





OPEN

Successful mating and hybridisation in two closely related flatworm species despite significant differences in reproductive morphology and behaviour

Pragya Singh¹  , Daniel N. Ballmer^{1,2}, Max Laubscher¹ & Lukas Schärer¹

Reproductive traits are some of the fastest diverging characters and can serve as reproductive barriers. The free-living flatworm *Macrostomum lignano*, and its congener *M. janickei* are closely related, but differ substantially in their male intromittent organ (stylet) morphology. Here, we examine whether these morphological differences are accompanied by differences in behavioural traits, and whether these could represent barriers to successful mating and hybridization between the two species. Our data shows that the two species differ in many aspects of their mating behaviour. Despite these differences, the species mate readily with each other in heterospecific pairings. Although both species have similar fecundity in conspecific pairings, the heterospecific pairings revealed clear postmating barriers, as few heterospecific pairings produced F1 hybrids. These hybrids had a stylet morphology that was intermediate between that of the parental species, and they were fertile. Finally, using a mate choice experiment, we show that the nearly two-fold higher mating rate of *M. lignano* caused it to mate more with conspecifics, leading to assortative mating, while *M. janickei* ended up mating more with heterospecifics. Thus, while the two species can hybridize, the mating rate differences could possibly lead to higher fitness costs for *M. janickei* compared to *M. lignano*.

The biological species concept defines species as groups of individuals that interbreed in nature to produce viable and fertile offspring^{1,2}. They are usually isolated from interbreeding with other species by reproductive barriers, though in some cases they remain capable of producing hybrid offspring with closely related species. Accordingly, an important step for the origin and maintenance of species is the evolution of reproductive barriers, which are usually split into prezygotic and postzygotic barriers^{3–6}. While prezygotic barriers involve the prevention of zygote formation, postzygotic barriers either lead to zygote mortality, or to inviable or sterile hybrid offspring that are unable to pass on their genes. Moreover, prezygotic barriers can be ecological, temporal, behavioural, mechanical or gametic, and can be further subdivided into premating barriers and postmating-prezygotic barriers. Premating barriers act to prevent the occurrence of heterospecific matings. For example, if a species has a mating preference for conspecific partners over heterospecifics, this mating preference can lead to assortative mating between conspecifics and thereby function as a premating barrier^{7–9}. Postmating-prezygotic barriers often involve conspecific sperm precedence due to postcopulatory processes, such as sperm competition and cryptic female choice, or they can result from an incompatibility of female reproductive organs with heterospecific male ejaculate^{10–15}.

Species in the early stages of divergence will often not have complete reproductive isolation between them, but as they diverge in their traits, reproductive barriers usually accumulate over time, as these divergent traits can

¹Department of Environmental Sciences, Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland. ²Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. ✉email: pragya.singh42019@gmail.com

function as additional barriers. Reproductive traits may diverge particularly quickly, since they are the primary targets of sexual selection, often leading to rapid accumulation of phenotypic differences between species^{16–19}. Therefore, sexual selection can play an important role in evolutionary diversification, reproductive isolation and speciation^{20,21}, but see²². This is supported by the fact that reproductive traits, such as mating behaviour and genital morphology, have been shown to diversify faster than other traits^{23–27} and can differ markedly even between recently diverged species^{25,27–30}, and sometimes even between populations of the same species^{26,31,32}. Moreover, some studies have shown that mating behaviour might evolve even more quickly than genital morphology²⁶. Thus, a rapidly evolving reproductive trait like reproductive behaviour can represent a pre-mating barrier by being involved in mate recognition and assortative mating^{32,33}, while a difference in genital morphology can prevent successful mating and thus represent a mechanical barrier^{34,35}.

In recently diverged species that occur in sympatry, selection may occur to reduce the likelihood of heterospecific reproductive interactions, whenever such interactions lower individual fitness (either directly or via low fitness hybrids). This selection can cause greater divergence in reproductive traits, leading to reproductive character displacement^{36–40} and reinforcement of reproductive isolation. An interesting question that arises then is whether differences in reproductive traits correlate in recently diverged species. For instance, do differences in reproductive morphology go along with differences in reproductive behaviour⁴¹? And are these differences sufficiently large to function as prezygotic reproductive barriers, leading to reproductive isolation⁴²? Under a scenario of reinforcement in sympatry, we might expect that divergent reproductive traits will serve as effective reproductive barriers (though not all sympatric species will necessarily be completely reproductively isolated). In contrast, species that have speciated in allopatry may lack (complete) reproductive isolation due to incomplete pre- or postzygotic barriers, despite having diverged in their reproductive traits. Secondary contact between such previously allopatric species may then result in the production of viable and potentially even fertile hybrid offspring.

Even in the absence of successful hybridization during secondary contact, both heterospecific mating attempts and actual heterospecific matings can result in wastage of energy, resources, time and/or gametes. This can lead to reproductive interference, which is defined as heterospecific reproductive activities that reduce the fitness of at least one of the species involved¹⁶. Interestingly, reproductive interference may be asymmetric, in that the fitness of one species is affected to a greater extent than that of the other, and can have effects ranging from reproductive character displacement to species exclusion^{16,43–45}.

In our study, we investigated reproductive barriers and reproductive interference in two species of the free-living flatworm genus *Macrostomum*, namely *M. lignano*, an established model for studying sexual reproduction in hermaphrodites⁴⁶, and the recently described *M. janickei*, the most closely related congener known²⁸. While *M. lignano* has previously been collected from the Northern Adriatic Sea (Italy) and the Aegean Sea (Greece), *M. janickei* has to date only been collected from the Gulf of Lion (France), so the detailed geographic distribution of both species is currently poorly known^{28,47,48}. Specifically, we examined if differences in the stylet morphology between these species correlated with differences in their mating behaviour and if they had similar fecundity. Furthermore, we investigated the potential for hybridization between the two species, and tested whether the resulting hybrids were fertile. Next, using geometric morphometrics we compared the stylet morphology of the parental species and the hybrids. Finally, we performed a mate choice experiment to test if individuals preferentially mated with conspecifics over heterospecifics, since this form of assortative mating could serve as a pre-mating barrier between these two closely related species in a putative zone of sympatry.

Materials and methods

Study organisms. *Macrostomum lignano* Ladurner, Schärer, Salvenmoser and Rieger 2005 and *M. janickei* Schärer 2019 are free-living flatworm species (Macrostomorpha, Platyhelminthes) found in the upper intertidal meiofauna of the Mediterranean Sea^{28,46–48}. Despite being very closely related sister species²⁸, the morphology of their stylet is clearly distinct (see Fig. 4 and results). *M. lignano* has a stylet that is "slightly curved, its distal opening [having a] slight asymmetric thickening"⁴⁶, while *M. janickei* has a more complex stylet that is a "long and a gradually narrowing funnel that includes first a slight turn (of ~40°) and then a sharp turn (of >90°) towards the distal end [...], giving the stylet tip a hook-like appearance."²⁸

Previous studies have shown that *M. lignano* is an outcrossing, reciprocally copulating species with frequent mating (on average about 6 copulations per hour)⁴⁹. Specifically, reciprocal copulation consists of both partners mating in the male and female role simultaneously, with reciprocal insertion of the stylet into the female antrum (the sperm-receiving organ) of the partner, and transfer of ejaculate consisting of both sperm and seminal fluids. Copulation is then often followed by a facultative postcopulatory suck behaviour^{49–51}, during which the worm bends onto itself and places its pharynx over its own female genital opening, while appearing to suck. This behaviour is thought to represent a female resistance trait that has evolved due to sexual conflict over the fate of received ejaculate. Specifically, it is likely aimed at removing ejaculate components from the antrum, and sperm is often seen sticking out of the antrum after a suck^{49–52}.

The individuals of *M. lignano* used in this experiment were either from the outbred LS1 culture⁵² or from the transgenic outbred BAS1 culture, which was created by backcrossing the GFP-expressing inbred HUB1 line^{53–55} onto the LS1 culture⁵⁶, subsequently cleaned from a karyotype polymorphism that segregates in HUB1^{47,48}, and finally bred to be homozygous GFP-positive⁵⁷. The LS1 culture is a genetically outbred metapopulation, established from worms collected from a site in Bibione and a site on Isola di Martignano, Northern Adriatic Sea, Italy⁵². The *M. janickei* worms used were from a culture that was established using individuals collected from Palavas-les-Flots, Gulf of Lion, near Montpellier, France^{28,47,48}. Both species are kept in mass cultures in the laboratory at 20 °C in glass Petri dishes containing either f/2 medium⁵⁸ or 32‰ artificial sea water (ASW) and fed with the diatom algae *Nitzschia curvilineata*.

Reproductive behaviour and hybridization. On day 1, for each species, we distributed 240 adult worms over four petri dishes with algae and ASW (using the transgenic BAS1 culture for *M. lignano*). On day 4, we removed the adults, such that the eggs were laid over a 3-day period, and the age of the resulting hatchlings did not differ by more than 3 days. On day 9 (i.e. well before the worms reach sexual maturity), we isolated these hatchlings in 24-well tissue culture plates (TPP, Switzerland) in 1 ml of ASW with ad libitum algae. Starting on day 34 and spread over 3 subsequent days, we then examined the mating behaviour by pairing these previously isolated and by then adult worms (as judged by their visible testes and ovaries) in one of three pairing types, namely *M. lignano* pairs (*M. lignano* x *M. lignano*, n = 57), *M. janickei* pairs (*M. janickei* x *M. janickei*, n = 57), or heterospecific pairs (*M. lignano* x *M. janickei*, n = 57).

Each observation chamber⁴⁹ was assembled by placing 9 mating pairs (3 pairs of each pairing type) in drops of 3 μ l of ASW each between two siliconized microscope slides separated by 257 μ m, for a total of 19 observation chambers (i.e. 7, 4, and 8 chambers on the three consecutive days, respectively). The observation chambers were filmed under transmitted light for 2 h at 1 frame s⁻¹ with digital video cameras (DFK 41AF02 or DFK 31BF03, The Imaging Source) in QuickTime format using BTV Pro 6.0b7 (<https://www.bensoftware.com/>), and the resulting movies were scored manually frame-by-frame using QuickTime player. We used two different movie setups for filming the mating and they differed slightly in the cameras and light sources used.

After the two-hour mating period, we isolated both individuals of the heterospecific pairs, and one randomly chosen individual each of the *M. lignano* and *M. janickei* pairs, respectively, in 24-well plates and subsequently transferred them weekly to new plates. To obtain an estimate of the (female) fecundity resulting from these pairings the offspring production of these maternal individuals was followed and counted for 14 days (since worms eventually run out of stored sperm)⁵⁹. For each heterospecific pair, the number of (hybrid F1) offspring produced was averaged over both maternal individuals. And by confirming that all maternal offspring of the GFP-negative *M. janickei* were GFP-positive, we could ascertain that the GFP-positive BAS1 *M. lignano* had indeed sired these F1 hybrids. Moreover, previous experiments had shown that neither species self-fertilizes over a comparable observation period^{60,61}, thus any offspring produced in the heterospecific pairs must have resulted from outcrossing with the partners.

