

Current Biology

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

Highlights

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

Authors

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

Correspondence

abderrahman.khila@ens-lyon.fr

In Brief

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



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

Antonin J.J. Crumière,¹ M. Emilia Santos,¹ Marie Sémon,² David Armisen,¹ Felipe F.F. Moreira,³ and Abderrahman Khila^{1,4,*}

¹Institut 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

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

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

⁴Lead Contact

*Correspondence: abderrahman.khila@ens-lyon.fr
<http://dx.doi.org/10.1016/j.cub.2016.09.061>

SUMMARY

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

RESULTS AND DISCUSSION

Phylogeny and Ancestral Habitat Reconstruction

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

Correlation between Locomotion Speed and Niche Preference

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

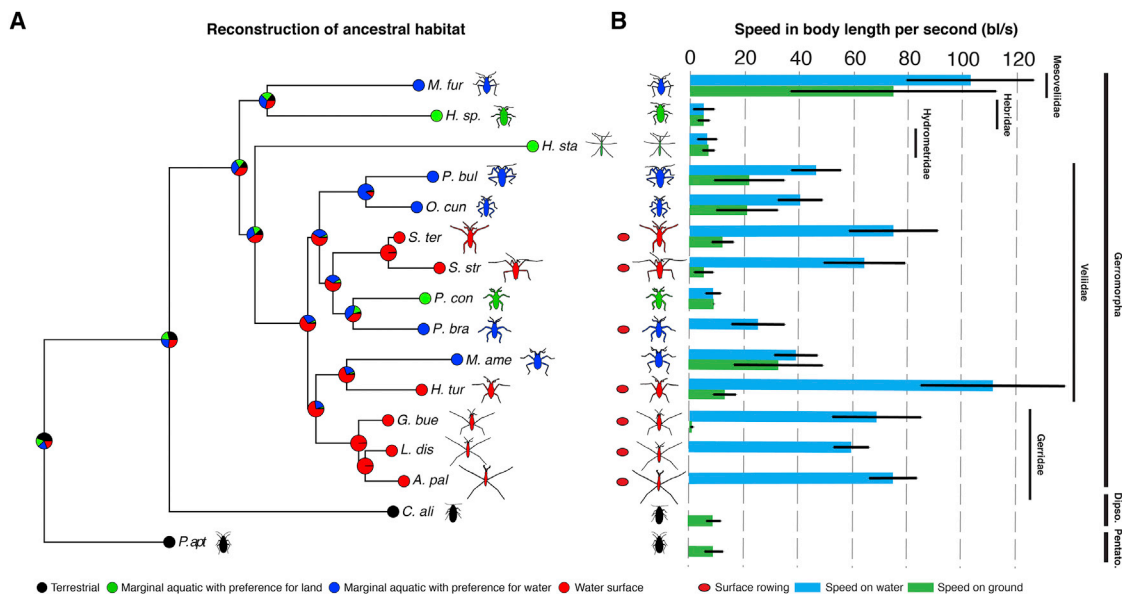


Figure 1. Ancestral Habitat Reconstruction and the Evolution of Maximum Speed in Relation to Habitat Preference

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

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

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

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

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

Correlation between Speed, Midleg Length, and Niche Preference

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

Increased leg length is known to contribute to generating faster movement through increasing the amplitude of strokes [20, 21]. When we analyzed how the legs act during locomotion on fluid compared to solid substrates, we failed to detect any difference in the amplitude of leg strokes between water and ground locomotion in species with short midlegs, preference

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

The Tripod Gait Employs High Frequency of Leg Strokes on Water

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

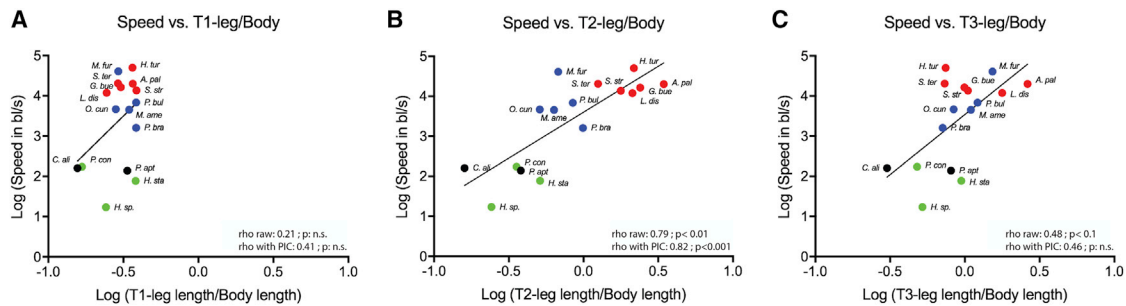


Figure 2. Correlation Tests between Leg Length, Speed, and Habitat Preference

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

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

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

See also [Figures S1](#), [S2](#), and [S4](#) and [Tables S1](#) and [S2](#).

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

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

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

the mode of locomotion, leg length, patterns of leg motion, and stroke frequency.

