Evolution of mating behaviour and sex allocation plasticity in simultaneous hermaphrodites

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Pragya Singh

aus India

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Fakultätsverantwortlicher:	Prof. Dr. Dieter Ebert, University of Basel
Dissertationsleiter:	PD Dr. Lukas Schärer, University of Basel
Korreferent:	Prof. Dr. Hanna Kokko, University of Zurich

Basel, den 18.02.2020

Prof. Dr. Martin Spiess

Dekan

Evolution of mating behaviour and sex allocation plasticity in simultaneous hermaphrodites

Pragya Singh, PhD thesis

To my mother and father who explained

विद्वत्वं च नृपत्वं च न एव तुल्ये कदाचन्।

स्वदेशे पूज्यते राजा विद्वान् सर्वत्र पूज्यते॥

And to my nephew Yuvraj,

To introduce him to the wondrous lifeforms all around

Table of Contents

Abstract

Sexual conflict can arise due to disagreement between the mating partners over the usage of received ejaculate post-mating and can give rise to traits of one sex being costly to the other sex. In simultaneously hermaphroditic organisms, both the male and female sexes are combined in the same individual. This could lead to unique conflict over the sex role taken during mating and its resolution via different mating strategies. These different mating strategies could also influence the sex allocation (SA), defined as the resource allocation towards the male and female reproductive function, and its plasticity, leading to interspecific variation. In my PhD thesis, I examine the evolution of mating behaviour and SA plasticity in simultaneously hermaphroditic flatworms of the genus *Macrostomum*. This genus contains species exhibiting at least two different mating strategies. One such mating strategy is reciprocal mating, in which both partners mate in both the male and female role simultaneously, with reciprocal transfer of sperm. Another possible strategy is hypodermic insemination, in which species presumably exhibit forced unilateral mating, with sperm being injected via the male copulatory organ into the partner.

In my first chapter, using two reciprocally mating species, *M. lignano* and its congener *M. janickei,* I could show that even closely related *Macrostomum* species can exhibit substantial variation in reproductive traits, including mating behaviour and genital morphology. Interestingly, this variation does not necessarily lead to reproductive isolation, and I showed that while these two species lack premating barriers, there exist some postmating barriers, since in spite of heterospecific matings, very few hybrids were produced. Moreover, using a mate preference assay I demonstrated that the nearly two-fold higher mating rate of *M. lignano* compared to *M. janickei*, led to *M. lignano* engaging predominantly in conspecific matings, while *M. janickei* ended up mating more often with heterospecific individuals. Thus, mating rate can be an important determinant for shaping heterospecific interactions.

The high mating rates that are often associated with reciprocal mating can lead to intense postcopulatory sexual selection or/and conflict and evolution of female resistance traits. A potential example of a female resistance trait is the postcopulatory ‘suck’ behaviour in *Macrostomum*, which is hypothesised to be used for removing received ejaculate, during which a worm bends down and places its pharynx on top of its female antrum (sperm receiving organ) and appears to suck. In my second chapter, I provided conclusive evidence

that the suck behaviour removes ejaculate out of the antrum in a reciprocally mating species, *M. hamatum*. I also, examined the evolution of the suck behaviour by documenting behavioural interactions of 64 species, and showed that there is a correlation between the presence, frequency and duration of reciprocal mating and suck, providing support for the hypothesis that the suck behaviour co-evolves with reciprocal mating. Moreover, I showed that the mating strategy of a species can be inferred from a combination of reproductive morphological traits, so called syndromes, in *Macrostomum*.

Hermaphroditic individuals also face a unique challenge in terms of needing to decide about the allocation of resources towards their male and female function. SA models predict that the amount of resources invested into the male function should vary with local sperm competition (LSC), in which related sperm compete for fertilizing a given set of eggs. In small groups, where there is high LSC, low allocation of resources to the male function is favoured. As the group size increases, LSC decreases and that favours increased allocation towards the male function. Moreover, LSC can vary temporally or spatially during an individual's lifetime, leading to plasticity in SA. In our experiments, I raised worms in three group sizes (isolated, pairs and octets) and measured their SA, calculated as testis size/(testis size+ovary size). In my third and fourth chapter, in line with the above predictions, I showed plasticity in SA in four *Macrostomum* species in response to variation in group size, with the species usually exhibiting lower SA in smaller groups. Moreover, theory predicts that any process, such as self-fertilization, which affects the strength of LSC can have an effect on the optimal SA, by changing the shape of the male fitness gain curve. In my fourth chapter, using data for seven *Macrostomum* species, I showed that there was substantial interspecific variation in SA and SA plasticity. I also, calculated standardized effect sizes for SA plasticity due to the presence of mating partners (i.e. isolated worms vs. worms with partners) and the strength of LSC (i.e. worms in pairs vs. octets) for each species. I showed that while the mating strategy does not have any effect on SA plasticity, there is a significant effect of self-fertilization such that self-fertilizing species had a lower SA plasticity with respect to presence of mating partners. To summarize, my thesis shows that both mating behaviour and SA plasticity can evolve rapidly in the *Macrostomum* genus.

Introduction

Hermaphroditism, which combines both male and female sexes in the same individual, either sequentially or simultaneously, is an understudied but widespread sexual system in animals (occurring in over 70% animal phyla, Jarne and Auld 2006) and offers a unique context for studying different biological processes. Using hermaphrodites as study systems can improve our understanding of biological processes, by providing a different perspective, and help us formulate general principles that would hold across all sexual systems. In my thesis I focus on simultaneous hermaphrodites. Classically, simultaneous hermaphroditism has been associated with low population density, and a sessile or parasitic life-history trait, such that the presence of both sexes in an individual allows the possibility of self-fertilization in absence of mating partners, and every encountered individual to be a potential mating partner (Darwin 1876; Altenburg 1934; Tomlinson 1966; Ghiselin 1969; Borgia and Blick 1981; Manning 1983). Studies have now suggested that simultaneous hermaphroditism does not necessarily only occur at low densities, and can also be stable if one of the sex functions has a strongly diminishing fitness returns on investment, such that it pays more to invest resources into the other sex function (Charnov 1982; Schärer 2009). Moreover, there is also strong phylogenetic inertia that might explain the current taxonomic distribution of simultaneous hermaphroditism (Michiels 1998; Jarne and Auld 2006; Eppley and Jesson 2008; Anthes et al. 2010), and 'maladaptive' seeming hermaphroditic systems (Michiels et al. 2009). The biology of simultaneous hermaphrodites, leads to two routes to fitness in an individual, i.e. via the male and the female function. This can lead to unique pre- and post-mating sexual conflicts. In addition, presence of both sex functions in an individual leads to the possibility and challenge of strategically allocating the available resources to the two sex functions such as to maximise the individual's fitness, as predicted by sex allocation theory.

Sexual conflict in simultaneously hermaphroditic animals

Sexual conflict is defined as the conflict between the two sexes, and can occur over the usage of the received ejaculate. It is rooted in anisogamy, with member of one sex producing smaller and more numerous gametes (called sperm in animals) than the other sex (called eggs in animals) (Charnov 1979; Parker 1979, 2006; Arnqvist and Rowe 2005). This conflict can in turn give rise to traits of one sex being costly to the other and an ensuing evolutionary arms race between the two sexes (Arnqvist and Rowe 2005). Sexual conflict concepts, which are typically defined in the literature using gonochoristic

terms, can be easily extended to simultaneously hermaphroditic animals, by considering the mating partners as sperm donors and sperm recipients, respectively (instead of as males and females) (Schärer et al. 2015). Despite the similarities with gonochoristic species, there are also some fundamental differences in the nature of the sexual conflict processes that operate in simultaneous hermaphrodites (hermaphrodites hereafter), and the responses it elicits. One such unique challenge that they face is over sex role preferences while mating.

Mating conflict in simultaneously hermaphroditic animals

Hermaphrodites face a unique conflict during mating interactions over sex role preferences, as each partner can mate as either sperm donor and/or sperm recipient, while the costs and benefits of mating in each sex role might differ (Charnov 1979; Michiels 1998; Anthes 2010; Schärer et al. 2015). The possible mating interactions between two such interacting individuals can be visualised via a matrix of compatible and incompatible sex role preferences, with a mating conflict resulting whenever the individuals want to adopt incompatible sex roles (Figure 1). An important assumption of sexual conflict theory in hermaphrodites is that Bateman's principle (Bateman 1948) also applies to hermaphrodites. In his seminal paper, Charnov (1979) argued that hermaphroditic individuals mate more to give away sperm than to receive sperm, resulting in a preference to mate as a sperm donor and hence a resulting mating conflict. Note that there is some debate on whether Bateman's principle applies to simultaneous hermaphrodites (Leonard and Lukowiak 1984; Michiels 1998; Leonard 2005; Anthes 2010; Schärer et al. 2015), with some suggesting that the sperm recipient role might be preferred in internally fertilizing species, as it allows control over the fate of the received ejaculate (Leonard and Lukowiak 1984; Leonard 2005), although the evidence for this has usually been descriptive studies (Anthes 2010). Studies on measuring the male and female gradients in hermaphrodites will allow us to better understand the sex role preferences and the resultant mating conflicts (Schärer et al. 2015), although note that Bateman gradients only estimate the fitness benefits from additional mating without taking the costs associated with it into account (Jennions and Kokko 2010; Kokko et al. 2012).

Assuming a preference for donating sperm, there can be diverse mating strategies to resolve the mating conflict. In the following, I focus on two possible mating strategies: reciprocal mating and forced unilateral insemination (Figure 1).

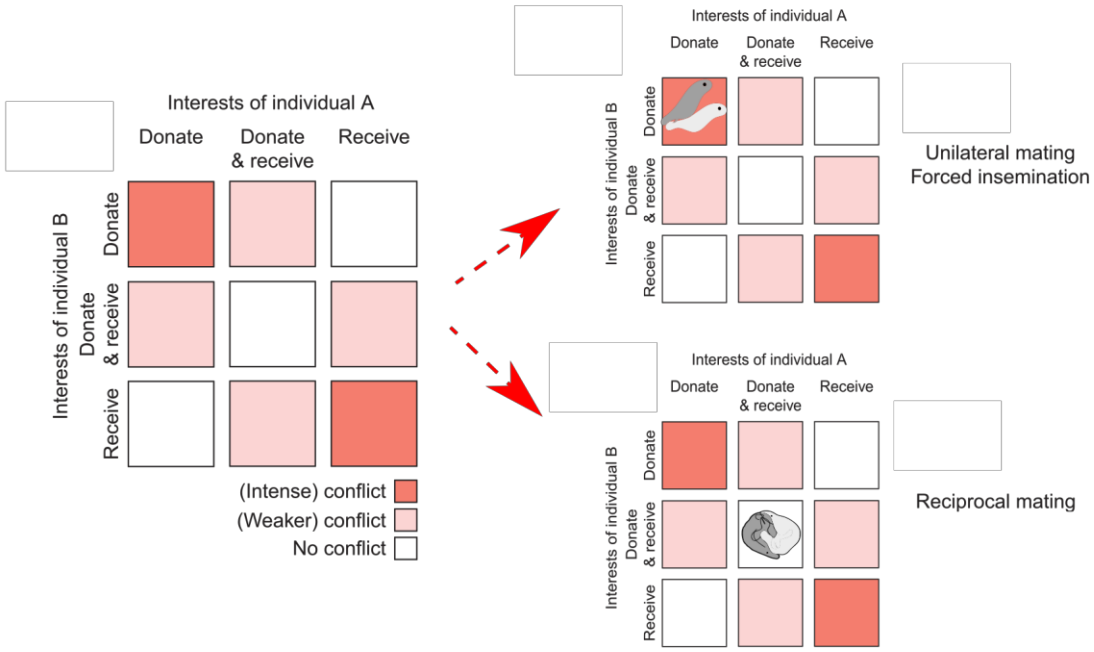


Figure 1. Possible mating interactions between two interacting individuals, where each individual can donate, receive, or both donate and receive sperm. Incompatible preference for sex roles can lead to mating conflict, visualised by red or pink squares. Assuming a preference for donating sperm, there can be different mating strategies to achieve a mating (from the perspective of individual A) of which I visualise two potential strategies indicated by the red dashed arrows. Note that there are other potential mating strategies that I do not visualise here (Figure modified from Schärer et al., 2015).

In reciprocal mating (also called reciprocal copulation), both partners donate and receive sperm simultaneously (Michiels 1998; Anthes 2010; Schärer et al. 2015) (Figure 1). This could, on the one hand, lead to a reduction in pre-copulatory sexual conflict, as both partners may be willing to engage in matings to be able to donate sperm. On the other hand, it may lead to an increase in the mating rate and consequently more intense post-copulatory sexual conflict or/and sexual selection (note that these are not necessarily exclusive, Kokko and Jennions 2014). Such multiple mating can lead to receipt of excessive ejaculate that could be harmful to the sperm recipient, e.g. if it leads to polyspermy, exposure to sexual infections, or contains manipulative substances that affect the sperm recipient detrimentally (Schärer et al. 2015). Moreover, receipt of sperm

from multiple partners can lead to evolution of cryptic female choice (Eberhard 1996). Thus, there could be selection for postcopulatory female resistance traits that allow manipulation of received ejaculate in the sperm recipient, such as sperm digestion (Sluys 1989; Baur 1998; Michiels 1998; Westheide 1999; Angeloni et al. 2003; Dillen et al. 2009; Koene et al. 2009).

Such postcopulatory female resistance traits can lead to counterselection for mating strategies or male persistence traits in the sperm donor that allow the ejaculate to circumvent the female resistance traits and thereby gain access to the eggs for fertilization. An example of such a mating strategy is the evolution of forced unilateral insemination (Figure 1). In this mating strategy, an individual insists on unilateral donation of sperm, usually against the partners interest, e.g. in hypodermic insemination or traumatic mating (Charnov 1979; Michiels 1998; Lange et al. 2013; Reinhardt et al. 2015; Schärer et al. 2015). Interestingly, traumatic insemination is common in simultaneous hermaphrodites (Lange et al. 2013), and one possible explanation suggested for this pattern is the tendency for mating conflict to escalate into damage/harm to the mating partner in hermaphrodites (Michiels and Koene 2006; Preece et al. 2009). An example of a potential persistence trait causing mate harm is the love dart in the hermaphroditic land snail, *Bradybaena pellucida*, in which an individual stabs its mating partner with a love dart that benefits the sperm donor, but has costs for the sperm recipient (Kimura and Chiba 2015). Despite these costs, the love dart is present in multiple species of land snails, though experimental data on the costs associated with the dart in other species is still lacking (Lodi and Koene 2015).

In summary, disagreement over the sex role and the usage of gametes post-mating can lead to a diversity of mating strategies. Even within the same mating strategy, antagonistic co-evolutionary arms race between resistance and persistence traits could lead to rapidly evolving reproductive behaviour and morphology.

