



Towards an adaptive and genomic understanding of an exaggerated secondary sexual trait in water striders

William Toubiana

► To cite this version:

William Toubiana. Towards an adaptive and genomic understanding of an exaggerated secondary sexual trait in water striders. Biodiversity. Université de Lyon, 2019. English. NNT : 2019LYSEN058 . tel-02360159

HAL Id: tel-02360159

<https://theses.hal.science/tel-02360159>

Submitted on 12 Nov 2019

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

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



Numéro National de Thèse : 2019LYSEN058

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 : Sciences de la vie

Soutenue publiquement le 20/09/2019, par :
William Toubiana

**Towards an adaptive and genomic
understanding of an exaggerated
secondary sexual trait in water striders**

**Vers une compréhension adaptative et génomique d'un trait sexuel secondaire chez une
espèce d'insecte semi-aquatique.**

Devant le jury composé de :

| | | | |
|-----------------------|--------------|-----------------------|--------------------|
| Dr. Leulier François | Group leader | ENS de Lyon | Examineur |
| Dr. Immonen Elina | Group leader | University of Uppsala | Rapporteuse |
| Dr. Perry Jennifer | Group leader | University of Oxford | Rapporteuse |
| Dr. Lüpold Stefan | Group leader | University of Zurich | Examineur |
| Dr. Khila Abderrahman | Group leader | ENS de Lyon | Directeur de thèse |

Table of contents

| | |
|---|----|
| Abstract | 6 |
| Résumé | 7 |
| Introduction | 9 |
| I. The history of evolutionary thought | 9 |
| II. The process of selection in evolution | 12 |
| A. Natural selection | 12 |
| B. Genetic drift | 13 |
| III. The theory of sexual selection | 13 |
| IV. Fitness | 16 |
| V. Sexual dimorphism | 17 |
| VI. Trait exaggeration | 18 |
| A. Trait exaggeration and sexual selection | 18 |
| B. Other types of trait exaggeration | 19 |
| VII. Evo-Devo of exaggerated sexually-selected traits | 21 |
| A. Growth pathways and their role in variable and non-variable exaggerations | 22 |
| B. The role of patterning genes in the development of exaggerated, sexually-selected traits | 23 |
| VIII. Towards a more genomic understanding of trait exaggeration | 23 |
| IX. Intra-species variation in sexually selected traits | 25 |
| X. The semiaquatic insects (Gerromorpha) as a model to study the role of natural and sexual selections on trait evolution | 26 |
| Chapter 1: Fluctuating selection strength and intense male competition underlie variation and exaggeration of a water strider's male weapon | 31 |
| I. Abstract: | 32 |
| II. Introduction | 33 |
| III. Material & methods | 34 |
| A. Population sampling and culture | 34 |
| B. Measurement of Microvelia species and statistics | 34 |
| C. Behavioral observations and video acquisition | 35 |
| D. Microvelia phylogenetic reconstruction | 35 |
| E. Male competition assay | 35 |
| F. Fight frequency assay | 35 |
| G. Artificial selection experiment | 36 |
| H. Condition-dependence experiment | 36 |
| I. Microsatellite development | 36 |
| J. Paternity test | 37 |
| IV. Results and discussion | 37 |

| | | |
|--|--|-----|
| A. | Sexual Dimorphism and scaling relationships in <i>Microvelia</i> species | 37 |
| B. | Mating systems in <i>Microvelia</i> species. | 39 |
| C. | Intensity of male competition in <i>M. longipes</i> compared to <i>M. pulchella</i> | 42 |
| D. | Effect of exaggerated leg length on male reproductive fitness in <i>M. longipes</i> | 43 |
| E. | Environmental and genetic contributions to male rear leg variation..... | 46 |
| V. | Conclusions | 49 |
| VI. | Supplementary material..... | 51 |
| Chapter 2: The genome of the water strider <i>Microvelia longipes</i> provides insights into the role of sexual selection on gene expression and genome architecture. | | 62 |
| I. | Abstract | 63 |
| II. | Introduction | 64 |
| III. | Material and methods | 66 |
| A. | Population sampling and culture | 66 |
| B. | Measurement and statistics | 66 |
| C. | Sample collection, assembly and annotation of the <i>M. longipes</i> genome | 66 |
| D. | Sample collection and preparation RNA-sequencing | 67 |
| E. | Transcriptome assembly, mapping and normalization | 67 |
| F. | Comparative transcriptomics: analyses of variance | 68 |
| G. | Sex-biased and leg-biased expression..... | 68 |
| H. | Gene Ontology analysis..... | 70 |
| I. | Sex-biased gene distribution | 70 |
| J. | Detection of large sex-biased gene clusters | 71 |
| K. | Detection of clusters of consecutive sex-biased genes | 72 |
| IV. | Results | 73 |
| A. | Assembly and annotation of the <i>M. longipes</i> genome | 73 |
| B. | Male rear-legs experience a burst of growth during the last nymphal instars. | 73 |
| C. | Variation in gene expression explains differences in exaggeration between legs and sexes. | 75 |
| D. | Male-biased genes show unique features of trait exaggeration | 77 |
| E. | Divergence in male leg homology is associated with a set of male-specific leg-biased genes | 78 |
| F. | Male exaggerated legs are enriched in genes with leg- and sex-biased expressions. | 81 |
| G. | The influence of sexual selection on <i>M. longipes</i> genome architecture | 83 |
| V. | Conclusion and discussion..... | 86 |
| VI. | Supplementary material..... | 90 |
| Summary and general conclusions | | 108 |
| General discussion and future directions | | 111 |
| I. | The influence of selection on sexual trait evolution..... | 111 |
| A. | Moving from reproductive fitness to a more integrative view of fitness | 111 |
| B. | How does variation in sexual selection strength influence the evolution of trait exaggeration? | 112 |
| II. | Sexual selection and the evolution of gene expression | 114 |

| | |
|--|-----|
| III. Understanding the genetics underpinning of leg exaggeration through genetic mapping. | 115 |
| Summary of other scientific contributions | 117 |
| I. Fitness Effects of Cis-Regulatory Variants in the <i>Saccharomyces cerevisiae</i> TDH3 Promoter..... | 117 |
| II. Predator strike shapes antipredator phenotype through new genetic interactions in water striders. | 118 |
| III. The <i>mlpt/Ubr3/Svb</i> module comprises an ancient developmental switch for embryonic patterning. ... | 119 |
| IV. The genome of the water strider <i>Gerris buenoi</i> reveals expansions of gene repertoires associated with adaptations to life on the water. | 120 |
| Acknowledgements | 121 |
| References..... | 122 |

Abstract

From the DNA molecule to the more complex phenotypes, variation is a universal process in life and living organisms. The innumerable differences that exist between species are probably one of the most manifest examples. Yet, all this diversity would never have occurred in nature without some pre-existing divergence within species. One of the most striking examples of intraspecies variation appears in sexual organisms, between males and females. Understanding the environmental and genetic factors influencing sexual divergence is a longstanding question in evolutionary biology. To this end, I focus here on a new insect model system, *Microvelia longipes*, which has the particularity to have evolved an extreme case of sexual dimorphism in the rear legs. Males display exaggerated long rear legs compared to females but also an extreme variability in these leg lengths from one male to another. We identified that *M. longipes* males use their exaggerated legs as weapons during male-male competition. Males with longer legs have more chance to access females on egg-laying sites and therefore increase their reproductive success. Moreover, fitness assays and comparative studies between *Microvelia* species revealed that the intensity of male competition was associated with the exaggeration and hypervariability of the rear legs in *M. longipes* males. In a second approach, we studied the developmental and genomic basis of this sexual dimorphism through a comparative transcriptomic analysis and identified genes and genomic regions associated with male exaggerated legs and ultimately with sexual selection. Overall, the integrative approach used in this work allows to establish *Microvelia longipes* as a promising new model system to study the influence of sexual selection in adaptive evolution.

Résumé

De la simple molécule d'ADN aux caractères les plus complexes, nous pouvons observer que le processus de variation est un phénomène universel en biologie. Les nombreuses différences qui existent entre les organismes peuplant notre planète en est certainement l'exemple le plus manifeste. Ces différences entre espèces n'auraient cependant pas pu avoir lieu au cours de l'évolution si certaines d'entre elles n'étaient pas déjà présentes au sein même d'une espèce. Ces variations intra-espèces sont particulièrement observables chez les espèces sexuées, entre mâles et femelles. Comprendre les différents facteurs biologiques, environnementaux et génétiques, à l'origine de ce dimorphisme sexuel est le cœur de mon sujet de thèse. Pour cela, j'ai établi au cours de ma thèse un nouveau modèle d'étude, l'insecte semi-aquatique *Microvelia longipes*. Ces insectes ont la particularité d'avoir évolué un dimorphisme sexuel spectaculaire par lequel les mâles présentent une croissance extrême et spécifique au niveau de la troisième paire de pattes. Pour étudier ce phénomène, j'ai choisi une approche intégrative visant à comprendre les causes environnementales et génétiques liées à la croissance exagérée des pattes arrières chez les mâles. En premier lieu, nous avons émis l'hypothèse que cette croissance exagérée était associée à des pressions de sélection sexuelle. Une caractérisation plus détaillée de la taille des pattes a également montré un degré de variabilité extrême entre les mâles d'une même population. Nous avons mis en évidence la présence de compétition intense entre mâles, qui utilisent leurs pattes arrière comme arme, pour s'accoupler avec les femelles. Les mâles à pattes plus longues gagnent souvent dans ces combats, expliquant l'importance adaptative de ces pattes exagérées chez les mâles. De plus, nous montrons que les variations de taille de pattes chez les mâles, de la même espèce ou d'espèces différentes, sont associées à une fluctuation de la sélection sexuelle, qui peut être plus ou moins intense en fonction de l'ardeur que mettent les mâles à se battre. Dans un deuxième temps, nous avons développé une approche transcriptomique comparative entre les sexes et les pattes afin d'identifier les gènes responsables de cette croissance exagérée. Combiné au développement d'un génome de *M. longipes* de haute qualité, ceci a permis de dresser une liste de gènes dont l'intensité d'expression corrèle de manière significative avec l'exagération de la croissance des pattes chez les mâles et d'identifier des régions génomiques soumises à la sélection sexuelle.

Introduction

I. The history of evolutionary thought

What is the place of human beings in the universe?

How did we come up living with such diversity of life?

Are we that different from other living organisms?

Those philosophical questions have crossed people's mind at least once. They have shaped our society for centuries and became the foundations of extant religions. However the belief in a certain divinity, that no one can sense or even agree on, was for the most sceptical ones a proof of its irrationality. The idea of a certain form of evolution, where species change over time, dates from the Antiquity but received a scientific attention only between the 18th and 19th century with the emergence of palaeontology [1]. The comparison of extinct and existing forms of life brings one of the first scientific evidences of a dynamic process in nature, where species go extinct and emerge over time [1]. The attestation to species extinction brought notably the field of biology into a new area, the evolutionary history of life.

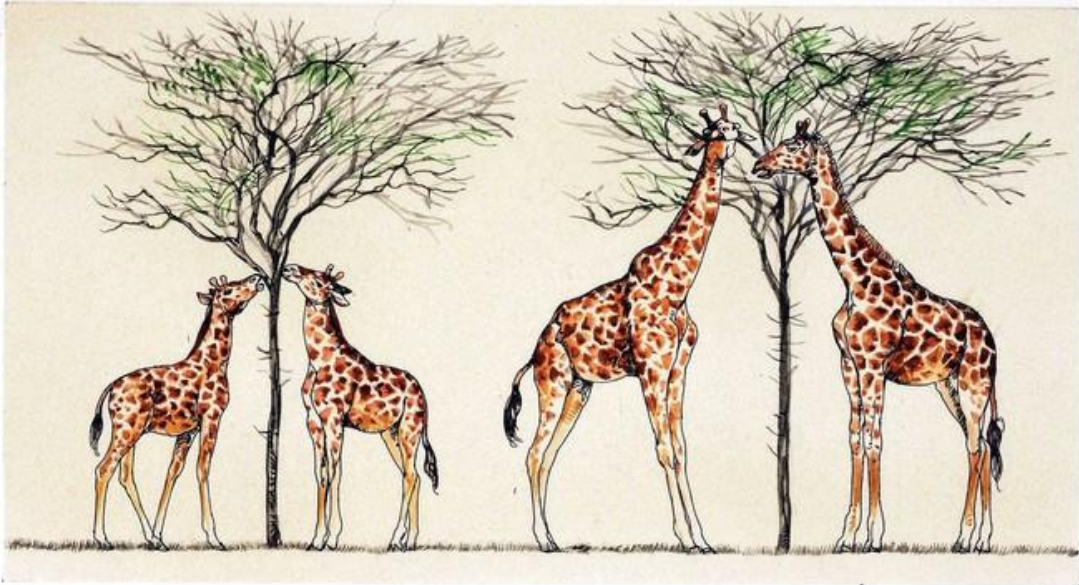
Jean-Baptiste Lamarck was the first biologist to formulate a fully coherent theory on biological evolution. In the early 19th century, Lamarck hypothesized that the new forms of life were continuously generated through modifications acquired by the individuals during their lifetime. These modifications would then be transmitted to the next generation that would in turn acquire new modifications, or not, according to their environment [2]. He named his theory *Transformisme* in regards to the two main forces acting on organismal evolution; one evolutionary force driving organisms from less complex to more complex forms, and a second force driving the diversification and adaptation of organisms to their local environments. A famous example illustrating his theory is the evolution of giraffe's neck. Lamarck believed that giraffes were stretching their neck in order to access food on high trees and that this neck elongation would be inherited at the next generation (Figure 1).

At this period, Lamarck's thought on the evolution of species contrasted with a new theory, published by Charles Darwin and Alfred Russel Wallace in 1858 [3, 4]. This theory, that will later become a reference to the science of evolution, is fully formulated in 1859 by Charles Darwin in his book *On the Origin of Species* [4]. Lamarck and Darwin's theories share a lot in common, especially regarding the concept of evolution *per se*. Both agree, for example, on

the concept of common ancestry where all existing species are linked by an ancestor that was a rather simpler form that complexified during the course of evolution [2, 4]. Their main discordance relied on the process by which species evolve. Instead of believing in the theory of transformation, Darwin argued that the variation observed between individuals of a same species was the source of species diversity through millions of years of variation accumulation (Figure 1). Under this scenario, human beings are notably considered as evolved as any other species. The human race is simply perceived as a species that has taken a certain evolutionary path that is different from the other existing species. The place of human beings in his environment was further detailed by Charles Darwin thirteen years later in his book *The Descent of Man, and Selection in Relation to Sex* [5].

Darwin acknowledged, nonetheless, that for his theory of evolution to hold, the factors generating phenotypic variation must be inherited to the next generation. The question of heritability became consequently one of the main requirements to validate his theory. It first started with Mendel's law of inheritance in 1866, demonstrating that some "invisible" factors were able to predictably determine the acquisition of some traits at the next generation in plants [6]. For some time ignored and criticized, Mendel's genetics will be revived in the early 20th century from a series of works from Walter Sutton, Theodor Boveri, Thomas Hunt Morgan and Ronald Aylmer Fisher. Walter Sutton and Theodor Boveri first developed in 1902 what will be called the "chromosomal theory of heredity" where they propose that the chromosomes are the heredity factors predicted in Mendel's laws of inheritance [7]. This theory was definitely approved after Morgan's work on *Drosophila* mutants that also led to the ideas of genetic linkage, crossing over and to the establishment of the first genetic map by Morgan's student, Alfred Sturtevant, in 1913 [8, 9]. Later, Ronald Fisher mathematically reconciled Mendel's model of inheritance

A



B

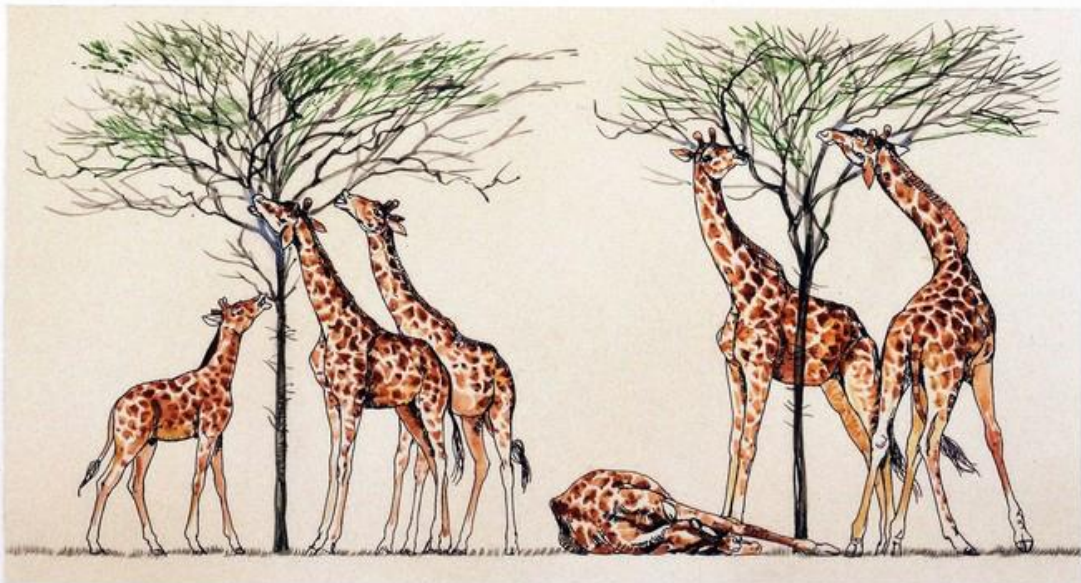


Figure 1: Representation of Lamarck and Darwin's theories of evolution
 The top panel illustrates Lamarck's idea by which two giraffes that cannot access food will stretch their neck over their lifetime to adapt to the environment. The bottom panel illustrates Darwin's idea where giraffes with different neck lengths occur in the population but only the ones with the longest neck survive.

that was working on discrete phenotypic variations with the evolution of more continuous phenotypes [10]. According to his mathematical predictions, the combined actions of many heritable factors could recapitulate the continuous variations observed in some traits. His work sets the basis of what we will later call the field of population genetics and led to the synthesis of Darwin's theory and Mendel's laws of inheritance, also named modern synthesis by Julian Huxley in 1942 [11].

Despite the accumulating evidence regarding the chromosomal support of heritability, this field of research was rather conceptual as it lacked strong molecular approval. It is only in the mid 20th century that James D. Watson and Francis Crick provided the scientific proof and answer to the mechanism of heredity with the discovery of the DNA double helix structure and its nucleotide composition [12]. The later founding finally brought the field of evolutionary biology to its most recent aspects, associating variations in phenotypes to variations in DNA sequences to ultimately trace back the evolution of phenotypes and nucleotide sequences [13].

II. The process of selection in evolution

A. Natural selection

Evolution is an endless process in which changes in DNA sequence, e.g. mutations, inexorably accumulate over time and increase overall phenotypic diversity. Yet phenotypic diversity does not seem to escalate as time goes. Instead, ecologists and palaeontologists revealed that the biodiversity varies according to the period of the history. At the early discovery of fossils, palaeontologists soon noticed that some of these ancient living forms do not resemble the existing species [14]. The evidence of a certain succession in the forms of life prompted naturalists to ask why some species get maintained while some others go extinct. In 1858, Alfred Wallace and Charles Darwin suggested that the natural environment in which species live could actively select for or against some of the diversity that occurs in natural populations over time (Figure 1). As a consequence, the variants conferring more survival advantage to the individuals would tend to be more present in the population at the next generation than the more deleterious ones. They called this active force of the environment Natural Selection [3, 4]. Although the theory of evolution by common ancestry was rapidly accepted, the mechanisms by which Nature would select and shape evolutionary

processes were highly debated until the emergence of the field of population genetics. Ronald Fisher in his book *The Genetical Theory of Natural Selection*, became a pioneer to use mathematics to conciliate Mendelian genetics with the principle of natural selection [10]. This contributed to the revival of Darwin's theory in the early 20th century that was revised as the modern synthesis [11].

B. Genetic drift

Following the establishment of the modern synthesis and the discovery of the DNA structure in 1953, the field of evolutionary biology integrated a new discipline, namely the molecular evolutionary genetics [15]. It describes the process by which molecules such as DNA, RNA and proteins can change over generations. Applied to studies of population genetics, it led to a new evolutionary theory named the neutral theory of molecular evolution. This theory, mostly introduced by Motoo Kimura, argues that most molecular changes confer no advantage or disfavour at the individual level and therefore evolve neutrally, without any environmental pressures [16]. Evolutionarily speaking, it implies that most genetic variation present in a population results from mutations that arise and increase in frequency by chance, due to their neutral effect on individuals that carry them. In a population, the fluctuation in allele frequency is called genetic drift and is a strong driver of evolution as it can randomly select and get to fixation some alleles present in the population. In some occasions, these neutral changes can even lead to changes in phenotypes [17].

III. The theory of sexual selection

“The sight of a feather in a peacock's tail, whenever I gaze at it, makes me sick!” (*Letter 2743 – Darwin, C. R. to Gray, Asa, 3 April (1860), Darwin Correspondence Project*).

Even after publishing his theory of evolution by means of natural selection, Charles Darwin was not convinced that selection through differential in survival could explain all the surrounded phenotypic diversity. For some examples, such as the peacock's tail, he was even thinking that natural selection should have acted against such phenotypes that look detrimental for the survival of the individual bearing it. For these traits that were challenging his theory of natural selection, Darwin imagined that they could be used for a certain function

that would overcome their viability costs. He reasoned that in sexually reproducing animals, one major goal for an individual is to reproduce in order to transmit his characters to the next generation. Heritable characters that would improve the ability of an individual to mate with the opposite sex would therefore be selected and spread in the population over time. This form of sexual selection would here favor the characters increasing the mating success of an individual. Under this assumption, both males and females should evolve elaborated structures to increase their number of partners. Yet, in nature it is generally the male that is the most modified. Darwin argues that “this seems to lie in the males of almost all animals having stronger passions than the females” [5]. The fact that males will be more eager to mate than females, will lead males to evolve features such as weapons to better compete with other opponent of the same sex or acquired more sedulous characters such as ornaments, that will better charm the females (Figure 2). Thus within the sexual selection process, Charles Darwin distinguishes the mechanism of male competition where males will fight between each other to access the females, from the mechanism of female choice where the female will actively choose the male she wants to mate with in the population, based on some attributes (Figure 2). Although these two mechanisms can act independently in some species, they are often linked as female choice usually implies competition in the other sex, either directly (e.g. male fights) or indirectly (e.g. hurrying the finding and seduction of the female before another male) [5]. The fact that males compete between each other to access females was generally acknowledged by other evolutionary biologists to be an important process in evolution. However, the idea that female choice could drive evolution was hardly acceptable for the evolutionary community [18]. Once again, this theory had to wait for the work of Ronald Fisher to renew under the mechanism of positive feedback by which the female preference for a certain male trait and the male trait elaboration will get heritably associated to finally advance together. Fisher called this scenario *runaway sexual selection* to illustrate that the female preference for a certain elaborated structure in the male trait will lead to the evolution of a more elaborated trait in males that will in turn affect the female preference toward this more elaborated structure.



Figure 2: Examples of male weapons and ornaments. Top picture shows the large antlers of male bull moose engaged in a fight for territory dominance. Bottom picture illustrates the deployment of a male peacock tail to seduce a female.

This endless process is thought to cause the extreme male elaborations observed in nature and can only be stopped when counter selective pressures (e.g. natural selection for survival) will act against the increase in male trait elaboration or female preference [10, 19].

IV. Fitness

The term fitness is central in the field of evolutionary biology. It represents a concept to describe “the ability of organisms to survive and reproduce in the environment in which they find themselves” [20]. In other words, the term fitness predicts the ability of an individual to transmit its genes to the next generation. During its lifetime, an individual faces various obstacles that will directly impact its fitness. After the zygote is produced, the later needs to survive until adulthood and then reproduce to generate a new progeny. This whole process can be disentangled into three major components, namely survival, mating and fertility. Combined, these three fitness components form the total fitness of an individual [20]. Any variations in the viability, mating success or fecundity between different individuals of a population will directly affect their total fitness and therefore be selected for or against by the environment in which they live. As a consequence, fitness can be considered as the ultimate phenotype on which natural selection can act. Although this definition of fitness encompasses the major components of an individual’s life, it cannot be universal as it can vary from species to species according to their life history and evolutionary strategies. For example, some fitness components can be inexistent in some organisms such as asexual organisms that have no mating success. Some fitness components also cannot be clearly comparable between the different taxa. The fitness of survival, for example, can be subdivided into distinct developmental stages that are hardly comparable between insects and mammals. Even between insects, hemimetabolous and holometabolous insects develop differently and therefore may be considered as having different components of survival fitness [21]. Another way to think about fitness has therefore been suggested and involves a more mathematical and probabilistic definition. This definition has been mainly developed by population geneticists that see fitness as a probabilistic value that monitors the frequency of a certain allele or genotype in the population over one to many generations [22]. In large populations, the effect of drift is minimized and the fate of an allele is mainly determined by how its resulting phenotype fits to the environment [23]. Mutations that confer the fittest phenotypes, which are those that are best adapted to a given environment, will be selected and increase in frequency

in the population. By comparing the frequency and the becoming of several alleles or genotypes segregating in a population, evolutionary geneticists can estimate a coefficient of selection that gives the relative “advantage” of one allele or genotype over one or several other alleles or genotypes [20].

In the following, I will focus mainly on the concept of fitness as the process describing the interaction between the three major components of an individual’s reproductive success; namely survival, reproduction and fecundity. The mathematical definition of fitness is nonetheless of relevant consideration for the future directions of my research project and will be evoked later as a perspective.

V. Sexual dimorphism

The theory of sexual selection by Charles Darwin was partly able to explain the evolution of male and female differences as it sheds light on the origin of secondary sexual trait evolution and diversification by male-male competition or female choice. However, it hardly explains why evolution has favored males and females to differ in their mating strategies. Most secondary sexual characters have evolved in males, although some exceptions exist [19, 24, 25]. Females usually retain more juvenile or ancestral characters but tend to be more choosy for the mating partner than males do, even though exceptions exist here too [5, 19, 26]. In 1948, Angus John Bateman published what will become the reference study to explain the evolution of sexual dimorphism. In this study, he placed an equal number of males and females fruit flies *Drosophila melanogaster* in bottles, and showed that females’ fertility was limited by the number of eggs they produce whereas males’ fertility was mainly restricted by the number of females they reproduce with [27]. This study sheds light on the different strategies favored by each sex and that will be later related to their asymmetric investment in gametes. Males are capable of producing sperm in excess compared to female eggs but do not have the guarantee to fertilize all the eggs produced by the mated female. Therefore, the number of offspring sired by a male is dependent on the number of females he mates with. In contrast, females have a limited number of eggs to fertilize, making them choosier towards the mating partners [19]. Later, the discovery of sperm competition between males reinforced the idea that males suffer more to secure fertilization than females do, causing males to invest into characters that improve their mating and fertilization success. On the other hand, it has

been shown that above a certain number, mating becomes detrimental for females fitness. These differences in sex roles lead to a conflict between the two sexes where males try to increase their number of mating to increase their fitness, whereas females require fewer numbers of mates to achieve their optimal reproductive fitness. This process called sexual conflict over mating rate can favor the evolution of some traits in one sex that could be harmful to the other sex. Such antagonistic co-evolution between the two sexes can drive the two sexes away from their fitness optima and engender sexual dimorphism as a resolution of the conflict [28-30].

VI. Trait exaggeration

A. Trait exaggeration and sexual selection

Extravagant or exaggerated phenotypes have fascinated naturalists for decades and have become one of the main foci in the field of evolutionary ecology [31-34]. Some of these phenotypes can reach degrees of expression so high that the individuals bearing them become almost unrecognizable as members of the same species [5, 19].

Charles Darwin was one of the first to put the concept of trait exaggeration in an evolutionary framework by explaining how such phenotypes can emerge under sexual selection [5]. Since Darwin's description, exaggerated traits have remained associated with two main components: a comparative reference allowing the trait to be defined as exaggerated *sensu stricto* and an evolutionary component explaining the origin and maintenance of such exaggeration [19, 35]. In the context of sexual selection, exaggeration describes traits that look extravagant in one sex compared to the other and that present seemingly inevitable costs, in terms of survival, to the bearer of the exaggerated trait [5, 19]. These survival costs are theorized to be largely balanced by opposing benefits due to increased mating success [19, 28, 36, 37], thus allowing the trait to evolve and maintain exaggerated degrees of expression (Figure 3A). In this case, trait exaggeration is a reflection of the imbalance, in terms of costs and benefits, between survivorship and mating success, which represent two of the three main components of fitness (Figure 3) [20]. Although survival costs and mating benefits are seldom clear, the cost-based definition of trait exaggeration is widely used in the field of sexual selection [5, 19, 38, 39], (but see [40]). Exaggerated sexually-selected traits can represent modifications in any kind of phenotype (Figure 4), although most famous examples are on structural exaggerations, including horns and mandibles in a number of beetles, stalks in

stalk-eyed flies, the claw of fiddler crabs, the long tail of peacocks or even the deer antlers [32, 41-46]. Exaggerations affecting behaviour and coloration are also quite widespread [19]. In many spiders, for example, males ‘offer’ themselves as food to the female immediately after mating [37, 47, 48]. This suicide behaviour represents an extreme and obvious case of exaggeration, as it ends the male’s survivorship on the spot. However, this male suicide behaviour increases the chances that the egg clutch, laid by the female (after having consumed the male), is fertilised by his sperm. This provides a hypothesis to explain the evolution and maintenance of such an extreme case of exaggeration [37, 47-49].

Bird nuptial parades are also famous examples of behavioural exaggeration, especially in the bird of paradise where males of some species have highly ritualised mating dances to attract females. Other cases include bright coloration found in many arthropods, birds, and fishes [50-52] (Figure 4). These extreme colors are often used by males to attract females that will choose them over less colorful males. In guppies for example, males from the same population develop various color patterns along their body and the less frequent phenotypes are often favored by females that are colorless [53, 54]. This female choice for rare coloration patterns creates a frequency-dependent selection that favors the maintenance of phenotypic variation in the population. Male coloration is often associated with heightened predation risk [52, 55-57] but balanced by benefits due to increased mating success [19, 52, 57] (Figure 3A).

B. Other types of trait exaggeration

Ever since the description of exaggeration in the context of sexual selection [5], the concept has evolved to include additional cases of exaggeration that are thought to be under natural and social selection [35] (Figure 3B-C). The comparative reference in these cases of exaggeration can be defined as the ancestral state of the trait, the expression of the trait in another caste (in the case of social insects), or as the state of trait expression in a homologous counterpart along the body axis [35, 42, 58-60]. For example, in the reproductive caste of social insects, females, also known as queens, exhibit extreme ovarian activity at the expense of their ability to forage or escape predation, and therefore at the expense of survivorship [61-64] (Figure 3B). Extreme cases are known in termites and army ants [61, 64], where females are entirely dependent on the care provided by their nestmates. Such exaggerated degrees of fecundity can only evolve in a social context where other colony members are able to fulfill survivorship-related tasks (Figure 3B) [61]. The insect family Gerridae, also known as water striders, have significantly longer legs compared to member of their immediate sister group the Veliidae [65-69].

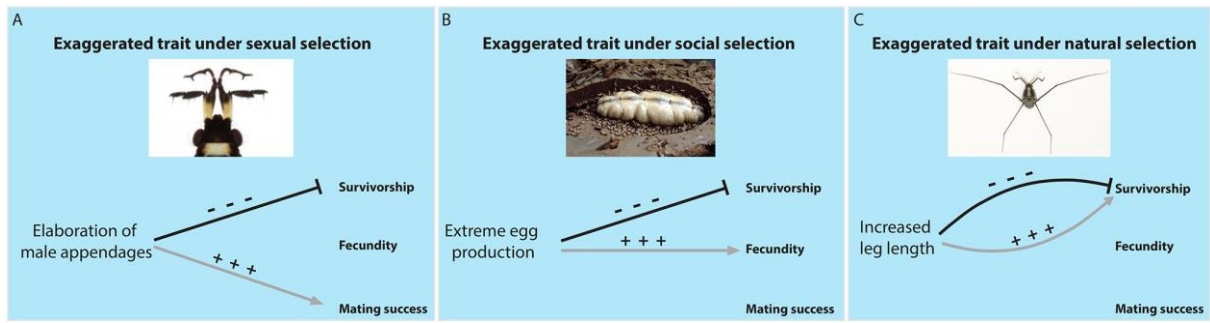


Figure 3. Effect of various types of trait exaggeration on fitness components. Usually, negative effects on one fitness components are compensated by gains from another component (A, B) or from a sub-section of the same component (C). Plus signs indicate benefits and minus signs costs. Images: (A) elaborated male antennae in the water strider *Rheumatobates riley* [36], (B) enlarged abdomen of termite queen (Photo credit: Judith Korb) [70], and (C) elongated legs of the water strider *Metrobates hesperius* [68].

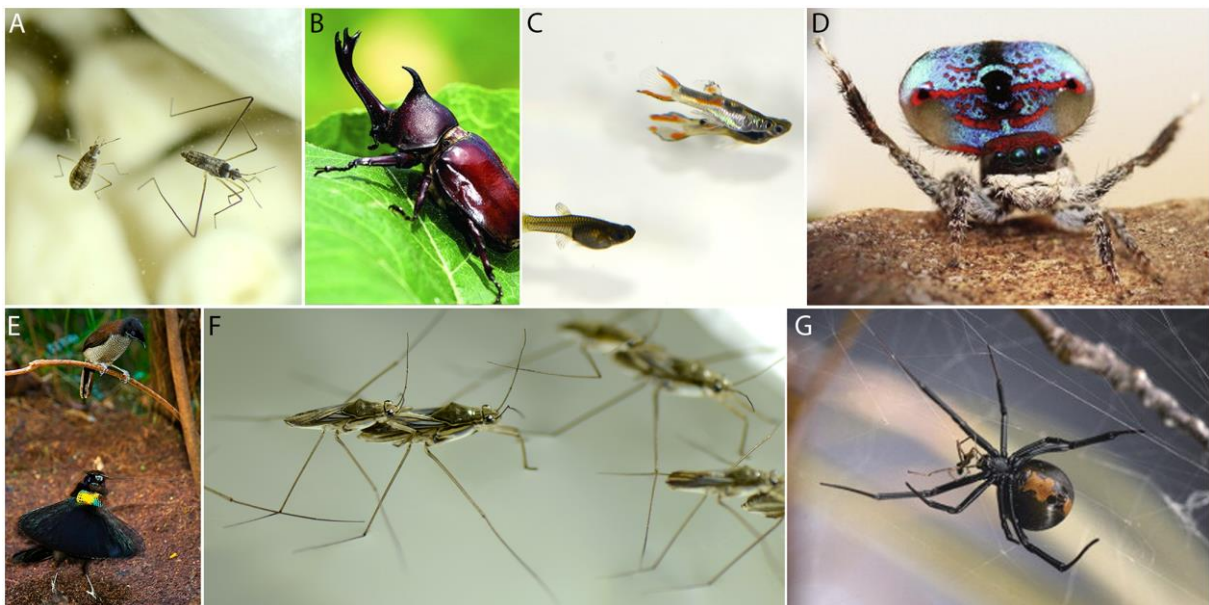


Figure 4. Examples of morphological, behavioural and physiological exaggerations. (a) Males of the pond skater *Microvelia longipes* exhibit exaggerated leg length relative to females. (b) Males of the Rhinoceros beetle *Trypoxylus dichotomus* develop exaggerated horns size (modified from Ref. [1] with permission from Science). (c) Male-specific color pattern in Guppies. (d, e) Male peacock spider (d, modified from Ref. [36] with permission from Current Biology) and bird of paradise (e, modified from Ref. [39], with permission from PNAS) exhibiting bright coloration and performing nuptial dance. (f) Guarding behaviour in males of the water strider *Limnogonus franciscanus*. (g) Male suicide behaviour in the red back spider.

Leg length in water striders can be considered a case of phylogenetic exaggeration. In this case, natural selection alone (except in cases where leg length is under sexual selection) appears to explain increased expression of this trait due to the requirement of locomotion on fluid surfaces (Figure 3C) [67, 71-73].

VII. Evo-Devo of exaggerated sexually-selected traits

The field of Evolutionary Developmental Biology (Evo-Devo) emerged with the development of molecular genetics and the acquisition of recombinant DNA technology in some model organisms [74, 75]. It consists in inferring evolutionary processes, such as ancestral relationship between different organisms, based on the comparison of their developmental processes [76]. The relevance of this developmental comparison was suggested when evolutionary and developmental biologists realized that the phenotypic variation observed within or between species involved features that are acquired during development [76]. With the emergence of molecular technologies, the field quickly started to ask questions about how developmental processes and genetic networks can evolve to give rise to phenotypic diversity. One of the main discoveries of Evo-Devo studies was to show that evolution in phenotypes was possible through the reuse of ancient, highly conserved genes that can change in time and space their expression during development. This has led notably to the concept of toolkit genes, describing the pleiotropic function of certain genes that are capable of forming a complex cascade of control, switching on and off various regulatory genes, according to their expression patterns.

Sexually selected exaggerated traits have also become one of the main foci in the modern field of Evolutionary Developmental Biology [36, 41, 44, 77, 78]. They offer an unequalled opportunity to understand not only the development of exaggeration but also the developmental processes underlying intra-species phenotypic variability [38, 39, 42, 79-81]. Yet this field of study is relatively recent and the bias of Evo-Devo towards the study of the developmental genetic mechanisms underlying morphological evolution directed research studies towards exaggeration of growth-related morphological traits. For this reason, I will introduce this new area of research by summarizing works on the mechanisms of growth and scaling relationships in sexually-selected exaggerated traits but I will also evoke the roles of

some key developmental genes in organ growth that have been overlooked in the context of sexual trait elaborations.

A. Growth pathways and their role in variable and non-variable exaggerations

Most recent studies of the developmental genetic mechanisms underlying trait exaggeration have focused on sexually-selected growth related traits that are hypervariable between individuals of the same sex and that are used as signals [35, 41, 42, 44, 60] [1, 8, 14, 28, 44]. This focus quickly steered the field towards analyzing various growth pathways [35, 42, 43, 60, 82-86], especially in horned and stag beetles. Several recent studies have identified pathways such as the Insulin-like or Juvenile Hormone (JH) pathways as being involved in the development and hyper-variability of morphological cases of sexually selected trait exaggeration [41, 87-89]. Increased tissue sensitivity to these pathways is argued to present a major way for growth-related exaggerated traits to develop and evolve. However the contribution of these pathways to exaggerated traits with low variability between individuals is still not clearly defined. Whether variable or not, tissue growth requires the action of growth pathways, such as Insulin-like signalling, and differences in growth rates across tissues are modulated through differences in their sensitivity to hormonal input [41, 86, 90]. In the rhinoceros beetle *Trypoxylus dichotomus* for example, males developed exaggerated head horns that are highly variables in males of the same population. The knock down of the Insulin receptor reveals that male horns are the organ that decreased the most in size, suggesting that they are hypersensitive to the Insulin pathway compared to other body parts such as wings or the genitalia. In the super-soldiers of the ant genus *Pheidole*, the development of exaggerated head size depends on the Juvenile Hormone pathway despite the low variation observed among individuals of this caste [91]. In this case, growth pathways would be important for exaggerated morphological trait regardless of their degree of variability.

Another highly conserved pathway for sex determination in insects was also shown to play a critical role in the development of sex-specific trait exaggeration. In some holometabolous insects like horned beetles, *doublesex* has been shown to encode for male and female specific isoforms that, when repressed, have opposite functions in the two sexes. In males, knock-down of *doublesex* generally decreases trait exaggeration whereas it enhances it in females. Overall, *doublesex* repression seems to decrease the degree of dimorphism between the two

sexes, suggesting that this gene represses female trait exaggeration whereas it activates it in males [92, 93]. A study in stag beetles has also shown that *doublesex* seems to interact with JH pathway in order to regulate mandible sensitivity to nutritive growth in a sex-specific manner [87].

B. The role of patterning genes in the development of exaggerated, sexually-selected traits

The role of patterning genes in the development of exaggerated traits that are subject to natural selection (such as the long legs of water striders or crickets) is now accepted in the Evo-Devo field [35, 66]. However their importance in the development and evolution of exaggerated traits that are under sexual selection remains unclear. Although rare, some studies have shown the importance of patterning genes in exaggerated sexually selected traits such as the modified antennae of the water strider *Rheumatobates rileyi* or the elongated tail of the swordtail fish *Xiphophorus helleri* [94]. Genes like *msxC*, which is a transcription factor involved in cell differentiation, have restricted temporal and spatial expression, in addition to being controlled by hormonal pathways, and may therefore contribute to the quantitative intra-sexual variation observed in tail length of swordtail fish [94]. Furthermore, a recent study implicates the Hedgehog signalling pathway in the polymorphic horn size in the beetle *Ontophagus taurus* [95]. Processes of growth and pattern formation are tightly linked and their relative contribution could depend on the function and/or the selective pressures acting on the trait. Exploring this question using different model systems and different kinds of traits may expand our understanding of the developmental mechanisms controlling trait exaggeration.

VIII. Towards a more genomic understanding of trait exaggeration

Although recent, evolutionary developmental studies on trait exaggeration are growing relatively fast, thanks notably to the large body of research already established on holometabolous morphological growth [82, 90, 96-101]. The primary focus on the genes and developmental pathways regulating sexually selected exaggerated traits led to overlook their evolution at the genomic scale. Yet, the genomic basis of sexual traits has already been documented by several theoretical and experimental studies [102-105].

The evolution of sex-biased genes and their influence on the evolution of sexual dimorphism is a growing field of research that has been poorly explored in the context of development and trait exaggeration. Sex-biased expression is often assumed to encode, at least to some extent, for the differences observed between males and females [106]. Different experimental studies have shown that variation in sex-biased expression at different developmental stages, in different organs or species is associated with variation in sexual dimorphism at the morphological, physiological or behavioral levels [102, 106-110]. Given the primary role of sexual selection in shaping male and female differences, the evolution of sex-biased gene expression has been hypothesized to result from the selection of male and female reproductive interests. Accordingly, many studies have now shown that sex-biased genes –notably male-biased genes that are supposed to be under strong sexual selection– are fast-evolving both at the coding and expression levels [102]. In a comparative study of birds, for example, the authors found an association between the rapid evolution of the proportion and expression of sex-biased genes, especially male-biased genes, and the increase of sexual dimorphism at the phenotypic level. However they did not find any association with their sequence evolution, although sex-biased genes were generally faster evolving than unbiased genes across species [111]. These results were observed only in adult gonad tissues as in somatic tissues only one locus displayed sex-biased expression [111].

In species where a good genome assembly is available and the sex chromosome(s) can be identified, sex-biased genes can also provide some insights into sex chromosome evolution. A general feature among sexually dimorphic characters lies in their antagonistic selection between males and females. Selected phenotypes favored in male through increased reproductive success are often costly for female fitness, the reverse being also true for females selected phenotypes [28]. Such sexual antagonism has notably led to the theory that sex-biased genes could preferentially accumulate on the X chromosome where the selection is most effective [112]. Because of the hemizyosity of the X chromosome in males, X-linked genes are constantly exposed to selection and male beneficial recessive mutations are more likely to accumulate relatively to autosomal recessive mutations. On the other hand, female beneficial dominant mutations are also more likely to accumulate on the X chromosome, which is twice as often present in females than males [112]. This expectation is however reverted in systems where the females are the heterogametic individuals. Altogether, these assumptions predict an accumulation of sex-biased genes on the X chromosome and an elevated rate of protein evolution among the X-linked genes (also known as faster-X

evolution) relative to autosomal genes [102]. So far, experimental evidence supports this sexualization of the X chromosome across taxa, although male-biased genes turned generally to be under-represented on the X chromosome [112-117]. It is, however, important to note that the general pattern described so far may also result from several non-adaptive processes (e.g. relaxed purifying selection, sex chromosome dosage compensation...) and lead to some misinterpretations regarding the possible effects of sexual selection on genome evolution [102, 118]. Moreover, studies assessing genes with sex-biased expression are often performed in adults where sexual dimorphism has already been established, at least in part, during development [106]. This implies that 1) a large proportion of relevant sexually dimorphic genes have been overlooked [107, 108], 2) adult sex-biased genes have thus far unclear functions in sexual dimorphism, especially when they are identified from whole adult transcriptomes [106]. Addressing similar questions with a greater attention to the developmental stage(s) and tissue(s) selected will help refining the type of selective pressures acting on sex-biased genes and their general influence on genome evolution.

