Mating behavior and the evolution of sperm design

Lukas Schärer\textsuperscript{a,1}, D. Timothy J. Littlewood\textsuperscript{b}, Andrea Waeschenbach\textsuperscript{b}, Wataru Yoshida\textsuperscript{c}, and Dita B. Vizoso\textsuperscript{a}

\textsuperscript{a}Evolutionary Biology, Zoological Institute, University of Basel, 4051 Basel, Switzerland; \textsuperscript{b}Department of Zoology, Natural History Museum, London SW7 5BD, United Kingdom; and \textsuperscript{c}Department of Biology, Faculty of Agriculture and Life Science, Hiroasaki University, Aomori 036-8561, Japan

Edited by John C. Avise, University of California, Irvine, CA, and approved December 1, 2010 (received for review September 16, 2010)

Sperm are the most diverse of all animal cell types, and much of the diversity in sperm design is thought to reflect adaptations to the highly variable conditions under which sperm function and compete to achieve fertilization. Recent work has shown that these conditions often evolve rapidly as a consequence of multiple mating, suggesting a role for sexual selection and sexual conflict in the evolution of sperm design. However, very little of the striking diversity in sperm design is understood functionally, particularly in internally fertilizing organisms. We use phylogenetic comparative analyses covering 16 species of the hermaphroditic flatworm genus Macrostomum to show that a complex sperm design is associated with reciprocal mating and that this complexity is lost secondarily when hypodermic insemination—sperm injection through the epidermis—evolves. Specifically, the complex sperm design, which includes stiff lateral bristles, is likely a male persistence trait associated with sexual conflicts over the fate of received ejaculates and linked to female resistance traits, namely an intriguing postcopulatory sucking behavior and a thickened epithelium of the sperm-receiving organ. Our results suggest that the interactions between sperm donor, sperm, and sperm recipient can change drastically when hypodermic insemination evolves, involving convergent evolution of a needle-like copulatory organ, a simpler sperm design, and a simpler female genital morphology. Our study documents that a shift in the mating behavior may alter fundamentally the conditions under which sperm compete and thereby lead to a drastic change in sperm design.

Platyhelminthes | sexually antagonistic coevolution | simultaneous hermaphrodite | sperm morphology | traumatic insemination

Parkinson’s (1–3) far-reaching extension of Darwin’s (4) narrow focus on precopulatory mating interactions highlighted that sexual selection continues to operate after mating partners have agreed to mate and that considering postcopulatory sexual selection therefore is crucial to understanding the evolution of many reproductive traits (5, 6). This insight has led to extensive research in evolutionary biology that focused on understanding the biology of sperm (7), the most diverse of all animal cell types (8, 9). From this research emerged an apparent consensus that the diversity in sperm design—the strikingly variable ways of constructing a sperm—reflects the highly variable physiological and morphological environments in which sperm have to survive, function, and compete for fertilization (5, 10–14). Moreover, recent studies have documented clearly that these environments can evolve rapidly, probably because of coevolutionary interactions linked to multiple mating and the resulting sexual selection and sexual conflicts (15–21). However, the bewildering diversity in sperm design is poorly understood at the functional level, particularly in internally fertilizing organisms (9, 12, 22). A recent review on the evolution of sperm morphological diversity concluded that we “currently have only a rudimentary understanding of the adaptive significance of [the awe-inspiring variation in sperm form]” and that we “know very little about sperm behavior, particularly within the female reproductive tract” (9). The reason for this lack of understanding is that the “goings-on” inside the female reproductive tract often are difficult to study (but see refs. 23–25).

In organisms with internal fertilization, multiple mating by females generally is expected to lead to sexual conflicts over the fate of received ejaculates (19). Females may mate multiply for different reasons: because it allows them to choose among sperm from different males based on compatibility (26) or male quality (including the competitive quality of their sperm) (27, 28), because they can obtain material benefits from the males (e.g., nuptial gifts or an adequate sperm supply) (26), or because, even though multiple mating may be costly to females, resisting male harassment is even costlier (18). After multiple mating, females may attempt to control the fate of the received ejaculates, either to exert choice or to lower direct costs [e.g., resulting from polyspermy (18, 29)], thereby removing some sperm from the fertilization set. In contrast, males always should prefer that their sperm, rather than that of a competitor, be used for fertilization (19), leading to sexual conflict. This definition of sexual conflict is a broad one (19, 30) and follows the reasoning of Parker (19) that “sexual conflict is present in all forms of female choice involving the rejection of some males, whether rejection occurs because they are not attractive enough or because of the costs they impose” (p. 236). In this scenario males therefore are expected to develop male persistence traits that allow them to overcome such female control, even if that persistence occurs at a cost to the females (19). This situation may lead to sexually antagonistic coevolution between the male persistence traits and the female resistance traits [as has been pointed out before (31), female resistance is equivalent to female preference in that both lead to a bias in the reproductive success toward persistent males] and is expected to affect the evolution of both the sperm and their environment.

In species with separate sexes, females often may be able to evade some of these postcopulatory sexual conflicts by avoiding males—either altogether or via precopulatory female choice—whenever the costs of mating exceed the benefits. This outcome is reflected nicely in the classical (precopulatory) mating roles, namely, eager males and choosy females (32–34). In contrast, simultaneous hermaphrodites cannot resolve these conflicts so easily, because each individual is both male and female at the same time. Here mating should be attempted whenever an individual can gain a net benefit from mating. For example, when its male benefits minus its female costs are positive (35). This inequality will be fulfilled even if the costs of mating to an individual’s female fitness are substantial, as long as they are compensated by sufficiently large benefits to its male fitness. Multiple mating in the female role therefore may occur not only for the reasons mentioned above but also as a consequence of behaviors that primarily serve the interests of an individual’s male sex function, such as actively approaching mating partners.

Author contributions: L.S. and D.B.V. designed research; L.S., D.T.J.L., A.W., W.Y., and D.B.V. performed research; L.S., D.T.J.L., A.W., and D.B.V. analyzed data; and L.S. wrote the paper. L.S. established the online Macrostomorpha Taxonomy and Phylogeny database; and D.B.V. made the drawings of the reproductive structures.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. FJ715315–FJ715350).

1To whom correspondence should be addressed. E-mail: lukas.scharer@unibas.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1013892108/-/DCSupplemental.
and attempting to engage in sperm donation. Matings that are disadvantageous to the female function therefore are more likely to occur in simultaneous hermaphrodites, shifting sexual selection and sexual conflict from the pre- to the postcopulatory stage. This argument assumes that multiple mating in simultaneous hermaphrodites is primarily male-driven and that all individuals in a population therefore tend to show a general willingness to donate but not necessarily to receive sperm in most matings, leading to conflicting mating interests (36). The extent to which the biology of many simultaneous hermaphrodites supports this assumption is an important focus of empirical research and is the subject of an ongoing debate (37–39), but in the following discussion we will assume that it approximately holds.

One solution to such conflicting mating interests is reciprocal mating in which mating partners simultaneously assume both the male and the female role, with individuals accepting an ejaculate from the partner in exchange for an opportunity to donate an ejaculate (36, 40). Although this behavior initially appears to be a cooperative solution, it offers ample opportunity for postcopulatory sexual selection and sexual conflicts. Specifically, we expect the evolution of female resistance traits that allow the fate of any unwanted ejaculates to be controlled (36, 40) [e.g., sperm digestion (41, 42)] and male persistence traits that defend against such control [e.g., love darts and chemical manipulation (42, 43)]. A suite of reproductive characters in the free-living flatworm *Macrostomum lignano* is highly suggestive of such a scenario (Fig. 1A). Based on detailed behavioral (44), in vivo light microscopic, and transmission electron microscopic observations (23), we here define the “reciprocal mating syndrome.” First, mating is reciprocal, with the stylet (the male intromittent genitalia) of each mate inserted into the partner via the female genital opening and sperm deposited into the female antrum (the sperm-receiving organ) of its partner (23). Moreover, mating often is followed directly by a postcopulatory sucking behavior (44) during which the worm bows down onto itself, places its pharynx over its own female genital opening, and appears to suck. After this behavior, sperm shafts often stick out of the female genital opening (Movie S1). We have hypothesized that this behavior is a female resistance trait involved in manipulating the fate of the received ejaculate (23, 44). Second, the aflagellate sperm also are highly motile (Movie S6) but are smaller, simpler, and carry no bristles, presumably facilitating their movement through tissue. Third, the female antrum is simple, without a thickened epithelium, and lacks a prominent cellular valve. The female antrum is involved only in egg laying.