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

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

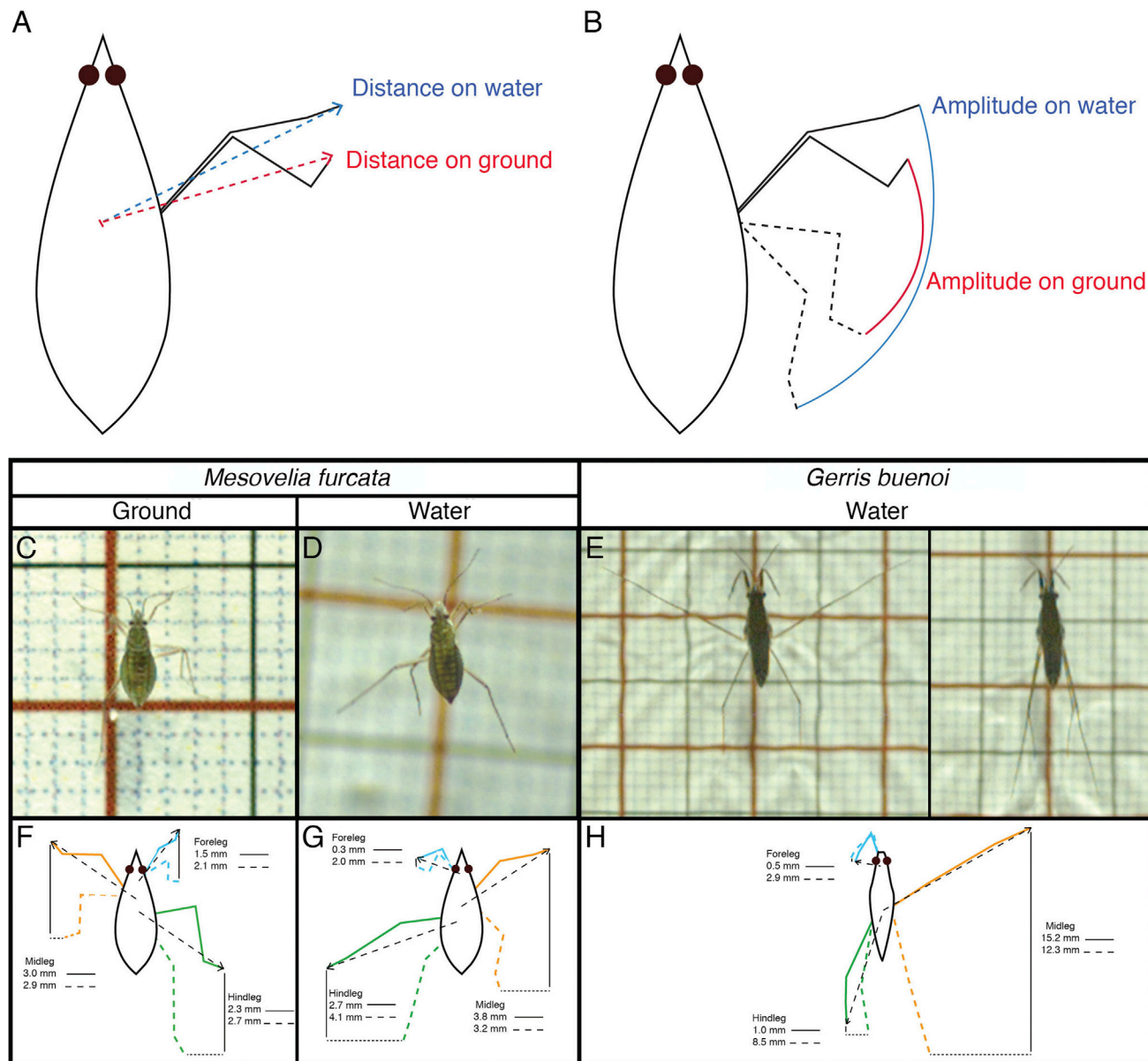


Figure 3. Analysis of Leg Movement Parameters between Ground and Water Surface Locomotion

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

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

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

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

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

Another important trade-off in walking systems, including arthropods, exists between body size and stroke frequency, such that stroke frequency decreases with increasing body size [21, 26].