Sex allocation in simultaneously hermaphroditic animals

Interestingly, the diversity in mating strategies exhibited in hermaphrodites can also influence the sex allocation, as I discuss below. Hermaphrodites face a unique challenge in that they need to decide on the allocation of resources towards their male and female function. Sex allocation (SA) theory, which has been successfully applied and demonstrated in different sexual systems, predicts the relative investment into the male

versus the female sex function, such as to maximise reproductive success (Charnov 1979, 1982, 1996).

One of the scenarios that leads to hermaphroditism being favoured is when the investment into a sex function shows diminishing fitness returns, such that it is more fruitful to reallocate the reproductive resources into the other, non-saturating, sex function to maximise the reproductive success. Usually, it is the male fitness gain curve that is thought to be saturating, while the female fitness gain curve is thought to have more linear returns (Charnov 1979, 1982; Schärer 2009). A possible situation that could lead to the male fitness gain curve saturating is local sperm competition (LSC) (Schärer 2009), and predictions of multiple SA models deal with how an individual allocates it resources in response to variation in LSC. In LSC, related sperm (usually from the same individual) compete for fertilizing partner's eggs, and it is in some way analogous to local mate competition in gonochorists (Schärer 2009). Thus, the prediction is that when LSC is high, low allocation of resources to male function, and hence a low SA (here defined as resources allocated to male function relative to both male and female function) is favoured, while a decrease in LSC should favour higher allocation of resources to male function and hence a higher SA, up to the asymptote of 50% (Schärer 2009).

In general, any process that affects the strength of LSC can have an effect on the optimal SA, by changing the shape of the male fitness gain curve (Schärer 2009). Such processes can include the mating group size, which depends on the number of sperm donors from which a sperm recipient receives sperm at the time its eggs are fertilized (Charnov 1980, 1982), the selfing-rate (Charlesworth and Charlesworth 1981), post-copulatory sexual selection processes, such as cryptic female choice (van Velzen et al. 2009) and sperm displacement (Charnov 1996; Schärer 2009; Schärer and Pen 2013). To my knowledge, no study has explored the effect of post-copulatory sexual selection processes on the evolution of SA in hermaphroditic animals, while there has been some work on the effect of mating group size and the selfing rate (Schärer 2009).

While the usual models of SA theory predict optimal SA on evolutionary timescales, LSC can vary temporally or spatially during an individual's lifetime leading to plasticity in SA (Schärer 2009). SA plasticity is thought to be particularly relevant for hermaphrodites, where a change in SA affects an individual during its own lifetime rather than in the next generation, as is the case for gonochorists (Michiels 1998; Schärer 2009). Moreover, studies on interspecific variation in SA can only be reliably interpreted

if we have an estimate of the intraspecific variation in SA, though note that both plasticity and standing genetic variation in SA can lead to variation in SA.

Model system

An excellent model system for studying SA, sexual selection, and sexual conflict in hermaphrodites is the free-living flatworm genus *Macrostomum* (Macrostomorpha, Platyhelminthes). It contains species that are usually highly transparent and small, which allows us to study their internal reproductive structures in detail (Schärer and Ladurner 2003; Vizoso et al. 2010; Giannakara and Ramm 2017; Winkler and Ramm 2018; Singh et al. 2019). The general morphology of *Macrostomum* flatworms is given in Figure 2a (represented by the species *M. lignano*). Briefly, an individual has a pair of testes and ovaries, with the paired testes located anterior to the paired ovaries. The sperm produced in the testes is transported to the seminal vesicle via the *vasa deferentia*. The seminal vesicle is the intermediate sperm storage organ from which sperm is later transferred during mating to the partner via the male intromittent organ, the stylet. The ovaries produce eggs, which develop posterior to the ovaries in the growth zone, and then pass on to the female antrum, located posterior to the growth zone, from which egg laying occurs.

Interestingly, *Macrostomum* species exhibit a striking diversity in their genital morphology and sperm design, which is associated with one of two kinds of mating strategies; one involves reciprocal mating and the other hypodermic insemination (Schärer et al. 2011) (Figure 2b). On the one hand, reciprocal mating is associated with a complex sperm design (usually including stiff lateral bristles on the sperm), a thickened epithelium of the female antrum and a post-copulatory suck behaviour (Schärer et al. 2011). Moreover, the female antrum is also involved in mating in reciprocally mating species like *M. lignano*, and received sperm can generally be found in the female antrum of such species (Brand et al. in preparation). During the postcopulatory suck behaviour, the worm places its pharynx over its female antrum opening and appears to suck (Schärer et al. 2004; Vizoso et al. 2010; Marie-Orleach et al. 2013) (Figure 2c). Suck has been proposed to be involved in ejaculate and sperm removal, with sperm often seen sticking out of the female antrum after the suck (Schärer et al. 2004). On the other hand, hypodermic insemination, which involves sperm injection through the partners' epidermis, is associated with convergent evolution of a needle-like stylet, a simpler

bristle-less sperm morphology, a simpler female antrum and no documented presence of the suck behaviour (Schärer et al. 2011). The hypodermically inseminating species do not generally have sperm in their female antrum, which is presumably involved only in egg laying (Brand et al. in preparation). Hypodermic insemination might also be associated with self-fertilization, with the needle-like stylet permitting self-insemination of sperm into the anterior region of the body (Ramm et al. 2012; Schärer et al. 2015; Giannakara and Ramm 2017). This association between the mating strategy, and genital and sperm morphology (defined as mating syndrome) has likely resulted from coevolution.

Within the *Macrostomum* genus, *M. lignano* has been a model organism for research in a variety of fields, including stem cell research (Ladurner et al. 2008), ageing (Mouton et al. 2009), karyology (Zadesenets et al. 2016, 2017), and sexual conflict and sexual selection (Schärer 2009; Schärer et al. 2015). Its closely related sister species is *M. janickei* (Schärer et al. 2019) has been collected from Italy, and recent field trips to Australia have permitted the collection of two additional species, that are also closely related, *M. cliftonensis* and *M. mirumnovem*. All four species are being successfully cultured in the laboratory (along with other species), and provide an excellent opportunity to explore variation in reproductive biology between closely-related reciprocally mating species.

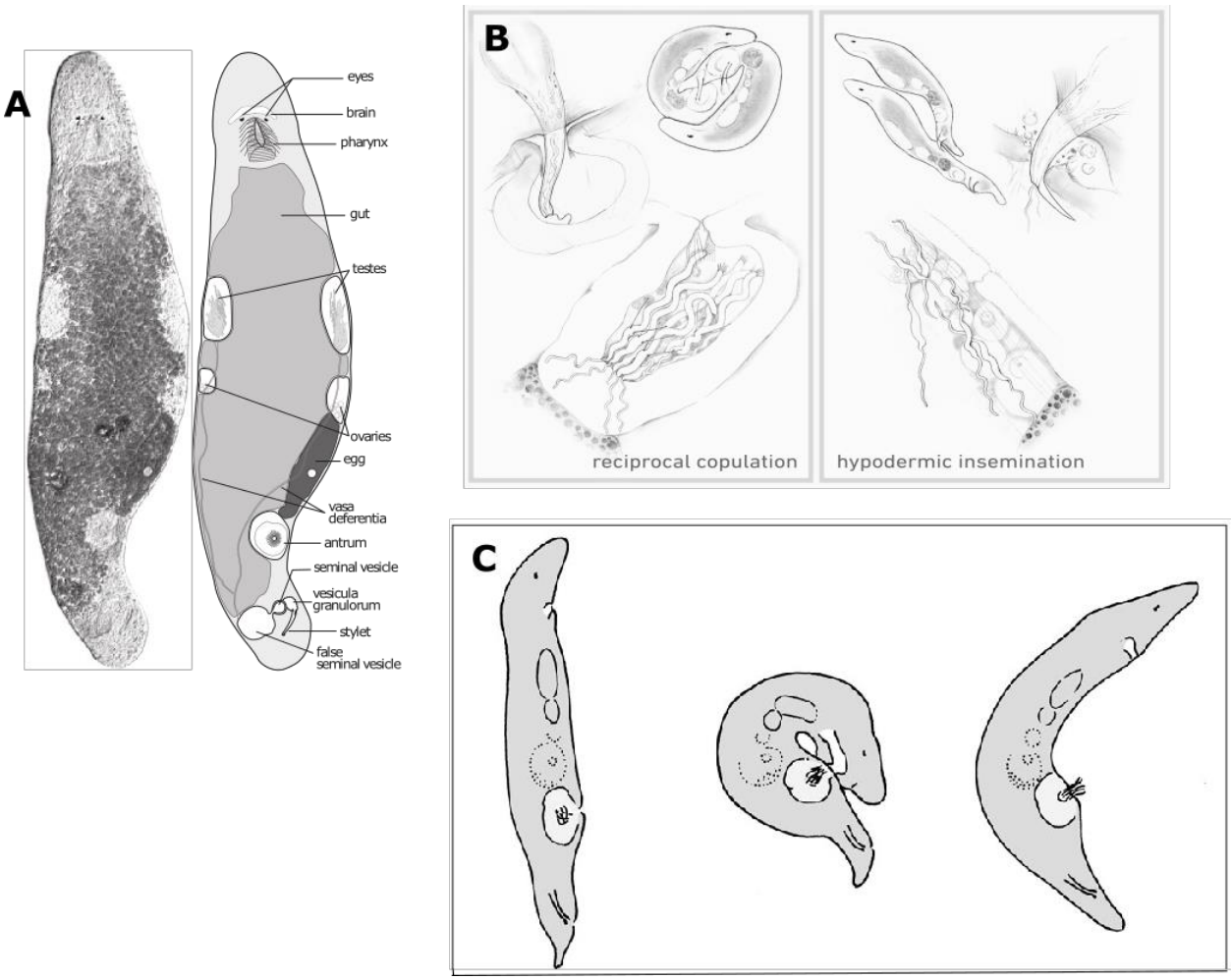


Figure 2. A) Micrograph and line drawing of a live adult specimen of *M. lignano* (~1.8 mm) (Figure from Vizoso et al., 2010). B) An artistic rendering of the mating postures, the processes of sperm transfer, and the location and morphology of the received sperm in the mating partner in a reciprocally mating (left) and a hypodermically inseminating (right) *Macrostomum* species (Artwork by Dita B. Vizoso). C) Depiction of the stages of the suck behaviour in *M. lignano*, in which the worm, after copulation, places its pharynx on top of its female antrum and appears to suck (Figure from Vizoso et al., 2010). Sperm can often be seen sticking out of the female antrum after the suck.

Thesis outline

In my thesis I examine the evolution of mating behaviour and sex allocation plasticity in the *Macrostomum* genus. For two of my chapters, I focus on species that are closely related to *M. lignano*, as this provides an excellent opportunity to explore variation in reproductive biology between closely-related hermaphroditic species. For the other chapters, I use species across the genus, including both reciprocally mating and hypodermically inseminating species to examine copulatory and postcopulatory behaviour, and sex allocation plasticity and its predictors.

In my **first chapter**, I study how differences in reproductive traits may evolve rapidly, by comparing two closely related *Macrostomum* species, and ask whether these differences serve as reproductive barriers. Specifically, I used the congeneric species pair, *M. lignano* and *M. janickei*, which despite being closely related, differ substantially in their stylet morphology. I examined 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. Remarkably, despite significant interspecific morphological and behavioural differences, some of the heterospecific matings produced hybrids. I examined the fertility and genital morphology of these resulting hybrids. Finally, using a mate choice experiment, I tested if *M. lignano* and *M. janickei* preferentially mated with conspecific individuals over heterospecifics, since such a preference could represent a premating barrier in putative zones of sympatry.

In my **second chapter**, I expand the comparison beyond *M. lignano* and its sister species, by aiming to collect data on mating behaviour in 64 species across the *Macrostomum* genus. Using this data, I examine, in a phylogenetic context, the evolution of the postcopulatory 'suck' behaviour, which is hypothesized to be a female resistance trait involved in manipulating the fate of the received ejaculate. Excitingly, as the genus contains both reciprocally mating and hypodermically inseminating species, this allows us to investigate the association of the suck behaviour with reciprocal mating. First, I provide evidence that ejaculate is indeed removed during the suck behaviour in the reciprocally mating species, *M. hamatum*. Next, I examine if there is a correlation between the presence, duration and frequency of the reciprocal mating and suck behaviour, as this would suggest that the suck behaviour co-evolves with reciprocal

mating. Finally, I test if the mating syndrome (deduced from the reproductive morphology) is a good proxy for the mating strategy.

In my **third chapter**, I examined variation in sex allocation plasticity in the three closely related *Macrostomum* species, namely *M janickei, M cliftonensis*, and *M mirumnovem*. Specifically, I tested Charnov's mating group size model (Charnov 1980, 1982), which predicts an influence of the group size on SA, defined as testes size/(testes size+ovaries size), mediated via changes in LSC such that SA increases with increase in group size. In my study, for each species, I experimentally raised worms in three group sizes (isolated, pairs, and octets) and two enclosure sizes (small and large) in all factorial combinations and studied the effects of these factors on different estimates of SA (Figure 3). In preparation for a larger comparative study (see chapter 4), I compared the different species by calculating standardized effect sizes for SA plasticity for (a) the presence/absence of mating partners (i.e. isolated worms vs. worms with partners) and (b) the strength of LSC (i.e. paired worms vs. octet worms). In addition, I also evaluated whether isolated worms can engage in self-fertilization (Figure 3).

In my **fourth chapter**, I first assessed the effect of three group sizes (isolated, pairs or octets) and two enclosure sizes (small and large) in all factorial combinations, on estimates of SA in a new and currently still undescribed species, *Macrostomum* sp. 22, and examined if isolated worms can self-fertilize (Figure 3). Then, I examined whether and how interspecific differences in the mating strategy or the presence of self-fertilization influence the effect size of SA plasticity via changes in LSC. For this, I, across seven *Macrostomum* species, obtained standardized effect sizes for SA plasticity due to i) the presence of mating and ii) the strength of LSC, and tested if the mating strategy or the ability to self-fertilize had an effect on these standardized effect sizes, while controlling for the phylogeny. Furthermore, I assessed how consistent these effect sizes were across multiple published experiments of *M. lignano*.

