qualitative assessment of coloration by human observers is limited and substantially biased as predators may display a different vision system (e.g., tetrachromatic vision of birds). However, more accurate and objective quantifications (e.g., using spectrograms) were beyond the scope of this study as they would require calibrated pictures and/or live material for a large number of species, and this is not possible at present. We intentionally kept the color categories broad, and assigned any ambiguous taxon to multiple categories, to avoid misclassifications.

The phasmid morphospace. To build a multidimensional morphospace, we performed a Principal Component Analysis (PCA) mixing continuous and categorical data (R package “PCAmixdata”, v3.1)(Carvalho and Cardoso, 2020; Chavent et al., 2017). The analysis was performed on \log_{10} -transformed continuous trait values, except for wing length and wing area (see below). We included body volume and body area as body size measurements. Then, we included the rest of the quantitative traits after controlling for the effect of size by substituting original trait values with the residuals calculated from a phylogenetically-corrected linear regression against body volume (*pgl*s: “caper”). In these regressions, branch length transformations were optimized using maximum likelihood (lambda=’ML’). Because wing length and wing area included zeros for wingless species, we divided the original measurements by body length or body length squared respectively, to obtain measures of relative wing length and area. We mean-centered and scaled to unit variance the continuous variables. In total, we included 23 continuous and three categorical variables (Figure S1).

PCA results can be misleading when the data considered have a phylogenetic structure (Revell, 2009). Phylogenetic correction is not available for the PCAmix procedure, so we used another method to assess the effect of phylogenetic bias on some of our estimates of habitat hypervolumes on the morphospace (see below in subsection “Association between morphology and habitat”). The coordinates from the PCAmix (including both quantitative and qualitative variables) and the corresponding standard PCA (including only quantitative variables) exhibited very similar phylogenetic signals (average lambda over the first 15 PCAs = 0.84 and 0.82 respectively) which suggested that using a phylogenetic PCA (*phyl.pca*: ”phytools”, method=”lambda”) on only the continuous variables should provide a satisfactory estimate of phylogenetic bias in our analyses.

Definition of morphotypes. We next used our multidimensional morphospace data to cluster species into distinct morphotypes by running a hierarchical clustering algorithm (*hclust*: “stats”) using the Ward’s method to define morphotypes based on overall proximity on the morphospace (defined by the first 15 PC axes). Using the resulting dendrogram, we arbitrarily defined a height threshold to define separate groups of morphologically similar taxa (thereafter called “morphotypes”) (**Figure 4**).

To describe the morphological niche occupied by each morphotype, we identified the morphospace axes that best distinguished them (Table S4). Using the full dataset, we trained random forest models to classify a species in either a given morphotype of interest or in a different one, given the first ten axes of the morphospace. This machine learning method classifies species into either belonging to a morphotype of interest or not while making minimal assumptions regarding the shape of the regions occupied by these two groups on the morphospace. It also accounts for interactions between trait axes. The algorithm uses decision trees that partition the morphospace into a set of non-overlapping rectangular hypervolumes inside which the group heterogeneity is minimized. Internal nodes in a decision tree correspond to a split along one dimension of the morphospace and terminal nodes correspond to a unique non-overlapping hypervolume. The group of each species is then “voted” by each decision tree in the random forest based on its coordinates on the morphospace. Species are then assigned to a group by counting the votes of the forest of decision trees. After training on the full dataset, the algorithm evaluates its classification accuracy by comparing its predictions relative to the observed group of each species. It then estimates the relative importance of each dimension to distinguish a given morphotype by quantifying the decrease in accuracy followed by the omission of a given dimension. Thus, we could identify the axes of the morphospace that were relatively more important in distinguishing each morphotype from the rest.

Ancestral state reconstruction of morphotypes. Morphotypes, as defined by our hierarchical clustering analysis, were mapped on the MCC tree to establish whether they had single or multiple origins. We ran ancestral state reconstructions using stochastic character mapping as implemented in the R package “phytools” (Revell, 2012). Given the large number of morphotypes ($n=20$), only the “equal rate” transition model (assuming a single transition rates between morphotypes) could be run. The transition matrix was calculated using MCMC, the prior distribution on the root node of the tree was estimated and 1,000 stochastic character maps were subsequently simulated and summarized to get posterior probabilities for each state at each node.

Visualizing morphological convergence. To visualize how putatively convergent lineages evolved similar trait values, we reconstructed ancestral trait values at each node on the MCC tree (*fastAnc*, “phytools”)(Revell, 2012) and extrapolated trait values from node to node to estimate ancestral values along all the branches of the tree at 500,000 year time intervals (Pigot et al., 2020). Using these ancestral morphospace coordinates, we visualized the trajectories of convergent lineages on the morphospace, for each morphotype. Because different morphotypes occupy regions of the morphospace that are best characterized by different axes (e.g., wide-leaf mimics differ from other morphotypes primarily in their relatively wide [PC1] and flat [PC2] bodies, while large-headed sticks are defined primarily by their large [PC5] and elongated [PC10] heads), we used the two trait axes that best distinguished each morphotype (Table S4). We visualized these trajectories for each morphotype with multiple origins, including a representative taxon per independent origin (maximum four). If more than four lineages were available, we chose the most distantly related ones.

After estimating the position of each lineage on the morphospace at each 500,000 year time step starting with the most recent common ancestor of Phasmatodea, we built a dynamic 2-D phylomorphospace (including PC1 and PC2) to show how lineages diversified morphologically over time (Video S1). At each time step, the morphotype of each lineage was predicted given its position on the morphospace using a random forest model, initially trained with the full extant dataset and the first five dimensions of the morphospace.

To quantify the cumulative distance travelled across the morphospace over time by pairs of convergent lineages for a given morphotype, we used the position of each lineage along the two axes for each time step and summed the Euclidean distance between successive time points. In order to visualize if and how much putatively convergent taxa phenotypically diverged before coming back

together, we calculated the trait gap between each possible pair of convergent lineages at each time step after they split (i.e., the phenotypic distance along the two axes between the ancestral states of the two separated lineages).

These methods allowed us to visualize whether the daughter lineages diverged first and then converged (true convergence), whether they diverged initially and then subsequently evolved in tandem (parallel shift), or whether they stayed in a same region after splitting (stasis) (**Figure 1**).

Quantifying convergence: pattern-based tests. For each morphotype for which we identified at least two independent origins, we first calculated the Wheatsheaf index (w), corresponding to the ratio of the average pairwise phenotypic distance between all taxa in the tree, to the average pairwise distance between all putatively convergent taxa, correcting for phylogenetic relatedness (Arbuckle et al., 2014). We used the R package “windex” to calculate w for each morphotype (Arbuckle and Minter, 2015). The first four morphospace axes (PC1-PC4) were used. For each convergent morphotype, we compared the obtained index to the ones calculated from random distributions of phenotypic values on the same tree topology obtained from 1,000 bootstrap resamples. The associated p-values test the hypothesis that convergence is significantly stronger for that morphotype than would be expected by chance.

We then calculated four convergence metrics, C1-C4, using the “convevol” R package (Stayton, 2015) for each convergent morphotype. In contrast with w , C1-C4 measure the increase in similarity between the convergent taxa through time and are therefore useful to distinguish convergent evolution from stasis, which the Wheatsheaf index is incapable of doing. C1-C4 use ancestral state reconstruction via a Brownian model of trait evolution to compare the distances between phylogenetic tips in the phylomorphospace and the distances between ancestral nodes. C1 is calculated as $C1 = 1 - \frac{D_{tip}}{D_{max}}$, where D_{tip} is the present Euclidean distance between the taxa of interest on the phylomorphospace and D_{max} the maximum distance between any two pairs of taxa along those two lineages (extant or ancestors). C1 therefore ranges from 0 to 1 and quantifies the phenotypic distance that has been reduced by convergent evolution over time. C1 close to 1 indicates strong convergence as it suggests that the extant taxa are much more phenotypically similar than were their ancestors. C2, calculated as $C2 = D_{max} - D_{tip}$, measures the absolute magnitude of convergence. C3 and C4 are standardized versions of C2. C3 measures the magnitude of convergence relative to the sum of phenotypic distances travelled on the morphospace by all the lineages starting from the common ancestor of the two taxa

of interest, and is calculated as $C3 = \frac{C2}{L_{tot,clade}}$. And finally, $C4 = \frac{C2}{L_{tot,tree}}$ quantifies the amount of convergence scaled by the total phenotypic evolution in the entire tree.

It should be noted that C1-C4 can only be calculated for pairs of taxa. Therefore, for each morphotype, we chose one taxon per convergent lineage (i.e., per independent origin as determined by the ancestral state reconstruction of morphotypes). Chosen species were the closest on the morphospace. C1-C4 were calculated for all possible lineage pairs from PC1-PC4 and then averaged for each morphotype.

Finally, in addition to quantifying the magnitude of convergence, we also assessed its frequency into specific regions of the morphospace (i.e., PC1-PC4), using the C5 metric (*convnum*: “convevol”) (Stayton, 2015). C5 measures the number of lineages that enter/cross from outside into a hypervolume of the morphospace defined by the convergent taxa. To assess significance of our values of C1-C5, we ran 1,000 simulations of character evolution along the phylogeny using a Brownian Motion model with the variance–covariance matrix determined from the data. Returned p-values test the hypothesis that convergence is more frequent in a certain region of a morphospace than would be expected randomly (i.e., with no constraint on the direction of evolution).

Quantifying convergence: Process-based tests. As an alternative to running a clustering method to define morphotypes followed by quantitative measures of convergence, we ran the SURFACE algorithm (R package “surface”) (Ingram and Mahler, 2013), a method that does not require *a priori* designation of convergent taxa. It identifies instances of convergent evolution by fitting Ornstein-Uhlenbeck (OU) models of character evolution and reconstructing selective regimes acting on different branches of the phylogeny (Hansen, 1997; Ingram and Mahler, 2013). It therefore explicitly assumes that selective processes are driving convergence and identifies adaptive peaks towards which different lineages are evolving. The algorithm first fits increasingly complex OU models by adding regime shifts along the branches of the tree based on the stepwise reduction of the Akaike Information Criterion (AICc). This ‘forward’ step ends when no additional regime shift can further improve the model. Then, in the ‘backward’ step, the algorithm tries to collapse all pairwise regimes and only merge those that further improve the model in a similar stepwise manner based on AICc. Thus, it determines which of the selective regimes identified in the forward phase are better considered identical –i.e., convergent. We ran SURFACE on the first four axes of the morphospace (PC1-PC4) as advised by Ingram & Mahler (2013) to maximize performance (see Supplementary results).

Association between morphology and habitat. To quantify morphospace occupation by species living in the different vegetation layers, we estimated 12-dimensional hypervolumes (the first 12 PC axes account for 95% of the total variation) using dynamic range boxes (R package "dynRB") (Junker et al., 2016). This non-parametric quantile-based approach does not assume any particular distribution of the data but rather considers the distribution of the observed values and is robust to outliers (Junker et al., 2016). To calculate volumes, it calculates the range sizes of each habitat for each dimension and aggregates them by either calculating their product, arithmetic or geometric mean. The overlap between hypervolumes is calculated as the portion of the hypervolume of habitat A covered by the hypervolume of habitat B ($\text{port}(A,B)$) and vice versa ($\text{port}(B,A)$). Hypervolumes and overlap values are bounded between 0 and 1. Using heatmaps, we represented pairwise overlaps between habitats for the full hypervolumes and for the first six dimensions individually, in order to visualize the morphospace axes best differentiating the three habitat zones. Estimated range size for each habitat was also plotted for the full niche and the first 12 dimensions individually.

Given that the PCAmix procedure used to build the morphospace does not allow for phylogenetic correction, we assessed the effect of phylogenetic bias on our volume estimates for each habitat following the methodology of Nordén et al. (2019). We compared habitat hypervolumes (calculated with all principal component axes) estimated by a standard PCA (*prcomp*: "stats") with those estimated using a phylogenetic PCA (*phyl.pca*: "phytools", method="lambda"), excluding the qualitative variables texture and color. Overall, volume estimates were very similar between the standard and phylogenetic PCAs (Figure S28).

To test whether we could predict the habitat of a species given its position on the morphospace, we used machine learning random forest models (R package "randomForest" v4.6-14) (Liaw and Wiener, 2002). The algorithm was first trained (number of trees=500, number of variables tried at each split= 2, maximum number of terminal nodes= 50) on a randomly downsampled dataset including 75% of all species with available morphological and habitat data (n=274 taxa including duplicated species showing color polymorphism), and then used to predict habitat for the remaining 25% of the data. We compared the predictive performance of models (i.e., classification accuracy) including one to 15 morphospace axes and repeated this step for 1,000 randomly sampled datasets for each set of the morphospace axes, and averaged prediction accuracies (Figure S24).

Supplementary results

Quantifying convergence: Process-based tests. The SURFACE algorithm ran on the first four axes of the morphospace, and found 35 different selective regimes on the MCC tree, 20 of which were identified as convergent (Figure S29). Inferred convergent adaptive peaks broadly corresponded to the different morphotype classes, and, overall, SURFACE confirmed the large amount of convergent evolution in the group. Incongruences may be attributed to the fact that SURFACE only considered the first four axes of the morphospace, while the hierarchical clustering approach considered the first 15 axes. Furthermore, the SURFACE algorithm incorrectly assumes independence between the morphological axes which can lead to model overfitting and misspecification, so these results should be interpreted with caution (Adams and Collyer, 2018).

SI references

- Adams, D. C. and Collyer, M. L.** (2018). Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations. *Systematic Biology* **67**, 14–31.
- Arbuckle, K. and Minter, A.** (2015). Windex: Analyzing convergent evolution using the wheatsheaf index in R. *Evolutionary Bioinformatics* **2015**, 11–14.
- Arbuckle, K., Bennett, C. M. and Speed, M. P.** (2014). A simple measure of the strength of convergent evolution. *Methods in Ecology and Evolution* **5**, 685–693.
- Bank, S., Buckley, T. R., Büscher, T. H., Bresseel, J., Constant, J., de Haan, M., Dittmar, D., Dräger, H., Kahar, R. S., Kang, A., et al.** (2021). Reconstructing the nonadaptive radiation of an ancient lineage of ground-dwelling stick insects (Phasmatodea: Heteropterygidae). *Systematic Entomology* **46**, 487–507.
- Bouckaert, R. R. and Drummond, A. J.** (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology* **17**, 1–11.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A. and Drummond, A. J.** (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology* **10**, e1003537.
- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.
- Bradler, S., Robertson, J. A. and Whiting, M. F.** (2014). A molecular phylogeny of Phasmatodea with emphasis on Necrosciinae, the most species-rich subfamily of stick insects. *Systematic Entomology* **39**, 205–222.
- Bradler, S., Cliquennois, N. and Buckley, T. R.** (2015). Single origin of the Mascarene stick insects: ancient radiation on sunken islands? *BMC Evolutionary Biology* **15**, 1–10.
- Brock, P. D. and Hasenpusch, J. W.** (2009). *The complete field guide to stick and leaf insects of Australia*. Collingwood, Victoria,

- Australia: CSIRO publishing.
- Brock, P. D., Büscher, T. H. and Baker, E.** (2021). Phasmida Species File Online. *Version 5.0/5.0*.
- Carvalho, J. C. and Cardoso, P.** (2020). Decomposing the Causes for Niche Differentiation Between Species Using Hypervolumes. *Frontiers in Ecology and Evolution* **8**, 1–7.
- Chavent, M., Kuentz-Simonet, V., Labenne, A. and Saracco, J.** (2017). Multivariate Analysis of Mixed Data: The R Package PCAmixdata. *arXiv*.
- Engel, M. S., Wang, B. and Alqarni, A. S.** (2016). A thorny, “anareolate” stick-insect (Phasmatidae s.l.) in Upper Cretaceous amber from Myanmar, with remarks on diversification times among Phasmatodea. *Cretaceous Research* **63**, 45–53.
- Forni, G., Martellosi, J., Valero, P., Hennemann, F. H., Conle, O., Luchetti, A. and Mantovani, B.** (2021). Macroevolutionary Analyses Provide New Evidences of Phasmids Wings Evolution as a Reversible Process. *bioRxiv*.
- Glaw, F., Hawlitschek, O., Dunz, A., Goldberg, J. and Bradler, S.** (2019). When giant stick insects play with colors: Molecular phylogeny of the Achriopterini and description of two new splendid species (Phasmatodea: Achrioptera) from Madagascar. *Frontiers in Ecology and Evolution* **7**, 1–18.
- Hansen, T. F.** (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**, 1341–1351.
- Ingram, T. and Mahler, D. L.** (2013). SURFACE: Detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. *Methods in Ecology and Evolution* **4**, 416–425.
- Junker, R. R., Kuppler, J., Bathke, A. C., Schreyer, M. L. and Trutschnig, W.** (2016). Dynamic range boxes – a robust nonparametric approach to quantify size and overlap of n-dimensional hypervolumes. *Methods in Ecology and Evolution* **7**, 1503–1513.
- Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. and Calcott, B.** (2017). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution* **34**, 772–773.
- Liaw, A. and Wiener, M.** (2002). Classification and regression by randomForest. *R News* **2**, 18–22.
- Miller, M. A., Pfeiffer, W. and Schwartz, T.** (2011). The CIPRES Science Gateway: A Community Resource for Phylogenetic Analyses. In *Proceedings of the 2011 TeraGrid Conference on Extreme Digital Discovery - TG '11*, p. New York, New York, USA: ACM Press.
- Misof, B., Liu, S., Meusemann, K., Peters, R. S., Flouri, T., Beutel, R. G., Niehuis, O. and Petersen, M.** (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767.
- Nel, A. and Delfosse, E.** (2011). A new Chinese Mesozoic stick insect. *Acta Palaeontologica Polonica* **56**, 429–432.
- Nordén, K. K., Faber, J. W., Babarović, F., Stubbs, T. L., Selly, T., Schiffbauer, J. D., Peharec Štefanić, P., Mayr, G., Smithwick, F. M. and Vinther, J.** (2019). Melanosome diversity and convergence in the evolution of iridescent avian feathers—Implications for paleocolor reconstruction. *Evolution* **73**, 15–27.
- Paradis, E. and Schliep, K.** (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528.
- Pigot, A. L., Sheard, C., Miller, E. T., Bregman, T. P., Freeman, B. G., Roll, U., Seddon, N., Trisos, C. H., Weeks, B. C. and Tobias, J. A.** (2020). Macroevolutionary convergence connects morphological form to ecological function in birds. *Nature Ecology & Evolution* **4**, 230–239.

- Poinar, G.** (2011). A walking stick, *Clonistria dominicana* n. sp. (phasmatoidea: Diapheromeridae) in Dominican amber. *Historical Biology* **23**, 223–226.
- R Core Team** (2021). R Development Core Team. R: *A Language and Environment for Statistical Computing*.
- Rambaut, A., Drummond, A. J., Xie, D., Baele, G. and Suchard, M. A.** (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* **67**, 901–904.
- Ranwez, V., Douzery, E. J. P., Cambon, C., Chantret, N. and Delsuc, F.** (2018). MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. *Molecular Biology and Evolution* **35**, 2582–2584.
- Revell, L. J.** (2009). Size-correction and principal components for interspecific comparative studies. *Evolution* **63**, 3258–3268.
- Revell, L. J.** (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223.
- Robertson, J. A., Bradler, S. and Whiting, M. F.** (2018). Evolution of oviposition techniques in stick and leaf insects (Phasmatoidea). *Frontiers in Ecology and Evolution* **6**, 1–15.
- Schliep, K., Jombart, T., Kamvar Namir, Z., Archer, E. and Harris, R.** (2020). apex: Phylogenetic Methods for Multiple Gene Data.
- Schneider, C. A., Rasband, W. S. and Eliceiri, K. W.** (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.
- Sellick, J. T. C.** (1994). Phasmida (stick insect) eggs from the Eocene of Oregon. *Palaeontology* **37**, 913–921.
- Seow-Choen, F.** (2005). *Phasmids of Peninsular Malaysia and Singapore*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F.** (2016). *A Taxonomic Guide to the Stick Insects of Borneo: Including New Genera and Species*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F.** (2017). *A Taxonomic Guide to the Stick Insects of Singapore*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F.** (2018). *A Taxonomic Guide to the Stick Insects of Sumatra, Volume 1*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Stayton, C. T.** (2015). The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* **69**, 2140–2153.
- Tihelka, E., Cai, C., Giacomelli, M., Pisani, D. and Donoghue, P. C. J.** (2020). Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatoidea) reveal a Permian–Triassic co-origination with insectivores. *Royal Society Open Science* **7**, 201689.
- Wedmann, S., Bradler, S. and Rust, J.** (2007). The first fossil leaf insect: 47 Million years of specialized cryptic morphology and behavior. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 565–569.
- Wright, E. S.** (2016). Using DECIPHER v2. 0 to analyze big biological sequence data in R. *The R Journal* **8**, 1–8.

Table S1: Fossil calibrations used in the divergence time estimates. Fossils are numbered according to **Figure 2**.

Fossil	Life stage	Reference	Formation	Minimum age (Ma)	Calibration Node	Comments
<i>Renphasma sinica</i>	Adult	(Nel and Delfosse, 2011)	Yixian formation, Liaonign, China	122	Root of tree, divergence of Embioptera and Phasmatodea (fossil 1)	Unambiguous crown group phasmatodean due to presence of the vomer between a pair of unsegmented cerci.
<i>Echinosomiscus primoticus</i>	Adult	(Engel et al., 2016)	Upper Cretaceous Amber, Myanmar	98.8	Divergence of Euphasmatodea and Timematodea (fossil 2)	Euphasmatodea with uncertain affinity (presumed male Neophasmatodea)
<i>Eophyllum messelensis</i>	Adult	(Wedmann et al., 2007)	Messel Germany deposits	47	Phyllinae (fossil 3)	Unambiguous crown leaf insect
<i>Eophasma spp.</i>	Eggs	(Sellick, 1994)	Eocene Clarno Formation Nut Beds of Oregon	44	Pseudophasmatinae (fossil 4)	Euphasmatodean eggs most similar to Anisomorphini (Pseudophasmatinae)
<i>Clonistria sp.</i>	Eggs	(Poinar, 2011)	Dominican amber, Hispaniola	20-40 [40]	Diapheromerinae (fossil 5)	Unambiguous Diapheromerinae egg due to vesicular/matrix capitulum
<i>Malacomorpha sp.</i>	Eggs	(Poinar, 2011)	Dominican amber, Hispaniola	20-40 [20]	<i>Malacomorpha</i> (fossil 6)	Egg most similar to extant genus <i>Malacomorpha</i> (Pseudophasmatinae)

Table S2: List of taxa represented in **Figure 2**.

1. *Timema californicum*
2. *Abrosoma festinatum*
3. *Dajaca napolovi*
4. *Dinophasma saginatum*
5. *Oreophoetes peruana*
6. *Diapheromera femorata*
7. *Cladomorphus phyllinus*
8. *Trychopeplus laciniatus*
9. *Cranidium gibbosum*
10. *Agathemera crassa*
11. *Metriophasma diocles*
12. *Paraprisopus antillarum*
13. *Anisomorpha buprestoides*
14. *Prisopus ariadne*
15. *Pseudophasma rufipes*
16. *Malacomorpha jamaicana*
17. *Peruphasma schultei*

18. *Epidares nolimentangere*
19. *Heteropteryx dilatata*
20. *Haaniella dehaanii*
21. *Aretaon asperrimus*
22. *Bacillus rossius*
23. *Spathomorpha adefa*
24. *Parectatosoma mocquerysi*
25. *Achrioptera punctipes*
26. *Cryptophyllum celebicum*
27. *Pulchriphyllum giganteum*
28. *Spinohirasea bengalensis*
29. *Kalocorinnis wegneri*
30. *Pseudodiacantha macklottii*
31. *Diesbachia tamyris*
32. *Phaenopharos struthioneus*
33. *Carausius morosus*
34. *Thaumatobactron guentheri*
35. *Eurycantha calcarata*
36. *Bactrododema hecticum*
37. *Taraxippus perezgelaberti*
38. *Diapherodes gigantea*
39. *Lamponius guerini*
40. *Phobaeticus serratipes*
41. *Parapachymorpha spiniger*
42. *Sceptrophasma hispidulum*
43. *Medaura sabriuscula*
44. *Clonaria conformans*
45. *Macrophasma biroi*
46. *Eurycnema goliath*
47. *Tropidoderus childrenii*
48. *Phasma gigas*
49. *Megacrania batesii*
50. *Extatosoma tiaratum*
51. *Dryococelus australis*
52. *Rhaphiderus spiniger*
53. *Apterograeffea reunionensis*
54. *Argosarchus horridus*
55. *Cnipsus rachis*

Table S3: List of taxa represented in **Figure 4**.

1. *Pulchriphyllum pulchrifolium*
2. *Nanophyllum frondosum*
3. *Microcanachus matileorum*
4. *Timema californicum*
5. *Prisopus ariadne*

6. *Ommatopseudes harmani*
7. *Epidares nolimetangere*
8. *Orestes bachmaensis*
9. *Agathemera crassa*
10. *Paraprisopus antillarum*
11. *Abrosoma festinatum*
12. *Dinophasma saginatum*
13. *Peruphasma schultei*
14. *Anisomorpha buprestoides* (brown morph)
15. *Macrophasma biroi*
16. *Diapherodes gigantea*
17. *Heteropteryx dilatata*
18. *Dryococelus australis*
19. *Eurycantha calcarata*
20. *Trychopeplus laciniatus*
21. *Extatosoma popa*
22. *Spinohirasea bengalensis*
23. *Cnipsus rachis*
24. *Taraxippus perezgelaberti*
25. *Parectatosoma mocquerysi*
26. *Aretaon asperrimus*
27. *Dimorphodes mancus*
28. *Orthomeria kangii*
29. *Oreophoetes peruana*
30. *Anisomorpha buprestoides* (white morph)
31. *Tropidoderus childrenii*
32. *Kalocorinnis wegneri*
33. *Anarchodes annulipes*
34. *Pseudophasma rufipes*
35. *Diesbachia tamyris*
36. *Spathomorpha adefa*
37. *Ctenomorpha marginipennis*
38. *Phobaeticus serratipes*
39. *Clonaria conformans*
40. *Clitarchus hookeri*
41. *Hyrtaeus procerus*
42. *Phalces tuberculatus*
43. *Megacrania batesii*
44. *Graeffea leverii*
45. *Apterograeffea reunionensis*
46. *Phaenopharos struthioneus*
47. *Pharnacia ponderosa*
48. *Achrioptera spinosissima*
49. *Phasma gigas*
50. *Pseudodiacantha macklottii*
51. *Neopromachus wallacei*
52. *Pterinoxylus crassus*

- 53. *Argosarchus horridus*
- 54. *Monoignosis spinosa*
- 55. *Sceptrophasma hispidulum*
- 56. *Clonopsis gallica*
- 57. *Carausius morosus*
- 58. *Hyrtacus tuberculatus*
- 59. *Diapheromera femorata*

Table S4: Description of the main distinctive morphological features of each morphotype. The number of estimated independent origins on the phylogeny (see Figure S6) and the name of the genera included are indicated. The main morphospace axes best describing the morphological “niche” of each morphotype were identified using a random forest model predicting a given morphotype versus all the others, given the first ten morphological axes (see Supplementary Materials and Methods). The importance of a given dimension was assessed by looking at the decrease in prediction accuracy after dropping each corresponding axes (as implemented in the function *randomForest*, R package “randomForest”). Only the main predictors are reported here.

Morphotype	# of Origins	Genera included	Main predicting morphospace axes
Wide leaf mimics	1	<i>Phyllium, Pulchriphyllium, Nanophyllium, Trolicaphyllium, Cryptophyllium, Chitoniscus</i>	PC2(43%): very flat body PC1(17%): very wide body
Small crawlers	2	<i>Timema, Microcanachus</i>	PC3(26%): wingless PC1(20%): relatively wide and stocky (also remarkably small)
Winged bark huggers	3	<i>Prisopus, Neoclides, Epicharmus</i>	PC6(22%): relatively small PC3(20%): fully winged PC1(14%): relatively wide
Diminutive spiny morphs	3	<i>Orestes, Epidares, Dares, Hoploclonia, Pterobrimus, Labidiophasma, Ommatopseudes</i>	PC1 (23%): relatively wide and stocky. PC2(17%): cylindrical body PC4(14%): spiny and textured
Chinchemolles	1	<i>Agathemera</i>	PC6(27%): relatively massive and stocky insects PC8(19%): dark PC2(17%): cylindrical body
Apterous bark huggers	7	<i>Damasippoides, Pseudoleosthenes, Abrosoma, Dinophasma, Dajaca, Paraprisopus, Anisomorpha, Malacomorpha, Peruphasma, Pseudophasma, Thaumatolectron,</i>	PC1(16%): relatively wide and short legged PC2(14%): cylindrical bodies PC7(11%): dark
Lime green morphs	6	<i>Diapherodes, Venupherodes, Paramenexenus, Macrophasma, Monandroptera, Rhapsiderus, Cranidium.</i>	PC2(14%): relatively flattened PC6(13%): relatively large PC9(13%): long subgenital plate
Tree lobsters	4	<i>Eurycantha, Dryococelus, Canachus, Haaniella, Heteropteryx, Mearnsiana</i>	PC4 (25%): large and spiny PC1 (22%): relatively wide
Spiny leaf mimics	2	<i>Extatosoma, Trychopeplus</i>	PC9 (26%): irregular body edges PC4 (21%): spiny and textured
Spiny robust morphs	11	<i>Spinohirasea, Neohirasea, Oxyartes, Erinaceophasma, Neopromachus, Hyposcyrthus, Rhynchacris, Taraxippus, Lamponius, Dimorphodes, Micrarchus, Cnipsus, Creoxylus, Lobolibethra, Sungaya, Aretaon, Trachyaretaon, Brasidas, Parectatosoma</i>	PC1(27%): relatively wide PC4(19%): spiny and/or textured

Aposematic morphs	7	<i>Oreophoetes, Tithonophasma, Anisomorpha, Pseudophasma, Orthomeria, Megacrania, Ophicrania</i>	PC8(42%): contrasting and conspicuous color pattern
Elongated leaf mimics	3	<i>Tropidoderus, Malandania, Parapodacanthus, Podacanthus</i>	PC3(45%) : very large wings PC5(12%) : relatively small heads
Winged sticks	2	<i>Anarchodes, Diesbachia, Sipyloidea, Kalocorinnis, Trachythorax, Metriophasma, Malacomorpha, Pseudophasma</i>	PC3(32%): large wings PC2(15%): cylindrical PC6(15%): relatively small
Elongated sticks	8	<i>Phobaeticus, Phryganistria, Cuniculina, Staelonchodes, Lonchodes, Leprocaulinus, Baculofractum, Ctenomorpha, Bacteria, Spathomorpha</i>	PC1(46%) : extremely long and elongated
Thin sticks	10	<i>Rhamphophasma, Clonaria, Ramulus, Medauroidea, Chondrostethus, Clitarchus, Hyrtacus, Orxines, Rhamphosipyloidea, Lopaphus, Leiophasma, Phalces, Phanocloidea</i>	PC1(31%) : Very elongated PC4(14%): smooth PC3(12%): wingless
Large-headed sticks	2	<i>Megacrania, Ophicrania, Graeffea, Apterograeffea</i>	PC5(22%) : relatively large head PC10(20%) : relatively elongated head
Large sticks	5	<i>Pharnacia, Tirachioidea, Phasmotaenia, Alienobostra, Cladomorphus, Phaenopharos</i>	PC1(16%): elongated PC2(13%): tubular PC6(11%): relatively large
Giant winged sticks	4	<i>Anchiale, Paronchestus, Acrophylla, Eurycnema, Phasma, Cigarophasma, Lopaphus, Bactrododema, Achrioptera</i>	PC3(30%): brachypterous PC4(16%): spiny thorax PC1(12%): elongated
Prickly sticks	8	<i>Argosarchus, Acanthoxyla, Asprenas, Mauritiophasma, Onchestus, Parapachymorpha, Agamemnon, Pterinoxylus, Neopromachus, Manduria, Pseudodiacantha, Centrophasma, Anisacantha, Antongilia</i>	PC4(18%) : spiny and/or textured PC8(17%): often brown PC6(16%): often relatively small PC1(12%): relatively elongated
Small sticks	8	<i>Ocnophiloidea, Diapheromera, Pseudosermyle, Caribbiopheromera, Bacillus, Clonopsis, Leptynia, Pijnackeria, Xylica, Zehntneria, Carausius, Hyrtacus, Lonchodes, Neopromachus, Parapachymorpha, Medaura, Sceptrophasma, Monoignosis, Carlius.</i>	PC1(18%) : relatively elongated PC5(14%): relatively small head PC4(13%): smooth PC2(11%): tubular

Table S5: List of taxa included in photographs in Figures S3-5.

Picture #	Species	Origin	Source and photo credits
1	<i>Spathomorpha adefa</i>	Moramanga, Madagascar (wild)	CC-BY-NC-ND Davorka Kitonić & Josip Skejo
2	<i>Ctenomorpha marginipennis</i>	New South Wales, Australia (wild)	CC-BY-NC Paul Whittington
3	<i>Bacteria ploiaria</i>	Barro Colorado, Panama (wild)	CC-BY-NC John G. Phillips
4	<i>Phobaeticus serratipes</i>	Penang, Malaysia (wild)	CC-BY-NC Albert Kang
5	<i>Diapherodes martinicensis</i>	Martinique (captivity)	© Bruno Kneubühler, used by permission
6	<i>Paramenexenus laetus</i>	Tay Yen Tu NP, Vietnam (captivity)	© Bruno Kneubühler, used by permission
7	<i>Monandroptera acanthomera</i>	Réunion (wild)	© Nicolas Cliquennois, used by permission
8	<i>Cranidium gibbosum</i>	French Guiana (captivity)	CC-BY 3.0 Daniel Dittmar
9	<i>Orthomeria kangii</i>	Luzon, Philippines (wild)	CC-BY-NC Albert Kang
10	<i>Anisomorpha buprestoides</i>	Florida, USA (wild)	CC-BY-SA 3.0 Bugenstein
11	<i>Oreophoetes peruana</i>	Tarapoto, San Martín, Peru (wild)	© Romain Boisseau (author)
12	<i>Apterograeffea reunionensis</i>	Réunion (wild)	© Nicolas Cliquennois, used by permission
13	<i>Megacrania batesii</i>	Cape Tribulation, Queensland, Australia (wild)	© Romain Boisseau (author)
14	<i>Trachythorax maculliculis</i>	Thailand (captivity)	CC-BY-SA 3.0 Dragüs
15	<i>Pseudophasma phthisicum</i>	French Guiana (wild)	CC-BY-NC Manuel Ruedi
16	<i>Trychopeplus laciniatus</i>	Monteverde, Costa Rica (captivity)	© Bruno Kneubühler, used by permission
17	<i>Extatosoma tiaratum</i>	Queensland, Australia (captivity)	© Bruno Kneubühler, used by permission
18	<i>Diapheromera femorata</i>	Virginia, USA (wild)	CC-BY-2.0 Judy Gallagher
19	<i>Clonopsis gallica</i>	Jaén, Spain (wild)	CC BY 2.0 Ramón Portellano
20	<i>Dryococelus australis</i>	Ball's pyramid, Australia (captivity)	© Paul D. Brock, used by permission
21	<i>Haaniella dehaanii</i>	Sarawak, Borneo, Malaysia (captivity)	CC-BY-SA Drägüs
22	<i>Canachus alligator</i>	Mt Koghi, New Caledonia (captivity)	© Bruno Kneubühler, used by permission
23	<i>Eurycantha calcarata</i>	Kimbe, WNB, Papua New Guinea (wild)	© Romain Boisseau (author)
24	<i>Spinohirasea bengalensis</i>	Bach Ma, Vietnam (captivity)	© Bruno Kneubühler, used by permission
25	<i>Cnipsus rachis</i>	Mt Koghi, New Caledonia (wild)	CC-BY-NC Damien Brouste
26	<i>Parectatosoma sp.</i>	Moramanga, Madagascar (captivity)	© Bruno Kneubühler, used by permission
27	<i>Lamponius bocki</i>	Punta Cana, Dominican Republic (captivity)	© Bruno Kneubühler, used by permission
28	<i>Aretaon asperrimus</i>	Sabah, Borneo, Malaysia (wild)	CC-BY-NC Albert Kang
29	<i>Agathemera crassa</i>	Santiago, Chile (wild)	CC-BY-NC-SA Ariel Cabrera Foix
30	<i>Ommatopseudes harmani</i>	Tanah Rata, Peninsular Malaysia (captivity)	© Bruno Kneubühler, used by permission
31	<i>Epidares nolimetangere</i>	Sarawak, Borneo, Malaysia (captivity)	© Bruno Kneubühler, used by permission
32	<i>Labidiophasma rouxi</i>	Mt Humboldt, New Caledonia (wild)	CC-BY-NC Damien Brouste
33	<i>Timema boharti</i>	California, USA (wild)	CC-BY-NC John Christensen
34	<i>Microcanachus matileorum</i>	Yaté, New Caledonia (wild)	CC-BY-NC Damien Brouste
35	<i>Paraprisopus antillarum</i>	Guadeloupe (captivity)	© Bruno Kneubühler, used by permission

36	<i>Dinophasma saginatum</i>	Mulu NP, Sarawak, Borneo, Malaysia (captivity)	© Bruno Kneubühler, used by permission
37	<i>Pseudoleosthenes irregularis</i>	Ranomafana NP, Madagascar (wild)	© Paul Bertner, used by permission
38	<i>Anisomorpha buprestoides</i>	Florida, USA (wild)	CC-BY-NC Scott Ward
39	<i>Centrophasma hadrillum</i>	Bako NP, Sarawak, Borneo, Malaysia (captivity)	© Bruno Kneubühler, used by permission
40	<i>Argosarchus horridus</i>	Christchurch, New Zealand (wild)	CC-BY-NC Chris Morse
41	<i>Neopromachus sp.</i>	Popondetta, Oro Province, Papua New Guinea (wild)	© Romain Boisseau (author)
42	<i>Pulchriphyllium pulchrifolium</i>	Java, Indonesia (captivity)	© Romain Boisseau (author)
43	<i>Prisopus berosus</i>	Ocosingo, Chis, Mexico (wild)	CC-BY-NC Silvano LG
44	<i>Epicharmus marchali</i>	Mauritius (wild)	© Sylvain Hugel & Nicolas Cliquennois, used by permission
45	<i>Neoclides buescheri</i>	Bako NP, Sarawak, Borneo, Malaysia (captivity)	© Bruno Kneubühler, used by permission
46	<i>Tropidoderus childrenii</i>	New South Wales, Australia (wild)	CC-BY-NC-ND 2.0 David Midgley
47	<i>Malandania pulchra</i>	Queensland, Australia (wild)	CC-BY-NC Linda Rogan EntSocVic
48	<i>Phasmotaenia lanyuhensis</i>	Lanyuh Island (captivity)	© Bruno Kneubühler, used by permission
49	<i>Cladomorphus phyllinus</i>	Sao Paulo, Brasil (captivity)	© Bruno Kneubühler, used by permission
50	<i>Pharnacia ponderosa</i>	Mt Capotoan, Samar, Philippines (captivity)	© Bruno Kneubühler, used by permission
51	<i>Lopaphus sphalerus</i>	Qinnan, Qinzhou, Guangxi, China (captivity)	© Bruno Kneubühler, used by permission
52	<i>Achrioptera punctipes</i>	Madagascar (captivity)	© Bruno Kneubühler, used by permission
53	<i>Bactrododema hecticum</i>	Windhoek, Namibia (captivity)	© Bruno Kneubühler, used by permission
54	<i>Acrophylla wuelfingi</i>	Cairns, Queensland, Australia (wild)	CC-BY-NC Felix Fleck

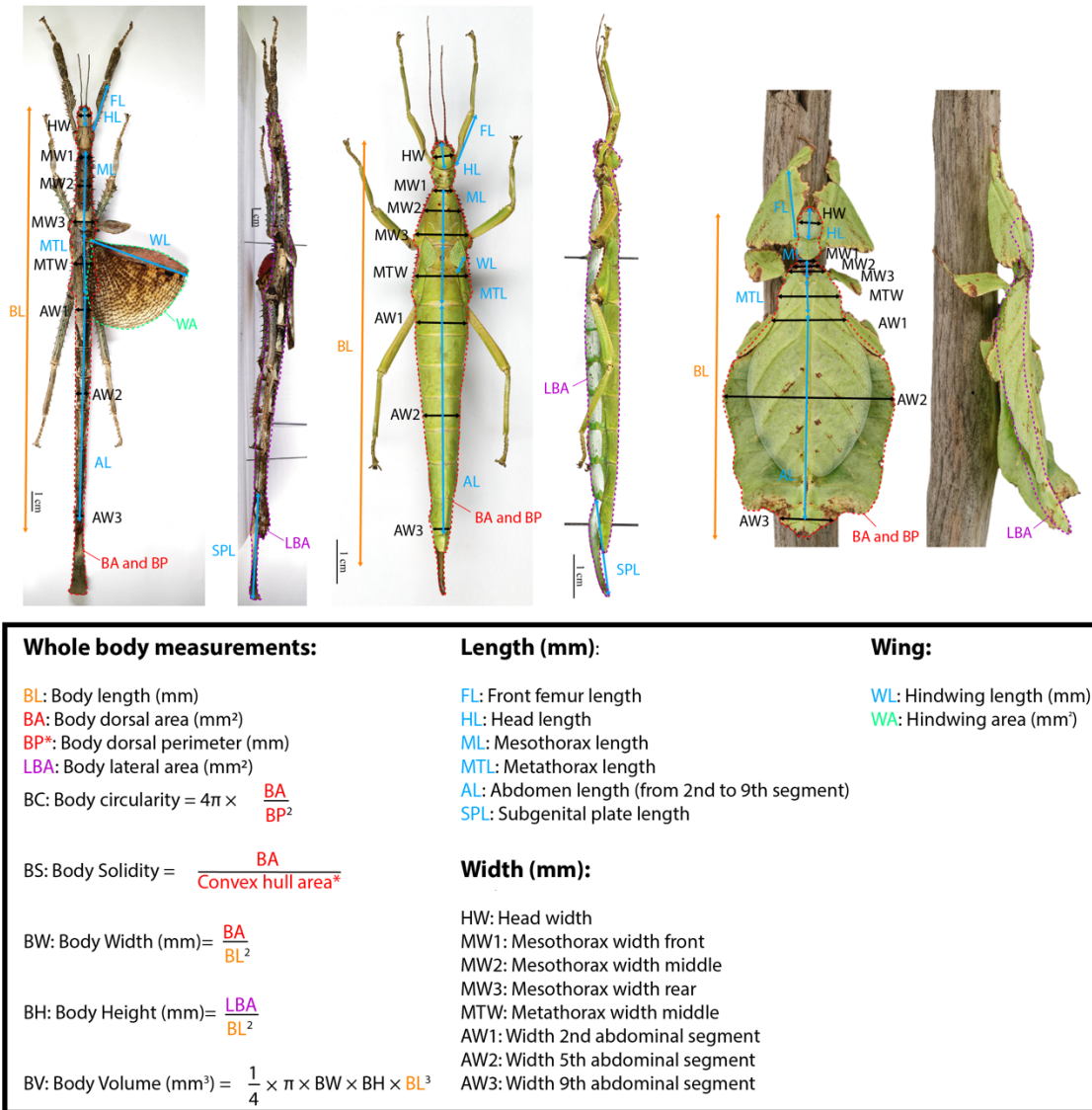


Figure S1: Quantitative measurements of phasmid specimens. **Left:** Adult female *Achrioptera punctipes cliquennoisi* Hennemann & Conle 2004 (Specimen MNHN-EO-PHAS127, project RECOLNAT (ANR-11-INBS-0004), photographs by Marion Depraetere, 2015, CC-BY-NC-ND). **Middle:** Adult female *Diapherodes martinicensis* Lelong & Langlois, 2005 (Specimen MNHN-EO-PHAS542, project RECOLNAT (ANR-11-INBS-0004), photographs by Marion Depraetere, 2015, CC-BY-NC-ND). **Right:** Adult female *Pulchriphyllium giganteum* Hausleithner, 1984 (Culture “Tapah Hills”, Bruno Kneubuehler, Switzerland, photographs by Bruno Kneubuehler, used with permission).

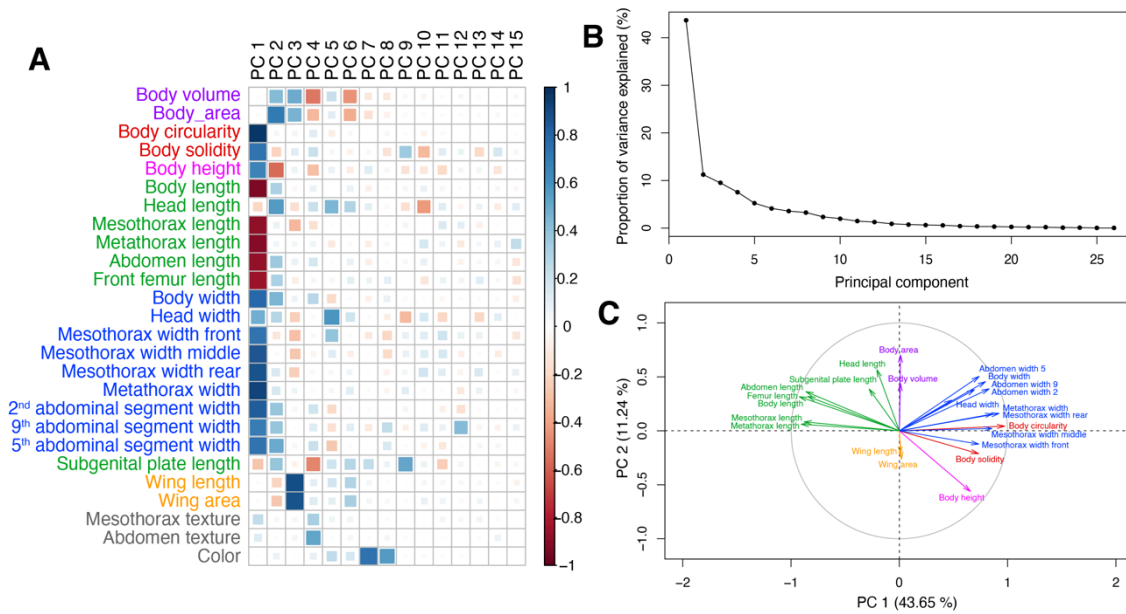


Figure S2 : Results of the PCAmix analysis of 24 quantitative measurements and three qualitative scores of body mesothorax and abdomen texture, and overall body coloration. **A**: correlation plot between the 15 first principal components and all the original dataset features. **B**: Proportion of total variance explained by each principal component. **C**: Loading plot for the first principal components and the quantitative dataset features.

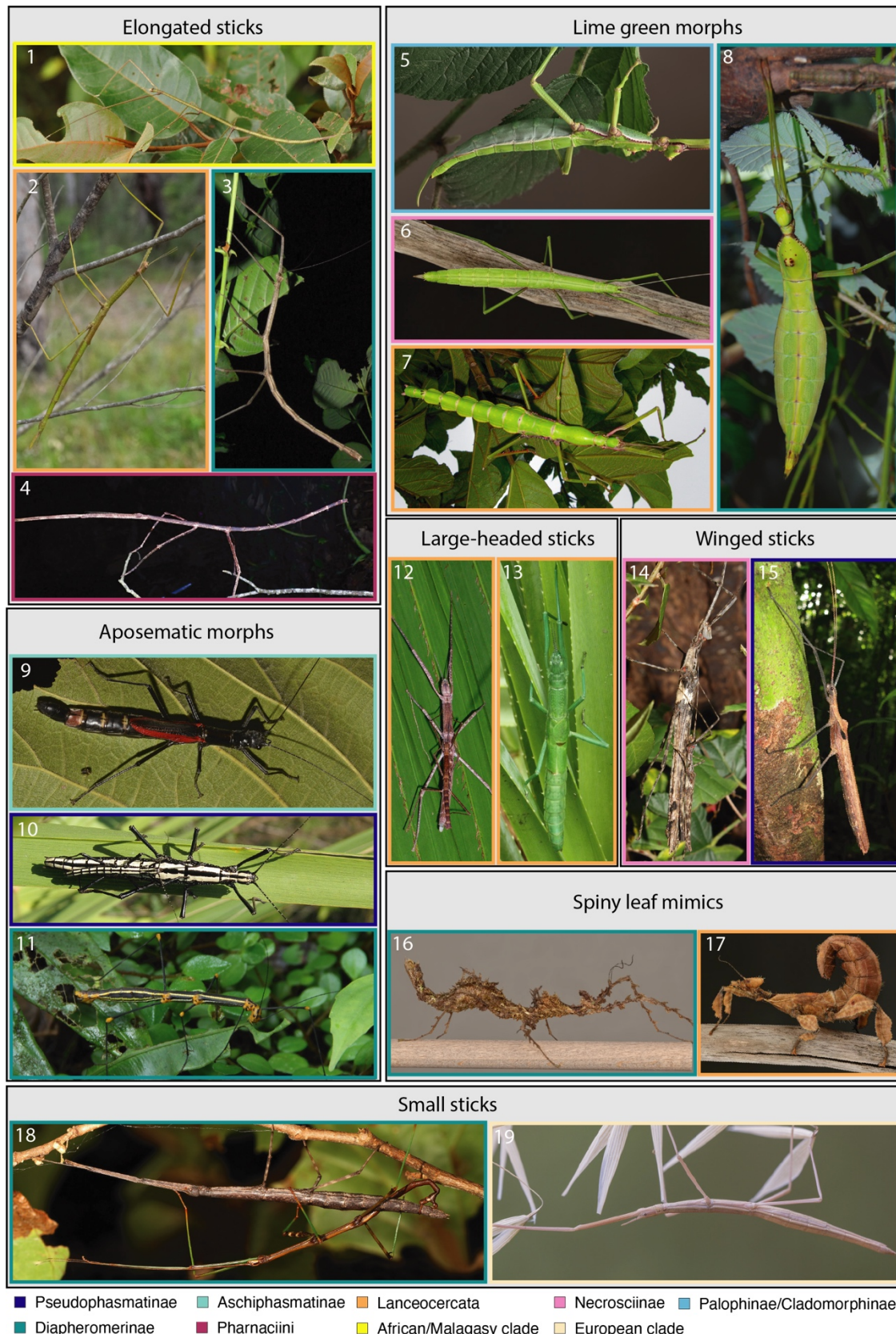


Figure S3: Photographs of adult females of each morphotype (part 1/3). The color surrounding each picture corresponds to a major phylogenetic clade. The list of species and photograph information can be found in Table S5.

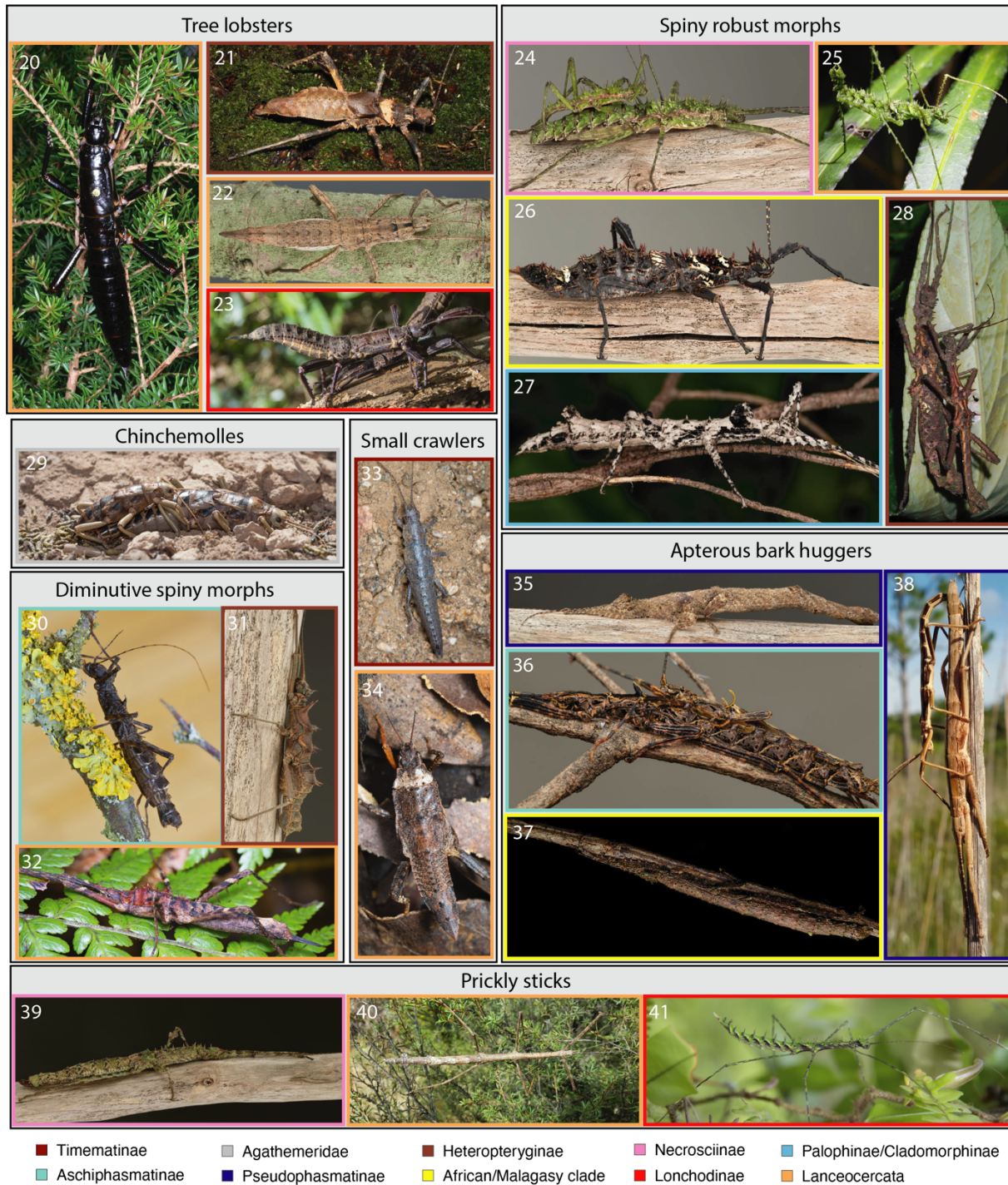


Figure S4: Photographs of adult females of each morphotype (part 2/3). The color surrounding each picture correspond to a major phylogenetic clade. The list of species and photograph information can be found in Table S5.

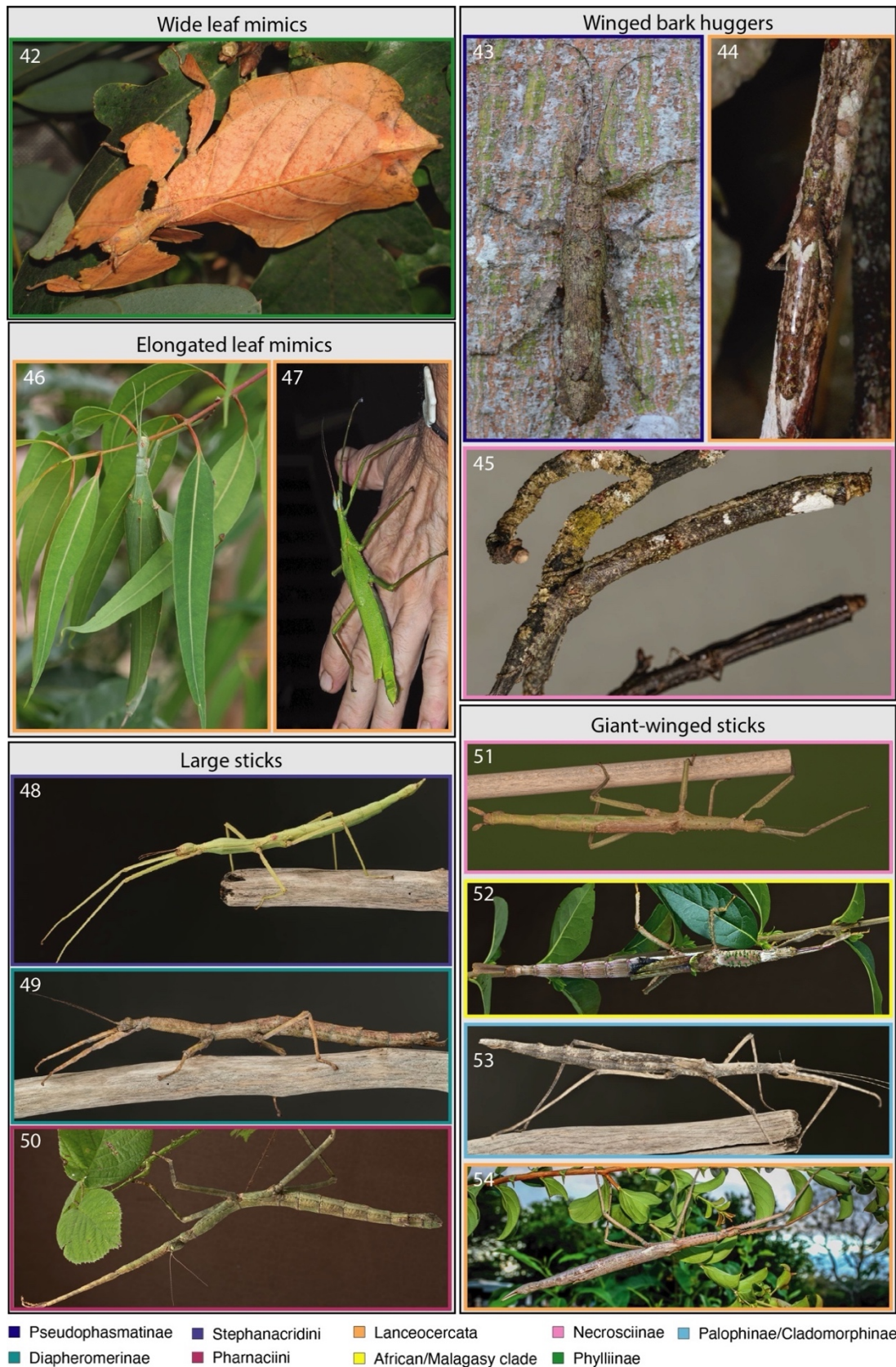


Figure S5: Photographs of adult females of each morphotype (part 3/3). The color surrounding each picture correspond to a major phylogenetic clade. The list of species and photograph information can be found in Table S5.

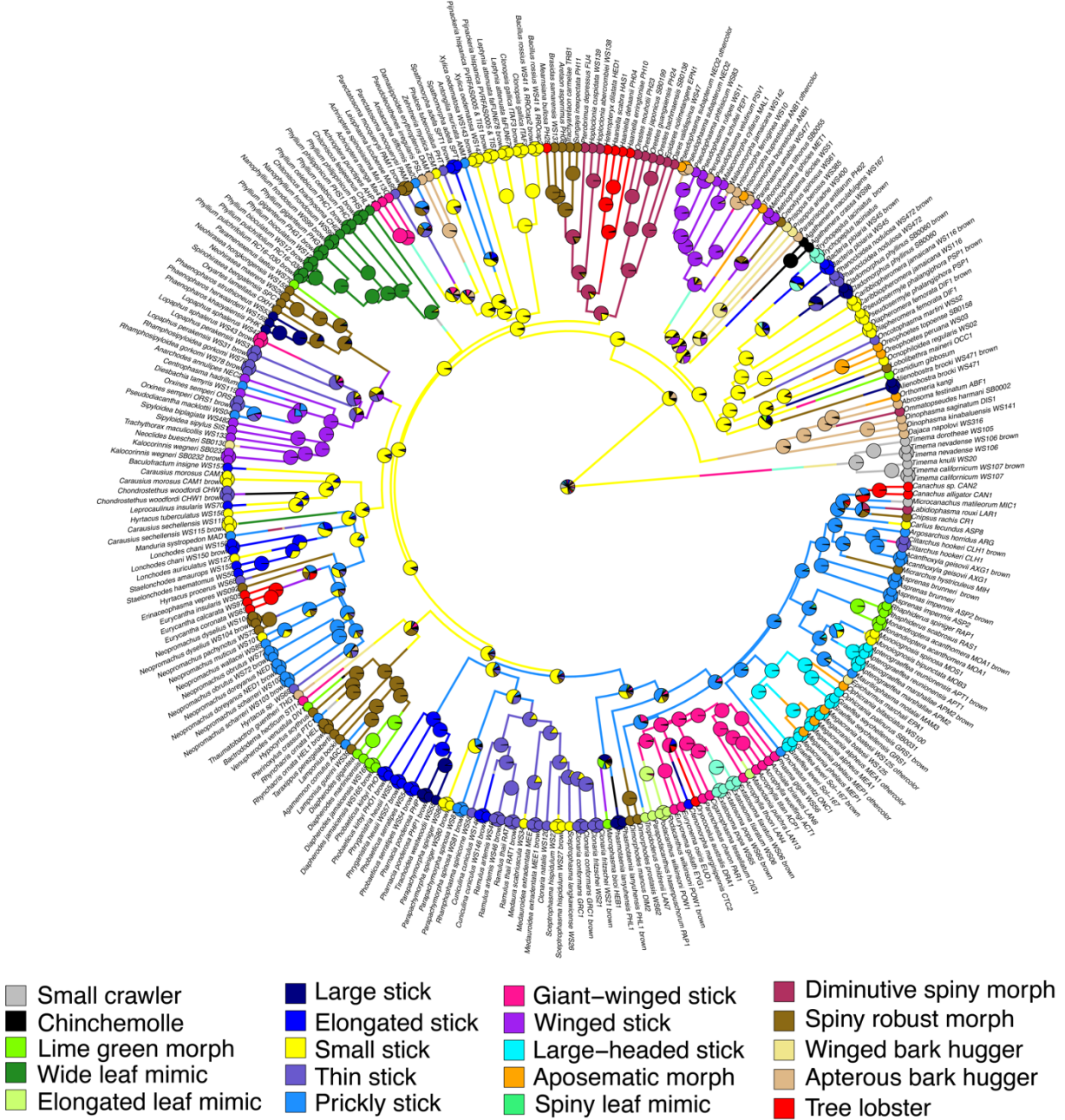


Figure S6: Ancestral state reconstruction of morphotypes using stochastic character mapping. The underlying transition matrix assumes a single rate of transition between morphotypes. The pie charts at each node represent the posterior probabilities that each internal node is in each state, and are overlaid over a random stochastic map.

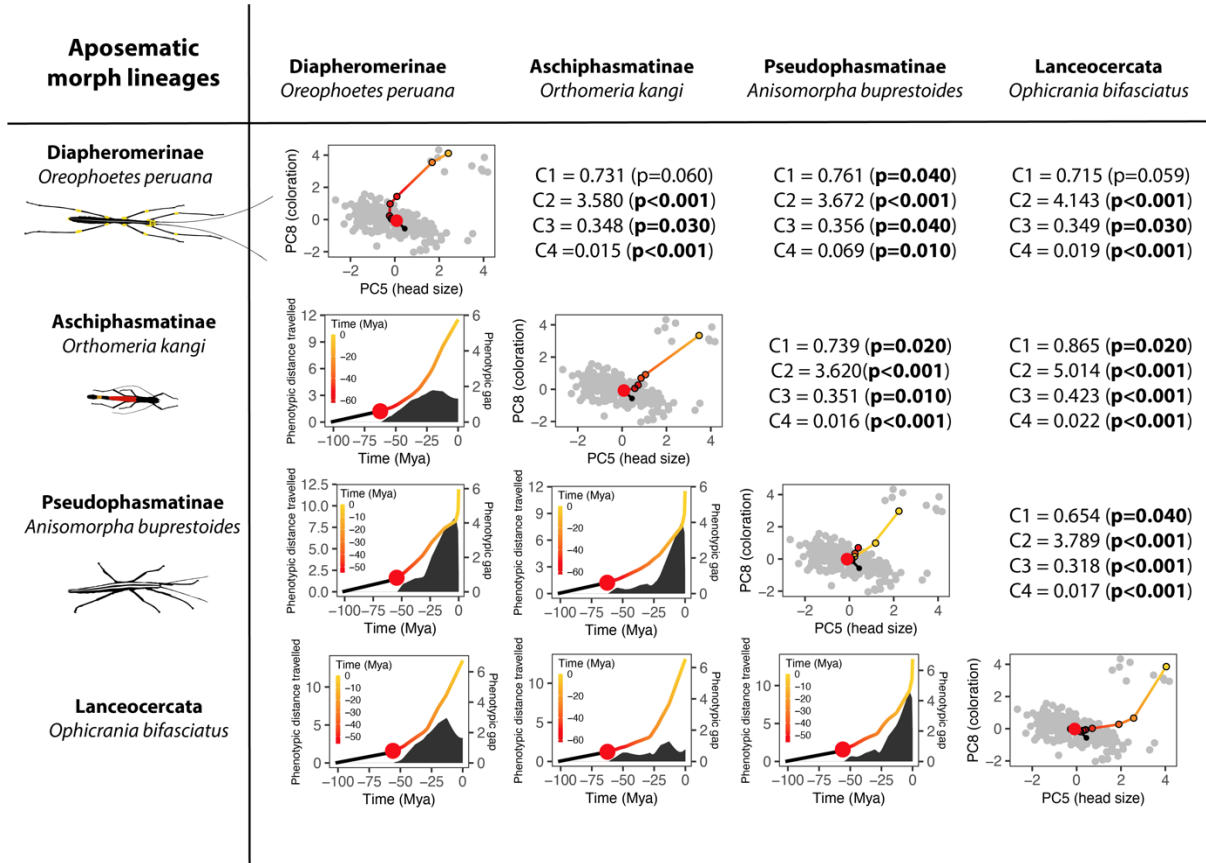


Figure S7: Convergent evolutionary trajectories of four lineages of aposematic morphs. The diagonal of the table show the position of ancestral nodes for each lineage on the morphospace along the two axes most distinguishing aposematic morphs. Colors transition from red to yellow as we go from the ancestral position of the most recent common ancestor of all four lineages (the big red dot) to present time. The lower table panels show the cumulative phenotypic distance travelled by each lineage pair through the 2-D morphospace and the corresponding trait gap over time. The numbers in the upper part of the table show the values and associated p-values of C1-C4 indexes calculated for each lineage pair only considering the two axes of the morphospace represented. Phasmid drawings by first author.

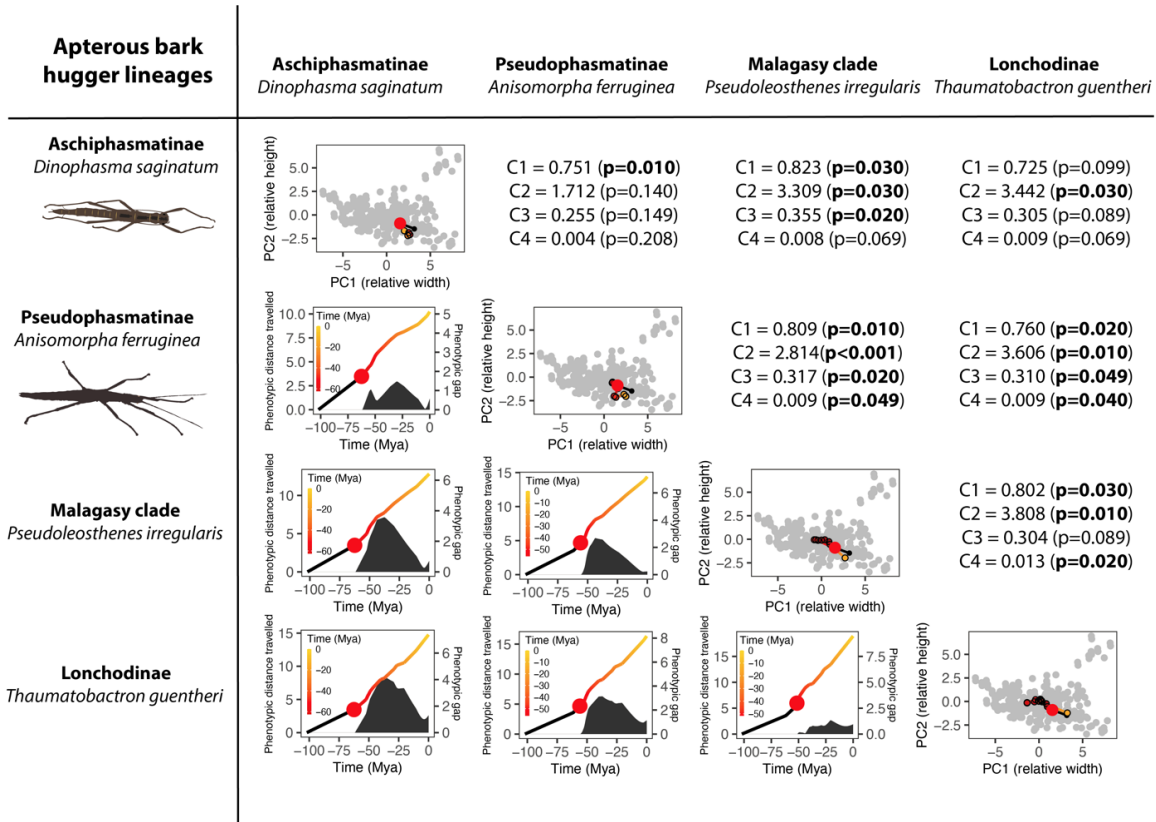


Figure S8: Convergent evolutionary trajectories of four lineages of apterous bark huggers. (see legend of figure S7)

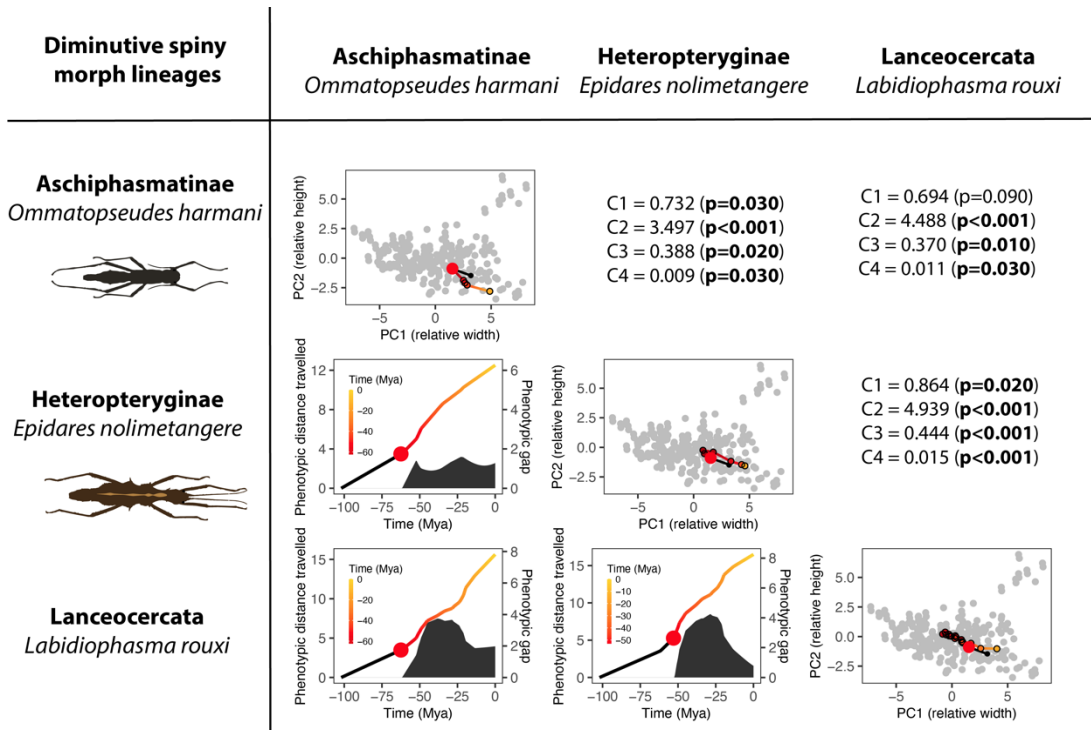


Figure S9: Convergent evolutionary trajectories of three lineages of diminutive spiny morphs. (see legend of figure S7)

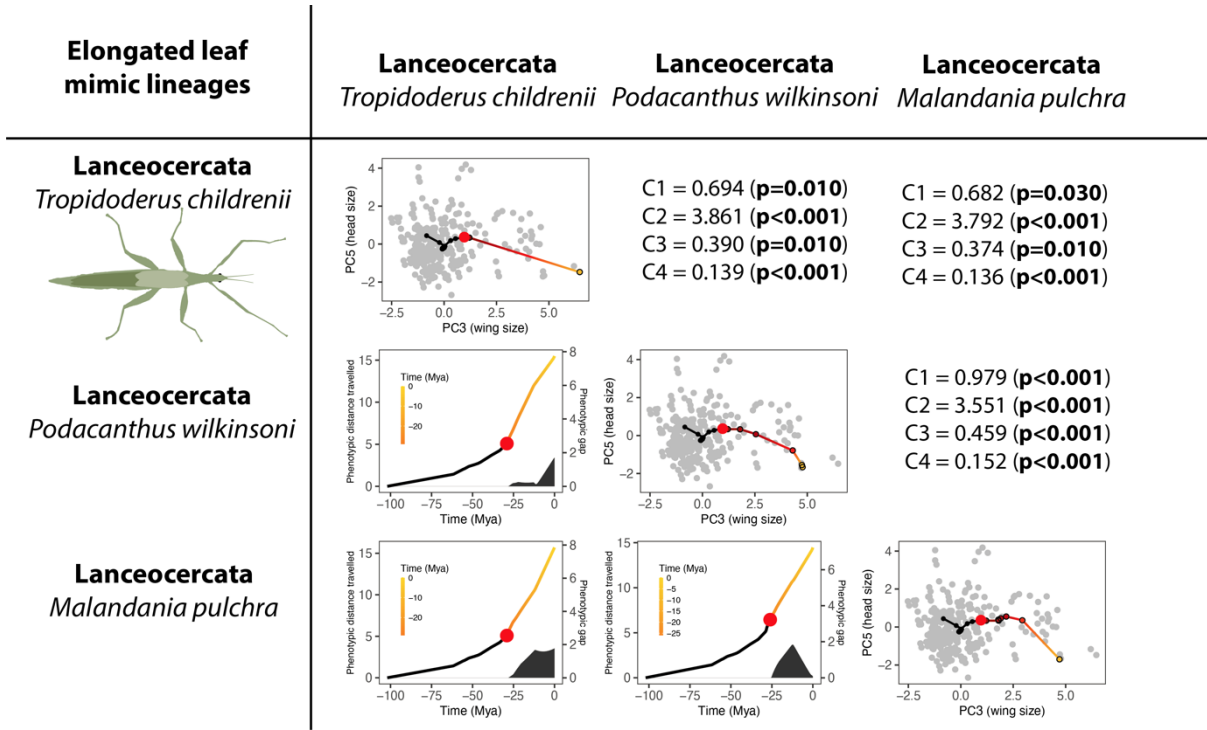


Figure S10: Convergent evolutionary trajectories of three lineages of diminutive spiny morphs. (see legend of figure S7)

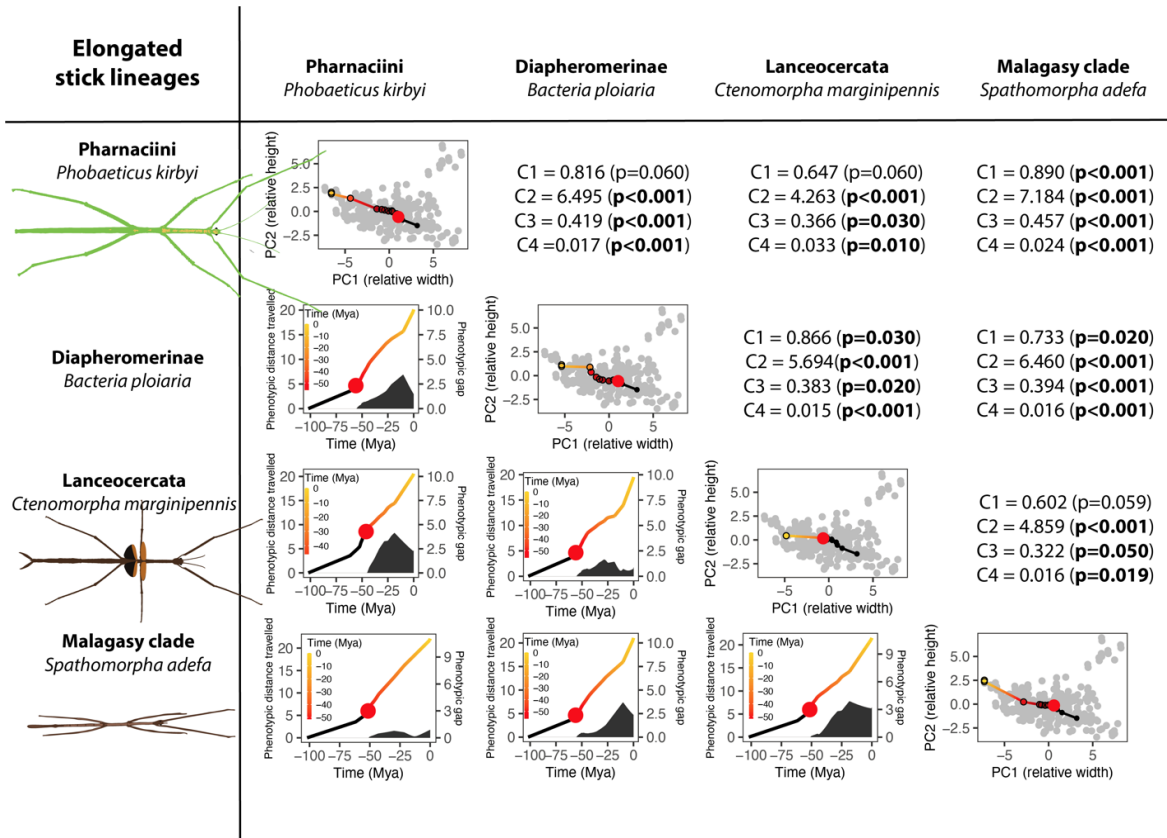


Figure S11: Convergent evolutionary trajectories of four lineages of elongated sticks. (see legend of figure S7)

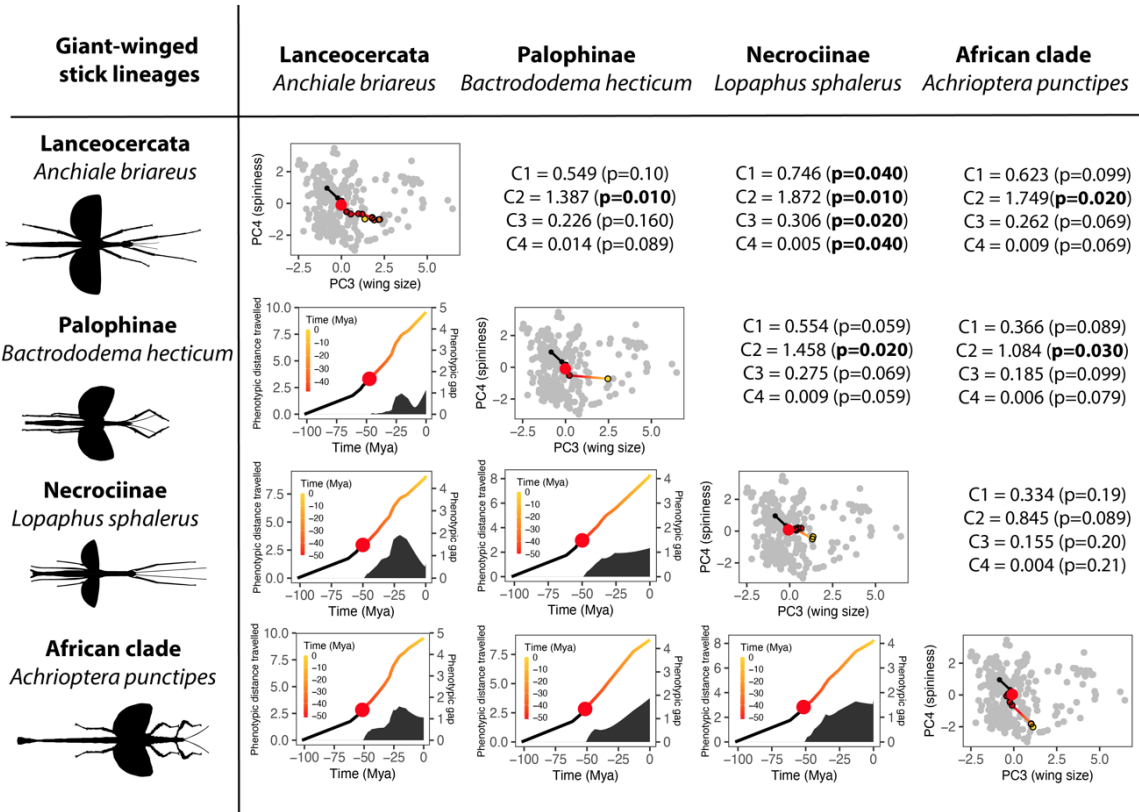


Figure S12: Convergent evolutionary trajectories of four lineages of giant-winged sticks. (see legend of figure S7)

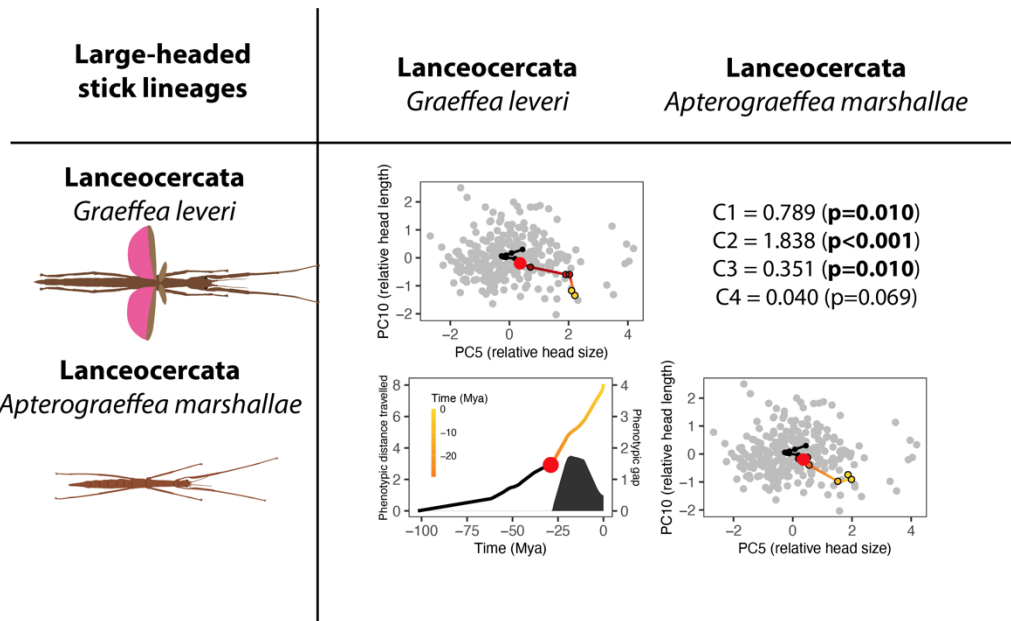


Figure S13: Convergent evolutionary trajectories of two lineages of large-headed sticks. (see legend of figure S7)

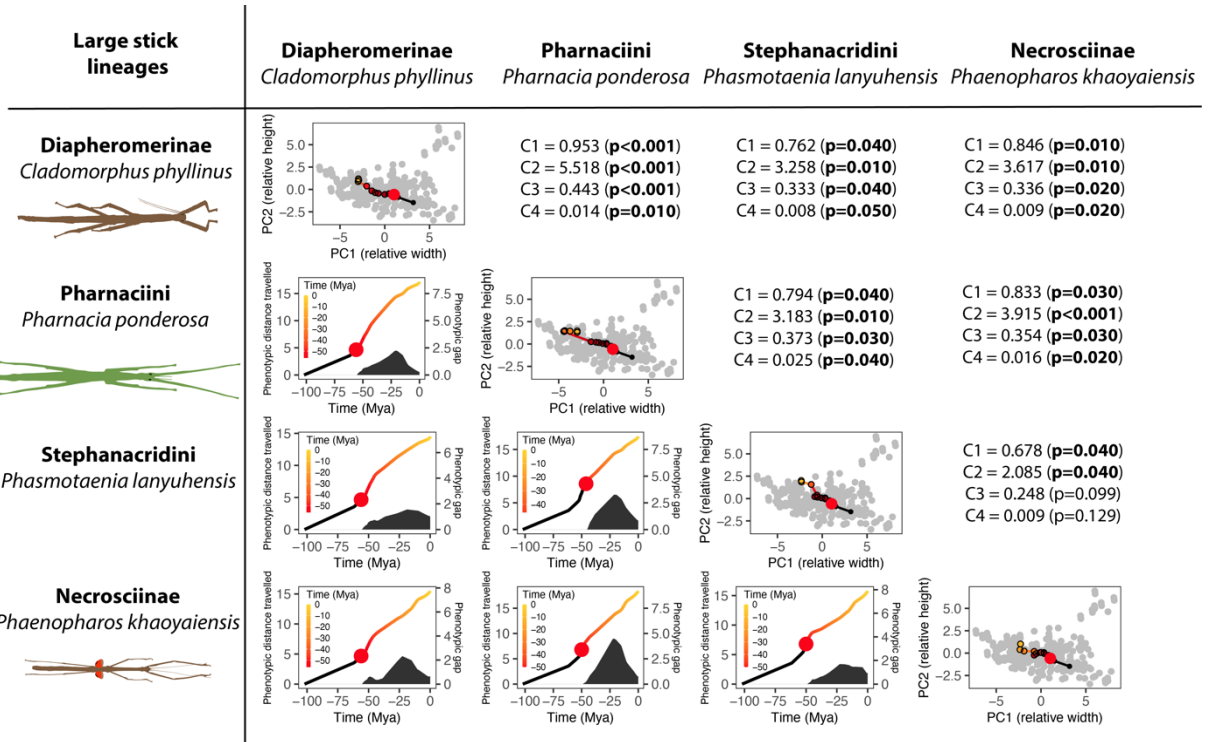


Figure S14: Convergent evolutionary trajectories of four lineages of large sticks. (see legend of figure S7)