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

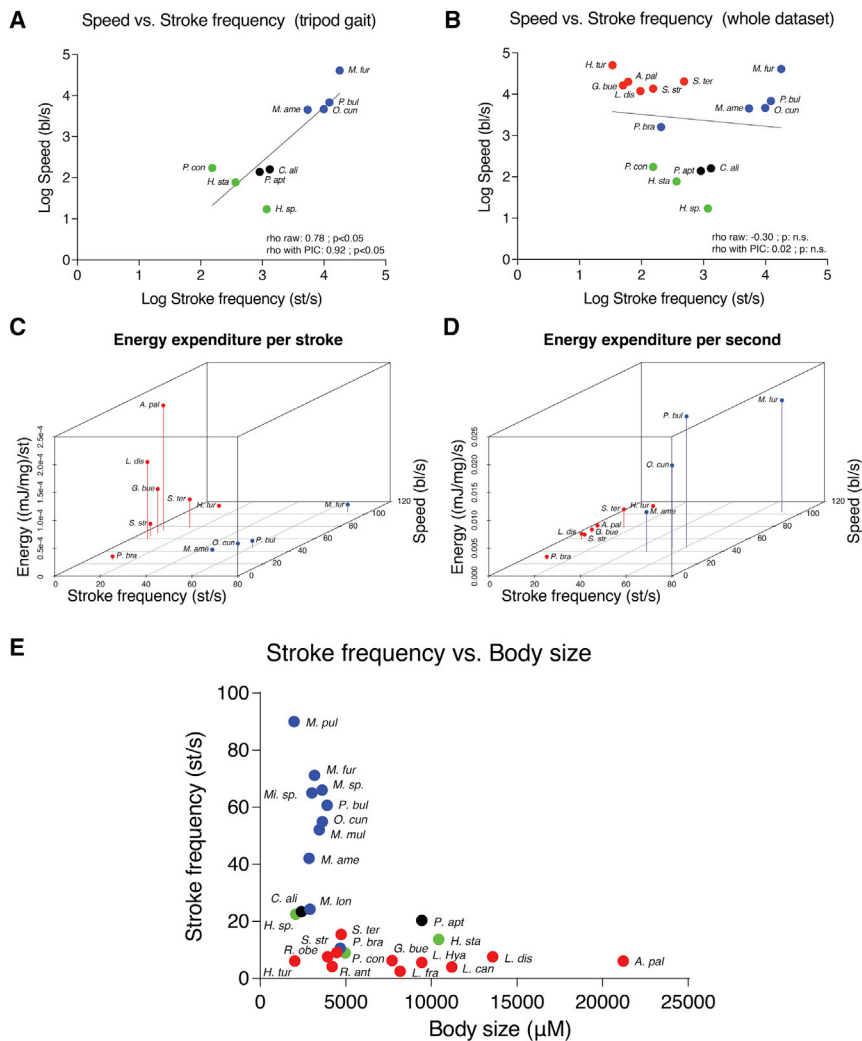


Figure 4. Correlation Tests between Stroke Frequency, Speed, and Habitat Preference with Associated Energy Expenditure and Relation with Body Size

(A) A significant positive correlation is observed between stroke frequency and speed in species employing the tripod gait using both Spearman correlation test on raw data (ρ raw: 0.78, adjusted p : 3.8×10^{-2}) and with phylogenetic correlation (ρ PIC: 0.92, adjusted p : 3.6×10^{-3}).

(B) There is no correlation between stroke frequency and speed when we add species employing the rowing gait to the dataset using both Spearman correlation test on raw data (ρ raw: -0.30 , adjusted p : 3.3×10^{-1} , n.s.) and with phylogenetic correlation (ρ PIC: 0.02, adjusted p : 0.97, n.s.). Sample sizes in term of number of videos are given in Table S1.

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

(D) However, when normalized to the number of stroke employed per second, species using the rowing gait spend less energy (Student's t test; p : 4.5×10^{-5}). Sample sizes and energy expenditure are given in the Supplemental Experimental Procedures and Table S1, respectively.

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

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

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

Conclusions

Understanding how the evolution of behavior and morphology can be associated with niche expansion and species diversification is a major challenge in evolutionary biology [1, 3, 17, 18]. We have shown that lineages that remained in the ancestral habitat, composed of solid substrates, produce low maximum speeds

and that lineages that specialize in open water surface niches generate significantly higher maximum speeds. This increase in locomotion speed is associated with the evolution of increased midleg length, changes in leg motion patterns, and increased frequency of leg strokes.

The subsequent evolution of rowing, as a novel mode of locomotion on the water, removed the requirement for high stroke frequency, reduced energy expenditure, and enabled derived lineages to specialize in open water zones. Finally, we have shown how the evolution of rowing, by removing the requirement for high stroke frequency, may have led to the evolution of larger Gerromorpha. By uncovering the link between the ecology of the semi-aquatic bugs and the variety of phenotypes they exhibit, this work draws a comprehensive picture of how a combination of behavioral and structural changes can impact the evolutionary trajectory of groups of animals and can be associated with niche expansion and lineage diversification.

EXPERIMENTAL PROCEDURES

Insect Sampling and Culture

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

Phylogenetic Reconstruction

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

with 100 bootstraps. More details can be found in the [Supplemental Experimental Procedures](#).

Habitat Classification

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

Video Acquisition, Quantification of Speed, and Stroke Frequency

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

Reconstruction of Ancestral Character State

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

Inference of Energy Consumption

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

Statistical Analyses

Details and R script used for statistical analyses are provided in the [Supplemental Experimental Procedures](#).

ACCESSION NUMBERS

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

SUPPLEMENTAL INFORMATION

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

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

We thank W. Salzburger, E. Abouheif, F. Bonneton, S. Viala, A. Decaras, and W. Toubiana for comments and discussion; H. Labrique and J.C. Streito for

help with the Dipsocoromorpha; and M. Burrows for advice on inferring energy expenditure. This work was funded by ERC-CoG #616346, ATIP-Avenir, and CNPq-PVE #400751/2014-3 to A.K. and a postdoctoral fellowship by the Swiss National Science Foundation to M.E.S. Specimens from Brazil were collected under SISBIO permit #43105-1.