For each mating pair, we scored the movie up to the fifth copulation and observed the following copulation traits: copulation latency (i.e. time to first copulation), copulation duration, copulation interval, time until suck (after copulation), suck duration, and the number of sucks, while being blind with respect to both the pairing type and the species identity of individuals in the heterospecific pairs (note that the GFP-status of a worm cannot be determined under normal transmitted light). The decision to observe the behaviour up to and including the fifth copulation was made *a priori*⁵², and was motivated by our desire to get accurate estimates for each behaviour, by averaging all traits (except copulation latency) over this period for each pair and to keep the total observation time manageable. Note that there was no clear change in the copulation duration over the five copulations (data not shown). The copulation behaviour was defined as in 49, and the copulatory duration was measured starting from the frame when the pair was first tightly interlinked (like two small interlocking G's) with the tail plates in close ventral contact, to the frame where their tail plates were no longer attached to each other. We scored a behaviour as a copulation only if the pair was in this interlinked position for at least 5 s. The copulation interval was measured as the duration between the end of a copulation to the start of the next copulation. The time until suck was measured (for sucks that followed a copulation, observed up to the fifth copulation) as the time elapsed between the end of the copulation preceding the suck and the start of the suck in question. The suck duration was measured from the frame where the pharynx was placed on the female genital opening, up to the frame where the pharynx disengaged. The number of sucks was measured as the number of sucks observed up to the fifth copulation. The copulation durations, copulation intervals, times until suck, and suck durations were each averaged over all occurrences in a replicate.

The final sample sizes varied for the different behavioural traits, depending on how many replicates exhibited the trait of interest. We, respectively, excluded 3, 7 and 2 replicates of the *M. lignano* pairs, heterospecific pairs and *M. janickei* pairs from all analyses, since these replicates showed no copulations. In addition, three replicates of *M. janickei* had only one copulation, so we could not calculate the copulation interval for these pairs. Moreover, in some replicates there were no sucks, which reduced our sample size for the time until suck and the suck duration. The suck is considered a postcopulatory behaviour, and we therefore might not expect an individual to exhibit the postcopulatory behaviour unless it copulates. Thus, to examine if the number of sucks differed between the pairing types, we considered only the subset of replicates in which we observed at least five copulations. Additionally, for offspring number we lost two replicates each for the *M. lignano* and *M. janickei* pairs. The final sample sizes are given in the respective figures.

Hybrid fertility. We assessed the fertility of the F1 hybrid offspring produced from the above experiment on reproductive behaviour and hybridization, by pairing for 7 days a subset of the virgin hybrids with, respectively, virgin adult *M. lignano* (n = 24) or virgin adult *M. janickei* (n = 24) partners (grown up under identical conditions as the parents, but using the wildtype LS1 culture for *M. lignano*) and then isolating both the hybrids and their partners for 14 days. We counted the offspring number produced both during the pairing period, and the isolation period. While it is clear who the maternal and paternal parent are for offspring produced in isolation, we cannot distinguish whether the F1 hybrid or its partner was the maternal or paternal parent for the offspring produced during the pairing period. But by confirming that at least some of the F2 offspring from the crosses between the GFP-heterozygote F1 hybrids and the GFP-negative parents were GFP-positive, we could ascertain that we were indeed seeing successful backcrosses. We did not statistically analyse if offspring number differed depending on which parental species the hybrid was backcrossed onto, as the hybrids used were not statistically

independent (e.g. some of them were siblings). Thus, we only descriptively examined the offspring number produced from the backcrossing.

Moreover, seventeen days after backcrossing onto the parental species, the F1 hybrids were also paired amongst themselves ($n = 25$ F1 x F1 pairs, one of these pairs had not been used in backcrossing experiment) for 7 days, and then also isolated for 14 days. Similar to above, we counted the offspring number produced both during the pairing period, and the isolation period. Note that while there is a small possibility of some sperm being carried over from the previous backcrossing with the parental species, since, although fecundity drops quite rapidly after isolation, sperm has been found to be stored for at least 20 days after mating⁵⁹, we can be more certain that the hatchling produced during isolation are from hybrid matings, since this followed a 7-day F1 x F1 pairing period during which previously received sperm is likely to have been displaced⁵⁴.

Hybrid and parental stylet morphology. To investigate the stylet morphology of the F1 hybrids, we compared the stylets of isolated virgin hybrids ($n = 29$; measured before the backcrossing experiment), to those of isolated *M. lignano* ($n = 25$, data from 62) and *M. janickei* ($n = 18$, data from 60), using a geometric morphometrics landmark-based method⁶³. This method permits measuring variation in stylet size and shape, using landmark and semi-landmark coordinates. We can then statistically analyse the information obtained from the landmark configurations to quantify morphological differences between these structures. For this, landmarks are placed on homologous anatomically recognizable points across all morphological structures, while semi-landmarks are used to capture information along a curvature where points might not be easily repeatable or identifiable.

Briefly, before measuring the stylet, the F1 hybrid worms were relaxed using a solution of $MgCl_2$ and ASW, and dorsoventrally squeezed between a glass slide and a haemocytometer cover glass using standardised spacers (40 μm). Stylet images were then obtained at 400 \times magnification (Fig. 4a–c), with a DM 2500 microscope (Leica Microsystems, Heerbrugg, Switzerland) using a digital camera (DFK41BF02, The Imaging Source, Bremen, Germany) connected to a computer running BTV Pro 6.0b7 (Ben Software). A similar procedure had been followed for the parental species in their respective studies too. For geometric morphometrics, we placed a total of 60 landmarks on each stylet, two fixed landmarks each on the tip and base of the stylet and 28 equally spaced sliding semi-landmarks each along the two curved sides of the stylet between the base and the tip (Fig. 4d–f), using tpsDig 2.31 (<https://life.bio.sunysb.edu/morph/>), while being blind to the identity of the individual. Note that this landmark placement differs somewhat from that used earlier in *M. lignano*⁶⁴ on account of the different morphology of the *M. janickei* stylet. Specifically, since landmarks should represent homologous points on a morphological structure, we here defined only four fixed landmarks that could be recognised in the F1 hybrids and both parental species (compared to six in *M. lignano* earlier), while more sliding semi-landmarks were used here to approximate the considerably more complex shape of the *M. janickei* stylet (i.e. 56 semi-landmarks now vs. 18 in *M. lignano* earlier). We always placed landmarks 1–30 on the stylet side that was further from the seminal vesicle (the sperm storage organ located near the stylet), while landmarks 31–60 were placed on the stylet side that was closer to the seminal vesicle (see Fig. 4d–f). Also, to ensure that the orientation of the seminal vesicle and stylet with respect to the viewer was similar across all images, we mirrored the images for some specimens (since the stylet can be imaged from both ventral or dorsal). We used tpsRelw 1.70 (<https://life.bio.sunysb.edu/morph/>) to analyse the resulting landmark configurations and extract the centroid size (an estimate of the size of the landmark configuration that can serve as a measure of the stylet size) and the relative warp scores (which decompose the total shape variation into major axes of shape variation). Our analysis yielded 71 relative warp scores, of which the first three relative warp scores explained 88% of all variation in stylet shape. For our statistical analysis, we here only focus on the first relative warp score (RWS1), as it explained 64% of the shape variation and captured the most drastic change in the stylet shape, including the extent of the stylet tip curvature (Fig. 4g–i).

Mate preference experiment. We assessed the mate preferences of *M. lignano* (BAS1) and *M. janickei* by joining two individuals of each species in 3 μl drops of ASW (for a total of 4 individuals per drop). In each of the four drops per observation chamber, the individuals of either one or the other species were dyed to permit distinguishing the species visually in the movies (i.e. *M. lignano* or *M. janickei* were dyed in two drops each per mating chamber). We dyed the worms by exposing them to a solution of the food colour Patent Blue V (Werner Schweizer AG, Switzerland, at 0.25 mg/ml of 32‰ ASW) for 24 h. Patent Blue V does not affect the mating rate of *M. lignano*⁵², or of *M. janickei*, as the mating rate of dyed and undyed worms was similar (see Supplementary Figure S2).

In total, we constructed 17 observation chambers and filmed them under transmitted light for 2 h at 1 frame s^{-1} (as outlined above), and the resulting movies were scored manually frame-by-frame using QuickTime player, while being blind to which species was dyed. For each drop, we determined the copulation type of the first copulation, i.e. conspecific *M. lignano*, conspecific *M. janickei* or heterospecific (*M. lignano* x *M. janickei*), and we also estimated the copulation frequencies of the three copulation types over the entire 2 h period. Out of the total 68 filmed drops we had to exclude 9 drops, 5 of which had an injured worm and 4 of which (one entire observation chamber) had dim lighting that made it difficult to distinguish the dyed worms. Thus, our final sample size was 59 drops.

Statistical analyses. In the experiment examining reproductive behaviour and hybridization, we constructed one-way ANOVAs with the pairing type (*M. lignano* pairs, heterospecific pairs, and *M. janickei* pairs) as the independent fixed factor, and using copulation latency, average copulation duration, average copulation interval, average time until suck, and average suck duration as the dependent variables, followed by post-hoc comparisons between the pairing types using Tukey's honest significant difference (HSD) tests. Note that all

conclusions remained unchanged if the two movie setups were included as a factor (data not shown). Data was visually checked for normality and homoscedasticity and log-transformed for all the above variables. For average time until suck, however, we added 1 to each data point before log-transformation, to avoid infinite values, since some sucks began immediately after copulation, leading to zero values. For the number of sucks and the offspring number we used Kruskal–Wallis tests (since these data could not be appropriately transformed to fulfil the assumptions for parametric tests), followed by post-hoc tests using Mann–Whitney–Wilcoxon tests with Bonferroni correction. Moreover, for all behaviours we calculated the coefficient of variation (CV) to evaluate how stereotypic the behaviour is for each pairing type. For all behaviours (except for the number of sucks), we calculated the CV for log-transformed data using the formula $CV = 100 \times \sqrt{e^{\text{standard deviation}^2} - 1}$ ⁶⁵, while for number of sucks we calculated the CV for raw data using $CV = \frac{\text{standard deviation}}{\text{mean}} \times 100$.

While studying the hybrid and parental stylet morphology, we constructed one-way ANOVAs with the types of worm (*M. lignano*, F1 hybrid, or *M. janickei*) as the independent fixed factor, and the centroid size and RWS1 as the dependent variables, followed by post-hoc comparisons using Tukey's HSD. Note that these analyses need to be interpreted with some caution, since the three groups we compared here were not grown and imaged as part of the same experiment (though using the same methodology).