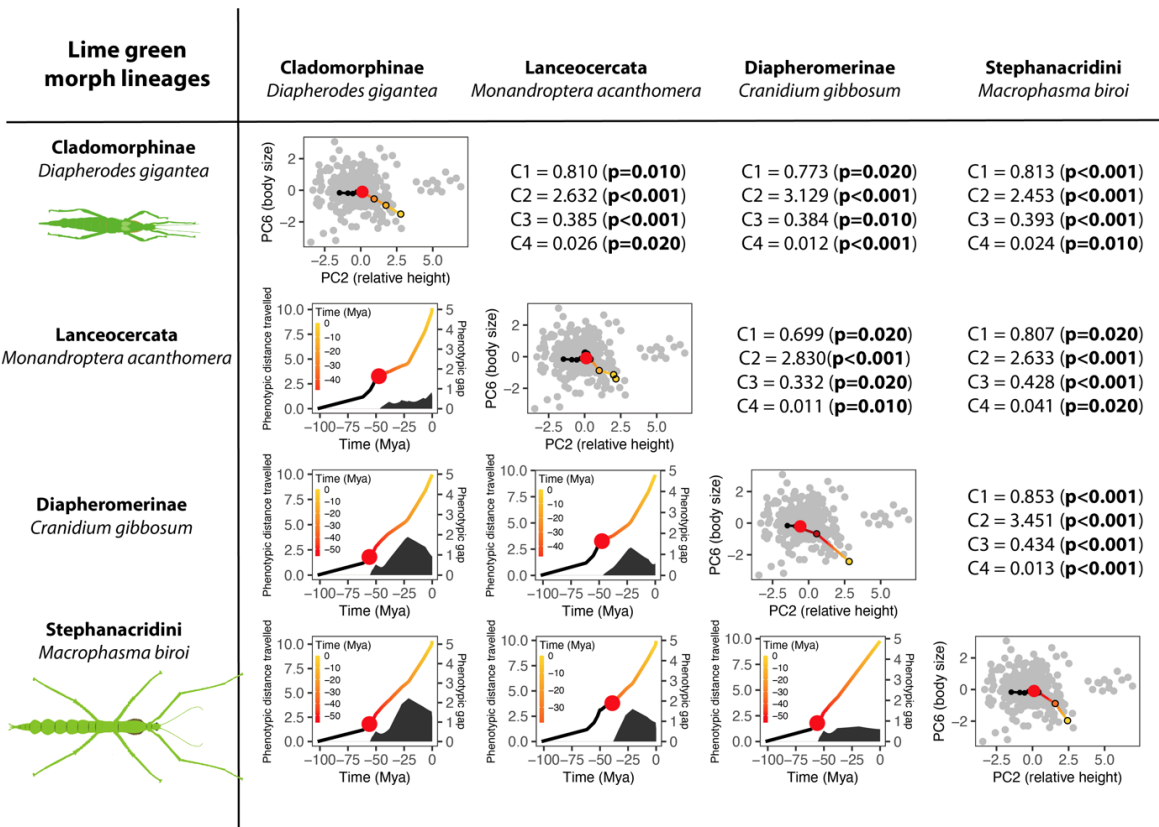


Figure S15: Convergent evolutionary trajectories of four lineages of lime green morphs. (see legend of figure S7)

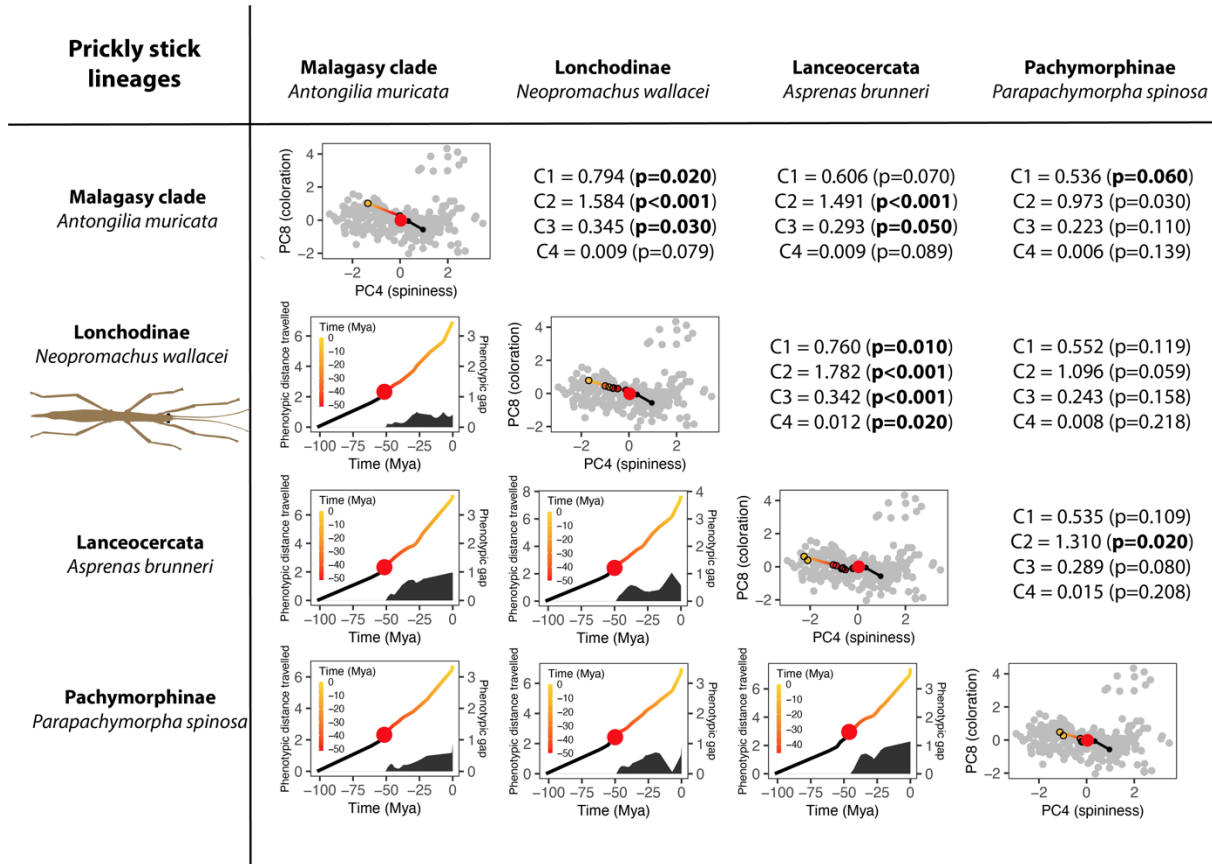


Figure S16: Convergent evolutionary trajectories of four lineages of prickly sticks. (see legend of figure S7)

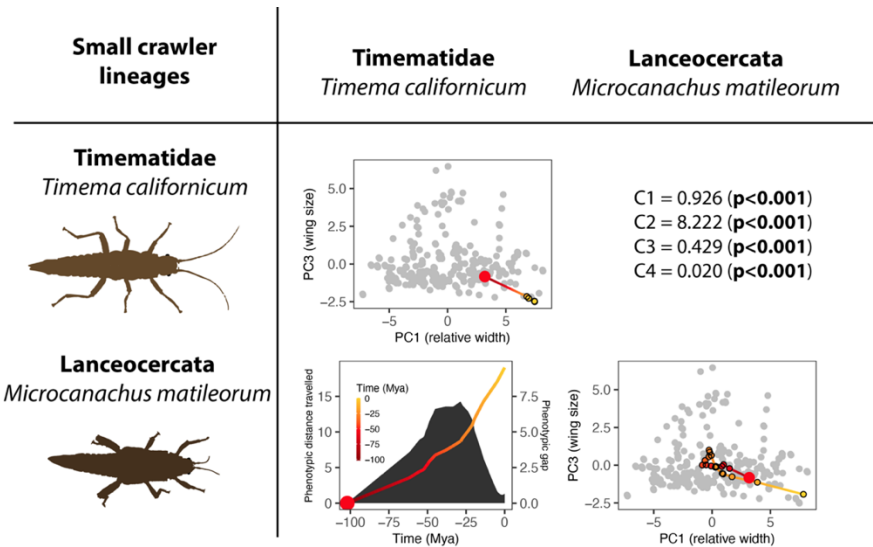


Figure S17: Convergent evolutionary trajectories of two lineages of small crawlers. (see legend of figure S7)

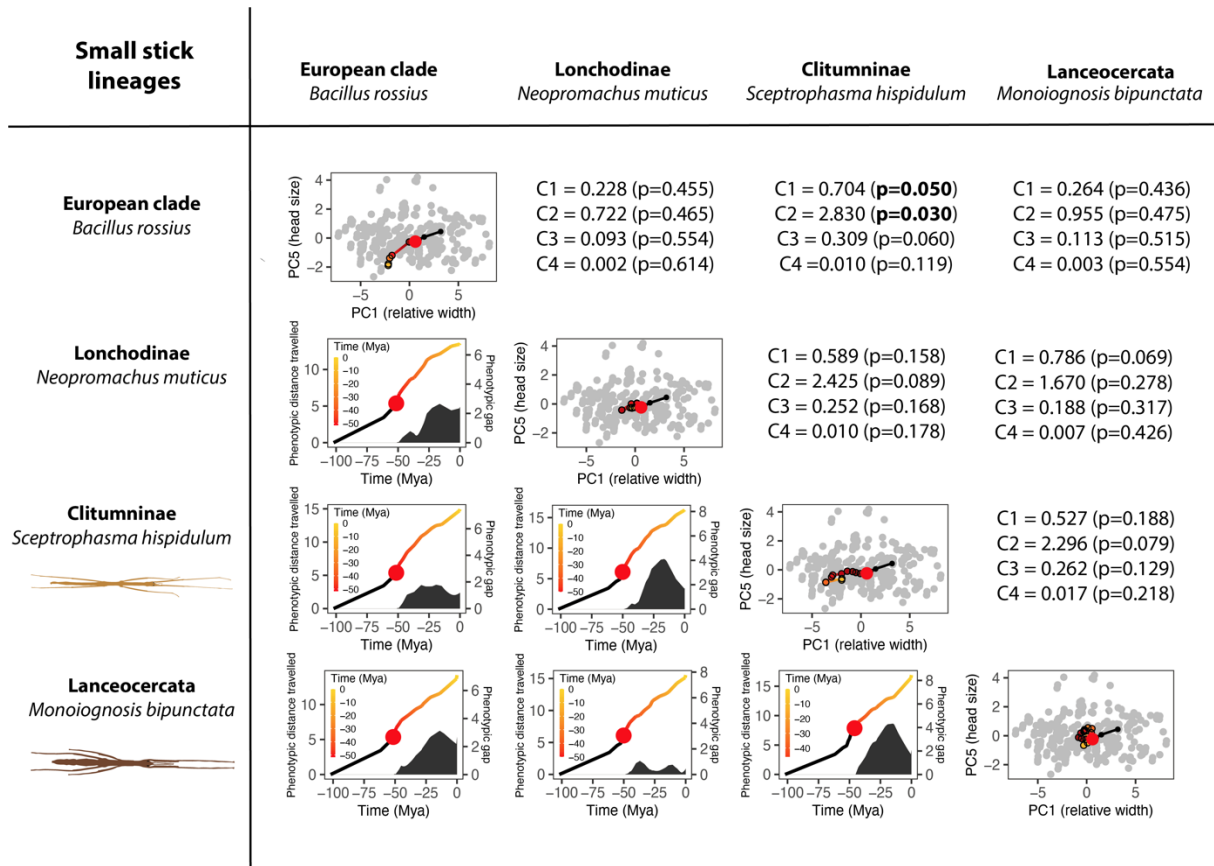


Figure S18: Convergent evolutionary trajectories of four lineages of small sticks. (see legend of figure S7)

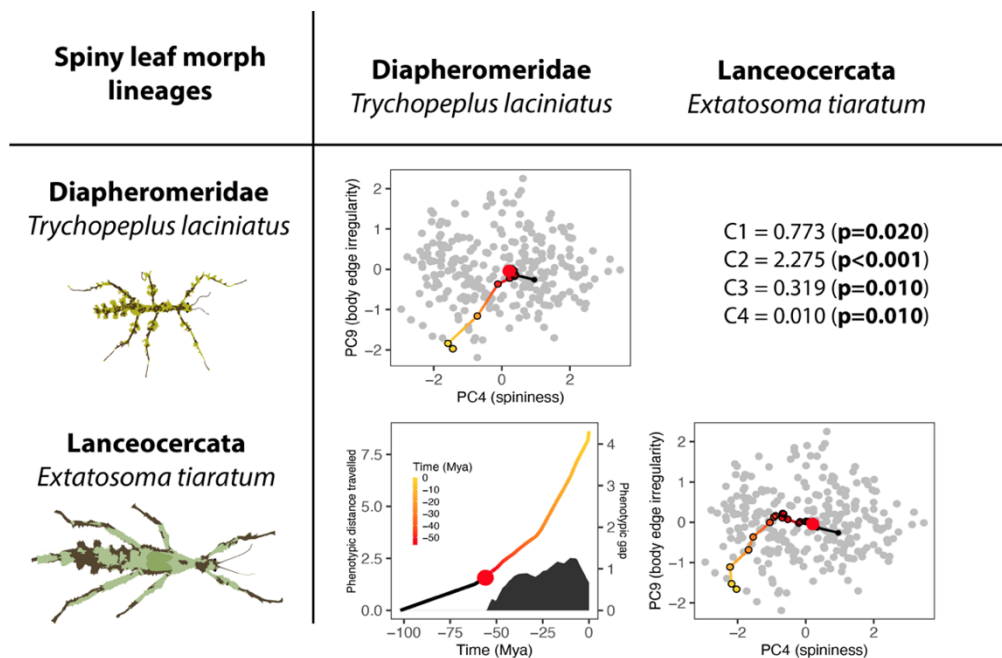


Figure S19: Convergent evolutionary trajectories of four lineages of spiny leaf mimics. (see legend of figure S7)

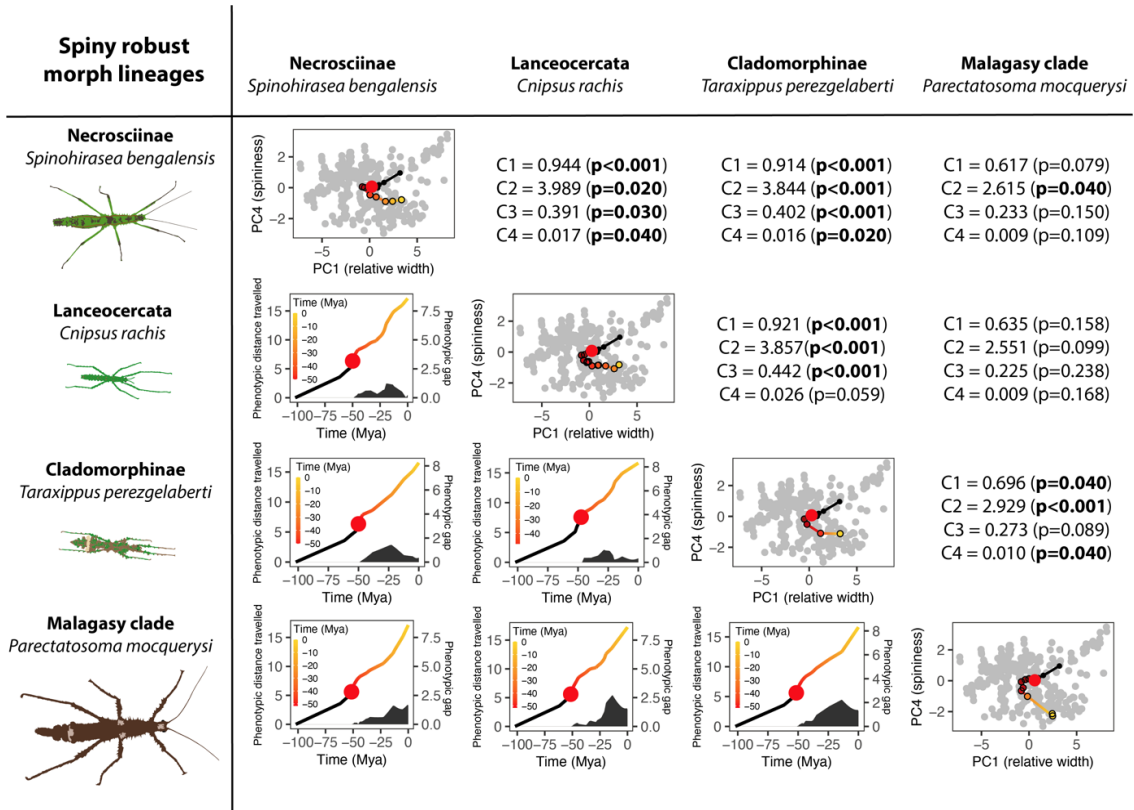


Figure S20: Convergent evolutionary trajectories of four lineages of spiny robust morphs. (see legend of figure S7)

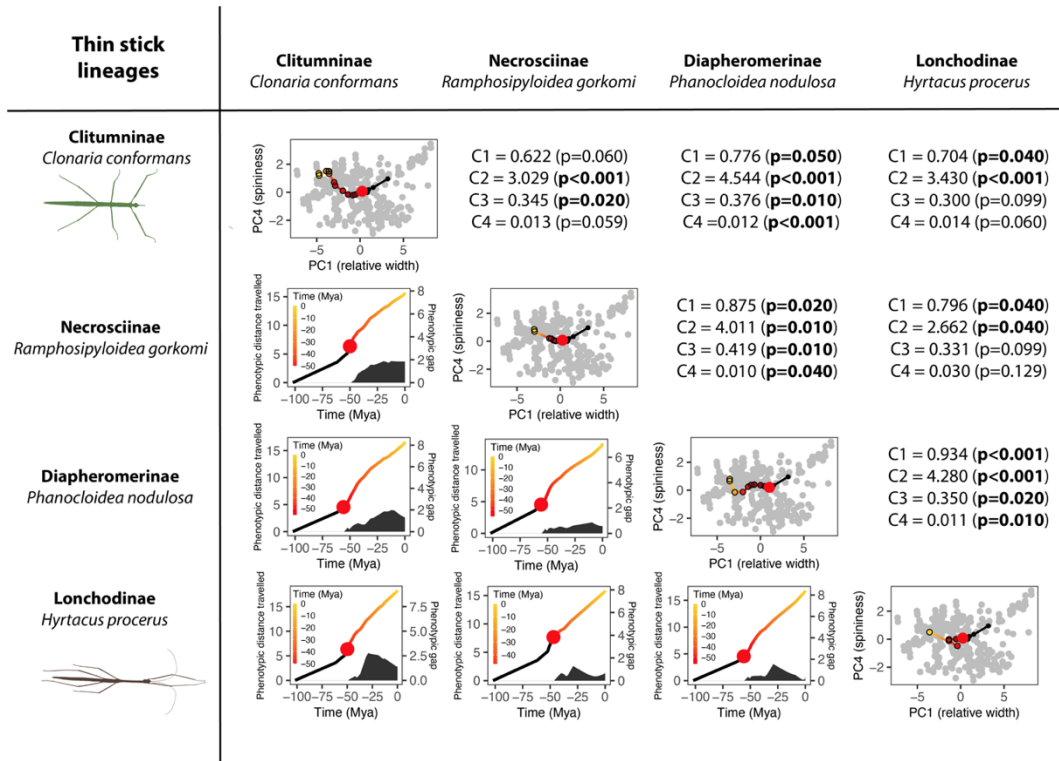


Figure S21: Convergent evolutionary trajectories of four lineages of thin sticks. (see legend of figure S7)

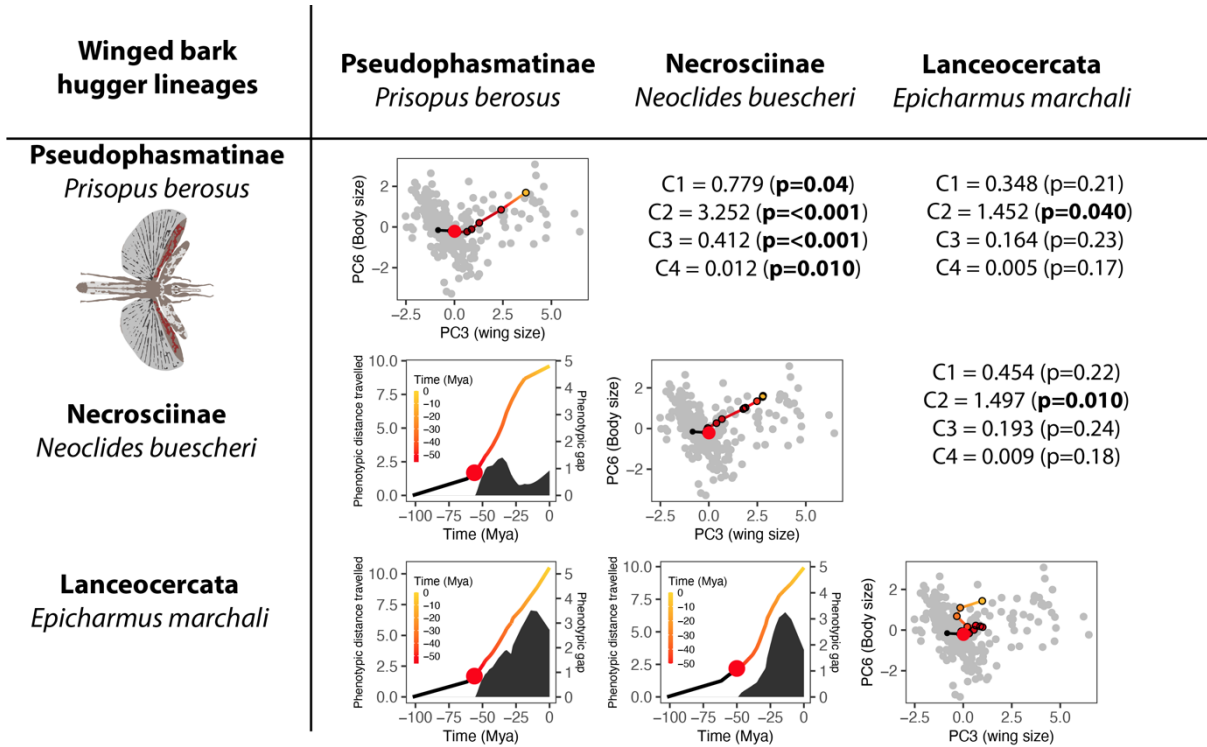


Figure S22: Convergent evolutionary trajectories of three lineages of winged bark huggers. (see legend of figure S7)

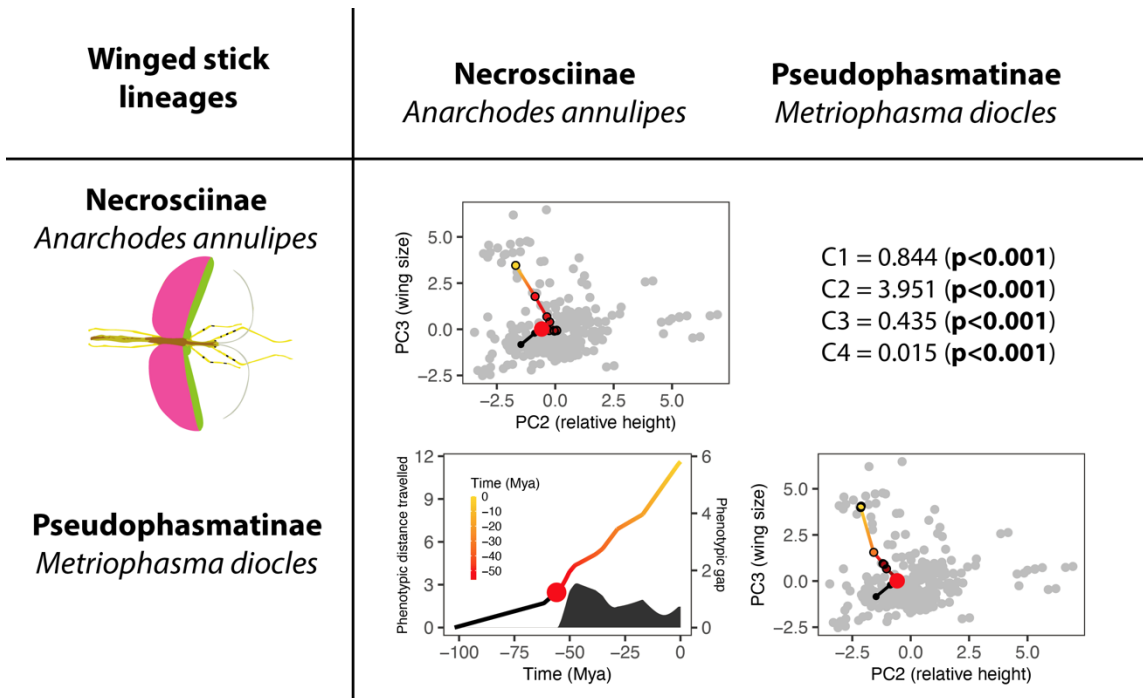


Figure S23: Convergent evolutionary trajectories of two lineages of winged sticks. (see legend of figure S7)

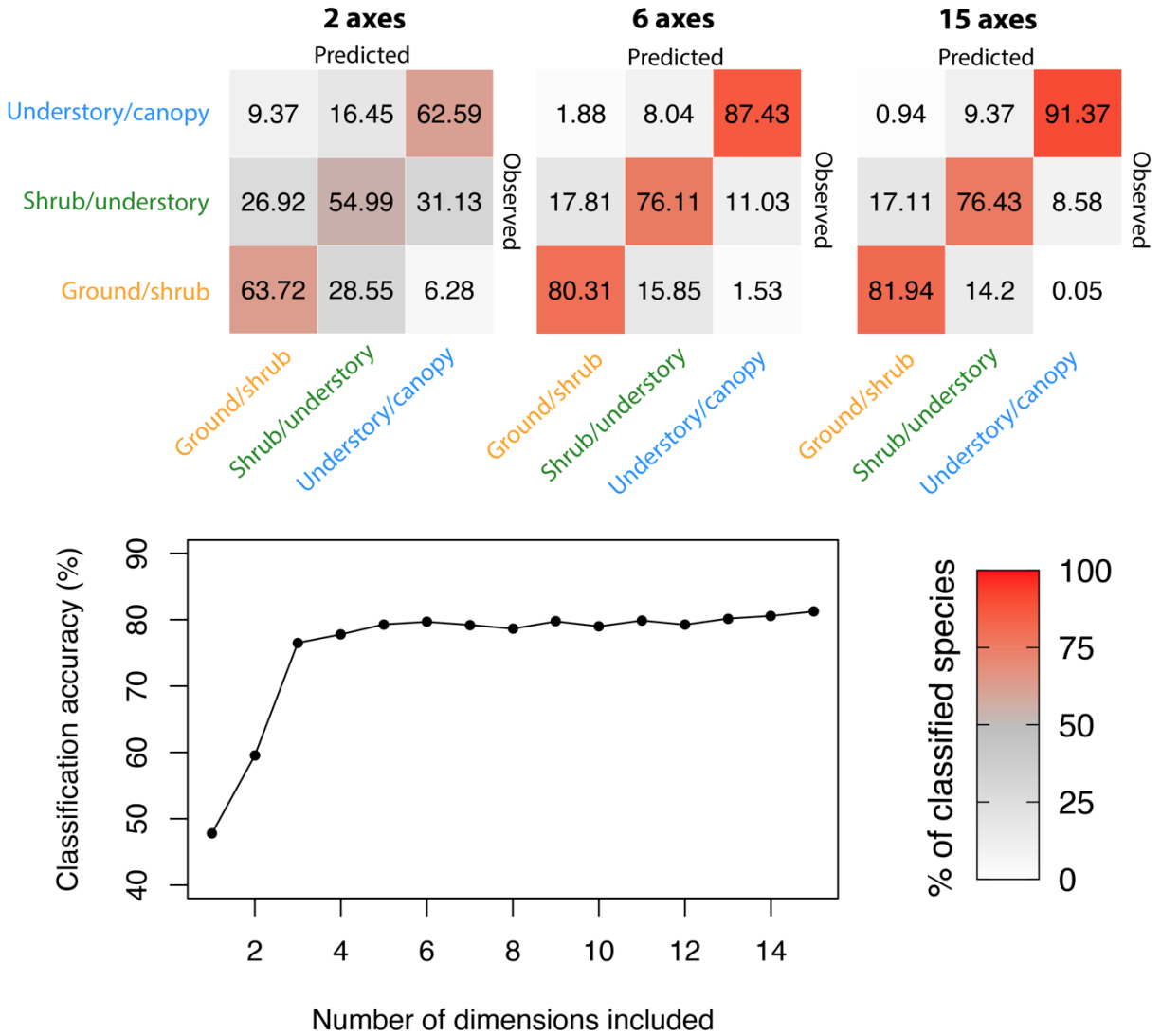


Figure S24: Habitat predictability based on morphological data. Heatmaps show prediction accuracy of random forest models for each habitat based on two, six or 15 morphospace axes. Predicted habitat states are displayed on the x axis and observed habitat states on the y axis. The line plot at the bottom shows the mean accuracy of the model at predicting habitat based on the number of morphospace axes provided.

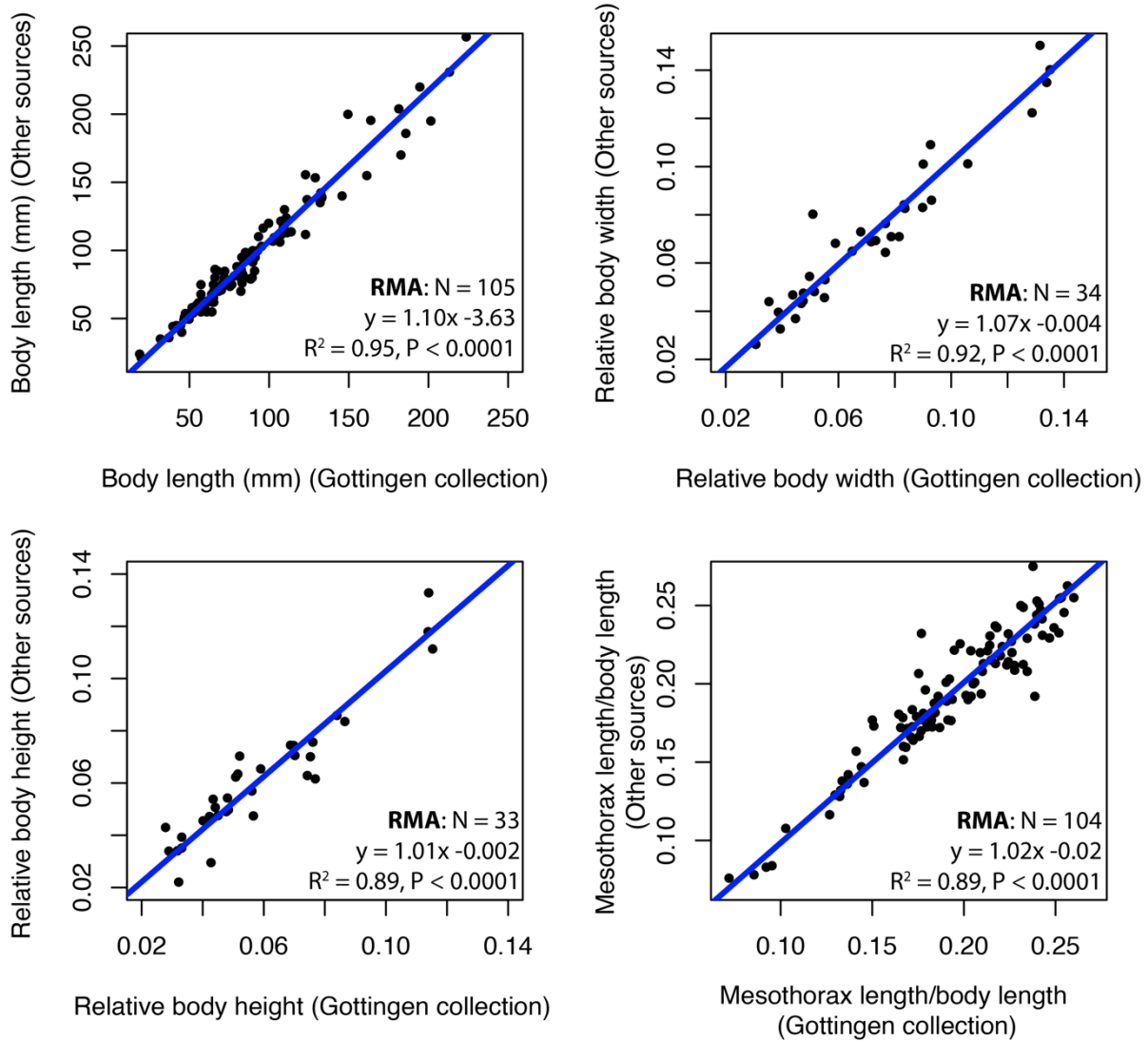


Figure S25 : Comparisons between measurements obtained from our own collection at the University of Göttingen and measurements obtained from all other sources. Dots correspond to individual species for which both data types were collected. Results from reduced major axis regressions (RMA) are included in each corresponding panel.

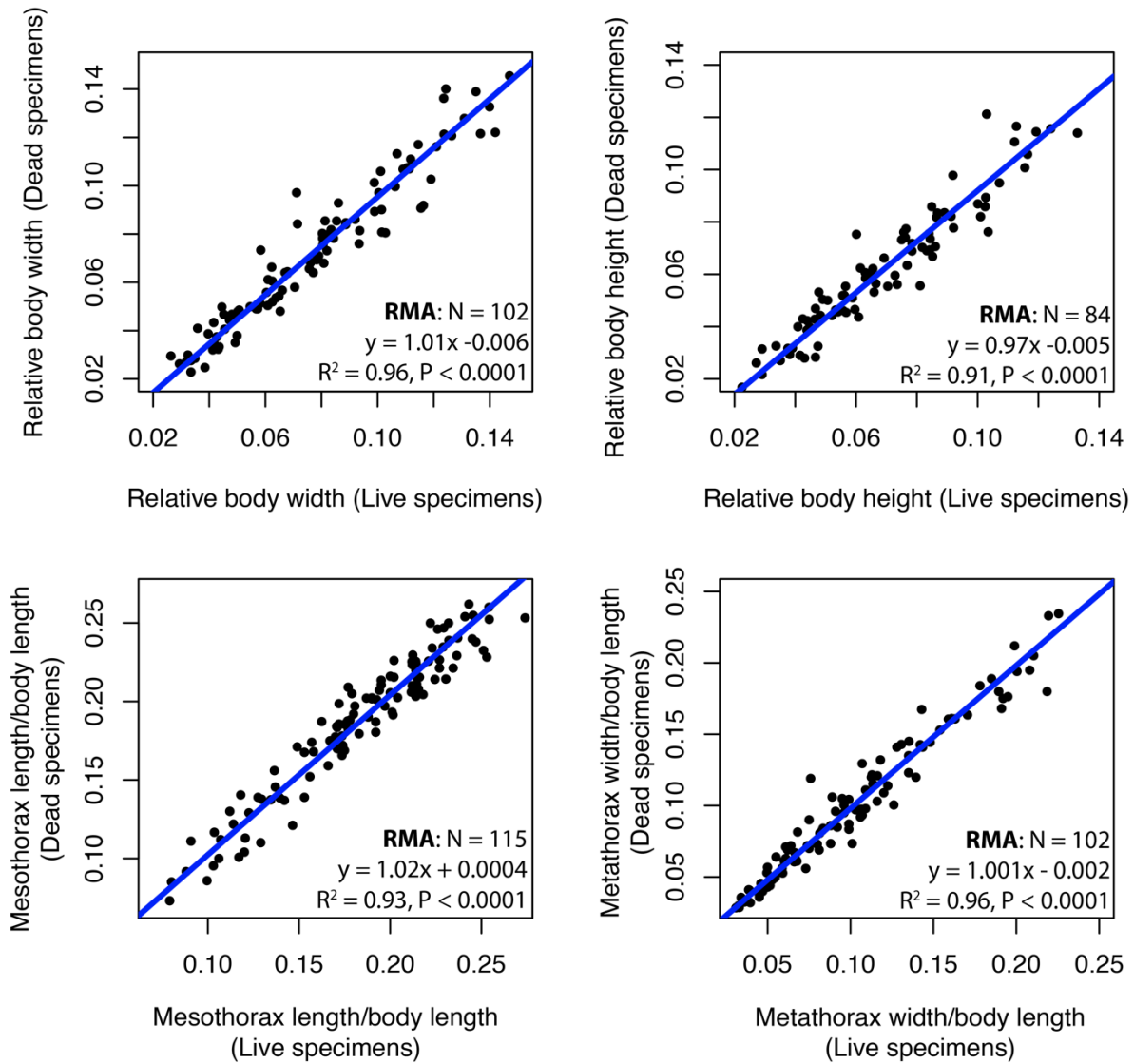


Figure S26: Comparisons between measurements obtained from pictures of dead specimens or live specimens. Dots correspond to individual species for which both data types were collected. Results from reduced major axis regressions (RMA) are included in each corresponding panel.

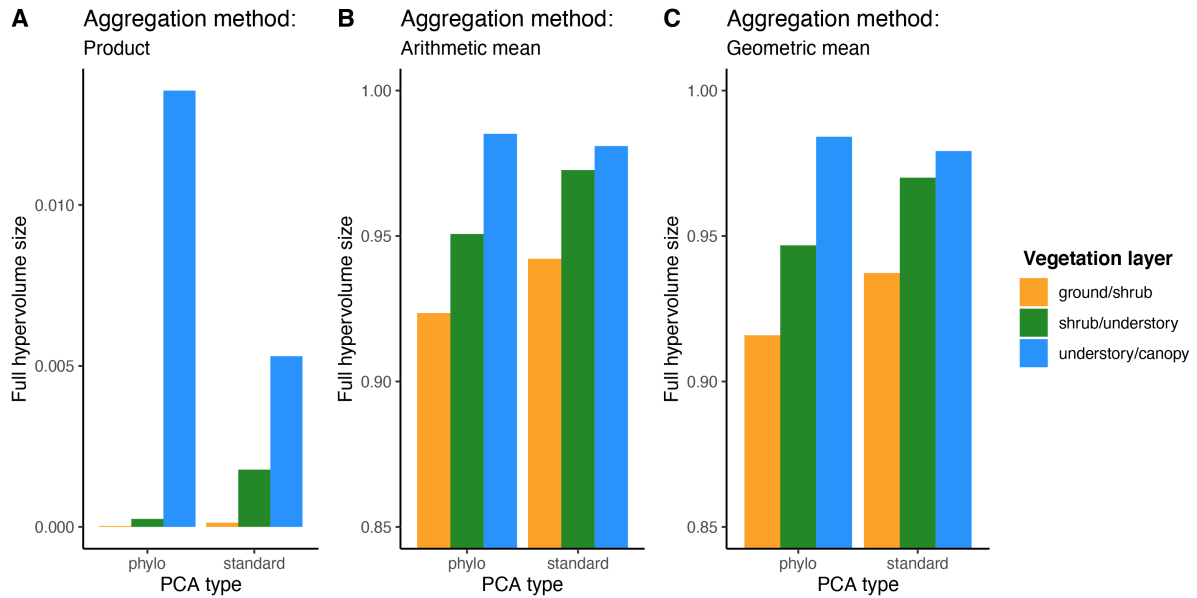


Figure S27: Habitat hypervolume size estimated from scores derived from a standard PCA and a phylogenetic PCA. Hypervolumes were calculated using the dynamic range boxes approach with the product (A), the mean (B) and geometric mean (C) aggregation methods.

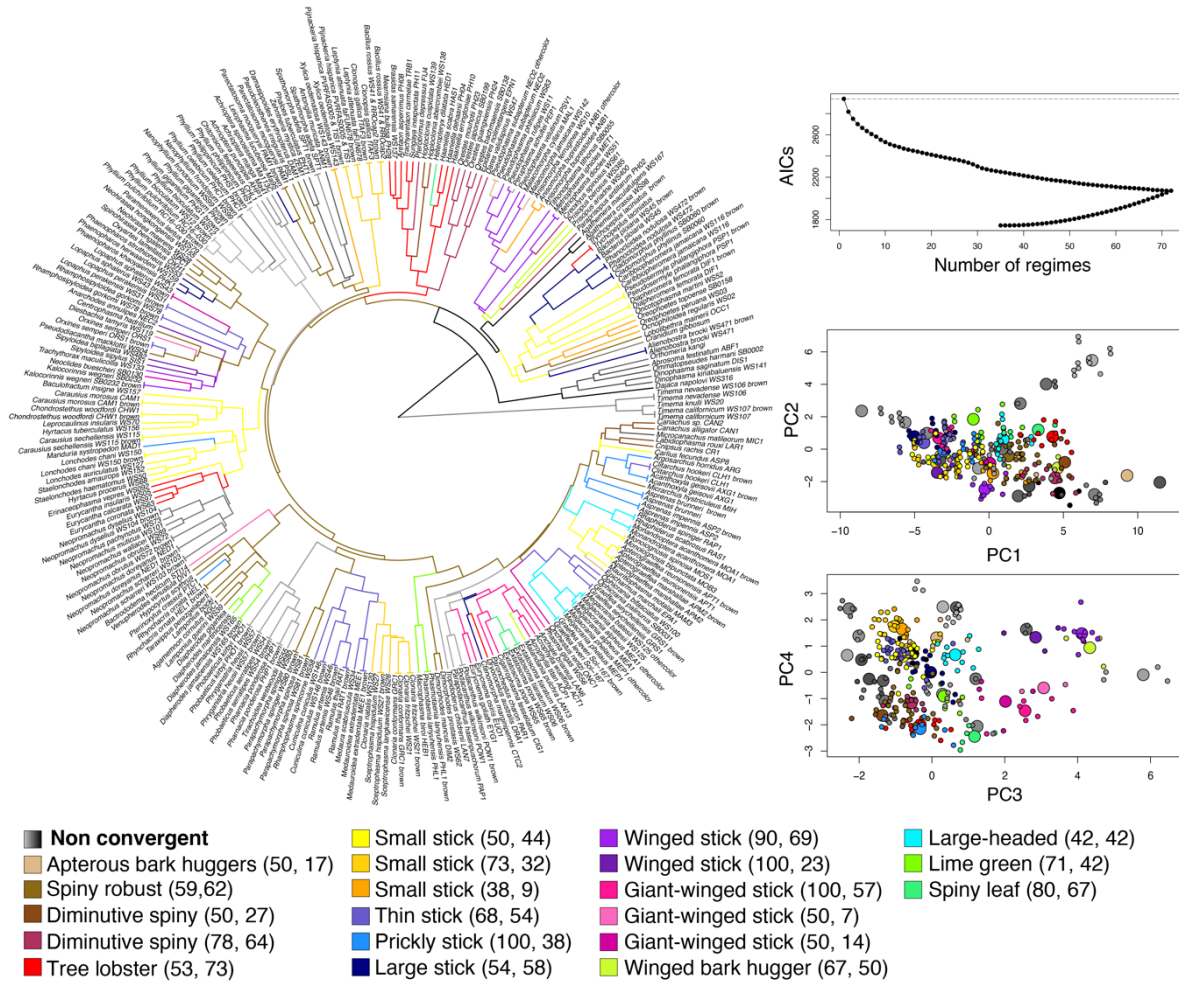


Figure S29: Results of the SURFACE analysis using the MCC phylogeny and the first four axes of the morphospace. MCC phylogeny (left panel) colored to show the different shifts in selective regimes. Convergent regimes (20 total) are colored while non-convergent ones are represented in gray. The top right panel shows the change in AIC of the OU models as they go through the two steps of the algorithm. In the first step, the algorithm iteratively adds the best regime shift to the model and as a second step tries to merge identical regime shifts. The two steps stop when the OU model cannot be further improved. Bottom right panels show how extant phasmid species (small circles) cluster near their inferred adaptive peak (large circles) on the morphospace. The inferred peaks broadly correspond to the morphotypes identified with the hierarchical clustering analysis. The color legend shows which SURFACE regime shift corresponds to which morphotype. The first number indicates the proportion of extant species corresponding to a given selective regime identified by SURFACE which were assigned to the named morphotype (in %). The second number indicates the proportion of species within a given morphotype also assigned to the given SURFACE selective regime.

Video S1: Dynamic 2-D phylomorphospace. At each time step starting with the ancestral node of Phasmatodea, the position of each lineage on the morphospace was reconstructed and the morphotype of each lineage was predicted using a trained random forest model (see methods).

CHAPTER 2

Sexual selection and flight as predictors of sexual size and shape dimorphism in stick and leaf insects

Sexual selection and flight as predictors of sexual size and shape dimorphism in stick and leaf insects

Romain P. Boisseau^{1,*}, Sven Bradler², Douglas J. Emlen¹

¹ Division of Biological Sciences, University of Montana, 32 Campus Dr, Missoula, MT 59812, United States of America

³ Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology, University of Göttingen, Göttingen, Germany

*Corresponding author: Romain P. Boisseau

Keywords: macroevolution | phasmatodea | Rensch's rule | protandry | body size | fecundity selection

Introduction

Phenotypic differences between the sexes are extremely widespread in animals and understanding their variation is a major question in evolutionary biology (Darwin, 1871; Fairbairn, 1997; Fairbairn et al., 2007). Body size differences in particular can be extreme (sexual size dimorphism; SSD), with mature females weighing hundreds of times more than their respective mates in some fishes, spiders, and marine invertebrates (Fairbairn, 2013). Interspecific variation in the extent of SSD is hypothesized to reflect differences in the relative intensity of three types of sex-dependent selective pressures (Blanckenhorn, 2005; Cox and Calsbeek, 2009; Littleford-Colquhoun et al., 2019). First, fecundity selection is expected to favor larger female sizes when their fitness increases with the number and size of eggs produced (i.e., reproductive output), and when larger body sizes enable greater reproductive outputs (Cox et al., 2003; Honěk, 1993; Pincheira-Donoso and Hunt, 2017; Shine, 1988). Second, sexual selection is expected to act primarily on male body size (Cox and Calsbeek, 2009; Horne et al., 2020), ubiquitously favoring relatively larger male sizes in species with resource-defense or lek mating systems (e.g., through male-male combat or female choice) (Andersson, 1994; Emlen, 2014; Hardy and Briffa, 2013) and sometimes favoring relatively smaller male sizes in species with scramble competition mating systems (e.g., through mobility selection) (Herberstein et al., 2017; Kelly et al., 2008). Finally, some aspects of natural selection may drive the evolution of sexual dimorphism through intersexual niche partitioning (Shine, 1989). In such cases, sex differences may arise initially via sexual or fecundity selection and subsequently be exacerbated by sex-dependent natural selection and adaptation to separate niches (secondary ecological dimorphism hypothesis) (Fairbairn, 1997; Stephens and Wiens, 2009; Svenson et al., 2016). Understanding adaptive interspecific variation in SSD therefore requires quantifying the relative contributions of these three types of selective forces (i.e., sexual, fecundity and ecological).

SSD is usually accompanied by sexual dimorphism in shape. Males and females may differ in body proportions (Butler et al., 2007), the shape of body structures (Berns and Adams, 2013), ornaments and weaponry (Emlen, 2008) and coloration (Dale et al., 2015). Interspecific variation in sexual shape dimorphism (SShD) is thought to result from the same three sex-dependent selective pressures (Adams et al., 2020), with fecundity selection primarily affecting female shape, sexual selection primarily affecting male shape, and some aspects of ecological selection acting differentially on both male and female shape.

We explored the selective drivers of interspecific variation in SSD and SShD across the stick and leaf insects (order Phasmatodea). Specifically, we tested whether (I) variation in fecundity selection was more strongly associated with variation in female than male size or shape; (II) variation in sexual selection was more strongly associated with variation in male than female size or shape; and (III) variation in ecological selection was differentially associated with variation in female or male size or shape.

Phasmids exhibit a spectacular diversity of body forms and sizes ranging from very elongated to much more robust bodies and from giant to more modest sizes (Bedford, 1978; Bradler and Buckley, 2018). For instance, the branch-like female *Phryganistria chinensis* (Zhao, informal name), considered the longest insect in the world, measures up to 37cm long (excluding legs and antennae)(Shi et al., 2019) while the sturdy female *Miniphasma prima* (Zompro, 1999) only grows to 1.7cm (Zompro, 1999). This diversity of body forms likely stems from the advergence of phasmid appearance towards various objects of their habitat (e.g., sticks, leaves, bark pieces, moss) which enables an astonishing camouflage through crypsis and masquerade. Even more notably, sexual dimorphism can be very strong in this group, to the point of confusing taxonomists into assigning males and females of the same species to separate genera (Cumming et al., 2020). As is most common in arthropods, males are usually (much) smaller and slenderer than females (Sivinski, 1979). In addition, specific traits such as relative leg size and shape, coloration, or the presence of wings or ocelli (i.e., simple eyes) are often dimorphic (Bank and Bradler, 2022; Boisseau et al., 2020). In many species, the much larger and sedentary females are flightless, either with shortened wings or no wings at all, while the mobile males can be fully winged and flight capable (Whiting et al., 2003). Phasmids also exhibit a remarkable array of reproductive modes and mating systems, from small mobile males searching for scattered sedentary females (Boisseau et al., 2022), to tiny monogamous males riding on the back of a female for life (Conle et al., 2009), large weapon-bearing males aggressively fighting for territories and females (Boisseau et al., 2020), and sometimes all the way to obligate thelytokous parthenogenesis (Schwander and Crespi, 2009). Therefore, stick and leaf insects constitute an ideal group in which to investigate the selective drivers of both SSD and SShD.

In the present study, we used morphological data on 196 phasmid species (6% of described species diversity and 27% described generic diversity) to uncover patterns of body size, shape, and wing sexual dimorphism in the group. We first mapped and quantified the variation of SSD and SShD across species. Then we tested for effects of female lifetime reproductive output (a proxy for

fecundity), mating system (a proxy for the direction and strength of sexual selection), and sex-specific flight capacity, climate, and habitat (proxies for dimorphic ecological selection) on the variation of each sex's size and shape, and on SSD and SShD. Finally, we also investigated proximate mechanisms contributing to SSD in this group by evaluating the relative importance of sex differences in growth and development time.

Materials and methods

The present study was carried out in tandem with a companion study focusing on ecomorphological convergence in female morphology across Phasmatodea (Chapter 1). As such, the two studies used the same phylogeny, methods and data on female morphology and habitat classification.

Taxon sampling and phylogeny. We used a phylogeny including a total of 314 phasmid taxa (9% of phasmid species diversity and 33% of generic diversity) and one embiopteran species as outgroup. The phylogenetic reconstruction was carried out using genetic data from 3 nuclear (*i.e.*, 18S rRNA, 28S rRNA and histone subunit 3) and 4 mitochondrial genes (*i.e.*, 12S rRNA, 16S rRNA, cytochrome-c oxidase subunit I and II), the basal topology of transcriptome-based trees (Simon et al., 2019; Tihelka et al., 2020) and Bayesian inferences (see Chapter 1 for details).

Morphological data. We measured 1,365 adult female (212 species) and 1,000 adult male specimens (200 species, 196 species with specimens of both sexes). Measurements from female specimens have already been used in the companion study focusing on female morphology (Chapter 1). We used ImageJ (v1.51)(Schneider et al., 2012) and high quality photographs of live or pinned specimens in dorsal and/or lateral view from our own collection at the University of Göttingen (Germany), field guides (Brock and Hasenpusch, 2009; Seow-Choen, 2005; Seow-Choen, 2016; Seow-Choen, 2017; Seow-Choen, 2018), the published literature, and various online databases (Brock et al., 2021) (Dataset S1). Between 1 and 23 different individuals per sex and per species (mean= 6.4 females and 5 males per species) were measured. Because photographs were sometimes unscaled, we measured each of the 23 continuous traits described in Figure S1 relative to body length. We measured body length directly (excluding ovipositor and subgenital plate) on properly scaled photographs and collected body length data from the literature (Dataset S1). The median of each relative measurement was then multiplied by

the sex- and species-specific median of body length to obtain absolute measurements for each sex and species. Further methodological details are available in Chapter 1.

Body volume was used as a proxy for body size and corresponded to the volume of an elliptic cylinder calculated from body length, average body width and average body height (Figure S1). SSD was quantified using a commonly used sexual dimorphism index (SDI) (Lovich and Gibbons, 1992):

$$\begin{cases} \text{SSD} = \left[\left(\frac{\text{volume of larger sex}}{\text{volume of smaller sex}} \right) - 1 \right] \times (-1), & \text{if male is larger} \\ \text{SSD} = \left[\left(\frac{\text{volume of larger sex}}{\text{volume of smaller sex}} \right) - 1 \right] \times (1), & \text{if female is larger} \end{cases} \quad \text{(Equation 1)}$$

Females were always the larger sex in our dataset so SSD was always positive. To quantify body shape while controlling for size, we performed a Principal Component Analysis (PCA, *R function*: “R package”; *prcomp*: “stats”) including all absolute measurements (except body volume), but original values were substituted with residuals calculated from ordinary least-squares (OLS) regressions against body volume (*lm*: “stats”). Wing length and area, which included zeros for wingless species, were divided by body length or body length squared respectively to obtain measures of relative wing length and area. These variables were all mean-centered and scaled to unit variance. PC1 accounted for 59% of the total variation and reflected body elongation, while PC2 accounted for 12% of the total variation and reflected wing size (Figure S2). We focused on these two aspects of shape to calculate SShD in body elongation (SShD_{elongation}) and relative wing size (SShD_{wings}) as the difference between male and female PC1 or PC2 respectively. A larger and positive SShD_{elongation} and SShD_{wings} respectively indicate that males are more elongated than females, or have larger wings. We also calculated overall SShD as the Euclidean distance between males and females considering their coordinates along PC1-PC4 (cumulatively accounting for 86% of the variation).

Life history and ecological data. Stick insects are elusive, nocturnal animals whose natural history in the wild is poorly known. However, many species are bred in captivity as pets, especially in Europe. Thus, some information about their natural history and behavior is available from amateur enthusiasts and breeders. We surveyed the literature, field guides, breeding guides and online databases to collect information about fecundity, mating system, flight capacities and habitat (Dataset S1).

Fecundity. We looked for data on average egg mass and number of eggs laid by a female during its lifetime for as many species as possible. We also collected egg pictures in dorsal and lateral view

from the scientific literature and private collections, in particular from the photographic database of F. Tetaert (available from phasmatodea.fr, accessed in July 2021, used with permission). Using these pictures, we measured egg length, width and height following the methods of (Church et al., 2019a; Church et al., 2019b). We then estimated egg volume from these measurements (Figure S3). After verifying the strong correlation between egg mass and volume data (Figure S4), we chose to use egg volume as our estimate of egg size as we had a larger sample size (84 species with mass data, 143 species with volume data). Finally, the fecundity of females was quantified as their lifetime reproductive output, calculated as the product of egg volume and the lifetime number of eggs laid (n=90 species with sufficient data). Under the hypothesis that variation in fecundity selection (i.e., favoring larger females) drives variation in SSD and SShD, we predicted that species where females have a relatively higher reproductive output would display a relatively higher SSD and body SShD.

Mating system. We classified mating systems of 101 species into three broad categories: “searching”, “guarding”, and “monogamy” based on the typical duration of the association between males and females, including the initial copulation and any post-mating guarding period. Males of species in the “searching” category typically spend less than a day paired up with a single female. In these species, male fitness is probably mostly determined by their capacity to locate and move efficiently between scattered females, not by direct male-male competition. In this context, we expect smaller and more mobile males to be favored as they likely benefit from a more efficient locomotion (walking or flying) (Boisseau et al., 2020; Boisseau et al., 2022; Kelly et al., 2008).

Males of species in the “guarding” category typically stay with the same female for more than a day after finding it. Intromission may occur only initially or be repeated intermittently. Either way, the male often spends most of his time mounted on the female but not actively transferring sperm. Males eventually leave the female after a few days, typically less than four days, to search for another mate. This mate guarding strategy suggests a higher risk of sperm competition and a higher level of sexual selection, and therefore male fitness may also depend on the ability of a male to defend and guard access to a female from rival males. Direct combats between a guarding and an intruder male have been reported in several species (Boisseau et al., 2020; Sivinski, 1979). In this case, we expect both male fighting and mobility abilities to be important. Given that, unlike mobility selection, direct male combat almost ubiquitously favors larger sizes, we expect male size in these species to reflect this balance and to be relatively larger than species with non-guarding males.

Finally, some phasmid males spend more than four days, and in this case often their lifetime, riding on one female's back, often staying in copula with their body essentially acting as a sperm plug. In these largely monogamous species, we expect male and female fitness interests to align and therefore smaller male sizes to be favored so as not to impair female locomotion (walking or flying), and to limit their need to forage (potentially away from the female).

Climatic data. We gathered information about the geographic range of each species based on sampling location of type specimens and observations on iNaturalist (available from <https://www.inaturalist.org>. Accessed July 2021). For each species, we then selected the median location with the most central latitude. From the GPS coordinates of the most central location for each species, we extracted data on annual mean temperature, mean diurnal range (i.e., mean of monthly (maximum - minimum temperature)), temperature seasonality (i.e., standard deviation $\times 100$), annual temperature range (i.e., maximum temperature of warmest month - minimum temperature of coldest month), annual precipitation, precipitation seasonality (i.e., coefficient of variation) from worldclim (Fick and Hijmans, 2017) (available from <https://www.inaturalist.org>. Accessed July 2021). We also extracted the length of the growing period zone (i.e., number of days during a year when temperatures are above 5°C and precipitation exceeds half the potential evapotranspiration (van Velthuisen et al., 2007), world map divided in 16 zones, available from <https://data.apps.fao.org/map/catalog>), net primary production of biomass (grams of dry matter per m² per year; Climate Research Unit, Univ. of East Anglia, period 1976-2000, available from <https://data.apps.fao.org/map/catalog>) from the FAO Map Catalog; and total annual growing degree days (i.e., a measure of the annual amount of thermal energy available for plant and insect growth; Climate Research Unit, Univ. of East Anglia, available from <https://sage.nelson.wisc.edu/data-and-models>). Such data was available for all species (n=196). Given that these nine climatic variables are highly correlated, we ran a principal component analysis (Figure S5). We kept the first component (climate PC1, explaining 60.5% of the total variation) to quantify climatic variation between species. Climate PC1 mainly reflected net primary productivity (i.e., food availability for herbivorous insects) and was positively correlated with annual precipitation and mean temperature, and negatively correlated with the annual and diurnal temperature range (i.e., seasonality). Thus, PC1 is overall high in tropical regions, and low in more temperate and seasonal regions. Season length and food availability may limit the growth of some species. Because females are larger than males, we expect this hypothetical growth limitation to affect primarily females and

therefore to reduce SSD. Therefore we predicted a higher SSD in tropical regions (high PC1, long growing season) than in temperate regions.

Flight dimorphism. We recently showed that large body sizes impair the flying performance of male phasmids (Boisseau et al., 2022). Larger males, with a relatively high wing loading, have more difficulties climbing and maneuvering in the air and are at higher risk of slipping and crashing when landing. This suggests that the mode of locomotion of males and females –i.e., flying and walking, or walking only—is likely to affect their size and shape evolution. We gathered observational data on the flight capacities of both sexes across species. Each sex was scored as either flight capable (including gliding) or flightless (including parachuting). Because flight can be reliably predicted from morphological characteristics and especially the size of the wings relative to the size of the animal, we trained random forest models (i.e., a classification machine learning algorithm) to predict flight in species with missing observations (*randomForest*: “randomforest” (Liaw and Wiener, 2002)). The algorithm was trained separately for each sex on the entire subset of the dataset for which flight data could be collected and was provided with data on sex-specific body volume, body length, relative body width, femur length, wing length, wing aspect ratio and wing area. Error rate was 1.18% for males and 2.16% for females. We used the trained algorithms to predict both male and female flight in species with missing flight data (n=8 species for female and n=19 species for male flight, over 196 species total). Finally, we qualitatively categorized flight dimorphism in each species as either both sexes being flightless, flight capable, or males being flight capable and females flightless. We expected species with only males able to fly to be more dimorphic as ecological selection related to flight would only affect males and favor relatively smaller sizes with larger wings.

Habitat. Finally, we classified each species’ habitat as the vegetation layer (i.e., ground/shrub, shrub/understory and understory/canopy) that they typically inhabit during the daytime (i.e., when they are exposed to visually hunting predators). The ground/shrub layer includes species mostly found resting below 1.5m (i.e., in the leaf litter, in the grass, in low shrubs, on logs, etc.). The shrub/understory layer corresponds to species resting at heights between 1.5m and 4m (i.e., in tall shrubs, on tree trunks, or in small trees). Finally, the understory/canopy layer corresponds to species typically resting in the canopy of tall trees (>4m). Habitat could be classified for 195 species. These strata differ in their structural and aerodynamic properties, predator communities and visual characteristics. For instance, unlike the ground/shrub layer, the canopy is characterized by substantial gaps, wind gusts, and a visual environment dominated by green leaves. Males and females often differ

in their locomotor needs: males need to move to find females while females may only need to move to feed and sometimes to lay eggs. Thus, we expect that the vegetation layer where a species mainly resides may affect sexual dimorphism through ecological selection on locomotor performance. We specifically predicted that higher strata would be associated with higher levels of sexual dimorphism as locomotion may be more challenging in high gappy habitats.

Ontogenetic data. We collected data on sex-specific postembryonic development duration (from hatching to final molt) from the literature and amateur breeding guides (Dataset S1). Despite such information often being reported qualitatively (± 2 weeks) without accounting for temperature, most phasmids are captive bred at room temperature (19-24°C) and exhibit extensive variation in development time across species (from 2 to 13.5 months) relative to measurement error, thus rendering the data informative. Sex-specific relative growth rate (RGR) was calculated as:

$$RGR = \frac{\log_{10}\left(\frac{\text{Adult body volume}}{\text{Egg volume}}\right)}{\text{Development duration}} \quad \text{(Equation 2)}$$

RGR therefore represents a growth rate averaged over the entire postembryonic development period (Tammaru and Esperk, 2007). Egg volume was used as a proxy for the volume of first-instar individuals, assuming that egg sizes do not differ between males and females, which has, to our knowledge, never been observed in phasmids. Development duration and RGR could be found or calculated in both sexes in a total of 67 species.

Because females always had a higher (or equal) RGR and development time than their respective males, we calculated sex differences in RGR (SDRGR) and duration of development (SDDD) as:

$$SDRGR = \frac{\text{Female RGR}}{\text{Male RGR}} \quad \text{(Equation 3)}$$

$$SDDD = \frac{\text{Female development time}}{\text{Male development time}} \quad \text{(Equation 4)}$$

Statistical analyses. Statistical analysis were carried out in R (v4.1.1) (R Core Team, 2021). SSD and overall SShD were \log_{10} -transformed to improve data distribution for the subsequent analyses. We first mapped the evolution of SSD and overall SShD on the Phasmatodean phylogeny using a fast estimation of maximum likelihood ancestral states (*fastAnc*: “phytools” (Revell, 2012)).

Allometry and SSD. Female-biased SSD in phasmids may have evolved primarily through negative directional selection acting on male size or positive directional selection on female size. Quantitative genetic theory predicts that the sex under historically stronger directional selection will exhibit greater interspecific variation. Consequently, as the extent of SSD and the allometry between \log_{10} male versus \log_{10} female body size reflects the history of selection on male and female size, we expect a correlation between the interspecific allometric slope (β) of \log_{10} male versus \log_{10} female body volume and the extent of SSD across related clades (De Lisle and Rowe, 2013; Zeng, 1988). A positive correlation between β and SSD among related clades would indicate that selection on average has acted more intensely on male body size, while a negative correlation would indicate more intense selection on female size as the primary explanation for why females are larger than males in phasmids. Therefore, for each phasmid clade with sufficient data ($n_{\text{species}} \geq 6$), we calculated β both as the phylogenetic reduced major axis (PRMA, *phyl.RMA*: “phytools”) and phylogenetic generalized least-squares (PGLS, *gls*: “nlme”, *corPagel*: “ape” (Paradis and Schliep, 2019), Pagel’s λ estimated by maximum likelihood) regression slopes of \log_{10} male versus \log_{10} female body volume. A phylogenetically controlled mean SSD was also computed for each clade by using an intercept-only PGLS model in which \log_{10} SSD was the response variable. Pharnaciini, the European clade, Stephanacridini and Agathemeridae (comprising less than 5 species included in the phylogeny) were grouped with their closest relatives, respectively Clitumninae + Pachymorphinae, the African/Malagasy clade, Lanceocercata and Pseudophasmatinae.

The correlation between β and mean SSD across major clades was tested using a PGLS model taking into account the phylogenetic relationships between the clades. Similarly, Rensch’s rule suggests that SSD will be negatively correlated with body size (average between males and females) in species in which females are larger than males, and that males are always the sex with the greatest interspecific variation in size (i.e., $\beta > 1$) (Rensch, 1959). To test Rensch’s rule across Phasmatodea, we first quantified the allometry of \log_{10} male versus \log_{10} female body volume across all phasmid species, using

both a PGLS and PRMA, as advised in a recent study (Meiri and Liang, 2021). We also regressed \log_{10} SSD on the mean body volume of males and females of each species using a PGLS.

Proximate causes of SSD. Females may be larger than males because they develop faster and/or for longer. We tested the association between SSD and sex difference in duration of post embryonic development and sex difference in relative growth rate using PGLS models. We tested the effect of \log_{10} (sex difference in duration of development) and \log_{10} (sex difference in relative growth rate) (explanatory variables) on \log_{10} SSD (response) using two single predictor and one multiple predictor model. \log_{10} (sex difference in duration of development) and \log_{10} (sex difference in relative growth rate) were mean-centered on zero and scaled to unit variance to allow meaningful comparison of their effect sizes and relative importance in determining SSD. In the multiple predictor model, the significance of the effect of each explanatory variable was assessed using a type III ANOVA.

Fecundity, sexual selection, and ecological selection drivers of SSD and SShD. To evaluate the predictions from the non-mutually exclusive ultimate hypotheses I-III laid out in the introduction, we tested the effect of variation in ecological variables and reproductive life-history characteristics on the variation of female and male body volume, body elongation (PC1) and relative wing size (PC2), and consequently on the extent of SSD, $SShD_{\text{elongation}}$ and $SShD_{\text{wings}}$. We used either \log_{10} male or female body volume, body elongation (PC1) or relative wing size (PC2), or \log_{10} SSD, $SShD_{\text{elongation}}$ and $SShD_{\text{wings}}$ as dependent variables and \log_{10} lifetime female reproductive output (i.e., fecundity), climate PC1, mating system, flight dimorphism and habitat as explanatory variables in single predictor PGLS models (including each explanatory variable alone to maximize sample size) and in multiple predictor PGLS models (including all five explanatory variables at once and no interaction). All continuous variables were mean-centered and scaled to unit variance to enable meaningful comparisons of effect sizes. In multiple predictor models, the significance of the effect of each explanatory variable was assessed using type II ANCOVAs. For categorical predictors showing significant effects, we subsequently performed post-hoc pairwise comparisons of estimated marginal means between groups using the Holm correction to account for multiple testing (*emmeans*: "emmeans"). We evaluated the proportion of variance explained by the PGLS models by calculating two R^2 coefficients (R^2_{resid} : "rr2" (Ives and Li, 2018)): R^2_{full} , the total variance explained by the full model, and R^2_{fixed} , the variance explained by the fixed effects only, after accounting for phylogeny.

Results

Variation in SSD and SShD across Phasmatodea. SSD and SShD varied extensively across clades and species of stick and leaf insects (**Figure 1-2**, Table S1). SSD was exclusively female-biased but varied from females only being 1.25 times more voluminous than males in *Epicharmus marchali* (Lanceocercata), all the way to 22.5 times larger in *Cranidium gibbosum* (Diapheromerinae). SSD and overall SShD were loosely but significantly correlated (Figure S6; PGLS: $\lambda = 0.59$, $F_{1,184} = 9.82$, $p = 0.002$, $R^2_{\text{full}} = 0.05$). Thus, strongly size dimorphic species tended to also exhibit strongly dimorphic body shapes, especially with much smaller males also being much slenderer and carrying larger wings.

A PGLS regression between \log_{10} male body volume and \log_{10} female body volume across all phasmid species showed an allometric slope β of 0.82 that significantly departed from 1 (**Figure 2A**, $\lambda = 0.79$, 95% confidence interval (CI): 0.76–0.89), thereby supporting the inverse of Rensch’s rule. However, the corresponding PRMA exhibited a β of 0.96, which was not significantly different from isometry (**Figure 2A**, $\lambda = 0.96$, 95% CI: 0.88–1.03). Finally, we found a positive relationship between SSD and the mean species body volume (**Figure 2B**, PGLS: $\lambda = 0.62$, slope = 0.16 ± 0.05 , $F_{1,194} = 10.78$, $p = 0.001$), again going against the definition of Rensch’s rule.

In accordance with the predictions of quantitative genetic theory (De Lisle and Rowe, 2013; Zeng, 1988), we found a significant negative relationship between clade-specific phylogenetically corrected mean SSD values and the associated β values calculated with the PRMA ($F_{1,9} = 5.16$, $p = 0.049$, **Figure 2C**) or PGLS ($F_{1,9} = 7.37$, $p = 0.024$, **Figure 2D**) methods. This suggests that, historically, the evolution of female-biased SSD in stick insects has been primarily driven by positive directional selection on female size.

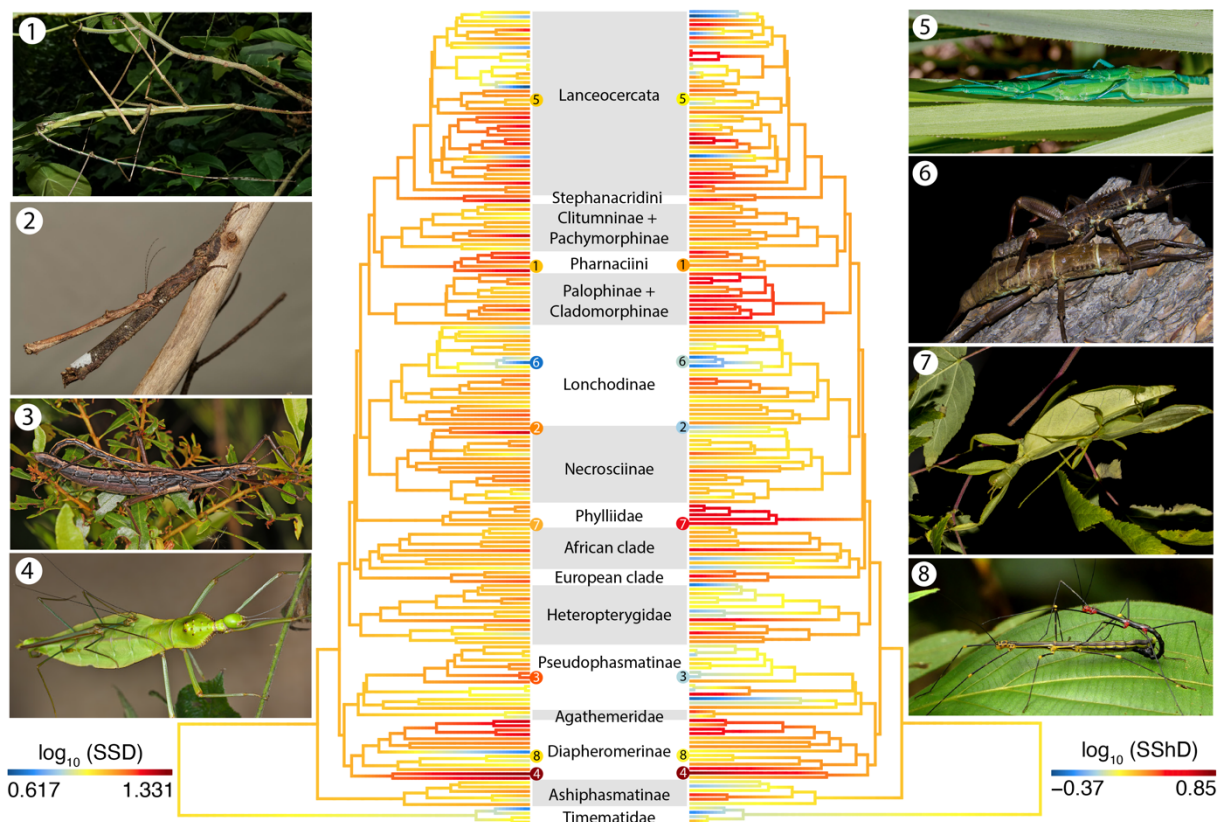


Figure 1: Diversity and phylogenetic pattern of sexual size dimorphism (SSD) and sexual shape dimorphism (SShD) in stick and leaf insects. The trees show the ancestral state reconstructions of SSD and overall SShD on the Phasmatodean tree. Numbered tips correspond to the pictures of mating pairs on the sides. **Picture 1:** *Phryganistria* sp. (Pharnaciini), photo by Rejoice Gassah (CC BY 4.0). **Picture 2:** *Neoclides buesheri* (Necrosciinae), photo by Bruno Kneubühler (used by permission). **Picture 3:** *Anisomorpha buprestoides* (Pseudophasmatinae), photo by Judy Gallagher (CC BY 2.0). **Picture 4:** *Cranidium gibbosum* (Diapheromerinae), photo by Bruno Kneubühler (used by permission). **Picture 5:** *Megacrania batesii* (Lanceocercata), photo by David White (CC BY-NC 4.0). **Picture 6:** *Eurycantha calcarata* (Lonchodinae), photo by Romain Boisseau. **Picture 7:** *Phyllium philippinicum* (Phylliidae), photo by Romain Boisseau. **Picture 8:** *Oreophoetes topoense* (Diapheromerinae), photo by Andreas Kay (CC BY-NC-SA 2.0).

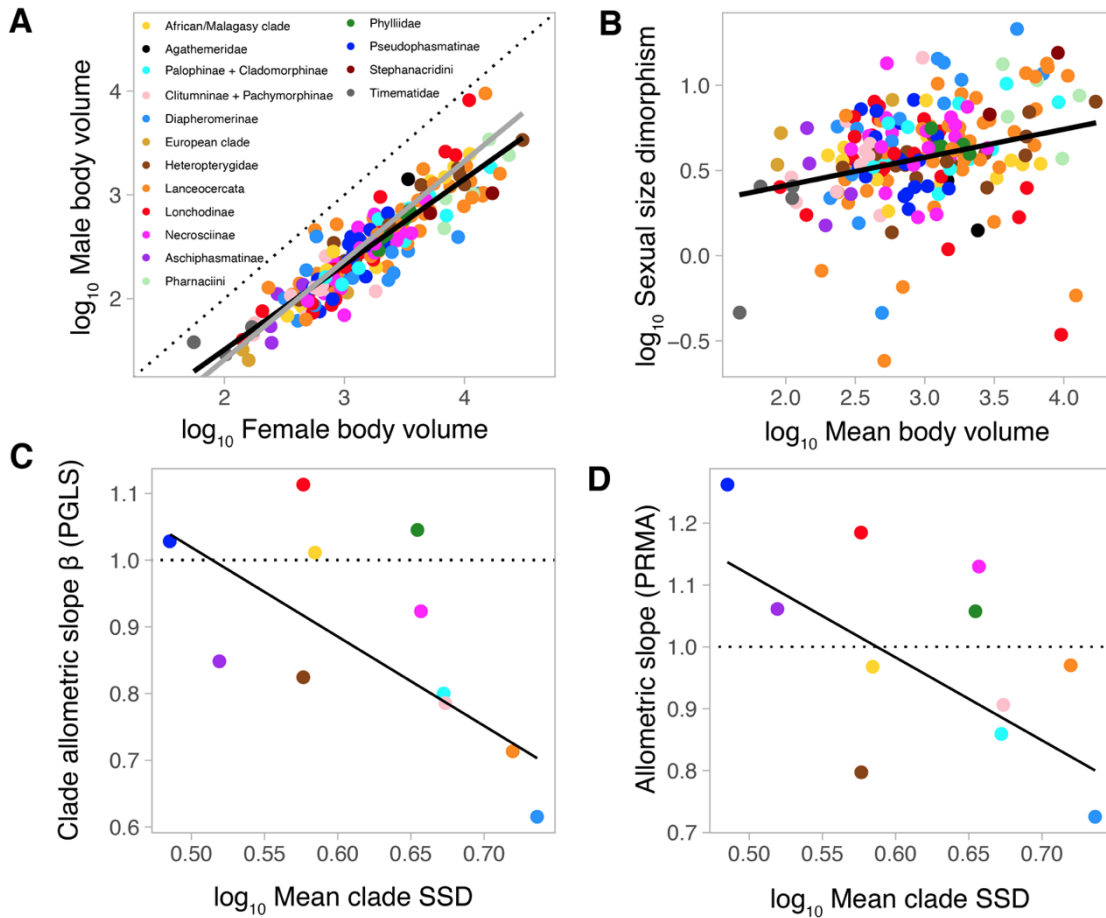


Figure 2: Sexual size allometry in stick insects (Phasmatodea). **A.** \log_{10} male body volume versus \log_{10} female body volume (mm^3) with PGLS (black) and PRMA (gray) regressions shown. Colors correspond to the main phasmid clades. **B.** \log_{10} Sexual size dimorphism (SSD) versus \log_{10} mean body volume of males and females. **C-D.** Allometric slopes β , calculated from clade-specific PGLS (**C**) or PRMA (**D**), versus phylogenetically controlled mean SSD by major phasmid clade (see color legend in **A**). The dashed lines correspond to isometry (1:1 relationship). Associated clade-specific data can be found in Table S1.

Proximate causes of SSD. SSD was positively correlated with sex difference in development duration and sex difference in relative growth rate (Table 1, Figure 3). The multiple-predictor PGLS further revealed that scaled sex differences in growth rate and in duration of development had similar effect sizes on the extent of SSD (Table 1), thus indicating that females are larger than males in phasmids because they develop faster and for longer, with equal importance.

Table 1: Sexual size dimorphism is caused by both sex differences in postembryonic development duration and relative growth rate. The table presents results of three phylogenetic generalized least squares models (PGLS) - and corresponding type III ANOVA-- including sexual size dimorphism (SSD) as the response variable (\log_{10} -transformed) and sexual dimorphism in duration of development (SDDD) and/or in relative growth rate (SDRGR). SDDD and SDRGR were \log_{10} -transformed, centered on zero and scaled to unit variance. Bold characters highlight significant effects ($p < 0.05$).

PGLS	Pagel's λ	Response	Explanatory variable	Effect size (\pm SE)	F _{df1, df2}	P
Single-predictor	0.30	\log_{10} (SSD)	\log_{10} (SDDD)	0.16 ± 0.03	30.0 _{1,65}	<0.0001
Single-predictor	0.02	\log_{10} (SSD)	\log_{10} (SDRGR)	0.14 ± 0.03	18.6 _{1,62}	0.0001
Multiple-predictor	0.18	\log_{10} (SSD)	\log_{10} (SDDD)	0.20 ± 0.02	84.1 _{1,60}	<0.0001
			\log_{10} (SDRGR)	0.18 ± 0.02	77.4 _{1,60}	<0.0001
			interaction	-0.02 ± 0.02	0.51 _{1,60}	0.48

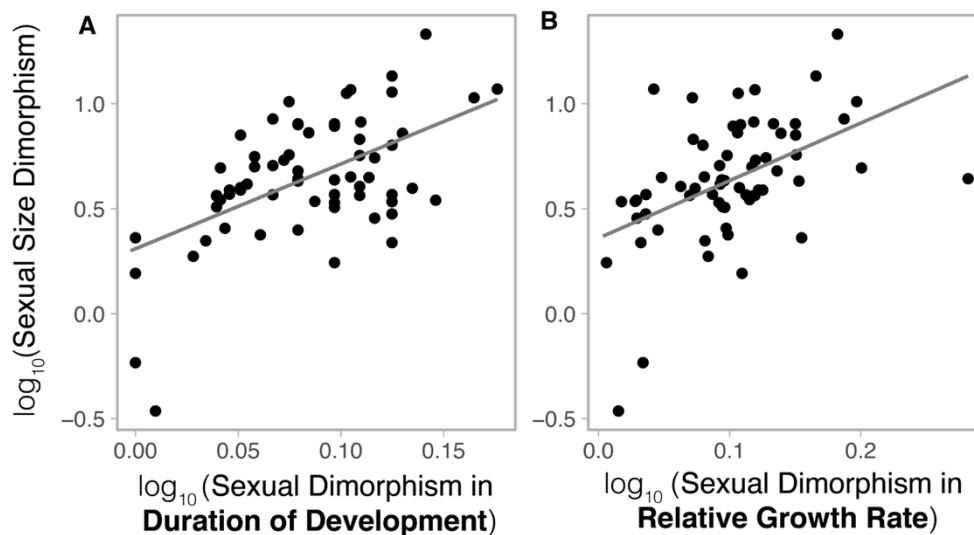


Figure 3: Sex differences in duration of postembryonic development (SDDD) and relative growth rate (SDRGR) explain sexual size dimorphism (SSD). \log_{10} (SSD) versus \log_{10} (SDDD) (A) and \log_{10} (SDRGR) (B), with PGLS regressions shown. Details of the associated statistical tests can be found in **Table 1**.

Selective drivers of variation in SSD. Single and multiple predictor PGLS models showed that female reproductive output (fecundity) was positively related to both female and male size, but not to SSD (Table S2, **Figure 4A-D**), thereby rejecting hypothesis I. Female fecundity alone explained 67% of the variation in female volume and 51% of the variation in male volume, but 0% of the interspecific variation in SSD (Table S2), after accounting for phylogenetic relatedness. This result seems contradictory with the previous finding that female-biased SSD in phasmids resulted from historically stronger positive selection on females. Therefore, it suggests that fecundity selection and overall positive selection on female size has driven an average female-biased SSD in phasmids but that variation in SSD around that average cannot be explained by variation in female fecundity, suggesting that other factors may be at play.

Both model types also showed that mating system significantly affected male but not female size, thus affecting SSD (Table S2, **Figure 4I-L**). These findings supported hypothesis II and were consistent with our predictions: guarding males are larger than non-guarding searching males and monogamous males. Mating system alone explained 0% of the variation in female volume and 5% of the variation in male volume, but 29% of the variation in SSD (Table S2), after accounting for phylogenetic relatedness.

We did not find any correlation between climate and male or female body volume or SSD (Table S2, **Figure 4E-H**), suggesting that ecological selection resulting from limited food availability or growing season length did not account for species differences in SSD.

Single predictor models revealed that males and females were significantly larger in species where only males are able to fly (Table S2, **Figure 4M**), which went against our prediction that flight-capable males would be smaller than flightless ones. Nevertheless, this effect was stronger in females than males resulting in a significantly higher SSD in species with flight capable males and flightless females (Table S2, **Figure 4N**), which was consistent with our original predictions. Flight dimorphism alone explained 9% of the variation in female volume, 4% of the variation in male volume, and 5% of the variation in SSD (Table S2), after accounting for phylogenetic relatedness. Our multiple predictor models similarly revealed that species with flying males and flightless females were more dimorphic than species with both sexes being flightless (Table S2, **Figure 4P**). However, they showed that this difference in SSD was mainly caused by flight-capable males being smaller in species with only males able to fly and not by females being larger (Table S2, **Figure 4O**), thus agreeing with our predictions.

Species with both sexes able to fly, for which we predicted a reduced size of males and females and a low SSD, did not differ from other categories in any of the models probably owing to a reduced sample size (n= 18 species, only including 9 with sufficient data to be included in the multiple predictor models).

Single predictor models showed that species dwelling in higher vegetation strata had a higher SSD, owing to a stronger effect on female than male size (Table S2, **Figure 4Q-R**). Habitat alone explained 16% of the variation in female volume, 8% of the variation in male volume and 7% of the variation in SSD (Table S2), after accounting for phylogenetic relatedness. In contrast, multiple predictor models did not show any significant effect of habitat on male or female volume or SSD. This may be partly explained by the correlation between habitat and flight dimorphism as volant forms (especially in males) are more commonly found higher up in the vegetation.

In the multiple predictor models including female volume, male volume and SSD as response variables, the variance explained by both phylogeny and predictors was higher ($R^2_{\text{full}}(\text{SSD})= 0.43$, $R^2_{\text{full}}(\text{female volume})= 0.82$, $R^2_{\text{full}}(\text{SSD})= 0.75$) than that explained by the predictors alone ($R^2_{\text{fixed}}(\text{SSD})= 0.40$, $R^2_{\text{fixed}}(\text{female volume})= 0.76$, $R^2_{\text{fixed}}(\text{SSD})= 0.71$, Table S2). But most of the variation was explained by fixed effects alone and not by phylogenetic relationships per se. This suggests that body size is quite phylogenetically labile.

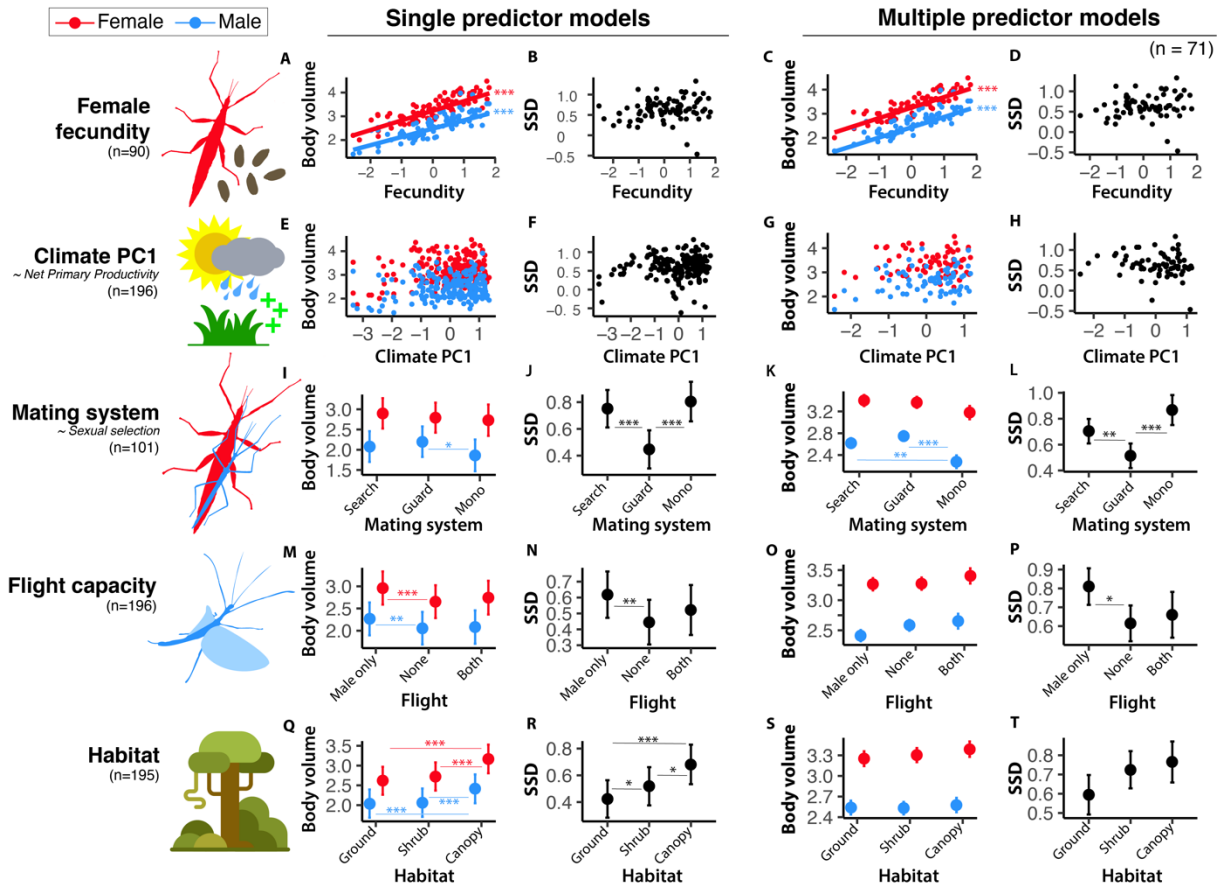


Figure 4: Ultimate predictors of male and female body size and sexual size dimorphism (SSD). Results of single predictor (left columns, sample size indicated below each predictor) and multiple predictor (right columns, $n=71$ species) PGLS models are presented. \log_{10} male body volume (blue, mm^3), \log_{10} female body volume (red, mm^3) and \log_{10} SSD (black) are shown versus \log_{10} female fecundity (i.e., lifetime reproductive output) (A-D), climate PC1 (E-H), mating system (I-L), flight dimorphism (M-P) and habitat (Q-T). PGLS regressions with significant slopes are shown for continuous predictors ($\beta \neq 0$, $p < 0.05$) (A-H). Estimated marginal means (\pm standard error) for each group are shown for categorical predictors (I-T). Significant pairwise group differences (assessed using the Holm correction) are shown. Asterisks indicate significant slopes/group differences (*: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$). Corresponding statistical analyses can be found in Table S2.

Selective drivers of variation in SShD. While single predictor models showed significant effects of mating system, flight dimorphism and habitat on male and female body elongation (PC1) (Table S3), multiple predictor models, accounting for the intercorrelation between predictors, only found a significant effect of mating system on PC1 in males and consequently on $\text{SShD}_{\text{elongation}}$ (Figure 5A-J, Table S3). Guarding males, whether they only temporarily guard females or guard them for their lifetime, tended to be stockier and less elongated than purely searching males, resulting in the most dimorphic overall body silhouettes being found in species with searching males and sedentary females.

