Numéro National de Thèse: **2017LYSEN095**

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée par

l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N°340

Biologie Moléculaire, Intégrative et Cellulaire (BMIC)

Discipline: Biologie
Soutenue publiquement le 14 Décembre 2017, par:

Antonin CRUMIERE

Developmental mechanisms of adaptive phenotypes and associated ecological relevance in the semiaquatic bugs

Mécanismes développementaux des phénotypes adaptatifs et leur importance écologique chez les insectes semi-aquatiques

---

Devant le jury composé de:

Dr. Jessica Abbott – Group leader, Lund University, Rapporteur et Examinateur
Dr. Nicola Nadeau – Group leader, University of Sheffield, Rapporteur et Examinateur
Dr. Anthony Herrel – Directeur de Recherche, MNHNP, Rapporteur et Examinateur
Dr. Charlie Scutt – Directeur de Recherche, ENS de Lyon, Examinateur
Dr. Abderrahman Khila – Directeur de Recherche, ENS de Lyon, Directeur de thèse

Acknowledgments

As the tradition, let's start this PhD thesis manuscript with acknowledgments. This part will be probably the hardest to write for me just because I consider that words are too weak to express what I really feel and also because I consider that acts are more important than words.

I want first to thank my PhD committee, Dr. Christen Mirth and Pr. Laurent Viriot who provided a lot of insight into my project. It was a pleasure to show you my progression and to discuss during these four years. I also thank the member of my jury, Dr. Nicola Nadeau, Dr. Jessica Abbott, Dr. Anthony Herrel and Dr. Charlie Scutt, who agreed to evaluate this work.

I thank my parents and family. Since I was young at school you pushed me to have good marks, not to be proud of me but just because "if you work well at school you will be able to do a job that you like". You also offered me the opportunity to open my mind and you transmitted moral and social values that build who I am today.

Then I want to thank Dr. Chantal Diaz, who gave me the opportunity to do my first research internship at the end of my Bachelor degree. At this time I was not really a hard worker and my marks were passable. You asked me whether I wanted to do a PhD. I said yes. You told me "you can, you are not stupid at all, but you are too lazy at learning your courses and you will not succeed during your Master to be able to do a PhD". I think it was exactly what I needed to hear at this time. This internship was more than helpful for me. I also want to thank Dr. Nathalie Depège-Fargeix, who was a great supervisor during my second internship. Even if I moved to the IGFL afterward, I absolutely do not regret my internship at the RDP.

Let's move to the Khila lab. Thank a lot Abdou for these years! I have to say that I am proud to have been part of the Khila lab. Since I started here I learned so many things related either to science or to human being, and I feel now that I am well prepared for my future. Of course we learn all along our life but it is really appreciable to feel confident to face the next steps.

Séverine, David and Peter, the three first persons in the lab when I started. I remember Séverine, you were happy to see a new person coming in the lab like a family that grows up. Actually I also remember my first morning as trainee when, David, you entered in the office, you growled an “hello” and then you sited in front of your screen in the “matrix” without speaking. I though that discussing with you would be a pain ☺. Peter, I remember this day where we started to inject bugs at 7pm on Friday, you were not complaining while I was simmering to go home.

Then come William “copain” and Emilia... my two companions in misery to control Abdou drives during fieldwork in French Guiana. I am sure you know what I mean ☺. Discussion about science are always long with William, you are too passionate for me ☺. Emilia I was really happy to work with you on projects, I learned a lot and we were kind a complementary to hold motivation.

Amélie and Augustin, I remember when you arrived Amélie, it was a pain to explain to you that a protocol is a kind of recipe that you can modify. Things have changed a lot! Augustin, the man from the moon living in the sky with who we can have such as stratospheric discussion about interesting or stupid things.

Aïda, Aïda, Aïda! We met the first time in Panama when I was doing fieldwork and when you just learned that you were going to join our lab. I really hope you will be successful after you PhD when you will be back in Panama. François, there is one important thing that I clearly learned or at least realized from you during these years: words have a meaning and use them well means transmit exactly the message we want to transmit. Thank you because this is something useful not only in science.

Cédric, Antoine and Roberto, my climbing brothers. What a pleasure to climb some walls (at least to try) with friends after a day of work. Antoine, I hope you will enjoy research in South Africa. Cedric, I am sure you will live in Asia at some point or maybe for the rest of your life ☺. Roberto, you just arrived in the lab but we met at several occasions before. Thanks a lot for your help for my postdoc fellowship application. I hope you will enjoy working with water striders.

Marie and Mathilde, it is a pleasure to see you working at the bench. You seem really enjoying to learn new things and I think it reminds me when I started in the lab. I do not forget Pauline and her cakes ☺. I also thank the different students that performed internships in the Khila lab during these years. I want also thank the IGFL members, especially the administrative staff, who was more than nice and efficient with orders, documents and travel organization.

I want also thank Princeton folks. Thanks Sarah for this opportunity of research stay in the Kocher lab at Princeton. Thanks to Ben, Eli, Paul, Luisa, Serge, Diego, Micah, Sudarshan, Alicia, and Julien. These three months, like a pre-postdoc, were determinant for my motivation during the last year of my PhD. I really hope to come back soon.

I want to thank you, Ludivine, for your support. I hope I was not too tiresome during these months of redaction (especially at the end).

Finally, I would like to acknowledge you, my friends who unconsciously have supported me during these years: Quentin (Gros), Kelly (Momo), Arnaud (Arnold), Marion (Poujlove), Gilles (Guilles), Leslie (Leslove), Maxime (Mox), Sylvain (Mousse), Cyril (Lep), Paul (Paulo), Benjamin (Benj), Marine (Patchone), Giovanni (Giovannoi), Pierre (Racaillou), Théo (Chef Nounou), Benoit (Ben), Nathan (Johnny), Alexandra (Barbie), and all persons I could have forgot. We shared so many precious moments digging holes in the cold, playing badminton, hiking in the snow, escaping zombies or sharing meals with drinks. Thanks a lot!

Developmental mechanisms of adaptive phenotypes and underlying ecological relevance in the semiaquatic bugs

How diversity arises is a major question in evolutionary biology. Since Darwin's expeditions during the 19th centuries and "The Origin of species" (Darwin, 1859) it is now widely accepted that any given organisms face a variety of ecological pressures along its lifespan. Genes, alleles and phenotypes that increase fitness are favored by natural selection and are transmitted to the next generation. Natural and sexual selection are intrinsically linked and explain why phenotypes are maintained while the genotype explains how phenotypes are formed (Figure 1). These processes directly impact the diversification of species and ultimately generate diversity.

The origin of diversity can be studied in different ways at different levels and timescales. We can study the diversity of species, morphologies, behaviors, physiologies, developmental processes, or molecules using different approaches such as evolution, development, and ecology across extinct and extant species (Figure 1). In the presented thesis I combined these three fields to understand the complexity of the relations between the different components of natural selection, namely survivorship, mating success and fecundity; with the genotype and their impact on the evolution of organism phenotypes (Figure 1). I used the Gerromorpha, also known as the semiaquatic bugs, a group of insects that transited from terrestrial life to life on water surface habitat and associated sub niches. In addition to their evolutionary trajectory, these insects have a large diversity of morphologies, behaviors and habitats and are easy to breed in laboratory condition, which make them powerful systems for Eco-Evo-Devo studies. Along the five chapters of this thesis, I will develop several examples that increase our understanding about how ecological pressures act on phenotype evolution and how genes participate to the development of adaptive traits.

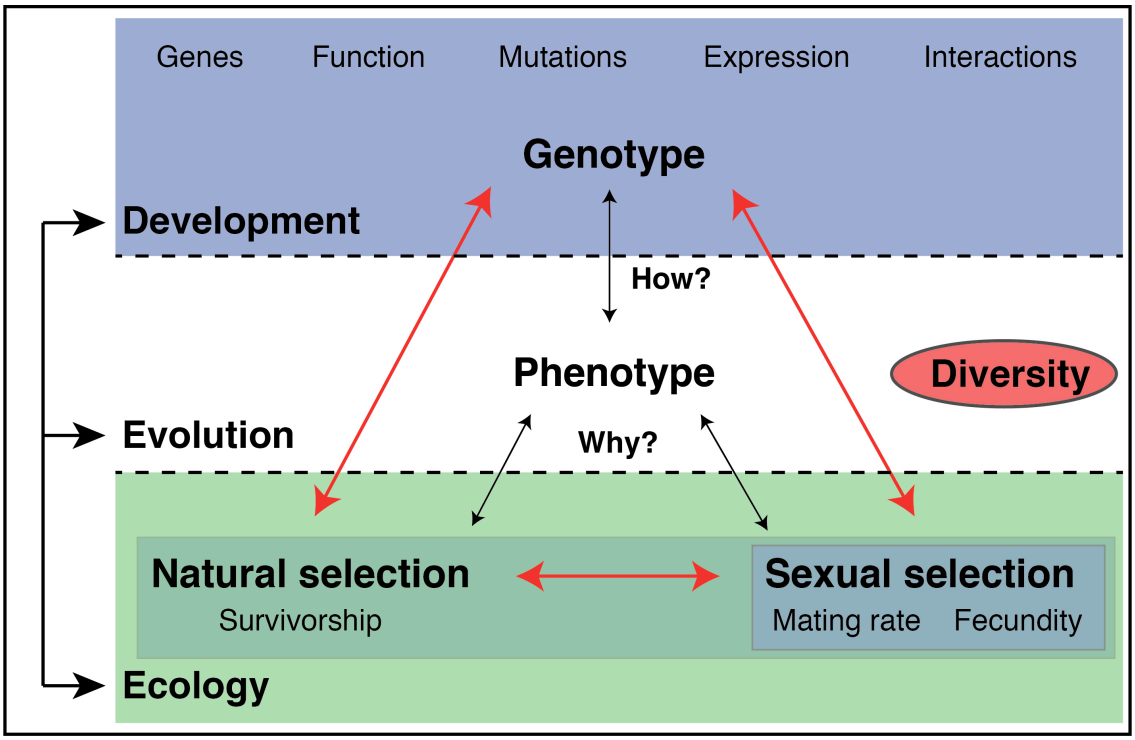


Figure 1: Representation of the concepts and fields used throughout this PhD work. The combination of Development, Evolution and Ecology fields allows us to understand how the relationships between the genotype and the selective pressures shape the phenotype and ultimately phenotypic diversity.

List of abbreviations

dsRNA: double stranded RNA

dsx: doublesex

Hox gene: Homeotic gene

RNAi: RNA interference

Scr: Sex combs reduced

T2: thoracic segment 2

T3: thoracic segment 3

Ubx: Ultrabithorax

List of figures

Figure 1: Representation of the concepts and fields used in the presented thesis. 7
Figure 2: Examples of adaptive traits. 14
Figure 3: Sexual conflict and antagonistic coevolution. 16
Figure 4: Evolution of the Hox cluster. 18
Figure 5: The sex determination gene *doublesex* involved in sexual dimorphism. 20
Figure 6: Fossils and phylogeny of the semiaquatic bugs. 22
Figure 7: Habitat, morphology and locomotion of the semiaquatic bugs. 24
Figure 8: Example of life cycle in the semiaquatic bugs (*Limnoporus dissortis*). 26
Figure 9: Animal rearing facility. 28
Figure 10: The different adaptive phenotypes and underlying genes. 184

Table of contents

General introduction and literature review **11**

I. Ecological pressures, fitness and adaptive phenotypes **12**
1.1 Selective pressures and survivorship **12**
1.2 Examples of adaptive traits and their impacts on evolution **12**

II. Sexual interactions, fitness and sexual dimorphism **15**
2.1 Sexual conflict **15**
2.2 Antagonistic coevolution and arms race **15**

III. Genetic basis of morphological evolution **17**
3.1 Molecular genetic mechanisms **17**
3.2 Examples of genes involved in morphological diversity **17**

IV. The semiaquatic bugs as model systems **21**
4.1 Evolution of semiaquatic bugs **21**
4.1.1 Fossil record and phylogeny of the semiaquatic bugs 21
4.1.2 Transition from land to water 23
4.1.3 Leg morphology and locomotion on water 23
4.1.4 Sexual interactions 25
4.2 Semiaquatic bugs as laboratory models **25**
4.2.1 Semiaquatic bug life cycle 26
4.2.2 Laboratory population rearing 27
4.2.3 Experimental advantages 29
4.2.3.1 Genomic, transcriptomic and phylogenetic analysis 29
4.2.3.2 Gene expression analysis 29
4.2.3.3 Gene function analysis 30
4.2.3.4 Trait function analysis 30

Objectives of the presented thesis **31**

Chapter 1 ..... **33**

Diversity in morphology and locomotory behavior is associated with niche expansion in the semiaquatic bugs. **33**

Chapter 2 ..... **59**

Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. **59**

Chapter 3 ..... **91**

Predator strike shapes antipredator phenotype through new genetic interactions in water striders. **91**

Chapter 4 ..... **111**

Taxon-restricted genes at the origin of a novel trait allowing access to a new environment. **111**

Chapter 5 ..... **161**

Dimorphism under sexual conflict controlled by the combination of sex determination and Hox genes. **161**

General discussion **181**

References **185**

General introduction and literature review

I. Ecological pressures, fitness and adaptive phenotypes

1.1 Selective pressures and survivorship

The ecological pressures of the habitat where species live directly challenge every organism, and by impacting the different components of fitness, select for the most adapted individuals. These environmental selective pressures include physical environmental factors (temperature, pressure, water current velocity, pH, mineral composition of the niche), types of food resources and availability, predator-prey interactions, density of population, and competition with other members (Chase and Leibold, 2003; Ghalambor et al, 2003). These diverse interactions result in pressures on species that select individuals with best-adapted phenotypes and underlying genotypes. Adaptive phenotypes increase the overall fitness of organisms under specific environmental conditions (Orr 2005a, Orr 2005b). Organisms that live longer generate more offspring, and therefore transmit genes involved in the development of selected adaptive traits to the next generations. The differences in selective pressures and genetic contents between organisms have a central role in generating the diversity of phenotypes observed in nature.

1.2 Examples of adaptive traits and their impacts on evolution

Adaptive phenotypes are quite diverse in nature and can have different magnitudes of impact by increasing the fitness of organisms in their habitat or by changing the evolutionary trajectory of a complete group of organisms through diversification (Schluter 2000; Losos 2009). In the first case, adaptive phenotypes increase the fitness of organisms without affecting the rate of species diversification. Some historical adaptive phenotypes are well known in the literature such as the dark coloration affecting moths and other insects following the industrial revolution (Kettlewell 1955; Lees & Creed 1975). The coloration directly impacts survival when the insect matches the color of the tree background thereby avoiding detection and predation by birds. A second example is the different morphologies of armor and spines in the three-spine sticklebacks (Barrett et al, 2008; Bell et al, 1993; Klepaker

1993) (Figure 2A) associated with the presence of predators and the calcium concentration of the water where sticklebacks live. A third example is the mimicry in *Heliconius* butterflies used to mimic toxic preys and intimidate predators (Walbank et al, 2016) (Figure 2B).

While they are spectacular, the three examples cited above have not impacted the evolutionary history of their respective groups as strongly as key innovations did. Key innovations represent a particular case of adaptive traits that had a huge impact on the diversification and evolution of a group of organisms (Schluter 2000). A key innovation is an adaptive phenotype that provides access to unexploited resources and that allowed a rapid burst of species diversification due to a release in selective pressures (Schluter 2000; Losos 2009). Examples include:

- The pharyngeal jaw of the cichlids fishes from the great African lakes that allow them to exploit the different types of resources (from plankton to snails) present in their respective sub niches (Muschik et al, 2012).

- The beak of birds with the historical example of the Darwin finches that use their beaks to exploit the different types of seeds present in the Galapagos Islands (Figure 2C) (Grant & Grant 2008 and 2014).

- The flower of plants that attracts insects or birds, which contribute to disseminate the pollen, and therefore increase the mix of gametes (Figure 2D) (Albert et al, 2002).

- Insect wings that open new unexploited resources for the groups and allowed escaping predators, dispersal and foraging behaviors (Averof & Cohen, 1997).

- Toepads and limb length of anoles lizards that allow the different species to exploit different sub niches in the forest of the Caribbean Islands (Losos 2009).

While a large number of phenotypes are considered as adaptive due to their association with specific evolutionary success, the adaptive value is difficult to evaluate but remains fundamental to understand how environmental factors select for specific phenotypes and how these phenotypes allow their bearer to face ecological pressures.

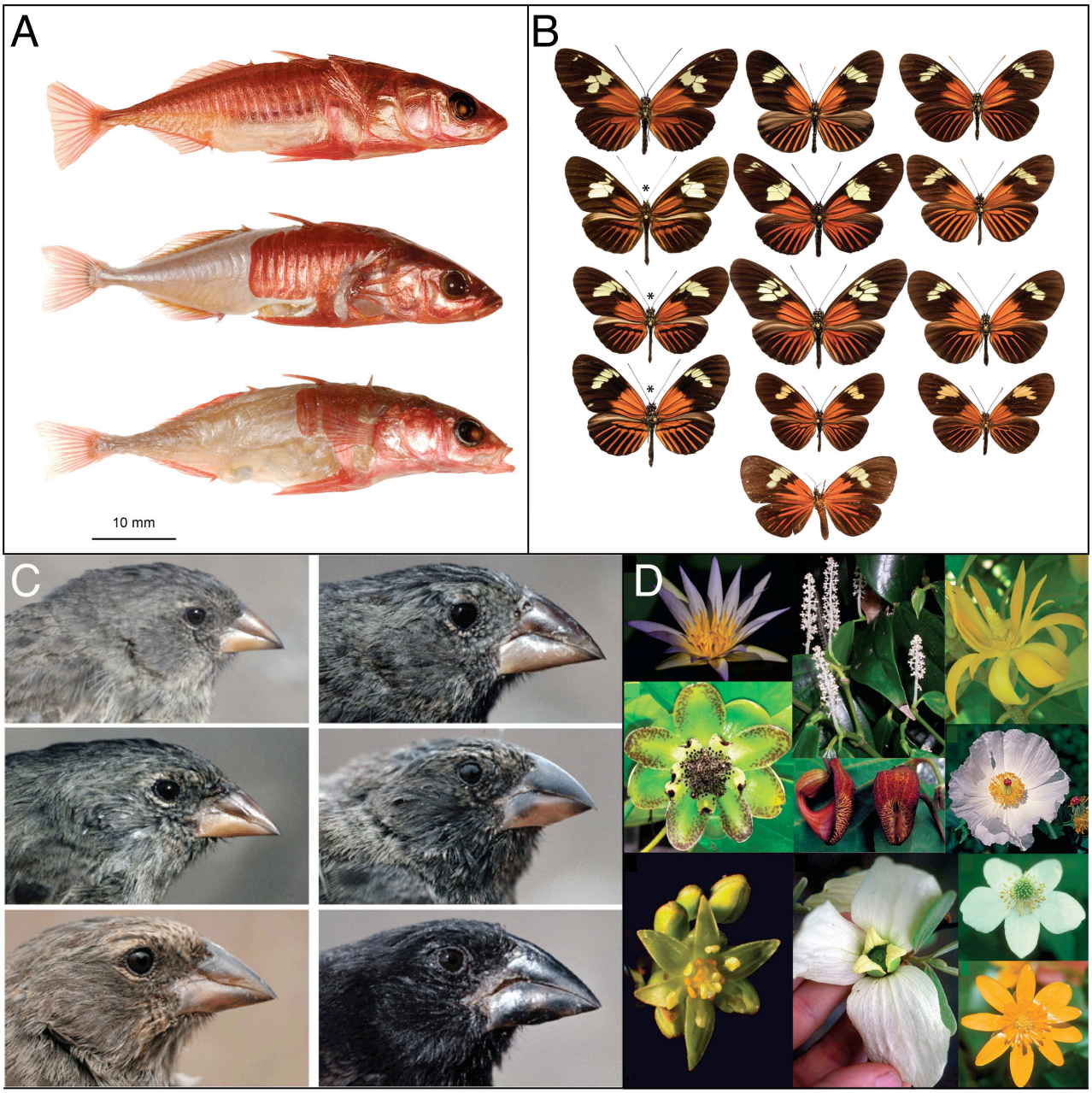


Figure 2: Examples of adaptive traits.
(A) the armor plates and spines in three-spine sticklebacks (from Barrett et al, 2008), (B) the mimicry of *Heliconius* butterflies (from Walbank et al, 2016), (C) The beak of Darwin finches (from Grant & Grant, 2014), (D) the flower of flowering plants (from Chamberdali et al, 2016).

II. Sexual interactions, fitness and sexual dimorphism

Sexual selection is part of natural selection and defines mating success, i.e. the ability to secure mating partners. Sexual selection operates via three different processes, namely male-male competition, mate choice and sexual conflict (Andersson 1994; Arnqvist & Rowe, 2005). These three processes contribute to generate diversity by driving sexual dimorphism; however, in the presented thesis I will only focus on sexual conflict.

2.1 Sexual conflict

Under sexual conflict, males and females have opposite benefits and costs in terms of mating rate. Females have a limited number of eggs and are favored by selection to mate less frequently than males. Additional superfluous matings are known to be unnecessary and even costly by exposing females to predation, injuries or infections (Figure 3A) (Arnqvist & Rowe, 2005). For males the situation is different. Even after mating, males are not sure to fertilize eggs due to sperm competition. Therefore males with high mating rate are favored by selection because they benefit to mate as much as possible with the maximum number of partners to sire the highest number of offspring possible (Figure 3A). This opposition in benefits and costs between males and females is the source of sexual conflict (Arnqvist & Rowe, 2005).

2.2 Antagonistic coevolution and arms race

Sexual conflict can drive the coevolution of antagonistic behavioral, morphological or physiological traits between males and females through interlocus conflict (Arnqvist & Rowe, 2005). If a trait appears in one sex, in response the second sex can also develop traits to counter the first adaptation. This process of antagonistic coevolution can lead to escalation and arms race (Dawkins & Krebs, 1979; Arnqvist & Rowe, 1995; Weigensberg and Fairbairn, 1996). Examples of traits that have evolved under sexual conflicts include the coevolution of the genitalia in ducks (Brennan et al, 2009; Brennan & Prum, 2015) and spiders (Kuntner et al, 2009 and 2016), the male intromittent organ and female paragenitalia structure in traumatically inseminating

bugs (Tatarnic and Cassis, 2010, Morrow et al, 2003; Reihart et al, 2003), the harming penis and resistance of female genitalia in beetles (Gay et al, 2011), male and female plugs (Kuntner et al, 2012), male sex-peptide and female microRNA resistance in *Drosophila* (Wigby et al, 2005; Fricke et al, 2014), antennae and the abdominal shape of semiaquatic bugs (Khila et al, 2012; Arnqvist & Rowe, 2002), male adhesive discs and the modification of female dorsal part in diving beetles (Bergsten et al, 2001; Bergsten & Miller, 2007). Sexual conflict therefore highly contributes to sexual dimorphism and represents an important source of morphological diversity. While arms races also exist in the context of predator-prey interaction, in the presented thesis I will only focus on sexual arms race and its impact on morphological diversity.

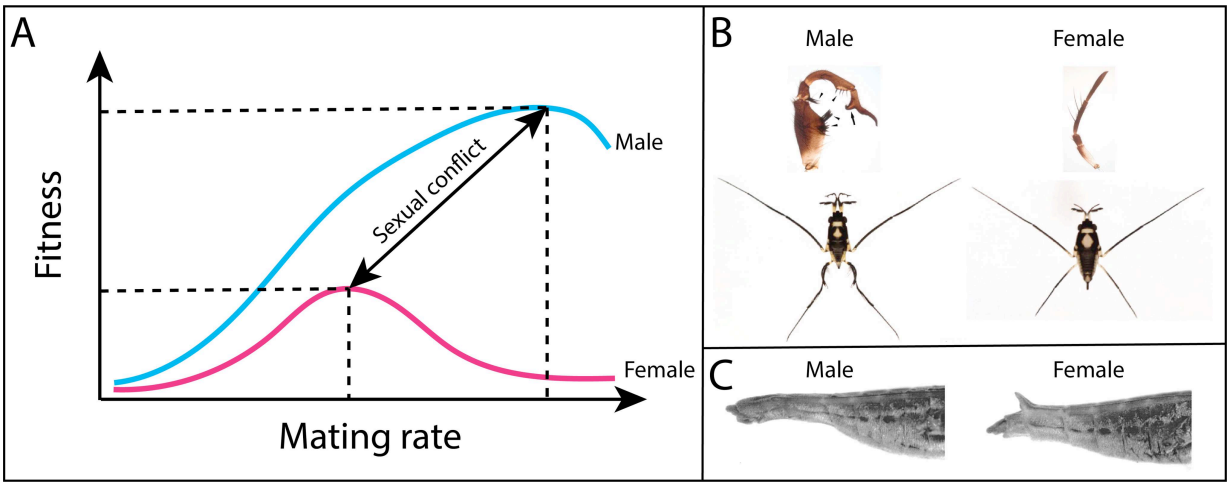


Figure 3: Sexual conflict and antagonistic coevolution.
(A) Schematic representation of the differences in fitness and mating rate between males and females at the origin of sexual conflict. (B) Modification of male antennae in *Rheumatobates rileyi* that increases mating success during pre-mating struggles (modified from Khila et al, 2012). (C) Modification of abdomen shape in *Gerris* with a spike in female abdomen to increase distance between male and female genitalia and counter adaptation in the male with curved abdomen allowing the male to avoid the spike (from Arnqvist & Rowe, 2002).

III. Genetic basis of morphological evolution

The phenotype of each individual is shaped during development and is determined by the combination of developmental genetic mechanisms and the action of hormones and metabolic pathways (Hughes & Kaufman, 2002; Jindra et al, 2015; Casasa et al, 2017). However in the presented thesis I will only focus on developmental genetic mechanisms.

3.1 Molecular genetic mechanisms

The genome of each individual and its expression during development contribute to shape the morphology of an organism. Genes can be lost (Albalat & Cañestro, 2016) or emerge *de novo* and acquire new function (Rastogi & Liberles, 2005). DNA variants resulting from mutations, gene isoforms resulting from alternative splicing, and epigenetics can influence gene activities or protein interactions and therefore change the pattern and the level of gene expression or create new genetic interactions involved in the evolution of traits (Sniegowski et al, 2000; Andreadis et al, 1987; True & Carroll, 2002).

3.2 Examples of genes involved in morphological diversity

Certain genes have a large impact on the development of organisms and highly influence the phenotype of species. This is the case of the Homeobox genes (Hox genes). These genes are organized in a cluster in the genome (Figure 4) and code for transcription factors characterized by the presence of the homeobox domain (homeodomain), a DNA binding domain conserved among the homeobox containing genes (Garcia-Fernandez 2005; Gehring 1985; McGinnis et al, 1984; Scott & Weiner, 1984). They regulate the expression of a high number of downstream targets (Pavlopoulos & Akam, 2011) through interaction with a variety of diverse co-factors (Merabet & Hudry, 2012). Hox genes are involved in establishing the identity of segments along the body axis during embryogenesis in a number of lineages (McGinnis & Krumlauf, 1992; Hughes & Kaufman, 2002; Angelini et al, 2005) and are involved in late development until adulthood (Chesebro et al, 2009). They are highly

conserved between vertebrates and invertebrates and the evolution of the cluster has contributed to the diversification of morphology (Figure 4) (Hughes & Kaufman, 2002; Akam et al, 1994; Averof & Akam 1993; Averof & Akam 1995; Krumlauf 1994; McGinnis et al, 1984; McGinnis & Krumlauf 1992; Scott & Weiner 1984; Pearson et al 2005; Garcia-Fernandez 2005). In addition to their roles in segment identity, the Hox genes are also involved in the fine tuning of specific structures such as the sex combs (Tanaka et al, 2011; Chesebro et al, 2009), the abdomen pigmentation (Kopp et al, 2000), the genitalia (Glassford et al, 2015), the eye spot in butterfly (Shirai et al, 2012), the halters in *Drosophila* (Hersh et al, 2007) and the leg length in crickets and semiaquatic bugs (Mahfooz et al, 2007; Khila et al, 2009).

Contrary to Hox genes that are highly central during development, other genes have more subtle effects, which generally allow a fine-tuning of the phenotype. These genes have been found to include transcription factors, enzymes, signaling proteins or transmembrane proteins. Examples include the level and pattern of expression of the transcription factor Alx3 and the transmembrane receptor Agouti in rodent color pattern involved in camouflage (Manceau et al, 2011; Mallarino et al, 2016), the level of expression of the receptor BMP4 in the diversification of beaks in Darwin finches (Abzhanov et al, 2004), the glycoprotein Wnt in the development of wings and halters in *Drosophila* (Sharma & Chopra, 1975) and the transcription factor Distal-less in the antennae of some water strider species (Khila et al, 2012) and wing spots in *Drosophila* (Arnoult et al, 2013) and butterflies species (Shirai et al, 2012).

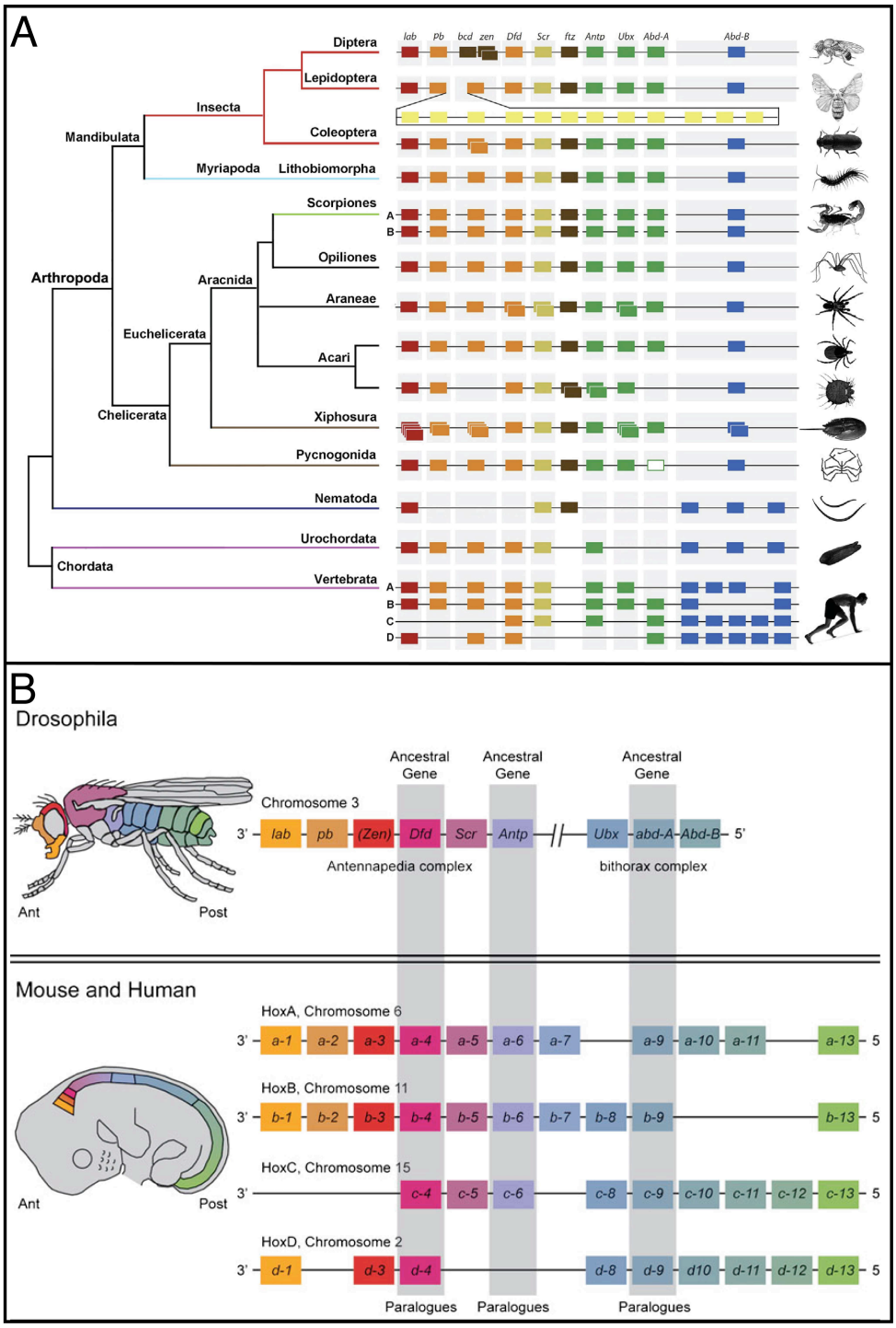


Figure 4: Evolution of the Hox cluster.
(A) Evolution of the Hox cluster across phylogeny (from Di et al, 2015). (B) Comparison of the Hox cluster between *Drosophila* and mammals (from Pang & Thompson 2010). In *Drosophila*, the Hox cluster is composed of the genes: *Labial*, *Proboscipedia*, *Deformed*, *Sex Combs Reduced*, *Antennapedia*, *Ultrabithorax*, *Abdominal-A* and *Abdominal-B*. The Hox genes are organized in this specific order along a unique chromosome and are expressed along the anterior-posterior axis of the embryo in the same order (Garcia-Fernandez 2005; McGinnis et al, 1984; Scott & Weiner 1984; McGinnis & Krumlauf 1992). In vertebrates, there is a total of 39 Hox genes organized into four chromosomal complexes (Duboule 1992; Kessel & Gruss 1990; Krumlauf 1994; McGinnis & Krumlauf 1992).

Another class of genes involved in morphological diversity is the sex determination genes. These genes code for transcription factors with isoforms specific to males and females that contribute to shape sexual dimorphism (Williams & Carroll 2009). Examples include *doublesex* in the development of horns in beetles (Kijimoto et al, 2012; Gotoh et al, 2016; Ledon-Rettig et al, 2017), the sex combs and abdomen pigmentation in *Drosophila* species (Tanaka et al, 2011; Kopp et al, 2000), mimicry in butterflies (Kunte et al, 2014), the proboscis in mosquitoes (Mysore et al, 2015). Other examples include *sox9* in the sex determination in mice (Sekido & Lovell-Badge 2008) or *Transformer 1* in *Caenorhabditis elegans* (Zarkower & Hodgkin 1992). These genes differentially controls downstream targets in both sexes and contribute to generate sexual dimorphism and ultimately morphological diversity.

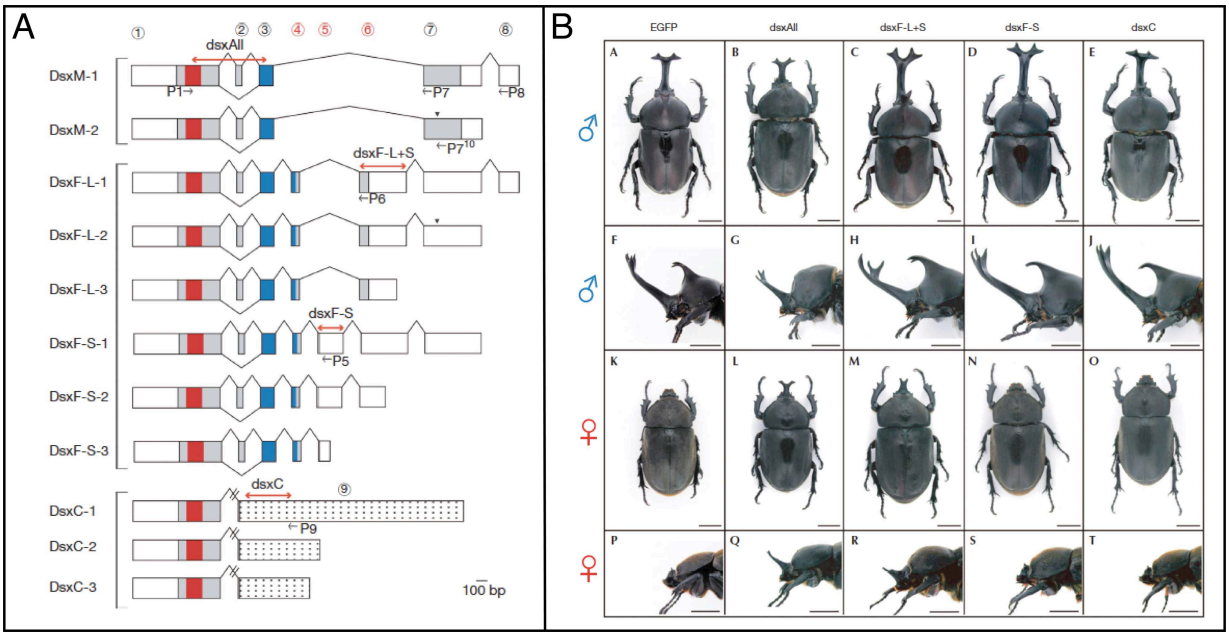


Figure 5: The sex determination gene *doublesex* involved in sexual dimorphism.
(A) The different isoforms of *dsx* gene in the Japanese rhinoceros beetle with male isoforms (M), female isoforms (F) and common isoforms (C). (B) The dimorphism between males and females and the effect of specific isoform knockdown on horn development (from Ito et al, 2013).

IV. The semiaquatic bugs as model systems

4.1 Evolution of semiaquatic bugs

In the presented thesis, I used the Gerromorpha (order Hemiptera, suborder Heteroptera), also known as semiaquatic bugs to investigate how biodiversity emerges. These insects are an important model for the study of sexual selection, ecology, behavior, morphology, biogeography, physics, and development. The Gerromorpha therefore represent a suitable model lineage to combine several approaches in order to understand the relationship between selective forces and genetic mechanisms that contribute to generate biodiversity.

4.1.1 Fossil record and phylogeny of the semiaquatic bugs

The group of the Gerromorpha is a monophyletic group of insects composed of about 2200 species split in eight families (Mesoveliidae, Hebridae, Macroveliidae, Hydrometridae, Paraphrynoveliidae, Vellidae, Gerridae, Hermatobatidae) present worldwide except above a certain latitudes corresponding to the poles (Andersen 1982). The fossil record of the Gerromorpha is not very rich although we can still find some interesting specimens. The most ancient confirmed fossil of Gerromorpha dates around 150 million years ago from the Upper Jurassic (Damgaard 2008a). All the different families of the Gerromorpha group seem to be present in the Mesozoic (Damgaard 2008a; Damgaard 2008b). These fossils give interesting information about the morphology of the extinct species such as their leg ground plan, the length of their different legs but also their possible diets (Figure 6A-D) (Perrichot et al, 2005; Damgaard 2008a). The morphology of the extant species was also used to reconstruct the phylogenetic relationships between the different species of the semiaquatic bugs. While the 8 different families remain well defined using morphological analysis, the utilization of the new sequencing technologies have highlighted numerous nodes of the phylogeny that are not well supported and required more genetic information to clearly define them (Daamgard 2008a; Daamgard 2008b).

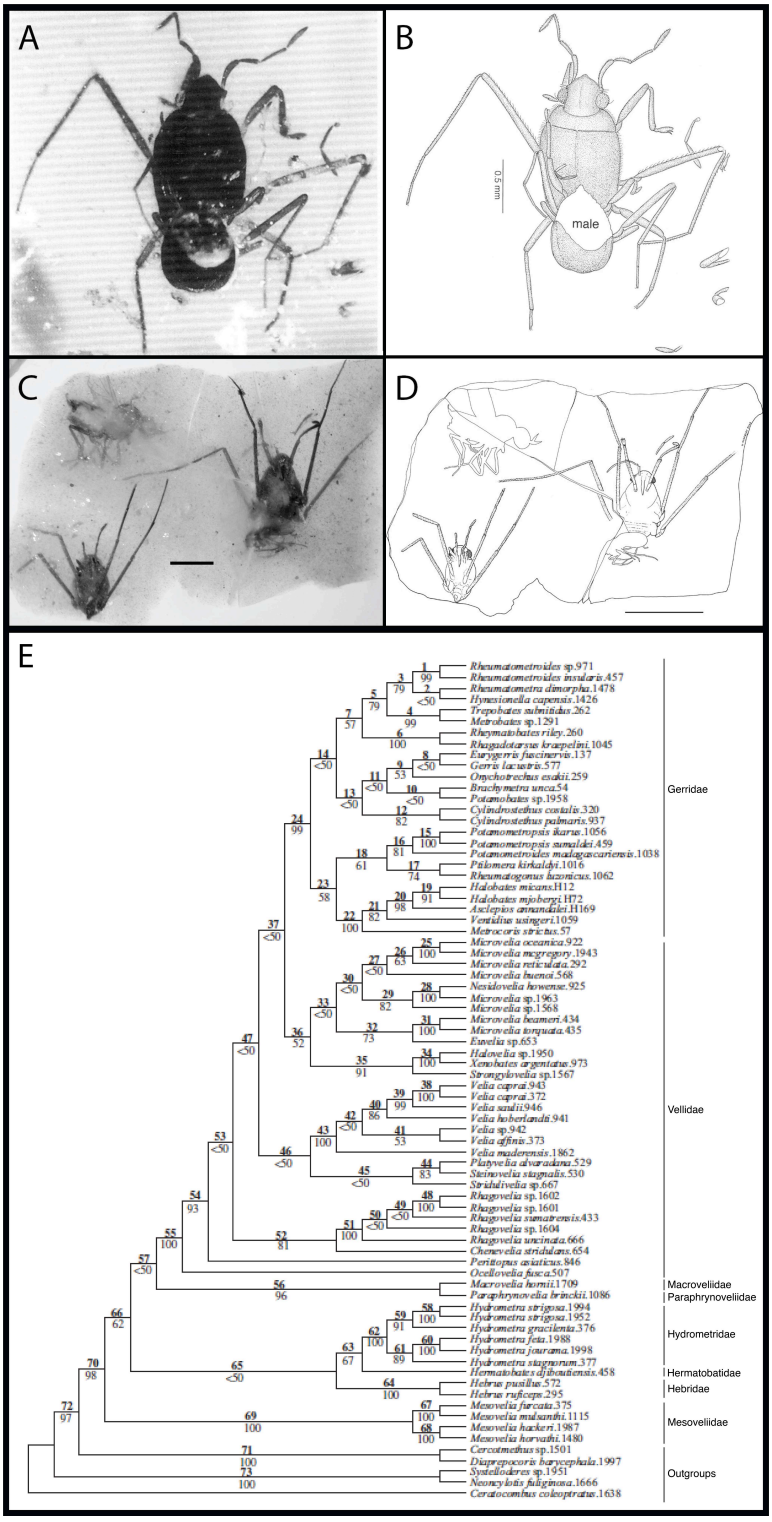


Figure 6: Fossils and phylogeny of the semiaquatic bugs
(A-B) Specimen and drawing of *Halovelia electrodominica* (Veliidae) specimen from Dominican amber (from Andersen & Poinar, 1998). (C-D) Specimen and drawing of *Cretogerris albianus* (Gerridae) from Late Albian amber (from Perrichot et al, 2005). This fossil is the oldest representative of the family Gerridae with drosophila as potential prey. (E) Phylogenetic reconstruction of the Gerromorpha group with annotation of the different families (adapted from Damgaard 2008b). The phylogeny was reconstructed using a combination of morphological traits and four genetic mitochondrial markers. Numerous nodes are weakly supported.

4.1.2 Transition from land to water

The Gerromorpha share a common terrestrial ancestor that is thought to have transited to humid environments such as mud (Andersen 1982). The bugs became adapted to this environment by an increase of the bristle density on the legs and on the body that allows them to stay dry and to avoid becoming trapped in the water surface film. This trait is shared with their immediate and exclusively terrestrial sister group, the Dipsocoromorpha that inhabit environments such as riverbanks or humid litter, and that are also able to float on water (Heiss & Péricart 2007). It seems conceivable that the increase in bristle density was the primary adaptation that allowed the ancestor of the Gerromorpha to stand on water. A second adaptation to life on the water involves the elongation of the legs (Andersen 1982). Legs diversified across species according to the natural niches where the species live. Now Gerromorpha occupy a variety of different sub niches at the water-air interface such as humid soils or rocks, puddles, shores, water surface covered by plants, stagnant water such as ponds, running water such as stream and even the open ocean (Andersen 1982).

4.1.3 Leg morphology and locomotion on water

The environment where the semiaquatic bugs lives is a source of a variety of selective pressures that are distinct from those acting on terrestrial insects. These bugs are well adapted to life on water surface in many aspects including feeding behavior, the presence of winged morphs, their biology of reproduction and egg laying behaviors, predator-prey interactions, and finally appendage morphologies and associated mode of locomotion (Andersen 1976; Andersen 1982). The Gerromorpha exhibit a huge variety of morphologies, particularly for the legs, which are associated with the sub niches that distinct species occupy. Species that live on the margin of water bodies share the ancestral ground plan, where the first leg is shorter than the second leg which in turn is shorter than the third leg (Figure 7). These species are able to move on both land and water using the tripod gait, which is known to be the ancestral locomotion mode (Andersen 1976). Derived species that specialize exclusively in open water zones such as ponds, streams or oceans possess a derived ground plan with the second leg longer than the third that allow them to move by rowing or jumping (Figure 7) (Andersen 1976). The *Rhagovelia* genus, a genus on the

Veliidae family, has even evolved a fan on the second legs and sometimes on the third legs. This structure is specifically associated with the life in the middle of fast flowing stream water (Andersen 1976). During locomotion on water the first leg is used to balance the body, the second leg generates the movement and the third leg orientates the body (Andersen 1976; Hu & Bush, 2010). More details about the different modes of locomotion and their impact on the diversification of the group will be developed in the coming chapters.

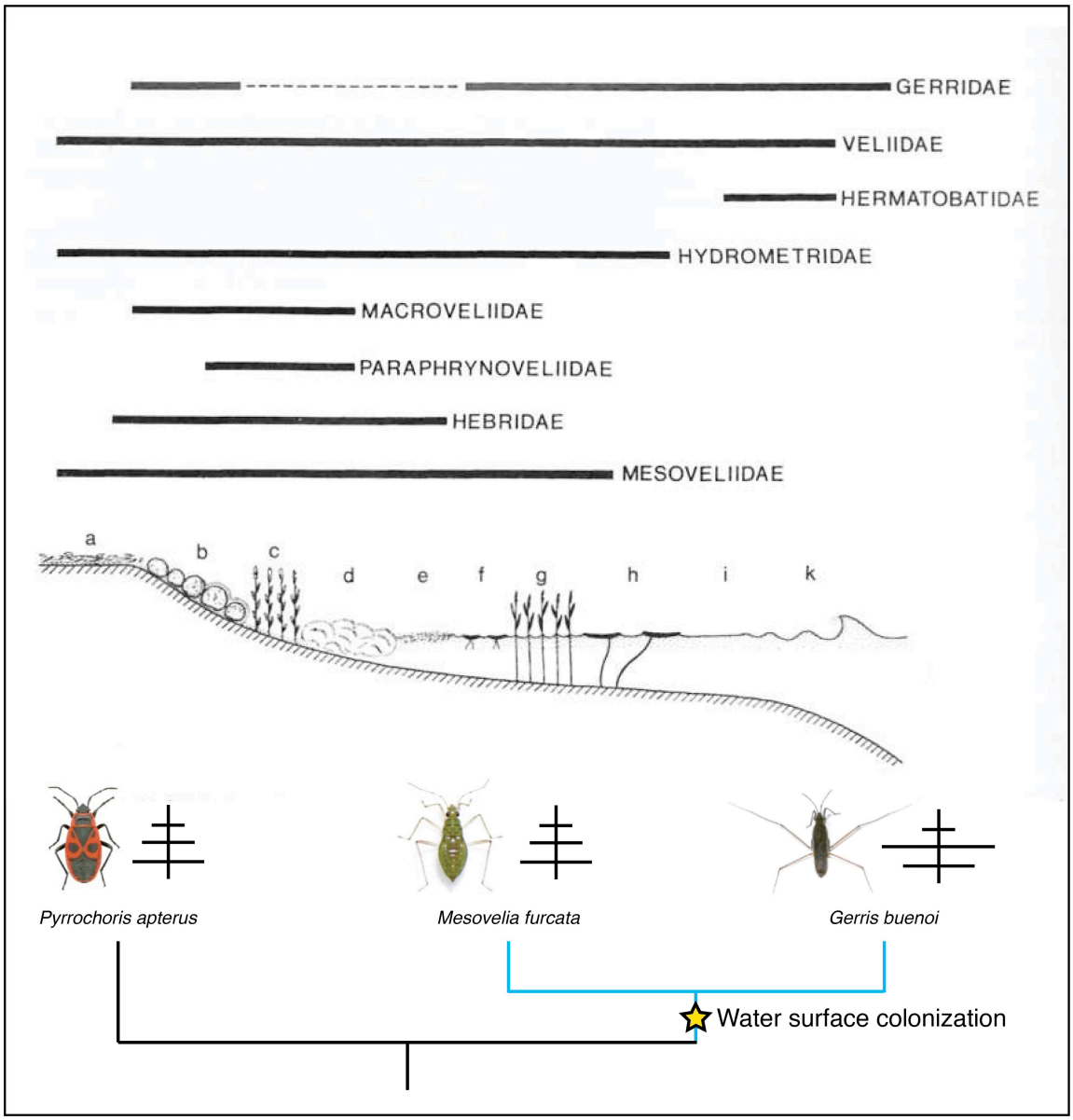


Figure 7: Habitat, morphology and locomotion of the semiaquatic bugs.
The upper part represents the distribution of the eight families of Gerromorpha in their respective sub niches (from Andersen 1982) in association with the lower part that represents the leg ground plan of the terrestrial *Pyrrhocoris* (Pentatomomorpha, Pyrrhocoridae), the basally branching species *Mesovelia* (Mesoveliidae), and the derived *Gerris* (Gerridae). Both *Pyrrhocoris* and *Mesovelia* have the ancestral leg ground plan and use the tripod gait while *Gerris* have the derived ground plan and use the rowing gait during locomotion.

4.1.4 Sexual interactions

Sexual interactions and associated behaviors are quite diverse across the Gerromorpha. We observe the three major behavioral mechanisms underlying sexual selection: male-male competition in which males fight each other to dominate a territory and mate with females (Hayashi 1985; Sih et al, 2002), female choice for which the female choses the best male (Hayashi 1985) and sexual conflict for which males and females fight during pre-mating struggles (Arnqvist & Rowe, 2005). Mate guarding, during which a male guards his mating partner to prevent other males from copulating with this partner, is also widespread in the group (Rubenstein 1989). Others more unusual behaviors related to sexual selection are also easily observable such as the attraction of predators to make females receptive (Han & Jablonski 2010) or the preparation of egg laying spot by the male for the female (Hayashi 1985). In the first strategy, *Gerris graciliocornis* males intimidate females that do not accept mating attempts in order to decrease their fitness by the attack of predators. As a result, females accept more often mating partners (Han & Jablonski 2010). The second strategy consists of *Gerris elongatus* males preparing egg-laying spot. Once the male considers the spot as ready, it vibrates his legs to attract and courtship females (Hayashi 1985). Before accepting the mating, the female evaluate the quality of the egg laying spot prepared by the male. The female accepts or rejects the mating depending of the spot quality (Hayashi 1985).

4.2 Semiaquatic bugs as laboratory models

The diversity of semiaquatic bug morphologies, the diversity of their habitat, their phylogenetic relations and the diversity of their behaviors and locomotion make them a suitable system to understand how ecological pressures act on the phenotype and impact the evolutionary trajectory of lineages. In order to understand the developmental genetic mechanisms contributing to the emergence of their morphological diversity, different species of the semiaquatic bugs had to brought and reared in laboratory conditions to be used as model systems for experiments.

4.2.1 Semiaquatic bug life cycle

Gerromorpha egg laying behavior (presented in the next paragraph) varies depending on the species and the microhabitat, but eggs have to be in constant contact with water to avoid desiccation (Andersen 1982). Gerromorpha are hemimetabolous insects that do not undergo metamorphosis (Andersen 1982), instead, appendages develop from limb buds during embryonic development (Khila et al, 2009). The embryo hatches into a nymph similar to the adult form with fully functional appendages (Figure 8). Then the nymphs undergo five subsequent molts while growing in size until the final adult stage (Figure 8) (Andersen 1982). Sexually mature adults will mate and reproduce closing the cycle.

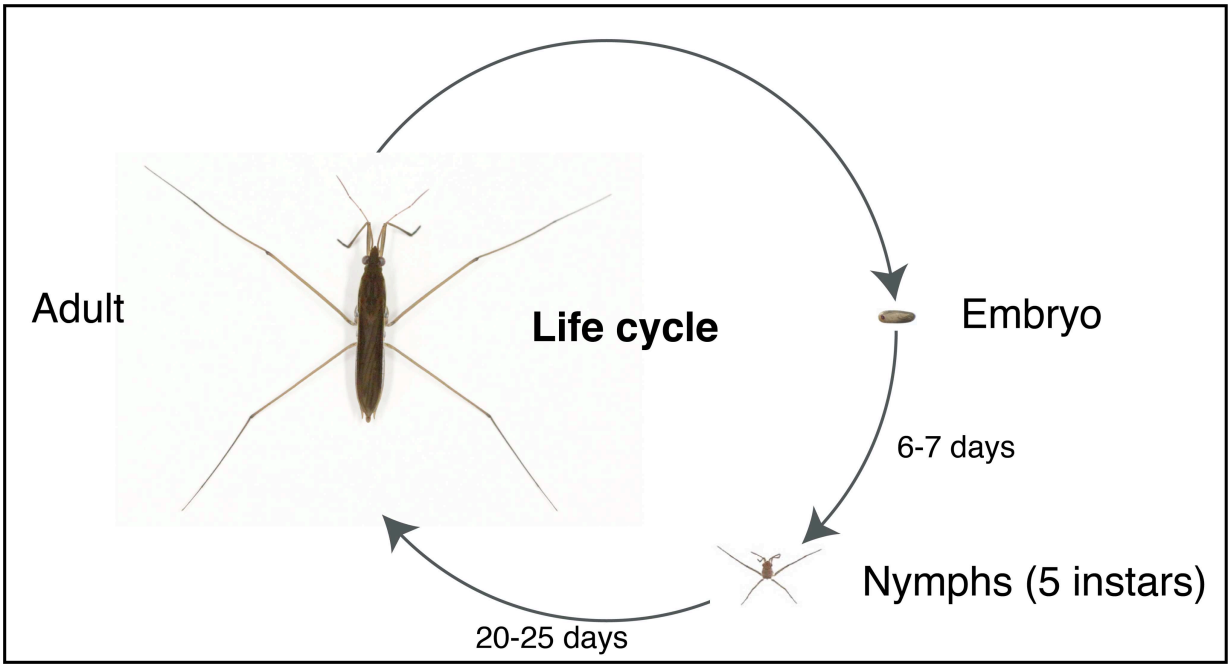


Figure 8: Example of life cycle in semiaquatic bugs (*Limnoporus dissortis*).
Embryonic development takes about 6-7 days. The newborn nymph then undergoes five successive molts (20-25 days) increasing in size until reaching adult form. Once sexually mature, adults mate and reproduce closing the cycle. The times of both embryonic and nymphal development, as the lifetime, are variable across species.

4.2.2 Laboratory population rearing

The Khila lab has been successful in collecting and rearing various species of the semiaquatic bugs in laboratory conditions. We mainly collect bugs in France, French Guiana, Brazil, United States, Canada and Puerto Rico. We keep the insects in containers with water changing every week (Figure 9 A and B). For some species living in stream water the containers are equipped with bubblers to mimic their environment and avoid the formation of bacterial film at the surface, which results in breaking surface tension and the drowning of the insects (Figure 9B). Insects are fed with either fresh or frozen crickets. The environment of the insect facility is controlled in order to optimize conditions for animal rearing. The temperature of the room is maintained at 25°C, the humidity at 50%, and the day and night cycle is set at 14 hours of day and 10 hours of night. We are currently able to maintain 27 species of semiaquatic bugs in the facility (excluding specific lines such as inbred lines used for genomic analysis). Some species collected during fieldwork are kept in the lab until we have collected what we need for a project or until they die. We provide adult females with floating Styrofoam strips (Figure 9C) that simulate the various substrate types used by females to lay eggs in the wild. Species that lay eggs on the edges of floating plant leaves lay their eggs underneath the edge of the Styrofoam strip (Figure 9C). Other species that embed their eggs inside the leaves embed their eggs in the Styrofoam strip, which provides a substrate soft enough to accommodate this behavior. Other species lay their eggs on top of the Styrofoam strip. For the last two examples, we add a wet tissue on top of the Styrofoam stripe to keep eggs constantly humid protecting them from desiccation.

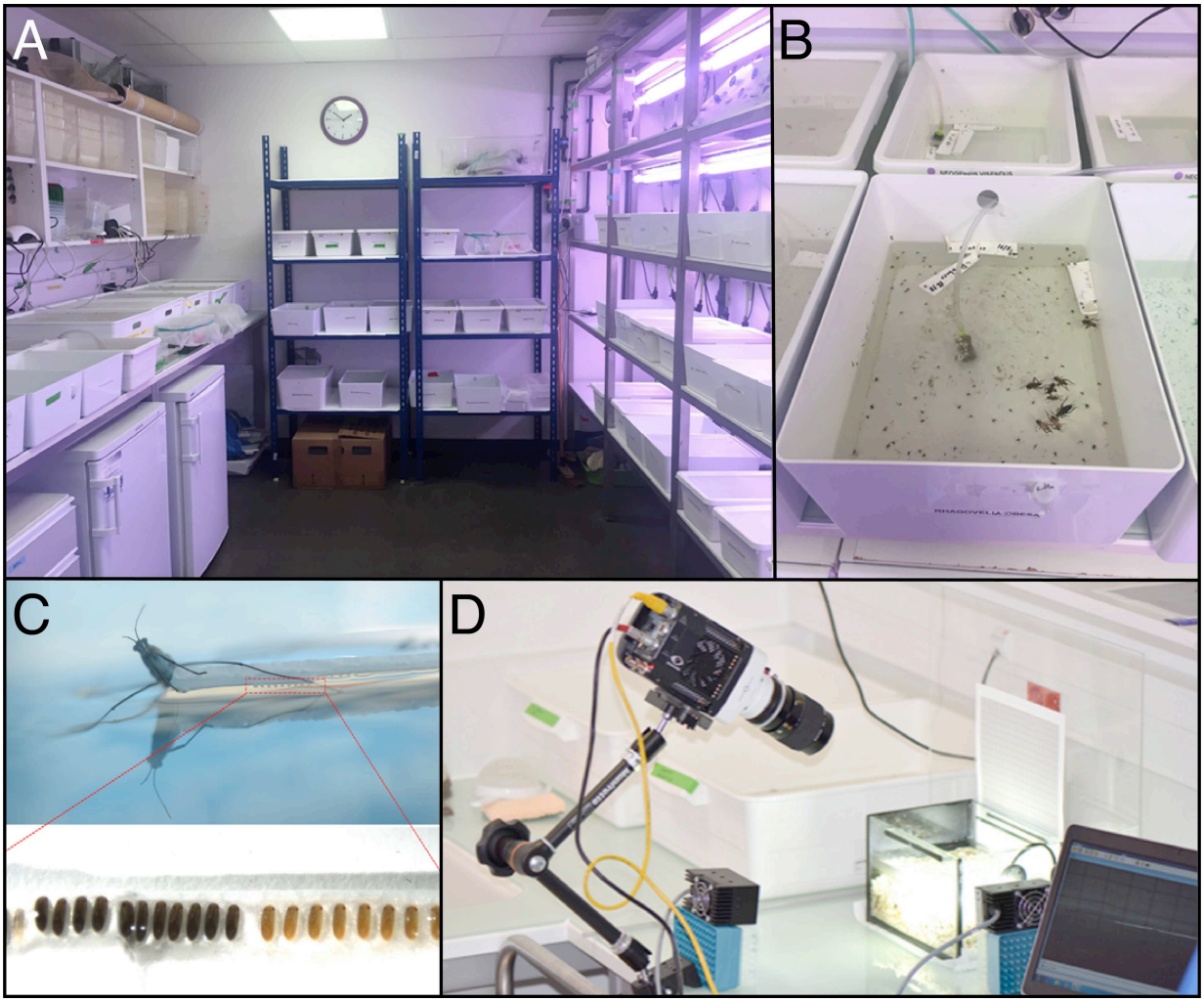


Figure 9: Animal rearing facility.
(A) Picture of our animal facility with (B) an example of bucket with bugs. (C) Female laying eggs on Styrofoam floater with focus on eggs. (D) Example of set-up with high-speed camera we used to quantify jumping performance.

4.2.3 Experimental advantages

In addition to their diversity of morphologies, behaviors, habitats and their ease of rearing under lab conditions, the semiaquatic bugs offer substantial experimental advantages to combine several different approaches in order to reveal the mechanisms underlying their diversity. This makes them an appropriate and powerful system organisms for Eco-Evo-Devo studies.

4.2.3.1 Genomic, transcriptomic and phylogenetic analysis

The development of the next generation sequencing technologies allows us to increase our resources in terms of genetic sequences. In our lab we are working on the assembly and the annotation of the genome of two species of semiaquatic bugs; *Gerris buenoi* (the first sequenced genome of Gerromorpha, paper in preparation) and *Microvelia longipes*. Additionally we have generated a transcriptomic dataset of about a hundred species of semiaquatic bugs from 6 of the 8 families of Gerromorpha. These data represent valuable resources for the Khila lab. It allows us to ask questions about the genetic mechanisms associated with the transition to life on water and gives access to genetic sequences for gene cloning to perform gene expression and function analysis. Finally our database furnishes the information required to conduct phylogenetic analysis. Ultimately this dataset will help to clarify relationships of the Gerromorpha species and solve the uncertain nodes of the phylogeny.

4.2.3.2 Gene expression analysis

Our lab has efficiently adapted and developed gene expression techniques from other model systems such as *in situ* hybridization to detect spatial expression of a given gene, and immuno-staining to reveal protein localization. Double staining is achievable to demonstrate gene interactions and quantitative PCR is also feasible to quantify gene expression.

4.2.3.3 Gene function analysis

Our lab has developed micro-injection of double-stranded RNA (dsRNA) to perform gene function analysis by RNA interference (RNAi) (Hannon 2002; Plasterk 2002; Khila et al, 2009; Khila et al, 2012; Khila et al, 2014; Refki et al, 2015). Both parental and nymphal RNAi work in our system with differences in efficiency depending on the species and genes tested. In parental RNAi, we inject females and recover the phenotypes in the progeny to study gene function during embryonic development. In nymphal RNAi, we inject individuals during nymphal development and we recover the effect in the adult to investigate gene function during late development.

4.2.3.4 Trait function analysis

The semiaquatic bugs represent an interesting model to perform trait function analysis. Their behaviors are quite often associated with their natural habitat and are highly stereotyped. The large numbers of individuals and species allow us to easily make several replicates for a given experiment and make comparison between individuals and species. We used either a normal camera or a high-speed camera to record, observe and measure the performance of each individual for a given behavior (Figure 9D). We also imagine and design specific set-ups to challenge the bugs in conditions that mimic the ecological parameters of their natural habitat.

Objectives of the presented thesis

The Gerromorpha represent an appropriate model to better understand how the relationships between natural/sexual selection and the genotype act on the phenotype and generate biodiversity. The semiaquatic bugs have acquired a number of adaptations associated with the ecological transition to water surface habitat and associated sub niches. These adaptations include a common leg elongation compared to terrestrial insects, different modes of locomotion associated with different leg ground plans and specific structures and behaviors. The Gerromorpha also exhibit a sexual dimorphism that could be more or less pronounced depending on the species, and associated with sexual conflict. However, while a number of traits have been described as adaptive, their adaptive value and their impact on the diversification of the group have never been properly assessed and their genetic basis remains unknown. This thesis aims at studying genetic developmental mechanisms underlying adaptive phenotypes and their ecological relevance in the semiaquatic bugs in order to understand the forces generating diversity. The study addresses this question through 5 chapters:

- **Chapter 1:** In this chapter, I investigated the adaptations associated with the transition to life on water surface habitat and their impact on the diversification of the Gerromorpha.

- **Chapter 2:** Here I analyzed the role of the Hox gene *Ultrabithorax* in the reversal of the leg ground plan associated with the difference in locomotion between basally branching and derived species.

- **Chapter 3:** Here I investigated the downstream targets of *Ubx* as a potential mediator controlling the opposite developmental effects on T2-leg and T3-leg growth and the impact on predator-prey interaction in derived species.

- **Chapter 4**: In this chapter I investigated the genetic mechanisms shaping an evolutionary novelty, the propelling fan of the *Rhagovelia* genus, and the importance of this structure in niche specialization.

- **Chapter 5:** Here I studied the roles of the sex determination gene *doublesex* and the Hox genes *Sex combs reduced* and *Ultrabithorax* involved in dimorphism under sexual conflict in the genus *Rhagovelia*.

Chapter 1

Diversity in morphology and locomotory behavior is associated with niche expansion in the semiaquatic bugs.

Paper published in *Current Biology*

First author

Synopsis of the paper:

The semiaquatic bugs invaded the water surface habitat and diversified to occupy a wide variety of sub niches. In this new habitat the Gerromorpha encountered new selective pressures imposed by the new substrate of locomotion and new interspecies interactions. Their adaptation has been possible through the diversification of their appendages and the different modes of locomotion. Basally branching species that retain the ancestral leg ground plan, similar to terrestrial insects with the T1-leg shorter than T2-leg shorter than T3-leg, use the tripod gait and can move on both water and ground. Derived species possess the derived leg ground plan with the T2-leg longer than the T3-leg and use the rowing gait and move only on water. While an association between leg morphology, mode of locomotion and habitat was suggested, we did not know the adaptive value conferred by the different types of legs and associated mode of locomotion as well as their impact on the diversification of the group. In this paper, we investigated the pattern of association between the different species living in different sub niches, their speed, their mode of locomotion and associated biomechanical characteristics in a phylogenetic context to understand the impact on the diversification of the semiaquatic bugs.

Summary of results:

Phylogenetic analysis showed correlation between speed and habitat and found that species living on the ground move with low speed while species living on water move with high speed. The increase in speed has been possible by different morphological and behavioral changes. The Gerromorpha have longer legs compared to terrestrial insects. They extend their legs more during locomotion on water than on the ground, which increases the amplitude of movements. Species using the tripod gait increase their speed by increasing their stroke frequency while species using the rowing gait are able to generate high speed with low stroke frequency. This derived mode of locomotion saves energy for species that live constantly on water. It has also resulted in an increase in body size.

Conclusion:

Our results show how the environmental selective pressures of the different sub niches of the water surface habitats have impacted the morphology and behavior of different Gerromorpha species. We observed an association between habitat, speed and mode of locomotion and we showed that life on the water surface required a high speed of locomotion. We also determined that the increase in speed has been possible for species using the tripod gait by the combination of increased leg length, increased amplitude of movement and increased stroke frequency. Species using the rowing gait are able to generate high speeds by the simultaneous sculling movement of their T2-leg. This new mode of locomotion could contribute to the evolution of Gerromorpha species with larger body sizes by overcoming the tradeoff that exists between stroke frequency and body size.

Author contributions:

Here are the contributions of the different authors:

Antonin Crumière: insect sampling (fieldwork), measurement of morphological characteristics, collection of DNA sequences, phylogenetic reconstruction, reconstruction of ancestral states and analysis, acquisition of videos, quantification of biomechanical parameters and energy, design of the study and redaction of the paper.

Emilia Santos: insect sampling (fieldwork), phylogenetic reconstruction, treatment of transcriptome dataset.

Marie Sémon: reconstruction of ancestral states and analysis.

David Armisén: treatment of transcriptome dataset.

Felipe Moreira: species identification.

Abderrahman Khila: insect sampling (fieldwork), measurement of morphological characteristics, design of the study and redaction of the paper.

Report

Current Biology

Diversity in Morphology and Locomotory Behavior Is Associated with Niche Expansion in the Semi-aquatic Bugs

Highlights

- Semi-aquatic bugs are adapted to life on water surface niches worldwide
- Life on the water surface requires high locomotory maximum speed
- Increased speed was achieved through changes in leg length and locomotion behavior
- Derived lineages evolved rowing, an energy-efficient mode of locomotion on water

Authors

Antonin J.J. Crumière, M. Emilia Santos, Marie Sémon, David Armisén, Felipe F.F. Moreira, Abderrahman Khila

Correspondence

abderrahman.khila@ens-lyon.fr

In Brief

During evolution, the semi-aquatic bugs colonized a variety of water surface niches from small puddles to open oceans. Crumière et al. show that access to this new habitat is associated with morphological and behavioral changes that determine locomotory speed. Variation in locomotion behavior in distinct lineages is correlated with niche specialization.

Crumière et al., 2016, Current Biology *26*, 3336–3342
December 19, 2016 © 2016 The Author(s). Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.cub.2016.09.061

CellPress

Current Biology
Report

Diversity in Morphology and Locomotory Behavior Is Associated with Niche Expansion in the Semi-aquatic Bugs

Antonin J.J. Crumière,[1] M. Emilia Santos,[1] Marie Sémon,[2] David Armisén,[1] Felipe F.F. Moreira,[3] and Abderrahman Khila[1,4,*]

[1]Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5242, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France

[2]Laboratoire de Biologie et de Modélisation de la Cellule, Ecole Normale Supérieure de Lyon, CNRS, Université de Lyon, 69007 Lyon, France

[3]Laboratório de Biodiversidade Entomológica, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21040-360, Brazil

[4]Lead Contact

*Correspondence: abderrahman.khila@ens-lyon.fr

http://dx.doi.org/10.1016/j.cub.2016.09.061

SUMMARY

Acquisition of new ecological opportunities is a major driver of adaptation and species diversification [1–4]. However, how groups of organisms expand their habitat range is often unclear [3]. We study the Gerromorpha, a monophyletic group of heteropteran insects that occupy a large variety of water surface-associated niches, from small puddles to open oceans [5, 6]. Due to constraints related to fluid dynamics [7–9] and exposure to predation [5, 10], we hypothesize that selection will favor high speed of locomotion in the Gerromorpha that occupy water-air interface niches relative to the ancestral terrestrial life style. Through biomechanical assays and phylogenetic reconstruction, we show that only species that occupy water surface niches can generate high maximum speeds. Basally branching lineages with ancestral mode of locomotion, consisting of tripod gait, achieved increased speed on the water through increasing midleg length, stroke amplitude, and stroke frequency. Derived lineages evolved rowing as a novel mode of locomotion through simultaneous sculling motion almost exclusively of the midlegs. We demonstrate that this change in locomotory behavior significantly reduced the requirement for high stroke frequency and energy expenditure. Furthermore, we show how the evolution of rowing, by reducing stroke frequency, may have eliminated the constraint on body size, which may explain the evolution of larger Gerromorpha. This correlation between the diversity in locomotion behaviors and niche specialization suggests that changes in morphology and behavior may facilitate the invasion and diversification in novel environments.

RESULTS AND DISCUSSION

Phylogeny and Ancestral Habitat Reconstruction

To reconstruct the ancestral habitat, we first built the phylogeny for our sample using 14 molecular markers (Figures 1A and S1A). We then assigned each species to one of four niches based on previous descriptions and our own field observations: terrestrial, marginal aquatic with preference for solid substrates, marginal aquatic with preference for water surface, and open water surface [5, 11–14]. Our ancestral character state reconstruction suggests a gradual transition from the ancestral terrestrial to the derived open water habitat (Figure 1A), consistent with a previous reconstruction by Andersen [5, 11].