In the mate preference experiment, three different copulation types could occur (i.e. *M. lignano* conspecific, heterospecific, and *M. janickei* conspecific), and to generate a null hypothesis of the expected proportions of each copulation type, we initially assumed random mating and hence no mating preference for either conspecific or heterospecific individuals in either species. Thus, under these assumptions the null hypothesis for the expected proportions of drops having as the first copulation these different copulation types was: *M. lignano* conspecific:heterospecific:*M. janickei* conspecific = 0.25:0.50:0.25. For each copulation type, we then determined the observed proportion of drops in which it was the first copulation, and examined if these proportions differed significantly from this null hypothesis, using a Chi-square goodness-of-fit test.

Next, we looked at the observed proportion of the three copulation types within each drop and across all drops, and as the null hypothesis we again used the same expected proportions as above. To test if the observed proportion of the three copulation types differed from this null hypothesis, we used repeated G-tests of goodness-of-fit⁶⁶, an approach that involves sequential tests of up to four different hypotheses, which, depending on the obtained results, will not all necessarily be carried out. The first hypothesis tests if the observed proportions within each drop fit the expectations. The second hypothesis examines if the relative observed proportions are the same across all drops by calculating a heterogeneity value. The third hypothesis examines if the observed proportion matches the expectation when the data is pooled across all drops. And finally, the fourth hypothesis examines if overall, the data from the individual drops fit our expectations using the sum of individual G-values for each replicate (obtained from testing the first hypothesis). Following this approach, we first calculated a G-test goodness-of-fit (with Bonferroni correction) for each drop. Second, this was followed by a G-test of independence on the data to obtain a 'heterogeneity G-value', which permits to evaluate if the drops differ significantly from each other. Since, this test revealed significant heterogeneity between the drops (see results), we did not pool the data or proceed with the remaining two tests, but instead drew our conclusion from the above G-tests of goodness-of-fit (corrected for multiple testing).

As we show in the results, in most drops the majority of copulations were of the *M. lignano* conspecific type, followed by the heterospecific type (Fig. 6a). To check whether this could be due to an intrinsically higher mating rate of *M. lignano* (see results), we generated a new null hypothesis that takes the observed mating rates of both *M. lignano* and *M. janickei* into account. For each drop, we therefore first calculated the mating rate of *M. lignano* as

$$p = \frac{2m_{LL} + m_{LJ}}{2m_T}$$

and similarly, the mating rate of *M. janickei* as

$$q = \frac{2m_{JJ} + m_{LJ}}{2m_T}$$

where, m_{LL} , m_{LJ} , and m_{JJ} , represent the observed numbers of *M. lignano* conspecific, heterospecific, and *M. janickei* conspecific copulations, and m_T represents the total number of copulations (i.e. summed across all copulation types). Thus, we obtained a p and q value for each drop and if both species had the same mating rate, then we would expect $p = q = 0.5$. However, the results of the above analysis showed that *M. lignano* and *M. janickei* differed greatly in their mating rates (Fig. 6b).

For each drop, we therefore calculated the expected numbers of the different copulation types, given the observed mating rates p and q as

$$e_{LL} = p^2 m_T$$

$$e_{LJ} = 2pqm_T$$

and

$$e_{JJ} = q^2 m_T$$

respectively, where e_{LL} , e_{LJ} , and e_{JJ} , represent the expected numbers of *M. lignano* conspecific, heterospecific, and *M. janickei* conspecific copulations. Using these we then tested whether the resulting expected proportions were

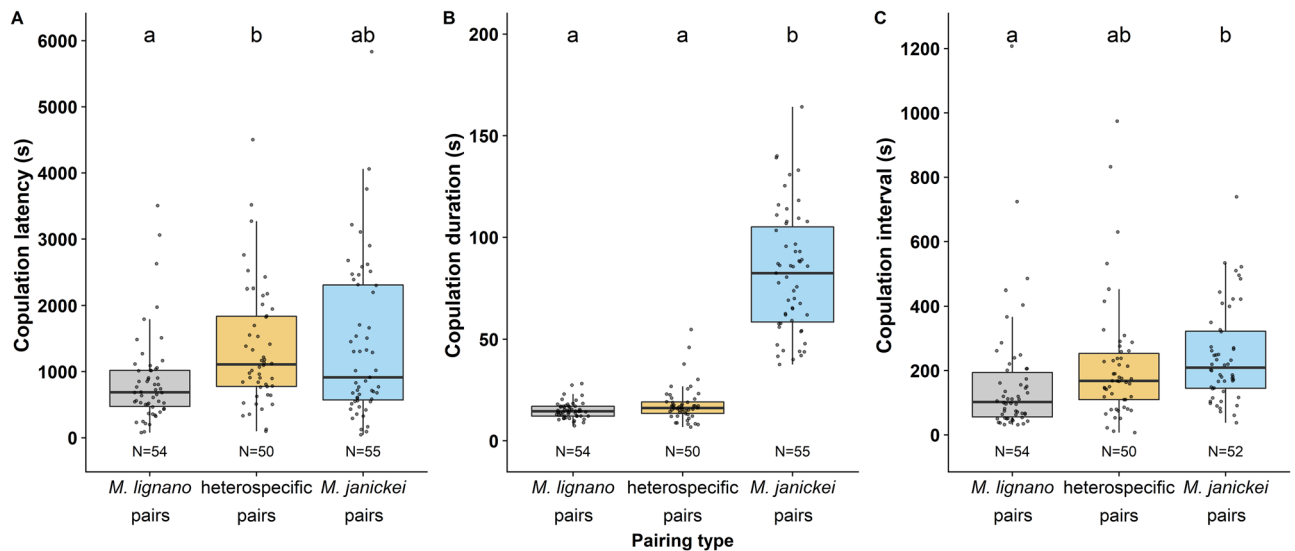


Figure 1. Boxplots of the (a) copulation latency, (b) (average) copulation duration, and (c) (average) copulation interval of the three pairing types. Different letters denote significantly different effects inferred from Tukey HSD post-hoc tests. The boxplots display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the raw data, but note that log-transformed data was used for statistical analysis of all variables shown here. Sample sizes are given at the bottom of the plots.

significantly different from the observed proportions for each drop, using a Chi-square goodness-of-fit test with Bonferroni correction for multiple testing. This allowed us to examine if the apparent preference of *M. lignano* for mating with conspecifics (i.e. the observed assortative mating) simply stemmed from the mating rate differences between the species, as opposed to a more explicit preference for conspecific partners.

All statistical analyses were carried out in R, version 3.2.4⁶⁷.

Ethical note. All animal experimentation was carried out in accordance to Swiss legal and ethical standards.

Results

Reproductive behaviour and hybridization. The three pairing types differed in their mating behaviour, though to varying degrees for the different copulation traits. Pairing type had a significant effect on copulation latency ($F_{2,156} = 4.688$, $P = 0.01$; Fig. 1a), with *M. lignano* pairs starting to copulate earlier than heterospecific pairs, while the *M. janickei* pairs had an intermediate copulation latency. The pairing type also had a significant effect on the copulation duration ($F_{2,156} = 370.6$, $P < 0.001$; Fig. 1b), with *M. janickei* pairs having a nearly five-fold longer copulation duration than *M. lignano* pairs and heterospecific pairs, which did not significantly differ amongst themselves. Moreover, the copulation interval was affected by the pairing type ($F_{2,153} = 8.124$, $P < 0.001$; Fig. 1c). *M. janickei* pairs had a significantly longer interval between copulations than *M. lignano* pairs, while the heterospecific pairs had intermediate copulation interval.

For the suck behaviour, very few heterospecific replicates exhibited the behaviour, leading to a reduction in our sample size for the time until suck and suck duration (Fig. 2). The time until suck (after copulation) differed between the pairing types ($F_{2,92} = 48.15$, $P < 0.001$; Fig. 2a), with *M. lignano* pairs usually sucking almost immediately after copulation, while the *M. janickei* pairs and heterospecific pairs took a longer time to start sucking. The suck duration was also significantly affected by the pairing type ($F_{2,92} = 7.80$, $P < 0.001$; Fig. 2b), with *M. janickei* pairs having a longer suck duration than *M. lignano* pairs, while the heterospecific pairs did not significantly differ from the other two pairing types. Interestingly, the number of sucks was significantly affected by the pairing type (Kruskal–Wallis test: $\chi^2 = 41.16$, $df = 2$, $P < 0.001$; Fig. 2c), with *M. lignano* pairs sucking most frequently, followed by the *M. janickei* pairs. The heterospecific pairs sucked least frequently.

Remarkably, for most behaviours the heterospecific pairs had the highest CV, suggesting that heterospecific behaviour was relatively variable and less stereotypic than conspecific behaviour (Table 1).

In addition, while heterospecific pairs were capable of producing hybrid offspring—a new finding for this genus—they produced significantly fewer offspring than conspecific pairs (Kruskal–Wallis test: $\chi^2 = 48.04$, $df = 2$, $P < 0.001$; Fig. 3a), which had a comparable fecundity. Out of the ten heterospecific replicates that produced hybrids, six replicates had only the *M. lignano* parent producing hybrids, while in the other four replicates only the *M. janickei* parent produced offspring. Thus, hybridization was symmetrical overall, with each species being capable of inseminating and fertilizing the other, but not symmetrical within any given pair.

Hybrid fertility. Most of the F1 hybrids were fertile and produced offspring in the wells while paired with worms from the parental species. Specifically, we found that 19/24 and 14/24 pairs of *M. lignano* x F1 hybrid and *M. janickei* x F1 hybrid produced backcrossed F2 offspring, respectively, when they were paired with an

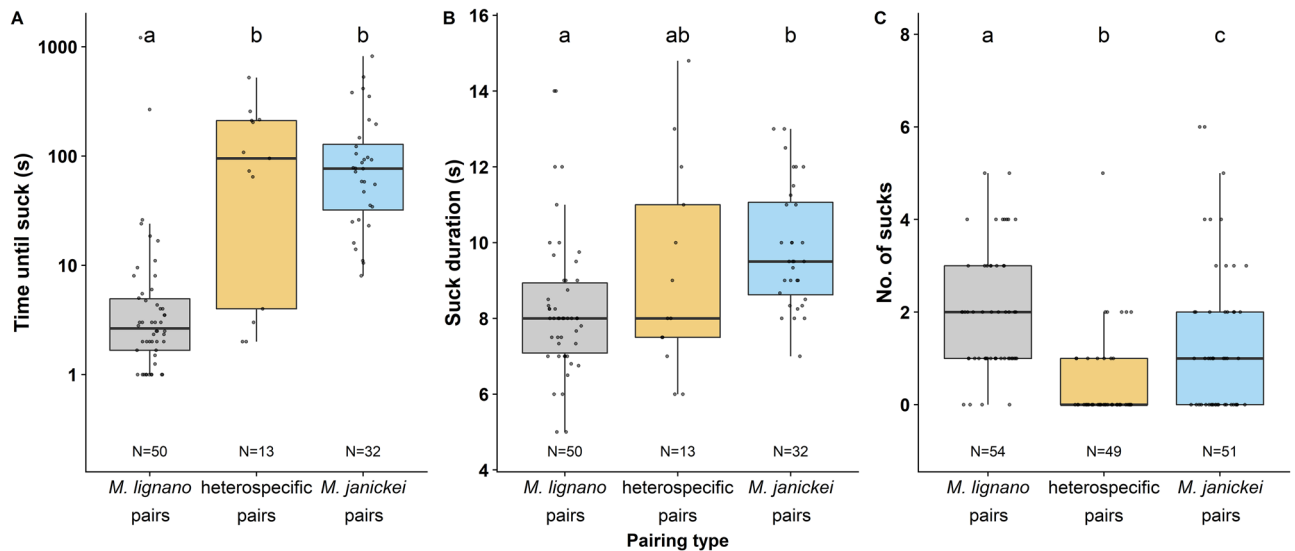


Figure 2. Boxplots of the (a) (average) time until suck (after copulation), (b) (average) suck duration, and (c) number of sucks of the three pairing types (recall that we here only consider pairs that copulated at least 5 times). Different letters denote significantly different effects inferred from Tukey HSD post-hoc tests [for (a) and (b)] or Mann–Whitney–Wilcoxon tests with Bonferroni correction [for (c)]. The boxplots display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the log-transformed data [for (a)] and the raw data for [(b) and (c)], but note that log-transformed data was used for statistical analysis [for (a) and (b)]. We added 1 to each data point for time until suck before log-transforming to avoid infinite values (see text for details). Sample sizes are given at the bottom of the plots.