In the multiple predictor models including female PC1, male PC1 and $SSD_{\text{elongation}}$ as response variables, the variance explained by both phylogeny and predictors was higher ($R^2_{\text{full}}(\text{SSD})=0.37$, $R^2_{\text{full}}(\text{female volume})=0.55$, $R^2_{\text{full}}(\text{SSD})=0.59$) than that explained by the predictors alone ($R^2_{\text{fixed}}(\text{SSD})=0.09$, $R^2_{\text{fixed}}(\text{female volume})=0.45$, $R^2_{\text{fixed}}(\text{SSD})=0.34$, Table S2). Here, the part of the variance explained by phylogeny was higher than for body size suggesting that body elongation is more phylogenetically conserved.

Similarly, single predictor models suggested a significant effect of mating system, flight dimorphism and habitat on male and female relative wing size (PC2)(Table S4), but multiple predictor models only found a significant effect of flight dimorphism (**Figure 5K-T**, Table S4). As expected, flying males and females had relatively larger wings than non-flying ones. Consequently, dimorphism in wing size was much greater in species where only males can fly. Finally, one of our multiple predictor models found a significant effect of climate on female PC2. It suggested that, after accounting for the other factors, females have relatively shorter wings in more tropical macrohabitats (i.e., with higher net primary productivity).

In the multiple predictor models including female PC2, male PC2 and SSD_{wings} as response variables, the variance explained by both phylogeny and predictors was sometimes lower ($R^2_{\text{full}}(\text{SSD})=0.68$, $R^2_{\text{full}}(\text{female volume})=0.86$, $R^2_{\text{full}}(\text{SSD})=0.82$) than that explained by the predictors alone ($R^2_{\text{fixed}}(\text{SSD})=0.63$, $R^2_{\text{fixed}}(\text{female volume})=0.93$, $R^2_{\text{fixed}}(\text{SSD})=0.87$, Table S2). This clearly suggests that relative wing size is not phylogenetically conserved and varies substantially between sister taxa.

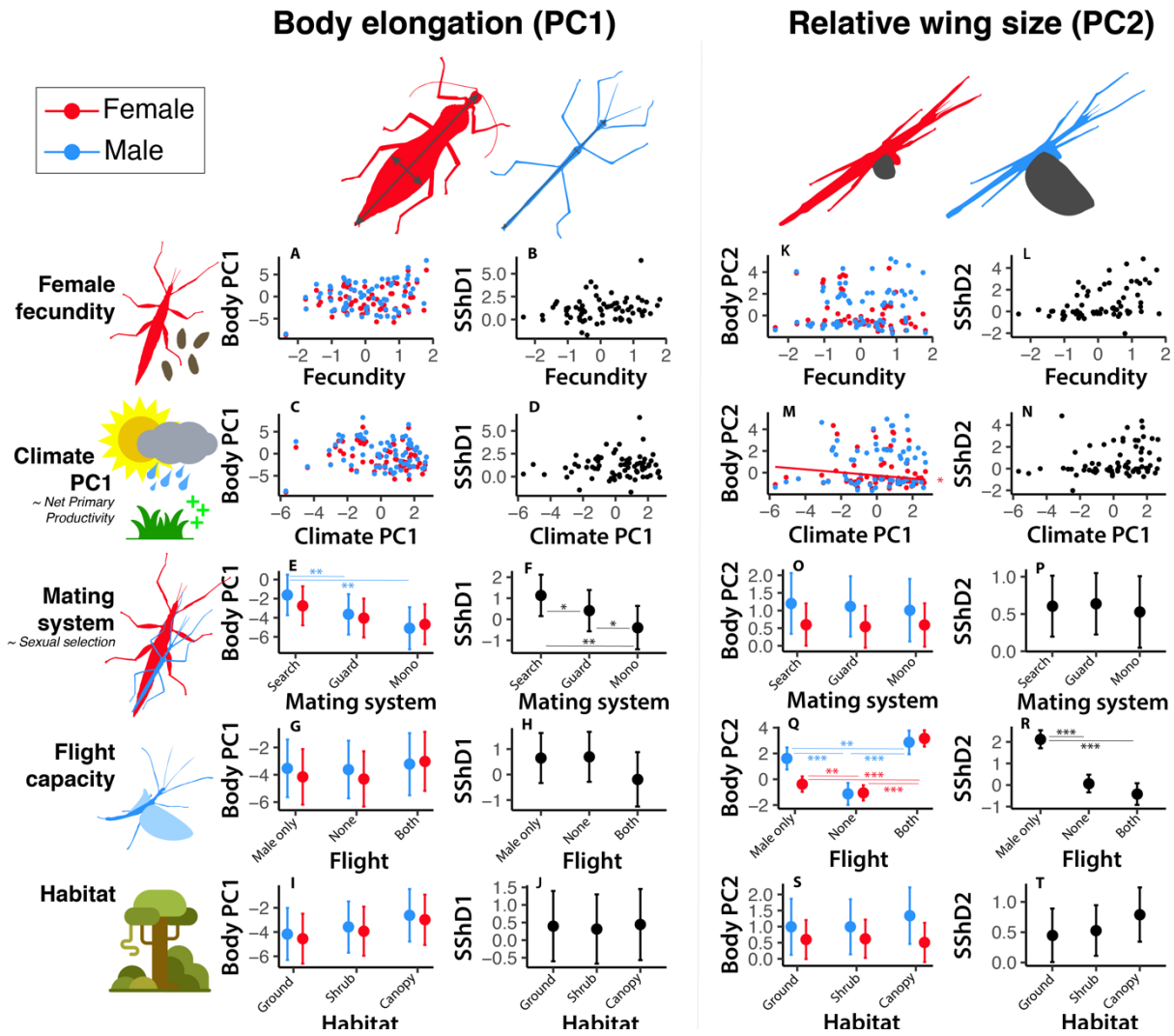


Figure 5: Ultimate predictors of male and female body elongation (PC1) and relative wing size (PC2), and associated sexual shape dimorphism (SShD_{elongation} and SShD_{wings}). Results of multiple predictor PGLS models (n= 69 species) are presented. Male (blue) and female (red) body PC1, SShD_{elongation} (A-J), male (blue) and female (red) body PC2 and SShD_{wings} (K-T) are shown versus log₁₀ female fecundity (i.e., lifetime reproductive output) (A-B, K-L), Climate PC1 (C-D, M-N), Mating system (E-F, E-P), Flight dimorphism (G-H, Q-R) and habitat (I-J, S-T). Significant slope coefficients are shown for continuous predictors ($\beta \neq 0$, $p < 0.05$) (A-D, K-N). Estimated marginal means (\pm standard error) for each group are shown for categorical predictors (E-J, O-T). Significant pairwise group differences (assessed using the Holm correction) are shown. Asterisks indicate significant slopes/group differences (*: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$). Corresponding statistical analyses can be found in Table S3-S4.

Discussion

The stick and leaf insects (Phasmatodea) display breathtaking interspecific variability in overall body size and in the extent and nature of sexual dimorphism. Here we tested three alternative hypotheses for the evolution of species differences in sexual dimorphism in both body size (SSD) and shape (SShD): (I) fecundity selection favoring larger and thicker females, (II) sexual selection favoring larger and stockier males, and (III) ecological variables differentially affecting male and female body size and shape. We show that positive selection on female size (likely stemming from fecundity selection) has been the primary historical driver of female-biased SSD in phasmids. However, in contrast with other studied female-biased taxa (e.g., amphibians (Monroe et al., 2015), fishes (Horne et al., 2020)), fecundity selection does not appear to have driven contemporary species differences in sexual dimorphism in these insects (rejecting hypothesis I). Instead, interspecific patterns of SSD and SShD likely resulted from selection acting on the males, either sexual selection (species differences in mating system, supporting hypothesis II) or ecological selection related to flight capability (supporting hypothesis III). Consequently we propose that (1) positive selection on female size has been the main historical driver of female-biased SSD in stick insects and that fecundity selection consistently favors larger female sizes across taxa but that (2) contemporary variation around the average female-biased SSD is mainly caused by selection acting on male size, whose magnitude and direction may be more variable across species.

Directional selection for larger female size is most often expected to result from the positive correlation between body size and fecundity (Fairbairn, 1997; Pincheira-Donoso and Hunt, 2017). Accordingly, we found a strong and positive correlation between female lifetime reproductive output and female size across stick insects (**Figure 4A**). However and contrary to our expectations, variation in female fecundity was not correlated with variation in SSD (**Figure 4B**), in contrast with other female-biased systems including fishes (Horne et al., 2020) and amphibians (Monroe et al., 2015). While we cannot predict the magnitude of SSD based on female fecundity, it is still likely a major factor maintaining large female sizes and driving SSD. Indeed, our data from 10 major phasmid clades showed that the magnitude of SSD of a clade increases as females become more size-variant than males (i.e., lower β , **Figure 2C-D**). Consistent with predictions from quantitative genetic theory (De Lisle and Rowe, 2013; Zeng, 1988), this pattern suggests that directional selection acting on female size played a major role in generating hypoallometry ($\beta < 1$) and female-biased SSD in stick insects. In

other words, this implies that positive selection on female size, rather than negative selection on male size, may have been the prominent driver of species size and female-biased SSD in phasmids. Therefore, the intensity of fecundity selection may not vary substantially across species and, while strong fecundity selection may have historically driven the evolution of much larger females, deviations observed today from the average female-biased SSD end up being mostly explained by variation in selective pressures acting on male size. Indeed, we start seeing males and females of comparable sizes (but never reaching male-biased SSD or monomorphism) only when pressures favoring larger male sizes become relatively intense, as exemplified by the weapon-bearing thorny devil stick insects. Unlike most other stick insect species, the relatively large male thorny devil stick insects (*Eurycantha* spp.) harbor extremely enlarged hindlegs that they use in violent combats with rival males over females or territories (Boisseau et al., 2020). Here, we found that female *E. calcarata* are only 1.34 times larger than males, relative to a phylogenetically-corrected average of 3.96 across all phasmids. We however recognize that the intensity of fecundity selection in a given taxa would be better quantified as the taxon-specific slope of the scaling relationship between female reproductive output and body size (Horne et al., 2020). Ultimately, our results suggest that although selection acting on male size best predicts variation in SSD in phasmids, fecundity selection acting on female size was likely the primary driver of the ubiquitous female-biased SSD seen in the group.

Our analyses pointed out to mating system (i.e., sexual selection intensity) and flight capacity as being the best predictors of variation in the extent of SSD and SShD across phasmids. In species where males do not typically stay with a female for extended periods of time, male reproductive success is likely determined by their searching and locomotor performance, which are predicted to be enhanced with smaller sizes (Boisseau et al., 2020; Boisseau et al., 2022; Herberstein et al., 2017; Kelly, 2020; Kelly et al., 2008). Consistently, males of such species are relatively smaller and slenderer (**Figure 4I-L & 5E-F**), even after accounting for their flight capacity. This suggests that even walking males may benefit from being smaller, maybe owing to a higher substrate attachment performance (Boisseau et al., 2022) or a lower energy need which would translate in a lower need to allocate time to foraging relative to mate searching (i.e., “Ghiselin–Reiss small-male hypothesis” (Blanckenhorn et al., 1995; Ghiselin, 1974; Reiss, 1989)). In contrast, species in which males search for and subsequently guard females have relatively larger and stockier males (**Figure 4I-L & 5E-F**). In such species, the post-copulatory mate guarding behavior of males suggests a high risk of sperm competition and a higher likelihood of one-on-one male encounters and physical contests (Boisseau et al., 2020; Sivinski, 1979).

In the prairie walkingstick (*Diaperomera velii*, Diaperomerinae) and other similar mate guarding systems, the largest and strongest males are almost invariably more successful at interrupting and displacing other guarding males and, when guarding themselves, are better able to resist takeovers (Bel-Venner and Venner, 2006; Borgia, 1980; Knox and Scott, 2006; Sivinski, 1979). Although male reproductive success is likely dependent on guarding ability in these species, in most cases, those males still extensively search for scattered females (Myers et al., 2015; Sivinski, 1977), and consequently their body morphology likely reflects the balance between selective pressures stemming from both searching and fighting (Kelly, 2014). The relative importance of these two factors may vary between species and also temporarily and spatially with the operational sex ratio and the relative densities of males and females (Kelly, 2015; Myers et al., 2015; Schöfl and Taborsky, 2002; Simmons et al., 2020). A more detailed quantification of the relative importance of searching and guarding would be desirable to better quantify variation in sexual selection intensity. Finally, monogamous males which stay on the back of a female for most of their lifetime exhibit much smaller sizes but a similar body shape compared to their respective females (**Figure 4I-L & 5E-F**). By staying in copula for very long periods of time, these males' body essentially acts as a sperm plug insuring exclusive paternity of the female's offspring. Thus, the reproductive interests of the two sexes should align and therefore selection should favor male traits —e.g., small body sizes— that help reduce the handicap of carrying its mate for the female. Dwarfed males may also benefit from lessening the disruption of camouflage while riding females and from keeping their maintenance cost to a minimum. Direct physical combat from intruding solitary rivals may be infrequent as solitary males approaching already coupled pairs have been observed to simply withdraw in such systems (e.g., *Anisomorpha buprestoides*, Pseudophasmatinae (Gunning, 1987)), suggesting a high owner's advantage and a low probability of takeovers potentially relieving selection for physical strength and larger sizes.

Differences in flight abilities between sexes was also a significant predictor of the magnitude of SSD and, unsurprisingly, $SShD_{wings}$. Indeed, after accounting for confounding factors such as mating system, flight appeared to favor smaller body sizes and relatively larger wings, leading to a lower wing loading. Wing loading —i.e., body mass divided by wing area— is often a good predictor of flight speed, wing beat frequency and therefore flight cost. Flapping animals with a high wing loading must fly fast to generate enough lift to stay in the air while animals with a low wing loading can fly slowly and inexpensively and benefit from an increased maneuverability (Biewener and Patek, 2018; Le Roy et al., 2019). Stick and leaf insects are slow fliers for which a higher wing loading translate into a reduced

climbing ability and maneuverability in flight, and therefore a reduced flight performance and capability (Boisseau et al., 2022; Zeng et al., 2020). Wings in stick insects can also be used in other contexts unrelated to flight including parachuting and predator deterrence through threatening displays of showy colorations. We could have expected flightless but winged species living in high vegetation layers to harbor relatively larger wings to slow down hypothetical falls, even if they cannot actively fly. But no other factor except flight capacity affected relative wing size.

Our data shows that female size is overall more variable than male size ($\beta < 1$) across Phasmatodea and that SSD increases with species size, thereby supporting the inverse of Rensch's rule. In addition, clade-specific allometries were never significantly hyperallometric and therefore never followed Rensch's rule (Table S1). These results are generally consistent with other groups with female-biased SSD, such as various insects (Blanckenhorn et al., 2007a) (including a previous study on stick insects using body length as body size metric (Sivinski, 1979; Sivinski and Dodson, 1992)), birds (Webb and Freckleton, 2007) and fishes (Horne et al., 2020).

SSD can arise from sex differences in post-embryonic development time and/or growth rate. The relative contributions of each of these proximate mechanisms has already received some attention, especially in insects (Blanckenhorn et al., 2007b; Teder, 2014; Teder et al., 2021) and fishes (Horne et al., 2020). We found that, in stick insects, species with the greatest relative difference in development time and in relative growth rate between sexes were the most size dimorphic. Females developed faster and for longer than males, both mechanisms conjointly generating female-biased SSD with nearly equal importance. Protandry –i.e., males maturing earlier than females– is ubiquitous in stick insects with only 3 out of 67 species with development time data exhibiting negligible differences in development duration. It is also the most common pattern of sex difference in development time in insects, and its evolution appears to be mainly driven as a by-product of the evolution of SSD (Teder et al., 2021). Nevertheless, it may still be advantageous for male stick insects to mature earlier. In taxa with discrete and nonoverlapping generations in seasonal environments (e.g., *Timema*, *Diapheromera*, *Bacillus*), males maturing early may have a mating advantage (Allen et al., 2011; Teder et al., 2021; Wiklund and Fagerström, 1977). For instance, early male maturation may be beneficial in systems where males guard pre-mature females (e.g., *Orthomeria kangi*, Aschiphasmatinae (Vallotto et al., 2016)) or where males pair with females for a very long time and, as discussed above, are unlikely to be overthrown by rivals once anchored (e.g., *Anisomorpha buprestoides*, Pseudophasmatinae (Gunning, 1987)). But such advantage should disappear in species with overlapping generations where mature females are continuously

available, as is the case for a lot of rainforest species. It should finally be noted that stick insects are generally sedentary and, despite some remarkable adaptations to dispersal (notably in the eggs (O’Hanlon et al., 2020)), may be at risk of inbreeding depression. Males maturing earlier, and sometimes dying before their sisters even reach maturity, may avoid this risk (Wiklund and Fagerström, 1977).

Conclusion

To date, the majority of large-scale comparative analyses investigating the extent to which SSD is driven by selection for increased size in males versus females, have consistently found correlation between the extent of SSD and the intensity of sexual selection acting on male size (Horne et al., 2020; Lislevand et al., 2009; Monroe et al., 2015; Serrano-Meneses and Székely, 2006), but, interestingly, have rarely offered support for the often-evoked fecundity selection hypothesis. However, our study exemplifies that, while variation in SSD may not be predictable based only on variation in female fecundity, it does not mean it is irrelevant to the evolution of SSD. Allometric patterns of SSD in stick insects suggest that positive selection on female size has been the most important historical driver of female-biased SSD in the group. Thus it suggests that fecundity may consistently favor larger female sizes across taxa but that variation around the average SSD is mainly caused by selection acting on male size, whose magnitude and direction may be more variable across species. By combining descriptions of allometric patterns of SSD and analyses of the effects of proximate and ultimate predictors on SSD and SShD, we hope our study can serve as a methodological template for future studies investigating the diversity of sexual dimorphism in other groups.

References

- Adams, D. C., Glynne, E. and Kaliontzopoulou, A.** (2020). Interspecific allometry for sexual shape dimorphism: macroevolution of multivariate sexual phenotypes with application to Rensch’s rule. *Evolution* **74**, 1908–1922.
- Allen, C. E., Zwaan, B. J. and Brakefield, P. M.** (2011). Evolution of sexual dimorphism in the Lepidoptera. *Annual Review of Entomology* **56**, 445–64.
- Andersson, M.** (1994). *Sexual Selection*. Princeton, NJ, USA: Princeton University Press.

- Bank, S. and Bradler, S.** (2022). A second view on the evolution of flight in stick and leaf insects (Phasmatodea). *BMC Ecology and Evolution* **22**, 1–17.
- Bedford, G. O.** (1978). Biology and ecology of the Phasmatodea. *Annual Review of Entomology* **23**, 125–149.
- Bel-Venner, M. C. and Venner, S.** (2006). Mate-guarding strategies and male competitive ability in an orb-weaving spider: results from a field study. *Animal Behaviour* **71**, 1315–1322.
- Berns, C. M. and Adams, D. C.** (2013). Becoming different but staying alike: patterns of sexual size and shape dimorphism in bills of hummingbirds. *Evolutionary Biology* **40**, 246–260.
- Biewener, A. A. and Patek, S. N.** (2018). *Animal Locomotion*. Oxford, United Kingdom: Oxford University Press.
- Blanckenhorn, W. U.** (2005). Behavioral causes and consequences of sexual size dimorphism. *Ethology* **111**, 977–1016.
- Blanckenhorn, W. U., Preziosi, R. F. and Fairbairn, D. J.** (1995). Time and energy constraints and the evolution of sexual size dimorphism - to eat or to mate? *Evolutionary Ecology* **9**, 369–381.
- Blanckenhorn, W. U., Meier, R. and Teder, T.** (2007a). Rensch's rule in insects : patterns among and within species. In *Sex, size, and gender roles: evolutionary studies of sexual size dimorphism*, pp. 60–70. Oxford, United Kingdom: Oxford University Press.
- Blanckenhorn, W. U., Dixon, A. F. G., Fairbairn, D. J., Foellmer, M. W., Gibert, P., Linde, K. van der, Meier, R., Nylin, S., Pitnick, S., Schoff, C., et al.** (2007b). Proximate causes of Rensch's rule: does sexual size dimorphism in arthropods result from sex differences in development time? *The American Naturalist* **169**, 245–257.
- Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G. and Emlen, D. J.** (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes* **181**, 104263.
- Boisseau, R. P., Büscher, T. H., Klawitter, L. J., Gorb, S. N., Emlen, D. J. and Tobalske, B. W.** (2022). Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect. *BMC Ecology and Evolution* **22**, 1–18.
- Borgia, G.** (1980). Sexual competition in *Scatophaga stercoraria*: size- and density-related changes in male ability to capture females. *Behaviour* **75**, 185–206.
- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.
- Brock, P. D. and Hasenpusch, J. W.** (2009). *The complete field guide to stick and leaf insects of Australia*. Collingwood, Victoria, Australia: CSIRO publishing.
- Brock, P. D., Büscher, T. H. and Baker, E.** (2021). Phasmida Species File Online. *Version 5.0/5.0*.
- Butler, M. A., Sawyer, S. A. and Losos, J. B.** (2007). Sexual dimorphism and adaptive radiation in *Anolis* lizards. *Nature* **447**, 202–205.
- Church, S. H., Donoughe, S., de Medeiros, B. A. S. and Extavour, C. G.** (2019a). Insect egg size and shape evolve with ecology but not developmental rate. *Nature* **571**, 58–62.
- Church, S. H., Donoughe, S., de Medeiros, B. A. S. and Extavour, C. G.** (2019b). A dataset of egg size and shape from more than 6,700 insect species. *Scientific Data* **6**, 1–11.
- Conle, O. V., Hennemann, F. H. and Dossey, A. T.** (2009). Survey of the color forms of the southern twostriped walkingstick (Phasmatodea: Areolatae: Pseudophasmatidae: Pseudophasmatinae: Anisomorphini), with notes on its

- range, habitats, and behaviors. *Annals of the Entomological Society of America* **102**, 210–232.
- Cox, R. M. and Calsbeek, R.** (2009). Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* **173**, 176–187.
- Cox, R. M., Skelly, S. L. and John-Alder, H. B.** (2003). A comparative test of adaptive hypotheses for sexual size dimorphism in lizards. *Evolution* **57**, 1653–1669.
- Cumming, R. T., Tirant, S. Le, Teemsma, S. N., Hennemann, F. H., Willemse, L. and Büscher, T. H.** (2020). Lost lovers linked at long last: elusive female *Nanophyllium* mystery solved after a century of being placed in a different genus (Phasmatodea, Phylliidae). *ZooKeys* **969**, 43–84.
- Dale, J., Dey, C. J., Delhey, K., Kempnaers, B. and Valcu, M.** (2015). The effects of life history and sexual selection on male and female plumage colouration. *Nature* **527**, 367–370.
- Darwin, C.** (1871). *The Descent of Man and Selection in Relation to Sex*. London, United Kingdom: John Murray.
- De Lisle, S. P. and Rowe, L.** (2013). Correlated evolution of allometry and sexual dimorphism across higher taxa. *The American Naturalist* **182**, 630–639.
- Emlen, D. J.** (2008). The Evolution of Animal Weapons. *Annual Review of Ecology, Evolution, and Systematics* **39**, 387–413.
- Emlen, D. J.** (2014). Reproductive contests and the evolution of extreme weaponry. In *The evolution of insect mating systems* (ed. Shuker, D. and Simmons, L.), pp. 92–105. Oxford, United Kingdom: Oxford University Press.
- Fairbairn, D. J.** (1997). Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. *Annual Review of Ecology and Systematics* **28**, 659–687.
- Fairbairn, D. J.** (2013). *Odd couples : extraordinary differences between the sexes in the animal kingdom*. Princeton, New Jersey, USA: Princeton University Press.
- Fairbairn, D. J., Blanckenhorn, W. U. and Székely, T.** (2007). *Sex, Size and Gender Roles*. Oxford, United Kingdom: Oxford University Press.
- Fick, S. E. and Hijmans, R. J.** (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315.
- Ghiselin, M.** (1974). *The economy of nature and the evolution of sex*. Berkeley, California, USA: University of California Press.
- Gunning, G. E.** (1987). Behavioral observations of the walking stick, *Anisomorpha buprestoides* (Phasmatodea: Phasmatidae). *The Florida Entomologist* **70**, 406–408.
- Hardy, I. C. W. and Briffa, M.** (2013). *Animal Contests*. Cambridge, United Kingdom: Cambridge University press.
- Herberstein, M. E., Painting, C. J. and Holwell, G. I.** (2017). Scramble competition polygyny in terrestrial arthropods. *Advances in the Study of Behavior* **49**, 237–295.
- Honěk, A.** (1993). Intraspecific variation in body size and fecundity in insects : a general relationship. *Oikos* **66**, 483–492.
- Horne, C. R., Hirst, A. G. and Atkinson, D.** (2020). Selection for increased male size predicts variation in sexual size dimorphism among fish species. *Proceedings of the Royal Society B* **287**, 20192640.
- Ives, A. R. and Li, D.** (2018). rr2: An R package to calculate R2s for regression models. *The Journal of Open Source Software* **3**, 1028.
- Kelly, C. D.** (2014). Sexual selection, phenotypic variation, and allometry in genitalic and non-genitalic traits in the

- sexually size-dimorphic stick insect *Micrarchus hystriculens*. *Biological Journal of the Linnean Society* **113**, 471–484.
- Kelly, C. D.** (2015). Male-biased sex ratios and plasticity in post-insemination behaviour in the New Zealand stick insect *Micrarchus hystriculens*. *Behaviour* **152**, 653–666.
- Kelly, C. D.** (2020). Sexual selection on size and shape in Japanese beetles (*Popillia japonica*). *Behavioral Ecology* **31**, 1073–1083.
- Kelly, C. D., Bussière, L. F. and Gwynne, D. T.** (2008). Sexual selection for male mobility in a giant insect with female-biased size dimorphism. *The American Naturalist* **172**, 417–423.
- Knox, T. T. and Scott, M. P.** (2006). Size, operational sex ratio, and mate-guarding success of the carrion beetle, *Necrophila americana*. *Behavioral Ecology* **17**, 88–96.
- Le Roy, C., Debat, V. and Llaurens, V.** (2019). Adaptive evolution of butterfly wing shape: from morphology to behaviour. *Biological Reviews* **94**, 1261–1281.
- Liaw, A. and Wiener, M.** (2002). Classification and regression by randomForest. *R News* **2**, 18–22.
- Lislevand, T., Figuerola, J. and Székely, T.** (2009). Evolution of sexual size dimorphism in grouse and allies (Aves: Phasianidae) in relation to mating competition, fecundity demands and resource division. *Journal of Evolutionary Biology* **22**, 1895–1905.
- Littleford-Colquhoun, B. L., Clemente, C., Thompson, G., Cristescu, R. H., Peterson, N., Strickland, K., Stuart-Fox, D. and Frere., C. H.** (2019). How sexual and natural selection shape sexual size dimorphism: evidence from multiple evolutionary scales. *Functional Ecology* **33**, 1446–1458.
- Lovich, J. E. and Gibbons, J. W.** (1992). A review of techniques for quantifying sexual size dimorphism. *Growth, Development and Aging* **56**, 269–281.
- Meiri, S. and Liang, T.** (2021). Rensch's rule—Definitions and statistics. *Global Ecology and Biogeography* **30**, 573–577.
- Monroe, M. J., South, S. H. and Alonzo, S. H.** (2015). The evolution of fecundity is associated with female body size but not female-biased sexual size dimorphism among frogs. *Journal of Evolutionary Biology* **28**, 1793–1803.
- Myers, S. S., Buckley, T. R. and Holwell, G. I.** (2015). Mate detection and seasonal variation in stick insect mating behaviour (Phamatodea: *Clitarchus booker*). *Behaviour* **152**, 1325–1348.
- O'Hanlon, J. C., Jones, B. R. and Bulbert, M. W.** (2020). The dynamic eggs of the Phasmatoidea and their apparent convergence with plants. *The Science of Nature* **107**, 1–12.
- Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. and Pearse, W.** (2018). caper: Comparative Analyses of Phylogenetics and Evolution in R.
- Paradis, E. and Schliep, K.** (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528.
- Pincheira-Donoso, D. and Hunt, J.** (2017). Fecundity selection theory: concepts and evidence. *Biological Reviews* **92**, 341–356.
- R Core Team** (2021). R Development Core Team. *R: A Language and Environment for Statistical Computing*.
- Reiss, M. J.** (1989). *The allometry of growth and reproduction*. Cambridge, United Kingdom: Cambridge University press.
- Rensch, B.** (1959). *Evolution above the species level*. New York, New York, USA: Columbia University Press.

- Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223.
- Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.
- Schöfl, G. and Taborsky, M. (2002). Prolonged tandem formation in firebugs (*Pyrrhocoris apterus*) serves mate-guarding. *Behavioral Ecology and Sociobiology* **52**, 426–433.
- Schwander, T. and Crespi, B. J. (2009). Multiple direct transitions from sexual reproduction to apomictic parthenogenesis in *Timema* stick insects. *Evolution* **63**, 84–103.
- Seow-Choen, F. (2005). *Phasmids of Peninsular Malaysia and Singapore*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F. (2016). *A Taxonomic Guide to the Stick Insects of Borneo: Including New Genera and Species*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F. (2017). *A Taxonomic Guide to the Stick Insects of Singapore*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Seow-Choen, F. (2018). *A Taxonomic Guide to the Stick Insects of Sumatra, Volume 1*. Kota Kinabalu, Sabah, Malaysia: Natural History Publications (Borneo).
- Serrano-Meneses, M.-A. and Székely, T. (2006). Sexual size dimorphism in seabirds: sexual selection, fecundity selection and differential niche-utilisation. *Oikos* **113**, 385–394.
- Shi, C., Shih, C., Chen, S. and Ren, D. (2019). Phasmatodea - Stick insects and leaf insects. In *Rhythms of Insect Evolution: Evidence from the Jurassic and Cretaceous in Northern China* (ed. Ren, D., Shih, C. K., Gao, T., Yao, Y., and Wang, Y.), pp. 165–173. Chichester, United Kingdom: John Wiley & Sons.
- Shine, R. (1988). The evolution of large body size in females: a critique of Darwin’s “fecundity advantage” model. *The American Naturalist* **131**, 124–131.
- Shine, R. (1989). Ecological Causes for the Evolution of Sexual Dimorphism: A Review of the Evidence. *The Quarterly Review of Biology* **64**, 419–461.
- Simmons, L. W., Parker, G. A. and Hosken, D. J. (2020). Evolutionary insight from a humble fly: sperm competition and the yellow dungfly. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 1–7.
- Simon, S., Letsch, H., Bank, S., Buckley, T. R., Donath, A., Liu, S., Machida, R., Meusemann, K., Misof, B., Podsiadlowski, L., et al. (2019). Old world and new world Phasmatodea: phylogenomics resolve the evolutionary history of stick and leaf insects. *Frontiers in Ecology and Evolution* **7**, 1–14.
- Sivinski, J. M. (1977). Factors affecting mating duration in the walkingstick *Diapheromera velii* (Walsh) (Phasmatodea: Heteronemiidae).
- Sivinski, J. (1979). Intrasexual aggression in the stick insects *Diapheromera veliei* and *D. covilleae* and sexual dimorphism in the Phasmatodea. *Psyche* **5**, 395–405.
- Sivinski, J. M. and Dodson, G. (1992). Sexual dimorphism in *Anastrepha suspensa* (Loew) and other tephritid fruit flies (Diptera: Tephritidae): possible roles of developmental rate, fecundity, and dispersal. *Journal of Insect Behavior* **5**, 491–492.
- Stephens, P. R. and Wiens, J. J. (2009). Evolution of sexual size dimorphisms in emydid turtles: Ecological dimorphism, rensch’s rule, and sympatric divergence. *Evolution* **63**, 910–925.

- Svenson, G. J., Brannoch, S. K., Rodrigues, H. M., O’Hanlon, J. C. and Wieland, F.** (2016). Selection for predation, not female fecundity, explains sexual size dimorphism in the orchid mantises. *Scientific Reports* **6**, 37753.
- Tammaru, T. and Esperk, T.** (2007). Growth allometry of immature insects: Larvae do not grow exponentially. *Functional Ecology* **21**, 1099–1105.
- Teder, T.** (2014). Sexual size dimorphism requires a corresponding sex difference in development time: A meta-analysis in insects. *Functional Ecology* **28**, 479–486.
- Teder, T., Kaasik, A., Taits, K. and Tammaru, T.** (2021). Why do males emerge before females? Sexual size dimorphism drives sexual bimaturism in insects. *Biological Reviews* **96**, 2461–2475.
- Tihelka, E., Cai, C., Giacomelli, M., Pisani, D. and Donoghue, P. C. J.** (2020). Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatodea) reveal a Permian–Triassic co-origination with insectivores. *Royal Society Open Science* **7**, 201689.
- Vallotto, D., Bresseel, J., Heitzmann, T. and Gottardo, M.** (2016). A black-and-red stick insect from the Philippines- observations on the external anatomy and natural history of a new species of *Orthomeria*. *ZooKeys* **559**, 35–57.
- van Velthuizen, H., Huddleston, B., Fischer, G., Salvatore, M., Ataman, E., Nachtergaele, F. O., Zanetti, M., Bloise, M., Gis, F., Antonicelli, A., et al.** (2007). *Mapping biophysical factors that influence agricultural production and rural vulnerability*. Rome, Italy: Food and Agriculture Organization of the United Nations and International Institute for Applied Systems Analysis.
- Webb, T. J. and Freckleton, R. P.** (2007). Only half right: species with female-biased sexual size dimorphism consistently break Rensch’s rule. *PLoS ONE* **2**, e897.
- Whiting, M. F., Bradler, S. and Maxwell, T.** (2003). Loss and recovery of wings in stick insects. *Nature* **421**, 264–267.
- Wiklund, C. and Fagerström, T.** (1977). Why Do Males Emerge before Females? A Hypothesis to Explain the Incidence of Protandry in Butterflies. *Oecologia* **31**, 153–158.
- Zeng, Z.-B.** (1988). Long-term correlated response, interpopulation covariation, and interspecific allometry. *Evolution* **42**, 363–374.
- Zeng, Y., Malley, C. O., Singhal, S., Rahim, F., Park, S., Chen, X. and Dudley, R.** (2020). A tale of winglets : evolution of flight morphology in stick insects. *Frontiers in Ecology and Evolution* **8**, 1–15.
- Zompro, O.** (1999). *Microphasma*, eine neue Stabschrecken-Gattung aus Sri Lanka (Phasmatodea: Pachymorphinae). *Entomologische Zeitschrift* **109**, 124–127.

Supplementary Information

Table S1: Clade-specific phylogenetically corrected mean SSD and allometric slope (calculated using a pRMA or PGLS) between \log_{10} male and female body volume, and associated 95% confidence intervals.

Clade	N _{species}	RMA slope	RMA 95% CI	PGLS slope	PGLS 95% CI	Mean SSD
Aschiphasmatinae	6	1.06	0.33-1.80	0.85	0.22-1.47	3.31
Diapheromerinae	14	0.73	0.45-1.00	0.62	0.37-0.86	5.45
Pseudophasmatinae + Agathemeridae	19	1.26	0.99-1.54	1.03	0.76-1.30	3.06
Heteropteryginae	17	0.80	0.66-0.94	0.82	0.72-0.93	3.77
African/Malagasy/ European clade	15	0.97	0.77-1.17	1.01	0.88-1.15	3.80
Phylliinae	6	1.06	0.63-1.49	1.05	0.89-1.20	4.51
Necroschiinae	18	1.13	0.75-1.51	0.92	0.61-1.24	4.54
Lonchodinae	23	1.18	0.94-1.43	1.11	0.93-1.30	3.77
Palophinae + Cladomorphinae	12	0.86	0.62-1.09	0.80	0.61 - 0.99	4.70
Pharnaciini + Clitumninae + Pachymorphinae	16	0.91	0.61-1.20	0.79	0.65-0.92	4.72
Stephanacridini + Lanceocercata	46	0.97	0.80-1.14	0.71	0.59-0.83	5.24

Table S2: Ultimate predictors of variation in female and male volume and SSD. The table presents results of single and multiple predictor PGLS models. The most likely value of Pagel's lambda (phylogenetic signal) is presented along with ANOVA outputs (using marginal sum of squares obtained by deleting a term from the model at a time) and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. The proportions of the variance explained by the full model R^2_{full} (phylogeny + fixed effects) and explained by fixed effects alone R^2_{fixed} (after accounting for phylogeny) are also reported.

# predictors	Response	Predictor	λ	R^2_{full}	R^2_{fixed}	$F_{df1,df2}$	P	Effect size \pm Standard Error (continuous) OR Post-hoc pairwise tests (Holm) (categorical)
1	log ₁₀ female volume	log₁₀ Female reproductive output	0.42	0.77	0.67	F_{1,88} = 185.8	<.0001	0.41 \pm 0.03
	log ₁₀ male volume	log₁₀ Female reproductive output	0.57	0.68	0.51	F_{1,88} = 94.9	<.0001	0.35 \pm 0.04
	log ₁₀ SSD	log ₁₀ Female reproductive output	0	0	0	F _{1,88} = 2.31	0.13	0.043 \pm 0.03
1	log ₁₀ female volume	Climate PC1	0.99	0.54	0.01	F _{1,194} = 3.38	0.07	0.065 \pm 0.036
	log ₁₀ male volume	Climate PC1	0.99	0.49	0	F _{1,194} = 1.17	0.28	0.037 \pm 0.034
	log ₁₀ SSD	Climate PC1	0.69	0.21	0	F _{1,194} = 0.85	0.36	0.023 \pm 0.025
1	log ₁₀ female volume	Mating system	0.95	0.48	0	F _{2,98} = 1.13	0.33	Search – Guard = 0.11 \pm 0.08, p = 0.62 Search – Mono = 0.17 \pm 0.14, p = 0.62 Guard – Mono = 0.06 \pm 0.14, p = 0.65
	log ₁₀ male volume	Mating system	0.96	0.44	0.05	F_{2,98} = 3.47	0.035	Search - Guard = -0.12 \pm 0.08, p = 0.24 Search – Mono = 0.22 \pm 0.14, p = 0.24 Guard – Mono = 0.34 \pm 0.14, p = 0.04
	log ₁₀ SSD	Mating system	0.60	0.38	0.29	F_{2,98} = 21.36	<.0001	Search - Guard = 0.30 \pm 0.05, p <.0001 Search – Mono = -0.05 \pm 0.09, p = 0.54 Guard – Mono = -0.36 \pm 0.09, p = 0.0002
1	log ₁₀ female volume	Flight dimorphism	0.98	0.58	0.09	F_{2,193} = 10.6	<.0001	Male only – None = 0.31 \pm 0.07, p <.0001 Male only – Both = 0.22 \pm 0.11, p = 0.08 None – Both = -0.09 \pm 0.09, p = 0.34
	log ₁₀ male volume	Flight dimorphism	0.98	0.52	0.04	F_{2,193} = 5.17	0.006	Male only – None = 0.21 \pm 0.07, p = 0.005 Male only – Both = 0.18 \pm 0.10, p = 0.15 None – Both = -0.02 \pm 0.091, p = 0.77
	log ₁₀ SSD	Flight dimorphism	0.52	0.21	0.05	F_{2,193} = 5.95	0.003	Male only – None = 0.17 \pm 0.05, p = 0.002 Male only – Both = 0.10 \pm 0.08, p = 0.50 None – Both = -0.08 \pm 0.08, p = 0.50
1	log ₁₀ female volume	Habitat	0.97	0.62	0.16	F_{2,192} = 19.0	<.0001	Ground - Shrub = -0.11 \pm 0.06, p = 0.06 Ground – Canopy = -0.55 \pm 0.09, p <.0001 Shrub – Canopy = -0.44 \pm 0.08, p <.0001
	log ₁₀ male volume	Habitat	0.98	0.54	0.08	F_{2,192} = 9.67	0.0001	Ground - Shrub = -0.02 \pm 0.06, p = 0.69 Ground – Canopy = -0.37 \pm 0.09, p = 0.0001 Shrub – Canopy = -0.35 \pm 0.08, p = 0.0001
	log ₁₀ SSD	Habitat	0.52	0.23	0.07	F_{2,192} = 8.62	0.0003	Ground - Shrub = -0.09 \pm 0.05, p = 0.05 Ground – Canopy = -0.26 \pm 0.06, p = 0.0002 Shrub – Canopy = -0.16 \pm 0.06, p = 0.01
5	log ₁₀ female volume	log₁₀ Female reproductive output	0.31	0.82	0.76	F_{1,62} = 137.3	<.0001	0.44 \pm 0.04
		Climate PC1				F _{1,62} = 1.76	0.19	-0.05 \pm 0.04
		Mating system				F _{2,62} = 2.69	0.08	Search – Guard = 0.03 \pm 0.06, p = 0.61 Search – Mono = 0.22 \pm 0.10, p = 0.08 Guard – Mono = 0.19 \pm 0.09, p = 0.08
		Flight dimorphism				F _{2,62} = 0.95	0.39	Male only – None = -0.01 \pm 0.08, p = 0.89 Male only – Both = -0.14 \pm 0.11, p = 0.58 None – Both = -0.13 \pm 0.10, p = 0.58
		Habitat				F _{2,62} = 1.07	0.35	Ground - Shrub = -0.05 \pm 0.07, p = 0.60 Ground – Canopy = -0.14 \pm 0.09, p = 0.45 Shrub – Canopy = -0.09 \pm 0.08, p = 0.59
5	log ₁₀ male volume	log₁₀ Female reproductive output	0.16	0.75	0.71	F_{1,62} = 104.5	<.0001	0.43 \pm 0.04
		Climate PC1				F _{1,62} = 0.08	0.78	-0.01 \pm 0.04
		Mating system				F_{2,62} = 10.9	0.0001	Search – Guard = -0.13 \pm 0.07, p = 0.09 Search – Mono = 0.34 \pm 0.11, p = 0.006 Guard – Mono = 0.47 \pm 0.10, p = 0.0001
		Flight dimorphism				F _{2,62} = 2.90	0.06	Male only – None = -0.17 \pm 0.09, p = 0.14 Male only – Both = -0.24 \pm 0.12, p = 0.14 None – Both = -0.07 \pm 0.12, p = 0.56
		Habitat				F _{2,62} = 0.12	0.88	Ground - Shrub = 0.009 \pm 0.09, p = 1 Ground – Canopy = -0.037 \pm 0.11, p = 1 Shrub – Canopy = -0.046 \pm 0.09, p = 1
5	log ₁₀ SSD	log ₁₀ Female reproductive output	0.26	0.43	0.40	F _{1,62} = 0.06	0.81	0.009 \pm 0.04
		Climate PC1				F _{1,62} = 2.03	0.16	-0.055 \pm 0.04
		Mating system				F_{2,62} = 9.83	0.0002	Search – Guard = 0.19 \pm 0.06, p = 0.009 Search – Mono = -0.16 \pm 0.10, p = 0.10 Guard – Mono = -0.35 \pm 0.09, p = 0.0007
		Flight dimorphism				F_{2,62} = 3.38	0.04	Male only – None = 0.20 \pm 0.08, p = 0.04 Male only – Both = 0.15 \pm 0.10, p = 0.31 None – Both = -0.04 \pm 0.10, p = 0.66
		Habitat				F _{2,62} = 1.99	0.15	Ground - Shrub = -0.13 \pm 0.07, p = 0.23 Ground – Canopy = -0.17 \pm 0.09, p = 0.23 Shrub – Canopy = -0.04 \pm 0.08, p = 0.62

Table S3: Ultimate predictors of variation in female and male body elongation (PC1) and SShD_{elongation}. The table presents results of single and multiple predictor PGLS models. The most likely value of Pagel's lambda (phylogenetic signal) is presented along with ANOVA outputs (using marginal sum of squares obtained by deleting a term from the model at a time) and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. The proportions of the variance explained by the full model R^2_{full} (phylogeny + fixed effects) and explained by fixed effects alone R^2_{fixed} (after accounting for phylogeny) are also reported.

# predictors	Response	Predictor	λ	R^2_{full}	R^2_{fixed}	$F_{df1,df2}$	P	Effect size \pm SE (continuous) OR Post-hoc pairwise tests (Holm) (categorical)
1	Female PC1	log ₁₀ Female reproductive output	0.95	0.47	0.01	$F_{1,85} = 2.14$	0.14	-0.36 \pm 0.25
	Male PC1	log ₁₀ Female reproductive output	0.96	0.50	0.03	$F_{1,85} = 3.34$	0.07	-0.49 \pm 0.27
	SShD _{elongation}	log ₁₀ Female reproductive output	0.93	0.33	0	$F_{1,85} = 0.20$	0.65	-0.06 \pm 0.12
1	Female PC1	Climate PC1	0.96	0.59	0	$F_{1,184} = 1.10$	0.29	-0.09 \pm 0.09
	Male PC1	Climate PC1	0.96	0.63	0	$F_{1,184} = 0.34$	0.56	-0.05 \pm 0.09
	SShD _{elongation}	Climate PC1	0.75	0.29	0	$F_{1,184} = 0.96$	0.33	0.04 \pm 0.04
1	Female PC1	Mating system	1	0.57	0.36	$F_{2,95} = 27.7$	<.0001	Search – Guard = 2.74 \pm 0.38, p <.0001 Search – Mono = 3.22 \pm 0.73, p = 0.0001 Guard – Mono = 0.48 \pm 0.69, p = 0.48
	Male PC1	Mating system	0.97	0.65	0.32	$F_{2,95} = 23.4$	<.0001	Search – Guard = 2.65 \pm 0.43, p <.0001 Search – Mono = 4.02 \pm 0.78, p <.0001 Guard – Mono = 1.37 \pm 0.73, p = 0.06
	SShD _{elongation}	Mating system	0.88	0.35	0.06	$F_{2,95} = 3.96$	0.02	Search – Guard = 0.25 \pm 0.22, p = 0.28 Search – Mono = 1.11 \pm 0.40, p = 0.02 Guard – Mono = 0.87 \pm 0.38, p = 0.05
1	Female PC1	Flight dimorphism	0.99	0.59	0.07	$F_{2,183} = 7.82$	0.0006	Male only – None = 0.40 \pm 0.41, p = 0.32 Male only – Both = -1.77 \pm 0.63, p = 0.01 None – Both = -2.17 \pm 0.55, p = 0.0003
	Male PC1	Flight dimorphism	0.97	0.64	0.03	$F_{2,183} = 3.56$	0.03	Male only – None = 0.70 \pm 0.43, p = 0.20 Male only – Both = -0.78 \pm 0.69, p = 0.26 None – Both = -1.49 \pm 0.63, p = 0.06
	SShD _{elongation}	Flight dimorphism	0.75	0.30	0.02	$F_{2,183} = 2.67$	0.07	Male only – None = 0.25 \pm 0.19, p = 0.24 Male only – Both = 0.72 \pm 0.32, p = 0.07 None – Both = 0.47 \pm 0.30, p = 0.24
1	Female PC1	Habitat	0.97	0.65	0.14	$F_{2,182} = 16.48$	<.0001	Ground – Shrub = -1.95 \pm 0.34, p <.0001 Ground – Canopy = -1.62 \pm 0.54, p = 0.006 Shrub – Canopy = 0.33 \pm 0.50, p = 0.52
	Male PC1	Habitat	0.96	0.68	0.12	$F_{2,182} = 13.40$	<.0001	Ground – Shrub = -1.84 \pm 0.36, p <.0001 Ground – Canopy = -1.77 \pm 0.56, p = 0.004 Shrub – Canopy = 0.07 \pm 0.53, p = 0.89
	SShD _{elongation}	Habitat	0.76	0.29	0	$F_{2,182} = 0.82$	0.44	Ground – Shrub = 0.03 \pm 0.18, p = 0.88 Ground – Canopy = -0.27 \pm 0.25, p = 0.63 Shrub – Canopy = -0.30 \pm 0.24, p = 0.63
5	Female PC1	log ₁₀ Female reproductive output	1	0.55	0.45	$F_{1,60} = 0.51$	0.48	-0.20 \pm 0.28
		Climate PC1				$F_{1,60} = 1.88$	0.18	-0.21 \pm 0.15
		Mating system				$F_{2,60} = 3.27$	0.04	Search – Guard = 1.28 \pm 0.57, p = 0.08 Search – Mono = 1.93 \pm 0.85, p = 0.08 Guard – Mono = 0.66 \pm 0.70, p = 0.36
		Flight dimorphism				$F_{2,60} = 1.07$	0.35	Male only – None = 0.15 \pm 0.71, p = 0.84 Male only – Both = -1.14 \pm 0.95, p = 0.48 None – Both = -1.28 \pm 0.89, p = 0.47
		Habitat				$F_{2,60} = 1.75$	0.18	Ground – Shrub = -0.62 \pm 0.57, p = 0.40 Ground – Canopy = -1.56 \pm 0.83, p = 0.20 Shrub – Canopy = -0.94 \pm 0.72, p = 0.40
5	Male PC1	log ₁₀ Female reproductive output	0.97	0.59	0.34	$F_{1,60} = 0.51$	0.48	-0.24 \pm 0.34
		Climate PC1				$F_{1,60} = 3.40$	0.07	-0.33 \pm 0.18
		Mating system				$F_{2,60} = 7.78$	0.001	Search – Guard = 2.04 \pm 0.64, p = 0.005 Search – Mono = 3.49 \pm 0.95, p = 0.015 Guard – Mono = 1.45 \pm 0.81, p = 0.08
		Flight dimorphism				$F_{2,60} = 0.07$	0.93	Male only – None = 0.09 \pm 0.79, p = 1 Male only – Both = -0.30 \pm 1.07, p = 1 None – Both = -0.39 \pm 1.02, p = 1
		Habitat				$F_{2,60} = 1.35$	0.27	Ground – Shrub = -0.57 \pm 0.65, p = 0.47 Ground – Canopy = -1.54 \pm 0.94, p = 0.32 Shrub – Canopy = -0.97 \pm 0.81, p = 0.47

5	SShD _{elongation}	log ₁₀ Female reproductive output	0.93	0.37	0.09	F _{1,60} = 0.32	0.57	-0.10 ± 0.17
		Climate PC1				F _{1,60} = 0.34	0.56	-0.05 ± 0.08
		Mating system				F_{2,60} = 5.55	0.006	Search – Guard = 0.72 ± 0.32, p = 0.05 Search – Mono = 1.53 ± 0.47, p = 0.006 Guard – Mono = 0.81 ± 0.41, p = 0.05
		Flight dimorphism				F _{2,60} = 1.61	0.21	Male only – None = -0.05 ± 0.39, p = 0.90 Male only – Both = 0.83 ± 0.53, p = 0.27 None – Both = 0.88 ± 0.51, p = 0.27
		Habitat				F _{2,60} = 0.07	0.94	Ground - Shrub = 0.08 ± 0.33, p = 1 Ground – Canopy = -0.04 ± 0.47, p = 1 Shrub – Canopy = -0.13 ± 0.40, p = 1

Table S4: Ultimate predictors of variation in female and male relative wing size (PC2) and SShD_{wings}. The table presents results of single and multiple predictor PGLS models. The most likely value of Pagel’s lambda (phylogenetic signal) is presented along with ANOVA outputs (using marginal sum of squares obtained by deleting a term from the model at a time) and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. The proportions of the variance explained by the full model R²_{full} (phylogeny + fixed effects) and explained by fixed effects alone R²_{fixed} (after accounting for phylogeny) are also reported.

# predictors	Response	Predictor	λ	R^2_{full}	R^2_{fixed}	$F_{df1,df2}$	P	Effect size \pm SE (continuous) Post-hoc pairwise tests (Holm) (categorical)
1	Female PC2	log ₁₀ Female reproductive output	0.69	0.24	0.02	$F_{1,85} = 2.65$	0.11	-0.26 \pm 0.16
	Male PC2	log ₁₀ Female reproductive output	0.78	0.34	0	$F_{1,85} = 1.28$	0.26	0.24 \pm 0.21
	SShD _{wings}	log₁₀ Female reproductive output	0.84	0.42	0.11	$F_{1,85} = 12.1$	0.0008	0.51 \pm 0.15
1	Female PC2	Climate PC1	0.83	0.37	0	$F_{1,184} = 0.66$	0.42	-0.04 \pm 0.05
	Male PC2	Climate PC1	0.88	0.47	0	$F_{1,184} = 0.85$	0.36	0.05 \pm 0.05
	SShD _{wings}	Climate PC1	1	0.31	0.01	$F_{1,184} = 3.39$	0.07	0.08 \pm 0.04
1	Female PC2	Mating system	0.69	0.22	0	$F_{2,95} = 1.04$	0.36	Search – Guard = 0.43 \pm 0.30, p = 0.46 Search – Mono = 0.21 \pm 0.52, p = 1 Guard – Mono = -0.22 \pm 0.51, p = 1
	Male PC2	Mating system	0.82	0.31	0.08	$F_{2,95} = 5.25$	0.007	Search – Guard = 1.01 \pm 0.35, p = 0.01 Search – Mono = 1.47 \pm 0.62, p = 0.04 Guard – Mono = 0.46 \pm 0.59, p = 0.44
	SShD _{wings}	Mating system	1	0.29	0.06	$F_{2,95} = 4.29$	0.02	Search – Guard = 0.49 \pm 0.21, p = 0.04 Search – Mono = 1.01 \pm 0.40, p = 0.04 Guard – Mono = 0.51 \pm 0.38, p = 0.18
1	Female PC2	Flight dimorphism	0.83	0.84	0.75	$F_{2,183} = 282.1$	<.0001	Male only – None = 0.59 \pm 0.11, p <.0001 Male only – Both = -3.49 \pm 0.18, p <.0001 None – Both = -4.08 \pm 0.17, p <.0001
	Male PC2	Flight dimorphism	0.95	0.88	0.78	$F_{2,183} = 330.4$	<.0001	Male only – None = 2.50 \pm 0.12, p <.0001 Male only – Both = -1.21 \pm 0.20, p <.0001 None – Both = -3.71 \pm 0.19, p <.0001
	SShD _{wings}	Flight dimorphism	0.84	0.70	0.54	$F_{2,183} = 110.2$	<.0001	Male only – None = 1.97 \pm 0.14, p <.0001 Male only – Both = 2.41 \pm 0.24, p <.0001 None – Both = 0.44 \pm 0.22, p = 0.05
1	Female PC2	Habitat	0.84	0.45	0.12	$F_{2,182} = 13.92$	<.0001	Ground – Shrub = -0.75 \pm 0.19, p = 0.0002 Ground – Canopy = -1.33 \pm 0.27, p <.0001 Shrub – Canopy = -0.58 \pm 0.26, p = 0.03
	Male PC2	Habitat	0.86	0.62	0.27	$F_{2,182} = 35.7$	<.0001	Ground – Shrub = -1.01 \pm 0.21, p <.0001 Ground – Canopy = -2.49 \pm 0.30, p <.0001 Shrub – Canopy = -1.49 \pm 0.28, p <.0001
	SShD _{wings}	Habitat	1	0.37	0.09	$F_{2,182} = 9.96$	0.0001	Ground – Shrub = -0.17 \pm 0.16, p = 0.29 Ground – Canopy = -1.21 \pm 0.27, p = 0.0001 Shrub – Canopy = -1.04 \pm 0.26, p = 0.0002
5	Female PC2	log ₁₀ Female reproductive output	1	0.86	0.93	$F_{1,60} = 0.08$	0.77	0.02 \pm 0.08
		Climate PC1				$F_{1,60} = 4.89$	0.03	-0.10 \pm 0.04
		Mating system				$F_{2,60} = 0.09$	0.91	Search – Guard = 0.06 \pm 0.17, p = 1 Search – Mono = 0.01 \pm 0.25, p = 1 Guard – Mono = -0.05 \pm 0.21, p = 1
		Flight dimorphism				$F_{2,60} = 132.1$	<.0001	Male only – None = 0.68 \pm 0.21, p = 0.002 Male only – Both = -3.54 \pm 0.28, p <.0001 None – Both = -4.22 \pm 0.26, p <.0001
		Habitat				$F_{2,60} = 0.15$	0.86	Ground – Shrub = -0.02 \pm 0.17, p = 1 Ground – Canopy = 0.09 \pm 0.24, p = 1 Shrub – Canopy = 0.11 \pm 0.21, p = 1
5	Male PC2	log ₁₀ Female reproductive output	0.99	0.82	0.87	$F_{1,60} = 0.59$	0.44	-0.09 \pm 0.12
		Climate PC1				$F_{1,60} = 1.05$	0.31	-0.07 \pm 0.07
		Mating system				$F_{2,60} = 0.14$	0.87	Search – Guard = 0.08 \pm 0.24, p = 1 Search – Mono = 0.19 \pm 0.36, p = 1 Guard – Mono = 0.11 \pm 0.30, p = 1
		Flight dimorphism				$F_{2,60} = 72.2$	<.0001	Male only – None = 2.75 \pm 0.31, p <.0001 Male only – Both = -1.25 \pm 0.41, p = 0.003 None – Both = -4.0 \pm 0.38, p <.0001
		Habitat				$F_{2,60} = 0.64$	0.53	Ground – Shrub = 0.001 \pm 0.24, p = 1 Ground – Canopy = -0.34 \pm 0.36, p = 0.81 Shrub – Canopy = -0.34 \pm 0.31, p = 0.81
5	SShD _{wings}	log ₁₀ Female reproductive output	0.38	0.68	0.63	$F_{1,60} = 0.17$	0.68	0.06 \pm 0.14
		Climate PC1				$F_{1,60} = 0.32$	0.58	0.03 \pm 0.06
		Mating system				$F_{2,60} = 0.05$	0.95	Search – Guard = -0.03 \pm 0.24, p = 1 Search – Mono = 0.08 \pm 0.36, p = 1 Guard – Mono = 0.11 \pm 0.34, p = 1
		Flight dimorphism				$F_{2,60} = 32.5$	<.0001	Male only – None = 2.05 \pm 0.29, p <.0001 Male only – Both = 2.54 \pm 0.39, p <.0001 None – Both = 0.49 \pm 0.39, p = 0.21
		Habitat				$F_{2,60} = 0.47$	0.63	Ground – Shrub = -0.08 \pm 0.28, p = 1 Ground – Canopy = -0.34 \pm 0.37, p = 1 Shrub – Canopy = -0.26 \pm 0.31, p = 1

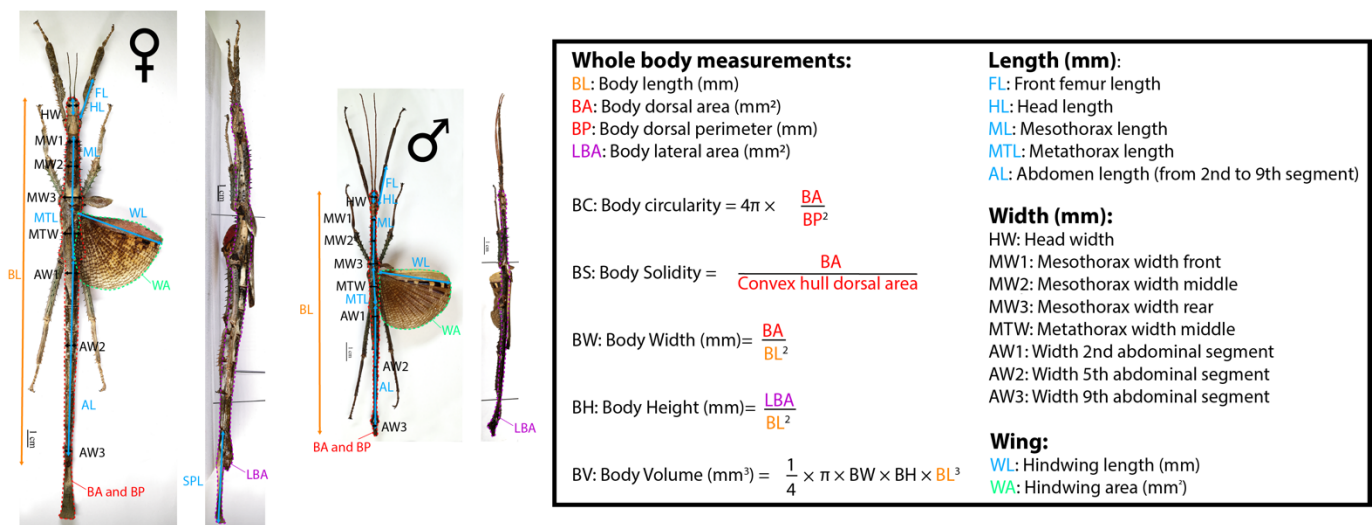


Figure S1: Measurements of phasmid specimens. **Left:** Adult female *Achrioptera punctipes cliquennoisi* Hennemann & Conle 2004 (Specimen MNHN-EO-PHAS127, project RECOLNAT (ANR-11-INBS-0004), photographs by Marion Depraetere, 2015, CC-BY-NC-ND) in dorsal and lateral view. **Middle:** Adult male *Achrioptera punctipes cliquennoisi* Hennemann & Conle 2004 (Specimen MNHN-EO-PHAS126, project RECOLNAT (ANR-11-INBS-0004), photographs by Marion Depraetere, 2015, CC-BY-NC-ND) in dorsal and lateral view. Body perimeter was only used to calculate body circularity and was not included in the principal component analysis described in Figure S2. Orange corresponds to body length (measured in absolute). Light blue corresponds to body part length measurements (measured relative to body length). Black corresponds to body part width measurements (measured relative to body length). Red corresponds to body dorsal perimeter and area (measured relative to body length). Purple corresponds to body lateral area (measured relative to body length). Green corresponds to wing area (measured relative to body length).

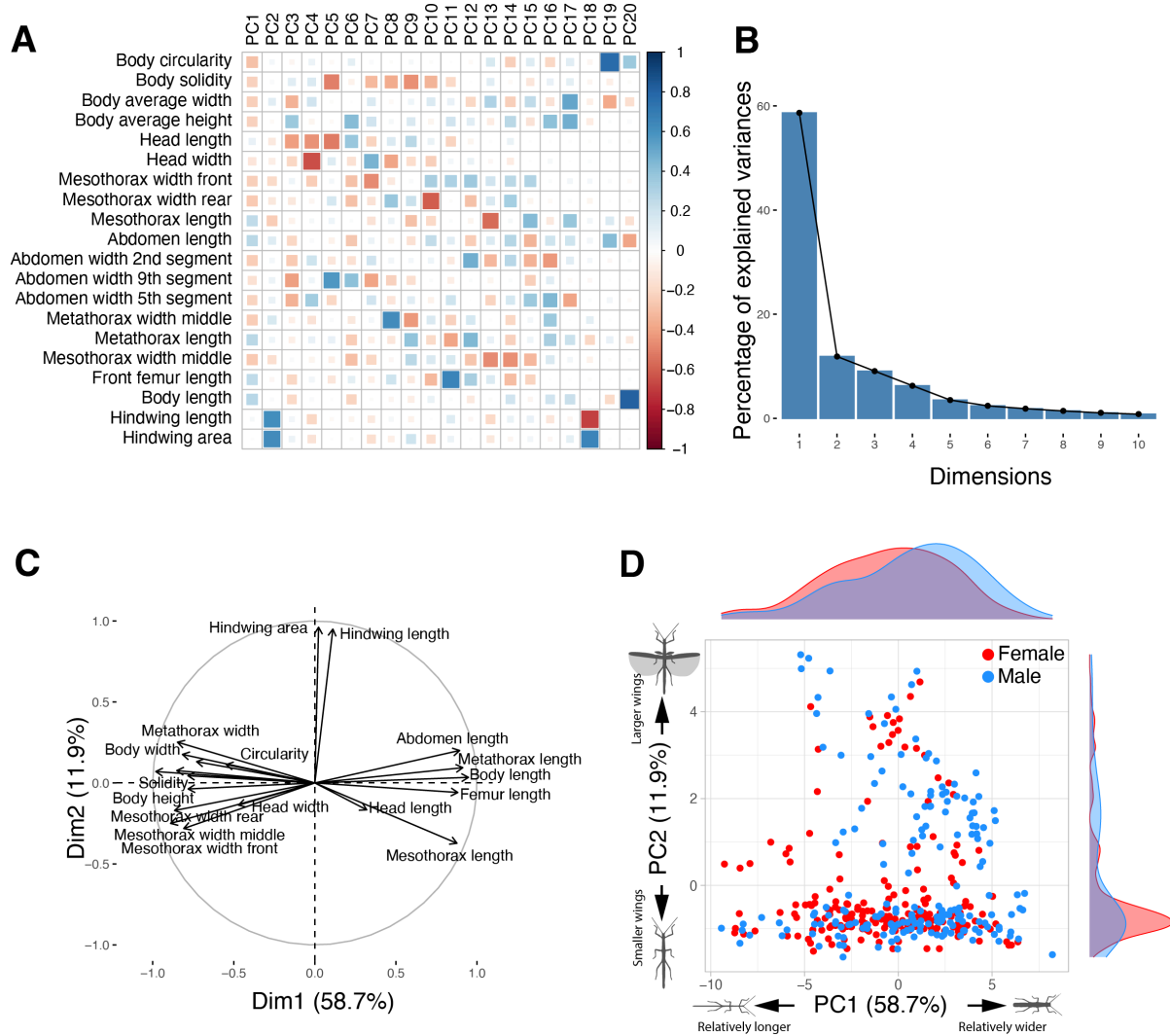


Figure S2: Principal component analysis (PCA) summarizing body shape of males and females. **A:** Correlation plot showing the extent and direction of the correlations between each principal component (PC) and the body shape variables included (i.e., residuals of PGLS regressions against body volume). **B:** Proportion of the total variance explained by each PC. **C:** PCA loading plot showing how strongly each characteristic influences PC1 and PC2. **D:** Morphospace showing PC2 against PC1 for males (blue) and females (red) of each species included.

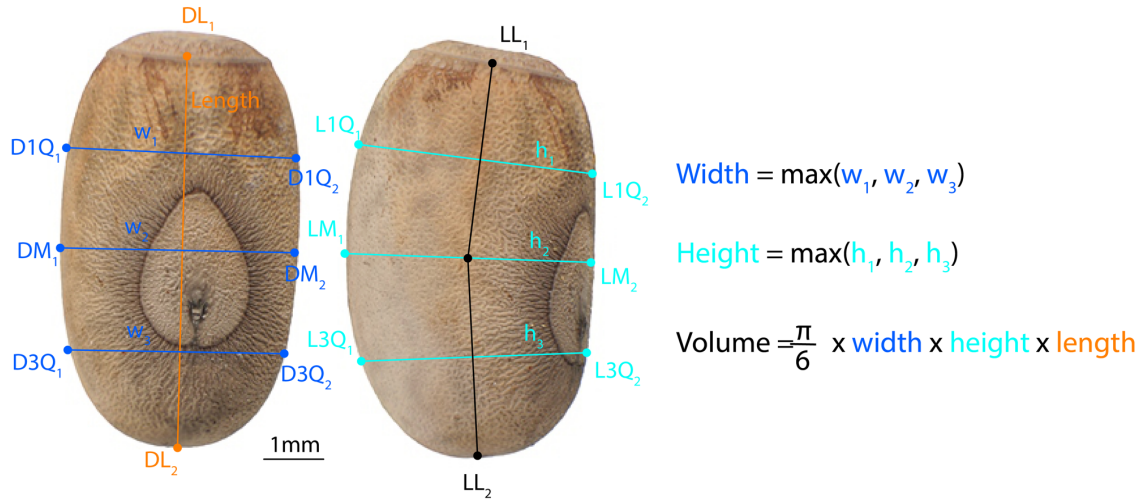


Figure S3: Guided landmark-based measurements of phasmid egg traits and calculation of volume, following the method of Church et al. (2019b). Photographs show the dorsal (left) and lateral (right) view of an egg of *Eurycantha insularis* Lucas, 1869 (collection F. Tetaert, phasmatodea.fr, used by permission). Egg length was measured between the middle of the base of the capitulum (DL₁) and the opposite pole of the egg where the curvature of the egg margin is steepest (DL₂) in dorsal view. From the segment [DL₁, DL₂], we built its perpendicular bisector [DM₁, DM₂] (= w₂). We then traced the two perpendicular bisectors of the segments connecting the midpoint of w₂ with DL₁ and DL₂, defining [D1Q₁, D1Q₂] (w₁) and [D3Q₁, D3Q₂] (w₃) respectively. The same approach was used on the lateral picture to define h₁, h₂ and h₃. Egg width, height and volume were then calculated as indicated on the right.

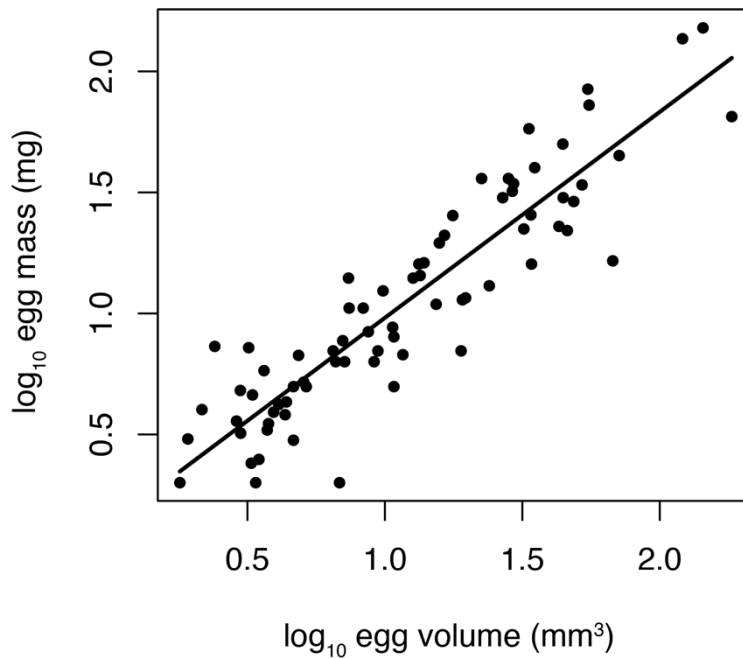


Figure S4: Correlation between log₁₀ egg mass and egg volume. Each point represents the average of a species. The corresponding linear regression is shown ($\beta=0.85 \pm 0.05$, $F_{1,73} = 354.6$, $p<.00001$, $R^2=0.83$).

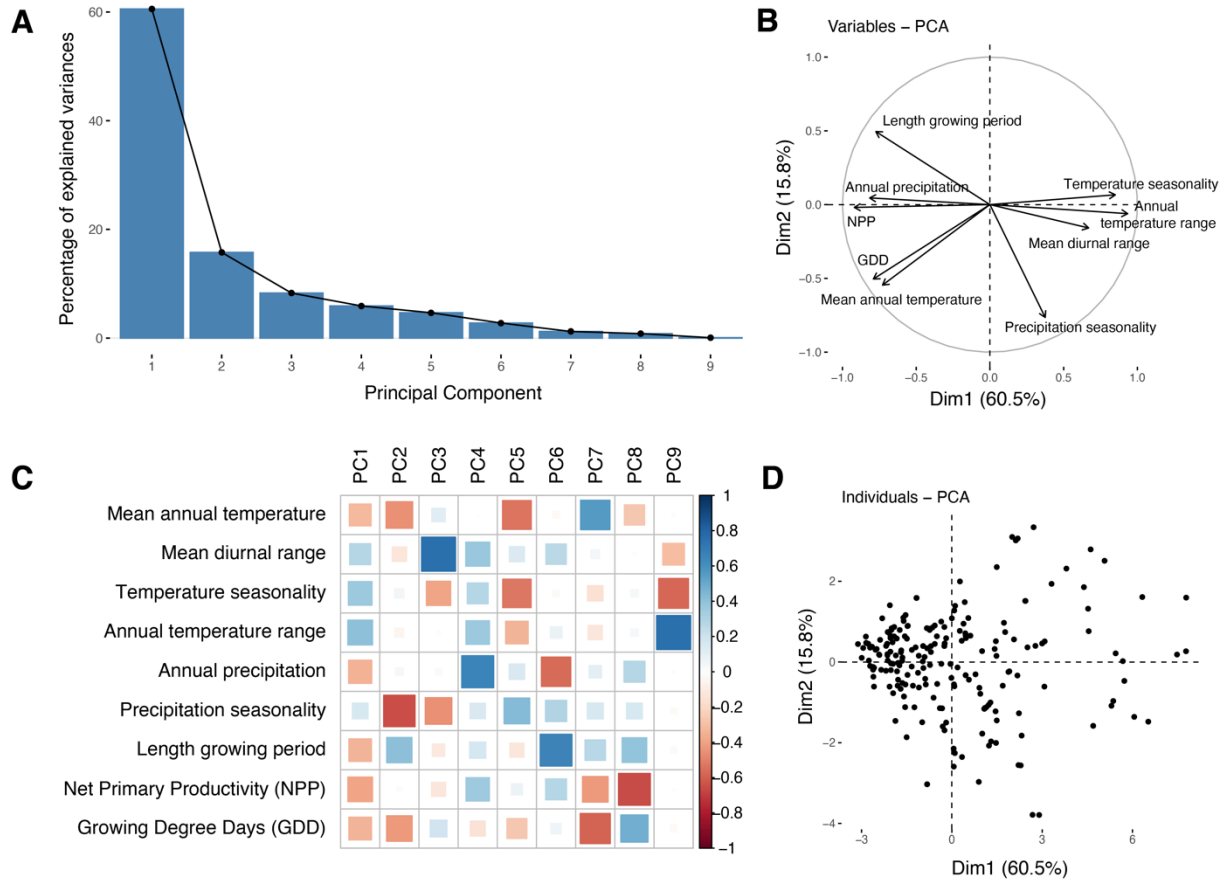


Figure S5: Principal component analysis (PCA) summarizing climatic conditions encountered by phasid species. **A:** Proportion of the total variance explained by each principal component (PC). **B:** PCA loading plot showing how strongly each climatic variable influences PC1 and PC2. **C:** Correlation plot showing the extent and direction of the correlations between each principal component (PC) and the climatic variables included. **D:** PC2 against PC1 for each species included.

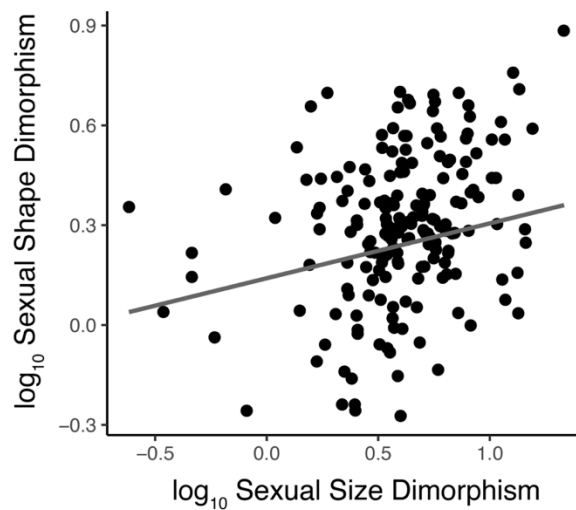


Figure S6 : Phylogenetically controlled correlation between overall SShD and SSD (PGLS regression: $\lambda = 0.59$, $F_{1,184} = 9.82$, $p = 0.002$, $r^2 = 0.05$).

CHAPTER 3

Physiology, mechanical constraints, and ecology drive the diversification of stick and leaf insect eggs

Physiology, mechanical constraints, and ecology drive the diversification of stick and leaf insect eggs

Romain P. Boisseau^{1,*} and H. Arthur Woods¹

¹Division of Biological Sciences, University of Montana, 32 Campus Dr, Missoula, MT 59812, United States of America

*Corresponding author: Romain P. Boisseau

Keywords: allometry | body size | tradeoff | life history evolution | oviposition | water loss

Abstract

The morphological diversity of insect eggs is astounding. Apart from the effect of some broad ecological drivers, however, this diversity is still largely unexplained. Here, we apply phylogenetic analyses to 200 species of stick and leaf insects (order Phasmatodea) coupled with physiological measurements of metabolic rate and water loss to evaluate a set of hypothesized drivers of egg diversification. We show that egg traits were most influenced by egg size and oviposition mode. Females that bury their eggs in the soil or glue them to substrates lay relatively larger but fewer eggs. Larger eggs exhibit longer developmental durations supported by disproportionately low metabolic rates. In parallel, egg burial and gluing is associated with egg elongation, most likely allowing an easier passage of these relatively large eggs through the oviduct in the slenderest species. In addition, flightless females display a higher reproductive output and consequently lay relatively larger eggs compared to flight-capable females. Surprisingly, local climatic conditions had only a weak effect on egg size and shape despite strong disparities between species distributed across various ecosystems worldwide. Overall, our results suggest that variation in life history, oviposition technique, mechanical constraints in the female body, and flight capacity, and to a lesser extent macroclimate, underlie the morphological diversification of stick insect eggs.

Introduction

Insect eggs share a common set of characteristics defined by their function: they are propagules -- finite packages of resources and information that support embryos as they grow from single cells into complex individuals that hatch into free-living juveniles (Hinton, 1981). Across taxa, eggs are recognizable as such because they share deep homologies. Egg diversity, by contrast, is more difficult to explain. Eggs diverge enormously in size, shape, composition, structure, physiology, and duration of development (Donoughe, 2022). What rules structure this diversity? How universal or taxon-specific are they? These questions have a long history of study in other taxa, especially birds (Stoddard et al., 2017), reptiles (Kratochvíl and Frynta, 2006), fish (Sargent et al., 1987), and marine invertebrates (Marshall and Keough, 2007; Moran and McAlister, 2009), though most studies have focused just on egg size. Below, we briefly outline three non-exclusive classes of hypotheses on the major causes of egg diversification in insects (Church et al., 2019a).

Size, life history, and pace of life. An organism's size can predict many other aspects of its life history, including length of development, potential fecundity, metabolic rate, and adult lifespan. In eggs from insect and non-insect taxa, for example, larger adults typically lay larger eggs (Blackburn, 1991; García-Barros and Munguira, 1997; Sargent et al., 1987), and large eggs typically develop more slowly than small eggs (Balon, 1984; García-Barros & Munguira, 1997; Gillooly & Dodson, 2000; Maino, Pirtle, & Kearney, 2017; Paine, 1985; Rahn & Ar, 1974; Steele & Steele, 1975; but for an alternative finding on insects, see Church et al., 2019b). In addition, like scaling relationships in other taxa and other life stages (Glazier, 2010; White and Seymour, 2005), metabolic rates of eggs scale hypoallometrically with egg mass (Maino and Kearney, 2014). Larger adults thus tend to lay larger eggs that develop over longer periods of time, supported by disproportionately low metabolic rates. In addition, egg size may trade off with egg number; females have access to finite resources, and even if they acquire nutrients in adulthood, they face the proximate problem of whether to construct smaller numbers of larger eggs or greater numbers of smaller eggs. Egg size-number tradeoffs in insects have been demonstrated both intraspecifically (Czesak and Fox, 2003; Gibbs et al., 2005; Koch and Meunier, 2014) and interspecifically (Berrigan, 1991; Blackburn, 1991; García-Barros, 2000).

Mechanical constraints on egg size and shape. Eggs may face strong geometric constraints on shape as a function of size (Church et al., 2019a; Stoddard et al., 2017). For example, because eggshell materials are costly (typically rich in protein), females laying more or larger eggs may minimize costs

by producing more spherical eggs (Blackburn, 1991; Kratochvíl and Frynta, 2006). On the other hand, larger eggs may need to have higher aspect ratios so that they can pass more easily through the reproductive canal during oviposition (Blackburn, 1991; Kratochvíl and Frynta, 2006). Larger eggs may also require higher ratios of surface area to volume, allowing them to obtain oxygen at rates high enough to support embryonic metabolism and to minimize diffusion distances between surface and central tissues (Woods and Hill, 2004). In addition, supporting higher metabolic rates may require higher-conductance eggshells, which in turn can lead to higher rates of water loss (Rahn and Ar, 1974; Woods, 2010).

Response to ecological circumstances. Egg traits may evolve in response to local suites of ecological conditions that they experience, including: (1) patterns of predation in relation to size, placement, or camouflage of the eggs, or as functions of the mechanical and chemical defenses that they deploy (Eisner et al., 2000; Guerra-Grenier, 2019); (2) whether females are flight-capable or flightless (Stoddard et al., 2017); and (3) local environmental conditions in their microsites, including whether eggs are aquatic or aerial (Church et al., 2019a).

Church et al. (2019a) recently examined a subset of these possibilities using data on egg and adult traits derived from literature on over 6,700 species in 526 families distributed across all extant orders, although not all traits were available for all species. Using phylogenetic analyses, they reached three main conclusions. First, geometric constraints on egg diversification were detectable but weak; larger eggs, for example, tended to have higher aspect ratios. Second, larger females indeed tended to lay larger eggs, but, contrary to other studies, those eggs did not have systematically longer developmental periods. Third, the most important drivers of changes in egg size and shape were the ecological circumstances of oviposition. In particular, evolutionary shifts from aerial oviposition (onto surfaces exposed to the atmosphere) to aqueous oviposition (into water or body fluids) led to systematic reductions in egg size and changes in shape.

Because of the large data set used, the analyses of Church et al. (2019) are uniquely comprehensive and powerful. Nevertheless, some conclusions – e.g., that larger eggs do not take longer to develop – are surprising and may reflect that some analyses were constrained by lack of data on focal traits to a relatively small number of species distributed across multiple major clades of insects. Such a situation may increase the probability that patterns in focal traits, even if present, are obscured by other evolved differences among major clades (i.e., orders). A complementary approach would be

to focus on individual clades with extensive sampling of species from across the group phylogeny. Here we present such an analysis for stick and leaf insects (order Phasmatodea), which appear to have originated in the early Cretaceous and diversified extensively following the K-T boundary (Robertson et al., 2018; Simon et al., 2019). Worldwide there are about 3,400 described species (Brock et al., 2021), for which we have compiled data on morphological and developmental traits of eggs and adults on nearly 210 species from approximately 27% of the ~520 described genera.

As masters of crypsis and masquerade, stick and leaf insects show remarkably diverse body morphologies, ranging from elongated stick-like silhouettes to robust, stocky forms (Chapter 1). They also produce remarkably diverse eggs, which span wide ranges in size, shape and structure. Eggs of some Heteropteryginae, for example, are among the largest produced by any insect, reaching 300 mg and 12 mm in length in *Haaniella echinata* (Hennemann et al., 2016). By contrast, eggs of *Spinoparapachymorpha spinosa* (Clitumninae) are only 2 mg and 1.6 mm in length (Seiler et al., 2000).

In Euphasmatodea, the evolution of a hardened egg chorion is a key innovation and may have contributed to the group's extreme diversification by opening up multiple routes of unusual dispersal (O'Hanlon et al., 2020; Robertson et al., 2018). Hard shells and their associated structures can withstand hazardous falls from the forest canopy, float for extended periods of time on sea water (hydrochory, Kobayashi et al., 2014), bear ant-attractive capitula (analogous to plants' eliasomes, myrmecochory, Stanton, Dias, & O'Hanlon, 2015), and even survive passage through the guts of birds (endozoochory, Suetsugu, Funaki, Takahashi, Ito, & Yokoyama, 2018). Accordingly, female phasmids employ a variety of egg-laying strategies, including passively dropping or actively flicking eggs to the forest floor, burying them in soil or other soft substrates, and gluing them to plant surfaces individually, in groups, or inside complex ootheca (Goldberg et al., 2015; Robertson et al., 2018). Since the late 1800s (Henneguy, 1890; Severin, 1910; Sharp, 1898), biologists have suggested that the morphological resemblance of phasmid eggs to seeds is a form of mimicry or masquerade. While this resemblance is indeed impressive, its ecological and evolutionary significance is still largely unknown.

Owing to their extreme diversity, relatively large size, and ease of breeding in captivity, stick insects are commonly kept as pets by amateur breeders or in classrooms, which has facilitated the compilation of a large data set on egg size, shape and physiology (Brock, 2000; Brock and Hasenpusch, 2009; Seiler et al., 2000). We leverage this comparative data set in a phylogenetic context, along with additional data on rates of metabolism and water loss by eggs of five of the species, to evaluate the

relative importance of the diversifying factors proposed above. Overall, we find strong positive associations among egg size, developmental duration, and adult characteristics such as female body size and abdominal size. Egg size trades off with egg number. In addition, repeated evolutionary transitions in style of oviposition, from the ancestral condition of egg flicking/dropping to more derived conditions of gluing eggs to plants or burying them underground, are associated with increases in egg size and increases in aspect ratio. Finally, egg size and shape appear to be associated with local climate conditions and female flight capacity.

Methods

Phylogenetic reconstruction. For the present study, we used a phylogeny from a companion study dealing with morphological convergence in adult female stick insects (Chapter 1), which includes a total of 314 phasmid taxa (9% of phasmid species diversity and 33% of currently recognized generic diversity) and one embiopteran species as outgroup (i.e., from the sister clade of Phasmatodea). The phylogenetic analyses were originally performed using genetic data from 3 nuclear (i.e., 18S rRNA, 28S rRNA and histone subunit 3) and 4 mitochondrial genes (i.e., 12S rRNA, 16S rRNA, cytochrome-c oxidase subunit I and II) and Bayesian inferences (BEAST2, v. 2.6.3, Bouckaert et al., 2014). The basal backbone topology of the tree was constrained to match that of transcriptome-based studies that could confidently infer the deep relationships between all the major clades of Phasmatodea (Simon et al., 2019; Tihelka et al., 2020). The resulting maximum clade credibility tree was overall strongly supported and congruent with previous phylogenetic reconstructions (Chapter 1, Bank & Bradler, 2022; Bradler, Cliquennois, & Buckley, 2015; Robertson et al., 2018).

Egg morphology. We collected high quality photographs of eggs in dorsal and lateral view for a total of 144 different species included in the phylogeny mainly from the egg picture database of F. Tetaert (Office pour les insectes et leur environnement OPIE, France, retrieved from phasmatodea.fr in August 2021) and from the published literature and other online databases (see Dataset S1 for details). We applied the guided landmark-based methodology of Church et al. (Church et al., 2019b) to quantify egg shape traits (Figure S1A). We used the R software (v4.1.1, R Core Team, 2021) and the package “raster” (Hijmans, 2021), to measure egg length (L) from the base of the operculum to the posterior end of the egg, and width and height along three different latitudinal lines at $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{3}{4}$ (respectively

w_1, w_2, w_3 and h_1, h_2, h_3) of the egg longitudinal axis (Figure S1A). Egg width (w) and height (h) were considered as the maximum values of w_1, w_2, w_3 and h_1, h_2, h_3 respectively. Egg volume was then estimated using the equation for the volume of an ellipsoid ($\frac{\pi}{6}lwh$). We verified the relevance of using this estimate of volume as our measure of egg size by regressing egg volume on egg mass (after \log_{10} -transformation) for species where that information could be collected ($n = 76$ species, see Dataset S1 for details). We found a strong correlation between the two ($\beta = 0.87 \pm 0.04$, $p < 0.0001$, $R^2 = 0.84$, Figure S2) and therefore chose to use egg volume as our measure of egg size as it was available for more taxa ($n = 144$ species). Egg surface area (SA) was calculated as the surface area of an ellipsoid using the approximate formula:

$$SA = 4\pi \sqrt[1.6]{\frac{L^{1.6}w^{1.6} + L^{1.6}h^{1.6} + w^{1.6}h^{1.6}}{3}}$$

To characterize egg shape variation while controlling for size, we performed a phylogenetic Principal Component Analysis (pPCA, R *function*: “R package”; *phyl.pca*: “phytools” (Revell, 2012)) including $L, w_1, w_2, w_3, h_1, h_2$ and h_3 , but original values were substituted with residuals calculated from a phylogenetically-corrected generalized least-squares (PGLS) regression against egg volume (*gls*: “nlme” and “ape” (Paradis and Schliep, 2019; Pinheiro et al., 2021), $\lambda = \text{'ML'}$).