Correlation between Locomotion Speed and Niche Preference

To test whether adaptation to locomotion on water has favored high speed in the Gerromorpha, we first measured maximum speed in 16 species with clear differences in niche preference (Figure 1). We used body length per second (bl/s) as a speed unit to account for differences in body size found among species. The terrestrial outgroups and the Gerromorpha that specialize in marginal aquatic niches with preference for solid substrates [5, 11, 12, 14] produce maximum speeds ranging between 6 and 9 bl/s (Figure 1; Table S1). In contrast, Gerromorpha with preference for water surface (marginal aquatic and open water; [5, 11]) produce higher maximum speeds ranging between 27 and 113 bl/s (Figure 1; Table S1). Formal correlation tests, after taking phylogeny into account, revealed a strong correlation between maximum speed and niche preference such that species that occupy solid substrates are slow, whereas those that occupy the water surface are remarkably faster and their maximum speeds were invariably higher on water than on ground (Figures 1, S1B, and S4A; Table S1; Spearman correlation test with phylogenetic independent contrast [PIC] rho = 0.88, adjusted p = 6.5e−5; similar results were obtained with Pearson correlations) [15, 16]. The increase in locomotion speed cannot be explained merely by the plastic nature of animal behavior [2, 17, 18], as lineages with preference for solid

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

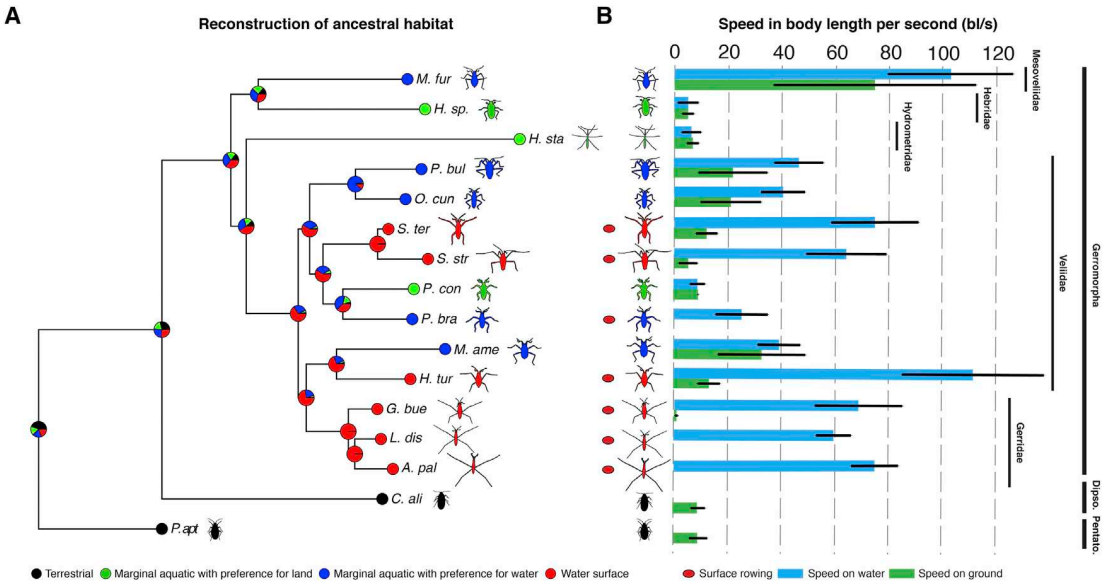


Figure 1. Ancestral Habitat Reconstruction and the Evolution of Maximum Speed in Relation to Habitat Preference

(A) The pies represent the probability of ancestral habitat.

(B) Speeds, in body lengths per second (bl/s), are given for both water (blue bars) and ground (green bars) locomotion, except for the terrestrial species, which cannot move on water.

Abbreviations are as follows: Pentato, Pentatomomorpha; Dipso, Dipsocoromorpha; *P.apt*, *Pyrrhocoris apterus*; *C.ali*, *Cryptostemma alienum*; *M.fur*, *Mesovelia furcata*; *H.sp*, *Hebrus* species; *H.sta*, *Hydrometra stagnorum*; *S.str*, *Stridulivelia strigosa*; *S.ter*, *Stridulivelia tersa*; *P. bra*, *Platyvelia brachialis*; *O.cun*, *Oiovelia cunucunumana*; *P.bul*, *Paravelia bullialata*; *P.con*, *Paravelia conata*; *H.tur*, *Husseyella turmalis*; *M.ame*, *Microvelia americana*; *L.dis*, *Limnoporus dissortis*; *G.bue*, *Gerris buenoi*; *A.pal*, *Aquarius paludum*.

Samples sizes in terms of number of videos are given in Table S1. Error bars represent standard deviation. See also Figures S1 and S4 and Tables S1 and S2.

substrates were not able to deliver high maximum speeds when tested on water. Altogether, these findings corroborate the hypothesis that water surface lifestyle has selected in favor of increased locomotion speed.

Correlation between Speed, Midleg Length, and Niche Preference

Because speed depends on leg length [19], we tested whether there is a correlation between leg length, speed, and niche preference in the Gerromorpha. The terrestrial outgroups and the species that prefer solid substrates [5, 11, 12, 14] exhibit low values of leg length to body length (Figures 2 and S2; Table S2). Conversely, species that prefer water surface, including the basally branching *Mesovelia* and the derived Veliidae and Gerridae [5, 11], exhibit higher values of leg length to body length (Figures 2 and S2; Table S2). Pairwise correlation tests detected a strong and significant correlation between the length of the midlegs and both speed and niche preference, but not between the forelegs or hindlegs and these two variables (Figures 2 and S4A; Spearman correlation tests with PIC; midleg and speed: rho = 0.82, adjusted p = 5.3e−4; midleg and habitat: rho = 0.88, adjusted p = 6.5e−5).

Increased leg length is known to contribute to generating faster movement through increasing the amplitude of strokes [20, 21]. When we analyzed how the legs act during locomotion on fluid compared to solid substrates, we failed to detect any difference in the amplitude of leg strokes between water and ground locomotion in species with short midlegs, preference for solid substrates, and that use the tripod gait (*Hebrus*) (Figures 3A and 3B; Table S3). However, in species with elongated midlegs, with preference for water surface, and that also use the tripod gait (*Microvelia* and *Mesovelia*; Movie S1), we detected a significant increase in the amplitude of midleg strokes when these species moved on water compared to when they moved on solid substrates (Figures 3C, 3D, 3F, and 3G; Table S3). Finally we tested *Gerris*, a species that employs rowing (Movie S1) and that generates movement through distortion of the water surface and generation and shedding of vortices [8, 22]. In this species, we found that the midlegs stretched in a straight shape and executed considerably large amplitudes of stroke, whereas the amplitudes of the forelegs and hindlegs were minimal if any (Figures 3E–3H; Table S3). Altogether, these results suggest that the preference for various water surface niches have favored increased speed through increasing the length and the amplitude of stroke of the midlegs.

The Tripod Gait Employs High Frequency of Leg Strokes on Water

Stroke frequency is another key biomechanical parameter that determines speed [21]. During locomotion, leg motion pattern includes two timescales representing a stance phase (leg pushing against substrate) and a swing phase (leg loses contact with substrate; [23, 24]). We therefore compared stroke frequency (number of steps per second) and pattern of leg motion (stance and swing phases separately) between 16 extant species (Figures 4A, 4B, and S3). Exclusively terrestrial species in addition to all

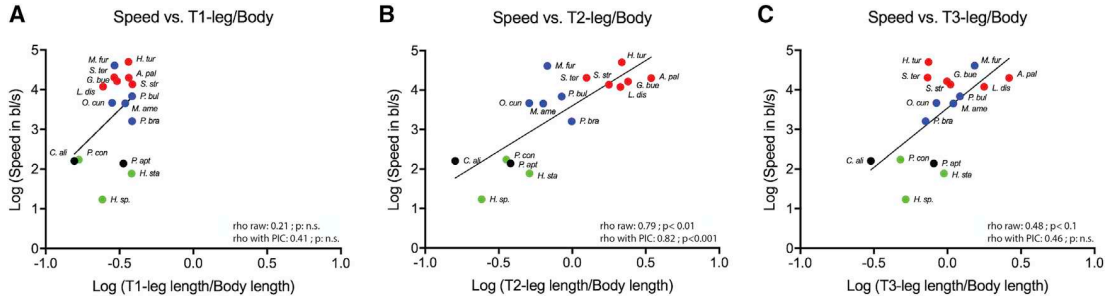


Figure 2. Correlation Tests between Leg Length, Speed, and Habitat Preference

The ratio of leg length by body length is used to take into account the differences in size across species.

(A and C) There is no significant correlation between foreleg length and speed (rho raw: 0.21, adjusted p: 4.9e−1, n.s.; rho PIC: 0.41, adjusted p: 1.8e−1, n.s.) (A) and between hindleg length and speed (rho raw: 0.48, adjusted p: 9.7e−2; rho PIC: 0.46, adjusted p: 1.5e−1, n.s.) (C) using respectively both Spearman correlation test on raw data and with phylogenetic correction.

(B) There is a significant positive correlation between midleg length and speed (rho raw: 0.79, adjusted p: 1.6e−3; rho PIC: 0.82, adjusted p: 5.3e−4) using respectively both Spearman correlation test on raw data and with phylogenetic correlation.

See also Figures S1, S2, and S4 and Tables S1 and S2.

Gerromorpha that occupy solid substrates move with a highly similar stroke frequency (ranging between 14 and 24 strokes per second [st/s]), and their pattern of motion is characterized by a longer stance phase (Figures 4A, 4B, and S3). Some of these species (*Hebrus* and *Hydrometra*) employed the same pattern of leg motion on water as the pattern they used on ground (Figures 4A, 4B, and S3).

The Gerromorpha that have preference for water surface habitats and that also use the tripod gait move both on water and on ground with a much higher stroke frequency (varying between 42 and 71 st/s; Figures 4A, 4B, and S3A). Their motion pattern on ground resembled that of species that prefer solid substrates, except for *Mesovelia furcata* (Figure S3B). Interestingly, when they moved on water, these species were able to change their motion pattern such that the stance and the swing phases were executed with the same duration (Figure S3C). In addition, we detected a strong correlation between stroke frequency, speed, and preference for water surface in the Gerromorpha that retain the ancestral tripod gait (Figures 4A and S4B; Spearman correlation tests with PIC; stroke frequency and speed: rho = 0.92, adjusted p = 3.6e−3; stroke frequency and habitat: rho = 0.83, adjusted p = 0.013). These findings are surprising because there is a trade-off between leg length and stroke frequency such that stroke frequency decreases with increasing leg length [25]. The Gerromorpha that move on water using the ancestral tripod gait show a significant increase in both leg length and stroke frequency, thus indicating that this trade-off was overcome during the transition to water surface-dwelling lifestyle.

In species that employ rowing as a mode of water surface locomotion (three Gerridae and three Veliidae), stroke frequency was dramatically low (varying between 6 and 16 st/s) (Figures 4B and S3A; Table S1) despite the high speed that they were able to generate (varying between 27 and 113 bl/s; Figures 1B and 4B; Table S1). In addition, the pattern of leg motion was reversed in most rowing species compared to species that specialize in solid substrates, such that the stance phase is now shorter than the swing phase (Figure S3C). Therefore, niche specialization across the Gerromorpha was accompanied by changes in the mode of locomotion, leg length, patterns of leg motion, and stroke frequency.

Rowing Gait Is Significantly More Efficient on the Water Surface Than Tripod Gait

Species that use the derived rowing gait show a substantially lower stroke frequency when compared to those using the ancestral tripod gait. Therefore, we tested whether the evolution of rowing could be associated with increased efficiency during water surface locomotion. First, we inferred (see Supplemental Experimental Procedures) the amount of energy spent in the two veliids *Stridulivelia strigosa* and *Paravelia bullialata*, which have a similar body mass (Table S1), generate comparable speeds, but differ in their mode of locomotion (Figure 1). *Stridulivelia strigosa* stroked 9 times to generate a speed of 64 body lengths (28 cm) in 1 s, whereas *Paravelia bullialata* stroked 61 times to generate a speed of 47 body lengths (18 cm) in 1 s (Figures 1B and 4D). We calculated that *Stridulivelia strigosa* spent 2.21e−5 mJ/mg/st (millijoule per milligram per stroke), and *Paravelia bullialata* spent 1.17e−5 mJ/mg/st (Figure 4C; Table S1). Therefore, by taking into account the number of strokes per unit of time, *Stridulivelia strigosa* (rowing) spent 2.9e−4 mJ/mg/s (millijoule per milligram per second), and *Paravelia bullialata* (tripod) spent 2.35e−2 mJ/mg/s, which is over 80-fold higher in the latter (Figure 4D; Table S1). When we extended this analysis to the entire sample of the Gerromorpha that transited to the derived water surface habitat, we found that the rowing species consistently spent less energy to generate movement than species that employ the ancestral tripod gait (Figure 4D; Table S1; Student's t test; p = 4.6e−4). Species using the ancestral tripod gait on water spend much of the time on aquatic plants and would only execute bursts of fast movement when crossing free water patches, presumably, to minimize the risk of capture by bottom-striking predators such as fish [5, 10]. Therefore, the maintenance of the tripod gait may have been advantageous, despite the high-energy demand, as it allows these animals to be versatile. Rowing species, however, spend much more time on the open water where they forage, mate, and interact with predators [5, 10]. This lifestyle may have increased demands on frequent,

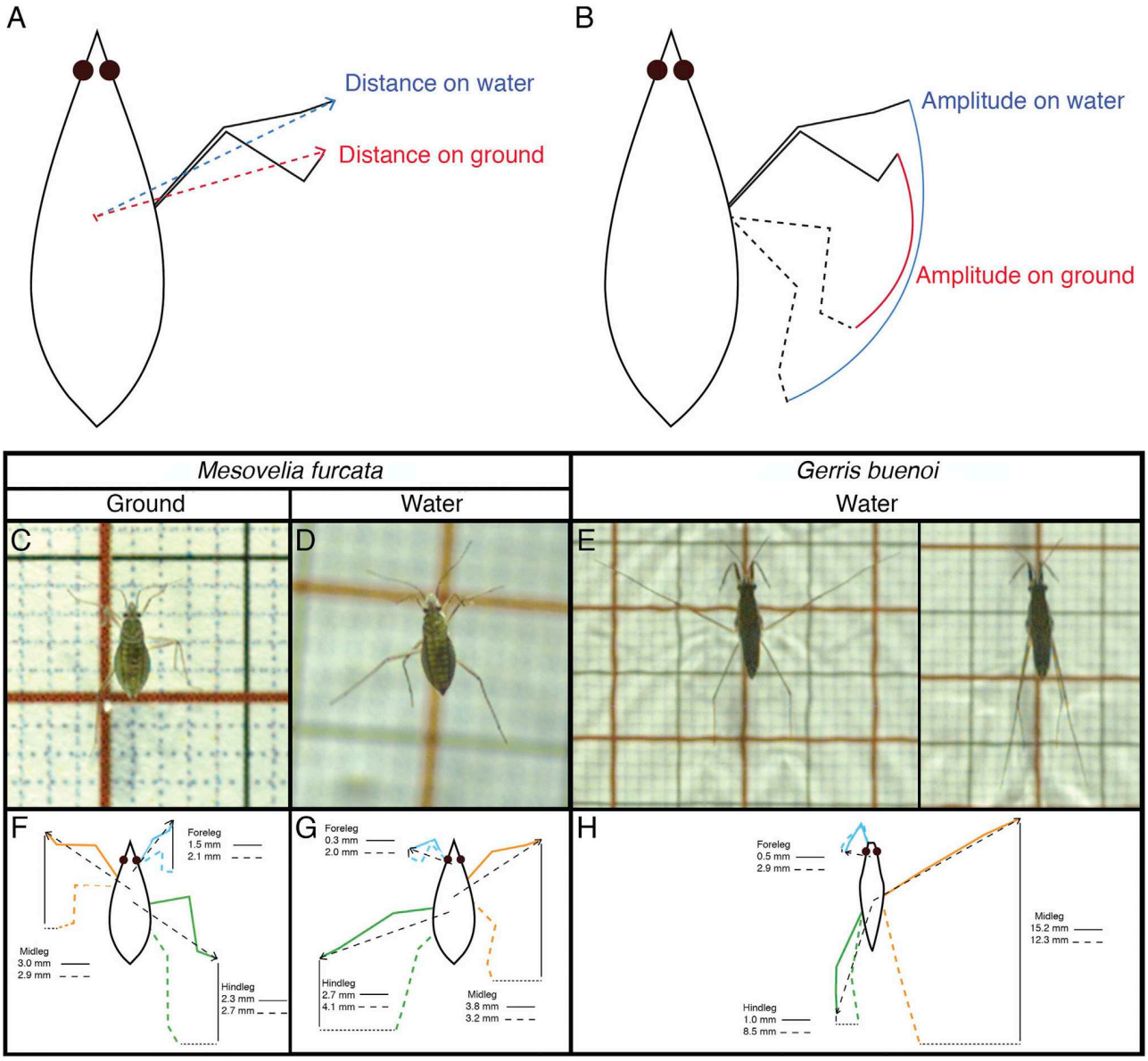


Figure 3. Analysis of Leg Movement Parameters between Ground and Water Surface Locomotion

(A and B) Representation of the measurements of leg deployment (A) and the amplitude of leg movements (B) during locomotion on ground and on water.

(C–H) Analysis of locomotion for *Mesovelia furcata* on ground (C and F) and on water (D and G) and for *Gerris buenoi* on water (E and H) with leg deployment (dashed line) and amplitude of movement (solid line).

Results are showed in Table S3. See also Figures S3 and S4, Table S3, and Movie S1.

fast, and energy-efficient locomotion. These results indicate that the derived mode of water surface locomotion through rowing gait, characteristic of derived species that specialize in open water, is more energy efficient when compared to the ancestral mode using the tripod gait.

Reduced Stroke Frequency Is Associated with Increased Body Size in Derived Gerromorpha

Another important trade-off in walking systems, including arthropods, exists between body size and stroke frequency, such that stroke frequency decreases with increasing body size [21, 26]. Because the evolution of rowing in derived lineages is associated with a dramatic decrease in stroke frequency, we hypothesized that this decrease may have removed the constraint on body size [27]. To test this hypothesis, we plotted body size by stroke frequency in a total of 25 species that exhibit large variation in these two traits (Figure 4E). We found that species with small body size (<5,000 μm) may exhibit either low or high stroke frequency (Figure 4E). Conversely, all species with larger body size (>5,000 μm) showed a significantly low stroke frequency (Figure 4E). In our dataset, we could not find any species with both high stroke frequency and large body size (Figure 4E). Litterbugs

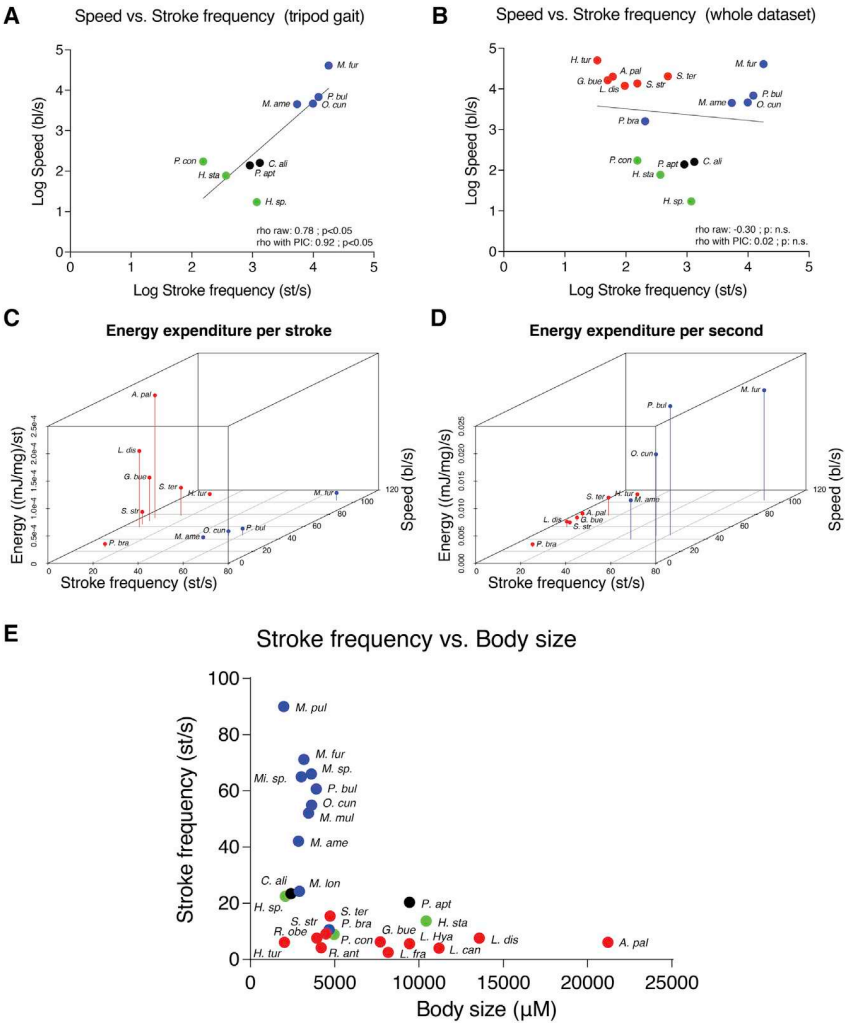


Figure 4. Correlation Tests between Stroke Frequency, Speed, and Habitat Preference with Associated Energy Expenditure and Relation with Body Size

(A) A significant positive correlation is observed between stroke frequency and speed in species employing the tripod gait using both Spearman correlation test on raw data (rho raw: 0.78, adjusted p: 3.8e−2) and with phylogenetic correlation (rho PIC: 0.92, adjusted p: 3.6e−3).

(B) There is no correlation between stroke frequency and speed when we add species employing the rowing gait to the dataset using both Spearman correlation test on raw data (rho raw: −0.30, adjusted p: 3.3e−1, n.s.) and with phylogenetic correlation (rho PIC: 0.02, adjusted p: 0.97, n.s.). Sample sizes in term of number of videos are given in Table S1.

(C) Species using tripod gait (blue dots) spend less energy per stroke (p: 0.08, n.s.; Student's t test).

(D) However, when normalized to the number of stroke employed per second, species using the rowing gait spend less energy (Student's t test; p: 4.5e−5). Sample sizes and energy expenditure are given in the Supplemental Experimental Procedures and Table S1, respectively.

(E) High stroke frequency can be observed only in animals with small body size.

See also Figures 3, S3, and S4 and Tables S1 and S2.

(Dipsocoromorpha; Figure 1A; [28]) are characteristically small [29], and similarly the ancestral state of body size in the Gerromorpha has been undoubtedly small to very small [5]. It is possible that because of their small body size, the ancestors of the Gerromorpha were free from the constraint imposed on larger bugs, such that stroke frequency increased and facilitated the transition to water surface life. Subsequently, the evolution of rowing combined with the significant increase in the length of the driving midlegs [10, 30–33], and most likely changes in the associated innervation and musculature [22, 34, 35], may have contributed to reducing stroke frequency while maintaining high speeds. These structural and behavioral changes may have contributed to the evolution of larger bodies including the largest semi-aquatic bug known, the Gerrid *Gigantometra gigas*, whose body length is over 3 cm and leg span over 25 cm [36].

Conclusions

Understanding how the evolution of behavior and morphology can be associated with niche expansion and species diversification is a major challenge in evolutionary biology [1, 3, 17, 18]. We have shown that lineages that remained in the ancestral habitat, composed of solid substrates, produce low maximum speeds and that lineages that specialize in open water surface niches generate significantly higher maximum speeds. This increase in locomotion speed is associated with the evolution of increased midleg length, changes in leg motion patterns, and increased frequency of leg strokes. The subsequent evolution of rowing, as a novel mode of locomotion on the water, removed the requirement for high stroke frequency, reduced energy expenditure, and enabled derived lineages to specialize in open water zones. Finally, we have shown how the evolution of rowing, by removing the requirement for high stroke frequency, may have led to the evolution of larger Gerromorpha. By uncovering the link between the ecology of the semi-aquatic bugs and the variety of phenotypes they exhibit, this work draws a comprehensive picture of how a combination of behavioral and structural changes can impact the evolutionary trajectory of groups of animals and can be associated with niche expansion and lineage diversification.

EXPERIMENTAL PROCEDURES

Insect Sampling and Culture

Extant specimens were collected during fieldwork in the locations indicated in Table S2. All species were kept in water tanks at 25°C, 50% humidity, and 14 hr of daylight and fed on live crickets.

Phylogenetic Reconstruction

Phylogeny reconstruction was conducted using 14 molecular markers with the software Geneious version 7.1.9 using plugins MrBayes version 3.2.6 [37] (one million generations; 25% burnin) and PhyML version 3.0 [38], using GTR model

with 100 bootstraps. More details can be found in the Supplemental Experimental Procedures.

Habitat Classification

To enable reconstruction, we consolidated Andersen's habitat classes [5, 11] into the following four: terrestrial, marginal aquatic with preference for solid substrates, marginal aquatic with preference for water surface, and open water surface. Each species was assigned to one of these four classes based on previous descriptions [5, 11–14] and on the environment where we caught them.

Video Acquisition, Quantification of Speed, and Stroke Frequency

A set of adult individuals for each species were filmed at 2,000 frames per second, both on water surface and on a solid substrate with a grid paper in the background as a calibration scale. To calculate speed, a mean value was extracted from a defined interval plateau phase from velocity curve along each video. This interval represents the maximum speed during the run of the individual. Stroke frequency was determined as the number of steps performed by the individuals during a given locomotion duration and converted into number of strokes per second. Details of sample sizes and calculations can be found in the Supplemental Experimental Procedures.

Reconstruction of Ancestral Character State

Reconstruction of ancestral character states was performed in Rstudio version 0.99.486 using a maximum likelihood method adapted to discrete characters (ace, package ape; [16]) and represented using phytools [39]. Details about methods of reconstruction and scripts used can be found in the Supplemental Experimental Procedures.

Inference of Energy Consumption

We inferred the amount of energy spent per stroke based on the procedure from [40]. Kinetic energy (E_k in joules) used during a stroke is determined using the following expression: $E_k = 0.5mv^2$, where m is the mass of the insect in grams, and v is the velocity generated during one stroke in meters per second. The analyses were performed using speed data extracted from the high-speed movies. Calculations and sample sizes are detailed in the Supplemental Experimental Procedures.

Statistical Analyses

Details and R script used for statistical analyses are provided in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

Accession numbers for the sequences and alignment of concatenated sequences reported in this paper are available in Dryad Digital Repository (http://dx.doi.org/10.5061/dryad.134c4).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, three tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.09.061.

AUTHOR CONTRIBUTIONS

Conceptualization, A.J.J.C. and A.K.; Methodology, A.J.J.C. and A.K.; Investigation, A.J.J.C.; Formal Analysis, A.J.J.C., M.E.S., and M.S.; Resources, A.J.J.C., M.E.S., D.A., F.F.F.M., and A.K.; Data Curation, D.A. and M.E.S.; Writing – Original Draft, A.J.J.C. and A.K.; Writing – Review & Editing, A.J.J.C., M.E.S., M.S., D.A., F.F.F.M., and A.K.; Visualization, A.J.J.C. and A.K.; Supervision, A.K.; Funding Acquisition, A.K.

ACKNOWLEDGMENTS

We thank W. Salzburger, E. Abouheif, F. Bonneton, S. Viala, A. Decaras, and W. Toubiana for comments and discussion; H. Labrique and J.C. Streito for help with the Dipsocoromorpha; and M. Burrows for advice on inferring energy expenditure. This work was funded by ERC-CoG #616346, ATIP-Avenir, and CNPq-PVE #400751/2014-3 to A.K. and a postdoctoral fellowship by the Swiss National Science Foundation to M.E.S. Specimens from Brazil were collected under SISBIO permit #43105-1.

Received: July 13, 2016
Revised: September 23, 2016
Accepted: September 28, 2016
Published: December 8, 2016

REFERENCES

1. Losos, J.B. (2009). Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles (University of California Press).
2. Mayr, E. (1963). Animal Species and Evolution (Belknap Press of Harvard University Press).
3. Schluter, D. (2000). The Ecology of Adaptive Radiation (Oxford University Press).
4. Yoder, J.B., Clancey, E., Des Roches, S., Eastman, J.M., Gentry, L., Godsoe, W., Hagey, T.J., Jochimsen, D., Oswald, B.P., Robertson, J., et al. (2010). Ecological opportunity and the origin of adaptive radiations. J. Evol. Biol. *23*, 1581–1596.
5. Andersen, N.M. (1982). The Semiaquatic Bugs (Hemiptera: Gerromorpha). Entomonograph, *Volume 3* (Scandinavian Science Press LTD).
6. Ikawa, T., Okabe, H., and Cheng, L.N. (2012). Skaters of the seas - comparative ecology of nearshore and pelagic *Halobates* species (Hemiptera: Gerridae), with special reference to Japanese species. Mar. Biol. Res. *8*, 915–936.
7. Denny, M.W. (1993). Air and Water: The Biology and Physics of Life's Media (Princeton University Press).
8. Hu, D.L., and Bush, J.W.M. (2010). The hydrodynamics of water-walking arthropods. J. Fluid Mech. *644*, 5–33.
9. Suter, R., Rosenberg, O., Loeb, S., and Long, H. (1997). Locomotion on the water surface: propulsive mechanisms of the fisher spider. J. Exp. Biol. *200*, 2523–2538.
10. Armisén, D., Refki, P.N., Crumière, A.J., Viala, S., Toubiana, W., and Khila, A. (2015). Predator strike shapes antipredator phenotype through new genetic interactions in water striders. Nat. Commun. *6*, 8153.
11. Andersen, N.M. (1979). Phylogenetic inference as applied to the study of evolutionary diversification of semiaquatic bugs (Hemiptera, Gerromorpha). Syst. Zool. *28*, 554–578.
12. Dias-Silva, K., Moreira, F.F.F., Giehl, N.F.D., Nóbrega, C.C., and Cabette, H.S.R. (2013). Gerromorpha (Hemiptera: Heteroptera) of eastern Mato Grosso State, Brazil: checklist, new records, and species distribution modeling. Zootaxa *3736*, 201–235.
13. Heiss, E., and Pericart, J. (2007). Faune n° 91 – Hémiptères Aradidae, Piesmatidae et Dipsocoromorphes (Fédération Française des Sociétés de Sciences naturelles).
14. Schuh, R.T., and Slater, J.A. (1995). True Bugs of the World (Hemiptera: Heteroptera): Classification and Natural History (Comstock Publishing Associates).
15. Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., and Webb, C.O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics *26*, 1463–1464.
16. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. Bioinformatics *20*, 289–290.
17. Duckworth, R.A. (2009). The role of behavior in evolution: a search for mechanism. Evol. Ecol. *23*, 513–531.
18. West-Eberhard, M.J. (2003). Developmental Plasticity and Evolution (Oxford University Press).
19. Zaaf, A., Van Damme, R., Herrel, A., and Aerts, P. (2001). Spatio-temporal gait characteristics of level and vertical locomotion in a ground-dwelling and a climbing gecko. J. Exp. Biol. *204*, 1233–1246.

20. Baudouin, A., and Hawkins, D. (2002). A biomechanical review of factors affecting rowing performance. Br. J. Sports Med. *36*, 396–402, discussion 402.
21. Wu, G.C., Wright, J.C., Whitaker, D.L., and Ahn, A.N. (2010). Kinematic evidence for superfast locomotory muscle in two species of teneriffiid mites. J. Exp. Biol. *213*, 2551–2556.
22. Andersen, N.M. (1976). A comparative study of locomotion on the water surface in semiaquatic bugs (Insecta, Hemiptera, Gerromorpha). Vidensk. Meddr dansk naturh. Foren. 337–396.
23. Alexander, R.M. (2003). Principles of Animal Locomotion (Princeton University Press).
24. Mendes, C.S., Bartos, I., Akay, T., Márka, S., and Mann, R.S. (2013). Quantification of gait parameters in freely walking wild type and sensory deprived *Drosophila melanogaster*. eLife *2*, e00231.
25. Vanhooydonck, B., Van Damme, R., and Aerts, P. (2001). Speed and stamina trade-off in lacertid lizards. Evolution *55*, 1040–1048.
26. Heglund, N.C., Taylor, C.R., and McMahon, T.A. (1974). Scaling stride frequency and gait to animal size: mice to horses. Science *186*, 1112–1113.
27. Arnold, S.J. (1992). Constraints on phenotypic evolution. Am. Nat. *140* (*Suppl 1*), S85–S107.
28. Li, M., Tian, Y., Zhao, Y., and Bu, W. (2012). Higher level phylogeny and the first divergence time estimation of Heteroptera (Insecta: Hemiptera) based on multiple genes. PLoS ONE *7*, e32152.
29. Weirauch, C., and Schuh, R.T. (2011). Systematics and evolution of Heteroptera: 25 years of progress. Annu. Rev. Entomol. *56*, 487–510.
30. Khila, A., Abouheif, E., and Rowe, L. (2009). Evolution of a novel appendage ground plan in water striders is driven by changes in the Hox gene Ultrabithorax. PLoS Genet. *5*, e1000583.
31. Khila, A., Abouheif, E., and Rowe, L. (2014). Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. Evolution *68*, 2159–2170.
32. Refki, P.N., Armisén, D., Crumière, A.J.J., Viala, S., and Khila, A. (2014). Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. Dev. Biol. *392*, 441–453.
33. Refki, P.N., and Khila, A. (2015). Key patterning genes contribute to leg elongation in water striders. Evodevo *6*, 14.
34. Burrows, M. (1992). Local circuits for the control of leg movements in an insect. Trends Neurosci. *15*, 226–232.
35. Gleeson, T.T., and Harrison, J.M. (1988). Muscle composition and its relation to sprint running in the lizard *Dipsosaurus dorsalis*. Am. J. Physiol. *255*, R470–R477.
36. Tseng, M., and Rowe, L. (1999). Sexual dimorphism and allometry in the giant water strider *Gigantometra gigas*. Can. J. Zool. *77*, 923–929.
37. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.
38. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. *52*, 696–704.
39. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. *3*, 217–223.
40. Burrows, M., and Dorosenko, M. (2014). Jumping mechanisms in lacewings (Neuroptera, Chrysopidae and Hemerobiidae). J. Exp. Biol. *217*, 4252–4261.

Current Biology, Volume *26*

Supplemental Information

Diversity in Morphology and Locomotory Behavior Is Associated with Niche Expansion in the Semi-aquatic Bugs

Antonin J.J. Crumière, M. Emilia Santos, Marie Sémon, David Armisén, Felipe F.F. Moreira, and Abderrahman Khila

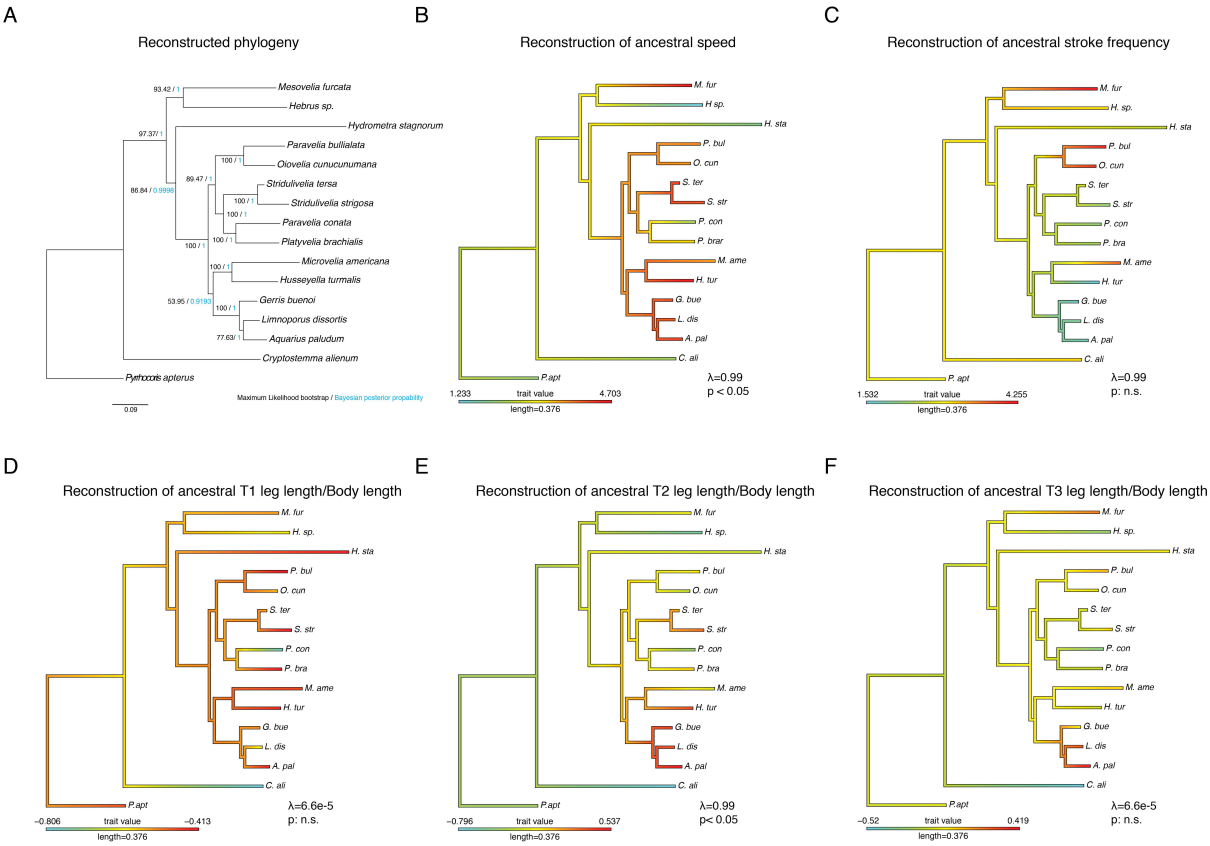


Figure S1: Phylogeny and ancestral character state reconstruction. Related to Figure 1, Figure 2 and Figure S4, Table S1, Table S2.

(A) Phylogenetic relationships between species sampled for this study. Phylogenetic construction of our sample using both Maximum Likelihood (bootstrap in black) and Bayesian methods (posterior probability in blue). (B) Phylogenetic reconstruction of ancestral state of speed of locomotion in the Gerromorpha and terrestrial outgroups. Increased speed is a derived state and correlates with preference for water surface habitat. (C) Phylogenetic reconstruction of ancestral state of stroke frequencies. Increased stroke frequency is a derived state and correlates with preference for water surface habitat for species using the tripod gait but not for species using the rowing gait. Phylogenetic reconstruction of ancestral state of the ratio leg length by body length for the foreleg (D), the midleg (E) and the hindleg (F). The elongated midleg is a derived state and correlates with preference for water surface habitat. Phylogenetic signals λ and associated significance are indicated for each ancestral reconstruction.

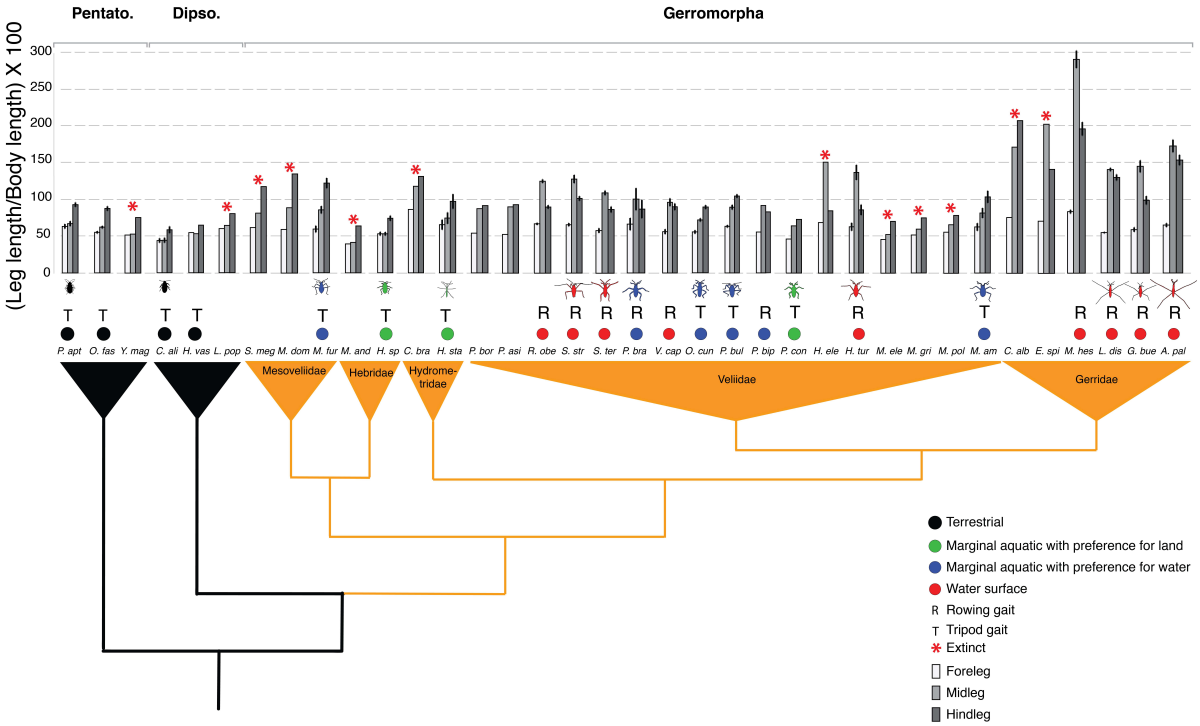


Figure S2: Ratio of leg length to body length across heteroptera including fossil record. Related to Figure 2, Table S2.

Phylogenetic relationships and ratios of leg length by body length across a sample of extant and extinct Heteropteran species. Extant Gerromorpha have generally longer legs than terrestrial relatives. *O.fas*, *Oncopeltus fasciatus*; *Y.mag*, *Yuripopovina magnifica* [S1]; *H.vas*, *Hypsipteryx vasarhelyii* [S2]; *L.pop*, *Libanohypselosoma popovi* [S3, S4]; *M.dom, Mesovelia dominicana* [S5]; *M.and*, *Miohebrus anderseni* [S5]; *C.bra*, *Cretaceometra brasiliensis* [S6]; *P.bor*, *Perittopus borneensis* [S7]; *P.asi*, *Perittopus asiaticus* [S7]; *R.obe, Rhagovelia obesa*; *V.cap*, *Velia caprai*; *H.ele, Halovelia electrodominica* [S8]; *M.ele, Microvelia electra* [S9]; *M.gri*, *Microvelia grimaldii* [S9]; *M.pol*, *Microvelia polhemi* [S10]; *C.alb, Cretogerris albianus* [S11]; *E.spi*, *Electrobates spinipes* [S12]; *M.hes, Metrobates hesperius*. Red asterisks indicate extinct species. Error bars represent the standard deviation.

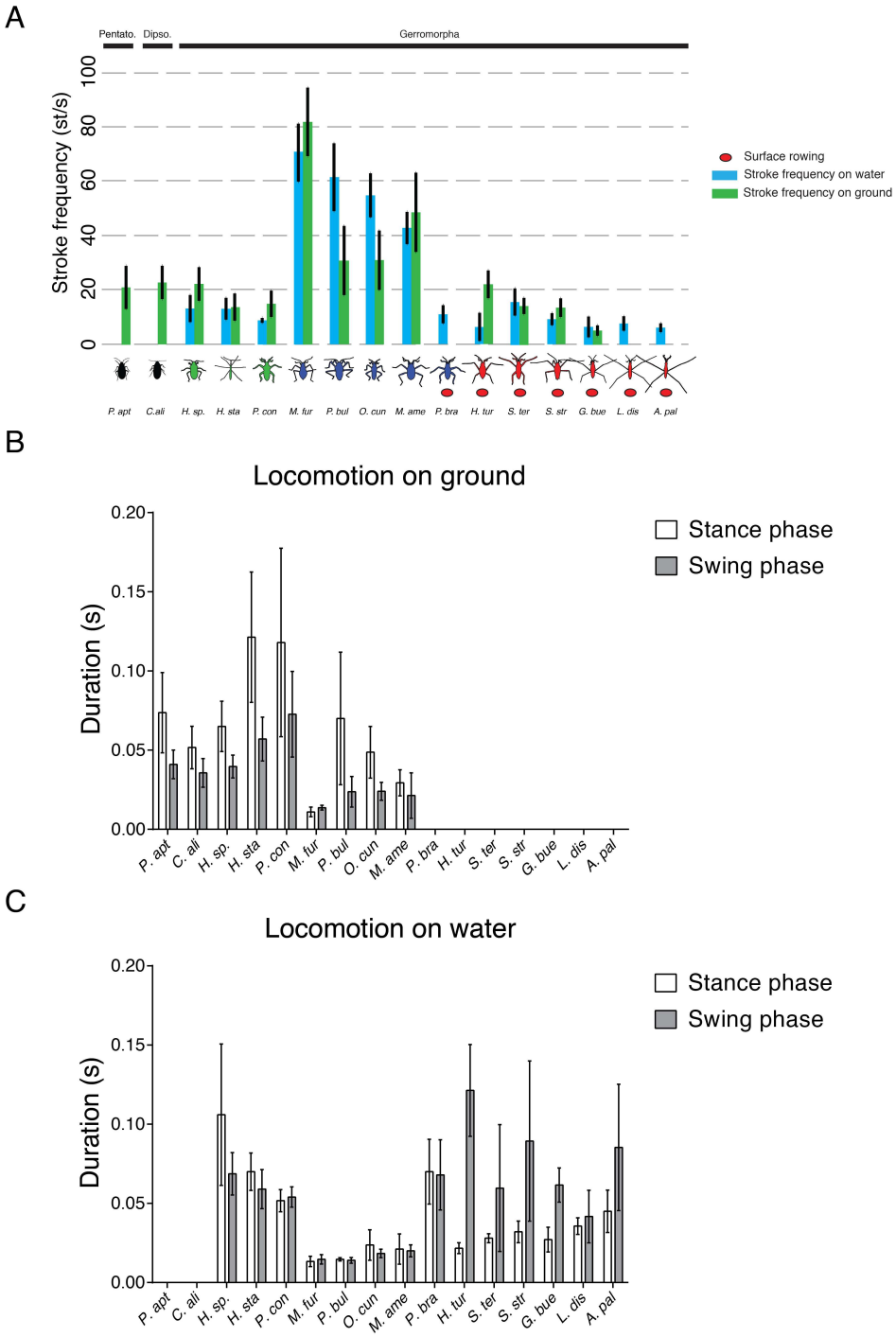


Figure S3: Stroke frequency and leg motion pattern across the Gerromorpha and terrestrial outgroups. Related to Figure 3, Figure 4 and Figure S4, Table S1, Table S3.
(A) Quantification of stroke frequencies during locomotion on ground (green bars) and on water (blue bars). Low stroke frequencies for species using the tripod gait are associated with life on ground. High stroke frequencies are associated with life on water surface for species using the tripod gait. Low stroke frequencies are associated with life on water surface only for species using the derived rowing gait. (B) Analysis of leg motion pattern on ground showing that the stance phase is longer than the swing phase for species using the tripod gait. (C) The same analysis on water showing that species with preference for water and using the tripod gait have similar stance and swing phases whereas species with preference for ground still have a longer stance phase than swing phase. Derived species using the rowing gait have a shorter stance phase compared to swing phase. This pattern is associated with life on water surface. Error bars represent standard deviation.

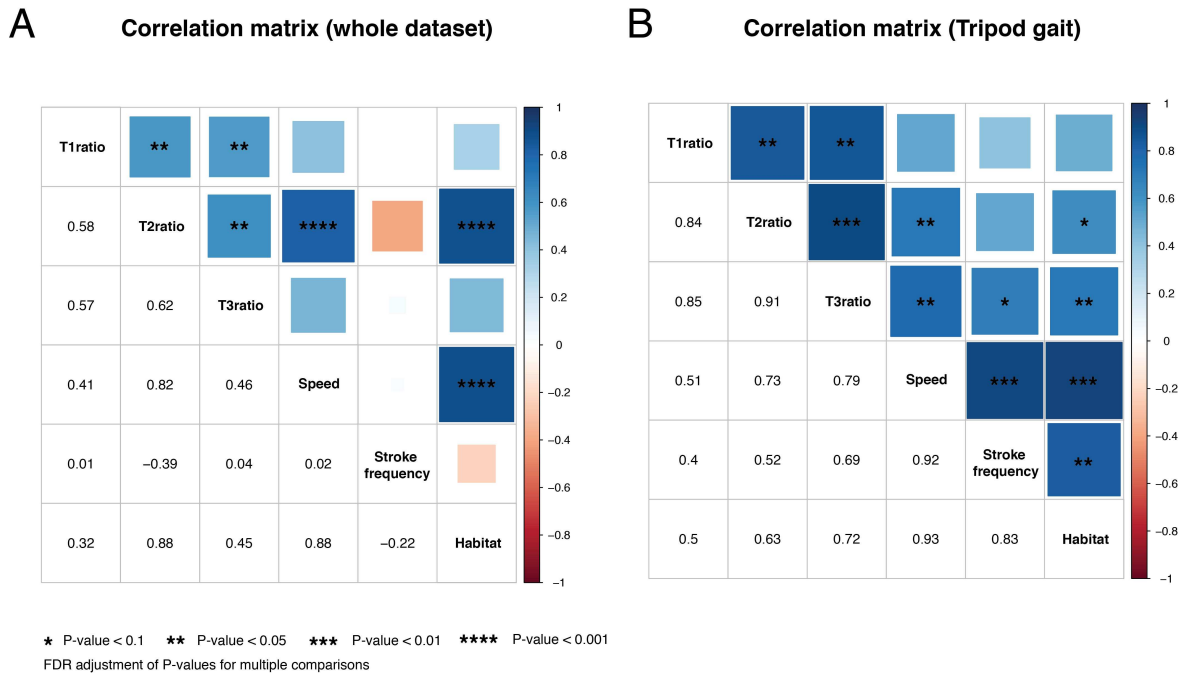


Figure S4: Matrix of correlation for the different correlation tests performed. Related to Figure 1, Figure 2, Figure 3, Figure 4 and Figure S1, Figure S3 and Table S1.

(A) Matrix of correlation for the whole dataset (tripod gait and rowing gait) and for (B) the dataset of species using only the tripod gait. Squares and number indicates the rho of the Sperman correlation test conducted with phylogenetic correction (PIC). FDR adjustment of P-values for multiple comparisons is applied. Stars indicate the P-values.

Species	Speed on ground (cm/s)	Speed on ground (bl/s)	Speed on water (cm/s)	Speed on water (bl/s)	Stroke frequency on ground (st/s)	Stroke frequency on water (st/s)	Mean weight (mg)	Energy per stroke ((mJ/mg)/st)	P-Value	Energy per second ((mJ/mg)/s)	P-Value
Pyrrhocoris apterus	8.6±3.2 (N=14)	9.1±3.4 (N=14)	NA	NA	20.4±7.7 (N=14)	NA	NA	NA	NA	NA	NA
Cryptostemma alienum	2.3±0.7 (N=14)	9.4±2.8 (N=14)	NA	NA	23.5±6.3 (N=14)	NA	NA	NA	NA	NA	NA
***Hebrus* sp.**	1.2±0.9 (N=10)	6.0±4.4 (N=10)	1.3±0.6 (N=10)	6.4±2.7 (N=10)	22.6±6.2 (N=10)	13.4±5.0 (N=10)	NA	NA	NA	NA	NA
Hydrometra stagnorum	8.0±4.4 (N=10)	7.6±4.2 (N=10)	8.1±2.6 (N=10)	7.8±2.5 (N=10)	13.7±4.9 (N=10)	13.1±3.9 (N=10)	NA	NA	NA	NA	NA
Paravelia conata	4.7±1.5 (N=10)	8.9±2.8 (N=10)	5.0±0.04 (N=3)	9.4±0.9 (N=3)	15.1±4.8 (N=10)	8.9±0.9 (N=6)	NA	NA	NA	NA	NA
Mesovelia furcata	26.1±9.3 (N=10)	74.4±37.6 (N=10)	33.1±7.5 (N=12)	103.6±23.6 (N=12)	82.2±12.7 (N=10)	71.2±10.7 (N=12)	1.16	1.3318e-05	0.08 n.s.	2.0058782e-2	4.6e-4 ***
Oiovelia cunucunumana	8.4±4.5 (N=10)	20.6±11.0 (N=10)	16.4±3.3 (N=8)	40.2±8.1 (N=8)	31.1±10.9 (N=10)	55.0±8.0 (N=10)	2.92	8.06773e-06		1.4839771e-2	
Paravelia bullialata	8.9±5.1 (N=10)	23.0±13.3 (N=10)	18.3±3.6 (N=10)	47.4±9.2 (N=10)	30.4±12.6 (N=10)	60.7±12.3 (N=10)	2.925	1.17074e-05		2.3499226e-2	
Microvelia americana	9.5±4.7 (N=9)	33.5±16.5 (N=9)	11.2±2.3 (N=8)	39.5±8.0 (N=8)	48.2±14.4 (N=10)	42.4±5.9 (N=9)	1.53	3.6114e-06		7.093716e-3	
Stridulivelia strigosa	3.8±1.2 (N=10)	6.8±4.5 (N=10)	28.1±6.5 (N=10)	64.4±15.0 (N=10)	13.3±3.4 (N=10)	9.1±2.2 (N=10)	2.733	2.21034e-05		2.92093e-4	
Stridulivelia tersa	5.3±1.7 (N=10)	11.0±3.5 (N=10)	37.1±8.1 (N=11)	76.3±16.6 (N=11)	14.1±2.9 (N=10)	15.5±4.9 (N=22)	3.65	5.06027e-05		3.276652e-3	
Platyvelia brachialis	NA	NA	12.4±4.7 (N=18)	26.5±10.1 (N=18)	NA	10.6±3.2 (N=35)	6.233	5.98386e-06		5.43207e-4	
Husseyella turmalis	2.8±0.9 (N=9)	13.8±4.3 (N=9)	23.2±5.5 (N=10)	113.0±26.8 (N=10)	21.2±4.9 (N=9)	6.1±5.0 (N=10)	0.5	8.38042e-07		3.00e-05	
Limnopor	NA	NA	80.6±8	59.3±6	NA	7.6±2.	31.1	1.38796		9.90729	

us dissortis			.8 (N=8)	.5 (N=8)		6 (N=8)	83	e-04		e-4	
Gerris buenoi	1.3±0.7 (N=8)	1.7±0.9 (N=8)	53.7±12.7 (N=8)	69.8±16.6 (N=8)	5.0±1.8 (N=8)	6.3±3.7 (N=8)	12.2	7.87642 e-05		5.82314 e-4	
Aquarius paludum	NA	NA	104.1±12.2 (N=8)	74.4±8.7 (N=8)	NA	6.1±1.6 (N=8)	34.667	2.23603 e-04		8.10206 e-4	

Table S1: Data of locomotion characteristics across the sample of species tested. Related to Figure 1, Figure 2, Figure 4, Figure S1, Figure S3. Mean values for speed in centimeters per second (cm/s) and body length per second (bl/s) and stroke frequency in number of strokes per second (st/s), on ground and water for species analyzed with sample size (N). ± Indicates the standard deviation. NA: not available. Energy expenditure per stroke and per second for each species with the associated mean weight. Highlighted in clear grey are species using the tripod gait and in dark grey are species using the rowing gait. Student t-tests are performed to compare the mean energy expenditure per stroke and per second between the two ways of locomotion.

Infraorder	Species	Habitat	T1-leg (µM)	T2-leg (µM)	T3-leg (µM)	Body (µM)	%T1-leg/Body	%T2-leg/Body	%T3-leg/Body	N	Location Or Reference
Pentatomomorpha	*Pyrrhocoris apterus*	Land	5869±353	6206 ±429	8596±508	9435 ±649	62±3	66±5	91±3	10	Lyon, France
Pentatomomorpha	*Oncopeltus fasciatus*	Land	6776±250	7628 ±273	10730± 500	12453±426	55±1.6	61±1.8	87±3	10	Samples generously provided by Jeremy Lynch.
Pentatomomorpha	*Yuripopovina magnifica*		2530	2590	3690	4840	52	54	76	1*	[S1]
Dipsocoromorpha	*Cryptostemma alienum*	Land (can stand on water)	1078±61	1088 ±38	1436±104	2414 ±146	44±3	45±3.4	60±4.7	6	Rivière de l'Eyrieux, St Martin de Valamas, France
Dipsocoromorpha	*Hypsipteryx vasarhelyii*	Land	1088	1060	1292	1945	56	54	66	1	[S2]
Dipsocoromorpha	*Libanohypselosoma popovi*	--	560	601	749	906	62	66	83	1*	[S3]
Gerromorpha	*Sinovelia mega*	--	2780	3670	5300	4460	62	82	119	1*	[S4]
Gerromorpha	*Mesovelia dominicana*	--	1010	1510	2290	1700	59	89	135	1*	[S5]
Gerromorpha	*Mesovelia furcata*	Land/ Water (prefer water)	1859±94	2680 ±173	3822±299	3180 ±295	59±4.5	85±5	120±6.8	10	Plan d'eau de Saloniques, Vilette d'Anthon, France
Gerromorpha	*Miohebrus anderseni*		1310	1380	2130	3220	41	43	066	1*	[S5]
Gerromorpha	*Hebrus* sp.	Land/Water (prefer land)	1124±38	1124 ±42	1568±53	2081 ±47	54±2.5	54±2.5	75±3.6	9	Plan d'eau de Saloniques, Vilette d'Anthon, France
Gerromorpha	*Cretaceometra brasiliensis*		9600	13100	14600	11300	85	116	129	1*	[S6]
Gerromorpha	*Hydrometra stagnorum*	Land /water (prefer land)	6801±400	7725 ±398	10089± 567	10425±1538	66±6.8	75±7.8	98±9.9	10	Lac de Miribel, Lyon, France
Gerromorpha	*Perittopus borneensis*	Land/ Water	1511	2440	2559	2700	56	90	95	1	[S7]
Gerromorpha	*Perittopus asiaticus*	Land/ Water	1562	2665	2753	2900	54	92	95	1	[S7]
Gerromorpha	*Rhagovelia obesa*	Water	2660±76	4959 ±119	3569±124	3939 ±68	68±1.6	126±2.5	91±2.4	10	Rivière du Nord, Montréal, Québec, Canada
Gerromorpha	*Stridulivelia strigosa*	Water	2961±110	5742 ±133	4564±117	4474 ±189	66±2.2	128±5.6	102±3.1	8	French Guiana, N 04.31572°, W -052.15396°
Gerromorpha	*Stridulivelia tersa*	Water	2767±117	5211 ±158	4137±231	4729 ±119	59±3.2	110±3.1	87±3.9	10	Brazil, N -19,07218°, W-39,79617°
Gerromorpha	*Platyvelia brachialis*	Water	3087±165	4660 ±351	4037±346	4678 ±338	66±8.3	100±15	87±12.3	3	Brazil N -18,37143°, W-40,1416
Gerromorpha	*Velia caprai*	Water	4040±142	6926 ±171	6484±222	7193 ±378	56±3.3	96±5.3	90±4.5	6	Lac de Miribel, Lyon, France
Gerromorpha	*Oiovelia cunucunumana*	Land/ Water (prefer water)	2090±48	2704 ±83	3363±146	3627 ±171	58±2.4	75±2.3	93±2.8	10	Brazil, N -20,25294°, W-44,91584°
Gerromorpha	*Paravelia bullialata*	Land/ Water (prefer water)	2579±85	3637 ±124	4252±89	3905 ±125	66±1.8	93±3.4	109±2.5	5	French Guiana, N 04.29769°, W -052.14927°

Gerromorpha	*Paravelia bipunctata*	Land/ Water	2983	4911	4447	5348	56	92	83	1	Brazil N -20,60034°, W-45,81498
Gerromorpha	*Paravelia conata*	Land/ Water (prefer land)	2282	3167	3602	4960	46	64	73	1	Brazil N - 18.367815º , W - 40.139877º
Gerromorpha	*Halovelia electrodominica*	Water (sea)	1030	2260	1270	1500	69	151	85	1*	[S8]
Gerromorpha	*Husseyella turmalis*	Water	1296±78	2820 ±193	1769±98	2023 ±263	65±5.5	140±10.5	88±7.3	6	French Guiana, N 04.34235°, W - 052.10730°
Gerromorpha	*Microvelia electra*	--	620	710	950	1330	47	53	71	1*	[S9]
Gerromorpha	*Microvelia grimaldii*	--	910	1050	1320	1750	52	60	75	1*	[S9]
Gerromorpha	*Microvelia polhemi*	--	1060	1250	1490	1850	57	68	81	1*	[S10]
Gerromorpha	*Microvelia americana*	Land/ Water (prefer water)	1798±76	2344 ±11	2975±139	2859 ±178	68±5.2	88±6.8	112±8.2	10	Rivière du Nord, Montréal, Québec, Canada
Gerromorpha	*Cretogerris albianus*	--	1390	3150	3810	1840	76	171	207	1*	[S11]
Gerromorpha	*Electrobates spinipes*	--	1820	5210	3630	2580	71	202	141	1*	[S12]
Gerromorpha	*Metrobates hesperius*	Water	3828±80	13315±397	8990±298	4596 ±154	83±2.5	290±11.8	196±9	10	Rivière du Nord, Montréal, Québec, Canada
Gerromorpha	*Limnoporus dissortis*	Water	7371±332	18843±844	17439± 1172	13580±563	54±0.9	139±2.3	128±4	10	Rivière de l'Acadie, Montréal, Québec, Canada
Gerromorpha	*Gerris buenoi*	Water	4584±140	11247±587	7679±417	7697 ±428	60±3	146±8.2	100±5.7	6	Toronto, Ontario, Canada
Gerromorpha	*Aquarius paludum*	Water	9003±721	23840±1821	21206± 1793	13986±1621	65±2.6	171±8.5	152±6.9	8	Plan d'eau de Saloniques, Vilette d'Anthon, France

Table S2: Mean values for leg length and body length with corresponding ratios. Related to Figure 1, Figure 2, Figure 4, Figure S1, Figure S2. Habitat preference, number of individuals and collecting locations or references for each extant or fossil (*) species used in this study are provided. ± indicates the standard deviation. These data are plotted in Figure S2.

	Distance body to leg tip (in mm)						Amplitude of movement (in mm)					
	T1-leg		T2-leg		T3-leg		T1-leg		T2-leg		T3-leg	
	Ground	Water	Ground	Water	Ground	Water	Ground	Water	Ground	Water	Ground	Water
Hebrus sp.	1.2 ± 0.1	1.1 ± 0.1	1.2 ± 0.2	1.2 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	0.9 ± 0.2	0.7 ± 0.1	1.0 ± 0.2	0.9 ± 0.2	0.8 ± 0.2	0.7 ± 0.2
	p: 0.01 *		p: 0.63 n.s.		p: 0.49 n.s.		p: 0.03 *		p: 0.40 n.s.		p: 0.08 n.s.	
Mesovelia furcata	2.1 ± 0.1	2.0 ± 0.1	2.9 ± 0.2	3.2 ± 0.2	2.7 ± 0.3	4.1 ± 0.2	1.5 ± 0.2	0.3 ± 0.3	3.0 ± 0.5	3.8 ± 0.4	2.3 ± 0.2	2.7 ± 0.5
	p: 0.03 *		p: 6.8 e-4 ***		p: 2.2 e-16 ***		p: 4.89 e-16 ***		p: 2.64 e-6 ***		p: 0.003 **	
Microvelia americana	1.5 ± 0.3	1.7 ± 0.1	2.4 ± 0.3	2.7 ± 0.2	2.2 ± 0.2	3.3 ± 0.2	0.7 ± 0.3	0.3 ± 0.2	1.1 ± 0.5	2.8 ± 0.6	1.3 ± 1.6	0.9 ± 0.5
	p: 0.22 n.s.		p: 1.8 e-4 ***		p: 2.2 e-16 ***		p: 5.09 e-5 ***		p: 8.99 e-11 ***		p: 0.41 n.s.	
Gerris buenoi	NA	2.9 ± 0.7	NA	12.3 ± 0.1	NA	8.5 ± 0.8	NA	0.5 ± 0.4	NA	15.2 ± 0.6	NA	1.0 ± 0.7

Table S3: Comparison of leg deployment and amplitude of movements during ground locomotion and locomotion on water surface. Related to Figure 3.

The locomotion of *Hebrus sp.*, *Mesovelia furcata*, *Microvelia americana* and *Gerris buenoi* are analysed. There is no important differences between ground and water locomotion for *Hebrus sp.*. The hindleg and mostly the midleg become more important during water surface locomotion in *Mesovelia* and *Microvelia*. Midleg is the most important leg for locomotion in *Gerris*. Each number is a mean of 18 measurements for *Hebrus sp.*, *Mesovelia furcata* and *Microvelia americana* and a mean of 6 values for *Gerris buenoi.* (±) indicates standard deviation. Student t- tests were performed.

Movie S1: Comparison between tripod gait and rowing gait during water surface locomotion.
Mesovelia furcata using the tripod gait (top) and *Gerris buenoi* using the rowing gait (bottom) during water surface locomotion. Videos were taken on a grid for size reference.

Supplemental experimental procedures

Insect sampling and culture

Extant specimens were collected during fieldwork in the locations indicated in Table S2. All species were kept in water tanks at 25°C, 55% humidity, 14 hours of day light and fed on live crickets.

Phylogenetic reconstruction

Sequences were retrieved from in house transcriptome and genomic sequence databases for the following markers: *12S RNA*; *16S RNA*; *18S RNA*; *28S RNA*; *Cytochrome Oxydase subunit I* (*COI*); *Cytochrome Oxydase subunit II* (*COII*); *Cytochorme Oxydase subunit III* (*COIII*); *Cytochrome b* (*cyt b*); *NADH-ubiquinone oxidoreductase chain 1* (*ND1*); *Ultrabithorax* (*Ubx*); *Sex combs reduced* (*Scr*); *Gamma interferon inducible thiol reductase* (*gilt*); *Antennapedia* (*Antp*); *Distal-less* (*dll*). All these markers were submitted to GenBank and their accession numbers can be found in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4. Sequences were aligned with MAFFT version 7.017 [S13] using default parameters. The alignments were visualized, corrected and concatenated in Geneious Version 7.1.9. Phylogenetic analysis was performed with MRBAYES version 3.2.6 [S14] (1 million generations; 25% burnin) and PhyML version 3.0 [S15], using GTR model with 100 bootstraps. Concatenation of sequence alignments and phylogenetic tree in Newick format are also available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4.

Habitat classification

Andersen [S16, S17] classified the various habitats of the Gerromorpha into eight classes: (1) terrestrial habitat far from water, (2) humid terrestrial habitat such as litter or humid gravel not necessary close to water, (3) marginal aquatic comprising moss, plants or rocks close to water, (4) plant-covered water surface, (5) water surface (with some plants), (6) streaming water with hood debris and foam, (7) stagnant water (such as ponds), (8) flowing water (such as streams). For simplicity, we consolidated these classes into the following four: terrestrial (Andersen's 1 and 2); marginal aquatic with preference for solid substrates (Andersen's 3); marginal aquatic with preference for water surface (Andersen's 4 and 5); and open water surface (Andersen's 6, 7 and 8). Each species was assigned to one of these four classes based on previous descriptions [S16-S20] and on the environment where we caught them.

Video acquisition, quantification of speed and stroke frequency

A set of adult individuals for each species were filmed at 2000 frames per second, both on water surface and on a solid substrate with a grid paper in the background as a calibration scale. Video acquisition was performed using the Phantom *Miro M310* Digital High Speed *Camera* and PCC Software (Vision research, Ametek). Videos were analyzed using TEMA 3.7 software (Images system) to extract speed values. To calculate speed, a mean value was extracted from a defined interval plateau phase from velocity curve along each video. This interval represents the max speed during the run of the individual. For speed, sample sizes were (on land/on water): *P. apt* (n=14/Not Applicable (NA)); *C. ali* (n=14/NA); *H.* sp. (n=10/10); *H. sta* (n=10/10); *M. fur* (n=10/12); *P. bul* (n=10/10); *M. ame* (n=9/8); *O. cun* (n=10/8); *S. str* (n=10/10); *S. ter* (n=10/11); *P. bra* (n=NA/18); *P. con* (n=10/3); *H. tur* (n=9/10); *G. bue* (n=8/8); *L. dis* (n=NA/8); *A. pal* (n=NA/8). Stroke frequency was determined as the number of strokes performed by the individuals during a given locomotion duration and converted into number of strokes per second. For stroke frequency, sample sizes were (on land/on water): *P. apt* (n=14/NA); *C. ali* (n=14/NA); *H.* sp. (n=10/10); *H. sta* (n=10/10); *M. fur* (n=10/12); *P. bul* (n=10/10); *M. ame* (n=10/9); *O. cun* (n=10/10); *S. str* (n=10/10); *S. ter* (n=10/22); *P. bra* (n=NA/35); *P. con* (n=10/6); *H. tur* (n=9/10); *G. bue* (n=8/8); *L. dis* (n=NA/8); *A. pal* (n=NA/8).

Measurements of leg length and body length

Measurements of the legs and body were performed using a SteREO Discovery V12 (Zeiss) with ZEN 2011 software (Zeiss). Body and leg length of the fossil species in (Figure S3) were extracted from the references in Table S1 and in the supplementary online information. Sample sizes used to perform these measurements were: *P. apt:* (n=10); *O. fas:* (n=10); *Y. mag:* (n=1); *C. ali:* (n=6); *H. vas:* (n=1); *L. pop:* (n=1); *S. meg:* (n=1); *M. dom:* (n=1); *M. fur:* (n=10); *M. and:* (n=1); *H. sp:* (n=9); *C. bra:* (n=1); *H. sta:* (n=10); *P. bor:* (n=1); *P. asi:* (n=1); *R. obe:* (n=10); *S. str:* (n=8); *S. ter:* (n=10); *P. bra:* (n=3); *V. cap:* (n=6); *O. cun:* (n=10); *P. bul:* (n=5); *P. bip:* (n=1); *P. con:* (n=1); *H. ele:* (n=1); *H. tur:* (n=6); *M. ele:* (n=1); *M. gri:* (n=1); *M. pol:* (n=1); *M. ame:* (n=10); *C. alb:* (n=1); *E. spi:* (n=1); *M. hes:* (n=10); *L. dis:* (n=10); *G. bue:* (n=6); *A. pal:* (n=8).