Such genome-wide analysis have already been initiated in horned beetles, for example, where Ledón-Rettig *et al.* used the sex-specific functions of *doublesex* isoforms in horn development to characterize their genome-wide sex-specific interactions [119]. They found especially that *doublesex* has a general regulatory influence on genes with sex-biased expression. These regulations seem to be often sex-specific, either by regulating a sex-specific target or by regulating the same gene in opposite directions in males and females [119]. Another study in horned beetles has also found signatures of selection in conditionally expressed genes during horn development [120]. Several other comparative transcriptomics studies have been performed in the developing exaggerated sexually selected traits but with the intention to characterize the genes and pathways involved in their developmental growth [121-125]. More efforts need therefore to be done on the genomic basis of sexual trait exaggeration.

IX. Intra-species variation in sexually selected traits

One striking feature of sexually selected exaggerated traits compared to other traits is their phenotypic variability within individuals of the same population. These traits can sometimes become so variable between males that naturalists were first thinking that extreme morphs of a population were in fact members of different species [5, 19]. As described by Charles

Darwin, and by many other evolutionary biologists after him, males with the highest exaggeration have more chance to secure partners either by male-male competition or by female choice. These observations suggest that exaggerated sexual traits are under directional selection and should therefore present low genetic and phenotypic variance in the population. Yet, these traits have been described to have unexpectedly high genetic variance and to be phenotypically hypervariable compared to other homologous traits in the same or different species [38]. How variation is maintained in sexually selected traits is probably one of the main puzzling questions in the field of sexual selection. Several theoretical studies have been formulated to explain the maintenance of phenotypic and genetic variance in traits under sexual selection, (e.g. handicap theory, genic capture theory) however few studies have been able to experimentally test them [10, 38, 39, 126-130]. Among the different research avenues, identifying the selective pressures and genetic mechanisms controlling phenotypic variance seems to be one of the most urgent questions to understand this paradox. This research topic is the core of my PhD study and will be further detailed in the following sections.

X. The semiaquatic insects (Gerromorpha) as a model to study the role of natural and sexual selections on trait evolution

Gerromorpha is a monophyletic group of insects in the “true bug” order Hemiptera, thought to have arisen in the Triassic about 200 million years ago [131]. Commonly named the semiaquatic insects, they are composed of about 2000 species classified in eight families, living all on the water surface and distributed worldwide except in the poles (Figure 5) [72]. The Gerromorpha share a common ancestor that is thought to have transited from terrestrial to more humid substrate surfaces. Extant species occupy diverse habitats including small water drops on leaves, lakes, streams or even oceans. The adaptation and diversification to this novel environment is associated with various phenotypic changes in morphology (e.g. appendage morphology), physiology (e.g. tolerance to salinity) and behaviour (e.g. locomotion) [72]. The modification of their legs is nonetheless one the most striking and most important changes associated with their habitat transition. All semi-aquatic bugs present longer legs with higher hair density compared to terrestrial insects [68, 71, 72, 132]. These modifications are thought to increase the general floatability of the bugs on the water surface by improving their mass distribution and leg hydrophobicity.

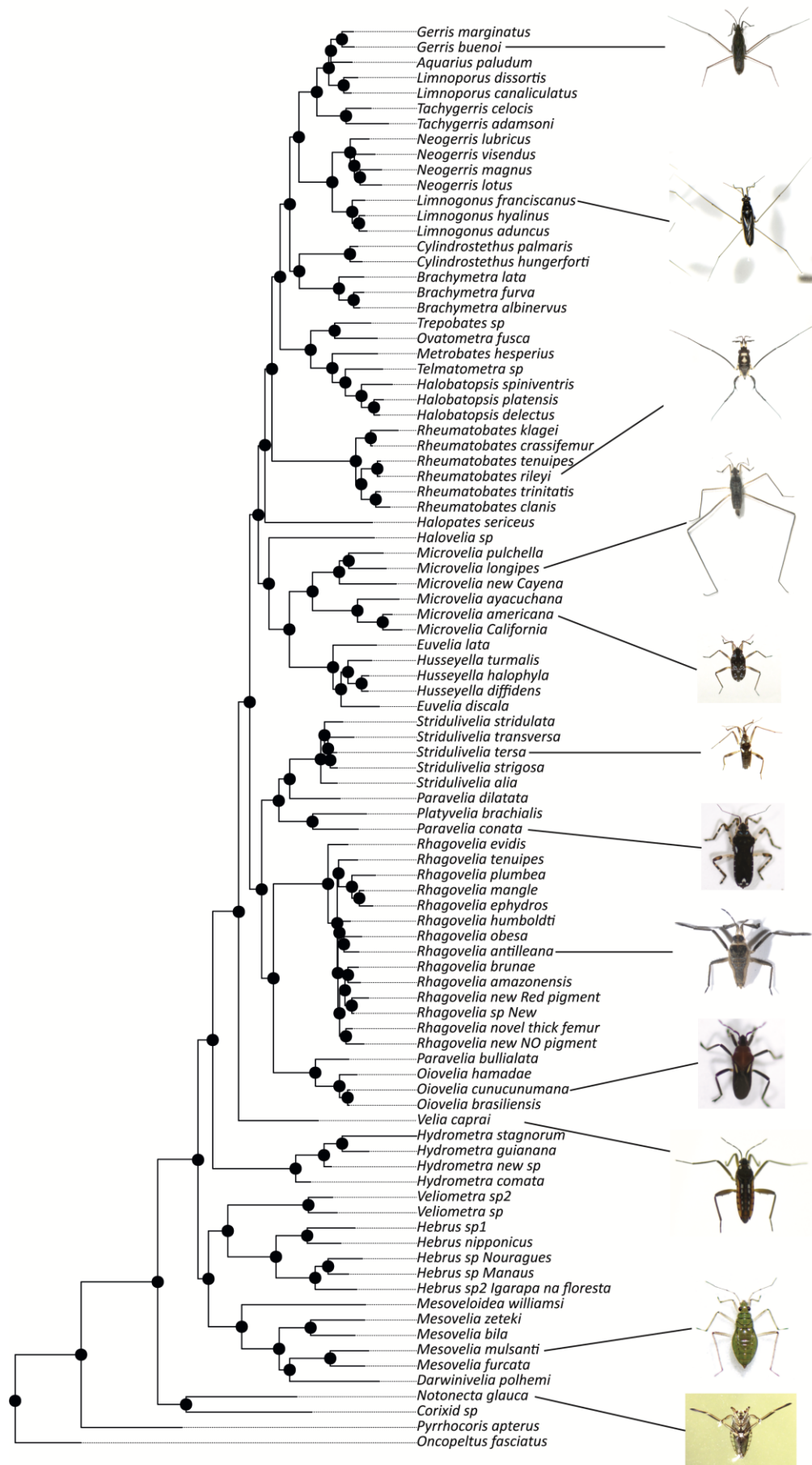


Figure 5: Gerromorpha (water striders) phylogeny. Phylogenetic relationships of 89 species of water striders and 4 other Hemiptera species used as outgroups.

The transition of the semi-aquatic bugs to the life on water follows a certain continuum where the more basally branching groups like the Hydrometridae, Hebridae and Mesoveliidae have the tendency to live near shore, at the interphase between terrestrial and water surfaces such as wet soil. To move on water, these species use the ancestral mode of locomotion, the walking gait, that consists of alternating the movements of the first and rear legs on one side with the mid-leg on the other side of the body [133]. The more derived species have specialized in more open water surfaces and their mode of locomotion has now changed into a rowing movement where the mid-legs are moved simultaneously from the anterior to the posterior part of the individual in order to propel it forward. This new mid-leg function is associated with the evolution of a new leg ground plan where the mid-legs are now longer than the rear legs. In the new mode of locomotion, the first and rear legs act as balancers [133]. In the recent years, Gerromorpha has become a new model in the field of Evo-Devo, allowing a more integrative understanding on the adaptive function of leg diversity. The development of genetic tools such as *in-situ* hybridization and knockdown using RNA interference (RNAi) sheds light on the evolution of gene expressions and genetic networks controlling the leg elongation and the development of a new ground plan in these insects [65-69]. Despite a growing interest in understanding the adaptive and developmental mechanisms associated with the transition to the water surface habitat [134], these insects are primarily studied as models for sexual selection. A large body of literature is dedicated to the interaction between males and females and the associated evolution of mating behaviour. During the breeding season, their mating activity can become very intense and easy to observe as it often occurs in restricted areas, on the water surface [135, 136]. It has been shown that a conflict could emerge between males and females during mating. In many species, males display harassing behaviour by chasing, grasping and/or leaping upon females in an attempt to mate. Females generally try to escape these mating harassments but when a male succeeds in grasping the female from her back, a vigorous pre-mating struggle ensues. As often, females will be reluctant to mate and will initiate a rejecting behaviour consisting of roles, somersaults and leverage flips in order to dislodge the grasping male. In the occasions when the male persists, copulation occurs. Interestingly, males have evolved all sorts of adaptive sex-specific phenotypes, especially morphological ones, to enhance their ability to grasp and maintain themselves during mating struggles. For example, males of several species have developed spikes on their rear legs to grasp female legs and/or abdomen [72, 137]. One very striking example occurs in the genus *Rheumatobates*, where males of some species have developed extremely complex leg and antennae structures that serve them to hold the female's

head when the latter is wrestling [36]. In the semi-aquatic insects, females possess a spermatheca that makes them capable of storing viable sperm for several days. Females therefore need few matings over their lifetime to secure the fertilization of their eggs [72, 136]. As a consequence, a prolonged or an additional number of matings will at some point become costly for the fitness of the female. As a consequence, females of some species will be favoured if they also evolved, in addition to their resistant behaviour, some anti-grasping structures that will act against the adaptive secondary sexual traits developed by the males. The most famous examples of such antagonistic traits are found in species of *Gerridae*, where females have developed abdominal spines near their genitalia to increase the efficiency of male dislodgement [138-140]. Interestingly, when we compare the degree of abdominal spines erection in different species of the genus *Gerris*, we observe that the more exaggerated the female's spines are and more the males of the same species have evolved flattened distal part of the abdomen in association with prolonged genital and pregenital segments. A strong correlation exists therefore between grasping and anti-grasping structures in males and females, highlighting a certain correlated evolution of these sex-specific structures [138]. This process is described as an evolutionary arms race where a positive feedback takes place and leads to the co-evolution and escalation of grasping and antigrasping structures involved in the sexual conflict over mating rate between the two sexes [28].

Other mating strategies have been described in Gerromorpha but with much less attention than the sexual conflict [141]. Some studies report for example that males of another Gerrid, *Gerris elongatus*, adopt different alternative mating strategies according to the seasonal period [142]. In the early breeding seasons, males are very active and produce a series of surface waves that are used as signals to attract or court the female near or far away from their egg-laying sites (e.g. piece of fallen bamboo). In this species, males can establish a territory around the oviposition sites and can sometimes fight with their midlegs for territory dominance against other intruder males. In midseason, the females will tend to lay eggs just after mating and the male will initiate a guarding behaviour to protect her from other male disturbances. Another study in *Aquarius remigis* described that in addition to male-male competition and male-female conflict, females of this species display some kind of female choice by favouring smaller males that have shorter mating duration [143]. A large body of literature on natural and sexual selection has been documented in the Gerridae family, a derived lineage from the Gerromorpha group. However, much of the diversity in mating systems and behaviours still needs to be explored in more basally branching species. A large

number of secondary sexual characters have already been described anatomically across the semiaquatic insects, but few of them have a characterized function (except in the Gerridae family). In addition, the establishment of key molecular technics (e.g. RNAi interference) associated with the rapid acquisition of new genomic data make this group of species a good emerging system to study the adaptive and genomic basis of sexual dimorphism.

Chapter 1: Fluctuating selection strength and intense male competition underlie variation and exaggeration of a water strider's male weapon

Authors: William Toubiana¹ and Abderrahman Khila¹

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

Keywords: Weapons, scaling relationships, phenotypic variation, mating systems, reproductive fitness

I. Abstract:

Sexually selected traits can reach high degrees of phenotypic expression and variation under directional selection. A growing number of studies suggest that such selection can vary in space, time and form within and between populations. However, the impact of these fluctuations on sexual trait evolution is poorly understood. In the water strider *Microvelia longipes*, males display striking trait exaggeration and phenotypic variation manifested as extreme differences in the rear leg length. To study the origin and maintenance of this exaggerated trait, we conducted comparative behavioral, morphometric and reaction norm experiments in a selection of *Microvelia* species. We uncovered differences both in the mating behavior and the degree of sexual dimorphism across these species. Interestingly, *M. longipes* evolved a specific mating behavior where males compete for egg-laying sites, consisting of small floating objects, to intercept and copulate with gravid females. Through male-male competition assays, we demonstrated that male rear legs are used as weapons to dominate egg-laying sites and that intense competition is associated with the evolution of rear leg length exaggeration. Field observations revealed rapid fluctuation in *M. longipes* habitat stability and the abundance of egg-laying sites. Paternity tests using genetic markers demonstrated that small males could only fertilize about 5% of the eggs when egg-laying sites are limiting, whereas this proportion increased to about 20% when egg-laying sites become abundant. Furthermore, diet manipulation and artificial selection experiments also showed that the exaggerated leg length in *M. longipes* males is influenced by both genetic and nutritional factors. Collectively, our results highlight how fluctuation in the strength of directional sexual selection, through changes in the intensity of male competition, can drive the exaggeration and phenotypic variation in this weapon trait.

II. Introduction

Sexually selected traits represent some of the primary examples of intra- and interspecies phenotypic variation [5, 19]. Males in both vertebrates and invertebrates are known to display extravagant phenotypes that differ in their nature, location, size, and shape [5, 19, 42, 144]. Such degrees of trait expression and variation among individuals are often associated with the type of mating strategy employed. In some systems, conflict over mating rate between the sexes can drive profound morphological modifications, such as male antennae in water striders that are used to grasp the females during pre-mating struggles [36]. In other examples, such as in horned beetles or ruffs, males occur in discrete morphs associated with alternative mating strategies. Large males dominate territories whereas small males tend to sneak [145-147].

An extreme form of variation lies in sexual characters displaying a continuum of trait expression with no distinguishable discrete morphs in the population [42, 144]. A central prediction for these exaggerated traits to evolve is that only large individuals can afford to bear them as they are good indicators of body size, thus representing an honest signal for male quality [39, 41, 148]. Under this prediction, females will favor males with the highest trait expression, which imposes strong directional selection in favor of trait exaggeration [19]. In other situations, the exaggerated trait is used as a weapon in male-male competition, and its size is a good predictor for the outcome of the contest over access to females [149, 150].

In these examples, sexual selection is thought to be directional and persistent over time [39, 126]. These traits are also known to be subject to survivorship costs, which constrain their degree of expression resulting in a net stabilizing selection [144]. These observations raise important questions regarding the maintenance of phenotypic variation in natural populations [19, 38, 39, 126, 128]. A growing number of studies suggest that selection may fluctuate over time and space, and that environmental changes may influence the strength, direction, and form of sexual selection [151-156]. These fluctuations in selection may, in turn, favor genetic variation and elevated plastic response observed in sexual traits, consequently influencing their variation and evolution [39, 151, 152]. Nonetheless, empirical studies have shown that the genetic and plastic influence on phenotypic variation could be highly variable from one species to another. For example, the three male discrete morphs of ruffs are mostly controlled genetically by different alleles in an inverted chromosomal region [157, 158]. Major and minor morphs in dung beetles or wild turkeys, however, can mostly be

recapitulated by changes in the environment such as nutritional intake or male competition [159, 160]. Studies assessing the interplay between selection, genetics and plasticity, within the context of a changing environment are therefore crucial to further the general understanding of the origin and maintenance of highly variable exaggerated sexual traits.

Here we focus on a novel model system, the water strider *Microvelia longipes*, that displays a strong sexual dimorphism where males have evolved both longer and more variable rear legs than females [72]. The genus *Microvelia* (Heteroptera, Gerromorpha, Veliidae) comprises some 170 species of small water striders distributed worldwide and occupying various fresh water habitats including temporary rain puddles and stable large water bodies [72]. First, we reconstructed phylogenetic relationships of five *Microvelia* species and compared their degrees of dimorphism, scaling relationships between leg and body length, and various aspects of mating behavior. We report a clear association between the intensity of male competition and the evolution of trait exaggeration in *M. longipes* males. We then determined the fitness advantages of these exaggerated legs through fertilization success performed under selective conditions reflecting fluctuations in their natural environment. Finally, we assessed the contribution of the strength of sexual selection, genetic variation, and phenotypic plasticity to the variation of exaggerated rear legs in *M. longipes* males.

III. Material & methods

A. Population sampling and culture

Microvelia populations were collected during fieldwork in French Guyana in Crique Patate near Cayenne. The bugs were maintained at 25°C and 50% humidity and fed on crickets.

B. Measurement of *Microvelia* species and statistics

Rear leg and body lengths of all *Microvelia* species were measured with a SteREO Discovery V12 (Zeiss) using the Zen software. All statistical analyses were performed in RStudio 0.99.486. Comparisons for mean trait size and trait distributions were performed on raw data whereas log-transformed data were used for scaling relationship comparisons. We used Standardized Major-Axis (SMA) regression to assess differences in scaling relationships (“smatr” package in R, [161]). Differences in intercepts were estimated using a Wald statistic test and we used Likelihood ratio test for differences in slopes [161].

C. Behavioral observations and video acquisition

Male and female interactions of all *Microvelia* species were observed in a recreated small puddle, using local mud, and were filmed with a Nikon digital camera D7200 with an AF-S micro nikkor 105mm lens. Observations and video acquisitions were taken a couple of hours after the bugs were transferred to the puddle. In *M. longipes* and *M. pulchella* male and female interactions were also observed in the field.

D. *Microvelia* phylogenetic reconstruction

The phylogenetic relationships between the five *Microvelia* species used in the behavioral assays was generated using the Geneious software version 7.1.9 using plugins MrBayes version 3.2.6 and PhyML version 3.0, as described in [133]. The phylogenetic reconstruction was performed using 14 molecular markers. Supplementary table 1 presents the identity and Genbank accession numbers for these markers. Phylogenetic reconstruction was performed using MrBayes version 3.2.6 and PhyML version 3.0 in Geneious 7.1.9 as described in [133].

E. Male competition assay

We generated five groups of three independent males from *M. longipes* lab population (N=15 individuals). The males of each group were chosen for their differences in rear leg length (large, intermediate and small legs) and painted on their back. A male from each category fought five times with a male from the two other categories (total number of fights per male=10).

F. Fight frequency assay

To compare the number of fights between males of *M. longipes* and *M. pulchella*, we isolated twenty-five adult males and females over a period of two days. Both sexes were then mixed together in the puddle during 30 minutes before observation. The number of fights on and outside floaters was counted during a period of one hour (Supplementary table 2). We repeated the experiment the following day with the same males and females kept together overnight (Supplementary table 2). In order to correct for size differences between the two species, we calculated the number of fights in a reduced sample of ten males and ten females

in *M. longipes* (Supplementary table 2). In all conditions, individuals were selected randomly (with respect to body and leg size) from the lab populations.

G. Artificial selection experiment

Individual males from the French Guyana natural population were selected for their absolute rear leg size and mated with a random female to initiate the successive sib-sib crosses. After fifteen generations of sib-sib inbreeding, two populations selected for extreme phenotypes were amplified over two generations before phenotyping.

H. Condition-dependence experiment

First instar nymphs were collected just after hatching and individuals were reared in either poor or rich nutritional condition. In the poor condition, a hundred first instar nymphs of the long-legged inbred line were fed with ten crickets per day during the first two nymphal instars, followed by only three cricket legs until adulthood. In the rich condition fifty individuals of the same line were fed with ten crickets per day over their entire nymphal development. In a second experiment we tested the effect of condition in an independent set of individuals from the lab population. This experiment was performed on three replicates per condition, with fifty individuals per condition. Replicates were then pooled for the analysis. We started the poor condition by feeding the first two nymphal instars with eight crickets per day and then switched to one small cricket every two days until they reached adulthood. Individuals from the rich condition were fed during their entire nymphal development with eight crickets per day.

I. Microsatellite development

DNA from *M. longipes* was extracted from ten males and females using the Genomic Genomic-tip 20/G DNA extraction kit from Qiagen. Using an Ion-Torrent Sequencer machine, we generated 3.7M reads with median size of 317 bp. These sequences were used to identify reads containing microsatellite repeats using Exact Tandem Repeat Analyzer 1.0 software [162]. The primers for microsatellite amplification can be found in (Supplementary table 3).

J. Paternity test

To assess the fertilization success of long and short-legged males, we collected six males from both the short- and the long-legged inbred lines, and put them together in an artificial puddle with twelve females from the long-legged inbred line. We conducted two treatments, each with four replicates, where we provided twenty floaters or three floaters in the puddle to create conditions with abundant and limiting egg-laying sites, respectively. On day 3, the parents were collected, their DNA extracted and the microsatellite of interest amplified using the protocol in supplementary table 10 and sent for genotyping to Genoscreen, Lille, France. We then isolated the floaters and genotyped the nymphs that hatched from the floaters and those that hatched from the mud after adults and floaters were removed.

IV. Results and discussion

A. Sexual Dimorphism and scaling relationships in *Microvelia* species

We found a considerable inter-species variation in the degree of sexual dimorphism within the *Microvelia* genus (Figure 1; Supplementary figure 1). Measurements of various body parts revealed dimorphism in average body length, leg length, and the scaling relationship between these two traits (Figure 1B; Supplementary table 4). In some species, such as *M. americana* and *M. paludicola*, the dimorphism in leg and body length is small, whereas in others such as *M. longipes*, the dimorphism is most striking (Figure 1A). The extreme leg elongation found in *M. longipes* males is associated with the evolution of hyperallometry where the allometric coefficient is significantly higher than 1 and reaches a value of 3.2 – one of the highest known (Figure 1B; Supplementary table 4) [163, 164]. In contrast, *M. longipes* females and both sexes of all other species show isometric or near-isometric scaling relationships between leg and body length (Figure 1B; Supplementary table 4). *M. longipes* male legs are both significantly longer and more variable than female legs (Figure 2A, B; see statistical tests in supplementary table 4). In contrast, *M. longipes* body size is significantly more variable in males than in females, although females show slightly longer average body length (Figure 2A, C; see statistical tests in supplementary table 4). Despite these differences, leg and body lengths in both sexes assumed normal distribution (Shapiro tests: male third legs, $W=0.99$; male bodies, $W=0.99$; female third legs, $W=0.98$; female bodies, $W=0.97$; all p -values > 0.05 ; Supplementary table 5).

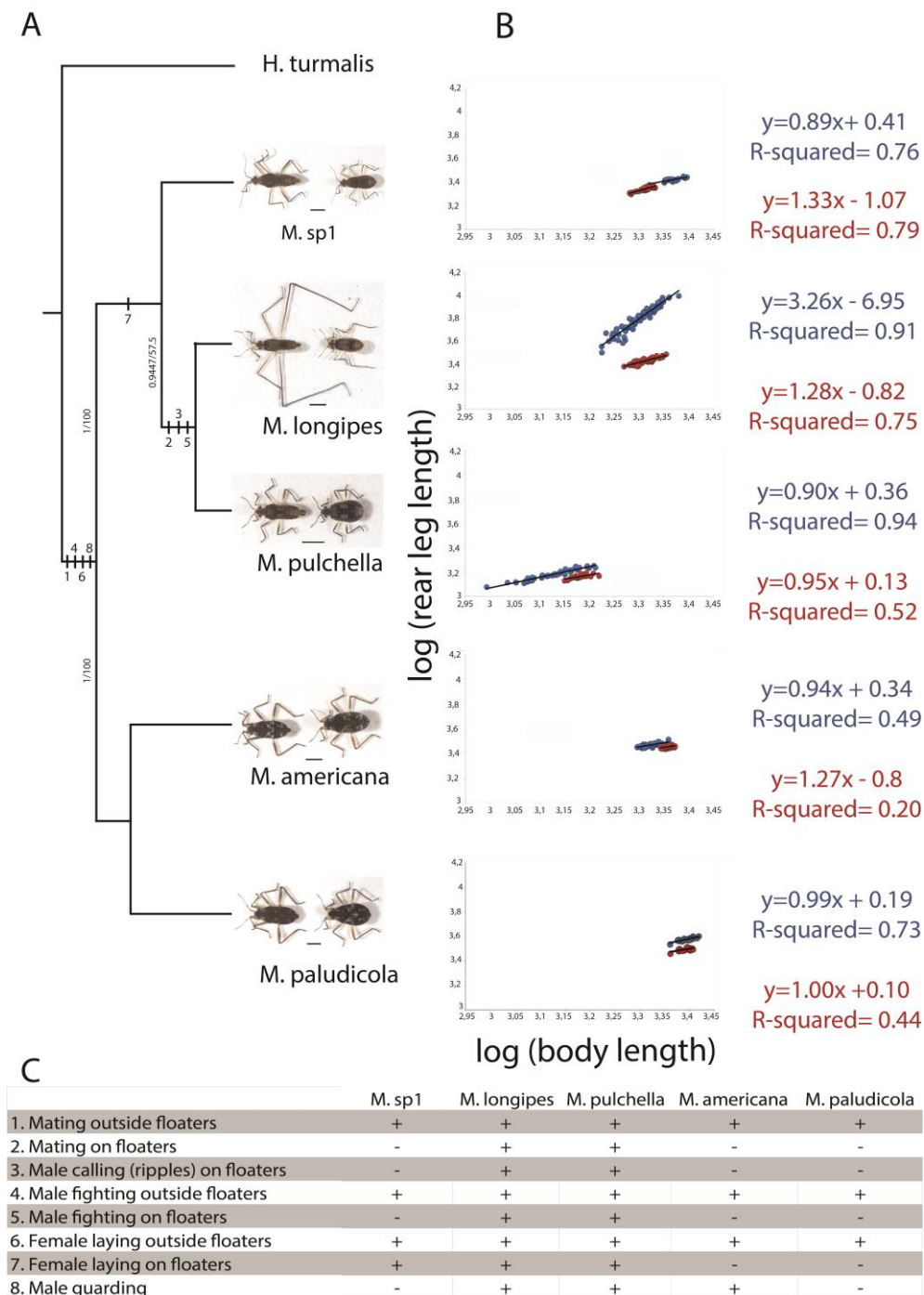


Figure 1: Diversity in leg sexual dimorphism and mating behaviours in *Microvelia*. A) Phylogenetic relationships between five *Microvelia* species using Maximum Likelihood and Bayesian analyses. Support values obtained after Bayesian posterior probabilities and 1000 bootstrap replicates, respectively, are shown for all branches. Pictures of males (right) and females (left) illustrate divergence in sexual dimorphism in the five *Microvelia* species. Scale bar represents 1mm. B) Scaling relationships of log-transformed data between rear legs and body lengths were estimated in males (blue) and females (red) of the five *Microvelia* species using Standardized Major Axis (SMA) regressions. The equations and fitting (R-squared) of the regressions in males and females were indicated using the same colour codes. C) Behavioural characters describing the mating system of the five *Microvelia* species. These characters were mapped onto the phylogeny based upon the parsimony criterion.

Finally, we found that the rear legs in males from three *Microvelia* species (*Microvelia* sp., *M. americana* and *M. paludicola*) bore prominent spikes that may function to grasp females during pre-mating struggles [28] (Figure 1A; supplementary Figure 1). Overall, these analyses indicate that the evolution of hypervariable and exaggerated legs in *M. longipes* males is a derived trait resulting from the high variance in body length and the associated hyperallometric relationship with leg length. In *M. pulchella*, despite the high variation in male body length, the near isometric relationship between leg and body length makes their legs less exaggerated and less variable than *M. longipes* males (Figure 1B; Supplementary table 4). Moreover, the diversity of sexual dimorphism in leg morphology between *Microvelia* species does not seem to follow any phylogenetic signal (Figure 1; Supplementary table 6), suggesting that variation in the ecology, behavior, or mating systems may play a role in the divergence of the sexes in these species.

B. Mating systems in *Microvelia* species.

We characterized mating systems and sexual interactions in all five species to better understand the differences in sexual dimorphism (Supplementary figure 2). In nature, the *Microvelia* genus comprises species that occupy a wide variety of habitats [72]. Most species live nearshore, in stagnant, large water bodies [72]. Some species, like *M. longipes*, *M. pulchella* or *Microvelia* sp. are gregarious and specialize in small temporary puddles filled with rainwater in tropical South America [72]. Behavioral observations both in the wild and in laboratory-recreated puddles revealed that *M. longipes* males are highly territorial and tend to aggressively guard floating objects such as small twigs or pieces of dead leaves (Supplementary figure 3). These are egg-laying sites where males signal to attract females, by vibrating their rear-legs and pounding with their genitalia on the water surface to generate ripples (Supplementary videos 1 and 2). We hereafter refer to these objects as egg-laying floaters. When a female approaches the floater, the dominating male switches from signaling to a courtship behavior. After inspecting the floater, she either leaves or mates without any resistance with the courting male, and immediately lays 1 to 4 eggs (n=26 mating events) (Supplementary video 2). The male then initiates an aggressive guarding behavior by turning around the egg-laying female and chasing other approaching males (Supplementary video 2). After the egg-laying female leaves and the male initiates another cycle of signaling on the same floater.

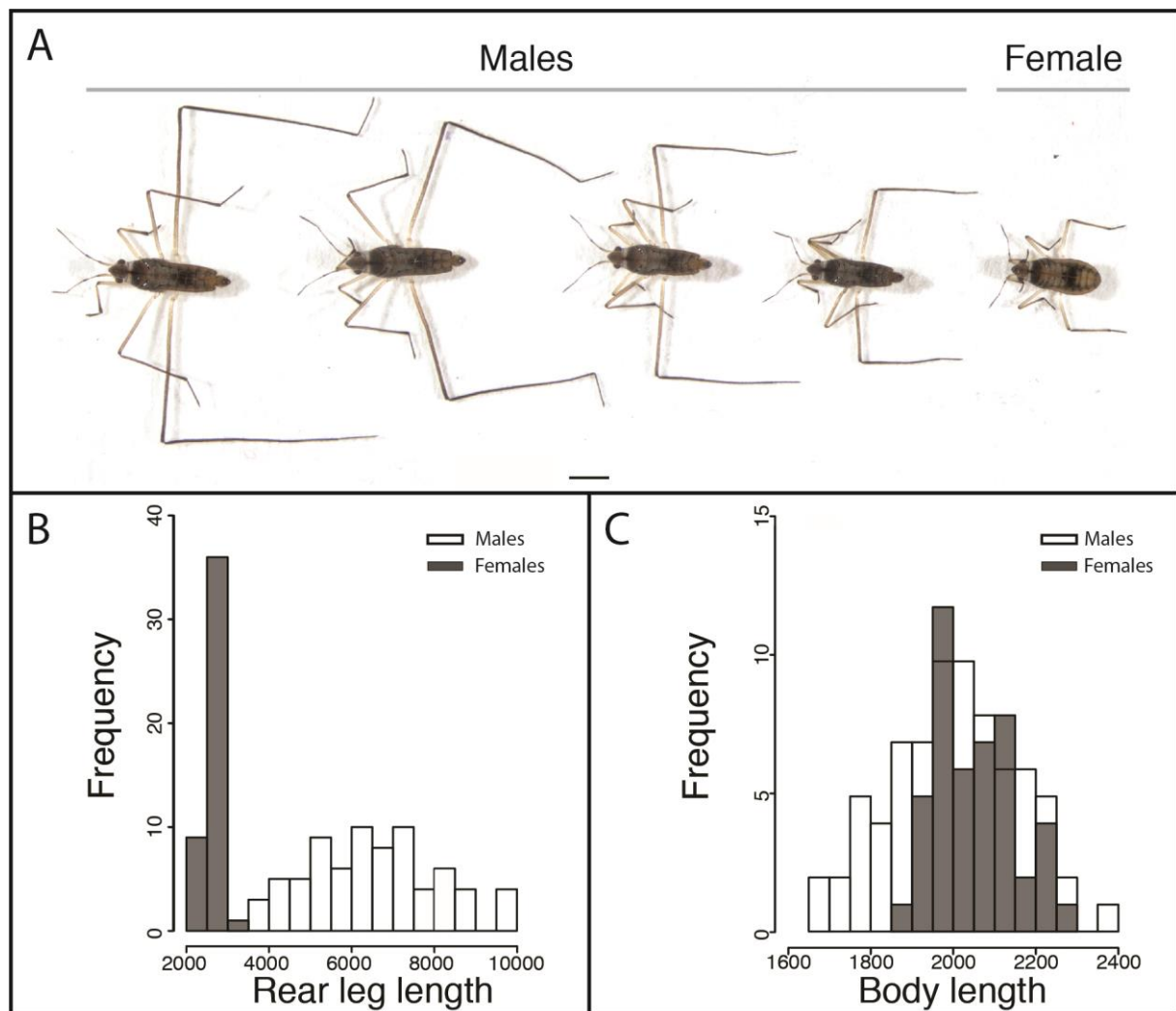


Figure 2: Phenotypic variation of rear leg exaggeration and body length in *M. longipes*. A) Phenotypic variation of rear leg length in males and in a female. (B) Rear leg length and (C) body length distributions of males (white) and females (grey) from a natural population collected in French Guyana. Leg and body measurements are in micrometres.

During this entire process, other males constantly challenge the signaling male in an attempt to dominate the floater. During these contests, the dominant and the challenging male fight back-to-back by kicking each other with their rear-legs until one of them is chased away (Supplementary video 2). We also observed that females could lay eggs in the mud at the margin of the puddle and that males attempt to mate outside floaters by jumping on females' back randomly in the puddle.

M. pulchella, the sister species of *M. longipes* (Figure 1A), is also found in small temporary puddles and displays a highly similar mating behavior despite the lack of rear-leg exaggeration (Figure 1C). Males of *M. pulchella* compete for egg-laying floaters, fight with their rear-legs, and generate ripples to attract females. Like *M. longipes*, females of *M. pulchella* also lay their eggs on floaters and in the mud (Supplementary video 3; Supplementary figure 2). Despite the similarities in their mating behavior, these two sister species display significant leg length differences, raising the question as to which factors drove the evolution of leg exaggeration in *M. longipes*.

In the three other species, *M. americana*, *M. paludicola*, and *Microvelia* sp., males possess grasping spines on their rear-leg femurs (Figure 1A; Supplementary Figure 1) and actively harass females in an attempt to mate. Females consistently struggle through vigorous shaking, frequently resulting in the rejection of the male. Males of these three species also fight occasionally but the fights do not seem to result in the dominance of any particular localized resource (Figure 1C; Supplementary video 4; Supplementary figure 2). *M. americana* and *M. paludicola* females lay eggs exclusively at water margins while *Microvelia* sp. females lay eggs either on floaters or at water margins, but not immediately after mating (Figure 1C; Supplementary video 4; Supplementary figure 2). Altogether, these data show that the behavior consisting of male contests using the rear-legs is plesiomorphic among *Microvelia* species, predating the origin of the derived exaggerated legs in *M. longipes*. Male contests seem therefore necessary but not sufficient for the evolution of exaggerated weapons. Moreover, differences in egg-laying habits may have driven the diversity in male mating strategies and sexual dimorphism in the *Microvelia* genus. In small temporary habitats, eggs laid in the mud are at high risk of desiccation when water levels go down, and nymphs tend to drown at hatching when water levels go up, something we frequently observe in laboratory conditions (data not shown). Laying eggs on floating objects, which remain on the surface

despite fluctuating water levels, is likely an adaptation to the fast-changing state of the habitat. Interestingly, male behavior consisting of dominating these egg-laying floaters is observed only in species where females lay eggs just after mating, indicative of the high fitness value for the males who fertilize these eggs. This behavior is also associated with a high body length variance in *M. longipes* and *M. pulchella* males (Figure 1B), suggesting a link between body size variation and competition for oviposition sites.

C. Intensity of male competition in *M. longipes* compared to *M. pulchella*

In order to evaluate the contribution of exaggerated leg length to male mating success, we tested whether a correlation existed between male leg length and their ability to dominate egg-laying sites. We found increased rear leg length to be strongly correlated with the favorable fighting outcome, where the males with longer legs won 97% of the fights ($n=75$ fights) and dominated the floater (GLM: $z\text{-value}=2.133$, $p\text{-value}<0.05$; Figure 3A; Supplementary table 7). We also observed this male dominance over egg-laying sites in *M. pulchella*, which did not evolve leg exaggeration. We therefore hypothesized that the phenotypic differences in male legs between *M. longipes* and *M. pulchella* could be driven by differences in the intensity of male competition. When we measured male competition in standardized space conditions, we found that *M. longipes* males on average fought 8 times more than *M. pulchella* males within one hour (t-test: $t=15.18$, $df=4$, $p\text{-value}<0.05$; Figure 3B, Supplementary table 2). Importantly, *M. longipes* males fought significantly more often on floaters than outside floaters, with 81% of the fights occurring on floaters (t-test: $t=9.37$, $df=4$, $p\text{-value}<0.05$; Figure 3C; Supplementary table 2), whereas *M. pulchella* males fought randomly on or away from floaters (t-test: $t=0.15$, $df=4$, $p\text{-value}=0.89$; Figure 3C, Supplementary table 2). Similar results were obtained when we repeated this experiment in standardized density conditions taking into account the size differences between the two species (Supplementary table 2). These data demonstrate that increased rear leg length in *M. longipes* males favors male dominance over egg-laying sites to better intercept gravid females. While both *M. longipes* and *M. pulchella* males intercept females and compete on those egg-laying sites, competition intensity for egg-laying sites is almost an order of magnitude higher in *M. longipes*. A primary difference between the ecology of these two species is that *M. longipes* specializes in rainwater-filled small puddles while *M. pulchella* is a generalist that can be found in both temporary and more stable water bodies ([165] and personal field observations). This difference in niche specialization has two major impacts on

M. longipes population structure. First, *M. longipes* populations can reach very high densities confined in a small space, something we observed frequently in the wild and which is not the case for *M. pulchella*. Second, because the water level in the puddle can change rapidly (Supplementary figure 4), floaters represent the safest substrate in terms of survival of the progeny. This may explain why females bounce the floater up and down before they copulate and lay eggs (Supplementary video 2), and why *M. longipes* males are particularly aggressive in dominating these floaters. In contrast, *M. pulchella* occupy a more stable habitat, making floaters less critical and the survival of eggs in the mud more likely. The ecological conditions favoring high-density populations and floating objects as the more suitable egg-laying substrate may have at least contributed to the high competitiveness observed in *M. longipes*, and thus acted as a driving force for the evolution of the exaggerated leg length for use as a weapon. Both empirical and theoretical models suggest that population density can influence aggressiveness and the intensity of sexual selection [166], and our data show how increased competitiveness can drive secondary sexual traits to reach dramatic levels of expression.

D. Effect of exaggerated leg length on male reproductive fitness in *M. longipes*

Post-mating competition is widespread in insects [167], including water striders [141, 168], and can strongly alter the outcome of pre-mating strategies [167, 169]. Field observations also indicate that the state of the habitat occupied by *M. longipes* can fluctuate rapidly and, sometimes, the water can evaporate entirely in days (Supplementary figure 4). Moreover, the available amount of resources that can serve as egg-laying substrates is highly variable from one puddle to another and therefore can also fluctuate with water level (personal observations from the field). We hypothesized that these rapidly changing conditions will influence competition and mating success across the distribution of male phenotypes. To test this hypothesis, we conducted paternity tests using *M. longipes* homozygous lines for distinct microsatellite markers that can reveal the identity of the parents (see methods for more details). We set the experiment such that heterozygous progeny could only originate from eggs fertilized by small males. Because egg-laying floaters represent the primary resource that males dominate to intercept gravid females, we designed a first treatment where floaters were limiting (3 floaters for 6 large and 6 small males) and another treatment where floaters were abundant (20 floaters for 6 large and 6 small males). We also genotyped the progeny from eggs laid in the mud to determine mating success of different male phenotypes in contexts other than the dominance of floaters.

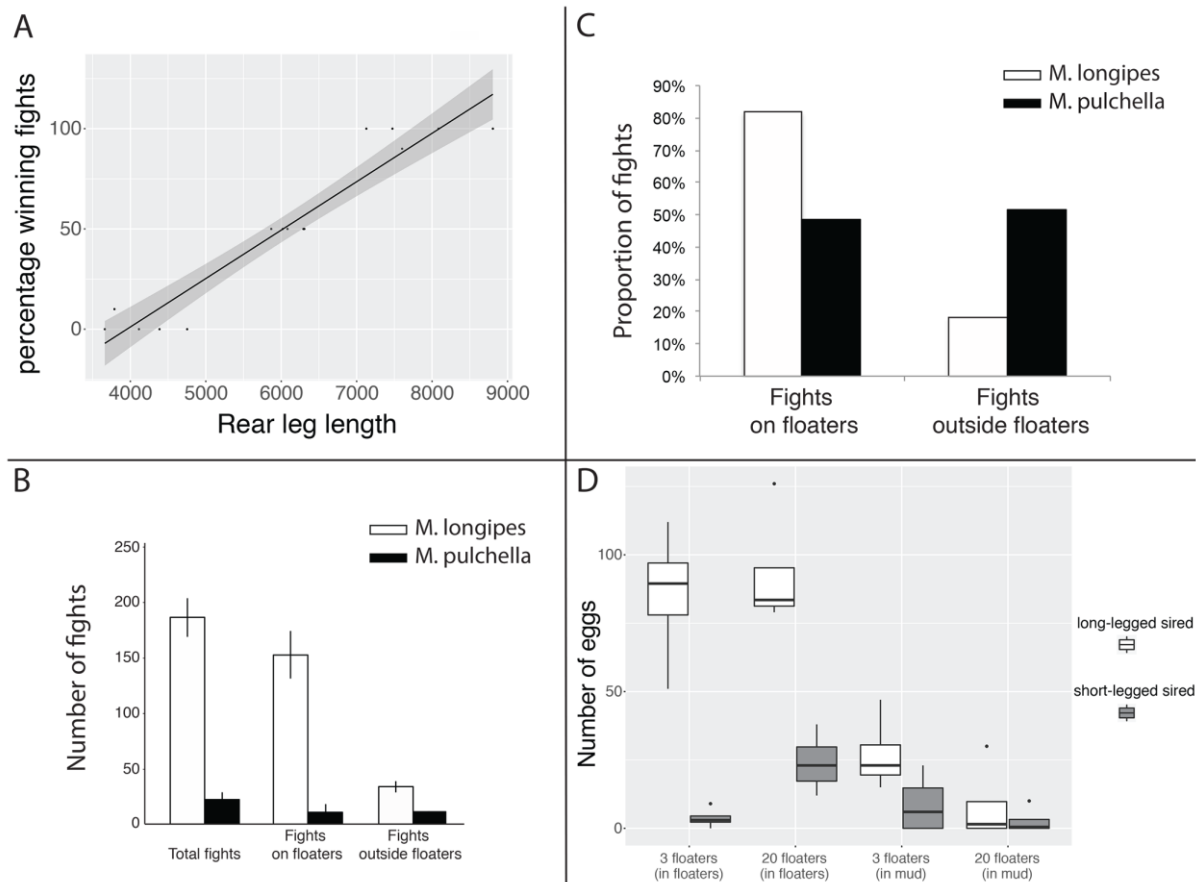


Figure 3: Selective pressures and reproductive fitness of leg exaggeration in *M. longipes* males. A) Relationships between fighting outcome and male rear leg length. Winners correspond to males keeping the access to the egg-laying sites after the fights. Solid line represents the fitted regression from a generalized linear model (statistics in Supplementary table 7), B) Frequency of fights between *M. longipes* and *M. pulchella* (N=50 individuals) on both floaters and outside floaters after two days of isolation. C) Proportion of male fights both on and outside floaters for *M. longipes* and *M. pulchella* (N=50 individuals) after two days isolation, D) Fertilization success of large and small males and the contribution of egg-laying sites. Heterozygous eggs result from the siring of short-legged males (short-legs selected line) and females (long-legs selected line). Homozygous eggs result from the siring of long-legged males (long-legs selected line) and females (long-legs selected line).

