

EVO-DEVO

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

M. Emília Santos,^{1*} Augustin Le Bouquin,^{1,2} Antonin J. J. Crumière,¹ Abderrahman Khila^{1*}

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

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

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

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

of these genes in the *Rhagovelia* lineage thus coincides with the emergence of the fan, suggesting that they are involved in the evolution of this trait.

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

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

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

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

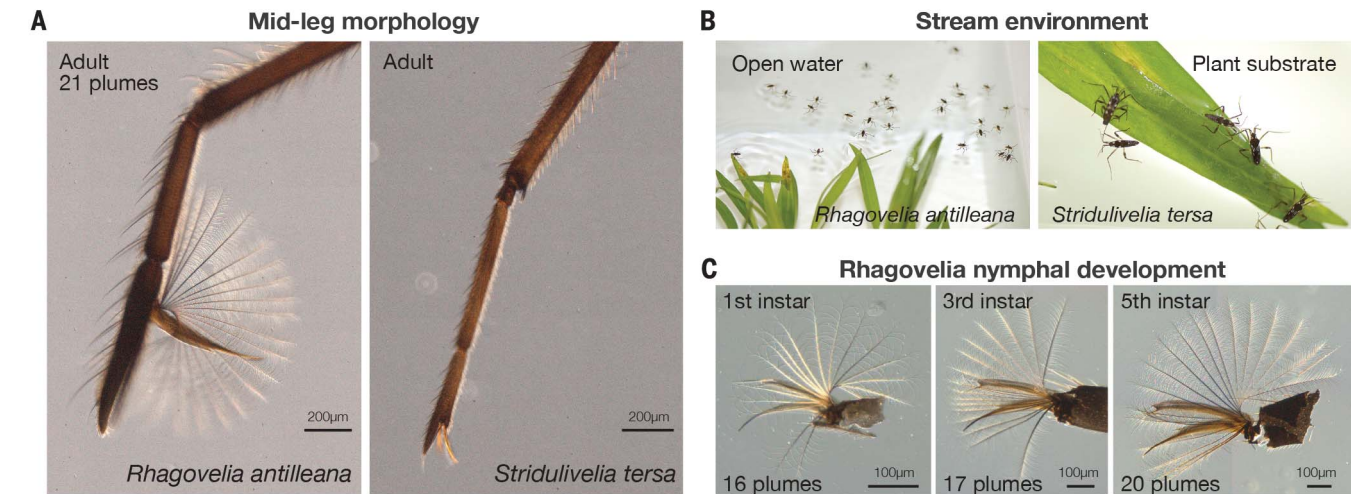


Fig. 1. Morphology, habitat correlation, and nymphal development of the *Rhagovelia* fan. (A) Midleg morphology comparison between *R. antilleana* and the outgroup species *S. tersa*. *Rhagovelia* tarsi have longer and wider claws and are forked to host the fan, which can be deployed and retracted following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

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

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

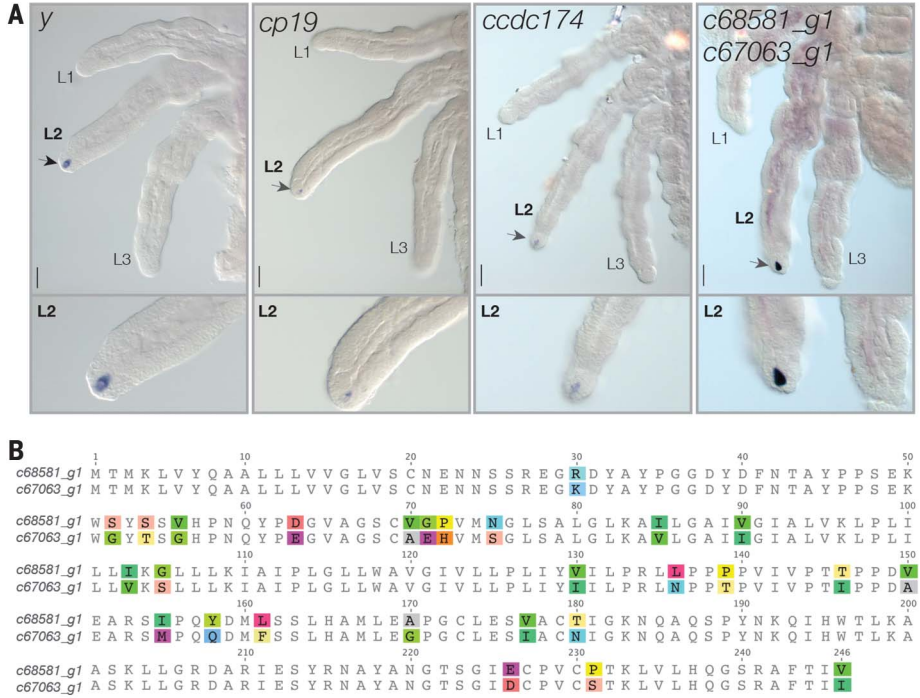


Fig. 2. Fan expression profiling and protein alignment of *gsha* and *mogsha*. (A) In situ hybridization showing the expression of all five genes at the tips of the midleg (leg 2, L2). Scale bar, 100 µm. (B) Protein alignment of c68581_g1 (*mogsha*) and c67063_g1 (*gsha*), with divergent amino acids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