Analysis of leg pattern during locomotion

The deployment of the leg (distance from the body to the tip of the leg) and the amplitude of leg movement (distance between the point where the leg starts to push on substrate and the point where it loses contact with substrate) were extracted from the high-speed videos. To measure these parameters we took 3 videos on ground and 3 videos on water for *Hebrus sp., Mesovelia furcata* and *Microvelia americana*. In each video we extracted 6 strokes (n=18) to obtain an average of leg deployment and the amplitude of leg movements for the three legs. For *Gerris buenoi* we used 6 videos with 1 stroke per video (n=6). Measurements were performed using Image J software [S21]. To measure the duration of stance phases and swing phases we extracted 6 gait cycles for each species from our high-speed videos during both locomotion on ground and locomotion on water. Then we recorded the duration of each phases using the PCC Software (Vision research, Ametek).

Reconstruction of ancestral trait

Ancestral reconstruction for habitat was performed in Rstudio version 0.99.486 using a maximum likelihood method adapted to discrete characters (ace, package ape, [S22]) and represented using phytools [S23]. The simplest model "ER", with equal transition rates across all 4 habitat categories, was the best both with AIC and likelihood comparisons (p value = 0.28 for comparison ER and SYM; failure of convergence of ARD model). The pies for ancestral nodes represent marginal ancestral states. We reconstructed the ancestral value of the quantitative characters on the internal nodes of the phylogenetic tree using contMap (ML reconstructions, package Phytools, [S23]). The resulting figures are maps of the observed and ancestral reconstructed phenotypic trait values onto the tree using a color gradient. This was done for the different variables (T1-leg/Body, T2-leg/Body, T3-leg/Body, speed, stroke frequency). The R script used and the data are available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4.

Inference of energy consumption

We inferred the amount of energy spent per stroke based on the procedure from [S24]. Kinetic energy (E_k in Joules) used during a stroke is determined using the following expression:

$$E_k=0.5mv^2$$

where m is the mass of the insect in grams and v the velocity generated during one stroke in meters per second. The analyses were performed using speed data extracted from the high-speed movies. We first extracted, using PCC software (Vision research, Ametek), the time of takeoff that corresponds to the interval of time when the leg starts to apply pressure on the substrate until the leg loses contact with the substrate. The distance travelled during this interval of time is recorded using ImageJ [S21] and a takeoff velocity is calculated. Pools of live insects were weighted to determine a mean weight for one individual for each species. Then the energy per stroke is calculated using the mass and the takeoff velocity. Because individuals from *Paravelia bullialata, Stridulivelia strigosa, Platyvelia brachialis,* and *Husseyella turmalis* died during the interval of time between video acquisition and weight recording, the samples were conserved in absolute ethanol and rehydrated using the procedure from [S25] to obtain body mass. Control of rehydration was performed on dead insects compare to live specimens from a control species to evaluate the accuracy of the protocol (data not shown). The numbers of individuals per species weighted are the following: *M. fur:* (n=10)*; M. ame:* (n=10)*; O. cun:* (n=10); *P. bul:* (n=4)*; S. str:* (n=6)*; S. ter:* (n=10); *P. bra:* (n=3); *H. tur:* (n=5)*; L. dis:* (n=6)*; G. bue:* (n=5); *A. pal:* (n=6). The numbers of videos used to measure the take-off velocity are the following: *M. fur:* (n=12)*; M. ame:* (n=10)*; O. cun:* (n=10); *P. bul:* (n=10)*; S. str:* (n=10)*; S. ter:* (n=22); *P. bra:* (n=12); *H. tur:* (n=10)*; L. dis:* (n=8)*; G. bue:* (n=8); *A. pal:* (n=8).

Statistical analyses

The quantitative variables (T1-leg/Body, T2-leg/Body, T3-leg/Body, speed, stroke frequency) did not follow a normal distribution (Shapiro tests), and hence were log-transformed and a mean value is calculated for each species and for each variable. The "habitat" variable is semi-quantitative and was not log transformed. Because habitat and speed did not follow a normal distribution we performed a classical non-parametric Spearman correlation test and non-parametric Spearman correlation test with Phylogenetic Independent Contrast (PIC) in order to take into account for the non independence of data points resulting from a phylogeny ([S26]; implemented in ape version 3.5 and Picante version 1.6-2 packages [S22, S27]). FDR P-value correction for multiple tests was applied (rho and P-value are indicated on figures). Parametric Pearson correlation test with and without PIC correction were also performed. These results are available in supplementary online information available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4. For the quantification of leg deployment and leg movement, statistical significance between ground and water locomotion was determined by performing a Student t-tests. P-values are indicated in Table S3. Comparison of energy expenditure between species using the tripod gait and species using the rowing gait is performed using a non-parametric Wilcoxon test and graphically represented using the scatterplot3d version 0.3-37 package. Results are indicated in Table

S1. Statistical analyses were performed using RStudio Version 0.99.486 [S28]. Graphs were made using both RStudio and GraphPad Prism (version 6.01). Correlation matrixes were made using Picante version 1.6-2 package for the calculation and Corrplot version 0.77 package for the graphs [S29]. The R script used and the data are available in the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4.

Online ressources

Genbank accession numbers as well as sequence alignments for phylogenetic and tree reconstruction, the R script and dataset table used for phylogenetic and statistical analyses can be found in Dryad. Data available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.134c4.

Supplemental references

S1. D. Azar, A. Nel, M. S. Engel, R. Garrouste, and Matocq, A. (2011). A new family of Coreoidea from the Lower Cretaceous Lebanese amber (Hemiptera: Pentatomomorpha). . Polish Journal of Entomology, 627-644.

S2. Redei, D. (2007). A new species of the family Hypsipterygidae from Vietnam, with notes on the hypsipterygid fore wing venation (Heteroptera, Dipsocoromorpha). Deut Entomol Z *54*, 43-50.

S3. Azar, D., and Nel, A. (2010). The earliest fossil schizopterid bug (Insecta: Heteroptera) in the Lower Cretaceous amber of Lebanon. Ann Soc Entomol Fr *46*, 193-197.

S4. Yao, Y.Z., Zhang, W.T., and Ren, D. (2012). The first report of Mesoveliidae (Heteroptera: Gerromorpha) from the Yixian Formation of China and its taxonomic significance. Alcheringa *36*, 107-116.

S5. Garrouste, R., and Nel, A. (2010). First semi-aquatic bugs Mesoveliidae and Hebridae (Hemiptera: Heteroptera: Gerromorpha) in Miocene Dominican amber. Insect Syst Evol *41*, 93-102.

S6. Nel, A., and Popov, Y.A. (2000). The oldest known fossil Hydrometridae from the Lower Cretaceous of Brazil (Heteroptera : Gerromorpha). J Nat Hist *34*, 2315-2322.

S7. Zettel, H. (2001). Five new species of Perittopus Fieber, 1861 (Hemiptera : Veliidae) from Southeast Asia. Raffles B Zool *49*, 109-119.

S8. Andersen, N.M., and Poinar, G.O. (1998). A marine water strider (Hemiptera : Veliidae) from Dominican amber. Entomol Scand *29*, 1-9.

S9. Andersen, N.M. (2000). Fossil water striders in the Oligocene/Miocene Dominican amber (Hemiptera : Gerromorpha). Insect Syst Evol *31*, 411-431.

S10. Andersen, N.M. (1999). Microvelia polhemi, n. sp (Heteroptera : Veliidae) from Dominican amber: The first fossil record of a phytotelmic water strider. J New York Entomol S *107*, 135-144.

S11. Perrichot, V., Nel, A., and Neraudeau, D. (2005). Gerromorphan bugs in Early Cretaceous French amber (Insecta : Heteroptera): first representatives of Gerridae and their phylogenetic and palaeoecological implications. Cretaceous Res *26*, 793-800.

S12. Andersen, N.M., and Poinar, G.O. (1992). Phylogeny and Classification of an Extinct Water Strider Genus (Hemiptera, Gerridae) from Dominican Amber, with Evidence of Mate Guarding in a Fossil Insect. Z Zool Syst Evol *30*, 256-267.

S13. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res *30*, 3059-3066.

S14. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754-755.

S15. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol *52*, 696-704.

S16. Andersen, N.M. (1979). Phylogenetic Inference as Applied to the Study of Evolutionary Diversification of Semiaquatic Bugs (Hemiptera, Gerromorpha). Syst Zool *28*, 554-578.

S17. Andersen, N.M. (1982). The semiaquatic bugs (Hemiptera: Gerromorpha). Volume - Entomonograph Vol. 3. , (Klampenborg, Denmark.: Scandinavian Science Press LTD.).

S18. Dias-Silva, K., Moreira, F.F.F., Giehl, N.F.D., Nobrega, C.C., and Cabette, H.S.R. (2013). Gerromorpha (Hemiptera: Heteroptera) of eastern Mato Grosso State, Brazil: checklist, new records, and species distribution modeling. Zootaxa *3736*, 201-235.

S19. Heiss, E., and Pericart, J. (2007). Faune n° 91 – Hémiptères Aradidae, Piesmatidae et Dipsocoromorphes, (Fédération Française des Sociétés de Sciences naturelles).

S20. Schuh, R.T., and Slater, J.A. (1995). True bugs of the world (Hemiptera: Heteroptera) : classification and natural history, (Ithaca ; London: Comstock Publishing Associates, a division of Cornell University Press).

S21. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat Methods *9*, 671-675.

S22. Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics *20*, 289-290.
S23. Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol *3*, 217-223.
S24. Burrows, M., and Dorosenko, M. (2014). Jumping mechanisms in lacewings (Neuroptera, Chrysopidae and Hemerobiidae). J Exp Biol *217*, 4252-4261.
S25. Ungureanu, E.M. (1972). Methods for dissecting dry insects and insects preserved in fixative solutions or by refrigeration. Bull. Wld Hlth Org. *47*, 239-244.
S26. Felsenstein, J. (1985). Phylogenies and the Comparative Method. Am Nat *125*, 1-15.
S27. Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P., and Webb, C.O. (2010). Picante: R tools for integrating phylogenies and ecology. Bioinformatics *26*, 1463-1464.
S28. Team, R. (2015). RStudio: Integrated Development for R. .
S29. Garland, T., Harvey, P.H., and Ives, A.R. (1992). Procedures for the Analysis of Comparative Data Using Phylogenetically Independent Contrasts. Syst Biol *41*, 18-32.

Chapter 2

Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait.

Paper published in *Developmental Biology*

Co-author

Synopsis of the paper:

In the Gerromorpha, basally branching species have retained the ancestral ground plan similar to terrestrial species where T1-leg is shorter than T2-leg which is shorter than T3-leg. This ground plan, associated with the tripod gait, allows species to move both on water and ground. Some derived lineages have reversed this leg ground plan with the T2-leg now longer than T3-leg. This novel leg ground plan is associated with the derived rowing gait that facilitated the invasion and specialization in the open water habitat. Previous work demonstrated that the Hox transcription factor *Ubx* is expressed in both T2- and T3-legs and controls this reversal in leg ground plan (Khila et al 2009). However, *Ubx* expression is conserved between basally branching and derived species of semiaquatic bugs, yet it has a divergent role in leg ground plan. In this paper, we investigated the dynamics and mechanisms of *Ubx* control over leg development across species with different different morphologies and inhabiting different ecological niches. We compared a basally branching species, *Mesovelia furcata,* that possesses the ancestral ground plan and that lives in near shore; a derived species, *Microvelia americana*, that retained the ancestral ground plan and that also lives in near shore; and a derived member of the Gerridae, *Limnoporus dissortis*, a specialist of open water with the novel ground plan.

Summary of results:

Ubx antibody staining indicates that the levels of Ubx protein in T3-legs are much higher than in T2-legs of the same embryo. Our transcriptomic data suggests that Ubx dose is approximately 6-7 folds higher in T3-legs than in T2-legs. Using incremental depletion of the *Ubx* transcript by RNAi knockdown we demonstrated that in *Limnoporus dissortis* a moderate depletion results in a shorter T2-leg and a longer T3-leg. The phenotype of T2-leg is aggravated by further depletion of *Ubx*, in contrast to T3-leg where the phenotype is partially rescued by further depletion of the *Ubx* transcript. On the other hand, in walking species, *Ubx* depletion results in the shortening of T2- and T3-legs in a linear manner. Our results also indicate that Ubx controls this leg length adjustment without disrupting the pattern of expression of canonical developmental and patterning genes.

Conclusion:

Our results indicate that Ubx expression in species using the tripod gait, with the ancestral leg ground plan, induce leg elongation for both T2- and T3-legs; however, derived rowing species have developed tissue sensitivity to Ubx levels such that Ubx elongates the leg until a certain threshold above which it will repress leg elongation. Given the difference in levels between T2- and T3-legs, Ubx levels in T2-legs do not reach this threshold, which results in T2-legs being longer than T3-legs. This bimodal response to Ubx emerged in derived rowers, which resulted in the morphological adaptation of their legs to facilitate open water colonization. We also demonstrated that this sensitivity to Ubx dose shapes the legs without disrupting the expression of early developmental genes that are essential for leg development and patterning.

Author contributions:

Here are the contributions of the different authors:

Peter Refki: insect collection, gene cloning, in situ hydridization, immunostainings, RNAi knockdown of the different genes in *Limnoporus dissortis*, quantification and characterization of the phenotypes, design of the study and redaction of the paper.

David Armisén: transcriptome analysis, statistical analysis, quantification and characterization of the phenotypes.

Antonin Crumière: *Ubx* RNAi knockdown in *Mesovelia furcata* and *Microvelia americana*, quantification and characterization of the phenotypes.

Séverine Viala: scanning electron microscopy.

Abderrahman Khila: insect collection, design of the study and redaction of the paper.

Developmental Biology 392 (2014) 441–453

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Evolution of Developmental Control Mechanisms

Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait

Peter Nagui Refki [a,b], David Armisén [a], Antonin Jean Johan Crumière [a,b], Séverine Viala [a], Abderrahman Khila [a,*]

[a] Institut de Génomique Fonctionnelle de Lyon, CNRS-UMR5242, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France
[b] Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex, France

ARTICLE INFO

Article history:
Received 28 November 2013
Received in revised form
20 May 2014
Accepted 24 May 2014
Available online 2 June 2014

Keywords:
Hox genes
Adaptation
Allometry
Morphological evolution
RNA interference

ABSTRACT

Growth control scales morphological attributes and, therefore, provides a critical contribution to the evolution of adaptive traits. Yet, the genetic mechanisms underlying growth in the context of specific ecological adaptations are poorly understood. In water striders, adaptation to locomotion on the water surface is associated with allometric and functional changes in thoracic appendages, such that T2-legs, used as propelling oars, are longer than T3-legs, used as steering rudders. The Hox gene *Ubx* establishes this derived morphology by elongating T2-legs but shortening T3-legs. Using gene expression assays, RNAi knockdown, and comparative transcriptomics, we demonstrate that the evolution of water surface rowing as a novel means of locomotion is associated with the evolution of a dose-dependent promoting-repressing effect of Ubx on leg growth. In the water strider *Limnoporus dissortis*, T3-legs express six to seven times higher levels of Ubx compared to T2-legs. *Ubx* RNAi shortens T2-legs and the severity of this phenotype increases with increased depletion of Ubx protein. Conversely, *Ubx* RNAi lengthens T3-legs but this phenotype is partially rescued when Ubx protein is further depleted. This dose-dependent effect of Ubx on leg growth is absent in non-rowing relatives that retain the ancestral relative leg length. We also show that the spatial patterns of expression of *dpp, wg, hh, egfr, dll, exd, hth*, and *dac* are unchanged in *Ubx* RNAi treatments. This indicates that the dose-dependent opposite effect of Ubx on T2- and T3-legs operates without any apparent effect on the spatial expression of major leg patterning genes. Our data suggest that scaling of adaptive allometries can evolve through changes in the levels of expression of Hox proteins early during ontogeny, and in the sensitivity of the tissues that express them, without any major effects on pattern formation.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

Introduction

Growth control properly scales different body parts and is therefore a critical process in organismal development and evolution (Abzhanov et al., 2004; Bronikowski, 2000; Emlen et al., 2012; Nijhout and Grunert, 2010; O'Farrell, 2003; Stern and Emlen, 1999). A major challenge for understanding how growth can shape animal diversity is to decipher the mechanisms of growth control in the context of the particular ecological environment underlying species-specific life histories (Abzhanov et al., 2004; Emlen et al., 2012; Moczek and Rose, 2009; Stern and Emlen, 1999; Stillwell et al., 2011). Growth is intricately tied to pattern formation and biological systems must integrate inputs from both processes (Baena-Lopez et al., 2012). Organisms have been facing this challenge throughout evolution, and a variety of routes to reconciling growth and pattern formation were taken depending on the trait. For example in many systems, modulation of the growth of some traits is delayed to near the end of ontogeny, after all other essential body parts are specified and their identity acquired. It is possible that such late fine-tuning strategy circumvents the constraints imposed by the tight and pleiotropic interactions of genetic processes during development (Carroll, 2008; Duboule et al., 2007; Gibson, 1996; Wilkins, 2002). However, this case is restricted to some traits especially those required only during adulthood and have no or minimal function during juvenile stages. Examples include the development and evolution of a variety of exaggerated secondary sexual traits in beetles and water striders (Emlen et al., 2012; Khila et al., 2012), or appendage differentiation in holometabolous insects (Loehlin and Werren, 2012; Nijhout and Grunert, 2010). A greater challenge facing organisms during development and evolution is when structures must be specified and the scaling of their allometry fine-tuned during early development. This is generally required for attributes that are essential for both juveniles and adults; such is the case of

* Corresponding author.
E-mail address: abderrahman.khila@ens-lyon.fr (A. Khila).

http://dx.doi.org/10.1016/j.ydbio.2014.05.021
0012-1606/© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

locomotory appendages in hemimetabolous non-metamorphosing systems.

The case of water striders—a monophyletic group of semi-aquatic heteropterans—is a striking example where animals must pattern, dramatically grow, and fine-tune the final allometry of their appendages before the end of embryonic development, to adapt to the challenging requirements of their habitat (Fig. 1) (Andersen, 1982; Hu and Bush, 2010). Throughout the evolution of the semi-aquatic insects, the transition to life on water surface seems to be associated with a gradual increase in leg length relative to body length. Basally branching semi-aquatic insects can walk both on water and land, and share a common relative leg length plan, where T3-legs are longer than T2-legs, with their close relatives that are exclusively terrestrial (Fig. 1). Conversely derived lineages, which have specialized in water-surface locomotion through rowing, have evolved a novel adaptive morphology where the length of locomotory appendages has been reversed such that T2-legs are now longer than T3-legs (Fig. 1) (Andersen, 1982; Damgaard et al., 2005; Tseng and Rowe, 1999). Water striders use their T2-legs as propelling oars and their T3-legs as steering rudders (Andersen, 1982; Hu and Bush, 2010). T1-legs are much shorter and function primarily for stability, grasping the female during mating, and for handling preys (Andersen, 1982). These changes in the allometry of posterior thoracic appendages facilitated their functional adaptation and enabled the animals to generate efficient propulsion on the water–air interface (Andersen, 1976; Hu et al., 2003). As such, water striders have diversified to dominate a range of water surface niches, including ponds, streams, lakes, and even oceans (Andersen, 1982).

We are beginning to understand some of the genetic mechanisms responsible for these adaptive changes in leg allometry in water striders (Khila et al., 2009). Hox genes are conserved transcription factors that impart segment identity along the body axis across animals, through establishing a set of characters that distinguish segments that express them from those that do not (Akam, 1998a). Hox mutations result in the loss of the distinguishing characters and cause homeotic transformation of the affected segment to the likeness of another (Struhl, 1982). The Hox gene *Ultrabithorax* (*Ubx*), in particular, manages a multitude of small-scale differences between the second and third thoracic segments across a variety of arthropod species (Akam, 1998c; Stern, 1998, 2003). Among these phenotypic differences between segments, Ubx protein regulates the growth of the third thoracic appendages in a number of species, including holometabolous and hemimetabolous insects (Mahfooz et al., 2007; Stern, 2003). In holometabolous insects, there is a temporal decoupling of segmentation along the anterior–posterior axis of the embryo and appendage differentiation during post-embryonic stages (Baker, 1988). Late during *Drosophila* pupal development, Ubx modulates the shape and size of T3-legs, and establishes inter-species differences in trichome patterns (Stern, 1998, 2003). High levels of Ubx protein repress trichome development on T2 femur, and variation in Ubx regulation underlies variation in trichome patterns across species (Stern, 1998). During wing development, T2 imaginal discs, which do not express Ubx, differentiate into functional adult wings. However T3 imaginal discs, which do express Ubx, differentiate into the much smaller halteres (Roch and Akam, 2000). Ubx controls haltere morphogenesis through regulation of a large number of genetic factors at distinct developmental stages, including signaling molecules, transcription factors, and growth pathways (Castelli-Gair and Akam, 1995; Crickmore and Mann, 2006, 2008; Pavlopoulos and Akam, 2011; Roch and Akam, 2000). This indicates that in flies, Ubx can regulate specific morphogenetic processes during post-embryonic stages. In the hemimetabolous water striders however, the dramatic growth of the legs and the requirement for Ubx in fine-tuning their allometry coincide with the requirement for patterning both the appendages and the segments along the anterior–posterior axis of the embryo. In the embryo of water striders, Ubx protein is expressed in both T2- and T3-legs, and functions to lengthen T2- but to shorten T3-legs (Khila et al., 2009). This fine-tuning of relative leg length by Ubx is driven by changes in the spatial expression and function of the protein in the second and third thoracic segments (Khila et al., 2009). The novel deployment of Ubx in T2-legs distinguishes water striders from close terrestrial relatives, such as the milkweed bug *Oncopeltus fasciatus* where Ubx expression is restricted to T3-legs (Mahfooz et al., 2007). Similarly, the reversed function of Ubx to shorten T3-legs in water striders distinguishes them from sister basally branching semi-aquatic insects taxa (Khila et al., 2014). In the water strider *Gerris buenoi*, Ubx seems to be expressed at higher levels in T3-legs compared to T2-legs (Khila et al., 2009). However whether these differential levels of expression have a relevant role in regulating leg allometry remains to be tested. Here we investigate a possible role of Ubx levels in shaping

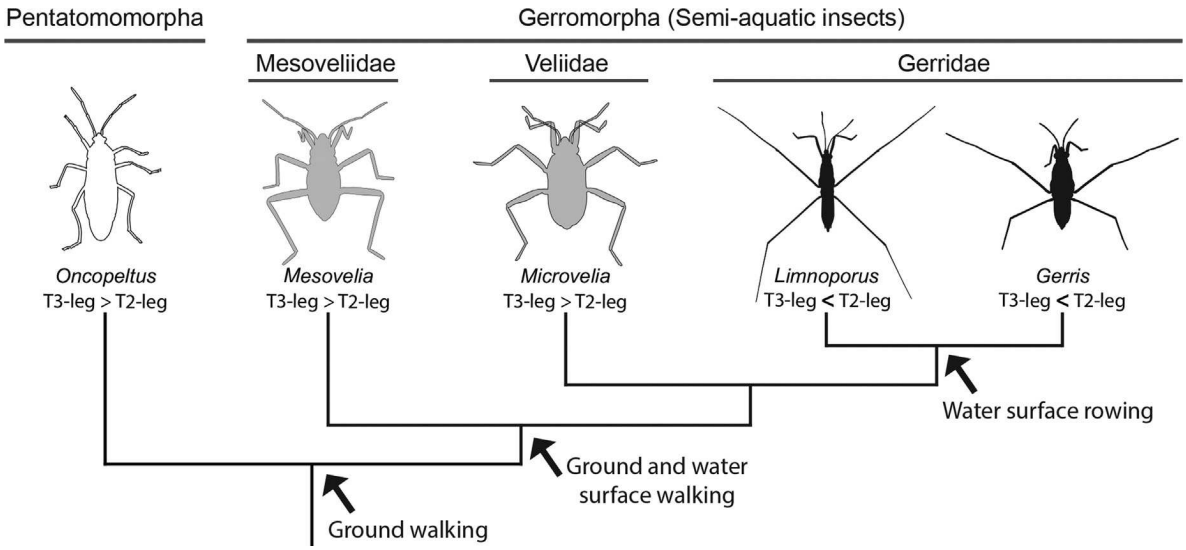


Fig. 1. Leg morphology and mode of locomotion across the semi-aquatic insects and the closely related terrestrial outgroup *Oncopeltus*. Both terrestrial insects (highlighted in white) and most semi-aquatic insects (highlighted in gray) retain the ancestral relative leg length where T2-legs are shorter than T3-legs. Derived lineages (highlighted in black), which specialize in open water rowing, have a reversed relative leg length where T2-legs are now longer than T3-legs.

the morphological differences that distinguish the second from the third thoracic segments during the evolution of the semi-aquatic insects. Although Ubx controls multiple morphological attributes in these segments, we focus on the morphology of the legs in the context of adaptation to life on the water surface. We use parental RNAi, which depletes Ubx from the start of embryogenesis and across all tissues expressing the gene. Although we obtain partial transformation of the third thoracic segment to the likeness of the second thoracic segment, our conclusions throughout the manuscript focus primarily on the response of the legs in each segment to *Ubx* RNAi. By manipulating the strength of RNAi knockdown and by analyzing the association between the responses of T2- and T3-legs to *Ubx* RNAi, we describe the evolution of tissue sensitivity to Ubx levels across a selection of semi-aquatic insects representing both basal and derived taxa. We also examine whether or not major leg patterning hierarchies are affected by the changes in *Ubx* regulation that led to the specialization of water striders in surface rowing.

Methods

Animal collection and rearing

Adult individuals of the water strider *Limnoporus dissortis* were collected from '*Rivière de l'Acadie*,' and *Microvelia americana* from '*Rivière du Nord*' at the vicinity of Montréal, Québec, Canada. *Mesovelia furcata* was collected in Vilette d'Anthon, '*plans d'eau de Salonique*' lake in Lyon, France. Animals were kept in aquaria at 25 °C with a 14-h light/10-h dark cycle, and were fed on live crickets. Pieces of floating Styrofoam were regularly supplied for the females to lay eggs.

Gene cloning

Limnoporus total RNA was extracted from different embryonic and nymphal stages. First strand cDNA synthesis was then performed, using total RNA as a template, according to Invitrogen manual instructions. To clone *Limnoporus dll*, *hh*, and *wg*, degenerate primers were designed based on sequence alignments of *Tribolium*, *Acyrthosiphon*, and *Pediculus*. Specific primers for *dac*, *dpp*, *egfr, hth*, and *Ubx* were designed based on sequences obtained from a whole transcriptome of *Limnoporus*. Primer sequences and Genbank accession numbers are presented in Table S1.

Embryo collection and dissection

Embryos were collected, treated with 25% bleach, and then washed with PTW 0.05% (1X PBS; 0.05% Tween-20). For image acquisition, late embryos were dissected out of the egg, fixed in 4% Formaldehyde, and their images captured using a Zeiss Discovery 12 scope. For staining, embryos of various early stages were dissected, cleaned from yolk, and kept briefly in PTW 0.05% on ice until fixation with the appropriate method, based on the type of subsequent staining (see below).

in situ hybridization

Dissected embryos were fixed in 200 μl 4% Paraformaldehyde (PFA)+20 μl Dimethyl Sulfoxide (DMSO), and 600 μl heptane for 20 min at room temperature with shaking. Embryos were then washed several times in cold methanol and rehydrated in decreasing concentrations of methanol in PTW 0.05%. These embryos were washed three times in PTW 0.05%, three times in PBT 0.3% (1X PBS; 0.3% Triton X100), and twice with PBT 1% (1X PBS; 1% Triton X100). Following these washes, embryos were transferred to 1:1 PBT 1%/hybridization solution (50% Formamide; 5% dextran sulfate; 100 μg/ml yeast tRNA; 1X salts). Embryos were pre-hybridized for 1 h at 60 °C, followed by addition of a Dig-labeled RNA probe overnight at 60 °C. Embryos were then transferred gradually from hybridization solution to PBT 0.3% through consecutive washes with 3:1, 1:1, 1:3 pre-warmed hybridization solution: PBT 0.3% gradient. A blocking step was performed with PAT (1X PBS; 1% Triton X100; 1% BSA) at room temperature followed by incubation with anti-DIG antibody coupled with alkaline phosphatase for 2 h at room temperature, or at 4 °C overnight. Embryos were washed several times in PBT 0.3% then in PTW 0.05% before color reaction is conducted with NBT/BCIP in alkaline phosphatase buffer (0.1 M Tris pH 9.5; 0.05 M $MgCl_2$; 0.1 M NaCl; 0.1% Tween-20).

Antibody staining

Dissected embryos were fixed in 4% Formaldehyde for 10–15 min at room temperature, washed three times in PTW 0.05% then three times in PBT 0.3%. Embryos were then incubated in a blocking solution (1X PBS; 0.1% Triton X100; 0.1% BSA; 10% NGS) for 1 h at room temperature, then incubated with a 1:5 dilution of the primary antibody (the FP6.87 mouse anti-UbdA or anti-Exd from *Developmental Studies Hybridoma Bank* (*DSHB*), rabbit anti-Mef2, and Alexa488-coupled anti-horseradish peroxidase) overnight at 4 °C. Both FP6.87 aliquots obtained from R. H. White and from DSHB cross-react in the semi-aquatic insects. Embryos were then washed four times in PTW 0.05%, three times in blocking solution and incubated in blocking solution for 30 min. Embryos were then incubated with the anti-mouse secondary antibody coupled with horseradish peroxidase at 1:1000 dilution for 2 h at room temperature. Several washes in PBT 0.3% and PTW 0.05% were performed before color reaction was conducted in DAB; 0.064% $NiCl_2$; 0.009% H_2O_2 for a black color, or without $NiCl_2$ for a brown color.

Quantification of expression using comparative transcriptomics

Limnoporus T2- and T3-legs from wild type and a sample of pooled *Ubx* RNAi treatments with a ~2 μg/μl ds-*Ubx* concentration, spanning mid- and late stages, were dissected and used for total RNA extraction. The RNA from these legs was used by *ProfilExperts* (Lyon-France) to conduct deep sequencing using Illumina technology. A total of ~50 million reads were generated per sample. These reads were aligned against a pre-existing full *Limnoporus* transcriptome, using BLASTn, to quantify *Ubx* transcript levels by determining the number of Reads Per Kilobase per Million mapped reads (RPKM) in each sample (Mortazavi et al., 2008).

Parental RNAi

Gene knockdown using parental RNAi was conducted following the protocol in (Khila et al., 2009). Control RNAi was conducted by injecting either ds-*yfp* (double stranded RNA against *yellow fluorescent protein*) or injection buffer. Template for *in vitro* transcription to produce ds-RNA for each gene was prepared using T7-tagged primers in Table S2. Embryos were collected from the Styrofoam stripes, treated with 25% bleach, and washed in 0.05% PTW. Embryos were routinely screened for phenotypes.

Leg measurements

A sample of 10 embryos ($N=10$) is used for each RNAi group in Fig. 3E and a total of 92 embryos in Fig. 4A–D. Two sets of measurements were recorded: egg length and the length of the

tarsus, tibia, and femur individually. Total leg length was calculated as the sum of the length of the three leg segments. After measuring egg length, late embryos prior to hatching were dissected and mounted on slides in Hoyer's medium. Measurements for each leg segment of each pair of legs were recorded on a Zeiss microscope using the Zen software.

Statistical analyses

Statistical significance in leg length between each RNAi group and controls was determined by performing an analysis of variance (ANOVA) with the mean value of each pair of leg segments corrected to the egg length to account for embryo size variations. Statistical analyses were performed using the SPSS (Statistical Package for the Social Sciences) software package (IBM Corporation). In the scatter plots in Fig. 4, the centered second order polynomial (quadratic) model, and the linear regression fit model were plotted using GraphPad Prism (version 6.01). R^2 values for the models are presented in Table S3.

Results

Ubx is expressed at higher levels in T3-legs relative to T2-legs across the semi-aquatic insects

To examine the evolution of differential levels of Ubx expression across the semi-aquatic insects we chose representatives from basally branching and derived lineages, some of which walk and others row on water (Fig. 1). In the surface rowing water strider *Limnoporus*, *in situ* hybridization revealed that the expression of *Ubx* mRNA is much stronger in T3- compared to T2-legs (Fig. 2A). *Ubx* expression is also detectable in abdominal segments, in addition to the thorax (Fig. 2A). To further confirm this difference in the levels of expression between the two legs, we detected Ubx protein using the FP6.87 antibody known to recognize both Ubx and Abdominal-A (Abd-A) proteins (Kelsh et al., 1994; White and Wilcox, 1984). Consistent with *in situ* hybridization, antibody staining is much stronger in T3-legs (Fig. 2B), supporting the conclusion that both *Ubx* mRNA and Ubx protein are expressed at higher levels in T3- relative to T2-legs in *Limnoporus* embryos. Interestingly, these clear differences in the levels of Ubx in the legs are not seen in a set of neurons located on T2 and T3 segments (Fig. S1). To verify that FP6.87 antibody reacts with Ubx protein only in the thorax, we stained embryos with *Abd-A* mRNA probe. *Abd-A* mRNA is restricted to abdominal segments A2–A8 and no staining was observed in the thorax (contrast Fig. 2A–C). Therefore, the strong signal of FP6.87 staining in T3-legs corresponds to Ubx alone, whereas the strong signal in the abdominal segments is a result of co-expression of Ubx and Abd-A there. Finally, to determine the differences in the levels of *Ubx* expression between T2- and T3-legs, we performed a comparative transcriptome between *Limnoporus* T2- and T3-legs using deep sequencing. This method indicated that the levels of *Ubx* transcript are over six to seven times higher in T3-legs compared to T2-legs, during both mid and late embryogenesis (Fig. 2D).

Apart from the Gerridae, most other semi-aquatic insects retain the ancestral leg allometry where T2-legs are shorter than T3-legs (Fig. 1) (Andersen, 1982; Damgaard et al., 2005). Therefore, we wanted to know whether this difference in Ubx levels between T2- and T3-legs is characteristic of lineages where relative leg length is reversed or whether it is a general feature of semi-aquatic insects regardless of relative leg length. We therefore chose *M. furcata*, a representative of the most basally branching Mesoveliidae, and *M. americana*, a representative of the derived Veliidae (Andersen, 1982; Damgaard et al., 2005). In contrast to *Limnoporus*, both *Mesovelia* and *Microvelia* retain ancestral leg length plan and can walk on water and land (Fig. 1) (Andersen, 1982). Surprisingly in both species, FP6.87 staining revealed a much stronger signal in T3- relative to T2-legs (Fig. 2E and F), just like in *Limnoporus*. These results suggest that across the semi-aquatic insects, Ubx is expressed at higher levels in T3- compared to T2-legs, regardless of phylogenetic position, relative leg length, or whether the legs are used for walking or rowing as a mode of water surface locomotion.

Stronger Ubx depletion results in stronger length phenotype in Limnoporus T2- but not T3-legs

By the end of embryo development, *Limnoporus* T2-legs reach a much longer size relative to T3-legs (compare control T2-leg and T3-leg in Figs. 3; S2). Because Ubx is expressed at higher levels in T3- relative to T2-legs, we tested whether these differences in expression levels contribute to the differential effect of Ubx in shaping the length of the two legs. To do this, we generated a series of mild to strong phenotypes by injecting low (0.1 μg/μl), moderate (1 μg/μl), or high (10 μg/μl) concentrations of *Ubx* double stranded RNA (ds-*Ubx*) to the mothers. First, we verified the specificity of *Ubx* RNAi, and whether *Ubx* depletion became stronger with increased amounts of injected ds-*Ubx*. Staining with the FP6.87 antibody allows us to assess the specificity of RNAi as ds-*Ubx* injection is expected to deplete Ubx but not Abd-A. In embryos obtained from mothers injected with a concentration of 0.1 μg/μl, FP6.87 staining still detects some Ubx protein in T3-legs in 66% of embryos examined, whereas no Ubx could be detected in T2-legs (Fig. 3B). In the 1 μg/μl and 10 μg/μl ds-*Ubx* injections, the percentage of embryos where Ubx was still detectable in T3-legs decreased to 39% and 26% respectively (Table S4) and the remaining embryo showed no Ubx expression in T3-legs (Fig. 3C). In embryos with strong Ubx depletion, the FP6.87 staining in the abdomen, which corresponds to Abd-A protein, became weaker because of the loss of Ubx expression there (Fig. 3C). Regardless of the injected concentration of ds-*Ubx*, Abd-A remains normally expressed in the abdomen (Fig. 3A–C). This indicates that there is a general correlation between the concentration of injected ds-*Ubx* and the severity of Ubx depletion. Therefore, increased ds-*Ubx* concentration generally results in stronger depletion of Ubx in the legs. Second, we wanted to know whether or not *Ubx* RNAi depletes *Ubx* transcripts in T2- and T3-legs with equal efficiency. We therefore analyzed *Ubx* transcript levels using the comparative transcriptomics data between legs dissected from wild type (WT) and *Ubx* RNAi dataset generated from a pool of samples injected with ~2 μg/μl concentration (Fig. 3D). In this dataset, the number of *Ubx* Reads Per Kilobase per Million (RPKM) is reduced from 32 in WT T2-legs to 17 in *Ubx* RNAi T2-legs, and from 162 in WT T3-legs to 92 in *Ubx* RNAi T3-legs (Fig. 3D). Therefore, RNAi depletes *Ubx* with equal efficiency in T2- and T3-legs, by 46.8% and 43.2% respectively. Combined, these experiments demonstrate that: (i) *Ubx* RNAi is specific and does not affect Abd-A; (ii) the efficiency of Ubx depletion generally increases with increased dose of injected ds-*Ubx*; and (iii) *Ubx* is depleted with equal efficiency in T2- and T3-legs.

To assess the effect of *Ubx* RNAi on the length of T2- and T3-legs, we dissected and measured each leg across a sample of embryos. In embryos laid by females injected with the lower 0.1 μg/μl concentration of ds-*Ubx*, T2-legs became shorter while T3-legs became longer compared to their respective control counterparts (Fig. 3E). In T2-legs of these embryos, we obtained a significant 38.6% and 16.3% shortening in the tarsus and tibia respectively, and a slight non-significant shortening effect in the femur (Fig. 3E; Table S5). In T3-legs of the same embryos, we obtained a significant 35.8% and 38.3% elongation in the tarsus and

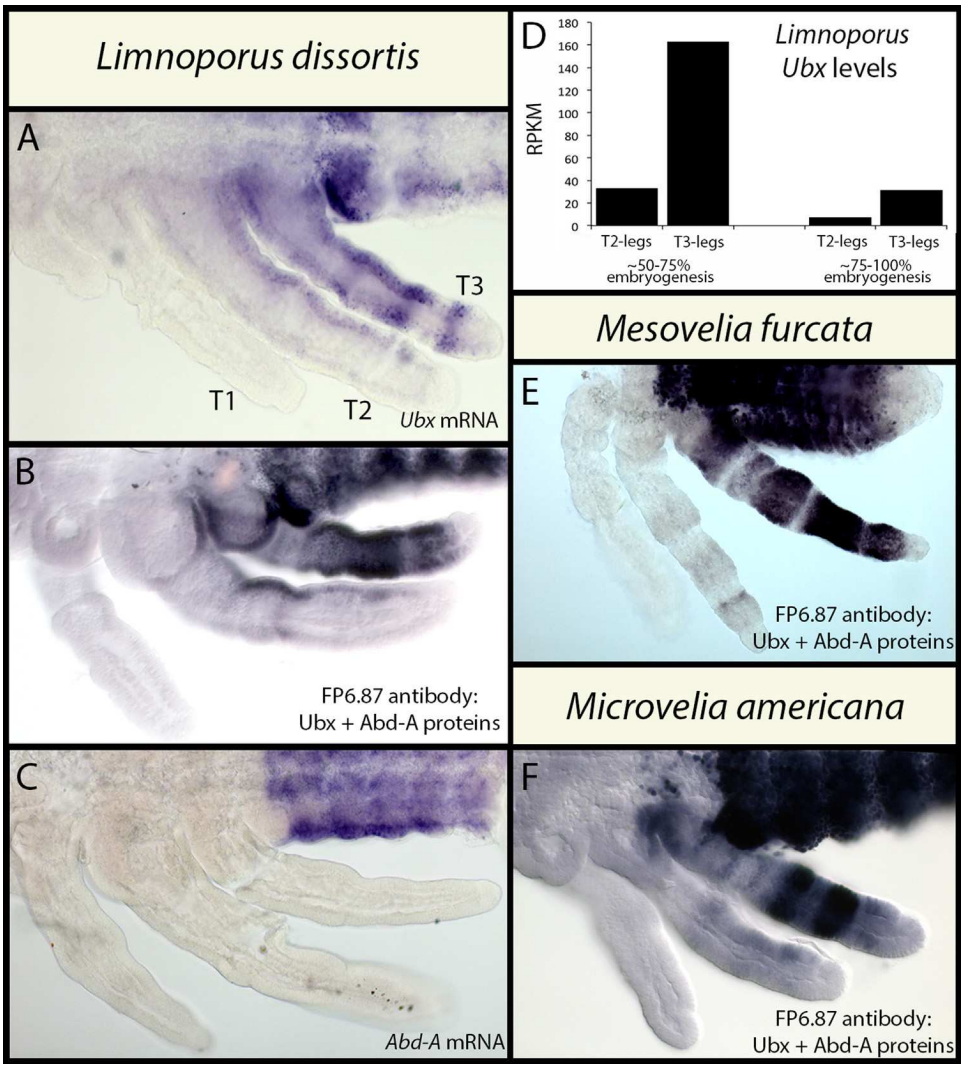


Fig. 2. Levels of *Ubx* mRNA and Ubx protein expression in T2- and T3-legs of three species of semi-aquatic insects. (A) Detection of *Limnoporus Ubx* mRNA using a Digoxygenin labeled probe. (B) Detection of both Ubx and Abd-A proteins using the FP6.87 anti-UbdA antibody in *Limnoporus* embryo. (C) Detection of *Limnoporus Abd-A* mRNA using a Digoxygenin labeled probe. Comparison between A, B, and C shows that Ubx is expressed in T2 and T3 thoracic as well as abdominal segments, whereas Abd-A is absent from thoracic segments. (D) Quantification, using deep sequencing, of the differences of *Ubx* transcript levels between T2- and T3-legs at mid- and late embryogenesis in *Limnoporus*. *Ubx* levels are consistently 6–7 times higher in T3- compared to T2-legs. (E and F) Detection of Ubx and Abd-A proteins using the FP6.87 anti-UbdA antibody in *Mesovelia* (E) and *Microvelia* (F) embryos. RPKM: Reads Per Kilobase per Million. T1–3: thoracic appendages 1–3.

tibia respectively, and a slight non-significant shortening in the femur (Fig. 3E; Table S5). This demonstrates that when a relatively low concentration of ds-*Ubx* is injected, *Ubx* RNAi shortens T2- and elongates T3-legs in a significant manner. In embryos laid by females injected with 1 μg/μl ds-*Ubx* concentration, the shortening effect on T2-legs became stronger, whereas the lengthening effect on T3-legs became milder (Fig. 3E; Table S5). In T2-legs of these embryos, all segments are now significantly shortened by 49.9%, 20.8%, and 8.3% in the tarsus, tibia, and femur respectively (Fig. 3E; Table S5). This suggests that the more Ubx is depleted the shorter T2-legs develop. In contrast, the elongating effect on T3-legs was attenuated, especially in the tarsus, by the injection of the 1 μg/μl dose of ds-*Ubx* such that the tarsus is now elongated by only 17.5% (instead of 35.8%). In the tibia and the femur of T3-legs, the 1 μg/μl dose of ds-*Ubx* affected the length in a similar manner as the lower 0.1 μg/μl dose (Table S5). When an even higher 10 μg/μl dose of ds-*Ubx* was injected, all three segments of T2-legs were further shortened, whereas in T3-legs the lengthening of the tarsus was further attenuated (10.1% instead of 17.5% obtained with 1 μg/μl and 35.8% obtained with 0.1 μg/μl ds-*Ubx* concentrations; Table S5). These results suggest that mild depletion of Ubx causes T3-legs to grow longer, but a stronger depletion reverses, at least partially, this lengthening effect.

Ubx dose-dependent effect on leg length is linear in walking but non-linear in surface rowing semi-aquatic insects

To determine the correlation between the responses of T2- and T3-legs to *Ubx* RNAi, we analyzed the association between T2- and T3-leg length in each individual using a larger dataset, composed of all injected concentrations and additional samples ($N=92$). This analysis completes the previous, average-based, method because it considers the values of leg length of each embryo rather than a mean, it controls for the variation in phenotype penetrance observed (Table S4) and that is characteristic to the RNAi technique (Jaubert-Possamai et al., 2007; Kitzmann et al., 2013), and finally it reveals how the lengths of the two legs change with respect to each other within each embryo. If Ubx operates to increase or decrease leg length in a dose-dependent manner, we should expect a non-linear association between the responses of T2- and T3-legs to *Ubx* RNAi. Testing this prediction is based on three consistent observations: (i) control legs express higher levels

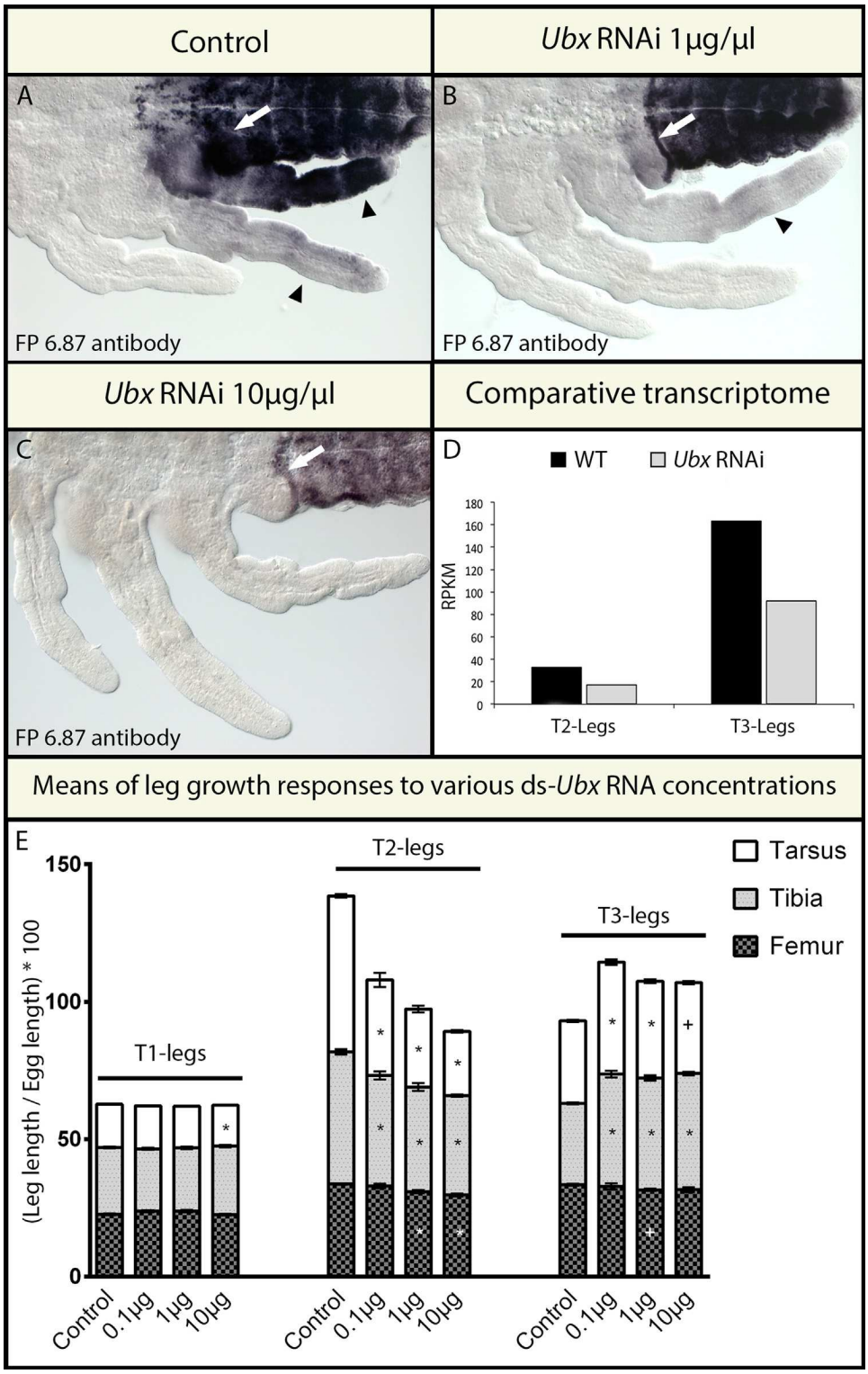


Fig. 3. Effect of RNAi on Ubx levels and on leg length in *Limnoporus* embryos. (A) Anti-UbdA staining is unaffected in control embryos and shows faint Ubx expression in T2-legs and stronger expression in T3-legs. In the abdomen, the anterior-most boundary of Abd-A is masked by the strong expression of Ubx (white arrow). (B) When a 1 µg/µl concentration of ds-*Ubx* was injected, Ubx is now undetectable in T2- and becomes faint in T3-legs. Note that the anterior-most boundary of Abd-A has now become sharp due to faint Ubx in abdominal segment A1 (white arrow). (C) When the even higher 10 µg/µl concentration of ds-*Ubx* was injected, Ubx is now undetectable in T2-legs, T3-legs, or in the abdomen. The anterior-most boundary of Abd-A remains intact (white arrow). Because Ubx is also expressed in abdominal segments and contributes to the strong signal there (Fig. 2A), the strong depletion of Ubx results in a lower signal in the abdomen, even when the reaction is left to develop longer. (D) Quantification, using deep sequencing, of the efficiency of *Ubx* transcript depletion in response to *Ubx* RNAi (pool of embryos from a ~2 µg/µl treatment). *Ubx* is depleted approximately by half and this depletion is uniform in T2- and T3-legs, such that the difference in Ubx levels remains between the two legs. (E) Growth response of T2- and T3-legs to the injection of increasing concentrations of ds-*Ubx*. * indicates statistical significance at $P \leq 0.01$ and + at $P \leq 0.05$. $N = 10$ in all samples.

of Ubx protein relative to their *Ubx* RNAi-treated counterparts (Fig. 3A–D); (ii) RNAi depletes *Ubx* transcripts equally in T2- and T3-legs (Fig. 3D); and (iii) the length of T2-legs consistently decreases with increased depletion of Ubx protein (Fig. 3E). Therefore, we can use the length of T2-legs as a proxy to infer Ubx levels, such that the shorter T2-legs are the lower Ubx protein

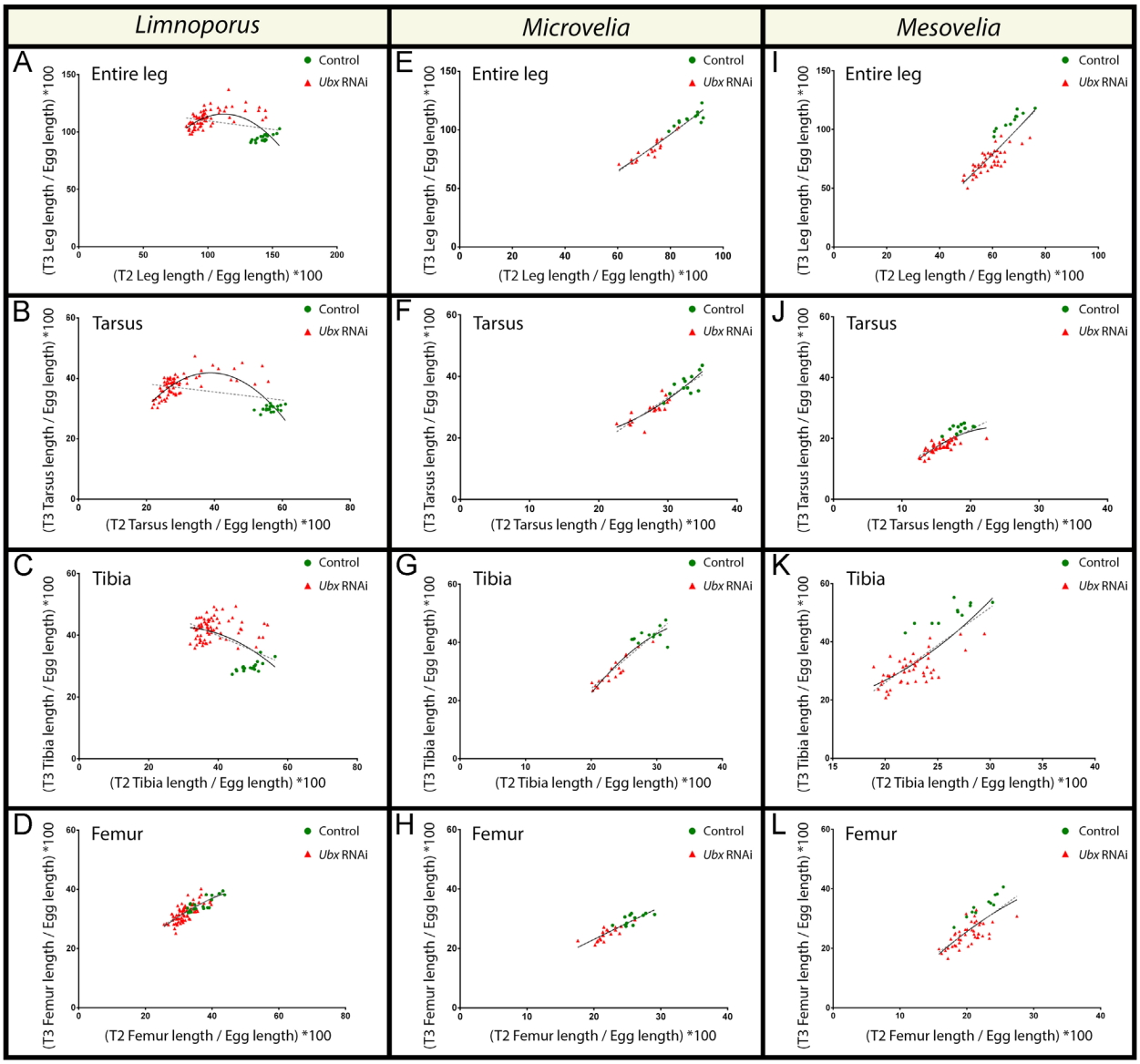


Fig. 4. Correlation between the responses of T2- and T3-legs to *Ubx* RNAi in the three semi-aquatic insects. Plotted values of leg length were corrected to egg length. Two fit models were plotted on the data for comparison: a centered second order polynomial (quadratic) curve shown as a solid black curve, and a linear regression fit shown as a dashed line. Each data point represents an individual; ds-*Ubx* treated individuals are represented by red triangles and control individuals are represented by green dots. (A) *Limnoporus* total T3-leg length relative to T2-leg length showing that the data points follow the curved model (R^2=0.3457) rather than the linear model (R^2=0.1197); N=92. (B–D) Individual leg segments of the T3-leg were plotted against their T2-leg counterparts to analyze the contribution of each to the global effect on the total curve. (B) *Limnoporus* T3-leg tarsus; (C) *Limnoporus* T3-leg tibia; and (D) *Limnoporus* T3-leg femur. The same analyses were conducted on *Microvelia* (E–H), N=31; and *Mesovelia* (I–L), N=56. In these two species, the centered second order polynomial follows the same curve as the linear model. Value of R^2 showing best fit for these models can be found in Table S3.

levels are. Based on this, we plotted T2-leg length by T3-leg length using a dataset including a larger sample size than the sample used in Fig. 3E, to determine the nature of this association (Fig. 4). This analysis revealed that in *Limnoporus*, when T2-leg length was mildly reduced in response to *Ubx* RNAi, the increase in T3-leg length was substantial (Fig. 4A). Therefore, a slight depletion may have shifted the levels of Ubx in T3-legs from those sufficiently high to decrease, to those that are low enough to increase leg length. The shift in the morphology of *Ubx* RNAi-treated T3-legs to the likeness of control T2-legs provides further support to this conclusion. In contrast, when T2-leg length was severely reduced, the increase in T3-leg length became attenuated (Fig. 4A). This further confirms that a strong depletion brings the levels of Ubx in T3-legs even lower such that T3-leg length is now reduced. To determine whether or not all the segments within T3-legs are responsible for this non-linear response to *Ubx* RNAi, we analyzed the association between the responses of the tarsi, tibias, and femora individually (Fig. 4B–D). We found that the response of T3-leg tarsus (Fig. 4B), in particular, and the tibia (Fig. 4C), to a lesser degree, to be non-linear dose-dependent, whereas that of the femur is largely linear (Fig. 4D). Therefore in *Limnoporus*, the tarsus is primarily responsible for the overall non-linear response of T3-legs to *Ubx* RNAi. Altogether, these data suggest that in *Limnoporus* embryos, the lengthening/shortening effect of Ubx is dose-dependent and that the lengthening of T2-legs is due to the lower levels of Ubx and the shortening of T3-legs is due to the higher levels of Ubx.

Next, we wanted to test whether the dose-dependent requirement of Ubx to regulate leg length is a general feature of the semi-aquatic insects, or whether it is characteristic of species where

relative leg length has been reversed. We therefore conducted a series of injection experiments with increasing doses of ds-*Ubx* in the two walking species, *Microvelia* and *Mesovelia* to determine if the strength of *Ubx* depletion would result in opposing effects on leg length. In both *Microvelia* (Fig. 4E–H) and *Mesovelia* (Fig. 4I–L), *Ubx* RNAi resulted in a consistent decrease of both T2- and T3-leg length in all injected concentrations. A plot of T2-leg length by T3-leg length in these two species revealed that, in contrast to *Limnoporus*, the lengths of the two legs decrease together in a linear manner in response to *Ubx* RNAi, regardless of the strength of this reduction. This linear relationship is present in all leg segments in the two species (Fig. 4E–L). Therefore, Ubx is required to increase the length of both T2- and T3-legs in these two species and no repressive role can be detected despite the apparent differences in Ubx levels between the two legs (Fig. 2E and F). Collectively, these findings suggest that the evolution of the dose-dependent activating-repressing effect of Ubx on leg growth is characteristic to the surface rowing water striders with derived relative leg length, but not other semi-aquatic insects that retain the ancestral state of leg length.

Ubx RNAi partially transforms T3-legs to the likeness of T2-legs

Because our experiment is based on parental RNAi, which depletes *Ubx* transcript starting early embryogenesis, we wanted to know the extent to which this experiment transforms one segment into the likeness of another. First, we looked for morphological and cellular attributes that could help differentiate thoracic segments from each other in normal animals. In *Limnoporus*, T1-legs are equipped with two distinctive combs; one sex comb and one grooming comb both located on the distal tibia (Fig. S3). Both T2- and T3-legs are similarly equipped with one grooming comb on the distal tibia (Fig. S3) (Khila et al., 2014). Therefore, while T2 and T3 can be clearly distinguished from T1 based on comb number, they are not distinguishable from each other based on this trait. We then stained *Limnoporus* embryos to reveal muscles, using the anti-Mef2 antibody (Sandmann et al., 2006), and nervous system, using the anti-HRP antibody (Jan and Jan, 1982; Paschinger et al., 2009), in an attempt to identify cellular characteristics that would help distinguish T2 from T3. Both antibodies revealed that muscular and nervous system structures are highly similar across all legs (Fig. S4). Therefore, the major observable difference between T2- and T3-legs in *Limnoporus* is their relative length and the length of the tarsi, tibias, and femora within each of them.

To examine the effect of *Ubx* RNAi on the development of these distinctive characteristics between segments across the three species, we quantified the degree of homeotic transformation in both mild and strong *Ubx* phenotypes. To do this, we selected a sample of five individuals representing the mildest and five representing the strongest *Ubx* phenotype presented in Fig. 4. In *Limnoporus*, because differences in length of the segments within legs are the only distinctive character that we could identify between T2-legs and T3-legs, we quantified the changes in the ratios of the tarsi, tibias, and femora in *Ubx* RNAi individuals in comparison with controls. We hypothesized that if the transformation of T3 towards T2 is complete, the ratio of each of the segments of *Ubx* RNAi T3-legs to control T2-legs should give an approximate value of 1. In normal *Limnoporus* embryos, the average ratio of T3-tarsus to T2-tarsus is 0.52, T3-tibia to T2-tibia is 0.58, and T3-femur to T2-femur is 0.9 (see Table S6A). We found in the sample representing the mildest *Ubx* RNAi phenotype that the average ratio of *Ubx*-T3-tarsus to control-T2-tarsus increases to the value of 0.7, the average ratio of *Ubx*-T3-tibia to control-T2-tibia increases to the value of 0.85 and the average ratio of *Ubx*-T3-femur to control-T2-femur decreases to the value of 0.86 (Table S6B). In the sample representing the strongest *Ubx* RNAi phenotype however, we found that the average ratio of *Ubx*-T3-tarsus to control-T2-tarsus decreases back to 0.57, the average ratio of *Ubx*-T3-tibia to control-T2-tibia decreases back to 0.78 and the average ratio of *Ubx*-T3-femur to control-T2-femur decreases to 0.71 (Table S6C). These changes in the ratios of leg segments with the differences in the strength of *Ubx* RNAi are consistent with the expression of Ubx in both legs. Mild *Ubx* RNAi probably brings the levels of Ubx protein in T3-legs down to the levels normally expressed in normal T2-legs, explaining the increase of the ratios towards the value of 1. This is the closest transformation of T3-legs towards T2-legs we observed. Conversely, strong *Ubx* RNAi depletes Ubx protein such that the remaining levels are even lower than those in normal T2-legs, explaining the decrease in the ratios of leg segments back away from 1. These data suggest that both mild and strong *Ubx* RNAi result in a partial transformation of T3-legs towards T2 identity. Aside from leg length, we were not able to detect any clear changes in the patterns of axons or muscles between T2- and T3-legs upon *Ubx* RNAi treatments in *Limnoporus* (Fig. S4).

In contrast to *Limnoporus*, we were able to unambiguously identify a set of morphological attributes that distinguish T2-and T3-legs in both *Microvelia* and *Mesovelia*. *Microvelia* (Fig. 5) and *Mesovelia* (Fig. S5) (Khila et al., 2014) have two combs in T1 (arrows in Figs. 5A; S5A) and one comb in T2 (arrow in Figs. 5B; S5B), just like *Limnoporus*. However, unlike *Limnoporus*, both species have no combs at all in T3-legs (Figs. 5C; S5C). Instead, T3-legs bear large bristles on the tibia that distinguish them from T2-legs (white arrowheads in Figs. 5C and S5C). In addition in *Microvelia*, T3-legs bear a particular bristle that is forked in its distal part and that is not found on T2-legs (red arrowhead in Fig. 5C). In both *Mesovelia* and *Microvelia*, we assessed the degree of transformation between T2-legs and T3-legs based on the

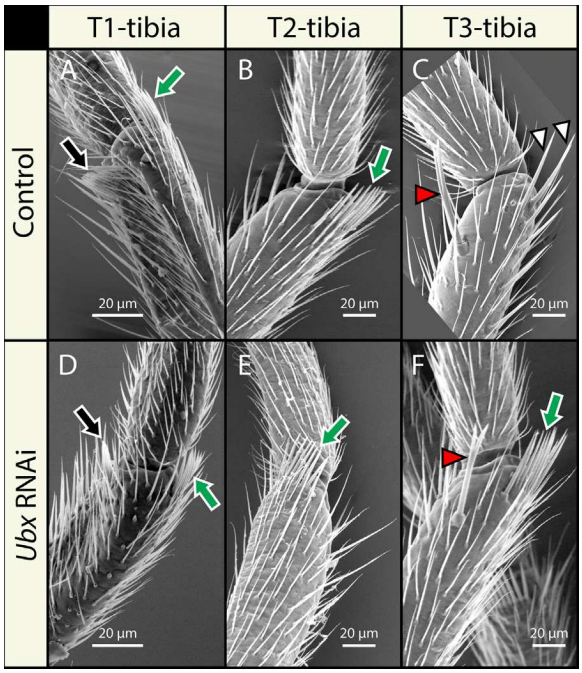


Fig. 5. *Ubx* RNAi causes a partial homeotic transformation of T3-legs to T2-legs. *Microvelia* forelegs bear one sex comb (black arrow in A) and one grooming comb (green arrow in A), the T2-legs bear a single grooming comb (green arrow in B), and the T3-legs do not have any combs (C). T3-legs however bear additional thick bristles (white arrowheads in C) and a characteristic forked bristle (red arrowhead in C). *Ubx* RNAi does not affect the combs either in the T1-legs (D) or in the T2-legs (E). T3-legs in *Ubx* RNAi treatment now develop an ectopic comb (green arrow in F), lose the thick bristles, but retain their characteristic forked bristle (red arrowhead in F).

presence or absence of the grooming comb and the morphology of T3 distinctive bristles. In no case have we observed the development of a sex comb, in addition to the grooming comb, on T2-legs in any of the species including *Limnoporus* (Figs. 5E; S3E; S5E; Table S7). This is consistent with the well-known expression of the Hox gene *sex combs reduced* in T1 but not in the other thoracic segments across insects (Rogers et al., 1997). Like in *Limnoporus*, we found that *Ubx* RNAi causes a partial transformation of T3-legs towards T2-legs in these two species. In *Microvelia*, *Ubx* RNAi occasionally causes T3-legs to develop a grooming comb similar to the one found on T2-legs (green arrow in Fig. 5F). The formation of this comb is observed in all individuals representing the strongest, but not all those representing the mildest *Ubx* RNAi phenotype (Table S7). However, *Ubx* RNAi, whether mild or strong, never resulted in the loss of the large forked bristle (red arrowhead in Fig. 5F; Table S7). We observed a similar dynamics in *Mesovelia* with regards to the presence/absence of the comb and the size of the bristles on T3-legs (green arrow and white arrowheads in Fig. S5F; Table S7). Altogether, these data suggest that in *Limnoporus* the weaker the *Ubx* RNAi the closer the morphology adopted by T3-legs to T2-legs, and conversely the stronger the *Ubx* RNAi the farther the morphology adopted by T3-legs from T2-legs. However in both *Mesovelia* and *Microvelia*, this relationship is reversed such that the stronger the *Ubx* RNAi the more the T3-legs resembles T2-legs.

The spatial expression of leg patterning genes is unaffected in Ubx RNAi treatments

Ubx is known to regulate a set of important leg patterning genes, including *decapentaplegic* (*dpp*), *hedgehog* (*hh*), *epidermal growth factor receptor* (*egfr*), and *wingless* (*wg*) in insects (Estella and Mann, 2008; Estella et al., 2008; Grossmann and Prpic, 2012). Because of the changes in the spatial pattern and the levels of Ubx expression as well as its opposite effect on leg length throughout the evolution of the semi-aquatic insects, we wanted to know whether or not these changes in Ubx affected the expression of a set of important leg patterning genes. First, we interrogated our comparative leg transcriptome datasets generated from normal and *Ubx* RNAi treatments to determine if the levels of expression of these key patterning genes are affected by Ubx depletion. In untreated embryos, the levels of expression of all genes tested, as revealed by the obtained RPKM, are comparable between T2- and T3-legs (Table S8). In *Ubx* RNAi treatments, we detected an increase in the levels of expression of these genes ranging from subtle (15–20%) to double the expression levels, as revealed by their RPKM (Table S8). In this analysis, *hh, egfr, exd*, and *wg* showed over 50% increase in both T2- and T3-legs upon *Ubx* RNAi (Table S8). These data suggest that Ubx may regulate at various degrees the levels of expression of many developmental genes, consistent with Ubx role in flies (Pavlopoulos and Akam, 2011).

Next, we wanted to determine whether this change in the transcriptional levels affected the spatial patterns of expression of these genes in the various legs of *Limnoporus* embryos. First, we examined their expression patterns in normal embryos. *in situ* hybridization and immuno-staining on untreated embryos revealed that all these developmental genes show specific patterns of expression in the developing legs (Fig. 6). However, unlike *Ubx* (Fig. 2), all these genes are expressed in an invariable pattern in all legs. At mid-embryogenesis, *wg* mRNA is distributed in a ventral segmental pattern and extends along the proximal–distal axis of all thoracic appendages (Fig. 6A). *dpp* mRNA is distributed in a punctuated pattern along both anterior and posterior edges of the appendages (Fig. 6B). *hh* is strongly expressed in the posterior compartment, and fades towards the anterior (Fig. 6C). In addition to this posterior-to-anterior distribution, *hh* mRNA forms four relatively faint stripes along the proximal–distal axis. *egfr* is expressed in five distinct rings along the proximal–distal axis, which seem to prefigure the junctions between the five leg segments. In addition, a sixth *egfr* ring separates the claw-bearing tip of the leg from the proximal part of the tarsus (Fig. 6D). mRNA of the gene *distal-less* (*dll*), required for the specification of distal appendages across arthropods (Panganiban et al., 1994), is expressed in a conserved pattern consisting of a proximal ring, and uniformly throughout the distal tarsus (Fig. 6I). *dachshund* (*dac*), which specifies proximal appendages (Abzhanov and Kaufman, 2000; Angelini and Kaufman, 2004; Prpic et al., 2003, 2001), is also expressed in a conserved pattern consisting of a single band extending from the mid-femur until the boundary between the tibia and the tarsus (Fig. 6J). Both *hth* and Exd, two known co-factors that bind and direct Hox proteins to their specific target DNA (Joshi et al., 2007; Mann et al., 2009) are expressed in a highly similar pattern and their domains span the proximal segments from the coxa to the proximal half of the femur (Fig. 6K and L). Second, we revealed the spatial patterns of expression of all these leg patterning genes in the developing legs in both 1 μg/μl (Fig. 6E–H and M–P) and 10 μg/μl (not shown) *Ubx* RNAi treatments. Following this treatment, we were not able to detect any changes in the spatial patterns of expression of any of the eight patterning molecules in the legs (Fig. 6). This result indicates that, despite the effect of *Ubx* RNAi on their levels of expression, neither the mild nor the severe depletion of *Ubx* resulted in any obvious effect on the spatial expression of these leg patterning genes in *Limnoporus*. Therefore, the spatial patterns of these genes are likely regulated independently from Ubx during *Limnoporus* leg development.

We also tested the possibility that *Ubx* might be downstream of some of the signaling molecules by examining the pattern of Ubx protein expression using the FP6.87 anti-UbdA antibody in *wg* and *hh* RNAi knockdown backgrounds. FP6.87 staining in *wg* and *hh* RNAi-treated embryos is unchanged in abdominal segments or in T2- and T3-legs despite the apparent developmental defects seen in these embryos (Fig. S6A–C). Altogether, these results suggest that Ubx may play a role in modulating the levels but not the spatial expression of leg patterning genes in *Limnoporus*.

Discussion

In this work, we present four main findings: (i) that generally in the semi-aquatic insects, Ubx is expressed at higher levels in T3- relative to T2-legs; (ii) that it is only in the derived Gerridae where the high levels of Ubx result in reduced T3-leg length; (iii) that in the surface rowing *L. dissortis*, both strong and mild depletion of Ubx result only in a partial transformation of T3 to the likeness of T2 segment; and (iv) that the changes in Ubx regulation and function have evolved in *Limnoporus* without any apparent effect on the spatial patterns of expression of a set of major leg patterning genes.