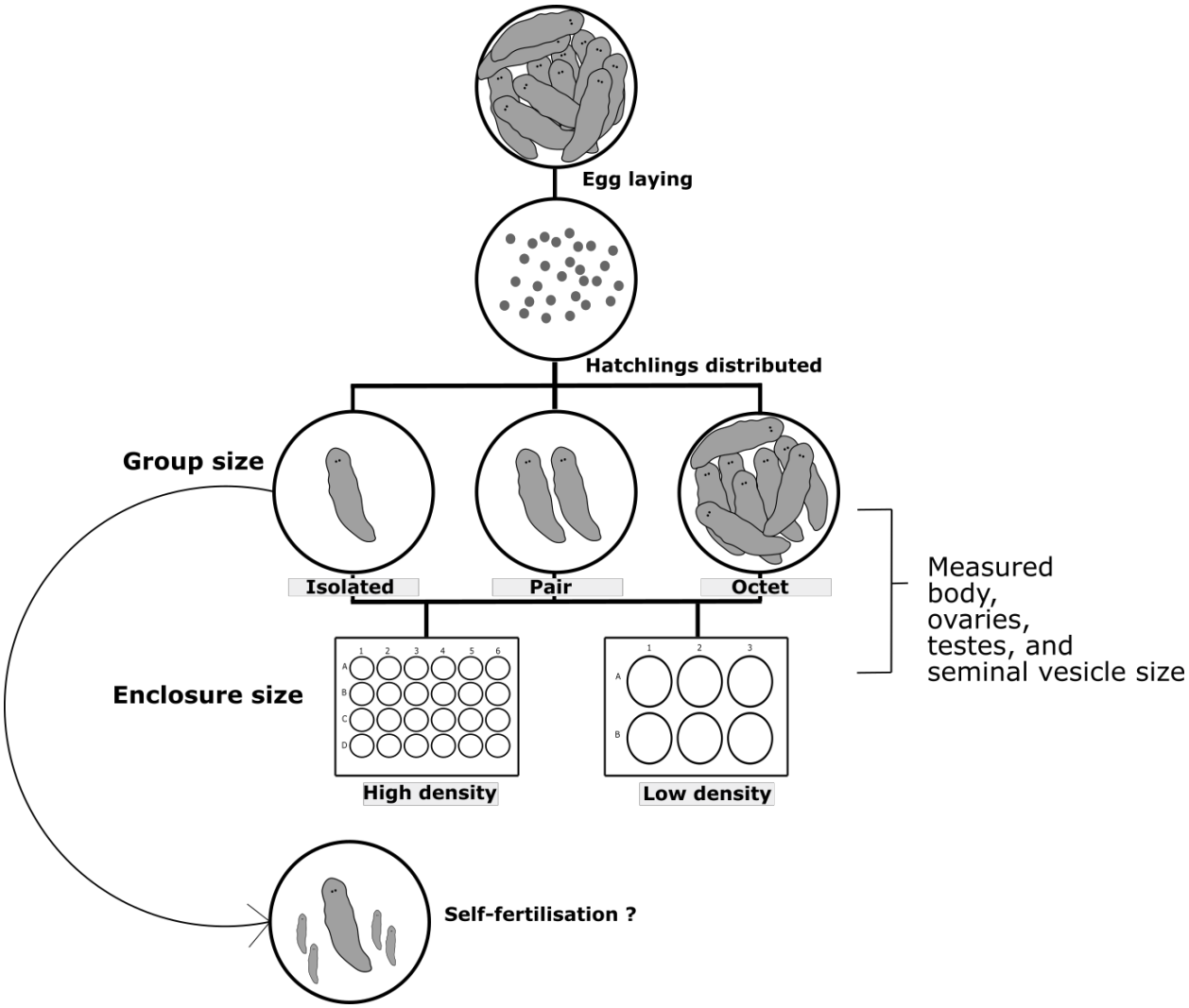


Figure 3. Graphic illustration of the experimental design for examining sex allocation plasticity in a species by varying group size and enclosure size (density), for Chapter III and IV.

Chapter I

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

Singh, P., D. Ballmer, M. Laubscher, and L. Schärer. In revision. Successful mating and hybridisation in two closely related flatworm species despite significant differences in reproductive morphology and behaviour.

Abstract

Speciation is usually a gradual process, in which premating or postmating reproductive barriers between two species accumulate over time. Reproductive traits, like genital morphology and mating behaviour, are some of the fastest diverging characters and can serve as reproductive barriers. The free-living flatworm *Macrostomum lignano*, an established model for studying sex in hermaphrodites, 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, with *M. janickei* having a five-fold longer copulation duration, copulating less frequently, and having a longer and more delayed suck behaviour (a postcopulatory behaviour likely involved in sexual conflict). Interestingly, despite these significant morphological and behavioural differences, the two species mate readily with each other in heterospecific pairings, often showing behaviours of intermediate duration. Although both species have similar fecundity in conspecific pairings, the heterospecific pairings revealed clear postmating barriers, as only few heterospecific pairings produced F1 hybrids. These hybrids had a stylet morphology that was intermediate between that of the parental species, and they could successfully backcross to both parental species. Finally, in a mate choice experiment we tested if the worms preferentially mated with conspecifics over heterospecifics, since such a preference could represent a premating barrier. Interestingly, the experiment showed 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*.

Introduction

The biological species concept defines species as groups of individuals that interbreed in nature to produce viable and fertile offspring (Mayr 1942; Coyne and Orr 2004). 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 (Butlin et al. 2012; Ostevik et al. 2016; Lackey and Boughman 2017; Sato et al. 2018). While prezygotic barriers involve the prevention of zygote formation, postzygotic barriers lead to zygote mortality, or 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 (Williams and Mendelson 2010; Ciccotto et al. 2013; Zhou et al. 2015). 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 (Manier et al. 2013; Soudi et al. 2016; Firman et al. 2017; Devigili et al. 2018; Garlovsky and Snook 2018; Turissini et al. 2018).

Species in the early stages of divergence will often not have complete reproductive barriers between them, but as they diverge in their traits, more reproductive barriers usually accumulate over time, since these divergent traits can function as barriers. Reproductive traits may diverge particularly quickly, since they are the primary targets of sexual selection, often leading to rapid accumulation of phenotypic differences (Eberhard 1985; Arnqvist 1997; Swanson and Vacquier 2002; Gröning and Hochkirch 2008). 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 (Arnqvist 1998; Gleason and Ritchie 1998; Puniamoorthy et al. 2009, 2010; Puniamoorthy 2014) and can differ markedly even between recently diverged species (Anthes and Michiels 2007; Puniamoorthy et al. 2009, 2010; Kelly and Moore 2016; Schärer et al. 2019), and sometimes even between populations of the same species (Herring and Verrell 1996; Klappert et al. 2007; Puniamoorthy 2014). Moreover, some studies have

shown that mating behaviour might evolve faster than genital morphology (Puniamoorthy 2014). Thus, a rapidly evolving reproductive trait like reproductive behaviour can represent a premating barrier by being involved in mate recognition and assortative mating (Herring and Verrell 1996; Ritchie et al. 1999), while a difference in genital morphology can prevent successful mating and thus represent a mechanical barrier (Masly 2012; Barnard et al. 2017).

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 (Brown and Wilson 1956; Blair 1974; Butlin and Ritchie 1994; Servedio and Noor 2003; Pfennig and Pfennig 2009) 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 correlate with differences in reproductive behaviour? And are these differences sufficiently large to function as prezygotic reproductive barriers, leading to reproductive isolation? 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 species may then result in the production of viable and potentially even fertile hybrid offspring.

Even in the absence of successful hybridization, 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 involved (Gröning and Hochkirch 2008). 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 (Gröning and Hochkirch 2008; Kyogoku 2015; Grether et al. 2017; Shuker and Burdfield-Steel 2017).

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 hermaphrodites (Ladurner et al. 2005), and the recently described *M. janickei*, the currently most closely related congener known (Schärer et

al. 2019). While *M. lignano* has previously been collected from locations in Greece and Italy, *M. janickei* has to date only been collected from France, though the geographic distribution of both species is yet unknown (Zadesenets et al. 2016, 2017; Schärer et al. 2019). 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 premating 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 (Ladurner et al. 2005; Zadesenets et al. 2016, 2017; Schärer et al. 2019). Despite being very closely related sister species (Schärer et al. 2019), the morphology of their stylet is clearly distinct (see Figure 4 and results). *M. lignano* has a stylet that is "slightly curved, its distal opening [having a] slight asymmetric thickening" (Ladurner et al. 2005), 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." (Schärer et al. 2019).

Previous studies have shown that *M. lignano* is an outcrossing, reciprocally copulating species with frequent mating (on average about 6 copulations per hour, Schärer et al. 2004). 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 (Schärer et al. 2004; Vizoso et al. 2010; Schärer et al. 2011), 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 (Marie-Orleach et al., 2013; Schärer et al., 2011; Schärer et al., 2004; Vizoso et al., 2010).

The individuals of *M. lignano* used in this experiment were either from the outbred LS1 culture (Marie-Orleach et al. 2013) or from the transgenic outbred BAS1 culture, which was created by backcrossing the GFP-expressing inbred HUB1 line (Janicke et al. 2013; Marie-Orleach et al. 2014; Wudarski et al. 2017) onto the LS1 culture (Marie-Orleach et al. 2016), subsequently cleaned from a karyotype polymorphism that segregates in HUB1 (Zadesenets et al. 2016, 2017), and finally bred to be homozygous GFP-positive (Vellnow et al. 2018). The LS1 culture is a genetically outbred metapopulation, established from worms collected from a site in Bibione and a site on Isola di Martignano, Italy (Marie-Orleach et al. 2013). The *M. janickei* worms used were from a culture that was established using individuals collected from Palavas-les-Flots, near Montpellier, France (Zadesenets et al. 2016, 2017; Schärer et al. 2019). Both species are kept in mass cultures in the laboratory at 20 °C in glass Petri dishes containing either f/2 medium (Andersen et al. 2007) or 32‰ artificial sea water (ASW) and fed with the diatom *Nitzschia curvilineata*.

Experimental design

Experiment 1: Reproductive behaviour and hybridization

On day 1, for each species, we distributed 240 adult worms over 4 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 chamber (Schärer et al. 2004) was assembled by placing 9 mating pairs (3 pairs of each pairing type) in drops of 3 µl of ASW each between two siliconized microscope slides separated by 257 µm, 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 2h 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 (http://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, Janicke et al. 2011). 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 (Schärer and Ladurner 2003; Singh et al. 2019); 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* (see also Marie-Orleach et al. 2013), 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 temporal pattern in the copulation duration. The copulation behaviour was defined as in Schärer et al. (2004), 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 seconds. 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 duration, copulation interval, time until suck, and suck duration was averaged over all occurrences in a replicate.

The final sample sizes varied for the different behavioural traits, depending on how many replicates exhibited the particular 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, 3 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 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 2 replicates each for the *M. lignano* and *M. janickei* pairs. The final sample sizes are given in the respective figures.

Experiment 2: Hybrid fertility

We assessed the fertility of the F1 hybrid offspring from experiment 1, 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. Note that we cannot distinguish the maternal parent for the offspring produced during the pairing period. 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.

Experiment 3: Hybrid and parental species 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, from Ramm et al. 2019) and *M. janickei* (n=18, from Singh et al. 2019), using a geometric morphometrics landmark-based method (Zelditch et al. 2004). Briefly, 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 400x magnification (Figure 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). 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 (Figure 4d-f), using tpsDig 2.31 (F. James Rohlf, 2006, Department of Ecology and Evolution, SUNY, http://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* (Janicke and Schärer 2009) on account of the different morphology of the *M. janickei* stylet. Specifically, landmarks should represent homologous points on a morphological structure, and 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 Figure 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. We used tpsRelw 1.70 (http://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 (Figure 4g-i).

Experiment 4: 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 in order 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 24h. Patent Blue V does not affect the mating rate of *M. lignano* (Marie-Orleach et al. 2013), or of *M. janickei*, as the mating rate of dyed and undyed worms was similar (see Supplementary Figure S1).

In total, we constructed 17 observation chambers and filmed them under transmitted light for 2h 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 2h 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 experiment 1, 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 $\mathrm{CV} = 100 \times \sqrt{e^{standard\ deviation^2} - 1}$ (Canchola 2017), while for number of sucks we calculated the CV for raw data using $CV = \frac{standard\ deviation}{mean} \times 100$.

In experiment 3, we constructed one-way ANOVAs with the types of worm (*M. lignano*, *M. janickei*, or hybrid) 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 care, since the three groups we compared were not grown and imaged as part of the same experiment (though using the same methodology).

In experiment 4, 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 proportion of drops having these different copulation types as the first copulation 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-fit (McDonald 2014), 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 in order 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 (Figure 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 (Figure 6b).

We, for each drop, 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 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 3.1.1 (R Development Core Team, 2016).

Ethical note

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

Results

Experiment 1: 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; Figure 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; Figure 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; Figure 1c). *M. janickei* pairs had a significantly longer interval between copulations than *M. lignano* pairs, while the heterospecific pairs had intermediate copulation interval.

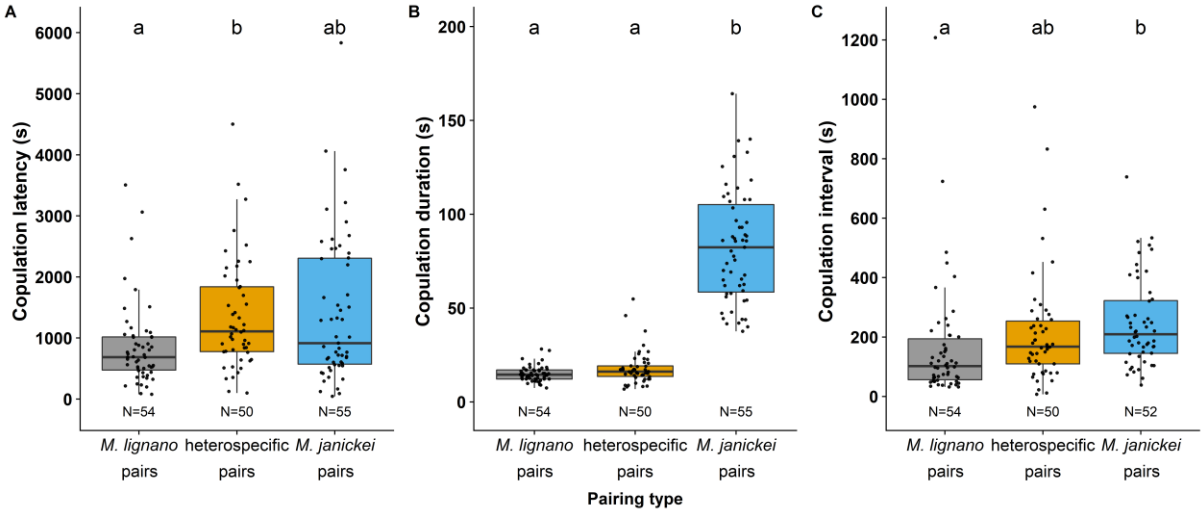


Figure 1. Boxplots of the a) copulation latency, b) (average) copulation duration, and c) (average) copulation interval of the three pairing types. Different letters denote significantly different effects inferred from Tukey HSD post-hoc tests. The boxplots display the 25th percentile, median, and 75th 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.