Received: July 13, 2016

Revised: September 23, 2016

Accepted: September 28, 2016

Published: December 8, 2016

REFERENCES

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

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

Current Biology, Volume 26

Supplemental Information

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

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

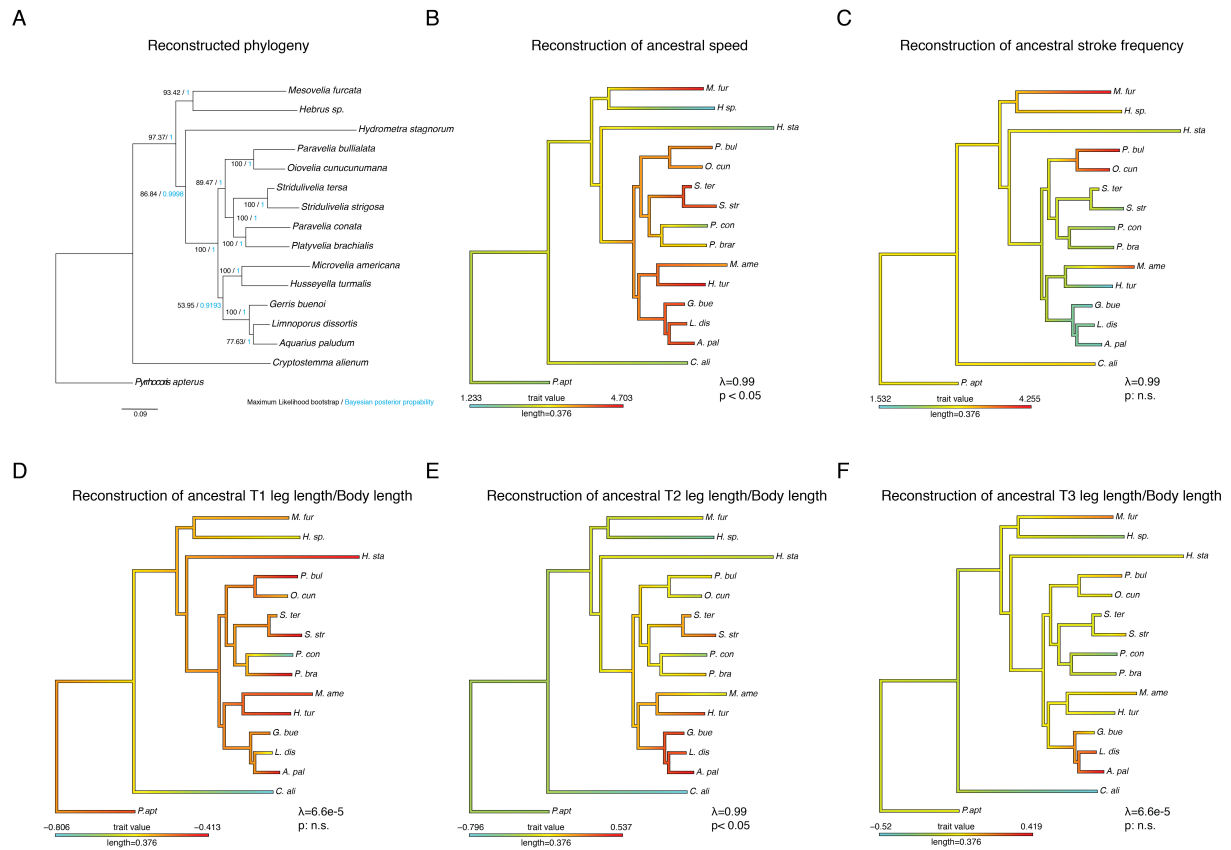


Figure S1: Phylogeny and ancestral character state reconstruction. Related to Figure 1, Figure 2 and Figure S4, Table S1, Table S2.