Changes in Ubx levels and their role in establishing leg allometry across the semi-aquatic insects

In the embryos of the water strider *Limnoporus*, Ubx dramatically increases the length of T2-legs whereas six to seven times higher levels of Ubx decrease the length of T3-legs, thereby reversing their relative length. *Ubx* RNAi embryos still exhibit basal levels of leg length indicating that a developmental ground state of leg growth is achieved without Ubx input. Therefore, this dose-dependent role of Ubx is required for fine-tuning relative length of T2- and T3-legs. Preliminary evidence indicates that most of the differences in leg length are due to differences in the

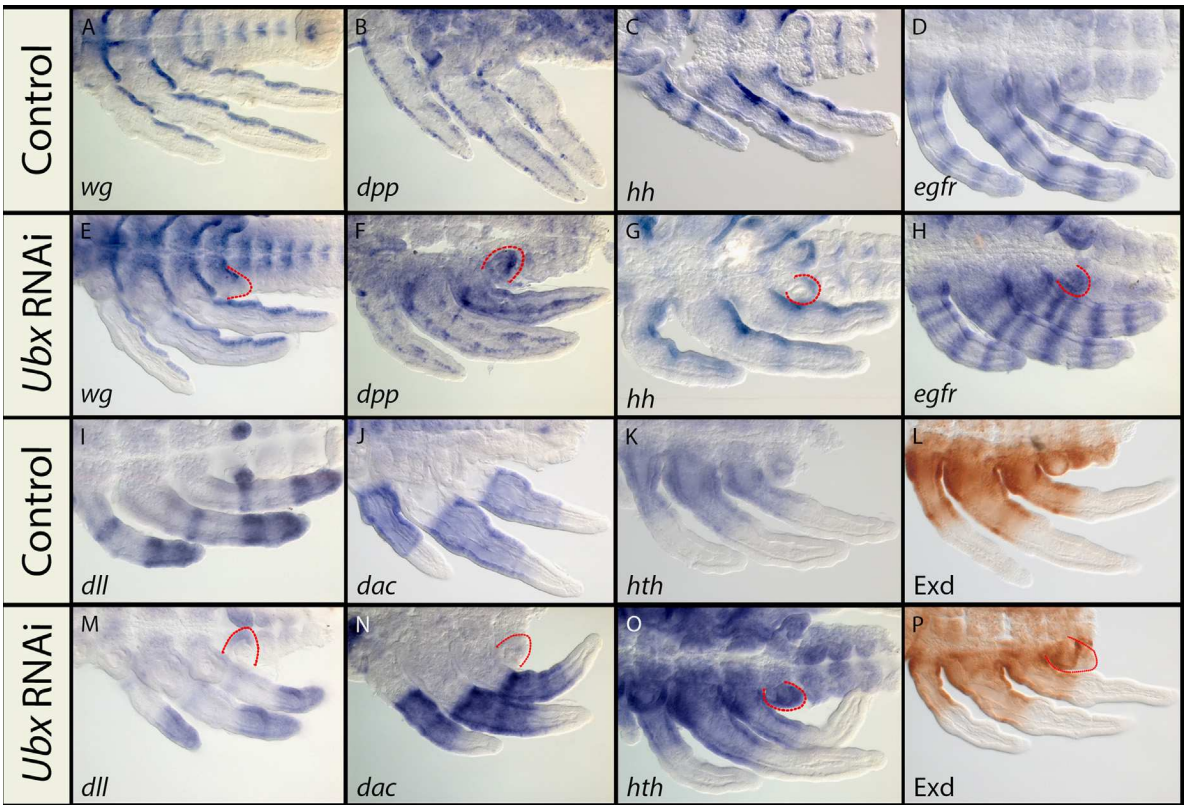


Fig. 6. Patterns of expression of *Limnoporus* developmental genes in the legs of control (A–D and I–L) and *Ubx* RNAi treatments (E–H and M–P). (A) *wg* is expressed along the posterior compartment of the three appendages. (B) *dpp* appears in a patched expression along the anterior posterior compartments of all legs. (C) *hh* is expressed as a stripe in the posterior compartment and fading towards the anterior side of all the legs, in addition to a weak ring expression. (D) *egfr* forms multiple proximo–distal stripes at the joints between the five leg segments. (E–H) patterns of the same genes in the 1 μg/μl *Ubx* RNAi treatments, showing no apparent change in the spatial patterns of these genes. (I) *dll* shows the classical 'ring' and 'sock' expression, in addition to the prominent expression in the first abdominal segment. (J) *dac* expression domain extends from the mid-femur until the end of the tibia. (K) *hth* and (L) Exd share the same pattern extending from the coxa until the mid-femur of all three legs. (M–P) patterns of the same genes in the 1 μg/μl *Ubx* RNAi treatments, showing no apparent change in the spatial patterns of these genes. All stainings represent mRNA (purple), except Exd, which represents the protein (brown). Highlighted in red is the ectopic limb bud that forms on the first abdominal segment, which is a phenotype characteristic to *Ubx* RNAi.

rate of cell division (in preparation), and suggest that Ubx mediates its effect on relative leg length through differential control of cell proliferation in each leg. In the terrestrial heteropteran relative *Oncopeltus* (Fig. 7A), Ubx expression is restricted to T3-legs and is required for increasing their length (Fig. 7B; Mahfooz et al., 2007). Although we have not measured the levels in *Oncopeltus*, histochemistry revealing mRNA (Angelini et al., 2005) or protein (Mahfooz et al., 2007) seems to indicate that Ubx levels in T3-legs are low. It appears that during the split between semi-aquatic and terrestrial lineages, both the novel deployment of Ubx in T2-legs and the increase of its levels in T3-legs emerged at the same time, as this feature is common to all semi-aquatic species tested (Fig. 7A). At this early stage of the evolution of the group, Ubx was still required to increase leg length and the increase in Ubx protein levels did not reverse this lengthening effect (Fig. 7C). This is clearly illustrated in the basally branching *Mesovelia* and the derived *Microvelia*, both of which retain the ancestral state of relative leg length (Andersen, 1982; Damgaard, 2008; Damgaard et al., 2005), and where the response of the legs to *Ubx* RNAi is linear (Fig. 7C). Although this first change in *Ubx* regulation at the base of the Gerromorpha resulted only in a subtle global increase in the length of both T2- and T3-legs, it has also provided a genetic potential that set the stage for further diversification of the group. This genetic potential was exploited by the Gerridae, such as *L. dissortis* (Fig. 7D), through the emergence of a novel sensitivity to the pre-existing high Ubx levels in T3-legs (Fig. 4). We conclude that this sensitivity of leg tissue to Ubx levels is a derived feature that emerged in surface-rowing species, and distinguishes them from the ancestral state inherited from the common ancestor of the semi-aquatic insects (Fig. 7A). Therefore, adaptive morphological traits can evolve through fine-tuning the levels of expression of key developmental genes early during embryogenesis and the emergence of tissue sensitivity to changes in these levels. This may have facilitated the change in leg function from walking to rowing structures, and has been instrumental in the specialization and radiation of the Gerridae in various open water niches, from small ponds to the vast surfaces of the great oceans.

Dose-dependent switch of Ubx role between increasing and decreasing leg length

The molecular interactions underlying the different responses of *Limnoporus* T3-legs and those of *Microvelia* and *Mesovelia* to high levels of Ubx remain to be investigated. The response of the legs in *Limnoporus* to *Ubx* RNAi suggests that there is a dose threshold, and when Ubx exceeds this threshold the growth-suppressing effect ensues. By knocking down *Ubx*, we have shifted the levels of Ubx protein in T3-legs close to or below this threshold, thereby increasing their length (Figs. 3E; 4A; 7D). In control T2-legs, Ubx levels are already below this threshold and RNAi brings these levels even further down, which explains the consistent decrease in T2-leg length with increased ds-*Ubx* injected concentration (Figs. 3E; 4A; 7D). A number of molecular mechanisms have been previously described and might explain these distinct responses of leg tissues to Ubx dose. In the fly *Drosophila*,

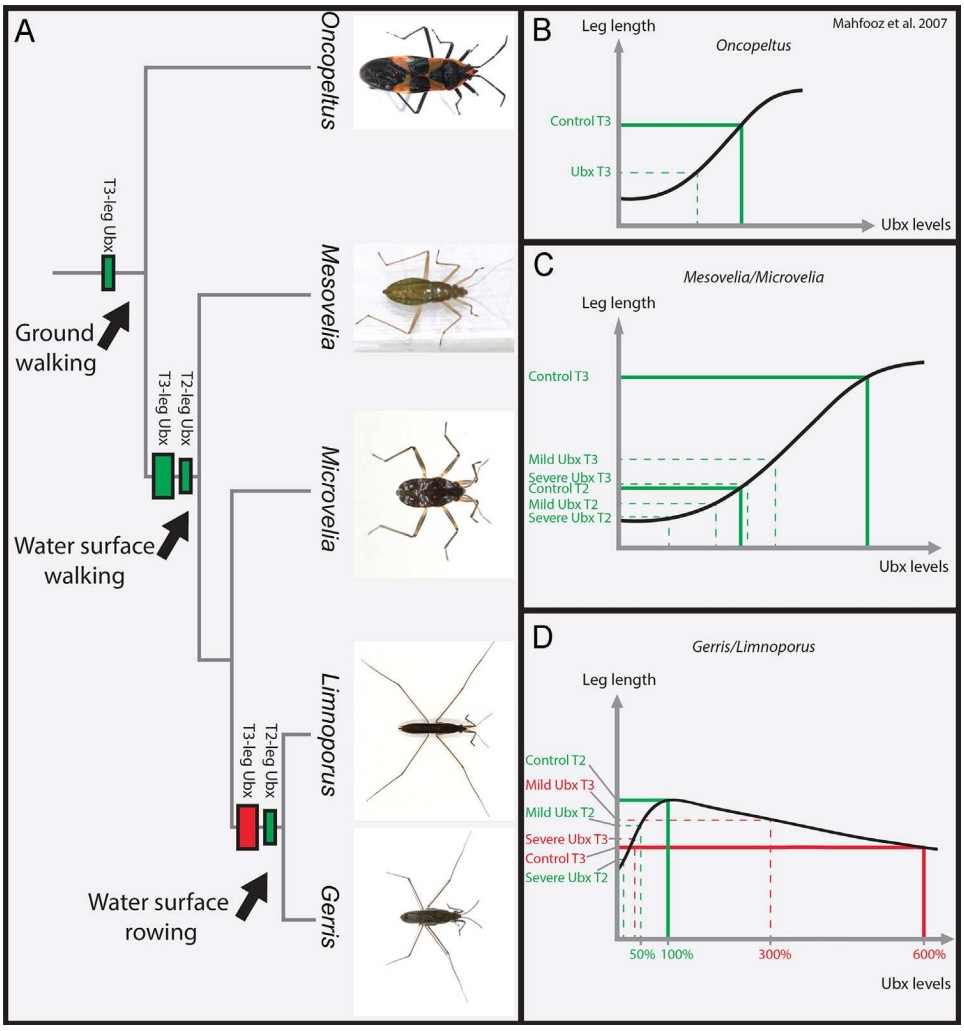


Fig. 7. Evolution of Ubx levels of expression and its role in shaping adaptive leg allometry. (A) Phylogenetic relationships between terrestrial Heteroptera, (*Oncopeltus*), water surface walking (*Mesovelia* and *Microvelia*), and water surface rowing (*Gerris* and *Limnoporus*) semi-aquatic insects. Narrow green rectangle represents the ancestral state where Ubx expression was confined to T3-legs and functioned to lengthen them. Large green rectangle in T3- and narrow green rectangle in T2-legs represent the change at the base of the semi-aquatic insects consisting of increased Ubx levels in T3- and deployment of Ubx in T2-legs. At this stage, Ubx lengthens both T2- and T3-legs. Large red rectangle indicates the emergence of T3-leg sensitivity to high Ubx levels, which now shortens T3-legs. This characteristic distinguishes surface-rowing semi-aquatic insects from their surface-walking relatives. (B) Ubx role and RNAi response in *Oncopeltus* T3-legs. (C) Ubx role in lengthening both T2- and T3-legs is linear in the basally branching *Mesovelia* and the derived *Microvelia*, both of which are surface walkers. (D) Model for the dose-dependent role of Ubx in increasing/decreasing leg length. Percentages of Ubx levels were extracted from a comparative transciptome between T2- and T3-legs in normal and *Ubx* RNAi-treatments. Approximate leg length values were taken from Fig. 3D. Levels of Ubx in T2-legs are used as a reference at 100%, and levels in T3-legs are at least six times higher (600%) based on the comparative transcriptome. At highest levels Ubx shortens whereas at lower levels Ubx lengthens the legs.

early gradient and gap genes can act as activators or repressors depending on their concentration (Sauer and Jackle, 1991; Schulz and Tautz, 1994). At low concentration, the transcription factor Kruppel (Kr), for example, remains a monomer and activates its target DNA, whereas at high concentration Kr forms homodimers and represses the same target (Sauer and Jackle, 1993). Many Hox proteins, including sex combs reduced (Papadopoulos et al., 2012), Abd-A (Hudry et al., 2011), and Ubx itself (Samir Merabet, personal communication) are known to form homodimers. It is therefore possible that a mechanism, similar to that of Kr, involving the formation of homodimers at high Ubx concentration results in the opposite effect on leg length relative to low Ubx concentration. The formation of heterodimers between Ubx and its known co-factors might also explain these differences in tissue response to Ubx levels. Our results, however, provide evidence that exclude possible heterodimer formation between Ubx and the two known co-factors Exd and Hth in regulating leg length. Both Exd and Hth patterns of expression are confined to the proximal segments, reminiscent to their pattern of expression in fly appendages (Casares and Mann, 2000). The non-linear bimodal response to *Ubx* RNAi concerns distal leg segments where these two co-factors are not expressed, and therefore, inconsistent with a possible cooperation between these factors and Ubx in this tissue. This is also the case in fly leg development, where Ubx acts without any cooperation with Exd and Hth (Casares and Mann, 1998; Galant et al., 2002). In addition to co-factors, Ubx is known to mediate its role in transcriptional regulation through cooperation with a number of transcription factors called Hox collaborators (Mann et al., 2009). These collaborators provide direct input into a Hox-regulated element without necessarily binding the element in cooperation with the Hox gene (Mann et al., 2009). It is conceivable that such collaborators may be part of the dose-dependent differential response of each leg to Ubx. Other additional mechanisms may be in play to mediate this dose-dependent positive-to-negative shift in Ubx effect. The degree of occupancy of Ubx binding sites within its target sequences, dependent on Ubx protein levels, might dictate the nature of transcriptional regulation of these target genes, as is the case in flies (Galant and Carroll, 2002; Hersh and Carroll, 2005).

Finally, it is possible that differences in Ubx affinity to its various targets may explain this bimodal response of the two legs to Ubx dose. Ubx may have high affinity to growth-promoting and low affinity to growth-suppressing targets. In this case, at low concentration, Ubx binding to the high-affinity growth-promoting partners prevails, and the outcome of this interaction increases cell proliferation, as is the case of T2-legs in *Limnoporus*. On the other hand, when Ubx levels exceed a certain dose the excess of Ubx protein is now free to interact with low affinity growth-suppressing partners. This type of interaction results in reduced cell proliferation and shortened legs, as is the case of *Limnoporus* T3. This interpretation is supported by the observation, in the fly *Drosophila*, that Ubx can act both as an activator and as a repressor, and that some target genes respond only at high Ubx concentration (Feinstein et al., 1995; Pavlopoulos and Akam, 2011).

The difference in length between T2 and T3-legs is primarily grounded in cell number, reflecting increased cell proliferation in T2- compared to T3-legs (in preparation). This is in contrast to the size difference in *Drosophila* halters compared to wing, which is mostly grounded on cell shape changes rather than cell proliferation or volume (Roch and Akam, 2000). Consistent with these cellular differences between T2- and T3-legs, we found in our comparative transcriptome that Ubx regulates a large number of genes involved in various cellular processes. These include cell cycle, cell death, cell differentiation, response to stress, and catabolic processes (not shown). Future work will examine the cellular nature of Ubx effect on leg length and dissect the role of its target genes in modulating the differential rate of cell proliferation between T2- and T3-legs.

Ubx, segment identity, and adaptive morphological evolution

It is striking that such substantive changes in *Ubx* regulation and requirement in shaping leg allometry have emerged throughout the evolution of the semi-aquatic insects without altering the spatial expression of major players of early pattern formation. The role of the transcription factors and signaling molecules we examined in growth and patterning is well documented (Baena-Lopez et al., 2012, 2003; Crickmore and Mann, 2006; Estella et al., 2008; Theisen et al., 1996; Zecca and Struhl, 2002). Leg patterning genes may contribute to the dramatic growth of *Limnoporus* legs and could have constituted a target of Ubx during the evolution of its novel role in reversing the length of T2- relative to T3-legs. It is possible that these genes generate the ground state of leg growth during development, and that Ubx comes to modulate this ground state and generate the adaptive relative leg length required for surface rowing. The interactions through which Ubx mediates this fine-tuning of relative leg length may involve a battery of genes that are required for growth but not pattern formation. It seems that there has been a change in the landscape of Ubx targets in repressing tissue growth, at least when this role is compared between *Limnoporus* T3-legs and *Drosophila* halteres where Ubx exerts a similar effect (Crickmore and Mann, 2006; Pavlopoulos and Akam, 2011; Roch and Akam, 2000). This observation raises an important question: why is Ubx's regulation of the spatial expression of this set of genes so subtle, if any? A plausible explanation is the early timing at which Ubx intervenes to set leg allometry in water striders. Embryogenesis is too early for the evolution of profound rearrangements in the network composed of major regulators of pattern formation (Davidson and Erwin, 2010; Wilkins, 2002). Alteration of these networks, by the novel changes in Ubx role and regulation, would have had disruptive effects that would not have been tolerated by natural selection (Davidson and Erwin, 2010; Wilkins, 2002). Throughout the semi-aquatic insects, either with the ancestral or the derived state of relative leg length (Fig. 7), the change in Ubx expression has not introduced a drastic change in morphology. Our results clearly outline yet another example, in addition to Ubx role in the crustacean *Parhyale* (Pavlopoulos et al., 2009), that the evolutionary transition to the derived state of relative leg length, adapted for water surface rowing, involved a gradual process that started with an initial change in Ubx expression already in the common ancestor of the semi-aquatic insects (Fig. 7). These observations come to further support the view that Ubx acts as a micromanager in establishing organ morphogenesis during development and evolution (Akam, 1998b, 1998c; Castelli-Gair and Akam, 1995; Stern, 1998, 2003). Hox genes may not only act as master binary switches of segment identity, but can rather accumulate changes throughout evolution, the sum of which leads to incremental changes in segment morphology, 'without the need to invoke the selective advantage of hopeful monsters (Akam, 1998b).' A similar observation involving the effect of Ubx levels on morphological diversification, at the micro-evolutionary scale, is known in flies. Across distinct fly species, the presence of trichomes on the femur is associated with low levels of Ubx, whereas the absence of trichomes is associated with high levels of Ubx (Stern, 1998). Our data show that, at a macro-evolutionary scale, this flexibility in the regulation, function, and most likely diversity of targets, can generate the remarkable phenotypic diversity that selection acts upon within specific ecological environments.

Acknowledgments

We thank Ehab Abouheif for hosting AK and PR in his lab during earlier steps of this project, Michalis Averof, Samir Merabet, Dali Ma, and Emilia Santos, and three anonymous reviewers for comments on the manuscript. We thank Kelly Langlais for technical support and E. Furlong, R. White, and *Developmental Studies Hybridoma Bank* for providing the anti-Mef2, anti-UbdA and Exd antibodies respectively, all of which cross-react in various semi-aquatic insect species. We thank John Colbourne for help with generating *Limnoporus* full transcriptome used as a reference, and Lois Taulelle and Hervé Gilquin for providing access to computing resources in the *Pôle Scientifique de Modélisation Numérique* (PSMN) at the ENS Lyon. Scanning electron microscopy was performed in the facility of the *Centre Technologique des Microstructures* at the Université Claude Bernard Lyon1. This work was funded through an ATIP-Avenir from CNRS and an ERC-CoG # 616346 to AK and a PhD Fellowship from Université Claude Bernard Lyon1 (0103/2010) to PR.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ydbio.2014.05.021.

References

Abzhanov, A., Kaufman, T.C., 2000. Homologs of *Drosophila* appendage genes in the patterning of arthropod limbs. Dev. Biol. 227, 673–689.
Abzhanov, A., Protas, M., Grant, B.R., Grant, P.R., Tabin, C.J., 2004. Bmp4 and morphological variation of beaks in Darwin's finches. Science 305, 1462–1465.
Akam, M., 1998a. Hox genes in arthropod development and evolution. Biol. Bull. 195, 373–374.
Akam, M., 1998b. Hox genes, homeosis and the evolution of segment identity: no need for hopeless monsters. Int. J. Dev. Biol. 42, 445–451.
Akam, M., 1998c. Hox genes: from master genes to micromanagers. Curr. Biol. 8, R676–R678.
Andersen, N.M., 1976. A Comparative Study of Locomotion on the Water Surface in Semiaquatic Bugs (Insecta, Hemiptera, Gerromorpha) (Vidensk. Meddr dansk naturh. Foren), pp. 337–396.
Andersen, N.M., 1982. The Semiaquatic Bugs (Hemiptera: Gerromorpha). Scandinavian Science Press Ltd., Klampenborg, Denmark.
Angelini, D.R., Kaufman, T.C., 2004. Functional analyses in the hemipteran *Oncopeltus fasciatus* reveal conserved and derived aspects of appendage patterning in insects. Dev. Biol. 271, 306–321.

Angelini, D.R., Liu, P.Z., Hughes, C.L., Kaufman, T.C., 2005. Hox gene function and interaction in the milkweed bug *Oncopeltus fasciatus* (Hemiptera). Dev. Biol. 287, 440–455.
Baena-Lopez, L.A., Nojima, H., Vincent, J.P., 2012. Integration of morphogen signalling within the growth regulatory network. Curr. Opin. Cell Biol. 24, 166–172.
Baena-Lopez, L.A., Pastor-Pareja, J.C., Resino, J., 2003. Wg and Egfr signalling antagonise the development of the peripodial epithelium in *Drosophila* wing discs. Development 130, 6497–6506.
Baker, N.E., 1988. Transcription of the segment-polarity gene wingless in the imaginal discs of *Drosophila*, and the phenotype of a pupal-lethal wg mutation. Development 102, 489–497.
Bronikowski, A.M., 2000. Experimental evidence for the adaptive evolution of growth rate in the garter snake *Thamnophis elegans*. Evolution 54, 1760–1767.
Carroll, S.B., 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134, 25–36.
Casares, F., Mann, R.S., 1998. Control of antennal versus leg development in *Drosophila*. Nature 392, 723–726.
Casares, F., Mann, R.S., 2000. A dual role for homothorax in inhibiting wing blade development and specifying proximal wing identities in *Drosophila*. Development 127, 1499–1508.
Castelli-Gair, J., Akam, M., 1995. How the Hox gene Ultrabithorax specifies two different segments: the significance of spatial and temporal regulation within metameres. Development 121, 2973–2982.
Crickmore, M.A., Mann, R.S., 2006. Hox control of organ size by regulation of morphogen production and mobility. Science 313, 63–68.
Crickmore, M.A., Mann, R.S., 2008. The control of size in animals: insights from selector genes. Bioessays 30, 843–853.
Damgaard, J., 2008. Phylogeny of the semiaquatic bugs (Hemiptera—Heteroptera, Gerromorpha). Insect Syst. Evol. 39, 30.
Damgaard, J., Andersen, N.M., Meier, R., 2005. Combining molecular and morphological analyses of water strider phylogeny (Hemiptera—Heteroptera, Gerromorpha): effects of alignment and taxon sampling. Syst. Entomol. 30, 289–309.
Davidson, E.H., Erwin, D.H., 2010. Evolutionary innovation and stability in animal gene networks. J. Exp. Zool. B: Mol. Dev. Evol. 314, 182–186.
Duboule, D., Tarchini, B., Zakany, J., Kmita, M., 2007. Tinkering with constraints in the evolution of the vertebrate limb anterior-posterior polarity. Novartis. Found. Symp. 284, 130–137 (discussion 138–141, 158–163).
Emlen, D.J., Warren, I.A., Johns, A., Dworkin, I., Lavine, L.C., 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. Science 337, 860–864.
Estella, C., Mann, R.S., 2008. Logic of Wg and Dpp induction of distal and medial fates in the *Drosophila* leg. Development 135, 627–636.
Estella, C., McKay, D.J., Mann, R.S., 2008. Molecular integration of wingless, decapentaplegic, and autoregulatory inputs into distalless during *Drosophila* leg development. Dev. Cell 14, 86–96.
Feinstein, P.G., Kornfeld, K., Hogness, D.S., Mann, R.S., 1995. Identification of homeotic target genes in *Drosophila* melanogaster including nervy, a proto-oncogene homologue. Genetics 140, 573–586.
Galant, R., Carroll, S.B., 2002. Evolution of a transcriptional repression domain in an insect Hox protein. Nature 415, 910–913.
Galant, R., Walsh, C.M., Carroll, S.B., 2002. Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. Development 129, 3115–3126.
Gibson, G., 1996. Epistasis and pleiotropy as natural properties of transcriptional regulation. Theor. Popul. Biol. 49, 58–89.
Grossmann, D., Prpic, N.M., 2012. Egfr signaling regulates distal as well as medial fate in the embryonic leg of *Tribolium castaneum*. Dev. Biol. 370, 264–272.
Hersh, B.M., Carroll, S.B., 2005. Direct regulation of knot gene expression by Ultrabithorax and the evolution of cis-regulatory elements in *Drosophila*. Development 132, 1567–1577.
Hu, D., Bush, J.W., 2010. The hydrodynamics of water-walking arthropods. J. Fluid Mech. 644, 5–33.
Hu, D.L., Chan, B., Bush, J.W., 2003. The hydrodynamics of water strider locomotion. Nature 424, 663–666.
Hudry, B., Viala, S., Graba, Y., Merabet, S., 2011. Visualization of protein interactions in living *Drosophila* embryos by the bimolecular fluorescence complementation assay. BMC Biol. 9, 5.
Jan, L.Y., Jan, Y.N., 1982. Antibodies to horseradish peroxidase as specific neuronal markers in *Drosophila* and in grasshopper embryos. Proc. Natl. Acad. Sci. USA 79, 2700–2704.
Jaubert-Possamai, S., Le Trionnaire, G., Bonhomme, J., Christophides, G.K., Rispe, C., Tagu, D., 2007. Gene knockdown by RNAi in the pea aphid *Acyrthosiphon pisum*. BMC Biotechnol. 7, 63.
Joshi, R., Passner, J.M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M.A., Jacob, V., Aggarwal, A.K., Honig, B., Mann, R.S., 2007. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell 131, 530–543.
Kelsh, R., Weinzierl, R.O., White, R.A., Akam, M., 1994. Homeotic gene expression in the locust Schistocerca: an antibody that detects conserved epitopes in Ultrabithorax and abdominal-A proteins. Dev. Genet. 15, 19–31.
Khila, A., Abouheif, E., Rowe, L., 2009. Evolution of a novel appendage ground plan in water striders is driven by changes in the Hox gene Ultrabithorax. PLoS Genet. 5, e1000583.
Khila, A., Abouheif, E., Rowe, L., 2012. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. Science 336, 585–589.
Khila, A., Abouheif, E., Rowe, L., 2014. Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. Evolution 〈http://dx.doi.org/10.1111/evo.12444〉 (published online).
Kitzmann, P., Schwirz, J., Schmitt-Engel, C., Bucher, G., 2013. RNAi phenotypes are influenced by the genetic background of the injected strain. BMC Genomics 14, 5.
Loehlin, D.W., Werren, J.H., 2012. Evolution of shape by multiple regulatory changes to a growth gene. Science 335, 943–947.
Mahfooz, N., Turchyn, N., Mihajlovic, M., Hrycaj, S., Popadic, A., 2007. Ubx regulates differential enlargement and diversification of insect hind legs. PloS One 2, e866.
Mann, R.S., Lelli, K.M., Joshi, R., 2009. Hox specificity unique roles for cofactors and collaborators. Curr. Top. Dev. Biol. 88, 63–101.
Moczek, A.P., Rose, D.J., 2009. Differential recruitment of limb patterning genes during development and diversification of beetle horns. Proc. Natl. Acad. Sci. USA 106, 8992–8997.
Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628.
Nijhout, H.F., Grunert, L.W., 2010. The cellular and physiological mechanism of wing-body scaling in *Manduca sexta*. Science 330, 1693–1695.
O'Farrell, P.H., 2003. How metazoans reach their full size: the natural history of bigness. In: Michael Hall, M.R., Thomas, George (Eds.), Cell Growth: Control of Cell Size. Cold Spring Harbor Press, NY.
Panganiban, G., Nagy, L., Carroll, S.B., 1994. The role of the distal-less gene in the development and evolution of insect limbs. Curr. Biol. 4, 671–675.
Papadopoulos, D.K., Skouloudaki, K., Adachi, Y., Samakovlis, C., Gehring, W.J., 2012. Dimer formation via the homeodomain is required for function and specificity of sex combs reduced in *Drosophila*. Dev. Biol. 367, 78–89.
Paschinger, K., Rendic, D., Wilson, I.B., 2009. Revealing the anti-HRP epitope in *Drosophila* and Caenorhabditis. Glycoconj. J. 26, 385–395.
Pavlopoulos, A., Akam, M., 2011. Hox gene Ultrabithorax regulates distinct sets of target genes at successive stages of *Drosophila* haltere morphogenesis. Proc. Natl. Acad. Sci. USA 108, 2855–2860.
Pavlopoulos, A., Kontarakis, Z., Liubicich, D.M., Serano, J.M., Akam, M., Patel, N.H., Averof, M., 2009. Probing the evolution of appendage specialization by Hox gene misexpression in an emerging model crustacean. Proc. Natl. Acad. Sci. USA 106, 13897–13902.
Prpic, N.M., Janssen, R., Wigand, B., Klingler, M., Damen, W.G., 2003. Gene expression in spider appendages reveals reversal of exd/hth spatial specificity, altered leg gap gene dynamics, and suggests divergent distal morphogen signaling. Dev. Biol. 264, 119–140.
Prpic, N.M., Wigand, B., Damen, W.G., Klingler, M., 2001. Expression of dachshund in wild-type and distal-less mutant Tribolium corroborates serial homologies in insect appendages. Dev. Genes Evol. 211, 467–477.
Roch, F., Akam, M., 2000. Ultrabithorax and the control of cell morphology in *Drosophila* halteres. Development 127, 97–107.
Rogers, B.T., Peterson, M.D., Kaufman, T.C., 1997. Evolution of the insect body plan as revealed by the sex combs reduced expression pattern. Development 124, 149–157.
Sandmann, T., Jensen, L.J., Jakobsen, J.S., Karzynski, M.M., Eichenlaub, M.P., Bork, P., Furlong, E.E., 2006. A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. Dev. Cell 10, 797–807.
Sauer, F., Jackle, H., 1991. Concentration-dependent transcriptional activation or repression by Kruppel from a single binding site. Nature 353, 563–566.
Sauer, F., Jackle, H., 1993. Dimerization and the control of transcription by Kruppel. Nature 364, 454–457.
Schulz, C., Tautz, D., 1994. Autonomous concentration-dependent activation and repression of Kruppel by hunchback in the *Drosophila* embryo. Development 120, 3043–3049.
Stern, D.L., 1998. A role of Ultrabithorax in morphological differences between *Drosophila* species. Nature 396, 463–466.
Stern, D.L., 2003. The Hox gene Ultrabithorax modulates the shape and size of the third leg of *Drosophila* by influencing diverse mechanisms. Dev. Biol. 256, 355–366.
Stern, D.L., Emlen, D.J., 1999. The developmental basis for allometry in insects. Development 126, 1091–1101.
Stillwell, R.C., Dworkin, I., Shingleton, A.W., Frankino, W.A., 2011. Experimental manipulation of body size to estimate morphological scaling relationships in *Drosophila*. J. Vis. Exp. 56, 1–4.
Struhl, G., 1982. Genes controlling segmental specification in the *Drosophila* thorax. Proc. Natl. Acad. Sci. USA 79, 7380–7384.
Theisen, H., Haerry, T.E., O'Connor, M.B., Marsh, J.L., 1996. Developmental territories created by mutual antagonism between wingless and decapentaplegic. Development 122, 3939–3948.
Tseng, M., Rowe, L., 1999. Sexual dimorphism and allometry in the giant water strider, *Gigantometra gigas*. Can. J. Zool. 77, 923–929.
White, R.A.H., Wilcox, M., 1984. Protein products of the bithorax complex in *Drosophila*. Cell 39, 163–171.
Wilkins, A.S., 2002. The Evolution of Developmental Pathways. Sinauer Associates, Inc., Sunderland.
Zecca, M., Struhl, G., 2002. Control of growth and patterning of the *Drosophila* wing imaginal disc by EGFR-mediated signaling. Development 129, 1369–1376.

Supplementary online material for:

Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait

Peter Refki[1, 2], David Armisen[1], Antonin Crumiere[1, 2], Severine Viala[1], and Abderrahman Khila[1*]

[1] Institut de Génomique Fonctionnelle de Lyon, CNRS-UMR5242, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364, Lyon Cedex 07, France

[2] Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex

[*] Corresponding author: abderrahman.khila@ens-lyon.fr

Supplementary tables:
Table S1: primers used to clone the various developmental genes in *Limnoporus,* and the respective Genbank accession numbers assigned to the sequences of these genes.

Gene	Forward primer	Reverse primer	GenBank Accession #
dll	Reference (Khila et al., 2012)	Reference (Khila et al., 2012)	Reference (Khila et al., 2012)
wg	CGGCATTCATMTAYGCRA TMACMAG	CCGCARCACATYARRTCGCA YCC	KF630597
hh	AAGACGAMGARGGMAG MGGWGCCGAT	TARTWGACCCAATCGAAWC CKGCTTC	KF630595
dac	CTACGCACCTCTCACAGC AAATCAA	CTGTTCGTCGATAAGCTGCC TTTCCA	KF630592
dpp	TCTACTACAGCTACT CGG CATGCCCA	TCTGCGTCGAATTTCTCCTC CTGCA	KF630593
egfr	AAACGCACTTGCCAACGA CTCAGAGTT	AGTGGTAGGGTTGTATCGCT GCATTG	KF630594
hth	GACAAGGATTCCATCTAC GGGCATC	TGCCGTTAGTATCGTCGTCT TCTTCG	KF630596
Ubx	AATGAACTCGTACTTCGA GCAGACGG	CTGTTCGTTGAGCTCCTTTAT CGCTTG	AFM55060

Table S2: Primers flanked by T7 promoter (small letters) used to construct double stranded RNA.

Gene	Forward primer	Reverse primer
wg	taatacgactcactatagggagaccacCGGC ATTCATMTATGCRATAACCAG	taatacgactcactatagggagaccacTTGTATCCCTA GGCT CGGGTTGCGTT
hh	taatacgactcactatagggagaccacAAGA CGAMGARGGCAGAGGWGCC GAT	taatacgactcactatagggagaccacTARTWGACCCA ATCGAAWCCGGCTTC
Ubx	taatacgactcactatagggagaccacATGA ATTCTTACTTTGAGCAGACGG GT	taatacgactcactatagggagaccacCTGCTATAGCC ATCCAGGGGTAGAA

Table S3: Values of R^2 indicating the best fit between the centered second order polynomial (quadratic) model and the linear regression fit model for each dataset plotted in Figure 4.

Species	Leg segment	Centered second order polynomial model R^2	Linear regression fit model R^2
Limnoporus: N=92	Entire leg	0,3457	0,1197
	Tarsus	0,5867	0,1527
	Tibia	0,2964	0,2870
	Femur	0,6162	0,6112
Microvelia: N=31	Entire leg	0,9204	0,9199
	Tarsus	0.8326	0.8225
	Tibia	0,8671	0,8524
	Femur	0.7972	0.7972
Mesovelia: N=56	Entire leg	0,7007	0,7004
	Tarsus	0.6555	0.6319
	Tibia	0,5948	0,5875
	Femur	0.5204	0.5179

Table S4: Penetrance of Ubx protein depletion in the legs as a result of various strengths of *Ubx* RNAi in *Limnoporus*.

		UbdA detected in T3 legs	UbdA undetected in T3 legs
Ubx RNAi 0.1µg	# Embryos	48	25
	Percentage	65.75%	34.25%
Ubx RNAi 1µg	# Embryos	13	20
	Percentage	39.39%	60.61%
Ubx RNAi 10µg	# Embryos	16	45
	Percentage	26.23%	73.77%

Table S5: Changes in leg length (by percentage) caused by the various strengths of *Ubx* RNAi in *Limnoporus*.

	T1-leg			T2-leg			T3-leg		
	Tarsus	Tibia	Femur	Tarsus	Tibia	Femur	Tarsus	Tibia	Femur
ds-*Ubx* 0.1µg	-0,7%	-6,5%	4,9%	-38,6%	-16,3%	-2,2%	35,8%	38,3%	-2%
ds-*Ubx* 1µg	-3,6%	-5,2%	5%	-49,9%	-20,8%	-8,3%	17,5%	37,5%	-5,6%
ds-*Ubx* 10µg	-5,3%	2,3%	-0,4%	-58,7%	-24,8%	-11,9%	10,1%	43,1%	-5,5%

Table S6: Changes of the ratios of various leg segments in controls and various strengths of *Ubx* RNAi in *Limnoporus*.

A:	**Controls**		
	T3-Tarsus/T2-Tarsus	T3-Tibia/T2-Tibia	T3-Femur/T2-Femur
Control 1	0,50	0,59	0,84
Control 2	0,53	0,57	0,94
Control 3	0,52	0,60	0,82
Control 4	0,53	0,57	0,94
Control 5	0,54	0,56	0,97
Average	0,524	0,58	0,90
B:	**Mildest *Ubx* RNAi**		
	Ubx-Tarsus/Control-Tarsus	*Ubx*-Tibia/ Control-Tibia	*Ubx*-Femur/ Control-Femur
Mildest 1	0,71	0,90	0,87
Mildest 2	0,77	0,76	0,87
Mildest 3	0,70	0,80	0,94
Mildest 4	0,70	0,96	0,8
Mildest 5	0,65	0,83	0,83
Average	0,70	0,85	0,86
C:	**Strongest *Ubx* RNAi**		
	Ubx-Tarsus/Control-Tarsus	*Ubx*-Tibia/ Control-Tibia	*Ubx*-Femur/ Control-Femur
Strongest 1	0,56	0,71	0,75
Strongest 2	0,53	0,84	0,62
Strongest 3	0,55	0,77	0,76
Strongest 4	0,63	0,73	0,74
Strongest 5	0,59	0,84	0,70
Average	0,57	0,78	0,71

Table S7: Counts of the responses of various attributes in the legs of *Microvelia* to various strengths of *Ubx* RNAi.

Yfp dsRNA injected controls								
T1 2 combs	**T2 comb**	**T2 bristle**	**T3 comb**	**T3 Bristle**	**Corrected T2 length**	**Corrected T3 length**	**T2 % Change**	**T3 % Change**
Yes	Yes	NO	NO	Yes	81,41	103,32	-	-
Yes	Yes	NO	NO	Yes	89,55	111,65	-	-
Yes	Yes	NO	NO	Yes	79,41	98,77	-	-
Yes	Yes	NO	NO	Yes	92,44	110,28	-	-
Yes	Yes	NO	NO	Yes	83,28	107,66	-	-
Mildest *Ubx* RNAi sample								
T1 2 combs	**T2 comb**	**T2 bristle**	**T3 comb**	**T3 Bristle**	**Corrected T2 length**	**Corrected T3 length**	**T2 % change**	**T3 % change**
Yes	Yes	NO	Yes	Yes	83,08	102,26	-4,72	-6,62
Yes	Yes	NO	NO	Yes	82,44	100,83	-5,45	-7,92
Yes	Yes	NO	NO	Yes	76,73	92,43	-12,01	-15,59
Yes	Yes	NO	Yes	Yes	74,95	91,90	-14,04	-16,08
Yes	Yes	NO	NO	Yes (small)	74,91	90,29	-14,09	-17,55
Strongest *Ubx* RNAi sample								
T1 2 combs	**T2 comb**	**T2 bristle**	**T3 comb**	**T3 Bristle**	**Corrected T2 length**	**Corrected T3 length**	**T2 % change**	**T3 % change**
Yes	Yes	NO	Yes	Yes	60,45	70,84	-30,67	-35,31
Yes	Yes	NO	Yes	Yes (small)	65,48	72,05	-24,91	-34,21
Yes	Yes	NO	Yes	Yes	65,24	74,55	-25,17	-31,92
Yes	Yes	NO	Yes	Yes	66,85	72,58	-23,33	-33,72
Yes	Yes	NO	Yes	Yes	67,29	74,13	-22,82	-32,30

Table S8: Changes in the levels of expression of the various leg patterning genes in T2- and T3-legs as a result of *Ubx* depletion, revealed by their RPKM.

	WT T2-leg RPKM	*Ubx* T2-leg RPKM	% Change	WT T3-leg RPKM	*Ubx* T3-leg RPKM	% Change
dpp	23,3	28,04	20,34	23,91	31,74	32,75
hh	211,07	340,56	61,35	223,41	349,86	56,60
egfr	28,63	45,08	57,46	30,5	51,64	69,31
dll	66,63	89,18	33,84	69,1	89,23	29,13
dac	15,08	22,21	47,28	18,19	24,94	37,11
hth	3,19	3,79	18,81	2,85	3,76	31,93
exd	36,27	67,58	86,32	32,28	68,08	110,90
wg	5,55	10,08	81,62	5,44	10,82	98,90

Supplementary figures:

Fig. S1: Ubx expression compared between legs and a subset of neurons on T2 and T3. In the legs, Ubx levels are much higher in T3 compared to T2 as revealed by the FP6.87 antibody. However, Ubx levels seem to be comparable between neurons located on T2 and T3. This suggests that the requirement for higher Ubx levels is a characteristic of the legs.

Fig. S2: Dynamics of leg growth during *Limnoporus* development. (A) By the end of *Limnoporus* embryogenesis, the differences in length between T2- and T3-legs are already established. **(B)** In the newly hatched first instar larvae, T2-legs are much longer than the T3-legs (see Figure 3E for measurements of the legs). **(C)** There is further differential growth between T2- and T3-legs during nymphal development, characteristic to this species, which narrows length differences between the two legs in the adult.

Fig. S3: The combs on the legs are unaffected by *Ubx* RNAi in *Limnoporus*. Just like *Mesovelia* and *Microvelia*, T1-legs have two combs and T2-legs have one comb (arrows in **A** and **B**). However, unlike *Mesovelia* and *Microvelia*, T3-legs of *Limnoporus* also have a comb (arrow in **C**). *Ubx* RNAi does not affect the combs in any of the legs (**D-F**; see also Khila et al. in press).

Fig. S4: (A) Pattern of axon projections forming the central nervous system scaffolds and peripheral nervous system, as revealed by the anti-HRP antibody, in control *Limnoporus* embryos. **(B)** *Ubx* RNAi does not seem to affect in any obvious way these axonal projections in the legs. Note the innervations on the ectopic leg that forms on the first abdominal segment as a result of *Ubx* RNAi (white arrow in **B**). **(C)** Pattern of expression of the Mef2 protein, a key regulator of muscle development, showing the distribution of muscle tissues in the legs of control *Limnoporus* embryos. **(D)** Mef2 expression in *Ubx* RNAi embryos showing no obvious defect in muscle tissues. Note the presence of the ectopic leg characteristic to *Ubx* RNAi on the first abdominal segment (arrow in **D**).

Fig. S5: *Ubx* RNAi causes a partial homeotic transformation of T3-legs to T2-legs. *Mesovelia* T1-legs bear one sex comb (black arrow in **A**) and one grooming comb (green arrow in **A**). T2-legs bear a single grooming comb (green arrow in **B**) in addition to a characteristic large bristle located close to the grooming comb (black arrowhead in **B**). T3-legs do not have any combs **(C)**, but rather bear a set of large bristles on the tibia that are different from the one on T2-legs (white arrowheads in **C**). *Ubx* RNAi does not affect the combs either in T1-legs (black and green arrow in **D**) or in T2-legs (green arrow in **E**), nor does it affect the large bristle on T2-legs (black arrowhead in **E**). T3-legs, however, now develop an ectopic comb (green arrow in **F**), but still retain their characteristic large bristles (white arrowheads in **F**).

Fig. S6: Ubx staining, as revealed by FP6.87 antibody, in *wg* and *hh* RNAi treatments. **(A)** Control embryo showing normal FP6.87 staining in the thorax and abdomen. This pattern is unchanged either after *wg* RNAi treatment **(B)**, or in *hh* RNAi treatment **(C)**, despite the clear patterning defects in these embryos.

Figure S1:

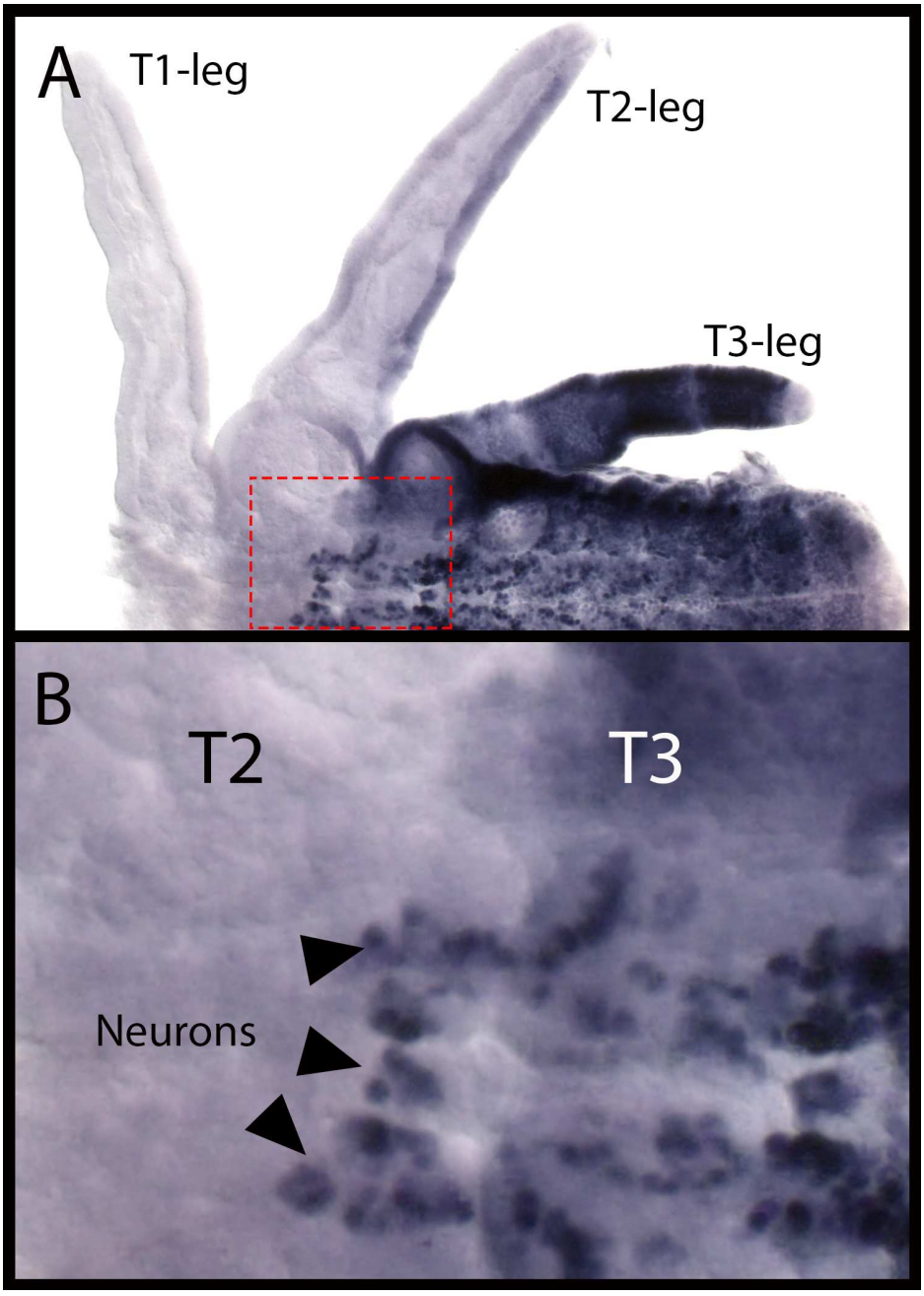


Figure S2:

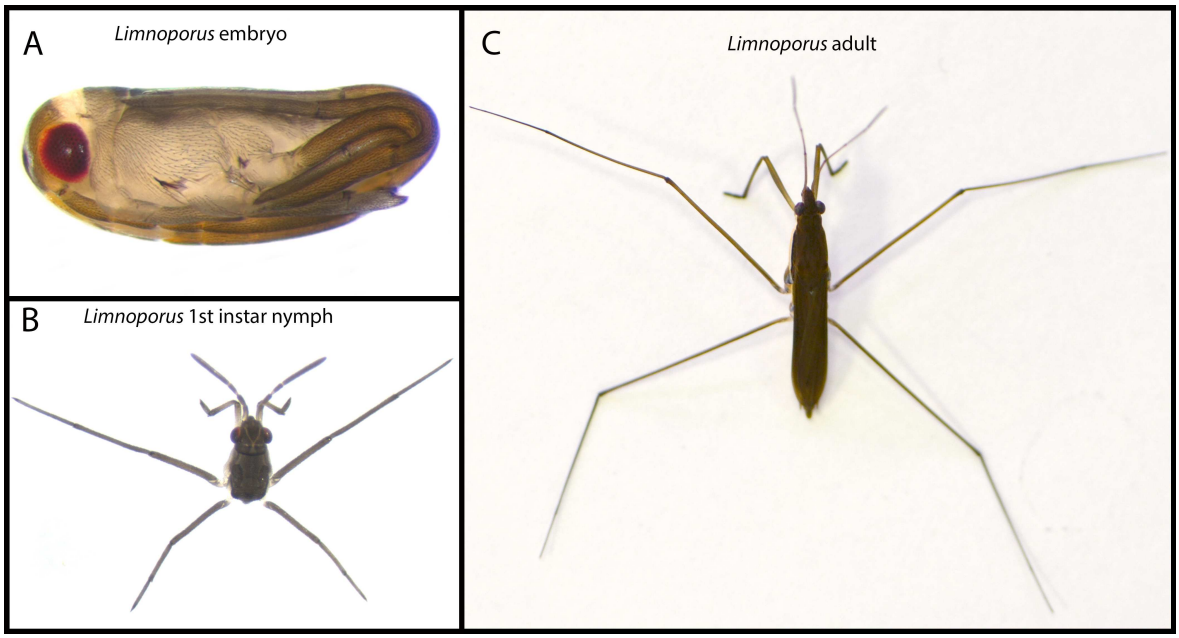

A
Limnoporus embryo
B
Limnoporus 1st instar nymph
C
Limnoporus adult

Figure S3:

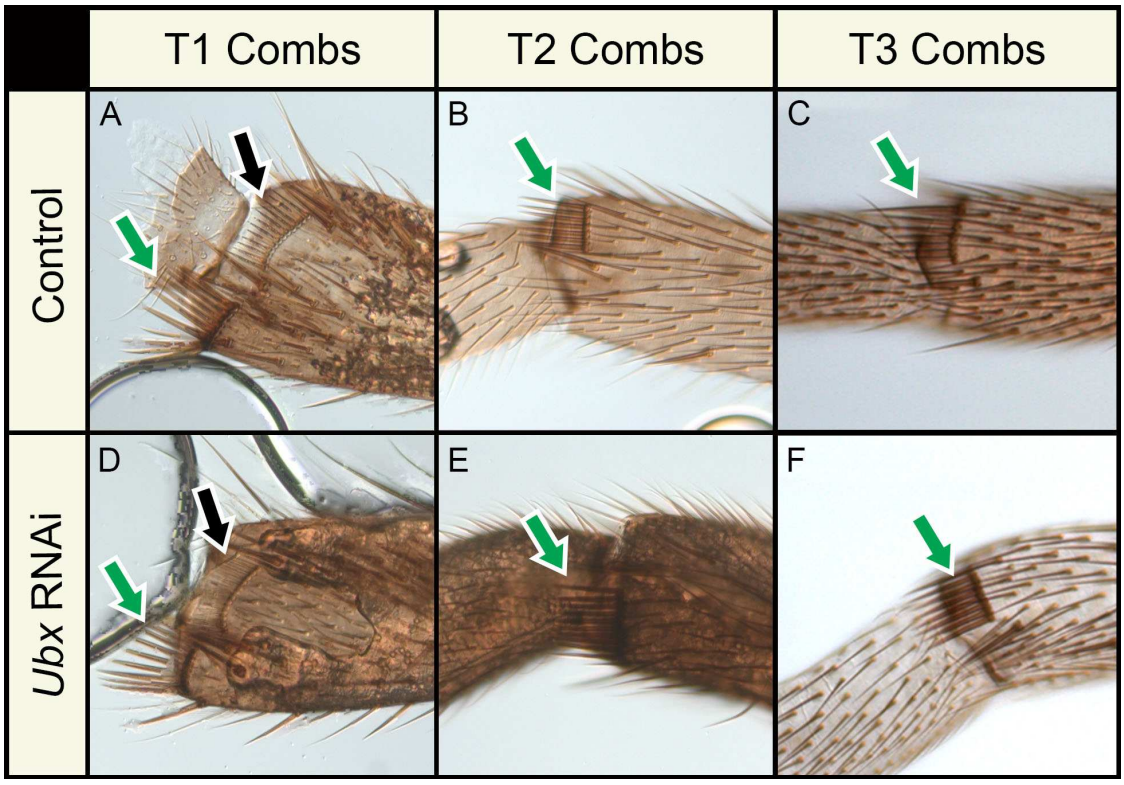


Figure S4:

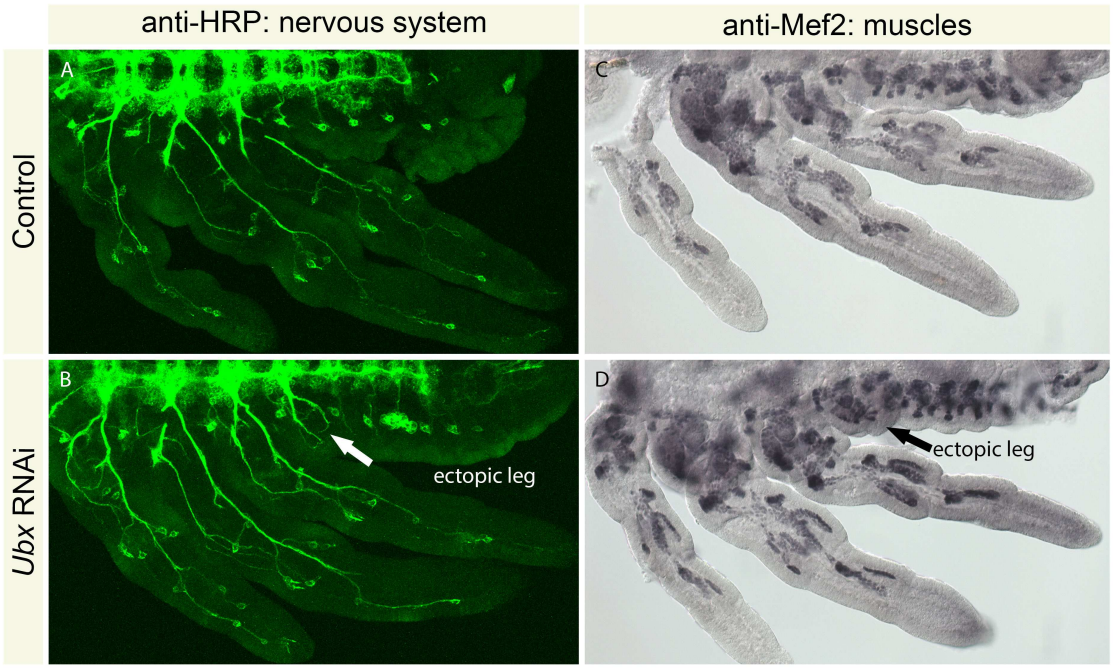

anti-HRP: nervous system
anti-Mef2: muscles
Control
Ubx RNAi
A
B
C
D
ectopic leg
ectopic leg

Figure S5:

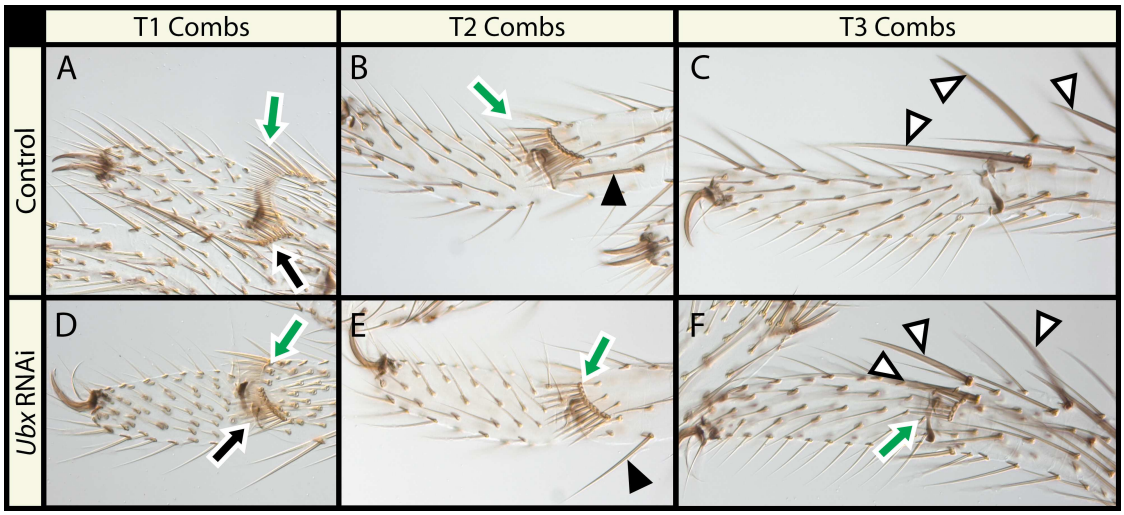
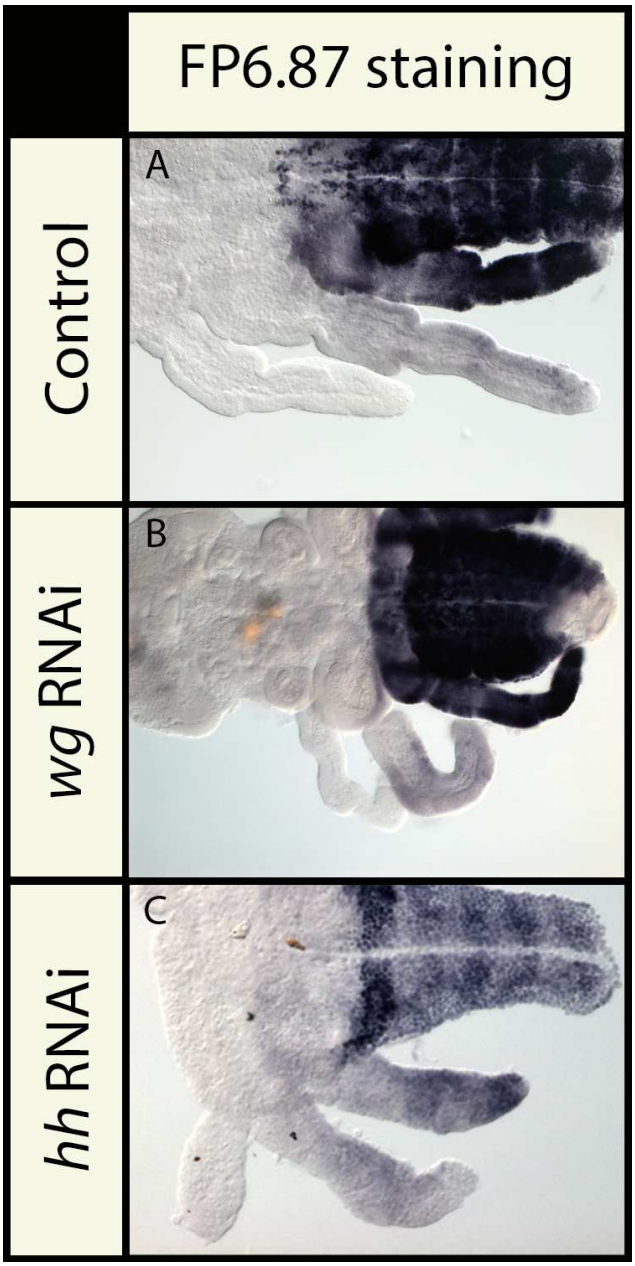


Figure S6:



Chapter 3

Predator strike shapes antipredator phenotype through new genetic interactions in water striders.

Paper published in *Nature communication*

Co-author

Synopsis of the paper:

The Hox gene *Ubx* participates to the establishment of the derived leg ground plan in derived species of the semiaquatic bugs with the T2-legs being longer than the T3-legs. Hox genes are known to control a large number of downstream targets however we did not know how *Ubx* was able to control differentially the growth of the T2- and T3-legs. The derived leg ground plan is characteristic to the species inhabiting in open water niches. In this new niche, water striders are exposed to new predators such as bottom striking fishes. In order to escape these attacks, water striders use a jumping behavior by pushing their T2-legs on the water surface to propel themselves in the air and escape. In this study we investigated the genetic mechanisms that allow the differential growth of T2- and T3-legs in combination with the ecological relevance of the trait by mimicking a natural situation of predation.

Summary of results:

Comparative transcriptomic analysis during embryonic development highlights the higher expression of the gene coding for the *gamma interferon inducible lysosomal thiol reductase* (*gilt*) in the T2-legs compared to the T3-legs. This gene was co-opted from the immune system and now has a role in the growth of the T2- and T3-legs in the derived water striders under the control of *Ubx*. Observation and quantification of predator-prey interactions show an absolute necessity for the jumping behavior by water striders to avoid predation by bottom striking fishes. A knockdown of *gilt* results in a decrease in leg length correlated with a decrease in the jumping performance.

Conclusion:

Our results indicate that the new predator-prey interactions in open water niche could have shaped new genetic interactions by the recruitment of an immune gene in developmental processes under the control of a Hox gene. This new interaction directly contributes to the derived leg ground plan and increases the performance of the water striders in escaping predators.

Author contributions:

Here are the contributions of the different authors (*first co-authors):

David Armisén*: transcriptome analysis, measurement of insect flight duration and fish strike duration, generated *gilt* RNAi knockdown, leg measurements, acquisition and analysis of jumping movies, statistical analysis, design of the study and redaction of the paper.

Peter Refki*: generated transcriptomes, in situ hybridization, generated *gilt* RNAi knockdown, leg measurements, acquisition and analysis of jumping movies, design of the study and redaction of the paper.

Antonin Crumière: cloning of *gilt* gene in the different species, generated *gilt* RNAi knockdown, leg measurements, acquisition and analysis of jumping movies.

Séverine Viala: in situ hybridization.

William Toubiana: in situ hybridization.

Abderrahman Khila: generated the transcriptome, design of the study and redaction of the paper.

nature COMMUNICATIONS

ARTICLE

Received 26 May 2015 | Accepted 23 Jul 2015 | Published 1 Sep 2015

DOI: 10.1038/ncomms9153 OPEN

Predator strike shapes antipredator phenotype through new genetic interactions in water striders

David Armisén[1,*], Peter Nagui Refki[1,*], Antonin Jean Johan Crumière[1], Séverine Viala[1], William Toubiana[1] & Abderrahman Khila[1]

How novel genetic interactions evolve, under what selective pressures, and how they shape adaptive traits is often unknown. Here we uncover behavioural and developmental genetic mechanisms that enable water striders to survive attacks by bottom-striking predators. Long midlegs, critical for antipredator strategy, are shaped through a lineage-specific interaction between the Hox protein Ultrabithorax (Ubx) and a new target gene called *gilt*. The differences in leg morphologies are established through modulation of *gilt* differential expression between mid and hindlegs under Ubx control. Furthermore, short-legged water striders, generated through *gilt* RNAi knockdown, exhibit reduced performance in predation tests. Therefore, the evolution of the new Ubx–*gilt* interaction contributes to shaping the legs that enable water striders to dodge predator strikes. These data show how divergent selection, associated with novel prey–predator interactions, can favour the evolution of new genetic interactions and drive adaptive evolution.

[1] Institut de Génomique Fonctionnelle de Lyon, CNRS-UMR5242, Ecole Normale Supérieure, Université Claude Bernard, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.K. (email: abderrahman.khila@ens-lyon.fr).

© 2015 Macmillan Publishers Limited. All rights reserved.

The discovery that distinct lineages share a similar genetic toolkit established that variations within genetic networks could shape the phenotype during development and evolution[1,2]. An important gap in understanding the origin of phenotypic diversity, however, resides in our poor knowledge about how selection can favour the emergence of novel developmental genetic interactions and ultimately shape phenotypic evolution[3–5]. Here we focus on the semi-aquatic bugs (Heteroptera, Gerromorpha) to study how divergent selection, associated with a new lifestyle on the water surface[6], can drive species diversification. The ancestors of the Gerromorpha transited from terrestrial life to life on the water surface over 200 Myr ago[6–8]. Acquisition of this lifestyle exposes the insects to new challenges including locomotion on fluid and escaping bottom-striking predators. Paramount to their success on water surfaces is the evolution of the shape and size of their legs in association with niche specialization and mode of locomotion[6,9,10]. Basal lineages, which occupy transitional zones between water and land, retain ancestral character states with midlegs shorter than hindlegs, and employ the ancestral mode of locomotion through alternating leg movements[6,7,9,11,12]. In contrast, water striders (Gerridae) that specialize in open water zones employ a derived mode of locomotion through simultaneous rowing motion of the midlegs. This rowing mode is enabled by a novel character state where midlegs are now longer than hindlegs[6,11–13]. A hallmark of the entire lineage of the semi-aquatic bugs, that distinguishes them from closely related terrestrial Heteroptera, is the deployment of the Hox protein Ubx in the second thoracic segment[11,13] and a notable increase of Ubx dose in the third thoracic segment[12]. These changes in Ubx expression and function are associated with changes in leg length across the semi-aquatic bugs[11–13]. While we are beginning to understand how morphological changes in the appendages are associated with specificities of locomotion and sexual interactions on the fluid water–air interface[6,10,14], little is known about how the legs are shaped under selection by aquatic predators. Here we investigated the evolutionary and developmental genetic mechanisms underlying antipredator strategies in water striders. We found that water striders survive attacks by aquatic predators using a quick jump enabled by their long and slender legs. We demonstrated that the shape of the legs, that are important for this behaviour, is modulated through the evolution of a lineage-specific interaction between Ubx and a new target gene called *gilt*. Finally, we showed that short-legged individuals, generated through *gilt* RNAi, exhibit reduced jump performance in predation tests.

Results and discussion

Water striders use quick jump, and not flight, to escape predators. First, we asked how water striders survive in the face of the various predators lurking under the surface[15–17] (Supplementary Movie 1). Using high-speed videography, we examined the interaction between the water strider *Limnoporus dissortis* and the stealth surface hunting halfbeak fish *Dermogenys sp*[18]. When attacked from beneath, the water strider uses an almost vertical jump in the air to avoid the strike (Fig. 1; Supplementary Movie 1 section 3). The animal pushes against the water surface with a characteristic bending of the mid and hindlegs, creating dimples without breaking surface tension (Fig. 1a; Supplementary Movie 1 section 3). The first legs do not seem to contribute. This behaviour results in movement amplification that propels the insect several times higher in the air than the length of its body. The time from the trigger of the strike until when the fish's jaws reach the water strider is 28.7 ± 3.9 ms, on average (Supplementary Table 1). As a result of the jump (Fig. 1b), the fish misses (Fig. 1c) and the water strider falls back and escapes (Fig. 1d; Supplementary Movie 1 section 3). To measure the duration of the jump, we developed a controlled set-up where the attack is simulated using a dead backswimmer (*Notonecta),* a natural predator of water striders[17], which we manoeuvred with a wire (Fig. 1e). This set-up reproduced the water strider jump, with 28.8 ± 1.8 ms from when the jump is triggered until when the legs of the animal lose contact with water surface

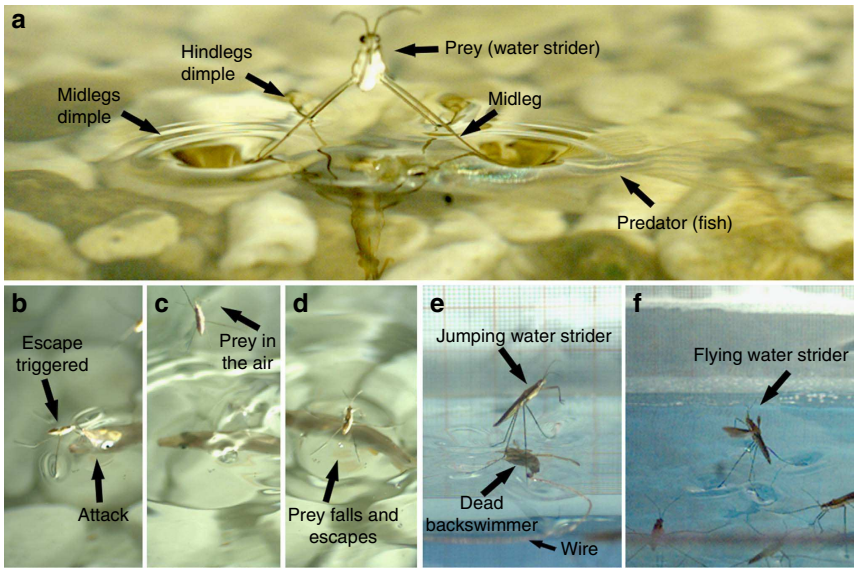


Figure 1 | Predation escape strategy and behavioural setups in water striders. (**a**) A water strider initiating the escape through a jump triggered by the approaching fish. The midlegs push downwards causing them to bend and form deep depressions (dimples). (**b**–**d**) Frames, taken from Supplementary Movie 1, showing the sequence of interaction between the water strider and the attacking halfbeak fish. The water strider dodges the attack (**b**) by moving away from the trajectory of the fish's strike through a quick jump in the air (**c**), while the fish is carried away by its own momentum, the water strider falls back and skates away (**d**). (**e**) Controlled set-up used to trigger the jump in adult water striders using a dead backswimmer manoeuvred by the experimentalist with a wire. (**f**) Flight in water striders triggered by increasing population density and exposure to high light intensity.

© 2015 Macmillan Publishers Limited. All rights reserved.