Here we examine the evolution and coevolution of these morphological and behavioral characters in the free-living flatworm genus *Macrostomum*. Specifically, we are interested in understanding how stable and widespread the reciprocal and hypodermic mating syndromes are and how they are distributed phylogenetically. To this end, we collected and morphologically described 16 different *Macrostomum* species, established the molecular phylogeny of this taxonomic group, and used this information to perform phylogenetic comparative analyses, including character mapping, constraint analyses, ancestral state reconstructions, and analyses of correlated evolution. Our results show that the majority of species fall into one of the two syndromes and that the molecular phylogeny does not fully match this morphological dichotomy, with the hypodermic mating syndrome having re-evolved independently at least once from the reciprocal mating syndrome. Moreover, we find strong evidence for correlated evolution between the mating behavior and both male and female morphological characters, shedding light on the evolution of sperm design.

### Results

#### Molecular Phylogeny, Character Mapping, and Constraint Analysis.

To study the phylogenetic distribution of the above-mentioned reproductive character states, we assembled a molecular phylogeny comprising 16 species of the free-living flatworm genus *Macrostomum*. By mapping the character states of the collected species (Table S1) onto this phylogeny, we can identify a first clade (clade 1 in Fig. 2), all members of which share the hypodermic mating syndrome. In addition, we can identify a second clade (clade 2 in Fig. 2), most members of which exhibit the reciprocal mating syndrome. However, clade 2 also includes

---

**Fig. 1.** Morphology of the sperm and stylet of two *Macrostomum* species. (A) *M. lignano*, a species that represents the reciprocal mating syndrome. (B) *M. hystrix*, a species that represents the hypodermic mating syndrome.
one species, *M. hystrix*, with the hypodermic mating syndrome 
(second species from the bottom in Fig. 2) (for a detailed tree 
with outgroups and exact nodal support values see 
Fig. S1). A 
Shimodaira–Hasegawa test strongly rejects a constrained tree 
with a single origin of the hypodermic mating syndrome (i.e., 
forcing *M. hystrix* within clade 1) as an alternative a priori hy-
pothesis to the current placement of this species ($P < 0.001$) 
(Fig. S2). The split between the two clades and the nodes at their 
base are well supported (Fig. S1), suggesting that the hypodermic 
mating syndrome has evolved independently twice, representing 
a clear case of convergent evolution. Indeed, the species exhib-
iting the hypodermic mating syndrome show striking similarities 
in the morphology of sperm, stylet, and female antrum, whereas 
species with the reciprocal mating syndrome exhibit considerable 
interspecific variation in these structures (Fig. 2). Additionally, 
*Macrostomum finlandense*, whose mating behavior currently is 
unknown, appears to be in transition between the two syndromes. 
Its sperm bristles are short but externally visible, and although the 
stylet is pointed and fairly rigid, it lacks a pointed distal thick-
ening and has an oblique rather than subterminal stylet opening.

**Ancestral State Reconstruction.** To understand further the origins 
and evolution of the studied reproductive characters, we per-
formed ancestral state reconstructions to assess the most probable 
character states at important nodes in the molecular phylogeny. 
Although the results of these analyses do not allow conclusions to 
be drawn about the ancestral states of any of the investigated 
characters at the base of the genus *Macrostomum*, they clearly 
support the presence of the reciprocal mating syndrome at the 
base of clade 2 (for trees with the estimated ancestral character 
states, see Fig. S3). This result further supports the independent 
origin of the hypodermic mating syndrome in *M. hystrix*.

**Correlated Evolution.** To assess statistically the degree to which 
there is coevolution between the morphological and behavioral 
characters, we performed two analyses of correlated evolution 
(analysis A: male morphology vs. copulation behavior; analysis B: 
female morphology vs. copulation behavior). The likelihoods re-
sulting from these analyses indicate strong evidence for correlated 
evolution, both for analysis A between the sperm design (and the 
stylet morphology) and copulation behavior ($\log H_{\text{dep}} = -9.82$, 
$\log H_{\text{indep}} = -13.01$, test statistic = 6.38), and for analysis B

---

**Fig. 2.** Phylogenetic mapping of reproductive character states. Variation in sperm and stylet morphology, mating behavior, and female antrum morphology among 16 species in the genus *Macrostomum*. (Details on character states are given in Table S1.) The character states are mapped on the maximum-likelihood tree of the genus (complete ssrDNA and partial IsrDNA), and nodes marked with a circle have Bayesian posterior probability >0.95 and maximum-likelihood bootstrap support of >70%. (A detailed tree with outgroups and exact nodal support values is given in Fig. S1.) Note that the stylet of *Macrostomum tuba* is depicted at one-third of the original size.
Correlated evolution of reproductive character states. Schematic diagram showing the most probable evolutionary routes for transitions between the four possible combinations of character states of two binary variables. For analysis A, the first variable represents the sperm morphology (0, bristles absent; 1, bristles present) or the stylet morphology (0, needle-like; 1, not needle-like) (SI Materials and Methods). For analysis B, the first variable represents female antrum morphology (0, simple; 1, thickened). For both analyses A and B, the second variable represents the copulation behavior (0, hypodermic; 1, reciprocal). Because the results of the two analyses are qualitatively similar, we show only one summary graph. The different arrows indicate three classes of transition rate parameters: transition rates with a high probability (>55%, thin arrows), intermediate probability (30–40%, medium arrows), and low probability (<15%, thick arrows) of being zero. (The actual distributions of the transition rate parameters are given in Fig. S4.) Thus, thicker arrows represent more likely transitions.

Discussion

Our study reveals an evolutionary link between mating behavior, male and female genital morphology, and sperm design. Specifically, our results suggest that the character states defining the reciprocal mating and the hypodermic mating syndromes in the genus *Macrostomum* are the result of concerted evolutionary changes and show that the majority of the studied species clearly exhibit one of the two syndromes. The reciprocal mating syndrome probably is driven by sexual conflict over the fate of received ejaculates; this conflict occurs as a result of reciprocal mating, leading to elaborate and highly variable female resistance and male persistence traits. In contrast, the hypodermic mating syndrome probably results from selection to by-pass female resistance traits (36, 40, 47, 48), leading to striking convergent evolution with little morphological variation between species.

Hypodermic insemination is expected to impose novel selection pressures related to this type of sperm transfer and simultaneously to relax selection on the character states defining the reciprocal mating syndrome. For example, sperm may be selected for efficient movement through tissue, perhaps favoring the loss of the bristles. As hypodermic insemination spreads through a population, the way sperm of different donors compete within the body of the recipient should resemble more closely a fair-ruffle sperm competition (2, 48), potentially favoring smaller and more numerous sperm. Furthermore, we expect selection on the stylet for efficient hypodermic delivery of sperm and that the evolution of the stylet no longer will be affected by the female antrum. Similarly, selection on the female antrum will be solely on the function of egg laying, because the antrum no longer interacts with the stylet or the sperm.