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

Next, we challenged these groups of individuals in a setup where they were forced to row against a slow (0.2 m/s) or fast (0.3 m/s) water

current (Fig. 4C and movie S3), simulating *Rhagovelia*'s stream environment. Although *S. tersa* animals were the fastest on still water, they failed to ascend the stream setup (Fig. 4C, fig. S9, and table S9). Normal *R. antilleana* ascended both currents without difficulty (Fig. 4C,

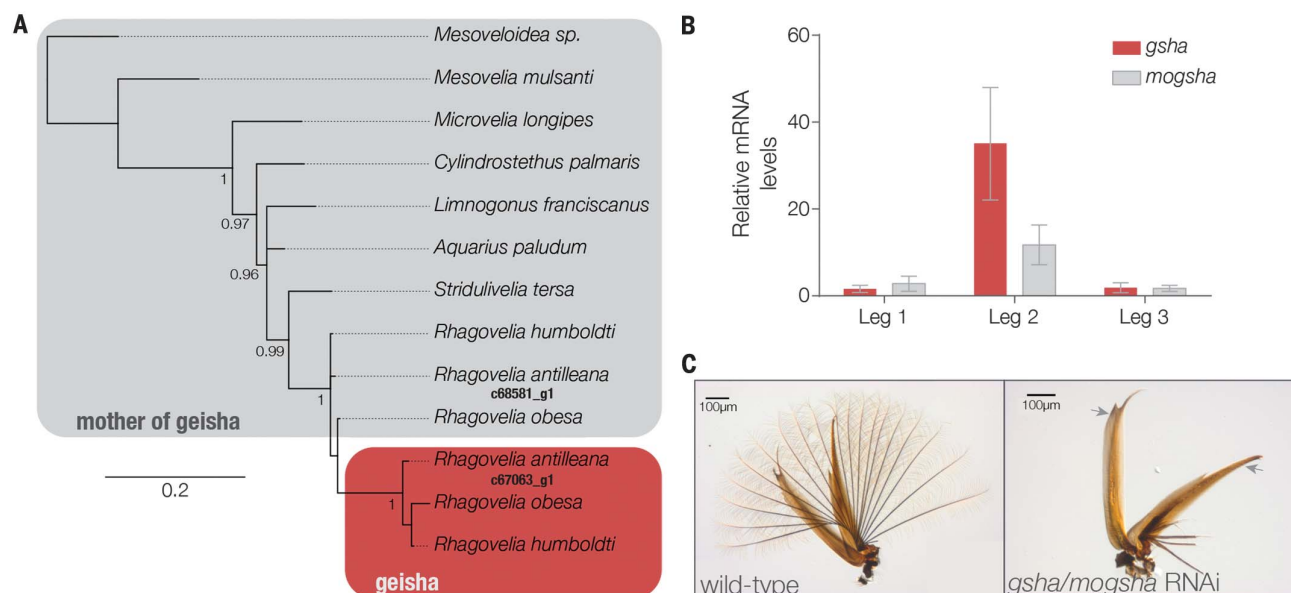


Fig. 3. Sequence evolution, expression, and functional characterization of *gsha* and *mogsha*. (A) Bayesian phylogenetic reconstruction showing that *mogsha* is more similar to the taxon-restricted unduplicated copy (gray box), whereas *gsha* is a *Rhagovelia*-specific duplicate that has evolved at a higher rate (red box; note the branch length). Node values represent the posterior probability, and the scale bar indicates the

genetic distance between sequences. (B) Gene-specific quantitative PCR showing *gsha* and *mogsha* expression in the midleg (leg 2; $n = 3$) of fourth instar nymphs. Error bars, SEM. (C) *gsha/mogsha* RNAi-mediated knockdown resulted in a severe reduction of the fan ($n = 27$), but the claws (arrows) and leg morphology (figs. S7 and S8) were unaffected.