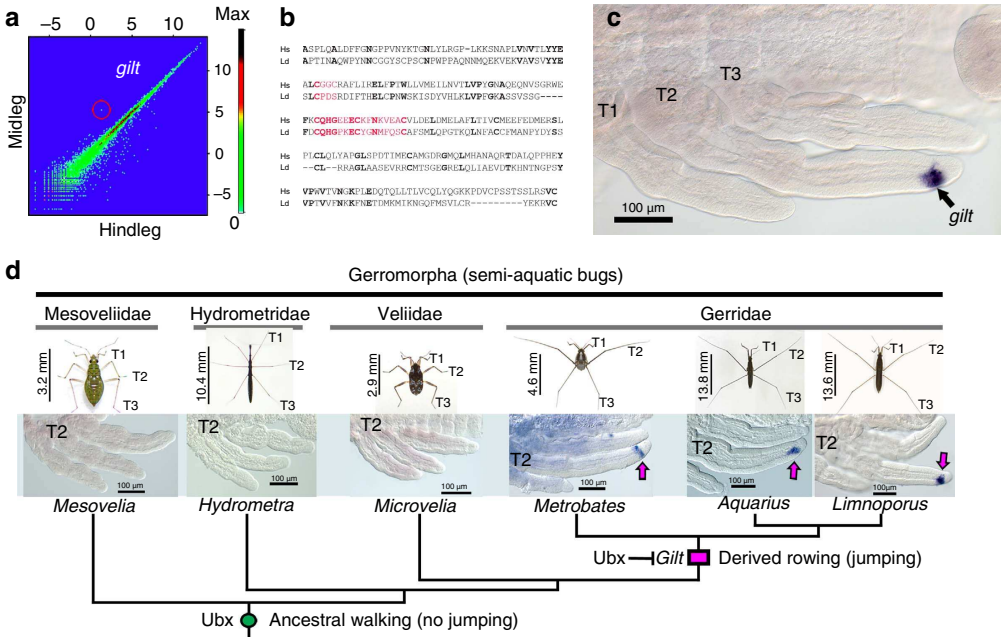


Figure 2 | *gilt* expression during development and evolution in the semi-aquatic bugs. (**a**) Comparative analysis of gene expression profiles between midlegs and hindlegs during early embryogenesis in *Limnoporus dissortis*. The transcript corresponding to *gilt* is encircled. (**b**) Sequence comparison between *L. dissortis* and *Homo sapiens* Gilt proteins. Gilt active site and signature motif are highlighted in red, identical amino acids are in bold. (**c**) *In situ* hybridization detecting *gilt* mRNA expression in the midlegs of an early *Limnoporus* embryo (48 h: ~30% development). T1, 2, 3: Thoracic segments 1, 2, 3. (**d**) Evolution of *gilt* expression across six semi-aquatic bugs representing four families and both the ancestral and derived leg-length plan. *gilt* expression in the midlegs is detected in species with derived relative leg length plan and absent in species with the ancestral leg plan. The gain of Ubx in midlegs at the base of the Gerromorpha lineage is indicated with green circle and the evolution of *gilt* expression at the base of the Gerridae with pink rectangle. Scale bars indicated are body sizes of adult individuals.

(Supplementary Movie 1 section 3; Supplementary Table 1). Therefore, the water strider delivers an antipredator response that matches the quick strike of the predator. Interestingly, we observed that adult water striders invariably jump to dodge the strike even though they are capable of flight[6,19]. When we triggered flight (Fig. 1f, see details in Methods and SOM), we found that the water strider takes an average of 124.8 ± 18.6 ms to leave the water surface. Eighty per cent (99.4 ± 18.0 ms) of the flight time is accounted for by the animal flapping its wings, while the remaining twenty per cent (25.4 ± 12.8 ms) represent take-off supported by the legs and the wings (Supplementary Movie 2). This indicates that flight is four times significantly slower (one-way analysis of variance (ANOVA), $n = 10$, $P < 0.0001$) than the jump, that is, four times slower than the predatory strike, and thus not suitable as an antipredator strategy (Supplementary Table 1; Supplementary Movie 2). Therefore, the jump enabled by the slender long legs is critical for the water strider to survive aquatic predators.

Emergence of novel gene expression in water striders. The derived morphology where midlegs are longer than hindlegs is important for this antipredator behaviour. We therefore wanted to uncover the genetic mechanisms that shape the differences between the legs during development in water striders. A genome-wide comparative analysis of gene expression profiles uncovered 155 differentially expressed transcripts between midlegs and hindlegs of developing *Limnoporus* embryos (Fig. 2a). We further focused on a transcript the expression of which perfectly matched the differences in morphology between the legs, that is, high in the midlegs and very low in the hindlegs (Fig. 2a; Supplementary Fig. 1a–c). This transcript is homologous to a gene known as *gamma interferoninducible thiol reductase* (*gilt*) (Fig. 2b), a reducing enzyme required for antigen processing and presentation by the major histocompatibility complex II in mammals[20]. A role for Gilt in innate immunity has recently been suggested in flies[21], but no known developmental function has ever been associated with this protein. During water strider development, we found that *gilt* mRNA expression is specific to the midlegs. Its expression begins at ~25% of embryogenesis (Supplementary Fig. 1d), exclusively in a distinct cell population at the tip of midleg tarsi, and continues through later stages of embryogenesis (Fig. 2c; Supplementary Fig. 1e,f). *gilt* is exclusively expressed in the midlegs until 45% of embryogenesis, but also appears in the hindlegs during later embryogenesis (Supplementary Fig. 1b,c). While all Gerridae have the derived leg plan where midlegs are longer than hindlegs, many other Gerromorpha retain the ancestral leg plan where midlegs are shorter than hindlegs[6,7,11,12]. We therefore cloned *gilt* from a sample of six species representing the two morphologies (Supplementary Fig. 1g). We found that *gilt* is expressed in the tarsus of midlegs in all three species with derived leg plan and absent from all three species that retain an ancestral leg plan (Fig. 2d). This indicates that the expression of *gilt* in the midlegs is characteristic to the derived morphology and that its evolution coincides with the evolution of rowing characteristic of species specialized in open water zones.

Ubx modulates *gilt* expression in the legs. Previous work has shown a role for Ubx in the acquisition of the novel long midleg body plan, with moderate levels of Ubx lengthening midlegs and six to seven times higher levels of Ubx shortening hindlegs[11–13]. Interestingly, our comparative transcriptome shows an inverse correlation between *Ubx* and *gilt* levels—with low *Ubx* high *gilt* in midlegs, and high *Ubx* low *gilt* in hindlegs (Supplementary Fig. 2)

© 2015 Macmillan Publishers Limited. All rights reserved.

suggesting a regulatory interaction between Ubx and *gilt*. To test this possible interaction, we contrasted a transcriptome of legs extracted from *Ubx* RNAi embryos to that of legs from untreated embryos (Fig. 3a). We found that in *Ubx* RNAi embryos *gilt* expression doubles in the midlegs and appears ectopically in the hindlegs (Fig. 3a). Furthermore, *in situ* hybridization in embryos treated with *Ubx* RNAi confirms the appearance of *gilt* expression in the hindlegs (Fig. 3b), demonstrating that *gilt* is no longer repressed when we knock *Ubx* down. Therefore, Ubx represses *gilt* partially in the midlegs and entirely in the hindlegs of *Limnoporus* early embryos, and the extent of this repression is dependent on Ubx levels.

Gilt is required to lengthen the midlegs in water striders. We further tested the role of *gilt* during embryonic leg development using parental RNA interference[11–13]. Embryos with depleted *gilt* transcript ($n=12$) grow significantly shorter midlegs relative to controls (YFP-injected controls $n=7$ and untreated controls $n=10$, one-way ANOVA $P<0.0001$; Fig. 3c,d) (The various controls for this experiment are detailed in the Methods and Supplementary Fig. 3). Statistical analyses of measured leg length revealed that *gilt* RNAi embryos develop significantly (12%) shorter midlegs compared with controls (Fig. 3e). This shortening affects the tarsus, tibia and femur (Supplementary Fig. 4), suggesting that the role of *gilt* in leg allometry is cell non-autonomous. *gilt* expression increases in *Ubx* RNAi treatments (Fig. 3a), whereas either *Ubx*[11–13] or *gilt* RNAi results in shorter legs (Fig. 3d,e). To test whether increased *gilt* expression impacts leg length in the absence of Ubx, we performed a double *Ubx* + *gilt* RNAi experiment (Fig. 3e,f; details and controls can be found in Methods and Supplementary Fig. 5). Midlegs in double *Ubx* + *gilt* RNAi ($n=21$) are 7% significantly shorter (one-way ANOVA, $P<0.0046$) than midlegs of single *Ubx* RNAi ($n=13$; Fig. 3e). Conversely in the hindlegs, double *Ubx* + *gilt* RNAi produces an intermediate effect between single *gilt* (one-way ANOVA, $P<0.0001$) and single *Ubx* (one-way ANOVA, $P<0.0001$) RNAi (Fig. 3e). Therefore, part of the role of Ubx in lengthening the midlegs allows some expression of *gilt*, and part of its role in shortening the hindlegs requires the complete silencing of *gilt*. Altogether, our data demonstrate that *gilt* represents a downstream effector of Ubx, whose expression contributes to the long leg phenotype that is associated with the evolution of the antipredator jump behaviour.

Gilt-free water striders jump lower. Unlike Ubx, the depletion of *gilt* does not impair whole-body development and instead allows the emergence of viable adults without any apparent defects other than the reduction of leg length. *gilt* RNAi animals therefore

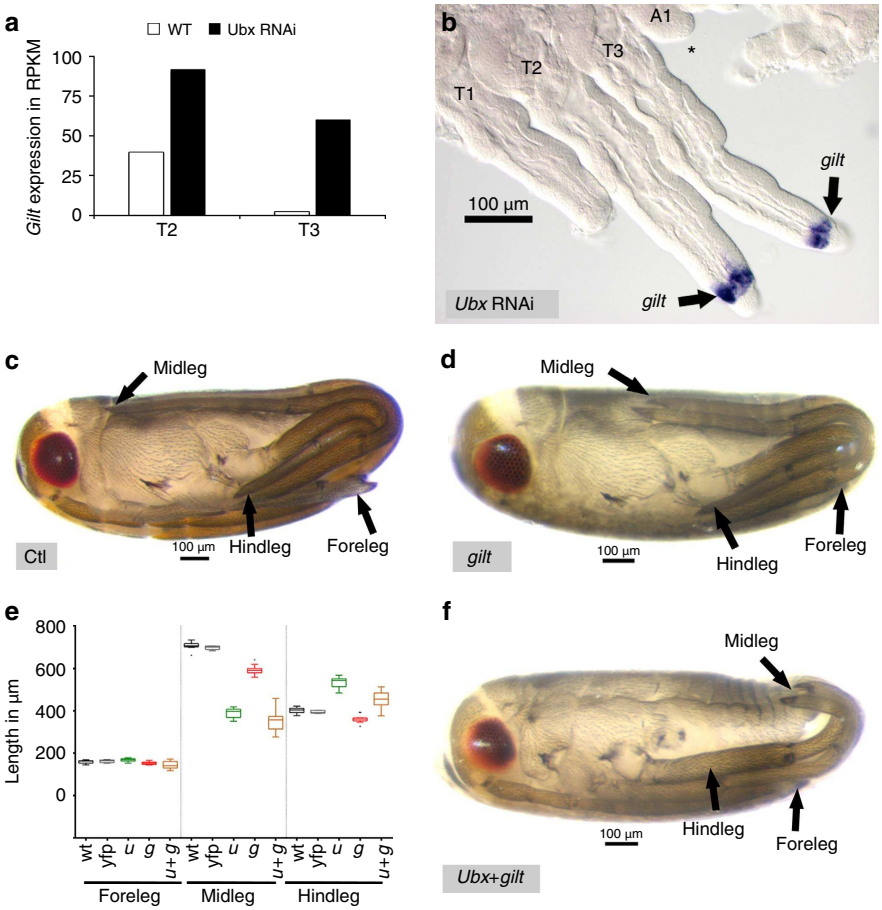


Figure 3 | *gilt* function and interaction with Ubx in *Limnoporus dissortis*. (**a**) Transcriptomics data show that *gilt* expression doubles in the midleg and appears in the hindlegs of *Ubx* RNAi embryos, as revealed by the number of reads per kilobase per million (RPKM). (**b**) *In situ* hybridization detecting *gilt* mRNA in an *Ubx* RNAi early embryo. Note the expansion of *gilt* in the midlegs and its ectopic expression in the hindleg. T1, T2 and T3: Thoracic segments 1, 2 and 3; A1: Abdominal segment 1; Asterisk (*) indicates an ectopic leg forming on A1 and that is characteristic to *Ubx* RNAi in water striders[12]. (**c**) Control late embryo. (**d**) *gilt* RNAi late embryos with shorter midlegs. (**e**) Tukey box plots showing measurements of the effect of *u* (*Ubx* RNAi $n=13$), *g* (*gilt* RNAi $n=12$), or *u* + *g* (double *Ubx* + *gilt* RNAi $n=21$) RNAi on leg length compared with controls (wt untreated $n=10$ and injected with YFP double-stranded RNA $n=7$). (**f**) Double *Ubx* + *gilt* RNAi late embryo showing drastically shorter midlegs. Scale bars are indicated.

© 2015 Macmillan Publishers Limited. All rights reserved.

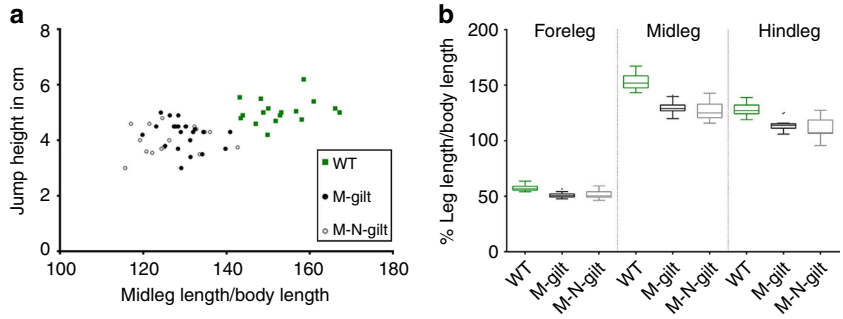


Figure 4 | Decreased performance of *gilt* RNAi water striders in predation test. (**a**) Tukey box plots showing leg lengths variations within normal and *gilt* RNAi individuals at the fifth nymphal instar used in this test. Wild types ($n=17$) are untreated, M-*gilt* (Maternal *gilt* RNAi; $n=21$) are fifth nymphal instar individual obtained by injecting mothers with *gilt* ds-RNA, and M-N-*gilt* (Maternal plus Nymphal *gilt* RNAi; $n=17$) received an additional injection at first nymphal instar. (**b**) Lower performance in simulated predation tests directly correlates with midleg length reduction across *gilt* RNAi individuals.

provided an exquisite opportunity to directly assay the importance of leg length for antipredator strategy. To do this, we measured the jump performance of water striders when attacked from beneath using our controlled set-up (Fig. 4). On average, control individuals ($n=17$) jumped 5.05 ± 0.46 cm (Fig. 4a), while *gilt* RNAi individuals ($n=38$) jumped 4.13 ± 0.51 cm (Fig. 4a), that is, a performance 18% significantly lower than that of normal individuals (one-way ANOVA, $P<0.0001$; Supplementary Movie 3). Interestingly, *gilt* RNAi animals display a varying degree of leg length reduction (in average 16 and 6% shorter mid and hindlegs, respectively; Fig. 4b) and there is a clear correlation between midleg length and jump performance (Fig. 4a). These results establish that the acquisition of *gilt* expression contributes to shaping the legs in association with the evolution of the jump behaviour, and that even a small reduction in leg length reduces the ability of water striders to deliver a powerful jump when the predator strikes.

Altogether, our findings show that the predation escape strategy employed by water striders relies primarily on a jump reflex, enabled by the long slender legs. We also show that the acquisition of this morphological character involves novel genetic interactions between the Hox protein Ubx and *gilt;* a previously unknown target, encoding a reducing enzyme. When compared with species that retain ancestral character states where midlegs are shorter than hindlegs, the gain of *gilt* expression contributes to increase the length of midlegs that is characteristic to species specialized in rowing in open waters. Therefore, the evolution of this novel interaction between Ubx and *gilt* contributes to shaping the legs in association with the efficient antipredator strategy that water striders employ to survive aquatic predators[22]. The jump, though representing a quick reaction to the strike, can only give advantage to the prey if coupled with a primary mechanism of predator detection[23]. Despite the stealth approach of the fish, water striders are exquisite sensors of vibration owing to a multitude of sensory bristles located in the contact surface between the legs and the water surface[24]. Therefore, the ability of water striders to detect predators, together with the adapted morphology of their legs may have been key to their success in open water surfaces worldwide. This rich ecological and evolutionary context, together with the amenability to experimental manipulation, make water striders a suitable system to study how genetic interactions evolve in association with adaptive phenotypic evolution.

Methods

Animal collection and rearing. *Limnoporus dissortis*, *Microvelia americana* and *Metrobates hesperius* were collected from 'Rivière de l'Acadie,' at the vicinity of Montréal, Québec, Canada. *Aquarius paludum* and *Hydrometra stagnorum* were collected in a pond near Lyon, France. *Mesovelia mulsanti* were collected in a pond near Cayenne, French Guiana. Halfbeak fishes (*Dermogenys sp.*) were purchased from a fish store in Lyon, France. Both water striders and fish were kept in aquaria at 25 °C with a 14-h light/10-h dark cycle, and both were fed on live crickets. Pieces of floating Styrofoam were regularly supplied to female water striders to lay eggs.

Measurement of fish strike duration. Ten high-speed movies were captured using a Miro M310 high-speed camera (Vision Research), at 2,000 frames per second, while halfbeak attacked water striders. In each movie, we manually identified the start frame where the animal first triggered the attack and the end frame where the fish's jaw reach the initial position of the body of the water strider. The strike duration represented by the time elapsed between the start and the end frames was calculated using PCC software (Vision Research). Details about software and method of calculation can be found in this tutorial by Vision Research: http://www.visionresearch.com/Service--Support/Tutorials/

Measurement of water strider jump duration. This experiment was conducted in adult water striders to compare with the duration of flight below. To measure jump duration in a controlled manner, we simulated the attack using a dead backswimmer (*Notonecta)* attached at the end of a wire and manoeuvred by a manipulator. This strategy reproducibly induced water striders to jump. We then filmed ten individuals at 2,000 frames per second. In each movie, we identified the start of the jump as the time elapsed between the frame where we see the first movement of the leg pushing downwards, and the frame where all legs lose contact with the water surface. Measurements of the time elapsed between start and end of the jump were performed using PCC software (Vision Research).

Measurement of flight duration. To measure flight duration, we stimulated flight by increasing the density of water striders in the aquaria and by exposing them to an extra source of light. These conditions originate from our own observations while breeding water striders in the lab. We then took ten high-speed movies (2,000 frames per second) of 10 individuals during flight. In each movie, we identified the start of the flight as the time elapsed between the frame where we first see wings moving and the frame where all legs lose contact with the water surface. Measurements of the time elapsed between start and end of the flight were performed using PCC software (Vision Research).

***Limnoporus dissortis* reference transcriptome.** A mixture of adult males and females, nymphal instars and various embryonic stages of *L. dissortis* were used to isolate total RNA using Trizol (Invitrogen). Transcriptome was sequenced using 454 Roche technology and assembled using Newbler program version 2.6 (Roche). A total of 26,237 isotigs defining 16,368 isogroups were assembled. Assembled isotigs were annotated by sequence similarity against the NCBI 'non-redundant' protein database using BLAST2GO. In addition to inferred annotations, for each annotated isogroup we predicted PFAM motifs using HMMER3 tools from Sanger institute. *L. dissortis* transcriptome can be retrieved in NCBI using the following accession number: PRJNA289202.

Quantification of gene expression using comparative transcriptomics. RNA from dissected *L. dissortis* legs was used by *ProfilExperts* (Lyon-France) to conduct deep sequencing using TruSeq RNA kit and Illumina HiSeq2000 technology. A total of ~50 million reads was generated per sample. These reads were trimmed and filtered to remove low-quality bases and aligned against the draft *Limnoporus* transcriptome and transcript levels were quantified by determining the number of Reads Per Kilobase per Million mapped reads (RPKM) in each sample[25]. The raw reads of *L. dissortis* leg transcriptomes can be retrieved here: PRJNA289202.

© 2015 Macmillan Publishers Limited. All rights reserved.

***gilt* cloning.** Total RNA from *L. dissortis, A. paludum*, *M. hesperius*, *M. americana*, *H stagnorum*, and *M. mulsanti* was extracted from different embryonic stages and nymphal instars. First strand cDNA synthesis was then performed for each species, using total RNA as a template, according to instructions outlined in the Invitrogen cDNA synthesis kit. Specific primers for *gilt* in *L. dissortis* were designed based on sequences obtained from the whole transcriptome of *L. dissortis*. Primers for each species used in PCR reaction to amplify a fragment of *gilt*, can be found in Supplementary Table 2. *gilt* sequence for each species can be retrieved in Genbank. *gilt* sequences from these various species can be retrieved in GenBank using the following accession numbers: KR704883, KR704884, KR704885, KR704886, KR704887, and KR704888.

Embryo dissection. Embryos were collected, treated with 25% bleach, and then washed with PTW 0.05% (1 × PBS; 0.05% Tween-20). For image acquisition, late embryos were dissected out of the eggshell, fixed in 4% formaldehyde, and their images captured using a Zeiss Discovery V12 scope. For staining, embryos of various stages were manually dissected, cleaned from yolk and fixed following the method below.

***In situ* hybridization.** Dissected embryos were fixed in 200 µl 4% Paraformaldehyde (PFA) + 20 µl dimethyl sulfoxide, and 600 µl heptane for 20 min at room temperature with shaking. Embryos were then washed several times in cold methanol and rehydrated in decreasing concentrations of methanol in PTW 0.05%. These embryos were washed three times in PTW 0.05%, three times in PBT 0.3% (1 × PBS; 0.3% Triton X100), and twice with PBT 1% (1 × PBS; 1% Triton X100). Following these washes, embryos were transferred to 1:1 PBT 1% / hybridization solution (50% formamide; 5% dextran sulfate; 100 µg − ml^{-1} yeast tRNA; 1 × salts). The composition for 100 ml 10 × salt is as follows: 17.5 g sodium chloride, 1.21 g tris-base, 0.71 g monosodium phosphate, 0.71 g sodium phosphate dibasic, 0.2 g Ficoll 400, 0.2 g Polyvinylpyrrolidone (PVP), 10 ml of 0.5 M EDTA, 0.2 g BSA (pH 6.8). Embryos were prehybridized for one hour at 60 °C, followed by addition of a Dig-labelled RNA probe overnight at 60 °C. Embryos were then transferred gradually from hybridization solution to PBT 0.3% through consecutive washes with 3:1, 1:1, 1:3 prewarmed hybridization solution: PBT 0.3% gradient. A blocking step was performed with PAT (1 × PBS; 1% Triton X100; 1% BSA) at room temperature followed by incubation with anti-DIG antibody coupled with alkaline phosphatase for 2 h at room temperature, or at 4 °C overnight. Embryos were washed several times in PBT 0.3% then in PTW 0.05% before colour reaction is conducted with NBT/BCIP in alkaline phosphatase buffer (0.1 M Tris pH 9.5; 0.05 M MgCl2; 0.1 M NaCl; 0.1% Tween-20).

Parental RNAi. Gene knockdown of *Ubx* and *gilt* using parental RNAi was conducted following the protocol described in references[11–13]. Control RNAi was conducted by injecting ds-*yfp* (double-stranded (ds) RNA from the *yfp* gene sequence). Template for *in vitro* transcription to produce ds-RNA for each gene was prepared using the T7-tagged primers in Supplementary Table 3.

Efficiency of *gilt* RNAi. We verify that *gilt* RNAi specifically depletes *gilt*, we stained embryos with a mix containing *gilt* probe and a 10-fold diluted *egfr* probe. *egfr* is expressed in the legs in five stripes prefiguring the joints between leg segments[12]. This probe mix allows *gilt* staining to become strong and easily visible before *egfr* staining. If our RNAi experiment successfully depletes *gilt*, then we expect to see *egfr* staining but not *gilt*. In control treatments, this probe mix stains the embryos such that *gilt* staining appears stronger than *egfr* staining, thus allowing us to assess *gilt* expression relative to the strength of *egfr* staining (Supplementary Fig. 3a). Staining embryos from *gilt* RNAi treatments with the same mix fails to detect *gilt* mRNA in midlegs tarsi even when we allow *egfr* staining to become strong (Supplementary Fig. 3b). This demonstrates that our RNAi experiment efficiently depletes *gilt* transcripts.

Specificity of *gilt* RNAi. To ensure that our RNAi does not suffer from off target effects, we designed double-stranded RNA from two non-overlapping fragments of *gilt* (Supplementary Fig. 3c). Injection of ds-RNA from these three fragments induced the same leg length phenotypes (Supplementary Fig. 3d–f). Furthermore, we repeated the same experiment on a second species, *A. paludum*, and obtained the same results (Supplementary Fig. 3g–j). Therefore, gilt RNAi depletes specifically *gilt* transcript. To make sure that the leg-shortening phenotype of *gilt* RNAi is not just a reflexion of natural variation in leg length, we measured the legs of individual that reach the first nymphal instar at successive days (Supplementary Fig. 3k). We found that the first individual that are laid by injected females (days 1 and 2) have either normal leg length or very subtle shortening. The shortening phenotype becomes stronger the following days (Supplementary Fig. 3k). This indicates that the effect we see on leg length is due to *gilt* knockdown.

Efficiency of double *gilt* + *Ubx* RNAi. We confirmed the efficiency of the double *gilt* + *Ubx* knockdown by staining embryos with the mix containing *gilt* probe and 10-fold diluted *egfr* probe (Supplementary Fig. 5). Control embryos show a clear *gilt* expression in the midlegs (Supplementary Fig. 5a), *Ubx* RNAi embryos show *gilt* expression in both mid- and hindlegs (Supplementary Fig. 5b), whereas single *gilt* RNAi (Supplementary Fig. 5c) and double *gilt*/*Ubx* RNAi (Supplementary Fig. 5d) show no *gilt* expression in either leg despite the clear *egfr* staining. This demonstrates that the double *gilt* + *Ubx* RNAi knockdown was successful.

Leg measurements. A sample of first instar nymphs was used for each RNAi group and control: untreated control $n = 10$; *yfp*-injected negative control $n = 7$; *gilt* RNAi $n = 12$; *Ubx* RNAi $n = 13$; and *Ubx* + *gilt* RNAi $n = 21$. Nymphs were dissected and mounted on slides in Hoyer's medium. Three sets of measurements were recorded: head width and the length of the tarsus, tibia and femur individually. Measurements for each leg segment of each pair of legs were recorded on a Zeiss microscope using the Zen software.

Predation test experiment. We conducted predation performance in a controlled set-up that induces the water striders to jump near a grid such that the height of the jump can be accurately measured. We triggered the jump by approaching a dead backswimmer under the water striders. Individuals from the different categories (WT, Maternal-*gilt* RNAi and Maternal–Nymphal-*gilt* RNAi) were allowed to develop until fifth nymphal instar. We used fifth instars because they are large enough for recording and also because they easily induced to jump. We took a total of seven movies for each individual and extracted jump height. We then killed the individuals and measured their leg length and body length. Finally, we plotted the highest jump from each individual by its leg length corrected to body length.

Statistical analyses. All statistical analyses and plots were performed using GraphPad Prism (version 6.00). For differences in leg length between controls and the various RNAi experiments (*gilt, Ubx, Ubx* + *gilt*), we performed a one-way ANOVA test to determine statistical significance. We normalized leg length by head width to correct for differences in body size. We did not use body length because body length at the first nymphal instar changes considerably with changes in food intake (well fed individuals tend to swell up). Statistical significance for the various legs and segment of the legs are summarized in Supplementary Table 4.

For the predation performance experiment, we performed a one-way ANOVA test to determine statistical significance in leg length and jump height between untreated (WT) and various RNAi experiments (Maternal-*gilt* and Maternal–Nymphal-*gilt*). We normalized leg length by body length to correct for differences in body size. We used body size because it is a reliable reference at the fifth nymphal instar. Statistical significance for differences in leg length and jump height are summarized in Supplementary Table 5.

References

1. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134,** 25–36 (2008).
2. Davidson, E. H. *The Regulatory Genome: Gene Regulatory Networks in Development And Evolution* (Academic, 2006).
3. Abouheif, E. *et al.* Eco-evo-devo: the time has come. *Adv. Exp. Med. Biol.* **781,** 107–125 (2014).
4. Barrett, R. D. & Hoekstra, H. E. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* **12,** 767–780 (2011).
5. Emilia Santos, M., Berger, C. S., Refki, P. N. & Khila, A. Integrating evo-devo with ecology for a better understanding of phenotypic evolution. *Brief. Funct. Genomics*. doi: 10.1093/bfgp/elv003 (8 March 2015).
6. Andersen, N. M. *The Semiaquatic Bugs (Hemiptera: Gerromorpha)* Vol. 3. (Scandinavian Science Press LTD., 1982).
7. Damgaard, J. Phylogeny of the semiaquatic bugs (Hemiptera-Heteroptera, Gerromorpha). *Insect. Syst. Evol.* **39,** 30 (2008).
8. Li, M., Tian, Y., Zhao, Y. & Bu, W. Higher level phylogeny and the first divergence time estimation of Heteroptera (Insecta: Hemiptera) based on multiple genes. *PLoS ONE* **7,** e32152 (2012).
9. Andersen, N. M. A comparative study of locomotion on the water surface in semiaquatic bugs (Insecta, Hemiptera, Gerromorpha). *Videnskab. Meddr. Dansk. Naturh. Foren.* **139,** 337–396 (1976).
10. Hu, D. L., Chan, B. & Bush, J. W. The hydrodynamics of water strider locomotion. *Nature* **424,** 663–666 (2003).
11. Khila, A., Abouheif, E. & Rowe, L. Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. *Evolution* **68,** 2159–2170 (2014).
12. Refki, P. N., Armisen, D., Crumiere, A. J. J., Viala, S. & Khila, A. Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. *Dev. Biol.* **392,** 441–453 (2014).
13. Khila, A., Abouheif, E. & Rowe, L. Evolution of a novel appendage ground plan in water striders is driven by changes in the Hox gene Ultrabithorax. *PLoS Genet.* **5,** e1000583 (2009).
14. Khila, A., Abouheif, E. & Rowe, L. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. *Science* **336,** 585–589 (2012).
15. Cooper, S. D. The effects of trout on water striders in stream pools. *Oecologia* **63,** 376–379 (1984).

© 2015 Macmillan Publishers Limited. All rights reserved.

16. Han, C. S. & Jablonski, P. G. Male water striders attract predators to intimidate females into copulation. *Nat. Commun.* **1,** 52 (2010).
17. Rowe, L. The costs of mating and mate choice in water striders. *Anim. Behav.* **48,** 1049–1056 (1994).
18. Reckel, F. & Melzer, R. R. Regional variations in the outer retina of atherinomorpha (Beloniformes, Atheriniformes, Cyprindontiformes: Toleostei): Photoreceptors, cone patterns, and cone densities. *J. Morphol.* **257,** 270–288 (2003).
19. Fairbairn, D. J. & King, E. Why do Californian striders fly? *J. Evol. Biol.* **22,** 36–49 (2008).
20. West, L. C. & Cresswell, P. Expanding roles for GILT in immunity. *Curr. Opin. Immunol.* **25,** 103–108 (2013).
21. Kongton, K., McCall, K. & Phongdara, A. Identification of gamma-interferon-inducible lysosomal thiol reductase (GILT) homologues in the fruit fly Drosophila melanogaster. *Dev. Comp. Immunol.* **44,** 389–396 (2014).
22. Suter, R. B. & Gruenwald, J. Predator avoidance on the water surface? Kinematics and efficacy of vertical jumping by Dolomedes (Araneae, Pisauridae). *J. Arachnol.* **28,** 201–210 (2000).
23. Edmunds, M. *Defence in Animals: A Survey of Anti-Predator Defences* (Longman, 1974).
24. Perez Goodwyn, P., Katsumata-Wada, A. & Okada, K. Morphology and neurophysiology of tarsal vibration receptors in the water strider Aquarius paludum (Heteroptera: Gerridae). *J. Insect. Physiol.* **55,** 855–861 (2009).
25. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5,** 621–628 (2008).

Acknowledgements

We thank F. Payre, M. Averof, F. Bonneton and F. Leulier for critical reading of the manuscript. We thank F. Moreira for help with species identification and PSMN-ENS de Lyon for providing access to servers and calculation power. This work was supported by an ATIP-Avenir from CNRS and an ERC-CoG # 616346 to A.K. and a PhD fellowship from Université Lyon1 to P.R.

Author contributions

D.A., P.R. and A.K. designed the study; P.R. and A.K. generated transcriptomics samples; D.A. performed transcriptome analyses; P.R., A.J.J.C. and S.V. cloned *gilt* from various species; P.R., A.J.J.C. and S.V. performed *in situ* hybridization; D.A., P.R. and A.J.J.C. did RNAi knockdown; D.A., P.R. and A.J.J.C. measured legs; D.A., P.R. and A.J.J.C. performed jumping movies; D.A. and A.K. performed flight movies; D.A. performed statistical analysis; and A.K., D.A. and P.R. wrote the paper and all authors commented on the draft.

Additional information

Accession codes: The transcriptome data generated in this study have been deposited in NCBI database under the accession code PRJNA289202. *gilt* sequences generated in this study have been deposited in GenBank nucleotide database under accession codes KR704883 to KR704888.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

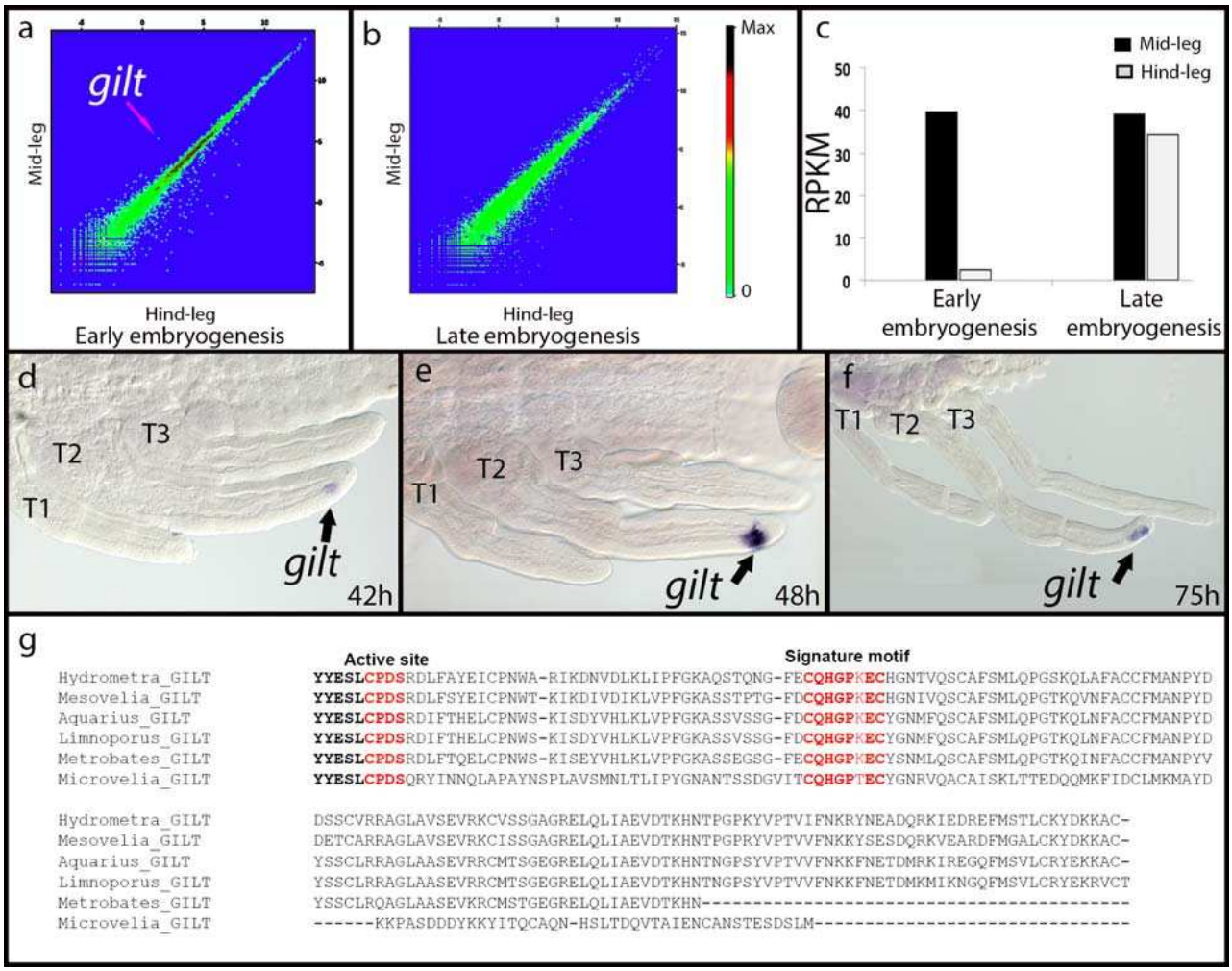
Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Armisén, D. *et al.* Predator strike shapes antipredator phenotype through new genetic interactions in water striders. *Nat. Commun.* 6:8153 doi: 10.1038/ncomms9153 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

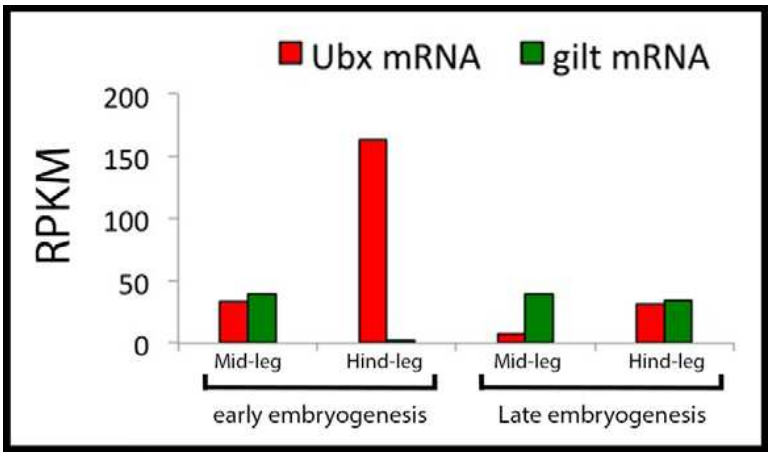
© 2015 Macmillan Publishers Limited. All rights reserved.

Supplementary figure 1



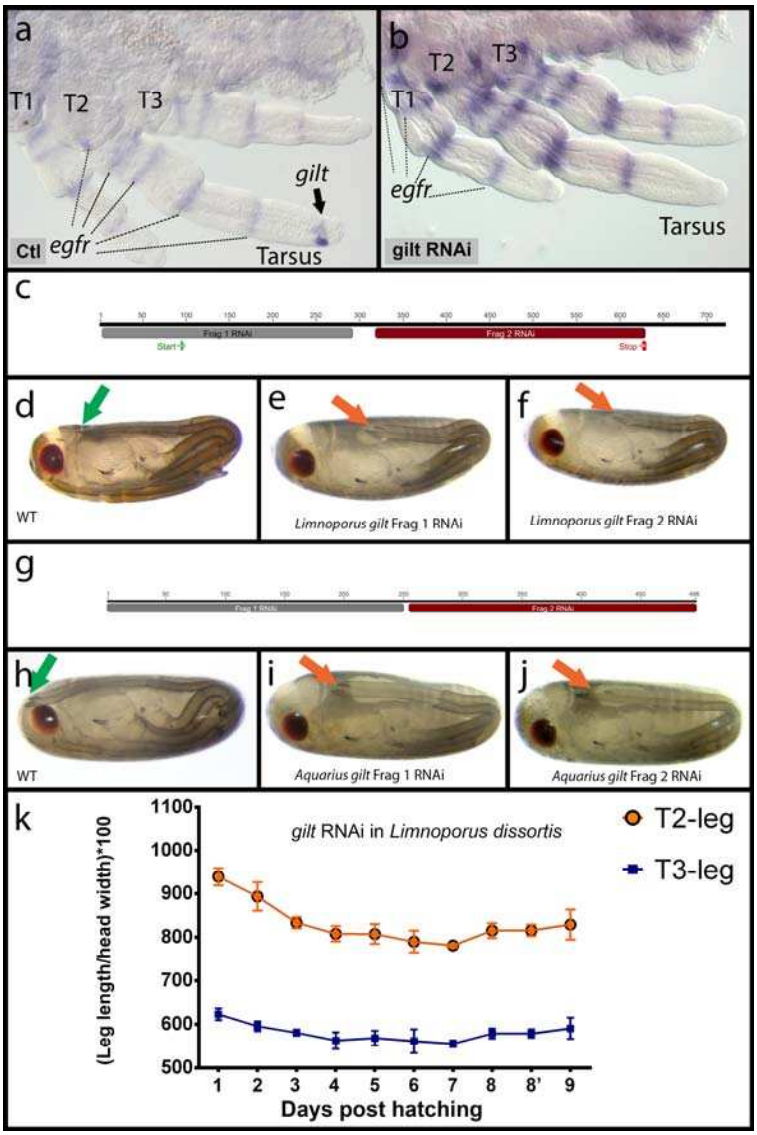
***gilt* expression during various stages of *Limnoporus* embryogenesis.** Comparative transcriptomics between mid-legs and hind-legs extracted from a pool of early (~50% to 75% embryogenesis, **a**) and late (~75% to 100% embryogenesis, **b**) embryos. *gilt* is highly expressed in mid-legs (arrow) in the early pool. (**c**) Levels of *gilt* transcript as reflected by the number of reads per kilobase per million (RPKM). Gilt expression is very high in the mid-legs early, but becomes equally expressed in both legs during late stages. (**d-f**) *gilt* mRNA appears first at 42 hours in the mid-legs, becomes strong at 48 hours, and continues in the mid-legs at 75 hours. We have not been able to detect *gilt* during late developmental stages using in situ hybridization, likely because of difficulties for probes to penetrate the cuticle that is deposited later during embryogenesis. (**g**) Alignment of Gilt protein across the six species of semi-aquatic bugs included in this study. Note the conservation of the active site and Gilt signature motif.

Supplementary figure 2:



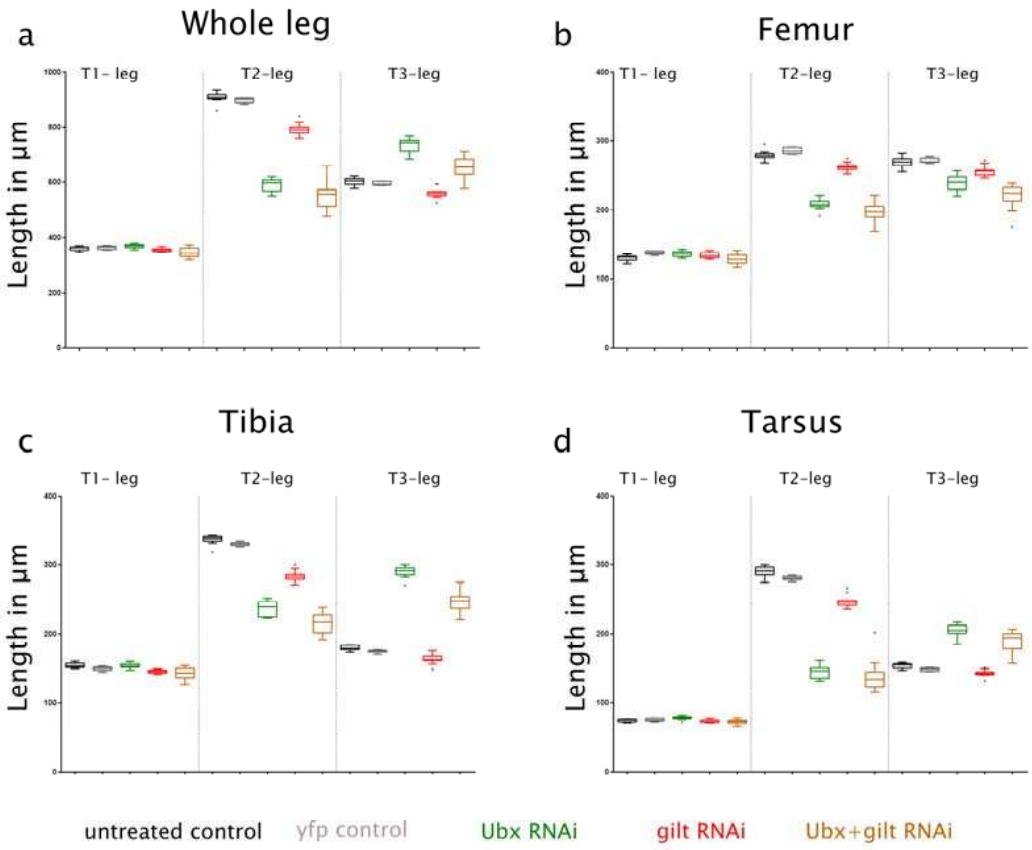
Changes in *gilt* levels (green) in relation with *Ubx* levels (red) as revealed by comparative transcriptomics between mid-legs and hind-legs during early and late embryogenesis. When Ubx level are high, *gilt* levels decrease.

Supplementary figure 3:



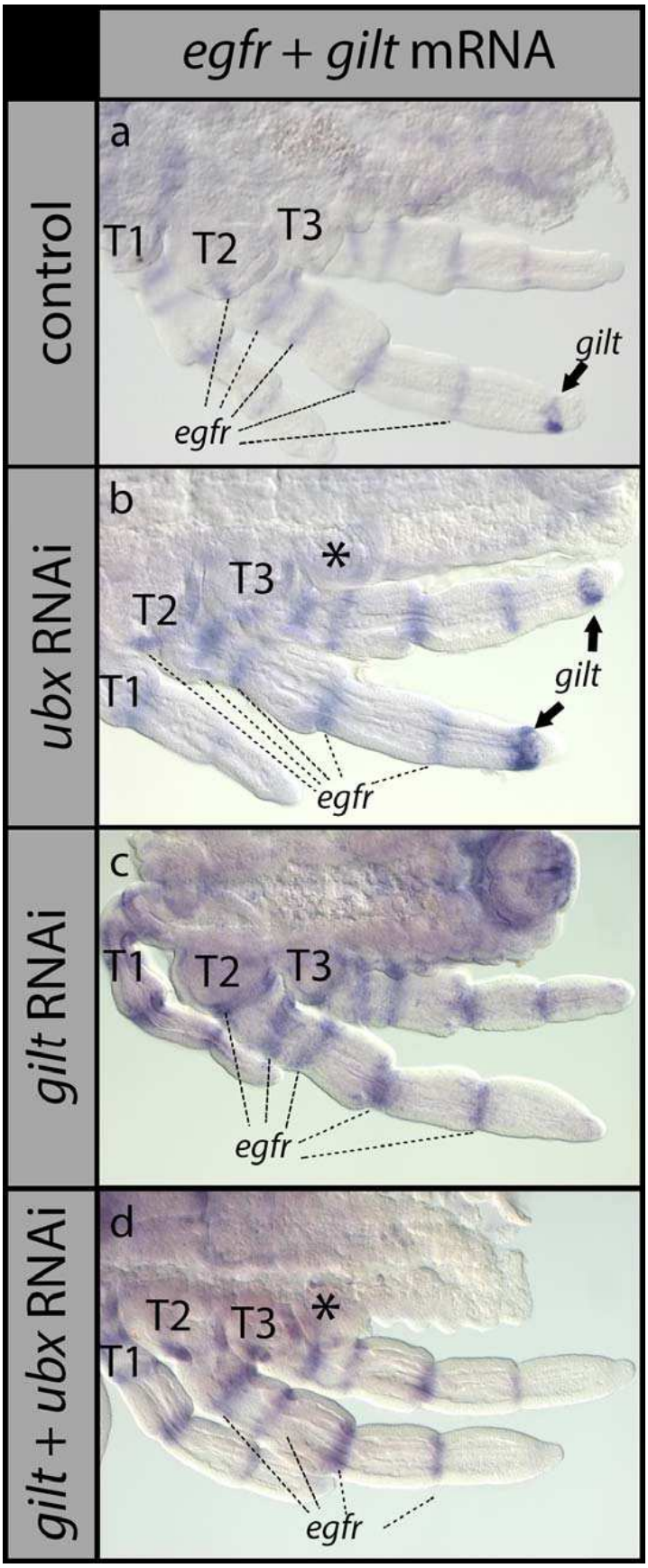
Controls for the *gilt* RNAi experiment. (**a**) Control embryo stained with a mix of *gilt* probe and ten times diluted *egfr* probe as an internal control. This probe mix detects *gilt* mRNA unambiguously in mid-leg tarsus and *egfr* mRNA in the boundaries between leg segments[12]. (**b**) *gilt* RNAi embryo stained with the same probe mix, which only detects *egfr* mRNA demonstrating that *gilt* has been successfully depleted. (**c**) A diagram presenting the two non-overlapping fragments of *Limnoporus gilt* used to synthesise double stranded RNA used in RNAi experiments. (**d**) Untreated *Limnoporus dissortis* embryo and (**e-f**) embryos obtained using the two different double stranded RNA regions of *gilt* showing the same phenotype, indicating that different fragments of *gilt* induce the same effect on leg length. (**g-h**) Two separate fragments of *gilt* in another water strider, *Aquarius paludum*, produce the same phenotype as in *Limnoporus dissortis*. (**k**) Dynamics of RNAi efficiency in *Limnoporus dissortis*. The sample size is 4 nymphs for each hatching day and error bars represent Standard Deviation. The first individuals produced after injection of double stranded RNA to the mothers tend to be either wild type or show a subtle phenotype. The strength of the phenotype increases during the following days.

Supplementary figure 4:



Tukey box plot showing the effect of *gilt* RNAi (n=12), *Ubx* RNAi (n=13), and *Ubx+gilt* (n=21) double RNAi on leg length and the length of the distal segments of each leg as compared to untreated controls (n=10) and yfp injected negative controls (n=7).

Supplementary figure 5: Efficiency of *gilt*, *Ubx*, and *gilt+Ubx* RNAi in *Limnoporus* embryos. (**a**) Control embryo (same as in Fig. S3a) stained with a mix of *gilt* probe and ten times diluted *egfr* probe as an internal control. This probe mix detects *gilt* mRNA unambiguously in mid-leg tarsus and *egfr* mRNA in the boundaries between leg segments[12]. (**b**) *Ubx* RNAi embryo stained with the same mix and showing unambiguous gilt mRNA expression in both mid-legs and hind-legs, along with *egfr* mRNA as an internal control. (c) *gilt* RNAi embryo (different than the one in Fig. S3b), stained with the mix of *gilt* and ten times diluted *egfr* probe, which only detects *egfr* mRNA demonstrating that *gilt* has been successfully depleted. (**d**) A double *gilt+Ubx* RNAi embryo, stained with the mix of *gilt* and ten times diluted *egfr* probe, showing a clear expression of the internal control *egfr* but no expression of *gilt* mRNA neither in the mid-legs or the hind-legs. Note in both *Ubx* RNAi and double *gilt+Ubx* RNAi the formation of an ectopic limb on the first abdominal segment (asterisk)[12].



Supplementary Table 1:

	Video	Duration in micro seconds
Halfbeak strike duration	Halfbeak strike 1	24000
	Halfbeak strike 2	32500
	Halfbeak strike 3	27000
	Halfbeak strike 4	30000
	Halfbeak strike 5	29000
	Halfbeak strike 6	25500
	Halfbeak strike 7	35500
	Halfbeak strike 8	33000
	Halfbeak strike 9	25000
	Halfbeak strike 10	26000
	Average	**28750**
***Limnoporus* jump duration**	Jump 1	30500
	Jump 2	29000
	Jump 3	27000
	Jump 4	27000
	Jump 5	32000
	Jump 6	29500
	Jump 7	28500
	Jump 8	29000
	Jump 9	26000
	Jump 10	30000
	Average	**28850**
***Limnoporus* flight duration**	Flight 1	112000
	Flight 2	152500
	Flight 3	96000
	Flight 4	131000
	Flight 5	124500
	Flight 6	126000
	Flight 7	145000
	Flight 8	110000
	Flight 9	107000
	Flight 10	144000
	Average	**124800**

Duration of strike in halfbeak fish, jump and flight in adult water striders calculated as an average of ten strikes, jumps and flight events respectively.

Supplementary table 2:

Species	Forward	Reverse
Limnoporus dissortis	AAGCCCCTACTATCAATGCCCAATG	TCAGCAGACTCTCTTCTCGTATCGG
Aquarius paludum	TACTACGAGTCGCTCTGTCCAGACAGC	CAGCAGACTCTCTTGTCGTATCGGCAGA
Hydrometra stagnorum	TACTACGAGTCACTTGCCCTGACA	TAGCAGGCCTTTTTGTCGTACTTGC
Metrobates hesperius	TACTACGAGTCGCTCTGTCCAGACAGC	GTGTTGTGTTTGGTGTCTACTTCAG
Microvelia americana	TTTGCGGGTTTCTACCAGCACTTT	GCCATGAGAGAATCAGATTCAGTGC
Mesovelia mulsanti	TACTACGAGTCACTCTGCCCTGACA	TAGCAGGCCTTTTTGTCGTACTTGC

Primer sequences used to clone *gilt* from the six semi-aquatic bugs species presented in this study.

Supplementary table 3:

		Forward	Reverse
Limnoporus dissortis	complete	taatacgactcactatagggagacca cAAGCCCCTACTATCAATGCCCAATG	M13-R_T7
	Frag 1	taatacgactcactatagggagacca cAAGCCCCTACTATCAATGCCCAATG	taatacgactcactatagggagacca cCATTCTTTCGGACCATGCTGACA
	Frag 2	taatacgactcactatagggagacca cCTTTCAGCATGCTGCAACCTGG	taatacgactcactatagggagacca cTCAGCAGACTCTCTTCTCGTATCGG
Aquarius paludum	Frag 1	taatacgactcactatagggagacca cTACTACGAGTCGCTATGCCCTGACA GC	taatacgactcactatagggagacca cGGTTTGCCATGAAACAGCAGGC
	Frag 2	taatacgactcactatagggagacca cCGACTACAGTTCATGTCTAAGAAG	taatacgactcactatagggagacca cCACAGGCTTTTTTCTCGTACCTGCA

Primers used to synthesize double stranded RNA used in RNAi experiments in *Limnoporus* and *Aquarius*. These primers are tagged with T7 promoter (small case letters) to enable in vitro transcription using the T7 polymerase enzyme.

Supplementary table 4:

		1way ANOVA (T1 leg Femur)	Mean 1	Mean 2	Mean Diff,	95% CI of diff,	Significant?	Summary	Adjusted P Value
Fore-legs	Femur	WT vs. Control	129,9	137,1	-7,191	-14,52 to 0,1379	No	ns	0,0569
		WT vs. Ubx+GILT	129,9	128,3	1,604	-4,110 to 7,319	No	ns	0,9323
		WT vs. GILT	129,9	133,6	-3,666	-10,03 to 2,702	No	ns	0,4905
		WT vs. Ubx	129,9	135,6	-5,636	-11,89 to 0,6196	No	ns	0,0965
		Control vs. Ubx+GILT	137,1	128,3	8,796	2,305 to 15,29	Yes	**	0,003
		Control vs. GILT	137,1	133,6	3,525	-3,548 to 10,60	No	ns	0,6281
		Control vs. Ubx	137,1	135,6	1,555	-5,417 to 8,527	No	ns	0,9699
		Ubx+GILT vs. GILT	128,3	133,6	-5,27	-10,65 to 0,1114	No	ns	0,0577
		Ubx+GILT vs. Ubx	128,3	135,6	-7,24	-12,49 to -1,992	Yes	**	0,0024
		GILT vs. Ubx	133,6	135,6	-1,97	-7,924 to 3,984	No	ns	0,8834
	Tibia	WT vs. Control	153,7	149,4	4,353	-3,028 to 11,73	No	ns	0,4661
		WT vs. Ubx+GILT	153,7	142,5	11,2	5,446 to 16,95	Yes	****	< 0,0001
		WT vs. GILT	153,7	145,4	8,355	1,942 to 14,77	Yes	**	0,0047
		WT vs. Ubx	153,7	153,7	0,0245	-6,275 to 6,324	No	ns	> 0,9999
		Control vs. Ubx+GILT	149,4	142,5	6,848	0,3109 to 13,38	Yes	*	0,0356
		Control vs. GILT	149,4	145,4	4,002	-3,121 to 11,13	No	ns	0,5148
		Control vs. Ubx	149,4	153,7	-4,328	-11,35 to 2,693	No	ns	0,4208
		Ubx+GILT vs. GILT	142,5	145,4	-2,845	-8,265 to 2,575	No	ns	0,5807
		Ubx+GILT vs. Ubx	142,5	153,7	-11,18	-16,46 to -5,890	Yes	****	< 0,0001
		GILT vs. Ubx	145,4	153,7	-8,33	-14,33 to -2,335	Yes	**	0,0022
	Tarsus	WT vs. Control	75,36	76,11	-0,7419	-4,227 to 2,744	No	ns	0,9747
		WT vs. Ubx+GILT	75,36	73,31	2,056	-0,6619 to 4,773	No	ns	0,2217
		WT vs. GILT	75,36	74,03	1,332	-1,697 to 4,360	No	ns	0,7293
		WT vs. Ubx	75,36	78,77	-3,404	-6,379 to -0,4289	Yes	*	0,0172
		Control vs. Ubx+GILT	76,11	73,31	2,797	-0,2895 to 5,884	No	ns	0,0934
		Control vs. GILT	76,11	74,03	2,074	-1,290 to 5,437	No	ns	0,4208
		Control vs. Ubx	76,11	78,77	-2,662	-5,978 to 0,6538	No	ns	0,1728
		Ubx+GILT vs. GILT	73,31	74,03	-0,7237	-3,283 to 1,836	No	ns	0,9307
		Ubx+GILT vs. Ubx	73,31	78,77	-5,459	-7,955 to -2,963	Yes	****	< 0,0001
		GILT vs. Ubx	74,03	78,77	-4,736	-7,567 to -1,904	Yes	***	0,0002
	Total	WT vs. Control	359,5	362,6	-3,073	-19,05 to 12,90	No	ns	0,9825
		WT vs. Ubx+GILT	359,5	344,1	15,37	2,911 to 27,82	Yes	**	0,0084
		WT vs. GILT	359,5	353	6,528	-7,354 to 20,41	No	ns	0,6776
		WT vs. Ubx	359,5	368	-8,508	-22,15 to 5,129	No	ns	0,4084
		Control vs. Ubx+GILT	362,6	344,1	18,44	4,291 to 32,59	Yes	**	0,0046
		Control vs. GILT	362,6	353	9,601	-5,818 to 25,02	No	ns	0,4105
		Control vs. Ubx	362,6	368	-5,435	-20,63 to 9,764	No	ns	0,8512
		Ubx+GILT vs. GILT	344,1	353	-8,839	-20,57 to 2,893	No	ns	0,2252
		Ubx+GILT vs. Ubx	344,1	368	-23,88	-35,32 to -12,43	Yes	****	< 0,0001
		GILT vs. Ubx	353	368	-15,04	-28,01 to -2,057	Yes	*	0,0154
Mid-legs	Femur	WT vs. Control	279,1	285,9	-6,752	-19,53 to 6,026	No	ns	0,5745
		WT vs. Ubx+GILT	279,1	197,9	81,15	71,19 to 91,12	Yes	****	< 0,0001
		WT vs. GILT	279,1	262	17,08	5,977 to 28,18	Yes	***	0,0006
		WT vs. Ubx	279,1	208,2	70,93	60,02 to 81,84	Yes	****	< 0,0001
		Control vs. Ubx+GILT	285,9	197,9	87,91	76,59 to 99,22	Yes	****	< 0,0001
		Control vs. GILT	285,9	262	23,83	11,50 to 36,16	Yes	****	< 0,0001
		Control vs. Ubx	285,9	208,2	77,68	65,53 to 89,84	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	197,9	262	-64,07	-73,46 to -54,69	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	197,9	208,2	-10,22	-19,37 to -1,072	Yes	*	0,0212
		GILT vs. Ubx	262	208,2	53,85	43,47 to 64,23	Yes	****	< 0,0001
	Tibia	WT vs. Control	335,8	330,8	5,018	-10,02 to 20,05	No	ns	0,8802
		WT vs. Ubx+GILT	335,8	214,9	120,9	109,2 to 132,7	Yes	****	< 0,0001
		WT vs. GILT	335,8	283,5	52,34	39,28 to 65,41	Yes	****	< 0,0001
		WT vs. Ubx	335,8	236,8	99,06	86,22 to 111,9	Yes	****	< 0,0001
		Control vs. Ubx+GILT	330,8	214,9	115,9	102,6 to 129,2	Yes	****	< 0,0001
		Control vs. GILT	330,8	283,5	47,33	32,82 to 61,84	Yes	****	< 0,0001
		Control vs. Ubx	330,8	236,8	94,04	79,73 to 108,3	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	214,9	283,5	-68,59	-79,63 to -57,55	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	214,9	236,8	-21,88	-32,65 to -11,12	Yes	****	< 0,0001
		GILT vs. Ubx	283,5	236,8	46,71	34,50 to 58,92	Yes	****	< 0,0001
	Tarsus	WT vs. Control	289,1	280,5	8,649	-9,181 to 26,48	No	ns	0,6517
		WT vs. Ubx+GILT	289,1	137,3	151,8	137,9 to 165,7	Yes	****	< 0,0001
		WT vs. GILT	289,1	247,3	41,88	26,39 to 57,37	Yes	****	< 0,0001
		WT vs. Ubx	289,1	144,4	144,7	129,5 to 160,0	Yes	****	< 0,0001
		Control vs. Ubx+GILT	280,5	137,3	143,2	127,4 to 159,0	Yes	****	< 0,0001
		Control vs. GILT	280,5	247,3	33,23	16,02 to 50,44	Yes	****	< 0,0001
		Control vs. Ubx	280,5	144,4	136,1	119,1 to 153,0	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	137,3	247,3	-110	-123,0 to -96,86	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	137,3	144,4	-7,102	-19,87 to 5,666	No	ns	0,5248
		GILT vs. Ubx	247,3	144,4	102,9	88,37 to 117,3	Yes	****	< 0,0001
	Total	WT vs. Control	904,1	897,2	6,915	-34,28 to 48,11	No	ns	0,9895
		WT vs. Ubx+GILT	904,1	550,1	353,9	321,8 to 386,0	Yes	****	< 0,0001
		WT vs. GILT	904,1	792,8	111,3	75,51 to 147,1	Yes	****	< 0,0001
		WT vs. Ubx	904,1	589,4	314,7	279,6 to 349,9	Yes	****	< 0,0001
		Control vs. Ubx+GILT	897,2	550,1	347	310,5 to 383,5	Yes	****	< 0,0001
		Control vs. GILT	897,2	792,8	104,4	64,63 to 144,1	Yes	****	< 0,0001
		Control vs. Ubx	897,2	589,4	307,8	268,6 to 347,0	Yes	****	< 0,0001

		Ubx+GILT vs. GILT	550,1	792,8	-242,6	-272,9 to -212,4	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	550,1	589,4	-39,21	-68,71 to -9,707	Yes	**	0,0037
		GILT vs. Ubx	792,8	589,4	203,4	169,9 to 236,9	Yes	****	< 0,0001
Hind-legs	Femur	WT vs. Control	268,5	271,5	-3,009	-18,55 to 12,53	No	ns	0,9821
		WT vs. Ubx+GILT	268,5	220	48,55	36,43 to 60,66	Yes	****	< 0,0001
		WT vs. GILT	268,5	255,9	12,64	-0,8627 to 26,14	No	ns	0,077
		WT vs. Ubx	268,5	239,2	29,29	16,03 to 42,55	Yes	****	< 0,0001
		Control vs. Ubx+GILT	271,5	220	51,56	37,80 to 65,32	Yes	****	< 0,0001
		Control vs. GILT	271,5	255,9	15,65	0,6513 to 30,64	Yes	*	0,0367
		Control vs. Ubx	271,5	239,2	32,3	17,52 to 47,08	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	220	255,9	-35,91	-47,32 to -24,50	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	220	239,2	-19,26	-30,39 to -8,130	Yes	****	< 0,0001
		GILT vs. Ubx	255,9	239,2	16,65	4,030 to 29,28	Yes	**	0,0041
	Tibia	WT vs. Control	178,7	174	4,692	-9,355 to 18,74	No	ns	0,8799
		WT vs. Ubx+GILT	178,7	248,6	-69,85	-80,80 to -58,90	Yes	****	< 0,0001
		WT vs. GILT	178,7	163	15,66	3,456 to 27,87	Yes	**	0,0055
		WT vs. Ubx	178,7	289,8	-111,1	-123,1 to -99,08	Yes	****	< 0,0001
		Control vs. Ubx+GILT	174	248,6	-74,54	-86,98 to -62,10	Yes	****	< 0,0001
		Control vs. GILT	174	163	10,97	-2,587 to 24,53	No	ns	0,1668
		Control vs. Ubx	174	289,8	-115,8	-129,1 to -102,4	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	248,6	163	85,51	75,20 to 95,83	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	248,6	289,8	-41,22	-51,28 to -31,16	Yes	****	< 0,0001
		GILT vs. Ubx	163	289,8	-126,7	-138,1 to -115,3	Yes	****	< 0,0001
	Tarsus	WT vs. Control	153,1	148,4	4,713	-8,410 to 17,84	No	ns	0,8493
		WT vs. Ubx+GILT	153,1	188,5	-35,39	-45,62 to -25,15	Yes	****	< 0,0001
		WT vs. GILT	153,1	142,6	10,5	-0,8973 to 21,91	No	ns	0,0848
		WT vs. Ubx	153,1	204,9	-51,83	-63,03 to -40,63	Yes	****	< 0,0001
		Control vs. Ubx+GILT	148,4	188,5	-40,1	-51,72 to -28,48	Yes	****	< 0,0001
		Control vs. GILT	148,4	142,6	5,792	-6,873 to 18,46	No	ns	0,6998
		Control vs. Ubx	148,4	204,9	-56,55	-69,03 to -44,06	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	188,5	142,6	45,89	36,25 to 55,53	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	188,5	204,9	-16,45	-25,84 to -7,050	Yes	****	< 0,0001
		GILT vs. Ubx	142,6	204,9	-62,34	-73,00 to -51,68	Yes	****	< 0,0001
	Total	WT vs. Control	600,3	594	6,395	-31,17 to 43,96	No	ns	0,989
		WT vs. Ubx+GILT	600,3	657	-56,69	-85,97 to -27,40	Yes	****	< 0,0001
		WT vs. GILT	600,3	561,5	38,8	6,168 to 71,44	Yes	*	0,012
		WT vs. Ubx	600,3	734	-133,6	-165,7 to -101,6	Yes	****	< 0,0001
		Control vs. Ubx+GILT	594	657	-63,08	-96,35 to -29,82	Yes	****	< 0,0001
		Control vs. GILT	594	561,5	32,41	-3,842 to 68,66	No	ns	0,1008
		Control vs. Ubx	594	734	-140	-175,7 to -104,3	Yes	****	< 0,0001
		Ubx+GILT vs. GILT	657	561,5	95,49	67,91 to 123,1	Yes	****	< 0,0001
		Ubx+GILT vs. Ubx	657	734	-76,92	-103,8 to -50,02	Yes	****	< 0,0001
		GILT vs. Ubx	561,5	734	-172,4	-202,9 to -141,9	Yes	****	< 0,0001

Supplementary table 4: Statistical analyses and significance of the effect of various RNAi experiments on leg length.

Statistical analyses and significance of the effect of various RNAi experiments on leg length.

Supplementary table 5:

1way ANOVA (Jumps)	Mean 1	Mean 2	Mean Diff,	95% CI of diff,	Significant?	Summary	Adjusted P Value
WT vs. M-*gilt*	5,050	4,200	0,8500	0,4261 to 1,274	Yes	****	< 0,0001
WT vs. M-N_*gilt*	5,050	4,050	1,000	0,5544 to 1,446	Yes	****	< 0,0001
M-*gilt* vs. M-N-*gilt*	4,200	4,050	0,1500	-0,2739 to 0,5739	No	ns	0,7859

One-way ANOVA analyses and significance of the effect of *gilt* RNAi on the jump performance.

Chapter 4

Taxon-restricted genes at the origin of a novel trait allowing access to a new environment.

Paper published in *Science*

Co-author

Synopsis of the paper:

Morphological innovations that allow organisms to adapt to new niches and access unexploited opportunities are important triggers of organismal diversification. Theory predicts that these innovations may originate either from the co-option of pre-existing developmental programs or through the emergence of taxon-restricted genes. While the role of gene co-option in the development of morphological innovation has been described, the implication of taxon-restricted genes remains rare. Furthermore the adaptive value conferred by morphological innovations is poorly documented. In this study we investigated the genetic mechanisms and the environmental pressures underlying the evolution of the propelling fan, a novel structure at the tip of the leg exclusively found in the water strider genus *Rhagovelia*.

Summary of results:

Comparative transcriptomics between the three legs in *Rhagovelia antilleana* embryos gave us a list of candidate genes. Five of these genes showed an expression pattern in the part of the leg where the fan develops. The RNAi knockdown of two of these genes resulted in a loss of the fan. Further analysis using genetic databases established that one of these two protein-coding genes where restricted to the Hemiptera and the two copies originated from a recent duplication restricted to the *Rhagovelia* lineage. These newly identified genes were named *geisha* and *mother-of-geisha* in reference to the hand fan of the Japanese Geisha. Finally we tested the adaptive value of the fan using a toboggan set-up mimicking the natural stream water environment where *Rhagovelia* species live. We showed that individuals without the fan are not able to face the current of the stream because the loss of the fan does not allow the bug to efficiently transform energy into movement.

Conclusion:

Our results indicate that taxon-restricted genes, through gene duplication, have acquired new developmental functions in the genus *Rhagovelia* and have contributed to shape the propelling fan, a morphological innovation that allows Rhagovelia species to adapt to the stream water environment.

Author contributions:

Here are the contributions of the different authors:

Emilia Santos: generated transcriptomes, transcriptome analysis, gene and duplication identification, rate of evolution, domain prediction, homologue search, gene cloning, in situ hybridization screening, RNAi knockdown screening, quantitative RT-PCR, acquisition of videos for biomechanical analysis and data analysis, design of the study and redaction of the paper.

Augustin Le Bouquin: gene cloning, in situ hybridization screening, RNAi knockdown screening, acquisition of videos for biomechanical analysis.

Antonin Crumière: gene cloning, reconstruction of semiaquatic bugs phylogeny, acquisition of videos for biomechanical analysis and data analysis, calculation of energy expenditure.

Abderrahman Khila: acquisition of videos for biomechanical analysis, design of the study and redaction of the paper.