A more detailed look at the stylet and female antrum morphologies associated with the reciprocal mating syndrome may help identify possible starting points for the evolution of hypodermic insemination, because these morphologies suggest that wounding might occur during reciprocal mating. In several species the stylets carry blunt distal thickenings, indicative of selection on the male function against wounding of the female antrum. Furthermore, at least two species of *Macrostomum* have sclerotized regions in the female antrum (49, 50) that might represent female adaptations against wounding. However, when the costs of wounding are small for the sperm donor, accidental wounding during copulation could lead eventually to the evolution of hypodermic insemination.

The evolutionary origin of the sperm bristles currently is unclear (51). A study of sperm ultrastructure suggests a homology between the bristles in *Macrostomum* and lateral ledges in the sperm of *Bradynectes sterreri* (one of our outgroup species) and *Psammomacrostomum turbanelloides* (51) (a likely relative of our Gen. nov. 1, sp. nov. 1). These morphological structures do not protrude outside the sperm, however, and therefore cannot have an anchoring function. Moreover, it has been shown that *Macrostomum pusillum* (a member of clade 1 in Fig. 2) has rudimentary bristles that are visible only at the ultrastructural level (52). Together, these findings suggest that the morphological structures that gave rise to the bristles were already present at
the base of the genus *Macrostomum*. The externally visible bristles thus may have arisen at the base of the genus and been lost twice, or they may have arisen at the base of clade 2 and been lost once. Future studies on a wider range of taxa may arbitrate between these two scenarios and may uncover more cases of gains and losses in this structure.

It has been suggested that sexual selection and sexual conflicts can generate rapid evolutionary diversification in reproductive traits, and this notion is reflected by the use of these traits for taxonomy in many groups of organisms (5, 9, 11, 53, 54). Our results suggest that the degree to which this morphological diversification occurs may depend on the nature of the mating behavior. The striking convergent evolution between the species exhibiting the hypodermic mating syndrome (i.e., *M. hystrix* and the members of clade 1) helps explain why these species are difficult to distinguish based on reproductive morphology alone and has considerable impact on a number of pressing taxonomical questions in this genus (*SI Materials and Methods*). In contrast, the reciprocal mating syndrome appears to generate extensive diversification in reproductive traits, thus providing ample morphological information to distinguish species. It is interesting to speculate that the greater diversification within the reciprocal mating syndrome may make it less evolutionarily stable than the hypodermic mating syndrome, as suggested by the correlated evolution analyses (Fig. 3). If so, we would expect that expanding the taxon sampling will reveal additional departures from the reciprocal mating syndrome in clade 2. Alternatively, because clade 1 currently contains few studied species, it also is possible that with larger taxon sampling we will reveal departures from the hypodermic mating syndrome.

Most studies that have investigated variation in sperm morphology in internally fertilizing organisms have aimed at explaining quantitative variations such as in sperm length within one particular sperm design and have identified sperm competition (9, 12, 55) or the interactions between the sperm and the female genital tract (11, 17, 21, 56) as the main evolutionary forces. Our results suggest that a change in the mating behavior may lead to strikingly different sperm designs within a single genus, and we thus have identified conditions under which drastic shifts in sperm design may arise rapidly and repeatedly. Shifts in mating behavior from normal copulation to hypodermic or traumatic insemination have occurred many times in both hermaphrodite (e.g., acel flatworms (57), polyclad flatworms (58), and sea slugs (59)) and separate-sexed organisms (e.g., bed bugs (47), plant bugs (60), and drosophilids (61)). Therefore it will be interesting to investigate further the consequences of these shifts for the evolution of sperm design, morphology, and function. The striking diversification in sperm and genital morphology in the genus *Macrostomum*, combined with the transparent nature of these worms, offers a veritable ‘window of opportunity’ to begin to understand the evolution of sperm morphological diversity.

**Materials and Methods**

**Specimens.** We collected specimens in a range of countries, habitats, and water bodies, using a range of extraction techniques (i.e., decantation with MgCl₂, oxygen deterioration, direct extraction). For each species we give information (*SI Materials and Methods*) on specimens, collection sites, species identification, taxonomic notes, and accession numbers to digital reference material deposited on the online Macrostomorpha Taxonomy and Phylogeny database, an EDIT scratchpad (62) available at http://macrostromorpha.info. Given the highly variable quality of taxonomic descriptions of the >100 species in the genus, only data on *Macrostomum* species that we have collected and observed ourselves were included in our data set, but data for the outgroups were taken in part from the literature.

**Digital Documentation and DNA Samples.** We observed live specimens in squeeze preparations using a Diaplan or a DM2500 compound microscope (Leica Microsystems) with bright-field, differential interference contrast, or phase-contrast illumination and at magnifications of 40×–1000× and documented their morphology with spatially calibrated micrographs and video, using a digital videocamera (DFW-X700 (SONY) or DFK 418BF02 ( Imaging Source)) and image capture software (BTV Pro 5.4.1 or 6.0.1β, available at http://www.bensoftware.com). Because these worms are transparent, we can observe and document internal structures in detail without histological sections (23, 63). Stylers were scored as “needle-like” if they had a rigid and pointed distal thickening and a subterminal stylist opening, and the female antrum was scored as “thickened” if the epithelium was clearly visible (Movie 53) and/or if there was a clear cellular valve (23) (Table S1). To document sperm morphology, we recovered worms from squeeze preparations, amputated the tail plate (64), and ruptured it by placing it on a microscope slide in 1–2 μL of liquid and covering it with a coverslip. Fully formed sperm emerged from the ruptured seminal vesicle and were documented (65). Because sperm are sensitive to low osmotic pressure, we added small amounts of diluted seawater for freshwater species to prevent osmotic damage to the sperm (artifacts thereof sometimes are reported as bona fide sperm morphology). Sperm were scored as “bristles present” if they had externally visible bristles (Table S1). To prepare a DNA sample, we placed the frontal part of the worm into absolute ethanol and stored it at 4 °C. We documented the mating behavior with our established technique (44) for a total of 6–52 h of observation of 4–29 individuals per field-collected species (in pairs and larger groups) and analyzed the videos by frame-by-frame observation. However, not all field-collected species were under these conditions, leading to some missing data in these character states (Table S1). Moreover, extensive observational data were derived from observations of the laboratory-cultured species (i.e., *M. pusillum* (Lignano), *M. spirale*, *M. lignano*, and *M. hystrix*).

**PCR, Sequencing, and Phylogenetic Analysis.** For each species we sequenced (in one to three specimens) the complete small subunit ribosomal DNA (srrDNA) and part (D1–D3) of the large subunit ribosomal DNA (IsrDNA) resulting in a total of >2,850 base pairs. (Details on primers, PCR, sequencing, and position of sequences are given in Tables S2 and S3 and *SI Materials and Methods*.) The resulting sequences were used to generate a phylogenetic hypothesis for the interrelationships between the studied taxa using both Bayesian inference and maximum likelihood analyses. (Details on alignments and parameters used in analyses are given in *SI Materials and Methods*.)

**Constraint Analysis, Ancestral State Reconstruction, and Analysis of Correlated Evolution.** We performed a constraint analysis (66) that tested whether the data support the a priori hypothesis of a single origin of the hypodermic mating syndrome. (Details are given in Fig. S2 and *SI Materials and Methods*.) Next we performed ancestral state reconstructions (67) to infer the character states at the base of the genus *Macrostomum* and the base of clade 2, both nodes that are well supported in the maximum likelihood and Bayesian inference analyses. (Details are given in Table S1, Fig. S3, and *SI Materials and Methods*.) Finally, we tested for correlated evolution between pairs of discrete binary character states (i.e., male morphology vs. mating behavior and female morphology vs. mating behavior) using BayesDiscrete implemented in BayesTraits (68). The analyses were performed on a reduced taxon set containing only the genus *Macrostomum*, the main focus of our study. (Details are given in Fig. 3, Fig. S4, and *SI Materials and Methods*.)