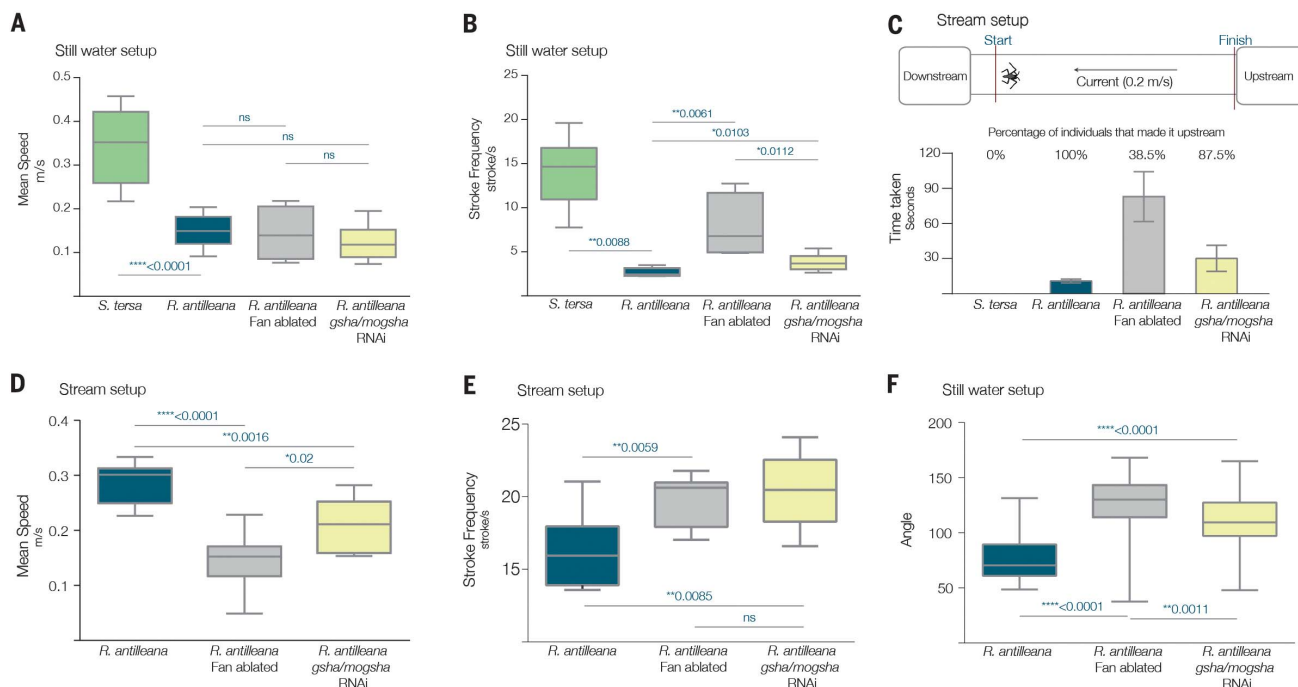


Fig. 4. Comparative analysis of locomotion parameters. (A and B) Boxplots for the still water setup, comparing (A) mean speed and (B) stroke frequency among *S. tersa* (green; $n = 10$), *R. antileana* (blue; $n = 7$), *R. antileana* with ablated fans (gray; $n = 4$), and *gsha/mogsha*-RNAi *R. antileana* (yellow; $n = 9$). The range of severity of *gsha/mogsha*-RNAi phenotypes can be found in fig. S7. (C) Scheme of the stream setup (top) (movie S3), with the percentage of individuals succeeding in rowing against the current, together with their corresponding arrival time, shown below. Error bars, SEM. (D and E) Boxplots for the stream setup,

comparing (D) mean speed and (E) stroke frequency among *R. antileana* (blue; $n = 10$), *R. antileana* with ablated fans (gray; $n = 8$) and *gsha/mogsha*-RNAi *R. antileana* (yellow; $n = 8$). (F) Angles of *R. antileana* turns in still water (*R. antileana*, $n = 8$; *gsha/mogsha*-RNAi, $n = 8$; *R. antileana* with fan ablated, $n = 8$). Each boxplot spans the minimum and maximum values (whiskers), with medians shown by the horizontal lines and first and third quartiles delineated by the box. Student's *t* test or the Mann-Whitney test was performed for each pairwise comparison. *P* values are shown next to asterisks. ns, not significant.

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

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

Collectively, our data provide important insights into the developmental genetic, evolu-

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

REFERENCES AND NOTES

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

ACKNOWLEDGMENTS

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

SUPPLEMENTARY MATERIALS

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

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

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

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

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

Fans enable water strider adaptation

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

Science, this issue p. 386

ARTICLE TOOLS

<http://science.sciencemag.org/content/358/6361/386>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2017/10/19/358.6361.386.DC1>

REFERENCES

This article cites 51 articles, 4 of which you can access for free
<http://science.sciencemag.org/content/358/6361/386#BIBL>

PERMISSIONS

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

Use of this article is subject to the [Terms of Service](#)

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2017 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works