In all replicates of each treatment, females laid significantly more eggs on floaters regardless of whether floaters were abundant (91% of a total of 512 eggs) or limiting (71% of a total of 500 eggs) (abundant floaters: $t=5.63$, $df=6$, $p\text{-value} < 0.05$; limiting floaters: $t=3.02$, $df=6$, $p\text{-value} < 0.05$; Figure 3D; Supplementary table 8). However, they also laid on average three times more eggs in the mud when floaters were limiting, although this difference was not statistically significant ($t=1.56$, $df=6$, $p\text{-value}= 0.17$; Supplementary table 8). In the condition where floaters were limiting, small males fertilized 4.6% (15 eggs of a total of 357 eggs) of the eggs laid on floaters and 25% of the eggs laid in the mud (35 eggs of a total of 143 eggs) on average (GLM: $z\text{-value}= 5.903$, $p\text{-value} < 0.05$; Figure 3D; Supplementary table 8). This suggests that when the dominance of floaters by small males is limited, they primarily achieve egg fertilization by mating outside floaters. In the condition of abundant floaters, the proportion of eggs fertilized by small males on floaters increased significantly to 19% (96 eggs of a total of 468 eggs) (Figure 3D; Supplementary table 8), while that outside floaters remained unchanged (11 eggs of a total of 44 eggs) (Figure 3D; Supplementary table 8). In contrast to the treatment with limiting floaters, here the number of eggs fertilized by small males is almost nine times higher on floaters than in the mud (GLM: $z\text{-value}= -3.547$, $p\text{-value} < 0.05$; Figure 3D; Supplementary table 8). These results show that small males can sire significantly more progeny when egg-laying sites are abundant but can also mate outside these egg-laying sites when floaters are limiting. Although we cannot exclude the possible effect of assortative mating, these results indicate that sexual selection is strong in favor of large males with long legs but can become relaxed in conditions where egg-laying sites are abundant. Rapid changes in water level and high heterogeneity between puddles are intrinsic to the life history of this species and are expected to cause variation in the amount of accessible egg-laying floaters over time and space. This fluctuating selection is therefore likely to influence the strength of competition and mating success and contribute to the high phenotypic variation found in *M. longipes* natural populations.

E. Environmental and genetic contributions to male rear leg variation

To test the relative contributions of genetic variation and phenotypic plasticity to male phenotypic variation, we artificially selected large and small male lines, generated through 15 sib-sib successive crosses from a natural population. The large and small male lines showed a shifted distribution of male leg length towards the respective extreme phenotypes of the

distribution (Figure 4A-B). In these two lines, there is a significant difference in absolute and relative leg length (t-test: $t=22.21$, $df=85.266$, $p\text{-value} < 0.05$; Wald statistic test: $W=16.52$, $df=1$, $p\text{-value} < 0.05$; Supplementary table 9), but the allometric coefficient remained unchanged (Likelihood ratio test: Likelihood ratio statistic= 0.17, $df=1$, $p\text{-value}= 0.68$; Figure 4B; Supplementary table 9). This shows that genotypic variation contributes to the variation in both rear leg length and body size.

Next, we tested the reaction norm of one of these inbred lines and a laboratory population in poor and rich nutritional condition. Despite the near identical genotype, individuals from the inbred line reared in poor condition developed shorter legs than individuals reared in rich condition such that the distributions of the two treatments were almost non-overlapping (t-test rear leg length: $t = 15.374$, $df = 39.232$, $p\text{-value} < 0.05$; Figure 4C-D, Supplementary table 9). Importantly, this difference in leg length between the two treatments resulted mostly from differences in overall body size (t-test body length: $t = 10.5643$, $df = 25.274$, $p\text{-value} < 0.05$) but not in the scaling relationship as we failed to detect any significant difference in the allometric coefficient or the intercept between rich and poor conditions (Likelihood ratio test: Likelihood ratio statistic= 1.932, $df=1$, $p\text{-value}= 0.16$; Wald statistic test: $W=1.69$, $df=1$, $p\text{-value}= 0.19$; Figure 4D; Supplementary table 9). A similar result was obtained when we tested condition dependence in a laboratory population where no specific selection has been applied, although a small but significant difference in intercept was detected between the two conditions (Wald statistic test: $W=7.214$, $df=1$, $p\text{-value} < 0.05$; Supplementary figure 5; Supplementary table 9). This difference was nonetheless not significant when using a linear model (ANCOVA, $F(1,88)= 2.6202$, $p\text{-value}= 0.1076$). We therefore conclude that, in *M. longipes*, body length is highly dependent on nutritional condition. However, the scaling relationship between leg length and body length shows little, if any, condition-dependence. Altogether, these results suggest that male leg length variation in nature results from the contribution of both genetic variation and strong condition dependence. The fluctuations in the number of egg-laying floaters, combined with phenotypic plasticity, is expected to result in the maintenance of a certain degree of genetic variation in the population through the incomplete removal of alleles of small leg and body size. However, episodes of relaxed selection are not only known to increase genetic variation in the population, but also to favor the evolution of reaction norms and therefore increase phenotypic plasticity [170, 171].

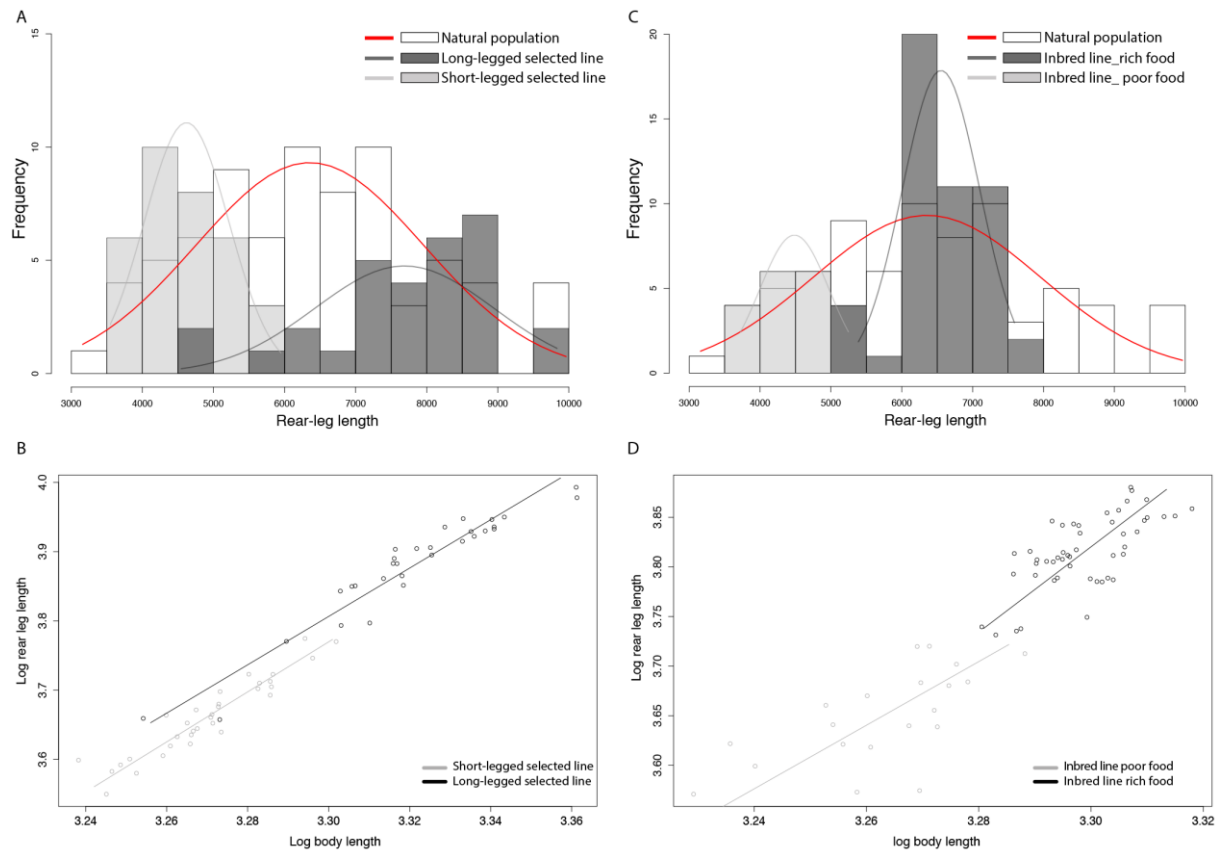


Figure 4: Environmental and genetic contributions to rear leg length variation in males *M. longipes*. A) Rear leg length distributions of adult males from natural population (white) and from two inbred lines that were selected for short (light grey) or long (dark grey) rear legs under rich condition. Normal curves were fitted to each distribution after testing for normality of each condition (Supplementary table 5). B) Scaling relationships of log-transformed data between rear legs and body lengths estimated in males from two inbred lines selected for short (light grey) or long (dark) rear legs under rich condition, using Standardized Major Axis (SMA) regressions. C) Rear leg length distributions of adult males from natural population (white) and from an inbred line that developed under poor (light grey) and rich (dark grey) conditions. Normal curves were fitted to each distribution after testing for normality of each condition (Supplementary table 5). D) Scaling relationships of log-transformed data between rear legs and body lengths estimated in males from an inbred line that developed under poor (light grey) and rich (dark) conditions, using Standardized Major Axis (SMA) regressions.

V. Conclusions

This study illustrates how various ecological factors influence the intensity of sexual selection and ultimately the mechanisms and patterns of phenotypic variation. In the genus *Microvelia*, mating systems are diverse and are likely to influence the diversification of male-specific secondary sexual traits used in pre-mating copulatory strategies. The intense male competition to dominate egg-laying sites in *M. longipes*, unlike other *Microvelia* species, underlies the evolution of exaggerated leg length used as a weapon. Dominating males that intercept and copulate with gravid females on egg-laying sites gain a significant increase in their reproductive fitness by siring the majority of the eggs. This intense selection on increased leg length can, however, be relaxed when egg-laying sites are abundant thus allowing small males to fertilize a significant number of eggs.

We have also shown that plasticity in response to nutritional condition along with genetic variation both contribute to the high phenotypic variation we observed in body and leg length. It is possible that fluctuating selection, combined with phenotypic plasticity, facilitates the dramatic increase and maintenance of phenotypic variation in *M. longipes* compared to other *Microvelia* species. It is also important to note that the fluctuating selection described here (availability of egg-laying floaters) is independent of the individual condition. Therefore, its influence on phenotypic variation cannot be the consequence of a pre-existing increase of condition-dependence, as it would be the case for fluctuating selection based on food resources for example. Altogether, these results point to two ways in which alleles for small male body and leg size will be maintained in the population. First, because small males can sire a significant number of progeny due to possible episodes of relaxed selection. Second, because males with allelic combinations for low trait expression can develop larger body and leg size if they experience higher nutritional condition during development. Therefore, condition dependence causes a non-linear relationship between genotypes and phenotypes, making directional selection less efficient in depleting genetic variation. This fits what Cornwallis and Uller [152] refer to as a “feedback loop between heterogeneity, selection and phenotypic plasticity”.

The findings outlined here open important research avenues to gain a general understanding of how sexual selection can impact phenotypic evolution. *Microvelia longipes*

as a new hemimetabolous insect model with an exaggerated secondary sexual trait offers the opportunity to complete the substantial literature in holometabolous insects such as beetles or various flies [41, 144, 172, 173]. Males of many species of water striders employ water surface ripples as mating calls, and it is unknown whether females can deduce the size of the male from the ripple pattern and whether this would influence female choice [142, 174]. The ease of rearing and the relative short generation time make *Microvelia longipes* a powerful future model to study the extent to which genetic variation and environmental stimuli influence gene expression and ultimately phenotypic variation.

Data accessibility: Data is available from the Dryad Digital Repository at: doi:10.5061/dryad.d4s8357

Authors' contributions: A.K. and W.T. designed research; W.T. performed experiments; A.K. and W.T. analyzed data; and A.K. and W.T. wrote the paper

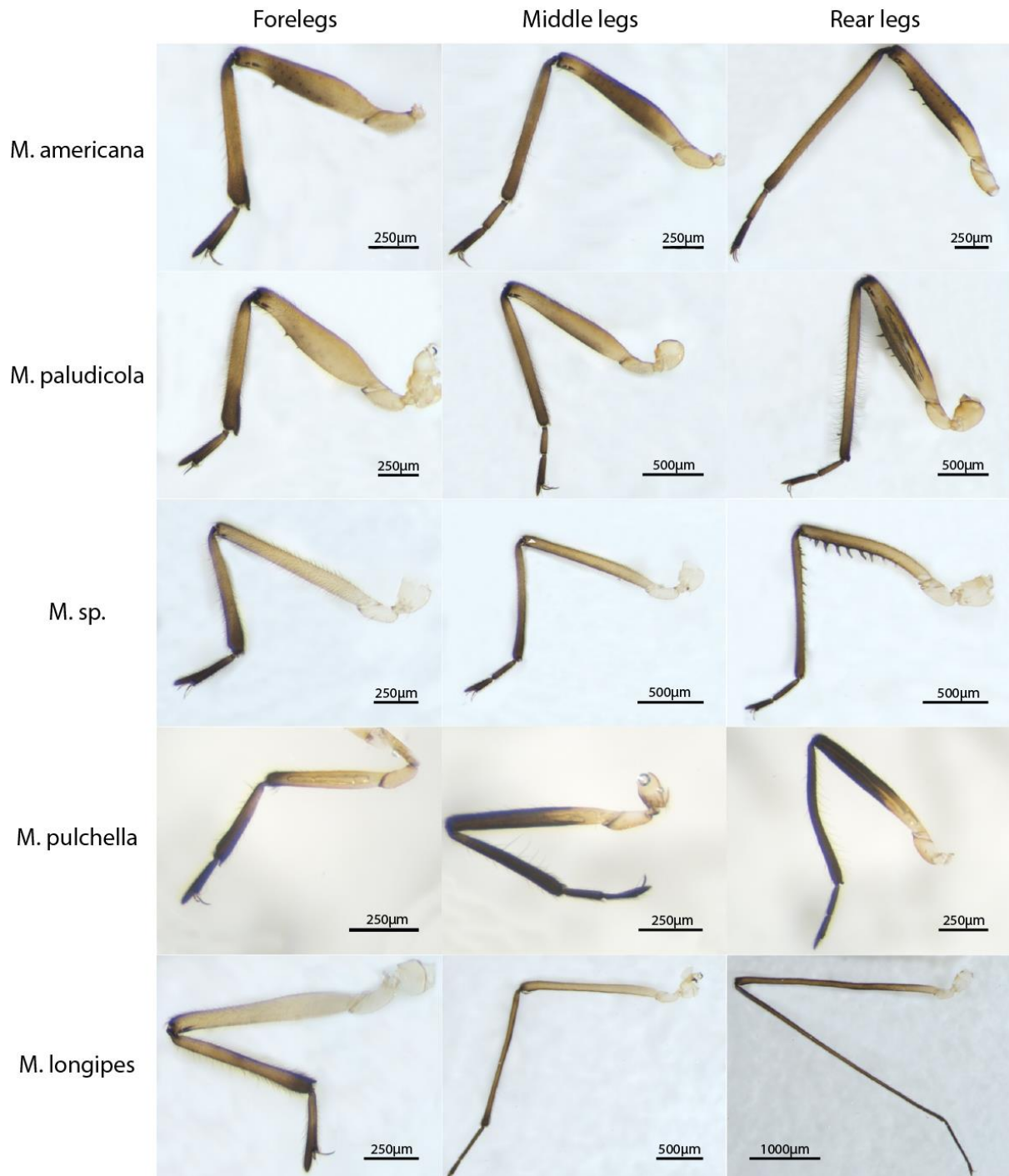
Competing interests: We declare we have no competing interests.

Funding: This work was funded by an ERC-CoG no. 616346 to and Labex SEBA A.K and a PhD fellowship from Ecole Doctorale BMIC de Lyon to W.T.

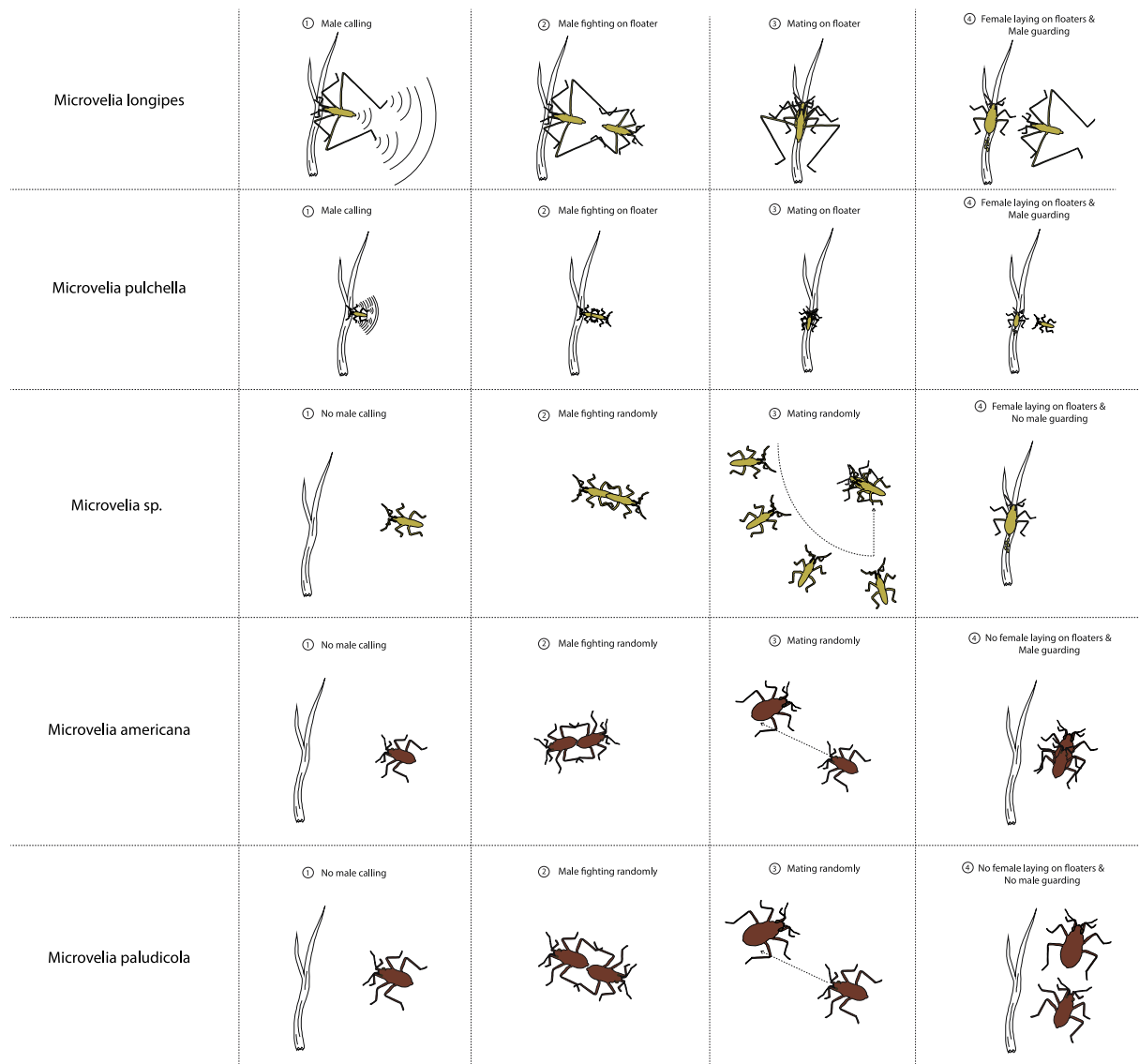
Acknowledgements

We thank Emília Santos and Antonin Crumière for help with collecting bugs, Felipe Moreira for help with species identification, Russell Bondurianski, Locke Rowe, Kevin Parsons, Gaël Yvert, François Leulier, Dali Ma, Augustin Le Bouquin, Amélie Decaras, Cédric Finet, Aidamalia Vargas, Roberto Arbore for helpful discussions and comments on the manuscript, the IGFL Plateforme de Séquençage, Benjamin Gillet, Sandrine Hughes and David Armisen for help with genetic markers identification and Antoine Melet for data measurements. This work was supported by an ERC-CoG# 616346 and labex CEBA to AK, and a PhD fellowship from Ecole Doctorale BMIC de Lyon to W.T.

VI. Supplementary material



Supplementary figure 1: Comparative morphology of the legs across *Microvelia* species sample in figure 1. Note the presence of grasping traits on male legs in *Microvelia americana*, *Microvelia paludicola* and *Microvelia sp.* These traits are absent in *Microvelia longipes* and *Microvelia pulchella*.



Supplementary figure 2: Schematic summary representation of the mating systems in the five *Microvelia* species



Supplementary figure 3: *M. longipes* natural habitat. Top panel: Example of rain-filled puddle in French Guyana in Crique Patate near Cayenne where *M. longipes* population was collected. Middle panel: Zoom on the floating substrates deposited on the water surface of the puddle. Bottom panel: Example of floater full of *M. longipes* eggs. Scale bar represents 5mm.

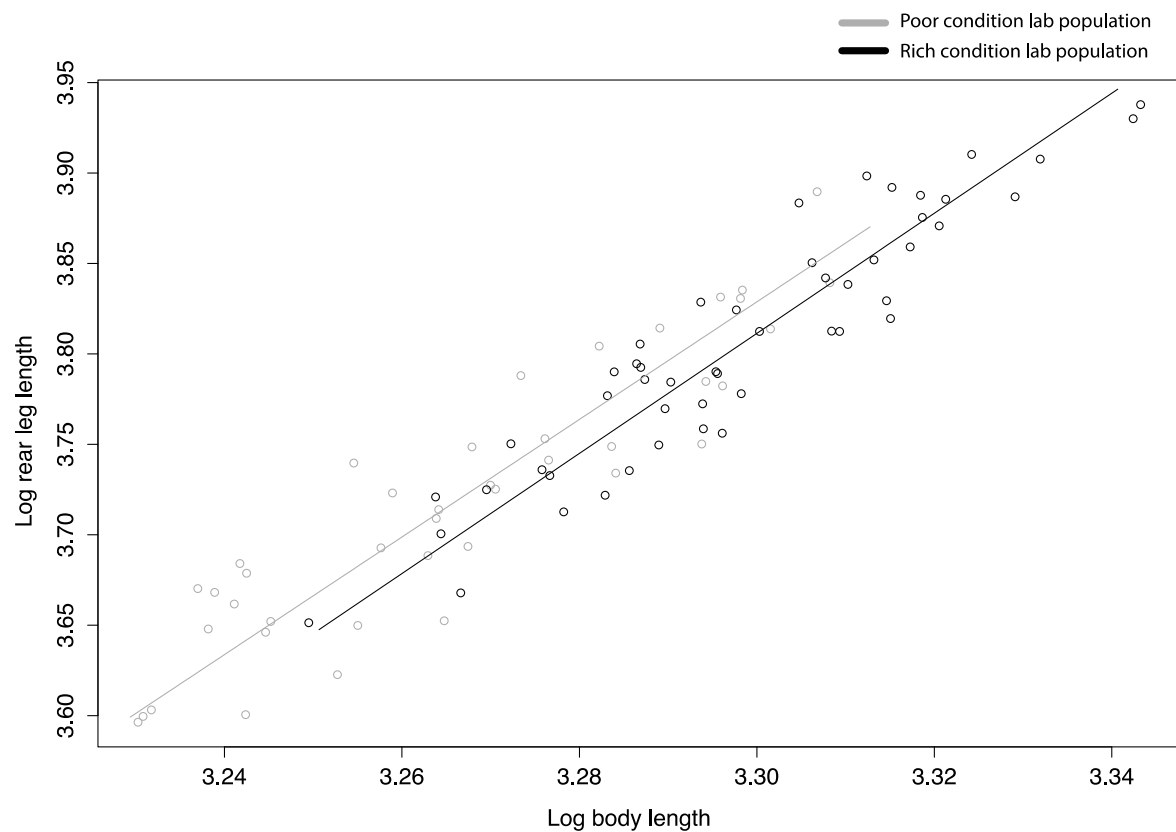
March 25, 2018



March 30, 2018



Supplementary figure 4: Fluctuating *Microvelia longipes* environment. Pictures of a rain-filled puddle in Rio de Janeiro where a *M. longipes* natural population was collected. The puddle dried out entirely in a period of five days.



Supplementary figure 5: Nutrition effects on body-leg scaling relationships in the lab population. Static allometry on log-transformed data between rear leg and body lengths for unselected adult individuals fed on rich (black) and poor (grey) diets.

| Species | 12S ribosomal rRNA gene | 16S ribosomal rRNA gene | 18S ribosomal rRNA gene | 28S ribosomal rRNA gene | Cytochrome b cytochrome b | Cytochrome b cytochrome b | Cytochrome b cytochrome b | Cytochrome b cytochrome b | Distal-less | Doublesex | Gamma-interferon-inducible lysosomal thiol reductase | Sex comb reduced | Ultrabithorax |
|----------------------|-------------------------|-------------------------|-------------------------|-------------------------|---------------------------|---------------------------|---------------------------|---------------------------|-------------|-----------|--|------------------|---------------|
| <i>H. trimalis</i> | MH591924 | MH591930 | MH591936 | MH591942 | KX821858 | KX821873 | KX821888 | KX821903 | KX821989 | MK559406 | KX821961 | KX821948 | KX821933 |
| <i>M. americana</i> | MH591925 | MH591931 | MH591937 | MH591943 | KX821860 | KX821875 | KX821890 | KX821905 | KX821991 | MK559407 | KX821963 | KX821950 | KX821935 |
| <i>M. paludicola</i> | MH591926 | MH591932 | MH591938 | MH591944 | MK256978 | MK256982 | MK256986 | MK256990 | MK256994 | MK256998 | MK257002 | MK257006 | MK257010 |
| <i>M. longipes</i> | MH591927 | MH591933 | MH591939 | MH591945 | MK256979 | MK256983 | MK256987 | MK256991 | MK256995 | MK256999 | MK257003 | MK257007 | MK257011 |
| <i>M. pulchella</i> | MH591928 | MH591934 | MH591940 | MH591946 | MK256980 | MK256984 | MK256988 | MK256992 | MK256996 | MK257000 | MK257004 | MK257008 | MK257012 |
| <i>M. sp.</i> | MH591929 | MH591935 | MH591941 | MH591947 | MK256981 | MK256985 | MK256989 | MK256993 | MK256997 | MK257001 | MK257005 | MK257009 | MK257013 |

Supplementary table 1: Accession numbers of the 14 molecular marker sequences used in the phylogeny.

| Condition | Population size | Replicates | Fights on floaters (FOF) | Fight outside floaters | Total number of fights |
|---|------------------------------|------------|--------------------------|------------------------|------------------------|
| M. pulchella (after 2 days isolation) | N= 50 (25 males, 25 females) | 1 | 16 | 11 | 27 |
| | | 2 | 14 | 11 | 25 |
| | | 3 | 2 | 12 | 14 |
| M. pulchella (after overnight together) | N= 50 (25 males, 25 females) | 1 | 2 | 1 | 3 |
| | | 2 | 1 | 1 | 2 |
| | | 3 | 5 | 4 | 9 |
| M. longipes (after 2 days isolation) | N= 50 (25 males, 25 females) | 1 | 128 | 39 | 167 |
| | | 2 | 168 | 33 | 201 |
| | | 3 | 161 | 29 | 190 |
| M. longipes (after overnight together) | N= 50 (25 males, 25 females) | 1 | 74 | 54 | 128 |
| | | 2 | 170 | 56 | 226 |
| | | 3 | 216 | 25 | 241 |
| M. longipes (after 2 days isolation) | N= 20 (10 males, 10 females) | 1 | 28 | 40 | 68 |
| | | 2 | 66 | 34 | 100 |
| | | 3 | 57 | 10 | 67 |
| M. longipes (after overnight together) | N= 20 (10 males, 10 females) | 1 | 12 | 66 | 78 |
| | | 2 | 27 | 1 | 28 |
| | | 3 | 171 | 21 | 192 |

| | | | | | |
|---|---|--|---|--|--|
| Standardized space conditions (after 2 days isolation) | t-test MP vs ML on floaters t = 10.8265 df = 4 P value= 0.0004 | t-test MP vs ML outsiders t = 7.6354 df = 4 P value= 0.0016 | t-test MP vs ML total t = 15.1834 df = 4 P value= 0.0001 | t-test ML floaters vs outsiders t = 9.3652 df = 4 P value= 0.0007 | t-test MP floaters vs outsiders t = 0.1521 df = 4 P value= 0.8865 |
| | | | | | |
| | | | | | |
| Standardized space conditions (after overnight together) | t-test MP vs ML on floaters t = 3.6004 df = 4 P value= 0.0228 | t-test MP vs ML outsiders t = 4.2716 df = 4 P value= 0.0129 | t-test MP vs ML total t = 5.4555 df = 4 P value= 0.0055 | t-test ML floaters vs outsiders t = 2.5186 df = 4 P value= 0.0655 | t-test MP floaters vs outsiders t = 0.4264 df = 4 P value= 0.6918 |
| | | | | | |
| | | | | | |
| Standardized density conditions (after 2 days isolation) | t-test MP vs ML on floaters t = 3.2328 df = 4 P value= 0.0319 | t-test MP vs ML outsiders t = 1.8173 df = 4 P value= 0.1433 | t-test MP vs ML total t = 4.8705 df = 4 P value= 0.0082 | t-test ML floaters vs outsiders t = 1.5215 df = 4 P value= 0.2028 | |
| | | | | | |
| | | | | | |
| Standardized density conditions (after overnight together) | t-test MP vs ML on floaters t = 1.3281 df = 4 P value= 0.2549 | t-test MP vs ML outsiders t = 1.4201 df = 4 P value= 0.2286 | t-test MP vs ML total t = 1.9487 df = 4 P value= 0.1231 | t-test ML floaters vs outsiders t = 0.7502 df = 4 P value= 0.4948 | |
| | | | | | |
| | | | | | |

Supplementary table 2: Comparison of fight frequency between *M. longipes* and *M. pulchella* males. Summary table of the number of fights in *M. longipes* and *M. pulchella* males in the different conditions for a period of one hour. Below are the associated statistical tests for differences in fight frequency.

| MicrosatelliteRef | MicrosatelliteNumber | SequenceLeftPrimer | SequenceRightPrimer | MicrosatelliteMotif | AmpliconSize(bp) |
|-------------------|----------------------|--------------------------------|-------------------------------|---------------------|------------------|
| 00252:02042 | microsatellite1 | TCT-CTC-TGA-CCA-ATG-GAC-CTC | ACC-CAA-AGA-GAT-GTT-TCA-GAC-C | AGT | 238 |
| 00455:00472 | microsatellite2 | TGT-TTA-CCG-TTG-GAG-TAG-CTC | CGC-GAA-TGA-CTG-TCT-CCA-AC | ACT | 231 |
| 00694:01783 | microsatellite3 | CAC-ACT-GGG-AGA-AGC-GAT-G | TTG-TTG-GGA-TTT-CAA-GAC-CAA-C | GTT | 134 |
| 02107:02724 | microsatellite4 | TTT-GGA-ACT-GGG-AAC-GCT-TC | CAG-GGT-GGC-CTA-CTA-AAC-GG | AGT | 298 |
| 00192:00384 | microsatellite5 | CCA-TCT-CGA-CCA-GGT-TAG-TGT-CC | CAA-CGT-GGA-CTA-CTT-GCC-TG | ACT | 115 |
| 00186:01709 | microsatellite6 | GTT-GCT-ACT-ACC-ACC-CGA-GG | CTA-GGA-GAA-GGC-TAG-GAA-CAA-C | AAT | 265 |
| 00005:01482 | microsatellite7 | AGT-CAT-TGC-GAA-GTG-TCA-GC | AGT-TGG-TTG-GTT-GAC-TTG-TTT-C | ACT | 286 |
| 00014:00678 | microsatellite8 | AAT-GGA-TTG-CAC-GGA-AAC-AC | CTG-AAG-GCC-GTA-GTG-AAT-CC | ACT | 314 |
| 00077:00950 | microsatellite9 | CAT-CCA-CGT-TCT-TAC-CTC-GC | CTA-GTC-GCA-GTG-CAA-ATG-GG | TAA | 174 |
| 00105:00329 | microsatellite10 | ACG-TGA-TTT-CTG-CAT-GTG-GG | GAA-TTT-ACC-CAA-TCC-TGC-CTT-G | AAT | 306 |
| 00036:01342 | microsatellite11 | AGC-CCT-TAC-ACC-CAG-GTA-TG | TTT-ACC-CGA-CCA-CAG-CCT-AG | CATT | 220 |
| 00037:01460 | microsatellite12 | TCT-CGT-TTC-CAT-GAC-ACA-CG | ACG-TCA-TCC-ACA-GCC-ATA-GG | ACTT | 361 |
| 00137:00588 | microsatellite13 | ACG-ACT-TAG-CAG-AGT-GAG-GC | ATG-GTC-CTA-ATT-CTG-CGT-GG | AATC | 198 |
| 01575:02722 | microsatellite14 | GCG-TGG-TCA-GTG-CTA-TAT-GG | TGT-TAG-CCT-AGT-AAG-CCT-TGC | AATG | 168 |
| 01756:00735 | microsatellite15 | ACA-GTT-TCA-AGC-CAT-CAA-CAA-G | AGA-CAA-ATA-CCG-CAG-CTT-TGG | CATT | 215 |
| 00153:02256 | microsatellite16 | ACC-TAT-TTG-CCT-GGC-TTT-GC | CCT-CAG-GTG-GGT-TGC-ATC-C | AATT | 158 |
| 00151:01753 | microsatellite17 | ACA-TGA-CTT-CCT-GAT-GCA-ACG | ATG-CAG-AGT-TCC-CTT-ACT-TGT-G | TTTA | 114 |
| 00336:02418 | microsatellite18 | TGG-CCT-CTA-CCT-CCT-GTA-AAC | CAG-CTA-GGC-ACA-TTC-CAC-TC | TAAA | 271 |
| 00214:02519 | microsatellite19 | CCT-CGG-TGG-TTA-TGC-TGA-TAT-G | GGG-TCA-TGG-AAG-AGG-AGG-G | ATAC | 201 |
| 00380:01139 | microsatellite20 | AGG-AAC-CTG-AGC-GAA-GTC-TC | ATG-GTC-CTA-ATT-CTG-CGT-GG | TTGA | 182 |
| 00368:00570 | microsatellite21 | ACT-TCG-ACT-CAG-TGC-CGC | GTA-GTT-CAC-AGT-GCT-GCG-AC | AAAT | 221 |
| 00144:00750 | microsatellite22 | ATC-CAC-TCG-AGA-TGG-TTC-CC | GTC-TGT-TAT-AAT-CGT-TGC-ACC-C | TTA | 129 |
| 00111:00274 | microsatellite23 | CAT-GAG-TGT-GAA-CGC-AGG-C | TGA-CCT-AAG-AGT-TGA-GCA-CCT-C | TTTA | 280 |
| 00059:02107 | microsatellite24 | TTG-TTG-CTA-GAA-GTT-GCA-GAA-C | GGA-TCC-CAA-CAA-GCA-AGA-AAT-G | ATAC | 185 |
| 00103:01645 | microsatellite25 | TAA-ATT-ACG-GCA-GCG-CAC-C | ATA-ATG-AGA-GGG-ACG-CAG-CC | TCAA | 198 |
| 00136:00891 | microsatellite26 | TCG-GCT-CAG-TCA-GAC-TAT-CC | TGG-GCC-GCA-TTA-TGT-TGA-AG | TATT | 188 |
| 00076:01917 | microsatellite27 | ACC-CGA-GGT-CAC-CAC-AGG | TTT-CCT-TGC-ACT-TTC-TTC-CC | CT | 189 |
| 00064:01180 | microsatellite28 | TAC-CCA-GAC-GTT-TAA-GGG-CC | GCT-TAA-TCT-GTT-GCT-TTA-GGG-C | CTA | 216 |
| 00104:00237 | microsatellite29 | CCC-TTG-CAC-TCC-CTA-TGG-TC | AGT-CGC-AGT-GCA-AAT-GTA-GTC | TTA | 150 |
| 00076:02233 | microsatellite30 | TTG-TTG-CTA-CCA-CCC-GAG-G | GCT-TGA-GAC-AGA-CAG-CAT-GC | CT | 206 |
| 00107:00591 | microsatellite31 | ACC-GAA-CAC-TTC-TTC-CCA-CC | ACC-CTG-ATG-AGC-TAC-AAG-GC | TAA | 157 |
| 00101:00348 | microsatellite32 | TTT-CCC-GCC-ATT-GGT-ATT-GC | TAT-AAT-ACT-GGA-CTT-CGG-CAC-C | TAA | 266 |
| 00076:00552 | microsatellite33 | ACA-CCA-TCG-CAT-TGC-TAT-CTT-C | AGC-AAG-GGT-CAA-CAT-TAG-AGT-G | CT | 148 |
| 00092:00489 | microsatellite34 | TCG-CTG-AGT-CAT-TAT-TGC-GG | TCC-GAA-ACC-GAA-ATG-AAA-CCG | AC | 222 |
| 00092:01561 | microsatellite35 | CGG-GAT-GAA-GCC-GGA-ATT-ATC | AGC-ATA-AAT-ACA-CTG-GGC-GC | TG | 135 |
| 00053:01814 | microsatellite36 | CGA-AAC-CGG-GAT-GTC-AAG-G | ACT-TAG-AGC-TAC-GAT-GAC-AAC-G | GT | 129 |
| 00019:00907 | microsatellite37 | GAC-CTA-TGA-CAC-CCA-CGG-AG | TCG-GAT-TGA-GAT-TAA-GTT-CGC-G | TTA | 351 |
| 00041:01546 | microsatellite38 | ACG-ATC-GGC-ACC-TCT-CTT-AG | TGC-CTT-CCC-TAC-TTT-CTC-ACC | TAA | 188 |
| 00023:02048 | microsatellite39 | TTC-GCA-AAG-TTG-TCT-GTC-TGG | GCT-ACA-TCC-GTG-CCT-GAT-TTC | GT | 219 |
| 00047:02069 | microsatellite40 | GAG-AAG-TTC-CCA-GAT-ACA-CCG | CGA-AGG-ACA-CAG-TTT-GCC-ATC | GT | 261 |

Supplementary table 3: Primer sequences, the full sequence, the motif and the length of each tested microsatellite.

| Species | Sex | Rearleglength | T-testrearleglength | Coefficientsof variationrearleglength | F-testrearleglength | Bodylength | T-testbodylength | Coefficientsof variationbodylength | F-testbodylength | Testforallometry | Likelihoodratio test (allometric coefficient) | Waldstatistic test (intercept) |
|------------------------------|---------|-------------------|--|---------------------------------------|-----------------------------------|--------------------|--|------------------------------------|---------------------------------|-------------------|---|--|
| <i>Microvelia longipes</i> | Males | 6448.66±1588.287 | t=1.1755, df=104.27, p-value=2.2e-16 | 0.25 | F=0.010632, df=104, p-value=0.94 | 2001.913±157.007 | t=2.0991, df=118.91, p-value=0.04 | 0.08 | F=0.39489, df=104, p-value=0.74 | r=0.9801, df=73 | Likelihoodratio test=2.07, df=1, p-value=0.15 | |
| | Females | 2668.63±163.7731 | | 0.06 | | 2050.705±98.66367 | | 0.05 | | | | |
| <i>Microvelia pulchella</i> | Males | 1524.735±162.0358 | t=2.3581, df=56.554, p-value=0.02185 | 0.10 | F=0.12042, df=56, p-value=0.73 | 1363.622±152.8648 | t=6.0103, df=49, p-value=0.871e-08 | 0.11 | F=0.16283, df=56, p-value=0.69 | r=0.9391, df=40 | Likelihoodratio test=1.026, df=1, p-value=0.31 | Waldstatistic=14.1, df=1, p-value=2.2e-16 |
| | Females | 2611.927±56.22933 | | 0.03 | | 1517.7485±56.22933 | | 0.04 | | | | |
| <i>Microvelia Sp1</i> | Males | 2631.544±97.7944 | t=4.188, df=37.988, p-value=2.2e-16 | 0.04 | F=0.0355, df=37, p-value=0.9402 | 2328.971±97.36657 | t=0.69, df=19, p-value=0.50 | 0.04 | F=0.51556, df=37, p-value=0.48 | r=0.2216, df=18 | Likelihoodratio test=7.75, df=1, p-value=0.016255 | |
| | Females | 2188.883±99.51558 | | 0.05 | | 2042.451±69.9115 | t=1.732e-12 | 0.03 | | | | |
| <i>Microvelia americana</i> | Males | 2952.615±104.8273 | t=8.8439, df=47.806, p-value=0.0003571 | 0.04 | F=0.38263, df=47, p-value=0.03235 | 2108.040±78.98845 | t=10.487, df=45.749, p-value=9.392e-14 | 0.04 | F=0.26881, df=47, p-value=0.61 | r=0.08294, df=28 | Likelihoodratio test=35.2, df=1, p-value=2.4498 | Waldstatistic=91.83, df=1, p-value=2.2e-16 |
| | Females | 2860.320±64.84288 | | 0.02 | | 2287.186±40.95299 | | 0.02 | | | | |
| <i>Microvelia paludicola</i> | Males | 3762.755±123.6174 | t=9.061, df=34.125, p-value=2.2e-16 | 0.03 | F=0.49589, df=34, p-value=0.1352 | 2472.490±81.61301 | t=4.9963, df=36.936, p-value=0.6203 | 0.03 | F=0.70982, df=34, p-value=0.19 | r=0.005372, df=18 | Likelihoodratio test=0.0002621, df=1, p-value=0.98708 | Waldstatistic=46.1, df=1, p-value=2.2e-16 |
| | Females | 3118.335±87.05043 | | 0.03 | | 2460.567±68.75955 | | 0.03 | | | | |

Supplementary table 4: Morphometric data and associated statistical tests. Summary table of the adult measurements and statistical tests for all *Microvelia* species.

| Conditions | Shapiro tests | p-values |
|--|---------------|----------|
| <i>M. longipes</i> males natural population | W(2)=0.98768 | 0.685 |
| <i>M. longipes</i> females natural population | W(2)=0.9759 | 0.4508 |
| <i>M. longipes</i> males inbred line selected long legs | W(2)=0.9346 | 0.06522 |
| <i>M. longipes</i> males inbred line selected short legs | W(2)=0.9759 | 0.658 |
| <i>M. longipes</i> males lab stock rich condition | W(2)=0.9784 | 0.4875 |
| <i>M. longipes</i> males lab stock poor condition | W(2)=0.9575 | 0.1197 |
| <i>M. longipes</i> males inbred line rich condition | W(2)=0.9639 | 0.1372 |
| <i>M. longipes</i> males inbred line poor condition | W(2)=0.951 | 0.3833 |

Supplementary table 5: Tests for normal distribution in all *M. longipes* conditions. Values of Shapiro tests and associated p-values for each *M. longipes* adult population reared in different conditions.