**ACKNOWLEDGMENTS.** We are indebted to the late Reinhard Rieger for introducing us to the taxonomy of the Macrostromorpha, and thank Seth Tyler and colleagues for maintaining the Turbellarian Taxonomic Database (http://turbellaria.umaine.edu). We also thank many volunteers and colleagues who helped us with field collections, especially Gregor Schulte, Peter Ladurner, Floriano Papli, and Werner Armonies. We are grateful to Göran Arngqvist, Dieter Ebert, Tim Janicke, Joris Koene, Nico Michiels, Markus Pfenninger, Steve Ramn, Gunde Rieger, and Walter Salzburger for discussion and/or comments on earlier versions of the manuscript.


1494 | www.pnas.org/cgi/doi/10.1073/pnas.1013892108

Schärer et al.


44. Schärer et al. (2007) EVOLUTION
Supporting Information

Schärer et al. 10.1073/pnas.1013892108

SI Materials and Methods

PCR and Sequencing. For each species we sequenced (in one to three specimens) the complete small subunit ribosomal DNA (ssrDNA) and part (D1–D3) of the large subunit ribosomal DNA (lsrDNA) resulting in a total of ~2,850 base pairs. Total genomic DNA (gDNA) was extracted from the DNA samples using the DNeasy tissue kit (QIAGEN) following the manufacturer’s instructions; gDNA was eluted in 2× 100-μL volumes. PCR reactions were carried out in 25-μL volumes using Illustra puReTaq Ready-To-Go PCR beads (GE Healthcare), 1 μL of 10 μM of each primer (a list of primers is given in Table S2), and 1–2 μL gDNA extract. Partial lsrDNA (1,142–1,189 base pairs) was amplified using ZX-1 (1) + 1500R (2); difficult templates were amplified with nested PCR using ZX-1 + 1200R and 300F + 1500R, ssrDNA (1,706–1,711 base pairs) was amplified using WormA + WormB; difficult templates were amplified with nested PCR using Macro_18S_200F + Macro_18S_1640R. Cycling conditions for lsrDNA were as follows: denaturation for 5 min at 95 °C followed by 40 cycles of 30 s at 95 °C, 30 s at 55 °C, 2 min at 72 °C, and 7 min extension at 72 °C. Cycling conditions for ssrDNA were as follows: denaturation for 2 min at 94 °C followed by 40 cycles of 30 s at 94 °C, 30 s at 54 °C, 2 min at 72 °C, and 7 min extension at 72 °C. PCR amplicons were gel-extracted using the QIAquick Gel Extraction Kit (QIAGEN) or purified directly using the QIAquick PCR Purification Kit (QIAGEN) following the manufacturer’s instructions. Cycle-sequencing from both strands was carried out on an ABI 3730 DNA Analyzer, Big Dye version 1.1 using ABI BigDye chemistry. Contiguous sequences were assembled and edited using Sequencher version 4.6 (Genecodes Corp.), and sequence identity was checked using BLAST (http://www.ncbi.nlm.nih.gov/BLAST). New sequences have been deposited with GenBank under accessions FJ715295–FJ715334 inclusive (Table S3).

Phylogenetic Analysis. Alignments were performed in ClustalX (3) using default settings and were improved by eye in MacClade (4). Regions that could not be aligned unambiguously were excluded from the analysis. The full alignments for lsrDNA and ssrDNA gene partitions (with an indication of exclusion sets) are available upon request. MODELTEST version 3.7maxX (5) was used to select a model of evolution using the Akaike Information Criterion. Phylogenetic trees were constructed using Bayesian inference (BI) with MrBayes version 3.1 (6) and using maximum likelihood (ML) with PAUP* version 4.0b10 (7). For BI, likelihood settings were set to number of substitution types (nst) = 6, rates = invgamma, ngammacat = 4 (equivalent to the general time-reversible plus proportion of invariant sites plus gamma-distributed rate variation across sites (i.e., GTR+1+G) model of nucleotide evolution; parameters were estimated separately for each gene. Four chains (temp = 0.2) were run for 5 × 10⁶ generations and sampled every 10⁹ generations; 5 × 10⁶ generations were discarded as burn-in. ML analyses were performed using successive approximation: Model parameters were estimated based on a starting tree determined by neighbor joining. A heuristic search was performed implementing the estimated model parameters using nearest-neighbor-interchange branch swapping. Model parameters were estimated on the best tree, and a heuristic search was performed using subtree-pruning-regrafting branch swapping. After model parameters were estimated, heuristic searches using tree-bisection-reconnection (TBR) branch swapping were performed until the topology remained unchanged. In addition to posterior probability values from BI analyses, nodal support was estimated using ML bootstrapping (100 replicates) as implemented in GARLI version 0.942 (8) using default settings, except setting Genthreshfortopoterm to 10⁸ generations. Clades were considered to have high nodal support if BI posterior probability was ≥95% and ML bootstrap resampling was ≥70%.

Constraint Analysis. We performed a constraint analysis that tested whether the data support the a priori hypothesis of a single origin of the hypodermic mating syndrome. A constrained tree holding the five taxa with this feature as monophyletic was loaded as a backbone constraint before ML analysis under the same model as the unconstrained tree (Fig. S2). Log likelihood scores of constrained and unconstrained trees were used in a Shimodaira–Hasegawa test (9), as implemented in PAUP*, with 10⁵ RELL bootstrap replicates.

Ancestral State Reconstruction. We performed ancestral state reconstructions to infer the character states at the base of the genus *Macrostomum* and the base of clade 2 (both nodes are well supported in the ML and BI analyses). We used Mesquite version 2.5 (10) to estimate ancestral states of characters illustrated in Fig. 2 and listed in Table S1, under a ML continuous-time Markov model (Mk1) on the ML tree shown in Fig. S1. Ancestral states were reported as proportional likelihoods at each node for both character states. Missing data resulted in some ancestral character states being reported as equivocal (Fig. S3).

Analysis of Correlated Evolution. We used BayesDiscrete in BayesTraits (available at http://www.evolution.reading.ac.uk/BayesTraits.html) to test for correlated evolution between pairs of discrete binary character states (11). BayesTraits uses a reversible-jump Markov chain Monte Carlo (RJ MCMC) approach to search among possible models of character state evolution while sampling from a set of trees derived from a Bayesian phylogenetic analysis, thus taking phylogenetic uncertainty into account (11). The analysis is run twice for each pair of character states, once allowing for dependent (or correlated) evolution (dep), and once restricting the models to the null hypothesis of independent evolution (indep). The rationale is that under independent evolution the transition rate of character 1 from one state to the other should be independent of the state of character 2 (11). The statistical inference compares the two analyses by the harmonic means (H) of the resulting likelihoods using the test statistic 2δ(logHdep – logHindep). By convention, values for this test statistic >2 are taken as positive evidence that the dependent model is favored, and values >5 and >10 represent strong and very strong evidence, respectively (11).

We performed the analyses on a reduced taxon set containing only the genus *Macrostomum*, the main focus of our study. Specifically, we tested if the character states for copulation behavior (0, hypodermic; 1, reciprocal) correlate with those for sperm bristles (0, absent; 1, present), or female antrum morphology (0, simple; 1, thickened). (Fig. 2, Fig. S3, and Table S1 give details on character states.) Note that the character states for sperm bristles and stylet morphology (0, needle-like; 1, not needle-like) are fully congruent among the available *Macrostomum* species, so the results we present for sperm bristles also are valid for stylet morphology. Moreover, the character states for the copulation behavior and the sucking behavior (0, never observed; 1, present) are nearly congruent, but the latter character has more uncertainty about its states (Table S1); here we focus only on the
copulation behavior. Using this method we performed analysis A (sperm bristles vs. copulation behavior) and analysis B (female antrum morphology vs. copulation behavior), each with a separate run for dependent and independent models.