EVO-DEVO

Taxon-restricted genes at the origin of a novel trait allowing access to a new environment

M. Emília Santos,[1]* Augustin Le Bouquin,[1,2] Antonin J. J. Crumière,[1] Abderrahman Khila[1]*

Taxon-restricted genes make up a considerable proportion of genomes, yet their contribution to phenotypic evolution is poorly understood. We combined gene expression with functional and behavioral assays to study the origin and adaptive value of an evolutionary innovation exclusive to the water strider genus *Rhagovelia*: the propelling fan. We discovered that two taxon-restricted genes, which we named *geisha* and *mother-of-geisha*, specifically control fan development. *geisha* originated through a duplication event at the base of the *Rhagovelia* lineage, and both duplicates acquired a novel expression in a specific cell population prefiguring fan development. These gene duplicates played a central role in *Rhagovelia*'s adaptation to a new physical environment, demonstrating that the evolution of taxon-restricted genes can contribute directly to evolutionary novelties that allow access to unexploited ecological niches.

Morphological innovations—i.e., lineage-restricted traits that perform evolutionarily new functions—are important triggers of organismal diversification (*1*, *2*). Theory predicts that the evolution of such innovations allows organisms to adapt to new niches and therefore provides access to unexploited ecological opportunities (*3*, *4*). Examples include the evolution of plant flowers, insect wings, butterfly color patterns, turtle shells, lizard dewlaps and toepads, and bird feathers, each of which is thought to have shaped the evolutionary trajectory of the corresponding lineages (*5*–*10*). Such innovations are predicted to originate either through the co-option of conserved preexisting developmental programs or through the emergence of taxon-restricted genes. Although the bulk of evidence supports the role of developmental gene co-option (*11*–*15*), clear examples of how taxon-restricted genes contribute to the evolution of such traits remain scarce (*16*–*19*). Furthermore, how morphological innovations are associated with the potential to occupy new ecological niches is poorly documented (*3*, *4*). To address these questions, we studied the genetic mechanisms and the environmental pressures underlying the evolution of the propelling fan, a structure exclusively found in the water strider genus *Rhagovelia* (Heteroptera, Gerromorpha, Veliidae) (*20*) (Fig. 1A). Our results highlight a central role for taxon-restricted genes, along with conserved gene co-option, in the evolution of the fan. Moreover, using biomechanics assays, we provide empirical evidence that the fan has played an essential role in the adaptation to a new environment, thus acting as an evolutionarily key innovation.

[1]Université de Lyon, Institut de Génomique Fonctionnelle de Lyon, CNRS UMR 5242, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, 46 allée d'Italie, 69364 Lyon, France. [2]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada.
***Corresponding author: emilia.p.santos@gmail.com (M.E.S.); abderrahman.khila@ens-lyon.fr (A.K.)**

The lifestyle of all ~200 known *Rhagovelia* species is characterized by movement on the surface of fast-running water, a niche that is not accessible to most other semiaquatic insects (*20*, *21*) (movie S1). Movement is generated through the rowing motion of the pair of midlegs, which harbor the fans (*20*, *22*, *23*). The fan consists of roughly 20 primary branches, each with thinner secondary branches (Fig. 1A). Unlike *Rhagovelia*, the closely related *Stridulivelia* genus (Fig. 1A and fig. S1) does not have fans, and its members occupy the interface between land and water (*20*). Although *Stridulivelia* are sympatric with *Rhagovelia* and occupy the same streams, they are mostly static on plant leaves and only occasionally perform bursts of fast movement (Fig. 1B and movie S1).

Fan development begins between 120 and 144 hours of embryonic development (~34 to 38% of embryogenesis duration) at 26°C. First instar nymphs hatch with fully functional fans, which are entirely renewed at each of the five molts, indicating persistence of the fan developmental program throughout postembryonic stages (Fig. 1C and fig. S2). To identify the genes underlying fan development, we performed transcriptomic profiling of *Rhagovelia antilleana* embryonic legs (144 hours at 26°C) (table S1). Comparison of mRNA expression identified 87 transcripts that were highly expressed in the midlegs relative to the two other pairs of legs, which do not develop fans (table S1). In situ hybridization screening identified five genes whose expression is restricted to the midleg tip, where the fan develops (Fig. 2A and table S1). These genes are *y* (*yellow*), *cp19* (*cuticular protein 19*), *ccdc174* (*coiled-coil domain-containing protein 174*), and two highly similar unknown genes (c67063_g1 and c68581_g1; Fig. 2, A and B). None of these five genes were expressed in the forelegs or hindlegs in *R. antilleana* or in any of the legs of the fanless *Stridulivelia tersa* and *Oiovelia cunucunumana* outgroup species (Fig. 2A and figs. S1 and S3). The gain of expression of these genes in the *Rhagovelia* lineage thus coincides with the emergence of the fan, suggesting that they are involved in the evolution of this trait.

A homology search detected a unique sequence related to both c67063_g1 and c68581_g1 in five Hemiptera and one Isoptera (tables S3 and S4) but failed to detect putative orthologs in the other major insect lineages such as Diptera, Hymenoptera, and Lepidoptera. Within Gerromorpha, a search using a set of in-house available transcriptomes detected a single putative homolog in nine fanless species (table S2). However, we consistently detected two distinct genes in all three *Rhagovelia* species analyzed (fig. S4 and table S2), suggesting that these two protein-coding genes originated from a recent duplication restricted to the *Rhagovelia* lineage. Phylogenetic analysis clustered the two *Rhagovelia* duplicates together, with c68581_g1 being more similar to the unduplicated gene state found in fanless Gerromorpha, whereas c67063_g1 underwent faster evolution (Fig. 3A and fig. S5). We named c67063_g1 *gsha* (*geisha*) and its paralog c68581_g1 *mogsha* (*mother-of-geisha*), these names being inspired from the hand fan of the Japanese geisha.

Examination of the protein sequences revealed that the duplicates share a signal peptide and two helical regions characteristic of transmembrane proteins. However, we observed divergence in other protein regions, possibly reflecting functional separation (table S5). The high sequence similarity between the two genes makes it impossible to discriminate between their spatial expression patterns using in situ hybridization alone; therefore, we quantified their respective expression levels using quantitative real-time polymerase chain reaction (PCR). Both genes were expressed in the midlegs, although *gsha* was expressed more than three times as highly as *mogsha* (Fig. 3B). Altogether, these results indicate that *mogsha* is a taxon-restricted gene, *gsha* is a *Rhagovelia*-restricted gene, and these genes diverge in protein sequence and expression levels.

RNA interference (RNAi) against *y* resulted in adults with yellowish instead of dark brown fans and with overall light-pigmented bodies (fig. S7A and table S6). Hence, the function of *y* in dark pigment formation (*24*) is conserved in *Rhagovelia*, suggesting that this gene was co-opted as part of the network to darken the cuticle of the fan. RNAi against *cp19* and *ccdc174* did not cause detectable defects despite repeated attempts, perhaps owing to functional redundancy between these genes and/or other members of the cuticular gene family (*25*). For *gsha* and *mogsha,* their high sequence similarity prevented the design of RNAi reagents specific to the mRNA of one or the other (table S7). Simultaneous knockdown of both *gsha* and *mogsha* resulted in the loss of the fan (Fig. 3C, fig. S7B, and tables S6 and S7). *gsha/mogsha* RNAi individuals were viable, and we failed to detect defects in any other structures, including the claws that are directly connected to the fan (Fig. 3C and fig. S7B). High-speed videography revealed that the *gsha/mogsha* RNAi individuals still deployed and retracted the

Downloaded from http://science.sciencemag.org/ on October 19, 2017

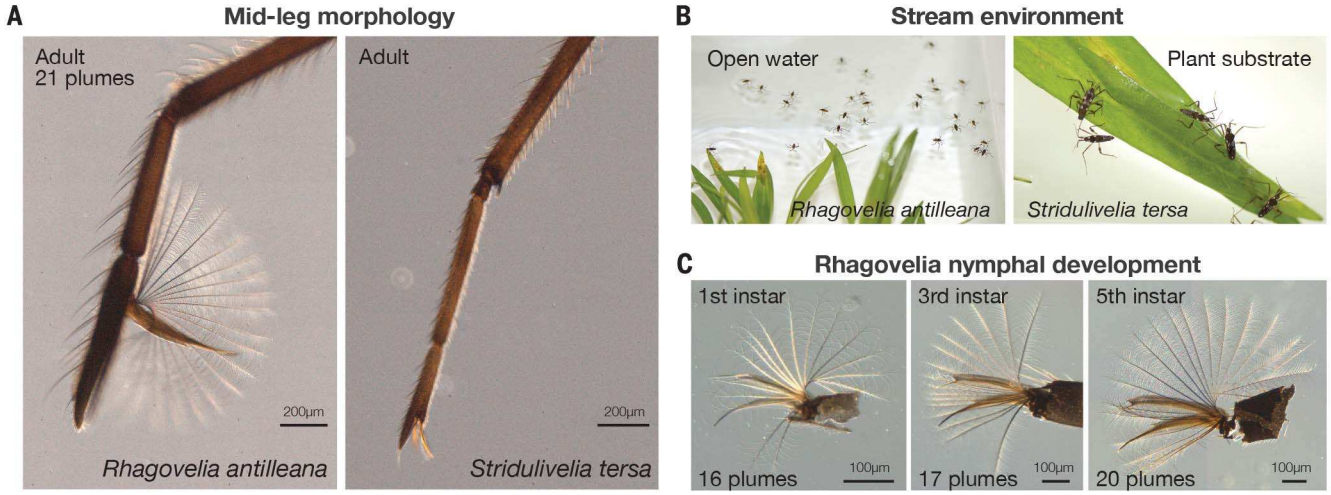


Fig. 1. Morphology, habitat correlation, and nymphal development of the *Rhagovelia* fan. (**A**) Midleg morphology comparison between *R. antilleana* and the outgroup species *S. tersa*. *Rhagovelia* tarsi have longer and wider claws and are forked to host the fan, which can be deployed and retracted following leg movements (movie S2). (**B**) Habitat and behavior of *Rhagovelia* (left) and *Stridulivelia* (right). (**C**) Nymph developmental series, showing an increase in fan size and in the number of branches across molting stages.

claws and the remnants of the fan, just like control individuals did (movie S2). These results suggest that *gsha* and *mogsha* function in nymphs solely for fan formation, consistent with their specific expression in the cell population prefiguring fan development. The association between the emergence of the fan, the gain of expression of a set of conserved genes, and the gene duplication implies that these molecular events participated in fan evolution. In addition, our data directly connect two taxon-restricted genes, *gsha* and *mogsha*, to the development and evolution of a taxon-restricted structure.

A major goal of evolutionary biology is not only to identify the genetic mechanisms underlying novel traits, but also to determine the impact of evolutionary innovations on organismal evolutionary history (*26*). To investigate whether the emergence of the fan and the duplication of *gsha/mogsha* contributed to evolutionarily novel adaptive functions, we tested the impact of different fan phenotypes on locomotory performance. First, we measured performance on still water, mimicking the riparian environment with no current. We compared individuals of *Stridulivelia tersa* (a fanless closely related species that inhabits this riparian habitat), *Rhagovelia antilleana* (normal fan), *gsha/mogsha*-RNAi *R. antilleana* (reduced fan), and *R. antilleana* individuals with surgically ablated fans. *S. tersa* animals were more than three times as fast as *R. antilleana* animals, with no significant differences between the three *R. antilleana* groups (Fig. 4A and table S8). In contrast, we detected a significant inverse correlation between fan and stroke frequency, with normal *R. antilleana* (normal fan) using the lowest stroke frequency, followed by *gsha/mogsha*-RNAi *R. antilleana* (reduced fan), ablated *R. antilleana*, and finally *S. tersa* (Fig. 4B and table S8). These results show that, in the still-water environment, the fan does not increase speed but rather allows

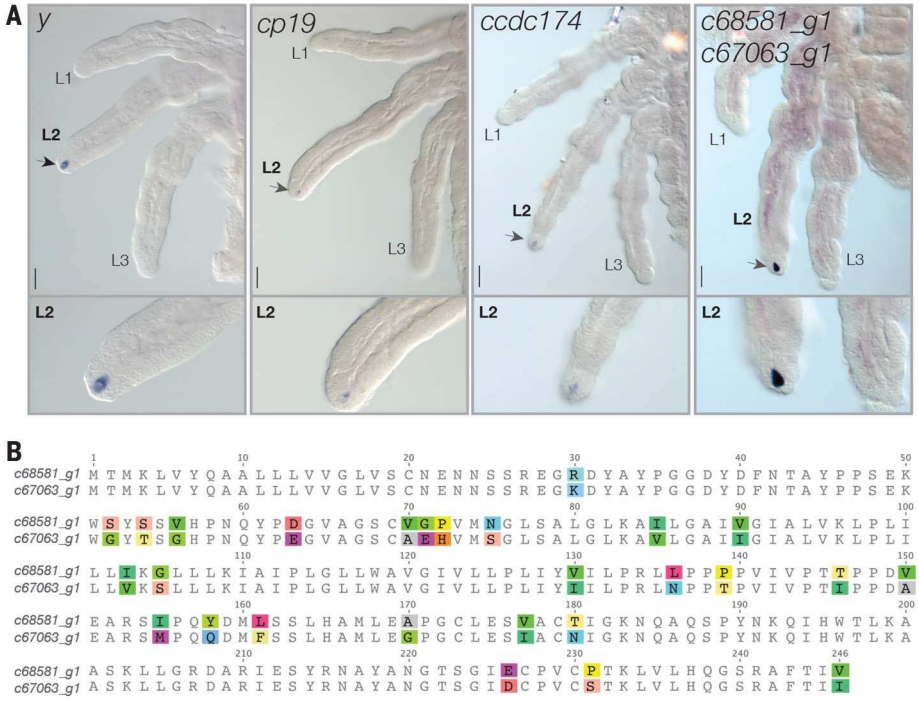


Fig. 2. Fan expression profiling and protein alignment of *gsha* and *mogsha*. (**A**) In situ hybridization showing the expression of all five genes at the tips of the midleg (leg 2, L2). Scale bar, 100 μm. (**B**) Protein alignment of c68581_g1 (*mogsha*) and c67063_g1 (*gsha*), with divergent amino acids highlighted in color (nucleotide alignment in fig. S4). Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

the insect to maintain fast speed while using a lower stroke frequency, therefore using less energy over a given distance (*22*).

Next, we challenged these groups of individuals in a setup where they were forced to row against a slow (0.2 m/s) or fast (0.3 m/s) water current (Fig. 4C and movie S3), simulating *Rhagovelia*'s stream environment. Although *S. tersa* animals were the fastest on still water, they failed to ascend the stream setup (Fig. 4C, fig. S9, and table S9). Normal *R. antilleana* ascended both currents without difficulty (Fig. 4C,

Downloaded from http://science.sciencemag.org/ on October 19, 2017

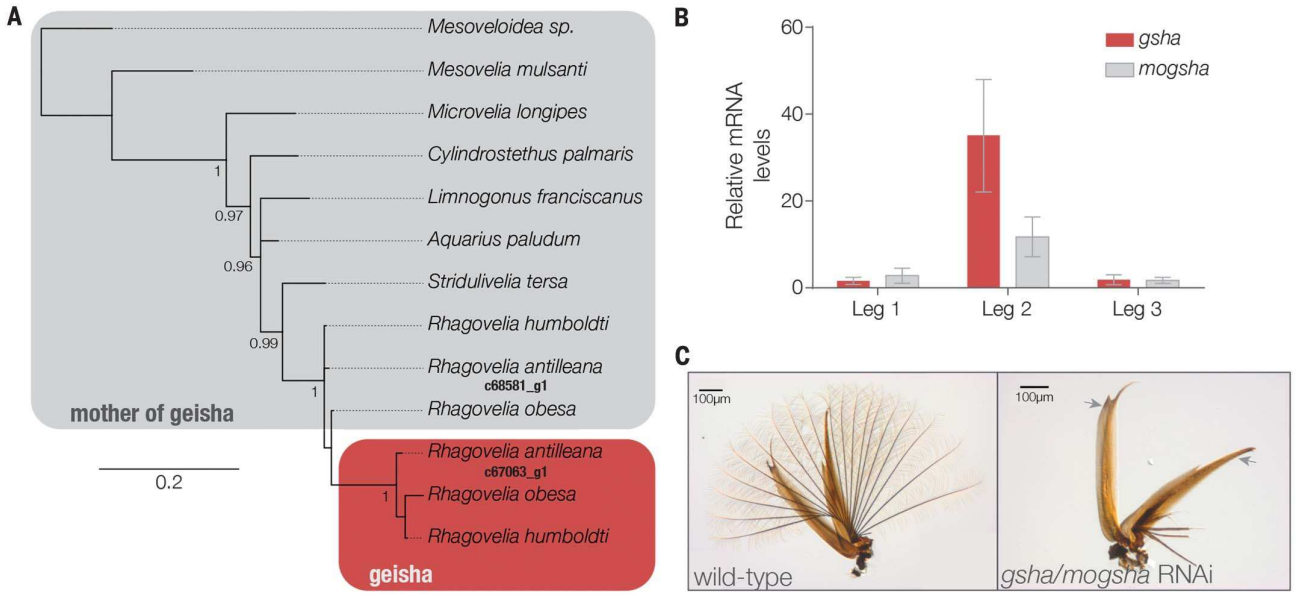


Fig. 3. Sequence evolution, expression, and functional characterization of *gsha* and *mogsha*. (**A**) Bayesian phylogenetic reconstruction showing that *mogsha* is more similar to the taxon-restricted unduplicated copy (gray box), whereas *gsha* is a *Rhagovelia*-specific duplicate that has evolved at a higher rate (red box; note the branch length). Node values represent the posterior probability, and the scale bar indicates the genetic distance between sequences. (**B**) Gene-specific quantitative PCR showing *gsha and mogsha* expression in the midleg (leg 2; $n = 3$) of fourth instar nymphs. Error bars, SEM. (**C**) *gsha/mogsha* RNAi-mediated knockdown resulted in a severe reduction of the fan ($n = 27$), but the claws (arrows) and leg morphology (figs. S7 and S8) were unaffected.

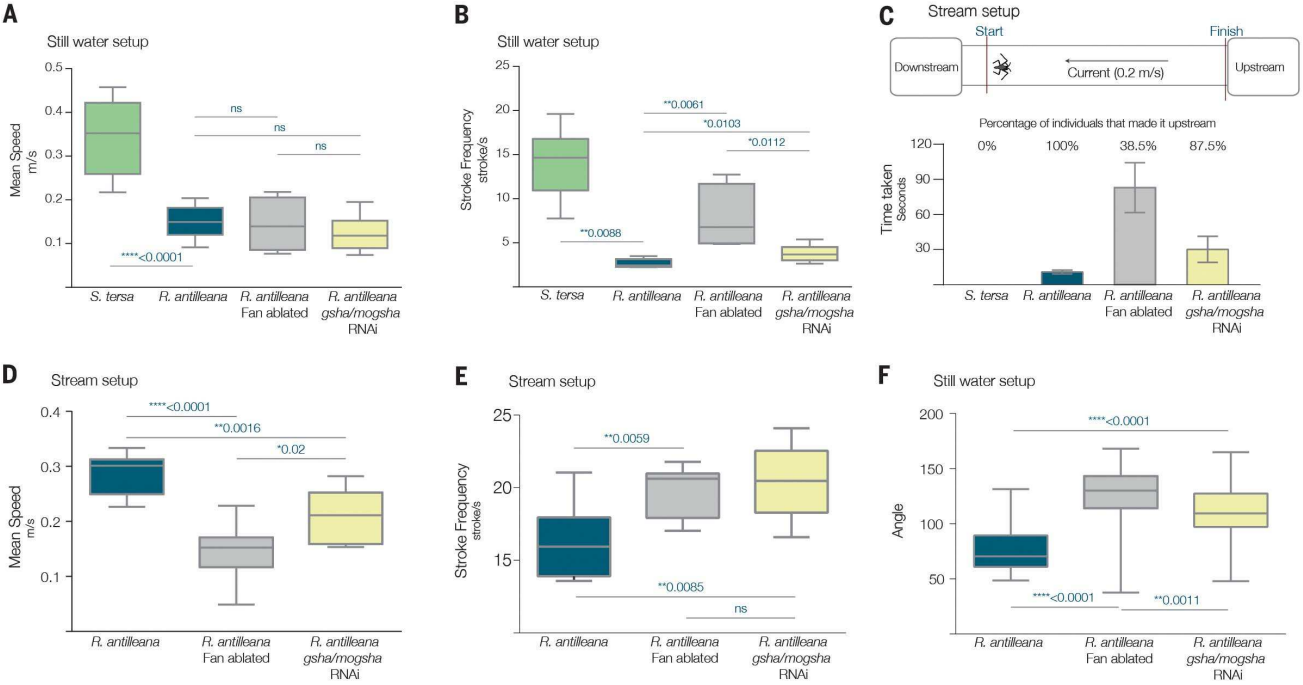


Fig. 4. Comparative analysis of locomotion parameters. (**A** and **B**) Boxplots for the still water setup, comparing (A) mean speed and (B) stroke frequency among *S. tersa* (green; $n = 10$), *R. antilleana* (blue; $n = 7$), *R. antilleana* with ablated fans (gray; $n = 4$), and *gsha/mogsha*-RNAi *R. antilleana* (yellow; $n = 9$). The range of severity of *gsha/mogsha*-RNAi phenotypes can be found in fig. S7. (**C**) Scheme of the stream setup (top) (movie S3), with the percentage of individuals succeeding in rowing against the current, together with their corresponding arrival time, shown below. Error bars, SEM. (**D** and **E**) Boxplots for the stream setup, comparing (D) mean speed and (E) stroke frequency among *R. antilleana* (blue; $n = 10$), *R. antilleana* with ablated fans (gray; $n = 8$) and *gsha/mogsha*-RNAi *R. antilleana* (yellow; $n = 8$). (**F**) Angles of *R. antilleana* turns in still water (*R. antilleana*, $n = 8$; *gsha/mogsha*-RNAi, $n = 8$; *R. antilleana* with fan ablated, $n = 8$). Each boxplot spans the minimum and maximum values (whiskers), with medians shown by the horizontal lines and first and third quartiles delineated by the box. Student's *t* test or the Mann-Whitney test was performed for each pairwise comparison. *P* values are shown next to asterisks. ns, not significant.

Downloaded from http://science.sciencemag.org/ on October 19, 2017

fig. S9, and table S9), suggesting a major role for the fan in running waters. Consistently, *R. antilleana* animals with ablated fans failed to ascend the fast stream and displayed a severely decreased performance in the slow stream (Fig. 4C, fig. S9, and table S9). *geisha/mogsha*-RNAi individuals (reduced fans) delivered an intermediate performance between those of the normal and fan-ablated *Rhagovelia*; the majority of the individuals failed in the fast stream but ascended the slow stream (Fig. 4C, fig. S9, and table S9). Furthermore, we found differences both in speed and stroke frequency between the three *R. antilleana* groups (Fig. 4, D and E). As observed in still water, *R. antilleana* with surgically or genetically altered fans used higher stroke frequencies. However, this increase in stroke frequency was not sufficient to compensate for the lack or reduction of the fan (Fig. 4, D and E). These data show that although the fan does not increase speed in still water, it is required to row against the current in stream environments. The fan increases the midleg area that is in contact with water, which increases the efficiency of each stroke by transforming the same amount of energy into more movement, thereby reducing stroke frequency (fig. S10). Sexual dimorphism did not influence the function of the fan, and its effect on the locomotory performance was the same in both sexes (fig. S11).

Last, we evaluated the ability to perform sharp turns, a characteristic behavior of *Rhagovelia* species. We found that normal *R. antilleana* turns were consistently initiated by fan deployment, with an average angle of 77° (Fig. 4F and table S8). *gsha/mogsha*-RNAi and fan-ablated individuals turned less efficiently, with an average angle of 111° and 126°, respectively (Fig. 4F and table S8). Taken together, our results show that the fan sustains frequent fast movement and increases maneuverability in an energy-demanding environment, thus playing a key role in the lifestyle of *Rhagovelia* species.

Collectively, our data provide important insights into the developmental genetic, evolutionary, and ecological mechanisms underlying the emergence of evolutionary innovations. We show that the genetic changes that participated in the evolution of the fan provided *Rhagovelia* with distinctive locomotory abilities required to adapt to fast-flowing water environments. Our results indicate that this trait evolved through the redeployment of conserved developmental programs (pigmentation and cuticle deposition), but also through a *Rhagovelia*-specific gene duplication of a taxon-restricted gene. Although the phylogenetic order of occurrence of these molecular events remains unknown, we provide strong evidence that evolutionary novelties can appear through multiple genetic events in both conserved and taxon-restricted genes. Animals lacking *gsha* and *mogsha*, which display fan rudiments, consistently delivered intermediate performance in stream environments, suggesting that fan rudiments are advantageous. We hypothesize that, initially, the gain of a preexisting developmental program in the midleg led to the emergence of fan rudiments that became later optimized by a gene duplication event, lending experimental support to the long-lasting hypothesis that complex evolutionary novelties emerge gradually from simpler structures (*27*–*30*).

REFERENCES AND NOTES

1. G. P. Wagner, V. J. Lynch, *Curr. Biol.* **20**, R48–R52 (2010).
2. M. Pigliucci, *Philos. Sci.* **75**, 887–898 (2008).
3. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).
4. J. B. Losos, *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (Univ. of California Press, 2009).
5. M. Averof, S. M. Cohen, *Nature* **385**, 627–630 (1997).
6. V. A. Albert, D. G. Oppenheimer, C. Lindqvist, *Trends Plant Sci.* **7**, 297–301 (2002).
7. R. O. Prum, A. H. Brush, *Q. Rev. Biol.* **77**, 261–295 (2002).
8. S. B. Carroll, J. K. Grenier, S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Science, ed. 1, 2001).
9. R. W. R. Wallbank *et al*., *PLOS Biol.* **14**, e1002353 (2016).
10. J. E. Moustakas-Verho, J. Cebra-Thomas, S. F. Gilbert, *Curr. Opin. Genet. Dev.* **45**, 124–131 (2017).
11. S. B. Carroll, *Cell* **134**, 25–36 (2008).
12. V. J. Lynch, G. P. Wagner, *Evolution* **62**, 2131–2154 (2008).
13. A. Monteiro, M. D. Gupta, in *Genes and Evolution*, V. Orgogozo, Ed. (Elsevier, ed. 1, 2016), vol. 119, pp. 205–226.
14. W. J. Glassford *et al*., *Dev. Cell* **34**, 520–531 (2015).
15. A. Khila, E. Abouheif, L. Rowe, *Evolution* **68**, 2159–2170 (2014).
16. K. Khalturin, G. Hemmrich, S. Fraune, R. Augustin, T. C. G. Bosch, *Trends Genet.* **25**, 404–413 (2009).
17. H. Kaessmann, *Genome Res.* **20**, 1313–1326 (2010).
18. M. Margulies *et al*., *Nature* **437**, 376–380 (2005).
19. W. C. Jasper *et al*., *Mol. Biol. Evol.* **32**, 334–346 (2015).
20. N. M. Andersen, *The Semiaquatic Bugs (Hemiptera: Gerromorpha)* (Scandinavian Science Press, 1982).
21. D. N. Padilla-Gil, F. F. F. Moreira, *Zootaxa* **3640**, 409–424 (2013).
22. A. J. J. Crumière *et al*., *Curr. Biol.* **26**, 3336–3342 (2016).
23. M. E. Santos, C. S. Berger, P. N. Refki, A. Khila, *Brief. Funct. Genomics* **14**, 384–395 (2015).
24. P. J. Wittkopp, P. Beldade, *Semin. Cell Dev. Biol.* **20**, 65–71 (2009).
25. R. Kafri, M. Springer, Y. Pilpel, *Cell* **136**, 389–392 (2009).
26. A. M. Dean, J. W. Thornton, *Nat. Rev. Genet.* **8**, 675–688 (2007).
27. A. Monteiro, *BioEssays* **34**, 181–186 (2012).
28. D.-E. Nilsson, S. Pelger, *Proc. Biol. Sci.* **256**, 53–58 (1994).
29. W. Gehring, M. Seimiya, *Ital. J. Zool.* **77**, 124–136 (2010).
30. C. Darwin, *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life* (J. Murray, London, 1859).

ACKNOWLEDGMENTS

We thank W. Salzburger, F. Payre, and the Khila laboratory for discussion and useful comments on the manuscript; F. Moreira for help with species identification; and S. Viala for technical help. Central processing unit resources were provided by Pôle Scientifique de Modélisation Numérique at Ecole Normale Supérieure de Lyon. This work was supported by a European Research Council Consolidator Grant to A. Khila (ERC-CoG #616346), a Swiss National Foundation Early Postdoc Mobility Fellowship to M.E.S., and a Seventh Framework Programme Marie Curie Intra-European Fellowship to M.E.S. DNA sequences were deposited in National Center for Biotechnology Information databases, and the accession numbers are available in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/358/6361/386/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S12
References (*31*–*57*)
Movies S1 to S3

21 March 2017; accepted 13 September 2017
10.1126/science.aan2748

Downloaded from http://science.sciencemag.org/ on October 19, 2017

Science

Taxon-restricted genes at the origin of a novel trait allowing access to a new environment

M. Emília Santos, Augustin Le Bouquin, Antonin J. J. Crumière and Abderrahman Khila

Science **358** (6361), 386-390.
DOI: 10.1126/science.aan2748

Fans enable water strider adaptation

Genomes of closely related organisms are similar but contain variations that enable different phenotypes and lifestyles. The origin of evolutionary innovations, such as insect wings and bird feathers, poses a challenge to evolutionary biology because the de novo emergence of complex traits cannot easily be explained by natural selection. Water-walking *Rhagovelia* insects evolved a propelling fan on the middle leg that is associated with life on fast-flowing streams. Santos *et al.* discovered that the *geisha* and *mother-of-geisha* genes underlie fan development and evolution and that this evolutionary innovation is essential to the adaptation of *Rhagovelia* to its environment. Thus, the evolution of taxon-restricted genes can contribute directly to taxon-restricted novelties that allow access to unexploited ecological niches.

Science, this issue p. 386

ARTICLE TOOLS http://science.sciencemag.org/content/358/6361/386

SUPPLEMENTARY MATERIALS http://science.sciencemag.org/content/suppl/2017/10/19/358.6361.386.DC1

REFERENCES This article cites 51 articles, 5 of which you can access for free
http://science.sciencemag.org/content/358/6361/386#BIBL

PERMISSIONS http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. 2017 © The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. The title *Science* is a registered trademark of AAAS.

Downloaded from http://science.sciencemag.org/ on October 19, 2017

www.sciencemag.org/content/358/6361/386/suppl/DC1

Supplementary Materials for

Taxon-restricted genes at the origin of a novel trait allowing access to a new environment

M. Emília Santos,* Augustin Le Bouquin, Antonin J. J. Crumière, Abderrahman Khila*

*Corresponding author. Email: emilia.p.santos@gmail.com (M.E.S.); abderrahman.khila@ens-lyon.fr (A.K.)

Published 20 October 2017, *Science* **358**, 386 (2017)
DOI: 10.1126/science.aan2748

This PDF file includes:

Materials and Methods
Figs. S1 to S11
Tables S2 to S12
Caption for Table S1
Captions for Movies S1 to S3
References

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/358/6361/386/suppl/DC1)

Table S1
Movies S1 to S3

Materials and Methods

Animal rearing

Rhagovelia antilleana, *Stridulivelia tersa* and *Oiovelia cunucunumana* are reared in water tanks at 26°C (14h light/10h dark cycle) and fed with live crickets. Each water tank contained floating pieces of Styrofoam (*R. antilleana* and *O. cunucunumana*) or plastic plants (*S. tersa*) providing substrate for females to lay eggs, which were then used for dissections and RNA extractions.

Reconstruction of semi-aquatic insects phylogeny

Phylogenetic reconstruction was conducted using 14 molecular markers (*22*) with the software Geneious version 7.1.9. These sequences were previously shown to provide robust semi-aquatic insect phylogenies (*22*), here we added two species (*Velia caprai* and *Rhagovelia antilleana*) to this dataset. These sequences from these two species can be found under the GenBank accession numbers MF44870 to MF448723. Sequences were aligned with MAFFT version 1.3.1 plugin (*31*) using default parameters; the corrected alignment is available upon request. The species trees were constructed with the plugins MrBayes version 3.2.6 (*32*) (one million generations and 25% burnin) and PhyML version 3.0 (*33*) using the GTR model with 100 bootstraps.

Differential gene expression using RNAseq

Leg dissections, RNA extractions and sequencing

To identify genes required for fan development we generated leg transcriptomes of 144 hours old (26°C) *R. antilleana* embryos. We chose this stage because it marks the start of fan development and is likely to represent the timing during development where the genes required for the fan are active. Embryos were treated with 25% bleach for three minutes. The chorion and vitellin membranes were removed with fine forceps in 1X PBS buffer and 0.05% Tween 20. The three pairs of legs (Leg 1, Leg 2 and Leg 3) were dissociated from the thorax with the help of very fine needles; each pair of legs was incubated immediately in TRIzol (Invitrogen). For each leg sample we had three replicates containing legs of 100 individuals. RNA extractions were performed according to manufacturer protocol. The concentrations were assessed using the Qubit 2.0 Fluorometer (Invitrogen) and the quality of RNA samples with the Agilent Technologies 2200 TapeStation. Library construction and sequencing of RNA extracted from three replicated per leg pair was performed by Beckman Coulter (USA). The samples were sequenced using Illumina TrueSeq RNAseq technology with a single read length of 100bp. The sequences were deposited at Sequence Read Archive (SRA) under the BioSample accessions: SAMN06689961, SAMN06689962 and SAMN06689963.

Reference transcriptome assembly

We used a transcriptomic approach (RNAseq) to identify genes differentially expressed between the different developing leg pairs. For this purpose we assembled a leg reference transcriptome using a

concatenation of the reads of the three pair of legs of one replicate. Read filtering, cleaning and adaptor removal was done using Trimmomatic version 0.32 with sliding window of three base pairs with a minimum quality of 15 (*34*). Any base pair at the beginning or end of the read which quality would be under 15 was removed and only reads with a minimum length of 50bp were kept. The assembly of these reads was conducted with Trinity r20140717, the reads were normalized to a maximum coverage of 50 and the path reinforcement distance was set to 50 (*35*). Gene ontology (GO) annotation was conducted using Blast2GO version 4.0.7 (*36*). BlastX searches were performed against the invertebrate non-redundant protein NCBI database using default parameters whilst setting the e-value to $1.0x10^{-20}$ and the number of hits to 5 (*37*). The mapping and annotation steps were done using default settings. This Transcriptome Shotgun Assembly project has been deposited at GenBank under the accession GFOS00000000. The version described in this paper is the first version, GFOS01000000.

Differential gene expression analysis

The reads from each leg replicate were mapped against the leg reference transcriptome using Bowtie2 (with the following parameters: --no-mixed --no-discordant --gbar 1000 --end-to-end) (*38*). The mapping results were reported in SAM format. The abundance of reads per trinity "gene" (number of reads per "gene") was calculated with RSEM tool version 1.2.15 (*39*), and then concatenated into one expression matrix. The expression matrix was analyzed with the R packages EDGER version 3.4.2 (*40*) to test for significant differences in gene expression between legs (FDR>0,05). The differences in expression between the first and the second leg, and the third and the second leg were clustered in a heatmap showing the relationships among the differentially expressed transcripts using the script analyze_diff_expr.pl provided on the Trinity package version r20140717 (*41*). The identity, annotation and level of expression of differentially expressed transcripts together with the list of screened genes can be found in the auxiliary file Table S1.

Cloning and *in situ* hybridization

Total RNA from *R. antilleana*, *S. tersa* and *O. cunucunumana* was extracted from different embryonic and nymphal stages. cDNA synthesis was then performed for each species using the SuperScript III First-Strand kit according to manufacturer instructions (Invitrogen). The gene fragments were amplified by PCR using GoTaq G2 DNA polymerase (Promega) following the manufacturer's guidelines. Specific primers for each gene screened were designed using the primer3 (*42*) plugin in Geneious version 7.0.6 (*43*) using the *R. antilleana* leg reference transcriptome described above. Primer sequences can be found in Table S1. The PCR fragments were cloned into pGEM-T vector using the pGEM-T cloning kit (Promega). Plasmid extractions were done with QIAprep spin Miniprep Kit (Qiagen). RNA probes were synthetized with the DIG RNA labeling kit (SP6/T7) (Roche) using M13 PCR products purified with MinElute PCR purification Kit (Qiagen). The sequence and direction of the insert was confirmed with Sanger sequencing using the M13 forward primer. *In situ* hybridization was performed on embryos as described in (*44*). The embryos with staining were mounted in 80% Glycerol and imaged with an AxioObserver Zeiss microscope. The identity of

the genes screened in *R. antilleana* and the description of their respective expression patterns can be found in Table S1. For *yellow, ccdc174* and *cp19* we confirmed their true identity with a reciprocal blast search between the NCBI non-redundant database and our leg reference transcriptome. For the genes cloned in *S. tersa* and *O. cunucunumana* primers can be found in Table S10.

***geisha* and *mother-of-geisha* phylogeny and domain prediction**

We constructed a gene tree in order to confirm the orthologous and paralogous relationships of both duplicates. Sequences from *geisha* and *mother-of-geisha* from semi-aquatic insects were retrieved from in house transcriptomic databases (species names, gene names and accession numbers are available in Table S2). Protein sequences were aligned with MAFFT version 1.3.1 (*31*) using default parameters. The alignment can be made available upon request. Phylogenetic analyses were performed with MrBayes (10^6 generations, Fig. 3) (*32*) and with PhyML (100 bootstraps, Fig. S5) (*33*) using default parameters. Both phylogenies show that the *geisha* and *mother-of-geisha* form two different groups and that *geisha* is evolving at a higher rate. *geisha* and *mother-of-geisha* domain predictions (Table S5) were determined using the software SMART sequence analysis (http://smart.embl-heidelberg.de/)(*45*).

Homologue search outside semi-aquatic insects

To identify the contigs (c68581_g1 and c67063_g1) for which the transcriptomic analysis retrieved no annotation, we performed blast searches against the NCBI non-redundant database (*37*). Only genes with an e-value lower than 10^{-4} were considered true homologs (*18*, *46*). BlastN and TBlastX searches did not retrieve homologous sequences. BlastX, TBlastN retrieved one homolog in *Cimex lectularius*, which also belongs to the Hemiptera order. The BlastP search revealed three homologs belonging to the Hemiptera (*Cimex lectularius, Acyrthosiphon pisum, Diaphorina citri*). All these are predicted uncharacterized proteins, and only one homolog is found per species (Table S3). When we use these Hemipteran sequences as a query for further BlastP searches we further found homologs of *Homalodisca liturata, Halyomorpha halys* and *Zootermopsis nevadensis*. These species belong to the order Hemiptera, except for *Z. nevadensis* that is part of the order Isoptera (Table S3). We performed a protein alignment of the identified protein sequences using MAFFT version 1.3.1 (*31*), the resulting sequence distance matrix is shown in Table S4. The protein alignment can be found in Fig. S6.

Rates of evolution analysis of gene duplication history

The protein alignments were performed using Clustal Omega (*47*) and the DNA codon based alignments calculated with PAL2NAL (*48*) using the DNA sequences shown in Table S2 and their respective protein sequences. The species tree was estimated with FastTree version 1.0 (*49*) using default parameters and the codon-based alignments. To test if the rates of evolution differed between the different branches of the phylogenetic tree we calculated the dN/dS value using the branch models from Codeml implemented in the PAML software (*50*). The dN/dS values were averaged across sites (NSsites=0). Three models of molecular evolution were computed, the one-ratio model (model=0), the free-ratio model (model=1) and the two-ratio

model (model=2). In the one-ratio model there is only one average dN/dS value for all the branches of the tree, therefore all lineages evolved at the same rate. The free-ratio model allows for different dN/dS ratios for each branch of the tree, lineages are evolving at different rates. The two-ratio model allows for two dN/dS values that are different between the branch of interest (in this case the *Rhagovelia* lineage) and the background branches. In all models, sites with ambiguous data were removed (cleandata=1). The free-ratio and the two-ratio models were both compared to the one-ratio model (null hypothesis) using a likelihood ratio test (LTR) using six and one degrees of freedom respectively.

To test if there were positively selected genes (dN/dS >1) in the *Rhagovelia* lineage, dN/dS values for each ortholog were calculated using the branch-site models. Branch-site models account for the fact that selection does not act evenly across the gene sequence but instead it acts in specific amino acids. We compared branch-site Model A (Model=2, NSsites=2, fixomega=0), that allows for a dN/dS value higher than one in the lineage of interest (*Rhagovelia*), to the Model A1 (Model=2, NSsites=2, fixomega=1) which allows only one dN/dS for all branches. For each ortholog the two models were compared using an LTR test with one degree of freedom. Despite the fact that *geisha* is more divergent in terms of coding sequence than *mother-of-geisha*, we detected no signs of positive selection either in the *Rhagovelia* lineage in general or in the specific duplicates.

Gene expression analysis using qPCR

The levels of expression of *geisha* and *mother-of-geisha* were characterized by means of quantitative Real-Time PCR (qPCR) using gene specific primers. Expression of both duplicates was measured in the three legs (three replicates each) of instar nymphs of *Rhagovelia antilleana*. For each leg replicate RNA was extracted from twenty 4th instar nymphs (40 legs per replicate). The comparative CT (cycle threshold) method (*51*) was used to calculate differences in expression between the different samples using the *ribosomal protein S18* (*rps18*) and the *ribosomal protein L13a* (*rpl13a*) as endogenous controls. RQ (relative quantities) values were calculated for each sample and the differential expression between legs was calculated using the first leg as the tissue reference. All reactions had a final cDNA concentration of 5ng/µl and a primer concentration of 200nM. The reactions were run on a CFX96 Real-Time PCR detection system (Biorad) using the iTaq Universal SYBR Green Supermix (Biorad) with an annealing temperature of 58ºC and following the manufacturer protocol. Primers were designed with the software GENSCRIPT REAL-TIME PCR primer design available at https://www.genscript.com/ssl-bin/app/primer (Table S11). Primer efficiencies of the experimental primers - *geisha* (100%) and *mother-of-geisha* (98%) - were comparable to the efficiency of the endogenous controls *rps18* (97%) and *rpl13a* (93%).

Nymphal RNAi

Double stranded RNA (dsRNA) was produced for *y, ccdc174, cp19* and *gsha/mogsha*. Forward and reverse primers both containing the T7 RNA polymerase promoter were designed and used to amplify the T7 PCR fragment using the previously cloned fragments as templates. The gene specific T7 PCR fragment was

purified and used as a template to synthesize the dsRNA as mentioned in (*44*, *52*). The resulting dsRNA was purified using an RNeasy purification kit (Qiagen) and eluted in Spradling injection buffer (*53*) at a concentration of 5µg/µl. For primer information please see Tables S12. A mix of developmental stages of *R. antilleana* nymphs (1st, 2nd and 3rd instar stages) were used for each injection (Table S6). Injection was performed as described in (*15*). Injected nymphs were placed in water tanks, fed daily and allowed to develop until adulthood. Adult legs and fans were dissected and mounted on slides in Hoyer's medium. Mounted fans were imaged in an AxioObserver Zeiss microscope. Measurements of body size, and of each leg segment of each pair of legs were recorded on a dissecting SteREO Discovery.V12 Zeiss stereomicroscope using the Zen Pro 2011 software.

Behavioral assays

Standard setup

Locomotion assays were performed using *Stridulivelia tersa* (n = 10), *Rhagovelia antilleana* (n = 7), *Rhagovelia antilleana* individuals where the fans were surgically removed with fine forceps without disrupting the claws (n = 4) and the *gsha/mogsha* dsRNA adults showing a fan phenotype (n = 9). Each individual was put in a still water tank and filmed at 2000 frames per second using the Phantom Miro M310 Digital High Speed Camera (Vision research) and PCC Software (Ametek). Three videos were recorded per individual and a reference scale of two centimeters was used. The videos were analyzed using the TEMA 3.7 software (Images system) and individuals were tracked to extract velocity and maneuverability (angle) values. Whilst tracking we counted the number of strokes for each individual and measured the total duration of the video in order to calculate stroke frequency. The average value for all these parameter is available in Table S8. Statistical significance of the results was determined with a Student's t-test or a Mann-Whitney test (if the variances were significantly different).

Stream setup

To subject the individuals to a directional steady current we designed a setup with a water source bucket and a recipient bucket connected by a water canal (150 cm long and 5 cm wide) where a unidirectional flow of water runs from the source bucket to the recipient bucket (Fig. 4C). Once in the recipient bucket the water is pumped, using aquaria filter, back to the source bucket providing the system with a continuous circular flow of water. The water current speed is controlled by changing the horizontal angle of the canal such that the steeper the angle the faster the water flow. The current speed is calibrated by the tracking, using high-speed videos, of a floating object across the water canal.

Stream setup: win-fail experiment

Assays were performed in two different current settings – 0.2m/s and 0.3m/s – using *Stridulivelia tersa* (n = 15), *Rhagovelia antilleana* (n = 10), *Rhagovelia antilleana* individuals where the fan was surgically ablated (n = 8) and the *gsha/mogsha* dsRNA adults showing a fan phenotype (n = 8). For this setup we have one

individual less in the *gsha/mogsha* dsRNA, since one female died in between behavioral experiments. In this setup we released all individuals on the same starting point (cm 60 on the canal) and scored the individuals that would ascend the current and arrive to the upstream bucket together with their arrival time. If the individuals would fall on the recipient bucket or fail to reach the source bucket after 5 minutes of struggle we would consider that they failed to ascend the current and would record the time (Table S9, Fig. S9).

Stream setup: speed, stroke frequency and energy expenditure experiment

Assays were performed in a current speed of 0.2m/s using *Rhagovelia antilleana* (n = 10), *Rhagovelia antilleana* individuals where the fan was surgically removed (n = 8) and the *gsha/mogsha* dsRNA individuals showing a reduced fan phenotype (n = 8). All individuals were released on the same starting point (cm 60) and filmed with a high-speed camera at 2000 frames per second. Three videos were recorded per individual a reference scale of three centimeters was used. Analysis of the video was performed as mentioned in the standard setup methodology. Estimates of energy expenditure were performed as described in (*22*). Results of the analysis can be found in Fig. 4, Table S8 and Fig. S10. Statistical significance of the results was determined with a Student's t-test or a Mann-Whitney test (if the variances were significantly different).

Supplementary tables

Table S1. Differential gene expression results with their respective annotation. Due to the large size of the file, the table is provided as an auxiliary file. The *Rhagovelia antilleana* transcript sequence, the annotation and the fold change in expression are provided in the first 10 columns. In addition, this table contains the information of the genes screened, the primers used to amplify them and the accession numbers of the sequences of the cloned fragments, together with a description of the *in situ* hybridization expression pattern. The assembly used to map the leg transcriptome reads is stored in the NCBI database under the accession number GFOS00000000, the respective transdecode peptide predictions file can be provided upon request. The file has three tabs for three different expression comparisons. On the first tab we can find the comparison between the three legs, where we screened for transcripts that were more expressed in the second leg when compared to the two other legs. On the second tab we screened for genes that were more expressed on the second leg when compared to the first. Finally, on the third tab are the genes that are more expressed on the second leg when compared to the third. We screened for genes that had a length longer than 500 base pairs and that showed a high expression on the second leg when compared to either legs.

Table S2. Semi-aquatic insects NCBI accession numbers for c67063_g1 and c68581_g1 homologs. These sequences derive from transcriptomes available in house, only the *Rhagovelia antilleana* and the *Stridulivelia tersa* sequences were cloned and sequenced using Sanger technology.

Species	Sequence	Accession Number
Mesoveloidea sp.	c68581_g1 like	MF448685
Mesovelia mulsanti	c68581_g1 like	MF448686
Microvelia longipes	c68581_g1 like	MF448687
Cylindrostethus palmaris	c68581_g1 like	MF448688
Aquarius paludum	c68581_g1 like	MF448689
Limnogonus franciscanus	c68581_g1 like	MF448690
Stridulivelia tersa	c68581_g1 like	MF448691
Oiovelia cunucunumana	c68581_g1 like	MF448692
Paravelia bullialata	c68581_g1 like	MF448693
Rhagovelia antilleana	c68581_g1	MF448694
Rhagovelia antilleana	c67063_g1	MF448695
Rhagovelia humboldti	c68581_g1 like	MF448696
Rhagovelia humboldti	c67063_g1 like	MF448697
Rhagovelia obesa	c68581_g1 like	MF448698
Rhagovelia obesa	c67063_g1 like	MF448699

Table S3 – BlastP of c67063_g1 and c68581_g1 homologs against the NCBI non-redundant protein database. To find all possible homologs, all protein sequences from the different clades were blasted against the non-redundant protein database. Only hits with an e-value lower than 10^{-4} were considered as true homologs. Homologs were only found within the Hemiptera order, except for one sequence that was found in *Zootermopsis nevadensis*, a species belonging to the Isoptera order. Conservation is low, but nonetheless present (Table S4, Fig. S5). Isoptera is the most derived branch of the Polyneoptera insects that includes 10 orders of insects (*54*). The fact that we could not find any other sequence within the Polyneoptera, despite the huge amount of sequence data (at least 28 transcriptomes) available for the group suggests that this gene was lost at least 6 independent times and it was kept only in the Isoptera order (*54*). Due to the unlikely events of 6 multiple independent losses in the Polyneoptera insects we deduce that the conservation between the proteins results from homoplastic events that led to the evolution of similar protein domains. The protein alignment of all homologs is depicted in Fig. S5.

BlastP using c68581_g1 as query

Organism name	Description	e-value	Accession nr.	Order
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X1 [Cimex lectularius]	2.00E-12	XP_014252242.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X2 [Cimex lectularius]	3.00E-12	XP_014252243.1	Hemiptera
Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	7.00E-07	XP_001946766.1	Hemiptera

BlastP using *Diaphorina citri* sequence as query

Organism name	Description	e-value	Accession nr.	Order
Diaphorina citri	PREDICTED: uncharacterized protein LOC103513139 [Diaphorina citri]	0	XP_008476174.1	Hemiptera
Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	3.00E-47	XP_001946766.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1026 [Homalodisca liturata]	5.00E-11	JAS91701.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1027 [Homalodisca liturata]	2.00E-10	JAS82778.1	Hemiptera

Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X1 [Cimex lectularius]	4.00E-05	XP_014252242.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X2 [Cimex lectularius]	5.00E-05	XP_014252243.1	Hemiptera

BlastP using *Acyrthosiphon pisum* sequence as query

Organism name	Description	e-value	Accession nr.	Order
Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	0	XP_001946766.1	Hemiptera
Diaphorina citri	PREDICTED: uncharacterized protein LOC103513139 [Diaphorina citri]	8.00E-51	XP_008476174.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1026 [Homalodisca liturata]	3.00E-11	JAS91701.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1027 [Homalodisca liturata]	4.00E-11	JAS82778.1	Hemiptera
Zootermopsis nevadensis	hypothetical protein L798_13671 [Zootermopsis nevadensis]	1.00E-09	KDR11725.1	Isoptera

BlastP using *Cimex lectularius* sequence as query

Organism name	Description	e-value	Accession nr.	Order
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X1 [Cimex lectularius]	1.00E-166	XP_014252242.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X2 [Cimex lectularius]	3.00E-164	XP_014252243.1	Hemiptera
Diaphorina citri	PREDICTED: uncharacterized protein LOC103513139 [Diaphorina citri]	6.00E-13	XP_008476174.1	Hemiptera
Halyomorpha halys	PREDICTED: uncharacterized protein LOC106684612 [Halyomorpha halys]	2.00E-08	XP_014282284.1	Hemiptera

Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	4.00E-08	XP_001946766.1	Hemiptera

BlastP using *Halyomorpha halys* sequence as query

Organism name	Description	e-value	Accession nr.	Order
Halyomorpha halys	PREDICTED: uncharacterized protein LOC106684612 [Halyomorpha halys]	9.00E-128	XP_014282284.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X1 [Cimex lectularius]	1.00E-08	XP_014252242.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X2 [Cimex lectularius]	2.00E-08	XP_014252243.1	Hemiptera

BlastP using *Homalodisca liturata* sequence as query

Organism name	Description	e-value	Accession nr.	Order
Homalodisca liturata	hypothetical protein g.1026 [Homalodisca liturata]	1.00E-75	JAS91701.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1027 [Homalodisca liturata]	9.00E-46	JAS82778.1	Hemiptera
Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	1.00E-13	XP_001946766.1	Hemiptera
Diaphorina citri	PREDICTED: uncharacterized protein LOC103513139 [Diaphorina citri]	2.00E-12	XP_008476174.1	Hemiptera
Zootermopsis nevadensis	hypothetical protein L798_13671 [Zootermopsis nevadensis]	3.00E-06	KDR11725.1	Isoptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X1 [Cimex lectularius]	5.00E-05	XP_014252242.1	Hemiptera
Cimex lectularius	PREDICTED: uncharacterized protein LOC106668226 isoform X2 [Cimex lectularius]	6.00E-	XP_014252243.1	Hemiptera

05

BlastP using *Zootermopsis nevadensis* sequence as query				
Organism name	**Description**	**e-value**	**Accession nr.**	**Order**
Zootermopsis nevadensis	hypothetical protein L798_13671 [Zootermopsis nevadensis]	0	KDR11725.1	Isoptera
Acyrthosiphon pisum	PREDICTED: uncharacterized protein LOC100165312 [Acyrthosiphon pisum]	2.00E-16	XP_001946766.1	Hemiptera
Diaphorina citri	PREDICTED: uncharacterized protein LOC103513139 [Diaphorina citri]	4.00E-10	XP_008476174.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1027 [Homalodisca liturata]	2.00E-06	JAS82778.1	Hemiptera
Homalodisca liturata	hypothetical protein g.1026 [Homalodisca liturata]	4.00E-06	JAS91701.1	Hemiptera

Table S4. Distance matrix showing the percentage of identity between the c67063_g1 and c68581_g1 homologs.

	H. liturata	*A. pisum*	*D. citri*	*R. antilleana* c68581_g1	*R. antilleana* c67063_g1	*A. paludum*	*Mesoveloidia sp.*	*C. lectularius*	*H. halys*	*Z. nevadensis*	Average distance
Homalodisca liturata	-	28.6	23.2	22.1	19.8	20.3	18.6	22.1	15.3	18.1	20.9
Acyrthosiphon pisum	-	-	28.7	15.6	15.3	14.5	13.3	14.9	14	14.8	17.7
Diaphorina citri	-	-	-	13.3	13.3	13.3	14.5	19.5	10.8	16	16.9
R. antilleana c68581_g1	-	-	-	-	88.2	85.3	61.6	26.5	20	15.9	38.7
R. antilleana c67063_g1	-	-	-	-	-	80	57.4	26.5	20.4	15.6	37.3
Aquarium paludum	-	-	-	-	-	-	64.7	26.1	20.8	16	37.8
Mesoveloidia sp.	-	-	-	-	-	-	-	26.6	18.2	16.6	32.3
Cimex lectularius	-	-	-	-	-	-	-	-	32.3	14.6	23.2
Halyomorpha halys	-	-	-	-	-	-	-	-	-	13.3	18.3
Zootermopsis nevadensis	-	-	-	-	-	-	-	-	-	-	15.6

Table S5. List of predicted domains, repeats, motifs and features determined with SMART sequence analysis (http://smart.embl-heidelberg.de/) for *geisha* and *mother-of-geisha*. Results in red depict the regions that show different signatures between the two duplicates according to the software prediction.

Predicted domains, repeats, motifs and features						
	mother-of-geisha			***geisha***		
Name	**Start**	**End**	**E-value**	**Start**	**End**	**E-value**
signal peptide	1	20	N/A	1	20	N/A
transmembrane region	84	103	N/A	84	103	N/A
transmembrane region	110	132	N/A	110	132	N/A
Pfam:DUF2931	*7*	*70*	*0.075*	*7*	*70*	*0.22*
low complexity	91	117	N/A	92	117	N/A
low complexity	122	148	N/A	122	148	N/A
DM4_12	*147*	*237*	*262000*	*152*	*237*	*262000*
RING	*170*	*230*	*1500*	N/A	N/A	N/A
G3P_acyltransf	N/A	N/A	N/A	*4*	*136*	*127000*
C4	N/A	N/A	N/A	*139*	*233*	*613*
STI1	N/A	N/A	N/A	*149*	*195*	*2550*

Notes: (a) results presented in bold show the domains that are confidently predicted, (b) results presented in italic have a significant score smaller than the required threshold

Table S6. Statistics of the dsRNA gene function assays for *yellow, gsha/mogsha* and *yfp* (yellow fluorescent protein) dsRNA was used as an injection control and synthetized according to (*44*, *52*).

Individuals	***yellow*** **(**5µg/µl**)**	***gsha/mogsha*** **(**5µg/µl**)**	***Yfp*** **(**5µg/µl**)**
Injected	35	72	67
Survived	12	27	27
Presence of phenotype	12	27	0

Table S7. Knockdown efficiency of *gsha/mogsha* RNAi treatment. We performed quantitative real time PCR for *gsha* and *mogsha* to determine if the expression of both duplicates would decrease with the injected dsRNA. One week after injection, 10 individuals (4th instar nymphs) from each treatment were isolated, their legs dissected and pooled for RNA extraction. qPCR measurements show that the expression (relative quantity using leg 1 as a reference) for both genes is drastically reduced when compared to the wild-type expression in Fig. 3B. This indicated that we are unable to target each duplicate separately using RNA interference technique, as expected from the sequence comparison analysis that shows a high extent of conservation between both duplicates.

Sample	**Treatment**	***geisha* expression**	***mother-of-geisha* expression**
Leg 1	*geisha/mother-of-geisha* RNAi	1	1
Leg 2		1.7	3.3
Leg 3		1.7	0.5

Table S8. Mean and standard deviation for the different locomotion parameters measured in the two water setups.

Still water setup	**Velocity (m/s)**	**Stroke frequency (stroke/s)**	**Angle°**	**Energy per stroke (mJ/mg)**
Stridulivelia tersa (n=10)	0.344 (0.085)	14.201 (3.518)	N/A	N/A
Rhagovelia antilleana (n=7)	0.146 (0.039)	2.672 (0.490)	76.60 (21.56)	N/A
R. antilleana ablated fan (n=4)	0.144 (0.062)	7.799 (3.699)	125.9 (23.79)	N/A
R. antilleana gsha/mogsha (n=9)	0.124 (0.039)	3.810 (0.916)	110.6 (24.06)	N/A
Stream setup (0.2 m/s)	**Velocity (m/s)**	**Stroke frequency (stroke/s)**	**Angle°**	**Energy per stroke (mJ/mg)**
Rhagovelia antilleana (n=10)	0.287 (0.0376)	16.27 (2.502)	N/A	0.0001138 (3.595E-005)
R. antilleana ablated fan (n=8)	0.1464 (0.05191)	19.80 (1.750)	N/A	0.00006228 (1.827E-005)
R. antilleana gsha/mogsha (n=8)	0.2095 (0.04906)	20.40 (2.519)	N/A	0.0000722 (1.744E-005)

Table S9. Time (seconds) spent in the stream setup. Red numbers depicts individuals that failed to row against the current, whereas black numbers depict individuals that manage to row against the current and reach the upstream bucket. *gsha/mogsha* individuals even when failing to reach the target, fight the current for longer than individuals where the fan was ablated. Data concerning the percentage of individuals that succeeded and the average time to reach the target is plotted in Fig. S9.

Species	Phenotype	Ind.	0.3m/s current	Average time to fail (0.3m/s)	0.2m/s current	Average time to fail (0.2m/s)
R. antilleana	dsRNA gsha/mogsha	Male 1	106.86	75.80	21.35	114.72
R. antilleana	dsRNA gsha/mogsha	Male2	7.39		18.59	
R. antilleana	dsRNA gsha/mogsha	Male 3	250.43		29.74	
R. antilleana	dsRNA gsha/mogsha	Male 4	53.78		19.96	
R. antilleana	dsRNA gsha/mogsha	Male 5	8.61		10.93	
R. antilleana	dsRNA gsha/mogsha	Male 6	11.83		15.16	
R. antilleana	dsRNA gsha/mogsha	Female 1	139		95.59	
R. antilleana	dsRNA gsha/mogsha	Female 3	6.54		114.72	
R. antilleana	Wild-type	Male 1	7.46	NA	16.5	NA
R. antilleana	Wild-type	Male 2	4.02		21.66	
R. antilleana	Wild-type	Female 3	7.64		6.47	
R. antilleana	Wild-type	Female 4	222.16		13.42	
R. antilleana	Wild-type	Male 5	16.1		11.69	
R. antilleana	Wild-type	Female 6	7.4		4.82	
R. antilleana	Wild-type	Female 7	7.92		10.46	
R. antilleana	Wild-type	Female 9	6.39		6.89	
R. antilleana	Wild-type	Female 10	5.52		8.18	
R. antilleana	Wild-type	Female 11	4.63		8.28	
R. antilleana	Fan ablated	Female 1	3.95	65.20	114.92	93.66
R. antilleana	Fan ablated	Female 2	149.07		54.34	
R. antilleana	Fan ablated	Female 3	7.93		85.68	
R. antilleana	Fan ablated	Female 4	3.44		4.08	
R. antilleana	Fan ablated	Male 5	124.66		143.63	
R. antilleana	Fan ablated	Female 6	3.23		5.09	
R. antilleana	Fan ablated	Female 7	39.91		84.49	
R. antilleana	Fan ablated	Male 8	6.57		76.57	
R. antilleana	Fan ablated	Male 10	300		49.54	
R. antilleana	Fan ablated	Male 11	12.37		53.4	
R. antilleana	Fan ablated	Male 12	188.29		168.11	

R. antilleana	Fan ablated	Male 13	5.36		100.66	
R. antilleana	Fan ablated	Male 14	2.85		300	
S. tersa	Wild-type	Male 1	Not tested	NA	3.94	9.97
S. tersa	Wild-type	Male 2	Not tested		3.74	
S. tersa	Wild-type	Male 3	Not tested		13.37	
S. tersa	Wild-type	Male 4	Not tested		6.62	
S. tersa	Wild-type	Male 5	Not tested		32	
S. tersa	Wild-type	Female 1	Not tested		3.15	
S. tersa	Wild-type	Female 2	Not tested		21.44	
S. tersa	Wild-type	Female 3	Not tested		3.8	
S. tersa	Wild-type	Female 4	Not tested		10.45	
S. tersa	Wild-type	Female 5	Not tested		7.05	
S. tersa	Wild-type	Female 6	Not tested		1.84	
S. tersa	Wild-type	Female 7	Not tested		2.64	
S. tersa	Wild-type	Female 8	Not tested		14.21	
S. tersa	Wild-type	Female 9	Not tested		23	
S. tersa	Wild-type	Female 10	Not tested		2.32	

Table S10 – Primer for candidate gene amplification in *Stridulivelia tersa* and *Oiovelia cunucunumana.*

For *O. cunucunumana* we have three primer pairs because the in situ was performed with a mix of three probes to increase the probability of signal detection.

Species	Primer	Sequence 5'3'	TºC
S. tersa	mogsha_ fw	GGCCGCTCTACTCCTAGTGA	52
S. tersa	mogsha_ rev	ACAAGCTTGGTAGGGCAGAC	52
S. tersa	coiledcoil_ fw	TGAATCTCACGTACGCCCTG	52
S. tersa	coiledcoil_ rev	TTGACTCCATCCTTGCCGAC	52
S. tersa	yellow_fw	ACACCGATATCAAGGCCGAC	52
S. tersa	yellow_rev	GGGTGTATCGCAAGGGGTAC	52
O. cunucunumana	mogsha_ fw1	AAYCAATATCCAGAYGGTGTTGCAGG	55
O. cunucunumana	mogsha_ rev1	TGA.TCCTTGCTGAAGAACAAGCTTT	55
O. cunucunumana	mogsha_ fw2	AAYCAATATCCAGAYGGTGTTGCAGG	55
O. cunucunumana	mogsha_ rev2	AAGCTTTGTTGGACAGACTTGACAT	55
O. cunucunumana	mogsha_ fw3	TYTRATCAAGGGMCTTCTGCTGA	55
O. cunucunumana	mogsha_ rev3	AGAATGCCTGATATCGTTTTTGACA	55

Table S11 – Primers used for the qPCR assay

Primer	Sequence 5'-3'	TºC
geisha_qpcr_fw	TGGGCTGTAGGAATCGTTCTC	58
geisha_qpcr_rev	CTGCGTCTGGAGGGATGGTA	58
mogsha_qpcr_fw	TCAATATCCAGATGGTGTTGCA	58
mogsha_qpcr_rev	CAGTCCCAAAGCAGATAATCCG	58
rps18_qpcr_fw	TGGGTACCAACATAGACGGTAA	58
rps18_qpcr_rev	CATCGGCCTTCTTTAGGACA	58
rpl13a_qpcr_fw	GGACACTTGTTGGGTCGTCT	58
rpl13a_qpcr_rev	CGGACGACGACAATCCTACT	58

Table S12 – Primers used for the dsRNA synthesis

Primer	Sequence 5'3'	TºC
yellow_T7_fw	TAATACGACTCACTATAGGGAG ACCACCCAGAGAACAACCTGCC TGT	52
yellow_T7_rev	TAATACGACTCACTATAGGGAG ACCACCCTTGCTCGTTAGTCTGC CA	52
cutprot19_T7_fw	TAATACGACTCACTATAGGGAG ACCACAAAGGTTGCGCTCTTTGC TG	52
cutprot19_T7_rev	TAATACGACTCACTATAGGGAG ACCACACCATCACGTTCCTCGCT TT	52
geisha_T7_fw	TAATACGACTCACTATAGGGAG ACCACTGCGATTCCCCTTGGTCT TC	52
geisha_T7_rev	TAATACGACTCACTATAGGGAG ACCACTCCAAACATCCTGGGGCT TC	52
coiledcoil_T7_fw	TAATACGACTCACTATAGGGAG ACCACAGCCAGCCAGAAAAGCA AGA	52
coiledcoil_T7_rev	TAATACGACTCACTATAGGGAG ACCACCCATGTGTCCTCGCCTCA TT	52

Supplementary Figures

Figure S1

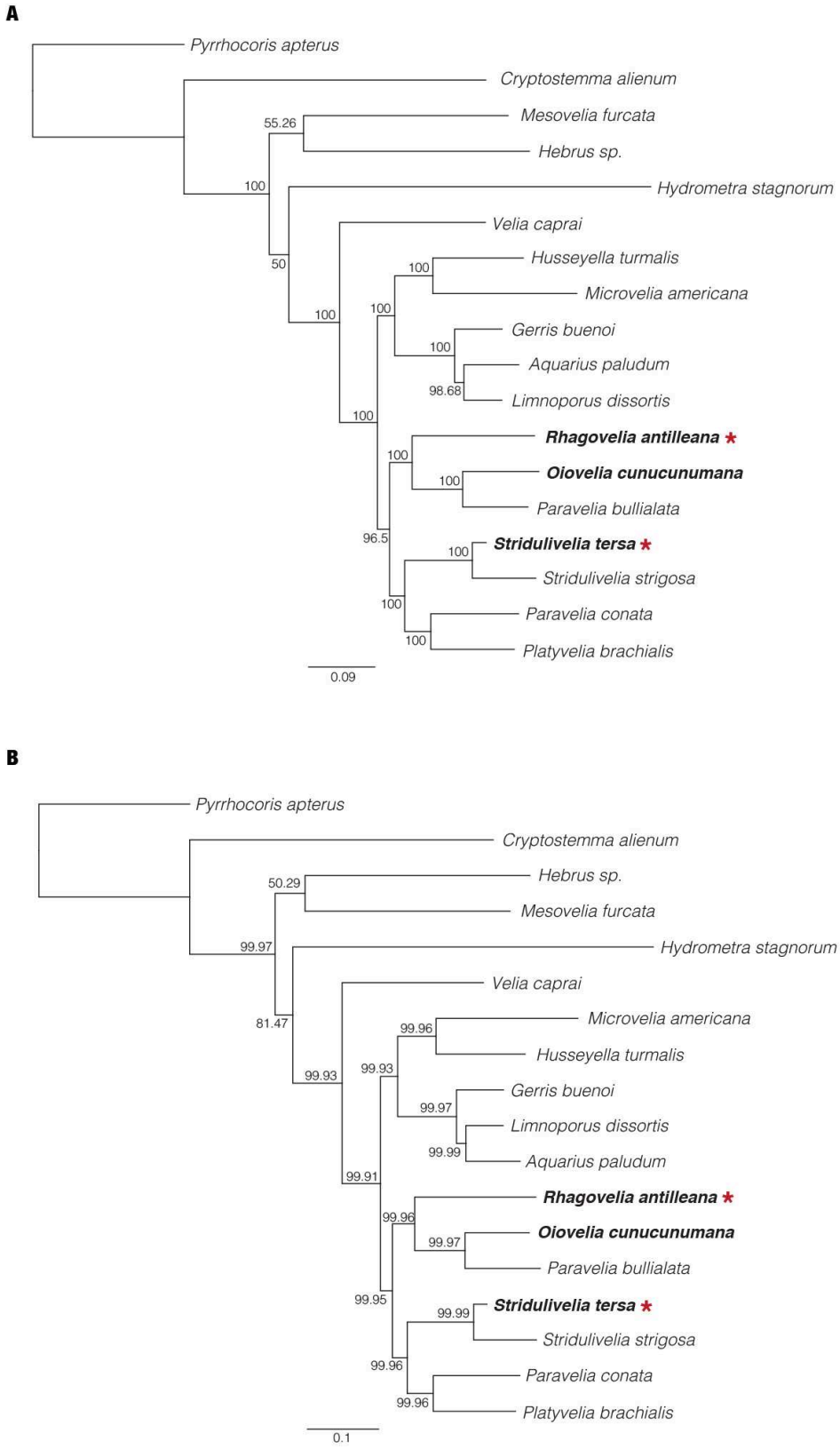


Fig. S1. Phylogeny of the semi-aquatic insects reconstructed using 14 molecular markers. (A) Maximum likelihood phylogeny**,** node values represent the bootstrap support. (B) Bayesian phylogeny, node values represent the posterior probabilities. The main clades of semi-aquatic insects are represented and trees are rooted with the terrestrial species *Pyrrhocoris apterus*. The three species in bold were experimentally tested for *gsha/mogsha* expression. Red asterisks highlight species using rowing as a mode of water surface locomotion. *S. tersa* was used as out-group for behavioral comparison to *R. antilleana* for four reasons: (1) it is a closest genus that uses the same mode of locomotion as *Rhagovelia,* contrary to *O. cunucunumana* and *P. bullialata* which live in stream margins and use the tripod gate; (2) it has the same leg morphology as *R. antilleana* (second leg longer than the third) ; (3) co-inhabits the stream environment with *Rhagovelia* and therefore face comparable environmental pressures*;* and finally (4) it is easily kept in the laboratory.