| Species | number of spikes | coefficient of variation legs | coefficient of variation bodies | slope | mating outside floaters | mating in floaters | male calling | male fight outside floaters | male fight in floaters | female laying outside floaters | female laying in floaters | male guarding |
|---------|------------------|-------------------------------|---------------------------------|-------|-------------------------|--------------------|--------------|-----------------------------|------------------------|--------------------------------|---------------------------|---------------|
| HT | 0 | | | | | | | | | | | |
| MA | 2 | 0,04 | 0,04 | 0,94 | + | - | - | + | - | + | - | + |
| MCAL | 4 | 0,03 | 0,03 | 0,99 | + | - | - | + | - | + | - | - |
| new_cay | 9 | 0,04 | 0,04 | 0,89 | + | - | - | + | - | + | + | - |
| MP | 0 | 0,1 | 0,11 | 0,9 | + | + | + | + | + | + | + | + |
| ML | 0 | 0,25 | 0,08 | 3,25 | + | + | + | + | + | + | + | + |

| STATISTICS | number of spikes | coefficient of variation legs | coefficient of variation bodies | slope | mating outside floaters | mating in floaters | male calling | male fight outside floaters | male fight in floaters | female laying outside floaters | female laying in floaters | male guarding |
|-------------------------|------------------|-------------------------------|---------------------------------|----------|-------------------------|--------------------|--------------|-----------------------------|------------------------|--------------------------------|---------------------------|---------------|
| Page's lambda value | 6,61E-05 | 0,12 | 0,99 | 6,61E-05 | NA | 0,99 | 0,99 | NA | 0,99 | NA | 0,99 | 6,61E-05 |
| p-value (Page's lambda) | 1 | 0,94 | 0,39 | 1 | NA | 0,3 | 0,3 | NA | 0,3 | NA | 0,051 | 1 |
| D statistic | | | | | NA | -0,48 | -0,56 | NA | -0,54 | NA | -3,2 | 0,67 |
| p-value (D=1) | | | | | NA | 0,1 | 0,1 | NA | 0,1 | NA | 0 | 0,39 |

| Interpretation | no phylogenetic signal | no phylogenetic signal | phylogenetic signal? More sampling required | no phylogenetic signal | | phylogenetic signal? More sampling required | phylogenetic signal? More sampling required | | phylogenetic signal? More sampling required | | phylogenetic signal | no phylogenetic signal |
|----------------|------------------------|------------------------|---|------------------------|--|---|---|--|---|--|---------------------|------------------------|
|----------------|------------------------|------------------------|---|------------------------|--|---|---|--|---|--|---------------------|------------------------|

Supplementary table 6: Phylogenetic signal. Summary table of all tested characters and associated statistical tests of the phylogenetic signals.

| individuals | rear leg length | % winning | groups |
|-------------|-----------------|-----------|--------|
| 1 | 7474,328 | 100 | 1 |
| 2 | 6083,038 | 50 | 1 |
| 3 | 4114,303 | 0 | 1 |
| 4 | 5867,771 | 50 | 2 |
| 5 | 7602,493 | 90 | 2 |
| 6 | 3789,154 | 10 | 2 |
| 7 | 6306,311 | 50 | 3 |
| 8 | 3661,332 | 0 | 3 |
| 9 | 7128,56 | 100 | 3 |
| 10 | 4386,102 | 0 | 4 |
| 11 | 8805,633 | 100 | 4 |
| 12 | 6293,644 | 50 | 4 |
| 13 | 4752,765614 | 0 | 5 |
| 14 | 6019,828478 | 50 | 5 |
| 15 | 8084,976151 | 100 | 5 |

| Generalized linear model | | | | |
|--------------------------|------------|----------------|---------|---------|
| | Estimate | Standard Error | z-value | p-value |
| (Intercept) | -6.7075935 | 3.2245137 | -2.0802 | 0.0375 |
| rear leg length | 0.0011091 | 0.0005201 | 2.133 | 0.0330 |

| Anova (glm) | | | | | |
|-----------------|----|----------|-----------|----|---------------|
| | df | Deviance | Residuals | df | Resid. Dev |
| NULL | | | | 14 | 12.5626 |
| rear leg length | 1 | 11.344 | | 13 | 1.2187 |
| | | | | | p-value (chi) |
| | | | | | 0.000757 |

Supplementary table 7: Summary table of male competition in *M. longipes*. Table indicating the leg lengths, winning success and the group of all males used in the experiment (see methods for more details). Associated statistics are also reported.

| number of eggs | floaters | genotypes | fertilizing male | location | day of experiment |
|----------------|----------|-----------|------------------|------------------|-------------------|
| 126 | 20 | homo | long legs | on floaters | 1 |
| 82 | 20 | homo | long legs | on floaters | 2 |
| 85 | 20 | homo | long legs | on floaters | 3 |
| 79 | 20 | homo | long legs | on floaters | 4 |
| 112 | 3 | homo | long legs | on floaters | 1 |
| 92 | 3 | homo | long legs | on floaters | 2 |
| 87 | 3 | homo | long legs | on floaters | 3 |
| 51 | 3 | homo | long legs | on floaters | 4 |
| 0 | 20 | homo | long legs | outside floaters | 1 |
| 3 | 20 | homo | long legs | outside floaters | 2 |
| 0 | 20 | homo | long legs | outside floaters | 3 |
| 30 | 20 | homo | long legs | outside floaters | 4 |
| 15 | 3 | homo | long legs | outside floaters | 1 |
| 21 | 3 | homo | long legs | outside floaters | 2 |
| 25 | 3 | homo | long legs | outside floaters | 3 |
| 47 | 3 | homo | long legs | outside floaters | 4 |
| 38 | 20 | hetero | short legs | on floaters | 1 |
| 12 | 20 | hetero | short legs | on floaters | 2 |
| 27 | 20 | hetero | short legs | on floaters | 3 |
| 19 | 20 | hetero | short legs | on floaters | 4 |
| 0 | 3 | hetero | short legs | on floaters | 1 |
| 3 | 3 | hetero | short legs | on floaters | 2 |
| 9 | 3 | hetero | short legs | on floaters | 3 |
| 3 | 3 | hetero | short legs | on floaters | 4 |
| 0 | 20 | hetero | short legs | outside floaters | 1 |
| 0 | 20 | hetero | short legs | outside floaters | 2 |
| 1 | 20 | hetero | short legs | outside floaters | 3 |
| 10 | 20 | hetero | short legs | outside floaters | 4 |
| 0 | 3 | hetero | short legs | outside floaters | 1 |
| 0 | 3 | hetero | short legs | outside floaters | 2 |
| 12 | 3 | hetero | short legs | outside floaters | 3 |
| 23 | 3 | hetero | short legs | outside floaters | 4 |

Supplementary table 8: Counts and genotypes of the total number of eggs laid by females in each condition. A second excel sheet reports the tables of the summary statistics.

| Condition | T-tests leg length | Allometric equation | Likelihood ratio test (allometric coefficient) | Wald statistic test (intercept) |
|--|--|---------------------|--|---|
| <i>Microvelia longipes</i> (selected long legs) | t-test: 22.2091, df=5.266, P-value=2.2e-16 | y=3.49x-7.71 | Likelihood ratio statistic=0.1699, df=1, P-value=0.68024 | Wald statistic: 16.52, df=1, P-value=0.8202e-05 |
| <i>Microvelia longipes</i> (selected short legs) | | y=3.61x-8.15 | | |
| <i>Microvelia longipes</i> (selected long legs) rich condition | t=5.374, df=9.232, P-value=3.2e-16 | y=4.3x-10.37 | Likelihood ratio statistic=1.932, df=1, P-value=0.16459 | Wald statistic: 1.69, df=1, P-value=0.19363 |
| <i>Microvelia longipes</i> (selected long legs) poor condition | | y=3.2x-6.81 | | |
| <i>Microvelia longipes</i> (laboratory unselected population) rich condition | t=5.5937, df=9.303, P-value=2.402e-07 | y=3.32x-7.14 | Likelihood ratio statistic=0.0723, df=1, P-value=0.78802 | Wald statistic: 7.214, df=1, P-value=0.0072332 |
| <i>Microvelia longipes</i> (laboratory unselected population) poor condition | | y=3.25x-6.89 | | |

Supplementary table 9: Statistical tests associated with differences in leg length and scaling relationships between artificially selected and nutritionally manipulated *M. longipes* populations.

| PCR protocol microsatellite identifications | | | PCR protocol single individual template | | |
|---|--------------|----------------------|---|--------------|----------------------|
| Reactants | Volumes (μL) | PCR program (cycles) | Reactants | Volumes (μL) | PCR program (cycles) |
| 10X buffer (Invitrogen) | 2.5 | 95°C for 30 min | 10X buffer (Invitrogen) | 2.5 | 95°C for 30 min |
| dNTPs (10mM) | 1 | 95°C for 30 sec | dNTPs (10mM) | 1 | 95°C for 30 sec |
| MgCl ₂ (50mM) (Invitrogen) | 1.5 | 55°C for 5 sec | MgCl ₂ (50mM) (Invitrogen) | 1.5 | 55°C for 5 sec |
| Taq DNA polymerase (5U/μL) (Invitrogen) | 0.25 | 72°C for 30 min | Taq DNA polymerase (5U/μL) (Invitrogen) | 0.25 | 72°C for 30 min |
| Left primer (10μM) | 1 | 72°C for 30 min | Left primer (2.5μM) | 2 | 72°C for 30 min |
| Right primer (10μM) | 1 | 4°C | Right primer (2.5μM) | 2 | 4°C |
| Template genomic DNA (25ng/μL) | 0.15 | | Template DNA extraction from floor tile | 1 | |
| millipore water | 17.6 | | millipore water | 14.75 | |
| Total volume | 25 | | Total volume | 25 | |

Supplementary table 10: Table PCR protocols for microsatellite amplifications and single individual genotyping.

Supplementary videos' caption

All supplementary videos are available on the Royal Society website: <https://royalsocietypublishing.org/doi/suppl/10.1098/rspb.2018.2400>.

Supplementary video 1: *M. longipes* male vibrations in slow motion.

Supplementary video 2: *M. longipes* mating system.

Supplementary video 3: *M. pulchella* mating system.

Supplementary video 4: *M. sp.* mating system.

Chapter 2: The genome of the water strider *Microvelia longipes* provides insights into the role of sexual selection on gene expression and genome architecture.

Authors: William Toubiana¹, David Armisén², Corentin Dechaud¹, Roberto Arbore³ and Abderrahman Khila¹

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 INRA-Université Clermont Auvergne, Clermont-Ferrand, France.

³ Instituto Gulbenkian de Ciência, Oeiras, Portugal.

Keywords: Sexual Selection, Sexual Dimorphism, Weaponry, RNA-seq, Genome Architecture

I. Abstract

How sexual selection promotes phenotypic divergence between individuals of the same species, sharing almost the same genome, is a longstanding question in evolutionary genetics. Exaggerated sexually-selected traits are unique examples to study intraspecies variation driven by sexual selection, as they display dramatic differences both within and between sexes. Much work has been conducted on the developmental genetics of these growth-related sexual traits, but relatively little is known regarding the genomic regulation underlying their exaggeration. The water strider *Microvelia longipes* is a new model system to study exaggerated sexually-selected traits. The males of this species have evolved exaggerated and hypervariable rear legs with an elevated allometric slope during development. To study the genomic basis of this exaggerated phenotype, we generated a high-quality genome of *M. longipes* and conducted a comparative transcriptomic analysis between pairs of legs with different degrees of exaggeration within and between sexes. We found that the development of male exaggerated legs was associated with specific expression patterns in both sex- and leg-biased genes. Furthermore, gene distribution along the genome revealed genomic regions with an enrichment of sex-biased genes specific to the exaggerated legs. Overall, our study offers one of the most complete genomic characterizations of an exaggerated sexual trait and sheds light on the regulatory mechanisms underlying its evolution.

II. Introduction

Phenotypic difference between males and females of the same species, also defined as sexual dimorphism, is one of the most common sources of phenotypic variation in nature and has fascinated biologists for centuries [5, 175]. Sexual selection is known to be a major driver of sexual dimorphism, influencing the evolution and maintenance of various sexual characters in a population [19]. Yet, how this process takes place at the genomic level, where the DNA sequence is almost identical between males and females, is still an important challenge in evolution [176]. Differences in gene expression have emerged as a common mechanism to explain phenotypic differences among individuals sharing almost the same genome [102, 106]. In the last decade, a large number of studies have characterized genes with sex-biased expression in a variety of species, leading to an emerging framework on how selective pressures shape genome evolution in the context of sexual dimorphism [102, 106, 177]. However, these studies have mostly focused on adult or whole body transcriptomic dataset, which remain unsuited to understand the developmental elaboration of sexual characters [106]. Insects represent a good model to study the regulatory mechanisms involved in the formation of sexual dimorphism as they develop relatively fast and their development can easily be synchronized in lab conditions.

Several insect species have notably evolved some extreme cases of sexual dimorphism by which males of some species develop such drastic phenotypes that they appear exaggerated compared to homologous traits in the other sex or in other body parts [35, 42, 79, 149]. These growth-related sexual traits have received lots of attention in developmental genetics but still lack a general understanding of the genomic regulation underlying their exaggeration [35, 41, 43, 77, 83, 87, 88, 92, 95, 122, 178-182] (but see [119-121, 123-125, 183, 184]). The recent advances in sequencing technologies allowed us to fill this gap by integrating genomic studies into ecologically relevant model systems and test for different models of evolutionary theory [80, 134]. This approach has been particularly employed for adaptive traits under natural selection but it is still at its infancy in the field of sexual selection [80, 185, 186]. Early genomic studies in mouse and flies and more recent works in the ruffs and zebra finches suggest, nonetheless, that sexual selection may drastically influence genome architecture [114, 117, 157, 158, 187-191].

We aimed here to assess the role of sexual selection on the regulation of *Microvelia longipes* genome, an emerging model in the field of sexual selection and trait exaggeration [192]. *M. longipes* is a hemimetabolous insect that displays a striking case of sex-specific exaggerated trait where some males have evolved extremely elongated rear legs compared to females. This sex-specific character is also extremely variable between males of the same population, ranging from very short to extremely long-legged males. We have identified that rear legs (L3) in males are used as weapons to kick opponents of the same sex away from the sites where females lay eggs. During these fights males with longer rear legs have more chance to win fights and therefore dominate egg-laying sites than males with shorter legs [192]. This type of directional selection is often associated with the evolution of hyperallometry that corresponds to the disproportional growth of the trait compared to the rest of the body [42, 149]. *M. longipes* males display the same feature, explaining both the hypervariability and the extreme values of some male rear legs compared to homologous traits in females or other *Microvelia* species [192]. However, how and when such extreme growth takes place during development is unknown. We took advantage of the hemimetabolous insect development, progressing through successive molts, to characterize in details the developmental elaboration of an exaggerated sexually-selected trait. We found that the leg exaggeration observed in adult *M. longipes* males was mostly established at the end of the nymphal development (5th nymphal instar) by a drastic increase in growth rate that was minimized in other non-sexually selected appendages.

To characterize the influence of such intense sexual selection on genome regulation, we first built a high-quality genome of *M. longipes*, with chromosome-scale resolution. A transcriptomic approach was also performed, comparing the expression of 5th nymphal instar males and females for the three pairs of legs. We chose to compare the three legs, as they are the most obvious homologous traits with distinct allometric slopes, underlying different types and degrees of sexual selection [192]. Combined, our approach first identified some signatures of trait exaggeration on gene expression patterns. Second, we shed light on chromosomes and genomic regions associated with the directional sexual selection applying to male exaggerated legs in *M. longipes*.

III. Material and methods

A. Population sampling and culture

Microvelia longipes population was collected during fieldwork in French Guyana in Crique Patate near Cayenne. The bugs were maintained at 25°C and 50% humidity and fed on crickets. To monitor leg development, we isolated first nymphal instars and recovered the molts of each individual until adulthood. Molts and adults of each individual were measured to construct growth curves. Inbred populations were generated as described in [192].

B. Measurement and statistics

Male and female appendages were measured with a SteREO Discovery V12 (Zeiss) using the Zen software. All statistical analyses were performed in RStudio 0.99.486. Growth curves were generated with the raw data whereas log-transformed data were used for ontogenetic allometry comparisons. Differences in ontogenetic allometries were estimated using generalized linear models and L1 as control variable and proxy of individual size.

C. Sample collection, assembly and annotation of the *M. longipes* genome

Hundreds of individuals (males and females mixed) were collected from three independent inbred populations and frozen in liquid nitrogen before DNA extraction. Genomic DNA was extracted and purified using the Genomic-tip 20/G DNA extraction kit from Qiagen. Genome sequencing was performed with the help of the Beijing Genomics Institute and the Dovetail Genomics Company. Supplementary table 1 summarizes the sequencing strategy employed. The genome assembly was conducted at the Beijing Genomics Institute and the Dovetail Genomics Company.

The genome sequence was polished using Illumina libraries (Supplementary table 1) and Pilon [193]. Three different automatic annotation strategies, namely Braker, Maker and StringTie were tested to annotate the genome [194-196]. These annotations were based on the leg transcriptomic dataset generated in this study (36 samples in total) as well as a transcriptome from whole body individuals collected at all developmental stages (1 sample) and a transcriptome from a third inbred population not mentioned in this study (18 samples). Braker and Maker pipelines also performed *de novo* automatic annotations.

Maker and Stringtie annotations yielded lower BUSCO quality and manual quality assessment using JBrowse revealed a relatively high number of gene fragmentations that were poorly supported by the alignments (data not shown). We therefore used Braker annotation for further analyses.

For Braker annotation, we used Hisat2 alignment files from each transcriptomic sample to train Augustus with UTR option. Final annotation includes 26,130 genes and 27,553 transcripts.

D. Sample collection and preparation RNA-sequencing

We collected leg tissues from male and female 5th nymphal instars (two days after molting with a time window of six hours) that belonged to two inbred populations that differ in average size (see [192]). All individuals were raised in the same laboratory condition and fed with nine fresh crickets daily until the 5th instar. Individuals from the same inbred population were raised in the same bucket. The three replicates of each condition (lines, sexes and legs) correspond to a pool of 20 individuals chose randomly (Supplementary figure 1). The dissection of the three pairs of legs, dissociated from the thorax, was performed in RNAlater (Sigma) using fine needles; each pair of legs was incubated immediately on ice, in tubes filled with TRIzol (Invitrogen). RNA extractions were performed according to manufacturer protocol. The concentrations were assessed using the Qubit 2.0 Fluorometer (Invitrogen). Quality of RNA samples, library construction and sequencing were performed by Beijing Genomics Institute. The samples were sequenced using HiseqXten sequencing technology with 50 million reads per sample and a paired-end read length of 150bp (Supplementary table 7).

E. Transcriptome assembly, mapping and normalization

Read quality was assessed with FASTQC version 0.10.1 (<http://www.bioinformatics.babraham.ac.uk/projects/download.html>), and trimmed with TRIMMO-MATIC version 0.32. Specifically, reads were trimmed if the sliding window average Phred score over four bases was <15 and only reads with a minimum length of 36bp were kept. Braker annotation was used as reference for read alignment and the transcriptome quantification. We obtained around 90% alignment rate on the genome and about 72% of uniquely mapped reads using Hisat2 method (Supplementary table 7) [197]. The latter

condition was used for the estimation of transcript abundances and the creation of count tables (raw counts, FPKM and TPM tables) were performed using the StringTie pipeline [196]. The abundance of reads per gene was finally calculated by adding the read counts of each predicted transcript isoforms.

F. Comparative transcriptomics: analyses of variance

Initially the transcriptomic approach was performed on three levels of comparisons; namely the lines, the sexes and the legs (Supplementary figure 1). The first three axes of variation in gene expression explained 57.1% of the total variation and separated the two inbred populations (Supplementary figure 2). This confirms the genetic similarity that exists between individuals of the same inbred population. In order to correctly assess the influence of sex and leg comparisons on gene expression variance, we corrected for the line effect using a Within-Class Analysis [198]. After correction, the first major axis of variation separated male and female conditions, while PC3 explained the variation between legs (Supplementary figure 2).

We also performed an analysis of variance on sex separately (Figure 3A). Here we identified both line and replicate effects. The latter effect matched the days where RNA was extracted from each sample. We corrected for both effects, using Within Class Analysis, in subsequent analyses (data not shown).

G. Sex-biased and leg-biased expression

The number of reads per “gene” was used to identify differences in expression among the different conditions of interest using DESeq2 [199]. We called differentially expressed genes any gene with a fold-change > 1.5 and a $P_{adj} < 0.05$. Other differentially expressed genes that do not fit such criteria (e.g. fold-change > 0 and a $P_{adj} < 0.05$) are specified in the result section. All differential expression analyses were performed on the two lines combined as we aimed to identify genes involved in allometric slope, which a common feature of both lines. We repeated the differential expression analyses in lines separately and generally found high expression similarities between the two lines (data not shown).

Sex biased expression across all leg tissues

We first filtered for lowly expressed transcripts by removing all transcripts for which the expression was lower than 2 FPKM in more than two-thirds of the samples (36 samples total). This filtering process leaves 9364 transcripts for the expression analysis. We performed the differential expression between sexes and corrected for the line effect.

Sex-biased expression in separated leg tissues

We first filtered transcripts for which expression was lower than 2 FPKM in more than half of the samples after combining the two inbred populations (12 samples total). Transcripts with average expression that was lower than 2 FPKM in both males and females were also discarded. Finally we ran the differential expression analysis by taking into account the line effect.

Leg-biased expression

To identify the leg-biased genes we used the same filtering process as for the identification of the sex-biased genes in the separated leg tissues. Male and female conditions were simply replaced by the different pairs of legs (e.g. L3 versus L1). We ran the differential expression analysis by applying a line and replicate correction.

Interaction model between legs and sexes with DESeq2

In order to detect a possible interaction between leg and sex regulations, we also used the interaction model implemented in DESeq2. In this analysis we used the same filtering process for low expressed genes as the one used to identify sex-biased genes across all legs. The interaction model between legs and sexes revealed 2 genes for the L3-L1 comparison and no gene for the L3-L2 and L2-L1 comparisons. When we looked at these two genes in the differential expression analyses without the interaction effect, we found that one of them (g7203) was detected as male-biased in L3 but not leg-biased whereas the second gene (g23967) was both male- and L3-biased.

Hierarchical clustering

Average expressions of sex- and leg-biased genes in the different tissues were clustered using Euclidean clustering in the R package PVCLUST version 1.3-2 [200] with 1000 bootstrap

resampling. Heatmaps and clustering were performed using the $\log_2(\text{TPM})$ average expression of each gene from each tissue. Heatmaps were generated using the R package GPLOTS version 3.0.1.1.

H. Gene Ontology analysis

Gene names and functions were annotated by sequence similarity against the NCBI ‘non redundant’ protein database using Blast2GO. The Blast2GO annotation was then provided to detect Gene Ontology terms enrichment ($p\text{-value} < 0.05$) using the default method of TopGO R package version 2.34.0.

I. Sex-biased gene distribution

Sex chromosome identification

The sex determination system in Gerromorpha is genetically determined and established by either the XX/XY or XX/X0 sex determination chromosomes [201, 202]. In *M. longipes*, Illumina genomic sequencing containing only males was used to align genomic reads against *M. longipes* genome and extract the genomic coverage of each scaffold. The scaffold 1893 was the only scaffold among the 13 biggest scaffolds (more than 90% of the genome) that was presenting twice less coverage than the other scaffolds. To finally assess the identity of the X chromosome in *M. longipes*, we monitored the gene expression and found that the scaffold 1893 included both male- and female-biased genes, excluding this scaffold to be the Y chromosome. We also looked for a possible Y chromosome by identifying scaffolds with similar genomic coverage as the X chromosome but containing genes with only male-biased expression. We did not find any among the fifty largest scaffolds, suggesting that *M. longipes* has a XX/X0 sex determination system or presents a highly degraded Y chromosome.

Genomic distribution of sex-biased genes

We identified the genomic location of each gene and selected genes with a fold change superior to 1.5 between males and females as sex-biased genes ($P_{adj} < 0.05$). Over- or under-representation of sex-biased genes in the X chromosome (scaffold 1893) compared to the autosomes (12 other largest scaffolds) was tested using Fisher’s exact tests.

Estimation of dosage compensation

To compare the average level of gene expression between males and females in the thirteen largest scaffolds we first selected expressed genes with FPKM > 2 in at least half of the samples. We also averaged the gene expressions between replicates and lines before testing for differences in expression (Wilcoxon tests on the $\log_2(\text{FPKM})$). For the analysis on all legs combined we averaged, in addition, the expression between legs.

J. Detection of large sex-biased gene clusters

To detect large chromosomal regions enriched in sex-biased genes we developed a bootstrapping method based on sliding windows of 2 Mb with a step size of 100 kb (Supplementary figure 3). Gene density calculation revealed that on average, genes are found every 20 kb in *M. longipes* genome. This pattern was homogeneous among chromosomes (Supplementary table 2). We therefore split each chromosome into bins of 100 kb and generated sliding windows of 2 Mb (20 bins) to include approximately 100 genes per window in the analysis (Supplementary figure 3B, Supplementary table 2). We used two scaffolds, one scaffold with two enriched regions (scaffold 2) and a scaffold with no enriched region (scaffold 1914), to repeat the analysis on smaller regions (1 Mb, 500 kb, 250 kb and 120 kb). We found similar results in both scaffolds, regardless of the size of the region, indicating that our analysis is statistically robust and is not missing information (data not shown).

Fold-change reassignment and gene position

From the DESeq2 analyses, all expressed genes were associated with a \log_2 fold change (Log2FC) and a p-value (P_{adj}). Unexpressed genes (FPKM < 2) were assigned a log2FC of 0 and a p-value of 1. Among the expressed genes, we switched the log2FC to 0 for the unbiased genes ($P_{adj} > 0.05$), in order to directly assess sex-biased genes based on log2FC values (Supplementary figure 3A).

In a second step, we merged the dataset on sex-biased expression with the gene positions (Supplementary figure 3A).

Genome-wide detection of sex-biased gene clusters

A mean log2FC was calculated for each window and reported along the chromosomes to reveal genome-wide clusters of sex-biased genes (Supplementary figure 3B).

Bootstrapping method

To test whether these clusters are significantly enriched in male or female-biased genes, we developed a bootstrap approach (Supplementary figure 3C). As the mean expression level of a gene influences the log2FC value (*i.e.* genes with low expressions are more likely to have high log2FC values and genes with high expression are more likely to be differentially expressed), we first decided to create 5 categories of genes, based on their expression levels (baseMean values from DESeq2 tables). We then reassigned randomly, within each category, the log2FC at each gene position in the genome. This step was performed 100 000 times, therefore generating 100 000 random log2FC profiles.

Finally, to test for the significant enrichment of gene clusters, we compared for each bin the observed log2FC values with the log2FC values generated from the bootstrap. To call for significantly enriched cluster of sex-biased genes, we identified regions for which the observed log2FC value was higher (male-biased) or lower (female-biased) than the 97500 (out of 100 000) bootstrap values generated randomly (Supplementary figure 3D). We finally applied a Bonferroni correction, correcting the bootstrap values by the total number of independent windows in the genome ($n=300$), leading to a bootstrap threshold of 99992.

K. Detection of clusters of consecutive sex-biased genes

This analysis was primarily inspired from Boutanaev *et al.* [188]. In short, we determined clusters by ordering genes along the genome and detecting regions of consecutive male- or female-biased genes ($P_{adj} < 0.05$). To avoid identifying clusters overlapping two different chromosomes, we performed this analysis on the thirteen largest scaffolds separately. We then test whether the observed distribution of genes differed from a stochastic distribution by randomly assigning a genomic position to unbiased, male-biased and female-biased genes respectively. The proportion of sex-biased genes found in clusters as well as the distribution of cluster sizes was calculated by averaging 1000 iterations (Supplementary figure 4). P-values were extracted from the 95% fluctuation intervals calculated from the 1000 randomized iterations.

IV. Results

A. Assembly and annotation of the *M. longipes* genome

To understand the molecular mechanisms underlying the males-specific extreme growth of the legs, we generated a *de novo* genome using inbred lines of a French Guiana *M. longipes* population [192]. Genome assembly combined multiple mate-pair Illumina libraries, long read PacBio libraries, and Dovetail Hi-C/Hi-Rise libraries (Table 1; Supplementary table 1; see Material and Methods) [203-205]. The final genome assembly generated a chromosome-length scaffolds with scaffold N50=54.155 Mb and contig N50= 216.72 kb (Table 1). The twelve longest scaffolds recapitulating 90% of the genome length (Table 1; see material and methods). We sequenced the genome of *M. longipes* using next generation sequencing methods and k-mer frequency distribution in raw sequencing reads yielded estimates of genome size to be 668.14 Mb (Table 1; see Material and Methods). The gene set of *M. longipes* was built by automatic genome annotation using *de novo* and transcriptome-based gene models and predicted 26,130 genes and 27,553 transcripts (see Material and methods). BUSCO analysis revealed 92% of completeness with few duplicated genes (4%) or missing data (2%) from the 2018 insect dataset (Supplementary figure 5) [206]. Overall, the use of inbred populations associated with the construction and assembly of multiple genomic libraries have produced a complete and well assembled genome with contig and scaffold sizes larger than most available insect genomes [207].

B. Male rear-legs experience a burst of growth during the last nymphal instars.

In *M. longipes*, rear-leg length is variable between males and is higher in males compared to females [192]. The developmental stage where these differences are established is unknown. To better characterize the development of this extreme phenotype, we followed the leg growth of developing males and females over their successive nymphal molts until adulthood. We identified that male leg exaggeration was mostly established at the end of the nymphal development, mostly at the fifth nymphal instar (Figure 1A). The growth rates of the other male and female legs were not as extreme, following the degree of exaggeration observed in adults (Figure 1A-B). Ontogenetic allometry comparisons also confirmed the exaggerated growth rate of L3 males during development (Supplementary table 3).

| Genome metrics | Value |
|------------------|---------------|
| Genome size | 668.14 Mb |
| Contig L50/N50 | 216.72 kb |
| Scaffold L50/N50 | 5 (54.155 Mb) |
| Scaffold L90/N90 | 12 (7.927 Mb) |
| # scaffolds | 1935 |
| % genome gaps | 1.26% |

Table 1: *Microvelia longipes* genome metrics

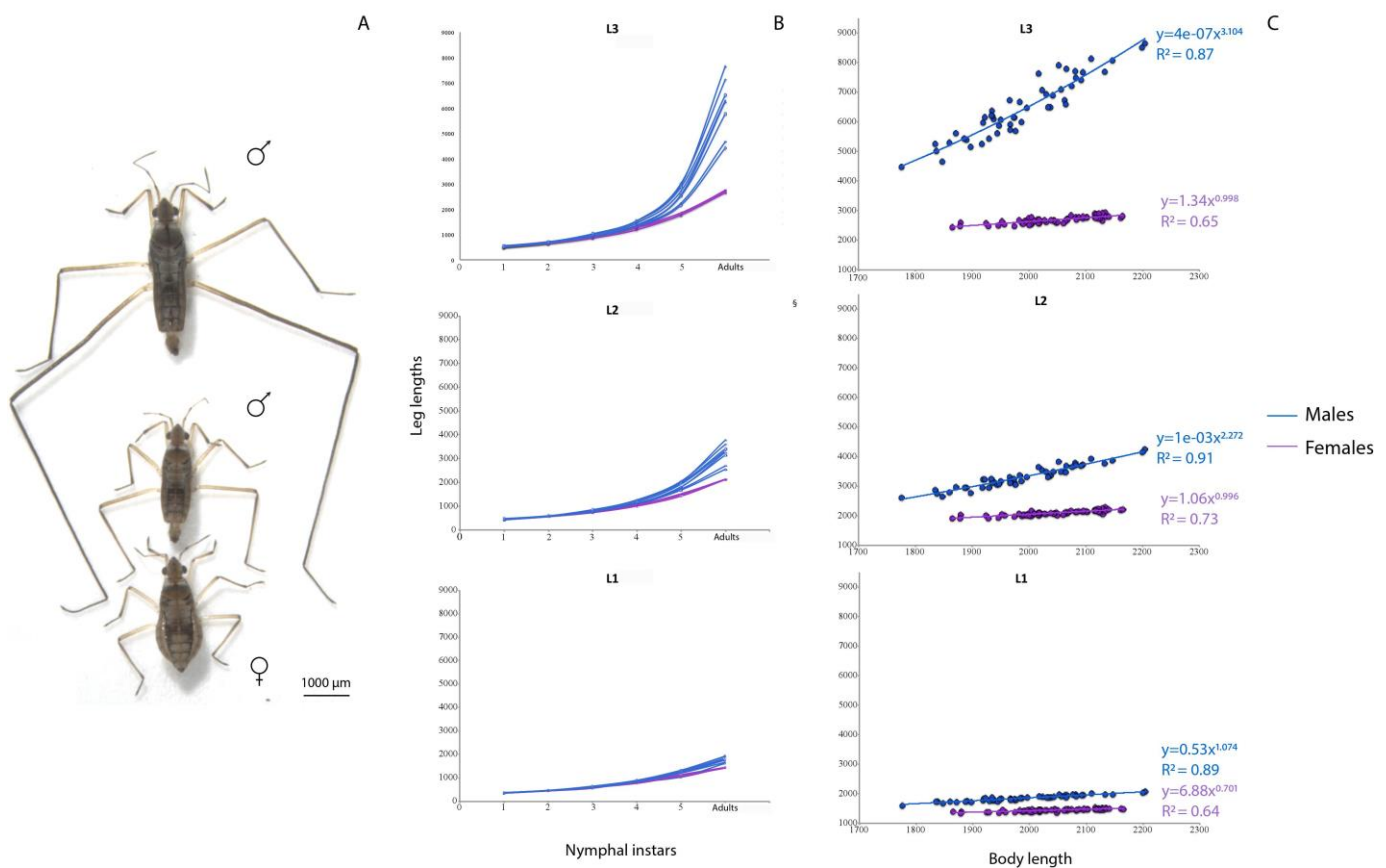


Figure 1: Association between developmental growth and static allometries in *M. longipes*. A) Pictures of extreme male phenotypes and a representative female from *M. longipes* population. B) Leg growth rates in *M. longipes* males and females over the nymphal development until adulthood. Curves were fitted from polynomial regressions. C) Static allometries of the three pairs of legs in adult *M. longipes* males and females. Allometric equations and R^2 values were extracted from a power regression model. Leg and body length values are in micrometers.

Finally, the hypervariability observed in L3 length between adult males of the same population, could also be explained by the differences in growth at the end of the nymphal development (Figure 1A-B). This difference is not associated with variation in developmental time to reach adulthood (ANCOVA, F-value= 1.965, p-value= 0.1712), which confirms a variation in growth rate between short and long-legged males (Supplementary figure 6). Overall, we demonstrate that male leg exaggeration is established at the end of the nymphal development through a drastic burst in growth compared to the other legs, sex or even other males from the same population. Combined, our results strongly suggest that L3 exaggeration in males results from developmental regulatory mechanisms that are, at least to some extent, independent from environmental factors such as individual condition.

C. Variation in gene expression explains differences in exaggeration between legs and sexes.

To understand the developmental and genetic mechanisms underlying leg exaggeration in L3 males, we performed a comparative transcriptomic approach in male and female legs at the fifth nymphal instar – the developmental stage where we observed the burst of growth (Figure 1B). This experiment used two lines artificially selected for large and small males respectively (Supplementary figure 1; see material and methods). A first principal component analysis separated primarily the samples based on line (Supplementary figure 2A). A multivariate analysis, after correcting for line effect, identified the first axis of variation in gene expression (PC1) explaining the differences between male and female conditions (28.3% of variation explained) (Supplementary figure 2B). Differences between legs were explained by the third principal component (PC3) and contributed to 10.3% of the total variation in gene expression (Supplementary figure 2B). Overall, we could identify that the main variations of gene expression in our dataset were associated with the phenotypic differences observed between lines, sexes and legs. Further, we focused on determining differences in gene expression between the three pairs of legs of the same sex and between the sexes, corresponding to the two major effects underlying differences in allometric slope and ultimately trait exaggeration (Figure 1; Supplementary figure 1; but see material and methods).

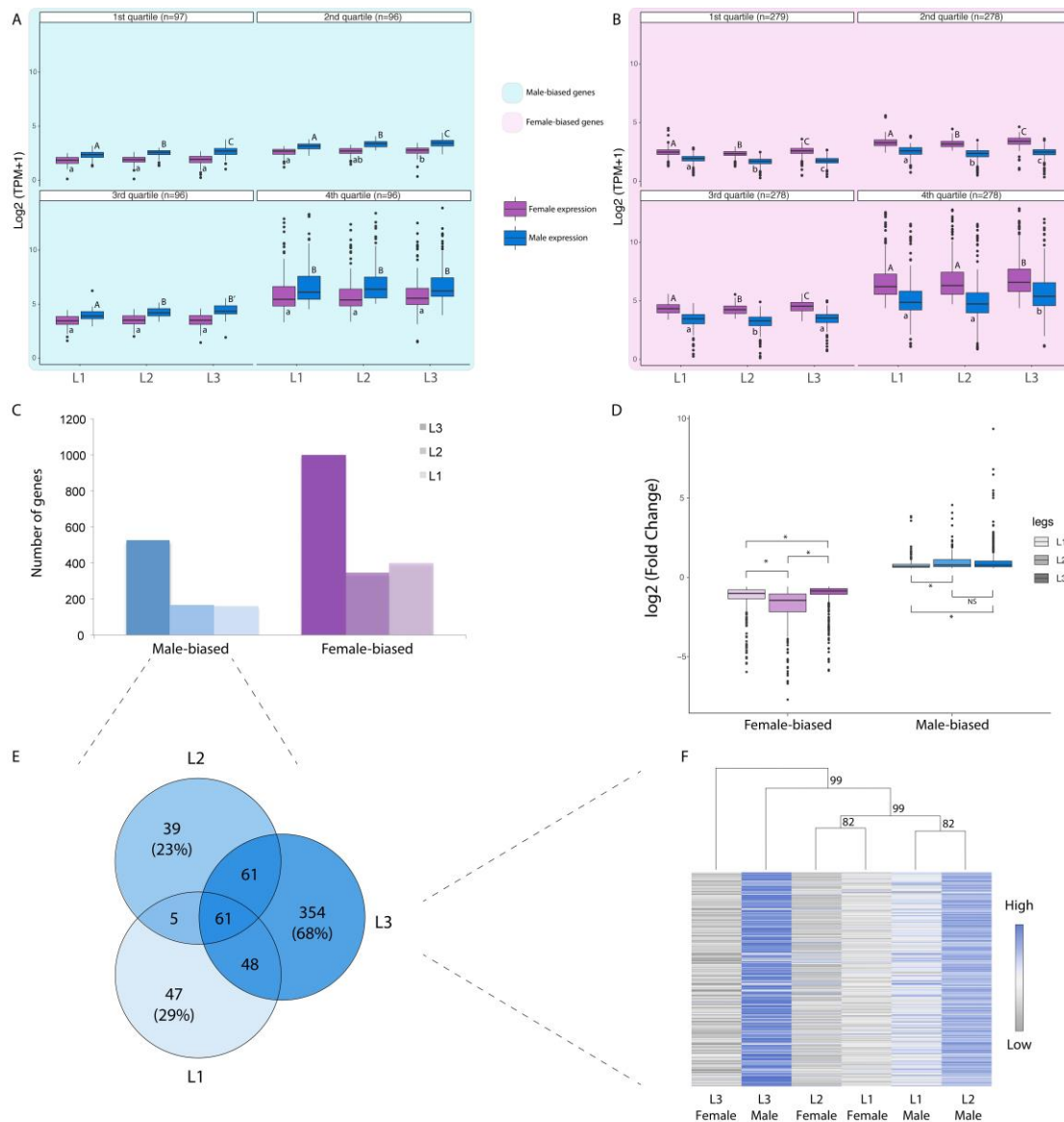


Figure 2: Signature of trait exaggeration among sex-biased genes. Expression in legs of male-biased genes (A) and female-biased genes (B) identified across all legs. Blue and purple colors indicate gene expression in male and female tissues respectively. Background colors represent the set of sex-biased genes, either male-biased ($N=385$, light blue) or female-biased ($N=1113$, light purple). Different letters indicate significant differences in expression (Wilcoxon tests). Data were divided into four quartiles based on expression level in males for panel (A) and females for panel (B). C) Number of male- and female-biased genes identified in each pair of legs independently. D) Differences in fold change (Wilcoxon tests) among the sex-biased genes identified in the three pairs of legs independently. E) Venn-diagrams of the male-biased genes identified from the three pairs of legs. F) Hierarchical clustering (1000 bootstraps) and heatmap based on average leg expression in males and females for the genes with significant male-biased expression specifically in L3 males ($n=354$).

D. Male-biased genes show unique features of trait exaggeration

In *M. longipes* males, the fore-legs are iso-allometric whereas the mid and rear-legs are hyper-allometric (Figure 1C). In females however, the fore-legs are hypo-allometric while both mid- and hind-legs are iso-allometric (Figure 1C). We therefore analyzed our dataset in search for patterns of sex-biased gene expression that correlate with the differences in scaling relationships between male and female legs. Sex-biased expression is often assumed to represent some of the molecular mechanisms underlying sexual dimorphism [102]. As a first approach, we identified sex-biased genes across the three pairs of legs and classified them into four quartiles, according to their average expression in males for the male-biased genes and female average expression for the female-biased genes (Figure 2A-B). We detected about three times more female-biased than male-biased genes (1 113 in females and 385 in males). Interestingly, male-biased genes from the three first quartiles displayed a progressive increase in male expression following the gradual increase in male leg lengths, while these genes were steadily expressed across female legs (Figure 2A). The highest expressed genes (4th quartile) were, however, not differentially expressed between male legs (Figure 2A). Conversely, female-biased genes displayed differences in expression in female legs that were inconsistent with the differences observed in female leg length (Figure 2B). Female-biased genes were always more expressed in female L3 compared to the two other legs but were also downregulated in female L2, except for the highest expressed genes (Figure 2B). Overall, *M. longipes* males show a progressive masculinization of male-biased gene expression in the legs, following the degree of sexual dimorphism in leg lengths.

To further correlate sex-biased genes expression with dimorphism in the leg lengths, we identified male- and female-biased genes in each leg separately. First, we found that the number of male- and female-biased genes follows the degree of sexual dimorphism in legs, with more sexually dimorphic genes in the L3 compared to the two others legs (Figure 2C). The degree of dimorphic expression (e.g. $\log_2(\text{Fold Change})$) also changes according to the legs (Figure 2C, D). In L3 and L2, we found a higher degree of male- biased expression compared to L1. The degree of female-biased expression was here again significantly different between the three pairs of legs, with on average the highest $\log_2(\text{Fold Changes})$ in L2 and the lowest in L3, although this pattern was inconsistent with the differences in leg lengths (Figure 2D). The development of exaggerated legs in *M. longipes* males is therefore associated with an increase in the number of male-biased genes and the degree of male-biased

expression. The number of female-biased genes was also higher in the exaggerated legs but their degree of differential expression remains to be elucidated, as it is inconsistent with the phenotypic differences observed in leg lengths. Interestingly, genes with male-biased expression largely overlap among the legs (77% for L2 and 71% in L1), although L3 presents 68% of unique genes (Figure 2E). A hierarchical clustering revealed that the expression of these genes in male L1 and L2 was more similar to female legs than L3 males. This indicates that the dimorphic expression observed is indeed specific to L3 males and not caused by a lack of statistical power in the two other legs (e.g. lower degree of sexually dimorphic expression) (Figure 2F).

Finally, we aimed to determine the molecular function of the sex-biased genes in our dataset. Gene ontology (GO) term analyses revealed enrichment in translation, metabolic processes and Wnt signaling pathways for the male-biased genes in L3 (Supplementary table 4). The “translation” GO term uncovered enrichment for several ribosomal proteins also known to play an essential role in cell proliferation in response to ribosomal stress [208]. We also identified enrichment in molecular functions such as the transferase activity indicative of some post-transcriptional regulations between the two sexes. Female-biased genes in L3 were enriched in various functions such as transcription factor, kinase, or GTPase activities that are probably involved in regulating biological processes such as transcription, metabolism, or signal transduction (Supplementary table 5).