Adult female morphology. To quantify adult female size, we used data on adult female volume originally collected in a previous study ($n = 212$ species) (Chapter 1), which estimated female body volume as an ellipsoid using the average body length, width and height of a species. Original measurements were obtained from digital images of live and dried specimens. In this study, we also used the species average width of the female’s ninth tergite (i.e., width the ninth abdominal dorsal plate) as it roughly corresponds to the location of the oviduct, from which eggs emerge during oviposition (Arai and Yago, 2015) and whose diameter may mechanically constrain egg morphology.

Ecological variables. Using primarily the dataset assembled by Robertson et al., (2018), we classified egg oviposition strategies into three categories: females drop or flick eggs to the ground from higher up in the local canopy, bury eggs into soil or other soft substrates, or glue eggs to substrate (including eggs encased in a ootheca) We then mapped oviposition strategies onto the phylogeny and ran an

ancestral state reconstruction using stochastic character mapping (*make.simmmap*: “phytools”). We calculated the transition matrix using MCMC and assumed different transition between all states (model “ARD”). We subsequently simulated and summarized 1,000 stochastic character maps to obtain posterior probabilities for each state at each node.

Female flight capacity was classified as either flight-capable (including gliding) or flightless (including parachuting) using the dataset compiled in a previous study (Chapter 2)(n = 214 species).

We also used the climate data gathered in the same dataset (Chapter 2, n = 214 species). These data were collected for each species based on current geographic distribution and corresponded to climate conditions experienced at the most central location of the range where the species have been collected (using the collection location of type specimens) or observed (using observations on iNaturalist, available from <https://www.inaturalist.org>. accessed July 2021). The dataset included annual mean temperature, mean diurnal range (i.e., mean of monthly (maximum - minimum temperature)), temperature seasonality (i.e., standard deviation \times 100), annual temperature range (i.e., maximum temperature of warmest month - minimum temperature of coldest month), annual precipitation, precipitation seasonality (i.e., coefficient of variation) from worldclim (Fick and Hijmans, 2017). The dataset also comprised the length of the growing period (number of days during a year when temperatures are above 5°C and precipitation exceeds half the potential evapotranspiration, available from <https://data.apps.fao.org/map/catalog> (van Velthuisen et al., 2007)), net primary production of biomass (grams of dry matter per m² per year; Climate Research Unit, Univ. of East Anglia, period 1976-2000, FAO Map Catalog, available from <https://data.apps.fao.org/map/catalog>), and total annual growing degree days (a measure of the annual amount of thermal energy available for plant and insect growth; Climate Research Unit, Univ. of East Anglia, available from <https://sage.nelson.wisc.edu/data-and-models>). We then ran a principal component analysis to summarize climatic variation across the geographic distribution of phasmids (*prcomp*: “stats”). We kept the first two axes of the PCA (climate PC1 and PC2, respectively, explaining 56% and 17% of the total variation) to quantify climatic variation between species (Figure S3). Climate PC1 reflected tropical versus temperate patterns differing in annual precipitation, annual temperature, and magnitude of diurnal and annual variation in temperature and consequently in net primary productivity (i.e., food availability for herbivorous insects). PC1 is overall high in tropical regions and low in more temperate and seasonal regions with restricted growing periods. Climate PC2 mostly represented variation in

precipitation seasonality with regions showing a strong shifts between dry and wet/monsoon seasons displaying a high PC2.

Life history variables. We used the lifetime fecundity data (i.e., average number of eggs laid by a female of a given species during its lifetime) compiled by Boisseau et al. (Chapter 2) from the literature and reports from amateur phasmid breeders (n= 100 species). Lifetime reproductive output was calculated as the product of egg volume and the fecundity (n= 96 species with sufficient data) and was used to quantify average female lifetime reproductive investment.

Finally, we collected data on average embryonic development duration (from oviposition to hatching) for 136 species from the published literature, amateur breeding guides and the phasmid breeder community (see references details for each species in Dataset S1). It should however be noted that such information is often quite approximate (± 2 weeks) and does not account for temperature. However, phasmids are typically bred at room temperature (20-24°C) and show higher variation in development time across species (from 1.25 to 11.5 months) relative to measurement error.

Phylogenetic analyses and hypothesis testing. We tested hypotheses on the drivers of egg diversification in phasmids following the framework outlined in the introduction and organizing the hypotheses into three classes. All analyses were performed in R (v4.1.1, R Core Team, 2021). Continuous variables were \log_{10} -transformed prior to the statistical analyses. PGLS models were run using the R packages “nlme” and “ape” (correlation= “corPagel”). Pagel’s lambda was estimated using maximum likelihood. Significance of the effect of each explanatory variable was evaluated using a type I (sequential) analysis of variance (ANOVA) or analysis of covariance (ANCOVA). The significance of interaction terms was systematically tested and non-significant interaction terms ($p < 0.05$) were removed from the final models to improve the estimation of intercepts and effect sizes. When appropriate, post hoc pairwise comparisons were ran to contrast the intercepts and/or slopes of different levels of a categorical explanatory variable (*emmeans* and *emtrends*: “emmeans” (Lenth, 2019)) using the Holm method to account for multiple testing.

First, we investigated hypotheses related to the effect of life history strategies on egg size. We tested whether egg size was dependent on adult female size, whether egg size traded off with egg number, and whether females varying in oviposition mode (i.e., parental care investment) varied in reproductive allocation strategy. We examined the scaling relationships between adult female size and egg volume and fecundity by running PGLS regressions including lifetime fecundity or egg volume as

the response variables and female body volume and oviposition mode as the explanatory variables. Then, to specifically test for a trade-off between egg number and size, we ran a PGLS model including fecundity as the response variable and female body and egg volume as explanatory variables, predicting a negative effect of egg volume on fecundity after accounting for female body size. We then compared the total reproductive output of females as a function of female body volume and oviposition mode, predicting that females investing more energy in parental care (by burying or gluing eggs to specific plants or substrates) would have relatively lower reproductive output. Finally, we tested whether larger eggs developed more slowly than smaller eggs by building a PGLS model including duration of embryogenesis as the response variable and egg volume and oviposition mode as predictors.

Next, we tested hypotheses related to mechanical and geometric constraints on egg size and shape. When egg size increases, eggs may become wider to save on costly eggshell material, leading to a hyperallometric relationship between egg width and length (i.e., slope greater than 1 on a log-log scale). Alternatively, as they get larger, eggs may become more elongated to be able to fit through a narrow opening during oviposition, to increase their surface area to volume ratio to obtain relatively more oxygen, and/or to minimize diffusion distances between the surface and the central tissues. In this case, we expect a hypoallometric relationship between egg width and egg length (i.e., slope lower than 1 on a log-log scale). We examined the scaling relationship between egg width and egg length using a PGLS regression including egg width as the response variable and egg length and oviposition mode as explanatory variables. Slopes were compared to isometry using 95% confidence intervals. We also tested the scaling relationship of egg surface area and egg size by including egg surface area as the response variable and egg volume and oviposition mode as explanatory variables in a PGLS model. Finally, we examined the relationship between egg width and the width of the female's ninth tergite (where the opening of the oviduct is located) by running a PGLS model including egg width as the response variable and female ninth tergite width and oviposition mode as the explanatory variables. To specifically investigate whether females with relatively narrower abdomen tips laid relatively more elongated eggs, we obtained measures of egg width relative to egg length, and of female ninth tergite width relative to female size by extracting the residuals of two PGLS regressions. One had egg width as a function of egg length and the other had ninth tergite width and female volume. Then we ran a PGLS model including residual egg width as a function of residual ninth tergite width and oviposition mode.

Finally, we examined hypotheses related to ecological circumstances. Oviposition mode is expected to affect egg size and shape as the conditions experienced by dropped, buried or glued eggs are likely to be very different in terms of exposure to predators, oxygen availability, and desiccation risk. Buried eggs may benefit from being more elongated so that females can insert them more easily into the substrate and so that they have relatively more surface area for gas exchange underground. Glued eggs, by contrast, are expected to have shapes that vary depending on the substrate to which they are glued (e.g., thin grass leaf, bark, broad leaf). Additionally, glued eggs are likely to be more exposed to predators and desiccation than eggs on or into soil and may therefore be selected to develop more rapidly. We also tested whether flight capacity affected egg size evolution, taking advantage of the numerous transitions between flight-capable and flightless species seen in stick insects (Bank and Bradler, 2022; Whiting et al., 2003). Because flight is costly (Biewener and Patek, 2018), we predicted that the reproductive investment by flying females would be relatively lower. We also looked at the effect of climate on egg size and shape. We hypothesized that drier climates would drive the evolution of rounder, larger eggs to increase total water stores and limit surface area to volume ratio and therefore relative rates of water loss. In order to test the effect of these ecological variables on egg size, we built a PGLS model including egg volume as the response variable and female body volume, fecundity, oviposition mode, female flight capacity, climate PC1 and 2 as predictors. Thus variation in fecundity and female size was accounted for to avoid confounding effects. Then we built similar PGLS models with either egg surface area, egg width, reproductive output and duration of embryogenesis as response variables and respectively egg volume, egg length, female body volume and egg volume as the first predictors to account for size effects. All models then included oviposition mode, female flight capacity, climate PC1 and 2 as predictors.

Study animals for physiological experiments. Sixteen freshly laid eggs were obtained from five unrelated phasmid species from culture stocks: *Eurycantha calcarata* (Lucas, 1869; subfamily Lonchodinae), *Extatosoma tiaratum* (Macleay, 1826; subfamily Phasmatinae: “Lanceocercata”), *Heteropteryx dilatata* (Parkinson, 1798; subfamily Heteropteryginae), *Medauroidea extradentata* (Brunner von Wattenwyl, 1907; subfamily Clitumninae) and *Phyllium philippinicum* (Hennemann, Conle, Gottardo & Bresseel, 2009; subfamily Phyllinae). All eggs were fertilized except *M. extradentata* eggs, which were obtained from a parthenogenetic culture. *P. philippinicum*, *M. extradentata* and *E. tiaratum* females typically drop or flick their eggs away during oviposition while *E. calcarata* and *H. dilatata* bury them (Robertson et al., 2018). Throughout embryonic development, eggs were kept individually in separate

wells of a plastic 24-well plate placed inside a plastic box whose bottom was filled with water to maintain nearly 100% relative humidity around the eggs. The eggs were kept at room temperature (21 - 23 °C) under a natural day/night cycle until hatching.

Egg metabolic rates. Egg metabolic rate was estimated as the rate of CO₂ production using flow-through respirometry. CO₂ concentrations were measured using a Licor LI-7000 CO₂/H₂O analyzer (Licor, Lincoln, NE, USA) in differential mode associated with a 16-chamber flow-through respirometry and data acquisition system (MAVEN-FT, Sable Systems International, North Las Vegas, NV, USA). The analyzer was calibrated frequently using CO₂-free air and 25 ppm CO₂ in N₂ (NorLab, Norco, Boise, ID, USA). Air flow rates were set to 25 ml.min⁻¹ (standard temperature and pressure) in the 16 experimental chambers. Gases circulated between the instruments in 3 mm inner-diameter plastic tubing (Bevaline-IV, Cole Parmer, Vernon Hills, IL, USA). Dry, CO₂-free air was first directed through a column containing Drierite and Ascarite to trap residual CO₂, then through the reference cell of the analyzer (referred to as cell A), which measured the fractional CO₂ concentration in incurrent air. From there, the air current was hydrated passing it through a bottle containing water (and several beads of Ascarite to keep CO₂ levels low). The hydrated stream was then directed through egg-containing chambers in the MAVEN and back into the measurement side of the analyzer (referred to as cell B), which measured the fractional CO₂ concentration in excurrent air. Data from the system was logged at 1 Hz using the MAVEN software.

Approximately every two weeks, each of the 16 eggs per species was weighed on an analytical balance (ME104TE/00, Mettler Toledo, Columbus, OH, USA) and subsequently placed individually in the 16 experimental chambers between the hours of 1600 and 2000 and left to run overnight until 0800-1000. Within each MAVEN cycle, air flow was first directed to the baseline channel for 5 min then to two experimental chambers sequentially for 20 min each and back to the baseline channel and so on through the remaining chambers.

Analysis of metabolic data. Data were extracted and analyzed using R (v4.1.1, R Core Team, 2021). The raw data files contained the relative concentration of CO₂ (parts per million, ppm) inside cell B compared with cell A (the reference cell) according to time (sampling frequency: 1 Hz). We converted raw measures (ppm) to molar rates of CO₂ production (MCO₂) using measured flow rates inside each chamber (which varied from 10 to 25 ml.min⁻¹) and the Ideal Gas Law:

$$\dot{M}CO_2 = \frac{P F \Delta[CO_2]}{RT}$$

where $\dot{M}CO_2$ is the rate of CO₂ production (mol.min⁻¹), $\Delta[CO_2]$ is the fractional concentration of CO₂ in cell B relative to cell A, F is the flow rate (L.min⁻¹), P is pressure (1 atm), R is the gas constant (0.08206 L.atm.K⁻¹.mol⁻¹) and T is the temperature measured by the MAVEn (K). Using custom scripts in R, we corrected the CO₂ traces for baseline drift by subtracting the mean baseline values that bracket each set of two experimental chambers.

In general, egg metabolic rate increased continuously during development (**Figure 4A**), which raises the questions of what representative values to use in other analyses. We approached this problem in two ways. First, for each egg, we estimated the metabolic rate halfway through the total developmental time from cubic regression splines and general additive models (GAM, *gam*: “mgcv”, k=5, bs= “cr”) fitted to the developmental trajectory of metabolism of each egg separately (Figure S4). Second, for each egg, we estimated the total CO₂ produced over its developmental period by integrating rates of CO₂ production at each time point estimated by the GAM and multiplying that value by the total development time. A measure of mean metabolic rate was then obtained by dividing total CO₂ produced by development time.

We tested the scaling relationships of mid-development metabolic rate and mean egg metabolic rate with egg mass by running linear mixed-effects regressions (*lmer*: “lme4”, Bates, Maechler, Bolker, & Walker, 2015) including species ID as a random effect. Variables were log₁₀-transformed prior to analyses. 95% confidence intervals were computed to compare the estimated slope to isometry (slope $\beta = 1$). We then investigated how the total CO₂ produced during embryogenesis (i.e., a proxy for the total energy for embryogenesis) varied with egg mass by running a similar linear mixed-effect model.

Egg rates of water loss. To assess rates of water loss, we held freshly laid eggs of the five species for approximately two months in a Tupperware container at 75% relative humidity (at room temperature 21 – 23 °C). Humidities were controlled by filling the bottom of each container with saturated solutions of sodium chloride in water (Winston and Bates, 1960), and eggs were held in 24-well plates positioned above the salt solutions. Depending on their size, eggs were weighed twice per week on either a Mettler Toledo ME54TE/00 (± 0.1 mg) or a Sartorius MC-5 microbalance (± 1 μ g). For each egg, we calculated water loss rate as the slope of the linear regression between egg mass and time. After two months, and before the eggs hatched, we transferred them to a drying oven (60 °C) for two weeks and

reweighed them dry. From these data and the initial fresh masses, we calculated initial water contents. We tested the scaling relationships of egg water loss rate and initial egg water content with initial egg mass by performing a linear mixed-effects regression after \log_{10} -transformation, including species ID as a random effect. The effect of oviposition mode was added in the model with water loss as it appeared to have a large effect. For controls, we used the eggs held in 100% relative humidity on which we measured metabolic rates (described above).

For each species, we calculated an average rate of mass loss at 100% relative humidity using a linear mixed-effects regression between egg mass and time, including egg ID as a random effect. For *E. calcarata* (the slope $\beta = -0.01 \pm 0.01$ mg.day⁻¹, Wald χ^2 test: $\chi^2 = 1.03$, df = 1, p = 0.3) and *M. extradentata* ($\beta = -0.0008 \pm 0.0007$ mg.day⁻¹, $\chi^2 = 1.54$, df = 1, p = 0.21), mass did not significantly decrease over time. In contrast, for *H. dilatata* ($\beta = -0.03 \pm 0.002$ mg.day⁻¹, $\chi^2 = 176.8$, df = 1, p < 0.0001), *P. philippinicum* ($\beta = -0.002 \pm 0.0006$ mg.day⁻¹, $\chi^2 = 9.38$, df = 1, p = 0.002) and *E. tiaratum* ($\beta = -0.005 \pm 0.0007$ mg.day⁻¹, $\chi^2 = 43.9$, df = 1, p < 0.0001), we found a significant decrease of mass over time. This decrease may be attributed to loss of organic matter through catabolism or to water loss during flow-through respirometry. However, mass loss rates estimated through the same method for eggs held at 75% relative humidity, showed much higher values (*H. dilatata*: $\beta = -0.11 \pm 0.002$ mg.day⁻¹, *P. philippinicum*: $\beta = -0.008 \pm 0.001$ mg.day⁻¹, *E. tiaratum*: $\beta = -0.013 \pm 0.001$ mg.day⁻¹), suggesting that most of the mass lost by eggs at 75% relative humidity were due to water loss.

In preliminary experiments on *E. calcarata*, we held freshly laid eggs at 100, 75, or 0% RH (held over water, saturated solution of sodium chloride, or Drierite, respectively, n = 20 per treatment) and assessed rates of water loss and of survival. No eggs held at 0% RH and only 5% of those held at 75% RH hatched. Of those held at 100% RH, 95% hatched. These data suggest that eggs of the species we measured need consistently high humidities to avoid fatal amounts of water loss. We used 75% RH in our fuller set of measurements in order to assess relative permeabilities to water at humidities low enough to induce measurable water loss.

Results

Phylogenetic analysis of transitions in oviposition style. Our ancestral state reconstruction unambiguously indicated that the ancestor lineages of Phasmatodea and Euphasmatodea dropped or flicked their eggs to the ground, like most extant phasmid species worldwide (**Figure 1**). The associated transition matrix identified asymmetric transition rates between oviposition modes and suggested that transitions from dropping/flicking to burying (at least 16 times) or gluing (at least 7 times), and from gluing to burying (at least 3 times) were the most frequent. Only one reverse transition back to dropping/flicking from burying was recovered, although equivocally, in the New Zealand/New Caledonia clade (belonging to Lanceocercata, **Figure 1**) highlighting that back transitions from more derived styles (i.e., burying or gluing) were very unlikely. Morphologically- and ecologically-diverse clades that colonized diverse habitats often exhibited multiple such transitions and diverse oviposition styles (e.g., Pseudophasmatinae, African/Malagasy clade, Necrosiinae, Lanceocercata) unlike morphologically- and ecologically- homogeneous clades (e.g., Heteropterygidae, Phylliidae) (Chapter 1).

Size, life history, and pace of life. Across phasmids, larger females laid more and larger eggs (**Table 1, Figure 2A-B**). After accounting for adult female body volume, we found a negative correlation (i.e., a trade-off) between lifetime egg number (fecundity) and egg volume (**Table 1, Figure 2C**). Oviposition mode significantly affected the size-number tradeoffs. Females burying or gluing their eggs in specific locations, i.e., those that invest relatively more in parental care, appeared to lay fewer but relatively larger eggs relative to females that drop or flick their eggs away. Overall, lifetime reproductive output (lifetime fecundity \times average egg volume) was strongly correlated with adult female volume but did not significantly differ between oviposition modes (**Table 1, Figure 2D**). Larger eggs developed slower than smaller eggs and, for a given egg volume, glued eggs developed faster than buried or dropped eggs (**Table 1, Figure 2E**).

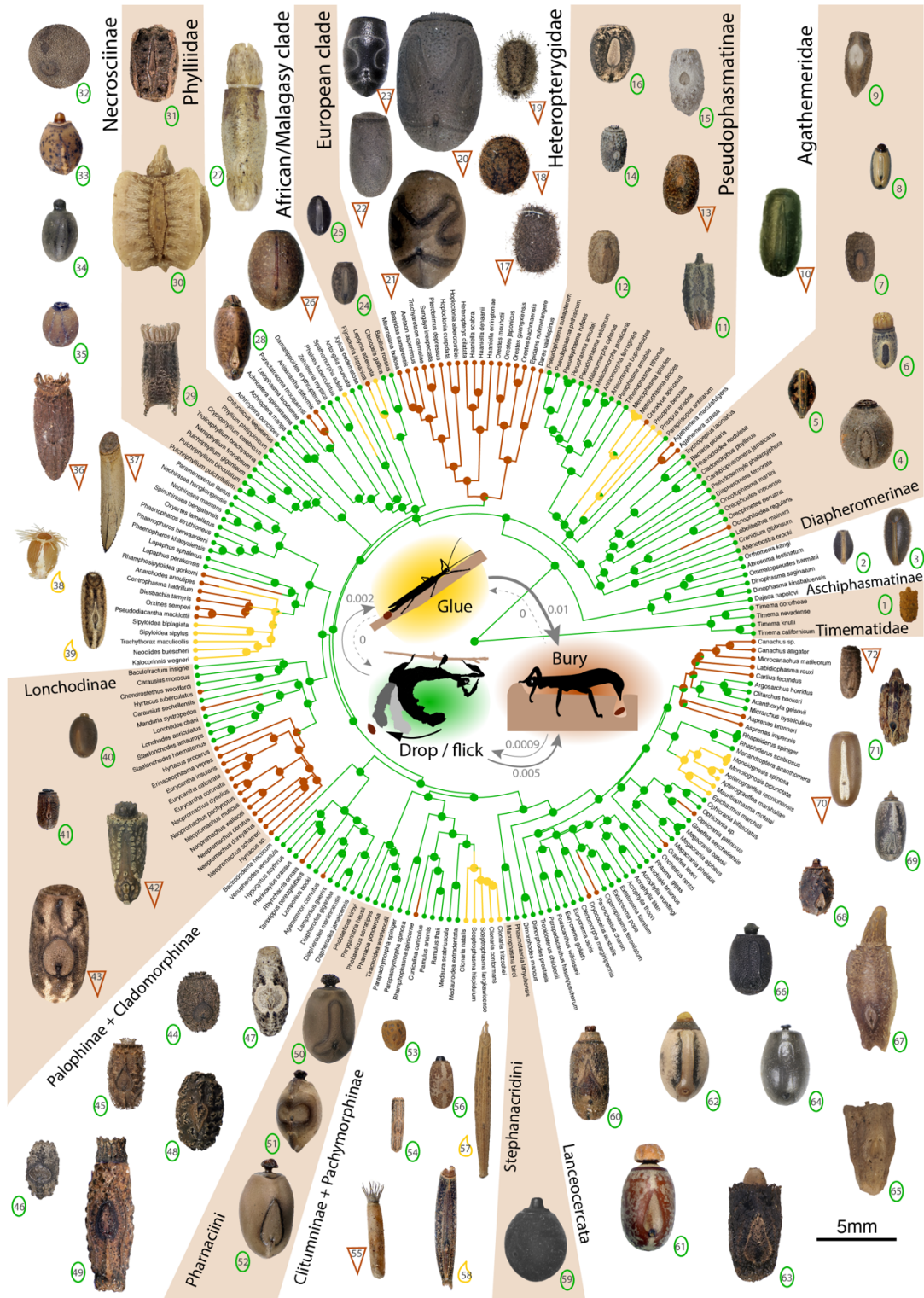


Figure 1: Egg morphological diversity and ancestral state reconstruction of oviposition mode in Phasmatodea. The ancestral state reconstruction used stochastic character mapping and a transition matrix (inset) estimated by maximum likelihood. Scaled egg pictures in dorsal view correspond to the species listed in table S1. Dropped or flicked eggs are represented by a green oval, buried eggs by a brown triangle and glued eggs by a yellow droplet.

Table 1: Life history correlates of egg size. The table presents the results of phylogenetic generalized least square models. The most likely value of Pagel’s lambda (phylogenetic signal) is presented along with type-I (sequential) ANOVA outputs and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. Significant effects are bolded ($p < 0.05$).

Response (n species)	Predictor	λ	$F_{df1,df2}$	P	Effect size \pm Standard Error (continuous)
					OR Post-hoc pairwise tests (Holm) (categorical)
\log_{10} Lifetime fecundity (= egg number) (n = 96)	\log_{10} Adult female body volume	0	$F_{1,92} = 44.27$	<.0001	0.30 ± 0.04
	Oviposition mode		$F_{2,92} = 23.29$	<.0001	Drop – bury = 0.35 ± 0.05 , $p < .0001$ Drop – glue = 0.29 ± 0.11 , $p = 0.017$ Bury – glue = -0.06 ± 0.12 , $p = 0.61$
\log_{10} Egg volume (n = 143)	\log_{10} Adult female body volume	0.57	$F_{1,139} = 192.2$	<.0001	0.67 ± 0.05
	Oviposition mode		$F_{2,139} = 15.3$	<.0001	Drop – bury = -0.28 ± 0.05 , $p < .0001$ Drop – glue = 0.23 ± 0.08 , $p = 0.005$ Bury – glue = 0.05 ± 0.08 , $p = 0.55$
\log_{10} lifetime fecundity (n = 96)	\log_{10} Adult female body volume	0.02	$F_{1,93} = 40.1$	<.0001	0.64 ± 0.07
	\log_{10} Egg volume		$F_{1,93} = 33.9$	<.0001	-0.53 ± 0.09
\log_{10} Reproductive output (egg number x volume) (n = 96)	\log_{10} Adult female body volume	0.12	$F_{1,92} = 290.9$	<.0001	0.91 ± 0.06
	Oviposition mode		$F_{2,92} = 2.56$	0.08	Drop – bury = 0.04 ± 0.06 , $p = 0.50$ Drop – glue = 0.28 ± 0.13 , $p = 0.08$ Bury – glue = 0.24 ± 0.14 , $p = 0.15$
\log_{10} Duration of embryogenesis (n = 121)	\log_{10} Egg volume	1	$F_{1,117} = 45.3$	<.0001	0.19 ± 0.03
	Oviposition mode		$F_{2,117} = 6.53$	0.002	Drop – bury = -0.04 ± 0.04 , $p = 0.26$ Drop – glue = 0.19 ± 0.06 , $p = 0.008$ Bury – glue = 0.23 ± 0.06 , $p = 0.001$

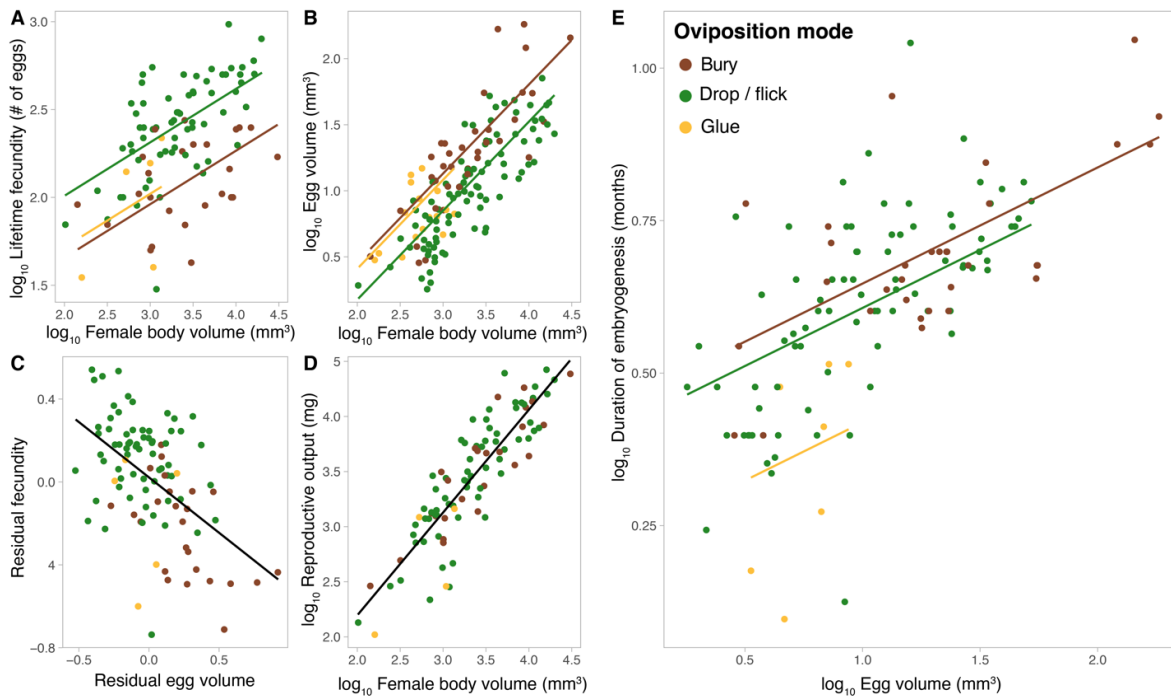


Figure 2: Life history predictors of egg size. Lifetime fecundity (A) and egg volume (B) compared to adult female body volume. C, Relationship between relative lifetime fecundity and relative egg volume after accounting for adult female body volume. D, Lifetime reproductive output (fecundity x egg volume) as a function of adult female body volume. E, Duration of embryogenesis, from egg laying to hatching, compared to egg volume. Colors represent different oviposition modes (see legend in E). Phylogenetic least square regressions are represented and correspond to the analyses reported in Table 1.

Mechanical constraints on egg size and shape. A phylogenetic principal component analysis revealed that most of the variation in egg capsule shape pertained to variation in aspect ratio (i.e., relative elongation): PC1 (66% of the total variation) reflected variation in elongation, PC2 (21% of the total variation) reflected variation in axis of elongation (anterior-posterior or dorsal-ventral) (**Figure 3A**, S1B-C). Most phasmid eggs were clustered in the center of the resulting egg morphospace, exhibiting a generic barrel-shaped egg capsule (**Figure 3A**; e.g., *Anisomorpha buprestoides*, egg #12; *Diapherodes martinicensis*, egg #48; *Acrophylla wuelfingi*, egg #66 in **Figure 1**). In contrast, other species had very elongated shapes (i.e., high PC1; e.g., *Sceptrophasma hispidulum*, egg #58 in **Figure 1**) or flattened lentil-shaped capsules (i.e., low PC2; e.g., *Oreophoetes peruana*, egg #5 in **Figure 1**).

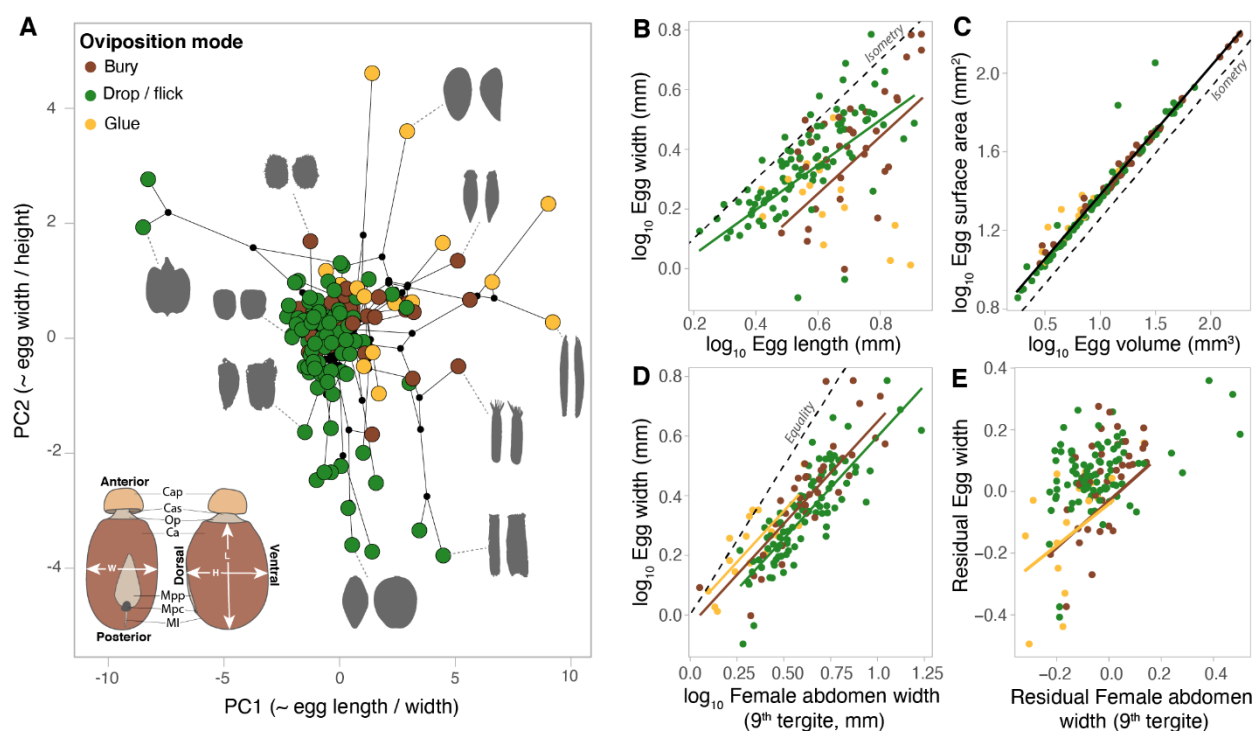


Figure 3: Egg shape and allometric scaling. **A**, Phasmid egg morphospace showing the two first axes of a phylogenetic principal component analysis. Black lines between points represent the underlying phylogenetic relationships between species. Black dots represent the inferred position of internal nodes. Egg silhouettes are represented in dorsal (left) and side (right) view. Bottom left inset shows the drawing of a phasmid egg (*Eurycnema osiris*, Lanceocercata): Cap, capitulum; Cas: capitulum stalk; Op, operculum; Ca, egg capsule; Mpp, micropylar playe; Mpc, micropylar cup; Ml, median line; L, egg length; W, egg width; H, egg height. **B**, Egg width as a function of egg length. **C**, Egg surface area as a function of egg volume. **D**, Egg width as a function of the width the adult female ninth tergite (i.e., where the egg is released). **E**, Residual egg width (i.e., residuals from the PGLS regression between egg width and egg length) compared to residual width of female ninth tergite (i.e., residuals from the PGLS regression between width of ninth tergite and adult female body volume). Phylogenetic least square regressions are represented in **B-E** and correspond to the analyses reported in **Table 1**. Colors correspond to oviposition mode (see legend in **A**). Dashed lines represent isometric slopes (arbitrary intercept) in **B** and **C**, or the equality line in **D**.

Egg width was positively correlated with egg length in dropped and buried eggs but not in glued eggs (**Table 2, Figure 3B**). The scaling relationship between egg width and egg length was significantly hypoallometric in dropped eggs ($\beta = 0.78 \pm 0.07$, 95% confidence interval = [0.64; 0.92], isometric slope = 1) but did not differ from isometry in buried eggs (1.03 ± 0.18 , 95% confidence interval = [0.68; 1.39], isometric slope = 1). Buried eggs were relatively more elongated than dropped eggs (**Table 2, Figure 3B**) while glued eggs ranged from the most elongated eggs (e.g., *Clonaria conformans*, egg #57 in **Figure 1**) to more spherical eggs, notably in species where females glue their eggs in batches (**Figure 3A**; e.g., *Trachythorax maculicollis*, egg #38 in **Figure 1**). This morphological diversity suggests a strong influence of substrate properties on which eggs are glued (e.g., bark, grass, leaves) and laying strategy (e.g., eggs laid singly or in batches) on the shape of glued eggs.

Overall, egg surface area scaled isometrically with egg volume ($\beta = 0.65 \pm 0.01$, 95% confidence interval = [0.633; 0.674], isometric slope = 0.667) and this relationship was not affected by oviposition mode (**Table 2, Figure 3C**).

Egg width scaled hypoallometrically with the width of the adult female's ninth tergite, where the opening of the oviduct is located (i.e., the putative maximum egg width for the egg to be able to come out of the oviduct) (**Table 2, Figure 3D**; $\beta = 0.68 \pm 0.04$, 95% confidence interval = [0.60; 0.77], isometric slope = 1). The width of almost all eggs was smaller than that of the ninth tergite (**Figure 3D**). Glued and buried eggs were relatively wider and their width was closer to the width of the female ninth tergite, sometimes slightly larger suggesting potential dilatation of the oviduct opening (**Table 2, Figure 3D**). Residual egg width after accounting for egg length (i.e., reflecting aspect ratio) was positively correlated with residual female ninth tergite width, after accounting for female volume, in species that bury or glue their eggs but not significantly in species that drop or flick them away (**Table 2, Figure 3E**). Thus, females with a relatively narrower abdominal extremity (i.e., with a relatively more elongated body) lay relatively more elongated eggs but this relationship was only true for species that bury or glue their eggs, which tend to be relatively larger (**Figure 2B**) and closer to the limit in egg width imposed by the oviduct opening diameter (**Figure 3D**). This suggests that, in species with more derived oviposition modes and relatively larger eggs, the width of these eggs is mechanically constrained by the diameter of the female oviduct opening.

Table 2: Allometric scaling of egg shape and mechanical constraints. The table presents the results of phylogenetic generalized least square models (PGLS). The most likely value of Pagel’s lambda (phylogenetic signal) is presented along with type-I (sequential) ANOVA outputs and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. Residual egg width corresponds to the residuals of a PGLS regression between egg width and egg length and therefore represents a measure of aspect ratio with lower values indicating more elongated eggs. Residual female ninth tergite width (i.e., the putative maximum diameter of the oviduct opening) corresponds to the residuals of a PGLS regression between ninth tergite width and female volume and therefore represents the relative elongation of the apex of the female’s abdomen, with lower values indicating a narrower abdomen. Significant effects are bolded ($p < 0.05$).

Response (n species)	Predictor	λ	$F_{df1,df2}$	P	Effect size \pm Standard Error (continuous) OR Post-hoc pairwise tests (Holm) (categorical)
log ₁₀ Egg width (n = 143)	log ₁₀ Egg length	0.97	$F_{1,137} = 107.6$	<.0001	0.95 \pm 0.16
	Oviposition mode		$F_{2,137} = 7.53$	0.0008	Intercept pairwise comparisons: Drop – bury = 0.10 \pm 0.03, p=0.005 Drop – glue = 0.11 \pm 0.04, p=0.005 Bury – glue = 0.02 \pm 0.04, p=0.63
	Interaction		$F_{2,137} = 8.65$	0.0003	Slope pairwise comparisons: Drop – bury = -0.20 \pm 0.18, p=0.25 Drop – glue = 0.87 \pm 0.23, p=0.0005 Bury – glue = 1.08 \pm 0.27, p=0.0003
log ₁₀ Egg surface area (n = 143)	log ₁₀ Egg volume	1	$F_{1,137} = 3913$	<.0001	0.64 \pm 0.02
	Oviposition mode		$F_{2,137} = 2.78$	0.07	Intercept pairwise comparisons: Drop – bury = -0.02 \pm 0.01, p=0.16 Drop – glue = -0.02 \pm 0.02, p=0.84 Bury – glue = 0.005 \pm 0.02, p=0.84
	Interaction		$F_{2,137} = 1.08$	0.34	Slope pairwise comparisons: Drop – bury = 0.03 \pm 0.02, p=0.73 Drop – glue = 0.06 \pm 0.05, p=0.73 Bury – glue = 0.04 \pm 0.05, p=0.73
log ₁₀ Egg width (n = 142)	log ₁₀ Female 9 th tergite width	0.57	$F_{1,138} = 242.8$	<.0001	0.68 \pm 0.04
	Oviposition mode		$F_{2,138} = 6.83$	0.0015	Drop – bury = -0.05 \pm 0.02, p=0.03 Drop – glue = -0.09 \pm 0.03, p=0.004 Bury – glue = -0.04 \pm 0.03, p=0.35
Residual egg width (after accounting for egg length) (n = 142)	Residual female 9 th tergite width (after accounting for female volume)	0.95	$F_{1,136} = 29.55$	<.0001	0.76 \pm 0.19
	Oviposition mode		$F_{1,136} = 7.22$	0.001	Intercept pairwise comparisons: Drop – bury = 0.07 \pm 0.02, p=0.01 Drop – glue = 0.07 \pm 0.04, p=0.07 Bury – glue = 0.005 \pm 0.04, p=0.88
	Interaction		$F_{1,136} = 4.41$	0.014	Slope pairwise comparisons: Drop – bury = -0.51 \pm 0.21, p=0.044 Drop – glue = -0.43 \pm 0.19, p=0.046 Bury – glue = 0.08 \pm 0.24, p=0.73

Response to ecological circumstances. After accounting for the effects of adult female volume, lifetime fecundity and oviposition mode, we found significant effects of female flight capacity and climate PC1 on egg volume (**Table 3**). Flight capable females and females living in more temperate and seasonal regions (reflected by climate PC1) laid relatively smaller eggs (**Table 3**).

After accounting for egg volume, egg surface area was not significantly affected by any of our ecological predictors. Therefore, species experiencing drier conditions, either seasonally (i.e., high PC2) or constantly (i.e., low PC1) do not seem to minimize egg surface to volume ratio to limit water loss.

After accounting for egg length, relative egg width (i.e., reflecting aspect ratio) was significantly affected by oviposition mode and climate PC1 (**Table 3**). Glued and buried eggs were more elongated than dropped eggs and eggs of species found in drier and more temperate and seasonal regions were slightly rounder (**Table 3**).

Total female reproductive output was significantly affected by female flight capacity and climate PC1, even after accounting for female body volume (**Table 3**). Flightless females were able to invest relatively more in egg production than were flight capable ones. This suggests either a physiological cost of the flight apparatus or a cost of carrying more and/or larger eggs for flying, and potentially explains why even after accounting for relative fecundity, flightless females lay relatively larger eggs. Similarly, females in regions with a higher net primary productivity (and likely benefitting from greater food availability) were able to invest relatively more in egg production, also suggesting that females in temperate and less productive regions may be resource-limited compared to those from tropical regions. This could also potentially explain why tropical females lay relatively larger eggs even after accounting for relative fecundity. However, we note that our fecundity data were mostly obtained from breeding cultures, in which insects are fed *ad libitum*. Alternatively, differences in total reproductive investment may reflect reduced adult life spans in seasonal regions with short periods of favorable conditions.

Finally, after accounting for variation in egg volume, the duration of embryogenesis was affected only by oviposition mode, with glued eggs developing relatively faster than buried and dropped eggs (**Table 1,3**). The other ecological variables had no significant effects (**Table 3**). However, our data on the duration of egg development were obtained from breeding cultures in which eggs are usually incubated in stable room temperatures (20-24°C) and high humidities (>80%). Therefore, the effects of climate variables may be masked, especially if duration of embryogenesis is highly sensitive to temperature. Nevertheless, this outcome suggests that adaptation to local macroclimatic conditions may not significantly affect the evolution of egg chorion properties (e.g., conductance to water and respiratory gases) or other egg physiological properties. Indeed, such genetic adaptations should lead to differences in egg development duration among species adapted to contrasting climatic conditions once they are held in common conditions in the laboratory. In contrast, microclimatic conditions are likely to play a much larger role. Despite being held in very similar conditions, eggs from species that oviposit onto or into the soil developed more slowly than those glued onto substrates, that likely experience very different microclimatic conditions in the wild.

Table 3: Effect of ecological variables on egg size and shape. The table presents the results of phylogenetic generalized least square models (PGLS). The most likely value of Pagel's lambda (phylogenetic signal) is presented along with type-I (sequential) ANOVA outputs and either estimated effect sizes or post-hoc pairwise comparisons between estimated marginal means using the Holm method to account for multiple testing, respectively for continuous or categorical explanatory variables. Significant effects are bolded ($p < 0.05$).

Response (n species)	Predictor	λ	$F_{df1,df2}$	P	Effect size \pm Standard Error (continuous) OR Post-hoc pairwise tests (Holm) (categorical)
log ₁₀ Egg volume (n = 96)	log ₁₀ Female body volume	0.44	F_{1,88} = 182.6	<.0001	0.66 \pm 0.05
	log ₁₀ lifetime fecundity		F_{1,88} = 34.1	<.0001	-0.32 \pm 0.09
	Oviposition mode		F_{2,88} = 6.47	0.002	Drop – bury = -0.14 \pm 0.06, p=0.09 Drop – glue = -0.07 \pm 0.11, p=1 Bury – glue = 0.07 \pm 0.11, p=1
	Adult female flight capacity		F_{1,88} = 7.27	0.008	Flying – flightless = -0.21 \pm 0.08, p=0.007
	Climate PC1 (~Net primary production, annual temperature and precipitation)		F_{1,88} = 10.2	0.002	0.04 \pm 0.01
	Climate PC2 (~Precipitation seasonality)		F _{1,88} = 2.45	0.12	-0.03 \pm 0.02
log ₁₀ Egg surface area (n = 96)	log ₁₀ Egg volume	0.72	F_{1,88} = 1801.0	<.0001	0.65 \pm 0.02
	log ₁₀ lifetime fecundity		F _{1,88} = 0.03	0.86	0.006 \pm 0.02
	Oviposition mode		F _{2,88} = 0.31	0.73	Drop – bury = -0.009 \pm 0.02, p=1 Drop – glue = -0.015 \pm 0.03, p=1 Bury – glue = -0.006 \pm 0.11, p=1
	Adult female flight capacity		F _{1,88} = 0.04	0.84	Flying – flightless = 0.004 \pm 0.02, p=0.86
	Climate PC1 (~Net primary production, annual temperature and precipitation)		F _{1,88} = 0.79	0.38	0.003 \pm 0.003
	Climate PC2 (~Precipitation seasonality)		F _{1,88} = 1.16	0.28	-0.005 \pm 0.005
log ₁₀ Egg width (n = 143)	log ₁₀ Egg length	0.98	F_{1,136} = 104.7	<.0001	0.65 \pm 0.07
	Oviposition mode		F_{2,136} = 6.94	0.001	Drop – bury = 0.08 \pm 0.03, p=0.009 Drop – glue = 0.11 \pm 0.04, p=0.009 Bury – glue = 0.03 \pm 0.04, p=0.39
	Adult female flight capacity		F _{1,136} = 1.39	0.24	Flying – flightless = -0.03 \pm 0.03, p=0.23
	Climate PC1 (~Net primary production, annual temperature and precipitation)		F_{1,136} = 7.54	0.007	0.01 \pm 0.005
	Climate PC2 (~Precipitation seasonality)		F _{1,136} = 2.76	0.10	-0.01 \pm 0.007
	log ₁₀ Female body volume		F_{1,89} = 319.4	<.0001	0.86 \pm 0.06
log ₁₀ Reproductive output (egg number \times volume) (n = 96)	Oviposition mode	0.13	F _{2,89} = 2.85	0.06	Drop – bury = 0.096 \pm 0.065, p=0.43 Drop – glue = 0.15 \pm 0.14, p=0.57 Bury – glue = 0.05 \pm 0.14, p=0.72
	Adult female flight capacity		F_{1,89} = 4.40	0.04	Flying – flightless = -0.22 \pm 0.10, p=0.025
	Climate PC1 (~Net primary production, annual temperature and precipitation)		F_{1,89} = 8.64	0.004	0.04 \pm 0.01
	Climate PC2 (~Precipitation seasonality)		F _{1,89} = 0.20	0.66	-0.01 \pm 0.02
	log ₁₀ Egg volume		F_{1,114} = 44.2	<.0001	0.19 \pm 0.03
	log ₁₀ Duration of embryogenesis (under 20-24°C) (n = 121)		Oviposition mode	1	F_{2,114} = 6.38
Adult female flight capacity		F _{1,114} = 0.08	0.77		Flying – flightless = 0.009 \pm 0.03, p=0.75
Climate PC1 (~Net primary production, annual temperature and precipitation)		F _{1,114} = 0.01	0.91		0.0009 \pm 0.006
Climate PC2 (~Precipitation seasonality)		F _{1,114} = 0.15	0.70		-0.003 \pm 0.007

Egg metabolic rate, energy use and water loss. The five sampled species differed extensively in egg size and consequently in duration of embryogenesis, from 3.7 mg (average 4.07 ± 0.07 mg) and 94 days (average 98 ± 0.9 days) in *Medauroidea extradentata*, to 173 mg (average 155.9 ± 2.4 mg) and 377 days (average 341 ± 6.8 days) in *Heteropteryx dilatata*. Egg metabolic rate increased exponentially during embryonic development in all five species (**Figure 4A**). Mid-development and mean metabolic rate scaled hypoallometrically with egg mass across species (respectively $\beta = 0.70 \pm 0.16$, 95% confidence interval: [0.36 ; 0.98], and $\beta = 0.78 \pm 0.10$, [0.58 ; 0.97]; **Table 4, Figure 4B**), indicating that larger egg have a relatively lower metabolic rate than smaller eggs. Estimated cumulative CO₂ produced during embryogenesis scaled isometrically with egg mass (**Table 4, Figure 4C**), suggesting that larger eggs use proportionately more energy than smaller eggs for development. In all five species, eggs lost water at a constant rate during embryogenesis (**Figure 4D**), and this rate increased hypoallometrically with egg mass after accounting for oviposition mode (**Table 4, Figure 4E**). The recovered scaling exponent (0.29 ± 0.07 , [0.15 ; 0.43]) was significantly lower than the expected scaling exponent of egg surface area with egg mass (0.66) suggesting that larger eggs have reduced chorion permeability, possibly from increased thickness or changes in layer properties (Woods et al., 2005). Interestingly, water loss rate was relatively higher in the two largest species, which bury their eggs. The three others drop their eggs (**Table 4, Figure 4E**). While the effect of size and oviposition mode may be confounded, these patterns suggest that buried eggs are relatively more permeable to water. Finally, we found a strong isometric relationship between egg water content and egg mass (**Table 4, Figure 4F**), suggesting that small and large eggs contain the same proportion water ($49 \pm 0.01\%$).

Table 4: Allometry of egg energy use and water loss. The table presents the results of linear mixed-effects models accounting for species as a random effect. Effect sizes (i.e., scaling exponents) are indicated along with type II ANOVA (Wald χ^2) outputs including p values.

Response	Fixed effect	Random effect	Effect size \pm Standard Error [95% Confidence interval]	Wald χ^2_{df}	P
log ₁₀ Mid embryonic development metabolic rate	log ₁₀ Egg mass	species	0.70 ± 0.16 [0.36 ; 0.98]	$\chi^2_1 = 19.9$	<.0001
log ₁₀ Mean embryonic development metabolic rate	log ₁₀ Egg mass	species	0.78 ± 0.10 [0.58 ; 0.97]	$\chi^2_1 = 62.9$	<.0001
log ₁₀ Total CO ₂ produced during embryogenesis	log ₁₀ Egg mass	species	0.97 ± 0.09 [0.76 ; 1.16]	$\chi^2_1 = 108.3$	<.0001
log ₁₀ Water loss rate	log ₁₀ Egg mass	species	0.29 ± 0.07 [0.15 ; 0.43]	$\chi^2_1 = 16.0$	<.0001
	Oviposition mode		Drop – Bury : -0.81 ± 0.08 [-0.97 ; -0.65]	$\chi^2_1 = 95.7$	<.0001
log ₁₀ Water content	log ₁₀ Egg mass	species	0.99 ± 0.02 [0.96 ; 1.03]	$\chi^2_1 = 2867$	<.0001

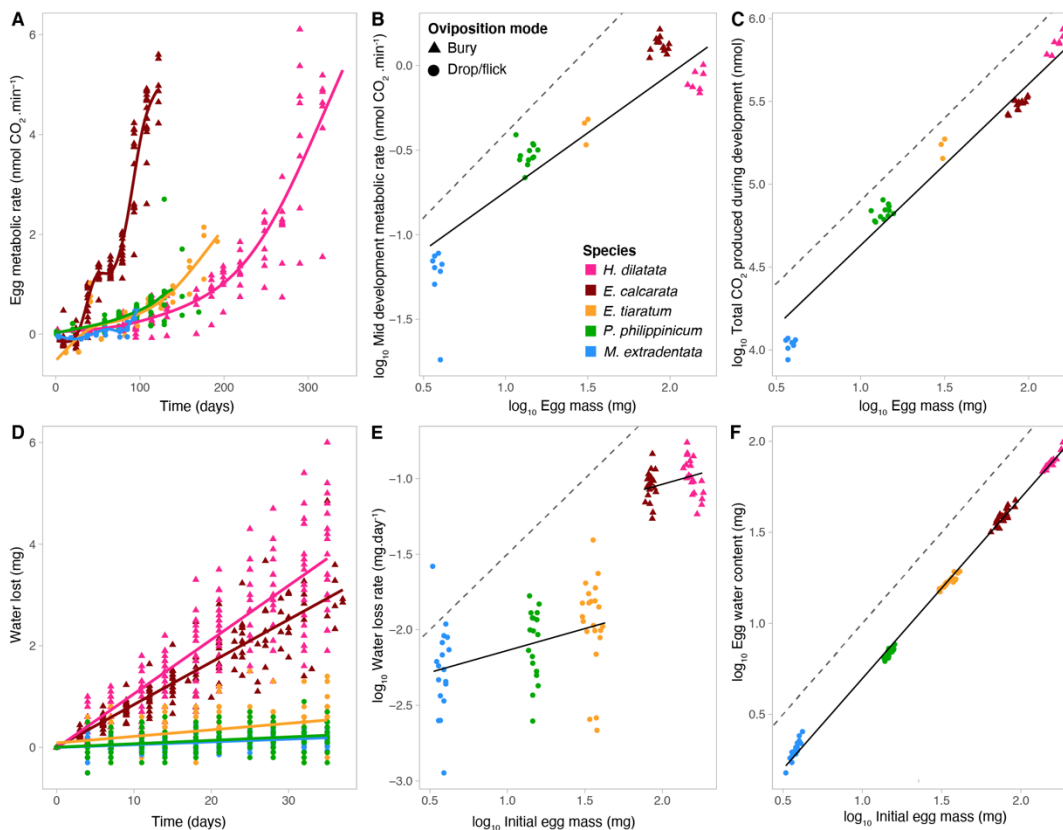


Figure 4: Allometry of egg energy use and water loss. **A.** Egg metabolic rate as a function of time (until hatching). Lines depict smoothing generalized additive models for each species. **B.** Estimated mid embryogenesis metabolic rate compared to egg mass. **C.** Estimated total CO₂ produced during embryogenesis as a function of egg mass. **D.** Cumulative egg water loss as a function of time (until hatching). Lines depict linear regressions for each species. **E.** Water loss rate compared to initial egg mass (right after being laid). **F.** Initial egg water content as function of initial egg mass. Colors represent the five different species and shapes represent oviposition modes (see legend in **B**). Dashed lines represent isometry (arbitrary intercept). Solid black lines represent linear mixed-effects regressions (see **Table 4** for details on the statistical analyses).

Discussion

Using a single mesodiverse clade of insects – the stick and leaf insects (order Phasmatodea, comprising approximately 3,400 described species) – we assessed a set of hypothesized drivers of egg diversification by combining phylogenetic analyses with physiological measurements of rates of metabolism and water loss. Together, the data suggest that egg morphological diversification is driven by a complex suite of factors including broad patterns of life history variation and covariation, constraints arising from eggs interacting mechanically with various functions of the female body, the

effects of female oviposition and locomotor (flight) behavior, and, to a lesser extent, the influences of climatic context in which lineages have evolved.

Distinguishing drivers from responses is not always possible, given that the data permit us to evaluate primarily correlations among traits in a phylogenetic context. Phylogenetic generalized least squares (PGLS) methods enable the analysis of the covariation between traits or traits and environmental characteristics while controlling for data points not being statistically independent due to shared ancestry (Martins and Hansen, 1997; Mundry, 2014; Revell, 2010; Symonds and Blomberg, 2014). But, like traditional linear models, they allow the analysis of only one response variable at a time, which is limiting when analyzing multiple variables as the underlying causal structure. Moreover, multivariate relationships may be much more complex and involve traits that can be responses and predictors. Egg size, for example, can be viewed both as a response (it is highly variable across the clade, variation in need of explanation) whose evolution integrates multiple influences. But it is also a predictor, and in some contexts likely even a driver, of other traits like fecundity, egg developmental duration, and egg metabolic rate. One proposed solution to disentangle cause-effect relationships among variables is the use of phylogenetic path analysis (Gonzalez-Voyer and von Hardenberg, 2014; Hardenberg and Gonzalez-Voyer, 2013) which will be applied to our data in the future. Our analyses thus focus on identifying patterns of correlation, with a secondary focus on inferring causation. Therefore, it should be kept in mind that, while the correlations we report may support some of the causal hypotheses laid out in the introduction, another underlying causal structure may be at play.

In our analyses, the two most important predictors of other egg traits were egg size and oviposition mode. In general, larger eggs were laid by larger females, and those females typically laid relatively fewer eggs, providing strong evidence of egg size-number tradeoffs, consistent with previous studies on insects (Berrigan, 1991; García-Barros, 2000). Larger eggs also consistently required longer periods of time to develop. This result stands in striking contrast to the data reported by Church et al. (2019b), who found no relationship between egg size and developmental duration. The difference likely reflects the phylogenetic scope of analysis. In their analysis, Church et al. (2019b) used data on 64 species distributed across seven major insect orders, such that order-specific differences may have obscured direct effects of egg size, if any, on duration of development. By contrast, our study, which analyzed 121 species distributed within a single order, finds strong evidence that larger eggs require longer developmental periods. This result is consistent with other reports on both insects (García-

Barros and Munguira, 1997; Gillooly and Dodson, 2000; Maino et al., 2017) and vertebrates (Balon, 1984; Paine, 1985; Rahn and Ar, 1974).

Consistent with our observation that large eggs require longer developmental periods, egg size also had systematic effects on rates of egg metabolism and water loss. Specifically, both rates scaled hypoallometrically, such that large eggs had disproportionately low rates. Although the confidence intervals are wide due to the low number of species used (five), the values of the metabolic scaling coefficient ($\beta = 0.70$, 95% confidence interval: [0.36 ; 0.98] for mid-development metabolic rate; $\beta = 0.78$ [0.58 ; 0.97] for mean metabolic rate) lie in the range expected from a wide variety of other metabolic scaling studies (insect eggs: $\beta = 0.92$ [0.86; 0.98] (Maino and Kearney, 2014), non-avian reptile eggs: $\beta = 0.82$ [0.77; 0.87] (Vleck and Hoyt, 1991), avian eggs: $\beta = 0.714 \pm 0.09$ (SE) (Hoyt and Rahn, 1980; Mortola, 2009), ectotherm eggs: $\beta = 0.66$ [0.65; 0.80] (Pettersen et al., 2022)).

The emerging picture is of large eggs generally having long developmental durations supported by low mass-specific metabolic rates. There is scope, however, for additional covariation in metabolic rates and developmental duration beyond just the effects of size. For example, the two largest eggs in our experiment, *E. calcarata* (84.2 ± 0.8 mg) and *H. dilatata* (155.9 ± 2.4 mg), had relatively similar metabolic rates at the temporal midpoint of development (**Figure 4B**). However, *E. calcarata* eggs develop much faster (121 ± 2 days versus 335 ± 5 days for *H. dilatata*), which is supported by a much more rapid time course of increase in metabolism during the course of development (**Figure 4A**). Finally, the total energy devoted to development, as measured by estimated total CO₂ emission over the entirety of development, scaled isometrically with egg size (**Figure 4C**). Collectively, these observations suggest that egg size, developmental duration, rates of energy (and water) expenditure, and total energy reserves are all tightly correlated (Rahn et al., 1974). This outcome likely reflects the division of eggs into finite packets of energy and materials that exchange only water, carbon dioxide, and oxygen with the surrounding habitat.

The second major predictor of egg traits was oviposition mode – whether females drop/flick their eggs (the ancestral state; see **Figure 1**) or glue or bury them (two derived states that each have evolved at least five times within the clade). The shift to gluing or burying is associated with major shifts in many other traits. In particular, females of gluing or burying species typically lay fewer, larger eggs, reflecting greater parental investment per offspring associated with the greater time investment involved in placing the eggs in specific locations. Similar patterns have also been found in other

ectotherms including fishes (Kolm and Ahnesjö, 2005) and frogs (Summers et al., 2005). However, a recent study failed to find a relationship between egg size and the extent of parental care across 287 species of insects from 16 orders, and only found an effect on egg number (Gilbert and Manica, 2010). In addition, glued (but not buried) eggs evolved significantly shorter developmental durations for their size. Glued eggs present the advantage that offspring will hatch directly on the host plant, saving the time and effort associated with finding a suitable host plant and climbing it (e.g., *Spinosipylloidea doddi* females attach their hairy eggs on the hairy leaves of their unique and very tall host tree (Brock and Hasenpusch, 2009)). However, they may be more vulnerable to egg predators and parasitoids which, we hypothesize, could select for shorter development times (Schwabl et al., 2007).

Gluing and burying are also associated with the evolution of elongated eggs. This effect may reflect one or more of several pressures: ease of oviposition – elongated eggs may be easier to bury within soil; camouflage – elongated eggs laid on leaves may be more difficult for visual predators to spot if they look less like the typically round eggs of most insect species or more like plant seeds (Guerra-Grenier, 2019); enhanced oxygen access – oxygen can be less concentrated in wet soils and elongated eggs may maximize their uptake by minimizing diffusion distances between surface and central tissues (Silver et al., 1999; Woods and Hill, 2004); contact surface area – glued elongated eggs may attach more strongly to substrates thanks to a larger contact area (Büscher et al., 2020); and passage through the oviduct – elongated eggs may pass through the female oviduct more easily. Females of species that glue or bury tend to produce relatively larger eggs, the diameters of which can approach or even exceed the diameters of the ninth abdominal segment through which they pass during oviposition (**Figure 3D-E**). Because glued or buried eggs are typically larger, there may be more pressure to evolve shapes that still will allow easy passage through the oviduct (Iverson and Ewert, 1991; Stoddard et al., 2017).

One surprising outcome of our analyses was the minimal predictive effects of climate variables. Phasmids are distributed across a broad range of tropical and temperate habitats (Bedford, 1978; Günther, 1953), with significant variation in mean conditions and patterns of variation in major climate variables such as temperature, humidity, and rainfall (Figure S3). A priori, one might expect that eggs in habitats with low annual precipitation would be larger and rounder to maximize water retention through a reduced surface to volume ratio; that eggs in more seasonal and potentially unfavorable environments would be larger to favor offspring survival and performance (Fischer et al., 2003); or that eggs in less productive environments (i.e., with a lower resource availability) would be smaller

because of a reduced reproductive investment (Kojima and Lin, 2022). Indeed, we found evidence that females from temperate and drier areas associated with lower net primary productivity tended to lay slightly smaller and rounder eggs, even after accounting for fecundity. Thus, we show that these females invest relatively less in egg production (reproductive output), compared to females of tropical species. Overall, however, the effect sizes of the climate variables (collapsed into principal component axes) were small compared to the effect sizes of other predictors examined in this study (**Table 3**). This outcome likely reflects that egg experience is poorly captured by the large-scale gridded climate data we used in our analysis (bin size 4.6 km on a side). Eggs experience microclimates rather than macroclimates, and conditions in or on the surface of the soil may differ strongly from conditions characterizing that grid cell (Potter et al., 2013). Likewise, eggs that are glued onto leaves experience conditions at least partly determined by leaf physiology (Pincebourde and Woods, 2012).

Finally, we found a relatively strong effect of flight capacity on egg size (**Table 3**): flightless females laid relatively larger eggs and had a relatively higher reproductive output compared to flight capable ones. Stick and leaf insects are ideally suited to address this specific question as they have undergone numerous shifts between winged forms and wingless forms (Bank and Bradler, 2022; Whiting et al., 2003). In accordance with Church et al. (2019a), who also focused on Phasmatodea for this analysis, we did not find any effect of flight capacity on egg shape, in contrast with birds (Stoddard et al., 2017). The relatively reduced egg size of flight capable females suggests that the physiological cost of wings and associated flight muscles may tradeoff with reproductive output (Biewener and Patek, 2018). Alternatively, large eggs may impair the flight performance of females by increasing wing loading (Boisseau et al., 2022).

Overall, our results on Phasmatodea align with the broad findings of Church et al. (2019a), notably in identifying a strong association between oviposition ecology and egg size and shape. In their study, “oviposition ecology” referred primarily to whether insects laid their eggs in aerial environments or in aquatic environments, which characterize both insects that lay eggs in freshwater or parasitoids that lay eggs in the tissues of other animals. They found that eggs laid in aquatic environments were systematically smaller and more elongated. We were unable to test the consequences of evolutionary transitions between aerial and aquatic oviposition sites, as no phasmids oviposit in water or are parasitoids. Nevertheless, like Church et al., we found that one of the most powerful forces shaping egg size, shape, and developmental characteristics is where females lay their eggs. Oviposition mode – whether they drop/flick, glue, or bury their eggs – has pervasive effects on the evolution of egg traits

per se (size, shape, developmental physiology) and also on the relationship between egg and other adult female traits. These results suggest that renewed attention should be directed toward understanding how ovipositing females experience interactions between eggs and their substrates, and the physical and biotic experiences that eggs have in their microsities.

Conclusion

In conclusion, the staggering diversity of insect eggs is obvious, but we have lacked clear understanding of the major processes that underpin that diversification. Recent studies that leverage larger data sets, including this one, are starting to reveal correlates and drives of the diversity. In particular, the ecology of oviposition – and for phasmids this means mode of oviposition (flick/drop, bury, glue) – profoundly influences the size, shape, and developmental dynamics of eggs.

Acknowledgements

We thank François Tetaert for allowing us to use an enormous number of photos of stick insect eggs on his website phasmatodea.fr.; Garret Jolma for help in the lab developing early water-loss protocols; Brenna Shea and The Missoula Insectarium for providing fresh eggs of several phasmid species.

References

- Arai, M. and Yago, M.** (2015). Curious Oviposition Behavior in *Phyllium westwoodii* (Phasmatodea: Phylliidae): Preliminary Observations. *Journal of Insect Science* **15**, 135.
- Balon, E. K.** (1984). Patterns in the evolution of reproductive styles in fishes. In *Fish reproduction: strategies and tactics* (ed. Potts, G. W. and Wootton, R. J.), pp. 35–53. London, UK: Academic Press.
- Bank, S. and Bradler, S.** (2022). A second view on the evolution of flight in stick and leaf insects (Phasmatodea). *BMC Ecology and Evolution* **22**, 1–17.
- Bates, D., Maechler, M., Bolker, B. and Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48.

- Bedford, G. O.** (1978). Biology and ecology of the Phasmatodea. *Annual Review of Entomology* **23**, 125–149.
- Berrigan, D.** (1991). The Allometry of Egg Size and Number in Insects. *Oikos* **60**, 313–321.
- Biewener, A. A. and Patek, S. N.** (2018). *Animal Locomotion*. Oxford, United Kingdom: Oxford University Press.
- Blackburn, T. M.** (1991). An interspecific relationship between egg size and clutch size in bird. *The Auk* **108**, 973–977.
- Boisseau, R. P., Büscher, T. H., Klawitter, L. J., Gorb, S. N., Emlen, D. J. and Tobalske, B. W.** (2022). Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect. *BMC Ecology and Evolution* **22**, 1–18.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A. and Drummond, A. J.** (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLOS Computational Biology* **10**, e1003537.
- Bradler, S., Cluquennois, N. and Buckley, T. R.** (2015). Single origin of the Mascarene stick insects: ancient radiation on sunken islands? *BMC Evolutionary Biology* **15**, 1–10.
- Brock, P. D.** (2000). *A Complete Guide to Breeding Stick and Leaf Insects*. Neptune City, New Jersey, USA: TFH Publications.
- Brock, P. D. and Hasenpusch, J. W.** (2009). *The complete field guide to stick and leaf insects of Australia*. Collingwood, Victoria, Australia: CSIRO publishing.
- Brock, P. D., Büscher, T. H. and Baker, E.** (2021). Phasmida Species File Online. *Version 5.0/5.0*.
- Büscher, T. H., Quigley, E. and Gorb, S. N.** (2020). Adhesion performance in the eggs of the Philippine leaf insect *Phyllium philippinicum* (Phasmatodea : Phylliidae). *Insects* **11**, 400.
- Church, S. H., Donoughe, S., de Medeiros, B. A. S. and Extavour, C. G.** (2019a). Insect egg size and shape evolve with ecology but not developmental rate. *Nature* **571**, 58–62.
- Church, S. H., Donoughe, S., de Medeiros, B. A. S. and Extavour, C. G.** (2019b). A dataset of egg size and shape from more than 6,700 insect species. *Scientific Data* **6**, 1–11.
- Czesak, M. E. and Fox, C. W.** (2003). Evolutionary ecology of egg size and number in a seed beetle: genetic trade-off differs between environments. *Evolution* **57**, 1121–1132.
- Donoughe, S.** (2022). Insect egg morphology: evolution, development, and ecology. *Current Opinion in Insect Science* **50**, 100868.
- Eisner, T., Eisner, M., Rossini, C., Iyengar, V. K., Roach, B. L., Benedikt, E. and Meinwald, J.** (2000). Chemical defense against predation in an insect egg. *Proceedings of the National Academy of Sciences* **97**, 1634–1639.
- Fick, S. E. and Hijmans, R. J.** (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315.
- Fischer, K., Brakefield, P. M. and Zwaan, B. J.** (2003). Plasticity in butterfly egg size: why larger offspring at lower temperatures? *Ecology* **84**, 3138–3147.
- García-Barros, E.** (2000). Body size, egg size, and their interspecific relationships with ecological and life history traits in butterflies (Lepidoptera: Papilionoidea, Hesperioidea). *Biological Journal of the Linnean Society* **70**, 251–284.
- García-Barros, E. and Munguira, M. L.** (1997). Uncertain branch lengths, taxonomic sampling error, and the egg to body size allometry in temperate butterflies (Lepidoptera). *Biological Journal of the Linnean Society* **61**, 201–221.

- Gibbs, M., Lace, L. A., Jones, M. J. and Moore, A. J.** (2005). Egg size-number trade-off and a decline in oviposition site choice quality: female *Pararge aegeria* butterflies pay a cost of having males present at oviposition. *Journal of Insect Science* **5**, 1–9.
- Gilbert, J. D. J. and Manica, A.** (2010). Parental care trade-offs and life-history relationships in insects. *American Naturalist* **176**, 212–226.
- Gillooly, J. F. and Dodson, S. I.** (2000). The relationship of egg size and incubation temperature to embryonic development time in univoltine and multivoltine aquatic insects. *Freshwater Biology* **44**, 595–604.
- Glazier, D. S.** (2010). A unifying explanation for diverse metabolic scaling in animals and plants. *Biological Reviews* **85**, 111–138.
- Goldberg, J., Bresseel, J., Constant, J., Kneubühler, B., Leubner, F., Michalik, P. and Bradler, S.** (2015). Extreme convergence in egg-laying strategy across insect orders. *Scientific reports* **5**, 1–7.
- Gonzalez-Voyer, A. and von Hardenberg, A.** (2014). An introduction to phylogenetic path analysis. In *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice* (ed. Zsolt Garamszegi, L.), pp. 201–229. Berlin Heidelberg: Springer-Verlag.
- Guerra-Grenier, E.** (2019). Evolutionary ecology of insect egg coloration: a review. *Evolutionary Ecology* **33**, 1–19.
- Günther, K.** (1953). Über die taxonomische Gliederung und die geographische Verbreitung der Insektenordnung der Phasmatodea. *Beiträge zur Entomologie* **3**, 541–563.
- Hardenberg, A. von and Gonzalez-Voyer, A.** (2013). Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution* **67**, 378–387.
- Henneguy, L. F.** (1890). Note sur la structure de l'enveloppe de l'oeuf des Phyllies. *Bulletin de Société Philomantique de Paris* **2**, 18–25.
- Hennemann, F. H., Conle, O. V., Brock, P. D. and Seow-Choen, F.** (2016). Revision of the Oriental subfamily Heteropteryginae kirby, 1896, with a re-arrangement of the family Heteropterygidae and the descriptions of five new species of *Haaniella* kirby, 1904. (Phasmatodea: Areolatae: Heteropterygidae). *Zootaxa* **4159**, 1–219.
- Hijmans, R. J.** (2021). raster: Geographic Data Analysis and Modeling.
- Hinton, H. E.** (1981). *Biology of Insect Eggs*. Oxford, UK: Pergamon Press.
- Hoyt, D. F. and Rahn, H.** (1980). Respiration of avian embryos - A comparative analysis. *Respiration Physiology* **39**, 255–264.
- Iverson, J. B. and Ewert, M. A.** (1991). Physical characteristics of reptilian eggs and a comparison with avian eggs. In *Egg Incubation: Its Effect on Embryonic Development in Birds and Reptiles*. (ed. Ferguson, M. W. J. and Deeming, D. C.), pp. 87–100. New York, New York, USA: Cambridge University press.
- Kobayashi, S., Usui, R., Nomoto, K., Ushirokita, M., Denda, T. and Izawa, M.** (2014). Does egg dispersal occur via the ocean in the stick insect *Megacrania tsudai* (Phasmida: Phasmatidae)? *Ecological Research* **29**, 1025–1032.
- Koch, L. K. and Meunier, J.** (2014). Mother and offspring fitness in an insect with maternal care: Phenotypic trade-offs between egg number, egg mass and egg care. *BMC Evolutionary Biology* **14**, 1–9.
- Kojima, W. and Lin, C. P.** (2022). Non-linear latitudinal cline of egg size and its consequence for larval survival in the rhinoceros beetle. *Biological Journal of the Linnean Society* **136**, 375–383.
- Kolm, N. and Ahnesjö, I.** (2005). Do egg size and parental care coevolve in fishes? *Journal of Fish Biology* **66**, 1499–1515.

- Kratochvíl, L. and Frynta, D.** (2006). Egg shape and size allometry in geckos (Squamata: Gekkota), lizards with contrasting eggshell structure: why lay spherical eggs? *Journal of Zoological Systematics and Evolutionary Research* **44**, 217–222.
- Lenth, R.** (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means.
- Maino, J. L. and Kearney, M. R.** (2014). Ontogenetic and interspecific metabolic scaling in insects. *American Naturalist* **184**, 695–701.
- Maino, J. L., Pirtle, E. I. and Kearney, M. R.** (2017). The effect of egg size on hatch time and metabolic rate: theoretical and empirical insights on developing insect embryos. *Functional Ecology* **31**, 227–234.
- Marshall, D. J. and Keough, M. J.** (2007). The evolutionary ecology of offspring size in marine invertebrates. *Advances in Marine Biology* **53**, 1–60.
- Martins, E. P. and Hansen, T. F.** (1997). Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *American Naturalist* **149**, 646–667.
- Moran, A. L. and McAlister, J. S.** (2009). Egg size as a life history character of marine invertebrates: is it all cracked up to be? *Biological Bulletin* **216**, 226–242.
- Mortola, J. P.** (2009). Gas exchange in avian embryos and hatchlings. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **153**, 359–377.
- Mundry, R.** (2014). Statistical issues and assumptions of phylogenetic generalized least squares. In *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice* (ed. Zolt Garamszegi, L.), pp. 131–153. Berlin Heidelberg: Springer-Verlag.
- O’Hanlon, J. C., Jones, B. R. and Bulbert, M. W.** (2020). The dynamic eggs of the Phasmatodea and their apparent convergence with plants. *The Science of Nature* **107**, 1–12.
- Paine, M. D.** (1985). Ecological and evolutionary aspects of the early ontogeny of darters (Percidae: Etheostamini).
- Paradis, E. and Schliep, K.** (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528.
- Pettersen, A. K., Schuster, L. and Metcalfe, N. B.** (2022). The evolution of offspring size: a metabolic scaling perspective. *Integrative and Comparative Biology* 1–11.
- Pincebourde, S. and Woods, H. A.** (2012). Climate uncertainty on leaf surfaces: the biophysics of leaf microclimates and their consequences for leaf-dwelling organisms. *Functional Ecology* **26**, 844–853.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team** (2021). nlme: linear and nonlinear mixed effects models.
- Potter, K. A., Woods, H. A. and Pincebourde, S.** (2013). Microclimatic challenges in global change biology. *Global Change Biology* **19**, 2932–2939.
- R Core Team** (2021). R Development Core Team. R: *A Language and Environment for Statistical Computing*.
- Rahn, H. and Ar, A.** (1974). The Avian Egg: Incubation Time and Water Loss. *The Condor* **76**, 147.
- Rahn, H., Paganelli, C. V. and Ar, A.** (1974). The avian egg: air-cell gas tension, metabolism and incubation time. *Respiration Physiology* **22**, 297–309.
- Revell, L. J.** (2010). Phylogenetic signal and linear regression on species data. *Methods in Ecology and Evolution* **1**, 319–329.

- Revell, L. J.** (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223.
- Robertson, J. A., Bradler, S. and Whiting, M. F.** (2018). Evolution of oviposition techniques in stick and leaf insects (Phasmatodea). *Frontiers in Ecology and Evolution* **6**, 1–15.
- Sargent, R. C., Taylor, P. D. and Gross, M. R.** (1987). Parental care and the evolution of egg size in fishes. *The American Naturalist* **129**, 32–46.
- Schwabl, H., Palacios, M. G. and Martin, T. E.** (2007). Selection for rapid embryo development correlates with embryo exposure to maternal androgens among passerine birds. *American Naturalist* **170**, 196–206.
- Seiler, C., Bradler, S. and Koch, R.** (2000). *Ratgeber Phasmiden: Pflege und Zucht von Gespenstschrecken, Stabschrecken und Wandelnden Blättern im Terrarium*. Ruhmannsfelden, Germany: Bede-Verlag.
- Severin, H. H. P.** (1910). A study on the structure of the egg of the walking-stick, *Diapheromera femorata* Say; and the biological significance of the resemblance of phasmid eggs to seeds. *Annals of the Entomological Society of America* **3**, 83–92.
- Sharp, D.** (1898). Account of the Phasmidae, with notes on the eggs. In *Zoological results based on material collected in New Britain, New Guinea, Loyalty Islands and elsewhere*. (ed. Willey, A.), p. Cambridge, UK: Cambridge University Press.
- Silver, W. L., Lugo, A. E. and Keller, M.** (1999). Soil oxygen availability and biogeochemistry along rainfall and topographic gradients in upland wet tropical forest soils. *Biogeochemistry* **44**, 301–328.
- Simon, S., Letsch, H., Bank, S., Buckley, T. R., Donath, A., Liu, S., Machida, R., Meusemann, K., Misof, B., Podsiadlowski, L., et al.** (2019). Old world and new world Phasmatodea: phylogenomics resolve the evolutionary history of stick and leaf insects. *Frontiers in Ecology and Evolution* **7**, 1–14.
- Stanton, A. O., Dias, D. A. and O’Hanlon, J. C.** (2015). Egg Dispersal in the Phasmatodea: Convergence in Chemical Signaling Strategies Between Plants and Animals? *Journal of Chemical Ecology* **41**, 689–695.
- Steele, D. H. and Steele, V. J.** (1975). Egg size and duration of embryonic development in Crustacea. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* **60**, 711–715.
- Stoddard, M. C., Yong, E. H., Akkaynak, D., Sheard, C., Tobias, J. A. and Mahadevan, L.** (2017). Avian egg shape: Form, function, and evolution. *Science* **356**, 1249–1254.
- Suetsugu, K., Funaki, S., Takahashi, A., Ito, K. and Yokoyama, T.** (2018). Potential role of bird predation in the dispersal of otherwise flightless stick insects. *Ecology* **99**, 1504–1506.
- Summers, K., McKeon, C. S. and Heying, H.** (2005). The evolution of parental care and egg size: a comparative analysis in frogs. *Proceedings of the Royal Society B: Biological Sciences* **273**, 687–692.
- Symonds, M. R. E. and Blomberg, S. P.** (2014). A primer on phylogenetic generalised least squares. In *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice* (ed. Garamszegi Zsolt, L.), pp. 105–130. Berlin Heidelberg: Springer-Verlag.
- Tihelka, E., Cai, C., Giacomelli, M., Pisani, D. and Donoghue, P. C. J.** (2020). Integrated phylogenomic and fossil evidence of stick and leaf insects (Phasmatodea) reveal a Permian–Triassic co-origination with insectivores. *Royal Society Open Science* **7**, 201689.
- van Velthuisen, H., Huddleston, B., Fischer, G., Salvatore, M., Ataman, E., Nachtergaele, F. O., Zanetti, M., Bloise, M., Gis, F., Antonicelli, A., et al.** (2007). *Mapping biophysical factors that influence agricultural production and rural vulnerability*. Rome, Italy: Food and Agriculture Organization of the United Nations and International Institute for Applied Systems Analysis.

- Vleck, C. M. and Hoyt, D. F.** (1991). Metabolism and energetics of reptilian and avian embryos. In *Egg incubation: its effects on embryonic development in birds and reptiles* (ed. Deeming, D. C. and Ferguson, M. W. J.), pp. 285–306. Cambridge, UK: Cambridge University Press.
- White, C. R. and Seymour, R. S.** (2005). Allometric scaling of mammalian metabolism. *Journal of Experimental Biology* **208**, 1611–1619.
- Whiting, M. F., Bradler, S. and Maxwell, T.** (2003). Loss and recovery of wings in stick insects. *Nature* **421**, 264–267.
- Winston, P. W. and Bates, D. H.** (1960). Saturated solutions for the control of humidity in biological research. *Ecology* **41**, 232–237.
- Woods, H. A.** (2010). Water loss and gas exchange by eggs of *Manduca sexta*: Trading off costs and benefits. *Journal of Insect Physiology* **56**, 480–487.
- Woods, H. A. and Hill, R. I.** (2004). Temperature-dependent oxygen limitation in insect eggs. *The Journal of experimental biology* **207**, 2267–2276.
- Woods, H. A., Bonnacaze, R. T. and Zrubek, B.** (2005). Oxygen and water flux across eggshells of *Manduca sexta*. *Journal of Experimental Biology* **208**, 1297–1308.

Supplementary information

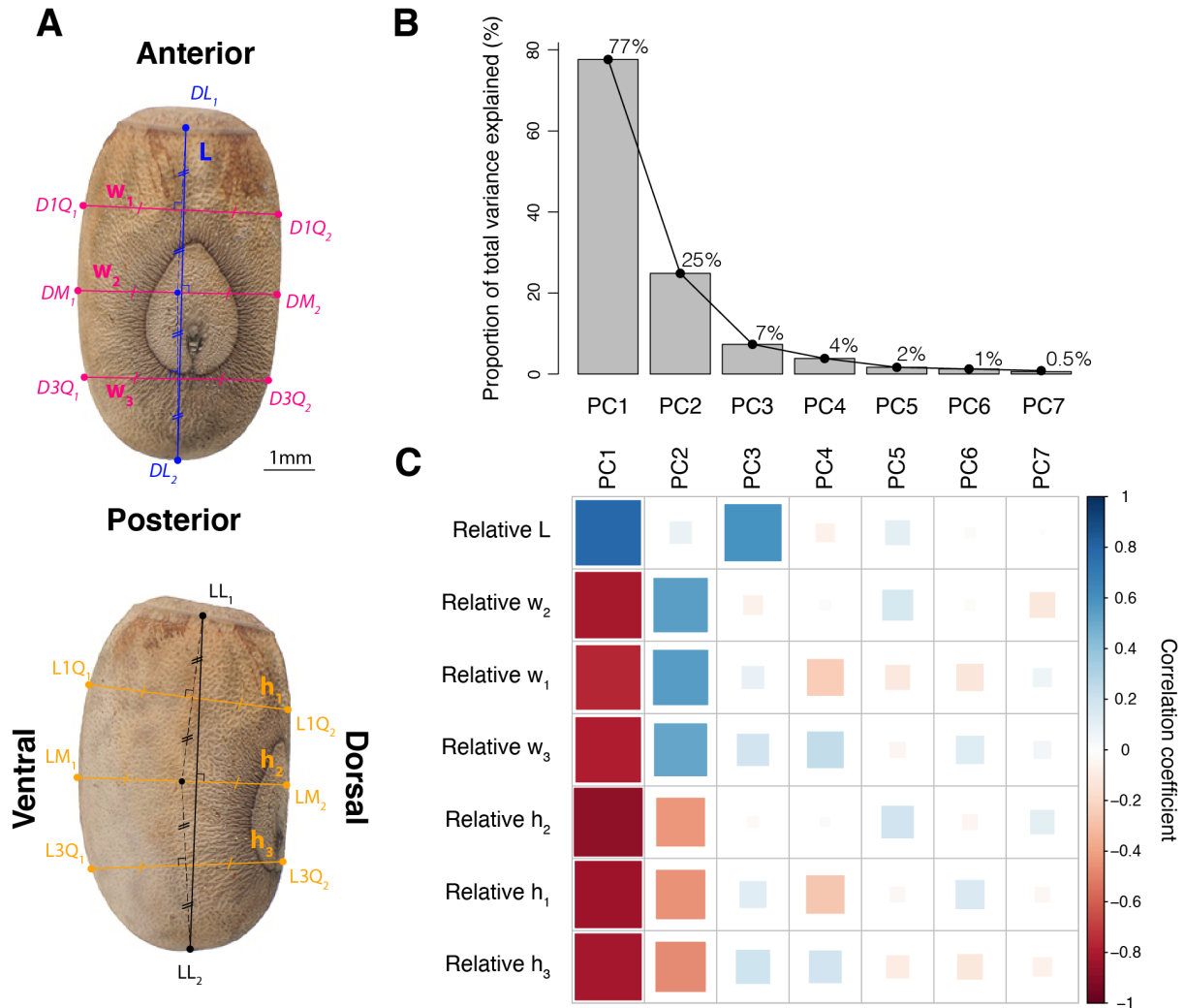


Figure S1: Phasmid egg measurements and principal component analysis (PCA) summarizing variation in egg shape. **A**, Example of guided landmark-based measurements of egg shape traits on two photographs of eggs of *Eurycantha insularis* in dorsal (top) and side (bottom) view following the method of Church et al. (2019b). **B**, Proportion of the total variance explained by each principal component of the egg shape PCA. **C**, Correlation plot showing the extent and direction of the correlations between each principal component (PC) and the egg shape variables included (i.e., residuals of PGLS regressions against egg volume).