| Behaviour | <i>M. lignano</i> pairs | Heterospecific pairs | <i>M. janickei</i> pairs |
|---------------------|-------------------------|----------------------|--------------------------|
| Copulation latency | 86 | 88 | 127 |
| Copulation duration | 27 | 44 | 39 |
| Copulation interval | 100 | 116 | 69 |
| Time until suck | 234 | 810 | 175 |
| Suck duration | 21 | 29 | 16 |
| No. of sucks | 66 | 209 | 120 |

Table 1. The coefficient of variation (CV) of each pairing type for all behaviours. For most behaviours the heterospecific pairs had the highest CV.

individual of one of their parental species for 7 days (Fig. 3b), while post-pairing, relatively few of the isolated F1 hybrids or parental individuals produced offspring (Fig. 3c). Interestingly, the F1 hybrids could both transfer and receive sperm from the parental species, such that F2 hybrids were produced by both the isolated F1 hybrids and parental worms (Fig. 3c).

From the 25 F1 hybrid x F1 hybrid pairs, 24 hybrid pairs had hatchlings in the wells where they were paired for 7 days (Supplementary Table 1). After being isolated, 23 hybrid individuals produced a total of 87 hatchlings, ranging from 1 to 7 hatchlings per individual over a period of 14 days.

Hybrid and parental stylet morphology. The stylet morphology was significantly different between *M. lignano*, the F1 hybrids, and *M. janickei* (Fig. 4). The centroid size, an estimate of stylet size, was different between the groups ($F_{2,69} = 33.26$, $P < 0.001$; Fig. 5a), with the F1 hybrids having a larger centroid size than *M. lignano* and *M. janickei*, which did not differ amongst themselves. The RWS1 of the stylets, which primarily seemed to capture variation in the curvature of the stylet tip and the width of the stylet base (Fig. 4g–i), was significantly different between all groups ($F_{2,69} = 238$, $P < 0.001$; Fig. 5b), with the RWS1 of the hybrids being intermediate between that of *M. lignano* and *M. janickei*, indicating that the shape of hybrid stylet was morphologically intermediate between the parental species. For visualization purposes, we show the thin plate splines of the mean RWS1 of *M. lignano*, the F1 hybrids, and *M. janickei*. This allows us to visualise how the stylet morphology varies across the parental species and F1 hybrids, although note that this is just the mean being visualised and there is variation for the RWS1 within each type (Fig. 5b).

Mate preference experiment. Out of the 59 analysed drops, we found that 34 (57.6%) drops had a copulation between two *M. lignano* individuals (i.e. *M. lignano* conspecific) as the first copulation, while only 18

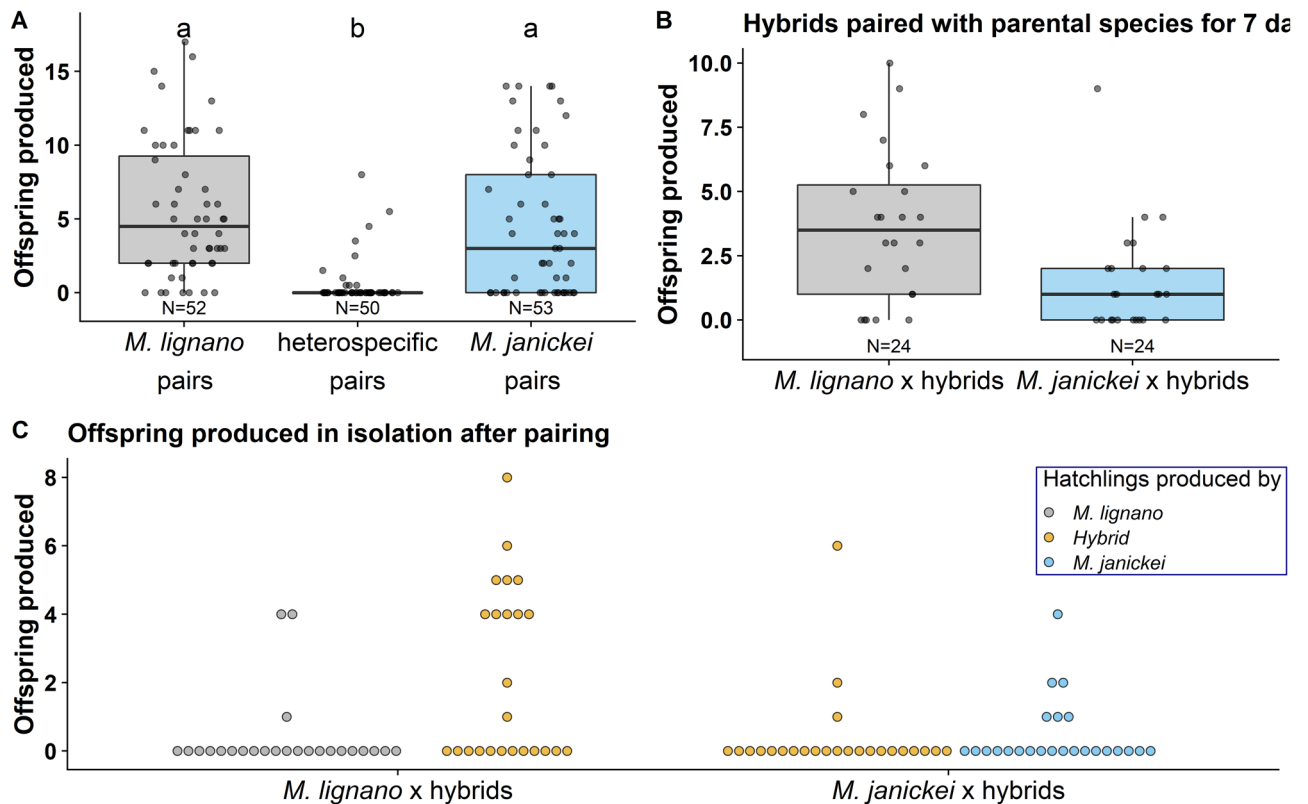


Figure 3. (a) Boxplot of the F1 hybrid offspring produced (female fecundity) by the three pairing types, (b) boxplot of the F2 hybrid offspring produced in the wells where the F1 hybrids were paired with an individual of one of their parental species for 7 days, and (c) dotplot of the F2 hybrid offspring produced post-pairing in both isolated F1 hybrids and parental individuals, respectively. The boxplots in (a) and (b) display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the raw data. Note that in (c) each backcrossed pair is represented twice as each pair comprises both a hybrid and a parental species individual, so the replicates are not independent and the figure is only for visualisation. Sample sizes are given at the bottom of the plots in (a) and (b).

(30.5%) drops had a copulation involving individuals of both species (i.e. heterospecific) as the first copulation. And, finally as few as 7 (11.9%) drops had a copulation between two *M. janickei* individuals (i.e. *M. janickei* conspecific) as the first copulation. These proportions differed significantly from our null hypothesis proportions under random mating (Chi-square goodness-of-fit test: $\chi^2 = 33.68$, $df = 2$, $P < 0.001$), which is *M. lignano* conspecific:heterospecific:*M. janickei* conspecific = 0.25:0.50:0.25.

With respect to the observed proportion of the different copulation types within drops, the data from 55 of the 59 drops differed significantly from the null hypothesis (without Bonferroni-correction $P < 0.05$, Supplementary Table S3), though after Bonferroni correction that number dropped to just 46 drops (Bonferroni-corrected $P < 0.05$, Supplementary Table S3). Interestingly, we found significant variation in the observed proportion between the drops ('heterogeneity G-value' = 358.55, $df = 116$, $P < 0.001$), as is also evident from Fig. 6a. The general trend was that *M. lignano* conspecific copulations were the most frequent, followed by heterospecific copulations, while we observed relatively few *M. janickei* conspecific copulations in most of the drops. In 51 drops, the *M. lignano* conspecific copulations were the most frequent, while in only one drop was the proportion of *M. janickei* conspecific copulations the highest (see colours in Fig. 6a). Moreover, in five drops, the highest proportion of copulations was of the heterospecific type, while in two drops, *M. lignano* conspecific and heterospecific copulations jointly had the highest proportion. Surprisingly, we found that in 52 drops there was a higher proportion of heterospecific copulations than of *M. janickei* conspecific copulations (with zero *M. janickei* conspecific copulations in 13 drops), indicating that under these conditions, the *M. janickei* worms mated more often with a *M. lignano* heterospecific than with a *M. janickei* conspecific individual. This could either represent a preference in *M. janickei* for mating with *M. lignano*, or it could potentially also result from *M. lignano* having an intrinsically higher mating rate, which we explore next.