E. Divergence in male leg homology is associated with a set of male-specific leg-biased genes

Understanding the regulatory mechanisms underlying divergence of serial homologues is a longstanding question in evolutionary developmental biology [209]. Insect legs are homologous structures that have received lots of attention in developmental genetics. Yet, the regulatory processes underlying homology divergence in males and females remain poorly understood [210, 211]. In *M. longipes*, leg divergence is more pronounced in males than females. We therefore hypothesized that gene expression between legs would be more similar in females than males. The analysis of expression variance on the whole dataset already revealed no major differences in leg expression variation between males and females (Supplementary figure 2B). In a second PCA analysis, where we separated male and female datasets, the first principal component separated the legs and contributed to about 35% of the

total variation in both males and females (Figure 3A). This indicates that leg divergence between sexes is not associated with a general change in leg expression.

Alternatively, such divergence could be encoded by the regulatory changes of a small set of genes. We therefore identified, in both sexes, genes with significant differences in expression between legs and found a higher number of leg-biased genes in males compared to females (Figure 3B,C; Supplementary figure 7). This suggests that heightened phenotypic divergence in male legs is associated with an increase of leg-biased genes. More specifically, when we compared the overlap of L3-biased genes (compared to L1) between males and females, only 56 genes (51%) were in common (Figure 3B). Even by changing the stringency, either more (Fold-Change > 1 , $P_{adj} < 0.05$) or less (Fold-Change > 0 , $P_{adj} < 0.05$) stringent, we still obtained about half of the genes in common (Figure 3B). Expression estimates and clustering of L3-biased genes limited to males, revealed more similarity in expression between sexes than legs (Figure 3B). This indicates that L3-biased genes specific to males tend to be similarly expressed in female legs. A similar pattern was observed in L3-biased genes restricted to females (Figure 3B).

To see whether this pattern was specific to L3, we performed the same analysis in L1 and found that up-regulated genes in these legs had up to 63% overlap between males and females (Figure 3C). Moreover, a hierarchical clustering of the L1-biased genes that were exclusive to one of the sexes, show more similar expression between sexes than between legs of the same sex (Figure 3C). This indicates that the non-overlap of some L1-biased genes, despite their similarity in expression between sexes, tends to result from a lack of statistical power to detect them in one of the sexes (e.g. lower differences in expression).

Upregulated genes in the second pair of legs (compared to the first legs) also displayed fewer overlap between males and females than L1 genes confirming that this pattern is associated with the development of exaggerated legs (Supplementary figure 7).

Finally, we found that the leg-biased genes shared among sexes do not show any significant differences in expression between males and females on average (Supplementary figure 8).

Taken together, our data indicate that the heightened divergence in male legs is associated with the use of a specific set of L3-biased genes that are not differentially expressed in female legs. Nonetheless, we detected a set of common genes between sexes that may encode key

developmental regulators for the general leg divergence, including allometric scaling, in *M. longipes*.

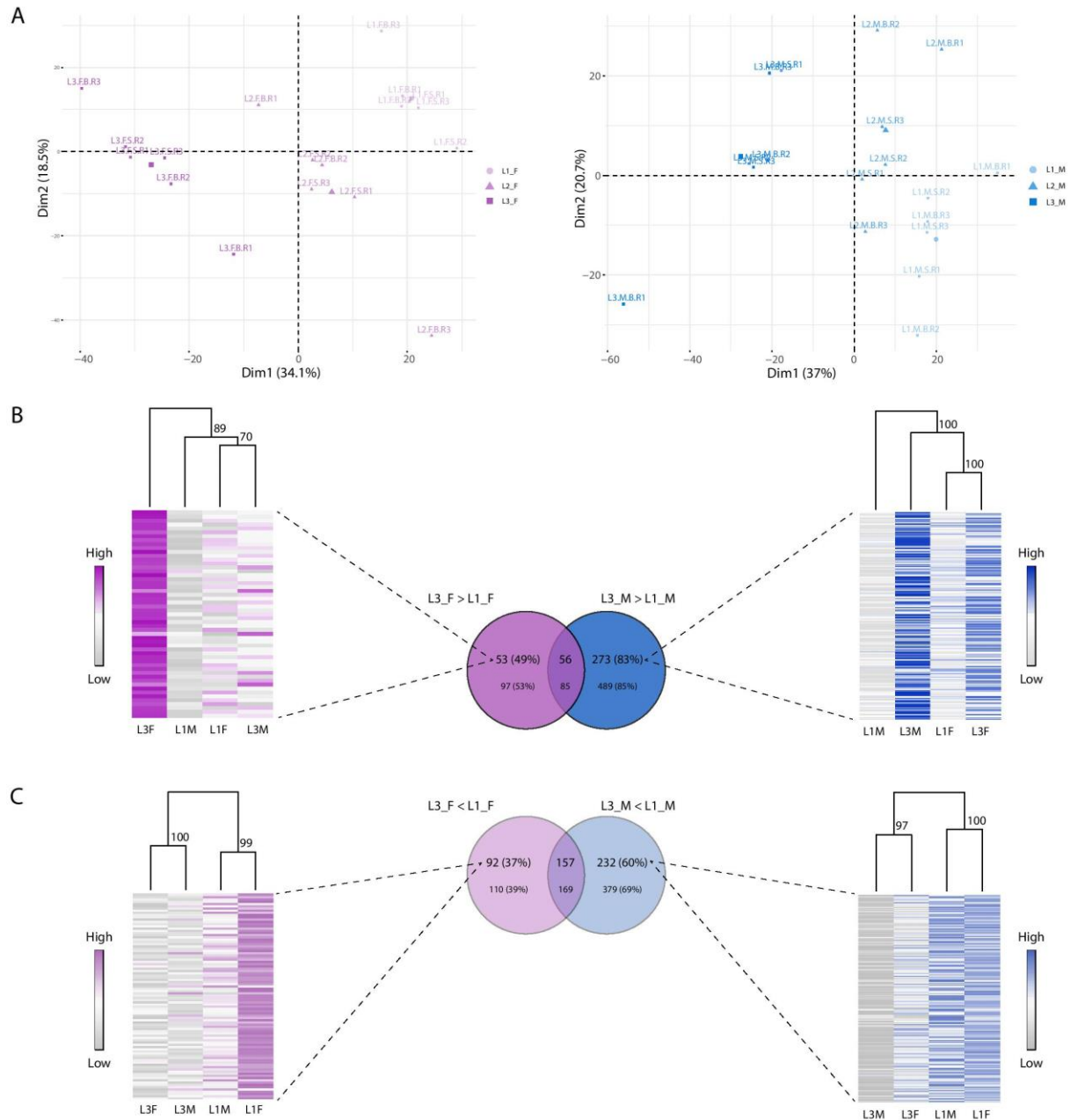


Figure 3: Signature of trait exaggeration among leg-biased genes. Venn-diagrams illustrate the number of leg-biased genes shared in males (grey) and females (brown), both for upregulated genes in L3 (B) and L1 (E). Higher font size indicates the number of leg-biased genes with a fold-change > 1.5. Smaller font size indicates the number of leg-biased genes using a lower threshold (fold-Change > 0). Heatmaps and hierarchical clustering show similarity in average expression between legs and sexes for the set of leg-biased genes (fold-Change > 1.5, $P_{adj} < 0.05$) specific to males (C, F) and females (A, D).

Among the differentially expressed genes identified in the legs, we recovered the HOX genes *Ultrabithorax* and *Sex-comb reduced*, known to be involved in the development of sexually dimorphic phenotypes in the third and first legs respectively [212, Pattatucci, 1991 #1379, 213]. Ontology terms for the upregulated genes in L3 male revealed enrichment in transmembrane transporter activity and related biological processes such as metabolic processes, membrane transport and signaling (Supplementary table 6). Among others, we could identify obvious candidate genes to tissue growth such as the *Insulin receptor 2*, the *growth differentiation factor 11* (*Bone Morphogenetic Protein 11*) or a duplicated copy of the *Gamma-interferon-inducible lysosomal thiol reductase* known to regulate leg growth in water striders [67].

F. Male exaggerated legs are enriched in genes with leg- and sex-biased expressions.

Our analyses in both sex- and leg-biased genes report a possible crosstalk between tissue and sex regulations in the context of trait exaggeration. To test for this hypothesis, we compared the fold-change (FC) of each gene from leg and sex comparisons, and searched for genes with both leg- and sex-biased expression in male rear legs. Interestingly, upregulated genes in male L3 (compared to L1) tend to also be male-biased in L3 (Fisher's exact test; $p\text{-value} < 0.05$) (Figure 4A). Similarly, downregulated genes in L3 male tend to be female-biased in the same legs (Fisher's exact test; $p\text{-value} < 0.05$). This association is, however, reduced when we looked at sex-biased genes in L1. Upregulated genes in L1 of males were deficient for male-biased expression (Fisher's exact test; $p\text{-value} < 0.05$) and did not show any tendency towards female-biased expression (Fisher's exact test; $p\text{-value} > 0.05$) (Figure 4B). These results suggest that a crosstalk between tissue and sex regulations may have occurred in association with the evolution of exaggeration in male rear legs.

To further test this hypothesis, we performed the same analyses in females and show that leg-biased genes do not present any tendency towards sex-biased expression in L3 (Fisher's exact tests; $p\text{-values} > 0.05$) (Figure 4C). We obtained similar results when we selected the sex-biased genes in L1, except for the L1-biased genes that tend to be female-biased (Fisher's exact test; $p\text{-value} < 0.05$) (Figure 4D).

On the second legs we also found enrichment for genes with both leg- and sex-biased expression in males (Fisher's exact tests; $p\text{-value} < 0.05$). Yet, we could observe that these genes were less numerous compared to the ones in L3 (Supplementary figure 9).

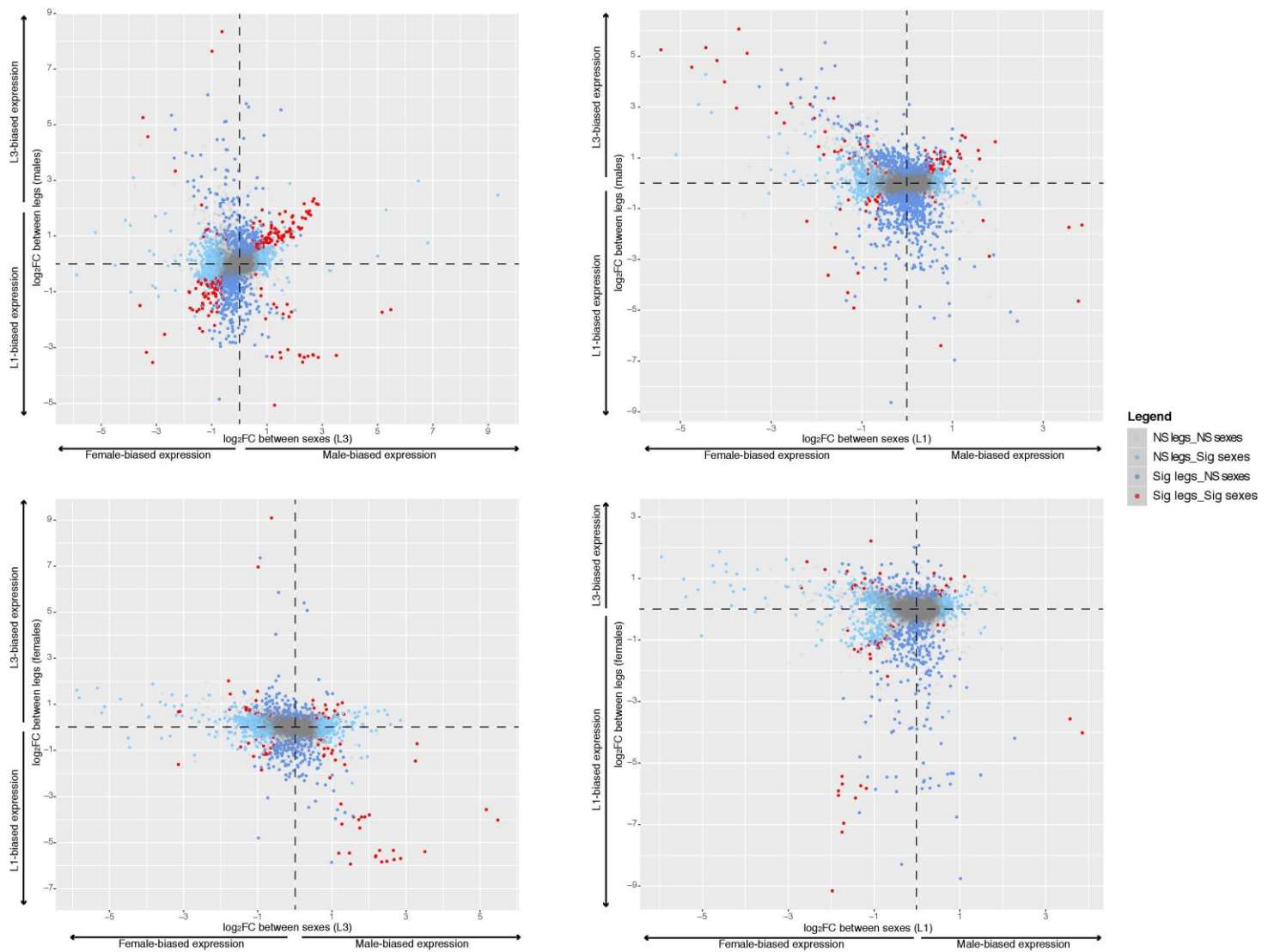


Figure 4: Correlation between leg and sex fold-changes for all expressed genes. Light blue dots are sex-biased genes ($FC > 0$, $P_{adj} < 0.05$), dark blue dots correspond to leg-biased genes ($FC > 0$, $P_{adj} < 0.05$) and red dots are genes with both leg- and sex-biased expressions. For the leg-biased genes, we only illustrated the comparisons between L3 and L1. In the figure legend, “NS” indicates, “Non Significant” and “Sig” indicates “Significant”. Data on the L2 are shown in the Supplementary figure 5. A) Comparison between sex-biased genes in L3 and leg-biased genes in males. B) Comparison between sex-biased genes in L1 and leg-biased genes in males. C) Comparison between sex-biased genes in L3 and leg-biased genes in females. A) Comparison between sex-biased genes in L1 and leg-biased genes in females.

Overall, we found that a crosstalk between tissue and sex regulations was particularly enhanced in male exaggerated legs. About 200 genes exhibit both leg- and sex-biased expressions in L3, including the *Insulin receptor 2*, a wingless-like protein, the *suppressor of cytokine signaling 2* gene or a small heat shock protein (HSP20 family), that are all good candidates to explain the exaggeration in male rear legs.

G. The influence of sexual selection on *M. longipes* genome architecture

The high-quality genome assembly of *M. longipes* allows us to ask more general evolutionary genomic questions such as how sexual selection, through male-male competition, can influence genome architecture. Theoretical studies predict that sexual selection can be a major driver of genome evolution [80, 214]. The X chromosome, for example, has been hypothesized to be a genomic hotspot for sexual selection where female beneficial dominant mutations and male beneficial recessive mutations are expected to accumulate [80, 114, 215]. Our genomic and transcriptomic dataset offers the opportunity to test more accurately the role of sexual selection on chromosome evolution by directly linking the development of sexual dimorphism with the genomic location of sex-biased genes in the three pairs of legs. Our approach tests for the non-random genomic distribution of sex-biased genes at different genomic resolutions, giving a general understanding of the link between sexual selection and genome evolution.

A) Accumulation of female-biased genes on the X chromosome

In a first approach, we tested whether selection could influence the accumulation of sex-biased genes on the X chromosome. We identified the scaffold 1893 as the X chromosome in *M. longipes* genome (see material & methods) and assessed accordingly the chromosomal location of the sex-biased genes from the three pairs of legs. Interestingly, female-biased genes in L3 were over-represented on the X chromosome compared to the autosomes and this pattern was absent in the two other legs (Figure 5A). The percentage of female-biased genes between the X chromosome and the autosomes in the different legs confirmed that the enrichment observed was caused by a L3-specific accumulation of female-biased genes on the X chromosome (Figure 5A). In contrast, we did not find any significant under- or over-representation of the male-biased genes on the X chromosome, regardless of the legs (Figure 5A).

The representation of sex-biased genes on the X chromosome is often influenced by dosage compensation [102, 114, 215].

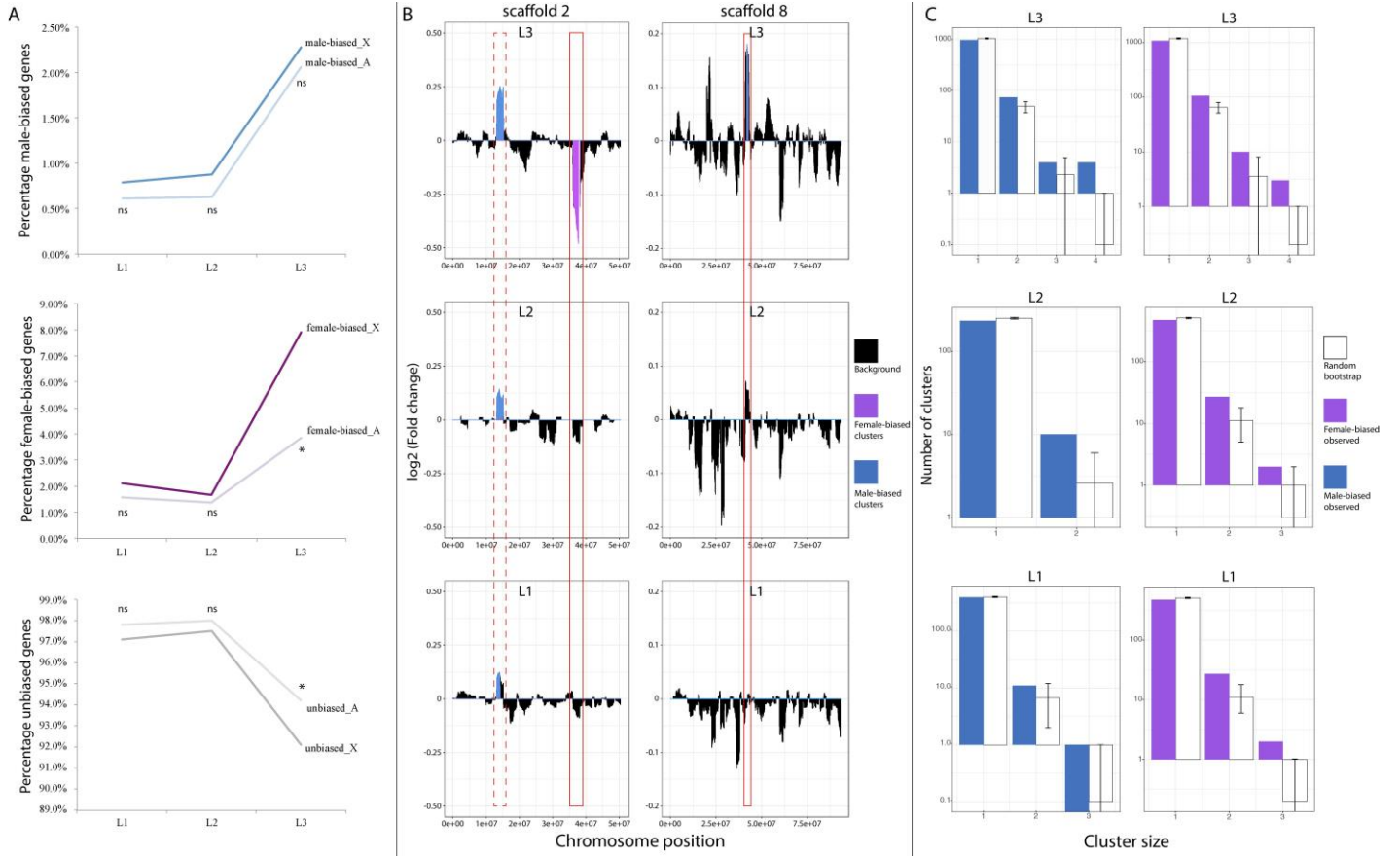


Figure 5: Male- and female-biased gene distributions in *M. longipes* genome. (A) Percentage of male-biased, female-biased and unbiased genes (from top to bottom) in the three pairs of legs was calculated in the X chromosome and the autosomes. Biased-distribution of sex-biased genes on the X chromosome was estimated using the Fisher's exact tests: * p-value < 0.05. (B) Large genomic clusters of sex-biased genes along the scaffold 2 and the scaffold 8. Clusters highlighted in blue represent genomic regions enriched in male-biased genes. Clusters highlighted in purple represent genomic regions enriched in female-biased genes. Solid red frame indicates genomic clusters enriched in male- and female-biased genes specifically in L3. Dotted red frame indicates genomic cluster enriched in male-biased genes in all three legs but with different degrees of fold-change recapitulating the degree of leg length exaggeration. (C) Genomic clusters of consecutive male- (blue) or female-biased (purple) genes in the three pairs of legs. Cluster size indicates the number of consecutive genes found in cluster. Note that the y axis, indicating the number of clusters, is log scaled. Error bars indicate fluctuation intervals.

In *Drosophila* for example, the deficiency of male-biased genes on the X chromosome was suggested to result, at least partially, from dosage compensation [118]. We therefore compared male and female gene expressions on the X chromosome and found no average differences in expression, in all three legs (Supplementary figure 10). This indicates that dosage compensation is a mechanism present in *M. longipes*, regardless of the tissue sampled and is therefore unlikely to be responsible for the over-representation of female-biased genes on the X chromosome. Overall, our results suggest that the resolution of sexual conflict during the evolution of extreme sexual dimorphism in *M. longipes* rear legs may have been favored by the accumulation of female-biased genes rather than male-biased genes on the X chromosome. It is, however, important to note that some X-linked genes, such as the wingless-like protein, are both leg- and sex-biased in L3 males and represent good candidates for the development of the sexually dimorphic exaggerated legs.

B) Large Genomic clusters of sex-biased genes

In the two other approaches, we hypothesized that sexual selection may also have influenced genome architecture through gene or chromosomal rearrangements, as it has been illustrated in some studies [157, 158, 188, 190, 191, 216]. We therefore performed a fine-scale visualization of the sex-biased gene distribution along the thirteen largest scaffolds in *M. longipes* genome (Figure 5B, Supplementary figure 9). We first tried to identify relatively large genomic regions of about 100 genes (2 Mb) that were significantly enriched in sex-biased genes. Only 11 such regions were found in the genome, across the legs. A total of 100 sex-biased genes were recovered in these clusters (about 2% of the total number of sex-biased genes in all three legs), indicating that these genes do not generally arrange in large genomic clusters (Figure 5B). We could nonetheless locate three large clusters, on scaffolds 2 and 8, composed of 36 sex-biased genes (11 female-biased and 25 male-biased) in L3 (Figure 5B). Interestingly, two of these clusters were specific to the third leg whereas the third one indicated an enrichment of male-biased genes that was common to the three pairs of legs but with different degrees of sex-biased expression (Figure 5B). In these clusters, we could notably identify several unknown genes (10 out of 36 genes), including a cluster of four consecutive, probably duplicated, genes that were all strongly male-biased. Protein motif prediction on these genes, using Pfam, revealed a conserved domain of several transmembrane regions.

C) Genomic clusters of consecutive sex-biased genes

Finally, the third approach consisted of identifying small clusters of consecutive male- or female-biased genes. For the male- and female-biased genes in L3, we detected an enrichment of clustered genes in the genome compared to a random distribution (p-value < 0.05; see material and method). More than 15% of sex-biased genes in L3 arrange in clusters of at least two genes whereas 10% are approximately expected by chance (Supplementary figure 4). More precisely, female-biased genes were enriched in clusters of two to four genes, whereas male-biased genes were enriched in clusters of two and four genes (Figure 5C). In the second pair of legs, we also revealed an enrichment of sex-biased gene clusters with about 10% of them arranging in clusters (p-value < 0.05; Supplementary figure 4). However, male- and female-biased clusters did not exceed two and three consecutive genes, respectively (Figure 5C). In comparison, male-biased genes in L1 did not show any enrichment in clusters in the genome (p-value > 0.05; Supplementary figure 4). Moreover, only one cluster of three male-biased genes was detected in L1 and no clusters of four genes (Figure 5C). We found, nonetheless, an enrichment of female-biased gene clusters, including 2 clusters of 3 consecutive genes (Figure 5C; p-value < 0.05; Supplementary figure 4).

Overall, our fine-scale approaches report that only few sex-biased genes in L3 significantly cluster over large genomic regions (about 1.4%) but more of them (about 15%) gather in smaller regions of two to four genes. It is, however, important to note that despite the larger number of small clusters detected in L3 compared to the two other legs, most sex-biased genes, regardless of the legs, are generally found randomly distributed in the genome.

V. Conclusion and discussion

Microvelia longipes is a new model system to study how sexual selection can drive morphological and genomic adaptation. In a previous study we found that the evolution of leg exaggeration was associated with intense competition between males over egg-laying substrates [192]. However, the developmental processes regulating the exaggeration of these legs were unknown. Hemimetabolous insects have an easy tractable post-embryonic development with individuals growing gradually through successive instars and each molt leaving an exuvia that can inform about earlier developmental processes [21]. Monitoring the growth of each individual molt throughout the nymphal development allowed to precisely

identify a developmental window, the fifth nymphal instar, where *M. longipes* males experience a burst of growth in their rear legs (Figure 1B, Supplementary table 3). We still however do not understand how this growth is regulated at the cellular level. Previous works in stalk-eyed flies and horned beetles have shown that cell size, cell numbers and even cell death through apoptosis were associated with differences in the size of these sexual traits [217, 218]. Understanding the cellular mechanisms controlling the exaggerated leg growth in *M. longipes* would ultimately narrow down the list of candidate genes already identified in the transcriptomic analysis.

An extensive literature has already been established regarding the developmental genetics underlying sex-specific trait exaggeration [35, 144]. Lavine et al. drew notably a comprehensive review on the developmental mechanisms and pathways likely to contribute to exaggerated trait elaboration [35]. Interestingly, many of these genes show differential expression between legs in our dataset. Some genes, such as *Ultrabithorax*, *dachshund*, the *insulin-like receptor 2* or a wingless-like protein, display significantly more expression in male rear legs compared to other male and female legs, consistent with male exaggerated leg growth. Some other genes from the same regulatory pathway, such as the *Epidermal Growth Factor Receptor*, *Distalless*, *Frizzled* or *Dachsous*, show surprisingly opposite patterns with a downregulation in male exaggerated legs. We also noticed that several of these genes present intriguing evolutionary histories with possible rounds of gene duplication, as it is the case for the wingless and wingless-like genes or the insulin-like receptors that were already identified in three copies in Gerromorpha [219]. Gene duplication is thought to be a common way to evolve sex-biased gene expression, with several male-biased genes containing paralogues in both worm and fly genomes [177, 220]. Duplication of sex-biased genes may therefore point to a common evolutionary path to induce sexual dimorphism and resolve sexual conflict [177]. Overall, these patterns highlight the accuracy of our comparative transcriptomic analyses but also suggest some complex regulatory mechanisms underlying the development of exaggerated phenotypes. Our data also shed light on new promising regulatory factors involved in allometric scaling. Among others, we identify the genes *BMP11* or the *pituitary homeobox 2* that are members of well-known signaling and developmental pathways and that display upregulated expression in male L3 [221, 222]. Functional assays through RNAi experiments, for example, will help confirming the potential role of these genes in the elaboration of exaggerated legs in *M. longipes* males.

Sex-biased gene comparisons also revealed some unexpected results, especially regarding the expression of female-biased genes (Figure 2). First, female-biased genes are followed an inconsistent pattern of expression between female legs in regard to the differences observed in leg length. This feature may underlie some physiological or morphological processes that we could not assess in our analyses. Unlike male-biased genes, for which the average expression in male and female tissues was decoupled, we could observe that the expression of female-biased genes in male tissues was following the same pattern as the one in female tissues (Figure 2A). This possibly highlights some on going sexual conflict at the expression level for female-biased genes, which may in turn influence the expression pattern observed between legs. Identifying more specifically the function of these genes and testing for their possible role in modulating *M. longipes* leg length would be an interesting question to follow in the future.

The three pairs of legs in *M. longipes* are unique comparisons to understand the genetic mechanisms of scaling relationships as they represent serial homologues that display different allometric slopes (Figure 1C). Moreover, male and female legs present different types and degrees of sexual dimorphism, likely reflecting distinct selective pressures (Figure 1). The development of sex-combs in *M. longipes* males, for example, is a common feature known to be under stabilizing sexual selection and specific to the first legs in insects [223, 224]. On the other hand, the exaggerated growth of male rear legs is an example of directional sexual selection that is absent, or at least strongly reduced, in the two other male legs [192]. We therefore used *M. longipes* as a comparative system to test the role of sexual selection on genome evolution by directly link the developmental regulation of sexual dimorphism with the genomic location of sex-biased genes in the three pairs of legs. Previous studies have generally reported large genomic clusters and profound genomic rearrangements (e.g. large chromosome inversions) in association with sexually dimorphic characters [157, 158, 188, 190, 191]. In contrast, we found relatively few and small clusters of sex-biased genes in *M. longipes* genome (Figure 5). It is, however, important to note that previous studies were primarily conducted on primary sexual organs, such as ovaries and testes [157, 158, 188, 190, 191, 216]. These tissues are highly complex, often expressing more sex-biased genes than secondary sexual traits and evolving under various selective pressures including survival, fertility and sexual selections [5, 19, 111, 225, 226]. More generally, many organisms acquire sexually dimorphic phenotypes during development, before becoming adults [36, 41, 106, 227]. This implies that sexual selection may primarily acts on developmental regulatory

processes and that adult sex-biased expressions may result from other selective pressures than sexual selection. To our knowledge, our work is the first to report the genomic distribution of sex-biased genes underlying the development of a sexually-selected dimorphic phenotype. We believe that more similar works are needed to thoroughly evaluate how sexual selection shapes genomic evolution and adaptive phenotypes.

Authors' contributions: A.K. and W.T. designed research; W.T. and R.A performed the DNA library extractions; D.A. did the automated annotation; W.T. performed the growth rate analysis; W.T. performed the RNA extractions and the transcriptomic analyses; W.T. and C.D. performed the analysis on sex-biased gene distributions; A.K. and W.T. analyzed data; and A.K. and W.T. wrote the paper.

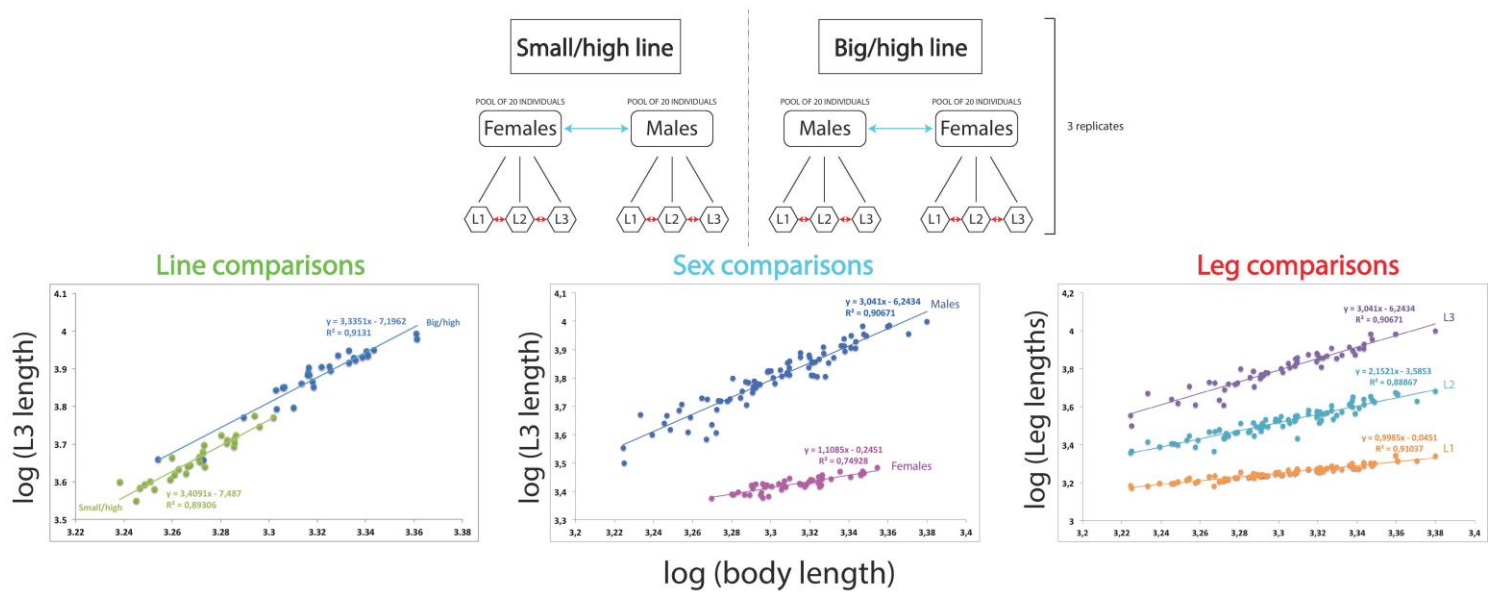
Competing interests: We declare we have no competing interests.

Funding: This work was funded by an ERC-CoG no. 616346 to A.K and a PhD fellowship from Ecole Doctorale BMIC de Lyon to W.T.

Acknowledgements

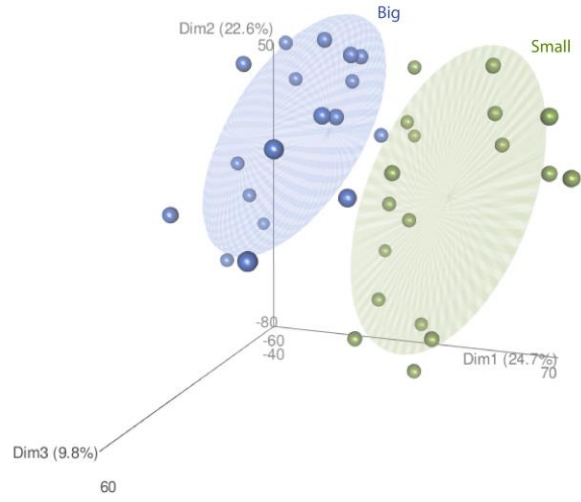
We thank Kevin Parsons, Gaël Yvert, François Leulier, Dali Ma, Mathilde Paris, Lucke Hayden, Augustin Le Bouquin, Amélie Decaras, Cédric Finet, Aidamalia Vargas, Séverine Viala, Vitoria Tobias Santos, Vincent Montbel for helpful discussions.

VI. Supplementary material

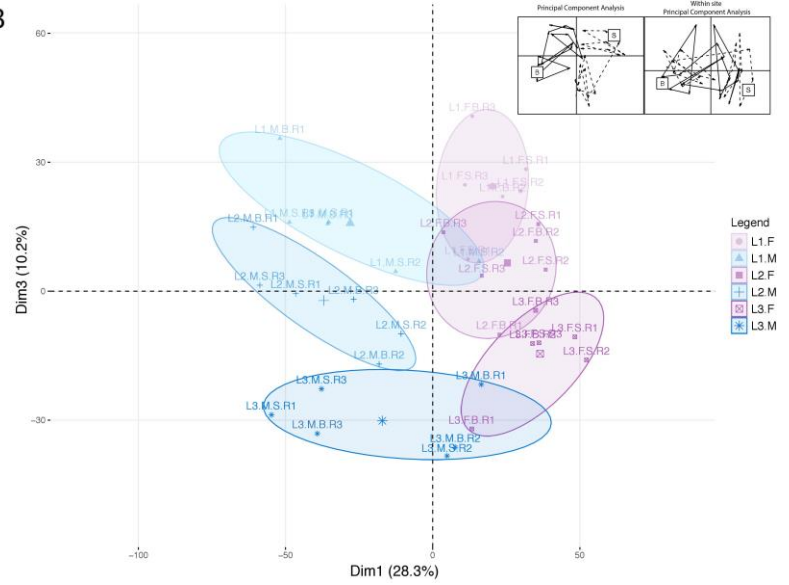


Supplementary figure 1: Experimental design for the comparative transcriptomic analysis.

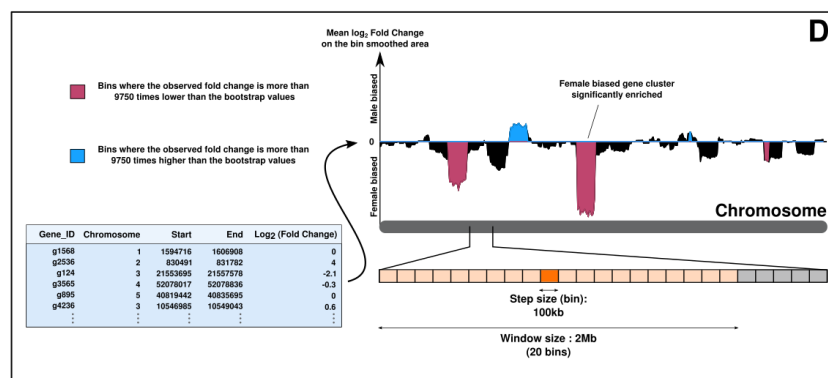
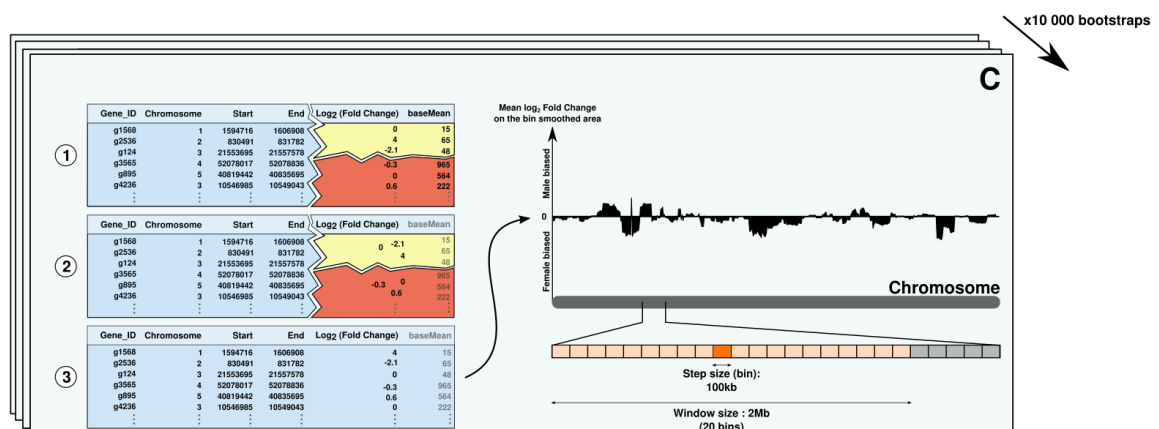
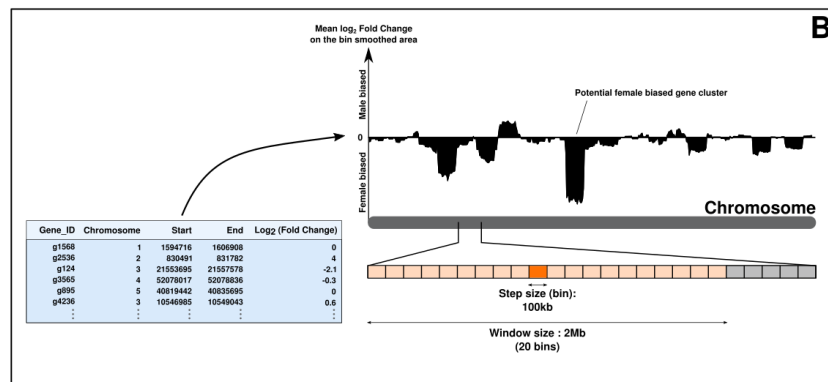
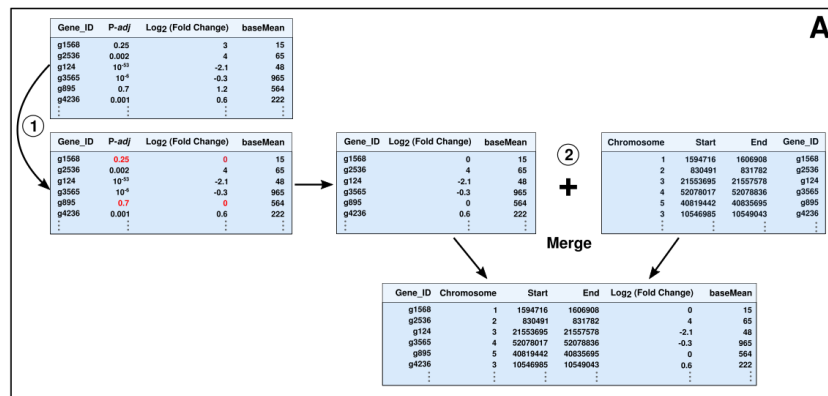
A



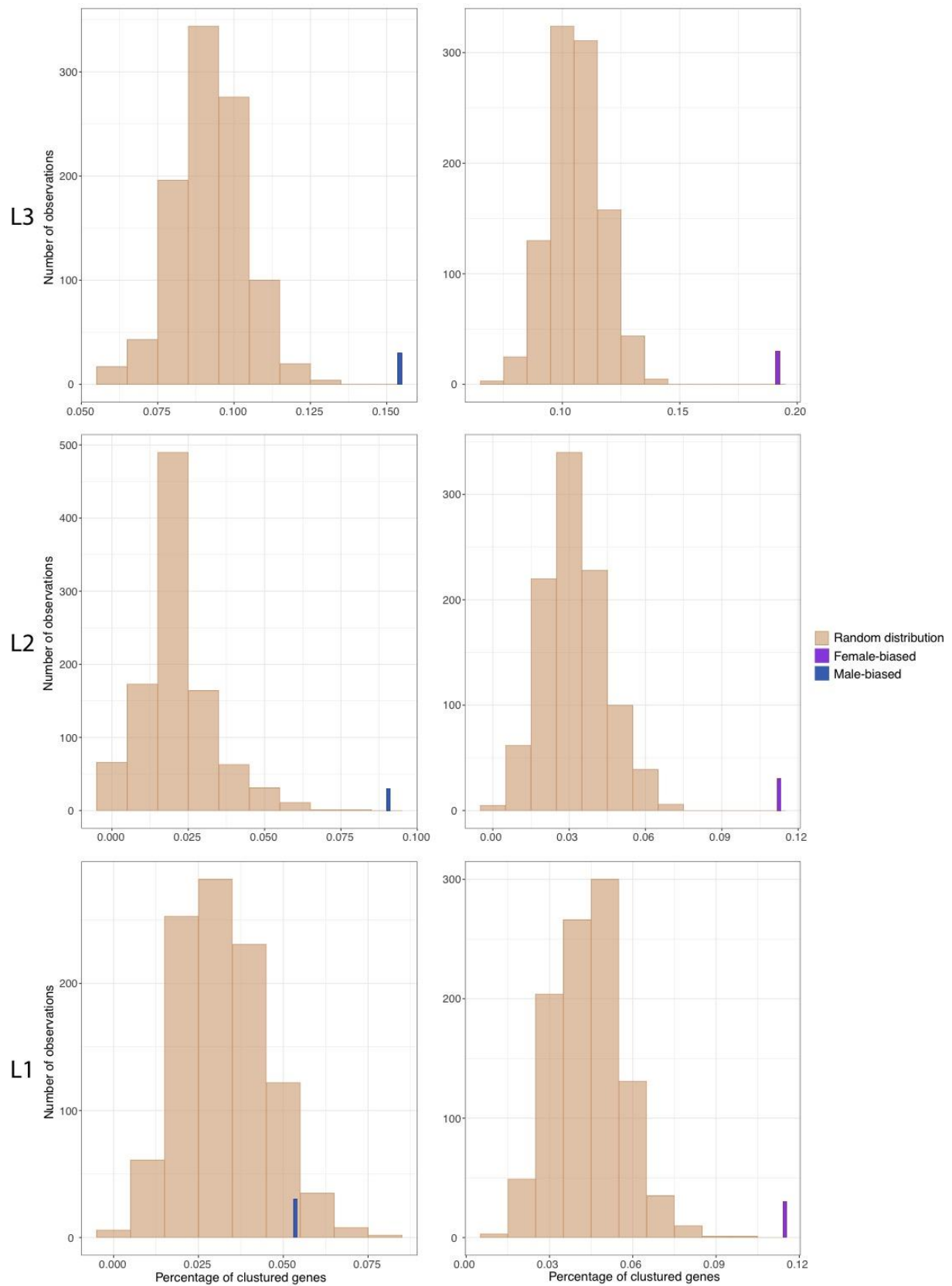
B



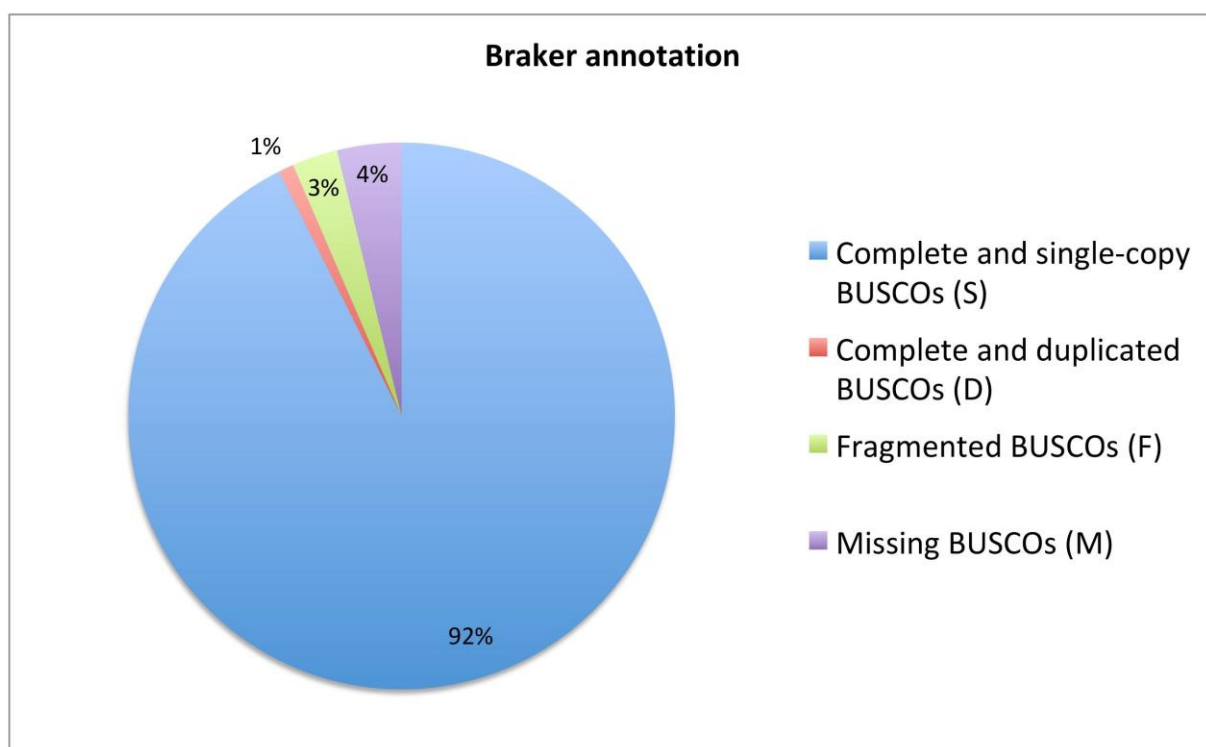
Supplementary figure 2: Principal Component Analysis (PCA) on the whole transcriptomic dataset. A) The three first PCAs (Dim1, 2, 3) recapitulate the variance between the Big (blue) and Small (green) lines. B) Within-Class analysis after correcting for the line effects. Dimension 1 separates sexes while Dimension 3 separates legs. The inset represents the Within-Class correction for the line effects.