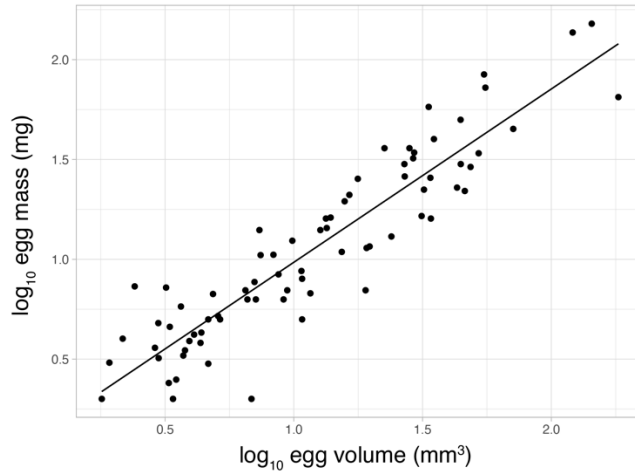


Figure S2: Egg mass as a function of egg volume across phasmid species (PGLS: $\beta=0.87 \pm 0.04$, $p<0.0001$, $R^2=0.84$)

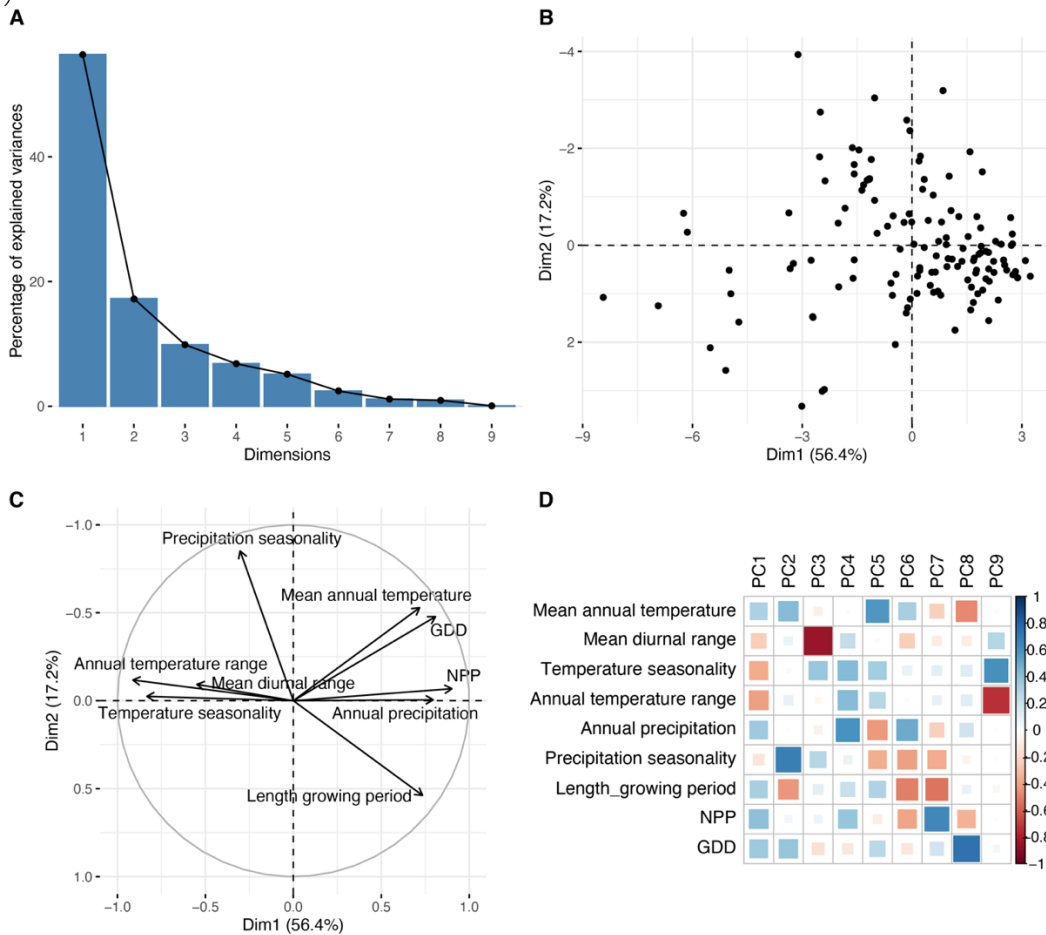


Figure S3: Principal component analysis (PCA) summarizing variation in climatic conditions experienced by phasmid species. **A**, Proportion of the total variance explained by each principal component. **B**, PC2 against PC1 for each species included. **C**, PCA loading plot showing how strongly each climatic variable influences PC1 and PC2. **D**, Correlation plot showing the extent and direction of the correlations between each principal component (PC) and the climatic variables included.

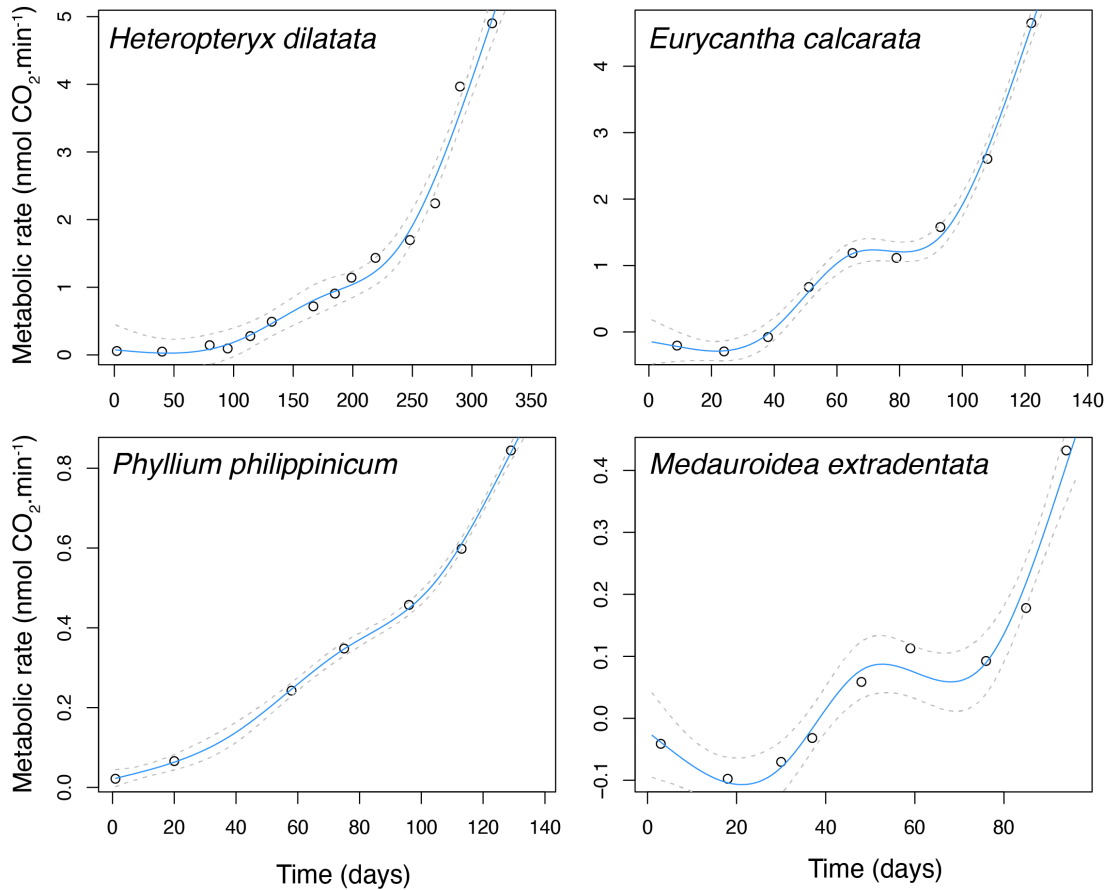


Figure S4: Changes in metabolic rate of individual eggs of four phasmid species during embryonic development (from oviposition to hatching). General additive models (cubic regression splines) are fitted to the developmental trajectory.

Table S1: List of taxa whose eggs are pictured in **Figure 1** and sources of the original pictures.

Egg number	Species	Picture source
1	<i>Timema californicum</i>	R. P. Boisseau
2	<i>Abrosoma festinatum</i>	F. Tetaert, phasmatodea.fr
3	<i>Dinophasma saginatum</i>	F. Tetaert, phasmatodea.fr
4	<i>Cranidium gibbosum</i>	F. Tetaert, phasmatodea.fr
5	<i>Oreophoetes peruana</i>	F. Tetaert, phasmatodea.fr
6	<i>Alienobostra brocki</i>	F. Tetaert, phasmatodea.fr
7	<i>Pseudosermyle phalangiphora</i>	F. Tetaert, phasmatodea.fr
8	<i>Diapheromera femorata</i>	F. Tetaert, phasmatodea.fr
9	<i>Ocnophiloidea regularis</i>	F. Tetaert, phasmatodea.fr
10	<i>Agathemera crassa</i>	Cubillos, C., & Vera, A. (2020). Comparative morphology of the eggs from the eight species in the genus <i>Agathemera</i> Stål (Insecta: Phasmatodea), through phylogenetic comparative method approach. <i>Zootaxa</i> , 4803(3), 523-543.
11	<i>Paraprisopus antillarum</i>	F. Tetaert, phasmatodea.fr
12	<i>Anisomorpha buprestoides</i>	F. Tetaert, phasmatodea.fr
13	<i>Creoxylus spinosus</i>	F. Tetaert, phasmatodea.fr
14	<i>Pseudophasma velutinum</i>	F. Tetaert, phasmatodea.fr
15	<i>Malacomorpha jamaicana</i>	F. Tetaert, phasmatodea.fr
16	<i>Peruphasma schultei</i>	F. Tetaert, phasmatodea.fr
17	<i>Epidares nolimentangere</i>	F. Tetaert, phasmatodea.fr
18	<i>Dares validispinus</i>	F. Tetaert, phasmatodea.fr
19	<i>Orestes mouhotii</i>	F. Tetaert, phasmatodea.fr

20	<i>Heteropteryx dilatata</i>	F. Tetaert, phasmatodea.fr
21	<i>Haaniella erringtoniae</i>	F. Tetaert, phasmatodea.fr
22	<i>Aretaon asperrimus</i>	F. Tetaert, phasmatodea.fr
23	<i>Mearnsiana bullosa</i>	F. Tetaert, phasmatodea.fr
24	<i>Clonopsis gallica</i>	F. Tetaert, phasmatodea.fr
25	<i>Bacillus rossius</i>	F. Tetaert, phasmatodea.fr
26	<i>Parectatosoma mocquersyi</i>	F. Tetaert, phasmatodea.fr
27	<i>Achrioptera punctipes</i>	F. Tetaert, phasmatodea.fr
28	<i>Achrioptera manga</i>	F. Tetaert, phasmatodea.fr
29	<i>Phyllium philippinicum</i>	F. Tetaert, phasmatodea.fr
30	<i>Pulchriphyllium bioculatum</i>	F. Tetaert, phasmatodea.fr
31	<i>Cryptophyllium celebicum</i>	F. Tetaert, phasmatodea.fr
32	<i>Spinohirasea bengalensis</i>	F. Tetaert, phasmatodea.fr
33	<i>Paramenexenus laetus</i>	F. Tetaert, phasmatodea.fr
34	<i>Phaenopharos herwardeni</i>	F. Tetaert, phasmatodea.fr
35	<i>Lopaphus sphalerus</i>	F. Tetaert, phasmatodea.fr
36	<i>Diesbachia tamyris</i>	F. Tetaert, phasmatodea.fr
37	<i>Anarchodes annulipes</i>	F. Tetaert, phasmatodea.fr
38	<i>Trachythorax maculicollis</i>	F. Tetaert, phasmatodea.fr
39	<i>Sipyloidea biplagiata</i>	F. Tetaert, phasmatodea.fr
40	<i>Carausius morosus</i>	F. Tetaert, phasmatodea.fr
41	<i>Lonchodes auriculatus</i>	F. Tetaert, phasmatodea.fr
42	<i>Manduria systropeidon</i>	F. Tetaert, phasmatodea.fr
43	<i>Eurycantha calcarata</i>	F. Tetaert, phasmatodea.fr
44	<i>Lamponius guerini</i>	F. Tetaert, phasmatodea.fr
45	<i>Hypocyrtus scythrus</i>	F. Tetaert, phasmatodea.fr
46	<i>Venuperodes venustula</i>	F. Tetaert, phasmatodea.fr
47	<i>Bactrododema hecticum</i>	F. Tetaert, phasmatodea.fr
48	<i>Diapherodes martinicensis</i>	F. Tetaert, phasmatodea.fr
49	<i>Pterinoxylus crassus</i>	F. Tetaert, phasmatodea.fr
50	<i>Pharnacia ponderosa</i>	F. Tetaert, phasmatodea.fr
51	<i>Phobaeticus serratipes</i>	F. Tetaert, phasmatodea.fr
52	<i>Phryganistria heusii</i>	F. Tetaert, phasmatodea.fr
53	<i>Parapachymorpha spiniger</i>	F. Tetaert, phasmatodea.fr
54	<i>Ramulus thaii</i>	F. Tetaert, phasmatodea.fr
55	<i>Rhamporphasma spinicorne</i>	F. Tetaert, phasmatodea.fr
56	<i>Medauroidea extradentata</i>	F. Tetaert, phasmatodea.fr
57	<i>Clonaria conformans</i>	F. Tetaert, phasmatodea.fr
58	<i>Sceptrorhasma hispidulum</i>	F. Tetaert, phasmatodea.fr
59	<i>Macrophasma biroi</i>	Hennemann, F. H., & Conle, O. V. (2006). Studies on New Guinean giant stick-insects of the tribe Stephanacridini Günther, 1953, with the descriptions of a new genus and three new species of Stephanacris Redtenbacher, 1908 (Phasmatodea: "Anareolatae"). Zootaxa, 1283(1), 1-24.
60	<i>Tropidoderus childrenii</i>	F. Tetaert, phasmatodea.fr
61	<i>Eurycnema osiris</i>	F. Tetaert, phasmatodea.fr
62	<i>Extatosoma tiaratum</i>	F. Tetaert, phasmatodea.fr
63	<i>Eurycnema goliath</i>	F. Tetaert, phasmatodea.fr
64	<i>Phasma gigas</i>	F. Tetaert, phasmatodea.fr
65	<i>Onchestus rentzi</i>	F. Tetaert, phasmatodea.fr
66	<i>Acrophylla wuelfingi</i>	F. Tetaert, phasmatodea.fr
67	<i>Megacrania phelaus</i>	F. Tetaert, phasmatodea.fr
68	<i>Rhaphiderus spiniger</i>	F. Tetaert, phasmatodea.fr
69	<i>Monandroptera acanthomera</i>	F. Tetaert, phasmatodea.fr
70	<i>Asprenas brunneri</i>	F. Tetaert, phasmatodea.fr
71	<i>Acanthoxyla geisovii</i>	F. Tetaert, phasmatodea.fr
72	<i>Carlius fecundus</i>	B. Kneubühler, phasmatodea.com

CHAPTER 4

Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf mimicking insect

Boisseau, R. P., Büscher, T. H., Klawitter, L. J., Gorb, S. N., Emlen, D. J. and Tobalske, B. W. (2022). Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect. *BMC Ecology and Evolution* **22**, 1-18

Multi-modal locomotor costs favor smaller males in a sexually dimorphic leaf-mimicking insect

Romain P. Boisseau^{1,*}, Thies H. Büscher², Lexi J. Klawitter¹, Stanislav N. Gorb², Douglas J. Emlen¹, Bret W. Tobalske¹

¹ Division of Biological Sciences, University of Montana, 32 Campus Dr, Missoula, MT 59812, United States of America

² Functional Morphology and Biomechanics, Zoological Institute, Kiel University, Am Botanischen Garten 9, D-24098 Kiel, Germany

*Corresponding author: Romain P. Boisseau

Keywords: Phasmatodea | scramble competition | flight | adhesion | Computational fluid | dynamics

Abstract

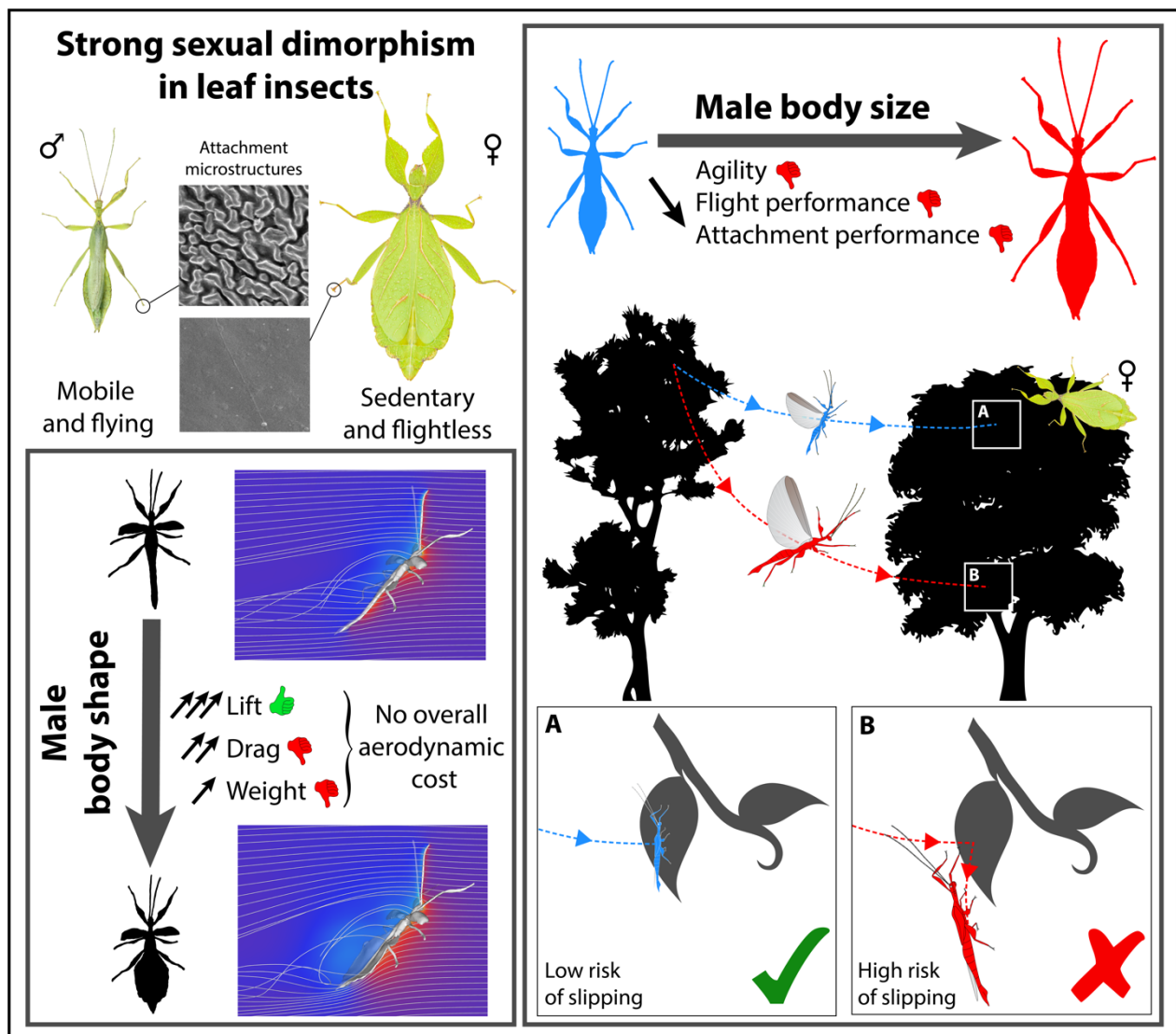
Background: In most arthropods, adult females are larger than males, and male competition is a race to quickly locate and mate with scattered females (scramble competition polygyny). Variation in body size among males may confer advantages that depend on context. Smaller males may be favored due to more efficient locomotion leading to higher mobility during mate searching. Alternatively, larger males may benefit from increased speed and higher survivorship. While the relationship between male body size and mobility has been investigated in several systems, how different aspects of male body morphology specifically affect their locomotor performance in different contexts is often unclear.

Results: Using a combination of empirical measures of flight performance and modelling of body aerodynamics, we show that large body size impairs flight performance in male leaf insects (*Phyllium philippinicum*), a species where relatively small and skinny males fly through the canopy in search of large sedentary females. Smaller males were more agile in the air and ascended more rapidly during flight. Our models further predicted that variation in body shape would affect body lift and drag but suggested that flight costs may not explain the evolution of strong sexual dimorphism in body shape in this species. Finally, empirical measurements of substrate adhesion and subsequent modelling of

landing impact forces suggested that smaller males had a lower risk of detaching from the substrates on which they walk and land.

Conclusions: By showing that male body size impairs their flight and substrate adhesion performance, we provide support to the hypothesis that smaller scrambling males benefit from an increased locomotor performance and shed light on the evolution of sexual dimorphism in scramble competition mating systems.

Graphical abstract



Introduction

Sexual dimorphism is the ultimate result of sex-dependent selection leading traits toward different optima in each sex (Blanckenhorn, 2005; Cox and Calsbeek, 2009; Fairbairn et al., 2007). Many key hypotheses for the ultimate drivers of sexual dimorphism, and specifically sexual size dimorphism (SSD), remain controversial or relatively poorly supported (Fairbairn et al., 2007). While factors favoring larger bodies are widely recognized –i.e., fecundity selection in females (Honěk, 1993; Pincheira-Donoso and Hunt, 2017) and sexual selection in males (Andersson, 1994) – the selective pressures favoring smaller sizes have received less attention (Blanckenhorn, 2000; Blanckenhorn, 2005; Fairbairn et al., 2007; Ghiselin, 1974; Kingsolver and Pfennig, 2004). In resource or female defense mating systems, the largest, most armored males are often favored against rivals over access to mating (Andersson, 1994; Emlen, 2014; Emlen and Oring, 1977; Hardy and Briffa, 2013; Shuker and Simmons, 2014). However, when females are sedentary and dispersed throughout the landscape and do not rely on an easily defensible resource, male competition unfolds instead as a race to locate females (i.e., scramble competition polygyny) (Herberstein et al., 2017). In this context, selection is predicted to favor male traits that increase the distance travelled during mate searching (i.e., mobility)(Herberstein et al., 2017).

Small and slender bodies are often assumed to be beneficial for the searching performance and mating success of scrambling males, as studies have shown in several systems (Blanckenhorn, 2000; Blanckenhorn et al., 1995; Boisseau et al., 2020; Ghiselin, 1974; Kelly, 2020; Kelly et al., 2008) – *e.g.*, by increasing endurance and enabling longer searching times, eventually leading to higher encounter rates with females (Husak and Fox, 2008). For instance, in aerial species, wider and heavier bodies, which produce more aerodynamic drag and have a higher wing loading (i.e., $\frac{\text{body weight}}{\text{wing area}}$), are usually thought to reduce flight performance as they require relatively more power to be maintained in the air, particularly during hovering or slow forward flight (Biewener and Patek, 2018). The “Ghiselin-Reiss small-male hypothesis” specifically suggests that males are usually smaller than females in scrambling systems because of the time allocation trade-off between foraging versus mate searching (Ghiselin, 1974; Reiss, 1989). Because small males require less energy to fuel their activities, they can spend relatively more time searching for mates and less time feeding. Such a trade-off has been shown empirically in water striders where smaller males have a greater mating success when food is limited (Blanckenhorn et al., 1995). However, directional selection for smaller sizes does not appear universal

among scrambling systems (Herberstein et al., 2017). In some species, it is the larger males that are more mobile and more successful than smaller counterparts, possibly owing to higher survivorship during mate searching, higher movement speeds or larger, more effective sensory structures (Barry, 2013; Hanks et al., 1996). Identifying how different aspects of locomotion are affected by body morphology is therefore critical to understand the variation of the effect of body size and shape on mobility. In the present study, we tested the hypothesis that smaller flying males have multiple locomotor advantages during mate searching in a scramble-competition insect species with pronounced sexual dimorphism. We also investigated the role of body shape *per se* in male flight performance to test the related hypothesis that males benefit from being slenderer in addition to being smaller.

We quantified the effects of variation in morphology (body size and shape) on flight and substrate attachment performance in male leaf insects (*Phyllium philippinicum*, Hennemann, Conle, Gottardo & Bresseel, 2009, Phylliidae, Phasmatodea). In this solitary canopy-dwelling species, large, sedentary adult females are outstanding leaf-mimics due to lateral ‘leaf-like’ expansions of the abdominal segments and legs (Bank et al., 2021; Bradler and Buckley, 2018; Hennemann et al., 2009). Adult males are nine times lighter and almost two times slenderer than females, and have relatively longer antennae (i.e., strong size [SSD] and shape dimorphism; **Figure 1**). Adult females lack hindwings but have extended forewings that lie flat on their dorsum, aiding in camouflage (Hennemann et al., 2009). These wings cannot flap but can contribute to parachuting if falling. In contrast, adult males have rudimentary forewings and long, fully-developed transparent hindwings allowing flapping flight. Males use their long antennae to detect pheromones and actively search for sedentary females widely scattered in the canopy (Missbach et al., 2014). Although the natural history of these nocturnal insects is largely unknown in the field, phylliid males are often observed flying to light traps (Srinivasan et al., 2017) and are very active at night in captivity – i.e., climbing on branches and readily taking off when reaching the top of a stem (Cumming et al., 2020a; Joly, 1871; Leigh, 1909; St Quentin, 1908). In our lab cultures, *P. philippinicum* males quickly initiate copulation after finding a female, which typically lasts for more than 3 hours, and then stay on the back of the female for the following day before resuming a period of high nocturnal activity. In contrast, females are sedentary, moving only to adjacent leaves to feed. These observations are consistent with a male-searching scramble competition mating system, and clearly suggest that mobility and specifically flight and substrate attachment performance (when walking and/or landing) may be critical for male fitness (Herberstein et al., 2017). The complex 3D

structure of the high canopy environment made of foliage, branches and substantial gaps (Zeng et al., 2015; Zeng et al., 2020a) and its aerodynamic characteristics –e.g., wind gusts (McCay, 2003)– likely select for greater agility and ascending performance in the air as well as a strong tarsal attachment performance to safely hold on to a wide variety of substrates when walking or landing.

We predicted that increases in body weight would impair flight performance (agility and ascending) (Maginnis, 2006; Zeng et al., 2020b), and reduce the ability of these insects to hold on to branches or leaves when they walk or land, increasing the risk of crashing and falling from high perches. We also hypothesized that selection for efficient locomotion in males (higher lift:drag ratios, lower power requirements) in parallel with fecundity selection in females (Pincheira-Donoso and Hunt, 2017), could account for the extensive sexual dimorphism in body shape (the “skinniness” of males relative to females). As these insects fly at high body angles of attack, we predicted that wide ‘leaf-like’ body silhouettes would have proportionally greater drag and, hence, a lower lift to drag ratio, requiring relatively more mechanical work for flying and putting wider males at a locomotor disadvantage. We tested these predictions using an integrative approach combining scaling of gross morphology, micro-structure descriptions, empirical measures of flight and attachment, and modeling of flight costs and landing forces.

Results

Small males were more agile and climbed faster during flight than larger males. Compared to females, male *P. philippinicum* are shorter, lighter, skinnier and have much longer antennae (Figure 1, S1-2). The large, leaf mimicking females are incapable of flight. For males, wing loading ($\frac{\text{body weight}}{\text{wing area}}$) increased significantly with body length (BL) (Figure S2H, Table S1) but relative flight muscle mass did not (Figure S2I, Table S1), indicating that large males were not compensating for their relatively heavier weight by building disproportionately large wings or flight muscles. This resulted in a substantial reduction in flight performance, which we empirically assessed using high-speed (500fps) video recordings of flight trajectories (Figure S3, Video S1-3).

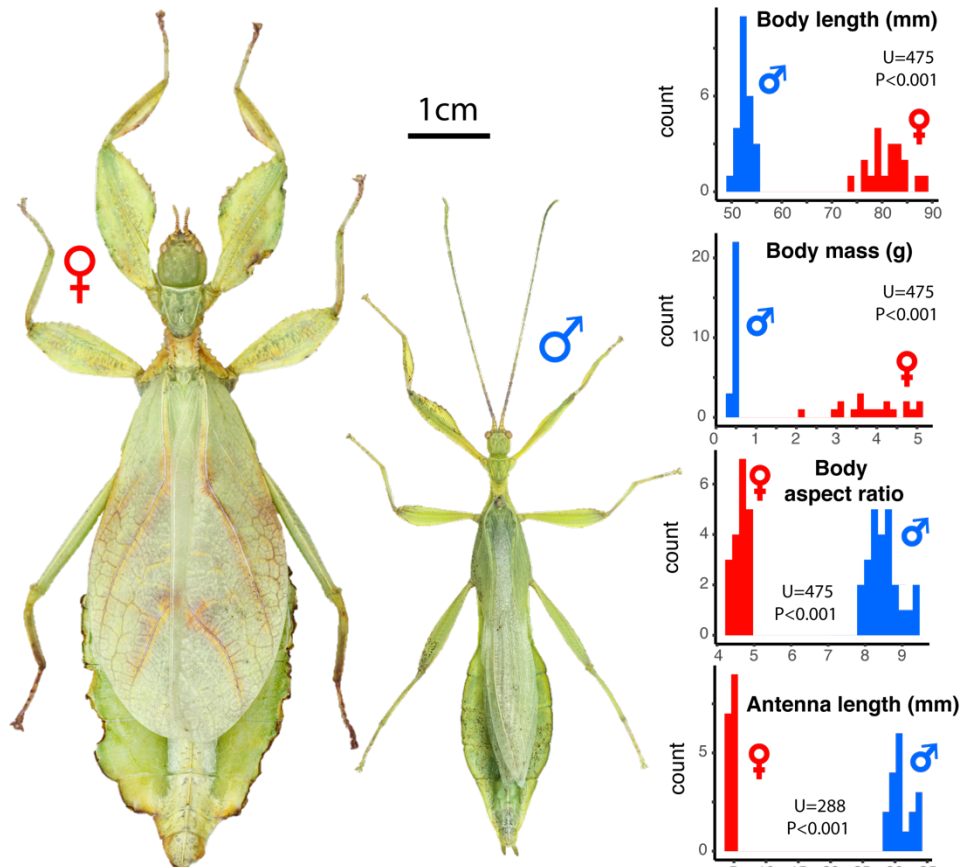


Figure 1: Sexual dimorphism in *P. philippinicum*. Pictures display an adult female (left) and male (right) in dorsal view. Distributions of male (blue) and female (red) body length, mass, aspect ratio and antenna length are shown on the right. Wilcoxon-Mann-Whitney tests are presented to compare the two sexes.

We first measured the rotational velocity of the insect long-axis body angle (i.e., pitch) when recovering from being dropped (see methods, Figure S4B)(Su et al., 2020). The rotation occurred over multiple wingbeats (Video S1-S2), so we interpret it to represent torsional agility, an aspect of aerial maneuverability that reflects how fast the animal can correct its body pitch in the air from a free falling, head first, position to a stable body pitch (Dudley, 2000; Dudley, 2002). This is distinct from oscillation in body pitch within and among wingbeats which may reflect longitudinal instability and a lack of control (Taylor and Thomas, 2003). We also measured mean horizontal and vertical velocity of the body center of mass to quantify the capacity of the individual to fly forward and ascend. Body mass and wing area had significant opposing effects on rotational velocity and on mean vertical velocity (Table 1). Consistently, wing loading negatively affected rotational velocity ($\chi^2=13.0$, $df=1$, $p=0.0003$, Figure 2B) and mean vertical velocity ($\chi^2=11.4$, $df=1$, $p=0.0007$, Figure 2C). Thus, lighter males

with relatively larger wings --i.e., with a lower wing loading -- were both more adept at changing body pitch angle in the air and had a greater capacity for ascending flight than heavier males with relatively smaller wings (**Figure 2A**). Despite flapping their wings at a higher frequency (**Table 1, Figure 2D**), larger males also decreased stroke amplitude (**Table 1, Figure 2E**). The kinematics of large males did not permit them to compensate their weight by flying faster horizontally and significantly increased their negative vertical (i.e., sinking) speed (Figure S5, **Table 1**). Thus, large males were impaired in their agility and ability to climb.

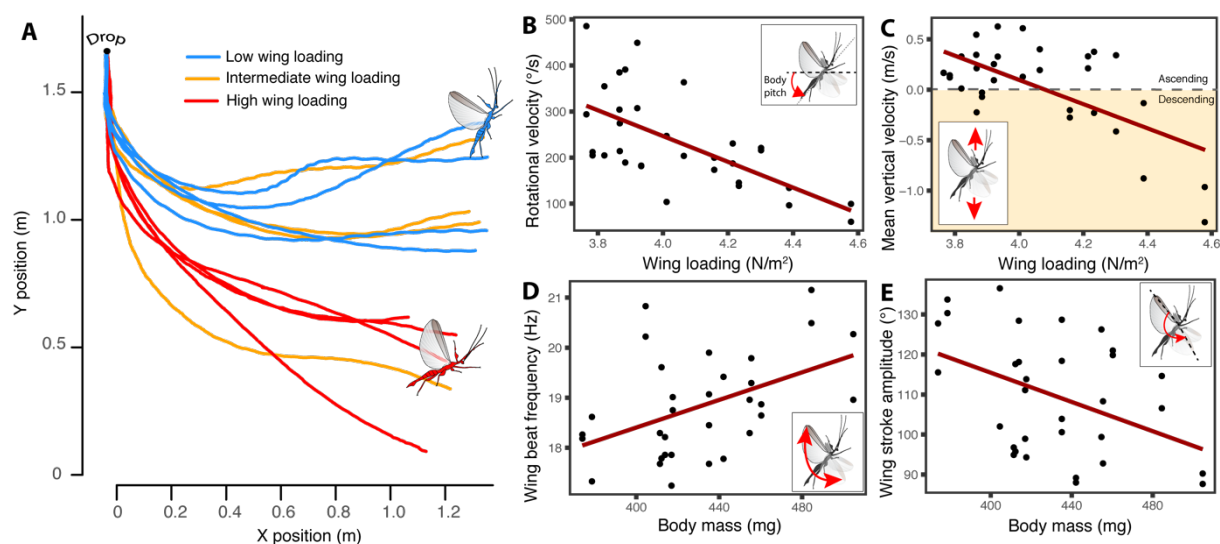


Figure 2: Effect of body mass and wing loading on male *P. philippinicum* flight performance. (A) Flight trajectories of a subset of males with low ($< 4 \text{ N}\cdot\text{m}^{-2}$), intermediate ($4 < \text{WL} < 4.22 \text{ N}\cdot\text{m}^{-2}$) and high ($> 4.22 \text{ N}\cdot\text{m}^{-2}$) wing loading. Rotational velocity (B) and mean vertical velocity (C) as a function of wing loading. Wingbeat frequency (D) and wing stroke amplitude (E) as a function of body mass. Mixed effect linear regressions including individual ID as a random factor are fitted in B, C, D and E. Each dot represent a flight trial (two per individual).

Small males attach to substrates better than larger males, reducing their risk of falling. Safely attaching to leaves and branches with their tarsi is critical for canopy-dwelling leaf insects. Females spend most of their time hanging from the undersides of leaves and therefore rely on friction attachment forces, which resists tarsal movement along (parallel to) the surface of the leaf, and adhesion forces, which resists falling from the underside of the leaf (i.e., perpendicular to the leaf surface). Male attachment performance also depends on both friction and adhesion forces, but males walk greater distances through the canopy in their search for females, requiring them to attach to a wider variety of plant surfaces (e.g., leaves, branches, trunks), and flying males must overcome impact forces to hold on to branches or leaves when they land.

Table 1: Analyses of the effects of body mass, wing area and body aspect ratio on various components of flight performance. The table shows a summary of linear mixed effect model outputs including individual ID as a random factor (N = 32 trials, 16 individuals, **Figure 2**). Fixed effects were mean-centered and standardized. Outputs include the estimated parameter value (\pm standard error) and a type-I likelihood ratio test to investigate the significance of each fixed effect sequentially. Fixed effects that were found to have a significant effect ($p < 0.05$) are bolded.

Response variable	Fixed effects	$\beta \pm SE$	Df	χ^2	p
Rotational velocity ($^{\circ}.s^{-1}$)	Body mass	-82.7 \pm 25.2	1	7.04	0.008
	Wing area	53.6 \pm 22.9	1	5.99	0.014
	Body aspect ratio	13.4 \pm 19.3	1	0.63	0.43
Mean vertical velocity ($m.s^{-1}$)	Body mass	-0.37 \pm 0.12	1	7.49	0.006
	Wing area	0.22 \pm 0.10	1	4.99	0.025
	Body aspect ratio	0.05 \pm 0.09	1	0.40	0.53
Mean horizontal velocity ($m.s^{-1}$)	Body mass	0.23 \pm 0.17	1	1.56	0.21
	Wing area	-0.10 \pm 0.15	1	0.49	0.49
	Body aspect ratio	0.07 \pm 0.13	1	0.44	0.51
Wing beat frequency (Hz)	Body mass	0.51 \pm 0.29	1	4.99	0.03
	Wing area	0.26 \pm 0.26	1	1.20	0.27
	Body aspect ratio	0.43 \pm 0.22	1	4.34	0.04
Wing beat amplitude ($^{\circ}$)	Body mass	-2.59 \pm 4.37	1	4.55	0.03
	Wing area	-6.12 \pm 3.97	1	2.93	0.09
	Body aspect ratio	-0.91 \pm 3.36	1	0.10	0.75

Friction (parallel to the substrate) and adhesion (perpendicular) forces are the product of maximum tarsal pad frictional or adhesive strength and pad area. Tarsi of male and female *P. philippinicum* are similar in overall morphology. Both sexes have five-segmented tarsi, each of the tarsomeres equipped with a euplantula (i.e., “heel” pads), and an arolium (i.e., “toe” pads) and two claws on the pretarsus (**Figure 3A**), as is typical for phasmids (Büscher et al., 2018a; Büscher et al., 2018b). The primary region of tarsal friction is the euplantula, and for adhesion is the arolium (Labonte and Federle, 2013). The overall size of these tarsal pads scales isometrically with BL and does not differ between sexes after accounting for size differences (**Table 2, Figure 3D**). However, scanning electron microscopy (SEM) revealed that the attachment microstructures on the euplantulae are sexually dimorphic. The euplantulae of females are smooth, without cuticular microstructures (**Figure 3B**), while in males this surface is covered with maze-like arrangements of ridges (**Figure 3C; sensu** (Büscher and Gorb, 2019; Büscher et al., 2018a; Büscher et al., 2018b)). These cuticular microstructures are likely to perform better, on average, on a broad range of substrate surfaces males experience from

active searching. In contrast, smooth euplantulae are specifically adapted to smooth substrates such as the surface of smooth leaves (Büscher and Gorb, 2019; Bußhardt et al., 2012).

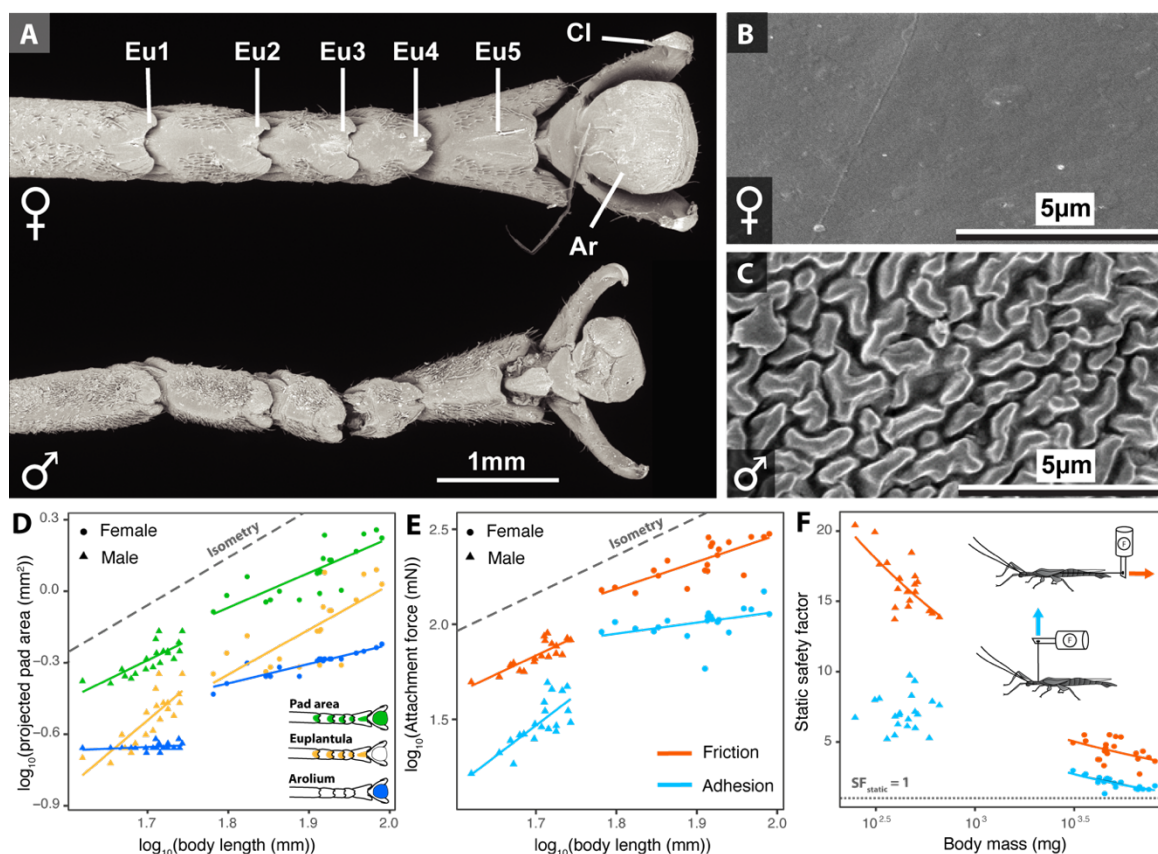


Figure 3: Sexual dimorphism in tarsal morphology and attachment forces in *P. philippicum*. (A) Morphology of female (top) and male (bottom) metathoracic tarsi. Eu. euplantula, Cl, claw, Ar, arolium. Euplantular attachment microstructures are smooth in females (B) and bear maze-like cuticular patterns in males (C). (D) Scaling relationships between body length and adhesive pad areas. (E) Friction and adhesion forces as a function of body length. Dots represent maximum force measurements averaged for each individual over three trials. Dashed lines in (D) and (E) represent the slope expected under isometry (slope = 2, arbitrary intercept). (F) Static safety factor as a function of body mass. The x-axis is on a \log_{10} scale. The dotted line in (F) represent a safety factor of one. Diagrams represent the experimental set-up to measure friction (orange) and adhesion (blue) forces.

To test for effects of body size on attachment performance, we measured maximal tarsal friction and adhesion using a force transducer mounted on a motorized micromanipulator. Static safety factors (SF_{static}) were calculated from measures of the force required to pull the insect horizontally off a glass plate (friction forces), or backwards off the plate (adhesion forces), divided by weight (see methods). This approximates how many times leaf insects can attach their own weight to smooth canopy surfaces when they are resting or hanging motionless, which is what females do most of the

time. Larger males displayed higher adhesion and friction attachment forces than smaller ones. This positive correlation between attachment force and BL was also found in females, but only when considering friction forces (**Figure 3E, Table 2**). Females also had relatively higher adhesion and friction attachment forces than males (**Figure 3E, Table 2**). However, females, with their large, leaf-mimicking bodies and mainly motionless behavior, had the lowest static safety factors (**Figure 3F, Table 2**). As they were much lighter, males had relatively larger friction SF_{static} than females (**Figure 3F, Table 2**), consistent with the greater attachment demands they experience from active searching and, especially, landing. While males displayed relatively higher adhesion SF_{static} , this difference was not significant ($p = 0.068$, **Figure 3F, Table 2**). In females, adhesion and friction SF_{static} negatively correlated with body mass, while in males, only friction SF_{static} significantly decreased with increasing body mass (**Figure 3F, Table 2**). Therefore, larger males are at higher risk of slipping off smooth substrates when resting or walking but the relatively high values of SF_{static} (>14) question the ecological relevance of such decreased attachment capacity.

Flying males, in contrast to females, experience additional attachment demands when they land, due to impact forces. Consequently, to investigate the potential order of magnitude of these forces and how they may affect safety factors, we built a model to predict the typical impact forces experienced by males (F_i) when landing on a leaf. This model computed dynamic safety factors ($SF_{dynamic} = \frac{F_{friction}}{F_i + Body\ weight}$) based on landing speed and leaf deflection estimated from body mass. Landing platforms (i.e., leaves) were modelled as cantilever beams spanning a range of size and flexural stiffness (reused from other studies (Higham et al., 2017)) as, in the field, landing platforms encountered by flying males during mate search may display extremely variable mechanical properties. Our goal was to estimate potential extreme values of impact force. Only frictional attachment forces (as opposed to adhesive forces) were used to compute these $SF_{dynamic}$ as they are the most important for accommodating deceleration and impact forces when landing (Higham et al., 2017).

Our landing model (**Figure 4A**) predicted heavier males to cause larger leaf deflections when landing as estimated by equation 3 (**Figure 4A,B**). Leaf deflection was greatest for the largest leaves (**Figure 4B**). We empirically found, in our flight experiments, that overall landing velocity was positively correlated with male body mass (Figure S6, likelihood ratio test: $\chi^2 = 4.12$, $df = 1$, $p = 0.042$). This is due to heavier males dropping faster as horizontal velocity was not significantly correlated with body mass (**Table 1**). In our landing model, we then used the fixed effect estimates of the

corresponding linear mixed model to predict the landing velocity of males given their body mass and estimate impact forces at landing (**Figure 4A**). Landing impact forces was predicted to increase with male body mass and to be relatively higher for smaller leaves (**Figure 4C, Table 3**). Finally, the model estimated that $SF_{dynamic}$ should decrease with male body mass and should be relatively lower for smaller leaves (**Figure 4D, Table 3**). Interestingly, predicted $SF_{dynamic}$ fell below 1.0 for body masses > 350 mg and small leaves (**Figure 4D**), therefore predicting slippage. In other words, large males were predicted to be more likely to slip and fall when landing on canopy substrates, especially on smaller and stiffer leaves.

Table 2: Scaling relationships between body size and attachment structures and forces. The table presents results of type I ANOVA from linear models contrasting the differences between sexes in terms of scaling relationships between body size and attachment pad areas, attachment forces and static safety factors (N = 20 males and 20 females, **Figure 3**). Scaling exponents β and the corresponding 95% confidence intervals are shown in comparison to isometric expectations. Significant effects (i.e., $p < 0.05$) are bolded.

Response variable (log ₁₀ transformed)	Explanatory variables	F	df1	df2	P	Isometric slope	Slope β [95% CI]
Pad morphology							
Arolium area (mm ²)	log₁₀(body length)	3460	1	36	<0.001	2	Males: 0.15 [-0.07, 0.37]
	sex	245.2	1	36	<0.001		Females: 0.81[0.63, 0.99]
	interaction	19.0	1	36	<0.001		
Euplantula area (mm ²)	log₁₀(body length)	225.6	1	36	<0.001	2	Males: 2.86 [1.87, 3.86]
	sex	0.42	1	36	0.52		Females: 1.88 [1.00, 2.77]
	interaction	1.81	1	36	0.19		
Combined pad area (mm ²)	log₁₀(body length)	483.0	1	36	<0.001	2	Males: 1.65 [1.01, 2.28]
	sex	2.41	1	36	0.13		Females: 1.46 [0.90, 2.02]
	interaction	0.17	1	36	0.68		
Attachment forces							
Adhesion forces (mN)	log₁₀(body length)	444.4	1	36	<0.001	2	Males: 3.22 [1.93, 4.51]
	sex	24.3	1	36	<0.001		Females: 0.58[-0.03, 1.18]
	interaction	16.5	1	36	<0.001		
Friction forces (mN)	log₁₀(body length)	803.4	1	36	<0.001	2	Males: 1.98 [1.34, 2.63]
	sex	18.3	1	36	<0.001		Females: 1.41 [0.86, 1.97]
	interaction	1.54	1	36	0.22		
Static adhesion safety factor	log₁₀(body mass)	562.2	1	36	<0.001	0	Males: 0.06 [-0.3, 0.43]
	sex	3.5	1	36	0.068		Females: -0.58 [-0.85, -0.31]
	interaction	9.11	1	36	0.005		
Static friction safety factor	log₁₀(body mass)	980.9	1	36	<0.001	0	Males: -0.34 [-0.51, -0.17]
	sex	5.31	1	36	0.027		Females: -0.33 [-0.64, -0.02]
	interaction	0.008	1	36	0.93		

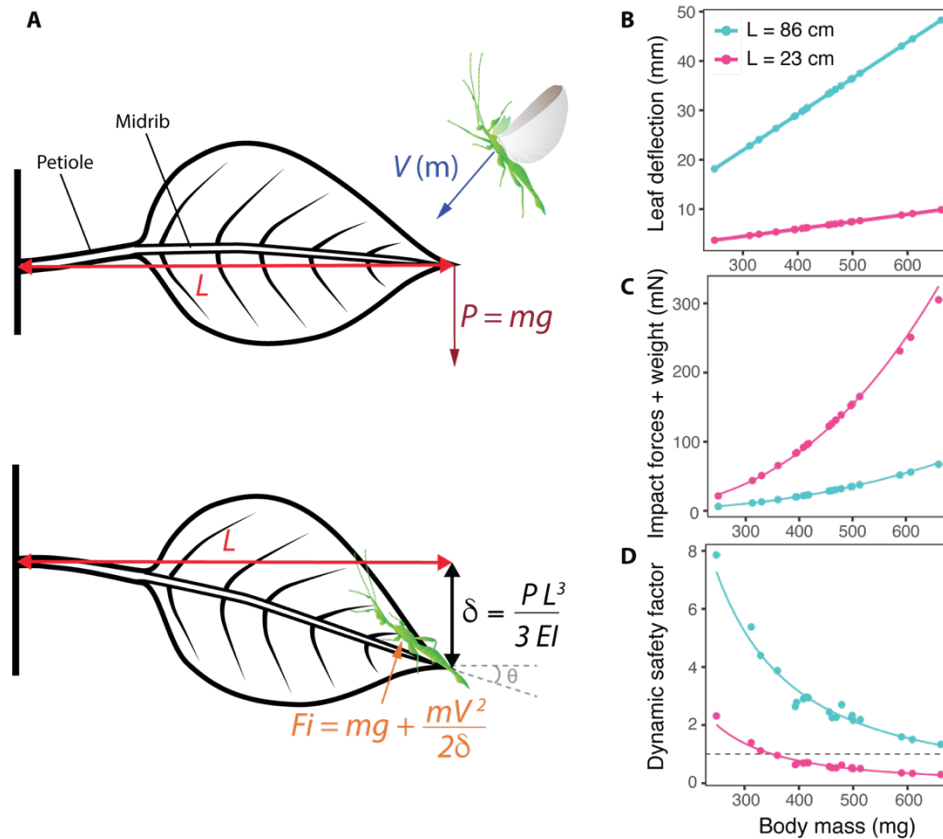


Figure 4: Landing model calculations and predictions. **(A)** Diagram of a male leaf insect landing on a leaf, including calculations used to estimate impact forces (F_i) and dynamic safety factors ($SF_{dynamic}$)— i.e., the ratio of frictional attachment forces and F_i . The male lands with a landing velocity V which depends on the body mass of the insect (m). The terminal load on the leaf (P) will be equal to the weight of the insect – i.e., body mass (m) multiplied by the acceleration due to gravity (g). The subsequent deflection of the leaf (δ) depends on P , the length of the leaf (L) and the leaf flexural stiffness (EI). The impact force experienced by the male leaf insect (F_i) will be determined by P , the kinetic energy of the insect ($mV^2/2$) and δ . See the text for more detailed information. Predicted relationships between male body mass (m) and δ **(B)**, F_i **(C)** and $SF_{dynamic}$ **(D)**. Predictions for large leaves ($L=86$ cm) are represented in blue, small leaves ($L=23$ cm) in pink. The dashed line in **(D)** represents a safety factor of one.

Table 3: Tests of the effect of body mass and leaf size on estimated (modelled) landing impact forces and dynamic safety factors in males ($N=20$ males, **Figure 4**). Type-I ANCOVAs were performed. Scaling exponents β and the corresponding 95% confidence intervals are shown for small and large leaves separately. Significant effects (i.e., $p < 0.05$) are bolded.

Response variable (\log_{10} transformed)	Explanatory variables	F	df1	df2	P	Slope β [95% CI]
Impact forces + weight (mN)	$\log_{10}(\text{body mass})$	20541	1	36	<0.001	Small leaf: 2.67 [2.6, 2.74]
	leaf size	31459	1	36	<0.001	Large leaf: 2.41 [2.38, 2.43]
	interaction	55.7	1	36	<0.001	
Dynamic friction safety factor	$\log_{10}(\text{body mass})$	892	1	36	<0.001	Small leaf: -2.02 [-2.21, -1.82]
	leaf size	2485	1	36	<0.001	Large leaf: -1.75 [-1.92, -1.57]
	interaction	4.40	1	36	0.043	

Body shape affects lift:drag ratio. To further understand how body size and shape affect aerodynamics in ways that could contribute to flight performance, we generated 3D models of the bodies of males of varying size and abdominal shape and estimated body lift and drag during steady horizontal flight using computational fluid dynamics modelling (CFD, **Figure 6A**, see methods) (Crandell et al., 2019; Goyens et al., 2015; Troelsen et al., 2019). Contrary to vertical velocity, mean horizontal velocity did not significantly correlate with body mass in our flight trials (**Table 1**). Therefore, the CFD simulations of horizontal flight were run at a constant average speed ($1.57 \text{ m}\cdot\text{s}^{-1}$). The models predicted that the male's flattened abdomen would produce a wide region of low-pressure behind the insect, the size of which largely being dependent on its shape (**Figure 5**). The males' poorly streamlined bodies would create high drag coefficients ($1.38 < C_D < 1.55$, **Figure 6C**) and only nominal lift coefficients ($0.85 < C_L < 1.03$, **Figure 5D**) resulting in relatively low lift to drag ratios ($0.56 < L/D < 0.74$, **Figure 6B**).

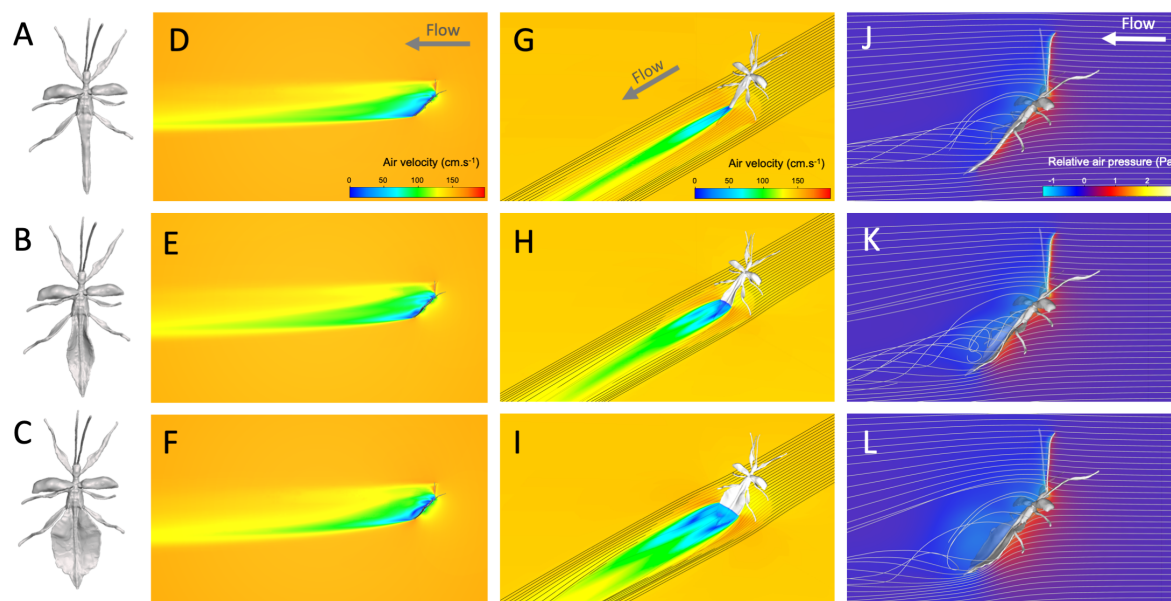


Figure 5: 3D models and computational fluid dynamics simulations. Dorsal view of phasmid models with no (**A**), natural (**B**) and widened (**C**) 'leaf-like' abdominal expansions. Air velocity in the mid-sagittal plane (**D-F**), air velocity and streamlines in a horizontal transverse plane (**G-I**), pressure (relative to ambient air static pressure) and streamlines in the mid-sagittal plane (**J-L**) around the bodies of males with no (**D, G, J**), natural (**E, H, K**) and widened (**F, I, L**) abdominal expansions.

Our simulations further suggested that flying males with wider abdomens generate more lift relative to drag, have a lower C_D , and a higher C_L (**Figure 6B-D**, **Table 4**). In contrast, body size did not significantly affect these parameters (**Table 4**). The drag and lift forces applied by the air on the

animal's body (expressed as proportion of body weight) were estimated to be higher in smaller males with wider bodies (**Figure 6E, Table 4**). Interestingly, body lift and drag were predicted to decrease at similar rates with body aspect ratio, suggesting that the gain in body lift provided by wider bodies (i.e., helping in weight support) may be offset by the gain in body drag (i.e., opposing the movement). Thus, contrary to our predictions, our models suggest that abdominal shape may not significantly affect the cost of flight and that selection for flight efficiency may not explain the relatively slenderer body shape of males in this species.

Table 4: The effects of body size and shape on various aerodynamic variables, as predicted by the CFD models. Results of type I ANOVA from linear models contrasting the effects of body length and body aspect ratio on various aerodynamic variables (**Figure 6**). Significant effects (i.e., $p < 0.05$) are bolded.

Response variable	Explanatory variables	$\beta \pm SE$	F	df1	df2	P
Lift to drag ratio (L/D)	Body length	0.001±0.0007	2.72	1	22	0.11
	Body aspect ratio	-0.02±0.001	221.7	1	22	<0.001
Drag coefficient (C_D)	Body length	-0.001±0.0008	2.42	1	22	0.13
	Body aspect ratio	0.015±0.001	103.1	1	22	<0.001
Lift coefficient (C_L)	Body length	0.001±0.001	0.60	1	22	0.45
	Body aspect ratio	-0.02±0.002	68.3	1	22	<0.001
Relative drag	log₁₀(body length)	-0.499±0.049	103.0	1	22	<0.001
	log₁₀(body aspect ratio)	-1.058±0.092	131.5	1	22	<0.001
Relative lift	log₁₀(body length)	-0.295±0.055	28.6	1	22	<0.001
	log₁₀(body aspect ratio)	-1.245±0.104	144.6	1	22	<0.001

The power required for flight increases with size faster than the available muscle power. Flight performance depends on the power available (the maximum amount of mechanical energy that can be provided by the flight muscles per unit of time, P_a) and on the power required for flight (the amount of mechanical energy required to fly per unit of time, P_r) (Ellington, 1991). We theoretically estimated the scaling exponents of these two variables with BL using our empirical data to uncover how the difference between them (ΔP), which represents the excess power available for demanding aerial activities, varies with size. We estimated that P_a increases with body size with a scaling exponent $\beta_{P_a} = 2$, as expected under isometry. In contrast, P_r increases with body size with a scaling exponent $\beta_{P_r} = 5.5$ (95% CI: [4, 7]) where isometry predicts $\beta = 3.5$ (Equation 5). Consequently, ΔP ($P_a - P_r$) decreases with body size more rapidly than would be expected under isometry, hypothetically accounting for the reduced flight climbing ability and maneuverability seen in larger males. Combined, our results suggest that selection for both flight and landing/attachment performance may help explain the relatively small size of males in this species.

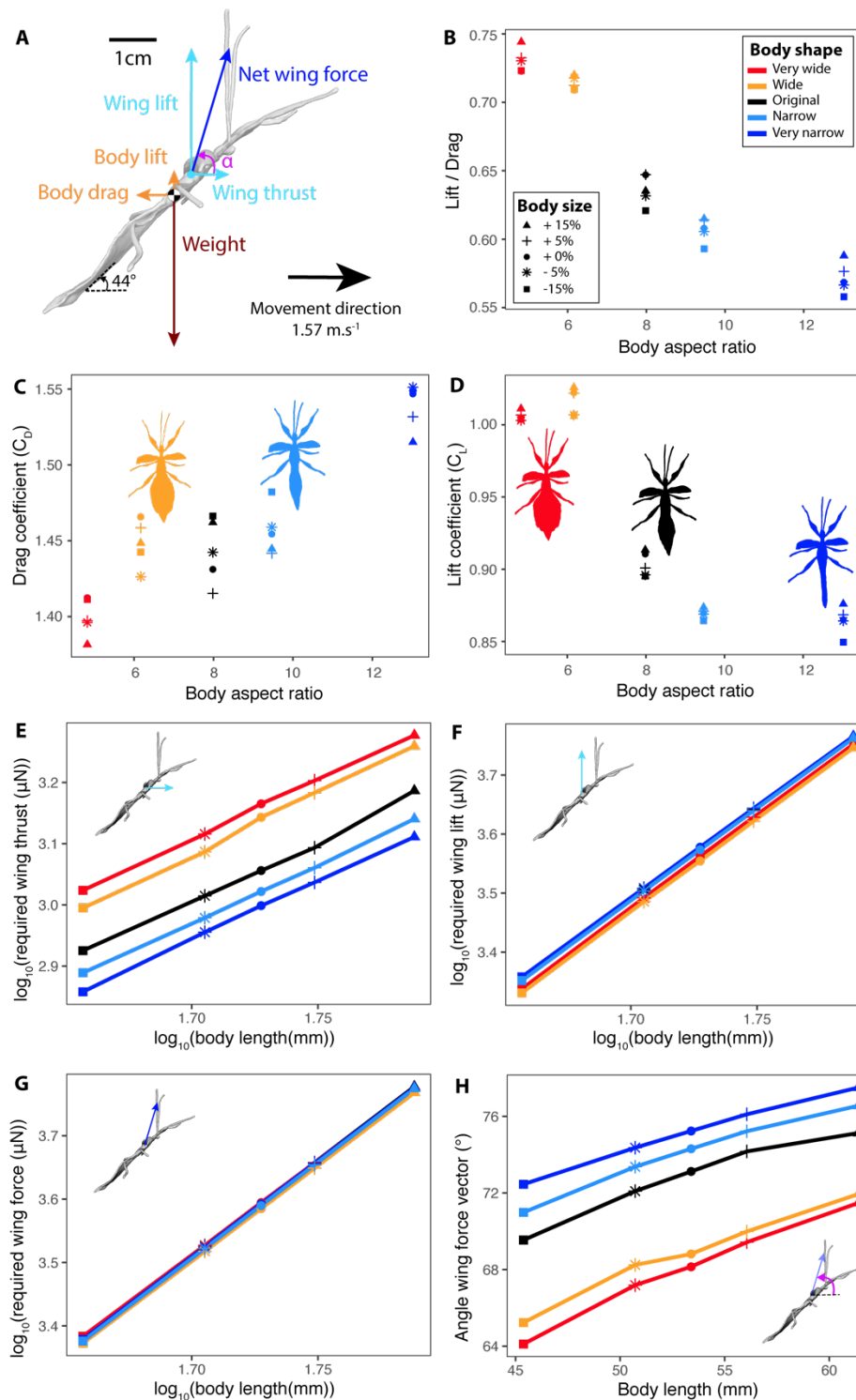


Figure 6: The effects of body shape on body aerodynamics as predicted by computational fluid dynamics simulations. **(A)** Diagram of the forces applying upon the modelled flying male leaf insect. Lift to drag ratio **(B)**, drag **(C)** and lift **(D)** coefficients, drag **(E, solid)** and lift **(E, dashed)** forces relative to body weight as a function of model aspect ratio as estimated from the CFD simulations. Colors correspond to different body sizes relative to the original model.

Discussion

As leaf masqueraders, leaf insects (Phylliidae) display some of the most extreme abdominal morphologies of the insect world, and sexual size and shape dimorphisms so spectacular that taxonomists have had difficulty associating males and females of the same species (Cumming et al., 2020a). Using these organisms, we provide support for the hypothesis that selection for locomotor performance favors small body sizes in males, contributing to the evolution of extreme sexual dimorphism in insects with scramble competition mating systems.

We show that small males have greater agility in the air, as they were able to stabilize their body angle faster after falling (Dudley, 2002), and they were better able to climb in flight than larger males. Flight is most likely essential to the mate searching performance of males, as females are sedentary and typically scattered in the rainforest canopy (Herberstein et al., 2017). A decreased agility in the air and a lower ability to maintain a horizontal trajectory or ascend during flight may be detrimental in that regard. It should however be noted that pitch control is only one aspect of maneuvering; roll and yaw likely also contribute (Cheng et al., 2016). Our momentum-jet analysis suggested that the mechanical power required for these males to fly steadily and horizontally (P_r) should increase at a faster rate than the power available from the flight muscles (P_a), and even faster than expected under isometry. The reduction of flight performance (agility and climb ability) with body size seen in leaf insects is therefore likely to result from the decrease of $\Delta P (=P_a - P_r)$, which represents the excess power available for more demanding aerial activities such as maneuvers and climbing in air (Ellington, 1991; Norberg and Norberg, 2012).

Not only were large males relatively poor flyers, they also were more likely to detach and fall from smooth substrates. Our measurements of attachment forces showed that static safety factors for friction forces decreased with body size in males but remained high (>14), questioning the ecological relevance of this decrease. However, the attachment capacity of males is likely to get challenged at the landing, during which friction forces are fundamental to accommodate impact forces (Higham et al., 2017). Our landing model predicted that the dynamic safety factors of the largest males were likely to come close or even below one (i.e., slippage) when landing on small and stiff leaves. However, it should be noted that the material properties of the landing platforms encountered by males in the field are still unknown. Therefore, the significance of the decreased attachment performance of large males, remains to be tested in the field. High attachment safety factors are crucial for canopy insects to avoid

falling, which can lead to predator exposure (Dudley et al., 2007; Yanoviak et al., 2005; Yanoviak et al., 2011) and energetic costs to return to the canopy (Zeng et al., 2020a). For searching males, it can also mean losing track of a female. We found that male leaf insects have specialized, ridged surfaces on the pads of their tarsi which are not present in females (**Figure 2B-C**), and which are probably adapted to gripping a broad range of plant surfaces (Büscher and Gorb, 2019; Büscher et al., 2019; Bußhardt et al., 2012; Gorb and Scherge, 2000; Varenberg and Gorb, 2007; Varenberg and Gorb, 2009). Flying males are likely to be confronted with unpredictable surfaces (e.g., branches, leaves) when walking and landing, and may benefit from generalist tarsal pads that adhere securely to a range of textured surfaces. In contrast, females move very little in the canopy and are strongly associated with the leaves of their food plants. Females likely use their claws and attachment pads on a narrower range of surfaces and for a narrower range of tasks -- primarily as anchors as they hang upside down from smooth leaf cuticle. In this context, adhesion (as opposed to friction) forces are essential. In stick insects, arolia are shear-sensitive attachment pads providing most of the adhesion (Büscher and Gorb, 2019; Labonte and Federle, 2013). Consistently, female leaf insects have larger arolia relative to body size than males and consequently produce larger adhesion forces. Nevertheless, as females weigh much more than males, their static safety factors for both adhesion and friction forces were still relatively lower than those of males.

Male and female leaf insects also exhibit a spectacular interspecific variation in body shape related to leaf mimicry (Bank et al., 2021; Cumming et al., 2021; Hennemann et al., 2009). This variation is likely driven by masquerade and the advergence of the insect appearance to resemble the size and shape of its host plants' leaves. Thus, predation is likely a major selective agent acting on male and female morphology that may greatly constrain or even oppose selection on male morphology imposed by locomotion, notably selection favoring smaller body sizes. As exemplified by *P. philippinicum*, flight-capable males have more elongated body shapes than their respective flightless females and relatively reduced abdominal leaf-like lobes. Our CFD models coupled with scaling of body weight predicted that despite the poorly streamlined body shapes of males and their high drag coefficients ($C_D > 1$), any increases in body drag due to wider body shapes would be offset by increased body lift. This leads us to predict a similar net cost of flight for thin or wide males. Therefore, flight performance may not constrain male body shape the way that it does body size, and may have provided male leaf insects the freedom to evolve a variety of body shapes (Brock et al., 2020; Conle et al., 2008; Cumming et al., 2018; Cumming et al., 2020b; Cumming et al., 2020c). However, given that body shape still significantly

affected the aerodynamic forces applying on the insect body, it is possible that it affects other relevant components of aerial performance beyond the scope of this study such as maneuverability and stability, or the net cost of flight in other contexts like ascending flight.

Our empirical and modelling approaches were designed to test how different aspects of leaf insect locomotor performance scaled with body size and shape in multiple relevant contexts. However, our models included a number of simplifying assumptions which we acknowledge, and future studies may wish to expand their focus accordingly. For instance, we chose a constant and average leg posture, flight direction (horizontal), flight speed and angle of attack in our CFD models, and we ignored the complexity of wing aerodynamics, all of which are likely relevant to the aerodynamic of the insect in flight. Similarly, our estimates of impact forces at landing are simplified as, for example, we did not account for the leg joint movement that may further reduce the impact force. More detailed modelling approaches and empirical tests of the predictions from our models (e.g., by measuring aerodynamic forces and slippage probability when landing) are worthwhile goals for future, more specific, studies.

Conclusion

In summary, our analyses suggest that large scrambling males are at a locomotor disadvantage in at least two aspects of mate searching in leaf insects: they have lower agility and ascending flight abilities likely stemming from lower power margins to fly, and they are at a higher risk of falling when landing on surfaces in search of rare and scattered females. Therefore we offer support for the hypothesis that large scrambling males suffer multiple locomotor costs in several critical aspects of mate searching. Our findings help to shed light on the repeated evolution of relatively small males in scramble competition mating systems by explicitly considering locomotor performance as an intermediate step between morphology and mobility and mating success.

Materials and Methods

Study animals. A first breeding population of *P. philippinicum* was obtained from the Audubon Insectarium in New Orleans, Louisiana, USA and shipped to the University of Montana, Missoula, Montana, USA. The insects were housed in a transparent plastic container (50x40x60cm) at 22°C, on 12h:12h light:dark cycles, sprayed with water daily (RH = 50 – 80%), and fed fresh *Rubus idaeus* leaves *ad libitum*. This population was used to investigate morphological scaling relationships and conduct the flight experiments described below.

A second culture stock of *P. philippinicum* was obtained from Kirsten Weibert (Jena, Germany) and captive bred in the department of Functional Morphology and Biomechanics at Kiel University, Germany. The specimens were kept in a large glass cage with proper ventilation at 20-22°C (RH = 50-80%; 16h:8h light:dark cycle) and fed with fresh blackberry (*Rubus* sp.) and common oak (*Quercus robur* L.) leaves *ad libitum*. This population was used to investigate attachment pad morphology and attachment forces.

Sexual dimorphism and scaling relationships. Photographs of the animals in dorsal view were taken using a DSLR camera (EOS 600D, Canon Inc., Tokyo, Japan). Using ImageJ software (v.1.52k) (Schneider et al., 2012), we measured body length (BL, mm), body area (mm²), body circularity (dimensionless, $\frac{4\pi \times Area}{Perimeter^2}$), body aspect ratio (dimensionless, $\frac{Body\ length}{Average\ body\ width}$), mean antenna length (mm), mean front femur length (mm), and total wing area (mm², including both forewings for females or both hindwings for males) in 25 adult males and 19 adult females (Figure S1). Antenna length could not be measured in seven males and three females because they were missing flagellomeres on both antennae. Wet body mass (g, measured immediately after flight trials) was obtained using an analytical balance (ME54TE/00, Mettler Toledo, Columbus, OH, USA). We calculated wing loading (N.m⁻²) as wet body mass multiplied by gravitational acceleration ($g=9.81m.s^{-2}$) and divided by total wing area. Male flight muscle mass (mg) was obtained by dissecting the muscles out of the metathorax of freshly dead males (n= 23) drying them at 70°C for 24h and weighing them with a more accurate analytical balance (UMT2, Mettler Toledo, Columbus, OH, USA).

Male flight performance. To evaluate male flight performance, here defined as a righting maneuver that required sustained control over long-axis body angle and climb ability, we dropped adult males in the air and recorded their flight trajectories in 2D (Video S1-2). Working in Missoula, MT (elevation

978 meters above sea level, average air density = $1.07 \text{ kg}\cdot\text{m}^{-3}$), adult males ($N=16$, $0.43 \pm 0.006 \text{ g}$) from our American culture population were held by the thorax and dropped by the experimenter at a horizontal body angle from a constant height above a floor (1.5m). A 2m-tall slab of ponderosa pine (*Pinus ponderosa*) with natural bark was placed vertically in front of the animal, 2m away, to serve as a target and landing site (Figure S3). The experimental room was largely featureless with white walls and at 26°C . The floor was covered with thick blankets to avoid injuries if crashing. We recorded the flight of the insects using a high-speed video camera (Photron FASTCAM SA-3, Photron USA, San Diego, CA, USA) sampling at 500 fps with a shutter speed of $1/5000\text{s}$ and 1024×1024 pixel resolution (Photron PFV v.3.20). Because the insects were induced to fly in a trajectory parallel to the plane of the imaging sensor (deviations $< 10^\circ$), we analyzed the trajectories in two dimensions (Figure S3, S4A). Pixels were scaled to metric coordinates using a 50cm bar held horizontally at the same level of the flight trajectories. We recorded a clay ball in free fall to calibrate the vertical direction. Average vertical ball acceleration was $9.816 \pm 0.14 \text{ m}\cdot\text{s}^{-2}$, less than 1% different from gravitational acceleration ($9.805 \text{ m}\cdot\text{s}^{-2}$) and the camera was oriented to gravity so that the vertical ball drop direction of acceleration was always $< 5\%$ of 90° . Each insect was dropped several times (maximum five times) until we obtained two straight trajectories per animal. A resting period of approximately 20min was left in between each flight to allow recovery. Males were frozen-killed at -80°C just after the experiments. After estimating the relative position of their center of mass (see below), males were pinned with their wings fully extended for morphometrics.

Video digitization was done by tracking morphological landmarks using the open source video analysis tool *DLTdv5* by T. Hedrick implemented in MATLAB (R2016b, MathWorks, Natick, MA, USA) (Hedrick, 2008) and the obtained data was analyzed using R (v 3.6.1)(R Core Team, 2019). In *DLTdv5*, we used autotracking mode (predictor tool: extended Kalman) and manual tracking when the autotracking mode was unreliable. We marked the position of the head and of the terminal abdominal segment on each frame. Body pitch ($^\circ$) was calculated for every frame by calculating the angle between the horizontal and the line linking the position of the head and that of the terminal segment (Figure S4B). Typically, in the first phase of the fall (free fall), body pitch decreases (i.e., the insect rotates forward, eventually diving head first), before the insect opens its wings (t_1) and actively corrects (t_2) and stabilizes its body pitch (phase 3) (Figure S4B). Body pitch was smoothed using a Savitzky-Golay filter with a polynomial order of 3 and a window size of 71 (*sgolayfilt*: ‘signal’, *function*: ‘R package’). The beginning of phase 2 (t_1) was determined as the time corresponding to the minimal body pitch (Figure

S4B). The end of phase 2 (t_2) corresponded to the time when the insect's pitch stabilized – *i.e.*, when the rotational velocity ($^{\circ} \cdot s^{-1}$) of body pitch reached a local minimum after a large peak corresponding to phase 2 (Figure S4B). We calculated the average rotational velocity during phase 2 (ω) as:

$$\textbf{Equation 1: } \omega = \frac{\textit{Body pitch}(t_2) - \textit{Body pitch}(t_1)}{t_2 - t_1}$$

We used ω to quantify torsional agility -- an important aspect of maneuverability -- as it reflects how fast the animal can rotate to correct its body pitch in the air from a free falling, head first, position to a stable flight body pitch. This correction occurred over several wingbeats and was therefore distinct from within and among-wingbeat oscillations which could have indicated a lack of longitudinal stability (Taylor and Thomas, 2003). The 2D position of the body center of mass was estimated using images taken in lateral view of the males (freshly dead) orthogonally balancing on a horizontal razorblade. Its position relative to the two landmarks (*i.e.*, head and terminal segment) was then calculated which enabled us to define the center of mass of the individual on each flight video. Trajectories of the center of mass were analyzed using the package “trajr” in R (McLean and Skowron Volponi, 2018). Raw trajectories were smoothed using a Savitzky-Golay filter with a polynomial order of 3 and a window size of 31 (*TrajSmoothSG*:‘trajr’; Figure S4A). Horizontal, vertical and composite velocities and accelerations were then computed on the smoothed trajectories (*TrajDerivatives*:‘trajr’; Figure S4C-E). For each trial, we defined transient and steady states for both vertical and horizontal velocities (Figure S4D-E). In both cases, the transient state corresponded to the free fall and maneuver of the insect in the air during which body velocity greatly varied. The steady state started when velocity stabilized and acceleration started oscillating around 0 $m \cdot s^{-2}$ (Figure S4D-E). We extracted the mean vertical and horizontal velocity ($m \cdot s^{-1}$) during their respective steady states. These measures were used to quantify the capacity of the insect to fly forward and ascend. On the videos, the position of the tip of the wing closest to the camera was also manually marked on frames corresponding to the end of upstroke and downstroke as this was sufficient to measure wing beat frequency (Hz) and stroke amplitude ($^{\circ}$). Average wing beat frequency was calculated after the animal reached a stable body pitch (*i.e.*, after t_2 ; Figure S4B). For each animal, we measured wing length and the position of the attachment of the wings relative to the head and tip of the abdomen using photographs and ImageJ. We then determined the position of the wing attachment point on each frame on the videos using this relative position between our two landmarks. The amplitude of wing strokes was calculated using trigonometry (Figure S7).

Adhesive pads and substrate attachment performance. We used scanning electron microscopy (SEM) to observe the tarsi of the metathoracic leg of adult males and females, measure attachment pad areas, and describe the microstructures on these pads. Tarsi of the right metathoracic leg were cut off from 20 adult males and 20 adult females and fixed in 2.5 % glutaraldehyde in PBS buffer for 24 h on ice on a shaker, dried in an ascending alcohol series, critical-point dried and sputter-coated with a 10nm layer of gold-palladium. To obtain overview images, we used a rotatable specimen holder (Pohl, 2010) and the scanning electron microscope (SEM) Hitachi TM3000 (Hitachi High-technologies Corp., Tokyo, Japan). The micrographs for visualization and measurements were taken at an acceleration voltage of 15kV. The attachment microstructures on the tarsi of both sexes were further examined using the SEM Hitachi S4800 (Hitachi High-Technologies Corp., Tokyo, Japan) at 7 kV of acceleration voltage. Processing of the raw micrographs and measurements of projected attachment pad area (mm²) -- i.e., the surface area of the tarsus specialized for adhesion and friction (Bullock et al., 2008; Labonte et al., 2016) -- were done using Photoshop CS6 (Adobe Systems Inc., San José, CA, USA).

To measure attachment forces (mN) in both pull-off (adhesion) and traction (friction) directions, we used 20 adult males (Mean \pm S.D.= 0.46 \pm 0.02g) and 20 adult females (5.22 \pm 0.31g). A horsehair was glued to the metanotum of each insect and then attached to a 100g force transducer (FORT100, World Precision Instruments, Sarasota, USA, linearity error: <0.1%, resolution: 0.01%), connected to a BIOPAC model MP100 and TCI-102 system (BIOPAC Systems, Inc., Goleta, CA, USA), and mounted on a motorized micromanipulator (DC 3001R, World Precision Instruments Inc.). Maximum adhesion forces were recorded by vertically pulling the insects off a horizontal glass plate until they detached from the glass plate (Büscher and Gorb, 2019; Wohlfart et al., 2014). The micromanipulator was moved upwards with a speed of 200 μ m/s at a step size of 10 μ m until the specimen was detached from the surface as indicated by an instantaneous drop in force. Maximum friction forces were recorded by horizontally pulling the insects backwards with the same retraction velocities as above, until detaching them from the glass plate (Büscher and Gorb, 2019; Wolff and Gorb, 2012). A glass plate was used as the substrate for the attachment force measurements, to eliminate mechanical interlocking of the claws with surface irregularities of rough substrates, or penetration of soft substrates. As glass is smooth on the microscopical level, this substrate enables estimation of the traction and pull-off performance of the attachment pads themselves on a standardized level, without the influence of substrate irregularities. Force-time curves were obtained

using Acqknowledge 3.7.0 (BIOPAC Systems Inc., Goleta, CA, USA) and the maximum peaks were extracted as maximum adhesion forces, or maximum friction forces respectively. Each of the 20 males and females were measured three times in both directions on a glass plate and the average of the three measurements was used as the individual maximum adhesion/friction force. The order of the individuals was randomized and the substrate was cleaned between every measurement. The experiments were conducted at 20-23°C and 50-60% relative humidity.

Static safety factors (*i.e.*, $SF_{\text{static}} = \frac{\text{Attachment force}}{\text{Body weight}}$) were computed for each individual.

Following the methods of Higham et al., 2017 (Higham et al., 2017), we estimated impact forces (F_i , in N) during landing using a model to eventually compute dynamic safety factors (*i.e.*, $SF_{\text{dynamic}} = \frac{\text{Friction force}}{F_i}$). In our landing model, we assumed that males would stop immediately after landing on the tip of a leaf without slippage. F_i was calculated using the work-energy principle:

$$\text{Equation 2: } F_i = mg + \frac{mv^2}{2\delta}$$

where m = mass of the insect (kg), g = acceleration of gravity ($9.81 \text{ m}\cdot\text{s}^{-2}$), v = landing speed of the insect ($\text{m}\cdot\text{s}^{-1}$) and δ = deflection of the leaf (m) (**Figure 4**).

We used a simplified model of a leaf to estimate a range of values for the deflection of a leaf upon landing of a male leaf insect and explore potential values of SF_{dynamic} . The leaf and petiole were considered as a uniform cantilever beam with impact forces applying at the tip at length L . We considered the leaf to be initially horizontal. Deflection was calculated as a function of the weight of the insect. Given the relatively light weight of male leaf insects, and in contrast with (Higham et al., 2017), who were considering geckos (*i.e.*, roughly 28 times heavier), we only accounted for small leaf deflections (deflection angle $\theta < 5^\circ$) and therefore used the following equation to estimate the deflection of the tip of the leaf δ (Goodno and Gere, 2018):

$$\text{Equation 3: } \delta = \frac{mgL^3}{3EI}$$

where L is the total length of the leaf (m) and EI is its flexural stiffness ($\text{N}\cdot\text{m}^2$). We used the range of values estimated by (Higham et al., 2017), for leaf length and corresponding EI (*i.e.*, $L_1 = 23 \text{ cm}$, $EI_1 = 2.67 \times 10^{-3} \text{ N}\cdot\text{m}^2$ and $L_2 = 86 \text{ cm}$, $EI_2 = 28.48 \times 10^{-3} \text{ N}\cdot\text{m}^2$) to account for the diversity of leaf mechanical properties and thus explore the possible range of SF_{dynamic} .

Finally, landing speed was estimated as a function of the body mass of the individual. Using our experimental flight videos (see above), we extracted the instantaneous speed of the experimental males right before making contact with the wood slab (i.e., the landing target) and built a linear regression between landing speed and body mass, including individual ID as a random factor (*lmer*: “lme4”, Bates et al., 2015). We found a significant positive effect of body mass on landing speed (see results, Figure S6) and used the fixed-effect estimates (intercept and slope) from this model to predict the landing speed of the males for which we empirically measured attachment forces, given their body mass. This estimated landing speed was used in equation 2.

Computational fluid dynamic simulations. To investigate how lift and drag produced by the insect were affected by size and shape, we generated 3D models of males of varying size and shape and predicted these forces during a steady and horizontal flight using computational fluid dynamics (CFD). We created a reference 3D surface model of an adult male body using photogrammetry. We pinned the body of a dead specimen in a flight posture (i.e., legs extended, forewings opened perpendicularly, antennae oriented 50° up, hindwings removed). We did not model flapping aerodynamics as fully integrating the complex three dimensional trajectories and aeroelastic deformations of the wings into a CFD model (e.g., (Young et al., 2009)) was beyond the scope of this study. The aerodynamic interactions of the flapping wings with the body of insects in slow flight has been shown to be negligible (~5%) in (Liang and Sun, 2013). Once dried, the individual was vertically mounted onto a pin on a custom-made turntable. 2D images using a DSLR camera (EOS 600D, Canon Inc., Tokyo, Japan) equipped with a macro lens (Canon EF 100mm f/2.8 Macro USM), were then obtained from 100 different orientations (Figure S8). The 3D model was then reconstructed from these multiple images using Autodesk ReCap Pro 2019 (v5.0.4.17, Autodesk Inc., San Rafael, CA, USA) and subsequently smoothed and rewrapped using Autodesk Meshmixer 2017 (v11.5.474, Solid accuracy: 402, cell size: 0.202, density: 219, offset: 0.25, min thickness: 0.14mm). From this reference model, we built four additional and artificial models using the “Move” tool in Meshmixer to either manually extend or shrink the abdominal lobes and therefore manipulate abdominal shape. These artificial shapes purposely spanned a wider range of body aspect ratios than the one found in actual males *P. philippinicum* (male natural range: 4.89 - 6.42, female natural range: 2.28 - 2.84, model range: 2.28 - 9.47). The model with the lowest aspect ratio displayed a female-like abdominal shape while the model with the highest aspect ratio had no abdominal expansions. The models were further scaled to a body length of 53 mm (mean male BL in our Montana population = 52.6 +/- 0.25mm) using Autodesk fusion 360

(v2.0.8335). From each of these five meshes, we created four additional models (25 models in total) respectively scaled to a factor 0.85, 0.95, 1.05 and 1.15. The insect models were tilted at a 44° body pitch in our control volume. This angle was determined using our flight experiments and a LMM with stable body pitch as the response variable, horizontal and vertical body velocity as main effects and individual ID as a random factor. Using the parameters estimated by this model, we predicted a body pitch of 43.9° for a vertical velocity of zero and a mean horizontal velocity of 1.57 m.s⁻¹ (i.e., the average horizontal velocity calculated from our flight trials, after the animal had stabilized its body pitch: 157 +/- 6.8 cm.s⁻¹).

We constructed a control volume around these body meshes in Autodesk CFD 2019 (v19.2), that provided numerical solutions to the Reynolds-averaged Navier-Stokes equations (Rahman, 2017; Troelsen et al., 2019). A fluid volume was built around the mesh with walls far enough from the model mesh to avoid any reflection effects (1x0.5x0.5m) (Figure S9A). The fluid was assigned the default properties of air in CFD 2019 (density at sea level = 1.205 kg.m⁻³, viscosity at 20°C = 18.2 μPa.s⁻¹). The phasmid models were then assigned properties of hardwood which, for mass-less and stationary models, should have no impact on results. The input flow on the anterior end of the control volume was set to 1.57 m.s⁻¹. We held the air velocity around in the insect constant as we were only considering horizontal flight and average horizontal velocity after reaching a steady state did not significantly increase with size in our flight trials (**Table 1**, Figure S5A). We applied a zero-pressure condition on the opposing end of the volume. A slip/symmetry condition was applied to all other fluid boundaries. We automatically meshed the domain around the phasmid model, applied a surface refinement, and locally defined a non-uniform mesh refinement region (0.7x0.2x0.2m) with a mesh size reduced to 75%, around and behind the model to better capture the resulting wake. We ran steady-state simulations using the turbulence model k-epsilon. The maximum number of iterations was set to 3,000 although the simulations were stopped when they reached convergence according to the default convergence detection parameters of the CFD software (mean = 849 ± 50 interactions). The adaptive meshing tool was used to insure mesh optimization for our models and mesh independence of the results. The simulation was first run with the meshing parameters described previously. Then, the solution results of this simulation were automatically used to refine the mesh in high velocity gradient regions and rerun the simulation. We enabled the ‘flow angularity’ option to improve mesh resolution in areas with a lot of flow separation, the ‘free shear layers’ and ‘external flow’ options to refine the mesh in areas of strong velocity gradients. We ran 3 such cycles for each model. Final mesh sizes

averaged $865,446 \pm 74,771$ nodes and $4,411,596 \pm 379,780$ elements (Figure S9). Finally, to help evaluate the validity of our simulations, we placed a sphere with the same Reynold's number as the one calculated for the original phasmid model ($Re = 5558$) in a similar control volume with the exact same settings as our insect simulations. We found a drag coefficient (C_d) of 0.643. This is very close to the value predicted from experimental data ($C_d = 0.652$) and which was determined using the equation 8.83 in Morrison, 2013.

The weight of the models (mN) with a non-modified abdominal shape (reference models) was determined from their BL using the linear regression built between male body mass and BL in our American population. To estimate the weight of the models with artificial abdominal shapes, we first measured their abdominal area relative to that of the reference model of identical BL and measured the average weight of the leaf-like abdominal expansions by cutting these extensions from 5 freshly frozen-killed males and measuring their areas and mass ($89.7 \pm 0.6 \text{ g.m}^{-2}$). The CFD simulations estimated the aerodynamic forces (drag and lift, mN) that apply to the rigid insect body flying horizontally and steadily.

For each model, we measured the projected frontal area on a plane perpendicular to the air flow. We then calculated their coefficient of drag ($C_D = \frac{2 F_{drag}}{\rho v^2 A}$) and lift ($C_L = \frac{2 F_{lift}}{\rho v^2 A}$) using the model frontal area (A), the mass density of air ($= 1.20473 \text{ kg.m}^{-3}$), the velocity of the insect ($v = 1.57 \text{ m.s}^{-1}$) and the drag or lift force estimated from the CFD simulations. Lift to drag ratios ($\frac{C_L}{C_D}$) were also calculated for each model.