In our mate preference assays, the mating rate of *M. lignano* and *M. janickei* was indeed different, with *M. lignano* having a much higher mating rate than *M. janickei* (Fig. 6b). When we took the mating rate differences between the two species into account, the Chi-Square goodness-of-fit test showed that in 55 out of 59 drops the observed and expected copulation frequencies were not significantly different (Bonferroni-corrected $P > 0.05$, Supplementary Table S4). This suggests that the difference in the copulation frequencies of the different copulation types, including the high frequency of heterospecific copulations in *M. janickei*, is largely explained

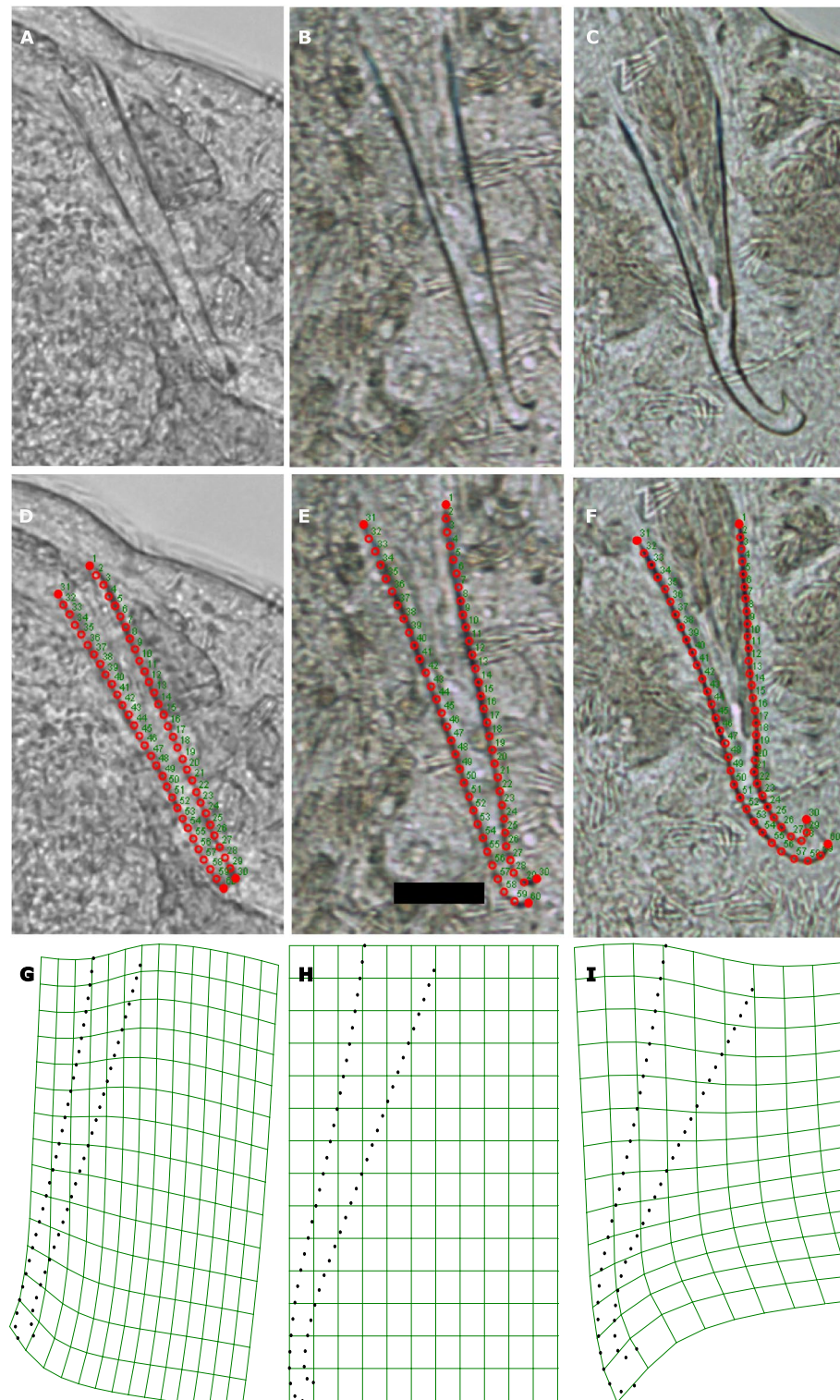


Figure 4. Morphology and geometric morphometrics of the stylet. Micrographs of the stylet of an individual (a) *M. lignano*, (b) F1 hybrid, and (c) *M. janickei*. The placement of 60 landmarks along the stylet of an individual (d) *M. lignano*, (e) F1 hybrid, and (f) *M. janickei*. Note that we placed four fixed landmarks (filled red circles), two on the stylet base and two on the stylet tip, and 28 equally spaced sliding semi-landmarks (empty red circles) along each curved side of the stylet. The numbers indicate the order in which the landmarks were placed (note that the seminal vesicles always are to the left of the stylet). Visualization of thin-plate splines of the stylet derived from relative warp score analysis. Each panel shows the visualization of the mean relative warp score 1 (RWS1) observed for individuals of (g) *M. lignano*, (h) the F1 hybrids and (i) *M. janickei*. Thus, in general *M. lignano* has a relatively straight stylet tip and *M. janickei* has a stylet tip that curves drastically, while the hybrids have intermediate curvature. The scale bar in (e) represents 20 μm , and is applicable to all photomicrographic images.

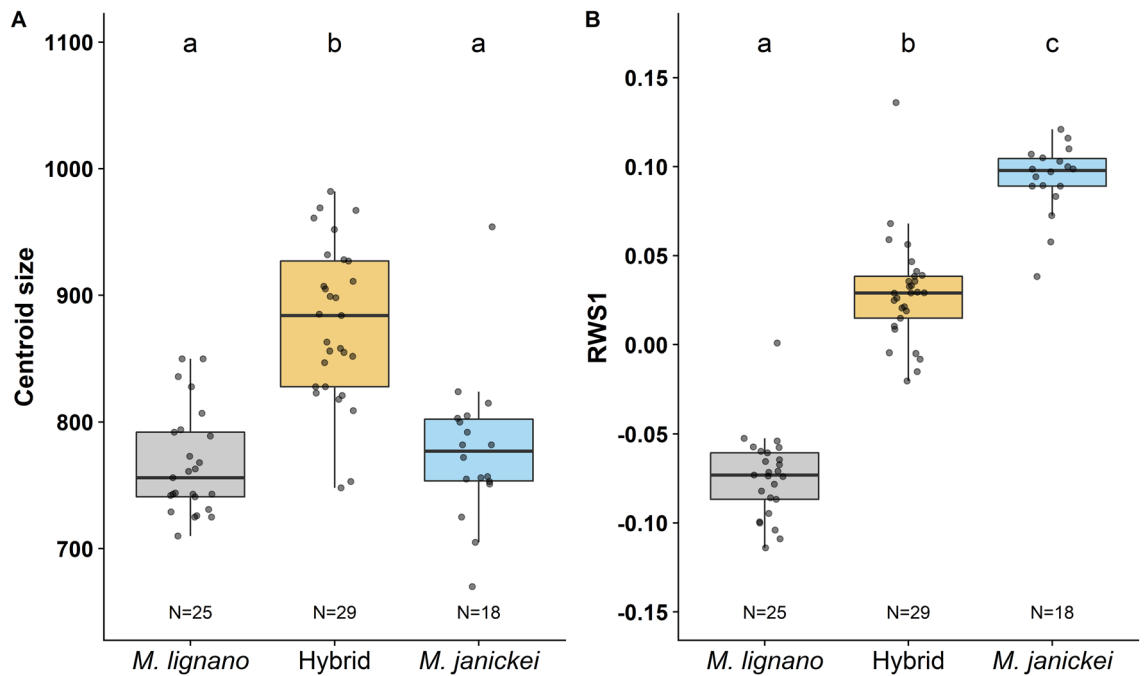


Figure 5. Boxplot for the (a) centroid size and (b) relative warp score 1 (RWS1) of the stylets of *M. lignano*, F1 hybrid and *M. janickei* worms. Different letters denote significantly different effects inferred from Tukey HSD post-hoc tests. The boxplots display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the raw data. Sample sizes are given at the bottom of the plots.

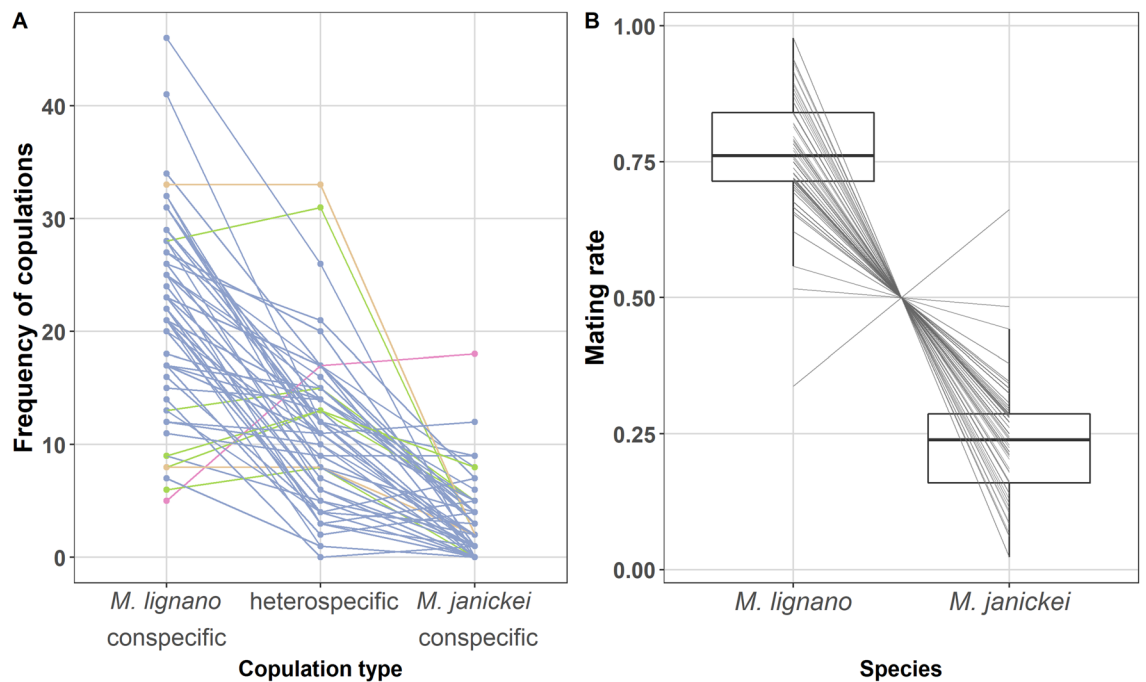


Figure 6. (a) Frequency of *M. lignano* conspecific, heterospecific, and *M. janickei* conspecific copulations. Each line connects values obtained from the same drop. The different colours help to visualise which copulation type had the highest frequency in a drop (blue, *M. lignano* conspecific; green, heterospecific; pink, *M. janickei* conspecific; orange, same in *M. lignano* conspecific and heterospecific), (b) boxplot of mating rate of *M. lignano* and *M. janickei*. The boxplots display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the raw data. Each line connects values obtained from the same drop.

by the intrinsic differences in mating rate of the two species, rather than stemming from an explicit preference for heterospecific partners.

Discussion

Our study shows that the closely related species *M. lignano* and *M. janickei* differ significantly, not only in their stylet morphology, but also in several aspects of their mating behaviour. These considerable morphological and behavioural differences do not, however, appear to represent strong premating barriers, since the worms were readily able to engage in heterospecific matings. In contrast, there seem to be significant postmating barriers between these two species, as only relatively few hybrid offspring were produced from these heterospecific matings. Moreover, the resulting F1 hybrids were fertile, showing a stylet morphology that was intermediate between that of the parental species, and these hybrids were capable of backcrossing to both parental species. Interestingly, the data from our mate preference assay revealed distinct asymmetries in the mating patterns between the two species. While *M. lignano* engaged predominantly in conspecific matings, thereby exhibiting assortative mating, *M. janickei* ended up mating more often with heterospecific individuals, and we suggest that both likely occurred as a result of the higher mating rate of *M. lignano* compared to *M. janickei*. In the following, we discuss these results in some more detail.