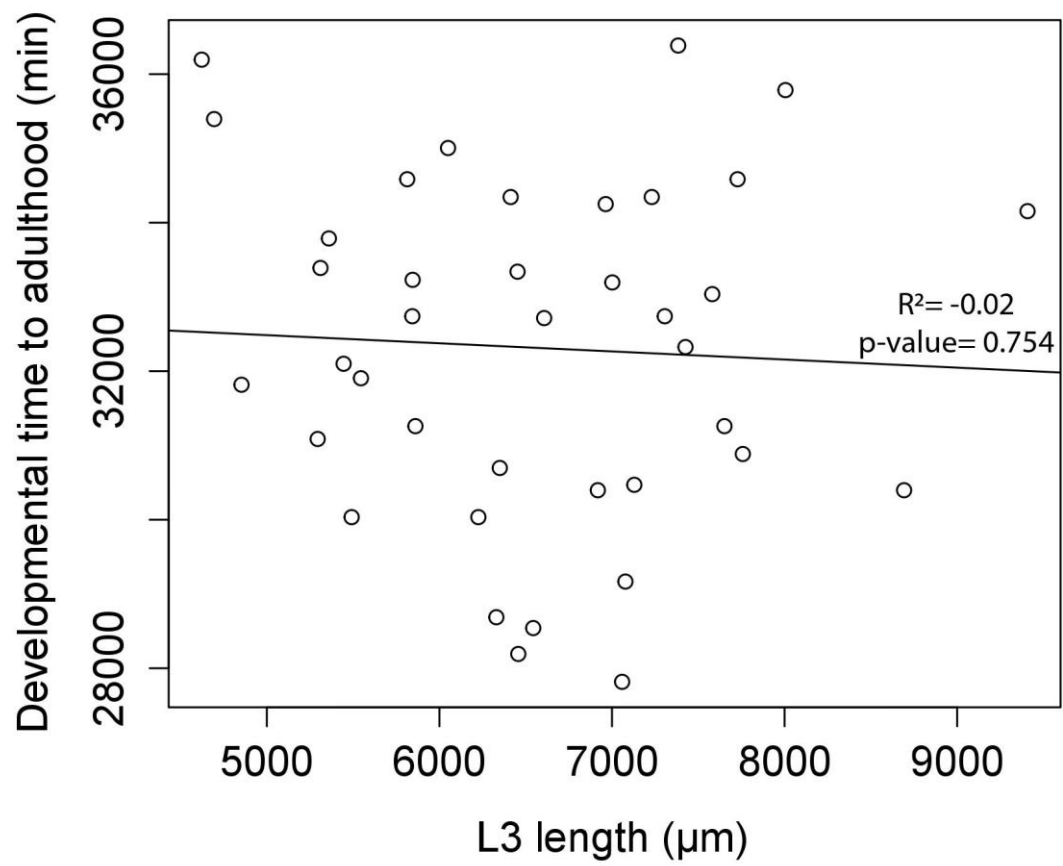
Supplementary figure 5: General pipeline of the bootstrap analysis to detect large genomic clusters of sex-biased genes.



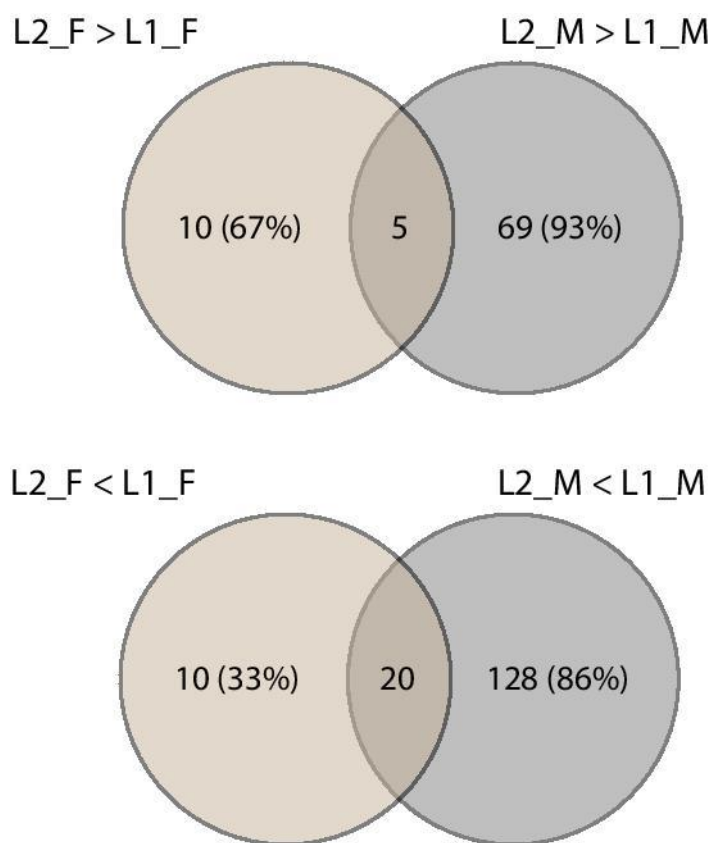
Supplementary figure 4: Distributions of the proportion of clustered sex-biased genes (at least two consecutive sex-biased genes) generated through 1000 random iterations in each leg separately. Blue and purple bars correspond to the observed proportions of male- and female-biased genes, respectively, in the three legs.



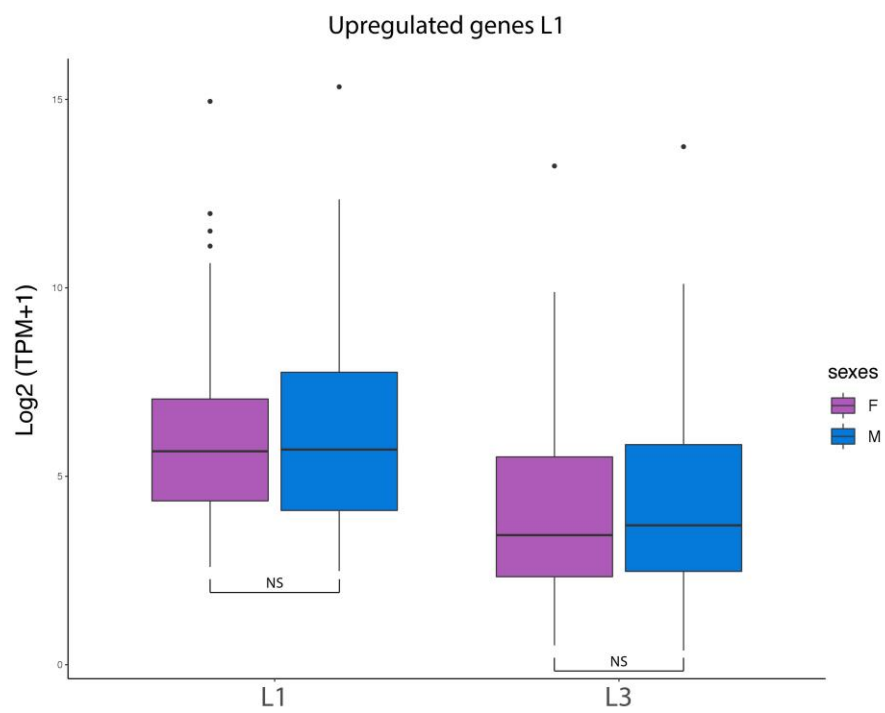
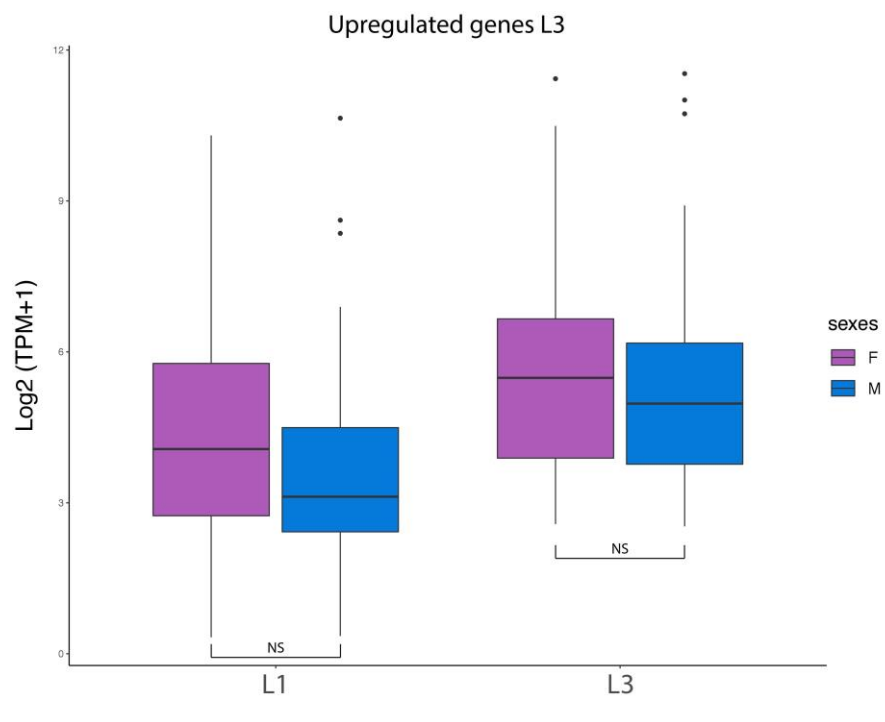
Supplementary figure 5: Diagram of BUSCO analysis run for the 26130 genes identified from Braker annotation.



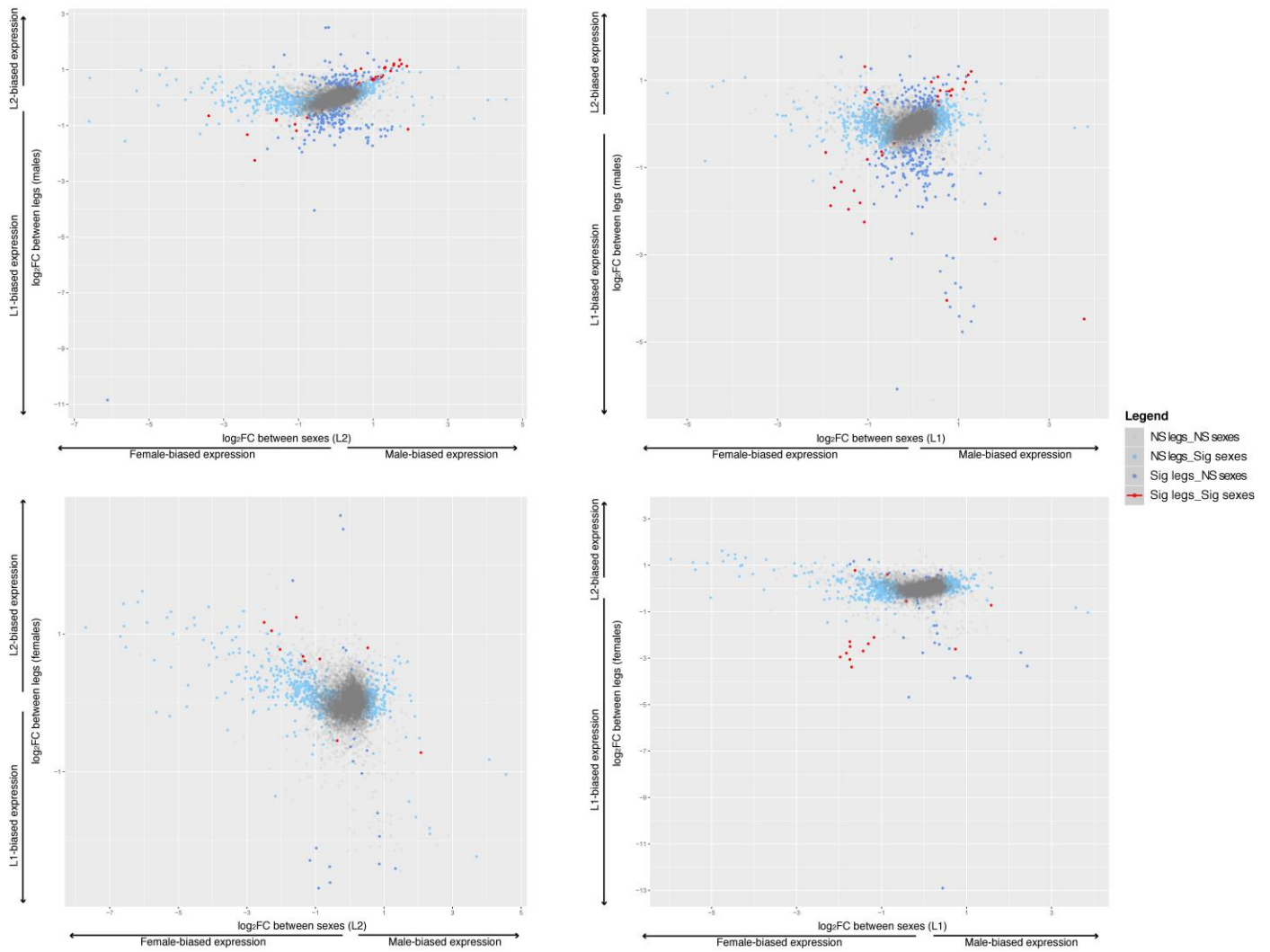
Supplementary figure 6: Relationships between L3 lengths and nymphal development time.



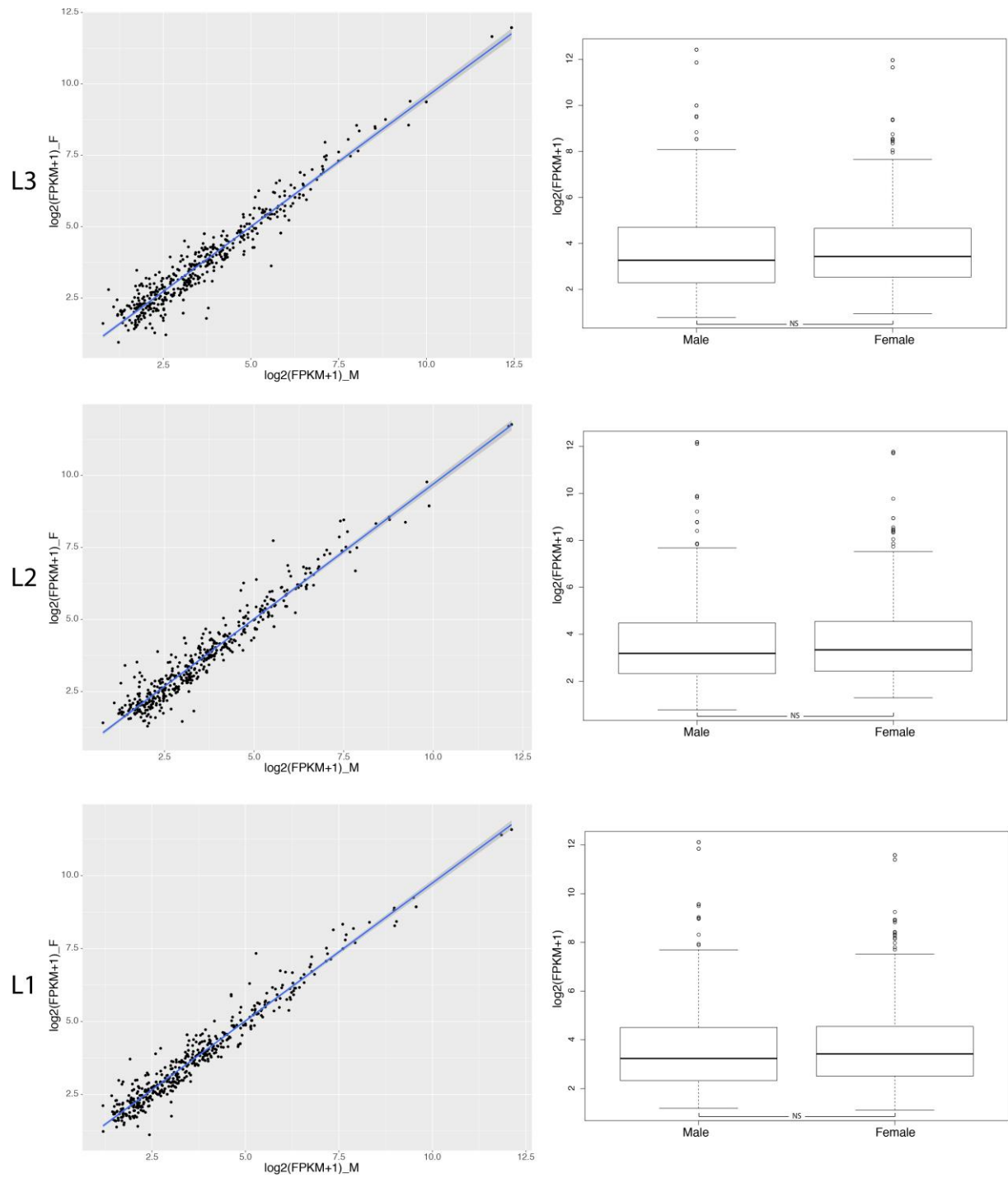
Supplementary figure 7: Venn-diagrams for the leg-biased genes between L2 and L1 in males and females.



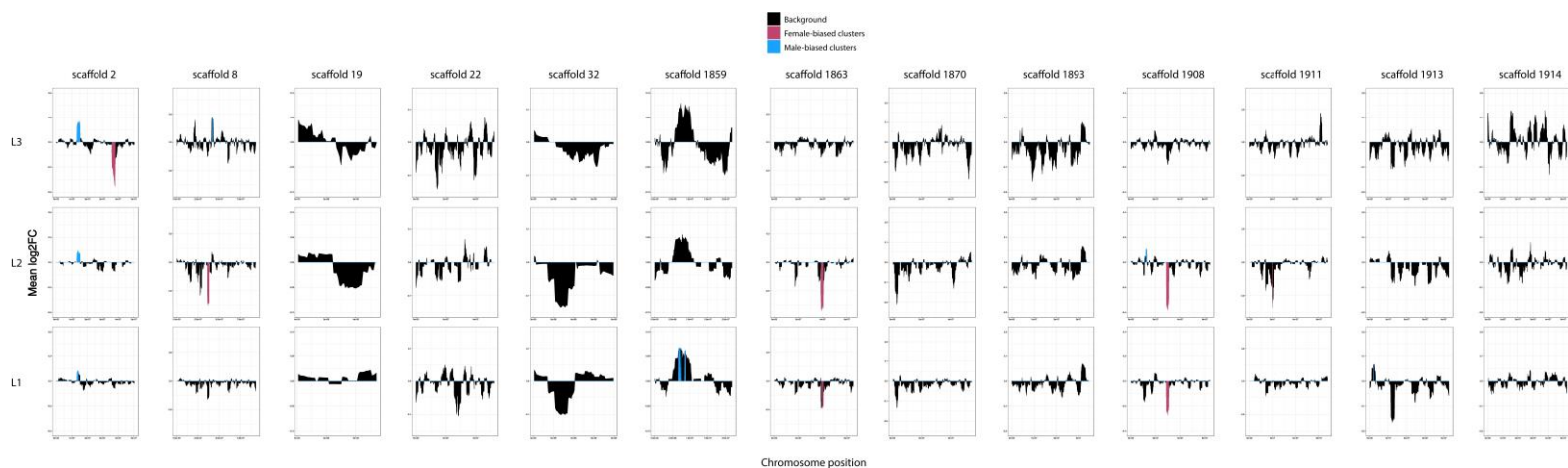
Supplementary figure 8: Boxplot leg-biased genes in common between L3 and L1 in males and females.



Supplementary figure 9: Correlation between leg and sex fold-changes using L2 as reference.



Supplementary figure 10: Average gene expression differences between males and females on the X chromosome. Regressions were fitted from a linear model.



Supplementary figure 11: Genome-wide characterization of large genomic clusters

| library type | library number | sample used | sequencing strategy |
|--------------|----------------|---|-------------------------|
| 270bp | 1 | Big/low 2 | HS4000 PE150 |
| 500bp | 1 | Big/low 2 | HS2500 PE 126 V4 |
| 800bp | 1 | Big/low 2 | HS2500 PE 126 V4 |
| 2kb | 1 | Females Big/low 2 | HS4000 PE50 |
| | | Small/high line 4 | HS4000 PE50 |
| | | Small/high line 8 | HS4000 PE50 |
| | | Small/high line 2 | HS4000 PE50 |
| | | Big/low 1 | HS4000 PE50 |
| | | Small/high line 3 | HS4000 PE50 |
| 5kb | 1 | Big/high | HS4000 PE50 |
| 10kb | 1 | Big/high | HS4000 PE50 |
| 20kb | 1 | Small/high line 5 | pacbio 20kb |
| | 1 | Small/high line 6 | pacbio 20kb |
| | 1 | Small/high line 7 | pacbio 20kb |
| | 1 | Small/high line 1 | pacbio 20kb |
| Hi-C | 1 | crossing populations from Big/high and Small/high (F2 and F5 generations) | Dovetail Hi-C libraries |

Supplementary table 1: Genomic libraries used to sequence and assemble *M. longipes* genome. Big/low, Big/high and Small/high being the three inbred populations used for the genome sequencing.

| scaffold | scaffold_length | gene number | gene density |
|------------|-----------------|-------------|--------------|
| chr1863_pp | 65320744 | 2365 | 27620 |
| chr8_pp | 92932883 | 4158 | 22350 |
| chr1870_pp | 54210254 | 2043 | 26535 |
| chr1908_pp | 64166213 | 2579 | 24880 |
| chr1859_pp | 21720206 | 1035 | 20986 |
| chr1913_pp | 42858894 | 1666 | 25726 |
| chr1914_pp | 52350239 | 1944 | 26929 |
| chr1911_pp | 67310210 | 2809 | 23962 |
| chr2_pp | 50305444 | 1907 | 26379 |
| chr22_pp | 52946816 | 1978 | 26768 |
| chr1893_pp | 35373809 | 1140 | 31030 |
| chr19_pp | 5395109 | 236 | 22861 |
| chr32_pp | 7944149 | 314 | 25300 |

Supplementary table 2: Summary table on the average gene density per scaffold.

A

Generalized linear model_L3-L2 male comparison (glm1)

| | Estimate | Standard Error | t value | p-value | |
|----------------|------------|----------------|---------|----------|--|
| (Intercept) | 5.79 | 3.123e-02 | 185.403 | < 2e-16 | |
| L1_length | 1.328e-03 | 3.046e-05 | 43.594 | < 2e-16 | |
| legs | -4.716e-04 | 4.417e-02 | -0.011 | 0.992 | |
| L1_length:legs | 3.131e-04 | 4.307e-05 | 7.270 | 1.15e-10 | |

Anova (glm1)

| | df | Deviance Residuals | df | Resid. Dev | p-value (chi) |
|-----------------|----|--------------------|----|------------|---------------|
| NULL | | | 95 | 63.642 | |
| L1_lengths | 1 | 60.174 | 94 | 3.468 | < 2e-16 |
| legs | 1 | 1.846 | 93 | 1.622 | < 2e-16 |
| L1_lengths:legs | 1 | 0.556 | 92 | 1.066 | 2.69e-12 |

B

Generalized linear model_male-female L3 comparison (glm2)

| | Estimate | Standard Error | t value | p-value | |
|------------------|------------|----------------|---------|---------|--|
| (Intercept) | 5.865 | 4.210e-02 | 139.329 | < 2e-16 | |
| L1_length | 1.464e-03 | 4.787e-05 | 30.586 | < 2e-16 | |
| sexes | -7.510e-02 | 4.945e-02 | -1.519 | 0.13343 | |
| L1_lengths:sexes | 1.767e-04 | 5.414e-05 | 3.264 | 0.00172 | |

Anova (glm2)

| | df | Deviance Residuals | df | Resid. Dev | p-value (chi) |
|-----------------|----|--------------------|----|------------|---------------|
| NULL | | | 71 | 46.994 | |
| L1_length | 1 | 46.308 | 70 | 0.686 | < 2e-16 |
| sexes | 1 | 0.072 | 69 | 0.614 | 0.002478 |
| L1_length:sexes | 1 | 0.078 | 68 | 0.536 | 0.001622 |

Supplementary table 3: Statistical analyses for differences in ontogenetic allometries between L3 and L2 in males (A) as well as male and female L3 (B). Allometric comparisons were performed using T1-legs as control variable and proxy for individual size.

| GO terms | Biological processes | Fisher test |
|------------|---|-------------|
| GO:0006412 | translation | 1.1e-16 |
| GO:0000028 | ribosomal small subunit assembly | 0.0041 |
| GO:0000398 | mRNA splicing, via spliceosome | 0.0067 |
| GO:0000413 | protein peptidyl-prolyl isomerization | 0.0074 |
| GO:0000462 | maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) | 0.0075 |
| GO:0001732 | formation of cytoplasmic translation initiation complex | 0.0142 |
| GO:0042274 | ribosomal small subunit biogenesis | 0.0167 |
| GO:0070816 | phosphorylation of RNA polymerase II C-terminal domain | 0.0169 |
| GO:0042733 | embryonic digit morphogenesis | 0.0169 |
| GO:0033499 | galactose catabolic process via UDP-galactose | 0.0169 |
| GO:0006069 | ethanol oxidation | 0.0169 |
| GO:2000211 | regulation of glutamate metabolic process | 0.0169 |
| GO:0006014 | D-ribose metabolic process | 0.0169 |
| GO:0009448 | gamma-aminobutyric acid metabolic process | 0.0169 |
| GO:1903850 | regulation of cristae formation | 0.0169 |
| GO:0021532 | neural tube patterning | 0.0169 |
| GO:1904491 | protein localization to ciliary transition zone | 0.0169 |
| GO:0006850 | mitochondrial pyruvate transmembrane transport | 0.0169 |
| GO:0019243 | methylglyoxal catabolic process to D-lactate via S-lactoyl-glutathione | 0.0169 |
| GO:0006294 | nucleotide-excision repair, preincision complex assembly | 0.0169 |
| GO:0006351 | transcription, DNA-templated | 0.0184 |
| GO:0051028 | mRNA transport | 0.0227 |
| GO:0006096 | glycolytic process | 0.0259 |
| GO:0051171 | regulation of nitrogen compound metabolic process | 0.0334 |
| GO:0033387 | putrescine biosynthetic process from ornithine | 0.0336 |
| GO:0035735 | intraciliary transport involved in cilium assembly | 0.0336 |
| GO:0006529 | asparagine biosynthetic process | 0.0336 |
| GO:0016260 | selenocysteine biosynthetic process | 0.0336 |
| GO:0008612 | peptidyl-lysine modification to peptidyl-hypusine | 0.0336 |
| GO:0006109 | regulation of carbohydrate metabolic process | 0.0336 |
| GO:0030490 | maturation of SSU-rRNA | 0.0494 |
| GO:0006012 | galactose metabolic process | 0.0497 |
| GO:0007596 | blood coagulation | 0.0499 |
| GO:0006488 | dolichol-linked oligosaccharide biosynthetic process | 0.0499 |
| GO:0006446 | regulation of translational initiation | 0.0499 |
| GO:0016925 | protein sumoylation | 0.0499 |
| GO:0031591 | wybutosine biosynthetic process | 0.0499 |
| GO:0032436 | positive regulation of proteasomal ubiquitin-dependent protein catabolic process | 0.0499 |
| GO:0006379 | mRNA cleavage | 0.0499 |
| GO:0007602 | phototransduction | 0.0499 |
| GO:0016055 | Wnt signaling pathway | 0.0527 |
| GO:0080090 | regulation of primary metabolic process | 0.0657 |
| GO:0042749 | regulation of circadian sleep/wake cycle | 0.0660 |
| GO:0018298 | protein-chromophore linkage | 0.0660 |
| GO:0019509 | L-methionine salvage from methylthioadenosine | 0.0660 |
| GO:0007064 | mitotic sister chromatid cohesion | 0.0660 |
| GO:0006729 | tetrahydrobiopterin biosynthetic process | 0.0660 |
| GO:0006376 | mRNA splice site selection | 0.0660 |
| GO:0008643 | carbohydrate transport | 0.0810 |
| GO:0006450 | regulation of translational fidelity | 0.0818 |
| GO terms | Molecular functions | Fisher test |
| GO:0003735 | structural constituent of ribosome | 7.7e-22 |
| GO:0004499 | N,N-dimethylaniline monooxygenase activity | 0.0033 |
| GO:0016740 | transferase activity | 0.0036 |
| GO:0003899 | DNA-directed 5'-3' RNA polymerase activity | 0.0061 |
| GO:0003755 | peptidyl-prolyl cis-trans isomerase activity | 0.0094 |
| GO:0042132 | fructose 1,6-bisphosphate 1-phosphatase activity | 0.0185 |
| GO:0030628 | pre-mRNA 3'-splice site binding | 0.0185 |
| GO:0008108 | UDP-glucose:hexose-1-phosphate uridylyltransferase activity | 0.0185 |
| GO:0017061 | S-methyl-5-thioadenosine phosphorylase activity | 0.0185 |
| GO:0004045 | aminoacyl-tRNA hydrolase activity | 0.0185 |
| GO:0008010 | structural constituent of chitin-based larval cuticle | 0.0185 |
| GO:0034386 | 4-aminobutyrate:2-oxoglutarate transaminase activity | 0.0185 |
| GO:0004416 | hydroxyacylglutathione hydrolase activity | 0.0185 |
| GO:0050833 | pyruvate transmembrane transporter activity | 0.0185 |
| GO:0047631 | ADP-ribose diphosphatase activity | 0.0185 |
| GO:0002196 | Ser-tRNA(Ala) hydrolase activity | 0.0185 |
| GO:0051903 | S-(hydroxymethyl)glutathione dehydrogenase activity | 0.0185 |
| GO:0005506 | iron ion binding | 0.0312 |
| GO:0004497 | monooxygenase activity | 0.0316 |
| GO:0003978 | UDP-glucose 4-epimerase activity | 0.0366 |
| GO:0048027 | mRNA 5'-UTR binding | 0.0366 |
| GO:0004066 | asparagine synthase (glutamine-hydrolyzing) activity | 0.0366 |
| GO:0052856 | NADHX epimerase activity | 0.0366 |
| GO:0000384 | first spliceosomal transesterification activity | 0.0366 |
| GO:0000386 | second spliceosomal transesterification activity | 0.0366 |
| GO:0017147 | Wnt-protein binding | 0.0366 |
| GO:0030366 | molybdopterin synthase activity | 0.0366 |
| GO:0031624 | ubiquitin conjugating enzyme binding | 0.0366 |
| GO:0004756 | selenide, water dikinase activity | 0.0366 |
| GO:0004332 | fructose-bisphosphate aldolase activity | 0.0366 |
| GO:0008124 | 4-alpha-hydroxytetrahydrobiopterin dehydratase activity | 0.0366 |
| GO:0008273 | calcium, potassium:sodium antiporter activity | 0.0366 |
| GO:0004586 | ornithine decarboxylase activity | 0.0366 |
| GO:0042302 | structural constituent of cuticle | 0.0414 |
| GO:0016705 | oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen | 0.0418 |
| GO:0050661 | NADP binding | 0.0427 |
| GO:0020037 | heme binding | 0.0530 |
| GO:0004634 | phosphopyruvate hydratase activity | 0.0544 |
| GO:0009881 | photoreceptor activity | 0.0544 |
| GO:0004596 | peptide alpha-N-acetyltransferase activity | 0.0544 |
| GO:0004869 | cysteine-type endopeptidase inhibitor activity | 0.0544 |
| GO:0008569 | ATP-dependent microtubule motor activity, minus-end-directed | 0.0544 |
| GO:0070181 | small ribosomal subunit rRNA binding | 0.0544 |
| GO:0005549 | odorant binding | 0.0556 |
| GO:0045503 | dynein light chain binding | 0.0718 |
| GO:0031386 | protein tag | 0.0718 |
| GO:0019843 | rRNA binding | 0.0766 |
| GO:0051959 | dynein light intermediate chain binding | 0.0889 |
| GO:0051920 | peroxiredoxin activity | 0.0889 |
| GO:0003723 | RNA binding | 0.0996 |

Supplementary table 4: GO terms male-biased genes in L3

| GO terms | Molecular functions | Fisher test |
|------------|--|-------------|
| GO:0003924 | GTPase activity | 2.1e-07 |
| GO:0005525 | GTP binding | 9.3e-07 |
| GO:0005524 | ATP binding | 1.0e-06 |
| GO:0005509 | calcium ion binding | 1.8e-06 |
| GO:0003779 | actin binding | 4.4e-06 |
| GO:0042302 | structural constituent of cuticle | 1.2e-05 |
| GO:0004674 | protein serine/threonine kinase activity | 2.7e-05 |
| GO:0030552 | cAMP binding | 3.6e-05 |
| GO:0008603 | cAMP-dependent protein kinase regulator activity | 3.6e-05 |
| GO:0003700 | DNA-binding transcription factor activity | 7.0e-05 |
| GO:0017048 | Rho GTPase binding | 0.00067 |
| GO:0005219 | ryanodine-sensitive calcium-release channel activity | 0.00068 |
| GO:0017075 | syntaxin-1 binding | 0.00110 |
| GO:0048037 | cofactor binding | 0.00153 |
| GO:0019905 | syntaxin binding | 0.00321 |
| GO:0042813 | Wnt-activated receptor activity | 0.00323 |
| GO:0047555 | 3',5'-cyclic-GMP phosphodiesterase activity | 0.00323 |
| GO:0004675 | transmembrane receptor protein serine/threonine kinase activity | 0.00323 |
| GO:0001228 | DNA-binding transcription activator activity, RNA polymerase II-specific | 0.00633 |
| GO:0015293 | symporter activity | 0.01030 |
| GO:0004683 | calmodulin-dependent protein kinase activity | 0.01031 |
| GO:0015299 | solute:proton antiporter activity | 0.01031 |
| GO:0004252 | serine-type endopeptidase activity | 0.01099 |
| GO:0005506 | iron ion binding | 0.01368 |
| GO:0004715 | non-membrane spanning protein tyrosine kinase activity | 0.01513 |
| GO:0043565 | sequence-specific DNA binding | 0.01624 |
| GO:0005096 | GTPase activator activity | 0.01862 |
| GO:0003682 | chromatin binding | 0.02052 |
| GO:0031683 | G-protein beta/gamma-subunit complex binding | 0.02072 |
| GO:0046983 | protein dimerization activity | 0.02667 |
| GO:0003707 | steroid hormone receptor activity | 0.02731 |
| GO:0015297 | antiporter activity | 0.03304 |
| GO:0042562 | hormone binding | 0.03316 |
| GO:0030274 | LIM domain binding | 0.03324 |
| GO:0008521 | acetyl-CoA transmembrane transporter activity | 0.03324 |
| GO:0008139 | nuclear localization sequence binding | 0.03324 |
| GO:0015412 | ATPase-coupled molybdate transmembrane transporter | 0.03324 |
| GO:0019799 | tubulin N-acetyltransferase activity | 0.03324 |
| GO:0004367 | glycerol-3-phosphate dehydrogenase [NAD+] activity | 0.03324 |
| GO:0030215 | semaphorin receptor binding | 0.03324 |
| GO:0004350 | glutamate-5-semialdehyde dehydrogenase activity | 0.03324 |
| GO:0004597 | peptide-aspartate beta-dioxygenase activity | 0.03324 |
| GO:0008531 | riboflavin kinase activity | 0.03324 |
| GO:0035100 | ecdysone binding | 0.03324 |
| GO:0017172 | cysteine dioxygenase activity | 0.03324 |
| GO:0045127 | N-acetylglucosamine kinase activity | 0.03324 |
| GO:0019187 | beta-1,4-mannosyltransferase activity | 0.03324 |
| GO:0003938 | IMP dehydrogenase activity | 0.03324 |
| GO:0015408 | ferric-transporting ATPase activity | 0.03324 |
| GO:0004349 | glutamate 5-kinase activity | 0.03324 |
| GO:0004719 | protein-L-isoaspartate (D-aspartate) O-methyltransferase activity | 0.03324 |
| GO:0047874 | dolichyldiphosphatase activity | 0.03324 |
| GO:0004511 | tyrosine 3-monooxygenase activity | 0.03324 |
| GO:0005483 | soluble NSF attachment protein activity | 0.03324 |
| GO:0045499 | chemorepellent activity | 0.03324 |
| GO:0001102 | RNA polymerase II activating transcription factor binding | 0.03324 |
| GO:0015183 | L-aspartate transmembrane transporter activity | 0.03324 |
| GO:0004709 | MAP kinase kinase activity | 0.03324 |
| GO:0015189 | L-lysine transmembrane transporter activity | 0.03324 |
| GO:0003842 | 1-pyrroline-5-carboxylate dehydrogenase activity | 0.03324 |
| GO:0005243 | gap junction channel activity | 0.03324 |
| GO:0070840 | dynein complex binding | 0.03324 |
| GO:0008832 | dGTPase activity | 0.03324 |
| GO:0004485 | methylcrotonoyl-CoA carboxylase activity | 0.03324 |
| GO:0004861 | cyclin-dependent protein serine/threonine kinase inhibitor activity | 0.03324 |
| GO:0008191 | metalloendopeptidase inhibitor activity | 0.03324 |
| GO:0036422 | heptaprenyl diphosphate synthase activity | 0.03324 |
| GO:0000064 | L-ornithine transmembrane transporter activity | 0.03324 |
| GO:0005385 | zinc ion transmembrane transporter activity | 0.03324 |
| GO:0047952 | glycerol-3-phosphate dehydrogenase [NAD(P)+] activity | 0.03324 |
| GO:0005313 | L-glutamate transmembrane transporter activity | 0.03324 |
| GO:0008349 | MAP kinase kinase kinase activity | 0.03324 |
| GO:0004672 | protein kinase activity | 0.03353 |
| GO:0042803 | protein homodimerization activity | 0.03399 |
| GO:0005089 | Rho guanyl-nucleotide exchange factor activity | 0.03514 |
| GO:0000166 | nucleotide binding | 0.03588 |
| GO:0061630 | ubiquitin protein ligase activity | 0.03924 |
| GO:0003774 | motor activity | 0.03964 |
| GO:0005267 | potassium channel activity | 0.04146 |
| GO:0005484 | SNAP receptor activity | 0.04157 |
| GO:0046872 | metal ion binding | 0.04561 |
| GO:0036459 | thiol-dependent ubiquitinyl hydrolase activity | 0.04880 |

Supplementary table 5: GO terms female-biased genes in L3

| GO terms | Biological processes | Fisher test |
|------------|---|-------------|
| GO:0006030 | chitin metabolic process | 5.9e-08 |
| GO:0055085 | transmembrane transport | 0.0015 |
| GO:0099133 | ATP hydrolysis coupled anion transmembrane transport | 0.0018 |
| GO:0008643 | carbohydrate transport | 0.0044 |
| GO:0006629 | lipid metabolic process | 0.0093 |
| GO:1900426 | positive regulation of defense response to bacterium | 0.0111 |
| GO:0009847 | spore germination | 0.0111 |
| GO:0046426 | negative regulation of JAK-STAT cascade | 0.0111 |
| GO:0010469 | regulation of signaling receptor activity | 0.0171 |
| GO:0016042 | lipid catabolic process | 0.0184 |
| GO:0090502 | RNA phosphodiester bond hydrolysis, endonucleolytic | 0.0198 |
| GO:0006564 | L-serine biosynthetic process | 0.0222 |
| GO:0042438 | melanin biosynthetic process | 0.0222 |
| GO:0023052 | signaling | 0.0328 |
| GO:0015718 | monocarboxylic acid transport | 0.0328 |
| GO:0030514 | negative regulation of BMP signaling pathway | 0.0330 |
| GO:0051592 | response to calcium ion | 0.0330 |
| GO:0015716 | organic phosphonate transport | 0.0330 |
| GO:0055114 | oxidation-reduction process | 0.0425 |
| GO:0032506 | cytokinetic process | 0.0438 |
| GO:0060179 | male mating behavior | 0.0545 |
| GO:0001558 | regulation of cell growth | 0.0545 |
| GO:0006635 | fatty acid beta-oxidation | 0.0545 |
| GO:0008152 | metabolic process | 0.0557 |
| GO:0080090 | regulation of primary metabolic process | 0.0639 |
| GO terms | Molecular functions | Fisher test |
| GO:0042302 | structural constituent of cuticle | 6.5e-10 |
| GO:0008061 | chitin binding | 3.9e-07 |
| GO:0004099 | chitin deacetylase activity | 0.0009 |
| GO:0004312 | fatty acid synthase activity | 0.0052 |
| GO:0008237 | metallopeptidase activity | 0.0052 |
| GO:0022857 | transmembrane transporter activity | 0.0072 |
| GO:0004806 | triglyceride lipase activity | 0.0108 |
| GO:0003796 | lysozyme activity | 0.0124 |
| GO:0004647 | phosphoserine phosphatase activity | 0.0124 |
| GO:0035673 | oligopeptide transmembrane transporter activity | 0.0124 |
| GO:0008269 | AK pathway signal transduction adaptor activity | 0.0124 |
| GO:0005159 | insulin-like growth factor receptor binding | 0.0124 |
| GO:0008113 | peptide-methionine (S)-S-oxide reductase activity | 0.0124 |
| GO:0004447 | iodide peroxidase activity | 0.0124 |
| GO:0102567 | phospholipase A2 activity (consuming 1,2-dipalmitoylphosphatidylcholine) | 0.0124 |
| GO:0102568 | phospholipase A2 activity consuming 1,2-dioleoylphosphatidylethanolamine) | 0.0124 |
| GO:0044548 | S100 protein binding | 0.0124 |
| GO:0004180 | carboxypeptidase activity | 0.0161 |
| GO:0008429 | phosphatidylethanolamine binding | 0.0246 |
| GO:0015419 | ATPase-coupled sulfate transmembrane transporter activity | 0.0246 |
| GO:0004613 | phosphoenolpyruvate carboxykinase (GTP) activity | 0.0246 |
| GO:0031177 | phosphopantetheine binding | 0.0246 |
| GO:0008138 | protein tyrosine/serine/threonine phosphatase activity | 0.0246 |
| GO:0048306 | calcium-dependent protein binding | 0.0246 |
| GO:0005200 | structural constituent of cytoskeleton | 0.0249 |
| GO:0016491 | oxidoreductase activity | 0.0254 |
| GO:0005507 | copper ion binding | 0.0298 |
| GO:0015416 | ATPase-coupled organic phosphonate transmembrane transporter activity | 0.0366 |
| GO:0004344 | glucose dehydrogenase activity | 0.0366 |
| GO:0005000 | vasopressin receptor activity | 0.0366 |
| GO:0008028 | monocarboxylic acid transmembrane transporter activity | 0.0366 |
| GO:0004198 | calcium-dependent cysteine-type endopeptidase activity | 0.0485 |
| GO:0008253 | 5'-nucleotidase activity | 0.0485 |
| GO:0004523 | RNA-DNA hybrid ribonuclease activity | 0.0492 |
| GO:0005179 | hormone activity | 0.0591 |

Supplementary table 6: GO terms L3-biased genes in males

| Sample | Reads | Hisat2 Overall Alignment rate | Hisat2 Aligned concordantly exactly 1 times | Hisat2 Aligned concordantly exactly 1 times |
|-----------|----------|-------------------------------|---|---|
| L1-F-B-R1 | 49041461 | 90.11 | 36228560 | 73.87 |
| L1-F-B-R2 | 47963890 | 90.31 | 35271572 | 73.54 |
| L1-F-B-R3 | 48099242 | 90.98 | 36902758 | 76.72 |
| L1-F-S-R1 | 48667273 | 88.80 | 34906909 | 71.73 |
| L1-F-S-R2 | 48484086 | 88.99 | 36083617 | 74.42 |
| L1-F-S-R3 | 48602231 | 89.60 | 35121277 | 72.26 |
| L1-M-B-R1 | 49233285 | 90.68 | 36152045 | 73.43 |
| L1-M-B-R2 | 48633078 | 91.47 | 34500256 | 70.94 |
| L1-M-B-R3 | 48512675 | 91.12 | 36441131 | 75.12 |
| L1-M-S-R1 | 48221191 | 90.29 | 34051512 | 70.62 |
| L1-M-S-R2 | 48840237 | 90.14 | 33977226 | 69.57 |
| L1-M-S-R3 | 48226775 | 89.34 | 34343220 | 71.21 |
| L2-F-B-R1 | 48952516 | 89.91 | 36891575 | 75.36 |
| L2-F-B-R2 | 48841398 | 90.34 | 38640035 | 79.11 |
| L2-F-B-R3 | 49065737 | 91.05 | 35395488 | 72.14 |
| L2-F-S-R1 | 47979225 | 88.87 | 33811193 | 70.47 |
| L2-F-S-R2 | 48786435 | 89.33 | 36592806 | 75.01 |
| L2-F-S-R3 | 48428508 | 89.55 | 34951145 | 72.17 |
| L2-M-B-R1 | 49067171 | 91.33 | 35823353 | 73.01 |
| L2-M-B-R2 | 48777017 | 91.13 | 34544156 | 70.82 |
| L2-M-B-R3 | 49099190 | 90.94 | 36829187 | 75.01 |
| L2-M-S-R1 | 48462648 | 90.22 | 34504457 | 71.2 |
| L2-M-S-R2 | 48625263 | 89.95 | 34783241 | 71.53 |
| L2-M-S-R3 | 49037014 | 90.38 | 33930732 | 69.19 |
| L3-F-B-R1 | 48968087 | 90.11 | 36829672 | 75.21 |
| L3-F-B-R2 | 48840162 | 90.44 | 37072881 | 75.91 |
| L3-F-B-R3 | 48703331 | 90.84 | 37624757 | 77.25 |
| L3-F-S-R1 | 48910387 | 88.60 | 36033625 | 73.67 |
| L3-F-S-R2 | 48710590 | 89.38 | 36662038 | 75.27 |
| L3-F-S-R3 | 48508285 | 89.34 | 36564719 | 75.38 |
| L3-M-B-R1 | 48756870 | 90.78 | 35760135 | 73.34 |
| L3-M-B-R2 | 48814129 | 91.27 | 34850291 | 71.39 |
| L3-M-B-R3 | 48911065 | 90.88 | 35754802 | 73.1 |
| L3-M-S-R1 | 48691572 | 90.95 | 32650548 | 67.06 |
| L3-M-S-R2 | 48400141 | 90.20 | 33402563 | 69.01 |
| L3-M-S-R3 | 48763992 | 90.66 | 34644144 | 71.04 |

Supplementary table 7: Transcriptome metrics with the number of reads per library and the alignment rate of these reads on the *M. longipes* genome.