Scaling of muscle power available and power required for flight. From theory, we estimated the scaling relationships of the power available (P_a) and the power required for flight (P_r) with body size using our empirical data. As leaf insects are slow flyers (average flight speed= $1.70 \pm 0.07 \text{ m.s}^{-1}$, mean \pm SE), P_r mostly corresponds to the induced power P_{ind} – i.e., the cost for producing lift (Biewener and Patek, 2018; Ellington, 1991). P_{ind} is the product of the net required force from the wings to maintain the animal in the air and of the induced velocity in the wake. Following (Epting and Casey, 1973), we assumed that, under isometry, the induced velocity in the wake and weight-specific power required for slow flight should be proportional to the square root of wing disc loading (DL, N.m^{-2}) -- i.e., body weight (W_B) divided by wing disc area (A_{WD}). Wing disc area --i.e., the area swept out by the

wing during a wing beat cycle and through which air is accelerated downward to develop lift force-- is determined by wing length (L_w) and stroke amplitude (θ , °) (Equation 4).

$$\text{Equation 4: } DL = \frac{W_B}{A_{WD}} = \frac{W_B}{2\pi L_w^2 \frac{\theta}{360}}$$

Thus, among geometrically similar animals, the power required for slow flight should scale as $BL^{7/2}$ (Equation 5).

$$\text{Equation 5: } P_r \propto W_B (DL)^{1/2} \propto W_B \left(\frac{W_B}{2\pi L_w^2 \frac{\theta}{360}}\right)^{1/2} \propto BL^3 \left(\frac{BL^3}{BL^2}\right)^{1/2} \propto BL^{7/2}$$

In slow flight, P_a is the product of the net wing force and of the tangential velocity of the tip of the wing --i.e., angular velocity (rad.s^{-1}) \times wing length (m)-- or the product of the muscle work --i.e., force \times distance of contraction-- and of flapping frequency (Hz). We accepted the assumption that, for geometrically and dynamically similar organisms, force is proportional to the cross-sectional area of the muscles, which scales as BL^2 , and distance of contraction scales as BL^1 . Thus, work scales as BL^3 and is therefore directly proportional to W_B (Ellington, 1991; Hill, 1950). Flapping frequency is predicted to scale as BL^{-1} when the animal is using maximal or near-maximal effort (Greenewalt, 1975). Therefore, under isometry, P_a is expected to scale as BL^2 .

To estimate the scaling exponent of the tangential velocity of the tip of the wing (i.e., angular velocity (rad.s^{-1}) \times wing length (m)) and of wing disc loading (Equation 4) with L in leaf insects, we built LMMs with \log_{10} tangential velocity or \log_{10} disc loading as the response variable, $\log_{10} BL$ as the fixed effect and male ID as a random factor. Tangential velocity of the tip of the wing did not significantly scale with BL or differ from isometrical expectations ($\beta = -1$) in our flight experiments ($\beta = -2.43 \pm 1.60$, $\chi^2 = 2.46$, $df = 1$, $p = 0.12$). Similarly, the observed scaling exponents of body and flight muscle mass did not significantly differ from isometric expectations (Table S1). Coupling this with the aforementioned assumptions about muscle force, we estimated P_a scaled isometrically in leaf insects ($\beta=2$). Wing disc loading (DL) positively scaled with BL ($\chi^2 = 8.99$, $df=1$, $p=0.003$). While, under isometry, we expected a scaling exponent of 1 (Equation 4), we found that DL increased disproportionately with BL ($\beta = 4.96$, 95% CI=[1.99,7.94]). This is a consequence of the reduced wing stroke amplitude seen in larger individuals (Table 2). As the induced velocity (V_i) in the wake scales proportionately with the square root of wing disc loading, we estimated that P_r scaled more steeply with size than expected under isometry ($\beta=5.48$ VS 3.5) (Equation 6)

$$\text{Equation 6: } P_r \propto W_B (DL)^{\frac{1}{2}} \propto BL^3 (BL^{4.96})^{\frac{1}{2}} \propto BL^{5.48}$$

Consequently, ΔP ($P_a - P_r$) decreased with body size more rapidly than would be expected under isometry.

Statistical analyses. All statistical analyses were run in R version 3.6.1 (R Core Team, 2019) and all statistical tests were two-sided. For all linear models, we systematically checked the normal distribution of the residuals and the absence of any specific patterns in their distribution.

We tested for sex differences in mean BL, body mass, body aspect ratio and antenna length using Wilcoxon-Mann-Whitney tests (*wilcox.test*: ‘stats’). To test for sex differences in the scaling relationships (i.e., in slope and intercept) of the various morphological traits, attachment and aerodynamic force measurements with BL, we built ordinary least square regressions (Kilmer and Rodríguez, 2017) including each \log_{10} -transformed trait as response variable and \log_{10} BL, sex and their interaction as predictor variables (*lm*: ‘stats’). Type-I ANCOVAs were used to determine significance of the fixed effects (*anova*: ‘stats’). Departure from isometry which corresponds, on a log-log scale, to a slope of 1 for linear measurements, 2 for areas and 3 for masses (Schmidt-Nielsen, 1984), was tested using 95% confidence intervals (CI) around the estimated regression slopes (*confint*: ‘stats’). We similarly built linear models to investigate the effect of body mass, leaf size (i.e., small or large) and their interaction on the estimated landing impact forces in males and dynamic safety factors.

To test for effects of body size, wing size and body shape on male flight performance, we built linear mixed models (LMM) (*lmer*: ‘lme4’). Response variables were either rotational velocity (ω), mean vertical or horizontal velocity, wingbeat frequency, or wing stroke amplitude. Body mass, wing area and body aspect ratio were mean-centered and standardized ($\mu = 0$ and $\sigma = 1$, *scale*: “base”) and were included as main fixed effects. Individual ID was included as a random factor to account for replications of each individual. Likelihood ratio tests were subsequently performed sequentially to assess the significance of the fixed effects (*anova*: ‘lme4’). For response variables significantly affected by body mass and wing area and to illustrate their combined effect, we built and plotted similar LMMs but with wing loading as the only fixed effect.

Following our CFD simulations, we tested for the effects of body size and shape on lift to drag ratio, C_D , C_L , and relative aerodynamic forces applying on the body using linear models including BL and body aspect ratio as explanatory variables. Sequential ANOVAs (type I) were subsequently

performed to assess significance (*anova*: 'stats'). In models including relative force estimates as response variables, variables were \log_{10} -transformed to compute scaling exponents (Schmidt-Nielsen, 1984).

List of abbreviations

SSD: Sexual size dimorphism

BL: Body length

ω : Body pitch rotational velocity

SEM: Scanning electron microscopy

SF_{static} : Static safety factors

F_i : Impact forces

SF_{dynamic} : Dynamic safety factors

CFD: Computational fluid dynamics

C_D: Drag coefficient

C_L: Lift coefficient

L/D: Lift to drag ratio

P_a: Power available for flight

P_r: Power required for flight

P_{ind}: Flight induced power

ΔP : Excess power available for flight ($P_a - P_r$)

δ : Deflection of the tip of the leaf upon insect landing.

EI: Flexural stiffness

m: Body mass

g: Gravitational acceleration

F_{drag} : Drag force applying on the body in flight

F_{lift} : Lift force applying on the body in flight

DL: Wing disc loading

W_B: Body weight

A_{WD}: Wing disc area

L_w: Wing length

θ : Wing stroke amplitude

V_i : induced velocity in the wake

LMM: Linear mixed model

CI: Confidence intervals

Ethics approval and consent to participate

The experimental use of leaf insect did not require ethic approval in the United States or in Germany. Leaf insects were shipped from the Audubon Insectarium in New Orleans, Louisiana, USA and housed at the University of Montana under a USDA-APHIS permit (P526P-17-03108).

Availability of data and material

Datasets and corresponding R scripts are publicly available on Figshare:

Boisseau, Romain; Büscher, Thies H.; J. Emlen, Douglas; Gorb, Stanislav N.; Tobalske, Bret W. (2021): Boisseau et al._Phyllium locomotion 2022_data and scripts.zip. figshare. Dataset.

<https://doi.org/10.6084/m9.figshare.14852904.v3>

Competing interests

The authors declare no competing interests.

Author contributions

R.P.B., T.H.B., S.N.G., D.J.E. and B.W.T. designed the study. R.P.B. and L.J.K. performed the flight experiments and subsequent video analyses. T.H.B. measured the attachment performance of the insects and produced the SEM images of the tarsi. R.P.B. performed the modelling approaches and ran the statistical analyses. R.P.B. wrote the initial manuscript, which was reviewed and edited by all authors.

Funding

We thank the NSF for funding (NSF award numbers, B. W. Tobalske: CMMI-1234737, D. J. Emlen: IOS 1456133 & 2015907).

Acknowledgements

The authors thank Art Woods for use of lab computer and space and Camille Thomas-Bulle and Anthony Lapsansky for useful discussions.

References

- Andersson, M.** (1994). *Sexual Selection*. Princeton, NJ, USA: Princeton University Press.
- Bank, S., Cumming, R. T., Li, Y., Henze, K., Le Tirant, S. and Bradler, S.** (2021). A tree of leaves: phylogeny and historical biogeography of the leaf insects (Phasmatodea: Phylliidae). *Communications Biology* **4**, 1–12.
- Barry, K. L.** (2013). You are what you eat: food limitation affects reproductive fitness in a sexually cannibalistic praying mantid. *PLoS ONE* **8**, e78164.
- Bates, D., Maechler, M., Bolker, B. and Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48.
- Biewener, A. A. and Patek, S. N.** (2018). *Animal Locomotion*. Oxford, United Kingdom: Oxford University Press.
- Blanckenhorn, W. U.** (2000). The evolution of body size: what keeps organisms small? *The Quarterly Review of Biology* **75**, 385–407.
- Blanckenhorn, W. U.** (2005). Behavioral causes and consequences of sexual size dimorphism. *Ethology* **111**, 977–1016.
- Blanckenhorn, W. U., Preziosi, R. F. and Fairbairn, D. J.** (1995). Time and energy constraints and the evolution of sexual size dimorphism - to eat or to mate? *Evolutionary Ecology* **9**, 369–381.
- Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G. and Emlen, D. J.** (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes* **181**, 104263.
- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.
- Brock, P. D., Büscher, T. and Baker, E.** (2020). Phasmida Species File Online. *Version 5.0/5.0*.
- Bullock, J. M. R., Drechsler, P. and Federle, W.** (2008). Comparison of smooth and hairy attachment pads in insects: friction, adhesion and mechanisms for direction-dependence. *Journal of Experimental Biology* **211**, 3333–3343.
- Büscher, T. H. and Gorb, S. N.** (2019). Complementary effect of attachment devices in stick insects (Phasmatodea). *Journal of Experimental Biology* **222**, jeb209833.
- Büscher, T. H., Kryuchkov, M., Katanaev, V. L. and Gorb, S. N.** (2018a). Versatility of Turing patterns potentiates rapid evolution in tarsal attachment microstructures of stick and leaf insects (Phasmatodea). *Journal of the Royal Society Interface* **15**, 20180281.
- Büscher, T. H., Buckley, T. R., Grohmann, C., Gorb, S. N. and Bradler, S.** (2018b). The evolution of tarsal adhesive microstructures in stick and leaf insects (Phasmatodea). *Frontiers in Ecology and Evolution* **6**, 1–11.

- Büscher, T. H., Grohmann, C., Bradler, S. and Gorb, S. N.** (2019). *Tarsal attachment pads in Phasmatodea (Hexapoda: Insecta)*. Vienna, Austria: Zoologica.
- Bußhardt, P., Wolf, H. and Gorb, S. N.** (2012). Adhesive and frictional properties of tarsal attachment pads in two species of stick insects (Phasmatodea) with smooth and nubby euplantulae. *Zoology* **115**, 135–141.
- Cheng, B., Tobalske, B. W., Powers, D. R., Hedrick, T. L., Wang, Y., Wethington, S. M., Chiu, G. T. C. and Deng, X.** (2016). Flight mechanics and control of escape manoeuvres in hummingbirds. I. Flight kinematics. *Journal of Experimental Biology* **219**, 3518–3531.
- Conle, O. V., Hennemann, F. H. and Perez-Gelabert, D. E.** (2008). Studies on neotropical Phasmatodea II: Revision of the genus *Malacomorpha* Rehn, 1906, with the descriptions of seven new species (Phasmatodea: Pseudophasmatidae: Pseudophasmatinae). *Zootaxa* **1748**, 1–64.
- Cox, R. M. and Calsbeek, R.** (2009). Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* **173**, 176–187.
- Crandell, K. E., Howe, R. O. and Falkingham, P. L.** (2019). Repeated evolution of drag reduction at the air–water interface in diving kingfishers. *Journal of the Royal Society Interface* **16**, 20190125.
- Cumming, R. T., Leong, J. V. and Lohman, D. J.** (2018). Leaf insects from Luzon, Philippines, with descriptions of four new species, the new genus *Pseudomicrophyllium*, and redescription of *Phyllium* (*Phyllium*) *geryon* Gray, 1843, (Phasmida: Phylliidae). *Zootaxa* **4365**, 101–131.
- Cumming, R. T., Tirant, S. Le, Teemsa, S. N., Hennemann, F. H., Willemse, L. and Büscher, T. H.** (2020a). Lost lovers linked at long last: elusive female *Nanophyllium* mystery solved after a century of being placed in a different genus (Phasmatodea, Phylliidae). *ZooKeys* **969**, 43–84.
- Cumming, R. T., Bank, S., Tirant, S. Le and Bradler, S.** (2020b). Notes on the leaf insects of the genus *Phyllium* of Sumatra and Java, Indonesia, including the description of two new species with purple coxae (Phasmatodea, Phylliidae). *ZooKeys* **913**, 89–126.
- Cumming, R. T., Thurman, J. H., Youngdale, S. and Le Tirant, S.** (2020c). *Walaphyllium* subgen. nov., the dancing leaf insects from Australia and Papua New Guinea with description of a new species (Phasmatodea, Phylliidae). *ZooKeys* **939**, 1–28.
- Cumming, R. T., Bank, S., Bresseel, J., Constant, J., Le Tirant, S., Dong, Z., Sonet, G. and Bradler, S.** (2021). *Cryptophyllium*, the hidden leaf insects - descriptions of a new leaf insect genus and thirteen species from the former celebicum species group (Phasmatodea, Phylliidae). *ZooKeys* **1018**, 1–179.
- Dudley, R.** (2000). *The biomechanics of insect flight: form, function, evolution*. Princeton, New Jersey, USA: Princeton University Press.
- Dudley, R.** (2002). Mechanisms and implications of animal flight maneuverability. *Integrative and Comparative Biology* **42**, 135–140.
- Dudley, R., Byrnes, G., Yanoviak, S. P., Borrell, B., Brown, R. M. and McGuire, J. A.** (2007). Gliding and the functional origins of flight: biomechanical novelty or necessity? *Annual Review of Ecology, Evolution, and Systematics* **38**, 179–201.
- Ellington, C. P.** (1991). Limitations on Animal Flight Performance. *Journal of Experimental Biology* **160**, 71–91.
- Emlen, D. J.** (2014). Reproductive contests and the evolution of extreme weaponry. In *The evolution of insect mating systems* (ed. Shuker, D. and Simmons, L.), pp. 92–105. Oxford, United Kingdom: Oxford University Press.
- Emlen, S. T. and Oring, L. W.** (1977). Ecology, sexual selection, and the evolution of mating systems. *Science* **197**, 215–

223.

- Epting, R. J. and Casey, T. M.** (1973). Power output and wing disc loading in hovering hummingbirds. *The American Naturalist* **107**, 761–765.
- Fairbairn, D. J., Blanckenhorn, W. U. and Székely, T.** (2007). *Sex, Size and Gender Roles*. Oxford, United Kingdom: Oxford University Press.
- Ghiselin, M.** (1974). *The economy of nature and the evolution of sex*. Berkeley, California, USA: University of California Press.
- Goodno, B. J. and Gere, J. M.** (2018). *Mechanics of Materials*. Ninth edit. Boston, Massachusetts, USA: Cengage Learning.
- Gorb, S. and Scherge, M.** (2000). Biological microtribology: anisotropy in frictional forces of orthopteran attachment pads reflects the ultrastructure of a highly deformable material. *Proceedings of the Royal Society B: Biological Sciences* **267**, 1239–1244.
- Goyens, J., Van Wassenbergh, S., Dirckx, J. and Aerts, P.** (2015). Cost of Flight and the Evolution of Stag Beetle Weaponry. *Journal of The Royal Society Interface* **12**, 20150222.
- Greenewalt, C. H.** (1975). The flight of birds: the significant dimensions, their departure from the requirements for dimensional similarity, and the effect on flight aerodynamics of that departure. *Transactions of the American Philosophical Society* **65**, 1–67.
- Hanks, L. M., Millar, J. G. and Paine, T. D.** (1996). Body size influences mating success of the eucalyptus longhorned borer (Coleoptera: Cerambycidae). *Journal of Insect Behavior* **9**, 369–382.
- Hardy, I. C. W. and Briffa, M.** (2013). *Animal Contests*. Cambridge, United Kingdom: Cambridge University press.
- Hedrick, T. L.** (2008). Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems. *Bioinspiration and Biomimetics* **3**, 034001.
- Hennemann, F. H., Conle, O. V., Gottardo, M. and Bresseel, J.** (2009). On certain species of the genus *Phyllium* Illiger, 1798, with proposals for an intra-generic systematization and the descriptions of five new species from the Philippines and Palawan (Phasmatodea: Phylliidae: Phylliinae: Phylliini). *Zootaxa* **2322**, 1–83.
- Herberstein, M. E., Painting, C. J. and Holwell, G. I.** (2017). Scramble competition polygyny in terrestrial arthropods. *Advances in the Study of Behavior* **49**, 237–295.
- Higham, T. E., Russell, A. P. and Niklas, K. J.** (2017). Leaping lizards landing on leaves: Escape-induced jumps in the rainforest canopy challenge the adhesive limits of geckos. *Journal of the Royal Society Interface* **14**, 20170156.
- Hill, A. V.** (1950). The dimensions of animals and their muscular dynamics. *Science Progress* **38**, 209–230.
- Honěk, A.** (1993). Intraspecific variation in body size and fecundity in insects : a general relationship. *Oikos* **66**, 483–492.
- Husak, J. F. and Fox, S. F.** (2008). Sexual selection on locomotor performance. *Evolutionary Ecology Research* **10**, 213–228.
- Joly, N.** (1871). Contributions à l'histoire naturelle et à l'anatomie de la mouche-feuille des Iles Seychelles: *Phyllium crurifolium*, (Audinet Serville), *Mantis sicciifolia* (Linné). *Mémoires de l'Académie des Sciences de Toulouse* **7**, 1–28.
- Kelly, C. D.** (2020). Sexual selection on size and shape in Japanese beetles (*Popillia japonica*). *Behavioral Ecology* **31**, 1073–1083.
- Kelly, C. D., Bussière, L. F. and Gwynne, D. T.** (2008). Sexual selection for male mobility in a giant insect with female-biased size dimorphism. *The American Naturalist* **172**, 417–423.

- Kilmer, J. T. and Rodríguez, R. L.** (2017). Ordinary least squares regression is indicated for studies of allometry. *Journal of Evolutionary Biology* **30**, 4–12.
- Kingsolver, J. G. and Pfennig, D. W.** (2004). Individual-level selection as a cause of Cope’s rule of phyletic size increase. *Evolution* **58**, 1608–1612.
- Labonte, D. and Federle, W.** (2013). Functionally different pads on the same foot allow control of attachment: stick insects have load-sensitive “heel” pads for friction and shear-sensitive “toe” pads for adhesion. *PLoS ONE* **8**, e81943.
- Labonte, D., Clemente, C. J., Dittrich, A., Kuo, C. Y., Crosby, A. J., Irschick, D. J. and Federle, W.** (2016). Extreme positive allometry of animal adhesive pads and the size limits of adhesion-based climbing. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 1297–1302.
- Leigh, H. R.** (1909). Preliminary account of the life-history of the leaf-insect, *Phyllium crurifolium* Serville. *Proceedings of the Zoological Society of London* **79**, 103–113.
- Liang, B. and Sun, M.** (2013). Aerodynamic Interactions Between Wing and Body of a Model Insect in Forward Flight and Maneuvers. *Journal of Bionic Engineering* **10**, 19–27.
- Maginnis, T. L.** (2006). Leg regeneration stunts wing growth and hinders flight performance in a stick insect (*Sipyloidea sipyilus*). *Proceedings: Biological Sciences* **273**, 1811–1814.
- McCay, M. G.** (2003). Winds under the rain forest canopy: the aerodynamic environment of gliding tree frogs. *Biotropica* **35**, 94–102.
- McLean, D. J. and Skowron Volponi, M. A.** (2018). trajr: An R package for characterisation of animal trajectories. *Ethology* **124**, 440–448.
- Missbach, C., Dweck, H. K. M., Vogel, H., Vilcinskas, A., Stensmyr, M. C., Hansson, B. S. and Grosse-Wilde, E.** (2014). Evolution of insect olfactory receptors. *eLife* **3**, e02115.
- Morrison, F. A.** (2013). *An introduction to fluid mechanics*. New York, New York, USA: Cambridge University press.
- Norberg, U. M. L. and Norberg, R. Å.** (2012). Scaling of wingbeat frequency with body mass in bats and limits to maximum bat size. *Journal of Experimental Biology* **215**, 711–722.
- Pincheira-Donoso, D. and Hunt, J.** (2017). Fecundity selection theory: concepts and evidence. *Biological Reviews* **92**, 341–356.
- Pohl, H.** (2010). A scanning electron microscopy specimen holder for viewing different angles of a single specimen. *Microscopy Research and Technique* **73**, 1073–1076.
- R Core Team** (2019). R: A language and environment for statistical computing.
- Rahman, I. A.** (2017). Computational fluid dynamics as a tool for testing functional and ecological hypotheses in fossil taxa. *Palaeontology* **60**, 451–459.
- Reiss, M. J.** (1989). *The allometry of growth and reproduction*. Cambridge, United Kingdom: Cambridge University press.
- Schmidt-Nielsen, K.** (1984). *Scaling: Why is animal size so important?* New York, New York, USA: Cambridge University press.
- Schneider, C. A., Rasband, W. S. and Eliceiri, K. W.** (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.

- Shuker, D. M. and Simmons, L. W.** (2014). *The evolution of insect mating systems*. Oxford, United Kingdom: Oxford University Press.
- Srinivasan, G., Surendar, C., Chatterjee, P. and Mukherjee, T. K.** (2017). Additional Records of Mantodea and Phasmida from Andaman and Nicobar Islands. *Records of the Zoological Survey of India* **117**, 264–273.
- St Quintin, W. H.** (1908). Notes on the life history of the leaf insects. *Naturalist* **618**, 235–238.
- Su, G., Dudley, R., Pan, T., Zheng, M., Peng, L. and Li, Q.** (2020). Maximum aerodynamic force production by the wandering glider dragonfly (*Pantala flavescens*, Libellulidae). *The Journal of experimental biology* **223**, jeb218552.
- Taylor, G. K. and Thomas, A. L. R.** (2003). Dynamic flight stability in the desert locust *Schistocerca gregaria*. *Journal of Experimental Biology* **206**, 2803–2829.
- Troelsen, P. V., Wilkinson, D. M., Seddighi, M., Allanson, D. R. and Falkingham, P. L.** (2019). Functional morphology and hydrodynamics of plesiosaur necks: Does size matter? *Journal of Vertebrate Paleontology* **39**, e1594850.
- Varenberg, M. and Gorb, S.** (2007). Shearing of fibrillar adhesive microstructure: Friction and shear-related changes in pull-off force. *Journal of the Royal Society Interface* **4**, 721–725.
- Varenberg, M. and Gorb, S. N.** (2009). Hexagonal surface micropattern for dry and wet friction. *Advanced Materials* **21**, 483–486.
- Wohlfart, E., Wolff, J. O., Arzt, E. and Gorb, S. N.** (2014). The whole is more than the sum of all its parts: collective effect of spider attachment organs. *Journal of Experimental Biology* **217**, 222–224.
- Wolff, J. O. and Gorb, S. N.** (2012). Surface roughness effects on attachment ability of the spider *Philodromus dispar* (Araneae, Philodromidae). *Journal of Experimental Biology* **215**, 179–184.
- Yanoviak, S. P., Dudley, R. and Kaspari, M.** (2005). Directed aerial descent in canopy ants. *Nature* **433**, 624–626.
- Yanoviak, S. P., Munk, Y. and Dudley, R.** (2011). Evolution and ecology of directed aerial descent in arboreal ants. *Integrative and Comparative Biology* **51**, 944–956.
- Young, J., Walker, S. M., Bomphrey, R. J., Taylor, G. K. and Thomas, A. L. R.** (2009). Details of insect wing design and deformation enhance aerodynamic function and flight efficiency. *Science* **325**, 1549–1552.
- Zeng, Y., Lin, Y., Abundo, A. and Dudley, R.** (2015). Visual ecology of directed aerial descent in first-instar nymphs of the stick insect *Extatosoma tiaratum*. *Journal of Experimental Biology* **218**, 2305–2314.
- Zeng, Y., Chang, S. W., Williams, J. Y., Nguyen, L. Y. N., Tang, J., Naing, G., Kazi, C. and Dudley, R.** (2020a). Canopy parkour: movement ecology of post-hatch dispersal in a gliding nymphal stick insect, *Extatosoma tiaratum*. *Journal of Experimental Biology* **223**, jeb226266.
- Zeng, Y., Malley, C. O., Singhal, S., Rahim, F., Park, S., Chen, X. and Dudley, R.** (2020b). A tale of winglets : evolution of flight morphology in stick insects. *Frontiers in Ecology and Evolution* **8**, 1–15.

Supplementary information

Table S1: Scaling relationships between body size and various morphological measurements. Results of type I ANOVA from linear models contrasting the differences between sexes in terms of scaling relationships between body size (body length) and body mass, area, circularity, aspect ratio, antenna length or front femur length (Fig. S2). The scaling relationships for male hindwing area, female forewing area, male wing loading and male flight muscle mass with body size are also presented. Scaling exponents β and the corresponding 95% confidence intervals are shown in comparison to isometric expectations. Significant effects (i.e., $p < 0.05$) are bolded.

Response variable (\log_{10} transformed)	Explanatory variables	F	df1	df2	P	Isometric slope	Slope β [95% CI]
Body mass	$\log_{10}(\text{body length})$	3075	1	40	<0.001	3	2.89 [1.61, 4.16]
	sex	14.2	1	40	<0.001		
	interaction	0.002	1	40	0.97		
Body area	$\log_{10}(\text{body length})$	12897	1	40	<0.001	2	1.82 [1.4, 2.24]
	sex	58.8	1	40	<0.001		
	interaction	1.56	1	40	0.22		
Body circularity	sex	1675	1	40	<0.001	0	0.11 [-0.27, 1.12]
	$\log_{10}(\text{body length})$	1.15	1	40	0.29		
	interaction	0.64	1	40	0.43		
Body aspect ratio	sex	1968	1	40	<0.001	0	0.36 [-0.09, 0.8]
	$\log_{10}(\text{body length})$	0.69	1	40	0.41		
	interaction	3.53	1	40	0.07		
Antenna length	$\log_{10}(\text{body length})$	15396	1	30	<0.001	1	1.22 [0.72, 1.72]
	sex	670	1	30	<0.001		
	interaction	0.01	1	30	0.92		
Front leg length	$\log_{10}(\text{body length})$	1047	1	40	<0.001	1	1.15 [0.7, 1.59]
	sex	1.06	1	40	0.31		
	interaction	0.32	1	40	0.57		
Male hindwing area	$\log_{10}(\text{body length})$	23.8	1	23	<0.001	2	1.86 [1.07, 2.65]
Female forewing area	$\log_{10}(\text{body length})$	79.4	1	17	<0.001	2	1.89 [1.44, 2.34]
Male wing loading	$\log_{10}(\text{body length})$	5.23	1	23	0.03	1	1.08 [0.1, 2.06]
Male flight muscle mass	$\log_{10}(\text{body length})$	9.98	1	21	0.005	3	5.2 [1.78, 8.65]

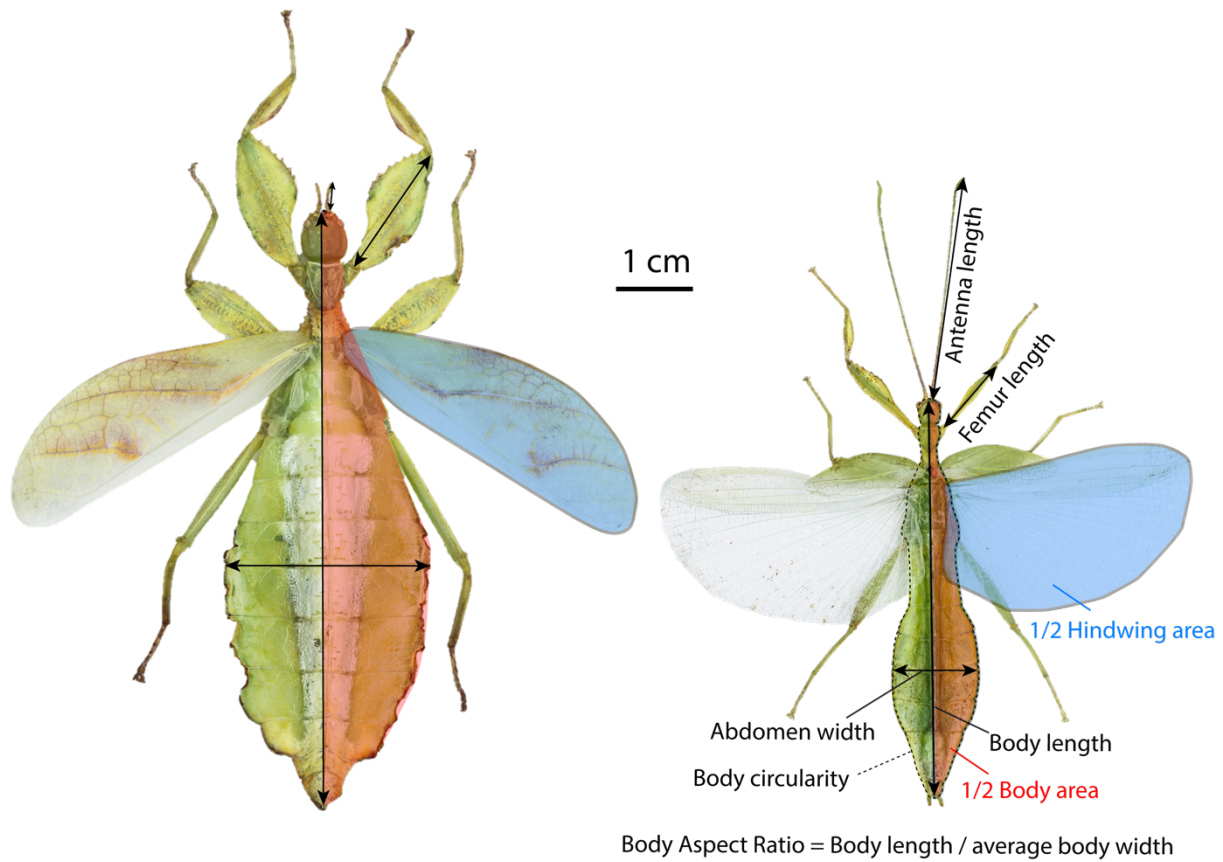


Figure S1: Morphological measurements for females (left) and males (right) *P. philippinicum*.

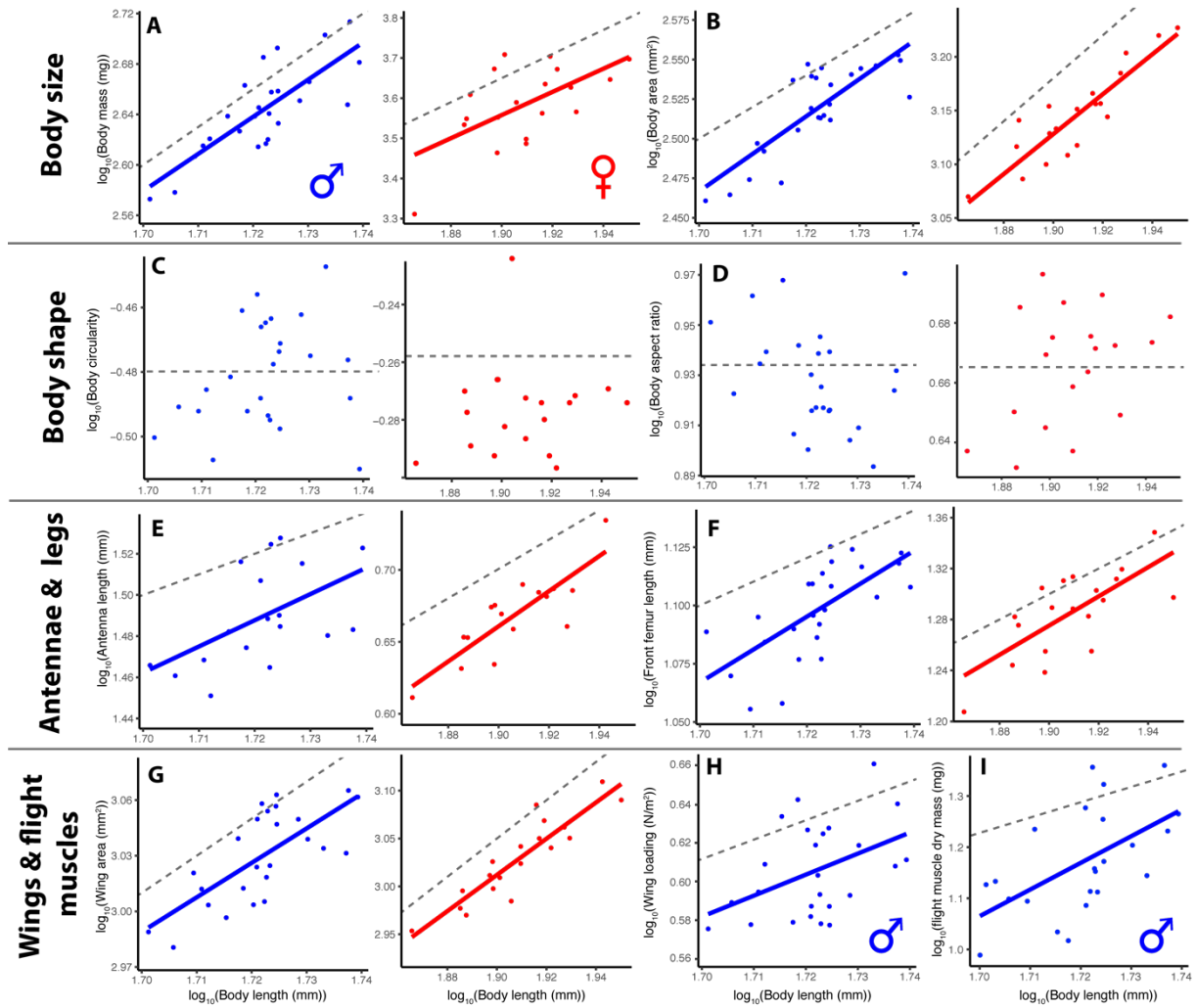


Figure S2: Scaling relationships between various morphological traits (body mass (A), body area (B), body circularity (C), body aspect ratio (D), antenna length (E), front femur length (F), total wing area (G), wing loading (H) and flight muscle dry mass (I)) and body length in males (blue) and females (red). Dashed lines show an isometric slope (arbitrary intercept). Wing area refers to hindwings for males, forewings for females. Wing loading and flight muscle mass are only shown for males as females are flightless.

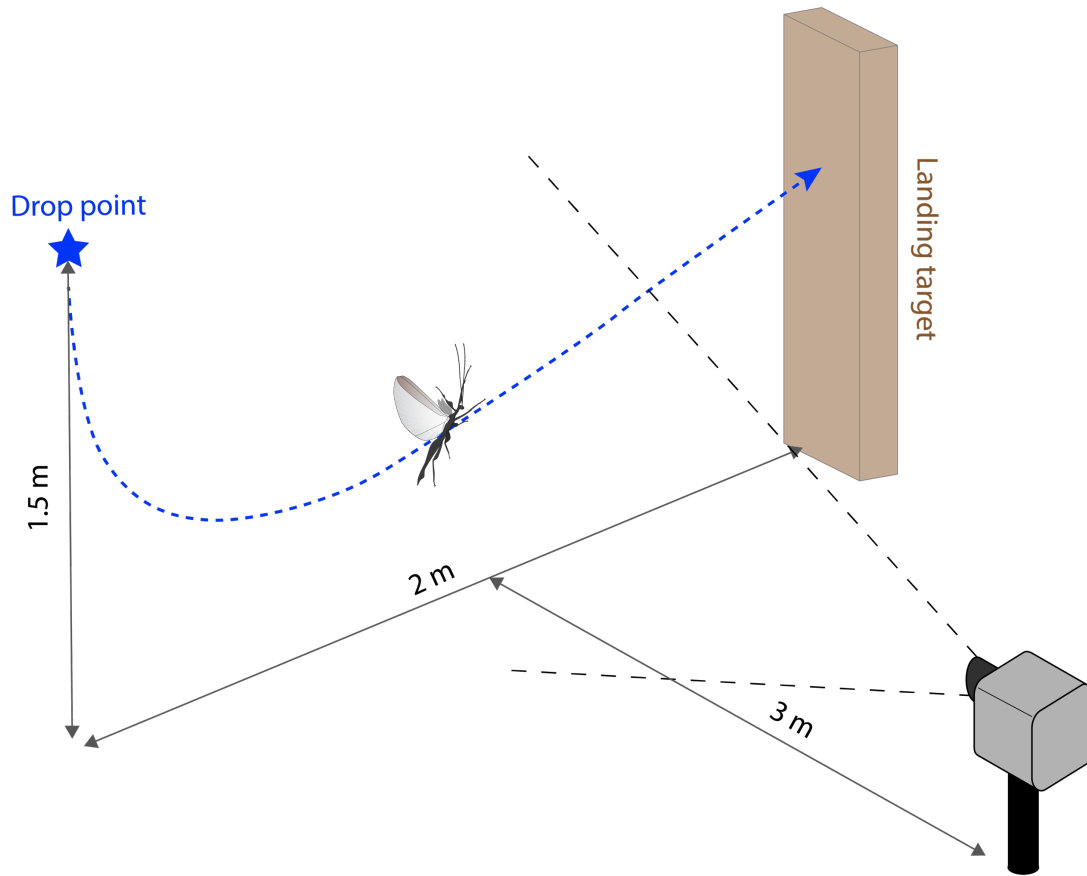


Figure S3: Set-up for flight trials.

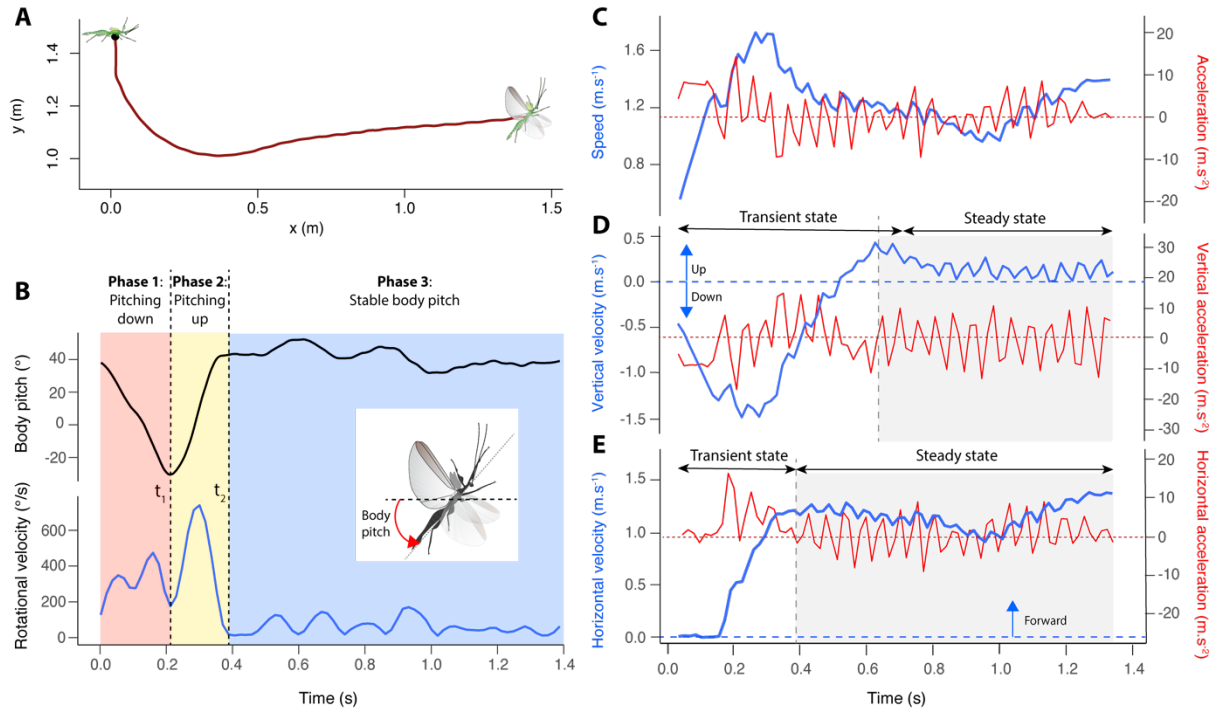


Figure S4: Analysis of a male *P. philippinicum* flight trial. **A:** Smoothed trajectory. Cartoons show the typical postures of the insect at the beginning and end of the trial. **B:** Body pitch and body pitch rotational velocity as a function of time. Colors indicate different phases during the trial defined by t_1 – i.e., to the time when body pitch reaches its absolute minimum – and t_2 – i.e., the time when body pitch stabilizes. **C:** Body speed and overall acceleration as a function of time. **D:** Vertical body velocity and acceleration as a function of time. Positive values indicate an upward velocity or acceleration. **E:** Horizontal body velocity and acceleration as a function of time. Positive values indicate a forward velocity or acceleration. In **D** and **E**, the boundary between the transient and steady state corresponds to the time when velocity (vertical or horizontal) stabilizes.

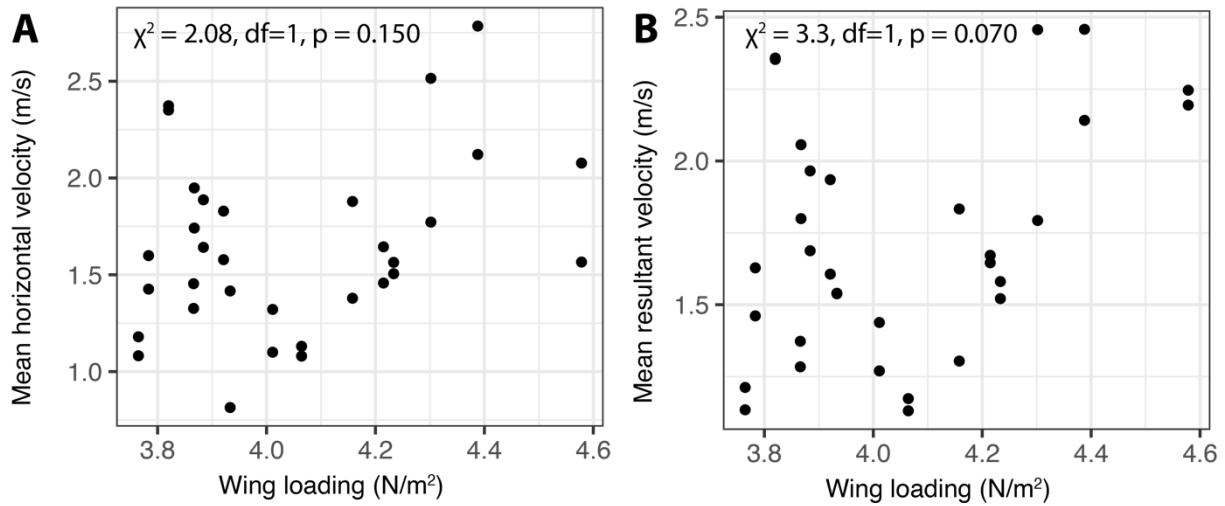


Figure S5: Mean horizontal velocity (**A**) and mean resultant velocity (**B**) as a function of wing loading. Mean horizontal velocity was calculated after reaching a steady state (Fig. S4E). Mean resultant velocity was calculated after stabilization of body angle (Fig. S4B). Each dot represents a trial (two per individual). Outputs from linear mixed effect models and corresponding likelihood ratio tests are shown on each graph.

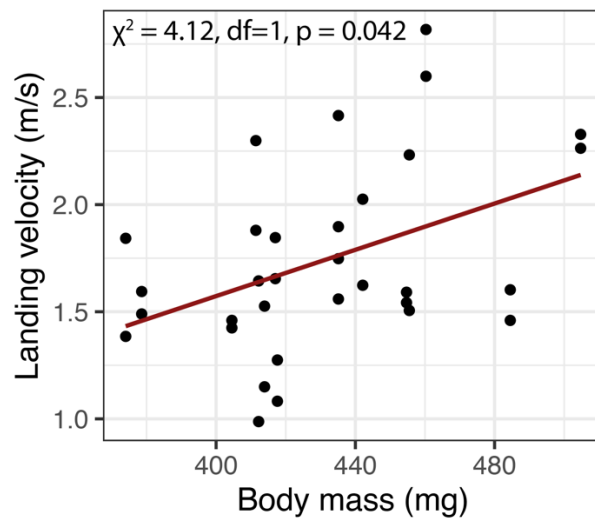


Figure S6: Instantaneous velocity at landing as a function of male body mass. Each dot represents a trial (two per individual). The output from a linear mixed effect model and the corresponding likelihood ratio test is shown on the graph (see methods).

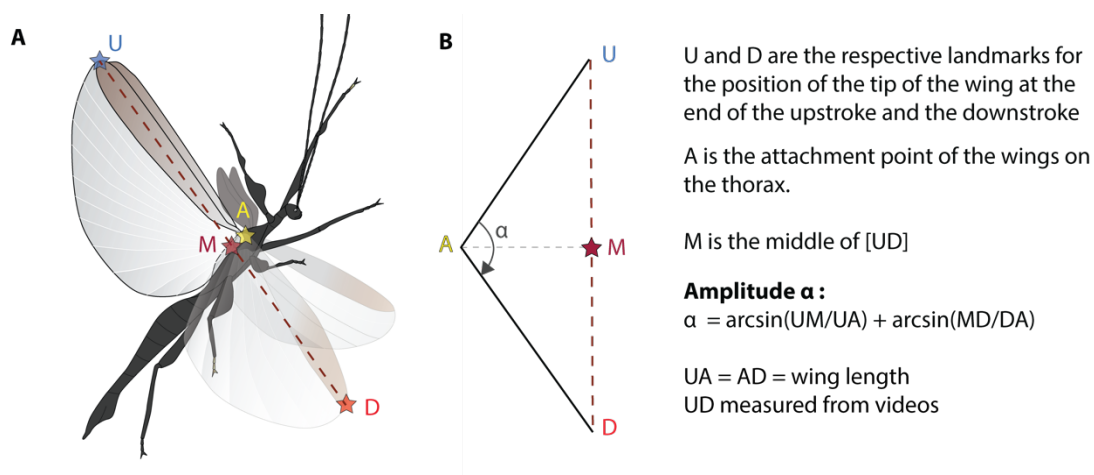


Figure S7: Calculation of wing stroke amplitude. **A:** side view of the animal with the relevant landmarks. **B:** landmarks in frontal view.

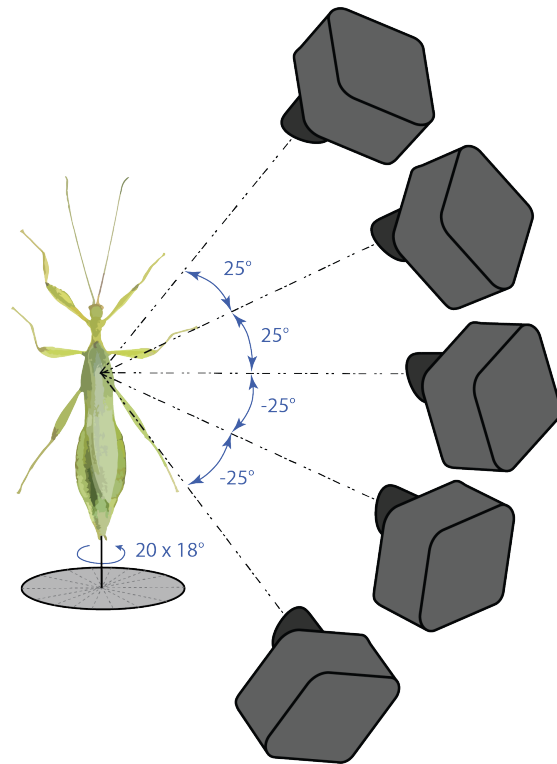


Figure S8: Acquisition of multiple 2D photographs from different angles of a male mounted on a pin to reconstruct a single 3D model using photogrammetry.

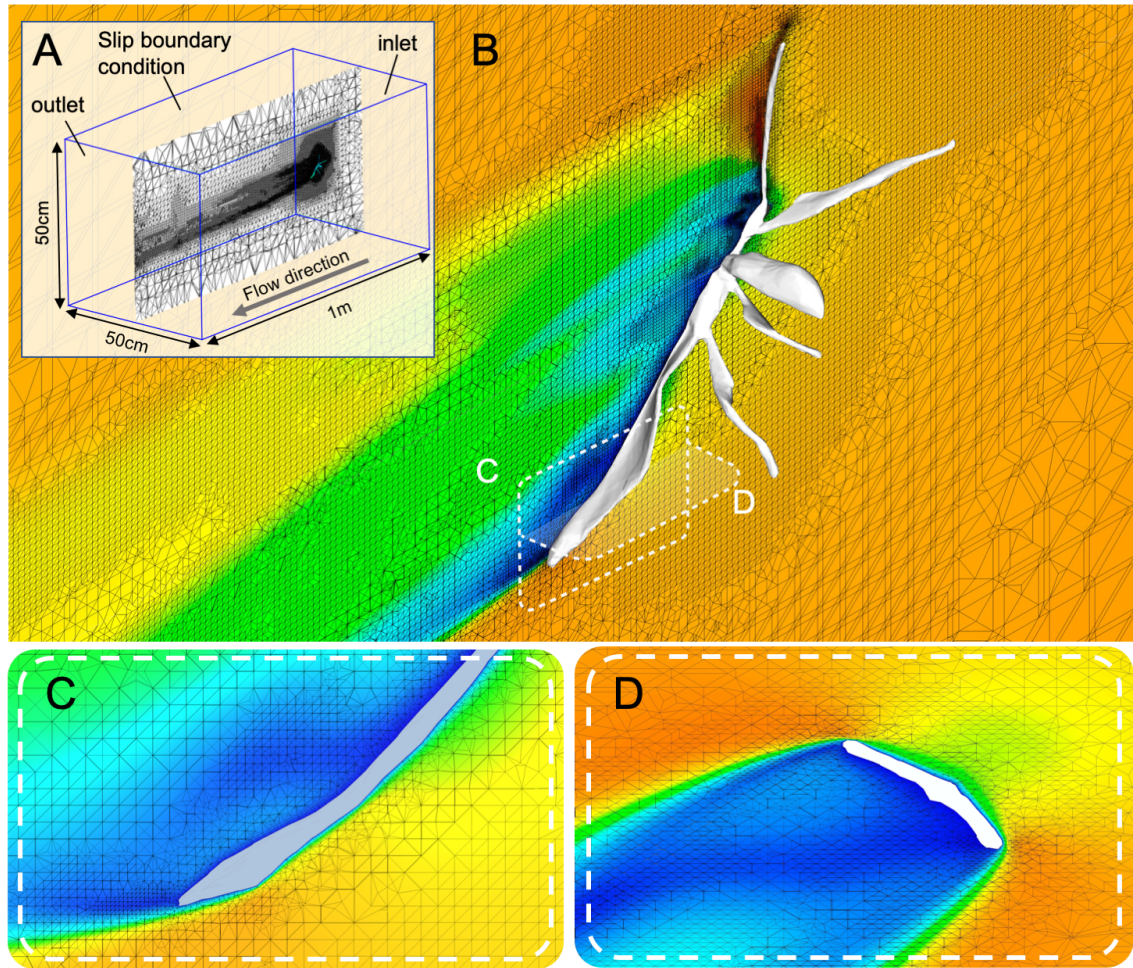


Figure S9: Leaf insect model and CFD simulation results. **A:** Meshed CFD domain. **B:** Detail of the mesh around the phasmid model along the median sagittal plane. **C:** Detail of the mesh around the tip of the abdomen along the median sagittal plane. **D:** Detail of the mesh around the middle of the abdomen along a cross-sectional plane. Colors in **B-D** represent air velocity (see scale in figure 6).

Video S1: Example of a flight trajectory from a relatively light male (374.0 mg)

Video S2: Example of a flight trajectory from a relatively heavy male (504.7mg)

Video S3: Detail of a male leaf insect ascending.

CHAPTER 5

Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects

Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G. and Emlen, D. J. (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes* **181**, 104263.

Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects

Romain P. Boisseau^a, Mark M. Ero^b, Simon Makai^b, Luc J.G. Bonneau^b, Douglas J. Emlen^a

^a Division of Biological Sciences, University of Montana, Missoula, Montana, United States of America

^b PNG Oil Palm Research Association Inc., Dami Research Station, PO Box 97, Kimbe, West New Britain, Papua New Guinea

Corresponding author: Romain P. Boisseau

Keywords: *Eurycantha* | Papua New Guinea | radiotelemetry | allometry | sexual selection | animal weapons

Highlights

- Male thorny devil stick insects are unusually large for male stick insects.
- They also display larger and more armored hindlegs than females.
- This lineage switched from solitary roosting in the canopy to aggregating inside cavities.
- This switch is associated with a change in their mating system from a searching to a defense polygyny.
- Males use their enlarged hindlegs in male competition for access to females.

Abstract

The habitat and resource use of females critically affects their pattern of distribution and consequently their monopolisability by males and the mating system of a species. Shifts in habitat use are therefore likely to be associated with changes in mating system and sexual selection acting on males' phenotypes, consequently affecting patterns of sexual dimorphism. Although sexual dimorphism is often correlated with shifts in habitat use at the macroevolutionary scale, the underlying microevolutionary processes involved are typically unclear. Here, we used the New Guinean stick insect genus *Eurycantha* to investigate how changes in habitat use and mating system were associated with a change in sexual dimorphism seen specifically in the thorny devil stick insects (*Eurycantha calcarata* and *Eurycantha horrida*). Male thorny devils display sexually dimorphic and enlarged hindlegs endowed with a sharp spine. Sexual size dimorphism is also very reduced in these species relative to other phasmids. Using field observations, morphological measurements and radiotelemetry, we investigated changes in mating system associated with the reduction of sexual dimorphism and tested predictions from the hypothesis that sexual selection drove the evolution of this unusual male morphology. We found that thorny devils switched from solitary roosting in the canopy during the day to communal roosting inside cavities of a few host trees, shifting the distribution of females from scattered to clumped. Male thorny devils used their large hindlegs to fight with rivals for positions on the tree close to cavities containing females, and larger males were associated with cavities containing relatively more females. In contrast, the sister species, *Eurycantha insularis*, displays relatively small and unarmored males (ancestral state). Adult female *E. insularis* were always scattered in the canopy, and this species displayed a scramble competition mating system typical of other phasmids, where mobility, rather than fighting ability, is probably critical to males' reproductive success. Overall, our study illustrates how a drastic change in sexual dimorphism can be associated with a switch from solitary to communal roosting and from a scramble competition to a defense-based polygyny mating system.

Introduction

Sexual dimorphism (*i.e.*, phenotypic differences between males and females) is widespread and widely variable across animals (Fairbairn *et al.* 2007). Dimorphism ultimately results from differential selection acting on traits that have sex-dependent benefits and costs, leading the same trait toward different optima in each sex (Blanckenhorn, 2005). Body size is typically sexually dimorphic as it influences multiple and potentially different aspects of fitness between the sexes (Blanckenhorn, 2000; Peters, 1986). Fecundity selection usually favors larger body sizes in females (Honěk, 1993; Pincheira-Donoso and Hunt, 2017; Preziosi *et al.*, 1996), whereas sexual selection, when present, typically favors larger sizes in males (Andersson, 1994; Kokko *et al.*, 2014). The ecology of females is a critical determinant of the mating system of a species and thus of the intensity and nature of competition between males for access to females (Emlen and Oring, 1977; Herberstein *et al.*, 2017; Shuker and Simmons, 2014; Thornhill and Alcock, 1983). Female habitat use in particular is a key parameter of the *environmental potential for polygamy* (EPP) because it affects the economical defensibility and monopolisability of females by males (Emlen and Oring, 1977). Switches in female habitat and resource use are therefore likely to critically influence the strength and/or direction of sexual selection acting on male phenotypes.

At the macroevolutionary scale, shifts in the habitat use of a lineage (e.g., from terrestriality to arboreality) are often associated with changes in sexual dimorphism (Lizards: Butler *et al.*, 2007, 2000; Kaliontzopoulou *et al.*, 2015; Turtles: Berry and Shine, 1980; Stephens and Wiens, 2009; Chameleons: Stuart-Fox and Moussalli, 2007; Snakes: Hendry *et al.*, 2014; Bats: Wu *et al.*, 2018; Grasshoppers: García-Navas *et al.*, 2017; Spiders: Moya-Laraño *et al.*, 2009; Ungulates: Pérez-Barbería *et al.*, 2002). However, in many of these clades, the specific selective pressures driving the evolution of dimorphism remain unclear, since elucidating these microevolutionary processes requires precise information on mating system and sex-specific habitat and resource use. Such insights can be gained by studying the sex-specific ecology of very closely related species or conspecific populations differing in the direction or extent of sexual dimorphism (e.g., Sepsis flies: Puniamoorthy *et al.*, 2012; Rohner *et al.*, 2016; Eyelid geckos: Kratochvíl and Frynta, 2007; Side-blotched lizards: Corl *et al.*, 2010; water striders: Toubiana and Khila, 2019; seed-feeding beetles: Fox *et al.*, 2007). The in-depth study of such systems can illuminate specific ways that recent shifts in habitat use or behavior may be associated with sex-specific phenotypic responses.

In this study, we studied how a recent shift in sexual dimorphism in thorny devil stick insects was associated with changes in habitat use and mating system. The New Guinean thorny devil stick insects (*Eurycantha calcarata* Lucas, 1869 and *Eurycantha horrida* Boisduval, 1835) and their close relatives, the oil palm stick insects (*Eurycantha insularis* Lucas, 1869) differ drastically in the extent of sexual dimorphism in body size and hindleg morphology. This genus of stick insects includes very large herbivorous, nocturnal and robust insects (**Figure 1**, Bedford, 1976; Bradler and Buckley, 2018) that are major oil palm pests where they occur in Papua New Guinea (PNG) (Kimsey et al., 2013; Monteith and Dewhurst, 2011). In *E. calcarata* and *E. horrida*, adult males have enlarged hindlegs endowed with a sharp spine and males are very similar in body size to females (**Figure 1, 2A**, Video S1), which is uncommon in the order Phasmatodea where sexual size dimorphism is usually extremely female-biased (Bedford, 1976; Bradler and Buckley, 2018; Buckley et al., 2009). This unusual “tree lobster” morphology seems to have evolved recently in the *Eurycantha* lineage as *E. insularis*, the sister species to *E. horrida* and *E. calcarata*, and *Eurycantha coronata* (Redtenbacher, 1908), a seemingly more basal lineage (Pacheco, 2018; Robertson et al., 2018), have relatively smaller males without enlarged hindlegs (**Figure 2A**). We used *E. insularis* as a control of the ancestral state of this clade to investigate potential changes in habitat use and mating system that happened in the *E. calcarata/horrida* lineage that may be associated with the evolution of large male body and hindleg sizes.

Except for two brief field notes (Bedford, 1976; Monteith and Dewhurst, 2011) and a few lab studies (Gottardo et al., 2015; Hsiung, 1987; Hsiung and Panagopoulos, 1998) on *E. calcarata*, the natural history of *Eurycantha spp.* is largely unknown and therefore it is especially unclear why species within this genus vary so drastically in sexual dimorphism, and what the primary function of male thorny devils’ exaggerated hindlegs is. Remarkably, *E. calcarata* have been documented to be ground dwellers and gregarious cavity nesters (Bedford, 1976), unlike most phasmids that are solitary canopy dwellers (Bradler and Buckley, 2018). However, the relationship between sexual dimorphism, habitat use and mating system within this clade is unresolved. We hypothesized that these behavioral and habitat use differences may be associated with changes in mating system and patterns of sexual dimorphism in this group. Indeed, the social behavior and pattern of distribution of females is likely to affect their monopolisability by males and the relative importance of fighting success for male fitness (Emlen and Oring, 1977; Thornhill and Alcock, 1983). Specifically, we predicted that clumped and gregarious females would be associated with a female-defense mating system favoring large and strong males, while scattered and solitary females would be associated with a scramble competition mating

system, favoring small and enduring males (Herberstein et al., 2017). In this context, sexual selection could explain the evolution large body sizes and enlarged hindlegs in *E. horrida* and *E. calcarata* (“sexual selection hypothesis”) (Andersson, 1994; Emlen, 2008; Rico-Guevara and Hurme, 2019; Shuker and Simmons, 2014). Males of these species may use their hindlegs to fight against rivals and defend groups of females inside diurnal cavities or near these cavities (Emlen, 2014). In many animals, such sexually selected weapons are typically hypervariable and disproportionately large in larger individuals (‘positive allometry hypothesis’ reviewed in O’Brien et al., 2018). While this positive allometry is not always true for all weapons (Bonduriansky, 2007; Fea and Holwell, 2018), recent studies suggest that it particularly applies to those that function as visual threat signals of individual body size (Eberhard et al., 2018; O’Brien et al., 2018). The allometry of hindleg size with body size in thorny devil stick insects could therefore offer insights into the function of these legs.

To provide a first critical test of the “sexual selection hypothesis”, we investigated whether changes in sexual size dimorphism were associated with changes in the mating system of these insects. We characterized the sex-specific habitat use and mating systems of these phasmid species in the field in relation to their pattern of sexual dimorphism in body and hindleg size. We used natural history observations, morphological scaling relationships and radiotelemetry to uncover intersexual differences in habitat use and interspecific differences in mating system. In *E. calcarata* and *E. horrida*, we specifically tested (1) whether the male hindlegs displayed a positively allometric scaling relationship with body size, (2) whether the movement patterns of males and females were consistent with a female-defense mating system (i.e., show aggregation of males around groups of females), (3) whether males used their hindlegs to fight against other males in the field, and (4) whether males with larger hindlegs were associated with larger groups of females. In contrast, we predicted a scramble competition mating system in *E. insularis*, where males would search for solitary females across the landscape (Herberstein et al., 2017), as is typically found in stick insects (Kelly, 2014; Myers et al., 2015). In such context, male mobility is likely to be critical for their reproductive success and may favor males with relatively smaller sizes and longer legs (Herberstein et al., 2017; Kelly et al., 2008). Consequently, using our observations and radiotelemetry data, we tested (1) whether females of *E. insularis* were solitary and scattered across the landscape, (2) whether males were more mobile than females, and (3) whether small males were more mobile than larger ones.

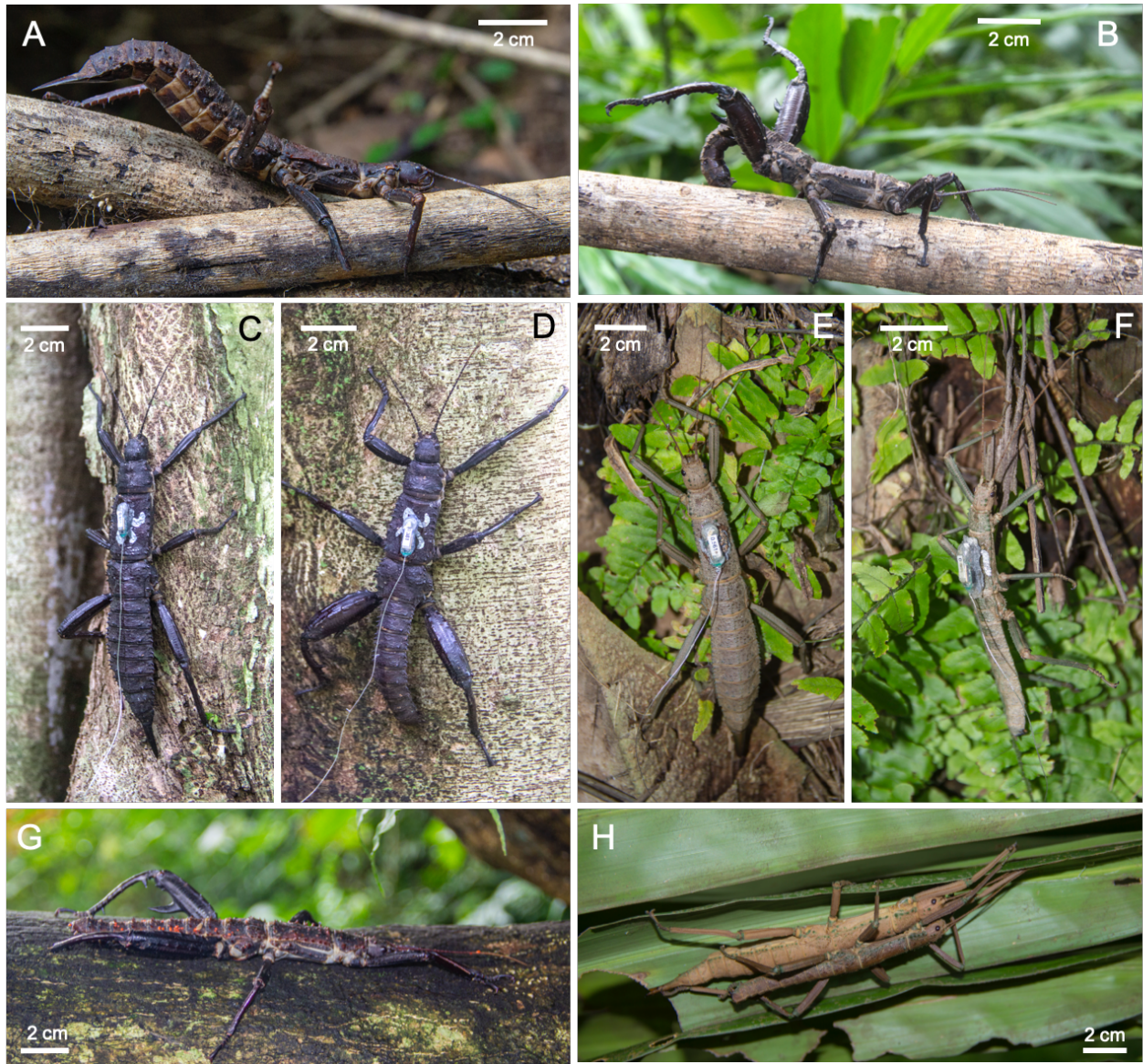


Figure 1: *Eurycantha* spp. adults *in situ*. Adult female (A) and male (B) *E. calcarata* in threatening posture. *E. calcarata* adult female (C) and adult male (D), and *Eurycantha insularis* adult female (E) and adult male (F) with paint markings and equipped with radio transmitters. (G) Adult male *E. horrida* in resting posture. (H) Adult pair *E. insularis*. Photos by first author.

Materials and methods

Study populations. We observed populations of *E. calcarata* near Kimbe, West New Britain, Papua New Guinea (PNG) and populations of *E. insularis* and *E. horrida* near Popondetta, Northern (Oro) Province, PNG, in October/November 2018. We surveyed tree cavities and vegetation during the daytime in search for phasmids at four field sites near Kimbe: Dami palm plantations (S5° 31.846' E150° 20.221'), Dami rainforest (S5° 31.129' E150° 20.077'), Tamabu palm plantations (S5° 19.757' E150° 01.212'), Gabuna rainforest (S5°26.787' E150°03.102'); and two sites near Popondetta: Koropata palm plantations (S8° 50.556' E148° 00.998'), and Tunana palm plantations (S8° 49.008' E148° 02.236'). To find *E. calcarata* and *E. horrida*, we visually searched for suitable trunk cavities and looked for piles of stick insect frass on the ground around trunks to potentially locate hidden cavities. At the same time, we visually searched for *E. insularis* in the canopy of trees and bushes as none were found inside cavities.

Morphology. We photographed adult males and females of each species in dorsal view and a side view of one of their hindlegs. Pronotum width, mesonotum length and width, front femur length, hind femur length, width and area, were then measured using ImageJ (v 1.51, Schneider, Rasband, & Eliceiri, 2012, Figure S1). *E. calcarata* individuals were photographed in the field, marked with quick drying silver paint markers (©Pilot, medium point) and then released back to their cavity (n = 73 males, n = 71 females). *E. horrida* (n = 9 females, n = 8 males) and *E. insularis* (n = 40 males, n = 25 females) individuals were brought to the entomology laboratory of the Higaturu PNGOPRA sub-center for photographing and were added to the Oil Palm Research Association (OPRA) captive populations. *E. calcarata* females' hindlegs were unfortunately not photographed in the field due to time constraints and security concerns (we had to work quickly, and our priority was relating male hindleg size with the number of females inside each cavity). Also, the defensive posture of both sexes of *E. calcarata* made it impossible to measure body length from pictures in the field as their abdomen could not be flattened. Morphological measurements for an additional 28 adult males and 40 adult females of *E. calcarata* were also obtained from individuals bred in captivity (Missoula, MT, USA) and fed maple leaves *ad libitum*. The initial population was provided by the Missoula Insectarium (MT, USA) but the exact origin of this culture is unknown.

All statistical analyses in this study were carried out using R version 3.3.1 (R Core Team, 2016), and for all linear models, we systematically checked for normal distributions of the residuals and the absence of any specific patterns in their distribution. For each species, body size measurements (*i.e.*, pronotum width, mesonotum width and length, and body length) for both sexes were used to run a principal component analysis in order to find the best linear proxy for body size (*function*: ‘R package’; *prcomp*:‘stats’, R Core Team, 2016). We determined mesothorax length to be the best proxy for body size as it was the main variable contributing to PC1 in all species (Figure S2). We then tested how the extent of sexual size dimorphism differed between species by building a linear mixed model (LMM, *lmer*:‘lme4’, Bates et al., 2015) with \log_{10} mesonotum length as the response variable and species, sex and their interaction as main fixed effects. Collection site was added as a random factor.

We then tested for sex and species differences in the scaling relationships between body size and front leg size or hind leg size. On a log scale, a steeper, and typically higher than one, slope between hindleg and body size in males than in females could be indicative of a sexually selected weapon serving as an intraspecific threat signal (O’Brien et al., 2018). The allometry of the putatively unspecialized front legs – used as a reference trait – with body size was compared to that of the hindlegs to potentially detect a unique scaling pattern in the hindlegs. We considered the scaling relationships of various hindleg measurements – *i.e.*, femur length, femur width, femur area and femoral spine area -- as only some aspects of hindleg morphology may show sex and/or species differences. We built five LMM (*lmer*:‘lme4’) including either \log_{10} front femur length, \log_{10} hind femur length, \log_{10} hind femur width, \log_{10} hind femur area or \log_{10} hind femoral spine area as response variable and \log_{10} mesonotum length, sex and species as predictor variables, as well as all two-way and three-way interactions. Collection site was also added as a random factor. Captive and wild *E. calcarata* populations were considered as different species in these analyses. AIC-based model selection was used to identify the best model. Only the fixed effects from the best models are included in the final analyses to provide the most accurate parameter estimates. Significance of the fixed effects was assessed using type III ANOVAs and the Satterthwaite's method to estimate degrees of freedom (*anova*:‘lmerTest’; Kuznetsova et al., 2017). Pairwise *post hoc* tests using estimated marginal means with Tukey contrasts were performed to further investigate interspecific and intersexual differences in these scaling relationships (*emmeans* and *emtrends*:‘emmeans’; Lenth, 2019). We tested for departure from isometry (null expectation) between each leg measurement and mesonotum length (*i.e.*, a linear measurement of body size) using 95% confidence intervals around the estimated regression slopes. In the case of front femur length, hind

femur length and width (i.e., linear measurements), isometry corresponds to a slope of 1 and in the case of hind femur and spine area (i.e., surface measurements) it corresponds to a slope of 2.

Roosting behavior. When a group of *E. calcarata* or *E. horrida* was found inside a cavity, we manually removed every individual and counted the number of adult males and females. In total, we found 87 adult female and 77 adult male *E. calcarata* in 44 different cavities; 9 adult female and 8 adult male *E. horrida* in 11 different cavities; 27 adult female and 40 adult male *E. insularis* that were never found in a cavity. Each cavity and tree was only sampled once to assess group size and sex ratio. We then tested the prediction that male *E. calcarata* with larger hindlegs would be associated with more females (i.e., found in more female-biased groups). We built a generalized linear mixed-effects model (*glmer*:*lme4*, Bates et al., 2015) with binomial errors including male-female proportion as the response variable, and \log_{10} group-averaged male hind femur area as fixed effect. Tree ID and field site were included as random factors. Significance was assessed using a Wald chisquare test (*Anova*:*car*, Fox and Weisberg, 2019).

Movement patterns and mobility. We used radiotelemetry to assess the habitat use and movement patterns of *E. calcarata* and *E. insularis* in smallholder oil palm blocks respectively near the Dami OPRA Station (Kimbe, West New Britain, PNG; S5° 31.846' E150° 20.221'), and at Ambogo New Britain Palm Oil (NBPOL) Plantation near the Higaturu PNGOPRA Sub-centre (Popondetta, Northern (Oro) Province, PNG; S8°44.468' E148°11.985'). The block in Dami contained 8-10m oil palms (*Elaeis guineensis* Jacq.) and *Kleinbovia hospita* L. trees (Malvaceae) offering suitable trunk cavities where high densities of *E. calcarata* have been found as well as low vegetation covering the ground (see picture in **Figure 4A**). The plantation near Popondetta was chosen for security reasons because it was close to the Higaturu center, and had oil palms measuring 5-6m in height allowing accurate localization of the insects in the canopy and contact between tree canopies (see picture in **Figure 4B**). Adult *E. calcarata* were collected inside different tree cavities at the study site and released on the same tree. Adult *E. insularis* were captured in smallholder oil palm blocks near the Tunana site, mostly in the canopies of oil palm trees, and transferred on the same day to the study site (Higaturu), 24 km away, which was safer to survey at night. The collection plantations were very similar to the survey plantation and unlike *E. calcarata*, adult *E. insularis* are never found inside cavities, hiding instead in the canopy. Therefore, *E. insularis* individuals were unlikely to be affected by this re-location.

We glued 0.5g transmitters (ATS, Model A2415, battery life: 12 - 24 days) onto the mesonotum of a total of 9 adult *E. calcarata* females, 9 *E. calcarata* males, 9 *E. insularis* females and 8 *E. insularis* males using cyanoacrylate glue (Loctite® Super Glue Gel Control™) (**Figure 1C-F**). To maximize sample size, the same transmitters were first glued onto *E. calcarata* individuals and were then transferred to *E. insularis* individuals. Transmitters represented on average 1.9% of a *E. calcarata* female's mass, 2.7% of a male's mass, 5.7% of a *E. insularis* female's mass and 15% of a male's mass. Radio-tagged individuals appeared to feed, mate and walk normally (Video S2). Radio transmitters had unique frequencies ranging from 150.341 to 150.621 Hz (pulse rate: 27ppm, pulse width: 15ms). Radio signals were detected using a three-element Yagi antenna connected to a portable scanning receiver (ATS, R410). We searched for the radio-tagged individuals once during the day (around 16:00h) and once during the night (around midnight) in both locations for three consecutive days. When an individual was located, we recorded its position with a global positioning system (Garmin GPSMAP 64S, accuracy ~5m). The position of each individual was marked with orange flag tape and the ground (*i.e.*, horizontal) distance travelled since the previous sighting was measured using a 50m measuring tape. Because the phasmids are mostly inactive and hiding during the daytime, the distances measured at midnight should reflect distances travelled during the first half of the night (18:00 – 0:00) while the distances measured in the afternoon should reflect distances travelled during the second half (0:00 – 6:00). Patterns of individual movements were visualized using ArcGIS Desktop (v. 10.7, Esri, Redlands, CA). GPS coordinates and individual trajectories were projected on the corresponding satellite map (2019; Bing, Microsoft corporation, Redmond, WA; Digitalglobe, Westminster, CO). Five male and five female *E. insularis* lost their radiotransmitter at some point during the survey period. Observations for which the transmitter was observed detached from the insect were not included in the analyses.

To examine whether sex-differences in mobility differed by mating system, we built LMMs (*mixed*:‘afex’, Singmann et al., 2019) for both species separately including distance travelled as the response variable. Sex, time of day (day- or nighttime) and their interaction were included as fixed effects while body size (\log_{10} mesonotum length) and front leg size (\log_{10} front femur length) were also included as covariates that we allowed to interact with sex. Individual ID, date and trunk ID where the individual was collected and released (for *E. calcarata* only) were finally added as random factors. We included body and leg size in the models as these variables are known to affect mobility in walking insects (Herberstein et al., 2017; Kelly, 2014; Kelly et al., 2008). All interactions including sex were

included to potentially uncover sex differences in the effect of time, body or leg size. Significance for fixed effects was determined using type III Likelihood Ratio Tests (LRT, *anova*:‘afex’). The LMM for *E. insularis* found a significant effect of the interaction terms between sex and body size and between sex and leg size. To investigate in details these sex-specific effects of body and leg size on distance travelled, beyond simply comparing slopes between sexes, we built additional LMMs for males and females separately including distance as the response variable, \log_{10} mesonotum length and \log_{10} front femur length as fixed effects, and individual ID and date as random factors.

Nightly activities. To explore the nocturnal behaviors of *E. calcarata* around cavities at night and further characterize their mating system, we placed two video cameras in time lapse mode (GoPro HERO 4©, interval: 0.5s) to film the trunk of a *Kleinbovia hospita* tree bearing 12 different cavities containing *E. calcarata* groups near the Dami research station (Kimbe, West New Britain, PNG) during four different nights in October and November 2018. We collected and marked some individuals ($n = 12$ females and 19 males) with a silver paint marker on the night prior to the first night of recording. Recordings were made under red illumination as red light did not interfere with the nocturnal behavior of captive animals (RB, personal observation). The red lamps and the cameras were powered by 12V 12AH SLA batteries placed inside dry bags. Recordings lasted from 4PM to 8AM. Sunrise was around 5:30AM and sunset around 5:50PM. Using the event logging software BORIS (v. 7.5.3, Friard and Gamba, 2016), we recorded the time after sunset when an identifiable individual would leave or enter a cavity, move to or return from the canopy of the tree to the trunk or move to or return from the ground. When a marked individual was observed both leaving and entering a cavity during the same night, we calculated the time it had spent outside. We compared the time of occurrence of each of these nocturnal behaviors and the duration spent outside cavities between males and females using LMMs (*mixed*:‘afex’) with time (or duration) as the response variable, sex as fixed effect and date and individual ID as random factors. Significance of the fixed effect was assessed using LRT (*anova*:‘afex’). The onset of copulations and male aggressive interactions was also recorded along with the duration of these copulations. Aggressive interactions between two males ranged from antennal contacts followed by one male walking away from the other, to kicks, pushes, mounting behaviors and hindleg squeezes (Video S3, S4). Finally, the predictability of female’s behavior in leaving or returning to cavities can have implications as to what mating behavior males should adopt throughout the night. Thus, we compared the variance of the times when marked females left cavities at the beginning of

the night and the variance of the times they re-entered them at the end of the night (combining the four nights) using a paired Pitman-Morgan test (*Var.test: 'PairedData'*, Champely, 2018).

In addition, the number of adult males and females visible on the trunk (marked or unmarked) was counted every 15 minutes from sunset for the whole night to evaluate how adult densities and sex ratios on the trunk changed over the course of a night. The temporal dynamic of male and female numbers was visualized by fitting a cubic smoothing spline to the data of all four nights combined ($df=10$, *smooth.spline: 'stats'*). We built a GLMM (family = Poisson, *mixed: 'afex'*) including the number of individuals visible on the trunk as the response variable, sex and time post sunset as fixed effects and date as a random effect. A quadratic term, $time^2$, was also added to account for the non-linearity of the effect of time. The interactions $sex:time$ and $sex:time^2$ were also included to uncover sex-specific effects of time on the number of individuals present on the trunk. Significance of fixed effects was assessed using Type III LRT (*anova: 'afex'*).

Finally, we examined the changes of male reproductive behaviors in relation to temporal changes in female number and sex ratio throughout the night. For each of the 15min time points, we also categorized whether each of the visible males present on the trunk was guarding -- i.e., immobile and likely defending a territory -- or searching -- i.e., walking and potentially searching for females or food. The temporal dynamic of the relative proportion of guarding and searching males was also visualized by fitting a cubic smoothing spline to the data of all four nights combined ($df=10$, *smooth.spline: 'stats'*). To formally test how this proportion changed across time, we ran a GLMM with binomial errors including male guarding/searching ratio as the response variable, time post sunset as a fixed effect and date as a random factor (*glmer: 'lme4'*). Significance was assessed using a Wald chisquare test (*Anova: 'car'*).

Results

Morphology. We first tested if species and sexes differed in body size (Table S1). We found a significant effect of species ($F_{3,2.2} = 155.7$, $p = 0.004$), sex ($F_{1,282.5} = 243.6$, $p < 0.0001$) and their interaction ($F_{3,282.6} = 34.6$, $p < 0.0001$) on mesonotum length. *Post hoc* tests revealed that mesonotum length differed significantly between males and females in all species ($p \leq 0.001$), except for *E. horrida* ($p=0.82$). Therefore, *E. horrida* was the only species to be non-size dimorphic.

We then explored the differences in scaling relationships between mesonotum length and front femur length between sexes and species, which were very similar overall across the three species (**Figure 2B**). The best linear model (*i.e.*, with the lowest AIC) showed that the small interspecific differences were not significant, but that males and females differed significantly in intercept in all species (**Table 1**). Specifically, males consistently had longer front femurs for a given mesonotum length than females ($p < 0.0001$, **Table 1, Figure 2B**). The slope of this scaling relationship did not differ significantly across species or sexes and was overall estimated as hypoallometric (*i.e.*, isometric slope = 1: $\beta = 0.82 \pm 0.04$, 95% CI: $0.75 < \beta < 0.90$; Table S2; **Figure 2B**).

Similarly, we investigated the differences in scaling relationships between mesonotum length and hind leg size between sexes and species (**Figure 2A, C-E, Table S1**). We found that the scaling relationships between mesonotum length and hind femur length, width and area differed between species and sexes in intercept but not in slope, as none of the interaction terms involving mesonotum length was informative and included in the best models (**Table 1**). For all of the hindleg measurements, the effect of sex was dependent on the species (species:sex, $p < 0.001$, **Table 1**). In *E. insularis*, the differences in intercepts between males and females were relatively small and not significant (**Figure 2C-E**; *post hoc* tests, $p > 0.1$). In *E. horrida* and *E. calcarata* (captive), the differences in intercepts between males and females for hind femur length, width and area were much more pronounced (**Figure 2C-E**) and all were highly significant (*post hoc* tests, $p < 0.001$). Males of these two species had longer, thicker and consequently larger hind femurs than their respective females for a given body size. The slope of the scaling relationship between mesonotum length and hind femur length did not differ significantly across species or sexes and was, as with the allometry of front leg length, significantly hypoallometric (isometric slope = 1: $\beta = 0.82 \pm 0.05$, 95% CI: $0.76 < \beta < 0.94$; Table S2; **Figure 2C**). Similarly, the slope of the scaling relationship between hind femur area and mesonotum length was

overall hypoallometric (isometric slope = 2: $\beta = 1.79 \pm 0.09$, 95% CI: $1.62 < \beta < 1.97$; Table S2; **Figure 2E**). Finally, the slope of the scaling relationship between hind femur width and mesonotum length did not differ significantly from isometry (isometric slope = 1: $\beta = 0.96 \pm 0.07$, 95% CI: $0.83 < \beta < 1.09$; Table S2; **Figure 2D**).

We found the same significant effects for the scaling relationship between femoral spine area and mesonotum length (**Table 1, Figure 2F**), with spines being much larger in male than in female *E. calcarata* (*post hoc* test, captive population, $p < 0.001$) and *E. horrida* ($p < 0.001$), and somewhat larger in male *E. insularis* than in females ($p < 0.001$). We also found a significant effect of the interaction between mesonotum length and sex: despite having smaller spines, females in all three species display a steeper slope between spine area and body size than males ($p = 0.04$, **Table 1** and S2). However, slopes of all of the scaling relationships were not different from isometry (*i.e.*, slope = 2) except for captive *E. calcarata* females which displayed a hyperallometric slope ($\beta = 3.27 \pm 0.47$, 95% CI: $2.33 < \beta < 4.21$, **Table S2; Figure 2F**).

Table 1: Results of type III analyses of variance (using the Satterthwaite's method) from linear mixed models contrasting the differences between sexes and species in terms of scaling relationships between body size (mesonotum length) and front leg size (front femur length) or hindleg size (hind femur length, width, area or femoral spine area). Only the best mixed models selected by AIC are presented.