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

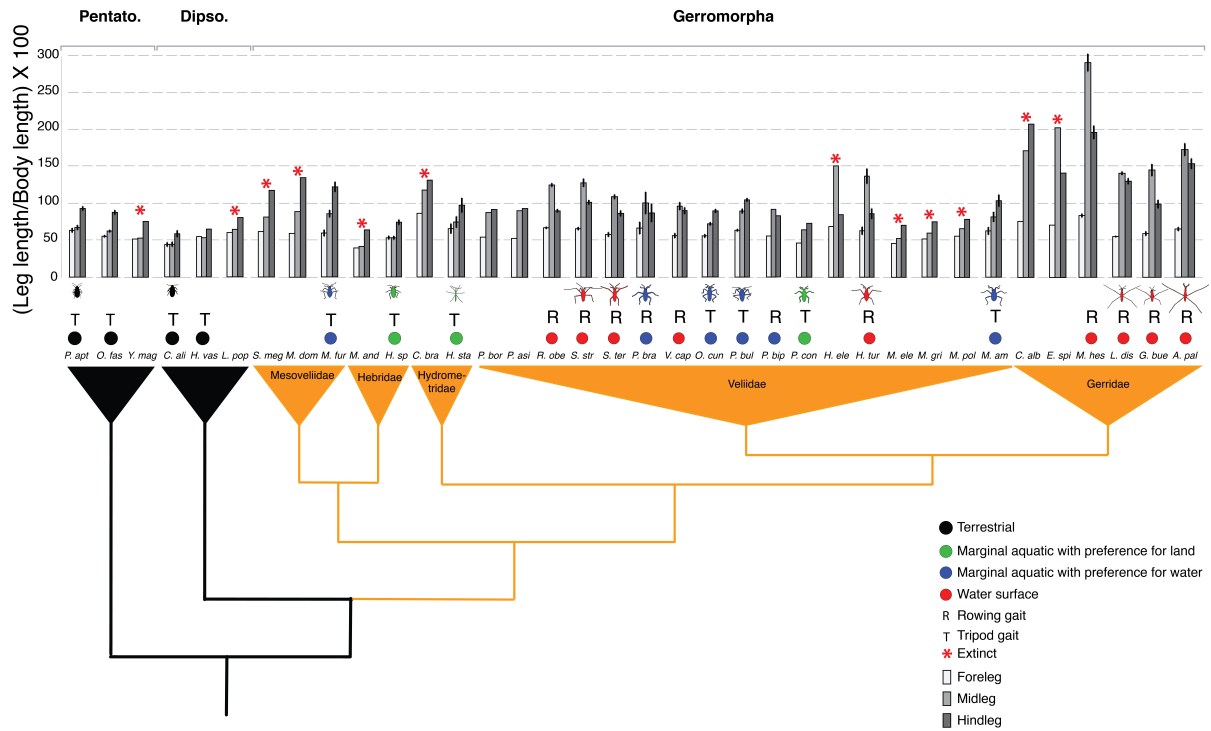


Figure S2: Ratio of leg length to body length across heteroptera including fossil record. Related to Figure 2, Table S2.

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

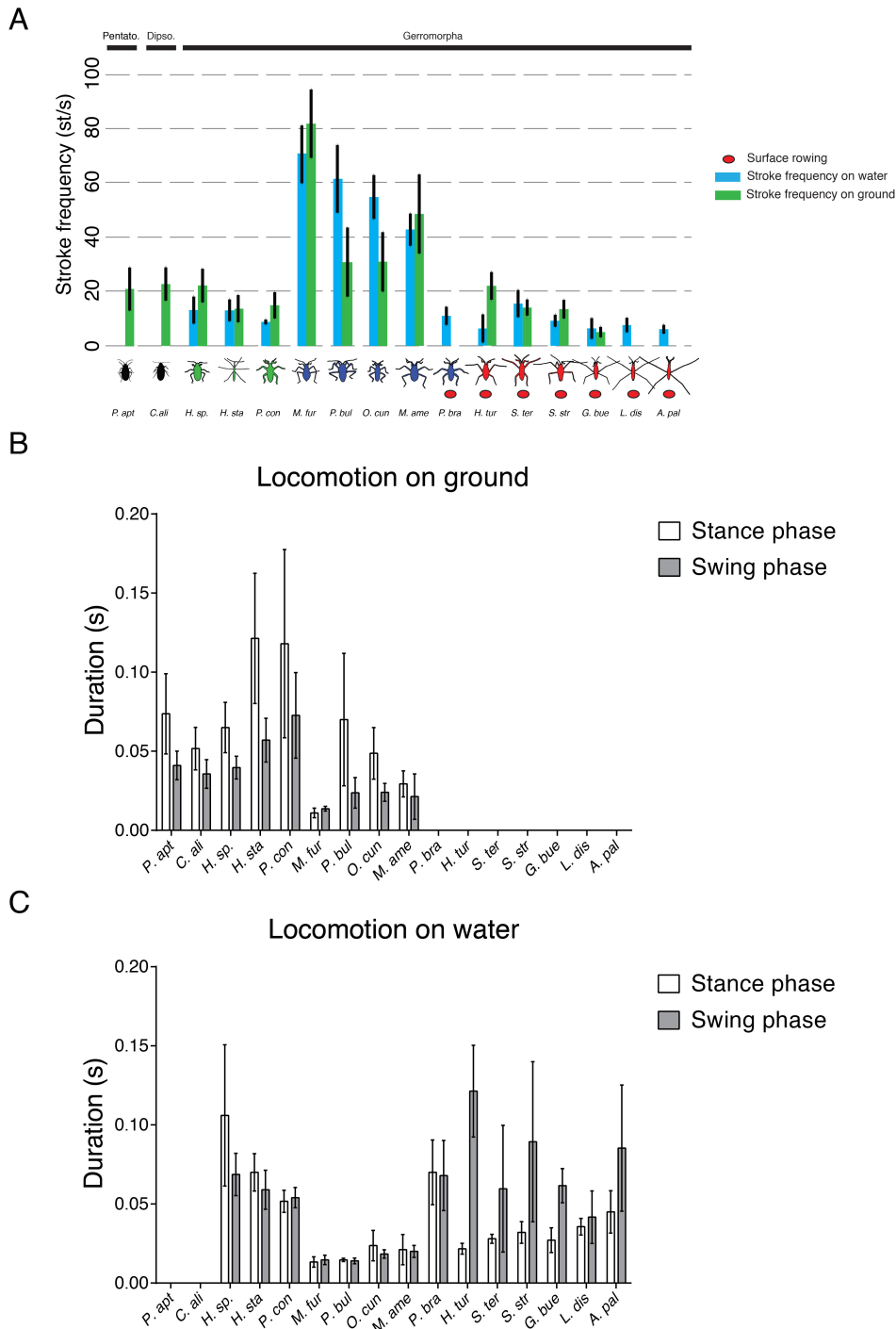
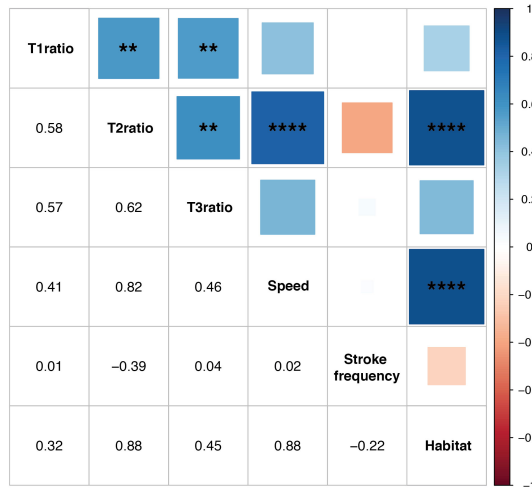