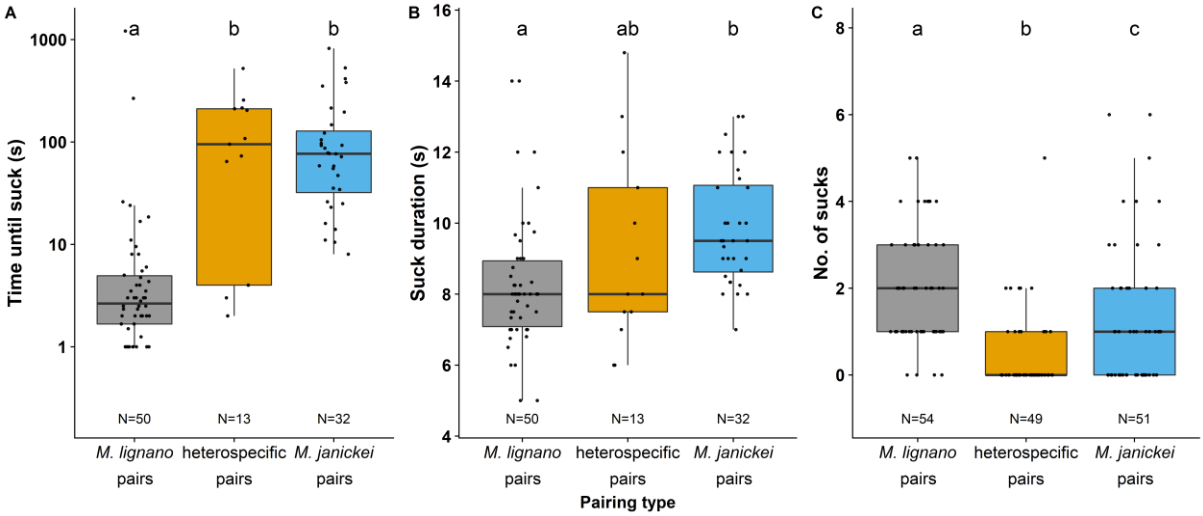


Figure 2. Boxplots of the a) (average) time until suck (after copulation), b) (average) suck duration, and c) number of sucks of the three pairing types (recall that we here only consider pairs that copulated at least 5 times). Different letters denote significantly different effects inferred from Tukey 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.

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 (Figure 2). The time until suck (after copulation) differed between the pairing types ($F_{2,92}$ = 48.15, P < 0.001; Figure 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; Figure 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: χ^2 = 41.16, df = 2, P < 0.001; Figure 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: χ^2 = 48.04, df = 2, P < 0.001; Figure 3a), which had a comparable fecundity. Out of the 10 heterospecific replicates that produced hybrids, in 6 replicates only the *M. lignano* parent produced hybrids while in the other 4 replicates only the *M. janickei* parent produced offspring. Thus, hybridization was symmetrical, with each species being capable of inseminating and fertilizing the other.

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

behaviour	*M. lignano* pairs	heterospecific pairs	*M. janickei* 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

Experiment 2: 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 hybrid and *M. janickei* x hybrid produced hybrid F2 offspring, respectively, while they were paired with an individual of one of their parental species for 7 days (Figure 3b), while post-pairing, relatively few individuals of either hybrids or parentals produced offspring in isolation (Figure 3c).

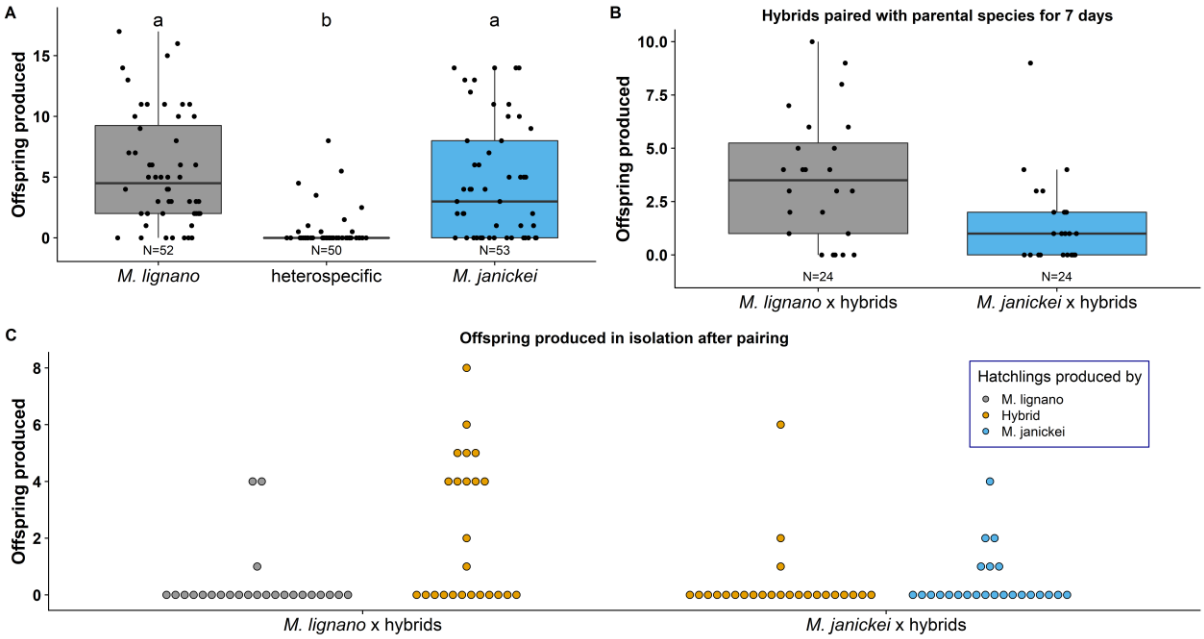


Figure 3. a) Plot of F1 hybrid offspring produced (female fecundity) by the three pairing types in Experiment 1. Plot of F2 hybrid offspring produced b) in the wells where the F1 hybrids were paired with an individual of one of their parental species for 7 days, and c) post-pairing isolated F1 hybrid and parental individuals in Experiment 2. 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, while c) is a dotplot. Note that in c) each backcrossed pair is represented twice as each pair comprises 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).

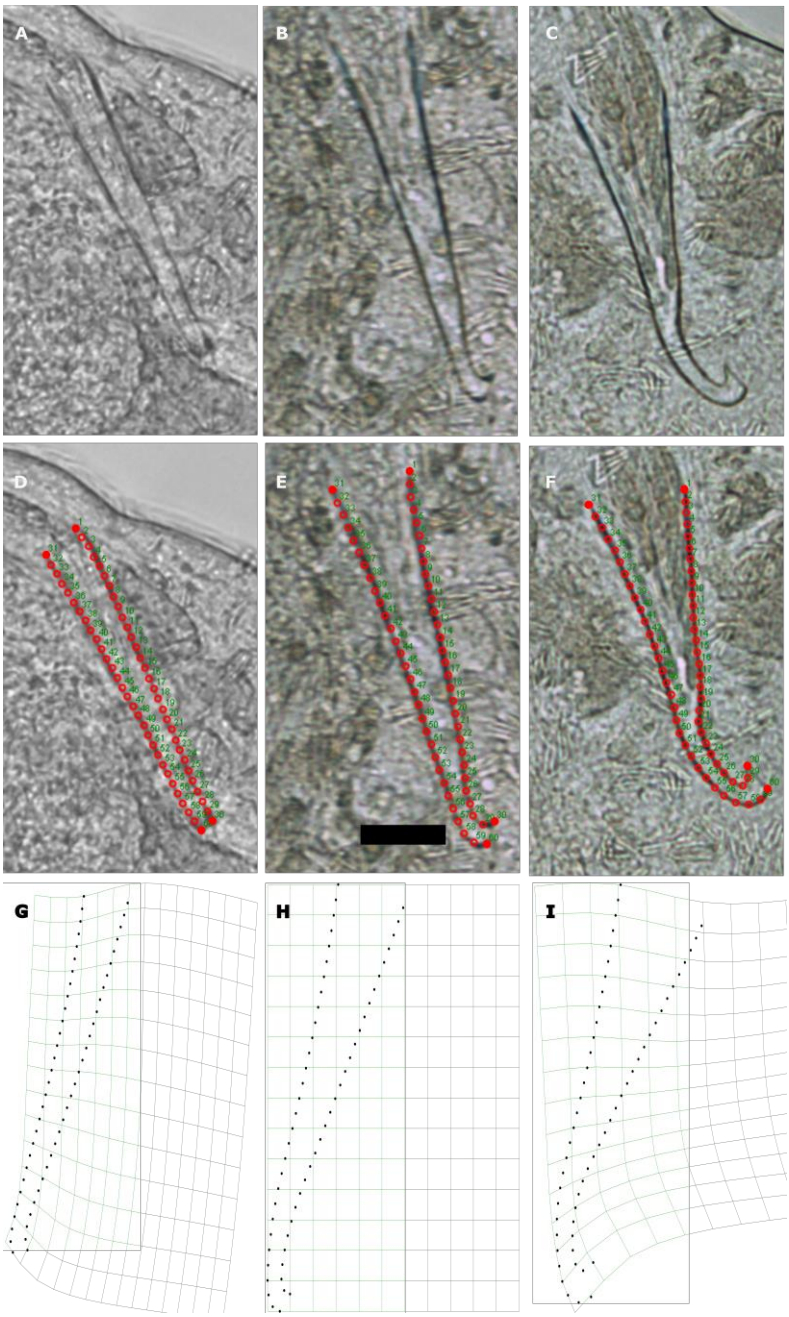


Figure 4. Morphology and geometric morphometrics of the stylet. Micrographs of the stylet of an individual a) *M. lignano*, b) F1 hybrid, and c) *M. janickei*. The placement of 60 landmarks along the stylet of an individual d) *M. lignano*, e) F1 hybrid and f) *M. janickei*. Note that we placed four 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 for the mean relative warp score 1 (RWS1) of 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.

Experiment 3: Hybrid and parental species stylet morphology

The stylet morphology was significantly different between *M. lignano*, *M. janickei* and the F1 hybrids (Figure 4). The centroid size, an estimate of stylet size, was different between the groups ($F_{2,69} = 33.26$, $P < 0.001$; Figure 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 (Figure 4g-i), was significantly different between all groups ($F_{2,69} = 238$, $P < 0.001$; Figure 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.

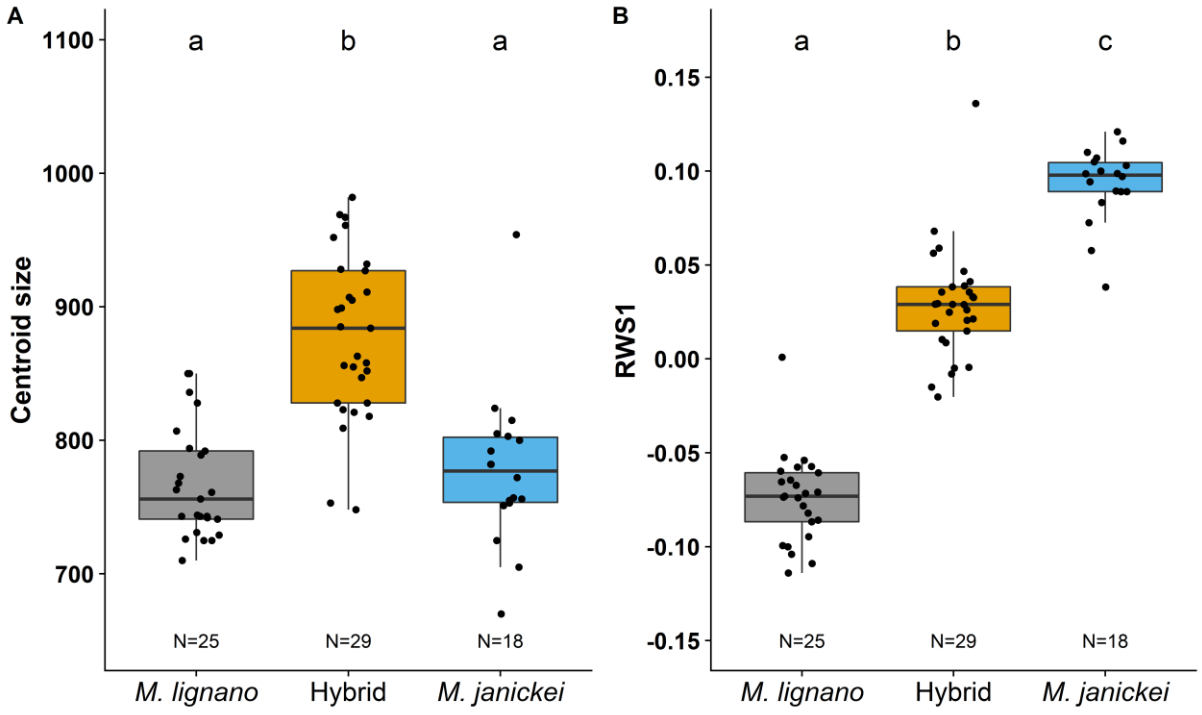


Figure 5 Boxplot for a) centroid size and b) relative warp score 1 (RWS1) of the stylets of *M. lignano*, F1 hybrid and *M. janickei* worms. Different letters denote significantly different effects inferred from Tukey HSD post-hoc tests. The boxplots display the 25th percentile, median, and 75th percentile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

Experiment 4: Mate preference experiment

Out of the 59 analysed drops, we found that 34 (57.6%) drops had a *M. lignano* conspecific copulation as the first copulation, while that was true for only 18 (30.5%) and 7 (11.9%) drops for heterospecifics and *M. janickei* conspecifics, respectively. These proportions differed significantly (Chi-square goodness-of-fit test: $\chi^2 = 33.68$, df = 2, $P < 0.001$) from our null hypothesis proportions under random mating, 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 (without Bonferroni-correction $P < 0.05$, Supplementary Table S2) differed significantly from the null hypothesis, though after Bonferroni correction that number dropped to just 46 drops (Bonferroni-corrected $P < 0.05$, Supplementary Table S2). 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 Figure 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 Figure 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* (Figure 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 S3). 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 by the intrinsic differences in mating rate of the two species, rather than stemming from an explicit preference for heterospecific partners.