Based on initial ML runs, we used a RJ hyperprior with a gamma distribution (rhp gamma 0 1 0 1) for the RJ MCMC analyses (11). Next we optimized the rate deviation (ratedev) parameters to achieve acceptance rates between 20–40% (11) and settled for 0.18 and 0.12, respectively, for the dependent and independent runs of analysis A and 0.3 and 0.25, respectively, for the dependent and independent runs of analysis B. We ran the analyses with a sample of 500 best trees taken from our Bayesian phylogenetic analysis (performed as outlined above but with the reduced taxon set) for 505 × 10^6 iterations with a burn-in of 5 × 10^6 iterations and sampling period of 10^5 for the independent models and for 1,020 × 10^6 iterations with a burn-in of 2 × 10^7 iterations and sampling period of 4 × 10^6 for the dependent models (to achieve reliable convergence stability).

Notes on Species Identification, Sampling Locations, and Taxonomic Status of the Studied Specimens

We have deposited extensive digital reference material for all specimens that we have sequenced to construct the molecular phylogeny on the online Macrostomorpha Taxonomy and Phylogeny database (available at http://macrostomorpha.info), an EDIT scratchpad (12), including images, videos, and maps. Each specimen carries a unique accession number (e.g., MTP LS 200, short for Macrostomorpha Taxonomy and Phylogeny, Lukas Schärer, specimen ID 200). In the following section we give detailed notes on species identification, sampling locations, and taxonomic status of the species and specimens studied.

*Doliocmacrostomum uniporum* Luther 1947 was described from Tvärminne, Finland (13). Rieger (14) and Ax (15) list the Baltic Sea, the east and west coasts of Sweden, and the Irish Sea as the distribution. Our specimen (MTP LS 222, not documented) was taken from a laboratory culture that we established from specimens collected on March 13, 2007, at low tide on an intertidal sand flat in the Königshafen, Sylt, Germany (55°02′51.0″N, 8°25′12.5″E). It matches the descriptions by Luther (13, 16) and Rieger (14) in every detail studied. Because the sequenced specimen was not documented, we deposited an additional specimen in GenBank (GenBank EF051329). Thus, these three genera and the close proximity of the male and female genital openings is similar to *Antromacrostomum arnaitum*. Thus, these three genera may be closely related. Our species (MTP LS 309) was collected at low tide from the typical location on July 22, 2007. From two additional specimens (MTP LS 55 and MTP LS 59), collected from the same location on April 9, 2006, we obtained partial sequences of 18S, which were identical to that of the main specimen. We plan to name this genus in honor of the late Reinhard M. Rieger, and a detailed taxonomic description will be presented elsewhere.

*Macrostomum* sp. nov. 1 has been collected repeatedly by our group from a single location in Lignano, Italy (45°41′28.7″N, 13°07′54.3″E). The sample location is among coarse algae-covered gravel, which lies on a strongly anoxic base of finer sediment. The species belongs to a group of *Macrostomum* species that are not easy to distinguish based on morphology alone, and therefore we cannot exclude the possibility that it corresponds to a previously described species. However, we can state clearly that it differs from the other species with a similar morphology that we have collected in Europe. It differs from *Macrostomum pusillum* Ax 1951 by the absence of long sensory cilia and droplets in the stylet, which lies on a strongly anoxic base of finer sediment. It differs from *Macrostomum hystrix* (GenBank AF051329 from ref. 26) stemmed from a laboratory culture of *Macrostomum equicaudatum*, which has a similar morphology. Finally, *Macrostomum hystrix* Rieger 1977 was described from the west coast of the United States and the Mediterranean from sheltered beaches and shallow subtidal fine sand flats with salinity above 25% (25). It belongs to a group of *Macrostomum* species that are not easy to distinguish based on morphology alone. Our specimen (MTP LS 278) was collected on July 16, 2007, at low tide on an intertidal sand flat near Grado, Italy (45°42′51.7″N, 13°23′06.0″E), where this species was collected previously together with R. M. Rieger. The published 18S sequence of *Macrostomum hystrix* (GenBank AF051329 from ref. 26) stemmed from a laboratory culture of *Macrostomum hystrix* from a population collected by R. M. Rieger on the west coast of the United States. That sequence, however, is somewhat distinct from that of our Mediterranean form.

*Macrostomum pusillum* Ax 1951 was described from fine sands in the intertidal zone of the North Sea coast of Germany (27). Ax and Armonies (28) list the distribution as North Sea, Baltic Sea, Atlantic Coast of Norway, Mediterranean, Black Sea, south-east Canada, and Alaska. Although it belongs to a group of *Macrostomum* species that are not easy to distinguish based on morphology alone, the identity of *Macrostomum pusillum* is
thought to be clear because of its long sensory cilia and droplets in the stylet. Our specimens were collected from two localities. One specimen (MTP LS 112, not documented) was taken from a laboratory culture that we are maintaining from specimens collected on April 8, 2006, at low tide on an intertidal sand flat in Lignano, Italy (around 45°41′30″ N, 13°07′52″ E). From an additional specimen (MTP LS 53) taken from the same sample location from which we founded the laboratory culture, we obtained a partial sequence of 18S, which was identical to that of the main specimen. Another specimen (MTP LS 132) was collected on March 7, 2007, at low tide on an intertidal sand flat near Rantum, Sylt, Germany (54°50′53.8″ N, 8°17′54.4″ E). Because the stylet of this sequenced specimen was not documented in much detail, we deposited an additional specimen (MTP LS 136) collected in the same sample. The specimens from Lignano and Sylt clearly are genetically distinct but currently cannot be distinguished based on their morphology. Because the type specimen of this species is from the North Sea, the Mediterranean form probably should be renamed.

**Macrostomum balticum** Luther 1947 was described from Tvärminne, Finland (13) based on material from Tor Karling. Luther (16) lists the Baltic Sea, the west coast of Sweden, and Sylt, Germany, as the distribution. Our specimen (MTP LS 144) was collected on March 7, 2007, at low tide on an intertidal sand flat near Rantum, Sylt, Germany (54°50′54.1″ N, 8°17′55.0″ E). It matches the descriptions of Luther (13, 16) in every detail studied.

**Macrostomum spirale** Ax 1956 was discovered by Schulz (29) from a sandy mudflat in Amrum, Germany. It was supposed to be described by Meixner, but his original account was never published because of the Second World War. The species later was described formally by Ax (30) from the Etang de Canet, near Perpignan, France. Ax (15) lists the distribution as North Sea, Baltic Sea, Channel Coast of England, Mediterranean, Black Sea, and Alaska. Our specimen (MTP LS 227, not documented) was taken from a laboratory culture that we established from specimens collected on March 7, 2007, from a water-covered salt marsh near Rantum, Sylt, Germany (54°50′50″ N, 8°17′53″ E). Because MTP LS 227 was not documented, we deposited an additional specimen (MTP LS 138) taken from the same sample location from which we founded the laboratory culture. We also obtained a partial 18S sequence from one additional specimen (MTP LS 1B) collected in Bibione, Italy (45°38′02″ N, 13°04′32″ E), which is identical to that of the main specimen.