Figure S3: Stroke frequency and leg motion pattern across the Gerromorpha and terrestrial outgroups. Related to Figure 3, Figure 4 and Figure S4, Table S1, Table S3.

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

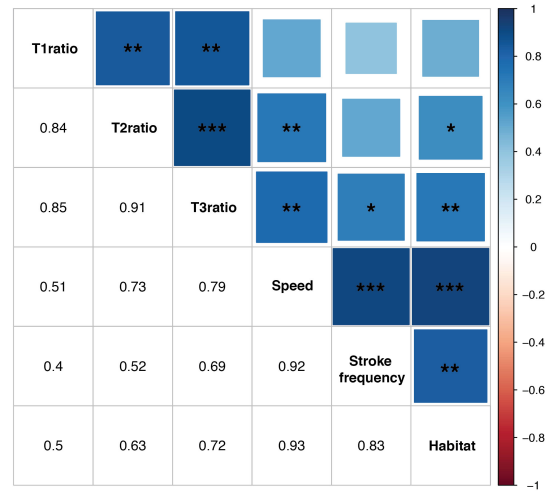
A

Correlation matrix (whole dataset)



B

Correlation matrix (Tripod gait)



* P-value < 0.1 ** P-value < 0.05 *** P-value < 0.01 **** P-value < 0.001
FDR adjustment of P-values for multiple comparisons