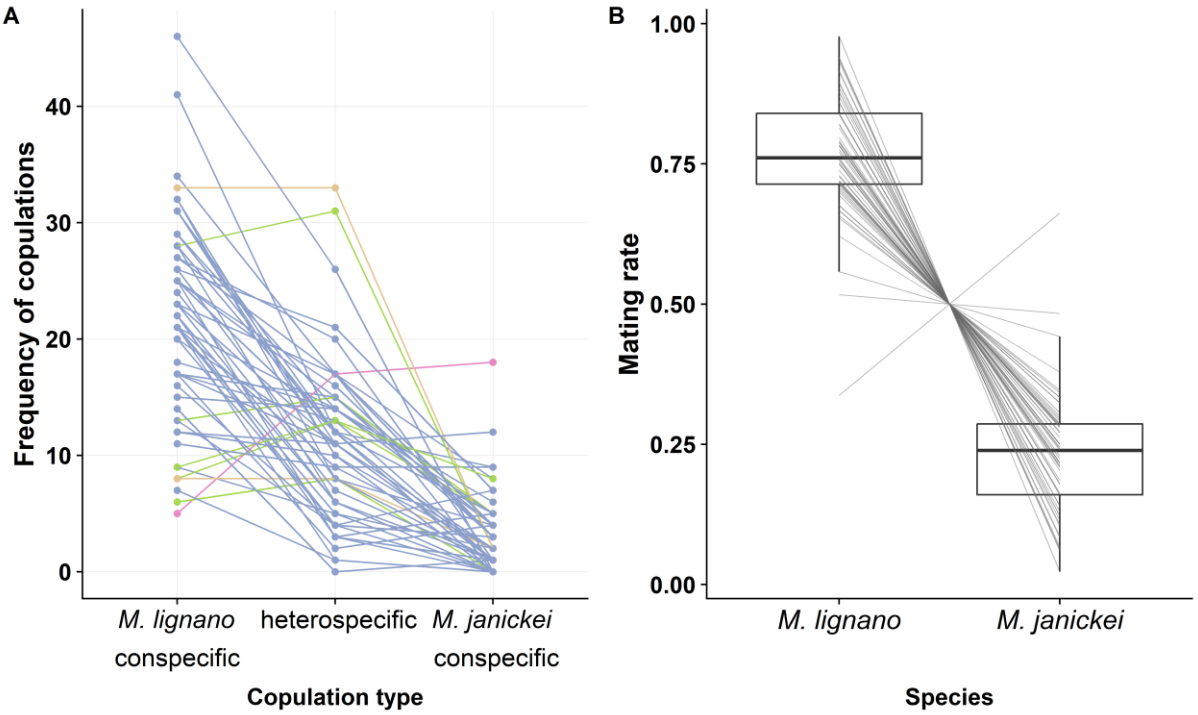


Figure 6. a) Frequency of *M. lignano* conspecific, heterospecific, and *M. janickei* conspecific copulations. Each line connects values obtained from the same drop. The different colours help to visualise which copulation type had the highest frequency in a drop (blue, *M. lignano* conspecific; green, 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.

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 few hybrid offspring were produced from these heterospecific matings. Moreover, the resulting hybrids were fertile, showing a stylet morphology that was

intermediate between the parental species, and 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* clearly 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.

Experiment 1: 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 natricine snakes (King et al. 2009), 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* (Schärer et al. 2019).

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 (Schärer et al. 2004; Vizoso et al. 2010). 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 1.6x 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 partners (Marie-Orleach et al. 2013). 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 sucking (Marie-Orleach et al. 2013). 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 (Chapman 2001; Perry et al. 2013; Schwenke et al. 2016; Billeter and Wolfner 2018). Efforts to elucidate the function of ejaculate components (like seminal-fluid proteins) in *M. lignano* have recently made considerable progress (Weber et al. 2018; Patlar et al. 2019; Ramm et al. 2019), and it will be interesting to see if these have similar functions.

Longer copulation intervals or temporal aspects of sucking (i.e. onset of sucking) could potentially also result from the action of some transferred ejaculate components that acts as a relaxant, 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 products (Lodi and Koene 2016).

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 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 *M. lignano* individuals, while the longer and delayed sucks were performed by *M. janickei* individuals. 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* (Schärer et al. 2004).

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 (Pala et al. 1982; Bullini 1985), while there is some more information about parasitic flatworms (Taylor 1970; Thèron 1989; Detwiler and Criscione 2010; Itagaki et al. 2011; Henrich et al. 2013).

Experiment 2: Hybrid fertility

While historically, hybrids have often been considered to be sterile and evolutionary dead-ends (see Mallet 2005), 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 (Mallet 2005; Bonnet et al. 2017), a view that has been reinforced by the widespread presence of allopolyploidy among plants (Soltis and Soltis 1995; Soltis et al. 2015; Wendel et al. 2016). 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. Though our study demonstrates hybridisation between the two species, we currently have no evidence for these species occurring in sympatry. Assuming their so far sampled locations indicates 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. 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 organization (Zadesenets et al. 2016), 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 levels (Zadesenets et al. 2017). Interestingly, individuals with unusual karyotypes do not show behavioural or morphological abnormalities and reproduce successfully, at least in *M. lignano* (Zadesenets et al. 2016). 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.

Experiment 3: Hybrid and parental species stylet morphology

The parental species differed significantly in the morphology of their stylet, though their overall stylet size was similar. In contrast, the hybrids 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 care, since the data used in this comparison stemmed from three separate experiments). 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 hybridize (Barnard et al. 2017).

Experiment 4: Mate preference experiment

Our mate preference experiment showed that there is some degree of assortative mating between *M. lignano* individuals, which appears to mostly stem from the higher mating rate of *M. lignano*. This is in line with our results from Experiment 1, where *M. lignano* conspecific pairs had shorter copulation latencies, shorter copulation durations and shorter copulation intervals compared to *M. janickei* conspecific pairs (Figure 1). Thus, mate choice in these two species seems to be governed mainly by behavioural characteristics, such as mating rate, rather than an explicit preference for a conspecific or heterospecific partner.

A potential factor affecting mating rate 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 (Gröning and

Hochkirch 2008; Burdfield-Steel and Shuker 2011; Kyogoku 2015). 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 (Snook et al. 2005; Debelle et al. 2014). A mate choice experiment showed that males from polyandrous populations had a higher probability of mating than those from monogamous populations (Debelle et al. 2016), potentially due to having evolved under strong male-male competition and hence initiating courtship faster and more frequently than monogamous males (Crudgington et al. 2010). 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 a congener species (*C. maculatus*) than beetles evolved under a monogamous regime (Kyogoku and Sota 2017; Kyogoku et al. 2019). In addition to the above examples, multiple empirical studies have proposed a role of sexual selection in occurrence of reproductive interference between species (Kyogoku and Sota 2015; Yassin and David 2016).

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. 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, though they 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.

Acknowledgements:

We would like to thank Tim Janicke and Georgina Rivera Ingraham for help with collecting specimens of the study species *M. janickei*, Steven A. Ramm for providing access to the images of the *M. 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 grant 31003A-162543 of the Swiss National Science Foundation (SNSF) to LS.

References

Andersen, R. A., J. A. Berges, P. J. Harrison, and M. M. Watanabe. 2007. Recipes for Freshwater and Seawater Media. Pp. 429–538 *in* Algal Culturing Techniques. Elsevier.

Anthes, N., and N. K. Michiels. 2007. Reproductive morphology, mating behavior, and spawning ecology of cephalaspid sea slugs (Aglajidae and Gastropteridae). Invertebr. Biol. 126:335–365.

Arnqvist, G. 1998. Comparative evidence for the evolution of genitalia by sexual selection. Nature 393:784–786.

Arnqvist, G. 1997. The evolution of animal genitalia: Distinguishing between hypotheses by single species studies. Biol. J. Linn. Soc. 60:365–379.

Barnard, A. A., O. M. Fincke, M. A. McPeek, and J. P. Masly. 2017. Mechanical and tactile incompatibilities cause reproductive isolation between two young damselfly species. Evolution (N. Y). 71:2410–2427.

Billeter, J.-C., and M. F. Wolfner. 2018. Chemical Cues that Guide Female Reproduction in *Drosophila melanogaster*. J. Chem. Ecol. 44:750–769.

Blair, W. F. 1974. Character Displacement in Frogs. Am. Zool. 14:1119–1125.

Bonnet, T., R. Leblois, F. Rousset, and P. A. Crochet. 2017. A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. Evolution (N. Y). 71:2140–2158.

Brown, W. L., and E. O. Wilson. 1956. Character Displacement. Syst. Zool. 5:49.

Bullini, L. 1985. Speciation by hybridization in animals. Bolletino di Zool. 52:121–137.

Burdfield-Steel, E. R., and D. M. Shuker. 2011. Reproductive interference. Curr. Biol. 21:R450–R451.

Butlin, R., A. Debelle, C. Kerth, R. R. Snook, L. W. Beukeboom, R. F. Castillo Cajas, W. Diao, M. E. Maan, S. Paolucci, F. J. Weissing, L. van de Zande, A. Hoikkala, E. Geuverink, J. Jennings, M. Kankare, K. E. Knott, V. I. Tyukmaeva, C. Zoumadakis, M. G. Ritchie, D. Barker, E. Immonen, M. Kirkpatrick, M. Noor, M. C. Garcia, T. Schmitt, and M. Schilthuizen. 2012. What do we need to know about speciation? Trends Ecol. Evol. 27:27–39.

Butlin, R. K., and M. G. Ritchie. 1994. Mating behaviour and speciation. Pp. 43–79 *in* P. Slater and T. Halliday, eds. Behaviour and evolution. Cambridge University Press, Cambridge.

Canchola, J. A. 2017. Correct use of percent coefficient of variation (%CV) formula for log-transformed data. MOJ Proteomics Bioinforma., doi: 10.15406/mojpb.2017.06.00200.

Chapman, T. 2001. Seminal fluid-mediated fitness traits in *Drosophila*. Heredity (Edinb). 87:511–521.

Ciccotto, P. J., J. M. Gumm, and T. C. Mendelson. 2013. Male association preference for conspecifics in the redband darter, *Etheostoma luteovinctum* (Teleostei: Percidae) based on visual cues. Copeia 2013:154–159.

Coyne, J. A., and H. A. Orr. 2004. Speciation. Massachusetts: Sinauer Associates Sunderland.

Crudgington, H. S., S. Fellows, and R. R. Snook. 2010. Increased opportunity for sexual conflict promotes harmful males with elevated courtship frequencies. J. Evol. Biol. 23:440–446.

Debelle, A., M. G. Ritchie, and R. R. Snook. 2014. Evolution of divergent female mating preference in response to experimental sexual selection. Evolution (N. Y). 68:2524–2533.

Debelle, A., M. G. Ritchie, and R. R. Snook. 2016. Sexual selection and assortative mating: an experimental test. J. Evol. Biol. 29:1307–1316.

Detwiler, J. T., and C. D. Criscione. 2010. An infectious topic in reticulate evolution:

introgression and hybridization in animal parasites. Genes (Basel). 1:102–123.
Devigili, A., J. L. Fitzpatrick, C. Gasparini, I. W. Ramnarine, A. Pilastro, and J. P. Evans. 2018. 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.
Eberhard, W. G. 1985. Sexual Selection and Animal Genitalia. Harvard University Press, Cambridge, MA and London, England.
Firman, R. C., C. Gasparini, M. K. Manier, and T. Pizzari. 2017. Postmating Female Control: 20 Years of Cryptic Female Choice. Trends Ecol. Evol. 32:368–382.
Garlovsky, M. D., and R. R. Snook. 2018. Persistent postmating, prezygotic reproductive isolation between populations. Ecol. Evol. 8:9062–9073.
Gleason, J. M., and M. G. Ritchie. 1998. Evolution of Courtship Song and Reproductive Isolation in the *Drosophila willistoni* Species Complex: Do Sexual Signals Diverge the Most Quickly? Evolution (N. Y). 52:1493.
Grether, G. F., K. S. Peiman, J. A. Tobias, and B. W. Robinson. 2017. Causes and Consequences of Behavioral Interference between Species. Trends Ecol. Evol. 32:760–772.
Gröning, J., and A. Hochkirch. 2008. Reproductive Interference Between Animal Species. Q. Rev. Biol. 83:257–282.
Henrich, T., D. P. Benesh, and M. Kalbe. 2013. Hybridization between two cestode species and its consequences for intermediate host range. Parasit. Vectors 6:33.
Herring, K., and P. Verrell. 1996. Sexual incompatibility and geographical variation in mate recognition systems: tests in the salamander *Desmognathus ochrophaeus*. Anim. Behav. 52:279–287.
Itagaki, T., M. Ichinomiya, K. Fukuda, S. Fusyuku, and C. Carmona. 2011. Hybridization experiments indicate incomplete reproductive isolating mechanism between *Fasciola hepatic*a and *Fasciola gigantica*. Parasitology 138:1278–1284.
Janicke, T., L. Marie-Orleach, K. De Mulder, E. Berezikov, P. Ladurner, D. B. Vizoso, and L. Schärer. 2013. Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution (N. Y). 67:3233–3242.
Janicke, T., M. G. Ritchie, E. H. Morrow, and L. Marie-Orleach. 2018. Sexual selection predicts species richness across the animal kingdom. Proc. R. Soc. B Biol. Sci. 285:20180173.
Janicke, T., P. Sandner, and L. Schärer. 2011. Determinants of female fecundity in a simultaneous hermaphrodite: the role of polyandry and food availability. Evol. Ecol. 25:203–218.
Janicke, T., and L. Schärer. 2009. Determinants of mating and sperm-transfer success in a simultaneous hermaphrodite. J. Evol. Biol. 22:405–415.
Kelly, D. A., and B. C. Moore. 2016. The Morphological Diversity of Intromittent Organs: An Introduction to the Symposium. Integr. Comp. Biol. 56:630–634.
King, R. B., R. C. Jadin, M. Grue, and H. D. Walley. 2009. Behavioural correlates with hemipenis morphology in New World natricine snakes. Biol. J. Linn. Soc. 98:110–120.
Klappert, K., D. Mazzi, A. Hoikkala, and M. G. Ritchie. 2007. Male courtship song and female preference variation between phylogeographically distinct populations of *Drosophila montana*. Evolution (N. Y). 61:1481–1488.
Kraaijeveld, K., F. J. L. Kraaijeveld-Smit, and M. E. Maan. 2011. Sexual selection and speciation: the comparative evidence revisited. Biol. Rev. 86:367–377.
Kyogoku, D. 2015. Reproductive interference: ecological and evolutionary consequences of interspecific promiscuity. Popul. Ecol. 57:253–260.
Kyogoku, D., M. Kondoh, and T. Sota. 2019. Does past evolutionary history under different