Response variable	N	Explanatory variables	F	df1	df2	P
Log₁₀(front femur length)	294	log ₁₀ (mesonotum length)	447.3	1	71.3	<0.0001
		species	9.32	3	2.0	0.10
		sex	48.0	1	188.7	<0.0001
Log₁₀(hind femur length)	222	log ₁₀ (mesonotum length)	272.7	1	181.7	<0.0001
		species	36.3	3	2.0	0.03
		sex	187.5	1	215.0	<0.0001
		species:sex	48.5	2	209.5	<0.0001
Log₁₀(hind femur width)	222	log ₁₀ (mesonotum length)	199.6	1	214	<0.0001
		species	1016.1	3	214	<0.0001
		sex	481.8	1	214	<0.0001
		species:sex	122.2	2	214	<0.0001
Log₁₀(hind femur area)	222	log ₁₀ (mesonotum length)	387.4	1	82.1	<0.0001
		species	650.5	3	2.0	0.0015
		sex	516.3	1	173.8	<0.0001
		species:sex	125.8	2	185.0	<0.0001
Log₁₀(femoral spine area)	222	log ₁₀ (mesonotum length)	160.6	1	213	<0.0001
		species	994.7	3	213	<0.0001
		sex	12.6	1	213	0.0005
		species:sex	132.1	2	213	<0.0001
		log ₁₀ (mesonotum length):sex	4.31	1	213	0.04

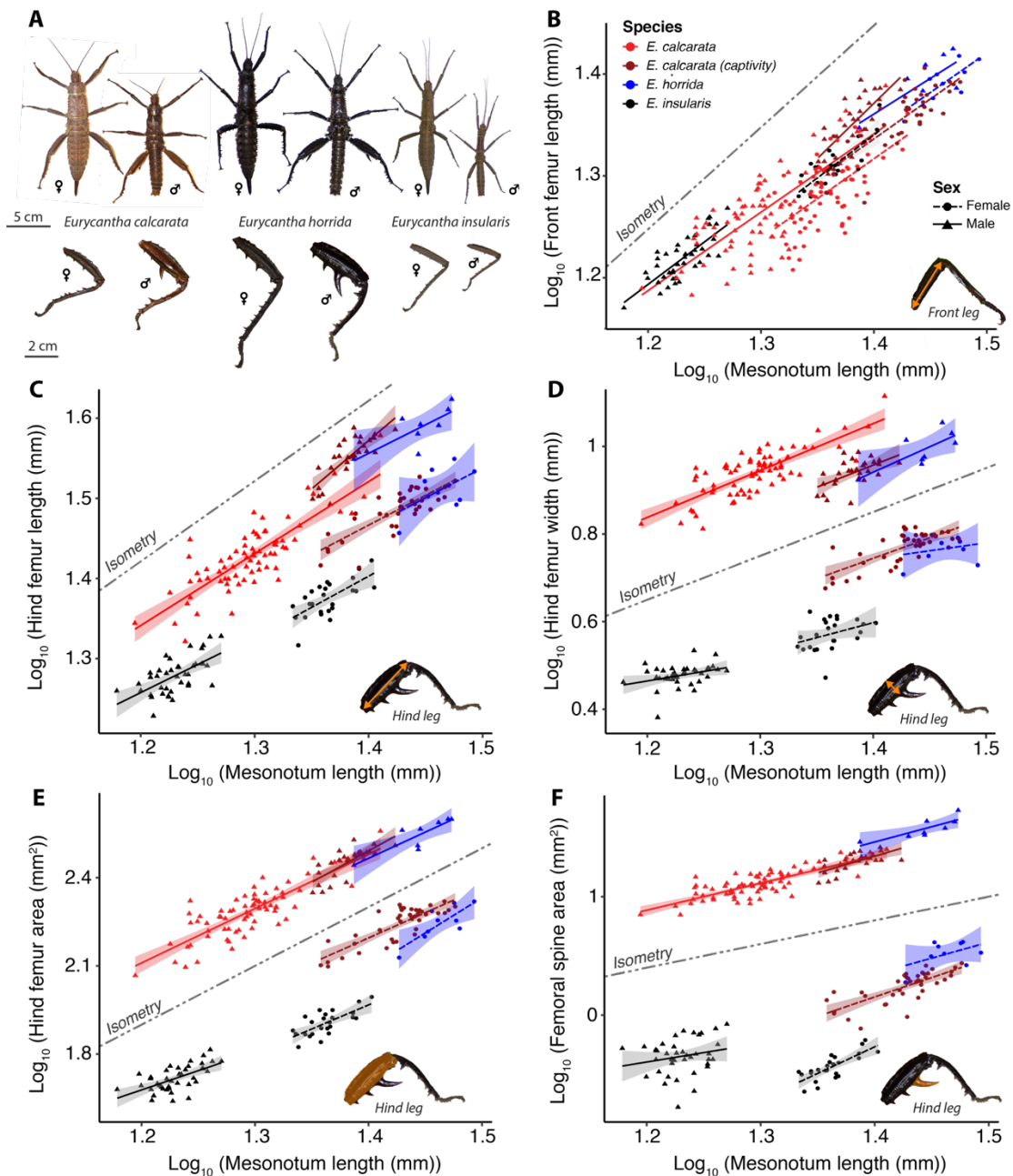


Figure 2: Sexual dimorphism and scaling relationships between body size and front or hind leg size across *Eurycantha* species. (A) Photographs in dorsal view of adult pairs of *E. calcarata*, *E. horrida* and *E. insularis*. Side views of their respective hindlegs are shown below. Scaling relationships between front femur length (B), hind femur length (C), hind femur width (D), hind femur area (E) and hind femoral spine area (F) and mesonotum length (\sim body size) for males and females of each species. Leg photographs in the bottom right corner illustrate the variable measured. The grey dashed line represents an isometric slope (intercept = arbitrary). Shaded areas represent 95% confidence intervals. They are not represented in (B) for clarity.

Roosting behavior. Aggregations of *E. calcarata* were typically found inside vertical tree cavities, especially inside live *K. hospita*, where they are especially hard to reach (**Figure 3A,B**). The largest group was composed of 26 individuals and the average group size was 5.29 ± 0.94 (mean \pm SE) if ignoring solitary individuals, 3.73 ± 0.67 otherwise (**Figure 3C**). Average sex ratio (= number of females / number of adults) inside the 44 sampled cavities was 0.39 ± 0.05 and 0.51 ± 0.05 when only considering the 28 groups with at least two individuals (**Figure 3D**). Group sex ratio was positively correlated with the log-transformed mean male hind femur area of the group (**Figure 3E**, GLMM, $n=38$ groups, Wald $\chi^2 = 8.89$, $df = 1$, $p = 0.003$), such that males with larger hindlegs were found in relatively more female-biased groups.

For *E. horrida*, population density at our field sites near Popondetta was much lower than that of *E. calcarata* near Kimbe, and we mostly found solitary individuals (six males and five females) hiding inside cavities of living tree trunks (*Pterocarpus indicus* and *Gliricidia sepium* (Fabaceae)) or on the side of rotten logs (**Figure 1G**), but also two groups inside cavities (one male with one female and one male with three females) confirming the gregarious behavior of this species.

In contrast, adult *E. insularis* were only found living in the canopies of diverse tree species, including oil palms (**Figure 1H**), sometimes in the exact same tree that would host *E. horrida* inside trunk cavities. They were all solitary (25 females and 38 males) except for two mating pairs.

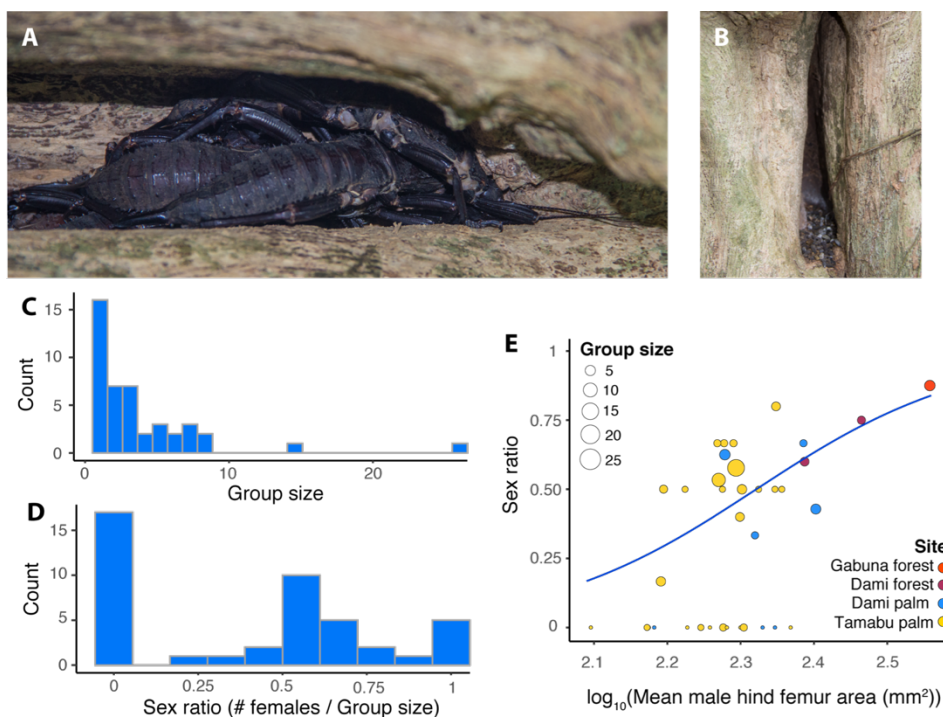


Figure 3: *E. calcarata* group composition inside cavities. (A) Group of adults inside a cavity. (B) Cavity entrance (photographs by first author). (C) Distribution of group sizes in the field. (D) Distribution of group sex ratios. (E) Binomial GLMM showing the relationship between group sex ratio and the average size of male hind femurs inside the group ($p = 0.003$). Field sites are distinguished by color and group size is indicated by the size of each circle.

Movement patterns and mobility. Radiotelemetry further confirmed that *E. calcarata* adult individuals mostly hide in groups inside tree cavities during the daytime. Relatively few *E. calcarata* were found out of a cavity during the day (11/49 sightings), and when that happened they were solitary and hiding under objects on the ground. *E. calcarata* seemed very faithful to a given cavity-bearing tree, hereafter referred to as “host tree” (Figure 4A). Only one male and one female (out of nine each) changed host tree during the 3-days survey period. Relatively few individuals were found venturing away from their host tree for more than 24h during the survey period (1/9 males, 2/9 females) (Figure 4A). These moving individuals were consistently found on or very close to the ground, suggesting dispersal on the ground rather than through the canopy. At night, adult males and females mostly fed on the canopy of their host tree (*K. hospita*) and that of neighbouring trees (<20m away), including oil palms, and females could be observed egg laying in the direct vicinity of the host tree. The maximum recorded ground distance travelled in half a night was 19.6m for males (mean= 1.29 ± 0.57 m), 30m for females (mean= 2.19 ± 0.78 m). Overall the movements of *E. calcarata* individuals on the landscape

appeared centered around their host tree as they aggregate in and on these cavity-bearing trees (**Figure 4A**). A LMM revealed that the effects of sex, body size, and time of day on distance travelled were non-significant ($p > 0.2$, **Table 2**). There was therefore no evidence that males and females differed in their mobility. However, we found that individuals with longer legs travelled longer distances ($p=0.04$, **Table 2**).

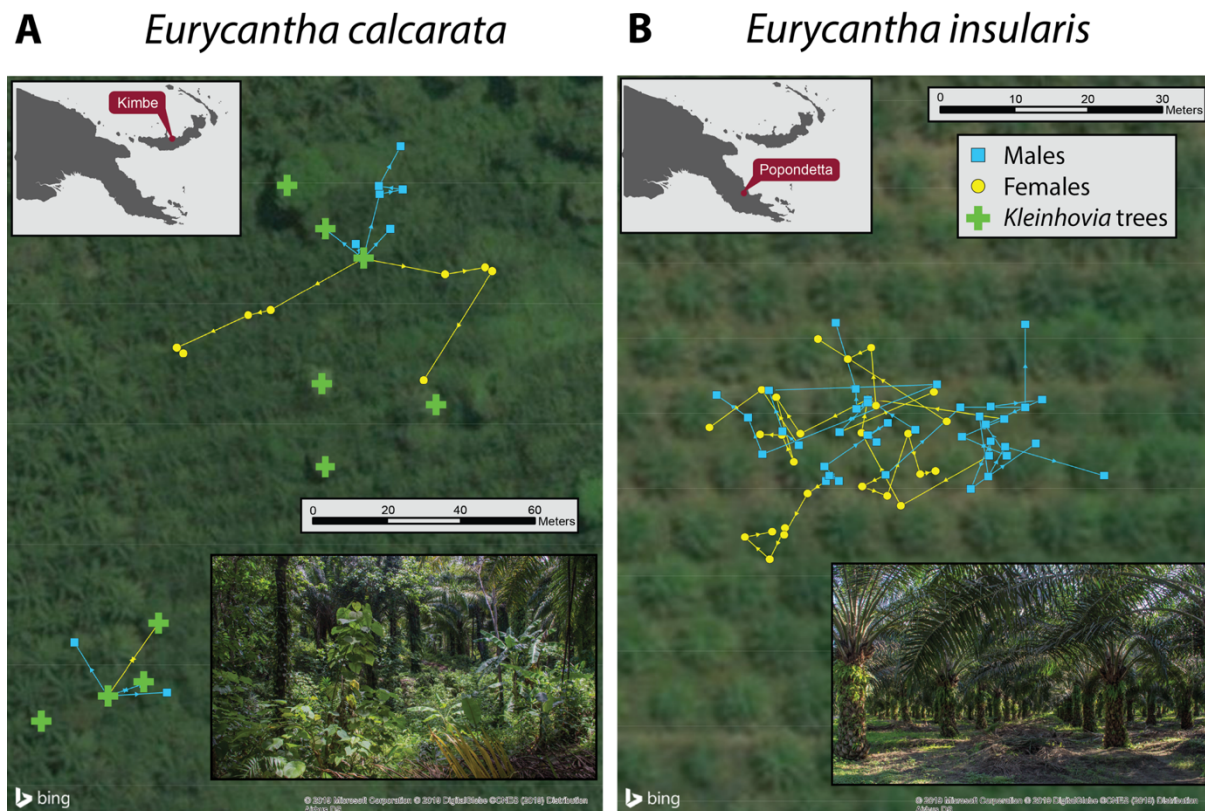


Figure 4: Horizontal movement patterns of the mobile *E. calcarata* individuals (**A**) and all the *E. insularis* individuals followed (**B**). Points represent the position of an individual after each half day period. Green crosses represent the position of potentially suitable cavity-bearing host trees. Positions of *E. calcarata* individuals that stayed inside or directly outside a *Kleinhowia* trunk are not represented. Trajectories are drawn on top of a satellite view of the study areas. Maps on the top left corners show the location of the two study areas. Pictures show the two study sites. Photos by first author.

In contrast, radiotelemetry showed that adult *E. insularis* exclusively hide in tree canopies during the day and do not aggregate, as we never found any individual roosting with others. Reconstruction of their trajectories from GPS coordinates over the course of the 3-day survey period revealed that the movements of both sexes on the landscape appeared very chaotic, and that individuals do not aggregate in specific locations (**Figure 4B**). We never found any individual on the ground, suggesting that they

moved mostly by walking through the canopy instead of walking on the ground. The maximum horizontal distance travelled during half a night was 25.6 m for males (mean= 6.59 ± 1.01 m), 19.3 m for females (mean= 5.57 ± 0.88 m). A LMM showed a significant effect of sex ($p = 0.002$, **Table 2**), of the interactions between sex and body size ($p = 0.001$) and between sex and leg size ($p = 0.01$) on distance travelled. All other fixed effects were not significant ($p > 0.17$, **Table 2**). Body size more negatively affected distance travelled by males than by females, and relative leg length more positively affected the distance travelled by males than by females (**Table 2**). A LMM only including males further showed that smaller males (LMM: \log_{10} (Mesonotum length), $\chi^2_{5,1} = 11.51$, $p < 0.001$) with relatively longer front legs (LMM: \log_{10} (Front femur length), $\chi^2_{5,1} = 9.51$, $p = 0.002$) walked significantly longer distances. In contrast, in females, we did not find any evidence that distance travelled was affected by mesonotum length (LMM: $\chi^2_{5,1} = 0.04$, $p = 0.85$) or front femur length (LMM: $\chi^2_{5,1} = 0.24$, $p = 0.63$).

Table 2: Analyses of the effects of body size, leg size and sex on ground distance travelled.

Species	Sample size	Explanatory variables	β	SE	χ^2	P
<i>E. calcarata</i>	110 observations 18 individuals	(Intercept)	-207.8	96.2		
		\log_{10} (Mesonotum length)	79.4	63.6	1.48	0.22
		\log_{10} (Front femur length)	78.1	36.3	4.18	0.04
		Sex (Contrast Male : Female)	127.8	106.9	1.38	0.24
		Time (Contrast Night : Day)	0.31	1.13	0.07	0.79
		Sex \times \log_{10} (Mesonotum length)	-52.4	78.4	0.44	0.51
		Sex \times \log_{10} (Front femur length)	-42.3	49.5	0.72	0.40
		Sex \times Time	0.34	1.59	0.04	0.83
<i>E. insularis</i>	59 observations 15 individuals	(Intercept)	-15.4	94.1		
		\log_{10} (Mesonotum length)	-5.25	87.9	0.004	0.95
		\log_{10} (Front femur length)	21.1	67.5	0.097	0.75
		Sex (Contrast Male : Female)	451.3	140.7	9.45	0.002
		Time (Contrast Night : Day)	1.16	1.62	0.40	0.53
		Sex \times \log_{10}(Mesonotum length)	-645.8	182.8	10.71	0.001
		Sex \times \log_{10}(Front femur length)	286.0	110.6	6.15	0.01
		Sex \times Time	-3.19	2.29	1.90	0.17

Summary of linear mixed effect model outputs investigating the association between the distance travelled (in meters) by each radio tracked individual between each sighting and body size, front leg size, sex and time at sighting (daytime or nighttime). The output includes the estimated parameter value for each continuous fixed effect and the difference between levels (treatment contrasts) of each categorical fixed effect, and the standard error. Chi-square statistics and associated p-values were obtained from type III Likelihood Ratio Tests. Explanatory variables that were found to have a significant effect ($p < 0.05$) are bolded.

Nightly activities. Males were estimated to leave the cavities on average 1h3min earlier than females at the beginning of the night ($\chi^2 = 11.33$, $p < 0.001$, **Figure 5A**). No other significant difference between sexes in time of occurrence of other behaviors was found: going to the canopy ($\chi^2 = 0.08$, $p = 0.78$), going to the ground ($\chi^2 = 1.50$, $p = 0.22$), returning from the canopy ($\chi^2 = 0.004$, $p = 0.95$), returning from the ground ($\chi^2 = 1.36$, $p = 0.24$) or entering a cavity ($\chi^2 = 0.99$, $p = 0.32$). The time when females exited cavities at the beginning of the night exhibited a lower variance, and was therefore relatively more predictable, than the time when the females returned to them at the end of the night ($t = -2.2$, $df=16$, $p = 0.046$). Males spent on average 1h42min longer outside cavities than females (5h23min VS 3h41min, $\chi^2 = 6.80$, $p = 0.009$). So males spent 46% more time outside cavities than females.

We found a significant effect of sex ($\chi^2 = 30.9$, $p < 0.0001$), time ($\chi^2 = 25.7$, $p < 0.0001$) and time² ($\chi^2 = 34.4$, $p < 0.0001$) on the number of individuals visible on the trunk (**Figure 5B**). The interactions between sex and time ($\chi^2 = 2.8$, $p = 0.09$), and sex and time² ($\chi^2 = 0.23$, $p = 0.23$) were not significant. These results indicate that the overall temporal dynamic of individual densities on the trunk did not differ between males and females except that males were consistently more numerous. The sex ratio (number of females/ total number of individuals) on the trunk was consistently and largely male biased throughout the night with an average of 0.19 ± 0.02 . The proportion of guarding males relative to searching ones decreased significantly throughout the night (GLMM, $\chi^2 = 63.0$, $p < 0.0001$) revealing a switch in overall male mating behavior during the night (**Figure 5C**).

After leaving their cavities at the beginning of the night, males typically stayed stationary at a specific location, such as the vicinity of cavity entrances or the base of large branches (**Figure 5D**, Video S4). They could be observed actively guarding these territories and/or fighting with rivals over direct access to females using kicks, pushes, mounting and hind leg squeezes (Video S3, S4). Additionally, scars of punctures, most likely caused by the femoral spine of another male, could be observed on the side of the hind femurs of three males (**Figure 5E**). Males would intercept and copulate with females as they left their cavities, and/or as they walked either to the canopy or the ground (Video S5). Females were never observed resisting male copulation attempts. Upon contact with a male, females typically froze and the male initiated copulation rapidly by positioning itself on the side of the female (Video S3, S5), wrapping her abdomen with one hindleg and reaching her genitalia from underneath (Video S3). Copulations lasted on average 3.29 ± 0.19 min ($n = 64$). Finally,

males could be observed actively preventing other males from entering a cavity at the end of the night (Video S6).

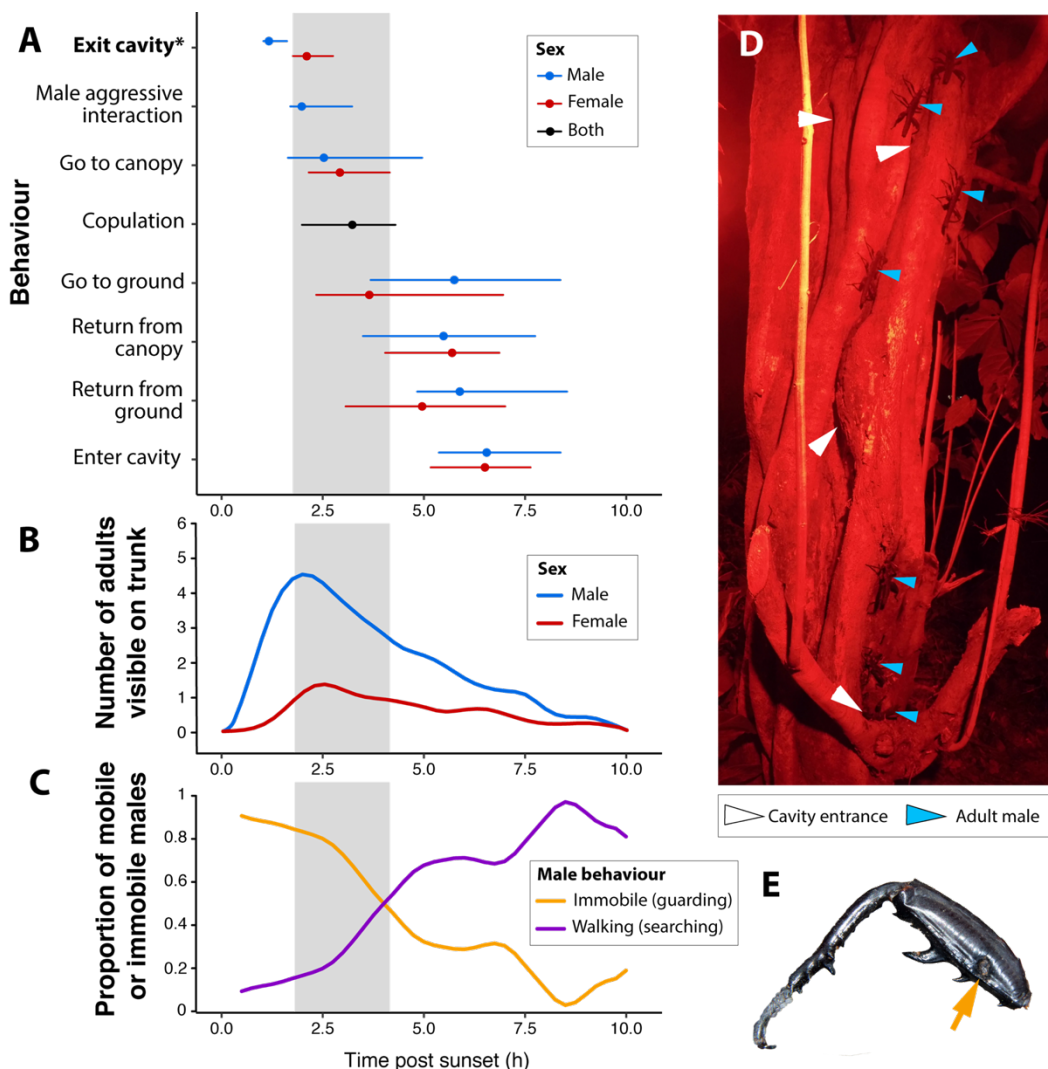


Figure 5: Timeline of *E. calcarata* adult males and females' behaviors, obtained by combining behavioral data from the four nights of recording. **(A)** Time post-sunset of sex-specific behaviors, copulations, and male-male aggressive interactions. Points represent median times. Horizontal bars represent interquartile ranges. The gray box only serves illustrative purposes and highlights the period encompassing the interquartile ranges of females exiting cavities and females going to the canopy. It is during this period that most male aggressive interactions occur as well as most copulations. The asterisk and bold font indicate a significant difference between males and females ($p < 0.001$, see result section). **(B, C)** Cubic smoothing splines showing the number of males and females visible on the trunk **(B)** and the relative proportion of immobile/guarding males relative to walking/searching males **(C)** as a function of time post-sunset. **(D)** Snapshot of one side of a host trunk, 2h post sunset: many males have already left their cavity but no females yet. Males are guarding and fighting over strategic locations close to cavity entrances that enable them to intercept females on their way to forage in the canopy. **(E)** Injury on a male hindleg likely caused by the femoral spine of a rival male.

Discussion

In this study, we show that the evolution of large armored males in thorny devil stick insects (*E. calcarata* and *E. horrida*) is associated with a switch in habit use and mating system. In most studied stick insects including *E. insularis*, solitary females are scattered across the vegetation (Kelly, 2014; Myers et al., 2015). Males are relatively small and display a scramble competition mating system as they actively search for dispersed females. In contrast with *E. insularis*, both of the thorny devil species roost in groups inside trunk cavities, and this change in habitat use and social behavior appears to be associated with a switch in mating system. We show that male thorny devil stick insects employ a defense-based polygyny mating system, with males actively fighting over direct control of clumped females or territories adjacent to roosting cavity entrances on the host tree trunk. Our data offer preliminary support for the hypothesis that sexual selection, rather than predator defense or sexual conflict, drove the evolution of large body and hindleg sizes in thorny devils: males use their enlarged and armored hindlegs during fights with rival males to squeeze and potentially puncture their opponent, and large males with the largest hindlegs are associated with relatively more females inside roosting cavities.

Our radiotelemetry data revealed that in *E. insularis*, which displays relatively small and unarmored males (i.e., the ancestral state of the *Eurycantha* clade), females remained dispersed in tree canopies foraging at night and hiding in the leaves during the day. As with other studied stick insects, males employed a scramble competition mating system in the form of a prolonged searching polygyny (Herberstein et al., 2017; Kelly, 2014; Myers et al., 2015). Both sexes were highly mobile in the canopy of palm trees and did not seem to associate for extended periods of time. We found that smaller males with relatively longer legs travelled longer distances at night than larger males. Mobility advantages to smaller males with relatively longer legs have also been found in other walking arthropods with a similar mating system (e.g., Cook Strait giant wētā: Kelly et al., 2008; crab spiders: Legrand and Morse, 2000). Mobility is expected to be critical in such mating systems as it increases the probability of encountering females (reviewed in Herberstein et al., 2017). We can therefore speculate that smaller *E. insularis* males are probably advantaged in this context because of increased mobility (the “mobility hypothesis”, Kelly et al., 2008), which may help explain the strong female-biased sexual size dimorphism in this species (Blanckenhorn, 2000; Blanckenhorn, 2005; Corcobado et al., 2010; Fairbairn et al., 2007; Kelly et al., 2008). A more comprehensive understanding of other selective forces potentially acting on male size,

as well as characterization of the selective forces acting on female size, will be needed to fully explain the pattern of sexual dimorphism seen in this species.

In contrast, in both *E. calcarata* and *E. horrida*, females descended from canopies before sunrise and took shelter inside cavities in the trunks of a small number of host trees. These animals collected in dense aggregations of mixed sex, suggesting the use of aggregation pheromones as is common in other insects (Wertheim et al., 2005). In *E. calcarata*, our primary focus of field observations, adult females aggregated inside these cavities and larger males with larger hindlegs were more often found roosting with relatively more females than smaller males. At dusk, males exited cavities earlier than the females and subsequently fought over direct access to these females as they left their cavities, or over control of strategic locations on the trunk for intercepting females, such as the vicinity of cavity entrances or the base of large branches. As females exited their cavities to forage, they were likely to walk past these locations and be intercepted by territory-guarding males. Interestingly, males only fought over territories at the beginning of the night. Later, after the majority of females had left the trunk, we observed a switch in male mating strategy from territory defense to scattering and active searching, with consequently fewer encounters between males. During the second half of the night, males found and copulated with females either on the ground or higher up in the canopy. Males never followed a female after copulation and therefore did not display any post-copulatory female guarding behaviors, unlike many other male phasmids known to spend extended periods of time *in copula* or on the female's back as a post-insemination mate guarding strategy (Bedford, 1978; Kelly, 2015; Sivinski, 1979). Males did not resume territory defense at the end of the night, which was consistent with the time of return of the females to a cavity being more variable, and therefore less predictable, than the time when they leave cavities. Thus, territory guarding is likely more beneficial at the beginning than at the end of the night as females are more likely to pass by these territories more frequently early in the night. The absence of post-copulatory mate guarding behaviors and the occurrence of male fights mainly at the beginning of the night when the females exit the cavities could also be consistent with a first male sperm precedence (Simmons, 2001; Simmons, 2014). However patterns of sperm precedence are still unknown in this insect lineage.

We clearly observed male *E. calcarata* using their hindlegs to fend off rivals and potentially injure them, as witnessed by the stereotypical puncture wounds found on several males' cuticles. Males consistently fought in three contexts: (1) over direct access to a nearby female with some males even stealing mates; (2) when defending strategic locations on the trunk where females are likely to walk by

when leaving their cavity; and (3) to prevent rivals from entering a cavity. In context 1, the benefit of winning a fight is clear as victors earn direct access to mating. In context 2, males successfully keeping rivals away from the likely path of females should also lead to increased mating opportunities. These males would typically rush at trespassing rival males and often fought to repel them. Successful males immediately mounted any females that passed by. The benefits of preventing rivals from entering a cavity (context 3) are less clear. Individuals are tightly packed one on top of the other inside these cavities, and mating in such crowded conditions is likely difficult. In our captive (Montana) *E. calcarata* population, for example, despite hundreds of hours of observation we have yet to observe even a single copulation inside our artificial cavities (which have a transparent side), even though males and females aggregate together for many hours each day and readily copulate outside these cavities. If the same is true in the field, then this may explain why males tolerate other males inside their cavities, with most groups containing several males, sometimes more than females.

In some ways the *E. calcarata* mating system appears similar to that of the Wellington tree wētā (*Hemideina crassidens*, order Orthoptera), for which males use their exaggerated mandibles to drive rivals away from groups of females who also reside inside tree cavities (Kelly, 2004, 2006a, 2006b, 2006c, 2007). However, in tree wētās, males do not tolerate other males inside cavities (Kelly, 2004, 2006). This is consistent with tree wētās copulating inside these cavities, and rarely outside (Kelly, 2006a). In contrast, copulation in thorny devils readily occurs directly outside the cavities, on the tree trunk, and probably only rarely inside cavities. This may reduce the benefit to males of fighting for exclusive control of harems inside cavities, and help explain why fights so often occurred just outside cavity entrances and along the trunks of host trees. Nevertheless, males could still be observed actively preventing some rivals from entering a cavity and large males were found inside cavities containing relatively more females. This behavior may reduce competition for the next night or guarantee large males a better spot right outside the cavity entrance after the next sunset to intercept females as they leave. However, these explanations are speculative and further experiments and observations are needed to test them.

Exaggeration of the male hind legs in thorny devils (*E. calcarata* and *E. horrida*) is the result of an increase in the intercept of the otherwise hypoallometric scaling relationship between hind leg size and body size. Front leg size, that we used as a reference trait, similarly scaled hypoallometrically with body size. In addition, the allometric slopes for our different metrics of hindleg size did not differ between males and females. Sexually selected traits often display steeper and positive static allometry

in males, but this is not universal (Bonduriansky, 2007) and our findings for *E. calcarata* and *E. horrida* add to our knowledge of those few species where exaggerated and potentially sexually selected male traits do not scale hyperallometrically (Bonduriansky, 2007; Fea and Holwell, 2018). Recent studies have proposed that hyperallometry -- *i.e.*, disproportionate trait size in large individuals -- is mostly expected when morphological traits function as intraspecific visual signals, as it can amplify individual differences in body size (Eberhard et al., 2018; O'Brien et al., 2018). Therefore, our finding suggests that the enlarged legs of thorny devils may function primarily as tools of combat rather than as intraspecific signalling displays, which is consistent with the fact that we never observed males sizing each other using their hindlegs or waving these legs at rival males. In the context of intrasexual fights, hind legs only appeared to be used to deliver powerful squeezes to the opponent and potentially injuring it by puncturing its cuticle with their sharp femoral spines. Hypoallometry may be remarkably advantageous in such context from a mechanical stand-point as lever components will stay in proportion, guaranteeing a conservation of mechanical advantage and avoiding the need for a compensatory investment in costly muscles (Levinton and Allen, 2005; O'Brien and Boisseau, 2018). Notably, the hindlegs of *E. calcarata* have already be shown to bear a significant energetic cost driven by muscle mass (O'Brien et al., 2019). Future studies will be needed to look at how hindleg size affects fighting success, how squeezing strength scales with hindleg size in this species, as well as assessment strategies during fights to shed further light on how these legs might function as tools versus signals during combat (McCullough et al., 2016).

Previous authors have repeatedly hypothesized that the enlargement of the hindlegs displayed by adult male *E. calcarata* was primarily driven by their function as antipredator defenses used in threat displays and active counterattacks (Bedford, 1976; Buckley et al., 2009; Carlberg, 1989). This idea stemmed from the observation that male and female *E. calcarata* and *E. horrida* aggressively wave their hindlegs when approached by a predator or a person (Bedford, 1976; Carlberg, 1989): when threatened, they adopt a spectacular startle posture, reaching out with their spiny hindlegs poised to strike them together (**Figure 1A,B**; Bedford, 1976; Carlberg, 1989). Upon further stimulation, these insects then rapidly swing these hindlegs together attempting to grasp imprudent attackers (Video S1). This defensive behavior has been observed in several large phasmid species (Bedford, 1978) and we also observed it in female *E. insularis*. Thus, defense against predators is likely an ancestral function of the spiny hindlegs in this group, in both males and females. It is possible that predators might have also selected for hind femora that are larger in males than in females (“antipredator hypothesis”). Data on

sex-specific predation risk and on the effect of hindleg size on survival and predator deterrence efficiency would be required to test this hypothesis. Nevertheless, this defensive function likely adds to the benefits of having enlarged hindlegs for males in addition to winning fights against rivals. Such dual benefits of a sexually selected weapon have also been found in fiddler crabs and elk, where males have been shown to benefit from their enlarged claw or antlers as it reduces predation risk (Bildstein, McDowell, & Brisbin, 1989; McLain, Pratt, & Berry, 2003; Metz et al., 2018).

A third hypothesis for the evolution of enlarged male hindlegs in thorny devils could be that males may use their hindlegs to overcome female resistance to copulation attempts (the “sexual conflict hypothesis”), as is found in other insects with sexually dimorphic hindlegs (*e.g.*, thick-legged flower beetles: Burrows, 2020; camel crickets: Haley and Gray, 2012; water striders: Rowe et al., 2006). In another phasmid species, the spiny leaf insect *Extatosoma tiaratum*, facultatively asexual females were observed readily resisting males, especially to avoid switching from parthenogenesis to sexual reproduction (Burke et al., 2015). In *E. calcarata*, females are also capable of facultative thelytokous parthenogenesis, raising the possibility that they might benefit from resisting male advances, especially if switching to sexual reproduction is costly or if parthenogenesis offers higher short-term fitness benefits. However, we never observed females resisting males, which was largely inconsistent with the hypothesis that males might have used their hindlegs primarily to overcome female resistance. Females always froze when approached by a male and were never observed kicking, shaking their abdomen or trying to walk away. We could speculate that this may be associated with higher benefits of sexual reproduction relative to parthenogenesis. Alternatively, given that, in thorny devils and contrary to most other phasmids, males occur in high densities around females, successively resisting many copulation attempts from several suitors may be extremely costly for the females and outweigh the cost of mating and thus favor sexual receptivity (“convenience polyandry”) (Arnqvist and Rowe, 2005; Cordero and Andrés, 2002; Rowe, 1992).

Large armored males are glaring exceptions among stick insects but occur in the New Guinean thorny devils, *E. calcarata* and *E. horrida* (Buckley et al., 2009). New Guinean insects generally, and *Eurycantha* species in particular, are poorly understood as few studies have ever observed these animals in the wild. Our field study characterized and compared aspects of the natural histories of *E. calcarata*, *E. horrida* and *E. insularis* for the first time. Our data notably suggest that the evolutionary increase in body and relative hindleg size seen in the thorny devils is associated with a shift in habitat use and roosting behavior of the females. This association is however unresolved as the reduction of sexual

size dimorphism could be both the cause or the consequence of this ecological shift in habitat use (Fryxell et al., 2019). Unlike most stick insects and unlike their sister species *E. insularis*, female *E. calcarata* and *E. horrida* descend from the canopy to roost in aggregate within cavities in the trunks of their host trees during the day. Males also collect in these cavities, but the huge size of these insects combined with the cramped space inside the cavities means that copulations are generally not possible within the cavities themselves. Instead, communal roosting may have increased the potential for polygyny in this lineage by concentrating multiple females within one location, and synchronizing their movements as females emerged each evening and began to walk up or down the trunk. We suspect that males benefitted from being "at the right place at the right time", intercepting and mating quickly with females as they emerged from cavities, and the resulting contests that ensued selected for increases in male body and weapon size. Interestingly, once females are no longer concentrated -- *i.e.*, after the majority have dispersed to the canopy or on the ground -- males appear to revert to what presumably was their ancestral mating behavior, abandoning their territories for the remainder of each night and instead searching for females in the vicinity of their host trunk (e.g., on the ground or in the canopy). In contrast with the ancestral scramble competition mating system found in *E. insularis*, the reproductive success of *E. calcarata* and *E. horrida* males may primarily be determined by weapon size and relative fighting success, rather than or in addition to mobility. Although our data offers preliminary support for the sexual selection hypothesis, further work relating male body and weapon size to fighting and mating success is required to fully test it.

Acknowledgements

We thank R. Dikrey, B. Sapau, P. Mana, S. Komda, G. Gumbira, R. Uker, T. Manjobie and T. Batari for assistance in the field; E. McCullough for kindly lending her telemetry equipment, C. Allen for help with insect rearing; the Missoula insectarium for providing our *E. calcarata* culture; C. Tobalske for help with ArcGIS; C. Thomas-Bulle for considerable feedback on the manuscript and D. Rozen-Rechels for useful discussions on the statistical analyses.

Authors' contributions

RPB, DJE: Conceptualization; RPB: Data curation; RPB: Formal analysis; RPB, DJE, MME, LJGB: Funding acquisition; RPB, SM: Investigation; RPB, MME, SM: Methodology; RPB, MME, LJGB: Project administration; MME, LJGB, DJE: Resources; MME, LJGB, DJE: Supervision; RPB, MME, SM, LJGB, DJE: Validation; RPB: Visualization; RPB: Roles/Writing – original draft; RPB, DJE, MME: Writing – review & editing

Funding

Funding was provided by the National Geographic Society through an Early career grant to R.P.B. (grant number: WW-255ER-17), by the Papua New Guinea Oil Palm Research Association (PNGOPRA) and by a NSF grant to D.J.E. (NSF IOS–1456133).

References

- Andersson, M.** (1994). *Sexual Selection*. Princeton, NJ, USA: Princeton University Press.
- Arnqvist, G. and Rowe, L.** (2005). *Sexual conflict*. Princeton, New Jersey, USA: Princeton University Press.
- Bates, D., Maechler, M., Bolker, B. and Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48.
- Bedford, G. O.** (1976). Defensive behaviour of the New Guinea stick insect *Eurycantha* (Phasmatodea: Phasmatidae: Eurycanthinae). *Proceedings of the Linnean Society of New South Wales* **100**, 218–222.
- Bedford, G. O.** (1978). Biology and ecology of the Phasmatodea. *Annual Review of Entomology* **23**, 125–149.
- Berry, J. F. and Shine, R.** (1980). Sexual Size Dimorphism and Sexual Selection in Turtles. *Oecologia* **191**, 185–191.
- Bildstein, K. L., McDowell, S. G. and Brisbin, I. L.** (1989). Consequences of sexual dimorphism in sand fiddler crabs, *Uca pugilator*: differential vulnerability to avian predation. *Animal Behaviour* **37**, 133–139.
- Blanckenhorn, W. U.** (2000). The evolution of body size: what keeps organisms small? *The Quarterly Review of Biology* **75**, 385–407.
- Blanckenhorn, W. U.** (2005). Behavioral causes and consequences of sexual size dimorphism. *Ethology* **111**, 977–1016.
- Bonduriansky, R.** (2007). Sexual selection and allometry: A critical reappraisal of the evidence and ideas. *Evolution* **61**,

838–849.

- Bradler, S. and Buckley, T. R.** (2018). Biodiversity of Phasmatodea. In *Insect Biodiversity: Science and Society* (ed. Footitt, R. G. and Adler, P. H.), pp. 281–313. Chichester, UK: Wiley-Blackwell.
- Buckley, T. R., Attanayake, D. and Bradler, S.** (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1055–1062.
- Burke, N. W., Crean, A. J. and Bonduriansky, R.** (2015). The role of sexual conflict in the evolution of facultative parthenogenesis: A study on the spiny leaf stick insect. *Animal Behaviour* **101**, 117–127.
- Burrows, M.** (2020). Do the enlarged hind legs of male thick-legged flower beetles contribute to take-off or mating? *Journal of Experimental Biology* **223**, jeb212670.
- Butler, M. A., Schoener, T. W. and Losos, J. B.** (2000). The relationship between sexual size dimorphism and habitat use in Greater Antillean anolis lizards. *Evolution* **54**, 259–272.
- Butler, M. A., Sawyer, S. A. and Losos, J. B.** (2007). Sexual dimorphism and adaptive radiation in *Anolis* lizards. *Nature* **447**, 202–205.
- Carlberg, U.** (1989). Aspects of Defensive Behaviour of *Eurycantha calcarata* Lucas Females and the Evolution of Scorpion Mimicry in the Phasmida (Insecta). *Biologisches Zentralblatt* **108**, 257–262.
- Champely, S.** (2018). PairedData: Paired Data Analysis.
- Corcobado, G., Rodríguez-Gironés, M. A., De Mas, E. and Moya-Larão, J.** (2010). Introducing the refined gravity hypothesis of extreme sexual size dimorphism. *BMC Evolutionary Biology* **10**.
- Cordero, A. and Andrés, J. A.** (2002). Male coercion and convenience polyandry in a calopterygid damselfly. *Journal of Insect Science* **2**, 1–7.
- Corl, A., Davis, A. R., Kuchta, S. R., Comendant, T. and Sinervo, B.** (2010). Alternative mating strategies and the evolution of sexual size dimorphism in the side-blotched lizard, *UTA stansburiana*: A population-level comparative analysis. *Evolution* **64**, 79–96.
- Eberhard, W. G., José, S., Rica, C., Lucas Rodríguez, R., Alexander, B. A. H., Speck, B., Miller, H., Rodríguez, R. L., Huber, B. A., Speck, B., et al.** (2018). Sexual Selection and Static Allometry: the Importance of Function. *The Quarterly Review of Biology* **93**, 207–250.
- Emlen, D. J.** (2008). The Evolution of Animal Weapons. *Annual Review of Ecology, Evolution, and Systematics* **39**, 387–413.
- Emlen, D. J.** (2014). Reproductive contests and the evolution of extreme weaponry. In *The evolution of insect mating systems* (ed. Shuker, D. and Simmons, L.), pp. 92–105. Oxford, United Kingdom: Oxford University Press.
- Emlen, S. T. and Oring, L. W.** (1977). Ecology, sexual selection, and the evolution of mating systems. *Science* **197**, 215–223.
- Fairbairn, D. J., Blanckenhorn, W. U. and Székely, T.** (2007). *Sex, Size and Gender Roles*. Oxford, United Kingdom: Oxford University Press.
- Fea, M. and Holwell, G.** (2018). Combat in a cave-dwelling wētā (Orthoptera: Rhaphidophoridae) with exaggerated weaponry. *Animal Behaviour* **138**, 85–92.
- Fox, J. and Weisberg, S.** (2019). *An R Companion to Applied Regression*. Second. Thousand Oaks CA: Sage.
- Fox, C. W., Stillwell, R. C. and Moya-Laraño, J.** (2007). Variation in selection, phenotypic plasticity, and the ecology of

- sexual size dimorphism in two seed-feeding beetles. In *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism* (ed. Fairbairn, D. J., Blanckenhorn, W. U., and Székely, T.), pp. 88–96. New York, USA: Oxford University Press.
- Friard, O. and Gamba, M.** (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* **7**, 1325–1330.
- Fryxell, D. C., Weiler, D. E., Kinnison, M. T. and Palkovacs, E. P.** (2019). Eco-Evolutionary Dynamics of Sexual Dimorphism. *Trends in Ecology and Evolution* **34**, 591–594.
- García-Navas, V., Nogueras, V., Cordero, P. J. and Ortego, J.** (2017). Ecological drivers of body size evolution and sexual size dimorphism in short-horned grasshoppers (Orthoptera: Acrididae). *Journal of Evolutionary Biology* **30**, 1592–1608.
- Gottardo, M., Vallotto, D. and Beutel, R. G.** (2015). Giant stick insects reveal unique ontogenetic changes in biological attachment devices. *Arthropod Structure and Development* **44**, 195–199.
- Haley, E. L. and Gray, D. A.** (2012). Mating behavior and dual-purpose armaments in a camel cricket. *Ethology* **118**, 49–56.
- Hendry, C. R., Guiher, T. J. and Pyron, R. A.** (2014). Ecological divergence and sexual selection drive sexual size dimorphism in new world pitvipers (Serpentes: Viperidae). *Journal of Evolutionary Biology* **27**, 760–771.
- Herberstein, M. E., Painting, C. J. and Holwell, G. I.** (2017). Scramble competition polygyny in terrestrial arthropods. *Advances in the Study of Behavior* **49**, 237–295.
- Honěk, A.** (1993). Intraspecific variation in body size and fecundity in insects : a general relationship. *Oikos* **66**, 483–492.
- Hsiung, C.-C.** (1987). Aspects of the biology of the Melanesian stick-insect *Eurycantha calcarata* Lucas (Cheleutoptera: Phasmatidae). *Journal of Natural History* **21**, 1241–1258.
- Hsiung, C.-C. and Panagopoulos, D.** (1998). Preliminary Observations on the Effects of Food Plant on the Stick Insect *Eurycantha calcarata* Lucas (Cheleutoptera : Phasmatidae). *Journal of Orthoptera Research* **7**, 93–98.
- Kaliontzopoulou, A., Carretero, M. A. and Adams, D. C.** (2015). Ecomorphological variation in male and female wall lizards and the macroevolution of sexual dimorphism in relation to habitat use. *Journal of Evolutionary Biology* **28**, 80–94.
- Kelly, C. D.** (2004). Allometry and sexual selection of male weaponry in Wellington tree weta , *Hemideina crassidens*. *Behavioral Ecology* **16**, 145–152.
- Kelly, C. D.** (2006a). Resource quality or harem size : what influences male tenure at refuge sites in tree weta (Orthoptera : Anostostomatidae)? *Behavioral Ecology and Sociobiology* **60**, 175–183.
- Kelly, C. D.** (2006b). The Relationship Between Resource Control , Association with Females and Male Weapon Size in a Male Dominance Insect. *Ethology* **112**, 362–369.
- Kelly, C. D.** (2006c). Movement patterns and gallery use by the sexually dimorphic Wellington tree weta. *New Zealand Journal of Ecology* **30**, 273–278.
- Kelly, C. D.** (2006d). Fighting for harems : assessment strategies during male - male contests in the sexually dimorphic Wellington tree weta. *Animal behaviour* **72**, 727–736.
- Kelly, C. D.** (2007). Identifying a causal agent of sexual selection on weaponry in an insect. *Behavioral Ecology* **19**, 184–192.
- Kelly, C. D.** (2014). Sexual selection, phenotypic variation, and allometry in genitalic and non-genitalic traits in the sexually

- size-dimorphic stick insect *Micrarchus hystriculens*. *Biological Journal of the Linnean Society* **113**, 471–484.
- Kelly, C. D.** (2015). Male-biased sex ratios and plasticity in post-insemination behaviour in the New Zealand stick insect *Micrarchus hystriculens*. *Behaviour* **152**, 653–666.
- Kelly, C. D., Bussière, L. F. and Gwynne, D. T.** (2008). Sexual selection for male mobility in a giant insect with female-biased size dimorphism. *The American Naturalist* **172**, 417–423.
- Kimsey, L. S., Dewhurst, C. F. and Nyaure, S.** (2013). New species of egg parasites from the Oil Palm Stick Insect (*Eurycantha insularis*) in Papua New Guinea (Hymenoptera, Chrysididae, Phasmatodea, Phasmatidae). *Journal of Hymenoptera Research* **30**, 19–28.
- Kokko, H., Klug, H. and Jennions, M. D.** (2014). Mating systems. In *The evolution of insect mating systems* (ed. Shuker, D. M. and Simmons, L. W.), pp. 42–58. Oxford, United Kingdom: Oxford University Press.
- Kratochvíl, L. and Frynta, D.** (2007). Phylogenetic analysis of sexual dimorphism in eye-lid geckos (Eublepharidae): the effects of male combat, courtship behavior, egg size, and body size. In *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism* (ed. Fairbairn, D. J., Blanckenhorn, W. U., and Székely, T.), pp. 154–162. New York, USA: Oxford University Press.
- Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B.** (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* **82**, 1–26.
- Legrand, R. S. and Morse, D. H.** (2000). Factors driving extreme sexual size dimorphism of a sit-and-wait predator under low density. *Biological Journal of the Linnean Society* **71**, 643–664.
- Lenth, R.** (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means.
- Levinton, J. S. and Allen, B. J.** (2005). The paradox of the weakening combatant: Trade-off between closing force and gripping speed in a sexually selected combat structure. *Functional Ecology* **19**, 159–165.
- McCullough, E. L., Miller, C. W. and Emlen, D. J.** (2016). Why sexually selected weapons are not ornaments. *Trends in Ecology and Evolution* **31**, 742–751.
- McLain, D. K., Pratt, A. E. and Berry, A. S.** (2003). Predation by red-jointed fiddler crabs on congeners: interaction between body size and positive allometry of the sexually selected claw. *Behavioral Ecology* **14**, 741–747.
- Metz, M. C., Emlen, D. J., Stahler, D. R., Macnulty, D. R., Smith, D. W. and Hebblewhite, M.** (2018). Predation shapes the evolutionary traits of cervid weapons. *Nature Ecology & Evolution* **2**, 1619–1625.
- Monteith, G. B. and Dewhurst, C. F.** (2011). Does the phasmid *Eurycantha calcarata* Lucas, 1869 (Phasmida: Phasmatidae) occur in Australia? *Australian Entomologist* **38**, 179–196.
- Moya-Laraño, J., Vinković, D., Allard, C. M. and Foellmer, M. W.** (2009). Optimal climbing speed explains the evolution of extreme sexual size dimorphism in spiders. *Journal of Evolutionary Biology* **22**, 954–963.
- Myers, S. S., Buckley, T. R. and Holwell, G. I.** (2015). Mate detection and seasonal variation in stick insect mating behaviour (Phasmatodea: *Clitarchus bookeri*). *Behaviour* **152**, 1325–1348.
- O'Brien, D. M. and Boisseau, R. P.** (2018). Overcoming mechanical adversity in extreme hindleg weapons. *Plos One* **13**, e0206997.
- O'Brien, D. M., Allen, C. E., Van Kleeck, M. J., Hone, D., Knell, R., Knapp, A., Christiansen, S. and Emlen, D. J.** (2018). On the evolution of extreme structures: static scaling and the function of sexually selected signals. *Animal Behaviour* **144**, 95–108.

- O'Brien, D. M., Boisseau, R. P., Duell, M., McCullough, E., Powell, E. C., Somjee, U., Solie, S., Hickey, A. J., Holwell, G. I., Painting, C. J., et al. (2019). Muscle mass drives cost in sexually selected arthropod weapons. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20191063.
- Pacheco, Y. M. (2018). Ecomorph Convergence in Stick Insects (Phasmatodea) with Emphasis on the Lonchodinae of Papua New Guinea.
- Pérez-Barbería, F. J., Gordon, I. J. and Pagel, M. (2002). The origins of sexual dimorphism in body size in ungulates. *Evolution* **56**, 1276–1285.
- Peters, R. H. (1986). *The Ecological Implications of Body Size*. Cambridge, United Kingdom: Cambridge University press.
- Pincheira-Donoso, D. and Hunt, J. (2017). Fecundity selection theory: concepts and evidence. *Biological Reviews* **92**, 341–356.
- Preziosi, R. F., Fairbairn, D. J., Roff, D. A. and Brennan, J. M. (1996). Body size and fecundity in the waterstrider *Aquarius remigis*: a test of Darwin's fecundity advantage hypothesis. *Oecologia* **108**, 424–431.
- Puniamoorthy, N., Schäfer, M. A. and Blanckenhorn, W. U. (2012). Sexual selection accounts for the geographic reversal of sexual size dimorphism in the dung fly, *sepsis punctum* (diptera: Sepsidae). *Evolution* **66**, 2117–2126.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*.
- Rico-Guevara, A. and Hurme, K. J. (2019). Introsexually selected weapons. *Biological Reviews* **94**, 60–101.
- Robertson, J. A., Bradler, S. and Whiting, M. F. (2018). Evolution of oviposition techniques in stick and leaf insects (Phasmatodea). *Frontiers in Ecology and Evolution* **6**, 1–15.
- Rohner, P. T., Blanckenhorn, W. U. and Puniamoorthy, N. (2016). Sexual selection on male size drives the evolution of male-biased sexual size dimorphism via the prolongation of male development. *Evolution* **70**, 1189–1199.
- Rowe, L. (1992). Convenience polyandry in a water strider: foraging conflicts and female control of copulation frequency and guarding duration. *Animal Behaviour* **44**, 189–202.
- Rowe, L., Westlake, K. P. and Currie, D. C. (2006). Functional significance of elaborate secondary sexual traits and their evolution in the water strider genus *Rheumatobates*. *Canadian Entomologist* **138**, 568–577.
- Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.
- Shuker, D. M. and Simmons, L. W. (2014). *The evolution of insect mating systems*. Oxford, United Kingdom: Oxford University Press.
- Simmons, L. W. (2001). *Sperm Competition and its Evolutionary Consequences in the Insects*. Princeton, New Jersey, USA: Princeton University Press.
- Simmons, L. W. (2014). Sperm competition. In *The evolution of insect mating systems* (ed. Shuker, D. M. and Simmons, L. W.), pp. 181–203. Oxford, United Kingdom: Oxford University Press.
- Singmann, H., Bolker, B., Westfall, J. and Aust, F. (2019). afex: Analysis of Factorial Experiments.
- Sivinski, J. (1979). Intrasexual aggression in the stick insects *Diapheromera veliei* and *D. covilleae* and sexual dimorphism in the Phasmatodea. *Psyche* **5**, 395–405.
- Stephens, P. R. and Wiens, J. J. (2009). Evolution of sexual size dimorphisms in emydid turtles: Ecological dimorphism, rensch's rule, and sympatric divergence. *Evolution* **63**, 910–925.

- Stuart-Fox, D. and Moussalli, A.** (2007). Sex-specific ecomorphological variation and the evolution of sexual dimorphism in dwarf chameleons (*Bradypodion* spp.). *Journal of Evolutionary Biology* **20**, 1073–1081.
- Thornhill, R. and Alcock, J.** (1983). *The evolution of insect mating systems*. Harvard University Press.
- Toubiana, W. and Khila, A.** (2019). Fluctuating selection strength and intense male competition underlie variation and exaggeration of a water strider's male weapon. *Proceedings of the Royal Society B: Biological Sciences* **286**,
- Wertheim, B., Van Baalen, E. J. A., Dicke, M. and Vet, L. E. M.** (2005). Pheromone-mediated aggregation in nonsocial arthropods: An evolutionary ecological perspective. *Annual Review of Entomology* **50**, 321–346.
- Wu, H., Jiang, T., Huang, X. and Feng, J.** (2018). Patterns of sexual size dimorphism in horseshoe bats: Testing Rensch's rule and potential causes. *Scientific Reports* **8**, 2616.

Supplementary information

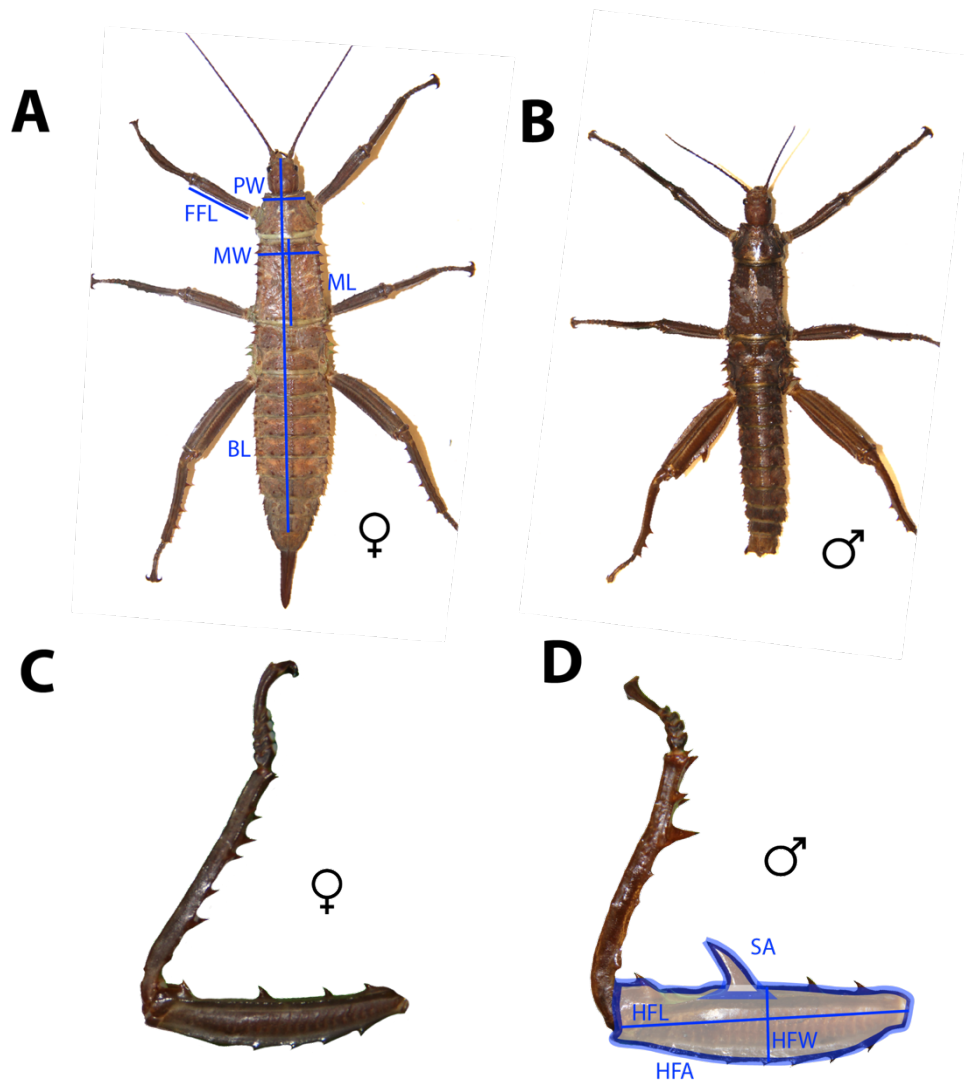


Figure S1: Sexually dimorphic morphology of *Eurycantha calcarata* (captive population). **A.** Adult female in dorsal view. BL: Body length (up the 9th abdominal segment); ML: Mesonotum length; MW: Mesonotum width; PW: Prothorax width; FFL: Front Femur length. **B.** Adult male in dorsal view. **C.** Adult female hind leg in lateral view. **D.** Adult male hind leg in lateral view. HFL: Hind Femur length; HFW: Hind Femur width; SA: Spine area; HFA: Hind Femur area. Photos by first author.

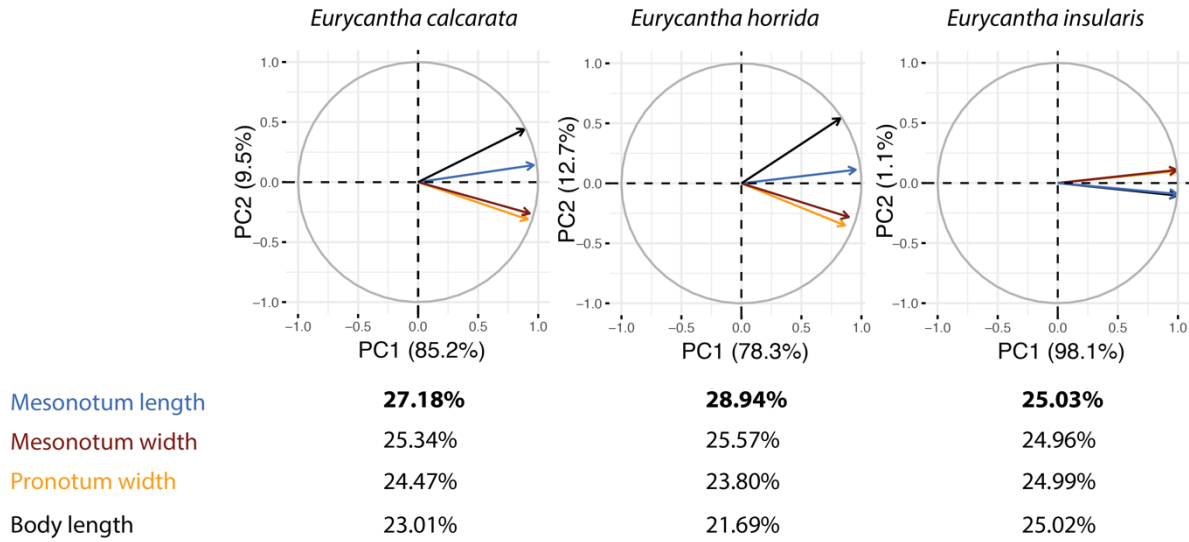


Figure S2: Results of a Principal Component Analysis (PCA) including different linear measurements of body size of adults *Eurycantha calcarata* (captive), *Eurycantha horrida* and *Eurycantha insularis*. Contributions of each variable to PC1 are indicated for each species.

Table S1: Summary of morphological measurements (mean \pm SE)

Species	Sex	Mesonotum Length (mm)	Front femur length (mm)	Hind femur length (mm)	Hind femur width (mm)	Hind femur area (mm ²)	Hind femoral spine area (mm ²)	N
<i>E. calcarata</i> (wild)	M	19.8 \pm 0.22	18.30 \pm 0.21	26.8 \pm 0.32	8.77 \pm 0.13	197.1 \pm 4.89	13.1 \pm 0.43	73
	F	22.9 \pm 0.16	19.33 \pm 0.17	NA	NA	NA	NA	71
<i>E. calcarata</i> (captive)	M	24.1 \pm 0.19	22.55 \pm 0.26	35.6 \pm 0.41	8.72 \pm 0.12	286.1 \pm 5.91	19.9 \pm 0.61	28
	F	26.6 \pm 0.31	22.78 \pm 0.24	30.5 \pm 0.33	5.89 \pm 0.08	175.3 \pm 3.75	1.8 \pm 0.08	40
<i>E. horrida</i> (wild)	M	27.9 \pm 0.6	24.68 \pm 0.52	38.8 \pm 0.80	9.86 \pm 0.34	353.1 \pm 15.65	38.4 \pm 2.83	8
	F	29.1 \pm 0.44	24.7 \pm 0.37	32.5 \pm 0.69	5.85 \pm 0.12	177.0 \pm 7.91	3.4 \pm 0.25	9
<i>E. insularis</i> (wild)	M	17.0 \pm 0.12	16.57 \pm 0.13	19.0 \pm 0.16	3.01 \pm 0.03	52.4 \pm 0.82	0.48 \pm 0.02	40
	F	23.1 \pm 0.21	20.43 \pm 0.2	23.7 \pm 0.25	3.74 \pm 0.06	80.8 \pm 1.61	0.37 \pm 0.01	25

Table S2: Summary of adult scaling relationships. Results corresponds to ordinary least squares regressions.

Trait	Species	Sex	Intercept	95% CI	Slope	95% CI	N	F _(df)	P
log ₁₀ Front femur length ~ log ₁₀ Mesonotum length (body size)	<i>E. calcarata</i> (wild)	Male	0.26	0.06 0.47	0.77	0.61 0.93	73	91.4 _(1,71)	<0.001
		Female	0.34	-0.11 0.66	0.70	0.46 1.02	71	28.5 _(1,66)	<0.001
	<i>E. calcarata</i> (captive)	Male	-0.15	-0.68 0.39	1.08	0.69 1.47	28	32.6 _(1,26)	<0.001
		Female	0.24	-0.02 0.5	0.78	0.60 0.97	40	73.6 _(1,38)	<0.001
	<i>E. borrida</i> (wild)	Male	0.38	-0.34 1.10	0.70	0.20 1.20	8	7.27 _(1,6)	0.04
		Female	0.26	-0.41 1.92	0.77	0.32 1.23	9	10.8 _(1,7)	0.01
	<i>E. insularis</i> (wild)	Male	0.24	-0.07 0.48	0.80	0.61 1.05	40	49.5 _(1,36)	<0.001
		Female	0.13	-0.21 0.50	0.87	0.59 1.11	25	46.2 _(1,22)	<0.001
log ₁₀ Hind femur length ~ log ₁₀ Mesonotum length (body size)	<i>E. calcarata</i> (wild)	Male	0.27	0.08 0.45	0.90	0.75 1.04	73	147.8 _(1,71)	<0.001
	<i>E. calcarata</i> (captive)	Male	-0.12	-0.61 0.38	1.21	0.85 1.57	28	47.5 _(1,26)	<0.001
		Female	0.43	0.18 0.67	0.74	0.57 0.91	40	76.2 _(1,38)	<0.001
	<i>E. borrida</i> (wild)	Male	0.56	-0.13 1.25	0.71	0.24 1.19	8	8.3 _(1,6)	0.03
		Female	0.42	-0.84 1.67	0.75	-0.11 1.60	9	2.82 _(1,7)	0.14
	<i>E. insularis</i> (wild)	Male	0.56	0.13 0.88	0.59	0.33 0.93	40	24.1 _(1,38)	<0.001
		Female	0.38	-0.10 0.81	0.73	0.41 1.08	25	19.7 _(1,22)	<0.001
	log ₁₀ Hind femur width ~ log ₁₀ Mesonotum length (body size)	<i>E. calcarata</i> (wild)	Male	-0.44	-0.69 -0.19	1.07	0.88 1.26	73	117.2 _(1,71)
<i>E. calcarata</i> (captive)		Male	-0.46	-1.25 0.33	1.01	0.44 1.58	28	13.2 _(1,26)	0.001
		Female	-0.56	-0.96 -0.16	0.93	0.65 1.21	40	45.4 _(1,38)	<0.001
<i>E. borrida</i> (wild)		Male	-0.68	-1.83 0.48	1.16	0.36 1.95	8	7.75 _(1,6)	0.03
		Female	0.24	-1.40 1.87	0.36	-0.75 1.48	9	0.39 _(1,7)	0.55
<i>E. insularis</i> (wild)		Male	-0.05	-0.50 0.39	0.43	0.07 0.79	40	5.49 _(1,38)	0.02
		Female	-0.08	-1.27 0.86	0.48	-0.22 1.35	25	1.67 _(1,22)	0.21

log ₁₀ Hind femur area ~	<i>E. calcarata</i> (wild)	Male	-0.11	-0.48 0.26	1.85	1.56 2.14	73	159.2 _(1,71)	<0.001
	<i>E. calcarata</i> (captive)	Male	-0.49	-1.46 0.48	2.13	1.43 2.83	28	38.81 _(1,26)	<0.001
		Female	-0.21	-0.74 0.33	1.71	1.34 2.09	40	84.33 _(1,38)	<0.001
log ₁₀ Mesonotum length (body size)	<i>E. borrida</i> (wild)	Male	-0.03	-1.22 1.16	1.78	0.96 2.61	8	17.2 _(1,6)	0.006
		Female	-1.27	-3.19 0.65	2.40	1.09 3.71	9	12.47 _(1,7)	0.01
	<i>E. insularis</i> (wild)	Male	0.26	-0.55 0.76	1.19	0.78 1.84	40	18.41 _(1,38)	<0.001
		Female	-0.08	-1.12 0.75	1.46	0.84 2.23	25	19.77 _(1,22)	<0.001
log ₁₀ Spine area ~ log ₁₀ Mesonotum length (body size)	<i>E. calcarata</i> (wild)	Male	-1.86	-2.39 -1.33	2.28	1.88 2.69	73	119.6 _(1,71)	<0.001
	<i>E. calcarata</i> (captive)	Male	-2.61	-4.30 -0.92	2.82	1.60 4.05	28	22.5 _(1,26)	<0.001
		Female	-4.43	-5.77 -3.08	3.27	2.33 4.21	40	49.3 _(1,38)	<0.001
	<i>E. borrida</i> (wild)	Male	-2.14	-4.21 -0.06	2.57	1.14 4.01	8	11.85 _(1,6)	0.01
		Female	-3.43	-8.61 1.76	2.69	-0.85 6.24	9	2.16 _(1,7)	0.19
	<i>E. insularis</i> (wild)	Male	-2.99	-5.31 -0.67	2.15	0.27 4.03	40	5.0 _(1,38)	0.03
		Female	-4.91	-6.46 -3.17	3.29	2.00 4.42	25	28.59 _(1,22)	<0.001

Video S1: The defensive behaviours of a male and female *E. calcarata* and of a male *E. borrida*. Video by first author.

Video S2: Males and females *E. calcarata* and *E. insularis* walking with radiotransmitters. Video by first author.

Video S3: Contest between two male *E. calcarata* over access to two females on a host tree trunk. Video by first author.

Video S4: Escalated male combat in *E. calcarata* over access to a female on a host tree trunk. Video by first author.

Video S5: Male *E. calcarata* guarding territories on a host tree trunk and intercepting and mating with a female. Video by first author.

Video S6: Male *E. calcarata* preventing a rival from entering a tree cavity. Video by first author.

CHAPTER 6

The function of male exaggerated weaponry in a gregarious stick insect

The function of male exaggerated weaponry in a gregarious stick insect

Romain P. Boisseau^{1,*} & Douglas J. Emlen¹

¹ Division of Biological Sciences, University of Montana, 32 Campus Dr, Missoula, MT 59812, United States of America

*Corresponding author: Romain P. Boisseau

Keywords: sexual selection | sexual conflict | phasmatodea | parthenogenesis | contest theory | assessment

Abstract

Sexually selected weapons often function as both tools of combat and threatening signals in contests against same-sex rivals. The way that weapons are used is dependent on the assessment strategy males use during fights to decide whether they should retreat or keep on fighting. Are males assessing each other when fighting and, if so, are they using their weapons as a signal of fighting ability? We investigated the function of the enlarged male hindlegs in the New Guinean thorny devil stick insects, *Eurycantha calcarata*, in the context of male-male and male-female interactions. Using field and lab experiments, we show that larger males with proportionately larger hindlegs are more likely to win fights over access to females and to subsequently mate. Analyses of contest behavioral sequences and predictors of contest costs support a mutual assessment strategy. Interestingly, males did not appear to use their hindlegs as signals of fighting ability but rather used other tactile and chemical cues for opponent assessment. Hindlegs were exclusively used to deliver powerful squeezes in rare but dangerous escalated fights. Males also used their hindlegs when copulating to stabilize their copulation position but, as females did not appear to resist male copulation attempts, we did not find any evidence to support the use of male hindlegs in coercion. Our results suggest that male hindlegs in *E. calcarata* are used as pure force-delivering combat tools and not as signals despite males assessing each other during contests.

Introduction

Sexually dimorphic structures bearing exaggerated proportions in a single sex are widespread in animals (Andersson, 1994; Emlen, 2008; Lavine et al., 2015; Rico-Guevara and Hurme, 2019). Such charismatic traits are often the result of sexual selection acting more strongly in the sex with the highest variance in reproductive success, typically males (Andersson, 1994; Shuker and Simmons, 2014). Sexual selection may result from female choice which drives the evolution of ornaments in males, like the remarkably elongated tails of widow birds (Andersson, 1982) or the flamboyant fins of guppy fishes (Kodric-Brown, 1985). Alternatively, sexual selection may result from direct male combats (i.e., intrasexual competition) for access to females or critical resources or territories (Emlen, 2014). This scenario may lead to the evolution of weapons through modifications of diverse types of appendages, such as the greatly elongated forelegs of harlequin beetles (Zeh et al., 1992), or novel structures such as the enormous horns of some scarab beetles (Emlen et al., 2007; Ohde et al., 2018).

Intra-sexually selected weapons are first and foremost tools of battle involving rivals of the same species and sex (Rico-Guevara and Hurme, 2019). However, many harbor additional functions sometimes unrelated to the competition for mating. For instance, enlarged claws of male fiddler crabs are effective predator deterrents (Bildstein et al., 1989; McLain et al., 2003), and male fallow and red deer use their antlers to harass and coerce females into mating (Clutton-Brock and Parker, 1995; Pradhan and Van Schaik, 2009). But more commonly, in addition to being fighting devices, weapons also serve as visual or tactile signals used to threaten rivals and advertise body size and fighting prowess or to attract females (Berglund et al., 1996; Muramatsu, 2011; Pratt et al., 2003; Rometsch et al., 2021). As such, these structures can be described along a weapon to signal continuum (McCullough et al., 2016). The position of intra-sexually selected structures along this continuum should determine the relative importance of biomechanical performance versus conspicuousness (McCullough and O'Brien, 2022). Weapons that function mainly as visual signals of male fighting abilities are expected to display more elaboration and positive static allometries (i.e., large individuals displaying disproportionately large versions of the trait) that exacerbate size differences between contestants and thereby facilitate opponent assessment (Eberhard et al., 2018; O'Brien et al., 2018; Rodríguez and Eberhard, 2019). In contrast, weapons used purely as tools applying force in combat may be biomechanically constrained to keep lever components in proportion to avoid the so-called “paradox of the weakening combatant” (i.e., weapons becoming relatively weaker as they get larger) (Levinton and Allen, 2005; O'Brien and

Boisseau, 2018). Therefore, the functional details of weapons and the different contexts in which they are used critically influence their evolution.

How weapons are used in combat is intrinsically dependent on assessment strategies, and therefore understanding the evolution of weaponry is reliant on understanding how fighting animals make decisions on when to start, escalate or retreat from a fight, how long to persist, how much energy to expend, and how much cost and damage to endure (Hardy and Briffa, 2013). The study of decision making during contests is central to the field of contest theory and is best understood in terms of game theory which investigates factors affecting optimal choices for decision makers when other deciding agents are present (Maynard Smith, 1974; Maynard Smith, 1982; Maynard Smith and Price, 1973). The decision to withdraw from a fight is made by the loser and the sources of information that weigh into that decision may be diverse and may include the value of the contested resource (Arnott and Elwood, 2008; Maynard Smith and Parker, 1976), past fighting experiences (Goubault and Decuignière, 2012; Hsu et al., 2006; Rutte et al., 2006) and, most importantly, the fighting abilities or resource holding potential (RHP) of the contestants (Arnott and Elwood, 2009; Maynard Smith and Parker, 1976; Parker, 1974).

Theoretical studies have modelled a variety of assessment strategies of RHP during contests which may be divided into two main types: self-assessment and mutual assessment models (Arnott and Elwood, 2009; Chapin et al., 2019; Taylor and Elwood, 2003). Under self-assessment strategies, contestants are only able to assess their own state and RHP endogenously during fights. These include the pure self-assessment models, in which each contestant can only access information about their own RHP or current physiological state and only endures costs resulting from their own actions (e.g., energetic) (Mesterton-Gibbons et al., 1996; Payne and Pagel, 1996). Under this scenario, fighting ceases when the weaker individual (i.e., with the lowest RHP) reaches its internal pre-set cost threshold first (Arnott and Elwood, 2009). Self-assessment strategies also include cumulative assessment models (CAM) for which costs can also be inflicted by the action of the opponent (e.g., through injuries) (Payne, 1998). Under this scenario, the weaker individual reaches its threshold first because it can bear fewer costs (i.e., lower internal threshold) and gets inflicted costs at a higher rate by the stronger opponent (Arnott and Elwood, 2009). In mutual assessment models, contestants are able to assess their own and their opponent's RHP and the loser gives up when it has assessed that it is inferior (Enquist and Leimar, 1983). These models make clear predictions that empiricists have tested in many systems (reviewed in Arnott and Elwood 2009; Elwood and Arnott 2012; Green and Patek 2018;

Chapin et al. 2019b; Pinto et al. 2019) (**Figure 1**). Mutual assessment presents the selective advantage that the weaker contestant can quickly terminate an unmatched confrontation and thus avoid paying a cost in time, energy or injury for a contest that it would indubitably lose. Yet, empirically, self-assessment strategies have received support in relatively more systems (Elwood and Arnott, 2013; Pinto et al., 2019).

Integrating studies of the diversity of functions of intra-sexually selected weapons, allometry and contest assessment strategies is critical to further our understanding of the evolution of these structures. For example, weapons that display a positive allometry and that putatively function as threat signals are only expected in systems where males assess their opponent RHP during combat. Here, we investigate the fighting behavior and weaponry function of males in the New Guinean thorny devil stick insects, *Eurycantha calcarata* (Lucas, 1869) (Lonchodinae, Phasmatodea) using laboratory and field data. In this nocturnal species, the relatively large males (compared to other phasmid species) harbor dramatically enlarged hind femora endowed with a large sharp spine (Buckley et al., 2009). Field observations revealed that this species displays a territory- and female-defense mating system where males fight with rivals, notably using their hindlegs, for strategic positions close to tree cavities where adults communally roost during the day. Territory holders are then able to intercept females as they leave their shelter at dusk, and copulate (Boisseau et al., 2020). Puncture injuries inflicted by the femoral spines of rivals are common on the bodies of males (mainly on their hind femora) highlighting the hazardous nature of these weapons. While field observations clearly suggest that these hindlegs are used during male combat for access to females, a quantitative link between body or weapon size and fighting success is yet to be established. Males also use their hindlegs as defenses against predators in an elaborate startle posture by reaching out with these armored legs, ready to strike them together (Bedford, 1976; Boisseau et al., 2020; Carlberg, 1989). Finally, males wrap one of their hindlegs around the female's abdomen when copulating (Hsiung, 1987), suggesting yet another potential function for these legs in female coercion, as is the case in many other insects with sexually dimorphic hindlegs (Burrows, 2020; Haley and Gray, 2012; Rowe et al., 2006).

In the present study, we tested the hypotheses that the male hindlegs of *E. calcarata* function as pure weapons, threat signals and coercive tools. We first investigated the effects of body and weapon size on fight outcomes and the link between fighting and mating success. In parallel, we tested the predictions of different assessment models to determine the most likely form of assessment that males are using during contests and whether they could be using their hindlegs as signals of RHP to rivals.

Finally, we investigated whether males could benefit from larger hindlegs by coercing females into copulating faster and whether females would benefit from resisting males and avoiding copulation to reproduce entirely through parthenogenesis, as was suggested in another phasmid species (Burke and Bonduriansky, 2022; Burke et al., 2015).

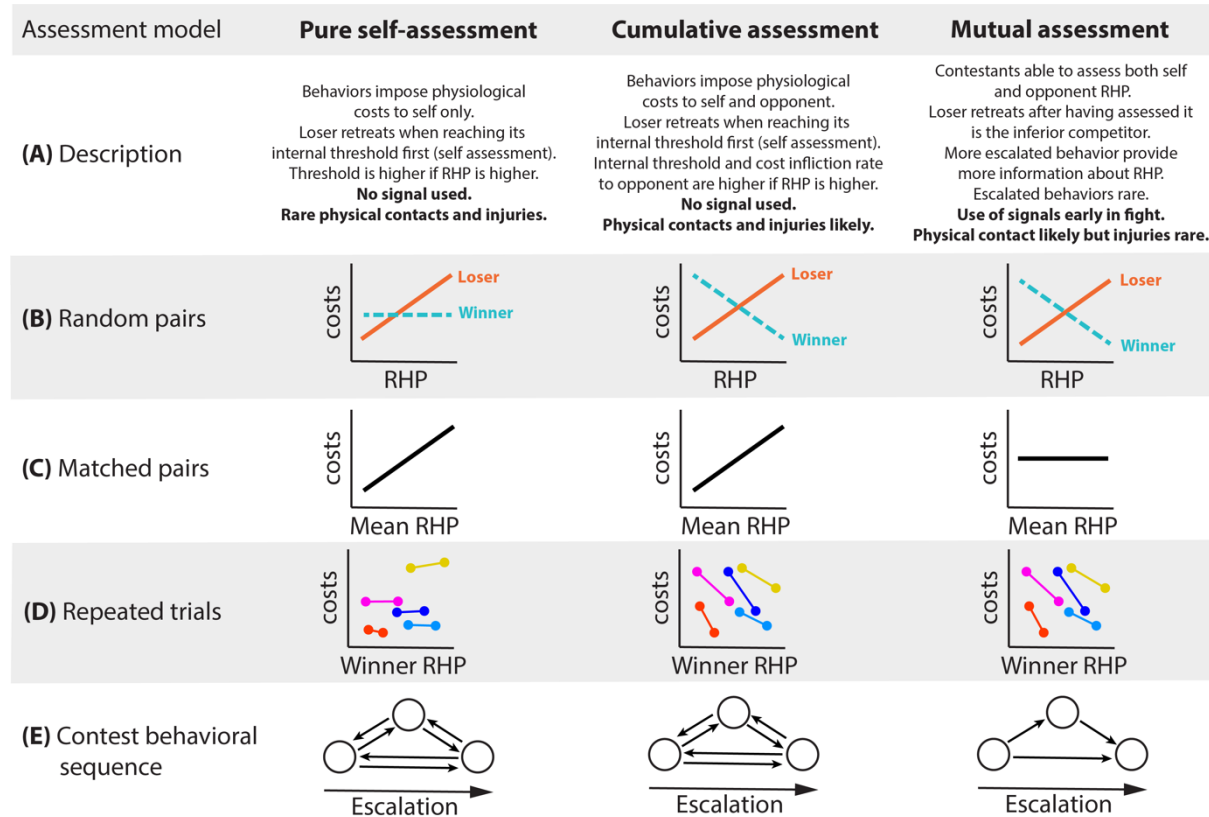


Figure 1: Main assessment models and associated predictions. The rationale of each model is described and predictions for the use of signals, physical contact, and injuries are indicated in bold (A). Models predict different relationships between contest costs and RHP for losers (orange line) and winners (turquoise line) of randomly matched contests (B), and for the averaged RHP of contestants in matched contests (black line) (C) (Arnott and Elwood, 2009; Green and Patek, 2018). For trials where focal animals with a low RHP (i.e., losers) are assigned to multiple opponents with higher RHP (i.e., winners), models also predict different relationships between winner RHP and contest cost for each focal animal (D) (Chapin et al., 2019). For each focal animal (represented by a different color), a negative slope indicates that winner RHP affects contest cost, which is only expected under the cumulative or mutual assessment models. No relationship is expected under a pure self-assessment strategy. This approach is used to discern heterogeneity in the strategy used by individuals within a population. Finally, models predict different trends in the sequence of behaviors during contests (circles correspond to behaviors, arrows represent most likely transitions between them) (E) (Green and Patek, 2018).

Materials and methods

Statistical analyses are detailed in each relevant section of the Materials and Methods. All analyses in this study were performed using R (v4.1.1, R Core Team 2021) and we systematically checked the normal distributions of the residuals and the absence of any specific patterns in their distribution in all linear models. R functions and packages used are indicated throughout the materials and methods as *function*: ‘package’.

Study animals and measurements. We used a breeding population of *E. calcarata* obtained from the Missoula Butterfly house and Insectarium (MT, USA) and originally collected around Kimbe (West New Britain, PNG) in the late 1970s. The insects were housed in Missoula, MT in transparent plastic containers (65 × 45 × 50cm) at 22°C, on 12h:12h light:dark cycles, sprayed with water daily (RH=50-80%) and fed maple leaves (*Acer platanoides*) *ad libitum*. We provided fresh leaves in the summer (June-September, fourth instar to adulthood) or dried summer leaves for the rest of the year (first to fourth instar). Males and females were reared in the same cages (~50 individual per container) from hatching to their final molt. Adult males and females were then kept separate.

To investigate the ontogenetic allometry of the hindlegs, we used photographs (EOS 600D, Canon, Tokyo, Japan) and ImageJ (v 1.51, Schneider et al. 2012) to measure mesothorax length (ML), right front femur length (FFL), right hind femur length (HFL) and width (HFW) in 47 males and 66 females at every instar starting at the fourth instar until adulthood. ML was shown to be the best proxy for body size in a previous study (Boisseau et al., 2020), and was also used as such in the present study. We calculated hind femur area (HFA) as the area of an ellipse of major diameter HFL and minor diameter HFW ($HFA = \pi \times \frac{HFL}{2} \times \frac{HFW}{2}$). We confirmed that this proxy was strongly correlated with the area measured by directly drawing the outline of the femur in adults (Figure S1). To investigate changes in scaling relationships between body size and front or hind leg size during postembryonic development, we ran linear mixed models (LMM, *lme*: ‘nlme’, Pinheiro et al. 2021) using either log₁₀-transformed FFL, HFL, HFW or HFA as response variable and log₁₀ ML, sex and instar as predictor variables as well as all two-way and three-way interactions. Individual ID was added as a random factor to account for repeated measurements of the same individuals throughout development. A sequential ANOVA (*anova.lme*: ‘nlme’) was used to assess the significance of the fixed effects. Non-significant interaction terms (p>0.05) were sequentially deleted from the model to provide the most accurate

parameter estimates. Only the final models are presented. Departures from isometry (i.e., slope= 1 for linear traits, slope= 2 for surface traits) were tested in adults using 95% confidence intervals around estimated regression slopes.

Field observations. We investigated male contest behaviors in the wild near Kimbe, WNB, PNG (Dami palm plantations, S5° 31.846' E150° 20.221') in parallel to a previously published study (Boisseau et al., 2020). The behaviors of adults around daily roosting cavities were recorded during four entire nights on a *Kleinhovia hospita* trunk bearing 12 different cavities sheltering groups of *E. calcarata*. Behaviors were video-recorded using two cameras in time lapse mode (HERO4, GoPro, San Mateo, CA, USA; interval: 0.5s) under red illumination, from 4PM to 8AM (sunrise ~5:30AM, sunset ~5:50PM). For more methodological details, see the methods of Boisseau et al. (2020).

We used the event-logging software BORIS (v. 7.5.3, Friard and Gamba 2016) to observe and record fight behaviors from the videos. An interaction between two males was considered a fight when the two individuals made contact (including with their antennae) and one individual clearly retreated from the fighting area at the end of the interaction (i.e., the loser) while the other did not (i.e., the winner). We observed a total of 33 fights involving 31 different pairs of contestants. Fights took place over territories on the trunk at the base of large branches, close to a cavity entrance, or directly over access to a female (Boisseau et al., 2020). For each fight, an individual was randomly designated as the focal individual, while the other was designated as the opponent. The outcome of a fight was classified as 0 if the focal male lost the fight, or 1 if it won. In addition, the focal male was designated as the owner of the territory or female if it was the resident male guarding the resource or the first present at the resource, or as the intruder otherwise. Recorded fighting behaviors corresponded to the list of behaviors in **Table 1** and included antennal contacts all the way to hind leg grasping and squeezing. Behaviors were recorded regardless of which contestant performed them. While some males were measured and marked prior to the days of recording, many males involved in observed agonistic interactions were unmarked and could not be measured. Therefore, as absolute contestant measurements were often unavailable, we only assessed relative body size difference between contestants by measuring the body length of both contestants directly on the videos (in pixels).

Table 1: Ethogram and description of the recorded fighting behaviors.

Behavior	Description
Start	Beginning of the fight: first interaction between the two contestants.
Antennate	The contestant touches its opponent (often repeatedly) with its antennae.
Kick	The contestant briefly kicks its opponent with either its front, middle or hind legs.
Push	The contestant lunges towards its opponent
Chase	The contestant chases its opponent who is running away
Mount	The contestant climbs on top of its opponent
Descend	The contestant dismounts its opponent
Turn	The contestant does a 180° rotation (often while on top of its opponent)
Back	The contestant moves backward towards its opponent
Grasp	The contestant grabs and squeezes its opponent with its hindlegs (often while on top of it)
Knock	The contestant repeatedly hammers the substrate with the end of its abdomen
End	End of the fight: the loser retreats away from the winner.

Set-up of lab trials. We performed fighting trials in a large glass container (90 × 45 × 45cm) in the laboratory (Figure S2). The bottom of the fighting arena was covered with moist brown paper towel, dried maple leaves and two Petri dishes filled with water serving as drinking troughs. To mimic a roosting cavity, we placed a piece of maple bark (*Acer platanoides*, 35 × 16cm) against the glass on one end of the arena. We drilled a hole (∅ 4cm) in the bark piece to allow the insects to enter the artificial cavity (between the glass wall and the bark). We glued a 3cm wide band of cardboard around the edges of the bark piece to prevent the insects from accessing the cavity from the sides. The container was lit with white fluorescent lights during the day (12h) or with red lights at night (12h). The inside of the cavity was constantly lit in red light and sheltered from white light during the day. Two cameras in time lapse mode (HERO4, GoPro, San Mateo, CA, USA; interval: 0.5s) were placed to film the inside and outside of the cavity through the glass (Figure S2). The females were not provided dirt to bury their eggs to prevent egg laying and to be able to measure their fecundity (see below). No eggs were laid during the trials.