**Macrostomum longituba** Papi 1953 was described from a small ditch near the sea in the San Rossore Park near Pisa, Italy (31). Our specimen (MTP LS 274) was collected on July 15, 2007, from a small drainage ditch in an agricultural area near Bibione, Italy (45°38′50.1″ N, 13°01′10.7″ E). These ditches are close to the mouth of the Tagliamento River and thus are quite variable in salinity. Our specimen matches the original description in every detail studied, except in the exact position of the opening in the female antrum, moved *Macrostomum gieysztori* to the genus *Promacrostromum*, which is characterized by two female genital openings (36). Finally, Ferguson (37) moved the species to the new genus, *Axis*, to distinguish it from *Promacrostromum paradoxa* An-der-Lan 1939, which has a connection between the female antrum and the gut, a feature that is absent in this species. Given that our specimen clusters well within the genus *Macrostomum*, we suggest that the old name *Macrostomum gieysztori* be reinstated. Moreover, we note that the genus *Axis* has been occupied by a genus of Lepidopatra since 1821 (38). Our specimen (MTP LS 264) was collected on July 1, 2007, from a captured source of the Rio Genal in Juzcar, Andalusia, Spain (36°37′37″ N, 5°10′27″ W). Because the female antrum of the genotyped specimen was not documented well and probably was in formation, we deposited an additional specimen (MTP LS 344) collected from the same site on March 30, 2008.

**Macrostomum quiritium** Kolasa 1973 originally was described from a basin in the Poznan palm house, in Poznan, Poland (39). In that paper Kolasa attributes the name *Macrostomum quiritium* to Beklemishev (40), who, however, never named such a species. Instead Beklemishev (40) described a variety of *Macrostomum japonicum* Okugawa 1930 from an aquarium of a Russian malaria research facility, which he called “*Macrostomum japonicum* var. *quiritium* Beklemishev 1951.” However, neither *Macrostomum japonicum* (by the shape of the stylet) nor *Macrostomum japonicum* var. *quiritium* (by the arrangement of the seminal vesicles and the vesicula granulorum with respect to the stylet) matches the species described by Kolasa. *Macrostomum quiritium* has been collected repeatedly by our group from a small pond in the Tropenhauz of the Botanical Garden of the University of Basel, Switzerland (47°33′31.0″ N, 7°34′54.5″ E). This pond contains a great diversity of aquatic plants of worldwide tropical origin. It appears likely that this species was introduced with the plants. Our specimen (MTP LS 102) was collected from this pond on November 6, 2006. Our specimens match *Macrostomum quiritium* in many aspects studied, including the type of the collection site, but because of the lack of collections in the natural habitat and the resulting wide range of possible origins, our species identification should be regarded with caution.

**Macrostomum tuba** von Graff 1882 was described from a pond in the Botanical Garden of Munich, Germany (18). Moreover, there are worldwide reports of this or similar species, mostly from artificial ponds or aquaria. The different published 18S sequences of *Macrostomum tuba* (GenBank U70080 from ref. 41, called “U70081” in their paper, and D85091 from ref. 42) and *Macrostomum* sp. (GenBank L41127 from ref. 43) are all somewhat distinct from that of our specimen. No reference material for these species is available, and in some cases it is not clear where they were collected. The taxon *Macrostomum tuba* therefore is best regarded currently as an assemblage of dorsoventrally flattened fresh-water *Macrostomum* species with very long and
slender stylets about 200–300 μm in length and a worldwide distribution, and it probably also includes the species described as *Macrostromum tuba* gigas Okugawa 1930, *Macrostromum gigas* Okugawa 1930, and *Macrostromum bulbostylum* Ferguson 1939. *Macrostromum tuba* has been collected repeatedly by our group from small ponds in the Victoriahaus of the Botanical Garden of the University of Basel, Switzerland (47°33′32.7″N, 7°34′54.1″E). Our specimen (MTP LS 261) was collected at this location on June 26, 2007. Our specimen matches *Macrostromum tuba* in many aspects studied, including the type of the collection site. However, because of the lack of collections in the natural habitat, and because these ponds contain a great diversity of aquatic plants of worldwide origin, it is difficult to judge if the collected specimens match the type species.

*Macrostromum finlandense* (Ferguson 1940) was described originally as “*Macrostromum viride* Luther 1905” from freshwater in Lohja (Lojo), Southern Finland (44). It later was transferred to *Macrostromum ruebushi finlandensis* by Ferguson (45), then to *Macrostromum appendiculatum finlandensis* by Luther (pp. 11–14 in ref. 13), and finally to its current designation by Luther (pp. 72–73 in ref. 13). Moreover, the (sub)species name has been referred to variably as “*finlandensis,*” “*finlandensis,*” or “*finlandensis.*” Luther (16) lists Finland and Italy as the distribution, but other authors have reported it from Holland and Germany (46–48) and from Romania (49). Our specimen (MTP LS 91) was collected by Peter Ladurner from the Schwarzwasser near Kitzbühel, Austria (47°27′22.9″N, 12°21′58.7″E), on July 4, 2006. Because this specimen was not documented in much detail, we deposited an additional specimen (MTP LS 515) collected in the same sample. The specimens match the description by Luther (13, 16) in every detail studied.

*Macrostromum kepneri* (Ferguson and Jones 1940) was described originally by Ferguson and Jones (50) as “*Macrostromum ruebushi var. kepneri*” from brackish water in Norfolk, VA, and later was transferred to its current designation by Ferguson (37). Our specimen (MTP LS 285) was collected on July 15, 2007, from a small drainage ditch in an agricultural area near Bibione, Italy (45°38′34.5″N, 12°58′52.5″E), which is close to the Adriatic Sea and thus is quite variable in salinity. The specimen matches the original description in every detail studied. However, given the large distance between the type locality and our collection site, the species identity needs to be regarded with some caution.

*Macrostromum lignano* Ladurner, Schärer, Sulvemmoser and Rieger 2005 was described from clean intertidal sand of the northern Adriatic Sea around Lignano, Italy (45°39′16.2″N, 13°04′10.2″E). This canal is close to the mouth of the Tagliamento River and therefore is highly variable in salinity. From an additional specimen (MTP LS 1G, not documented), collected on April 1, 2005, from a nearby sample location (45°38′33.5″N, 12°58′52″E), we obtained a partial sequence of 18S, which was identical to that of the main specimen. Because the main specimen was not documented in much detail, we deposited an additional specimen (MTP LS 292), collected from the second site on July 15, 2007.

*Macrostromum mystrophorum* Meixner 1926 was described briefly by Meixner (55) from moss in a freshwater spring in the Steiermark, Austria, based primarily on the morphology of the stilet. A more detailed description was given by Papi (31) from a flooded zone near the sea in the San Rossore park near Pisa, Italy (variable salinity, ~5%). Our specimen (MTP LS 64) was collected on April 10, 2006, from a small drainage ditch in an agricultural area near Bibione, Italy (45°39′16.5″N, 13°04′11.9″E). This ditch is close to the mouth of the Tagliamento River and therefore is highly variable in salinity. At the time of collection the salinity was about 12‰. Because the main specimen was not documented in much detail, we deposited an additional specimen (MTP LS 516) collected in the same sample. Our specimens match the description by Papi (31) in every detail studied.

Because the sequenced specimen was not documented, we deposited an additional specimen (MTP LS 517) taken from the PS location.