mating regimes influence the demographic dynamics of interspecific competition? Ecol. Evol. 9:8616–8624.
Kyogoku, D., and T. Sota. 2015. Exaggerated male genitalia intensify interspecific reproductive interference by damaging heterospecific female genitalia. J. Evol. Biol. 28:1283–1289.
Kyogoku, D., and T. Sota. 2017. The evolution of between-species reproductive interference capability under different within-species mating regimes. Evolution (N. Y). 71:2721–2727.
Lackey, A. C. R., and J. W. Boughman. 2017. Evolution of reproductive isolation in stickleback fish. Evolution (N. Y). 71:357–372.
Ladurner, P., L. Schärer, W. Salvenmoser, and R. M. Rieger. 2005. 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.
Lodi, M., and J. M. Koene. 2016. On the effect specificity of accessory gland products transferred by the love-dart of land snails. BMC Evol. Biol. 16:104.
Mallet, J. 2005. Hybridization as an invasion of the genome. Trends Ecol. Evol. 20:229–237.
Manier, M. K., S. Lüpold, J. M. Belote, W. T. Starmer, K. S. Berben, O. Ala-Honkola, W. F. Collins, and S. Pitnick. 2013. Postcopulatory sexual selection generates speciation phenotypes in *Drosophila*. Curr. Biol. 23:1853–1862.
Marie-Orleach, L., T. Janicke, and L. Schärer. 2013. Effects of mating status on copulatory and postcopulatory behaviour in a simultaneous hermaphrodite. Anim. Behav. 85:453–461.
Marie-Orleach, L., T. Janicke, D. B. Vizoso, P. David, and L. Schärer. 2016. Quantifying episodes of sexual selection: Insights from a transparent worm with fluorescent sperm. Evolution (N. Y). 70:314–328.
Marie-Orleach, L., T. Janicke, D. B. Vizoso, M. Eichmann, and L. Schärer. 2014. Fluorescent sperm in a transparent worm: validation of a GFP marker to study sexual selection. BMC Evol. Biol. 14:148.
Masly, J. P. 2012. 170 Years of "Lock-and-Key": Genital Morphology and Reproductive Isolation. Int. J. Evol. Biol. 2012:1–10.
Mayr, E. 1942. Systematics and the Origin of Species. New York: Columbia University Press.
McDonald, J. H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publ. Balt. Maryl., doi: 10.1017/CBO9781107415324.004.
Morrow, E. H., T. E. Pitcher, and G. Arnqvist. 2003. No evidence that sexual selection is an "engine of speciation" in birds. Ecol. Lett. 6:228–234.
Mouton, S., M. Willems, B. P. Braeckman, B. Egger, P. Ladurner, L. Schärer, and G. Borgonie. 2009. The free-living flatworm *Macrostomum lignano*: A new model organism for ageing research. Exp. Gerontol. 44:243–249. Elsevier Inc.
Ostevik, K. L., R. L. Andrew, S. P. Otto, and L. H. Rieseberg. 2016. Multiple reproductive barriers separate recently diverged sunflower ecotypes. Evolution (N. Y). 70:2322–2335.
Pala, M., S. Casu, and N. G. Lepori. 1982. Stabilized natural interspecific hybrid population of the fresh water planarians *Dugesia Gonocephala* S. L. (Turbellaria, Tricladida). Caryologia 35:247–256.
Patlar, B., M. Weber, and S. A. Ramm. 2019. Genetic and environmental variation in transcriptional expression of seminal fluid proteins. Heredity (Edinb). 122:595–611.
Perry, J. C., L. Sirot, and S. Wigby. 2013. The seminal symphony: how to compose an ejaculate. Trends Ecol. Evol. 28:414–422.
Pfennig, K. S., and D. W. Pfennig. 2009. Character displacement: ecological and reproductive responses to a common evolutionary problem. Q. Rev. Biol. 84:253–76.

Puniamoorthy, N. 2014. Behavioural barriers to reproduction may evolve faster than sexual morphology among populations of a dung fly (Sepsidae). Anim. Behav. 98:139–148. Elsevier Ltd.

Puniamoorthy, N., M. R. B. Ismail, D. S. H. Tan, and R. Meier. 2009. 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. 22:2146–56.

Puniamoorthy, N., M. Kotrba, and R. Meier. 2010. Unlocking the "Black box": internal female genitalia in Sepsidae (Diptera) evolve fast and are species-specific. BMC Evol. Biol. 10:275.

Ramm, S. A., B. Lengerer, R. Arbore, R. Pjeta, J. Wunderer, A. Giannakara, E. Berezikov, P. Ladurner, and L. Schärer. 2019. Sex allocation plasticity on a transcriptome scale: Socially sensitive gene expression in a simultaneous hermaphrodite. Mol. Ecol. mec.15077.

Ritchie, M. G., E. J. Halsey, and J. M. Gleason. 1999. *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.

Sato, Y., H. Sakamoto, T. Gotoh, Y. Saito, J.-T. Chao, M. Egas, and A. Mochizuki. 2018. 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.

Schärer, L., J. N. Brand, P. Singh, K. S. Zadesenets, C.-P. Stelzer, and G. Viktorin. 2019. 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. jzs.12344.

Schärer, L., G. Joss, and P. Sandner. 2004. Mating behaviour of the marine turbellarian Macrostomum sp.: these worms suck. Mar. Biol. 145:373–380.

Schärer, L., and P. Ladurner. 2003. Phenotypically plastic adjustment of sex allocation in a simultaneous hermaphrodite. Proc. R. Soc. B Biol. Sci. 270:935–941.

Schärer, L., D. T. J. Littlewood, A. Waeschenbach, W. Yoshida, and D. B. Vizoso. 2011. Mating behavior and the evolution of sperm design. Proc. Natl. Acad. Sci. 108:1490–1495.

Schwenke, R. A., B. P. Lazzaro, and M. F. Wolfner. 2016. Reproduction–Immunity Trade-Offs in Insects. Annu. Rev. Entomol. 61:239–256.

Servedio, M. R., and M. A. F. Noor. 2003. The Role of Reinforcement in Speciation: Theory and Data. Annu. Rev. Ecol. Evol. Syst. 34:339–364.

Shuker, D. M., and E. R. Burdfield-Steel. 2017. Reproductive interference in insects. Ecol. Entomol. 42:65–75.

Singh, P., N. Vellnow, and L. Schärer. 2019. Variation in sex allocation plasticity in three closely related flatworm species. Ecol. Evol. ece3.5566.

Snook, R. R., A. Robertson, H. S. Crudgington, and M. G. Ritchie. 2005. Experimental Manipulation of Sexual Selection and the Evolution of Courtship Song in *Drosophila pseudoobscura*. Behav. Genet. 35:245–255.

Soltis, D. E., and P. S. Soltis. 1995. The dynamic nature of polyploid. Proc. Natl. Acad. Sci. 92:8089–8091.

Soltis, P. S., D. B. Marchant, Y. Van de Peer, and D. E. Soltis. 2015. Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. 35:119–125.

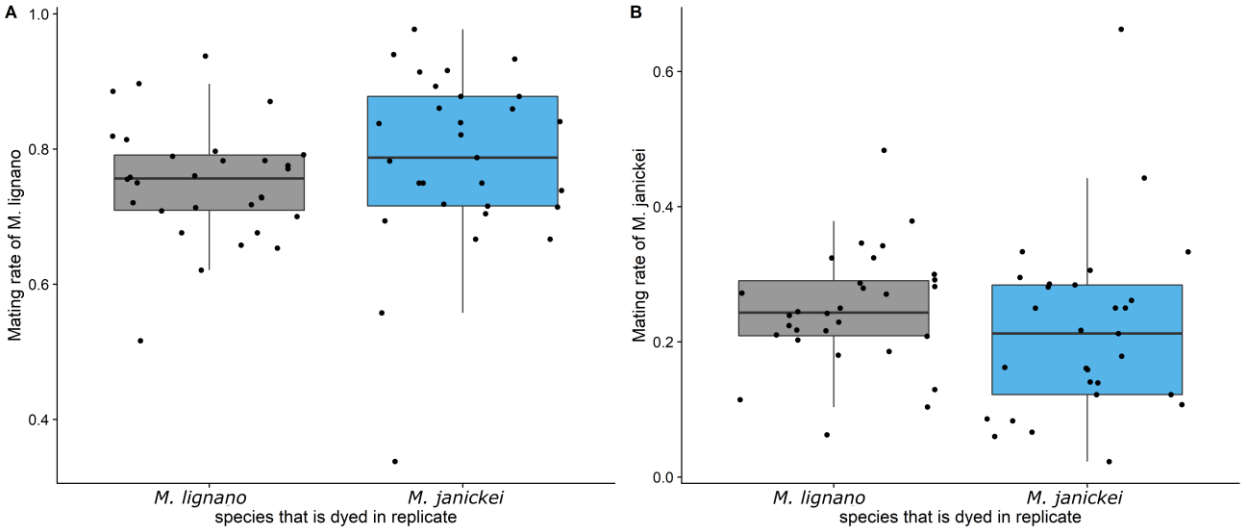
Soudi, S., K. Reinhold, and L. Engqvist. 2016. Strong cryptic prezygotic isolation despite lack of behavioral isolation between sympatric host races of the leaf beetle *Lochmaea capreae*. Evolution (N. Y). 70:2889–2898.

Swanson, W. J., and V. D. Vacquier. 2002. The rapid evolution of reproductive proteins. Nat.

Rev. Genet. 3:137–144.
Taylor, M. G. 1970. Hybridisation Experiments on Five Species of African *Schistosomes*. J. Helminthol. 44:253–314.
Thèron, A. 1989. Hybrids between *Schistosoma mansoni* and *S. rodhaini* : characterization by cercarial emergence rhythms. Parasitology 99:225–228.
Turissini, D. A., J. A. McGirr, S. S. Patel, J. R. David, and D. R. Matute. 2018. The rate of evolution of postmating-prezygotic reproductive isolation in *Drosophila*. Mol. Biol. Evol. 35:312–334.
Vellnow, N., L. Marie-Orleach, K. S. Zadesenets, and L. Schärer. 2018. Bigger testes increase paternity in a simultaneous hermaphrodite, independently of the sperm competition level. J. Evol. Biol. 31:180–196.
Vizoso, D. B., G. Rieger, and L. Schärer. 2010. Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. Biol. J. Linn. Soc. 99:370–383.
Weber, M., J. Wunderer, B. Lengerer, R. Pjeta, M. Rodrigues, L. Schärer, P. Ladurner, and S. A. Ramm. 2018. A targeted in situ hybridization screen identifies putative seminal fluid proteins in a simultaneously hermaphroditic flatworm. BMC Evol. Biol. 18:81.
Wendel, J. F., S. A. Jackson, B. C. Meyers, and R. A. Wing. 2016. Evolution of plant genome architecture. Genome Biol. 17:37.
Williams, T. H., and T. C. Mendelson. 2010. Behavioral Isolation Based on Visual Signals in a Sympatric Pair of Darter Species. Ethology 116:1038–1049.
Wudarski, J., D. Simanov, K. Ustyantsev, K. de Mulder, M. Grelling, M. Grudniewska, F. Beltman, L. Glazenburg, T. Demircan, J. Wunderer, W. Qi, D. B. Vizoso, P. M. Weissert, D. Olivieri, S. Mouton, V. Guryev, A. Aboobaker, L. Schärer, P. Ladurner, and E. Berezikov. 2017. Efficient transgenesis and annotated genome sequence of the regenerative flatworm model *Macrostomum lignano*. Nat. Commun. 8:2120.
Yassin, A., and J. R. David. 2016. Within-species reproductive costs affect the asymmetry of satyrization in *Drosophila*. J. Evol. Biol. 29:455–460.
Zadesenets, K. S., L. Schärer, and N. B. Rubtsov. 2017. New insights into the karyotype evolution of the free-living flatworm *Macrostomum lignano* (Platyhelminthes, Turbellaria). Sci. Rep. 7:6066.
Zadesenets, K. S., D. B. Vizoso, A. Schlatter, I. D. Konopatskaia, E. Berezikov, L. Schärer, and N. B. Rubtsov. 2016. Evidence for Karyotype Polymorphism in the Free-Living Flatworm, *Macrostomum lignano*, a Model Organism for Evolutionary and Developmental Biology. PLoS One 11:e0164915.
Zelditch, M. L., D. L. Swiderski, H. D. Sheets, and W. L. Fink. 2004. Geometric morphometrics for biologists: a primer. Elsevier Acad. Press San Diego.
Zhou, M., E. R. Loew, and R. C. Fuller. 2015. Sexually asymmetric colour-based species discrimination in orangethroat darters. Anim. Behav. 106:171–179.

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

S1. Effect of dye on mating rate of a) *M. lignano* and b) *M. janickei*.



Welch Two Sample t-test:

1) Effect of dye on *M. lignano* (comparing replicates in which *M. lignano* was dyed with replicates in which *M. lignano* was not dyed)

t = -1.0923, df = 48.152, p-value = 0.2801

2) Effect of dye on *M. janickei* (comparing replicates in which *M. janickei* was dyed with replicates in which *M. janickei* was not dyed)

t = 1.0923, df = 48.152, p-value = 0.2801

S2. Number of copulations of different types for each replicate, and the associated G-tests of goodness-of-fit with Bonferroni correction for multiple testing.

Movie	Drop no.	Total no. of copulations in drop	No. of copulations of *M. lignano* conspecific type	No. of copulations of heterospecific type	No. of copulations of *M. janickei* conspecific type	G	df	P	P (corrected for multiple testing)
A	1	20	13	4	3	14.45	2	0.001	0.043
A	2	14	9	5	0	13.64	2	0.001	0.065
A	3	44	32	10	2	45.75	2	<0.001	<0.001
A	4	39	21	14	4	15.82	2	<0.001	0.022
C	1	35	17	15	3	11.53	2	0.003	0.185
C	2	49	28	12	9	23.61	2	<0.001	<0.001
C	3	40	5	17	18	8.70	2	0.013	0.760
C	4	59	34	17	8	28.26	2	<0.001	<0.001
D	1	14	6	8	0	8.60	2	0.014	0.799
D	2	61	28	31	2	26.91	2	<0.001	<0.001
D	3	48	27	20	1	31.53	2	<0.001	<0.001
E	1	33	13	15	5	3.96	2	0.138	1.000
E	2	40	23	17	0	32.79	2	<0.001	<0.001
E	3	26	8	13	5	0.70	2	0.705	1.000
E	4	38	18	14	6	8.94	2	0.011	0.675
F	1	38	26	8	4	31.59	2	<0.001	<0.001
F	2	18	8	8	2	4.08	2	0.130	1.000
F	3	33	15	14	4	7.54	2	0.023	1.000
F	4	39	21	9	9	16.87	2	<0.001	0.013
G	1	48	29	14	5	27.33	2	<0.001	<0.001
G	3	32	20	11	1	24.25	2	<0.001	<0.001
G	4	32	21	4	7	27.57	2	<0.001	<0.001
H	1	31	17	13	1	18.04	2	<0.001	0.007
H	2	44	25	15	4	21.47	2	<0.001	0.001
H	4	68	33	33	2	33.25	2	<0.001	<0.001
I	1	37	25	12	0	39.32	2	<0.001	<0.001
I	2	30	22	3	5	33.64	2	<0.001	<0.001
I	3	54	31	11	12	28.96	2	<0.001	<0.001
I	4	25	22	3	0	46.81	2	<0.001	<0.001
J	1	44	23	17	4	17.07	2	<0.001	0.012
J	3	54	26	21	7	14.33	2	0.001	0.046
J	4	15	14	0	1	34.24	2	<0.001	<0.001
K	1	72	46	26	0	69.40	2	<0.001	<0.001
K	3	43	31	12	0	51.67	2	<0.001	<0.001
K	4	29	23	6	0	42.52	2	<0.001	<0.001
L	1	46	29	12	5	29.70	2	<0.001	<0.001
L	2	24	20	4	0	39.37	2	<0.001	<0.001

L	3	41	31	10	0	54.26	2	<0.001	<0.001
L	4	30	9	13	8	0.59	2	0.743	1.000
M	1	22	11	9	2	7.59	2	0.022	1.000
M	2	23	16	4	3	20.40	2	<0.001	0.002
M	3	47	27	17	3	25.73	2	<0.001	<0.001
M	4	23	17	2	4	26.96	2	<0.001	<0.001
N	1	35	28	6	1	47.95	2	<0.001	<0.001
N	2	20	13	4	3	14.45	2	0.001	0.043
N	3	41	32	8	1	53.15	2	<0.001	<0.001
N	4	29	17	11	1	18.94	2	<0.001	0.005
O	1	22	21	1	0	51.47	2	<0.001	<0.001
O	2	35	23	11	1	29.90	2	<0.001	<0.001
O	3	8	7	1	0	14.77	2	0.001	0.037
O	4	28	22	6	0	40.22	2	<0.001	<0.001
P	1	29	25	3	1	48.48	2	<0.001	<0.001
P	2	24	12	10	2	8.59	2	0.014	0.803
P	3	24	12	11	1	11.14	2	0.004	0.225
P	4	44	25	16	3	23.06	2	<0.001	0.001
Q	1	54	41	12	1	66.42	2	<0.001	<0.001
Q	2	32	24	7	1	37.00	2	<0.001	<0.001
Q	3	28	21	5	2	30.83	2	<0.001	<0.001
Q	4	36	23	11	2	26.31	2	<0.001	<0.001

S3. Observed and expected number of copulations of different types and mating rate of each species for each replicate, and the associated Chi-square goodness-of-fit test with Bonferroni correction for multiple testing.