Reproductive behaviour and hybridization. A potential factor that could lead to the observed differences in behavioural traits between the two species is genital morphology. For example, a positive correlation between copulation duration and structural complexity of the intromittent organ has been reported in New World naticine snakes⁶⁸, wherein the authors hypothesized that the evolution of elaborate copulatory organ morphology is driven by sexual conflict over the duration of copulation. Like the findings of that study, the nearly five-fold longer copulation duration of *M. janickei* pairs compared to *M. lignano* pairs could in part be dictated by its considerably more complex stylet. Moreover, similar to the male genitalia, the female genitalia are also more complex in *M. janickei* than *M. lignano*²⁸.

In addition to genital morphology, both copulatory and post-copulatory behaviour might also be influenced by the quantity and composition of the ejaculate transferred during copulation. For example, a larger quantity of ejaculate might be accompanied by a longer copulation duration, and possibly also a longer suck duration, since the hypothesised function of the suck behaviour is to remove ejaculate components^{49,50}. Moreover, a longer copulation duration might require longer phases of recovery during which spent ejaculate is replenished, leading to lower copulation frequency and a longer copulation interval. A previous study in *M. lignano* showed that pairs formed from virgin worms copulated approximately 1.6× longer than pairs formed from sexually-experienced worms, and also that individuals that had copulated with virgin partners had a lower suck frequency compared to individuals that had copulated with sexually-experienced partners⁵². This led the authors to hypothesize that virgin partners have more own sperm and seminal fluid available (which were both confirmed), and may thus transfer more ejaculate than sexually-experienced partners, and that some components of the ejaculate are aimed at manipulating the partner and preventing it from sucking⁵². Indeed, studies in *Drosophila* have shown the presence of non-sperm components in the ejaculate, which can alter the physiology, immunity, life history, and behaviour of the recipient, causing strong effects on the fitness of both the donor and the recipient^{69–72}. Efforts to elucidate the function of ejaculate components (like seminal-fluid proteins) in *M. lignano* have recently made considerable progress^{62,73–75}, with a recent study identifying two seminal fluid transcripts that cause mating partners to suck less often, and potentially function as male counter-adaptations to suck behavior⁷⁵.

Longer copulation intervals or temporal aspects of sucking (e.g. the time until suck) could potentially also result from the action of some transferred ejaculate components that act as relaxants, leading to inactivity and delayed re-mating or delayed sucking. Interestingly, we noticed that very few individuals in the heterospecific pairs exhibited the suck behaviour, which could simply result from low or absent ejaculate transfer. It is also conceivable that sucking is triggered by species-specific ejaculate components and their interaction with the female reproductive organ, and hence the absence or low amounts of such components could result in fewer sucks. Alternatively, individuals of one species might be more effective at preventing suck in heterospecific partners, as heterospecific partners may lack coevolved defences against such ejaculate substances. Similar to our observation, a cross-reactivity study in the land snail, *Cornu aspersum*, showed that its diverticulum (a part of the female reproductive system) only responded to the love-dart mucus of some, but not other, land snail species, pointing towards species-specific effects of accessory gland products⁷⁶.

Moreover, the different behavioural components might be correlated with each other. For example, there could be a trade-off between the suck duration and suck frequency for ejaculate removal, such that longer sucks or more frequent sucks serve the same purpose. Similarly, a longer copulation duration might be accompanied by a longer suck duration and copulation interval (as discussed above). In support of this, we did see that *M. lignano* pairs had both a short copulation and suck duration, but a high copulation and suck frequency, while the converse was true for *M. janickei* pairs. Thus, there can be correlations between different aspects of reproductive behaviour and morphology, and a large-scale comparative study of reproductive behaviour and morphology in *Macrostomum* species would help to improve our understanding of the complexity and evolution of reproductive traits.

Heterospecific pairs showed higher coefficients of variation (CVs) compared to the other two pairing types for both copulation duration and copulation interval, potentially suggesting disagreements over the optimal copulation duration and copulation frequency in these pairs. In addition, heterospecific pairs exhibited higher CVs compared to conspecific pairs for all suck related behaviours. Note that in these movies we could not visually distinguish the two species in the heterospecific pairs (as the GFP-status of a worm cannot be determined under normal transmitted light), but it appears likely that the short and immediate sucks were performed by the *M. lignano* individuals, while the longer and delayed sucks were performed by the *M. janickei* individuals

in these pairs. Interestingly, the suck duration seems to be a highly stereotypical behaviour, with its CV being lower than that of copulation duration for each of the mating pair types. This is similar to what was noted from earlier behaviour studies of *M. lignano*⁴⁹.

Whereas conspecific pairs of both species produced similar offspring numbers, heterospecific pairs gave rise to offspring relatively rarely, despite most pairs having copulated successfully, presumably due to postmating-prezygotic or postzygotic reproductive barriers. In our study, hybridization was symmetrical, with both species being able to inseminate and fertilize the other species. Interestingly, in none of the heterospecific replicates did both partners produce offspring. While this could point towards unilateral transfer of sperm during copulation, we cannot ascertain if this only occurs in heterospecific pairs or if conspecific pairs also show a similar pattern, as we collected only one partner for each conspecific pair. To the best of our knowledge this is the first study to have documented hybridization between species of the genus *Macrostomum*, and there is also very sparse information only about hybridization in free-living flatworms in general^{77,78}, while there is some more information about parasitic flatworms^{79–83}.

Hybrid fertility. While historically, hybrids have often been considered to be sterile and evolutionary dead-ends⁸⁴, hybridization sometimes leads to viable and fertile offspring. In such cases, hybridization can serve as a mechanism for generating diversification, by creating adaptive variation and functional novelty in morphology and genotypes^{84,85}, a view that has been reinforced by the widespread presence of allopolyploidy among plants^{86–88}. In our study, heterospecific matings between *M. lignano* and *M. janickei* resulted in the production of viable hybrids, which we could successfully backcross onto both parental species. Although our study clearly demonstrates hybridisation between these two species, we currently have no evidence about whether these species actually occur in sympatry, although both are from the Mediterranean Sea. If we assume that the currently known sampling locations indicate an absence of sympatric zones, it would follow that the observed reproductive trait divergence might not have occurred as a result of reinforcement of reproductive isolation, but rather independent processes in allopatry. Thus, the differences in reproductive characters will not necessarily serve as reproductive barriers, and this could potentially explain our observed results.

Remarkably, both of our study species exhibit an unusual karyotype organization⁴⁷, involving hidden tetraploidy and hexaploidy in *M. lignano* and *M. janickei*, respectively (likely as a result of a whole genome duplication event). Moreover, both species show additional chromosome number variation in the form of aneuploidies of the largest chromosome, also leading to other ploidy levels⁴⁸. Interestingly, individuals with unusual karyotypes do not show behavioural or morphological abnormalities and reproduce successfully, at least in *M. lignano*⁴⁷. The fact that we can obtain viable hybrids between the two species calls for studies of the resulting karyotypes of these F1 hybrids and the F2 backcrosses.

Hybrid and parental stylet morphology. The parental species differed significantly in the morphology of their stylet, though their overall stylet size was similar. A study in closely related species of damselflies had also shown that, despite differences in genitalia morphology, the species had incomplete mechanical isolation and could hybridize³⁵. In contrast to the parental species, the hybrid offspring in our study possessed a stylet that had a morphology that was intermediate between that of the parental species, but was distinctly larger in size, for which we currently have no explanation (as already mentioned above, these results need to be interpreted with some caution, since the data used in this comparison stemmed from three separate experiments).

A previous study in *M. lignano* showed that the stylet morphology can affect the sperm-transfer success⁶⁴ in conspecific matings and hence, potentially, the fitness of an individual. Thus, it is possible that individuals with certain stylet morphologies can more successfully transfer sperm to heterospecific partners and hybridise more readily. This could potentially explain some of the variation we observe in our experiment, where only few heterospecific pairs successfully produced offspring despite mating. Future studies could explore the explicit prezygotic or postzygotic barriers that lead to such a lack of F1 hybrid production. One possibility could be that despite successful mating, sperm is not successfully transferred in some matings, e.g. due to stylet morphology differences, and this can be examined by measuring the stylet morphology of all individuals before mating and by using GFP + cultures for *M. lignano* to easily visualise and track sperm in their partners. A study on hybridizations between three species in the *Drosophila simulans* species complex showed that, despite heterospecific copulations, the three species exhibited postmating-prezygotic reproductive isolation⁸⁹. Interestingly, each of the three species-pairs exhibited a different set of cryptic barriers to heterospecific fertilization, such that either the matings were too short for successful sperm transfer; or very few heterospecific sperm were transferred even during long copulations; or that despite abundant sperm being transferred, these sperm were lost rapidly from the female's reproductive tract.

Mate preference experiment. Our mate preference experiment showed that there is a degree of assortative mating between *M. lignano* individuals, which appears to mostly stem from the intrinsically higher mating rate of *M. lignano*. This is in line with our results from the first experiment, where *M. lignano* conspecific pairs had shorter copulation latencies, shorter copulation durations and shorter copulation intervals compared to *M. janickei* conspecific pairs (Fig. 1). Thus, mate choice in these two species seems to be governed mainly by behavioural characteristics, such as the mating rate, rather than representing an explicit preference for a conspecific or heterospecific partner. A potential factor affecting mating rate evolution could be sexual selection. For instance, in polygamous mating systems, sexual selection can select for persistent mating efforts, particularly in males, which in turn can lead to reproductive interference between the species^{16,44,90}. Interestingly, a similar phenomenon has been observed in experimentally evolved populations of *Drosophila pseudoobscura* that experienced different sexual selection intensity regimes of either monogamy or polyandry^{91,92}. A mate choice experi-

ment showed that males from polyandrous populations had a higher probability of mating than those from monogamous populations⁹³, potentially due to having evolved under strong male-male competition and hence initiating courtship faster and more frequently than monogamous males⁹⁴. Similarly, an experimental evolution study on a seed beetle, *Callosobruchus chinensis*, also showed that beetles evolved under a polygamous regime caused stronger reproductive interference on the congener, *C. maculatus*, than beetles evolved under a monogamous regime^{95,96}. In addition to the above examples, multiple empirical studies have proposed a role of sexual selection in occurrence of reproductive interference between species^{97,98}.

In our experiment, the over-representation of heterospecific matings in *M. janickei* could lead to asymmetric reproductive interference between these species. Though we did not explicitly investigate how fecundity is affected, it seems likely that *M. janickei* would pay a higher fitness cost compared to *M. lignano* in such a context. Future studies should explicitly investigate if and how mating rate differences can affect the fecundity of the species and whether the cost is symmetric for both species, or if *M. janickei* suffers more due to a reduced conspecific mating rate. Population level-demographic costs of reproductive interference have been documented in a recent study using *Caenorhabditis* species, which showed that species coexistence can be influenced by reproductive interference⁹⁹. Moreover, as we outlined above, while our study raises the interesting possibility of hybridization occurring in zones of secondary contact between the two species, we are currently not aware of any overlapping ranges of the two species, but this may largely be due to the lack of sampling effort. Considering their heterospecific interactions though, it might be difficult for the species to co-exist, since *M. lignano* might be expected to displace *M. janickei* from any overlapping zones due to potential asymmetric reproductive interference. Alternatively, selection for reinforcement of reproductive isolation might occur, leading to character displacement of the species in sympatric zones, such that heterospecific interactions are reduced.