*Macrostromum hystricatum* Ørsted 1843 sensu Luther 1905 was described by Ørsted (52) from the Baltic Sea and studied in detail by Luther (44) from samples collected in Tvärminne, Finland. It belongs to a group of *Macrostromum* species that are not easy to distinguish based on morphology alone, and as a result the species boundaries within this group often are unclear. Our molecular phylogeny, however, clearly reveals that *M. hystricatum* is not closely related to the other species in clade 1 (e.g., *M. pusillum* and *M. hystrix marinum*), despite its striking similarity in both male and female traits, which has caused even an expert like Luther to synonymize this species variably with *Macrostromum appendiculatum* (13) and *Macrostromum hystricatum* (16). However, Luther (44) gives an exquisite drawing of the stilet of his specimen (ref. 33, plate 4, fig. 1) that matches ours in every detail, and both the above names are taxonomically problematic. We therefore prefer to refer to our specimens as *Macrostromum hystricatum* Ørsted 1843 sensu Luther 1905. Convergent evolution evidently can lead to very similar outcomes and trait simplifications, and we thus argue strongly against the future use of the genera *Infra_macrostromum* and *Archnimacrostomum*, as has been stressed repeatedly (53, 54), or the unwarranted erection of new genera based on minor morphological differences (see also *Macrostromum gyczorii*). Our specimen (MTPLS 68) was collected on April 10, 2006, from a drainage canal in an agricultural area near Bibione, Italy (45°39′16.2″N, 13°04′10.2″E). This canal is close to the mouth of the Tagliamento River and therefore is highly variable in salinity. From an additional specimen (MTP LS 1G, not documented), collected on April 1, 2005, from a nearby sample location (45°38′33.5″N, 12°58′52″E), we obtained a partial sequence of 18S, which was identical to that of the main specimen. Because the main specimen was not documented in much detail, we deposited an additional specimen (MTP LS 292), collected from the second site on July 15, 2007.

44. Luther A (1905) Zur Kenntnis der Gattung Macrostoma. Festschrift für Palmén, Helsingfors, 5:1–61 (64 plates).
Fig. S1. Molecular phylogeny of 16 *Macrostomum* and four outgroup species. ML tree based on combined partial lsrDNA and complete ssrDNA sequences (for a total of ∼2,850 base pairs) from 16 *Macrostomum* and four outgroup species, covering members of all three families in the order Macrostomida (Platyhelminthes: Macrostomorpha). Values above branches are Bayesian posterior probabilities, and values below branches are ML bootstrap values. The topologies of trees derived from Bayesian and ML analyses are in broad agreement. Final ML model settings were as follows: nucleotide frequencies [π(A) = 0.2334; π(C) = 0.2233; π(G) = 0.2894; π(T) = 0.2538]; rate matrix [(A,C) = 0.6007; (A,G) = 3.6492; (A,T) = 2.1686; (C,G) = 0.4266; (C,T) = 7.9278; (G,T) = 1.0000]; invariable sites = 0.5247; γ shape parameter = 0.4968; log likelihood = −11679.817. The accession code identifies the morphological documentation of each sequenced specimen, which we have deposited as digital reference material at http://macrostomorpha.info (GenBank accession numbers are given in Table S3). Details on phylogenetic reconstruction are given in SI Materials and Methods.
Fig. S2. Unconstrained and constrained tree used for the Shimodaira–Hasegawa test. The test shows that the unconstrained tree (A) fits the data significantly better ($\Delta -\ln$ likelihood = 66.0; $P < 0.001$) than the constrained tree (B). Details of analysis are given in SI Materials and Methods.
Fig. S3. ML ancestral state reconstruction of character states. The small pie charts indicate the likelihoods of the black vs. white character states at each node, and gray nodes indicate equivocal character states. (A) Sperm bristles. (B) Stylet morphology. (C) Copulation behavior. (D) Sucking behavior. (E) Female antrum morphology. (F) Phylogeny with ancestral nodes numbered. (G) Probabilities (P values) for the black character state at each node. Details on analysis are provided in SI Materials and Methods.
Fig. S4. Posterior distributions of the rate parameters of the models of character state evolution. Graphs are based on $2.5 \times 10^5$ observations drawn from $10^9$ iterations of the Markov chain. (A) Correlated evolution between the sperm/stylet morphology and the copulation behavior. (B) Correlated evolution between the female antrum morphology and the copulation behavior. The percentages indicate the proportion of the rate parameter estimates that are zero, with higher values indicating less likely transitions. The panels are arranged so that vertical pairs correspond to rates that would be expected to be the same if the independent model of character evolution were true, which is never the case in our data. For example, in the first pair of analysis A, the transition from hypodermic to reciprocal mating is more likely when bristles are present than when they are not (i.e., the rate parameter estimate is zero in 14% and 59.9% of the iterations, respectively). Details on analysis are provided in SI Materials and Methods.
<table>
<thead>
<tr>
<th>Species</th>
<th>Bristles</th>
<th>Stylet</th>
<th>Copulation</th>
<th>Sucking</th>
<th>Antrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolichomacrostomum uniporum</td>
<td>0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;j&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;j&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;j&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;j&lt;/sup&gt;</td>
</tr>
<tr>
<td>Microstomum papillosum</td>
<td>0&lt;sup&gt;b&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;n&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;n&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;n&lt;/sup&gt;</td>
</tr>
<tr>
<td>Bradynectes sterreri</td>
<td>0&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1&lt;sup&gt;h&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;o&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;o&lt;/sup&gt;</td>
<td>ND&lt;sup&gt;o&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gen. nov. 1, sp. nov. 1</td>
<td>0</td>
<td>1&lt;sup&gt;l&lt;/sup&gt;</td>
<td>1</td>
<td>0</td>
<td>ND&lt;sup&gt;n&lt;/sup&gt;</td>
</tr>
<tr>
<td>Macrostromum sp. nov. 1</td>
<td>0</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum hystrichinum marinum</td>
<td>0</td>
<td>0</td>
<td>0&lt;sup&gt;o&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum pusillum (Lignano)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum pusillum (Sytl)</td>
<td>0&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum balticum</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum spirale</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum longituba</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum clavituba</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum gieysztori</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0&lt;sup&gt;p&lt;/sup&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum quiritium</td>
<td>1</td>
<td>1</td>
<td>ND</td>
<td>ND</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum tuba</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1&lt;sup&gt;q&lt;/sup&gt;</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum finlandense</td>
<td>1&lt;sup&gt;s&lt;/sup&gt;</td>
<td>1&lt;sup&gt;r&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum kepneri</td>
<td>1</td>
<td>1&lt;sup&gt;s&lt;/sup&gt;</td>
<td>ND</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum lignano</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Macrostromum hystrix</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Macrostromum mystrophorum</td>
<td>1</td>
<td>1</td>
<td>ND</td>
<td>ND</td>
<td>1</td>
</tr>
</tbody>
</table>

Sperm bristles (0, absent; 1, present), stylet morphology (0, needle-like; 1, not needle-like), copulation behavior (0, hypodermic; 1, reciprocal), sucking behavior (0, never observed; 1, present); female antrum morphology (0, simple; 1, thickened); ND, no data.

<sup>a</sup>ND for O. uniporum; character state based on data for other Dolichomacrostomidae, namely Paromalostomum fusculum (1, 2) and P. atratum (3).

<sup>b</sup>ND for M. papillosum; character state based on data for other Microstomidae, namely M. spiculifer (4) and Microstomidae (5).

<sup>c</sup>Sopott-Ehlers and Ehlers (2) state, “[T]he two lateral ledges found in spermatozoa of B. sterreri are discussed to correspond to the pair of ‘lateral bristles’ known from Macrostomum species,” but even if they were homologous, these structures do not protrude outside of the sperm and they cannot have a sperm anchoring function.

<sup>d</sup>Sperm ultrastructure suggests putative rudimentary bristles (6).

<sup>e</sup>Bristles are small but clearly visible.

<sup>f</sup>Homology is unclear; many Dolichomacrostomidae have two stylets, a penis stylet and a gland stylet, the latter of which can be needle-like.

<sup>g</sup>Homology is somewhat unclear (8).

<sup>h</sup>Homology is unclear; the Dolichomacrostomidae have a common (male and female) genital opening and atrium genitale (7).

<sup>i</sup>Homology is unclear; the Dolichomacrostomidae have a common (male and female) genital opening and atrium genitale (7).

<sup>j</sup>No data on mating behavior exist for any Dolichomacrostomidae.