Movie name	Drop no.	Total no. of copulations in drop of *M. lignano*	Total no. of copulations in drop of *M. janickei*	observed mating rate of *M. lignano* (p)	observed mating rate of *M. janickei* (q)	Expected no. of copulations of *M. lignano* conspecific type	Expected no. of copulations of heterospecific type	Expected no. of copulations of *M. janickei* conspecific type	chi-square	df	P	P (corrected for multiple testing)
A	1	30	10	0.75	0.25	11.25	7.5	1.25	4.36	2	0.11	1
A	2	23	5	0.82	0.18	9.45	4.11	0.45	0.66	2	0.72	1
A	3	74	14	0.84	0.16	31.11	11.77	1.11	1	2	0.61	1
A	4	56	22	0.72	0.28	20.1	15.79	3.1	0.5	2	0.78	1
C	1	49	21	0.7	0.3	17.15	14.7	3.15	0.01	2	0.99	1
C	2	68	30	0.69	0.31	23.59	20.82	4.59	8.79	2	0.01	0.73
C	3	27	53	0.34	0.66	4.56	17.89	17.56	0.1	2	0.95	1
C	4	85	33	0.72	0.28	30.61	23.77	4.61	4.79	2	0.09	1
D	1	20	8	0.71	0.29	7.14	5.71	1.14	2.24	2	0.33	1
D	2	87	35	0.71	0.29	31.02	24.96	5.02	3.57	2	0.17	1
D	3	74	22	0.77	0.23	28.52	16.96	2.52	1.54	2	0.46	1
E	1	41	25	0.62	0.38	12.73	15.53	4.73	0.04	2	0.98	1
E	2	63	17	0.79	0.21	24.81	13.39	1.81	2.91	2	0.23	1
E	3	29	23	0.56	0.44	8.09	12.83	5.09	0	2	1	1
E	4	50	26	0.66	0.34	16.45	17.11	4.45	1.25	2	0.53	1
F	1	60	16	0.79	0.21	23.68	12.63	1.68	5.11	2	0.08	1
F	2	24	12	0.67	0.33	8	8	2	0	2	1	1
F	3	44	22	0.67	0.33	14.67	14.67	3.67	0.07	2	0.97	1
F	4	51	27	0.65	0.35	16.67	17.65	4.67	9.37	2	0.01	0.54
G	1	72	24	0.75	0.25	27	18	3	2.37	2	0.31	1
G	3	51	13	0.8	0.2	20.32	10.36	1.32	0.12	2	0.94	1
G	4	46	18	0.72	0.28	16.53	12.94	2.53	15.27	2	0.0004	0.03
H	1	47	15	0.76	0.24	17.81	11.37	1.81	0.64	2	0.73	1
H	2	65	23	0.74	0.26	24.01	16.99	3.01	0.6	2	0.74	1
H	4	99	37	0.73	0.27	36.03	26.93	5.03	3.45	2	0.18	1
I	1	62	12	0.84	0.16	25.97	10.05	0.97	1.39	2	0.5	1
I	2	47	13	0.78	0.22	18.41	10.18	1.41	14.93	2	0.0006	0.03
I	3	73	35	0.68	0.32	24.67	23.66	5.67	15.46	2	0.0004	0.03
I	4	47	3	0.94	0.06	22.09	2.82	0.09	0.1	2	0.95	1
J	1	63	25	0.72	0.28	22.55	17.9	3.55	0.11	2	0.95	1
J	3	73	35	0.68	0.32	24.67	23.66	5.67	0.68	2	0.71	1
J	4	28	2	0.93	0.07	13.07	1.87	0.07	15	2	0.0005	0.03
K	1	118	26	0.82	0.18	48.35	21.31	2.35	3.5	2	0.17	1
K	3	74	12	0.86	0.14	31.84	10.33	0.84	1.13	2	0.57	1
K	4	52	6	0.9	0.1	23.31	5.38	0.31	0.39	2	0.82	1
L	1	70	22	0.76	0.24	26.63	16.74	2.63	3.69	2	0.16	1
L	2	44	4	0.92	0.08	20.17	3.67	0.17	0.2	2	0.91	1

L	3	72	10	0.88	0.12	31.61	8.78	0.61	0.79	2	0.67	1
L	4	31	29	0.52	0.48	8.01	14.98	7.01	0.53	2	0.77	1
M	1	31	13	0.7	0.3	10.92	9.16	1.92	0.01	2	1	1
M	2	36	10	0.78	0.22	14.09	7.83	1.09	5.5	2	0.06	1
M	3	71	23	0.76	0.24	26.81	17.37	2.81	0.02	2	0.99	1
M	4	36	10	0.78	0.22	14.09	7.83	1.09	12.75	2	0.002	0.1
N	1	62	8	0.89	0.11	27.46	7.09	0.46	0.82	2	0.66	1
N	2	30	10	0.75	0.25	11.25	7.5	1.25	4.36	2	0.11	1
N	3	72	10	0.88	0.12	31.61	8.78	0.61	0.32	2	0.85	1
N	4	45	13	0.78	0.22	17.46	10.09	1.46	0.24	2	0.89	1
O	1	43	1	0.98	0.02	21.01	0.98	0.01	0.01	2	0.99	1
O	2	57	13	0.81	0.19	23.21	10.59	1.21	0.05	2	0.97	1
O	3	15	1	0.94	0.06	7.03	0.94	0.03	0.04	2	0.98	1
O	4	50	6	0.89	0.11	22.32	5.36	0.32	0.4	2	0.82	1
P	1	53	5	0.91	0.09	24.22	4.57	0.22	3.42	2	0.18	1
P	2	34	14	0.71	0.29	12.04	9.92	2.04	0	2	1	1
P	3	35	13	0.73	0.27	12.76	9.48	1.76	0.62	2	0.73	1
P	4	66	22	0.75	0.25	24.75	16.5	2.75	0.04	2	0.98	1
Q	1	94	14	0.87	0.13	40.91	12.19	0.91	0.01	2	0.99	1
Q	2	55	9	0.86	0.14	23.63	7.73	0.63	0.29	2	0.87	1
Q	3	47	9	0.84	0.16	19.72	7.55	0.72	3.2	2	0.2	1
Q	4	57	15	0.79	0.21	22.56	11.88	1.56	0.2	2	0.91	1

Chapter II

Evolution of the suck behaviour, a postcopulatory female resistance trait in a hermaphroditic flatworm genus

Singh, P., J. N. Brand, and L. Schärer. In preparation.
Evolution of the suck behaviour, a postcopulatory female resistance trait in a hermaphroditic flatworm genus.

Abstract

Sexual conflict over the usage of received ejaculate post-mating can give rise to traits of one sex being costly to the other sex. While this conflict can sometimes be resolved pre-copulatorily by mate choice, reciprocally-mating hermaphrodites face a unique challenge as they mate in both the male and female role at the same time. Thus, while multiple mating can lead to more opportunities to donate sperm, it can also lead to receipt of unwanted ejaculate. This can give rise to postcopulatory female resistance traits that allow manipulation and/or rejection of the received ejaculate. A potential instance of such a trait is the 'suck' behaviour, observed in the flatworm genus *Macrostomum*, in which the sperm recipient places its pharynx over its own female antrum (sperm receiving organ) and appears to suck. This is hypothesised to remove received ejaculate after mating. Interestingly, this genus also contains hypodermically-inseminating species, which presumably exhibit forced unilateral mating, with sperm being injected via the male copulatory organ into the partner, and these have so far not been observed to suck. Moreover, the species also exhibit different combinations of reproductive morphological traits, so called mating syndromes, which could be associated with the mating strategy. Here, we examine the evolution of the suck behaviour across the genus *Macrostomum*, by aiming at documenting the behaviour of 64 species. First, we provide evidence that sperm is indeed removed during the suck behaviour in a reciprocally mating species, *Macrostomum hamatum*. Next, we show that there is a correlation between the presence, duration and frequency of the reciprocal mating and suck behaviour, providing evidence for our hypothesis that the suck behaviour co-evolves with reciprocal mating. Finally, we show that reproductive morphology is a good proxy for inferring the mating strategy, presumably as a result of coevolution.

Introduction

Sexual conflict is defined as the conflict between the two sexes over their evolutionary interests involving reproduction (Charnov 1979; Parker 1979, 2006; Arnqvist and Rowe 2005). The ultimate cause of the conflict is anisogamy, in which the male sex produces more but smaller gametes (called sperm in animals), whereas the female sex produces fewer but larger gametes (called eggs in animals). This often leads to the eggs being a limiting resource for obtaining reproductive success and a resulting divergence of interests between the two sexes (Bateman 1948; Charnov 1979). These conflicting interests can give rise to traits of one sex being costly to the other sex, and an ensuing evolutionary arms race between the two sexes (Arnqvist and Rowe 2005). Although work on sexual conflict has usually focussed on gonochoristic organisms (Parker 1970, 2006; Rice and Chippindale 2001; Chapman et al. 2003; Arnqvist and Rowe 2005; Rice et al. 2006; Dougherty et al. 2017; Perry and Rowe 2018; Sayadi et al. 2019; Svensson 2019), sexual conflict is also pervasive in the lesser-studied hermaphroditic organisms (Charnov 1979; Leonard 1991; Michiels 1998; Abbott 2011; Schärer et al. 2015; Beekman et al. 2016; Schenkel et al. 2018).

Of particular interest is the biology of simultaneous hermaphrodites (referred to as hermaphrodites hereafter), which involves interactions that can lead to unique sexual conflicts. For example, there can be conflicts between the mating partners over the sex role exhibited in a mating (i.e. mating as a sperm donor, a sperm recipient, or both), which, depending on the costs and benefits of mating in each role, may lead to sex role preferences (Charnov 1979; Michiels 1998; Anthes 2010; Schärer et al. 2015). Bateman's principle, term coined by Charnov (1979), proposes that there is a "greater dependence of males for their fertility on frequency of insemination" (Bateman 1948). This is qualitatively supported by multiple empirical studies across species (Pélissié et al. 2012; Collet et al. 2014; Janicke et al. 2016; Janicke and Morrow 2018), although studies recommend caution while measuring Bateman's principles (Ah-King 2011) and advocate examining underlying causality of the Bateman's principles (Kokko et al. 2012). In a seminal paper, Charnov (1979) proposed that the Bateman's principle also applies to simultaneous hermaphrodites, such that individuals mate more to give away sperm than to receive sperm, resulting in a mating conflict.

This conflict over the sex roles can be resolved via different strategies (Michiels 1998; Anthes 2010; Schärer et al. 2015). One such strategy is reciprocal mating (also called reciprocal copulation), in which both partners mate in the male and female role simultaneously. Thus, each sperm donor is also a sperm recipient, and while multiple mating can lead to more

opportunities to donate sperm, it may also lead to the receipt of unwanted ejaculate from the partner (Marie-Orleach et al. 2013; Nakadera et al. 2014; Weber et al. 2019; Patlar et al. 2020). While this strategy might seem to represent a cooperative conflict resolution, it may in turn lead to the conflict shifting from the precopulatory to the postcopulatory arena. In the presence of sperm competition, the partner may donate more sperm than is required for fertilisation in order to secure paternity in competition with other sperm donors, which in turn could potentially cause polyspermy (Frank 2000). Even in the absence of any direct risk posed by the received ejaculate, mating with multiple partners—which is probably the norm for most species (Kokko and Mappes 2013)—could lead to the evolution of cryptic female choice in the recipient (Eberhard 1996). Thus, receipt of unwanted or excessive ejaculate can lead to selection of female resistance traits that allow post-copulatory manipulation of the received ejaculate (e.g. via sperm digestion, Sluys 1989; Baur 1998; Michiels 1998; Westheide 1999; Angeloni et al. 2003; Dillen et al. 2009; Koene et al. 2009).

Such female resistance traits can in turn lead to the evolution of male persistence traits or other mating strategies as counter-adaptations that allow the partner's ejaculate to counteract or circumvent the female resistance traits and thereby retain access to the eggs (Michiels 1998; Anthes 2010; Schärer et al. 2015). One such mating strategy is forced unilateral insemination. In this strategy, one of the partners mates in the male role and donates sperm while the other mates in the female role, potentially against its interests, and receives sperm via traumatic or hypodermic insemination (Charnov 1979; Michiels 1998; Lange et al. 2013; Reinhardt et al. 2015; Schärer et al. 2015). Moreover, sexual conflict could lead to the evolution of multiple male persistence and female resistance traits (spanning behaviour, morphology and physiology) that act in sync to either gain access to eggs or to manipulate the received ejaculate, respectively, rather than a single persistence or resistance trait (Arnqvist and Rowe 2005; Dougherty et al. 2017). Thus, we might expect mating strategies (behaviour) to coevolve and be associated with morphological or physiological traits.

A potential example of a behavioural female resistance trait is the 'suck' behaviour, which has initially been documented in the free-living flatworm, *Macrostomum lignano.* Studies in this reciprocally mating simultaneous hermaphrodite, have shown that a mating is often followed by the suck behaviour, during which the worm bends down and places its pharynx over its own female genital opening (which is connected to the female antrum, the sperm receiving organ) and appears to suck (Schärer et al. 2004a; Vizoso et al. 2010; Marie-Orleach et al. 2013). This suck behaviour is hypothesised to be a postcopulatory behaviour used for

removing sperm or other ejaculate components received during mating and thus potentially functions as a female resistance trait (Schärer et al. 2004a; Vizoso et al. 2010; Schärer et al. 2011). While there have been multiple studies on the suck behaviour (Schärer et al. 2004a; Janicke and Schärer 2009; Marie-Orleach et al. 2013; Patlar et al. 2020), there has to date been no direct evidence for sperm and/or ejaculate being removed during the suck behaviour.