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

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

Species	Speed on ground (cm/s)	Speed on ground (bl/s)	Speed on water (cm/s)	Speed on water (bl/s)	Stroke frequency on ground (st/s)	Stroke frequency on water (st/s)	Mean weight (mg)	Energy per stroke ((mJ/mg)/st)	P-Value	Energy per second ((mJ/mg)/s)	P-Value
<i>Pyrrhocoris apterus</i>	8.6±3.2 (N=14)	9.1±3.4 (N=14)	NA	NA	20.4±7.7 (N=14)	NA	NA	NA	NA	NA	NA
<i>Cryptostemma alienum</i>	2.3±0.7 (N=14)	9.4±2.8 (N=14)	NA	NA	23.5±6.3 (N=14)	NA	NA	NA	NA	NA	NA
<i>Hebrus</i> sp.	1.2±0.9 (N=10)	6.0±4.4 (N=10)	1.3±0.6 (N=10)	6.4±2.7 (N=10)	22.6±6.2 (N=10)	13.4±5.0 (N=10)	NA	NA	NA	NA	NA
<i>Hydrometra stagnorum</i>	8.0±4.4 (N=10)	7.6±4.2 (N=10)	8.1±2.6 (N=10)	7.8±2.5 (N=10)	13.7±4.9 (N=10)	13.1±3.9 (N=10)	NA	NA	NA	NA	NA
<i>Paravelia conata</i>	4.7±1.5 (N=10)	8.9±2.8 (N=10)	5.0±0.04 (N=3)	9.4±0.9 (N=3)	15.1±4.8 (N=10)	8.9±0.9 (N=6)	NA	NA	NA	NA	NA
<i>Mesovelgia furcata</i>	26.1±9.3 (N=10)	74.4±37.6 (N=10)	33.1±7.5 (N=12)	103.6±23.6 (N=12)	82.2±12.7 (N=10)	71.2±10.7 (N=12)	1.16	1.3318e-05	0.08 n.s.	2.0058782e-2	4.6e-4 ***
<i>Oiovelia cunucunumana</i>	8.4±4.5 (N=10)	20.6±11.0 (N=10)	16.4±3.3 (N=8)	40.2±8.1 (N=8)	31.1±10.9 (N=10)	55.0±8.0 (N=10)	2.92	8.06773e-06		1.4839771e-2	
<i>Paravelia bullialata</i>	8.9±5.1 (N=10)	23.0±13.3 (N=10)	18.3±3.6 (N=10)	47.4±9.2 (N=10)	30.4±12.6 (N=10)	60.7±12.3 (N=10)	2.925	1.17074e-05		2.3499226e-2	
<i>Microvelia americana</i>	9.5±4.7 (N=9)	33.5±16.5 (N=9)	11.2±2.3 (N=8)	39.5±8.0 (N=8)	48.2±14.4 (N=10)	42.4±5.9 (N=9)	1.53	3.6114e-06		7.093716e-3	
<i>Stridulivelia strigosa</i>	3.8±1.2 (N=10)	6.8±4.5 (N=10)	28.1±6.5 (N=10)	64.4±5.0 (N=10)	13.3±3.4 (N=10)	9.1±2.2 (N=10)	2.733	2.21034e-05		2.92093e-4	
<i>Stridulivelia tersa</i>	5.3±1.7 (N=10)	11.0±3.5 (N=10)	37.1±8.1 (N=11)	76.3±6.6 (N=11)	14.1±2.9 (N=10)	15.5±4.9 (N=22)	3.65	5.06027e-05		3.276652e-3	
<i>Platyvelia brachialis</i>	NA	NA	12.4±4.7 (N=18)	26.5±0.1 (N=18)	NA	10.6±3.2 (N=35)	6.233	5.98386e-06		5.43207e-4	
<i>Husseyella turmalis</i>	2.8±0.9 (N=9)	13.8±4.3 (N=9)	23.2±5.5 (N=10)	113.0±26.8 (N=10)	21.2±4.9 (N=9)	6.1±5.0 (N=10)	0.5	8.38042e-07		3.00e-05	
<i>Limnopor</i>	NA	NA	80.6±8	59.3±6	NA	7.6±2.	31.1	1.38796		9.90729	

<i>us dissortis</i>			.8 (N=8)	.5 (N=8)		6 (N=8)	83	e-04		e-4
<i>Gerris buenoi</i>	1.3±0 .7 (N=8)	1.7±0. 9 (N=8)	53.7±1 2.7 (N=8)	69.8±1 6.6 (N=8)	5.0±1. 8 (N=8)	6.3±3. 7 (N=8)	12.2	7.87642 e-05		5.82314 e-4
<i>Aquarius paludum</i>	NA	NA	104.1± 12.2 (N=8)	74.4±8 .7 (N=8)	NA	6.1±1. 6 (N=8)	34.6 67	2.23603 e-04		8.10206 e-4

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

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

Gerromorpha	<i>Paravelia bipunctata</i>	Land/ Water	2983	4911	4447	5348	56	92	83	1	Brazil N -20,60034°, W-45,81498
Gerromorpha	<i>Paravelia conata</i>	Land/ Water (prefer land)	2282	3167	3602	4960	46	64	73	1	Brazil N - 18.367815°, W - 40.139877°
Gerromorpha	<i>Halovelgia electrodominica</i>	Water (sea)	1030	2260	1270	1500	69	151	85	1*	[S8]
Gerromorpha	<i>Husseyella turmalis</i>	Water	1296±7 8	2820 ±193	1769±9 8	2023 ±263	65±5.5	140±10. 5	88±7.3	6	French Guiana, N 04.34235°, W - 052.10730°
Gerromorpha	<i>Microvelgia electra</i>	--	620	710	950	1330	47	53	71	1*	[S9]
Gerromorpha	<i>Microvelgia grimaldii</i>	--	910	1050	1320	1750	52	60	75	1*	[S9]
Gerromorpha	<i>Microvelgia polhemi</i>	--	1060	1250	1490	1850	57	68	81	1*	[S10]
Gerromorpha	<i>Microvelgia americana</i>	Land/ Water (prefer water)	1798±7 6	2344 ±11	2975±1 39	2859 ±178	68±5.2	88±6.8	112±8.2	10	Rivière du Nord, Montréal, Québec, Canada
Gerromorpha	<i>Cretogerris albianus</i>	--	1390	3150	3810	1840	76	171	207	1*	[S11]
Gerromorpha	<i>Electrobates spinipes</i>	--	1820	5210	3630	2580	71	202	141	1*	[S12]
Gerromorpha	<i>Metrobates hesperius</i>	Water	3828±8 0	1331 5±39 7	8990±2 98	4596 ±154	83±2.5	290±11. 8	196±9	10	Rivière du Nord, Montréal, Québec, Canada
Gerromorpha	<i>Limnoporus dissortis</i>	Water	7371±3 32	1884 3±84 4	17439± 1172	1358 0±56 3	54±0.9	139±2.3	128±4	10	Rivière de l'Acadie, Montréal, Québec, Canada
Gerromorpha	<i>Gerris buenoi</i>	Water	4584±1 40	1124 7±58 7	7679±4 17	7697 ±428	60±3	146±8.2	100±5.7	6	Toronto, Ontario, Canada
Gerromorpha	<i>Aquarius paludum</i>	Water	9003±7 21	2384 0±18 21	21206± 1793	1398 6±16 21	65±2.6	171±8.5	152±6.9	8	Plan d'eau de Saloniques, Vilette d'Anthon, France