<sup>k</sup>No data on mating behavior exist for any Microstomidae.

<sup>l</sup>No data on mating behavior exist for any Bradynectes species.

<sup>m</sup>No data on mating behavior exist for any Bradynectes species.

<sup>n</sup>Not observed directly but inferred from the presence of sperm in the parenchyma.

<sup>o</sup>Homology is unclear; the Dolichomacrostomidae have a common (male and female) genital opening and atrium genitale (7).

<sup>p</sup>Homology is somewhat unclear (8).

<sup>q</sup>Lacks a vagina and female antrum (8, 9); mechanism of sperm transfer is unclear.

<sup>r</sup>Structure of female antrum is unclear.

<sup>s</sup>Homology is somewhat unclear (8).

<sup>t</sup>Sperm ultrastructure suggests putative rudimentary bristles (6).

<sup>u</sup>No data on mating behavior exist for any Microstomidae.

<sup>v</sup>No data on mating behavior exist for any Microstomidae.

<sup>w</sup>No data on mating behavior exist for any Bradynectes species.

<sup>x</sup>Not observed directly but inferred from the presence of sperm in the parenchyma.

<sup>y</sup>Homology is unclear; the Dolichomacrostomidae have a common (male and female) genital opening and atrium genitale (7).

<sup>z</sup>Lacks a vagina and female antrum (8, 9); mechanism of sperm transfer is unclear.

<sup>AA</sup>Structure of female antrum is unclear.
Table S2. Primers used for amplification and sequencing

<table>
<thead>
<tr>
<th>lsrDNA primers</th>
<th>PCR and sequencing primers</th>
<th>F</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZX-1*</td>
<td></td>
<td>F</td>
<td>ACCCGCTGAATTTAAGCATAT</td>
</tr>
<tr>
<td>1200R</td>
<td></td>
<td>R</td>
<td>GCATATGGTCACCACCTTTCGG</td>
</tr>
<tr>
<td>1500R</td>
<td></td>
<td>R</td>
<td>GCTATCCTGAGGGAAAACCTCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional sequencing primers</th>
<th>F</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>300F</td>
<td>F</td>
<td>CAATACCGTGAGGGAAAGTGT</td>
</tr>
<tr>
<td>ECD2</td>
<td>R</td>
<td>CTGGTGCCGTTTCAAAGACCGGG</td>
</tr>
<tr>
<td>1090F</td>
<td>F</td>
<td>TGAACACGGACCAAG</td>
</tr>
<tr>
<td>ssrDNA primers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PCR and sequencing primers</th>
<th>F</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>WormA</td>
<td>F</td>
<td>GCGAATGGCTCATAAATCAG</td>
</tr>
<tr>
<td>WormB</td>
<td>R</td>
<td>CTGTGAACGACTTTTAACCC</td>
</tr>
<tr>
<td>Macro_18S_200F</td>
<td>F</td>
<td>GCCGATTATTAGATCAAAACCA</td>
</tr>
<tr>
<td>Macro_18S_1640R</td>
<td>R</td>
<td>GCAAAGCCCAGATCCCTGTC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional sequencing primers</th>
<th>F</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>300F</td>
<td>F</td>
<td>AGGGTTCAGTTCCGGAG</td>
</tr>
<tr>
<td>600R</td>
<td>R</td>
<td>ACCGCAGGGGCAGCC</td>
</tr>
<tr>
<td>1270F</td>
<td>F</td>
<td>ACTAAAGGAATTGACG</td>
</tr>
<tr>
<td>1270R</td>
<td>R</td>
<td>CCAGTAAATCTCCTTAAGT</td>
</tr>
<tr>
<td>1200F</td>
<td>F</td>
<td>CAGGTCCTGAGTCCC</td>
</tr>
</tbody>
</table>

All primers are 5′–3′. F, forward; R, reverse.

*Modified from the original ZX-1 (1): ACCCGCTGAAAYTAAGCATAT; Y replaced with T.


Table S3. GenBank accession numbers for each taxon and gene

<table>
<thead>
<tr>
<th>Taxon</th>
<th>GenBank accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolichomacrostomum uniporum (MTP LS 222)</td>
<td>FJ715315</td>
</tr>
<tr>
<td>Microstomum papillosum (MTP LS 146)</td>
<td>FJ715316</td>
</tr>
<tr>
<td>Bradyoctes sterrei (MTP LS 180)</td>
<td>FJ715318</td>
</tr>
<tr>
<td>Gen. nov. 1, sp. nov. 1 (MTP LS 309)</td>
<td>FJ715317</td>
</tr>
<tr>
<td>Macrostomum sp. nov. 1 (MTP LS 302)</td>
<td>FJ715332</td>
</tr>
<tr>
<td>Macrostomum hystricinum marinum (MTP LS 278)</td>
<td>FJ715331</td>
</tr>
<tr>
<td>Macrostomum pusillum (Lignano) (MTP LS 112)</td>
<td>FJ715333</td>
</tr>
<tr>
<td>Macrostomum pusillum (Sylt) (MTP LS 132)</td>
<td>FJ715334</td>
</tr>
<tr>
<td>Macrostomum balticum (MTP LS 144)</td>
<td>FJ715330</td>
</tr>
<tr>
<td>Macrostomum spirale (MTP LS 227)</td>
<td>FJ715328</td>
</tr>
<tr>
<td>Macrostomum longituba (MTP LS 274)</td>
<td>FJ715329</td>
</tr>
<tr>
<td>Macrostomum clavituba (MTP LS 301)</td>
<td>FJ715324</td>
</tr>
<tr>
<td>Macrostomum gieysztori (MTP LS 264)</td>
<td>FJ715321</td>
</tr>
<tr>
<td>Macrostomum quirtitum (MTP LS 102)</td>
<td>FJ715319</td>
</tr>
<tr>
<td>Macrostomum tuba (MTP LS 261)</td>
<td>FJ715320</td>
</tr>
<tr>
<td>Macrostomum finlandense (MTP LS 91)</td>
<td>FJ715322</td>
</tr>
<tr>
<td>Macrostomum kepneri (MTP LS 285)</td>
<td>FJ715327</td>
</tr>
<tr>
<td>Macrostomum lignano (MTP LS 244)</td>
<td>FJ715326</td>
</tr>
<tr>
<td>Macrostomum hystrix (MTP LS 68)</td>
<td>FJ715323</td>
</tr>
<tr>
<td>Macrostomum mystrophorum (MTP LS 64)</td>
<td>FJ715325</td>
</tr>
</tbody>
</table>

Species are listed in the order in which they appear in the tree. The MTP accession code identifies the morphological documentation of each sequenced specimen (http://macrostomorpha.info). All sequences are new for this study.
**Movie S1.** A copulating pair of the flatworm *Macrostomum lignano*. Note that one individual performs the postcopulatory sucking behavior, after which a bundle of sperm shafts can be seen sticking out of the female genital opening.

**Movie S2.** A single sperm of the flatworm *Macrostomum lignano*. Note the highly motile feeler and shaft, which allow the sperm to perform complex movements.
Movie S3. Anchored received sperm in a live specimen of the flatworm *Macrostomum lignano*. Note the thickened epithelium of the female antrum (i.e., the translucent rim around the sperm) and the polarized nature of the sperm, most of which are anchored in the cellular valve (i.e., the part of the antrum epithelium closest to the forming oocyte, which is the dark area on the right).

Movie S3

Movie S4. Detail of the anchored sperm of the specimen in Movie S3. Note the undulating sperm feelers, which are deeply embedded in the cellular valve.

Movie S4
Movie S5. Focusing through the parenchyma of a live specimen of the flatworm *Macrostomum hystrix*. Note the abundant hypodermically inseminated and highly motile sperm.