Summary and general conclusions

About five years ago, during a field trip to French Guyana, we got the chance to come across a species of water striders with an extreme case of sexual dimorphism. Soon later, we decided to embark into a new project where *Microvelia longipes* would be used as a new model organism to study the process of phenotypic variation within and between the sexes. Its short generation time, the easiness to maintain large populations in lab conditions and its tractable development made us rapidly think that *M. longipes* could be a good model for a integrative studies of the evolution and maintenance of intraspecies variation. In this line, our first approach has been to characterize the biology of this species and to understand the possible adaptive significance of the exaggerated leg length in males. Recreating their natural habitats with artificial puddles turned out to be a key experiment to discover that *M. longipes* males are highly active sexually. Intense competition, long-term signaling on egg-laying substrates, vehement guarding over egg-laying females; males use all sorts of pre-copulatory strategies to access females and fertilize eggs. Under these conditions, males developing longer rear legs get an obvious advantage in accessing females as they become better fighters and have more chance to dominate egg-laying substrates. Such directional selection is, nonetheless, thought to erase genetic variation and consequently the variance in phenotypes. Yet, males with drastically reduced rear leg lengths are still present in the population. We found that, at the population level, selection over the egg-laying substrates may not be as intense as expected. In an experiment where we manipulated the number of egg-laying floaters we found that the intensity of sexual selection can fluctuate and sometimes allow males with shorter legs to fertilize a substantial number of eggs. Relaxed selection is known to strongly influence the evolution and maintenance of phenotypic plasticity [170, 171] and may have played a similar role in the rear leg lengths of *M. longipes* males that are highly dependent on nutrition intake.

In a second approach, we aimed to decipher the genetic and genomic architecture of this exaggerated trait under directional sexual selection. We knew from morphological analyses and developmental tracking that the hypervariability observed in male rear leg lengths resulted from the evolution of a hyperallometric growth of the legs that was taking place at the end of nymphal development. We therefore conducted a comparative transcriptomic approach on the fifth instars, using two major factors illustrating the differences in allometric scaling in *M. longipes*, namely the leg and sex comparisons. Interestingly, both comparisons

presented a signature of leg exaggeration among the differentially expressed genes. For the sex-biased genes, we found a higher number of sex-biased genes and a masculinization of male-biased gene expression in the exaggerated legs compared to the legs with isometric scaling relationship (non-exaggerated legs). A high proportion of leg-biased genes shows also a pattern of male-specific upregulation in the rear legs. Importantly, several of these genes also exhibit male-biased expression making a direct link between gene expression and the development of exaggerated legs. These analyses also revealed obvious candidate genes involved in cell metabolism and tissue growth. Finally, we aimed to assess how directional sexual selection in male rear leg length may have influenced the architecture of *M. longipes* genome. As predicted by theory, we found an excess of female-biased genes on the X chromosome but not male-biased genes. The sexualization of the X chromosome has often been a source of debate regarding its adaptive significance. Our comparison with the two other legs, which are not under the same type or degree of sexual selection, showed no overrepresentation of female-biased genes. This indicates that sexual selection may therefore actively influence the sexualization of the X chromosome. More generally, we found that sexual selection may have a broad influence on genome architecture by creating clusters of sex-biased genes across the genome, that are shared or not between legs.

Overall, *M. longipes* is an emerging model system that we use to get a more integrative understanding on how selection and genetics shape the evolution of sexual dimorphism (Figure conclusion).

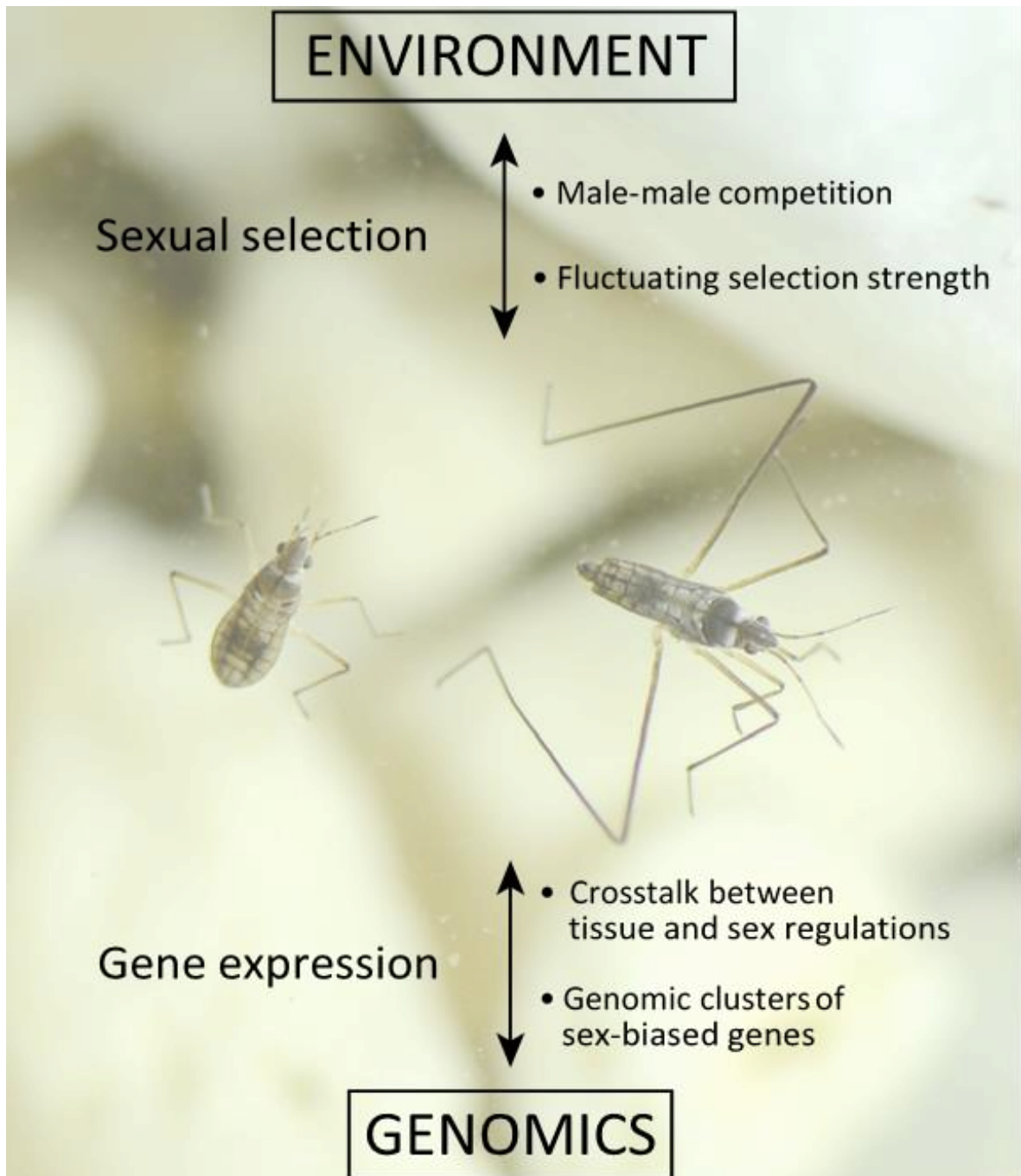


Figure conclusion: Conclusion figure illustrating the integrative approach to link the ecology, evolution and developmental genomics (Eco-Evo-Devo) to understand the elaboration of male exaggerated legs in the water strider *Microvelia longipes*

General discussion and future directions

I. The influence of selection on sexual trait evolution

A. Moving from reproductive fitness to a more integrative view of fitness

Fitness is a complex concept characterizing the overall ability of an individual to transmit its alleles to the next generation [20]. In nature, fitness is a multifaceted concept describing how multiple genotypes (or individuals) cope with various selective pressures during their lifetime. In *M. longipes*, for example, some genotypes in the population could be better at reproducing while some others would have better success in surviving. A trade-off between two fitness components, namely mating success and survival, could therefore dramatically influence the overall fitness (also named lifetime fitness) of each individual and ultimately lead to the maintenance of the phenotypic variation observed in natural populations. Given the sex-specificity of leg exaggeration and the lack of information about the biology of *M. longipes*, we started our approach by examining the possible reproductive advantage associated with the exaggerated leg. Today, it would be interesting to extend our study to a more integrative approach, involving several components of fitness. In this regard, we still don't know whether the variation in male leg length is associated with variation in survival. A mark-recapture experiment involving several male samples from different populations could help determine whether an association exists between leg length and adult survival in natural conditions. Developing disproportionate legs in large males may also be associated with a high metabolic cost and a constraint on the molting process that would ultimately lower the survival rate. A mark-recapture experiment would not be possible in this case, as juveniles renew their entire cuticle at every molt. Isolating juveniles from natural populations to follow their development could help answering this question. Our field observations also strongly suggest that the puddles in which *M. longipes* lives, are temporary habitats. A relatively large puddle in the field can evaporate entirely in a few days. Such instable environment is associated with wing polymorphism during nymphal development, allowing probably adults to disperse from one puddle to another. Wing polyphenism is known to be an adaptive mechanism used by individuals under stressful environments but it is also an energetic costly structure to develop. How wing polyphenism influences the resource allocation to rear leg exaggeration in *M. longipes* is unknown. Similarly, how leg exaggeration predisposes individual's ability to disperse is also a line of research that needs to be explored.

Much needs to be also explored on the third component of fitness, namely individual's fertility. It is now well established that sexual selection does not only act before mating but also after copulation [228]. This post-copulatory sexual selection, often manifested by sperm competition, where males develop various adaptations to increase their chance to fertilize eggs [229]. In some cases, this can even lead to some examples of exaggerations [230]. Males have also evolved several defensive strategies such as male-guarding or the release of some toxic seminal substances to cope with this post-copulatory competition. In *M. longipes* males also have evolved intense male-guarding by which they will turn around the female until she stops laying eggs. Such behavior is indicative of strong competition after mating and may also influence the maintenance of short-legged males in the population.

As detailed in the introduction, the definition of fitness can also be envisioned from a more probabilistic perspective. In this case, fitness would illustrate the variation in frequency of a certain allele or genotype in the population over one to many generations [22]. One way to investigate the fitness influence of leg exaggeration could also be by monitoring in the population the relative fitness of alleles underlying differences in relative rear leg length in males. Although such approach is not currently accessible in *M. longipes*, the development of a high quality genome and the relatively easy access to natural populations in the field, make feasible the identification of genetic variants associated with variation in relative leg length in males (e.g. Genome Wide Association Mapping, QTLs...). Identifying the genetic variants and following their change in frequency over several generations in natural populations would give a first insight into the relationship between genotype, phenotype and relative fitness.

B. How does variation in sexual selection strength influence the evolution of trait exaggeration?

The literature often describes sexual selection as a directional and persistent evolutionary force driving sexual traits to exaggeration [19, 38, 39, 126]. A growing number of studies, however, raise the idea that sexual selection may be more fluctuating than previously thought [151-154, 231]. This is the case of our study where the strength of sexual selection seems to vary within and between species. In *M. longipes*, the increase in egg-laying sites is associated with the increase in fertilization success of short-legged males in the population. Although selection is directional around egg-laying sites (favoring individuals with longer rear legs),

the variation in the number of accessible sites may lead in nature to periods of relaxed selection where small males have higher chance to mate. Similarly, *M. pulchella* has evolved the same mating behavior by which males that win contests over floaters, dominate them and increase their mating success. We found that the lack of exaggeration in this species is associated with a lower intensity of male fight, which ultimately results in a decrease of selection strength. Despite these evidences linking the strength of sexual selection to the evolution of trait exaggeration, we still lack a formal experimental study to test it.

The ease to culture *Microvelia* species in laboratory conditions makes possible to test different selective pressures in a long-term evolutionary experiment. Our study indicates that egg-laying sites could be an important ecological factor influencing the strength of sexual selection. Manipulating the number of floaters in controlled populations over many generations could shed light on the relationship between sexual selection and sexual trait elaboration in the context of male competition. An evolutionary experiment with different *Microvelia* species (e.g. *M. longipes* and *M. pulchella*) where the populations would evolve under either excess or restricted number of egg-laying sites would allow to compare the effects of selection on species with different levels of trait exaggeration. It is important to note that such experiment is possible when there is standing genetic variation in the population of interest. In the case of *M. longipes*, the difference in phenotypes observed between the selected inbred populations indicates that genetic variation is still present for both relative and absolute rear leg length in natural populations. Another approach to overcome this possible limitation would be to generate artificially mutagenized populations that could be treated with chemicals such as Ethyl methanesulfonate (EMS) or X-rays. In addition to increase the genetic variance of a population, this method could possibly generate extreme phenotypes in rear leg length that would otherwise be erased by natural selection. The choice of mutagenized populations would be particularly relevant in regards to relative leg length, which displays relatively low variation in natural population (*M. longipes* allometry, $R^2=0.92$; Figure 1B chapter 1).

Finally, the relatively good description of the reproductive behavior in *M. longipes* allows now the development of modeling that could simulate the evolution of male leg length under various selective pressures (e.g. numbers of floaters) and explore in more details the role of certain environmental factors on the evolution and maintenance of this sexual character.

II. Sexual selection and the evolution of gene expression

Secondary sexual traits fascinated naturalists for centuries and have become one of the main foci in the field of evolutionary ecology. In contrast, relatively little is known about the genetic and developmental regulation of these traits. However, the advancement in sequencing technologies, recently allowed evolutionary geneticists to fill this gap. This new field of research allowed notably the discovery of an unexpectedly large number of genes with sex-biased expressions, constituting sometimes more than half of the total number of expressed genes in an organism or specific tissue [106]. An increasing number of transcriptomic studies interested in understanding the evolution of sex-biased genes and their link to the phenotypic dimorphism in males and females emerged recently [102, 106]. Many of these studies, however, rely on adults or whole-body transcriptomes to assess the relationship between sexual selection, gene regulation and sexual dimorphism. Although these gene regulatory patterns may indeed underlie a certain difference between sexes, it is still unclear to which extent they contribute to it [106]. Our approach in *M. longipes* was therefore designed to directly link the observed patterns of sex-biased expression to the actual elaboration of a specific exaggerated tissue, known to be under sexual selection in males. Consequently, we could first shed light on the developmental regulatory processes associated with the elaboration of sexually selected traits and second decipher the link between sexual selection and the evolution of genome architecture in the context of a sexually dimorphic phenotype. Adding another layer of comparison, taking into account the evolutionary change in sex-biased gene expression, could improve our genetic understanding of sexual dimorphism. For example, how sex-biased genes in *M. longipes* are expressed in other *Microvelia* species lacking exaggeration in the legs is still unknown. In this respect, *M. pulchella* is a close relative species to *M. longipes* and seems therefore the most appropriate to use in a comparative transcriptomics approach.

Our data also revealed that the development of a secondary sexual character may not always be associated with the same gene regulatory patterns. In the first leg of *M. longipes*, a sex-comb also develops specifically in males. Yet, we found almost no genes with male-biased and L1-biased expression in the first legs. This suggests that the regulatory mechanism controlling the development of a secondary sexual trait may strongly depend on the nature of the trait. This hypothesis could be tested with the diversity of sexual characters that have evolved in *Microvelia* rear legs. A striking difference is, for example, observed in *M. sp.*

where males have invested in a series of spines along the femur and tibia. How genetically similar is this sexual character compared to the leg exaggeration in *M. longipes* is an open question. Comparative transcriptomics between the two species could help disentangle the set of genes common to the general establishment of sexual dimorphism and the genes involved in sexual diversification. Overall, extending the study of sex-biased gene expressions to other organisms will help to better characterize the genetic regulation of sexual dimorphism and its evolution.

III. Understanding the genetics underpinning of leg exaggeration through genetic mapping.

Understanding the genetic basis of exaggeration in sexually selected traits is challenging, as it is particularly difficult to cross species with different degrees of exaggeration. As a consequence there is no study, at least to my knowledge, which has ever reported any cross between species with different degrees of exaggerated weapons or ornaments. Selection experiments have not been very successful either, often reporting too small or transient changes in scaling relationship, for further genetic mapping [232-234]. The recent advances in high throughput sequencing technologies offer the opportunity to sequence the genome of entire natural populations, for which the extent of genetic variation and crossing overs is often more important than artificially crossed populations. These new techniques, also called Genome Wide Association Mapping (GWAS), allow a higher mapping resolution and can deal with relatively small phenotypic variance. A sampling of multiple males from the same or different *M. longipes* populations could be used to assess the relationship between genetic variation and variation in male relative rear leg length. This approach suffers, nonetheless, by the limited amount of variation present in the population. Random mutagenesis appears once again to be a possible alternative to generate unusual extreme phenotypes in the population. Treating *M. longipes* natural populations with UV radiation or mutagenic chemicals could generate a F2 population with higher variation in relative leg length, which could be later used in a mapping study.

Changes in scaling relationships are the main drivers of evolution of forms in nature, yet we know very little about the genetic mechanisms underlying their evolution [42]. In the future,

geneticists and developmental biologists should therefore put a large emphasis on this research aspect.

Summary of other scientific contributions

I. Fitness Effects of Cis-Regulatory Variants in the *Saccharomyces cerevisiae* TDH3 Promoter.

Fabien Duveau, William Toubiana, Patricia J. Wittkopp. *Molecular Biology and Evolution* 2017.

Variation in gene expression is one of the main sources of phenotypic evolution and is genetically determined by cis- and trans-regulatory mutations [235]. Although a large literature describes the role of gene expression in species divergence, little is known regarding the fitness consequence of such variation within species. We used mutations in the *Saccharomyces cerevisiae* TDH3 promoter to assess how changes in TDH3 expression affect cell growth. We found a non-linear relationship between the relative TDH3 promoter activity and the relative fitness. Moreover, we show that nearly all mutations and polymorphisms in the TDH3 promoter have no significant effect on fitness in the environment assayed, suggesting that the wild-type allele of this promoter is robust to the effects of most new cis-regulatory mutations [236].

I specifically contributed to this study by generating the mutant strains, developing and performing the fitness assay experiment, analyzing the data and commenting on the manuscript.

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

David Armisen, Peter Refki, Antonin Crumière, Severine Viala, William Toubiana, Abderrahman Khila. *Nature communications* 2015.

Understanding the ecological factors and genetic mechanisms underlying the evolution of adaptive phenotypes is a major challenge in evolutionary developmental biology (Evo-Devo). The use of the same genes in a different context (e.g. developmental stage or tissue) is often thought to be a mechanism favored by evolution to generate new adaptive phenotypes. However, relatively little is known about the genetic mechanisms underlying this process [237]. We show in this study that the gene *gilt*, previously characterized to play a role in immunity, has acquired a function in leg elongation in some water striders through its new interaction with the Hox gene *Ultrabithorax*. The differences in leg morphologies are established through modulation of *gilt* differential expression between mid and hindlegs under *Ubx* control. Finally, behavioral assays revealed that short-legged water striders generated through *gilt* RNAi, exhibit reduced performance in jumping to escape attacks by bottom-striking predators [67]. Overall, our results illustrate how divergence in selective pressures, through new prey-predator interaction, can shape new genetic interactions that influence adaptive evolution.

I specifically contributed to this study by monitoring *gilt* expression in water strider legs throughout embryonic development and commenting on the manuscript.

III. The mlpt/Ubr3/Svb module comprises an ancient developmental switch for embryonic patterning.

Suparna Ray, Miriam I Rosenberg, Hélène Chanut-Delalande, Amélie Decaras, Barbara Schwertner, William Toubiana, Tzach Auman, Irene Schnellhammer, Matthias Teuscher, Philippe Valenti, Abderrahman Khila, Martin Klingler, François Payre. *Elife* 2019.

The formation of body plans is a key process in animal development but it is also strikingly variable between species. A good example is the mode of segmentation where some insects pattern all segments nearly simultaneously (long germband) while others add them sequentially from a posterior growth zone (short germband) [238]. The molecular mechanisms underlying these different modes of segmentations remain an active field of research in evolutionary developmental biology. In this study we show that the molecular complex mille-pattes peptides (Mlpt), ubiquitin-ligase Ubr3 and Shavenbaby, identified in *Drosophila* epidermal differentiation, represents an ancient developmental module required for early insect embryo patterning. The knock down of this module in insect species that retained the ancestral short germband segmentation process exhibited segmentation defects. In early *Drosophila* embryos that develop through the derived mode of segmentation, *Shavenbaby* expression is restricted to the anterior pole. Over-expression of the Shavenbaby active form (processed by Mlpt and Ubr3) in early *Drosophila* embryos causes segmentation defects while the repressive form (unprocessed by Mlpt and Ubr3) did not cause segmentation defects. Our results therefore highlight the evolution of a key developmental switch and its consequences on the embryonic patterning modes in insects.

I specifically contributed to this study by cloning the genes from the complex in the water strider species and commenting on the manuscript.

IV. The genome of the water strider *Gerris buenoi* reveals expansions of gene repertoires associated with adaptations to life on the water.

David Armisén, Rajendhran Rajakumar, Markus Friedrich, Joshua B. Benoit, Hugh M. Robertson, Kristen A. Panfilio, SeungJoon Ahn, Monica F. Poelchau, Hsu Chao, Huyen Dinh, Harsha Vardhan Doddapaneni, Shannon Dugan, Richard A. Gibbs, Daniel S. T. Hughes, Yi Han, Sandra L. Lee, Shwetha C. Murali, Donna M. Muzny, Jiaxin Qu, Kim C. Worley, Monica Munoz-Torres, Ehab Abouheif, François Bonneton, Travis Chen, Li-Mei Chiang, Christopher P. Childers, Andrew G. Cridge, Antonin J. J. Crumière, Amelie Decaras, Elise M. Didion, Elizabeth J. Duncan, Elena N. Elpidina, Marie-Julie Favé, Cédric Finet, Chris G. C. Jacobs, Alys M. Cheattle Jarvela, Emily C. Jennings, Jeffery W. Jones, Maryna P. Lesoway, Mackenzie R. Lovegrove, Alexander Martynov, Brenda Oppert, Angelica Lillico Ouachour, Arjuna Rajakumar, Peter Nagui Refki, Andrew J. Rosendale, Maria Emilia Santos, William Toubiana, Maurijn vanderZee, Iris M. VargasJentsch, Aidamalia Vargas Lowman, Severine Viala, Stephen Richards and Abderrahman Khila. *BMC Genomics* 2018.

Water striders form a group of semi-aquatic insects that have conquered a variety of water surface habitats, from small drops on a leaf to open oceans. The diversity of this family is the subject of various scientific studies encompassing a large panel of research fields, from ecology and evolution to developmental genetics and hydrodynamics of fluid locomotion [239]. However, the lack of a representative genome hinders our progress to better understand the selective pressures and molecular mechanisms underlying the adaptation of water striders. We therefore conducted a collaborative work aiming to sequence, assemble and manually annotate the genome of the water strider *Gerris buenoi*. Manual annotation uncovered several gene duplications and expansion of certain gene families in association with water surface adaptation such as growth, bristle development, color vision or desiccation resistance. This study opens new areas of research to characterize the adaptive mechanisms associated with the life on water surface.

I specifically contributed to this study by manually annotating some of the predicted genes in the genome and commenting on the manuscript.

Acknowledgements

Let me end up this manuscript by acknowledging all the people that have been contributing, in so many different aspects, to this tough but so enriching period that was my PhD.

My first thanks will of course go to my Boss, as I like to name him. Abdou has been my scientific mentor all along these four years and none of it would have been possible without his enthusiasm, support, ideas, suggestions and most importantly his patience. Abdou gave me the opportunity to work of my passion and guided me through all the difficulties of a PhD, I would never be thankful enough for this.

Then comes the members of the lab, past and present, my companions in misery when I was frustrated. But I will mostly remember you for all the crazy laughs, all these interesting and stupid discussions around beers, all your advices, all these times brainstorming about science and life generally in the office or the bug room. And I am probably missing so many other things... Thank you for all.

I would also like to greatly thank all the collaborators that have contributed to the different projects mentioned in this manuscript. I would like to specifically mention in this section the people that have collaborated on the *M. longipes* project, the co-authors but also my PhD committee. Their contribution has been of tremendous help for the development of this new model system.

Of course I will not forget all the members of my institute. Thank you to all the research teams for their helpful discussions and advices, to the people from the research support team in IGFL who spent a lot of time and effort making my administrative and scientific life easier in the institut.

Finally, comes the family and friends who have been supportive not only during my PhD but all along my life. First, Amélie, thank you for all your patience and understanding. This has been a tough period but we went through it and nothing would have been possible without you by my side. My parents and sister, who made what I am today; no words can describe what I feel and what I owe you. I will end by my friends who have been my source of laughs and zest for life.

References

1. Bowler P. Evolution: The history of an idea. Berkeley, University of California Press ed1984.
2. Lamarck J-Bd. Zoological Philosophy, or Exposition with Regard to the Natural History of Animals: Museum d'Histoire Naturelle (Jardin des Plantes); 1809.
3. Darwin C., Wallace A. On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. Journal of the Proceedings of the Linnean Society of London. 1858.
4. Darwin C. On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life. London: J. Murray; 1859. ix, [1], 502, 32 , [1] fold. leaf of plates (32 at end advertisements) p.
5. Darwin C. The descent of man, and selection in relation to sex. London,: J. Murray; 1871. 2 v. p.
6. Mendel G. Experiments in Plant Hybridization. Brit Med J. 1965;1(5431):370-&. PubMed PMID: WOS:A19656118700022.
7. Sutton WS. On the morphology of the chromosome group in *Brachystola magna*. Biol Bull-US. 1902;4(1):24-39. doi: Doi 10.2307/1535510. PubMed PMID: WOS:000201839900003.
8. Morgan TH, Sturtevant A. H, Muller H. J, Bridges C. B. The Mechanism of Mendelian heredity. Company NYHHa, editor1915.
9. Sturtevant A. H. A History of Genetics. Cold Spring Harbor Laboratory Press US, editor1965.
10. Fisher RA. The genetical theory of natural selection. Oxford,: The Clarendon Press; 1930. xiv, 272 p. p.
11. Huxley J. Evolution: the modern synthesis. Limited GAU, editor1942.
12. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737-8. PubMed PMID: 13054692.
13. Carroll S. B, Grenier J. K, Weatherbee S. D. From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design. 2nd edition ed: Wiley-Blackwell; 2004.
14. Cuvier G, Jameson R, Mitchill S, L. . Essay on the Theory of the Earth: New York, Kirk & Mercein; 1818.
15. Masatoshi N. Molecular Evolutionary Genetics: Columbia University Press, New York; 1987.
16. Kimura M. The Neutral Theory of Molecular Evolution: Cambridge University Press; 1983.
17. Lynch M, Hill WG. Phenotypic Evolution by Neutral Mutation. Evolution. 1986;40(5):915-35. doi: 10.1111/j.1558-5646.1986.tb00561.x. PubMed PMID: 28556213.
18. Smith CH. Note on Sexual Selection (S459: 1892). Available from: <http://people.wku.edu/charles.smith/wallace/S459.htm>.
19. Andersson MB. Sexual selection. Princeton, N.J.: Princeton University Press; 1994. xx, 599 p. p.
20. Orr HA. Fitness and its role in evolutionary genetics. Nat Rev Genet. 2009;10(8):531-9. doi: 10.1038/nrg2603. PubMed PMID: 19546856; PubMed Central PMCID: PMC2753274.
21. Gilbert L. Insect Development: Morphogenesis, Molting and Metamorphosis: Academic Press; 2009.

22. Crow J. F, Kimura M. An introduction to population genetics theory: Minneapolis : Burgess Pub. Co.; 1970.
23. Ewens W. Mathematical Population Genetics; 1979.
24. Fritzsche K, Booksmythe I, Arnqvist G. Sex Ratio Bias Leads to the Evolution of Sex Role Reversal in Honey Locust Beetles. *Curr Biol*. 2016;26(18):2522-6. doi: 10.1016/j.cub.2016.07.018. PubMed PMID: WOS:000384512600031.
25. Bonduriansky R. The evolution of male mate choice in insects: a synthesis of ideas and evidence. *Biological Reviews*. 2001;76(3):305-39. doi: Doi 10.1017/S1464793101005693. PubMed PMID: WOS:000171032000002.
26. Edward DA, Chapman T. The evolution and significance of male mate choice. *Trends Ecol Evol*. 2011;26(12):647-54. doi: 10.1016/j.tree.2011.07.012. PubMed PMID: WOS:000297777500011.
27. Bateman AJ. Intra-sexual selection in *Drosophila*. *Heredity (Edinb)*. 1948;2(Pt. 3):349-68. PubMed PMID: 18103134.
28. Arnqvist G, Rowe L. Sexual conflict. Princeton, N.J.: Princeton University Press; 2005. xii, 330 p. p.
29. Rowe L, Day T. Detecting sexual conflict and sexually antagonistic coevolution. *Philos T R Soc B*. 2006;361(1466):277-85. doi: 10.1098/rstb.2005.1788. PubMed PMID: ISI:000235238200005.
30. Bonduriansky R, Chenoweth SF. Intralocus sexual conflict. *Trends Ecol Evol*. 2009;24(5):280-8. Epub 2009/03/25. doi: S0169-5347(09)00075-5 10.1016/j.tree.2008.12.005. PubMed PMID: 19307043.
31. Gould SJ. Allometry and Size in Ontogeny and Phylogeny. *Biological Reviews*. 1966;41(4):587-+. doi: DOI 10.1111/j.1469-185X.1966.tb01624.x. PubMed PMID: WOS:A19668557300003.
32. Gould SJ. Positive Allometry of Antlers in Irish Elk, *Megaloceros-Giganteus*. *Nature*. 1973;244(5415):375-6. doi: DOI 10.1038/244375a0. PubMed PMID: WOS:A1973Q347700047.
33. Gould SJ. Origin and Function of Bizarre Structures - Antler Size and Skull Size in Irish Elk, *Megaloceros-Giganteus*. *Evolution*. 1974;28(2):191-220. doi: DOI 10.1111/j.1558-5646.1974.tb00740.x. PubMed PMID: WOS:A1974T392500003.
34. Bonduriansky R. Sexual selection and allometry: A critical reappraisal of the evidence and ideas. *Evolution*. 2007;61(4):838-49. doi: 10.1111/j.1558-5646.2007.00081.x. PubMed PMID: WOS:000245819500010.
35. Lavine L, Gotoh H, Brent CS, Dworkin I, Emlen DJ. Exaggerated Trait Growth in Insects. *Annu Rev Entomol*. 2015;60:453-72. doi: 10.1146/annurev-ento-010814-021045. PubMed PMID: WOS:000348560200025.
36. Khila A, Abouheif E, Rowe L. Function, developmental genetics, and fitness consequences of a sexually antagonistic trait. *Science*. 2012;336(6081):585-9. Epub 2012/05/05. doi: 10.1126/science.1217258. PubMed PMID: 22556252.
37. Andrade MCB. Sexual selection for male sacrifice in the Australian redback spider. *Science*. 1996;271(5245):70-2. doi: DOI 10.1126/science.271.5245.70. PubMed PMID: WOS:A1996TP02200045.
38. Pomiankowski A, Moller AP. A Resolution of the Lek Paradox. *P R Soc B*. 1995;260(1357):21-9. doi: DOI 10.1098/rspb.1995.0054. PubMed PMID: WOS:A1995QX30200004.

39. Rowe L, Houle D. The lek paradox and the capture of genetic variance by condition dependent traits. *Proc R Soc B*. 1996;263(1375):1415-21. doi: DOI 10.1098/rspb.1996.0207. PubMed PMID: WOS:A1996VQ62200022.
40. McCullough EL, Emlen DJ. Evaluating the costs of a sexually selected weapon: big horns at a small price. *Anim Behav*. 2013;86(5):977-85. doi: 10.1016/j.anbehav.2013.08.017. PubMed PMID: WOS:000326272200017.
41. Emlen DJ, Warren IA, Johns A, Dworkin I, Lavine LC. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science*. 2012;337(6096):860-4. Epub 2012/07/28. doi: 10.1126/science.1224286. PubMed PMID: 22837386.
42. Emlen DJ, Nijhout HF. The development and evolution of exaggerated morphologies in insects. *Annual Review of Entomology*. 2000;45:661-708. PubMed PMID: ISI:000086173900026.
43. Gotoh H, Hust JA, Miura T, Niimi T, Emlen DJ, Lavine LC. The Fat/Hippo signaling pathway links within-disc morphogen patterning to whole-animal signals during phenotypically plastic growth in insects. *Developmental Dynamics*. 2015;244(9):1039-45. doi: 10.1002/Dvdy.24296. PubMed PMID: WOS:000360366700002.
44. Warren I, Smith H. Stalk-eyed flies (Diopsidae): modelling the evolution and development of an exaggerated sexual trait. *Bioessays*. 2007;29(3):300-7. doi: 10.1002/bies.20543. PubMed PMID: WOS:000244641300011.
45. Miller DC. Growth in *Uca*, 1. Ontogeny of Asymmetry in *Uca Pugilator* (Bosc) (Decapoda, Ocypodidae). *Crustaceana*. 1973. doi: doi:10.1163/156854073x00128
46. Andersson M. Female Choice Selects for Extreme Tail Length in a Widowbird. *Nature*. 1982;299(5886):818-20. doi: DOI 10.1038/299818a0. PubMed PMID: WOS:A1982PM85200043.
47. Andrade MCB. Risky mate search and male self-sacrifice in redback spiders. *Behav Ecol*. 2003;14(4):531-8. doi: DOI 10.1093/beheco/arg015. PubMed PMID: WOS:000184534600013.
48. Andrade MCB, Gu L, Stoltz JA. Novel male trait prolongs survival in suicidal mating. *Biol Letters*. 2005;1(3):276-9. doi: DOI 10.1098/rsbl.2005.0318. PubMed PMID: WOS:000232170300007.
49. Gage MJ. Evolution: sex and cannibalism in redback spiders. *Curr Biol*. 2005;15(16):R630-2. doi: 10.1016/j.cub.2005.08.003. PubMed PMID: 16111935.
50. Laman T, Scholes E. Birds of paradise: revealing the world's most extraordinary birds. 2012.
51. Girard MB, Endler JA. Peacock spiders. *Curr Biol*. 2014;24(13):R588-R90. doi: DOI 10.1016/j.cub.2014.05.026. PubMed PMID: WOS:000338799800003.
52. Maan ME, Sefc KM. Colour variation in cichlid fish: developmental mechanisms, selective pressures and evolutionary consequences. *Semin Cell Dev Biol*. 2013;24(6-7):516-28. doi: 10.1016/j.semcdb.2013.05.003. PubMed PMID: 23665150; PubMed Central PMCID: PMC3778878.
53. Hughes KA, Du L, Rodd FH, Reznick DN. Familiarity leads to female mate preference for novel males in the guppy, *Poecilia reticulata*. *Anim Behav*. 1999;58:907-16. doi: DOI 10.1006/anbe.1999.1225. PubMed PMID: WOS:000083347600023.
54. Hughes KA, Houde AE, Price AC, Rodd FH. Mating advantage for rare males in wild guppy populations. *Nature*. 2013;503(7474):108-10. doi: 10.1038/nature12717. PubMed PMID: 24172904.

55. Haskell DG. Do bright colors at nests incur a cost due to predation? *Evol Ecol.* 1996;10(3):285-8. doi: Doi 10.1007/Bf01237685. PubMed PMID: WOS:A1996UN97600006.
56. Godin JGJ, McDonough HE. Predator preference for brightly colored males in the guppy: a viability cost for a sexually selected trait. *Behav Ecol.* 2003;14(2):194-200. doi: DOI 10.1093/beheco/14.2.194. PubMed PMID: WOS:000181661100006.
57. Endler JA. Natural and Sexual Selection on Color Patterns in Poeciliid Fishes. *Environ Biol Fish.* 1983;9(2):173-90. doi: Doi 10.1007/Bf00690861. PubMed PMID: WOS:A1983RJ38100008.
58. Swanson BO, George MN, Anderson SP, Christy JH. Evolutionary variation in the mechanics of fiddler crab claws. *Bmc Evol Biol.* 2013;13:137. doi: 10.1186/1471-2148-13-137. PubMed PMID: 23855770; PubMed Central PMCID: PMC3716949.
59. Palmer AR. Symmetry breaking and the evolution of development. *Science.* 2004;306(5697):828-33. doi: 10.1126/science.1103707. PubMed PMID: 15514148.
60. Shingleton AW, Frankino WA. New perspectives on the evolution of exaggerated traits. *Bioessays.* 2013;35(2):100-7. doi: 10.1002/bies.201200139. PubMed PMID: 23255216.
61. Hölldobler B, Wilson E. *The Ants*: Harvard University Press; 1990.
62. Khila A, Abouheif E. Reproductive constraint is a developmental mechanism that maintains social harmony in advanced ant societies. *P Natl Acad Sci USA.* 2008;105(46):17884-9. Epub 2008/11/14. doi: 10.1073/pnas.0807351105. PubMed PMID: 19004767; PubMed Central PMCID: PMC2584687.
63. Khila A, Abouheif E. Evaluating the role of reproductive constraints in ant social evolution. *Philos Trans R Soc Lond B Biol Sci.* 2010;365(1540):617-30. Epub 2010/01/20. doi: 365/1540/617 10.1098/rstb.2009.0257. PubMed PMID: 20083637.
64. Matsuura K, Himuro C, Yokoi T, Yamamoto Y, Vargo EL, Keller L. Identification of a pheromone regulating caste differentiation in termites. *Proc Natl Acad Sci U S A.* 2010;107(29):12963-8. doi: 10.1073/pnas.1004675107. PubMed PMID: 20615972; PubMed Central PMCID: PMC2919916.
65. Refki PN, Khila A. Key patterning genes contribute to leg elongation in water striders. *Evodevo.* 2015;6. doi: 10.1186/S13227-015-0015-5. PubMed PMID: WOS:000354235300001.
66. Refki PN, Armisen D, Crumiere AJJ, Viala S, Khila A. Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. *Developmental Biology.* 2014;392(2):441-53. doi: 10.1016/j.ydbio.2014.05.021. PubMed PMID: WOS:000339534300026.
67. Armisen D, Refki PN, Crumiere AJJ, Viala S, Toubiana W, Khila A. Predator strike shapes antipredator phenotype through new genetic interactions in water striders. *Nature Communications.* 2015;6. doi: 10.1038/Ncomms9153. PubMed PMID: WOS:000363016300002.
68. 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.* 2014. doi: 10.1111/evo.12444. PubMed PMID: 24766229.
69. 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.* 2009;5(7):e1000583. Epub 2009/08/04. doi: 10.1371/journal.pgen.1000583. PubMed PMID: 19649305.