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

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

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

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

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

Supplemental experimental procedures

Insect sampling and culture

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

Phylogenetic reconstruction

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

Habitat classification

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

Video acquisition, quantification of speed and stroke frequency

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

Measurements of leg length and body length

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

Analysis of leg pattern during locomotion

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

Reconstruction of ancestral trait

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

Inference of energy consumption

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

$$E_k=0.5mv^2$$

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

Statistical analyses

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

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

Online resources

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

Supplemental references

- S1. D. Azar, A. Nel, M. S. Engel, R. Garrouste, and Matocq, A. (2011). A new family of Coreoidea from the Lower Cretaceous Lebanese amber (Hemiptera: Pentatomomorpha). *Polish Journal of Entomology*, 627-644.
- S2. Redei, D. (2007). A new species of the family Hypsipterygidae from Vietnam, with notes on the hypsipterygid fore wing venation (Heteroptera, Dipsocoromorpha). *Deut Entomol Z* 54, 43-50.
- S3. Azar, D., and Nel, A. (2010). The earliest fossil schizopterid bug (Insecta: Heteroptera) in the Lower Cretaceous amber of Lebanon. *Ann Soc Entomol Fr* 46, 193-197.
- S4. Yao, Y.Z., Zhang, W.T., and Ren, D. (2012). The first report of Mesoveliidae (Heteroptera: Gerromorpha) from the Yixian Formation of China and its taxonomic significance. *Alcheringa* 36, 107-116.
- S5. Garrouste, R., and Nel, A. (2010). First semi-aquatic bugs Mesoveliidae and Hebridae (Hemiptera: Heteroptera: Gerromorpha) in Miocene Dominican amber. *Insect Syst Evol* 41, 93-102.
- S6. Nel, A., and Popov, Y.A. (2000). The oldest known fossil Hydrometridae from the Lower Cretaceous of Brazil (Heteroptera : Gerromorpha). *J Nat Hist* 34, 2315-2322.
- S7. Zettel, H. (2001). Five new species of Perittopus Fieber, 1861 (Hemiptera : Veliidae) from Southeast Asia. *Raffles B Zool* 49, 109-119.
- S8. Andersen, N.M., and Poinar, G.O. (1998). A marine water strider (Hemiptera : Veliidae) from Dominican amber. *Entomol Scand* 29, 1-9.
- S9. Andersen, N.M. (2000). Fossil water striders in the Oligocene/Miocene Dominican amber (Hemiptera : Gerromorpha). *Insect Syst Evol* 31, 411-431.
- S10. Andersen, N.M. (1999). *Microvelia polhemi*, n. sp (Heteroptera : Veliidae) from Dominican amber: The first fossil record of a phytotelmic water strider. *J New York Entomol S* 107, 135-144.
- S11. Perrichot, V., Nel, A., and Neraudeau, D. (2005). Gerromorphan bugs in Early Cretaceous French amber (Insecta : Heteroptera): first representatives of Gerridae and their phylogenetic and palaeoecological implications. *Cretaceous Res* 26, 793-800.
- S12. Andersen, N.M., and Poinar, G.O. (1992). Phylogeny and Classification of an Extinct Water Strider Genus (Hemiptera, Gerridae) from Dominican Amber, with Evidence of Mate Guarding in a Fossil Insect. *Z Zool Syst Evol* 30, 256-267.
- S13. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30, 3059-3066.
- S14. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754-755.
- S15. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52, 696-704.
- S16. Andersen, N.M. (1979). Phylogenetic Inference as Applied to the Study of Evolutionary Diversification of Semiaquatic Bugs (Hemiptera, Gerromorpha). *Syst Zool* 28, 554-578.
- S17. Andersen, N.M. (1982). The semiaquatic bugs (Hemiptera: Gerromorpha). Volume - Entomonograph Vol. 3. , (Klampenborg, Denmark.: Scandinavian Science Press LTD.).
- S18. Dias-Silva, K., Moreira, F.F.F., Giehl, N.F.D., Nobrega, C.C., and Cabette, H.S.R. (2013). Gerromorpha (Hemiptera: Heteroptera) of eastern Mato Grosso State, Brazil: checklist, new records, and species distribution modeling. *Zootaxa* 3736, 201-235.
- S19. Heiss, E., and Pericart, J. (2007). Faune n° 91 – Hémiptères Aradidae, Piesmatidae et Dipsocoromorphes, (Fédération Française des Sociétés de Sciences naturelles).
- S20. Schuh, R.T., and Slater, J.A. (1995). True bugs of the world (Hemiptera: Heteroptera) : classification and natural history, (Ithaca ; London: Comstock Publishing Associates, a division of Cornell University Press).
- S21. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9, 671-675.

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