Figure S2

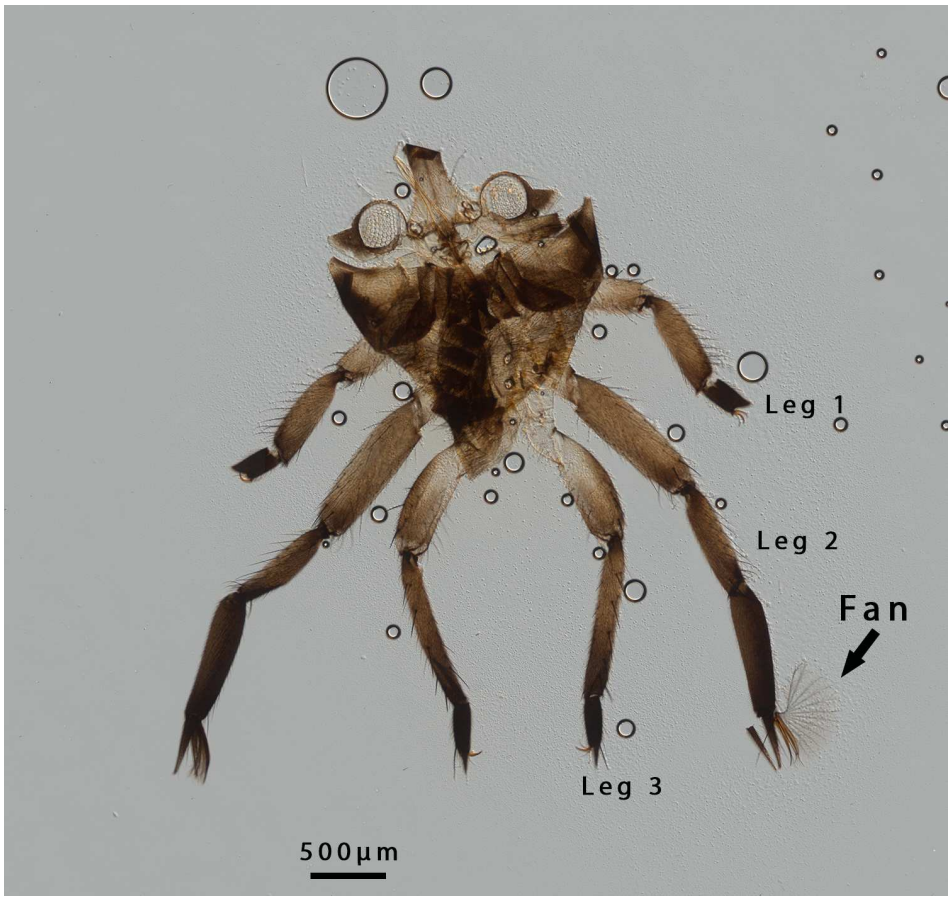


Fig. S2. Cuticle molt of a third instar *R. antilleana* nymph. Semi-aquatic insects are hemimetabolous insects that do not undergo metamorphosis; instead the embryo hatches as first instar nymph that already resembles the adult form. They undergo five molts, where at each stage the cuticle is renewed. Here we show that the fan (signaled with an arrow) is part of this cuticular exoskeleton and that it is entirely renewed during this process. At each molt its size and branch number increases until it reaches its final shape in the adults (Fig. 1C), this indicates that the program responsible for its development persists during post-embryonic stages.

Figure S3.

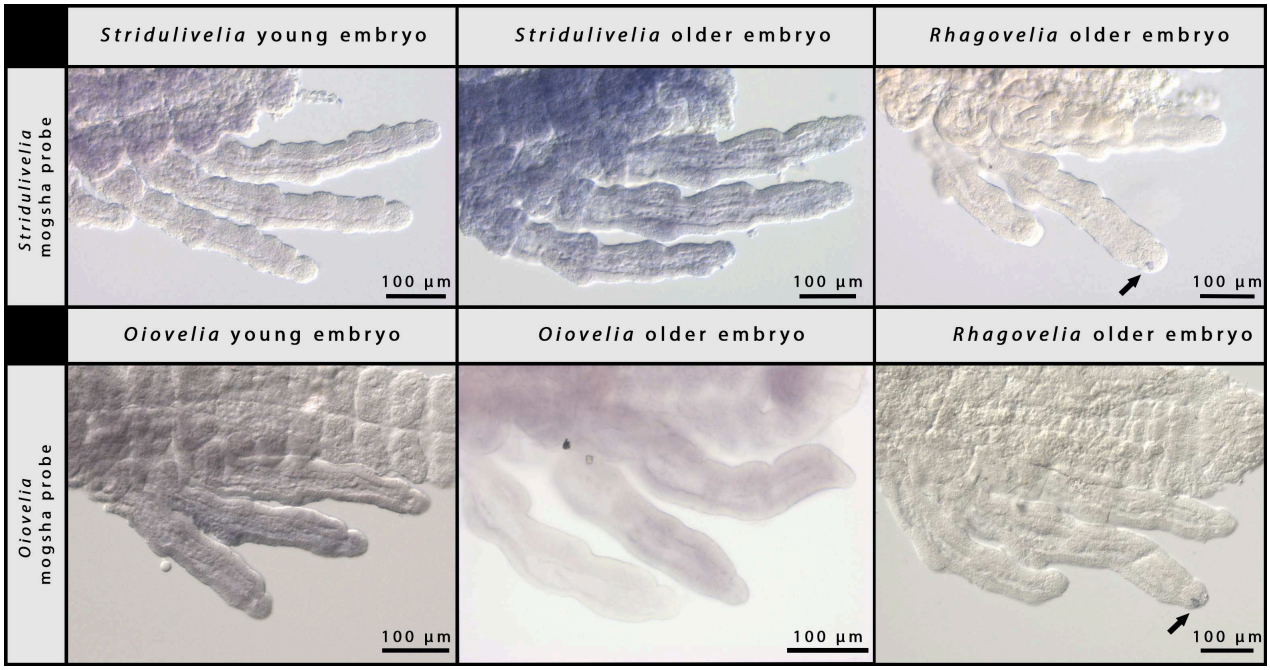


Fig. S3. *In situ* hybridization of c68581_g1 homolog (*mogsha*) in the embryos of *Stridulivelia tersa, Oiovelia cunucunumana and* and *Rhgovelia antilleana*. We failed to detect any expression in the embryonic legs of *S. tersa and O. cunucunumana,* neither in young embryos or older embryos using specific probes from these two species. We extended the revelation time to ensure that we are not missing expression due to low activity of the gene. However, the probes from both *S. tersa* and *O. cunucunumana* were able to reveal *mogsha* expression at the tip of *R. antilleana* embryos (arrows). This demonstrates that that the probe are functional and that the lack of staining in *S. tersa* and *O. cunucunumana* embryos is due to the lack of expression. We obtained the same results for *yellow* and *coiled-coil domain-containing protein 174*. Overall, our results suggest that there is a gain of expression of the studied genes in the *Rhagovelia* lineage, leading to the specification of a novel cell lineage.

Figure S4

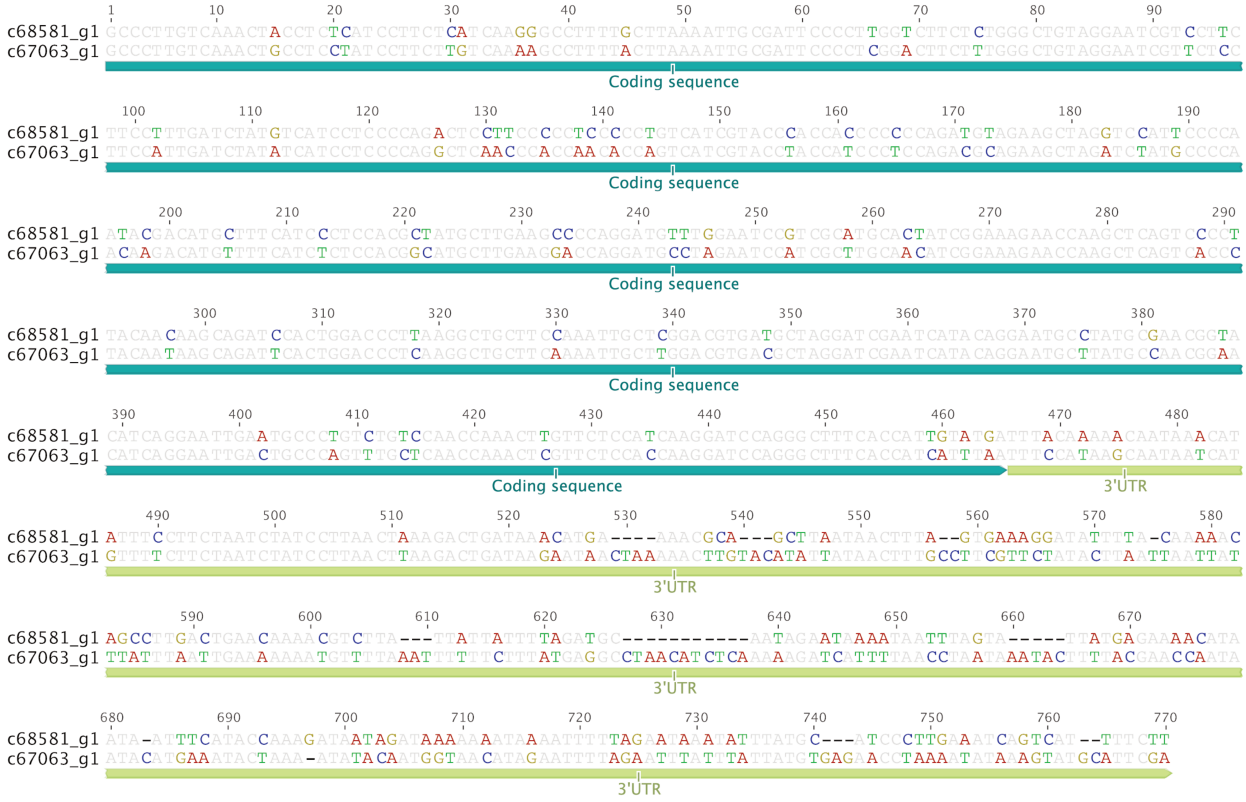


Fig. S4. Alignment of the region overlapping the c67063_g1 and c68581_g1 coding and 3'UTR sequences. The alignment shows that there are single nucleotide polymorphisms within the coding sequence (blue) that translate into differences in the amino acid sequence (Fig. 2B). Furthermore, the two sequences have different 3'UTRs (green) further confirming that they are products of two distinct duplicates.

Figure S5

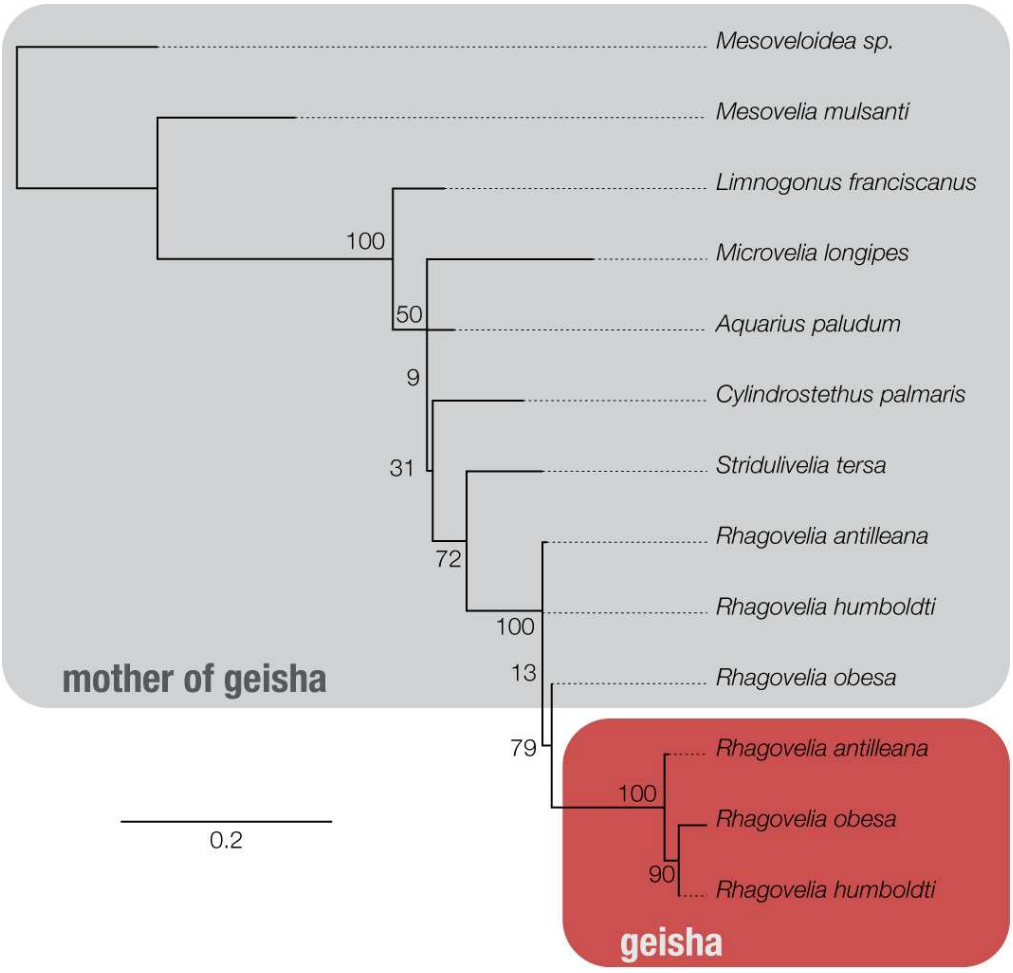


Fig. S5. Maximum likelihood phylogeny of *geisha* and *mother-of-geisha* within semi-aquatic insects. This phylogeny corroborates the Bayesian phylogeny shown in Fig. 3. It is a *Rhagovelia* specific duplication with *geisha* being the more derived gene in terms of sequence. Branch values represent the bootstrap support. All semi-aquatic species studied have only one duplicate (*mother of* geisha) including *Oiovelia cunucunumana* and *Paravelia bullialata* (Fig. S1 and Table S2). These sequences were not used in the gene phylogeny due to their short length. These results indicate that *gsha/mogsha* duplication event is specific to the *Rhagovelia* lineage and *gsha* is a *Rhagovelia* restricted gene.

Figure S6

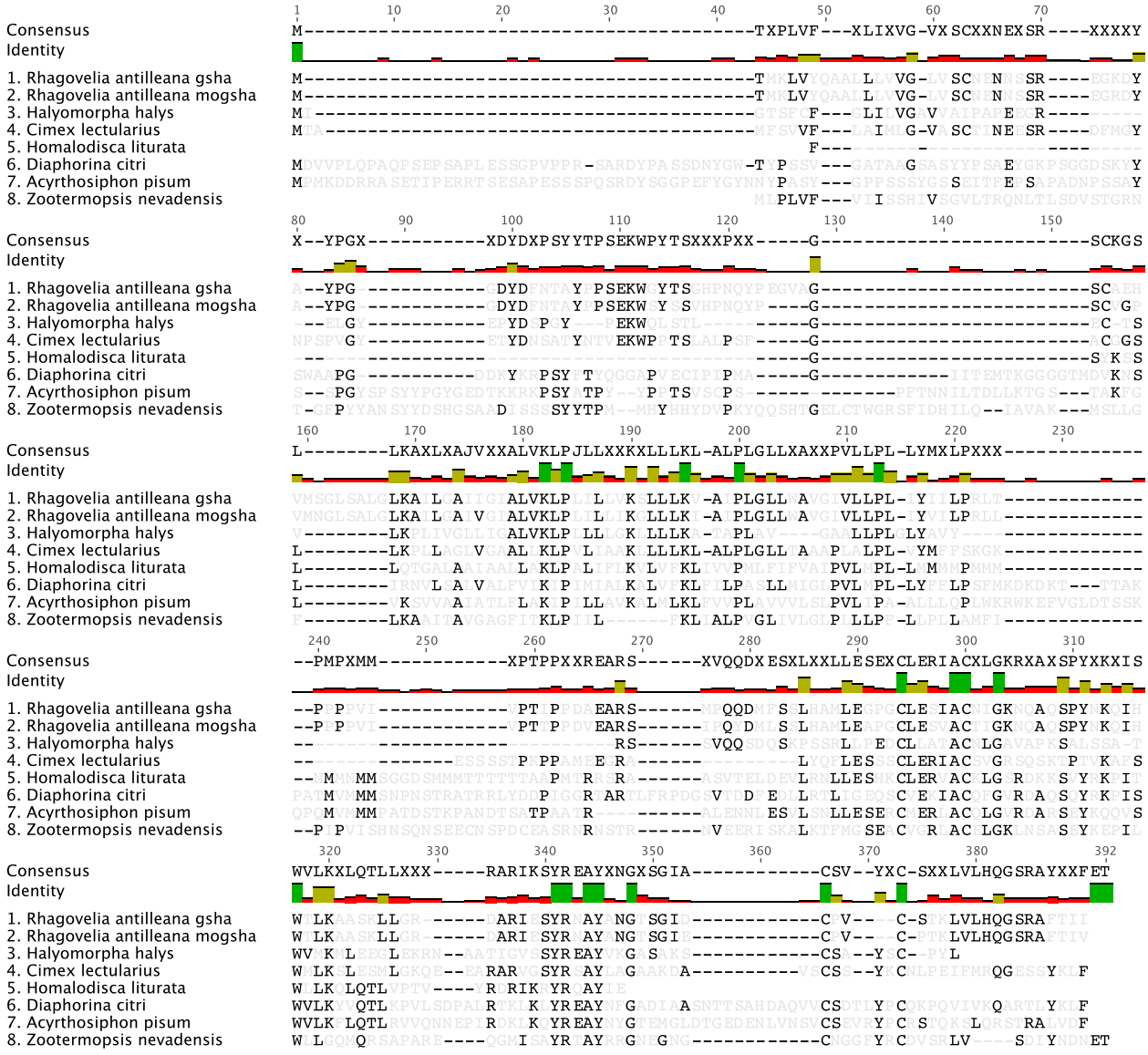


Fig. S6. Protein alignment of *geisha/mother-of-geisha* homologs. The Hemiptera and Isoptera protein sequences retrieved from the NCBI non-redundant protein database using BlastP (Table S2) were aligned with the *Rhagovelia antilleana geisha* and *mother-of-geisha* using MAFFT. Conserved amino acid residues are depicted in black showing that the extent of conservation between sequences is very low (Table S3 and S4).

Figure S7



Fig. S7. RNAi knockdown fan phenotypes. (A) Fan phenotype of *yellow* RNAi treated individuals. The plumose structure of the fan is not affected in *yellow* knockdown individuals. The *yellow* individuals have yellow fans instead of the usual darkbrown. The same phenotype is observed in the adult body showing that *yellow* acts to darken the body and fan cuticle. (B) Fan phenotype of *gsha/mogsha* RNAi treated individuals. Although all RNAi individuals had an abnormal fan, we obtained a range of phenotypes with different penetrance. Here are the phenotypes of each individual used to phenotype the trait and for the locomotion assays. The secondary branches of each filament disappear and the number and the length of branches are reduced. High-speed videography revealed that *gsha/mogsha* individuals still deploy and retract the claw and the remnants of the fan just like control individuals do (Movie S2).

Figure S8

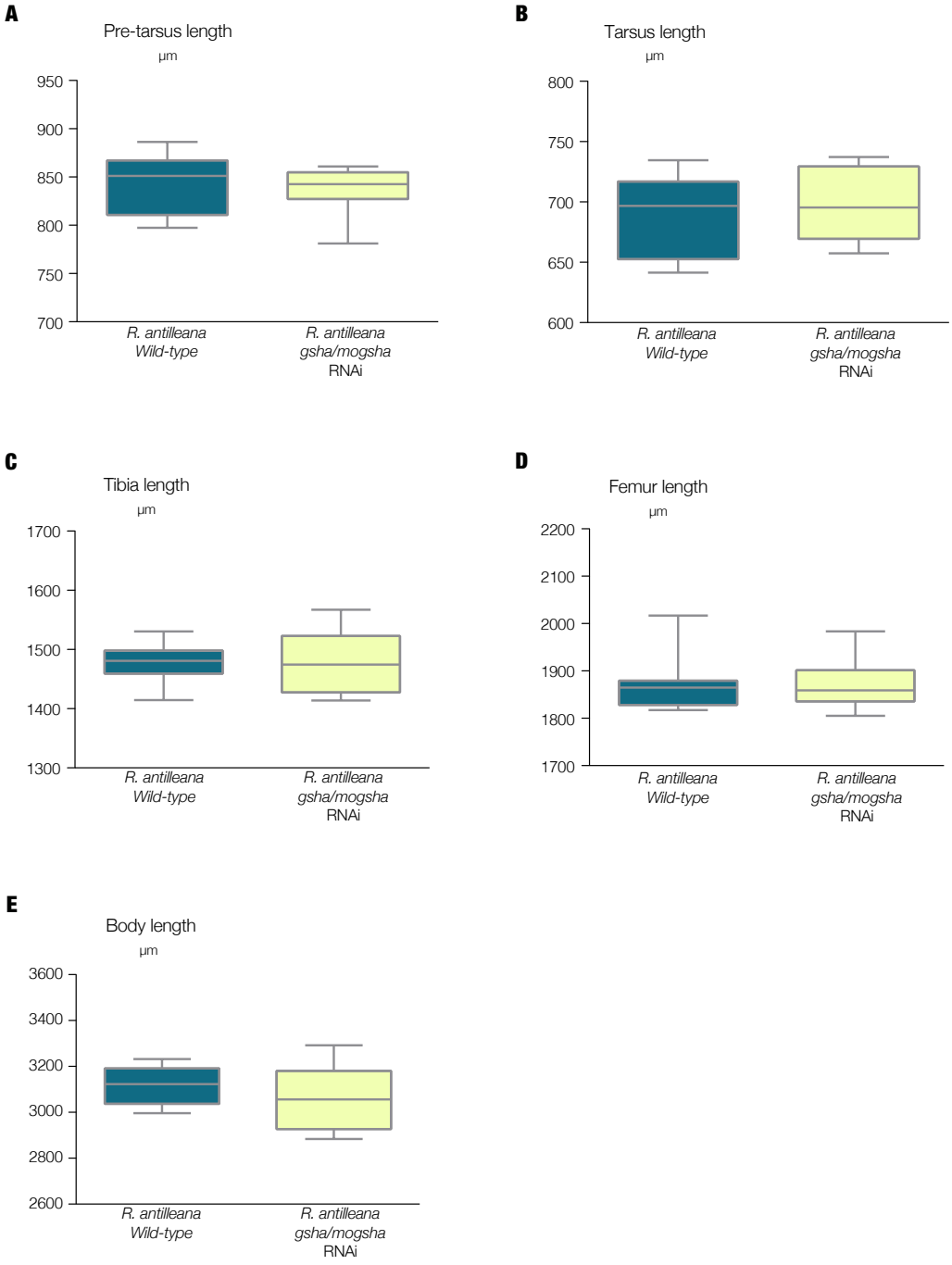


Fig. S8. Body length and second leg length comparison between wild-type individuals and *gsha*/*mogsha* RNAi individuals. The differences in locomotion performance described in Fig. 4 could be due to a phenotypic effect of the *gsha*/*mogsha* knockdown in body length and on the morphology of the second leg. To control for these effects we compared (A) Pre-tarsus length, (B) tarsus length, (C) tibia length, (D) femur length and (E) body length between the wild-type group (blue box, 10 individuals) and the *gsha*/*mogsha* knockdown group (yellow box, 8 individuals). The bars represent the minimum and maximum value. No significant difference (Student's t-test) was found for any of the comparisons, body and leg morphology is identical and therefore cannot explain the locomotion differences between the two groups.

Figure S9

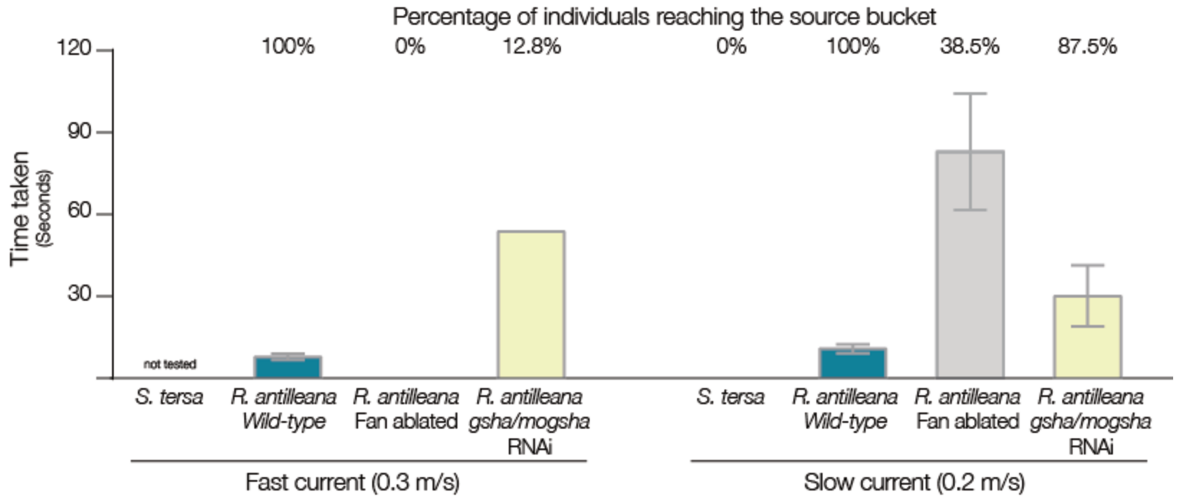


Fig. S9. Percentage of individuals that succeed in rowing against the current and the time taken to arrive to the upstream buckets (Movie S3). The toboggan setup consisted in two water buckets (upstream and downstream compartments) connected by a water canal where we can control the speed of water current. A scheme of the setup can be found in Fig. 4 in the main manuscript. The individuals jump from a fixed platform and invariably try to row against the current; this is a very well characterized behavior in species that inhabit streams, to avoid the current drift (*55*). The objective was to test if there was a performance difference between *Stridulivelia tersa* (a fan-less closely related species that inhabits this riparian habitat), *Rhagovelia antilleana* (normal fan), and *Rhagovelia antilleana* without the fan (surgically ablated or RNAi treated individuals) in an environment where the organisms are forced to face the current. For this experiment we used two water current settings, fast current at 0.3 m/s and slow current at 0.2 m/s. For each current setting we measured the number of individuals that reached the upstream bucket. The percentage on top of each bar represents the number of individuals that succeeded, whereas the bars represent the time that took them to reach the upstream bucket (bars).

Surprisingly, *Stridulivelia tersa*, which is the fastest in the standard setup, failed to ascend any current and all individuals (n = 15) drifted downstream. On the opposite extreme is *R. antilleana* (blue bar, n = 10) with a fully functional fan, which always makes it to the source bucket in an average of 7.8 seconds in the fast and 10.9 seconds and slow current. When we ablated the fan, all individuals (n = 13) failed ascend the fast current, whereas in the slow current 38,5% (5 out of 13) succeed although in a very slow average time of 83 seconds (grey bar). Finally, for the *gsha/mosha* RNAi treated individuals we obtained intermediate results. In the fast current only one individual (1 out of 8) managed to ascend with a timing of 53.8 seconds, whilst in the slow current most individuals made it upstream (87.5%, 7 out of 8) with an average time of 30.2 seconds. This is an intermediate timing between the normal (10.8 seconds) and the fan-ablated group (83 seconds). These results clearly mimic the severity of the phenotype, with the fan-less individuals performing worse

than the RNAi individuals that have rudimentary fans. Taken together, these results show that the fan helps sustaining permanent movement against the current allowing the insects to avoid the stream drift.

The fan is tightly associated with the lifestyle of *Rhagovelia* in running waters, and there are no known instances where it was lost, but rather an extra fan was gained in the hind-legs in *Rhagovelia (Tetraripis) zitteli* (*56*, *57*). Our locomotion assays results suggest that the association between the evolution and maintenance of the fan and the lifestyle of *Rhagovelia* is probably the result of strong selection due to locomotion requirements in running waters. Furthermore, we know of at least another genus called *Husseyella* (see phylogenies in Fig. S1), which is found in the margins of streams, where the claws of mid-tarsi were widened to act as blades during leg strokes and therefore fill a similar function as the *Rhagovelia* fan (*20*).

Figure S10

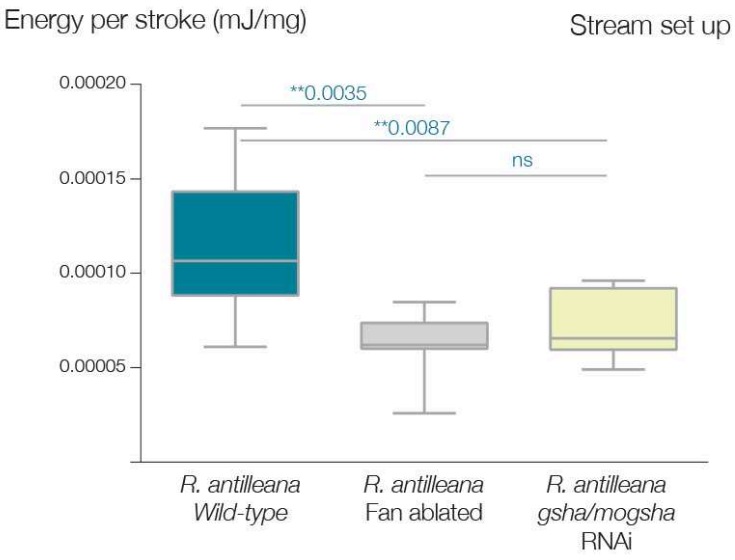


Fig. S10. Estimates of kinetic energy per stroke (millijoule/milligram). Our results show that, as expected wild-type *Rhagovelia antilleana* with a fully functional fan (blue box, n = 10) converts more energy into more movement when compared to the fan-less individuals (grey box, n = 7) and the RNAi treated individuals (yellow box, n = 8). The error bars represent the minimum and maximum value. Student's t-test or Mann-Whitney test was performed for each pair-wise comparison to test for significant differences. Overall, our data suggests that the fan allows the insect to transform more energy into movement on the water surface.

Figure S11

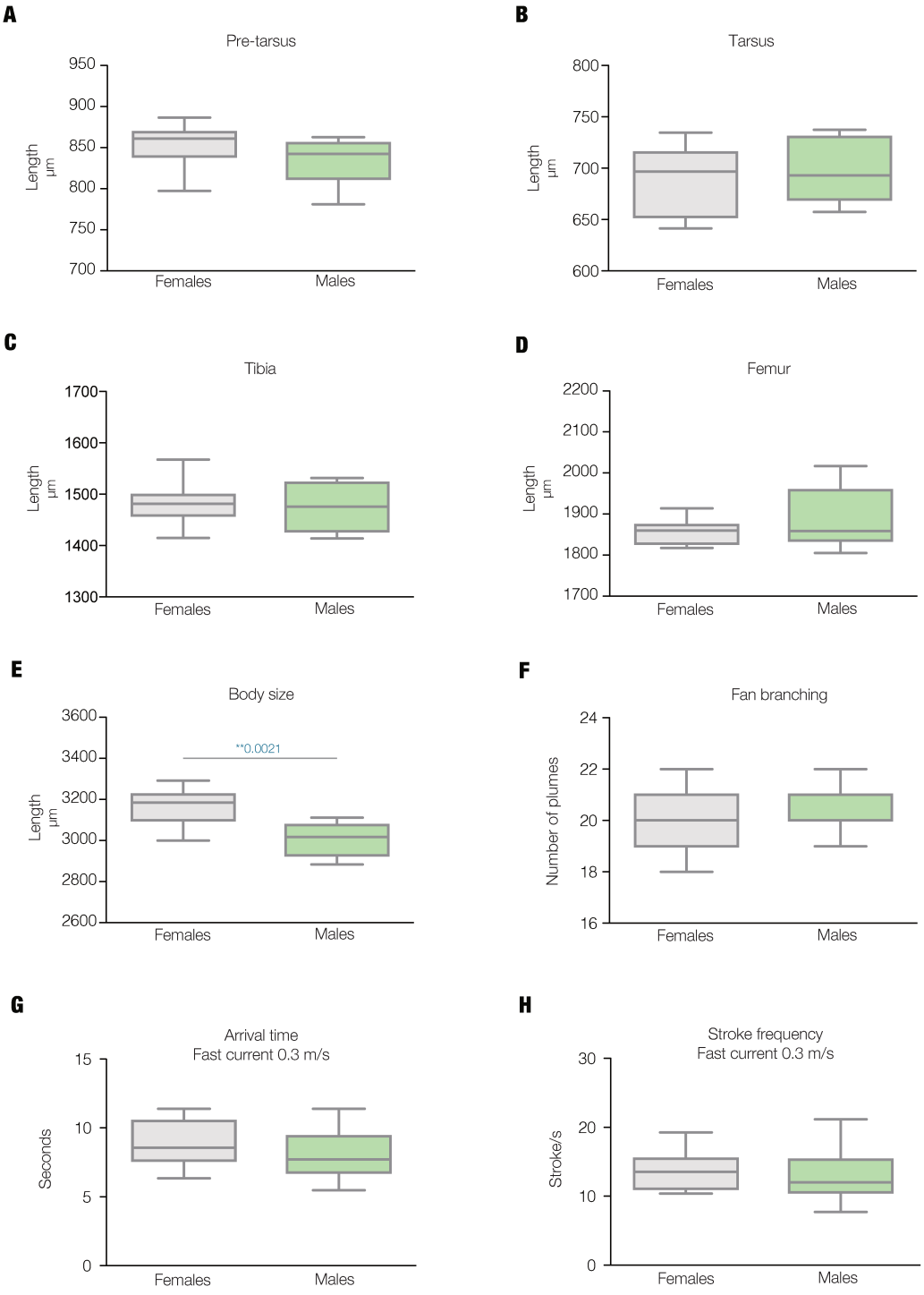


Fig. S11. Comparison between male and female *R. antilleana*. Differences in locomotion observed in our behavioral experiments could be influence by our female biased sampling (Table S10). To control for this effect we compared female and male (A, B, C, D) length of the different segments of the second leg, (E) body size and (F) fan branching (proxy for fan size). Our results show that there are no differences between males and females in leg length neither in the branching of the fan. *R. antilleana* females are significantly longer than males, but this does not translate to differences in locomotion in the stream set up. Males and

females show similar (G) arrival times and (H) stroke frequency. Over all our data suggests that there are no locomotion differences between males and females that could affect our conclusions concerning the *gsha/mogsha* dsRNA effect on *Rhagovelia* locomotion. The error bars represent the minimum and maximum value. Student's t-test or Mann-Whitney test was performed for each pair-wise comparison to test for significant differences.

Movie captions

Movie S1. Locomotion habits and niche preference of *Rhagovelia* and *Stridulivelia.*

Movie S2. Comparison of fan deployment between untreated and *gsha/mogsha* RNAi-treated *Rhagovelia*. Note that despite the reduction of the fan in RNAi-treated individuals, the remnants of the fan are still deployed and retracted just like in normal individuals.

Movie S3. Stream set up showing the condition for the ascension test against current and the performance of untreated compared to *gsha/mogsha* RNAi-treated individuals.

References and Notes

1. G. P. Wagner, V. J. Lynch, Evolutionary novelties. *Curr. Biol.* **20**, R48–R52 (2010). doi:10.1016/j.cub.2009.11.010 Medline

2. M. Pigliucci, What, if anything, is an evolutionary novelty? *Philos. Sci.* **75**, 887–898 (2008). doi:10.1086/594532

3. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, 2000).

4. J. B. Losos, *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (Univ. of California Press, 2009).

5. M. Averof, S. M. Cohen, Evolutionary origin of insect wings from ancestral gills. *Nature* **385**, 627–630 (1997). doi:10.1038/385627a0 Medline

6. V. A. Albert, D. G. Oppenheimer, C. Lindqvist, Pleiotropy, redundancy and the evolution of flowers. *Trends Plant Sci.* **7**, 297–301 (2002). doi:10.1016/S1360-1385(02)02300-2 Medline

7. R. O. Prum, A. H. Brush, The evolutionary origin and diversification of feathers. *Q. Rev. Biol.* **77**, 261–295 (2002). doi:10.1086/341993 Medline

8. S. B. Carroll, J. K. Grenier, S. D. Weatherbee, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Science, ed. 1, 2001).

9. R. W. R. Wallbank, S. W. Baxter, C. Pardo-Diaz, J. J. Hanly, S. H. Martin, J. Mallet, K. K. Dasmahapatra, C. Salazar, M. Joron, N. Nadeau, W. O. McMillan, C. D. Jiggins, Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. *PLOS Biol.* **14**, e1002353 (2016). doi:10.1371/journal.pbio.1002353 Medline

10. J. E. Moustakas-Verho, J. Cebra-Thomas, S. F. Gilbert, Patterning of the turtle shell. *Curr. Opin. Genet. Dev.* **45**, 124–131 (2017). doi:10.1016/j.gde.2017.03.016 Medline

11. S. B. Carroll, Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008). doi:10.1016/j.cell.2008.06.030 Medline

12. V. J. Lynch, G. P. Wagner, Resurrecting the role of transcription factor change in developmental evolution. *Evolution* **62**, 2131–2154 (2008). doi:10.1111/j.1558-5646.2008.00440.x Medline

13. A. Monteiro, M. D. Gupta, in *Genes and Evolution*, V. Orgogozo, Ed. (Elsevier, ed. 1, 2016), vol. 119, pp. 205–226.

14. W. J. Glassford, W. C. Johnson, N. R. Dall, S. J. Smith, Y. Liu, W. Boll, M. Noll, M. Rebeiz, Co-option of an ancestral hox-regulated network underlies a recently evolved morphological novelty. *Dev. Cell* **34**, 520–531 (2015). doi:10.1016/j.devcel.2015.08.005 Medline

15. A. Khila, E. Abouheif, L. Rowe, Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. *Evolution* **68**, 2159–2170 (2014). Medline

16. K. Khalturin, G. Hemmrich, S. Fraune, R. Augustin, T. C. G. Bosch, More than just orphans: Are taxonomically-restricted genes important in evolution? *Trends Genet.* **25**, 404–413 (2009). doi:10.1016/j.tig.2009.07.006 Medline

17. H. Kaessmann, Origins, evolution, and phenotypic impact of new genes. *Genome Res.* **20**, 1313–1326 (2010). doi:10.1101/gr.101386.109 Medline

18. M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P.

Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, J. M. Rothberg, Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380 (2005). Medline

19. W. C. Jasper, T. A. Linksvayer, J. Atallah, D. Friedman, J. C. Chiu, B. R. Johnson, Large-scale coding sequence change underlies the evolution of postdevelopmental novelty in honey bees. *Mol. Biol. Evol.* **32**, 334–346 (2015). doi:10.1093/molbev/msu292 Medline

20. N. M. Andersen, *The Semiaquatic Bugs (Hemiptera: Gerromorpha)* (Scandinavian Science Press, 1982).

21. D. N. Padilla-Gil, F. F. F. Moreira, Checklist, taxonomy and distribution of the *Rhagovelia* Mayr, 1865 (Hemiptera: Heteroptera: Veliidae) of the Americas. *Zootaxa* **3640**, 409–424 (2013). doi:10.11646/zootaxa.3640.3.5 Medline

22. A. J. J. Crumière, M. E. Santos, M. Sémon, D. Armisén, F. F. F. Moreira, A. Khila, Diversity in morphology and locomotory behavior is associated with niche expansion in the semi-aquatic bugs. *Curr. Biol.* **26**, 3336–3342 (2016). doi:10.1016/j.cub.2016.09.061 Medline

23. M. E. Santos, C. S. Berger, P. N. Refki, A. Khila, Integrating evo-devo with ecology for a better understanding of phenotypic evolution. *Brief. Funct. Genomics* **14**, 384–395 (2015). doi:10.1093/bfgp/elv003 Medline

24. P. J. Wittkopp, P. Beldade, Development and evolution of insect pigmentation: Genetic mechanisms and the potential consequences of pleiotropy. *Semin. Cell Dev. Biol.* **20**, 65–71 (2009). doi:10.1016/j.semcdb.2008.10.002 Medline

25. R. Kafri, M. Springer, Y. Pilpel, Genetic redundancy: New tricks for old genes. *Cell* **136**, 389–392 (2009). doi:10.1016/j.cell.2009.01.027 Medline

26. A. M. Dean, J. W. Thornton, Mechanistic approaches to the study of evolution: The functional synthesis. *Nat. Rev. Genet.* **8**, 675–688 (2007). doi:10.1038/nrg2160 Medline

27. A. Monteiro, Gene regulatory networks reused to build novel traits. *BioEssays* **34**, 181–186 (2012). doi:10.1002/bies.201100160 Medline

28. D.-E. Nilsson, S. Pelger, A pessimistic estimate of the time required for an eye to evolve. *Proc. Biol. Sci.* **256**, 53–58 (1994). doi:10.1098/rspb.1994.0048 Medline

29. W. Gehring, M. Seimiya, Eye evolution and the origin of Darwin's eye prototype. *Ital. J. Zool.* **77**, 124–136 (2010). doi:10.1080/11250001003795350

30. C. Darwin, *On the Origin of Species by Means of Natural Selection, Or, The Preservation of Favoured Races in the Struggle for Life* (J. Murray, London, 1859).

31. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi:10.1093/molbev/mst010 Medline

32. F. Ronquist, M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, J. P. Huelsenbeck, MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012). doi:10.1093/sysbio/sys029 Medline

33. S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010). doi:10.1093/sysbio/syq010 Medline

34. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014). doi:10.1093/bioinformatics/btu170 Medline

35. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011). doi:10.1038/nbt.1883 Medline

36. A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005). doi:10.1093/bioinformatics/bti610 Medline

37. K. D. Pruitt, T. Tatusova, D. R. Maglott, NCBI Reference Sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005). doi:10.1093/nar/gki025 Medline

38. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi:10.1038/nmeth.1923 Medline

39. B. Li, C. N. Dewey, RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011). doi:10.1186/1471-2105-12-323 Medline

40. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi:10.1093/bioinformatics/btp616 Medline

41. B. J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. LeDuc, N. Friedman, A. Regev, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013). doi:10.1038/nprot.2013.084 Medline

42. S. Rozen, H. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**, 365–386 (2000). Medline

43. M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, A. Drummond, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012). doi:10.1093/bioinformatics/bts199 Medline

44. P. N. Refki, D. Armisén, A. J. J. Crumière, S. Viala, A. Khila, Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. *Dev. Biol.* **392**, 441–453 (2014). doi:10.1016/j.ydbio.2014.05.021 Medline

45. J. Schultz, F. Milpetz, P. Bork, C. P. Ponting, SMART, a simple modular architecture research tool: Identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5857–5864 (1998). doi:10.1073/pnas.95.11.5857 Medline

46. D. Tautz, T. Domazet-Lošo, The evolutionary origin of orphan genes. *Nat. Rev. Genet.* **12**, 692–702 (2011). doi:10.1038/nrg3053 Medline

47. M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, D. G. Higgins, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007). doi:10.1093/bioinformatics/btm404 Medline

48. M. Suyama, D. Torrents, P. Bork, PAL2NAL: Robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–W612 (2006). doi:10.1093/nar/gkl315 Medline

49. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**, 1641–1650 (2009). doi:10.1093/molbev/msp077 Medline

50. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi:10.1093/molbev/msm088 Medline

51. M. W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45 (2001). doi:10.1093/nar/29.9.e45 Medline

52. A. Khila, E. Abouheif, L. Rowe, Evolution of a novel appendage ground plan in water striders is driven by changes in the *Hox* gene *Ultrabithorax*. *PLOS Genet.* **5**, e1000583 (2009). doi:10.1371/journal.pgen.1000583 Medline

53. G. M. Rubin, A. C. Spradling, Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982). doi:10.1126/science.6289436 Medline

54. B. Misof, S. Liu, K. Meusemann, R. S. Peters, A. Donath, C. Mayer, P. B. Frandsen, J. Ware, T. Flouri, R. G. Beutel, O. Niehuis, M. Petersen, F. Izquierdo-Carrasco, T. Wappler, J. Rust, A. J. Aberer, U. Aspöck, H. Aspöck, D. Bartel, A. Blanke, S. Berger, A. Böhm, T. R. Buckley, B. Calcott, J. Chen, F. Friedrich, M. Fukui, M. Fujita, C. Greve, P. Grobe, S. Gu, Y. Huang, L. S. Jermiin, A. Y. Kawahara, L. Krogmann, M. Kubiak, R. Lanfear, H. Letsch, Y. Li, Z. Li, J. Li, H. Lu, R. Machida, Y. Mashimo, P. Kapli, D. D. McKenna, G. Meng, Y. Nakagaki, J. L. Navarrete-Heredia, M. Ott, Y. Ou, G. Pass, L. Podsiadlowski, H. Pohl, B. M. von Reumont, K. Schütte, K. Sekiya, S. Shimizu, A. Slipinski, A. Stamatakis, W. Song, X. Su, N. U. Szucsich, M. Tan, X. Tan, M. Tang, J. Tang, G. Timelthaler, S. Tomizuka, M. Trautwein, X. Tong, T. Uchifune, M. G. Walzl, B. M. Wiegmann, J. Wilbrandt, B. Wipfler, T. K. F. Wong, Q. Wu, G. Wu, Y. Xie, S. Yang, Q. Yang, D. K. Yeates, K. Yoshizawa, Q. Zhang, R. Zhang, W. Zhang, Y. Zhang, J. Zhao, C. Zhou, L. Zhou, T. Ziesmann, S. Zou, Y. Li, X. Xu, Y. Zhang, H. Yang, J. Wang, J. Wang, K. M. Kjer, X. Zhou, Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014). doi:10.1126/science.1257570 Medline

55. T. F. Waters, The drift of stream insects. *Annu. Rev. Entomol.* **17**, 253–272 (1972). doi:10.1146/annurev.en.17.010172.001345

56. N. M. Andersen, A new species of *Tetraripis* from Thailand, with a critical assessment of the generic classification of the subfamily Rhagoveliinae (Hemiptera, Gerromorpha, Veliidae). *Tijdschr. Entomol.* **142**, 185–194 (1999). doi:10.1163/22119434-99900028

57. J. T. Polhemus, D. A. Polhemus, Zoogeography, ecology, and systematics of the genus Rhagovelia Mayr (Heteroptera: Veliidae) in Borneo, Celebes, and the Moluccas. *Insecta Mundi* **2**, 161–230 (1988).

Chapter 5

Dimorphism under sexual conflict controlled by the combination of sex determination and Hox genes.

Manuscript in preparation

First author

Synopsis of the paper:

Sexual conflict is an important driver of diversity in nature by affecting both sexual dimorphism and speciation. Sexual conflict can lead to an arms race, with the coevolution of sexually antagonistic structures such as male grasping structures facilitating the access to females or female anti-grasping structures to avoid superfluous matings. In the genus *Rhagovelia*, males and females interact during pre-mating struggles and possess dimorphic secondary sexual traits. However the evolution and the role of some of these structures has never been assessed, as well as the nature of the developmental genetic mechanisms involved in the coevolution of such traits.

Summary of results:

The reconstruction of phenotypic complexity shows the presence of an escalation in sexual conflict between males and females in the genus *Rhagovelia*. Males and females of two species, *Rhagovelia antilleana* and *Rhagovelia obesa*, possess dimorphic traits such as sex combs in the T1-legs and several modifications of the T3-legs in males; while females have modification of the abdomen shape and a pronotum projection. Males use their T3-legs, which have big femur containing muscles, to block female legs during pre-mating struggles in order to mate. Females with a pronotum projection are more efficient at rejecting males. Candidate gene approach shows that the combination of the sex determination gene *doublesex* and the Hox genes *Sex combs reduced* and *Ultrabithorax* participate differentially, in males and females, in the development of antagonistic traits.

Conclusion:

Our results indicate that different dimorphic structures in males and females of the genus *Rhagovelia* are actually secondary sexual traits resulting from an escalation in sexual conflict between the sexes. The development of these structures is orchestrated by the same combination of genes; *doublesex*, *Sex combs reduced* and *Ultrabithorax*, which act differentially in males and females.

Author contributions:

Here are the contributions of the different authors:

Antonin Crumière: insect sampling, description of behaviors and structures, imaging, quantification of traits functions, statistical analysis, muscle staining, phylogenetic reconstruction and analysis, gene cloning, RNAi knockdown, design of the study and redaction of the paper.

David Armisén: generation and treatment of transcriptomes.

Martha Koubarakos: description of behaviors, video acquisition.

Felipe Moreira: insect sampling and species identification.

Abderrahman Khila: insect sampling, design of the study and redaction of the paper.

Dimorphism under sexual conflict controlled by the combination of sex determination and Hox genes

Article in preparation

[1]A. J. J. Crumière, [1]D. Armisén, [1]M. Koubarakos, [2]F. F. F. Moreira and [1]A. Khila

Author affiliations:

[1]Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5242, Ecole Normale
Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France

[2]Laboratório de Biodiversidade Entomológica, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21040-360, Brazil

Abstract:

Sexual conflict drives the antagonistic coevolution of secondary sexual traits between males and females that can result in an arms race. However, only few studies have assessed the evolution of antagonistic secondary sexual traits of organisms, the function and the importance of these traits during sexual interactions and the genetic mechanisms underlying their development and evolution. Here we used the genus of water striders *Rhagovelia* (Heteroptera, Gerromorpha), in which males and females exhibit different degrees of sexual dimorphism associated with pre-mating struggles, characteristic of sexual conflict. We used the variability in traits between sexes and across species to assess the intensity of sexual conflict. Then we focused on *Rhagovelia antilleana*, a species that exhibits extreme sexual dimorphism with several grasping structures in males and anti-grasping structures in females to determine the role and the impact of these traits on mating success. We showed that males used their rear-legs as clamps to grasp females during pre-mating struggles while female use pronotum projection to resist harassing males. Finally we identified the developmental genetic mechanisms underlying the coevolution of antagonistic structures in *Rhagovelia antilleana* using RNA interference. We showed that the sex determination gene *doublesex* and the Hox genes *Sex combs reduced* and *Ultrabithorax* play opposing roles between males and females by their action in the development of antagonistic structures. Our integrative work highlights how the combination of sex determination genes and Hox genes can control dimorphism between the sexes under sexual conflict.

Introduction

The evolutionary interests of males and females often diverge and traits favored in one sex can be harmful to the other, thus fueling antagonistic co-evolution of the sexes (Arnqvist and Rowe, 2005). Such conflict can drag the sexes away from their fitness optima (Arnqvist and Rowe, 2002a; Pennell and Morrow, 2013) and drive conflict resolution through the evolution of sexual dimorphism. In cases of sexual conflict, the evolution of antagonistic traits in one sex, such as grasping traits in males, can be matched by the evolution of counter-adaptations in the other sex, such as anti-grasping traits in females (Arnqvist and Rowe, 2002b; Weigensberg and Fairbairn, 1996). This sex-specific armament can lead to episodes of escalation and ultimately arms race (Arnqvist and Rowe, 1995; Pennell and Morrow, 2013). Because males and females shared a common genetic basis, the majority of studies have focused on evolution of genomes, especially sex chromosomes, or sex biased gene expressions to understand the dynamic and the evolution of genetic mechanisms involved in intra- and interlocus sexual conflict (Dean et al., 2012; Mank, 2017; Mank et al., 2013; Snook et al., 2010). Intralocus sexual conflict can constrain sex-specific trait expression and, therefore, conflict resolution by preventing the independent evolution of the sexes and achievement of sex-specific fitness optima (Arnqvist and Rowe, 2002a). Although the study of antagonistic co-evolution of the sexes is attracting attention, the evolution and developmental genetic mechanisms involved in sexually antagonistic traits are poorly understood (Arnqvist and Rowe, 2002a; Bergsten and Miller, 2007; Kuntner et al., 2009). Here we studied the diversity and function of secondary sexual traits, in a lineage of water striders called *Rhagovelia* (Heteroptera, Gerromorpha), to understand how adaptations and counter adaptations evolved in each of the sexes. We uncovered and functionally tested loci underlying adaptations and counter-adaptations in each of the sexes thus informing the nature of genes undergoing sexual conflict. We provide evidence that highly pleiotropic genes can be involved in both intra and interlocus sexual conflict and armament escalation between the sexes.

Results and discussion

Sexual dimorphism driven by sexual conflict in the genus *Rhagovelia*

Sexual dimorphism is widespread in water striders (Arnqvist, 1992; Hayashi, 1985; Rowe et al., 1994; Weigensberg and Fairbairn, 1996), and is particularly manifest in the tropical genus *Rhagovelia sp* (Andersen, 1982; Moreira and Ribeiro, 2009; Polhemus, 1997). In this genus, sexual dimorphism ranges between subtle to spectacular differences between males and females (Figure 1A-D'). Generally various processes of sexual selection can drive dimorphism (Andersson, 1994; Arnqvist and Rowe, 2005), however the nature of the process acting on the sexes in *Rhagovelia* is unknown. We therefore observed mating system and sexual interactions in *Rhagovelia antilleana,* a representative species of the genus (Supplementary video 1). We detected vigorous pre-mating interactions, consisting of stereotypical phases of struggles between males and females, a behavior that characterizes sexual conflict (Arnqvist and Rowe, 1995; Arnqvist and Rowe, 2005). During the first phase the male grasps the female, with his fore- and mid-legs, and starts to climb on her back. Then the male blocks the female's mid- and rear-legs using his own rear-legs as clamps. During this step the female resists by shaking her body in an attempt to eject the male. In the second phase the male clasps the pronotum of the female using the forelegs while the rear-legs maintain tight grip on female legs. In the third phase the male tries to insert his genitalia while holding female legs. When successful, the male releases female legs and remains on her back by maintaining grip on her pronotum with his forelegs. These observations indicate that in *Rhagovelia*, females try to resist while males try to persist during mating interactions. This suggests that dimorphism in *Rhagovelia* may be driven by sexual conflict over mating frequency (Arnqvist, 1992; Arnqvist and Rowe, 2002a; Arnqvist and Rowe, 2005; Khila et al., 2012; Rowe et al., 1994).

Sex-specific armaments in *Rhagovelia antilleana*

Sexual conflict over mating rate is expected to drive the evolution of antagonistic traits that further the evolutionary interest of each of the sexes (Arnqvist and Rowe, 2002a; Arnqvist and Rowe, 2005; Bergsten and Miller, 2007; Kuntner et al., 2009). Specifically, males are expected to evolve clasping traits that increase persistence whereas females are expected to evolve anti-clasping traits that increase their ability to resist mating attempts (Arnqvist and

Rowe, 1995; Arnqvist and Rowe, 2002a). In *R. antilleana*, the ability of the male to clasp the pronotum of the female is made possible due to the sex combs located at the distal part of foreleg tibia (Figure 1E, E'). The femurs of the rear-legs are much larger in males and are equipped with a set of rows of large and small spines (Figure 1F, F'). In addition, the tibiae and trochanters of the rear-legs are also equipped with rows of teeth-like structures and that are not found in the females (Figure 1F, F', G, G'). These male structures confer the rear-legs with a strong clamp function that allows the male to clasp and neutralize female legs (Supplementary video 1). Females present both morphological and behavioral antagonistic traits that seemingly increase their ability to reject male mating attempts. When attacked by males, females struggle by vigorously shaking their body and by performing repeated somersaults (Supplementary video 1). Furthermore, in *R. antilleana* we found that the winged and wingless morphs of *R. antilleana* females present morph-specific morphological modifications that seem to counter mating attempts by males. The wingless females have a narrow abdomen that seems to make it difficult for the male to hold with his rear-legs (Figure 1C, C'). Winged females possess a prominent projection on the pronotum, which forms a barrier between the attacking male and the body of the female thus facilitating female resistance (Supplementary video 1, Figure 1H, H'). Wing polymorphism is widespread and variable across Gerromorpha species (Andersen, 1982), and could be due to different causes such as temperature (Fairbairn, 1988), density of population (Ditrich and Papacek, 2010), photoperiod (Spence, 1989; Zera et al., 1983), or genetically controlled (Braendle et al., 2005a; Braendle et al., 2005b; Spence, 1989; Zera et al., 1983). While we do not know the origin of wing polymorphism in *R. antilleana*, the presence of sexually antagonistic traits in both males and females reinforces our conclusion that dimorphism in this species is driven by sexual conflict.

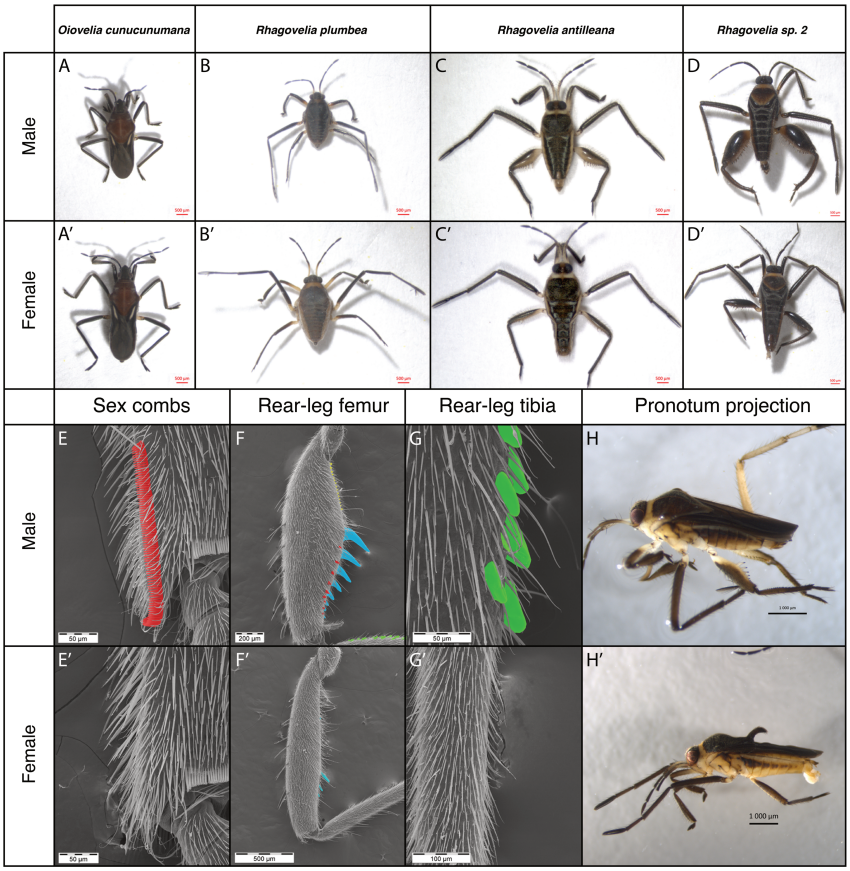


Figure 1: Sexual dimorphism and secondary sexual traits in the genus *Rhagovelia* genus. The upper panel (A-D') represents the variability in sexual dimorphism observed in our sample between males (A-D) and female (A'-D') of *Oiovelia cunucunumana* (A and A'), *Rhagovelia plumbea* (B and B'), *Rhagovelia antilleana* (C and C') and *Rhagovelia sp. 2* (D and D'). Male and female of species such as *O. cunucunumana* and *R. plumbea* are be relatively similar while species such as *R. antilleana* and *R. sp 2* are higly dimorphic. The lower panel represents the different sex-specific traits in *R. antilleana* with the male sex combs (E and E'), the modification of femur shape and the different row of spines in male rear-legs (F and F'), the spines on rear-leg tibia in males (G and G') and the pronotum projection in winged female (H and H').

Importance of male and female antagonistic traits during pre-mating struggles

Natural *Rhagovelia antilleana* populations generally exhibit variation in some of the antagonistic traits both in males and females (Figure 1; Figure S1). In males, we observe variation primarily in the size of rear-leg femurs whereas the sex combs seem to be stable (Figure S1). In females, only winged individuals develop pronotum projection and only wingless females have narrow abdomen (Figure 1C', H'). The variation in these traits could be indicative of unresolved genetic conflicts in the population. Furthermore, how these traits affect individual performance during pre-mating struggles remains unknown. We therefore designed a tournament set-up (Ivy and Sakaluk, 2007) to separate individuals based on the outcome (copulation for males and rejection for females) of premating struggles into successful, intermediate, and unsuccessful individuals (Figure 2A). We found that successful males have significantly more tibia spikes and thicker rear-leg femurs than unsuccessful males (Figure 2B, C), whereas body size, number of spikes on rear-leg femur, tibia and femur length do not seem to differ between the two groups (Figure 2D, E, F, G). Staining using a molecular marker showed that most of the rear-leg femur is occupied by muscles (Figure S2), suggesting that larger femurs can generate a higher amount of force. To test this hypothesis, we inferred male grasping ability using the femur width as proxy for adductor force (Gilliver et al., 2009) and the length of the tibia as outlever length in the groups of successful and unsuccessful males. These two leg segments are important for males blocking female legs. We found that successful males, with thicker femurs, have a significantly better grasping ability than unsuccessful males with thinner femurs (Figure 2H) (Wilcoxon test, p-val: 0,0056 **). Furthermore, we observed that winged males, which typically have the thinnest rear-leg femurs (Figure S1), were overrepresented in the unsuccessful group (Table 1, Chi2: 9.317; degree of freedom: 2; p-val: 0.0094806 **), suggesting a tradeoff between developing wings and thick rear-leg femurs. Our test revealed that females with pronotum projection are significantly over represented in the successful group and under represented in the intermediate and unsuccessful groups (Table 1, Chi2: 26.313; degree of freedom: 2; p-val: 1.93e-06 ***). Strikingly, 14 among a total of 22 females with a pronotum projection were successful in rejecting males whereas only 4 out of a total of 50 females without pronotum projection were successful (Table 1). This result indicates that the pronotum projection allows females to more efficiently resist male harassment.

Conditions	Successful (observed/expected)	Intermediate (observed/expected)	Unsuccessful (observed/expected)	Total
Winged males	1/3.75	6/7.5	8/3.75	15
Winged females	14/5.5	7/11	1/5.5	22
Wingless males	15/12.25	26/24.5	8/12.25	49
Wingless females	4/12.5	29/25	17/12.5	50

Table 1: Distribution of morphs after tournament.

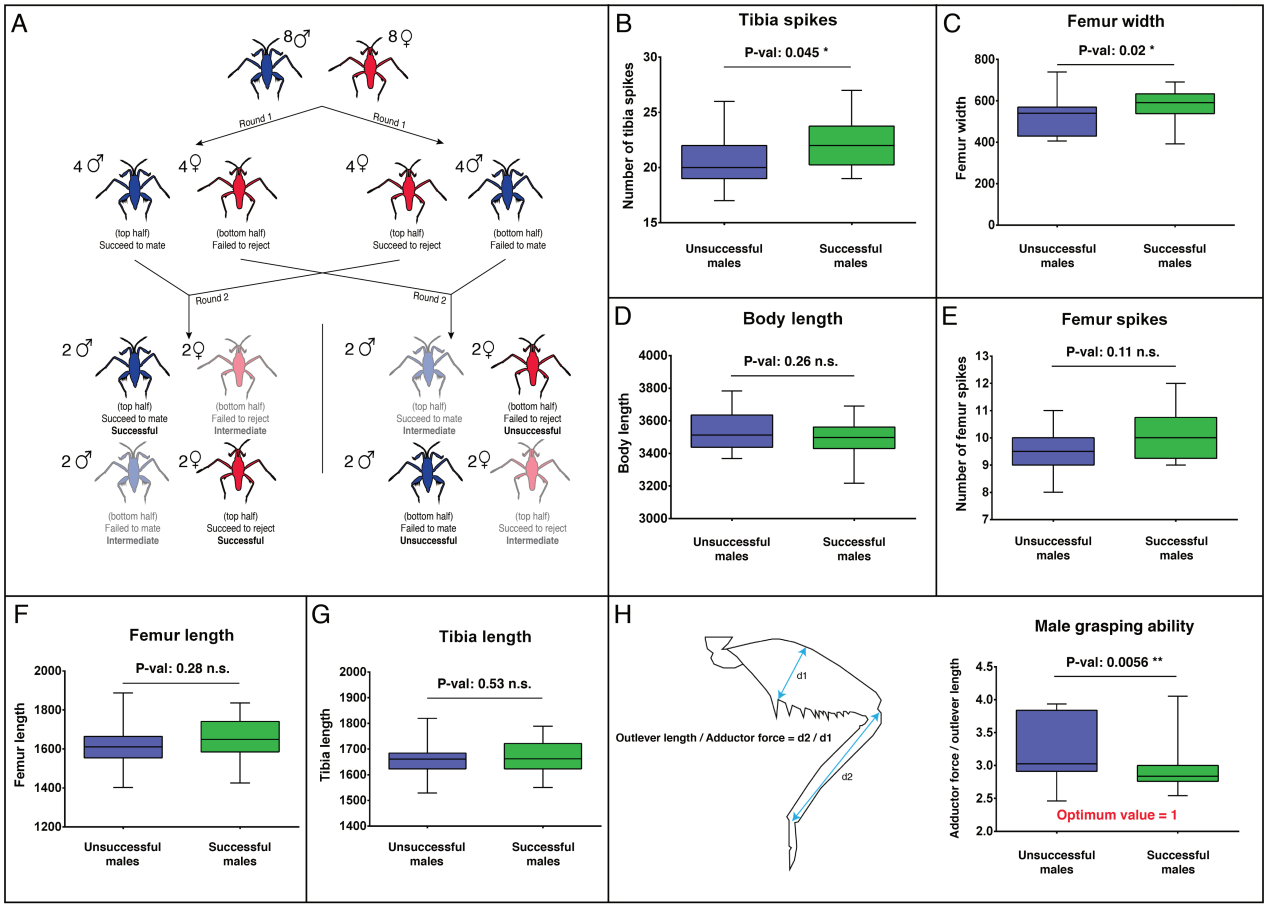


Figure 2: Tournament set-up and comparison of phenotypes between successful and unsuccessful individuals. (A) Representation of the set-up used to split individuals into successful and unsuccessful individuals based on their mating success. This set-up allows us to obtain 20 males for both groups from 10 different replicates. Different morphological characters have been quantified between successful and unsuccessful males. Successful males have significantly higher number of tibia spines (B) and larger femur width (C) than unsuccessful males. There were no significant differences for body length (D), number of femur spines (E), femur length (F) and tibia length (G). We also observed a significantly better grasping ability for successful males (H). T-tests were performed for all comparison except for the number of tibia spikes for which we performed a Wilcoxon test.

Escalation in sexual conflict has driven sexual dimorphism in the genus *Rhagovelia*

Coevolution of the sexes due to antagonistic interactions could lead to sexual conflict escalation and arms race that deeply shape the evolutionary trajectory of lineages in nature (Arnqvist and Rowe, 1995; Bergsten and Miller, 2007; Kuntner et al., 2009). We therefore decided to test the intensity of sexual conflict in the *Rhagovelia* genus. Intensity of sexual conflict can be detected by analyzing phylogenetic patterns of correlation of male and female phenotypic complexity in terms of secondary sexually antagonistic traits (Arnqvist and Rowe, 2002a; Bergsten and Miller, 2007; Kuntner et al., 2009). To test this hypothesis, we generated a matrix of male- and female-specific sexually antagonistic traits in a total of thirteen species including ten *Rhagovelia* and three closely related outgroups (Table 2, Table S1). Based on our behavioral observation, we found eight traits that are associated with clasping in males and two traits associated with anti-clasping in females (Table 2, Figure 1). We built the phylogeny of this sample and mapped the sexually antagonistic traits to determine the pattern of phenotypic complexity between males and females in this sample (Figure 3, Figure S3, Figure S4). Our reconstruction showed that there is an increase of phenotypic complexity along the male phylogeny in the *Rhagovelia* genus with some species having higher scores than others (Figure 3). This result indicates that males from *Rhagovelia* species have evolved more secondary sexually antagonistic traits than the out-groups. We also found that females of two species,

Rhagovelia obesa and *Rhagovelia antilleana*, matched the complexity of males through the evolution of anti-grasping traits (Figure 3). This result indicates an escalation of sexual conflict in the genus *Rhagovelia*, particularly in *R. antilleana* and *R. obesa*, which has shaped sexual dimorphism between the sexes across species.

	Sex combs	Rear-leg femur big spikes	Rear-leg femur interme-diate spikes	Rear-leg femur small spikes	Rear-leg femur shape	Rear-leg tibia spikes	Rear-leg curvature	Rear-leg trochanter spikes	**Number of male traits**	Pronotum projection	Abdomen shape	**Number of female traits**
Microvelia americana	1	1	0	0	0	0	0	0	2	0	0	0
Stridulivelia tersa	1	1	0	0	0	1	0	0	3	0	0	0
Rhagovelia mangle	1	1	0	0	0	1	1	0	4	0	0	0
Rhagovelia obesa	1	1	1	1	1	1	1	0	7	1	1	2
Rhagovelia antilleana	1	1	1	1	1	1	1	1	8	1	1	2
Rhagovelia humboldti	1	1	1	1	1	1	1	1	8	0	0	0
Rhagovelia tenuipes	1	1	1	0	0	1	0	0	4	0	0	0
Rhagovelia amazonensis	1	1	1	1	1	1	0	1	7	0	0	0
Rhagovelia sp. 1	1	1	1	1	1	1	1	1	8	?	0	0
Rhagovelia sp. 2	1	1	1	1	1	1	1	1	8	0	0	0
Rhagovelia plumbea	1	1	0	0	0	0	1	0	3	0	0	0
Oiovelia cunucunumana	1	0	0	0	0	0	0	0	1	0	0	0
Paravelia bullialata	1	0	0	0	0	0	0	0	1	0	0	0

Table 2: Matrix of secondary sexual traits in our sample. (1) indicates presence and (0) indicates absence.

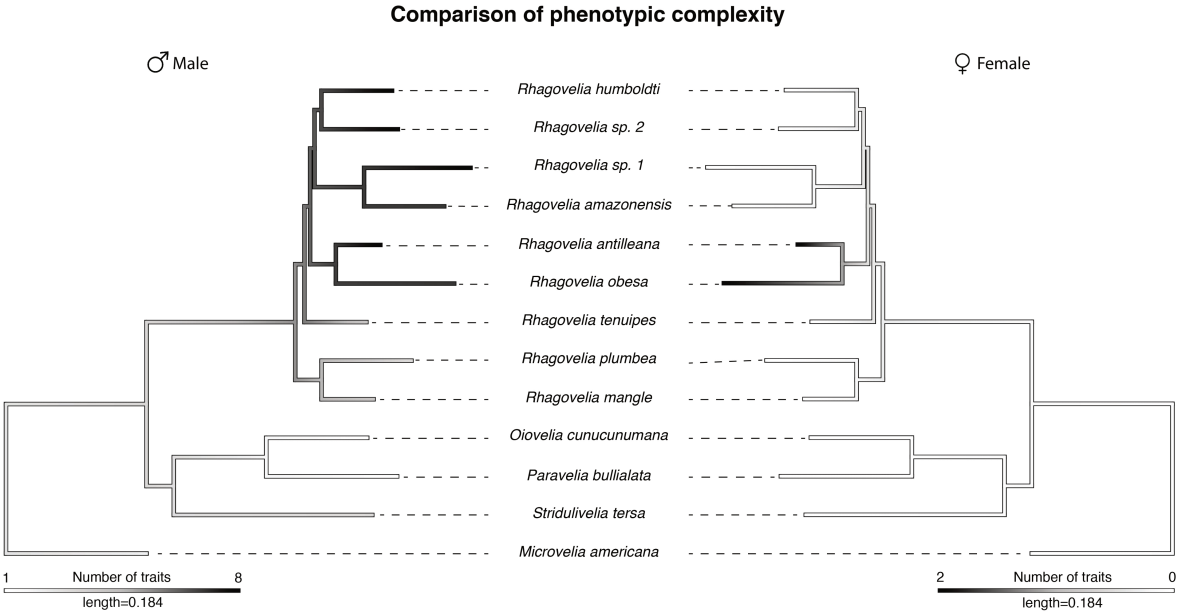


Figure 3: Phylogenetic reconstruction of phenotypic complexity. Males in the *Rhagovelia* genus have higher phenotypic complexity than out-groups based on our choice of secondary sexual traits. Females from two species, *R. antilleana* and *R. Obesa* have higher phenotypic complexity than the others species and they match the complexity of male phenotype. This indicates the presence of escalation in sexual conflict in the genus.