Conclusions

Our study shows that reproductive traits can evolve rapidly, even between closely related species, but that such diverged traits do not necessarily pose a reproductive barrier to hybridization. An interesting question that arises then is whether mating behaviour and genital morphology co-evolve or whether they diversify independently. A phylogenetic comparative study that looks at the evolution of these reproductive traits in more species across the *Macrostomum* genus would help us answer these open questions. Moreover, using hybridization and techniques like QTL mapping, we could aim at understanding the genetic basis of rapidly evolving and diversifying reproductive traits like mating behaviour and genitalia, and in turn broaden our understanding of speciation in free-living flatworms, a highly species-rich group of simultaneous hermaphrodites.

Data availability

All data in this manuscript is deposited online on Zenodo (<https://doi.org/10.5281/zenodo.3954023>).

Received: 2 April 2020; Accepted: 20 July 2020

Published online: 30 July 2020

References

1. Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer Associates Sunderland, Massachusetts, 2004).
2. Mayr, E. *Systematics and the Origin of Species* (Columbia University Press, New York, 1942).
3. Butlin, R. *et al.* What do we need to know about speciation?. *Trends Ecol. Evol.* **27**, 27–39 (2012).
4. Lackey, A. C. R. & Boughman, J. W. Evolution of reproductive isolation in stickleback fish. *Evolution (N. Y.)* **71**, 357–372 (2017).
5. Ostevik, K. L., Andrew, R. L., Otto, S. P. & Rieseberg, L. H. Multiple reproductive barriers separate recently diverged sunflower ecotypes. *Evolution (N. Y.)* **70**, 2322–2335 (2016).
6. Sato, Y. *et al.* Patterns of reproductive isolation in a haplodiploid—Strong post-mating, prezygotic barriers among three forms of a social spider mite. *J. Evol. Biol.* **31**, 866–881 (2018).
7. Ciccotto, P. J., Gumm, J. M. & Mendelson, T. C. Male association preference for conspecifics in the redband darter, *Etheostoma luteovinctum* (Teleostei: Percidae) based on visual cues. *Copeia* **2013**, 154–159 (2013).
8. Zhou, M., Loew, E. R. & Fuller, R. C. Sexually asymmetric colour-based species discrimination in orangethroat darters. *Anim. Behav.* **106**, 171–179 (2015).
9. Williams, T. H. & Mendelson, T. C. Behavioral isolation based on visual signals in a sympatric pair of darter species. *Ethology* **116**, 1038–1049 (2010).
10. Firman, R. C., Gasparini, C., Manier, M. K. & Pizzari, T. Postmating female control: 20 years of cryptic female choice. *Trends Ecol. Evol.* **32**, 368–382 (2017).
11. Manier, M. K. *et al.* Postcopulatory sexual selection generates speciation phenotypes in drosophila. *Curr. Biol.* **23**, 1853–1862 (2013).
12. Garlovsky, M. D. & Snook, R. R. Persistent postmating, prezygotic reproductive isolation between populations. *Ecol. Evol.* **8**, 9062–9073 (2018).
13. Turissini, D. A., McGirr, J. A., Patel, S. S., David, J. R. & Matute, D. R. The rate of evolution of postmating-prezygotic reproductive isolation in *Drosophila*. *Mol. Biol. Evol.* **35**, 312–334 (2018).
14. Soudi, S., Reinhold, K. & Engqvist, L. Strong cryptic prezygotic isolation despite lack of behavioral isolation between sympatric host races of the leaf beetle *Lochmaea capreae*. *Evolution (N. Y.)* **70**, 2889–2898 (2016).
15. Devigili, A. *et al.* Possible glimpses into early speciation: the effect of ovarian fluid on sperm velocity accords with post-copulatory isolation between two guppy populations. *J. Evol. Biol.* **31**, 66–74 (2018).
16. Gröning, J. & Hochkirch, A. Reproductive interference between animal species. *Q. Rev. Biol.* **83**, 257–282 (2008).
17. Eberhard, W. G. *Sexual Selection and Animal Genitalia*. (Harvard University Press, Harvard, 1985). <https://doi.org/10.4159/harvard.9780674330702>.
18. Arnqvist, G. The evolution of animal genitalia: Distinguishing between hypotheses by single species studies. *Biol. J. Linn. Soc.* **60**, 365–379 (1997).
19. Swanson, W. J. & Vacquier, V. D. The rapid evolution of reproductive proteins. *Nat. Rev. Genet.* **3**, 137–144 (2002).
20. Kraaijeveld, K., Kraaijeveld-Smit, F. J. L. & Maan, M. E. Sexual selection and speciation: the comparative evidence revisited. *Biol. Rev.* **86**, 367–377 (2011).

21. Janicke, T., Ritchie, M. G., Morrow, E. H. & Marie-Orleach, L. Sexual selection predicts species richness across the animal kingdom. *Proc. R. Soc. B Biol. Sci.* <https://doi.org/10.1098/rspb.2018.0173> (2018).
22. Morrow, E. H., Pitcher, T. E. & Arnqvist, G. No evidence that sexual selection is an 'engine of speciation' in birds. *Ecol. Lett.* **6**, 228–234 (2003).
23. Arnqvist, G. Comparative evidence for the evolution of genitalia by sexual selection. *Nature* **393**, 784–786 (1998).
24. Gleason, J. M. & Ritchie, M. G. Evolution of courtship song and reproductive isolation in the *Drosophila willistonii* species complex: Do sexual signals diverge the most quickly? *Evolution (N. Y.)* **52**, 1493 (1998).
25. Puniamoorthy, N., Kotrba, M. & Meier, R. Unlocking the 'Black box': Internal female genitalia in Sepsidae (Diptera) evolve fast and are species-specific. *BMC Evol. Biol.* **10**, 275 (2010).
26. Puniamoorthy, N. Behavioural barriers to reproduction may evolve faster than sexual morphology among populations of a dung fly (Sepsidae). *Anim. Behav.* **98**, 139–148 (2014).
27. Puniamoorthy, N., Ismail, M. R. B., Tan, D. S. H. & Meier, R. From kissing to belly stridulation: comparative analysis reveals surprising diversity, rapid evolution, and much homoplasy in the mating behaviour of 27 species of sepsid flies (Diptera: Sepsidae). *J. Evol. Biol.* **22**, 2146–2156 (2009).
28. Schärer, L. *et al.* A phylogenetically informed search for an alternative *Macrostomum* model species, with notes on taxonomy, mating behavior, karyology, and genome size. *J. Zool. Syst. Evol. Res.* **58**, 41–65 (2020).
29. Anthes, N. & Michiels, N. K. Reproductive morphology, mating behavior, and spawning ecology of cephalaspid sea slugs (Aglajidae and Gastropteridae). *Invertebr. Biol.* **126**, 335–365 (2007).
30. Kelly, D. A. & Moore, B. C. The morphological diversity of intromittent organs: an introduction to the symposium. *Integr. Comp. Biol.* **56**, 630–634 (2016).
31. Klappert, K., Mazzi, D., Hoikkala, A. & Ritchie, M. G. Male courtship song and female preference variation between phylogeographically distinct populations of *Drosophila montana*. *Evolution (N. Y.)* **61**, 1481–1488 (2007).
32. Herring, K. & Verrell, P. Sexual incompatibility and geographical variation in mate recognition systems: Tests in the salamander *Desmognathus ochrophaeus*. *Anim. Behav.* **52**, 279–287 (1996).
33. Ritchie, M. G., Halsey, E. J. & Gleason, J. M. *Drosophila* song as a species-specific mating signal and the behavioural importance of Kyriacou and Hall cycles in *D. melanogaster* song. *Anim. Behav.* **58**, 649–657 (1999).
34. Masly, J. P. 170 Years of "Lock-and-Key": Genital morphology and reproductive isolation. *Int. J. Evol. Biol.* **2012**, 1–10 (2012).
35. Barnard, A. A., Fincke, O. M., McPeck, M. A. & Masly, J. P. Mechanical and tactile incompatibilities cause reproductive isolation between two young damselfly species. *Evolution (N. Y.)* **71**, 2410–2427 (2017).
36. Pfennig, K. S. & Pfennig, D. W. Character displacement: Ecological and reproductive responses to a common evolutionary problem. *Q. Rev. Biol.* **84**, 253–276 (2009).
37. Brown, W. L. & Wilson, E. O. Character displacement. *Syst. Zool.* **5**, 49 (1956).
38. Blair, W. F. Character displacement in frogs. *Am. Zool.* **14**, 1119–1125 (1974).
39. Butlin, R. K. & Ritchie, M. G. Mating behaviour and speciation. in *Behaviour and Evolution* (eds. Slater, P. & Halliday, T.) 43–79 (Cambridge University Press, Cambridge, 1994).
40. Servedio, M. R. & Noor, M. A. F. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **34**, 339–364 (2003).
41. Butlin, R. K. & Smadja, C. M. Coupling, reinforcement, and speciation. *Am. Nat.* **191**(2), 155–172 (2018).
42. Germain, R. M., Williams, J. L., Schluter, D. & Angert, A. L. Moving character displacement beyond characters using contemporary coexistence theory. *Trends Ecol. Evol.* **33**, 74–84 (2018).
43. Grether, G. F., Peiman, K. S., Tobias, J. A. & Robinson, B. W. Causes and consequences of behavioral interference between species. *Trends Ecol. Evol.* **32**, 760–772 (2017).
44. Kyogoku, D. Reproductive interference: Ecological and evolutionary consequences of interspecific promiscuity. *Popul. Ecol.* **57**, 253–260 (2015).
45. Shuker, D. M. & Burdfield-Steel, E. R. Reproductive interference in insects. *Ecol. Entomol.* **42**, 65–75 (2017).
46. Ladurner, P., Schärer, L., Salvenmoser, W. & Rieger, R. M. A new model organism among the lower Bilateria and the use of digital microscopy in taxonomy of meiobenthic Platyhelminthes: *Macrostomum lignano*, n. sp. (Rhabditophora, Macrostomorpha). *J. Zool. Syst. Evol. Res.* **43**, 114–126 (2005).
47. Zadesenets, K. S. *et al.* Evidence for karyotype polymorphism in the free-living flatworm, *Macrostomum lignano*, a model organism for evolutionary and developmental biology. *PLoS ONE* **11**, e0164915 (2016).
48. Zadesenets, K. S., Schärer, L. & Rubtsov, N. B. New insights into the karyotype evolution of the free-living flatworm *Macrostomum lignano* (Platyhelminthes, Turbellaria). *Sci. Rep.* **7**, 6066 (2017).
49. these worms suck. Schärer, L., Joss, G. & Sandner, P. Mating behaviour of the marine turbellarian *Macrostomum* sp. *Mar. Biol.* **145**, 373–380 (2004).
50. Vizoso, D. B., Rieger, G. & Schärer, L. Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. *Biol. J. Linn. Soc.* **99**, 370–383 (2010).
51. Schärer, L., Littlewood, D. T. J., Waeschenbach, A., Yoshida, W. & Vizoso, D. B. Mating behavior and the evolution of sperm design. *Proc. Natl. Acad. Sci.* **108**, 1490–1495 (2011).
52. Marie-Orleach, L., Janicke, T. & Schärer, L. Effects of mating status on copulatory and postcopulatory behaviour in a simultaneous hermaphrodite. *Anim. Behav.* **85**, 453–461 (2013).
53. Janicke, T. *et al.* Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. *Evolution (N. Y.)* **67**, 3233–3242 (2013).
54. Marie-Orleach, L., Janicke, T., Vizoso, D. B., Eichmann, M. & Schärer, L. Fluorescent sperm in a transparent worm: Validation of a GFP marker to study sexual selection. *BMC Evol. Biol.* **14**, 148 (2014).
55. Wudarski, J. *et al.* Efficient transgenesis and annotated genome sequence of the regenerative flatworm model *Macrostomum lignano*. *Nat. Commun.* **8**, 2120 (2017).
56. Marie-Orleach, L., Janicke, T., Vizoso, D. B., David, P. & Schärer, L. Quantifying episodes of sexual selection: Insights from a transparent worm with fluorescent sperm. *Evolution (N. Y.)* **70**, 314–328 (2016).
57. Vellnow, N., Marie-Orleach, L., Zadesenets, K. S. & Schärer, L. Bigger testes increase paternity in a simultaneous hermaphrodite, independently of the sperm competition level. *J. Evol. Biol.* **31**, 180–196 (2018).
58. Andersen, R. A., Berges, J. A., Harrison, P. J. & Watanabe, M. M. Recipes for freshwater and seawater media. In *Algal Culturing Techniques* 429–538 (Elsevier, London, 2007). <https://doi.org/10.1016/b978-012088426-1/50027-5>
59. Janicke, T., Sandner, P. & Schärer, L. Determinants of female fecundity in a simultaneous hermaphrodite: the role of polyandry and food availability. *Evol. Ecol.* **25**, 203–218 (2011).
60. Singh, P., Vellnow, N. & Schärer, L. Variation in sex allocation plasticity in three closely related flatworm species. *Ecol. Evol.* **10**, 26–37 (2020).
61. Schärer, L. & Ladurner, P. Phenotypically plastic adjustment of sex allocation in a simultaneous hermaphrodite. *Proc. R. Soc. B Biol. Sci.* **270**, 935–941 (2003).
62. Ramm, S. A. *et al.* Sex allocation plasticity on a transcriptome scale: Socially sensitive gene expression in a simultaneous hermaphrodite. *Mol. Ecol. mec.* 15077, <https://doi.org/10.1111/mec.15077> (2019).

63. Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. *Geometric Morphometrics for Biologists: A Primer* (Elsevier Press, San Diego, 2004).
64. Janicke, T. & Schärer, L. Determinants of mating and sperm-transfer success in a simultaneous hermaphrodite. *J. Evol. Biol.* **22**, 405–415 (2009).
65. Canchola, J. A., Tang, S., Hemyari, P., Paxinos, E. & Marins, E. Correct use of percent coefficient of variation (%CV) formula for log-transformed data. *MOJ Proteomics Bioinform.* **6**(4), 316–317 (2017).
66. McDonald, J. H. *Handbook of Biological Statistics* (3rd ed.) (Sparky House Publ., Balt., 2014). <https://doi.org/10.1017/CBO9781107415324.004>
67. R Core Team. R: a language and environment for statistical computing. <https://www.R-project.org/> (2019).
68. King, R. B., Jadin, R. C., Grue, M. & Walley, H. D. Behavioural correlates with hemipenis morphology in New World naticine snakes. *Biol. J. Linn. Soc.* **98**, 110–120 (2009).
69. Perry, J. C., Sirot, L. & Wigby, S. The seminal symphony: How to compose an ejaculate. *Trends Ecol. Evol.* **28**, 414–422 (2013).
70. Chapman, T. Seminal fluid-mediated fitness traits in *Drosophila*. *Heredity (Edinb.)* **87**, 511–521 (2001).
71. Billeter, J.-C. & Wolfner, M. F. Chemical cues that guide female reproduction in *Drosophila melanogaster*. *J. Chem. Ecol.* **44**, 750–769 (2018).
72. Schwenke, R. A., Lazzaro, B. P. & Wolfner, M. F. Reproduction-immunity trade-offs in insects. *Annu. Rev. Entomol.* **61**, 239–256 (2016).
73. Weber, M. *et al.* A targeted in situ hybridization screen identifies putative seminal fluid proteins in a simultaneously hermaphroditic flatworm. *BMC Evol. Biol.* **18**, 81 (2018).
74. Patlar, B., Weber, M. & Ramm, S. A. Genetic and environmental variation in transcriptional expression of seminal fluid proteins. *Heredity (Edinb.)* **122**, 595–611 (2019).
75. Patlar, B., Weber, M., Temizyürek, T. & Ramm, S. A. Seminal fluid-mediated manipulation of post-mating behavior in a simultaneous hermaphrodite. *Curr. Biol.* **30**, 143–149.e4 (2020).
76. Lodi, M. & Koene, J. M. On the effect specificity of accessory gland products transferred by the love-dart of land snails. *BMC Evol. Biol.* **16**, 104 (2016).
77. Pala, M., Casu, S. & Lepori, N. G. Stabilized natural interspecific hybrid population of the fresh water planarians *Dugesia Gonocephala* S. L. (Turbellaria, Tricladida). *Caryologia* **35**, 247–256 (1982).
78. Bullini, L. Speciation by hybridization in animals. *Boll. Zool.* **52**, 121–137 (1985).
79. Théron, A. Hybrids between *Schistosoma mansoni* and *S. rodhaini*: Characterization by cercarial emergence rhythms. *Parasitology* **99**, 225–228 (1989).
80. Taylor, M. G. Hybridisation experiments on five species of African *Schistosomes*. *J. Helminthol.* **44**, 253–314 (1970).
81. Itagaki, T., Ichinomiya, M., Fukuda, K., Fussyuku, S. & Carmona, C. Hybridization experiments indicate incomplete reproductive isolating mechanism between *Fasciola hepatica* and *Fasciola gigantica*. *Parasitology* **138**, 1278–1284 (2011).
82. Detwiler, J. T. & Criscione, C. D. An infectious topic in reticulate evolution: Introgression and hybridization in animal parasites. *Genes (Basel)* **1**, 102–123 (2010).
83. Henrich, T., Benesh, D. P. & Kalbe, M. Hybridization between two cestode species and its consequences for intermediate host range. *Parasit. Vectors* **6**, 33 (2013).
84. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **20**, 229–237 (2005).
85. Bonnet, T., Leblois, R., Rousset, F. & Crochet, P. A. A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. *Evolution (NY)* **71**, 2140–2158 (2017).
86. Soltis, D. E. & Soltis, P. S. The dynamic nature of polyploid. *Proc. Natl. Acad. Sci.* **92**, 8089–8091 (1995).
87. Soltis, P. S., Marchant, D. B., Van de Peer, Y. & Soltis, D. E. Polyploidy and genome evolution in plants. *Curr. Opin. Genet. Dev.* **35**, 119–125 (2015).
88. Wendel, J. F., Jackson, S. A., Meyers, B. C. & Wing, R. A. Evolution of plant genome architecture. *Genome Biol.* **17**, 37 (2016).
89. Price, C. S. C., Kim, C. H., Gronlund, C. J. & Coyne, J. A. Cryptic reproductive isolation in the *Drosophila simulans* species complex. *Evolution (NY)* **55**, 81–92 (2001).
90. Burdfield-Steel, E. R. & Shuker, D. M. Reproductive interference. *Curr. Biol.* **21**, R450–R451 (2011).
91. Snook, R. R., Robertson, A., Crudgington, H. S. & Ritchie, M. G. Experimental manipulation of sexual selection and the evolution of courtship song in *Drosophila pseudoobscura*. *Behav. Genet.* **35**, 245–255 (2005).
92. Debelle, A., Ritchie, M. G. & Snook, R. R. Evolution of divergent female mating preference in response to experimental sexual selection. *Evolution (NY)* **68**, 2524–2533 (2014).
93. Debelle, A., Ritchie, M. G. & Snook, R. R. Sexual selection and assortative mating: An experimental test. *J. Evol. Biol.* **29**, 1307–1316 (2016).
94. Crudgington, H. S., Fellows, S. & Snook, R. R. Increased opportunity for sexual conflict promotes harmful males with elevated courtship frequencies. *J. Evol. Biol.* **23**, 440–446 (2010).
95. Kyogoku, D. & Sota, T. The evolution of between-species reproductive interference capability under different within-species mating regimes. *Evolution (NY)* **71**, 2721–2727 (2017).
96. Kyogoku, D., Kondoh, M. & Sota, T. Does past evolutionary history under different mating regimes influence the demographic dynamics of interspecific competition?. *Ecol. Evol.* **9**, 8616–8624 (2019).
97. Yassin, A. & David, J. R. Within-species reproductive costs affect the asymmetry of satyrization in *Drosophila*. *J. Evol. Biol.* **29**, 455–460 (2016).
98. Kyogoku, D. & Sota, T. Exaggerated male genitalia intensify interspecific reproductive interference by damaging heterospecific female genitalia. *J. Evol. Biol.* **28**, 1283–1289 (2015).
99. Ting, J. J. & Cutter, A. D. Demographic consequences of reproductive interference in multi-species communities. *BMC Ecol.* **18**, 46 (2018).

Acknowledgements

We would like to thank Tim Janicke and Georgina Rivera Ingraham for help with collecting specimens of the study species *Macrostomum janickei*, Steven A. Ramm for providing access to the images of the *Macrostomum lignano* stylets from their study, Jeremias N. Brand for help with imaging the hybrid morphology, Nikolas Vellnow and Axel Wiberg for helpful discussions, and Gudrun Viktorin, Lukas Zimmerman, Jürgen Hottinger, Urs Stiefel and Daniel Lüscher for technical support. This research was supported by grants 31003A-162543 and 310030-184916 of the Swiss National Science Foundation (SNSF) to L.S.

Author contributions

P.S. and L.S. designed the experiments. D.N.B. and M.L. collected the data for the mate preference experiment, while P.S. conducted and collected the data for the other experiments. P.S. analysed the data and wrote the first draft of the manuscript, with input from L.S. and D.N.B. All authors read and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69767-5>.

Correspondence and requests for materials should be addressed to P.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020