70. Korb J, Poulsen M, Hu HF, Li C, Boomsma JJ, Zhang GJ, et al. A genomic comparison of two termites with different social complexity. *Front Genet.* 2015;6. doi: 10.3389/fgene.2015.00009. PubMed PMID: WOS:000352764100001.
71. Andersen NM. A comparative study of locomotion on the water surface in semiaquatic bugs (Insecta, Hemiptera, Gerromorpha). *Vidensk Meddr dansk naturh Foren.* 1976;(139):337-96.
72. Andersen NM. The semiaquatic bugs (Hemiptera: Gerromorpha). Klampenborg, Denmark.: Scandinavian Science Press LTD.; 1982.
73. Hu DL, Chan B, Bush JW. The hydrodynamics of water strider locomotion. *Nature.* 2003;424(6949):663-6. Epub 2003/08/09. doi: 10.1038/nature01793. PubMed PMID: 12904790.
74. Jacob F. Evolution and tinkering. *Science.* 1977;196(4295):1161-6. PubMed PMID: 860134.
75. Gould SJ. *Ontogeny and Phylogeny*: Harvard University Press; 1977.
76. Carroll S. B. *Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom*: W. W. Norton; 2005.
77. Moczek AP, Rose DJ. Differential recruitment of limb patterning genes during development and diversification of beetle horns. *Proc Natl Acad Sci U S A.* 2009;106(22):8992-7. Epub 2009/05/20. doi: 0809668106 10.1073/pnas.0809668106. PubMed PMID: 19451631; PubMed Central PMCID: PMC2690047.
78. Warren IA, Fowler K, Smith H. Germline transformation of the stalk-eyed fly, *Teleopsis dalmanni*. *BMC Mol Biol.* 2010;11:86. doi: 10.1186/1471-2199-11-86. PubMed PMID: 21080934; PubMed Central PMCID: PMC2999598.
79. Emlen DJ. A developmental view of exaggerated growth and conditional expression in the weapons of sexual selection. In: Press PU, editor. *From Field Observations to Mechanisms: A Program in Evolutionary Biology* 2010.
80. Wilkinson GS, Breden F, Mank JE, Ritchie MG, Higginson AD, Radwan J, et al. The locus of sexual selection: moving sexual selection studies into the post-genomics era. *J Evolution Biol.* 2015;28(4):739-55. doi: 10.1111/jeb.12621. PubMed PMID: WOS:000353295200001.
81. Wilkinson GS, Taper M. Evolution of genetic variation for condition-dependent traits in stalk-eyed flies. *P R Soc B.* 1999;266(1429):1685-90. doi: DOI 10.1098/rspb.1999.0832. PubMed PMID: WOS:000082118200011.
82. Nijhout HF. The control of body size in insects. *Dev Biol.* 2003;261(1):1-9. PubMed PMID: 12941617.
83. Warren IA, Gotoh H, Dworkin IM, Emlen DJ, Lavine LC. A general mechanism for conditional expression of exaggerated sexually-selected traits. *Bioessays.* 2013;35(10):889-99. doi: 10.1002/bies.201300031. PubMed PMID: WOS:000327868600010.
84. Mirth CK, Frankino WA, Shingleton AW. Allometry and size control: what can studies of body size regulation teach us about the evolution of morphological scaling relationships? *Curr Opin Insect Sci.* 2016;13:93-8. doi: 10.1016/j.cois.2016.02.010. PubMed PMID: WOS:000376301800014.
85. Koyama T, Mendes C, Mirth CK. Mechanisms regulating nutrition-dependent developmental plasticity through organ-specific effects in insects. *Front Physiol.* 2013;4. doi: 10.3389/fphys.2013.00263. PubMed PMID: WOS:000346774000260.

86. Andersen DS, Colombani J, Leopold P. Coordination of organ growth: principles and outstanding questions from the world of insects. *Trends Cell Biol.* 2013;23(7):336-44. doi: 10.1016/j.tcb.2013.03.005. PubMed PMID: 23587490.
87. Gotoh H, Miyakawa H, Ishikawa A, Ishikawa Y, Sugime Y, Emlen DJ, et al. Developmental Link between Sex and Nutrition; doublesex Regulates Sex-Specific Mandible Growth via Juvenile Hormone Signaling in Stag Beetles. *Plos Genetics.* 2014;10(1). doi: 10.1371/journal.pgen.1004098. PubMed PMID: WOS:000336525000056.
88. Gotoh H, Cornette R, Koshikawa S, Okada Y, Lavine LC, Emlen DJ, et al. Juvenile Hormone Regulates Extreme Mandible Growth in Male Stag Beetles. *Plos One.* 2011;6(6). doi: 10.1371/journal.pone.0021139. PubMed PMID: WOS:000292033700038.
89. Fry CL. Juvenile hormone mediates a trade-off between primary and secondary sexual traits in stalk-eyed flies. *Evol Dev.* 2006;8(2):191-201. doi: DOI 10.1111/j.1525-142X.2006.00089.x. PubMed PMID: WOS:000235541700009.
90. Hietakangas V, Cohen SM. Regulation of Tissue Growth through Nutrient Sensing. *Annu Rev Genet.* 2009;43:389-410. doi: 10.1146/annurev-genet-102108-134815. PubMed PMID: WOS:000273580300017.
91. Rajakumar R, San Mauro D, Dijkstra MB, Huang MH, Wheeler DE, Hiou-Tim F, et al. Ancestral developmental potential facilitates parallel evolution in ants. *Science.* 2012;335(6064):79-82. Epub 2012/01/10. doi: 10.1126/science.1211451. PubMed PMID: 22223805.
92. Kijimoto T, Moczek AP, Andrews J. Diversification of doublesex function underlies morph-, sex-, and species-specific development of beetle horns. *P Natl Acad Sci USA.* 2012;109(50):20526-31. doi: 10.1073/pnas.1118589109. PubMed PMID: WOS:000312605600080.
93. Ito Y, Harigai A, Nakata M, Hosoya T, Araya K, Oba Y, et al. The role of doublesex in the evolution of exaggerated horns in the Japanese rhinoceros beetle. *Embo Rep.* 2013;14(6):561-7. doi: 10.1038/embor.2013.50. PubMed PMID: WOS:000319808600019.
94. Zauner H, Begemann G, Mari-Beffa M, Meyer A. Differential regulation of *msx* genes in the development of the gonopodium, an intromittent organ, and of the "sword," a sexually selected trait of swordtail fishes (*Xiphophorus*). *Evol Dev.* 2003;5(5):466-77. doi: DOI 10.1046/j.1525-142X.2003.03053.x. PubMed PMID: WOS:000185487200004.
95. Kijimoto T, Moczek AP. Hedgehog signaling enables nutrition-responsive inhibition of an alternative morph in a polyphenic beetle. *P Natl Acad Sci USA.* 2016;113(21):5982-7. doi: 10.1073/pnas.1601505113. PubMed PMID: WOS:000376779900065.
96. Nijhout HF, Riddiford LM, Mirth C, Shingleton AW, Suzuki Y, Callier V. The developmental control of size in insects. *Wires Dev Biol.* 2014;3(1):113-34. doi: 10.1002/wdev.124. PubMed PMID: WOS:000331053100006.
97. Nijhout HF, Grunert LW. The Cellular and Physiological Mechanism of Wing-Body Scaling in *Manduca sexta*. *Science.* 2010;330(6011):1693-5. doi: 10.1126/science.1197292. PubMed PMID: WOS:000285390500071.

98. Cohen SM. Controlling growth of the wing: Vestigial integrates signals from the compartment boundaries. *Bioessays*. 1996;18(11):855-8. doi: DOI 10.1002/bies.950181102. PubMed PMID: WOS:A1996VV00600001.
99. Baena-Lopez LA, Nojima H, Vincent JP. Integration of morphogen signalling within the growth regulatory network. *Current Opinion in Cell Biology*. 2012;24(2):166-72. doi: 10.1016/j.ceb.2011.12.010. PubMed PMID: WOS:000303144700003.
100. Baena-Lopez LA, Franch-Marro X, Vincent JP. Wingless Promotes Proliferative Growth in a Gradient-Independent Manner. *Sci Signal*. 2009;2(91). doi: 10.1126/scisignal.2000360. PubMed PMID: WOS:000275603600004.
101. Nijhout HF. The control of growth. *Development*. 2003;130(24):5863-7. PubMed PMID: ISI:000187400100001.
102. Grath S, Parsch J. Sex-Biased Gene Expression. *Annual Review of Genetics*, Vol 50. 2016;50:29-44. doi: 10.1146/annurev-genet-120215-035429. PubMed PMID: WOS:000389584800002.
103. Mank JE. Population genetics of sexual conflict in the genomic era. *Nature Reviews Genetics*. 2017;18(12):721-30. doi: 10.1038/nrg.2017.83. PubMed PMID: WOS:000415108900008.
104. Rowe L, Chenoweth SF, Agrawal AF. The Genomics of Sexual Conflict. *Am Nat*. 2018;192(2):274-86. doi: 10.1086/698198. PubMed PMID: WOS:000438972600016.
105. Chenoweth SF, McGuigan K. The Genetic Basis of Sexually Selected Variation. *Annual Review of Ecology, Evolution, and Systematics*, Vol 41. 2010;41:81-101. doi: DOI 10.1146/annurev-ecolsys-102209-144657. PubMed PMID: WOS:000284978200005.
106. Mank JE. The transcriptional architecture of phenotypic dimorphism. *Nat Ecol Evol*. 2017;1(1). doi: 10.1038/s41559-016-0006. PubMed PMID: WOS:000417168900004.
107. Perry JC, Harrison PW, Mank JE. The Ontogeny and Evolution of Sex-Biased Gene Expression in *Drosophila melanogaster*. *Mol Biol Evol*. 2014;31(5):1206-19. doi: 10.1093/molbev/msu072. PubMed PMID: WOS:000335914400013.
108. Mank JE, Nam K, Brunstrom B, Ellegren H. Ontogenetic Complexity of Sexual Dimorphism and Sex-Specific Selection. *Mol Biol Evol*. 2010;27(7):1570-8. doi: 10.1093/molbev/msq042. PubMed PMID: WOS:000279872000011.
109. Bloch NI, Corral-Lopez A, Buechel SD, Kotrschal A, Kolm N, Mank JE. Early neurogenomic response associated with variation in guppy female mate preference. *Nat Ecol Evol*. 2018;2(11):1772-81. doi: 10.1038/s41559-018-0682-4. PubMed PMID: WOS:000447964800020.
110. Zemp N, Tavares R, Muyle A, Charlesworth D, Marais GAB, Widmer A. Evolution of sex-biased gene expression in a dioecious plant. *Nat Plants*. 2016;2(11). doi: 10.1038/Nplants.2016.168. PubMed PMID: WOS:000395792200009.
111. Harrison PW, Wright AE, Zimmer F, Dean R, Montgomery SH, Pointer MA, et al. Sexual selection drives evolution and rapid turnover of male gene expression. *P Natl Acad Sci USA*. 2015;112(14):4393-8. doi: 10.1073/pnas.1501339112. PubMed PMID: WOS:000352287800059.
112. Vicoso B, Charlesworth B. Evolution on the X chromosome: unusual patterns and processes. *Nature Reviews Genetics*. 2006;7(8):645-53. doi: 10.1038/nrg1914. PubMed PMID: WOS:000239118400016.

113. Meisel RP, Malone JH, Clark AG. Disentangling the relationship between sex-biased gene expression and X-linkage. *Genome Res.* 2012;22(7):1255-65. doi: 10.1101/gr.132100.111. PubMed PMID: WOS:000306030000006.
114. Dean R, Mank JE. The role of sex chromosomes in sexual dimorphism: discordance between molecular and phenotypic data. *J Evolution Biol.* 2014;27(7):1443-53. doi: 10.1111/jeb.12345. PubMed PMID: WOS:000338032700013.
115. Magnusson K, Lycett GJ, Mendes AM, Lynd A, Papathanos PA, Crisanti A, et al. Demasculinization of the *Anopheles gambiae* X chromosome. *Bmc Evol Biol.* 2012;12. doi: 10.1186/1471-2148-12-69. PubMed PMID: WOS:000308077700001.
116. Sturgill D, Zhang Y, Parisi M, Oliver B. Demasculinization of X chromosomes in the *Drosophila* genus. *Nature.* 2007;450(7167):238-U3. doi: 10.1038/nature06330. PubMed PMID: WOS:000250746200045.
117. Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, et al. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science.* 2003;299(5607):697-700. doi: 10.1126/science.1079190. PubMed PMID: WOS:000180687700044.
118. Bachtrog D, Toda NRT, Lockton S. Dosage Compensation and Demasculinization of X Chromosomes in *Drosophila*. *Curr Biol.* 2010;20(16):1476-81. doi: 10.1016/j.cub.2010.06.076. PubMed PMID: WOS:000281328800028.
119. Ledon-Rettig CC, Zattara EE, Moczek AP. Asymmetric interactions between doublesex and tissue- and sex-specific target genes mediate sexual dimorphism in beetles. *Nature Communications.* 2017;8. doi: 10.1038/ncomms14593. PubMed PMID: WOS:000394797000001.
120. Pespeni MH, Ladner JT, Moczek AP. Signals of selection in conditionally expressed genes in the diversification of three horned beetle species. *J Evolution Biol.* 2017;30(9):1644-57. doi: 10.1111/jeb.13079. PubMed PMID: WOS:000409239400003.
121. Zinna R, Emlen D, Lavine LC, Johns A, Gotoh H, Niimi T, et al. Sexual dimorphism and heightened conditional expression in a sexually selected weapon in the Asian rhinoceros beetle. *Molecular Ecology.* 2018;27(24):5049-72. doi: 10.1111/mec.14907. PubMed PMID: WOS:000454600500008.
122. Warren IA, Vera JC, Johns A, Zinna R, Marden JH, Emlen DJ, et al. Insights into the Development and Evolution of Exaggerated Traits Using De Novo Transcriptomes of Two Species of Horned Scarab Beetles. *Plos One.* 2014;9(2). doi: 10.1371/journal.pone.0088364. PubMed PMID: WOS:000331714700015.
123. Ohde T, Morita S, Shigenobu S, Morita J, Mizutani T, Gotoh H, et al. Rhinoceros beetle horn development reveals deep parallels with dung beetles. *Plos Genetics.* 2018;14(10). doi: 10.1371/journal.pgen.1007651. PubMed PMID: WOS:000449328500010.
124. Kijimoto T, Snell-Rood EC, Pespeni MH, Rocha G, Kafadar K, Moczek AP. The nutritionally responsive transcriptome of the polyphenic beetle *Onthophagus taurus* and the importance of sexual dimorphism and body region. *P R Soc B.* 2014;281(1797). doi: 10.1098/rspb.2014.2084. PubMed PMID: WOS:000344594400020.

125. Ledon-Rettig CC, Moczek AP. The transcriptomic basis of tissue- and nutrition-dependent sexual dimorphism in the beetle *Onthophagus taurus*. *Ecol Evol*. 2016;6(6):1601-13. doi: 10.1002/ece3.1933. PubMed PMID: WOS:000372488300003.
126. Kirkpatrick M, Ryan MJ. The Evolution of Mating Preferences and the Paradox of the Lek. *Nature*. 1991;350(613):33-8. doi: DOI 10.1038/350033a0. PubMed PMID: WOS:A1991FA69300054.
127. Houle D. Comparing Evolvability and Variability of Quantitative Traits. *Genetics*. 1992;130(1):195-204. PubMed PMID: ISI:A1992GX54300018.
128. Radwan J. Maintenance of genetic variation in sexual ornaments: a review of the mechanisms. *Genetica*. 2008;134(1):113-27. doi: 10.1007/s10709-007-9203-0. PubMed PMID: WOS:000258371900012.
129. Pomiankowski A. Sexual Selection - the Handicap Principle Does Work Sometimes. *Proc R Soc Ser B-Bio*. 1987;231(1262):123-45. doi: DOI 10.1098/rspb.1987.0038. PubMed PMID: WOS:A1987J032400008.
130. Grafen A. Sexual Selection Unhandicapped by the Fisher Process. *J Theor Biol*. 1990;144(4):473-516. doi: Doi 10.1016/S0022-5193(05)80087-6. PubMed PMID: WOS:A1990DL42100003.
131. Damgaard J. Phylogeny of the semiaquatic bugs (Hemiptera-Heteroptera, Gerromorpha) *Insect Systematics & Evolution*. 2008;39(4):30. doi: <http://dx.doi.org/10.1163/187631208788784264>.
132. Finet C, Decaras A, Armisen D, Khila A. The achaete-scute complex contains a single gene that controls bristle development in the semi-aquatic bugs. *P R Soc B*. 2018;285(1892). doi: 10.1098/rspb.2018.2387. PubMed PMID: WOS:000452355700014.
133. Crumiere AJJ, Santos ME, Semon M, Armisen D, Moreira FFF, Khila A. Diversity in Morphology and Locomotory Behavior Is Associated with Niche Expansion in the Semi-aquatic Bugs. *Curr Biol*. 2016;26(24):3336-42. doi: 10.1016/j.cub.2016.09.061. PubMed PMID: WOS:000390666200027.
134. Santos ME, Berger CS, Refki PN, Khila A. Integrating evo-devo with ecology for a better understanding of phenotypic evolution. *Brief Funct Genomics*. 2015;14(6):384-95. doi: 10.1093/bfpg/elv003. PubMed PMID: WOS:000365706000002.
135. Arnqvist G. Multiple Mating in a Water Strider - Mutual Benefits or Intersexual Conflict. *Anim Behav*. 1989;38:749-56. doi: Doi 10.1016/S0003-3472(89)80107-1. PubMed PMID: WOS:A1989CA58200002.
136. Rowe L, Arnqvist G, Sih A, Krupa J. Sexual Conflict and the Evolutionary Ecology of Mating Patterns - Water Striders as a Model System. *Trends Ecol Evol*. 1994;9(8):289-93. doi: Doi 10.1016/0169-5347(94)90032-9. PubMed PMID: WOS:A1994NX31800006.
137. Crumiere AJJ, Armisen D, Vargas A, Kubarakos M, Moreira FFF, Khila A. Escalation and constraints of antagonistic armaments in water striders. *bioRxiv*. 2018. doi: <https://doi.org/10.1101/430322>.
138. Arnqvist G, Rowe L. Correlated evolution of male and female morphologies in water striders. *Evolution*. 2002;56(5):936-47. Epub 2002/07/03. PubMed PMID: 12093029.
139. Arnqvist G, Rowe L. Antagonistic coevolution between the sexes in a group of insects. *Nature*. 2002;415(6873):787-9. PubMed PMID: ISI:000173833900045.
140. Arnqvist G, Rowe L. Sexual Conflict and Arms Races between the Sexes - a Morphological Adaptation for Control of Mating in a Female Insect. *P R Soc B*. 1995;261(1360):123-7. doi: DOI 10.1098/rspb.1995.0126. PubMed PMID: WOS:A1995RN27900019.

141. Arnqvist G. The evolution of water strider mating systems: causes and consequences of sexual conflicts. In: Crespi B, editor. *The Evolution of Mating Systems in Insects and Arachnids*. Cambridge University Press ed1997.
142. Hayashi K. Alternative Mating Strategies in the Water Strider *Gerris-Elongatus* (Heteroptera, Gerridae). *Behav Ecol Sociobiol*. 1985;16(4):301-6. doi: 10.1007/Bf00295542. PubMed PMID: WOS:A1985AHR1900002.
143. Sih A, Lauer M, Krupa JJ. Path analysis and the relative importance of male-female conflict, female choice and male-male competition in water striders. *Anim Behav*. 2002;63:1079-89. doi: 10.1006/anbe.2002.2002. PubMed PMID: WOS:000177715700007.
144. Toubiana W, Khila A. The benefits of expanding studies of trait exaggeration to hemimetabolous insects and beyond morphology. *Curr Opin Genet Dev*. 2016;39:14-20. doi: 10.1016/j.gde.2016.05.015. PubMed PMID: WOS:000388922900004.
145. Emlen DJ. Alternative reproductive tactics and male-dimorphism in the horned beetle *Onthophagus acuminatus* (Coleoptera:Scarabaeidae). *Behav Ecol Sociobiol*. 1997;41(5):335-41. doi: DOI 10.1007/s002650050393. PubMed PMID: WOS:A1997YK46400004.
146. Moczek AP, Emlen DJ. Male horn dimorphism in the scarab beetle, *Onthophagus taurus*: do alternative reproductive tactics favour alternative phenotypes? *Anim Behav*. 2000;59(2):459-66. doi: 10.1006/anbe.1999.1342. PubMed PMID: 10675268.
147. Widemo F. Alternative reproductive strategies in the ruff, *Philomachus pugnax*: a mixed ESS? *Anim Behav*. 1998;56(2):329-36. doi: 10.1006/anbe.1998.0792. PubMed PMID: 9787023.
148. Zahavi A. Mate Selection - Selection for a Handicap. *J Theor Biol*. 1975;53(1):205-14. doi: DOI 10.1016/0022-5193(75)90111-3. PubMed PMID: WOS:A1975AR81900012.
149. Emlen DJ. The Evolution of Animal Weapons. *Annu Rev Ecol Evol S*. 2008;39:387-413. doi: 10.1146/annurev.ecolsys.39.110707.173502. PubMed PMID: WOS:000261725500019.
150. Miller CW. Sexual selection: Male-male competition. In: Losos JB, editor. *The Princeton Guide to Evolution*. Princeton: Princeton University Press ed2013.
151. Miller CW, Svensson EI. Sexual selection in complex environments. *Annu Rev Entomol*. 2014;59:427-45. doi: 10.1146/annurev-ento-011613-162044. PubMed PMID: 24160419.
152. Cornwallis CK, Uller T. Towards an evolutionary ecology of sexual traits. *Trends Ecol Evol*. 2010;25(3):145-52. doi: 10.1016/j.tree.2009.09.008. PubMed PMID: 19853321.
153. Robinson MR, Pilkington JG, Clutton-Brock TH, Pemberton JM, Kruuk LEB. Environmental heterogeneity generates fluctuating selection on a secondary sexual trait. *Curr Biol*. 2008;18(10):751-7. doi: 10.1016/j.cub.2008.04.059. PubMed PMID: WOS:000256047600028.
154. Gosden TP, Svensson EI. Spatial and temporal dynamics in a sexual selection mosaic. *Evolution*. 2008;62(4):845-56. doi: 10.1111/j.1558-5646.2008.00323.x. PubMed PMID: WOS:000254640900010.
155. Arnqvist G. Spatial Variation in Selective Regimes - Sexual Selection in the Water Strider *Gerris-Odontogaster*. *Evolution*. 1992;46(4):914-29. doi: DOI 10.1111/j.1558-5646.1992.tb00609.x. PubMed PMID: WOS:A1992JK16900005.
156. Martinossi-Allibert I, Rueffler C, Arnqvist G, Berger D. The efficacy of sexual selection under environmental change. *bioRxiv*. 2018. doi: doi: <http://dx.doi.org/10.1101/283457>.

157. Kupper C, Stocks M, Risse JE, dos Remedios N, Farrell LL, Mcrae SB, et al. A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics*. 2016;48(1):79-+. doi: 10.1038/ng.3443. PubMed PMID: WOS:000367255300017.
158. Lamichhaney S, Fan GY, Widemo F, Gunnarsson U, Thalmann DS, Hoepfner MP, et al. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics*. 2016;48(1):84-+. doi: 10.1038/ng.3430. PubMed PMID: WOS:000367255300018.
159. Emlen DJ. Environmental-Control of Horn Length Dimorphism in the Beetle *Onthophagus-Acuminatus* (Coleoptera, Scarabaeidae). *P R Soc B*. 1994;256(1346):131-6. doi: DOI 10.1098/rspb.1994.0060. PubMed PMID: WOS:A1994NN67200005.
160. Pointer MA, Harrison PW, Wright AE, Mank JE. Masculinization of Gene Expression Is Associated with Exaggeration of Male Sexual Dimorphism. *Plos Genetics*. 2013;9(8). doi: 10.1371/journal.pgen.1003697. PubMed PMID: WOS:000323830300048.
161. Warton DI, Duursma RA, Falster DS, Taskinen S. smatr 3-an R package for estimation and inference about allometric lines. *Methods Ecol Evol*. 2012;3(2):257-9. doi: 10.1111/j.2041-210X.2011.00153.x. PubMed PMID: WOS:000302538500005.
162. Karaca M, Bilgen M, Onus AN, Ince AG, Elmasulu SY. Exact Tandem Repeats Analyzer (E-TRA): A new program for DNA sequence mining. *J Genet*. 2005;84(1):49-54. doi: Doi 10.1007/Bf02715889. PubMed PMID: WOS:000228875600008.
163. Baker RH, Wilkinson GS. Phylogenetic analysis of sexual dimorphism and eye-span allometry in stalk-eyed flies (Diopsidae). *Evolution*. 2001;55(7):1373-85. PubMed PMID: WOS:000170699300010.
164. McCullough EL, Ledger KJ, O'Brien DM, Emlen DJ. Variation in the allometry of exaggerated rhinoceros beetle horns. *Anim Behav*. 2015;109:133-40. doi: 10.1016/j.anbehav.2015.08.013. PubMed PMID: WOS:000363268300017.
165. Taylor JS, McPherson EJ. Morphological Variation and Polyvoltinism of *Microvelia pulchella* (Heteroptera: Veliidae) in Southern Illinois, USA. *Acta Societatis Zoologicae Bohemicae*. 1999.
166. Kokko H, Rankin DJ. Lonely hearts or sex in the city? Density-dependent effects in mating systems. *Philos Trans R Soc Lond B Biol Sci*. 2006;361(1466):319-34. doi: 10.1098/rstb.2005.1784. PubMed PMID: 16612890; PubMed Central PMCID: PMC1569612.
167. Simmons LW. Sperm competition. In: Shuker DM, Simmons LW, editors. *The evolution of insect mating systems*. Oxford University Press ed2014.
168. Devost E, Turgeon J. The combined effects of pre- and post-copulatory processes are masking sexual conflict over mating rate in *Gerris buenoi*. *J Evolution Biol*. 2016;29(1):167-77. doi: 10.1111/jeb.12772. PubMed PMID: WOS:000368074900013.
169. Simmons LW. *Sperm Competition and Its Evolutionary Consequences in the Insects*. Princeton University Press, Princeton, NJ ed2001.
170. Hunt BG, Ometto L, Wurm Y, Shoemaker D, Yi SV, Keller L, et al. Relaxed selection is a precursor to the evolution of phenotypic plasticity. *P Natl Acad Sci USA*. 2011;108(38):15936-41. doi: 10.1073/pnas.1104825108. PubMed PMID: WOS:000295030000056.

171. Lahti DC, Johnson NA, Ajie BC, Otto SP, Hendry AP, Blumstein DT, et al. Relaxed selection in the wild. *Trends Ecol Evol.* 2009;24(9):487-96. doi: 10.1016/j.tree.2009.03.010. PubMed PMID: WOS:000270016800009.
172. Bonduriansky R. Condition dependence of developmental stability in the sexually dimorphic fly *Telostylinus angusticollis* (Diptera: Neriidae). *J Evol Biol.* 2009;22(4):861-72. Epub 2009/02/18. doi: JEB168610.1111/j.1420-9101.2009.01686.x. PubMed PMID: 19220651.
173. Chapman T, Pomiankowski A, Fowler K. Stalk-eyed flies. *Curr Biol.* 2005;15(14):R533-R5. doi: DOI 10.1016/j.cub.2005.07.015. PubMed PMID: WOS:000230903900006.
174. Wilcox RS. Ripple Communication in Aquatic and Semiaquatic Insects. *Ecoscience.* 1995;2(2):109-15. doi: Doi 10.1080/11956860.1995.11682274. PubMed PMID: WOS:A1995TF50800001.
175. Punzalan D, Hosken DJ. Sexual Dimorphism: Why the Sexes Are (and Are Not) Different. *Curr Biol.* 2010;20(22):R972-R3. doi: 10.1016/j.cub.2010.09.067. PubMed PMID: WOS:000284923700011.
176. Lande R. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters. *Evolution.* 1980;34(2):292-305. doi: Doi 10.2307/2407393. PubMed PMID: WOS:A1980JQ65100010.
177. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 2007;8(9):689-98. doi: 10.1038/nrg2167. PubMed PMID: 17680007.
178. Lavine LC, Emlen DJ, Duke A, Warren IA, Gotoh H. Mechanisms of Honest Signaling in Condition Dependent Growth in Exaggerated Traits. *Integr Comp Biol.* 2014;54:E119-E. PubMed PMID: WOS:000333049501008.
179. Shingleton AW, Frankino WA, Flatt T, Nijhout HF, Emlen DJ. Size and shape: the developmental regulation of static allometry in insects. *Bioessays.* 2007;29(6):536-48. doi: 10.1002/bies.20584. PubMed PMID: 17508394.
180. Emlen DJ, Szafran Q, Corley LS, Dworkin I. Insulin signaling and limb-patterning: candidate pathways for the origin and evolutionary diversification of beetle 'horns'. *Heredity.* 2006;97(3):179-91. Epub 2006/07/20. doi: 10.1038/sj.hdy.6800868. PubMed PMID: 16850039.
181. Lavine LC, Hahn LL, Warren IA, Garczynski SF, Dworkin I, Emlen DJ. Cloning and characterization of an mRNA encoding an insulin receptor from the horned scarab beetle *Onthophagus nigriventris* (Coleoptera: Scarabaeidae). *Arch Insect Biochem Physiol.* 2013;82(1):43-57. Epub 2012/11/09. doi: 10.1002/arch.21072. PubMed PMID: 23136112.
182. Shelby JA, Madewell R, Moczek AP. Juvenile hormone mediates sexual dimorphism in horned beetles. *J Exp Zool B Mol Dev Evol.* 2007;308(4):417-27. Epub 2007/03/23. doi: 10.1002/jez.b.21165. PubMed PMID: 17377953.
183. Kijimoto T, Costello J, Tang Z, Moczek AP, Andrews J. EST and microarray analysis of horn development in *Onthophagus* beetles. *Bmc Genomics.* 2009;10:504. Epub 2009/11/03. doi: 1471-2164-10-504 10.1186/1471-2164-10-504. PubMed PMID: 19878565; PubMed Central PMCID: PMC2777201.
184. Snell-Rood EC, Cash A, Han MV, Kijimoto T, Andrews J, Moczek AP. Developmental decoupling of alternative phenotypes: insights from the transcriptomes of horn-polyphenic beetles. *Evolution; international journal of organic evolution.* 2011;65(1):231-45. Epub 2010/08/25. doi: 10.1111/j.1558-5646.2010.01106.x. PubMed PMID: 20731717; PubMed Central PMCID: PMC3010270.

185. Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, et al. Genomics and the origin of species. *Nat Rev Genet.* 2014;15(3):176-92. doi: 10.1038/nrg3644. PubMed PMID: 24535286.
186. Savolainen O, Lascoux M, Merila J. Ecological genomics of local adaptation. *Nat Rev Genet.* 2013;14(11):807-20. doi: 10.1038/nrg3522. PubMed PMID: 24136507.
187. Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet.* 2004;36(6):642-6. doi: 10.1038/ng1368. PubMed PMID: 15156144.
188. Boutanaev AM, Kalmykova AI, Shevelyov YY, Nurminsky DI. Large clusters of co-expressed genes in the *Drosophila* genome. *Nature.* 2002;420(6916):666-9. doi: 10.1038/nature01216. PubMed PMID: WOS:000179751800045.
189. Catalan A, Macias-Munoz A, Briscoe AD. Evolution of sex-biased gene expression and dosage compensation in the eye and brain of *Heliconius* butterflies. *Mol Biol Evol.* 2018. doi: 10.1093/molbev/msy111. PubMed PMID: 29931127.
190. Kim KW, Bennison C, Hemmings N, Brookes L, Hurley LL, Griffith SC, et al. A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nat Ecol Evol.* 2017;1(8):1168-76. doi: 10.1038/s41559-017-0235-2. PubMed PMID: WOS:000417188600024.
191. Knief U, Forstmeier W, Pei YF, Ihle M, Wang DP, Martin K, et al. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nat Ecol Evol.* 2017;1(8):1177-84. doi: 10.1038/s41559-017-0236-1. PubMed PMID: WOS:000417188600025.
192. Toubiana W, Khila A. Fluctuating selection strength and intense male competition underlie variation and exaggeration of a water strider's male weapon. *Proc Biol Sci.* 2019;286(1901):20182400. doi: 10.1098/rspb.2018.2400. PubMed PMID: 30991924.
193. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *Plos One.* 2014;9(11):e112963. doi: 10.1371/journal.pone.0112963. PubMed PMID: 25409509; PubMed Central PMCID: PMC4237348.
194. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, et al. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 2008;18(1):188-96. doi: 10.1101/gr.6743907. PubMed PMID: WOS:000251965300019.
195. Hoff KJ, Lange S, Lomsadze A, Borodovsky M, Stanke M. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics.* 2016;32(5):767-9. doi: 10.1093/bioinformatics/btv661. PubMed PMID: WOS:000371693900017.
196. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature protocols.* 2016;11(9):1650-67. doi: 10.1038/nprot.2016.095. PubMed PMID: 27560171; PubMed Central PMCID: PMC45032908.
197. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357-60. doi: 10.1038/nmeth.3317. PubMed PMID: 25751142; PubMed Central PMCID: PMC4655817.
198. Baty F, Facompre M, Wiegand J, Schwager J, Brutsche MH. Analysis with respect to instrumental variables for the exploration of microarray data structures. *Bmc Bioinformatics.* 2006;7:422. doi: 10.1186/1471-2105-7-422. PubMed PMID: 17010189; PubMed Central PMCID: PMC1594581.

199. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PubMed Central PMCID: PMC4302049.
200. Suzuki R, Shimodaira H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics.* 2006;22(12):1540-2. doi: 10.1093/bioinformatics/btl117. PubMed PMID: 16595560.
201. Castanhole MMU, Pereira LLV, Souza HV, Itoyama MM. Spermatogenesis and karyotypes of three species of water striders (Gerridae, Heteroptera). *Genet Mol Res.* 2010;9(3):1343-56. doi: 10.4238/vol9-3gmr841. PubMed PMID: WOS:000280397000008.
202. Fairbairn DJ, Kiseliova O, Muir S. Variation in chromosome numbers and the sex determination system in the Gerromorpha with special reference to the family Gerridae (Hemiptera). *Aquat Insect.* 2016;37(2):127-44. doi: 10.1080/01650424.2016.1167222. PubMed PMID: WOS:000379733700004.
203. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol.* 2013;31(12):1119-25. doi: 10.1038/nbt.2727. PubMed PMID: 24185095; PubMed Central PMCID: PMC4117202.
204. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature.* 2008;456(7218):53-9. doi: 10.1038/nature07517. PubMed PMID: WOS:000260674000039.
205. Rhoads A, Au KF. PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinformatics.* 2015;13(5):278-89. doi: 10.1016/j.gpb.2015.08.002. PubMed PMID: 26542840; PubMed Central PMCID: PMC4678779.
206. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31(19):3210-2. doi: 10.1093/bioinformatics/btv351. PubMed PMID: 26059717.
207. Yin CL, Shen GY, Guo DH, Wang SP, Ma XZ, Xiao HM, et al. InsectBase: a resource for insect genomes and transcriptomes. *Nucleic Acids Res.* 2016;44(D1):D801-D7. doi: 10.1093/nar/gkv1204. PubMed PMID: WOS:000371261700113.
208. Zhou X, Liao WJ, Liao JM, Liao P, Lu H. Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol.* 2015;7(2):92-104. doi: 10.1093/jmcb/mjv014. PubMed PMID: WOS:000355232100002.
209. Abouheif E. Developmental genetics and homology: a hierarchical approach. *Trends Ecol Evol.* 1997;12(10):405-8. PubMed PMID: ISI:A1997XY60500022.
210. Akam M. Arthropods: Developmental diversity within a (super) phylum. *P Natl Acad Sci USA.* 2000;97(9):4438-41. doi: DOI 10.1073/pnas.97.9.4438. PubMed PMID: WOS:000086703000010.
211. Schram FR, Koenemann S. Developmental genetics and arthropod evolution: part 1, on legs. *Evol Dev.* 2001;3(5):343-54. PubMed PMID: 11710766.
212. Crumiere AJJ, Khila A. Hox genes mediate the escalation of sexually antagonistic traits in water striders. *Biol Lett.* 2019;15(2):20180720. doi: 10.1098/rsbl.2018.0720. PubMed PMID: 30958129; PubMed Central PMCID: PMC6405465.
213. Pattatucci AM, Otteson DC, Kaufman TC. A functional and structural analysis of the Sex combs reduced locus of *Drosophila melanogaster*. *Genetics.* 1991;129(2):423-41. PubMed PMID: 1743486; PubMed Central PMCID: PMC41204634.

214. Mank JE. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *Am Nat.* 2009;173(2):141-50. doi: 10.1086/595754. PubMed PMID: 20374139.
215. Vicoso B, Charlesworth B. Evolution on the X chromosome: unusual patterns and processes. *Nat Rev Genet.* 2006;7(8):645-53. doi: 10.1038/nrg1914. PubMed PMID: 16847464.
216. Dorus S, Busby SA, Gerike U, Shabanowitz J, Hunt DF, Karr TL. Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. *Nat Genet.* 2006;38(12):1440-5. doi: 10.1038/ng1915. PubMed PMID: 17099714.
217. Kijimoto T, Andrews J, Moczek AP. Programed cell death shapes the expression of horns within and between species of horned beetles. *Evol Dev.* 2010;12(5):449-58. Epub 2010/10/05. doi: 10.1111/j.1525-142X.2010.00431.x. PubMed PMID: 20883214.
218. Warren I. Transgenic and histological approaches to investigating the development of the stalk-eyed fly, *Teleopsis dalmanni*: University College London; 2009.
219. Armisen D, Rajakumar R, Friedrich M, Benoit JB, Robertson HM, Panfilio KA, et al. The genome of the water strider *Gerris buenoi* reveals expansions of gene repertoires associated with adaptations to life on the water. *Bmc Genomics.* 2018;19(1):832. doi: 10.1186/s12864-018-5163-2. PubMed PMID: 30463532; PubMed Central PMCID: PMC6249893.
220. Gnad F., J. P. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. *Bioinformatics.* 2006.
221. Dutko JA, Mullins MC. SnapShot: BMP signaling in development. *Cell.* 2011;145(4):636, e1-2. doi: 10.1016/j.cell.2011.05.001. PubMed PMID: 21565618.
222. Logan M, Pagan-Westphal SM, Smith DM, Paganessi L, Tabin CJ. The transcription factor *Pitx2* mediates situs-specific morphogenesis in response to left-right asymmetric signals. *Cell.* 1998;94(3):307-17. PubMed PMID: 9708733.
223. Ahuja A, Singh RS. Variation and evolution of male sex combs in *Drosophila*: nature of selection response and theories of genetic variation for sexual traits. *Genetics.* 2008;179(1):503-9. doi: 10.1534/genetics.107.086363. PubMed PMID: 18493067; PubMed Central PMCID: PMC62390627.
224. Maraqa MS, Griffin R, Sharma MD, Wilson AJ, Hunt J, Hosken DJ, et al. Constrained evolution of the sex comb in *Drosophila simulans*. *J Evol Biol.* 2017;30(2):388-400. doi: 10.1111/jeb.13015. PubMed PMID: 27859860; PubMed Central PMCID: PMC62324616.
225. Mank JE, Hultin-Rosenberg L, Webster MT, Ellegren H. The unique genomic properties of sex-biased genes: insights from avian microarray data. *Bmc Genomics.* 2008;9:148. doi: 10.1186/1471-2164-9-148. PubMed PMID: 18377635; PubMed Central PMCID: PMC62294128.
226. Ma WJ, Veltsos P, Touns MA, Rodrigues N, Sermier R, Jeffries DL, et al. Tissue Specificity and Dynamics of Sex-Biased Gene Expression in a Common Frog Population with Differentiated, Yet Homomorphic, Sex Chromosomes. *Genes (Basel).* 2018;9(6). doi: 10.3390/genes9060294. PubMed PMID: 29895802; PubMed Central PMCID: PMC6227210.
227. Badyaev AV. Growing apart: an ontogenetic perspective on the evolution of sexual size dimorphism. *Trends Ecol Evol.* 2002;17(8):369-78. Doi 10.1016/S0169-5347(02)02569-7. PubMed PMID: WOS:000176998300014.

228. Birkhead TR, Pizzari T. Postcopulatory sexual selection. *Nature Reviews Genetics*. 2002;3(4):262-73. doi: 10.1038/nrg774. PubMed PMID: WOS:000174739800013.
229. Simmons LW. *Sperm Competition and Its Evolutionary Consequences in the Insects*. Princeton University Press ed2001.
230. Lupold S, Manier MK, Puniamoorthy N, Schoff C, Starmer WT, Luepold SHB, et al. How sexual selection can drive the evolution of costly sperm ornamentation. *Nature*. 2016;533(7604):535-+. doi: 10.1038/nature18005. PubMed PMID: WOS:000376443100041.
231. Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, et al. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*. 2013;502(7469):93-+. doi: 10.1038/nature12489. PubMed PMID: WOS:000325106000037.
232. Wilkinson GS. Artificial Sexual Selection Alters Allometry in the Stalk-Eyed Fly *Cyrtodiopsis Dalmanni* (Diptera, Diopsidae). *Genetical research*. 1993;62(3):213-22. doi: Doi 10.1017/S001667230003192x. PubMed PMID: WOS:A1993MX71000006.
233. Emlen DJ. Artificial selection on horn length body size allometry in the horned beetle *Onthophagus acuminatus* (Coleoptera: Scarabaeidae). *Evolution*. 1996;50(3):1219-30. doi: Doi 10.2307/2410662. PubMed PMID: WOS:A1996UY55500024.
234. Bolstad GH, Cassara JA, Marquez E, Hansen TF, van der Linde K, Houle D, et al. Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *P Natl Acad Sci USA*. 2015;112(43):13284-9. doi: 10.1073/pnas.1505357112. PubMed PMID: WOS:000363458100054.
235. Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet*. 2011;13(1):59-69. doi: 10.1038/nrg3095. PubMed PMID: 22143240.
236. Dubeau F, Toubiana W, Wittkopp PJ. Fitness Effects of Cis-Regulatory Variants in the *Saccharomyces cerevisiae* TDH3 Promoter. *Mol Biol Evol*. 2017;34(11):2908-12. doi: 10.1093/molbev/msx224. PubMed PMID: 28961929; PubMed Central PMCID: PMC5850770.
237. Barrett RDH, Hoekstra HE. Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*. 2011;12(11):767-80. doi: 10.1038/nrg3015. PubMed PMID: WOS:000296538500012.
238. Peel AD, Chipman AD, Akam M. Arthropod segmentation: beyond the *Drosophila* paradigm. *Nat Rev Genet*. 2005;6(12):905-16. Epub 2005/12/13. doi: nrg1724 10.1038/nrg1724. PubMed PMID: 16341071.
239. Khila A, Abouheif E, Rowe L. Eco-Evo-Devo: Lessons from Semi-Aquatic Bugs. *Developmental Biology*. 2011;356(1):107-. doi: 10.1016/j.ydbio.2011.05.025. PubMed PMID: WOS:000292784400021.