Sexual dimorphism in *Rhagovelia* is controlled by sex determination and Hox genes

Our phylogenetic reconstruction detected a clear pattern of sexual conflict escalation and conflict resolution through the evolution of extreme dimorphism between the sexes (Bonduriansky and Chenoweth, 2009; Pennell and Morrow, 2013; Perry and Rowe, 2015). These patterns may involve intra- or inter-locus sexual conflict, or both, but the loci under conflict in *Rhagovelia* remain unknown. The dimorphism in *Rhagovelia* is established late during nymphal development and affects the morphology of different segments along the body axis. Because segment identity is established by the action of Hox genes (Averof and Akam, 1993, 1995; Hughes and Kaufman, 2002), we hypothesized that sexual dimorphism in *Rhagovelia* is under the influence of sex-specific Hox function. The sex combs in males and the pronotum projection in females are both located in the first thoracic segment whose morphology is known to be under the control of *Sex combs reduced* (Chesebro et al., 2009; Tanaka et al., 2011). Nymphal RNAi knockdown against *Scr* resulted in a decrease of the size of pronotum projection in females (Figure 4 F, G, M, N and I-L) and the reduction of the height of sex comb teeth in males (Figure 4 A, B, D, E). This demonstrates that the same locus, *Scr,* is responsible for the development of distinct sexually antagonistic traits in males and females of *Rhagovelia*. This result suggests resolved intralocus conflicts for these structures (Bonduriansky and Chenoweth, 2009; Pennell and Morrow, 2013). The modified rear-legs, used by males as clamps to hold the females, are located on the third thoracic segment known to be the domain of expression of the Hox gene *Ultrabithorax* (Davis et al., 2007; Khila et al., 2014; Refki et al., 2014; Rozowski and Akam, 2002; Stern, 2003). RNAi knockdown against *Ubx* strongly reduces the width of male rear-leg femur that now resembles the femurs of the females (Figure 4 S, T, W). In addition, *Ubx* knockdown results in the complete loss of the spines in all rear-leg segments (Figure 4 S, T). We also obtained some individuals with mild RNAi effect showing a milder reduction in rear-leg femur width and persistence of some spines. In the females, *Ubx* knockdown resulted in subtle reduction of the width and the loss of the small spines found on the rear-leg femur (Figure 4 V, W). These results demonstrate that *Ubx* is responsible for the development of all sexually antagonistic modifications found on the rear-leg femur in the males. Furthermore, these data show that the partial expression of some of these traits in the rear-leg femur in females is due to genetic correlation of *Ubx* expression suggesting an unresolved intralocus conflict (Bonduriansky and Chenoweth, 2009; Pennell and Morrow, 2013). Sex determination genes, such as *doublesex (dsx)*, are known to regulate the sex-specific action of Hox genes in some insect species through their sex-specific isoforms (Kopp et al., 2000; Tanaka et al., 2011). Nymphal RNAi knockdown against *dsx* strongly affected the organization and reduced the size of the sex combs on male fore-legs (Figure 4C). In male rear-legs, *dsx* RNAi decreased the width of the femur and the number of spikes in a manner strikingly similar to *Ubx* knockdown (Figure 4T, U). *dsx* knockdown also resulted in wingless males with narrow abdomen (Figure 4P, Q, R) and winged males with pronotum projection (Figure 4H, O). These data indicate that *dsx* promotes the development of the sex combs and the modification of the rear-legs and suppresses the development of the pronotum projection and abdomen modification in males. Our *dsx* RNAi in females did not result in ectopic sex combs or affect the pronotum projection in females, suggesting that female-specific antagonistic traits are not under *dsx* control. All together our results indicate that the sex determination gene *dsx* and the Hox genes *Scr* and *Ubx* are involved to shape antagonistic structures in males and females in the context of sexual conflict.

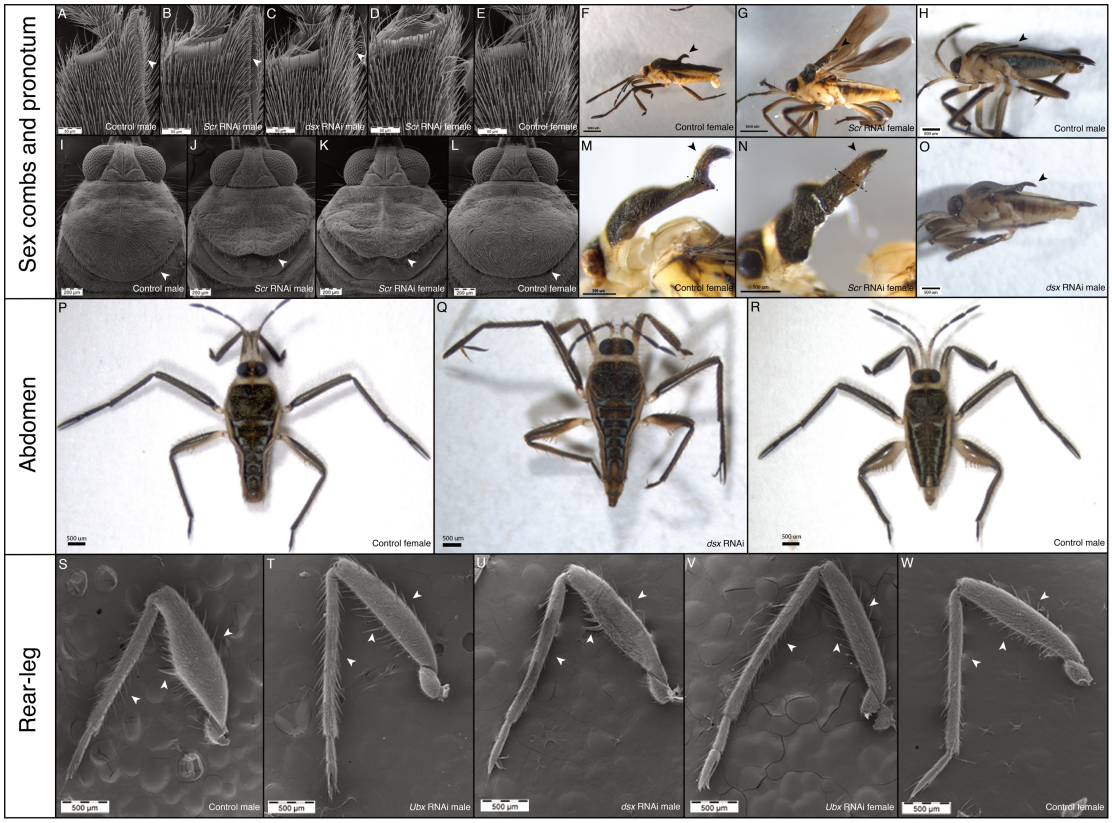


Figure 4: Effect of RNAi knockdown on secondary sexual traits in *R. antilleana*. (A-E) The sex combs in male (A) is affect by both *Scr* RNAi (B) and *dsx* RNAi (C) while we do not observe effect in female for *Scr* RNAi (D) due to the absence of this structure in female (E). The pronotum in wingless male (I) and wingless female (L) is also affected by *Scr* RNAi (J) and (K). In winged female (F) and (M), the *Scr* RNAi affects the development of the pronotum projection (G) and (N). In winged male (H) the *dsx* RNAi induces the development of the pronotum projection (O). *dsx* RNAi also affects the shape of the abdomen and genitalia (Q) generating a phenotype mixing female (P) and male aspects (R). For the male rear-leg (S), the *Ubx* RNAi affects the shape and secondary structures of the leg (T) that become similar to female rear-leg (W). However the *Ubx* RNAi also affects the few modifications present of female legs (V). The *dsx* RNAi generates a phenotype similar to *Ubx* RNAi (U).

Conclusion

Understanding the processes underlying the evolution of sexual dimorphism is a major challenge in evolutionary biology. We have shown how the evolutionary dynamics of sexual conflict escalation and intralocus conflict resolution may have shaped the evolutionary trajectory of a lineage of water striders. Our phylogenetic analysis revealed that increased complexity in traits increasing male persistence is generally matched by increased complexity in traits increasing female resistance. The development of these antagonistic traits in both sexes is under the control of shared loci with highly importance and pleiotropic developmental function in both sexes. Interestingly, some of these loci are associated with the adaptation in one sex and its counter-adaptation in the other. This is the case of *Scr*, which is responsible for the sex comb allowing males to clasp and the pronotum projection allowing females to resist. In these cases, intra-locus sexual conflict has been fully resolved through sex-restricted function that result in sexual dimorphism, such as males that develop without pronotum projection and females without sex combs. Other cases, such as *Ubx*, still show significant genetic correlation as the function is present in both sexes and *Ubx* knockdown removes rear-leg modification in both males and females. How, at the molecular genetic level, certain traits remain genetically correlated between the sexes and not others remain unknown. Altogether our data reveal how lineages can diversify under sexual conflict and the evolution of sex-specific action of a set of highly pleiotropic and conserved loci such as Hox and sex determination genes.

Materials and methods

Insect sampling and culture

Species were collected during fieldwork in the locations indicated in Table S1. Lab populations were established for some species. They were kept in water tanks at 25°C, 50% humidity, 14 hours of daylight and fed daily on crickets. Styrofoam floaters were provided for adult female eggs laying. Adults and nymphs were raised in independent tanks to decrease nymphs mortality by adults.

Video acquisition and imaging

Video acquisition to observe interactions between males and females during pre-mating struggles was performed using the Phantom *Miro M310* Digital High Speed *Camera* and PCC Software (Vision research, Ametek). Picture acquisition and observation of secondary sexual traits were performed using a SteREO Discovery V12 (Zeiss), a Stereomicroscope M205 FA (Leica) and using Scanning Electron Microscopy at Centre Technologique des Microstructures (UCBL).

Quantification of trait functions

Both non-virgin males and females were isolated into two different buckets, one for each sex, during 5 days, in order to increase their reproductory behavior. We set-up tournaments between 8 males and 8 females to split them in three different categories depending the ability in mating for males and the ability in rejecting male for females: successful, intermediate and unsuccessful. Both males and females were introduced at the same time in a same bucket of water. During the first round, for each mating we removed the couple and interrupted the mating by separating the male and the female in order to preserve the motivation to mate for the second round. The male was classified as successful and the female as unsuccessful. Inversely the males that did not succeed to mate were classified as unsuccessful and females as successful. The first round occurs until we got 4 successful males and 4 unsuccessful females. The different rounds of this experiment occurred for few minutes to several hours depending the success rate of the bugs. For the second round, we used two different buckets. In the first bucket we introduced, at the same time, the 4 successful males with the 4 successful females and in the second bucket we introduced, at the same time, the 4 unsuccessful males with the 4 unsuccessful females. At each mating we removed the couple, as in the first round. The second round occurs until we got 2 successful males and 2 unsuccessful females. At the end we were able to obtain the following rank for both males and females: successful for males that have mated twice and females that have never mated, intermediate for individuals that mated once, and unsuccessful for males that have never mated and for females that mated twice. Then using SteREO Discovery V12 (Zeiss) with ZEN 2011 software (Zeiss) we took pictures of each individual to quantify the individual phenotypes. We measured the length and the width of the rear-leg femur, the length of the rear-leg tibia and we counted the number of spikes on the rear-leg femur and the number of buds on the rear-leg tibia. We also recorded the presence of a pronotum projection in female. Using the measurement of femur width and tibia length we inferred the grasping ability of male rear-legs (Figure 2H).

Statistical analysis

Student t-test and Wilcoxon test between trait values of successful and unsuccesful males, depending whether variables followed normal distributions, were performed using RStudio version 1.0.153. Chi2 distribution tests of male and female morphs involved in tournaments were performed using Past3 version 3.14.

Muscle staining

Male rear-leg femurs were dissected and opened using forceps and scalpel to remove the cuticle. Femurs were fixed with Formaldehyde 4% in PTW 1% during 20 minutes. Femurs were then washed in 5 successive 10-minute baths of PTW 1%. The muscles were marked with 1/1000 Phalloïdine 488 for 1 hour followed by 5 successive 10-minute baths of PTW 1%. The femur were mounted in 50% glycerol + Dapi.

Phylogenetic reconstruction

Sequences were retrieved from in house transcriptome databases for the following markers: *12S RNA*; *16S RNA*; *18S RNA*; *28S RNA*; *Cytochrome Oxydase subunit I (COI); Cytochrome Oxydase subunit II (COII); Cytochorme Oxydase subunit III (COIII); Cytochrome b (cyt b); NADH-ubiquinone oxidoreductase chain 1 (ND1); Ultrabithorax (Ubx); Sex combs reduced (Scr); Gamma interferon inducible thiol reductase (gilt); Antennapedia (Antp); Distal-less (dll) ; Doublesex (dsx)*. All these markers were submitted to GenBank and their accession numbers can be found in the Dryad Digital Repository: xxxxx. Phylogenetic reconstruction was perfomed using MrBayes version 3.2.6 and PhyML version 3.0 in Geneious 7.1.9 as described in (Crumiere et al., 2016). Concatenation of sequence alignments and phylogenetic tree in Newick format are also available in the Dryad Digital Repository: xxxxx.

Test of sexual conflict intensity

We created a matrix of presence/absence of secondary sexual traits for both males and females (Table 2). Based on our observation, we found eight traits in males; the sex combs in the forelegs, the different rows of spines and the shape of the different segments of rear-legs. In females, these traits include the pronotum projection and the narrow abdomen shape. Then we mapped the sexually antagonistic traits individually on our phylogeny to reconstruct the ancestral state of each of them. Finally we performed a reconstruction using the combination of all traits, in males and females separately, to determine the pattern of phylogenetic complexity in this sample. Reconstruction of ancestral state was performed using Mesquite version 3.2 (Maddison and Maddison, 2017) and mapping of phenotypic complexity was performed using contMap (package phytools, (Revell, 2012)) in RStudio version 1.0.153.

Cloning of Scr, Ubx and Dsx genes and Nymphal RNA interference knockdown

Extraction of total RNA from *Rhagovelia antilleana* nymphs was performed using Trizol extraction protocol. cDNA was made using SuperScript III First-Strand kit according to manufacturer instructions (Invitrogen). Fragment of Scr, Ubx and Dsx genes were amplified by PCR using GoTaq G2 DNA polymerase (Promega) and the following primers (forward/reverse): For Scr, Rantilleana-scr-fw7 5' GAATACGCTCAGCTCGGTTC 3'/Rantilleana-scr-rev7 5' TTGCCGAGGTAGTGTTGTTG 3' ; for Ubx, SAB-Ubx-ATG-F 5' ATGAATTCTTACTTTGAGCAGACGGGT 3' /Ubx HDR6 5' CTGCTATAGCCATCCAGGGGTAGAA 3' ; for Dsx, Dsx-RA-F2 5' TTACAGCCTTGCATCGTGCACTCTA 3'/ Dsx-RA-R2 5' GTTGGCAACATTGACAGCTTCTGAC 3'. PCR products were checked by electrophoresis, purified using PCR Minelute kit (QUIAGEN) and were cloned inside pGEM-T vector kit (Promega) followed by bacteria transformation. Positive clones were sequenced using Genewiz sequencing facilities. Plasmid extractions were done with QIAprep spin Miniprep Kit (Qiagen). Then PCR on insert was performed using the following primers containing T7 promoter sequence (in bold) for T7 RNA polymerase (forward/reverse): for Scr, Rantilleana-scr-fw-T7 5' **TAATACGACTCACTATAGGGAGACCAC**GAATACGCTCAGCTCGGTTC 3'/ Rantilleana-scr-rev-T7 5' **TAATACGACTCACTATAGGGAGACCAC**TTGCCGAGGTAGTGTTGTTG 3' ; for Ubx, SAB-Ubx-ATG-F-T7 5' **TAATACGACTCACTATAGGGAGACCAC**ATGAATTCTTACTTTGAGCAGACGGGT 3'/ Ubx-HDR6-T7 5' **TAATACGACTCACTATAGGGAGACCAC**CTGCTATAGCCATCCAGGGGTAGAA 3' ; for dsx, dsx-RA-F2-T7 5' **TAATACGACTCACTATAGGGAGACCAC**TTACAGCCTTGCATCGTGCACTCTA 3'/dsx-RA-R2-T7 5' **TAATACGACTCACTATAGGGAGACCAC**GTTGGCAACATTGACAGCTTCTGAC 3'. PCR products were purified using PCR Minelute kit (QUIAGEN) and in vitro transcription with T7 RNA polymerase (Ambion) was performed to make double stranded RNA (dsRNA). DsRNA were purified using RNEasy kit (QUIAGEN), concentrated with speedvac and resuspend in 1X injection buffer (Rubin and Spradling, 1982) at the final concentration of 3 ug/uL for Scr dsRNA, 3 ug/uL for Ubx dsRNA, 2 ug/uL for Dsx dsRNA. Yfp dsRNA was used as a control at 1,8 ug/uL concentration. We performed injections in Rhagovelia antilleana first to fifth nymphal instars using a SteREO Discovery V8 (Zeiss), a Cell Tram Vario Oil Eppendorf injector and a Narishige micromanipulator under CO2 anesthesia. A total of 47, 348, 81 and 297 nymphs were injected respectively with Scr dsRNA, Ubx dsRNA, Dsx dsRNA, and Yfp dsRNA. The nymphs were reached as described in the insect culture paragraph below after injection.

References

Andersen, N.M., 1982. The semiaquatic bugs (Hemiptera: Gerromorpha). Scandinavian Science Press LTD., Klampenborg, Denmark.

Andersson, M.B., 1994. Sexual selection. Princeton University Press, Princeton, N.J.

Arnqvist, G., 1992. SPATIAL VARIATION IN SELECTIVE REGIMES: SEXUAL SELECTION IN THE WATER STRIDER, GERRIS ODONTOGASTER. Evolution 46, 914-929.

Arnqvist, G., Rowe, L., 1995. Sexual Conflict and Arms Races between the Sexes - a Morphological Adaptation for Control of Mating in a Female Insect. P Roy Soc B-Biol Sci 261, 123-127.

Arnqvist, G., Rowe, L., 2002a. Antagonistic coevolution between the sexes in a group of insects. Nature 415, 787-789.

Arnqvist, G., Rowe, L., 2002b. Correlated evolution of male and female morphologles in water striders. Evolution 56, 936-947.

Arnqvist, G., Rowe, L., 2005. Sexual Conflict. Princeton University Press.

Averof, M., Akam, M., 1993. HOM/Hox genes of Artemia: implications for the origin of insect and crustacean body plans. Curr Biol 3, 73-78.

Averof, M., Akam, M., 1995. Hox genes and the diversification of insect and crustacean body plans. Nature 376, 420-423.

Bergsten, J., Miller, K.B., 2007. Phylogeny of diving beetles reveals a coevolutionary arms race between the sexes. PLoS One 2, e522.

Bonduriansky, R., Chenoweth, S.F., 2009. Intralocus sexual conflict. Trends Ecol Evol 24, 280-288.

Braendle, C., Caillaud, M.C., Stern, D.L., 2005a. Genetic mapping of aphicarus - a sex-linked locus controlling a wing polymorphism in the pea aphid (Acyrthosiphon pisum). Heredity 94, 435-442.

Braendle, C., Friebe, I., Caillaud, M.C., Stern, D.L., 2005b. Genetic variation for an aphid wing polyphenism is genetically linked to a naturally occurring wing polymorphism (vol 272, pg 657, 2005). P Roy Soc B-Biol Sci 272, 2659-2659.

Chesebro, J., Hrycaj, S., Mahfooz, N., Popadic, A., 2009. Diverging functions of Scr between embryonic and post-embryonic development in a hemimetabolous insect, Oncopeltus fasciatus. Dev Biol 329, 142-151.

Crumiere, A.J.J., Santos, M.E., Semon, M., Armisen, D., Moreira, F.F.F., Khila, A., 2016. Diversity in Morphology and Locomotory Behavior Is Associated with Niche Expansion in the Semi-aquatic Bugs. Curr Biol 26, 3336-3342.

Davis, G.K., Srinivasan, D.G., Wittkopp, P.J., Stern, D.L., 2007. The function and regulation of Ultrabithorax in the legs of Drosophila melanogaster. Dev Biol 308, 621-631.

Dean, R., Perry, J.C., Pizzari, T., Mank, J.E., Wigby, S., 2012. Experimental evolution of a novel sexually antagonistic allele. PLoS Genet 8, e1002917.

Ditrich, T., Papacek, M., 2010. Effect of population density on the development of Mesovelia furcata (Mesoveliidae), Microvelia reticulata and Velia caprai (Veliidae) (Heteroptera: Gerromorpha). Eur J Entomol 107, 579-587.

Fairbairn, D.J., 1988. Adaptive Significance of Wing Dimorphism in the Absence of Dispersal - a Comparative-Study of Wing Morphs in the Waterstrider, Gerris-Remigis. Ecol Entomol 13, 273-281.

Gilliver, S.F., Degens, H., Rittweger, J., Sargeant, A.J., Jones, D.A., 2009. Variation in the determinants of power of chemically skinned human muscle fibres. Exp Physiol 94, 1070-1078.

Hayashi, K., 1985. Alternative mating strategies in the water strider Gerris clongntus (Heteroptera, Gerridae). Behavioral Ecology and Sociobiology 16, 301-306.

Hughes, C.L., Kaufman, T.C., 2002. Hox genes and the evolution of the arthropod body plan. Evol Dev 4, 459-499.

Ivy, T.M., Sakaluk, S.K., 2007. Sequential mate choice in decorated crickets: females use a fixed internal threshold in pre- and postcopulatory choice. Anim Behav 74, 1065-1072.

Khila, A., Abouheif, E., Rowe, L., 2012. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. Science (New York, N Y) 336, 585-589.

Khila, A., Abouheif, E., Rowe, L., 2014. Comparative functional analyses of ultrabithorax reveal multiple steps and paths to diversification of legs in the adaptive radiation of semi-aquatic insects. Evolution 68, 2159-2170.

Kopp, A., Duncan, I., Godt, D., Carroll, S.B., 2000. Genetic control and evolution of sexually dimorphic characters in Drosophila. Nature 408, 553-559.

Kuntner, M., Coddington, J.A., Schneider, J.M., 2009. Intersexual arms race? Genital coevolution in nephilid spiders (Araneae, Nephilidae). Evolution 63, 1451-1463.

Maddison, W.P., Maddison, D.R., 2017. Mesquite: a modular system for evolutionary analysis.

Mank, J.E., 2017. The transcriptional architecture of phenotypic dimorphism. Nat Ecol Evol 1, 6.
Mank, J.E., Wedell, N., Hosken, D.J., 2013. Polyandry and sex-specific gene expression. Philos Trans R Soc Lond B Biol Sci 368, 20120047.

Moreira, F.F.F., Ribeiro, J.R.I., 2009. Two new Rhagovelia (Heteroptera: Veliidae) and new records for twelve species in southeastern Brazil. Aquat Insect 31, 45-61.

Pennell, T.M., Morrow, E.H., 2013. Two sexes, one genome: the evolutionary dynamics of intralocus sexual conflict. Ecol Evol 3, 1819-1834.

Perry, J.C., Rowe, L., 2015. The Evolution of Sexually Antagonistic Phenotypes. Cold Spring Harb Perspect Biol 7.

Polhemus, D.A., 1997. SYSTEMATICS OF THE GENUS RHAGOVELIA MAYR (HETEROPTERA: VELIIDAE) IN THE WESTERN HEMISPHERE (EXCLUSIVE OF THE ANGUSTIPES COMPLEX). THOMAS SAY PUBLICATIONS.

Refki, P.N., Armisen, D., Crumiere, A.J.J., Viala, S., Khila, A., 2014. Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. Dev Biol 392, 441-453.

Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3, 217-223.

Rowe, L., Arnqvist, G., Sih, A., J Krupa, J., 1994. Sexual conflict and the evolutionary ecology of mating patterns: water striders as a model system. Trends Ecol Evol 9, 289-293.

Rozowski, M., Akam, M., 2002. Hox gene control of segment-specific bristle patterns in Drosophila. Genes Dev 16, 1150-1162.

Snook, R.R., Bacigalupe, L.D., Moore, A.J., 2010. The quantitative genetics and coevolution of male and female reproductive traits. Evolution 64, 1926-1934.

Spence, J.R., 1989. The Habitat Templet and Life-History Strategies of Pond Skaters (Heteroptera, Gerridae) - Reproductive Potential, Phenology, and Wing Dimorphism. Can J Zool 67, 2432-2447.

Stern, D.L., 2003. The Hox gene Ultrabithorax modulates the shape and size of the third leg of Drosophila by influencing diverse mechanisms. Dev Biol 256, 355-366.

Tanaka, K., Barmina, O., Sanders, L.E., Arbeitman, M.N., Kopp, A., 2011. Evolution of sex-specific traits through changes in HOX-dependent doublesex expression. PLoS Biol 9, e1001131.

Weigensberg, I., Fairbairn, D.J., 1996. The sexual arms race and phenotypic correlates of mating success in the waterstrider, Aquarius remigis (Hemiptera: Gerridae). J Insect Behav 9, 307-319.

Zera, A.J., Innes, D.J., Saks, M.E., 1983. Genetic and Environmental Determinants of Wing Polymorphism in the Waterstrider Limnoporus-Canaliculatus. Evolution 37, 513-522.

Supplementary information for

Dimorphism under sexual conflict controlled by the combination of sex determination and Hox genes

Article in preparation

[1]A. J. J. Crumière, [1]D. Armisén, [1]M. Koubarakos, [2]F. F. F. Moreira and [1]A. Khila

Author affiliations:

[1]Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5242, Ecole Normale
Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France

[2]Laboratório de Biodiversidade Entomológica, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21040-360, Brazil

Species	Sampling location
Microvelia americana	Rivière du Nord, Montréal, Québec, Canada
Paravelia bullialata	French Guiana, N 04.29769°, W -52.14927°
Oiovelia cunucunumana	Brazil, N -20.25294°, W -44.91584°
Stridulivelia tersa	Brazil, N -19.07218°, W -39.79617°
Rhagovelia mangle	Brazil, N -18,669098°, W -39.7643°
Rhagovelia plumbea	USA, N 26.80819°, W -80.05566°
Rhagovelia tenuipes	Brazil, N -19.481264°, W-39.925354°
Rhagovelia amazonensis	Brazil, N -02.93451°, W -59.95137°
Rhagovelia obesa	Washington DC, USA
Rhagovelia antilleana	El Yunque, Puerto Rico
Rhagovelia humboldti	French Guiana, N 04.26076°, W -52.13031°
Rhagovelia sp. 1	French Guiana, N 04.62387°, W -52.31003°
Rhagovelia sp. 2	French Guiana, N 04.62387°, W -52.31003°

Table S1: List of species use in this study and location of sampling.

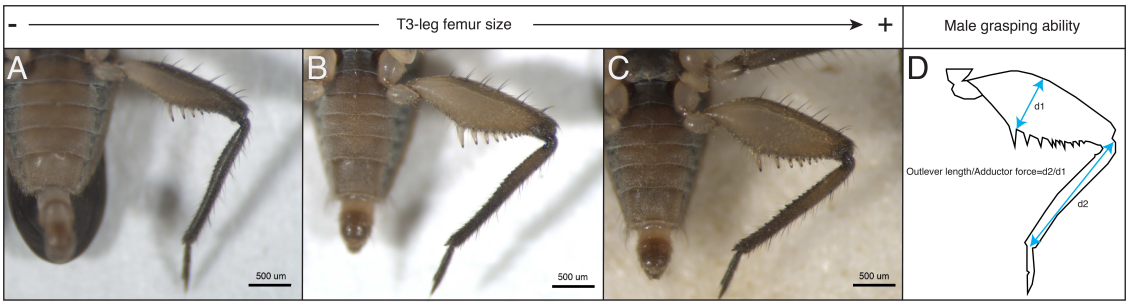


Figure S1: Variability in rear-leg femur size in males with representation of mechanical advantage calculation. Males have variability in femur width with male having small femur (A), medium femur (B) and big femur (C). Using the calculation in (D) we calculated a proxy of male grasping ability.

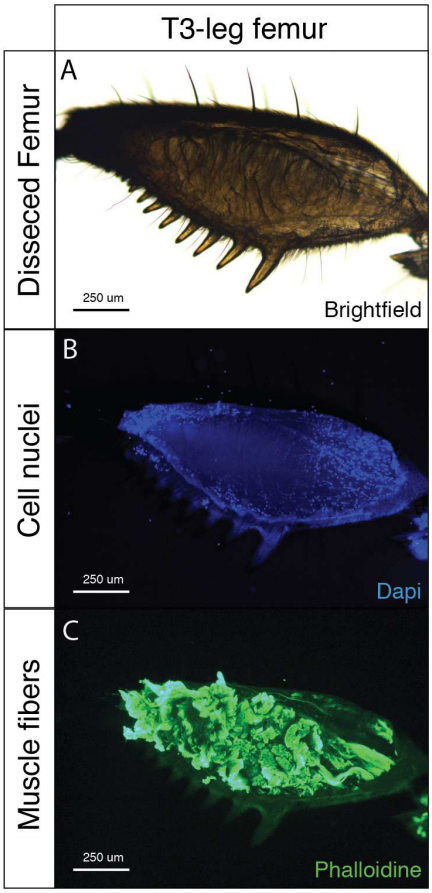


Figure S2: Muscle staining. The rear-leg femur (A) contains muscles that we can detect using Phaloidine (C). Dapi (B) is used as control of staining.

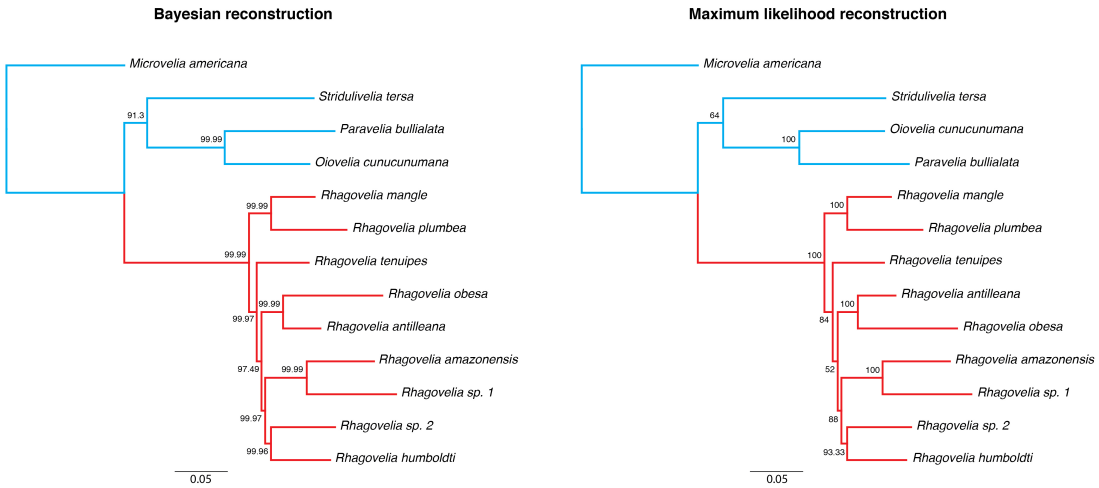


Figure S3: Phylogenetic reconstruction of our sample of species with both Bayesian (with posterior probability) and Maximum Likelihood (with bootstrap) methods.

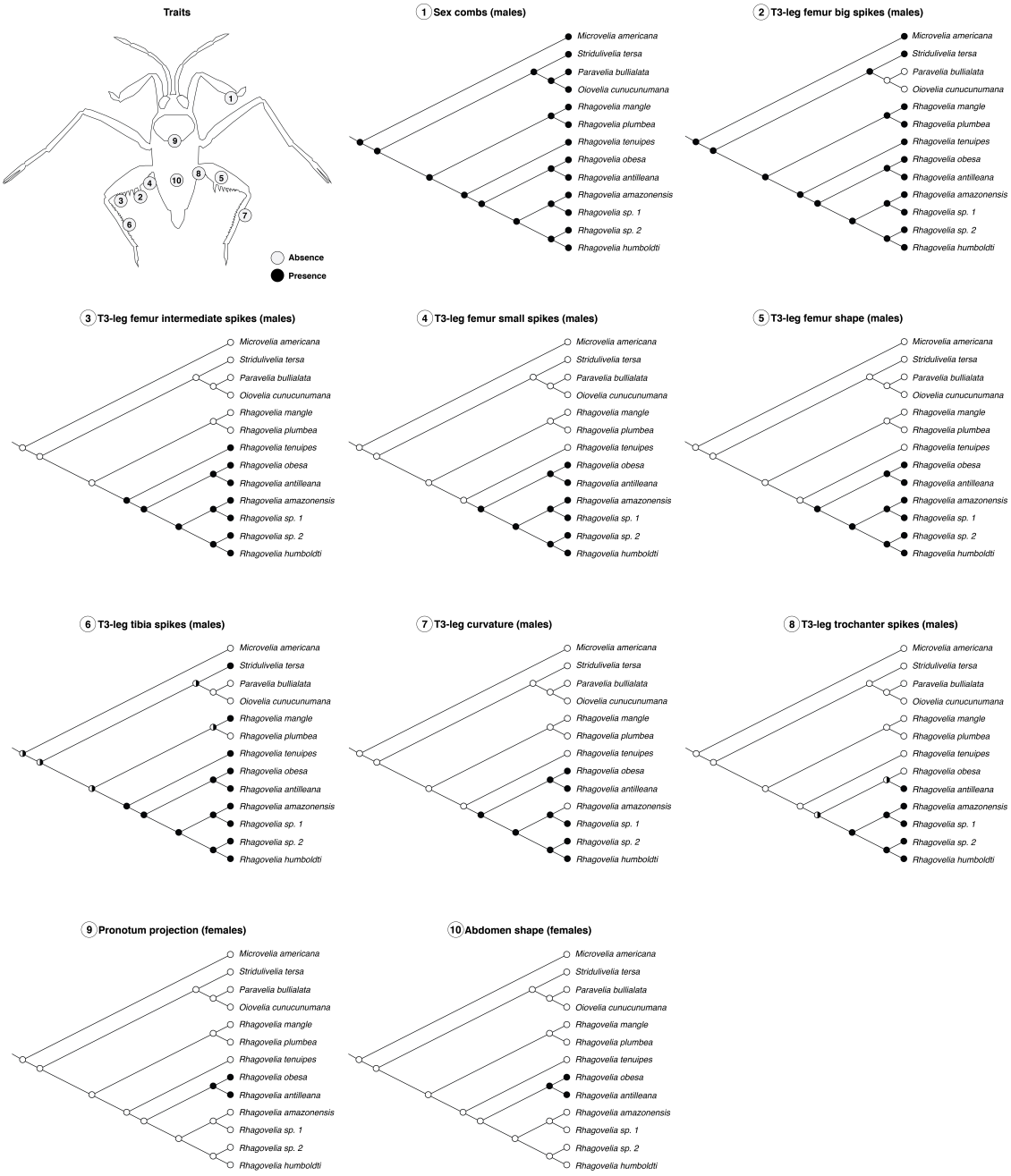


Figure S4: Reconstruction of ancestral trait for the different secondary sexual traits. Each tree represent the evolution of each trait we used in this study based on our matrix of traits and Bayesian phylogenetic reconstruction.

General discussion

Ecological pressures are important factors generating the diversity of morphologies we observe in nature. In the case of the semiaquatic bugs, the new predator-prey interactions and the physics of the substrate are the selective pressures associated with the colonization of water surface habitats. Interestingly species phenotypic divergence is often due to differences in lifestyle and sub niches rather than phylogenetic relationships. For example, species presumed to more closely resemble the common ancestor (basally branching), but with preference for water, are more adapted to life on water than more derived species with preference for solid substrate. Furthermore species can be part of the same family, sometimes even part of the same genus, and can possess different leg morphologies and use different modes of locomotion associated with their respective sub-niches. Ecological pressures also influence species diversification. Selective pressures, from both land and water, seem to constrain the diversification of the families inhabiting the shore, such as the families Mesoveliidae (46 species), Hydrometridae (130 species) or Hebridae (220 species). In contrast, in their new and exclusively aquatic niches the Gerridae family reaches about 710 species and the *Rhagovelia* genus (part of the Veliidae family with 900 species) about 200 species (Andersen 1982). Sexual selection represents another important factor that generates phenotypic diversity. Contrary to physical environmental conditions, sexual selection tends to drive the divergence of the sexes. In the presented thesis, males and females of a given species coevolved in phenotypic complexity as the result of strong sexual selection. Both sexes have evolved a set of antagonistic secondary sexual traits that contribute to sexual dimorphism. Altogether these observations highlight the importance of the ecological environment and interaction on diversification. The results obtained in this thesis highlight the importance of evaluating the adaptive value of phenotypes to improve our understanding of these processes. It would be interesting to study how an optimal phenotype is selected in response of both environmental pressures and sexual selection.

In the presented thesis I have also investigated the developmental genetic mechanisms underlying adaptive traits associated with the different steps of the ecological transition from land to water surface habitat and subsequent niche specialization in the Gerromorpha. These different adaptations have been possible through different genetic mechanisms such as new gene expression pattern of the Hox

gene *Ubx* that allowed leg elongation, the tissue sensitivity to Ubx protein level involved in the establishment of the derived leg ground plan, and new genetic interactions and neofunctionalization of duplicate genes that allowed respectively the elongation of the tarsus in Gerridae and the development of the propelling fan in *Rhagovelia*. It is also important to notice that the different adaptive phenotypes studied in the presented thesis support the idea that specific traits derived from previous adaptations that acted as precursors to subsequent adaptations (Figure 10). For example the hydrophobic bristles allowed to stand on water and then the increase in stroke frequency associated with leg elongation allowed access to water surface habitats (Figure 10). The initial leg elongation and modification of the underlying genetic network could act as a prerequisite to the strong leg elongation and the new leg ground plan that allowed the rowing gait in species inhabiting open water (Figure 10). Again the rowing gait has probably allowed colonization of niches such as river margins until the apparition of the propelling fan that allowed the colonization of stream water (Figure 10).

The role of the Hox genes in the diversification of morphologies is quite established as the role of the protein level ((Hughes & Kaufman, 2002; Akam et al, 1994; Averof & Akam 1993; Averof & Akam 1995; Krumlauf 1994;McGinnis et al, 1984; McGinnis & Krumlauf 1992; Scott & Weiner 1984; Pearson et al, 2005; Garcia-Fernandez 2005). It is also the case of gene co-option for which we can find many examples in the literature (Carroll 2008 ; Glassford et al, 2015 ; Khila et al, 2014 ; Lynch & Wagner 2008 ; Monteiro & Gupta 2016). Based on the results obtained along my PhD, the importance of these genes and processed are reinforced. However we have also highlighted an example of the importance of taxon restricted gene duplicates at the origin of adaptive traits associated with the expansion of habitat in a specific genus that then has diversified. Evidences for the role of taxon-restricted genes remained scarce in literature (Japser et al, 2015 ; Kaessmann 2010 ; Khalturin et al, 2009 ; Margulies et al, 2005) and our results provide evidences for the importance of such genes in the evolution of specific lineages. In all these cases environmental pressures select the traits that equally affect both males and females. In parallel, it is interesting to observe that the Hox gene *Ubx*, with *Scr*, also participate to generate sexual dimorphism. Their specific roles during the development of sex-specific secondary antagonistic traits seem to be dependent, in part, on sex determination

genes such as *doublesex* that contribute to the fine-tuning of structures. These results confirm the importance of the Hox genes during the development of traits selected by both environmental and sexual selective pressures.

To conclude, the work I performed along my PhD thesis has improved our understanding of how selective and developmental genetic mechanisms interact together to generate diversity. Hopefully, a lot of interesting questions remain to be elucidated and exciting stories remain to be shared.

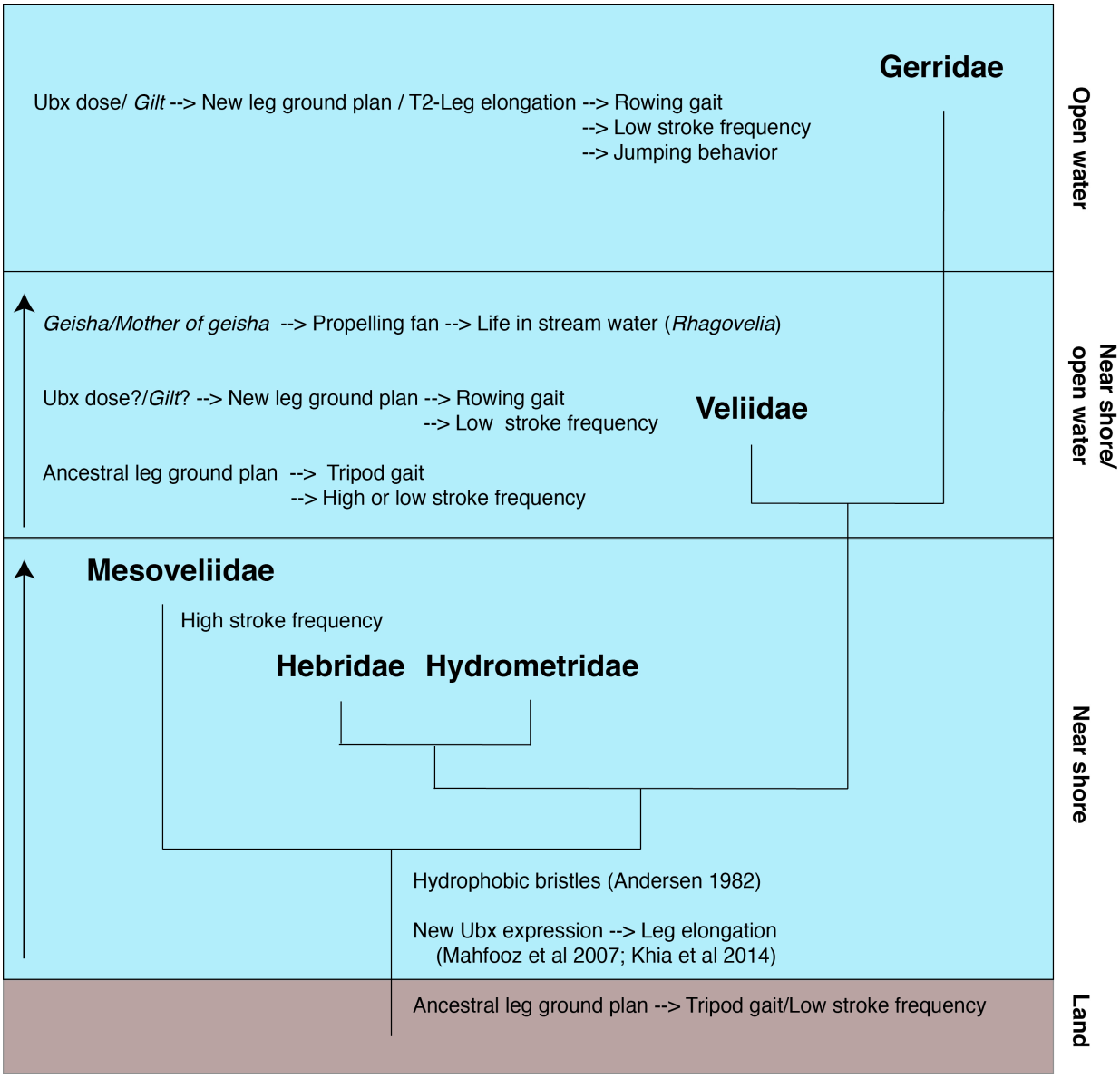


Figure 10: The different adaptive phenotypes and underlying genes.
This phylogeny summarizes the different Gerromorpha families used in this thesis with the sub niches where they live, the different adaptive phenotypes they have and associated underlying genes. The arrows indicate the degree of specialization inside the sub niche.

References

Abzhanov, Arhat, Meredith Protas, B. Rosemary Grant, Peter R. Grant, and Clifford J. Tabin. "Bmp4 and Morphological Variation of Beaks in Darwin's Finches." [In English]. *Science (New York, N Y)* 305, no. 5689 (2004): 1462-5.

Akam, M., M. Averof, J. Castelli-Gair, R. Dawes, F. Falciani, and D. Ferrier. "The Evolving Role of Hox Genes in Arthropods." [In English]. *Development (Cambridge, England) Supplement* (1994): 209-15.

Albalat, Ricard, and Cristian Canestro. "Evolution by Gene Loss." [In English]. *Nature reviews Genetics* 17, no. 7 (2016): 379-91.

Albert, Victor A., David G. Oppenheimer, and Charlotte Lindqvist. "Pleiotropy, Redundancy and the Evolution of Flowers." [In English]. *Trends in plant science* 7, no. 7 (2002): 297-301.

Andersen, N. M., and G. O. Poinar. "A Marine Water Strider (Hemiptera : Veliidae) from Dominican Amber." [In English]. *Entomologica Scandinavica* 29, no. 1 (1998): 1-9.

Andersen, N.M. "A Comparative Study of Locomotion on the Water Surface in Semiaquatic Bugs (Insecta, Hemiptera, Gerromorpha)." *Vidensk. Meddr dansk naturh. Foren.*, no. 139 (1976): 337-96.

Andersen, N.M. *The Semiaquatic Bugs (Hemiptera: Gerromorpha).* Vol. - Entomonograph Vol. 3. , Klampenborg, Denmark.: Scandinavian Science Press LTD., 1982.

Andersson, M. B. *Sexual Selection*. Monographs in Behavior and Ecology. Princeton, N.J.: Princeton University Press, 1994.

Andreadis, A., M. E. Gallego, and B. Nadal-Ginard. "Generation of Protein Isoform Diversity by Alternative Splicing: Mechanistic and Biological Implications." [In English]. *Annual review of cell biology* 3 (1987): 207-42.

Angelini, David R., Paul Z. Liu, Cynthia L. Hughes, and Thomas C. Kaufman. "Hox Gene Function and Interaction in the Milkweed Bug Oncopeltus Fasciatus (Hemiptera)." [In English]. *Developmental biology* 287, no. 2 (2005): 440-55.

Arnoult, Laurent, Kathy F. Y. Su, Diogo Manoel, Caroline Minervino, Justine Magrina, Nicolas Gompel, and Benjamin Prud'homme. "Emergence and Diversification of Fly Pigmentation through Evolution of a Gene Regulatory Module." [In English]. *Science (New York, N Y)* 339, no. 6126 (2013): 1423-6.

Arnqvist, G, and L Rowe. *Sexual Conflict*. Princeton University Press, 2005.

Arnqvist, G., and L. Rowe. "Sexual Conflict and Arms Races between the Sexes - a Morphological Adaptation for Control of Mating in a Female Insect." [In English]. *Proceedings of the Royal Society B-Biological Sciences* 261, no. 1360 (Jul 22 1995): 123-27.

Arnqvist, Goran, and Locke Rowe. "Correlated Evolution of Male and Female Morphologles in Water Striders." [In English]. *Evolution; international journal of organic evolution* 56, no. 5 (2002): 936-47.

Averof, M., and M. Akam. "Hom/Hox Genes of Artemia: Implications for the Origin of Insect and Crustacean Body Plans." [In English]. *Current biology : CB* 3, no. 2 (1993): 73-8.

Averof, M., and M. Akam. "Hox Genes and the Diversification of Insect and Crustacean Body Plans." [In English]. *Nature* 376, no. 6539 (1995): 420-3.

Averof, M., and S. M. Cohen. "Evolutionary Origin of Insect Wings from Ancestral Gills." [In English]. *Nature* 385, no. 6617 (1997): 627-30.

Barrett, Rowan D. H., Sean M. Rogers, and Dolph Schluter. "Natural Selection on a Major Armor Gene in Threespine Stickleback." [In English]. *Science (New York, N Y)* 322, no. 5899 (2008): 255-7.

Bell, Michael A., Guillermo Orti, Jeffrey A. Walker, and Jeffrey P. Koenings. "Evolution of Pelvic Reduction in Threespine Stickleback Fish: A Test of Competing Hypotheses." [In English]. *Evolution; international journal of organic evolution* 47, no. 3 (1993): 906-14.

Bergsten, J., A. Toyra, and A. N. Nilsson. "Intraspecific Variation and Intersexual Correlation in Secondary Sexual Characters of Three Diving Beetles (Coleoptera : Dytiscidae)." [In English]. *Biological Journal of the Linnean Society* 73, no. 2 (Jun 2001): 221-32.

Bergsten, Johannes, and Kelly B. Miller. "Phylogeny of Diving Beetles Reveals a Coevolutionary Arms Race between the Sexes." [In English]. *PloS one* 2, no. 6 (2007): e522.

Brennan, P. L. R., C. J. Clark, and R. O. Prum. "Explosive Eversion and Functional Morphology of the Duck Penis Supports Sexual Conflict in Waterfowl Genitalia." [In English]. *Proceedings of the Royal Society B-Biological Sciences* 277, no. 1686 (May 7 2010): 1309-14.

Brennan, Patricia L. R., and Richard O. Prum. "Mechanisms and Evidence of Genital Coevolution: The Roles of Natural Selection, Mate Choice, and Sexual Conflict." [In English]. *Cold Spring Harbor perspectives in biology* 7, no. 7 (2015): a017749.

Carroll, Sean B. "Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution." [In English]. *Cell* 134, no. 1 (2008): 25-36.

Casasa, Sofia, Daniel B. Schwab, and Armin P. Moczek. "Developmental Regulation and Evolution of Scaling: Novel Insights through the Study of Onthophagus Beetles." [In English]. *Current opinion in insect science* 19 (2017): 52-60.

Chanderbali, Andre S., Brent A. Berger, Dianella G. Howarth, Pamela S. Soltis, and Douglas E. Soltis. "Evolving Ideas on the Origin and Evolution of Flowers: New Perspectives in the Genomic Era." [In English]. *Genetics* 202, no. 4 (2016): 1255-65.

Chase, Jonathan M., and Mathew A. Leibold. *Ecological Niches : Linking Classical and Contemporary Approaches*. Interspecific Interactions. Chicago: University of Chicago Press, 2003.

Chesebro, John, Steven Hrycaj, Najmus Mahfooz, and Aleksandar Popadic. "Diverging Functions of Scr between Embryonic and Post-Embryonic Development in a Hemimetabolous Insect, Oncopeltus Fasciatus." [In English]. *Developmental biology* 329, no. 1 (2009): 142-51.

Damgaard, J. "Evolution of the Semi-Aquatic Bugs (Hemiptera: Heteroptera: Gerromorpha) with a Re-Interpretation of the Fossil Record." *ACTA ENTOMOLOGICA MUSEI NATIONALIS PRAGAE* 48, no. 2 (2008): 251-68.

Damgaard, J. "Phylogeny of the Semiaquatic Bugs (Hemiptera-Heteroptera, Gerromorpha)." [In English]. *Insect Systematics & Evolution* 39, no. 4 (2008): 431-60.

Darwin, Charles. *On the Origin of Species by Means of Natural Selection*. London,: J. Murray, 1859.

Dawkins, R., and J. R. Krebs. "Arms Races between and within Species." [In English]. *Proceedings of the Royal Society of London Series B, Biological sciences* 205, no. 1161 (1979): 489-511.

Di, Zhiyong, Yao Yu, Yingliang Wu, Pei Hao, Yawen He, Huabin Zhao, Yixue Li*, et al.* "Genome-Wide Analysis of Homeobox Genes from Mesobuthus Martensii Reveals Hox Gene Duplication in Scorpions." [In English]. *Insect biochemistry and molecular biology* 61 (2015): 25-33.

Duboule, D. "The Vertebrate Limb: A Model System to Study the Hox/Hom Gene Network During Development and Evolution." [In English]. *BioEssays : news and reviews in molecular, cellular and developmental biology* 14, no. 6 (1992): 375-84.

Fricke, Claudia, Darrell Green, Damian Smith, Tamas Dalmay, and Tracey Chapman. "Micrornas Influence Reproductive Responses by Females to Male Sex Peptide in Drosophila Melanogaster." [In English]. *Genetics* 198, no. 4 (2014): 1603-19.

Garcia-Fernandez, Jordi. "The Genesis and Evolution of Homeobox Gene Clusters." [In English]. *Nature reviews Genetics* 6, no. 12 (2005): 881-92.

Gay, Laurene, David J. Hosken, Paul Eady, Ram Vasudev, and Tom Tregenza. "The Evolution of Harm--Effect of Sexual Conflicts and Population Size." [In English]. *Evolution; international journal of organic evolution* 65, no. 3 (2011): 725-37.

Gehring, W. J. "The Homeo Box: A Key to the Understanding of Development?" [In English]. *Cell* 40, no. 1 (1985): 3-5.

Ghalambor, Cameron K., Jeffrey A. Walker, and David N. Reznick. "Multi-Trait Selection, Adaptation, and Constraints on the Evolution of Burst Swimming Performance." [In English]. *Integrative and comparative biology* 43, no. 3 (2003): 431-8.

Glassford, William J., Winslow C. Johnson, Natalie R. Dall, Sarah Jacquelyn Smith, Yang Liu, Werner Boll, Markus Noll, and Mark Rebeiz. "Co-Option of an Ancestral Hox-Regulated Network Underlies a Recently Evolved Morphological Novelty." [In English]. *Developmental cell* 34, no. 5 (2015): 520-31.

Gotoh, Hiroki, Robert A. Zinna, Ian Warren, Michael DeNieu, Teruyuki Niimi, Ian Dworkin, Douglas J. Emlen, Toru Miura, and Laura C. Lavine. "Identification and Functional Analyses of Sex Determination Genes in the Sexually Dimorphic Stag Beetle Cyclommatus Metallifer." [In English]. *BMC genomics* 17 (2016): 250.

Grant, Peter R., and B. Rosemary Grant. *40 Years of Evolution Darwin's Finches on Daphne Major Island*. Princeton: Princeton University Press, 2014.

Grant, Peter R., and B. Rosemary Grant. *How and Why Species Multiply : The Radiation of Darwin's Finches*. Princeton Series in Evolutionary Biology. Princeton: Princeton University Press, 2008.

Han, Chang S., and Piotr G. Jablonski. "Male Water Striders Attract Predators to Intimidate Females into Copulation." [In English]. *Nature communications* 1 (2010): 52.

Hannon, Gregory J. "Rna Interference." [In English]. *Nature* 418, no. 6894 (2002): 244-51.

Hayashi, K. "Alternative Mating Strategies in the Water Strider Gerris Clongntus (Heteroptera, Gerridae)." *Behavioral Ecology and Sociobiology* 16 (1985): 301-06.

Heiss, E., and J. Pericart. *Faune N° 91 – Hémiptères Aradidae, Piesmatidae Et Dipsocoromorphes*. Faune De France. Fédération Française des Sociétés de Sciences naturelles, 2007.

Hersh, Bradley M., Craig E. Nelson, Samantha J. Stoll, Jason E. Norton, Thomas J. Albert, and Sean B. Carroll. "The Ubx-Regulated Network in the Haltere Imaginal Disc of D. Melanogaster." [In English]. *Developmental biology* 302, no. 2 (2007): 717-27.

Hu, D. L., and J. W. M. Bush. "The Hydrodynamics of Water-Walking Arthropods." [In English]. *Journal of Fluid Mechanics* 644 (Feb 10 2010): 5-33.

Hughes, Cynthia L., and Thomas C. Kaufman. "Hox Genes and the Evolution of the Arthropod Body Plan." [In English]. *Evolution & development* 4, no. 6 (2002): 459-99.

Ito, Yuta, Ayane Harigai, Moe Nakata, Tadatsugu Hosoya, Kunio Araya, Yuichi Oba, Akinori Ito*, et al.* "The Role of Doublesex in the Evolution of Exaggerated Horns in the Japanese Rhinoceros Beetle." [In English]. *EMBO reports* 14, no. 6 (2013): 561-7.

Jasper, W. C., T. A. Linksvayer, J. Atallah, D. Friedman, J. C. Chiu, and B. R. Johnson. "Large-Scale Coding Sequence Change Underlies the Evolution of Postdevelopmental Novelty in Honey Bees." [In English]. *Molecular Biology and Evolution* 32, no. 2 (Feb 2015): 334-46.

Jindra, Marek, Xavier Belles, and Tetsuro Shinoda. "Molecular Basis of Juvenile Hormone Signaling." [In English]. *Current opinion in insect science* 11 (2015): 39-46.

Kaessmann, H. "Origins, Evolution, and Phenotypic Impact of New Genes." [In English]. *Genome Research* 20, no. 10 (Oct 2010): 1313-26.

Kessel, M., and P. Gruss. "Murine Developmental Control Genes." [In English]. *Science (New York, N Y)* 249, no. 4967 (1990): 374-9.

Kettlewell, H. B. "Recognition of Appropriate Backgrounds by the Pale and Black Phases of Lepidoptera." [In English]. *Nature* 175, no. 4465 (1955): 943-4.

Khalturin, K., G. Hemmrich, S. Fraune, R. Augustin, and T. C. G. Bosch. "More Than Just Orphans: Are Taxonomically-Restricted Genes Important in Evolution?" [In English]. *Trends in Genetics* 25, no. 9 (Sep 2009): 404-13.

Khila, Abderrahman, Ehab Abouheif, and Locke Rowe. "Comparative Functional Analyses of Ultrabithorax Reveal Multiple Steps and Paths to Diversification of Legs in the Adaptive Radiation of Semi-Aquatic Insects." [In English]. *Evolution; international journal of organic evolution* 68, no. 8 (2014): 2159-70.

Khila, Abderrahman, Ehab Abouheif, and Locke Rowe. "Evolution of a Novel Appendage Ground Plan in Water Striders Is Driven by Changes in the Hox Gene Ultrabithorax." [In English]. *PLoS genetics* 5, no. 7 (2009): e1000583.

Khila, Abderrahman, Ehab Abouheif, and Locke Rowe. "Function, Developmental Genetics, and Fitness Consequences of a Sexually Antagonistic Trait." [In English]. *Science (New York, N Y)* 336, no. 6081 (2012): 585-9.

Kijimoto, Teiya, Armin P. Moczek, and Justen Andrews. "Diversification of Doublesex Function Underlies Morph-, Sex-, and Species-Specific Development of Beetle Horns." [In English]. *Proceedings of the National Academy of Sciences of the United States of America* 109, no. 50 (2012): 20526-31.

Klepaker, T. "Morphological-Changes in a Marine Population of Threespine Stickleback, Gasterosteus-Aculeatus, Recently Isolated in Fresh-Water." [In English]. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* 71, no. 6 (Jun 1993): 1251-58.

Kopp, A., I. Duncan, D. Godt, and S. B. Carroll. "Genetic Control and Evolution of Sexually Dimorphic Characters in Drosophila." [In English]. *Nature* 408, no. 6812 (2000): 553-9.

Krumlauf, R. "Hox Genes in Vertebrate Development." [In English]. *Cell* 78, no. 2 (1994): 191-201.

Kunte, K., W. Zhang, A. Tenger-Trolander, D. H. Palmer, A. Martin, R. D. Reed, S. P. Mullen, and M. R. Kronforst. "Doublesex Is a Mimicry Supergene." [In English]. *Nature* 507, no. 7491 (2014): 229-32.

Kuntner, Matjaz, Ren-Chung Cheng, Simona Kralj-Fiser, Chen-Pan Liao, Jutta M. Schneider, and Mark A. Elgar. "The Evolution of Genital Complexity and Mating Rates in Sexually Size Dimorphic Spiders." [In English]. *BMC evolutionary biology* 16, no. 1 (2016): 242.

Kuntner, Matjaz, Jonathan A. Coddington, and Jutta M. Schneider. "Intersexual Arms Race? Genital Coevolution in Nephilid Spiders (Araneae, Nephilidae)." [In English]. *Evolution; international journal of organic evolution* 63, no. 6 (2009): 1451-63.

Kuntner, Matjaz, Matjaz Gregoric, Shichang Zhang, Simona Kralj-Fiser, and Daiqin Li. "Mating Plugs in Polyandrous Giants: Which Sex Produces Them, When, How and Why?" [In English]. *PloS one* 7, no. 7 (2012): e40939.

Ledon-Rettig, C. C., E. E. Zattara, and A. P. Moczek. "Asymmetric Interactions between Doublesex and Tissue- and Sex-Specific Target Genes Mediate Sexual Dimorphism in Beetles." [In English]. *Nature communications* 8 (2017): 14593.

Lees, D. R., and E. R. Creed. "Industrial Melanism in Biston-Betularia - Role of Selective Predation." [In English]. *Journal of Animal Ecology* 44, no. 1 (1975): 67-&.

Losos, Jonathan B. *Lizards in an Evolutionary Tree : Ecology and Adaptive Radiation of Anoles*. Organisms and Environments ;. Berkeley: University of California Press, 2009.

Lynch, Vincent J., and Gunter P. Wagner. "Resurrecting the Role of Transcription Factor Change in Developmental Evolution." [In English]. *Evolution; international journal of organic evolution* 62, no. 9 (2008): 2131-54.

Mahfooz, Najmus, Nataliya Turchyn, Michelle Mihajlovic, Steven Hrycaj, and Aleksandar Popadic. "Ubx Regulates Differential Enlargement and Diversification of Insect Hind Legs." [In English]. *PloS one* 2, no. 9 (2007): e866.

Mallarino, Ricardo, Corneliu Henegar, Mercedes Mirasierra, Marie Manceau, Carsten Schradin, Mario Vallejo, Slobodan Beronja, Gregory S. Barsh, and Hopi E. Hoekstra. "Developmental Mechanisms of Stripe Patterns in Rodents." [In English]. *Nature* 539, no. 7630 (2016): 518-23.

Manceau, Marie, Vera S. Domingues, Ricardo Mallarino, and Hopi E. Hoekstra. "The Developmental Role of Agouti in Color Pattern Evolution." [In English]. *Science (New York, N Y)* 331, no. 6020 (2011): 1062-5.

Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka*, et al.* "Genome Sequencing in Microfabricated High-Density Picolitre Reactors." [In English]. *Nature* 437, no. 7057 (Sep 15 2005): 376-80.

McGinnis, W., R. L. Garber, J. Wirz, A. Kuroiwa, and W. J. Gehring. "A Homologous Protein-Coding Sequence in Drosophila Homeotic Genes and Its Conservation in Other Metazoans." [In English]. *Cell* 37, no. 2 (1984): 403-8.

McGinnis, W., and R. Krumlauf. "Homeobox Genes and Axial Patterning." [In English]. *Cell* 68, no. 2 (1992): 283-302.

Merabet, Samir, and Bruno Hudry. "Hox Transcriptional Specificity Despite a Single Class of Cofactors: Are Flexible Interaction Modes the Key? Plasticity in Hox/Pbc Interaction Modes as a Common Molecular Strategy for Shaping Hox Transcriptional Activities." [In English]. *BioEssays : news and reviews in molecular, cellular and developmental biology* 35, no. 2 (2013): 88-92.

Monteiro, A., and M. D. Gupta. "Identifying Coopted Networks and Causative Mutations in the Origin of Novel Complex Traits." [In English]. *Current topics in developmental biology* 119 (2016): 205-26.

Morrow, Edward H., and Goran Arnqvist. "Costly Traumatic Insemination and a Female Counter-Adaptation in Bed Bugs." [In English]. *Proceedings Biological sciences* 270, no. 1531 (2003): 2377-81.

Muschick, Moritz, Adrian Indermaur, and Walter Salzburger. "Convergent Evolution within an Adaptive Radiation of Cichlid Fishes." [In English]. *Current biology : CB* 22, no. 24 (2012): 2362-8.

Mysore, Keshava, Longhua Sun, Michael Tomchaney, Gwyneth Sullivan, Haley Adams, Andres S. Piscoya, David W. Severson, Zainulabeuddin Syed, and Molly Duman-Scheel. "Sirna-Mediated Silencing of Doublesex During Female Development of the Dengue Vector Mosquito Aedes Aegypti." [In English]. *PLoS neglected tropical diseases* 9, no. 11 (2015): e0004213.

Orr, H. Allen. "The Genetic Theory of Adaptation: A Brief History." [In English]. *Nature reviews Genetics* 6, no. 2 (2005): 119-27.

Orr, H. Allen. "Theories of Adaptation: What They Do and Don't Say." [In English]. *Genetica* 123, no. 1-2 (2005): 3-13.

Pang, Dachling, and Dominic N. P. Thompson. "Embryology and Bony Malformations of the Craniovertebral Junction." [In English]. *Child's nervous system : ChNS : official journal of the International Society for Pediatric Neurosurgery* 27, no. 4 (2011): 523-64.

Pavlopoulos, Anastasios, and Michael Akam. "Hox Gene Ultrabithorax Regulates Distinct Sets of Target Genes at Successive Stages of Drosophila Haltere Morphogenesis." [In English]. *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 7 (2011): 2855-60.

Pearson, Joseph C., Derek Lemons, and William McGinnis. "Modulating Hox Gene Functions During Animal Body Patterning." [In English]. *Nature reviews Genetics* 6, no. 12 (2005): 893-904.

Perrichot, V., A. Nel, and D. Neraudeau. "Gerromorphan Bugs in Early Cretaceous French Amber (Insecta : Heteroptera): First Representatives of Gerridae and Their Phylogenetic and Palaeoecological Implications." [In English]. *Cretaceous Research* 26, no. 5 (Oct 2005): 793-800.

Plasterk, Ronald H. A. "Rna Silencing: The Genome's Immune System." [In English]. *Science (New York, N Y)* 296, no. 5571 (2002): 1263-5.

Rastogi, Shruti, and David A. Liberles. "Subfunctionalization of Duplicated Genes as a Transition State to Neofunctionalization." [In English]. *BMC evolutionary biology* 5 (2005): 28.

Refki, P. N., and A. Khila. "Key Patterning Genes Contribute to Leg Elongation in Water Striders." [In English]. *Evodevo* 6 (Apr 28 2015).

Reinhardt, Klaus, Richard Naylor, and Michael T. Siva-Jothy. "Reducing a Cost of Traumatic Insemination: Female Bedbugs Evolve a Unique Organ." [In English]. *Proceedings Biological sciences* 270, no. 1531 (2003): 2371-5.

Rubenstein, D. I. "Sperm Competition in the Water Strider, Gerris-Remigis." [In English]. *Animal Behaviour* 38 (Oct 1989): 631-36.

Schluter, Dolph. *The Ecology of Adaptive Radiation*. Oxford: Oxford University Press, 2000.

Scott, M. P., and A. J. Weiner. "Structural Relationships among Genes That Control Development: Sequence Homology between the Antennapedia, Ultrabithorax, and Fushi Tarazu Loci of Drosophila." [In English]. *Proceedings of the National Academy of Sciences of the United States of America* 81, no. 13 (1984): 4115-9.

Sekido, Ryohei, and Robin Lovell-Badge. "Sex Determination Involves Synergistic Action of Sry and Sf1 on a Specific Sox9 Enhancer." [In English]. *Nature* 453, no. 7197 (2008): 930-4.

Sharma, R. P., and V. L. Chopra. "Innate Metabolic Differences and Mutagen Sensitivity of Drosophila-Melanogaster .1. Radiosensitivity of Hyperkinetic Mutant." [In English]. *Mutation Research* 30, no. 1 (1975): 77-82.

Shirai, Leila T., Suzanne V. Saenko, Roberto A. Keller, Maria A. Jeronimo, Paul M. Brakefield, Henri Descimon, Niklas Wahlberg, and Patricia Beldade. "Evolutionary History of the Recruitment of Conserved Developmental Genes in Association to the Formation and Diversification of a Novel Trait." [In English]. *BMC evolutionary biology* 12 (2012): 21.

Sih, A., M. Lauer, and J. J. Krupa. "Path Analysis and the Relative Importance of Male-Female Conflict, Female Choice and Male-Male Competition in Water Striders." [In English]. *Animal Behaviour* 63 (Jun 2002): 1079-89.

Sniegowski, P. D., P. J. Gerrish, T. Johnson, and A. Shaver. "The Evolution of Mutation Rates: Separating Causes from Consequences." [In English]. *BioEssays : news and reviews in molecular, cellular and developmental biology* 22, no. 12 (2000): 1057-66.

Tanaka, Kohtaro, Olga Barmina, Laura E. Sanders, Michelle N. Arbeitman, and Artyom Kopp. "Evolution of Sex-Specific Traits through Changes in Hox-Dependent Doublesex Expression." [In English]. *PLoS biology* 9, no. 8 (2011): e1001131.

Tatarnic, Nikolai J., and G. Cassis. "Sexual Coevolution in the Traumatically Inseminating Plant Bug Genus Coridromius." [In English]. *Journal of evolutionary biology* 23, no. 6 (2010): 1321-6.

True, John R., and Sean B. Carroll. "Gene Co-Option in Physiological and Morphological Evolution." [In English]. *Annual review of cell and developmental biology* 18 (2002): 53-80.

Wallbank, Richard W. R., Simon W. Baxter, Carolina Pardo-Diaz, Joseph J. Hanly, Simon H. Martin, James Mallet, Kanchon K. Dasmahapatra*, et al.* "Evolutionary Novelty in a Butterfly Wing Pattern through Enhancer Shuffling." [In English]. *PLoS biology* 14, no. 1 (2016): e1002353.

Weigensberg, I., and D. J. Fairbairn. "The Sexual Arms Race and Phenotypic Correlates of Mating Success in the Waterstrider, Aquarius Remigis (Hemiptera: Gerridae)." [In English]. *Journal of Insect Behavior* 9, no. 2 (Mar 1996): 307-19.

Wigby, Stuart, and Tracey Chapman. "Sex Peptide Causes Mating Costs in Female Drosophila Melanogaster." [In English]. *Current biology : CB* 15, no. 4 (2005): 316-21.

Williams, Thomas M., and Sean B. Carroll. "Genetic and Molecular Insights into the Development and Evolution of Sexual Dimorphism." [In English]. *Nature reviews Genetics* 10, no. 11 (2009): 797-804.

Zarkower, D., and J. Hodgkin. "Molecular Analysis of the C. Elegans Sex-Determining Gene Tra-1: A Gene Encoding Two Zinc Finger Proteins." [In English]. *Cell* 70, no. 2 (1992): 237-49.