Randomly-matched trials. To investigate how body and weapon size affected fighting success, dominance among males and mating success, we staged trials involving 3 randomly picked adult males and 3 randomly picked adult females using our first generation of lab-bred insects. The insects were introduced simultaneously in the arena and were weighed prior and after the trial (ME104T/00, Mettler Toledo, Columbus, OH, USA). We filmed their interactions during four consecutive days, changing the memory cards in the cameras, spraying water and adding food as needed every 24h. 17 such trials could be run involving 22 different males and 17 females. Thus, most individuals were used in several

trials which was accounted for in the statistical analyses (see below). Individuals (males and females) were involved in a trial approximately every month (± 1 week), starting two weeks after their final molt (to insure sexual maturity) and until they died. Trials always involved different combinations of individuals. Some females ($n=12$) involved in these mixed trials and who all mated with the males were individually followed to record their fecundity (see section **Effect of mating on female fitness**), and corresponded to the ‘mated’ treatment. As for field observations, for every male-male fight outside the cavity on the bark or on the floor, we recorded the outcome, who was the owner of the resource and the detailed sequence of agonistic behaviors.

We also looked at male interactions inside the cavity during the daytime. Field observations previously revealed that larger males were associated with more female-biased groups inside cavities (Boisseau et al., 2020). Males were also observed excluding rivals from roosting cavities suggesting that male-male contests could also happen inside those cavities. Our video recordings inside the cavities during the lab trials revealed that one male would typically sit on or close to the females and exclude the other males to the opposite corner away from the females, sometimes by actively pushing them or using its hindlegs (Video S1). Individual contests were difficult to delineate as losers could not clearly retreat away from the winner. Thus we considered fights to last the entire time two males were found together in the cavity with at least one female. The male sitting the closest to the females and clearly successfully guarding them was considered the winner. Day periods for which only one male was present in the cavity or for which the winner was unclear were not considered.

Body size-matched trials. To be able to test the effect of relative weapon size on fight outcome and distinguish between the cumulative and mutual assessment models (**Figure 1D**), we used our second generation of insects (only originating from sexual reproduction) to perform body size-matched contests. During these trials ($n=21$), only 2 males and 1 female were simultaneously introduced and filmed in the arena for 3 days. Pairs of males were matched by mesothorax length (i.e., body size) and did not differ by more than 5%. We used 40 different males, among which two were involved in two trials, and 21 different 2-weeks old virgin females. In other respects, we followed the same procedures detailed above.

Fight outcome analyses. For all field, randomly- and size-matched lab fights, we calculated the size asymmetry between contestants using the following size asymmetry index:

$$\left\{ \begin{array}{l} \text{Size asymmetry index} = \frac{\text{Focal size}}{\text{Opponent size}} - 1, \text{ if Focal size} \geq \text{Opponent size} \\ \text{Size asymmetry index} = -\left(\frac{\text{Opponent size}}{\text{Focal size}} - 1\right), \text{ if Focal size} < \text{Opponent size} \end{array} \right. \quad \text{(Equation 1)}$$

where size corresponded to body length for field fights or mesothorax length, body mass and hind femur area for lab fights. This size asymmetry index is an analog of the sexual dimorphism index of Lovich and Gibbons (1992) which presents the advantage of being symmetrical and centered on zero regardless of which sex or, in our case contestant, is larger.

We tested the effect of body length asymmetry index on focal individual contest outcome (win or lose) using a generalized linear mixed-effects model (GLMM; family: binomial, logit link function; *glmer*, “lme4”). Body length asymmetry index was scaled to unit variance and was subsequently included as a fixed effect in the model, along with ownership (owner or intruder). Male pair ID was included as a random intercept to account for fight replications between the same two males. It should however be noticed that our analyses could not account for pseudoreplication of individual males involved in several fights with different opponents. Statistical significance of fixed effects was assessed sequentially using chi-square tests (*anova*: “stats”).

For lab fights, we calculated size asymmetry indexes using either mesothorax length (i.e., body size), body mass (i.e., average between mass measured before and after the trial) and hind femur area (i.e., weapon size) as size measurements (**Equation 1**). In total, we observed 140 randomly-matched and 188 size-matched contests outside the roosting cavity involving respectively 51 and 21 different male pairs, and 73 randomly-matched contests inside the roosting cavity involving 33 different male pairs. We compiled the contest data in a long format (sensu Briffa et al. 2013) for all types of contest separately, including two rows per fight (one for each contestant) to be able to run GLMMs accounting for the replication of contestants. To test for the effects of body size, relative weapon size, relative body mass and ownership on the probability of winning a fight, we ran GLMMs (binomial, logit link function; *glmer*: “lme4”) including whether the contestant won or not the contest as the response variable, mesothorax length, hind femur area, body mass and ownership as fixed effects and fight ID,

contestant ID and trial ID as random effects. Mesothorax length, hind femur area and body mass were \log_{10} -transformed, scaled to unit variance and centered on zero prior to the analysis. The significance of the fixed effects was sequentially assessed using chi-square tests (*anova*: “stats”).

Sequential behavioral analyses. Assessment models differ in their predictions regarding whether contest behaviors should occur in phases of increased escalation or more randomly (**Figure 1E**). We used a sequential analysis to test how male fighting behaviors fit these predictions. Following Green and Patek (2018) and using the R package “igraph” (Csardi and Nepusz, 2006), we combined the behavioral sequences from all observed field, randomly- and size-matched lab contests separately. We summarized the three compiled behavioral sequences into three separate adjacency matrices. Rows and columns in the matrices represented one of the 12 contest behaviors outlined in **Table 1**. Each cell in the matrix contained the number of times a transition occurred from behavior 1 (row) to behavior 2 (column). These numbers were then divided by the total number of transitions to obtain frequencies. We isolated transitions from each behavioral sequence that occurred more often than expected by chance by generating 10,000 adjacency matrices keeping the observed frequencies of behaviors but randomizing the transitions between them. These matrices were obtained by randomly resampling the second column of the two-column behavioral transition datasets containing the end behavior of each transition. Transitions from a fighting behavior to “start” or from “end” to a fighting behavior were disallowed as no behavior can happen before the fight starts or after it ends. Transitions from “mount” to “mount” or from “descend” to “descend” were also disallowed as a contestant cannot mount its opponent twice in a row without descending first. The same logic applies for descending. Thus, we obtained, for each behavioral sequence, a null distribution of the frequencies of each transition expected if behaviors transitioned randomly. We extracted the 95% quantile of the null distribution for each cell of the adjacency matrices (i.e., each transition) and kept observed values more frequent than their respective 95% null quantile. These “significant” transitions occurred more often than expected randomly. We visualized the three adjacency matrices deprived of “non-significant” transitions as a network graph. Each individual behavior was represented as a network vertex (i.e., circle) whose size was proportional to the relative frequency of the behavior. Each significant transition was represented as a network edge (i.e., arrow) whose width was proportional to the probability of the transition.

These networks were used to test the presence of phases during fights. Phases correspond to subsets of behaviors of similar frequencies that often transition into one another. A phase begins or

ends with a transition to a new subset of behaviors. After a phase ends, behaviors from that phase are unlikely to re-occur (Enquist et al., 1990; Green and Patek, 2018).

Correlational tests of assessment models. Assessment models also differ in their predictions regarding the correlation between contestant RHP and overall contest cost (**Figure 1B-D**). Contest duration and maximum escalation level are often used as proxies for contest cost (e.g., Fea and Holwell 2018a; Green and Patek 2018). We used a composite metric of contest cost integrating both the duration and intensity aspect of the fight. We assigned an escalation score to each behavior and calculated contest cost as the sum of the scores of all the behaviors used during the fight. Our behavioral networks revealed that fights in male *E. calcarata* unfold in phases of increasing escalation level: contestants start using low intensity phases with rare or brief physical contact and transition into more and more intense phases involving more physical contact and more dangerous behaviors (e.g., hind leg grasping) (see **Results**). Out of the 6 identified phases, behaviors from the first and the last phase were assigned an escalation score of one as they involved low intensity behaviors, respectively antennal contacts and descending and substrate knocking. Other behaviors were assigned an escalation score of 2 to 5 depending on which phase they belonged to (phase 2 to 5). Later phases involved more physical and dangerous contacts (i.e., going from kicking to mounting the opponent, all the way to grasping with the armored hindlegs). Therefore, our metric of contest cost is higher for longer and/or more intense fights as longer fights will include more individual behaviors and more intense fights will include behaviors with higher escalation scores. However, it should be mentioned that despite being based on the phases identified through our network analyses, defining degrees of escalation is relatively subjective and biased by the perception of human observers (Green and Patek, 2018).

First, we investigated the correlation between contest cost and RHP asymmetry between contestants. This relationship is expected to be negative in mutual assessment models as a larger RHP asymmetry between contestants should be more readily detected and lead to less escalated and shorter fights. However, a negative relationship can also incidentally appear in self-assessment systems as a result of the correlation between loser RHP and RHP difference (Taylor and Elwood, 2003). Therefore such a negative correlation between contest cost and RHP asymmetry is predicted by mutual assessment but not sufficient to rule out other assessment strategies. For this analysis, we combined fights from our randomly- and size-matched lab trials and ran a linear mixed-effects model (LMM, *lmer*: “lme4”) including \log_{10} -transformed contest costs as the response variable and the absolute value of the mesothorax length asymmetry index as fixed effect. Pair ID, trial ID and year (i.e., the two

generations of lab insects respectively used in the randomly- or size-matched trials) were including as random effects. Significance of the fixed effect was assessed using a chi-square test (*Anova*: “car”).

We then looked at the correlation between loser and winner RHP (here approximated by mesothorax length) and contest costs in our randomly-matched fights (**Figure 1B**). Similarly, we ran LMMs including \log_{10} -transformed contest costs as the response variable and either loser or winner \log_{10} mesothorax length as the only fixed effect. Loser (or winner) ID, pair ID, trial ID and day (i.e., if the fight occurred on the first, second, third or fourth day of the trial) were included as random effects.

For size-matched fights, we investigated the relationship between average contestant RHP and contest costs (**Figure 1C**). We built a LMM including \log_{10} -transformed contest costs as the response variable, \log_{10} average contestant mesothorax length as fixed effect and pair ID, trial ID and day as random effects.

Finally, as our randomly-matched trials involved three males, we were able to use the “repeated trial” approach suggested by Chapin et al. (2019), to test for heterogeneity in assessment strategies used by different individuals of the same population (i.e., some males may assess their opponent while others may only assess their own RHP). Out of the three males involved in a given trial, we designated a focal loser (i.e., with the smallest mesothorax length) and looked at the effect of winner RHP (i.e., mesothorax length) on contest cost for each focal loser individually (**Figure 1D**) (Chapin et al., 2019). We built a LMM including \log_{10} -transformed contest costs as the response variable, \log_{10} opponent (i.e., winner) mesothorax length, focal individual (i.e., loser) ID and their interaction as fixed effects. Opponent ID and trial ID were included as random effects. Significance of the fixed effects was assessed using sequential chi-square tests (*anova*: “stats”).

Correlation between fighting and mating success. We also used our first generation of lab insects (i.e., used in randomly-matched trials involving 3 males and 3 females) to investigate whether winning more fights translated into mating more often for males. Specifically, we recorded the number of copulations of each male (i.e., absolute mating success) and calculated each male’s relative mating success by dividing their absolute mating success by the average absolute mating success of the three males in the same trial. Fighting success was assessed as a male’s dominance rank among the triad. We assigned dominance ranks based on the overall outcome of contests of each male with the two others in the trial. A male was considered dominant over another one if it won more contests overall against that rival over the 4 days of the trial. The male dominating its two other rivals was assigned a dominance

rank of one and so on. In cases of draws between two males (i.e., same number of victories against one another or no fight), they were both assigned the rank one if they were both dominant over the third male (which was assigned the rank three), or the rank two if they were both submissive to the third male (which was assigned the rank one). For each male, we assigned two dominance ranks respectively based on contests outside or inside the cavity, averaged the two ranks and rounded to the nearest inferior integer to obtain an overall dominance rank used as a proxy for fighting success. Finally, we ran a cumulative link mixed model (*clmm*: “ordinal”) to test the relationship between fighting and mating success in these lab trials, including dominance rank (considered as an ordered factor) as the response variable, relative mating success as the only fixed effect and trial ID and male ID as random effects.

Sexual conflict and male-female interactions. In parallel to investigating male-male interactions, again, we used our first generation of insects (i.e., used in randomly-matched trials involving 3 males and 3 females) to investigate male-female interactions, and uncover a potential coercive function of the male’s hindlegs. Specifically, for each observed mating, we recorded the latency of males to mate (i.e., duration between the first contact between the male and the female and the onset of copulation) and copulation duration. If males use their enlarged hindlegs to force females into copulating, we predicted that larger males with larger hindlegs would be able to overcome female resistance (if present at all) and initiate copulation faster and copulate for longer. In addition, female reproductive status (i.e., virgin versus already mated) should also be considered as virgin females may either be more willing to mate to fertilize their eggs or be more resistant to male copulation attempts to avoid egg fertilization and reproduce parthenogenetically. We ran a LMM including either latency to mate or copulation duration as the response variable, male mesothorax length, male hind femur area, female mesothorax length and female mating status (i.e., mated or not) as fixed effects, and male ID, female ID and trial ID as random effects. Continuous variables were all \log_{10} -transformed, centered on zero and scaled to unit variance prior to the running the analyses.

Effect of mating on female fitness. In order to assess the fitness effects of reproducing sexually or parthenogenetically, a subset of females from our first generation was never used in the randomly-matched trials described above and instead were involved in analogous female-only trials during which 6 females were put together for 4 days in a similar arena and in similar conditions. Each of these females were submitted to such trials roughly once a month (± 1 week), starting 2 weeks after their final molt and until they died. These females that were never exposed to males at the adult stage

corresponded to the ‘parthenogenetic’ treatment (n=16). They were compared to females used in the mixed trials described above (‘mated’ treatment, n=12). In between trials, females from both treatments were kept individually in small containers (35 × 20 × 15cm) containing a dirt box (13 × 9 × 5cm) for egg laying and a small trough filled with water for drinking. They were fed with dried maple leaves *ad libitum*. We cleaned their container, collected, counted and weighed their eggs every week. Female survival was also monitored daily. Eggs were kept on moist sterilized compost at 22°C. Hatching rate was recorded for each female. Incubation duration was measured for each female as the duration between the laying of the first egg and the hatching of the first offspring. Upon hatching, nymphs were placed in small boxes (35 × 20 × 15cm) containing 50 individuals grouped by female treatment (mated or parthenogenetic). Offspring of different females were mixed up to insure the same density of individuals inside the boxes. They were fed dried maple leaves and were sprayed water daily. We monitored offspring survival until they successfully reached their second instar. We tested the effect of female reproductive treatment (i.e., parthenogenesis versus sex) on adult female survival by comparing the survival curves of both treatments (*survdiff*: “survival”). Then we tested its effect on fecundity by building linear models (*lmer*: “stats”) including lifetime egg number, average egg mass, egg incubation duration, egg laying rate (calculated after the first egg was laid) or egg hatching rate as response variable, and female mesothorax length and treatment as explanatory variables. Continuous variables were log₁₀-transformed prior to the analyses. A type I ANCOVA was used to assess the significance of the explanatory variables. Finally, to test the effect of female treatment on offspring first instar survival, we built a contingency table including the number of dead and live nymphs for each female treatment and a Fisher’s exact test (*fisher.test*: “stats”).

Results

Ontogenetic scaling relationships. Males had relatively longer front femurs and relatively longer and wider hind femurs than females and this difference increased in the final two instars (**Figure 2**, Table S1). As they reached later instars, both males and females harbored relatively longer front femurs and relatively longer and wider hind femurs (**Figure 2**, Table S1). Males and females did not differ in allometric slope and adults exhibited slopes consistent with isometry for all measured morphological traits (**Figure 2**, Table S1). Specifically, hind femur length and width did not scale more steeply than isometry or front femur length (**Figure 2**, Table S1). Here, front femur length was used as a reference

trait as front legs putatively function as unspecialized legs in this species (O'Brien et al., 2018). Therefore, exaggeration of the male hindlegs in *E. calcarata* is the result of an increase in intercept between hindleg and body size that arises late in development.

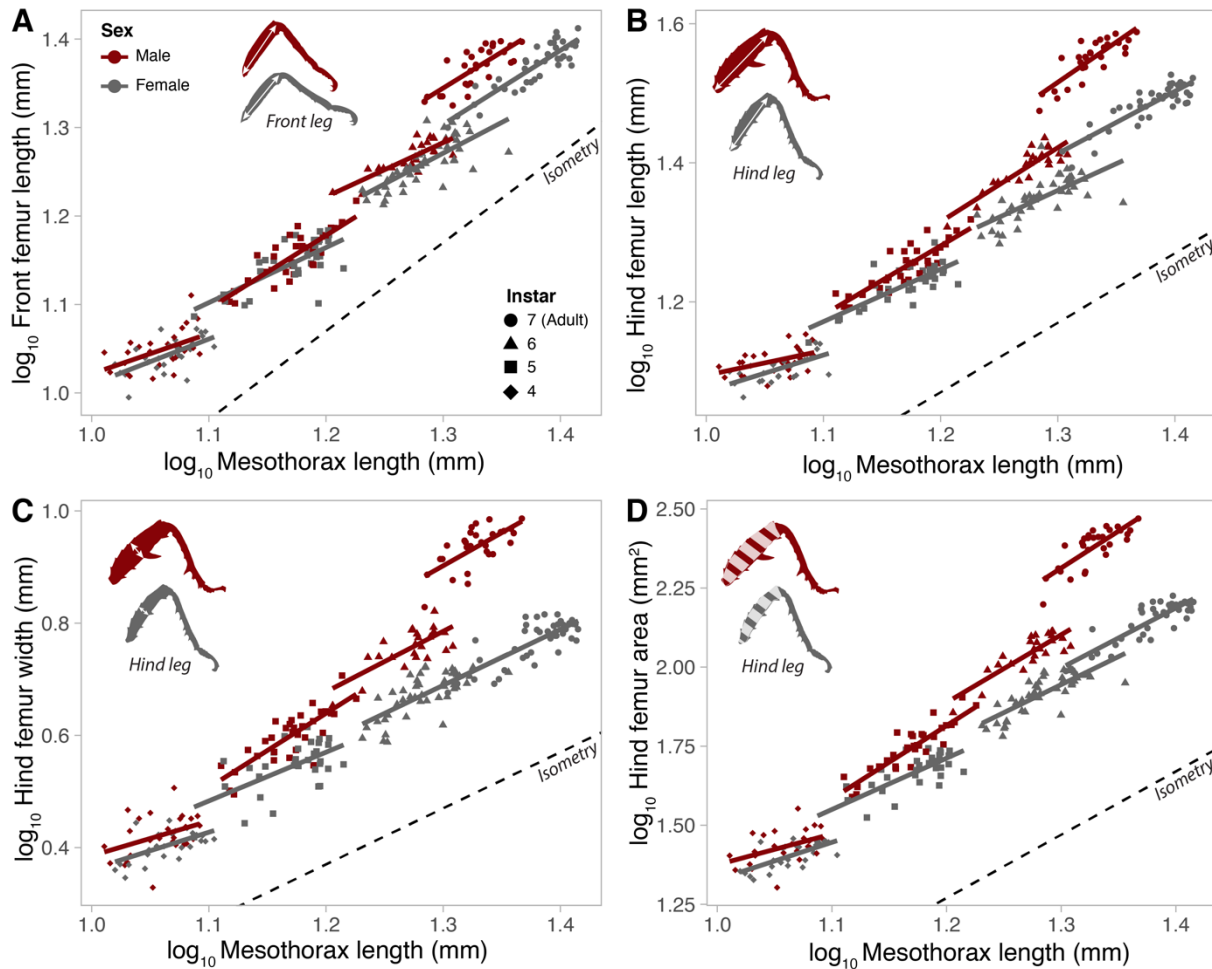


Figure 2: Sexual dimorphism and scaling relationships between body size and front or hind leg size across the second half of postembryonic development in lab-reared *E. calcarata*. Scaling relationships between front femur length (A), hind femur length (B), hind femur width (C) and hind femur area (D) and mesothorax length (~body size) for males and females across the last three nymphal instars and adults. The dashed line represents an isometric slope (arbitrary intercept). Leg drawings in the top left corner illustrate the trait measured. Corresponding statistical analyses are reported in Table S1.

Fight outcomes. In the field, we found that body length difference ($\chi^2= 17.3$, $df=1$, $p<0.0001$) and ownership ($\chi^2= 7.01$, $df=1$, $p=0.008$) significantly affected contest outcome (**Figure 3A**): the larger contestant had a higher probability of winning and owners were also relatively advantaged compared to intruders of similar size. The interaction between body length difference and ownership did not have a significant effect on contest outcome ($\chi^2= 0.47$, $df=1$, $p=0.49$).

In our randomly matched lab contests, we found a significant effect of mesothorax length ($\chi^2= 18.24$, $df=1$, $p<0.0001$) and ownership ($\chi^2= 40.03$, $df=1$, $p<0.0001$) on contest outcomes outside the roosting cavity. The effect of hind femur area ($\chi^2= 0.42$, $df=1$, $p=0.52$) and body mass ($\chi^2= 0.12$, $df=1$, $p=0.73$) were not significant after accounting for mesothorax length. Therefore, larger individuals were more likely to win and resident males were also relatively advantaged against intruders of similar size (**Figure 3B**). Similarly, mesothorax length positively affected the probability of winning contests inside a cavity ($\chi^2= 5.93$, $df=1$, $p=0.015$), but the effect of hind femur area ($\chi^2= 2.33$, $df=1$, $p=0.13$) and body mass ($\chi^2= 1.06$, $df=1$, $p=0.30$) were not significant after accounting for mesothorax length.

In our size-matched lab contests, mesothorax length ($\chi^2= 0.004$, $df=1$, $p=0.95$), hind femur area ($\chi^2= 0.14$, $df=1$, $p=0.71$) and body mass ($\chi^2= 1.63$, $df=1$, $p=0.20$) did not significantly affect contest outcome which was only affected by ownership ($\chi^2= 29.7$, $df=1$, $p<0.0001$) (**Figure 3C**). Therefore, when body size-matched, resident males are more likely to win and differences in relative weapon size or body condition do not seem to matter.

We found a negative relationship between fighting rank and relative mating success ($z= -2.10$, $p = 0.035$) (**Figure 3D**). More dominant males that won more fights (i.e., with a lower fighting rank) were thus more likely to mate with the females during the 4-day trials.

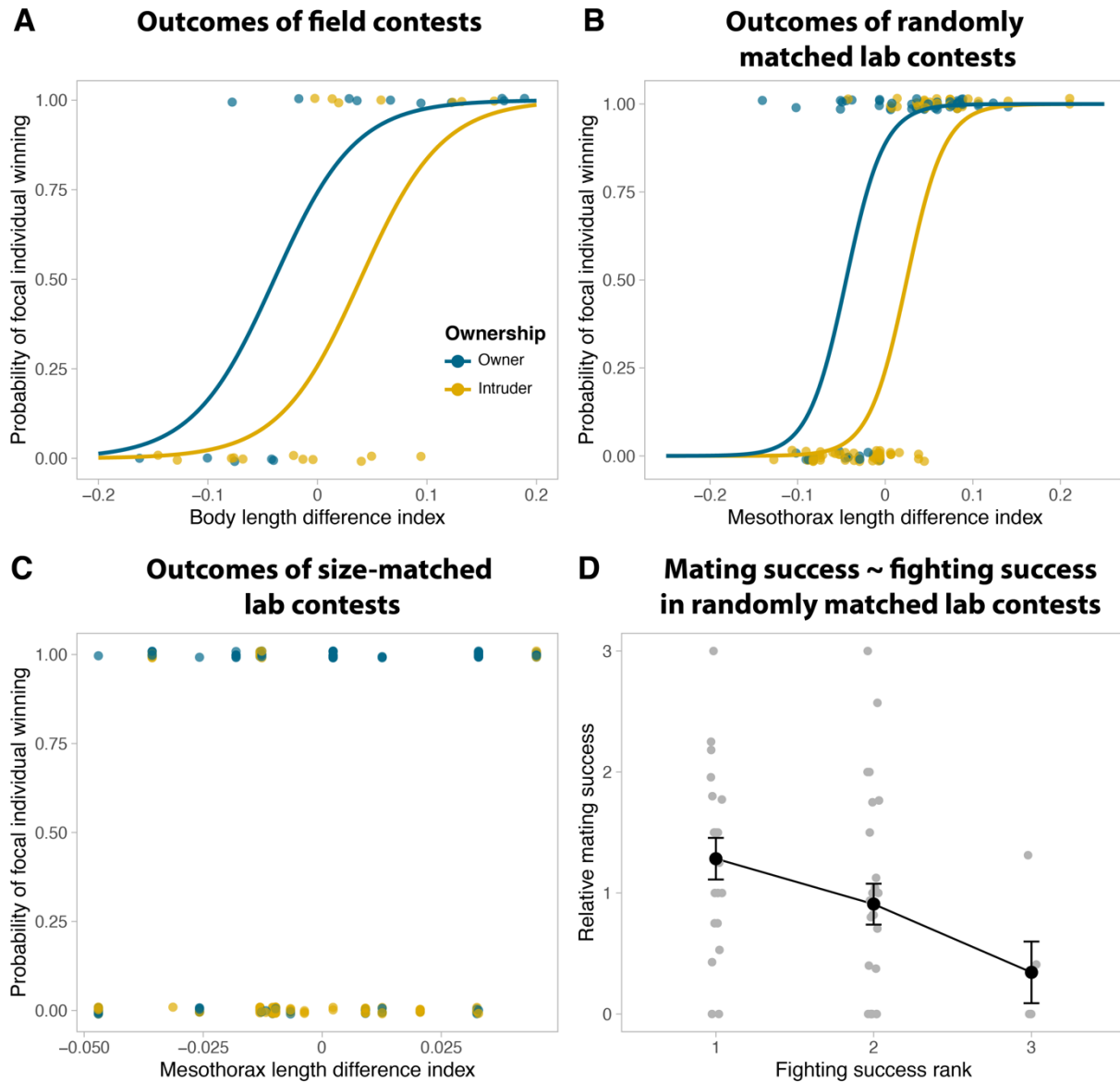


Figure 3: Male fight outcomes and correlation between fighting and mating success. Binary GLMM of (A) body length difference index in the field, mesothorax length difference index in randomly matched lab trials (B) and mesothorax length difference index in size-matched lab trials (C) against contest outcome. Colors represent the ownership status of the focal individual (resident/owner in blue, intruder in yellow). Correlation between relative mating success and male fighting rank in randomly-matched trials involving three males and three females. Black points correspond to means per fighting rank and error bars indicate standard errors to the mean.

Sequential behavioral analyses. Our analyses of behavioral sequences during fights showed that behaviors progressed in phases which matched the predictions of mutual assessment (**Figure 1E, 4**). In both randomly- and size-matched lab contests, and in field contests, fights started from (1) tactile and chemosensory antennal contacts between the contestants and transitioned to (2) brief kicks and pushes, to either (3) a chase if one contestant ran away, or to (4) one contestant mounting on the other one and putting itself in position (i.e., by turning and backing up) for (5) grasping with its weaponized hindlegs. Each phase was likely to deescalate and transition directly towards (6) the end of the fight. We often observed the winner thumping on the substrate (i.e., bark) with the tip of its abdomen at the end of fights (Video S2), while the loser was retreating. Within a phase, behaviors were used in similar frequencies. No significant transition occurred between one phase and previously-occurring phases (**Figure 4B-D**). Behaviors pertaining to lower, less escalated phases were used more often than behaviors of later, more escalated phases (**Figure 4B-D**).

More escalated behaviors (e.g., grasping) were more often observed in the lab than in the field. This suggests that the reduced space available in the lab trials may have artificially prevented losers from retreating quickly from winners (e.g., when stuck in the corner of the container), triggering the onset of more escalated phases. Field contests clearly suggest that the use of hindlegs is rare and that most fights terminate after antennal contacts or brief pushes (e.g., Video S3). Nevertheless, significant transitions between behaviors were largely similar in the field and in the lab.

In accordance with predictions from mutual assessment, size-matched lab contests showed more backwards movements and grasping (i.e., most escalated/dangerous behaviors) compared to randomly-matched contests (8.0 % versus 6.4%, and 4.4 % versus 4.0% of all behaviors respectively, **Figure 4B-C**).

When using their hindlegs, males usually sit on top of their opponent and squeeze their opponent's hind femora. This is consistent with the puncture injuries that have been observed in wild populations (Boisseau et al., 2020) and in the lab (this study). Fighting behaviors point to the use of tactile, chemical (via antennal contacts) and vibrational (via abdominal thumping) signals for assessment. Hindlegs were never waved or directly gauged visually or tactilely which suggests that these weapons only function as a tool of combat and not as a signal.

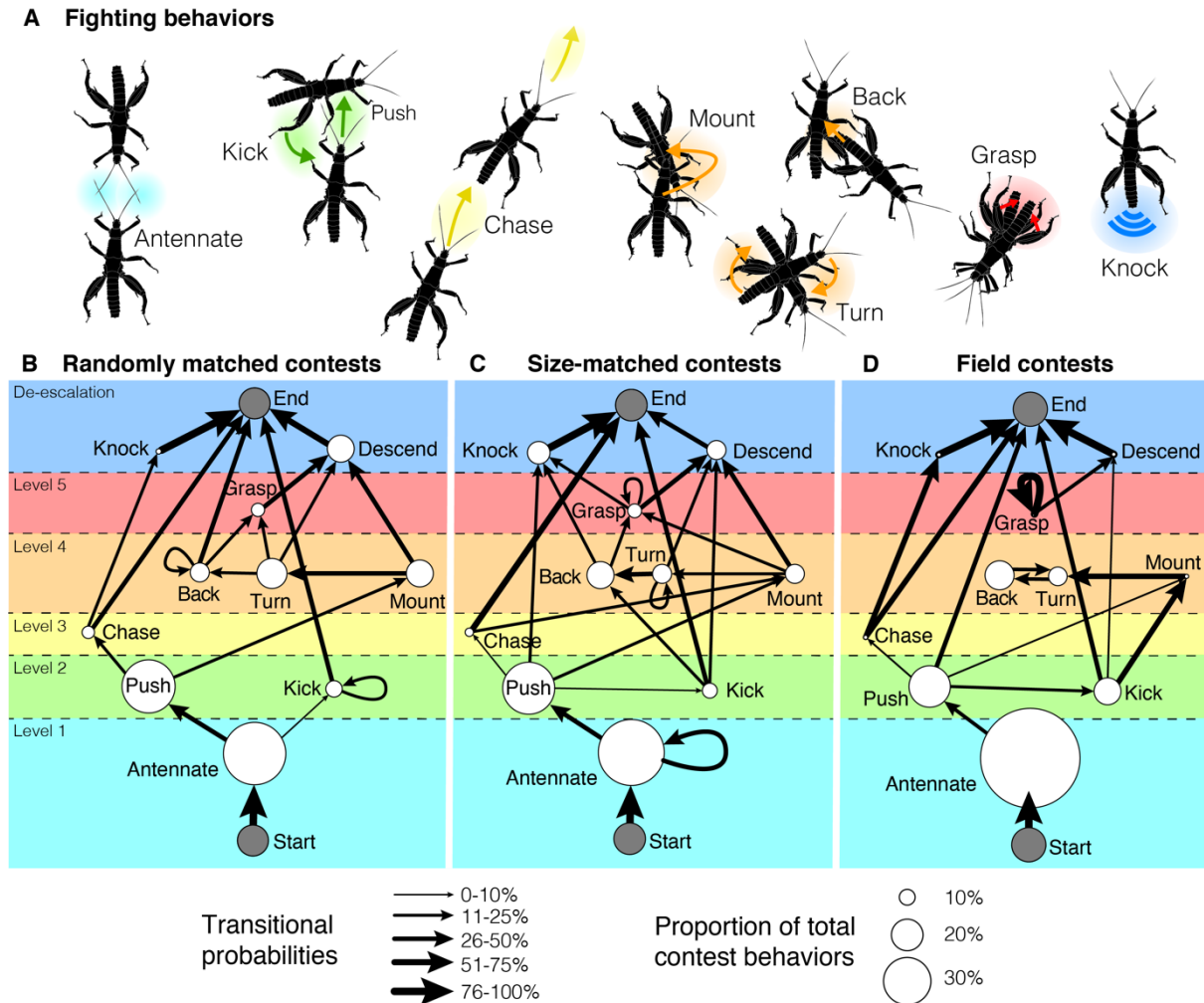


Figure 4: Contest behaviors (A) and sequential analysis in (B) randomly matched and (C) body size-matched lab contests, and (D) field contests. Contest behaviors progressed in phases (represented by different colors) of increasing intensity and escalation level (level 1 to 5) eventually leading to the resolution of the contest and its termination (de-escalation). Only significant transitions (i.e., occurring more often than expected randomly) are shown (arrows). The thickness of the arrows is proportional to transitional probability. No significant transition towards past phases that already occurred were observed. Individual behaviors are represented by circles, the size of which is scaled to the proportion of total contest behaviors represented by the given behavior.

Correlational tests of assessment models. We found a negative relationship between contest costs and contestant mesothorax length asymmetry ($\chi^2 = 23.3$, $df=1$, $p < 0.0001$, **Figure 5A**). While this relationship is predicted by mutual assessment, it can also appear under self-assessment models because of the correlation between loser RHP and RHP asymmetry, and therefore should not be used to rule out self-assessment (Taylor and Elwood, 2003). We further found a positive relationship between loser RHP (i.e., mesothorax length) and contest costs in our randomly matched contests ($\chi^2 = 6.87$, $df=1$, $p = 0.009$, **Figure 5B**) which was expected under all models. More interestingly, we found

a negative relationship between winner RHP and contest costs ($\chi^2= 4.85$, $df=1$, $p= 0.028$, **Figure 5B**) which ruled out a pure self-assessment strategy. Size-matched contests further revealed that average contestant RHP was only marginally correlated with contest costs ($\chi^2= 3.61$, $df=1$, $p= 0.06$, **Figure 5C**), thereby supporting mutual assessment to the detriment of cumulative assessment. Finally, our analyses of repeated contests for focal losers revealed that overall contest costs decreased with winner RHP (i.e., in accordance with mutual assessment) (effect of \log_{10} mesothorax length: $\chi^2= 8.40$, $df=1$, $p= 0.004$, **Figure 5D**) and that individuals differed in intercept (effect of focal loser ID: $\chi^2= 25.1$, $df=1$, $p= 0.009$) and slope (effect of interaction: $\chi^2= 37.9$, $df=1$, $p< 0.0001$). However, for 11 out of 12 focal losers, the estimated slope was negative suggesting that males unanimously used mutual assessment (**Figure 5D**).

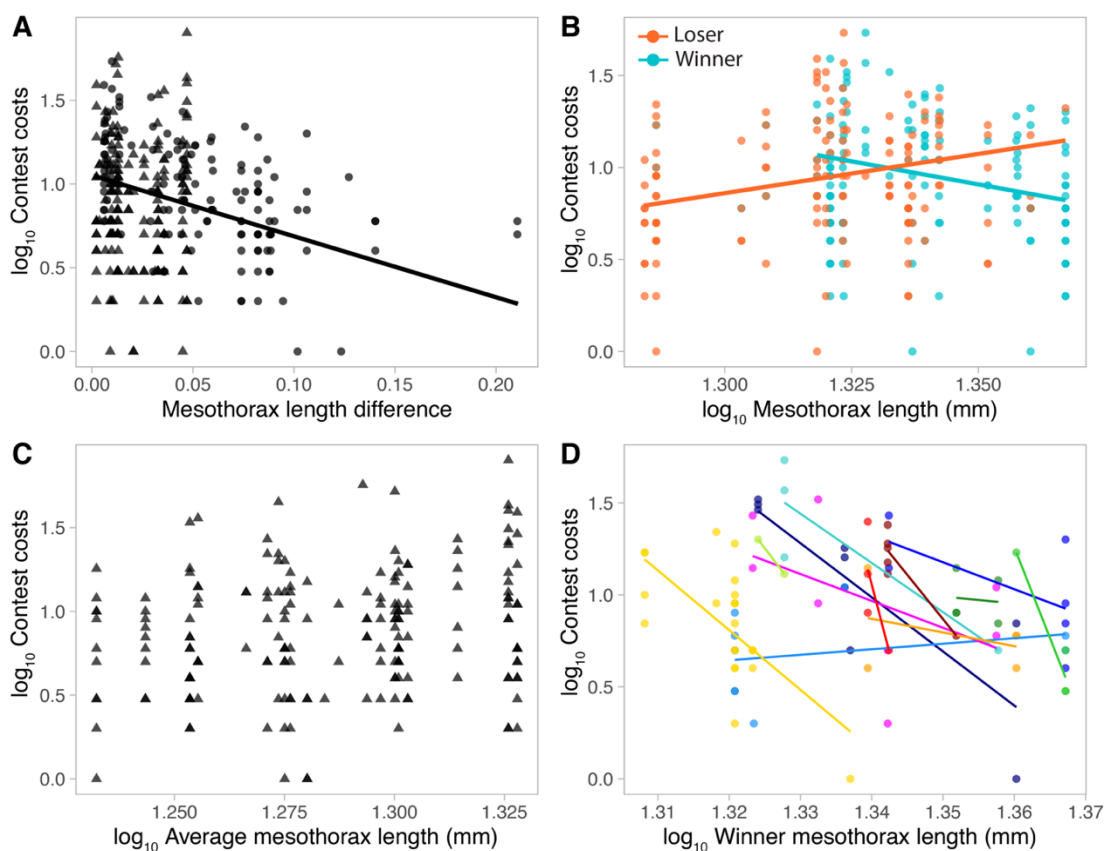


Figure 5: Correlational tests support mutual assessment during male-male contests. Relationships between contest costs and (A) mesothorax length asymmetry (i.e., absolute value of mesothorax length asymmetry index) in both randomly (circles) and size-matched contests (triangles), \log_{10} -corrected (B) loser (orange) and winner (blue) mesothorax length in randomly matched contests, (C) average contestant mesothorax length in size-matched contests. The repeated-testing approach is also presented for randomly matched contests (D): colors correspond to different focal losers which fought against several different contestants. Winner mesothorax length negatively affected contest costs for most focal losers.

Sexual conflict and male-female interactions. Before copulation, male *E. calcarata* usually made contact with females with their antennae and quickly mounted on their back, aligned themselves and slid their bodies to the right side of the female's body, holding the female's abdomen with the left hind leg (**Figure 6A**). They then passed the lower part of their abdomen under the female to reach and attach to the female's genitalia after some twisting and stretching (**Figure 6A**). Upon being mounted, females usually stopped moving and stayed largely immobile until the end of copulation. No specific female resistance behavior was observed. The hind leg of the male appeared to help bend the female's abdomen slightly to the right, making it easier for the male's genitalia to come into contact with the female's.

Latency to mate was not affected by male mesothorax length ($\chi^2 = 2.80$, $df=1$, $p = 0.09$), hind femur area ($\chi^2 = 1.85$, $df=1$, $p = 0.17$, **Figure 6B**), female mesothorax length ($\chi^2 = 0.10$, $df=1$, $p = 0.75$) or female reproductive status ($\chi^2 = 1.63$, $df=1$, $p = 0.20$, **Figure 6B**). Copulation duration was not correlated with male mesothorax length ($\chi^2 = 0.56$, $df=1$, $p = 0.46$) or hind femur area ($\chi^2 = 0.01$, $df=1$, $p = 0.92$) but was significantly affected by female mesothorax length ($\chi^2 = 9.41$, $df=1$, $p = 0.002$, **Figure 6C**) and female reproductive status ($\chi^2 = 6.32$, $df=1$, $p = 0.01$, **Figure 6C**).

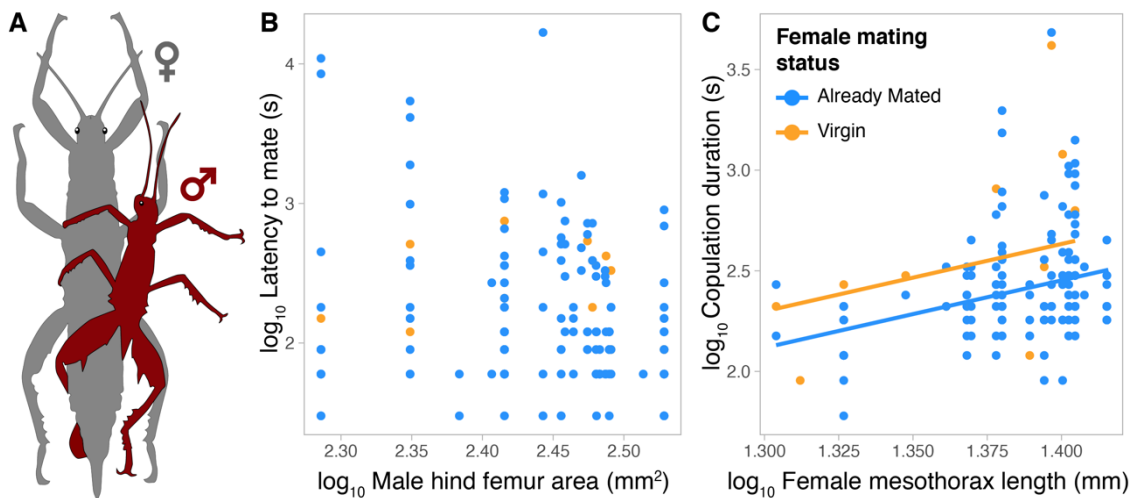


Figure 6: Male-female interactions in *E. calcarata*. (A) Typical copulation position. (B) Larger males with larger hind legs did not initiate copulation faster. (C) Males copulated for longer with both larger and virgin females.

Effect of mating on female fitness. Larger females did not lay significantly more eggs than smaller ones ($F_{1,25} = 0.79$, $p = 0.38$) but mated females laid significantly more eggs than parthenogenetic ones

($F_{1,25} = 6.72$, $p = 0.02$, **Figure 7A**). Average egg mass was neither affected by female mesothorax length ($F_{1,25} = 2.83$, $p = 0.10$) nor mating treatment ($F_{1,25} = 0.13$, $p = 0.72$). Incubation duration was not affected by female mesothorax length ($F_{1,23} = 0.23$, $p = 0.64$) but was shorter for mated females ($F_{1,23} = 6.43$, $p = 0.018$, **Figure 7B**). Egg laying rate and hatching rate were not correlated with female size (respectively: $F_{1,25} = 1.81$, $p = 0.19$ and $F_{1,23} = 0.04$, $p = 0.84$) or mating status (respectively: $F_{1,25} = 0.23$, $p = 0.63$, **Figure 7B** and $F_{1,23} = 2.73$, $p = 0.11$, **Figure 7C**). Parthenogenetic females died significantly sooner than mated females (on average 116 versus 180 days after the final molt, $\chi^2 = 9.7$, $df = 1$, $p = 0.002$, **Figure 7D**). Nymphs from parthenogenetic females also had a higher mortality rate during their first instar than nymphs from mated females (10.2% versus 4%, Fisher's exact test: $p = 0.0002$, **Figure 7E**).

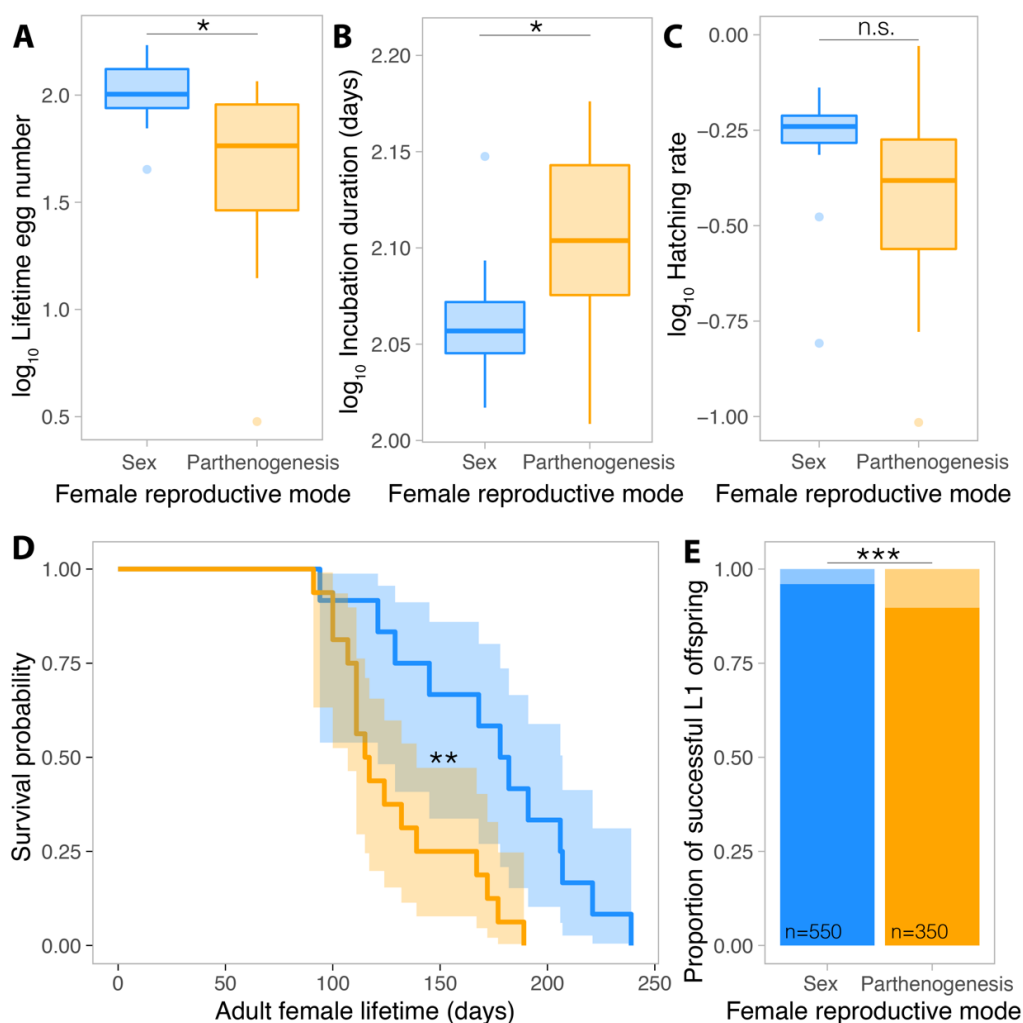


Figure 7: Effect of sex (blue) versus parthenogenesis (orange) on female fitness. Differences between mated and parthenogenetic females in terms of total lifetime egg number (**A**), average single egg mass (**B**), hatching rate (**C**), survival (**D**) and offspring survival during the first instar (**E**). Asterisks indicate significant differences (see **Results** for details)

Discussion

Our results supported the hypothesis that the exaggerated hindlegs of male *E. calcarata* primarily function as pure cost-delivering tools used in physical fights against rival males, and not as threat signals to rivals or coercive tools to force copulation from females. During aggressive male-male interactions, the largest male with proportionately the largest hindlegs was more likely to win and subsequently mate with females. Rival male *E. calcarata* appeared to mutually assess each other's RHP but we did not find any evidence that males sized up their rival by using the hindlegs as a visual or tactile threatening signal. Most likely, males used other chemical, mechanical and vibrational cues, given the pervasiveness of antennal contacts and abdomen thumping. Males exclusively made use of their weaponized hindlegs to deal powerful squeezes to their opponent. However, such instances were rare as contests followed a stereotyped pattern of escalating phases that often de-escalated before reaching the grasping phase.

The hindlegs of *E. calcarata* are secondary sexual traits that only become strongly sexually dimorphic upon reaching adulthood. Males use these hindlegs and attached sharp spine to grab, squeeze and puncture their rival's body parts, most often the hind femora, during fights for access to females (Boisseau et al., 2020). Males most often delivered these powerful and hazardous squeezes when sitting on top of their rival (**Figure 4A**). Similar insect systems with enlarged male hindlegs most often adopt a backwards-facing fighting style (e.g., Eberhard, 1998; Procter et al. 2012; O'Brien et al. 2017a; Fea and Holwell 2018; Rink et al. 2019), thereby contrasting with the fighting style reported here. Most of these systems, including the leaf-footed bugs (Miller et al., 2016; Allen and Miller, 2017; O'Brien and Boisseau, 2018) and frog-legged beetles (O'Brien et al., 2017b), also exhibit a hyperallometric scaling of hindleg (i.e., weapon) size as a function of body size. However, *E. calcarata* adds to weapon systems like cave wētās (Fea and Holwell, 2018) and monkey beetles (Rink et al., 2019) for which weapon size scales proportionately with body size, rather than positively allometrically (Bonduriansky, 2007). This is consistent with their function as pure tools of combat. Indeed, hyperallometry is specifically expected in systems where weapons are also used as signals addressed to rivals or females as it exacerbates size differences between males facilitating assessment of RHP (Eberhard et al., 2018; O'Brien et al., 2018). On the other hand, isometric scaling is often more advantageous from a mechanical stand-point as the lever components of the weapon would stay in proportion and maintain mechanical advantage (O'Brien and Boisseau, 2018). In contrast,

hyperallometric weapons are prone to decreasing mechanical advantage and weakening as they get larger because the out-lever arm --where force is delivered to the opponent-- becomes disproportionately longer relative to the in-lever arm where the associated musculature attaches and provides the input force to the system (i.e., the “paradox of the weakening combatant”; (Levinton and Allen, 2005; O’Brien and Boisseau, 2018). *E. calcarata* males have already been shown to harbor relatively large and costly femoral flexor muscles in those hindlegs (O’Brien et al., 2019), which further suggests that the squeezing strength of these weapons is critical to males.

Body size was the main predictor of contest outcome and therefore used as a proxy for RHP. We did not find any significant effect of weapon size or body condition on the probability of winning after accounting for body size in both randomly and size-matched contests. The effect of weapon size *per se* is usually difficult to disentangle from that of body size as both are often tightly correlated (e.g., Painting and Holwell 2014; del Sol et al. 2020). However, our body size-matched contests only found a strong effect of residency on contest outcome (i.e., the resident male was much more likely to win against a matched intruder) but did not show any effect of weapon size. This suggests that disproportionately large hind legs are not advantageous in combat, perhaps owing to their relatively higher cost of maintenance (O’Brien et al., 2019) or relatively lower mechanical performance (O’Brien and Boisseau, 2018). This is in contrast with other studies that found that body size was of little importance relative to the size of the sexually-selected weapons (e.g., Bridge et al. 2000; Lailvaux et al. 2005; Fea and Holwell 2018). In such situations, some other aspects of weapon performance uniquely related to weapon size, such as reach, may be more decisive than strength, which depends on associated musculature, stamina and lever mechanics.

While weapon size was not a better predictor of RHP than overall body size, it does not rule out a potential signalling function during combat. This would first and foremost require contestants to actually be able assess the RHP of their opponent during fights and presupposes an opponent-only or mutual assessment strategy of contest resolution. We found that fights escalated more when contestants were similar in size, and when losers were larger but less when winners were larger, thereby ruling out pure self-assessment strategies (Arnott and Elwood, 2009; Green and Patek, 2018) (**Figure 1B**). Contest costs and escalation levels were not different between size-matched contests involving large or small individuals, which supported mutual assessment rather than cumulative assessment (Pinto et al., 2019) (**Figure 1C**). Repeated trials involving small focal losers consistently showed more escalation the smaller (and more matched) the winner was, further suggesting mutual assessment

(Chapin et al., 2019) (**Figure 1D**). Finally, analyses of fight behavioral sequences revealed that contests unfolded in phases of increasing intensity, with more size-matched contests more often reaching the latest and most dangerous phases (Green and Patek, 2018) (**Figure 1E**). Combined, our results strongly suggest that *E. calcarata* males are able to assess their rival's size and RHP during combat and use that information to decide on whether to keep fighting or retreat. Our study illustrates how in combination a suite of recently developed methods can lead to strong empirical support for a specific contest assessment model.

Although male *E. calcarata* do appear to assess the RHP of their opponents, they likely are not using their hindleg weapons for this assessment. Instead, males seem to assess each other mainly through antennal contact, which is how contests invariably start. This suggests that males may be using mechanical and/or chemical cues to assess their rival's size early on. However, the specific cues and signals used by males are still unknown (see Arnott and Elwood 2009 for a review on such cues used in other systems). For example, cuticular hydrocarbon abundance and composition might be used to signal body size and dominance status (e.g., Thomas and Simmons 2011; Steiger et al. 2013; Lane et al. 2016). These antennal contacts did not appear directed towards the hindlegs and we did not observe any behavioral evidence that would suggest a signalling role of the hindlegs (e.g., no waving or backward facing), contrary to other weapon systems (Clutton-Brock et al., 1979; Jennions, 1996; Katsuki et al., 2014; McCullough et al., 2016; Miyatake, 1993). Finally, we repeatedly observed winners thumping on the substrate (i.e., bark) with the tip of their abdomen at the end of a fight as the loser was retreating. Similar tapping behaviors had already been anecdotally reported in males of other unrelated phasmid species (e.g., *Aretaon asperrimus* (Delfosse, 2003), *Anisomorpha buprestoides* (James, 1981)), also in association with agonistic interactions with other males. It is possible that some characteristics of this vibrational signal convey information about size and RHP (e.g., frequency, amplitude) (Cocroft and Rodríguez, 2005; De Luca and Morris, 1998).

As thumping was restricted to winners and almost exclusively performed after contests, it likely constitutes a victory display (Bower, 2005) analogous to that found in wētās and crabs (Chen et al., 2014; Kelly, 2006). Victory displays are relatively poorly understood but seem to be more common in gregarious animals where they could function as advertisement of recent victory to nearby bystanders, dissuading them from initiating a new contest with the freshly victorious male (Mesterton-Gibbons and Sherratt, 2006). In the field, male *E. calcarata* can be found in high densities on the same tree trunk which suggests that this signal could indeed reach a large audience (Boisseau et al., 2020). Alternatively,

this signal could also be directly addressed to the retreating loser to reduce the chance that the latter would re-initiate a new contest (Chen et al., 2017; Mesterton-Gibbons and Sherratt, 2006).

Exaggerated hindlegs are used by males to force copulation from resisting females in many insect systems (Burrows, 2020; Haley and Gray, 2012; Rowe et al., 2006). Male *E. calcarata* use their hindlegs to grasp females and position their abdomen during copulation (Hsiung, 1987), consistent with a potential role in overcoming female resistance behaviors and forcing copulation. However, mating latency was not affected by male size, female size, or female mating status, as would have been expected if large body and weapon sizes contributed to forced copulation success (Shine and Mason, 2005). Thus, larger males did not appear to take advantage of their larger hindlegs to force females to copulate faster. This is likely a consequence of females almost never behaviorally resisting copulation attempts in the field or lab: females froze upon contact with a male and never kicked or tried to walk away.

In addition, females do not appear to benefit in any way by avoiding mating with males, unlike some other phasmids where parthenogenetically reproducing females have the highest fitness (Burke et al., 2015). Non-mated *E. calcarata* females had a shorter adult survival and laid fewer eggs, and these eggs took longer to develop and suffered increased offspring mortality compared with eggs of mated females.

In the field, females were observed copulating with several males every night but never resisting. The absence of female resistance behaviors may be explained by the potential benefits of mating multiply or by the cost of resisting being higher than the cost of mating (i.e., “convenience polyandry”; Arnqvist and Rowe, 2005; Cordero and Andrés, 2002; Rowe, 1992). As males can be present in high densities on the trunks of host trees, it is likely that repeatedly repelling harassing males is very costly for females in this system, thereby favoring sexual receptivity. However, it is possible that in circumstances where male density and therefore harassment level is low (not observed here), females may exhibit resistance behaviors (Rowe, 1992). Therefore, we cannot completely rule out a potential coercive role of the male hindlegs in this particular context.

Conclusion

Large males bearing exaggerated weapons are extremely rare in Phasmatodea but convergently evolved in at least three lineages (Boisseau et al., 2020; Buckley et al., 2009; Emberts and Wiens, 2021). A previous field study showed that the evolution of exaggerated male hindlegs was associated with a switch from solitary to communal diurnal roosting and from a scramble competition to a defense-based polygyny mating system. The present study built upon this and showed that males benefit from being larger and consequently having larger hindlegs by increasing their fighting success against rivals and consequently their mating success. In parallel, we did not find any evidence for a coercive role of these hindlegs in overcoming female resistance, despite males wrapping a hindleg around the female's abdomen during copulation. Thus, hindleg exaggeration appears to have been driven primarily by sexual selection through direct male-male competition. Furthermore, these weapons were only used to deliver powerful and harmful squeezes to opponents in escalated contests. Although males do appear to mutually assess the relative size and fighting ability of opponents during contests, they never used their hindlegs for this assessment, which is consistent with their proportional scaling relationship with body size (Eberhard et al., 2018; McCullough and O'Brien, 2022; O'Brien et al., 2018). By combining natural history observations in the field, analyses of the predictors of male-male contest outcomes (to characterize determinants of resource holding potential), predictors of contest costs and contest behavioral sequences (to characterize assessment strategy), and predictors of the duration of male-female interactions (to investigate a potential role of sexual conflict), we were able to identify precisely the function of this sexually dimorphic trait and shed light on its evolution and pattern of scaling.

Acknowledgements

We thank M. Ero, L. Bonneau, S. Makai, R. Dikrey, B. Sapau, P. Mana, S. Komda, G. Gumbira, R. Uker, T. Manjobie and T. Batari for assistance in the field in Papua New Guinea; the Missoula Butterfly house and Insectarium for providing the initial captive population of *E. calcarata*; P. Green for kindly providing R scripts to run the behavioral sequence analyses; C. Allen for help with insect rearing and logistics; C. Thomas-Bulle for helping with the experiments in the lab and for helpful discussions and feedback on the manuscript. The field part of this project was funded by an early career grant from the National Geographic Society to RB (grant number: WW-255ER-17) and by the Papua New Guinea

Oil Palm Research Association (PNGOPRA). The lab part of the project was funded by a National Science Foundation grant to DJE (IOS–1456133).

Author contributions

RPB and DJE conceived of the study. RB conducted the experiments in the field and in the lab, reared the study animals, performed the statistical analyses and wrote the initial version of the manuscript. RPB and DJE contributed to editing and revising subsequent versions of the manuscript.

References

- Allen, P. E. and Miller, C. W.** (2017). Novel host plant leads to the loss of sexual dimorphism in a sexually selected male weapon. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20171269.
- Andersson, M.** (1982). Female choice selects for extreme tail length in a widowbird. *Nature* **299**, 818–820.
- Andersson, M.** (1994). *Sexual Selection*. Princeton, NJ, USA: Princeton University Press.
- Arnott, G. and Elwood, R. W.** (2008). Information gathering and decision making about resource value in animal contests. *Animal Behaviour* **76**, 529–542.
- Arnott, G. and Elwood, R. W.** (2009). Assessment of fighting ability in animal contests. *Animal Behaviour* **77**, 991–1004.
- Arnqvist, G. and Rowe, L.** (2005). *Sexual conflict*. Princeton, New Jersey, USA: Princeton University Press.
- Bedford, G. O.** (1976). Defensive behaviour of the New Guinea stick insect *Eurycantha* (Phasmatoidea: Phasmatidae: Eurycanthinae). *Proceedings of the Linnean Society of New South Wales* **100**, 218–222.
- Berglund, A., Bisazza, A. and Pilastro, A.** (1996). Armaments and ornaments: an evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society* **58**, 385–399.
- Bildstein, K. L., McDowell, S. G. and Brisbin, I. L.** (1989). Consequences of sexual dimorphism in sand fiddler crabs, *Uca pugilator*: differential vulnerability to avian predation. *Animal Behaviour* **37**, 133–139.
- Boisseau, R. P., Ero, M. M., Makai, S., Bonneau, L. J. G. and Emlen, D. J.** (2020). Sexual dimorphism divergence between sister species is associated with a switch in habitat use and mating system in thorny devil stick insects. *Behavioural Processes* **181**, 104263.
- Bonduriansky, R.** (2007). Sexual selection and allometry: A critical reappraisal of the evidence and ideas. *Evolution* **61**, 838–849.
- Bower, J. L.** (2005). The occurrence and function of victory displays within communication networks. In *Animal Communication Networks* (ed. McGregor, P.), p. 115–126. Cambridge, UK: Cambridge University press.

- Bridge, A. P., Elwood, R. W. and Dick, J. T. A.** (2000). Imperfect assessment and limited information preclude optimal strategies in male-male fights in the orb-weaving spider *Metellina mendei*. *Proceedings of the Royal Society B* **267**, 273–279.
- Briffa, M., Hardy, I. C. W., Gammell, M. P., Jennings, D. J., Clarke, D. D. and Goubault, M.** (2013). Analysis of animal contest data. In *Animal Contests* (ed. Hardy, I. C. W. and Briffa, M.), pp. 47–85. Cambridge, United Kingdom: Cambridge University press.
- Buckley, T. R., Attanayake, D. and Bradler, S.** (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. *Proceedings of the Royal Society B: Biological Sciences* **276**, 1055–1062.
- Burke, N. W. and Bonduriansky, R.** (2022). Sexually but not parthenogenetically produced females benefit from mating in a stick insect. *Functional Ecology* **36**, 2001–2014.
- Burke, N. W., Crean, A. J. and Bonduriansky, R.** (2015). The role of sexual conflict in the evolution of facultative parthenogenesis: A study on the spiny leaf stick insect. *Animal Behaviour* **101**, 117–127.
- Burrows, M.** (2020). Do the enlarged hind legs of male thick-legged flower beetles contribute to take-off or mating? *Journal of Experimental Biology* **223**, jeb212670.
- Carlberg, U.** (1989). Aspects of Defensive Behaviour of *Eurycantha calcarata* Lucas Females and the Evolution of Scorpion Mimicry in the Phasmida (Insecta). *Biologisches Zentralblatt* **108**, 257–262.
- Chapin, K. J., Peixoto, P. E. C. and Briffa, M.** (2019). Further mismeasures of animal contests: A new framework for assessment strategies. *Behavioral Ecology* **30**, 1177–1185.
- Chen, P. Z., Carrasco, L. R. and Ng, P. K. L.** (2014). Post-contest stridulation used exclusively as a victory display in mangrove crabs. *Ethology* **120**, 532–539.
- Chen, P. Z., Carrasco, R. L. and Ng, P. K. L.** (2017). Mangrove crab uses victory display to “browbeat” losers from re-initiating a new fight. *Ethology* **123**, 981–988.
- Clutton-Brock, T. and Parker, G. A.** (1995). Sexual coercion in animal societies. *Animal Behaviour* **49**, 1345–1365.
- Clutton-Brock, T. H., Albon, S. D., Gibson, R. M. and Guinness, F. E.** (1979). The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus* L.). *Animal Behaviour* **27**, 211–225.
- Cocroft, R. B. and Rodríguez, R. L.** (2005). The Behavioral Ecology of Insect Vibrational Communication. *BioScience* **55**, 323–334.
- Cordero, A. and Andrés, J. A.** (2002). Male coercion and convenience polyandry in a calopterygid damselfly. *Journal of Insect Science* **2**, 1–7.
- Csardi, G. and Nepusz, T.** (2006). The igraph software package for complex network research. *InterJournal, complex systems* **1695**, 1–9.
- De Luca, P. A. and Morris, G. K.** (1998). Courtship communication in meadow katydids: female preference for large male vibrations. *Behaviour* **135**, 777–794.
- del Sol, J. F., Hongo, Y., Boisseau, R., Berman, G., Allen, C. E. and Emlen, D. J.** (2020). Population differences in the strength of sexual selection match relative weapon size in the Japanese rhinoceros beetle, *Trypoxylus dichotomus* (Coleoptera: Scarabaeidae). *Evolution* **75**, 394–413.
- Delfosse, E.** (2003). Taxonomie, répartition, élevage, émission et réception de sons chez le Phasme épineux « marteleur » : *Aretaon* (*Aretaon*) *asperrimus* (Redtenbacher, 1906) (Insecta Orthopteroidea Phasmatodea Areolatae Bacillidae Heteropteryginae Obrimini). *Bulletin de Phyllie* **16**, 16–30.

- Eberhard, W. G.** (1998). Sexual behavior of *Acanthocephala declivis guatemalana* (Hemiptera: Coreidae) and the allometric scaling of their modified hind legs. *Annals of the Entomological Society of America* **91**, 863–871
- Eberhard, W. G., José, S., Rica, C., Lucas Rodríguez, R., Alexander, B. A. H., Speck, B., Miller, H., Rodríguez, R. L., Huber, B. A., Speck, B., et al.** (2018). Sexual Selection and Static Allometry: the Importance of Function. *The Quarterly Review of Biology* **93**, 207–250.
- Elwood, R. W. and Arnott, G.** (2012). Understanding how animals fight with Lloyd Morgan’s canon. *Animal Behaviour* **84**, 1095–1102.
- Elwood, R. W. and Arnott, G.** (2013). Assessments in contests are frequently assumed to be complex when simple explanations will suffice. *Animal Behaviour* **86**, e8–e12.
- Embergs, Z. and Wiens, J. J.** (2021). Do sexually selected weapons drive diversification? *Evolution* **75**, 2411–2424.
- Emlen, D. J.** (2008). The Evolution of Animal Weapons. *Annual Review of Ecology, Evolution, and Systematics* **39**, 387–413.
- Emlen, D. J.** (2014). Reproductive contests and the evolution of extreme weaponry. In *The evolution of insect mating systems* (ed. Shuker, D. and Simmons, L.), pp. 92–105. Oxford, United Kingdom: Oxford University Press.
- Emlen, D. J., Lavine, L. C. and Ewen-Campen, B.** (2007). On the origin and evolutionary diversification of beetle horns. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8661–8668.
- Enquist, M. and Leimar, O.** (1983). Evolution of Fighting Behaviour : Decision Rules and Assessment of Relative Strength. *Journal of Theoretical Biology* **102**, 387–410.
- Enquist, M., Leimar, O., Ljungberg, T., Mallner, Y. and Segerdahl, N.** (1990). A test of the sequential assessment game: fighting in the cichlid fish *Nannacara anomala*. *Animal Behaviour* **40**, 1–14.
- Fea, M. and Holwell, G.** (2018). Combat in a cave-dwelling wētā (Orthoptera: Rhaphidophoridae) with exaggerated weaponry. *Animal Behaviour* **138**, 85–92.
- Friard, O. and Gamba, M.** (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* **7**, 1325–1330.
- Goubault, M. and Decuignière, M.** (2012). Previous Experience and Contest Outcome: Winner Effects Persist in Absence of Evident Loser Effects in a Parasitoid Wasp. *The American Naturalist* **180**, 364–371.
- Green, P. A. and Patek, S. N.** (2018). Mutual assessment during ritualized fighting in mantis shrimp (Stomatopoda). *Proceedings of the Royal Society B: Biological Sciences* **285**, 20172542.
- Haley, E. L. and Gray, D. A.** (2012). Mating behavior and dual-purpose armaments in a camel cricket. *Ethology* **118**, 49–56.
- Hardy, I. C. W. and Briffa, M.** (2013). *Animal Contests*. Cambridge, United Kingdom: Cambridge University press.
- Hsiung, C.-C.** (1987). Aspects of the biology of the Melanesian stick-insect *Eurycantha calcarata* Lucas (Cheleutoptera: Phasmatidae). *Journal of Natural History* **21**, 1241–1258.
- Hsu, Y., Earley, R. L. and Wolf, L. L.** (2006). Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biological Reviews* **81**, 33–74.
- James, T.** (1981). *Anisomorpha buprestoides* noises. *The Phasmid Study Group Newsletter* **6**, 2.
- Jennions, M. D.** (1996). Residency and size affect fight duration and outcome in the fiddler crab *Uca annulipes*. *Biological Journal of the Linnean Society* **57**, 293–306.

- Katsuki, M., Yokoi, T., Funakoshi, K. and Oota, N.** (2014). Enlarged hind legs and sexual behavior with male-male interaction in *Sagra femorata* (Coleoptera : Chrysomelidae). *Entomological news* **124**, 211–220.
- Kelly, C. D.** (2006). Fighting for harems : assessment strategies during male - male contests in the sexually dimorphic Wellington tree weta. *Animal behaviour* **72**, 727–736.
- Kodric-Brown, A.** (1985). Female preference and sexual selection for male coloration in the guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology* **17**, 199–205.
- Lailvaux, S. P., Hathway, J., Pomfret, J. and Knell, R. J.** (2005). Horn size predicts physical performance in the beetle *Euoniticellus intermedius* (Coleoptera: Scarabaeidae). *Functional Ecology* **19**, 632–639.
- Lane, S. M., Dickinson, A. W., Tregenza, T. and House, C. M.** (2016). Sexual Selection on male cuticular hydrocarbons via male–male competition and female choice. *Journal of Evolutionary Biology* **29**, 1346–1355.
- Lavine, L., Gotoh, H., Brent, C. S., Dworkin, I. and Emlen, D. J.** (2015). Exaggerated trait growth in insects. *Annual Review of Entomology* **60**, 453–472.
- Levinton, J. S. and Allen, B. J.** (2005). The paradox of the weakening combatant: Trade-off between closing force and gripping speed in a sexually selected combat structure. *Functional Ecology* **19**, 159–165.
- Lovich, J. E. and Gibbons, J. W.** (1992). A review of techniques for quantifying sexual size dimorphism. *Growth, Development and Aging* **56**, 269–281.
- Maynard Smith, J.** (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology* **47**, 209–221.
- Maynard Smith, J.** (1982). *Evolution and the Theory of Games*. Cambridge, United Kingdom: Cambridge University press.
- Maynard Smith, J. and Parker, G. A.** (1976). The logic of asymmetric contests. *Animal Behaviour* **24**, 159–175.
- Maynard Smith, J. and Price, G. R.** (1973). Logic of animal conflict. *Nature* **246**, 15–18.
- McCullough, E. L. and O'Brien, D. M.** (2022). Variation in allometry along the weapon-signal continuum. *Evolutionary Ecology* **36**, 591–604.
- McCullough, E. L., Miller, C. W. and Emlen, D. J.** (2016). Why sexually selected weapons are not ornaments. *Trends in Ecology and Evolution* **31**, 742–751.
- McLain, D. K., Pratt, A. E. and Berry, A. S.** (2003). Predation by red-jointed fiddler crabs on congeners: interaction between body size and positive allometry of the sexually selected claw. *Behavioral Ecology* **14**, 741–747.
- Mesterton-Gibbons, M. and Sherratt, T. N.** (2006). Victory displays: a game-theoretic analysis. *Behavioral Ecology* **17**, 597–605.
- Mesterton-Gibbons, M., Marden, J. H. and Dugatkin, L. A.** (1996). On wars of attrition without assessment. *Journal of Theoretical Biology* **181**, 65–83.
- Miller, C. W., McDonald, G. C. and Moore, A. J.** (2016). The tale of the shrinking weapon: seasonal changes in nutrition affect weapon size and sexual dimorphism, but not contemporary evolution. *Journal of Evolutionary Biology* **29**, 2266–2275.
- Miyatake, T.** (1993). Male-male aggressive behavior is changed by body size difference in the leaf-footed plant bug, *Leptoglossus australis*, Fabricius (Heteroptera: Coreidae). *Journal of Ethology* **11**, 63–65.
- Muramatsu, D.** (2011). The function of the four types of waving display in *Uca lactea*: effects of audience, sand structure,

- and body size. *Ethology* **117**, 408–415.
- O'Brien, D. M. and Boisseau, R. P.** (2018). Overcoming mechanical adversity in extreme hindleg weapons. *Plos One* **13**, e0206997.
- O'Brien, D. M., Katsuki, M. and Emlen, D. J.** (2017a). Selection on an extreme weapon in the frog-legged leaf beetle (*Sagra femorata*). *Evolution* **71**, 2584–2598.
- O'Brien, D. M., Katsuki, M. and Emlen, D. J.** (2017b). Selection on an extreme weapon in the frog-legged leaf beetle (*Sagra femorata*). *Evolution* **71**, 2584–2598.
- O'Brien, D. M., Allen, C. E., Van Kleeck, M. J., Hone, D., Knell, R., Knapp, A., Christiansen, S. and Emlen, D. J.** (2018). On the evolution of extreme structures: static scaling and the function of sexually selected signals. *Animal Behaviour* **144**, 95–108.
- O'Brien, D. M., Boisseau, R. P., Duell, M., McCullough, E., Powell, E. C., Somjee, U., Solie, S., Hickey, A. J., Holwell, G. I., Painting, C. J., et al.** (2019). Muscle mass drives cost in sexually selected arthropod weapons. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20191063.
- Ohde, T., Morita, S., Shigenobu, S., Morita, J., Mizutani, T., Gotoh, H., Zinna, R. A., Nakata, M., Ito, Y., Wada, K., et al.** (2018). Rhinoceros beetle horn development reveals deep parallels with dung beetles. *PLoS Genetics* **14**, e1007651.
- Painting, C. J. and Holwell, G. I.** (2014). Exaggerated rostra as weapons and the competitive assessment strategy of male giraffe weevils. *Behavioral Ecology* **25**, 1223–1232.
- Parker, G. A.** (1974). Assessment strategy and the evolution of animal conflicts. *Journal of Theoretical Biology* **47**, 223–243.
- Payne, R. J. H.** (1998). Gradually escalating fights and displays : the cumulative assessment model. *Animal Behaviour* **56**, 651–662.
- Payne, R. J. H. and Pagel, M.** (1996). Escalation and time costs in displays of endurance. *Journal of Theoretical Biology* **183**, 185–193.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team** (2021). nlme: linear and nonlinear mixed effects models.
- Pinto, N. S., Palaoro, A. V and Peixoto, P. E. C.** (2019). All by myself? Meta-analysis of animal contests shows stronger support for self than for mutual assessment models. *Biological Reviews* **55**,
- Pradhan, G. R. and Van Schaik, C. P.** (2009). Why do females find ornaments attractive? The coercion-avoidance hypothesis. *Biological Journal of the Linnean Society* **96**, 372–382.
- Pratt, A. E., Kelly McLain, D. and Lathrop, G. R.** (2003). The assessment game in sand fiddler crab contests for breeding burrows. *Animal Behaviour* **65**, 945–955.
- Procter, D. S., Moore, a. J. and Miller, C. W.** (2012). The form of sexual selection arising from male-male competition depends on the presence of females in the social environment. *Journal of Evolutionary Biology* **25**, 803–812.
- R Core Team** (2021). R: A Language and Environment for Statistical Computing.
- Rico-Guevara, A. and Hurme, K. J.** (2019). Introsexually selected weapons. *Biological Reviews* **94**, 60–101.
- Rink, A. N., Altwegg, R., Edwards, S., Bowie, R. C. K. and Colville, J. F.** (2019). Contest dynamics and assessment strategies in combatant monkey beetles (Scarabaeidae: Hopliini). *Behavioral Ecology* **30**, 713–723.

- Rodríguez, R. L. and Eberhard, W. G.** (2019). Why the static allometry of sexually-selected traits is so variable: the importance of function. *Integrative and Comparative Biology* **59**, 1290–1302.
- Rometsch, S. J., Torres-Dowdall, J., Machado-Schiaffino, G., Karagic, N. and Meyer, A.** (2021). Dual function and associated costs of a highly exaggerated trait in a cichlid fish. *Ecology and Evolution* **11**, 17496–17508.
- Rowe, L.** (1992). Convenience polyandry in a water strider: foraging conflicts and female control of copulation frequency and guarding duration. *Animal Behaviour* **44**, 189–202.
- Rowe, L., Westlake, K. P. and Currie, D. C.** (2006). Functional significance of elaborate secondary sexual traits and their evolution in the water strider genus *Rheumatobates*. *Canadian Entomologist* **138**, 568–577.
- Rutte, C., Taborsky, M. and Brinkhof, M. W. G.** (2006). What sets the odds of winning and losing? *Trends in Ecology & Evolution* **21**, 16–21.
- Shine, R. and Mason, R. T.** (2005). Does large body size in males evolve to facilitate forcible insemination? A study on garter snakes. *Evolution* **59**, 2426–2432.
- Schneider, C. A., Rasband, W. S. and Eliceiri, K. W.** (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.
- Shuker, D. M. and Simmons, L. W.** (2014). *The evolution of insect mating systems*. Oxford, United Kingdom: Oxford University Press.
- Steiger, S., Ower, G. D., Stökl, J., Mitchell, C., Hunt, J. and Sakaluk, S. K.** (2013). Sexual selection on cuticular hydrocarbons of male sagebrush crickets in the wild. *Proceedings of the Royal Society B: Biological Sciences* **280**,
- Taylor, P. W. and Elwood, R. W.** (2003). The mismeasure of animal contests. *Animal Behaviour* **65**, 1195–1202.
- Thomas, M. L. and Simmons, L. W.** (2011). Short-term phenotypic plasticity in long-chain cuticular hydrocarbons. *Proceedings of the Royal Society B: Biological Sciences* **278**, 3123–3128.
- Zeh, D. W., Zeh, J. A. and Tavakilian, G.** (1992). Sexual selection and sexual dimorphism in the harlequin beetle *Acrocinus longimanus*. *Biotropica* **24**, 86–96.

Supplementary Information

Table S1 : Analyses of the scaling relationships between front femur length, hind femur length, hind femur width, hind femur area and mesothorax length (~body size) throughout development and across sexes. The table shows the results of best linear mixed-effects models (accounting for individual ID as a random effect) and associated type I ANOVA. Departure from isometry (i.e., slope= 1 for linear traits, slope= 2 for surface traits) was tested in adults using 95% confidence intervals (between brackets) around estimated regression slopes.

Response variable	N	Explanatory variables	F	df1	df2	P	Isometric slope	Adult slope
Log ₁₀ (Front femur length)	254	log ₁₀ (Mesothorax length)	18119	1	133	<.0001	1	0.83 [0.70; 0.96]
		instar	119	3	133	<.0001		
		sex	38	1	109	<.0001		
		instar:sex	9	3	133	<.0001		
		log ₁₀ (Mesothorax length):instar	4	3	133	0.01		
Log ₁₀ (Hind femur length)	253	log ₁₀ (Mesothorax length)	25047	1	132	<.0001	1	0.97 [0.82; 1.12]
		instar	345	3	132	<.0001		
		sex	261	1	109	<.0001		
		instar:sex	99	3	132	<.0001		
		log ₁₀ (Mesothorax length):instar	8	3	132	<.0001		
Log ₁₀ (Hind femur width)	253	log ₁₀ (Mesothorax length)	7821	1	135	<.0001	1	1.05 [0.81; 1.29]
		instar	167	3	135	<.0001		
		sex	447	1	109	<.0001		
		instar:sex	129	3	135	<.0001		
		log ₁₀ (Mesothorax length):instar	5	3	132	0.002		
Log ₁₀ (Hind femur length)	253	log ₁₀ (Mesothorax length)	17763	1	132	<.0001	2	2.02 [1.69; 2.36]
		instar	300	3	132	<.0001		
		sex	526	1	109	<.0001		
		instar:sex	164	3	132	<.0001		
		log ₁₀ (Mesothorax length):instar	5	3	132	0.002		

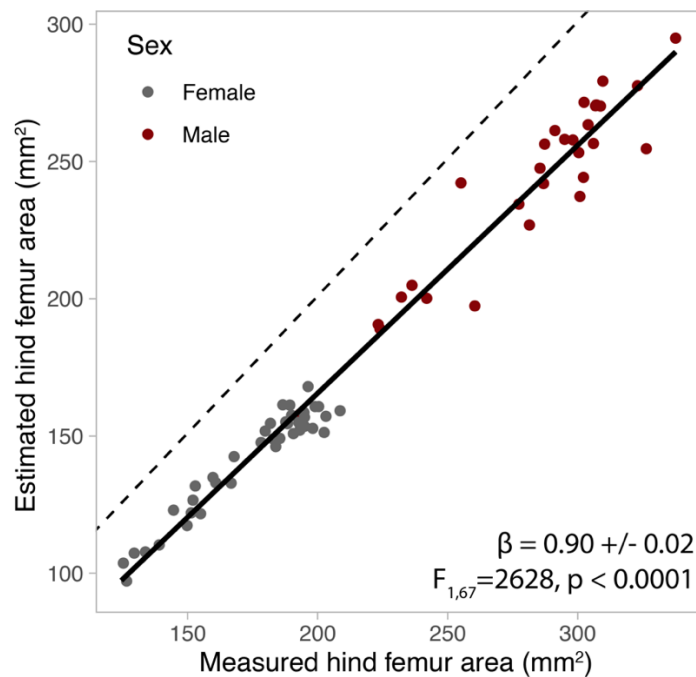


Figure S1: Correlation between estimated (i.e., the area of the ellipse of major diameter hind femur length and minor diameter hind femur width) and measured hind femur area (i.e., the area obtained by drawing the lateral outline of the femur). The dashed line represents the equality line.

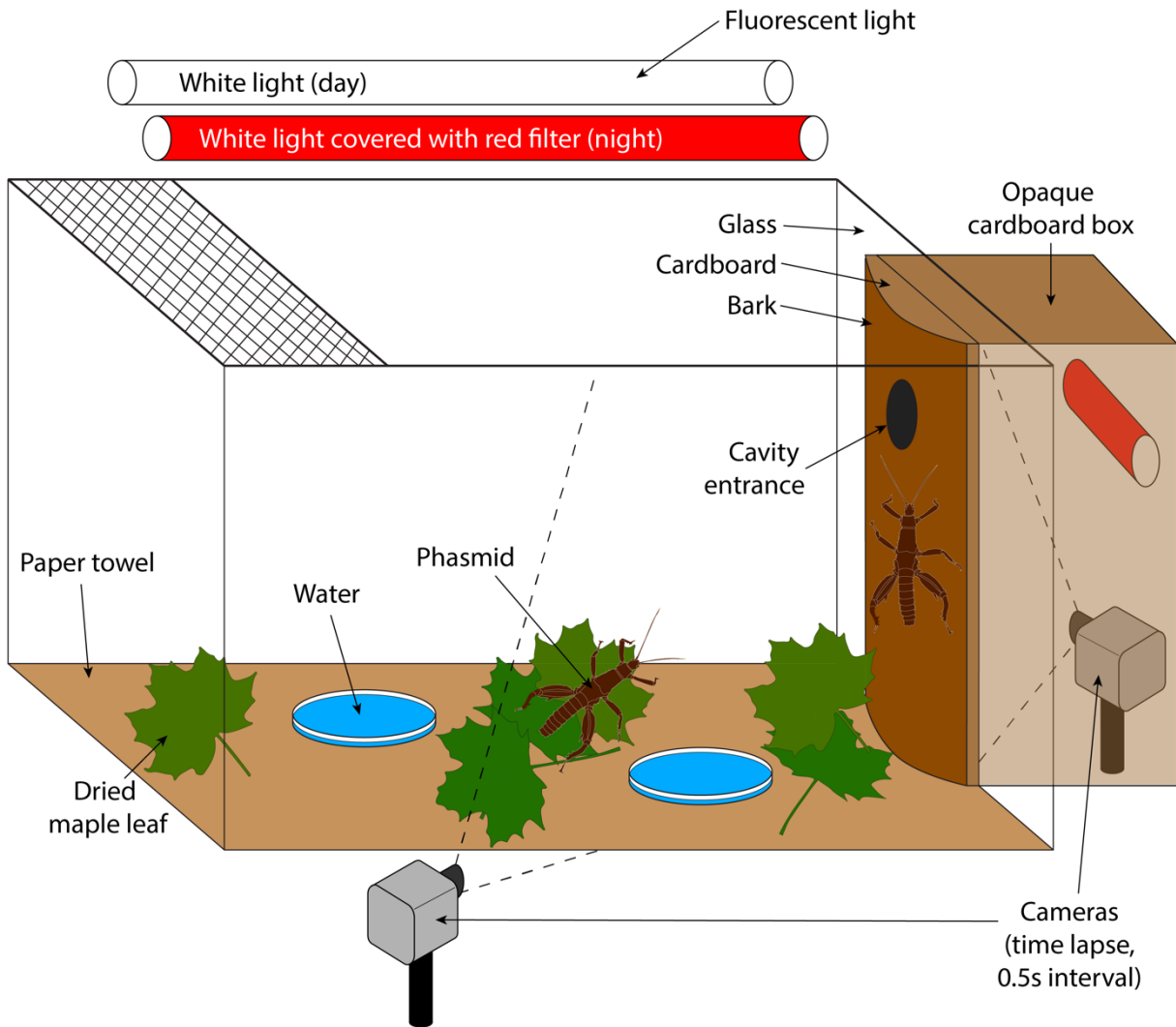


Figure S2: Experimental set up for fighting trials.

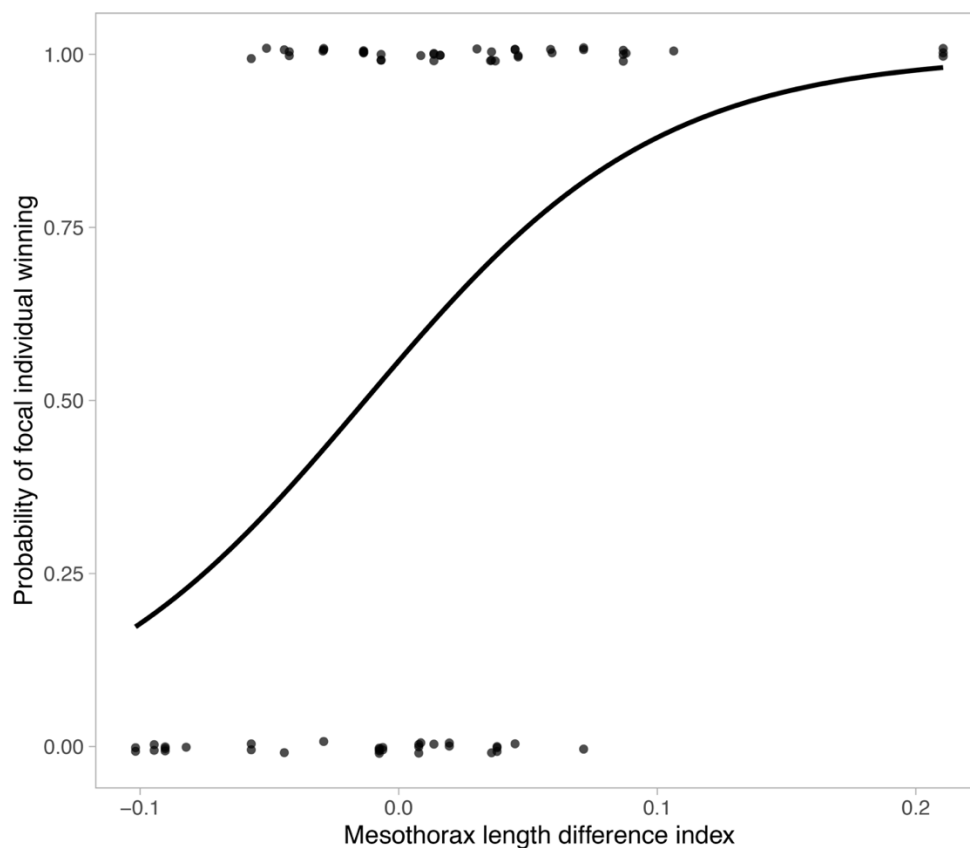


Figure S3: Binary GLMM of mesothorax length difference index in randomly matched lab trials against contest outcome inside the daily roosting cavity.

Video S1: Male 15 and male 17 fighting for control of two females inside an artificial cavity. Male 17 excludes male 15 from the vicinity of the females and eventually sits on top of them.

Video S2: Male thumping its abdomen on bark at the end of a fight.

Video S3: Contest between two male *E. calcarata* over access to two females on a host tree trunk in the wild.