Interestingly, species of the genus *Macrostomum* exhibit different combinations of reproductive morphological traits that could be associated with these two different mating strategies, i.e. reciprocal mating and hypodermic insemination. Indeed, a previous study—taking advantage of the fact that internal reproductive structures can be directly observed *in vivo* in these highly transparent organisms—had shown an association between the mating strategy and certain male and female reproductive traits in 16 *Macrostomum* species, calling these two alternative outcomes the reciprocal and hypodermic mating syndrome, respectively (Schärer et al. 2011). Moreover, a more recent study has used a composite measure derived from the observation of different components of the reproductive morphology, in an attempt to classify ~150 *Macrostomum* species as showing the reciprocal or hypodermic mating syndrome, respectively (Brand et al. in prep.). Briefly, the species with the reciprocal mating syndrome have a thick female antrum wall, a complex sperm design (including lateral bristles), a blunt or sharp stylet (male intromittent organ) tip and allosperm can be found in the female antrum. In contrast, the species with the hypodermic mating syndrome have a thin female antrum wall, a simpler sperm design (with absent or reduced bristles), a sharp stylet tip, and hypodermic allosperm can be observed in the parenchymal tissue of the worm. The stiff lateral bristles on the sperm in reciprocally mating species are hypothesized to represent a male persistence trait that allows the sperm to remain anchored in the female antrum, and not be removed during the suck behaviour (Vizoso et al. 2010; Schärer et al. 2011), while the thick female antrum wall might prevent injury resulting from the male genitalia during mating. The sharp needle-like stylet tip in hypodermically inseminating species potentially allows sperm to be injected though the partner's epidermis, while the simple sperm design presumably facilitates its movement through the partners body (Schärer et al. 2011; Ramm et al. 2015).

In our study, we examine the evolution of the suck behaviour by aiming to document reproductive behaviour in 64 *Macrostomum* species. We, for the first time, provide videographic evidence that ejaculate is indeed removed during the suck behaviour in one reciprocally mating species, namely *M. hamatum*, thus supporting the previously proposed

hypotheses for the function of this postcopulatory behaviour. Further, we show that the suck behaviour and other copulatory behaviours have likely coevolved, by documenting correlated evolution while accounting for phylogeny. Finally, we show that the reproductive morphology is a good proxy for inferring the mating strategy.

Materials and Methods

Study organisms

We obtained multiple specimens for a large number of *Macrostomum* species from a range of locations and habitats, using a variety of extraction techniques, which we will report on in more detail as part of a separate study on the phylogenetic interrelationships and character evolution in this genus (Brand et al. in prep). Briefly, most specimens were sampled directly from field sites, some were sampled in natural and artificial ponds, or in aquaria containing other organisms, and some were obtained from short- and long-term laboratory cultures maintained either by our group or by colleagues.

Observation methodology

We aimed at documenting the mating behaviour of a total of 64 *Macrostomum* species, by placing the worms in so-called mating chambers (Schärer et al. 2004a). Briefly, a mating chamber consists of the worms being placed between two microscope slides in a drop of a medium (i.e. either freshwater or water with different salinity depending on their habitat from which they are collected), with a certain number of spacers (between the slides) and sealed with pure white Vaseline (note that we generally also placed 4-6 empty drops around the drops containing worms to further avoid evaporation). The number of spacers and the volume of the drops we used varied depending on the size and number of worms in a drop, respectively. Usually, for a pair of worms of the size of *M. lignano* (~1.5 mm body length), we used 2 spacers (each spacer being ~105 μm) and a drop of ~3 μl. The movie recordings were done over different time-periods for the different species, depending on their availability during sampling trips and/or in lab cultures. The chambers were recorded under different setups that varied somewhat in the cameras or lighting used, though from an earlier study we know that different setups do not strongly influence mating behaviour, at least in some species (Singh et al., submitted). Usually, the movies were recorded in QuickTime Format using BTV Pro (http://www.bensoftware.com/) at 1 frame s^{-1}. All worms were checked for being adults (defined as having visible gonads or eggs) before or after being filmed.

Detailed observation of the mating and suck behaviour in M. hamatum

While we have been able to document ejaculate being removed during the suck behaviour or sperm sticking out of female antrum after the suck behaviour in multiple species, this could be seen most distinctly in detailed closeup movies of *M. hamatum*, since field-collected specimens of this species appeared to be more transparent. This allowed us to visualise the deposition and subsequent removal of ejaculate during the mating and suck behaviours, respectively, more clearly than in any other species to date. Here we examine the mating behaviour of *M. hamatum*, collected on 27. July 2017 directly in front of the Tvärminne Zoological Station, Finland (N 59.84452, E 23.24986), in detail under the microscope. For definition and scoring of mating and suck (see below). Note that while we describe and illustrate in detail one such instance of suck, we have observed ejaculate removal in other detailed closeup movies of *M. hamatum* as well, and these corroborated our finding as described here.

Scoring of reciprocal mating and suck behaviour across species

We scored the mating behaviours in these mating movies by visual frame-by-frame analysis (Supplementary Table S1). A reciprocal mating was scored when the tail plates of two worms were in ventral contact and intertwined, such that the female antrum was accessible to the partner's stylet and vice-versa, which would allow reciprocal transfer of ejaculate. In most species, this copulatory posture is usually accompanied by the pair being tightly inter-linked (like two interlocking G's, see figure 1), and is similar to how mating has been described in *M. lignano*, Schärer et al. 2004. In some species the mating posture can be slightly different from that observed in *M. lignano*, as described in Supplementary Table S2. The copulation duration was measured from the frame when the tail plates are in ventral contact and tightly intertwined, to the frame where the tail plates were no longer attached to each other. We defined observed behaviours as matings only if the pair was in the above described intertwined posture for at least 3 seconds. The suck duration was measured starting from the frame when an individual placed its pharynx on its female genital opening, up to the frame where the pharynx disengaged. For each replicate drop, we divided the total number of matings and sucks by the number of worms and the movie duration to obtain a standardized value (thus obtaining frequency estimates per worm and hour). We then averaged the frequency and duration estimates across all replicates for each species to obtain the species estimates of the respective behaviours.

While we tried to observe hypodermic insemination, we only saw some rare behavioural instances in certain species that could presumably be hypodermic insemination. Possible reasons for not observing hypodermic insemination could be if the mating occurs very rapidly or if these species mate rarely. Since we could not be sure of hypodermic insemination, we finally scored species as either exhibiting reciprocal mating or not. Note that absence of reciprocal mating does not necessarily imply presence of hypodermic insemination, e.g. it could also result from a species not mating under the laboratory condition.

Unless stated otherwise, we performed our analysis in R, version 3.6.1 (R Core Team 2017).

Evolution of the mating and suck behaviour across the Macrostomum genus

To perform phylogenetically corrected comparative analyses, we used a trimmed ultrametric phylogeny of the genus *Macrostomum*, which is based on large-scale phylogenomic analyses using a combination of de novo transcriptomes and 28S sequences.

We estimated the phylogenetic signal for the continuous traits (duration and frequency of reciprocal mating and suck) using Pagel's λ (*phytools* package version 0.6-99, Pagel 1999; Revell 2012). A λ value of 1 indicates a strong phylogenetic signal, while a value around 0 indicates no/low phylogenetic signal (Pagel 1999). We found that the suck frequency and duration, and the mating frequency exhibited phylogenetic signal that was significantly different from 0, while mating duration exhibited a phylogenetic signal that was marginally significantly different from 0 (Supplementary Table S3).

For each trait, we fitted four different models of trait evolution, i.e. Brownian motion, Ornstein–Uhlenbeck, early-burst and lambda model (*geiger* package version 2.0.6.2, Harmon et al. 2008) (Supplementary Table S4). We found that the lambda model had the largest Akaike Information Criteria (AIC_c) weights, and this hence was chosen for further PGLS analysis (Supplementary Table S4). All continuous traits were log-transformed for all the analysis.

Testing for coevolution between components of reproductive behaviour

Correlation between reciprocal mating and the suck behaviour: We used the DISCRETE model in BayesTraits V.3.0.1 to test for correlated evolution between reciprocal mating and the suck behaviour (scored as present/absent), using Reversible Jump Markov Chain Monte Carlo (RJ MCMC) (Pagel 1994; Pagel and Meade 2006; Meade and Pagel 2016). We tested support for correlated evolution by comparing dependent and independent models of

character evolution. Specifically, we compared marginal likelihood of a dependent model, in which the presence of suck depends on the presence of reciprocal mating, with an independent model, in which the presence of suck and the presence of reciprocal mating evolve independently. Each RJ MCMC chain was run for ten million iterations and the first one million iterations were discarded as burn-in, after which the chain was sampled every 1000^{th} iteration. We used a gamma hyperprior (gamma 0 1 0 1), and placed 1000 stepping stones (with each iterating 10000 times) to obtain the marginal likelihood values for the models. We performed three separate runs for both dependent and independent model to check for the stability of the likelihood values. Using the R package coda (Plummer et al. 2006), we confirmed that the chains had converged (using Gelman and Rubin's Convergence Diagnostic; Gelman and Rubin 1992) and that the Effective Sample Size >200 for all parameters. In addition, we also confirmed that the acceptance rate was between 20-40% (the ideal acceptance rate when the chain is at convergence and indicating good mixing). We evaluated the alternative models using the Log Bayes Factor (BF). We used the convention that a BF value greater than 2 is considered as positive support for the best-fit model, while a value between 5-10 and greater than 10 is considered as strong and very strong support for the model respectively (Pagel and Meade 2006). To examine the robustness of our co-evolutionary model, we repeated the analysis for a subset of species, by excluding six species that had been observed for less than 21 hours (~10% quantile). For the dependent models (of entire dataset), we estimated the rates of transition among the different trait states and calculated z-values. The z-value can be understood as the percentage of times the parameter of the relevant transition rate was zero, amongst all the sampled parameters. Thus, a high z-value implies that in many iterations the transition rate value was zero, indicating that the transition between these states is unlikely.

Reciprocal mating and the morphologically inferred mating syndrome: We checked for an association between behaviourally observed reciprocal mating and morphologically inferred mating syndrome, using the DISCRETE model in BayesTraits V.3.0.1 (with the specifications as described above for presence of reciprocal mating and suck). Briefly, the species were inferred as reciprocally mating or hypodermically inseminating using data on their male and female genital morphology, sperm morphology, and the location of received sperm (Brand et al. in prep.). One of the species, *Macrostomum* sp. 101 had a morphology intermediate between reciprocal and hypodermic, but since the discrete method in BayesTraits only allows binary states for each trait, we excluded this species. To examine the validity of our analysis,

we repeated it for a subset of species, by excluding six species that had been observed for less than 21 hours (~10% quantile). In addition, we calculated the rates of transition among the different trait states and calculated z-values (similar to earlier).

Correlation between the frequency and the duration of behaviours: For the species that exhibited both reciprocal mating and suck, we investigated if there was a correlation between the frequency and duration of mating and suck behaviours (continuous variables), using phylogenetic generalized least-squares (PGLS) regression implemented in the (*caper* package version 1.0.1, Orme et al. 2014). PGLS accounts for the nonindependence of the data by incorporating phylogenetic relationships between species into the error structure of the model. For each analysis, we used the phylogenetic signal (Pagel's λ) of the dependant variable that was previously estimated, as the species exhibiting both copulation and suck are a subset of species that exhibited only copulation or suck. Thus, using the Pagel's λ estimated from the latter is more precise. We examined the residuals of each model for normality and homogeneity (Mundry 2014). Additionally, we scrutinized for influential cases (species) in each of our model, by excluding one species at a time from the data and rerunning the analysis, and comparing the results obtained with the results for the entire dataset (Mundry 2014). Finally, we evaluated the validity of our model by repeating the PGLS analysis for a subset of data, which excluded species in which mating or suck had only been observed in one replicate.

Results

Detailed observation of reciprocal mating and suck behaviour in M. hamatum

In the detailed movie of *M. hamatum*, the worms were already interlinked in the copulatory position at the beginning of the clip (Figure 1), and we considered this as t = 0 s. We, hereafter, refer to the worm on right as P1 (white in Figure 1) and the worm on left as P2 (grey in Figures 1 and 2). The tail plates were "touching each other ventrally in opposite directions and their anterior ventral surface was in contact with the posterior dorsal surface of the partner", as previously described for *M. lignano* (Schärer et al. 2004a). Interestingly, in *M. hamatum* the copulatory position resembled a square with rounded corners, as opposed to *M. lignano,* which seemed more circular or G-shaped. There also was a striking difference in the position of the tail plate, which in *M. hamatum* stood at a 90° angle from the posterior body axis and appeared to poke into the anterior ventral surface of the partner. Moreover,

M. hamatum had a much more prominent erection (i.e. a translucent finger-like shape on the tail plate, likely formed by the everted male genital antrum), which poked into the posterior ventral surface of the partner in the region of the vagina (although it was unclear if the ejection entered the partner). The stylet of P2 performed poking movements towards P1's female antrum (moving inside of the relatively stationary erection), initially without any transfer of ejaculate. At t = 3-5 s, the stylet of P2 can be seen poking against the dorsal side of the partner's antrum wall several times, each time leading to a visible bulge on P1's dorsal side, and eventually began to deposit ejaculate (seen as a visible darkening of the antrum lumen starting at about t = 5s). During this the seminal vesicle emptied (seen at the base of the erection), while the female antrum of the partner filled up with ejaculate over the next ~21 seconds. At t = 28 s, P1 pushed out its female antrum and placed its pharynx on top of its own female genital opening and sucked. The ejaculate could be seen leaving P1's female antrum (i.e. the visible darkening of the antrum lumen moved towards the pharynx between t = 29-30s). In total the suck behaviour lasted for 7 seconds and interestingly, during the suck the stylet of P2 remained partly anchored in P1's female antrum. At t = 52 s, the mating was over with a total duration of ~64 s. At t = 56 s, only P2 was in the frame and the ejaculate in its female antrum was clearly visible. It continued to have a relatively small erection despite the mating being over. At t = 78 s, P2 pushed its female antrum out and some sperm was ejected from the female antrum at t = 80 s, notably before the pharynx made contact (Figure 2). At t = 81 s, P2 put its pharynx on top of its female genital opening and sucked for 10 s. After the suck, some sperm could still be seen sticking out of the female antrum, especially at 92 s, but most of the ejaculate had been removed from the antrum. The female antrum remained slightly everted and the erection visible until at least 108 s.

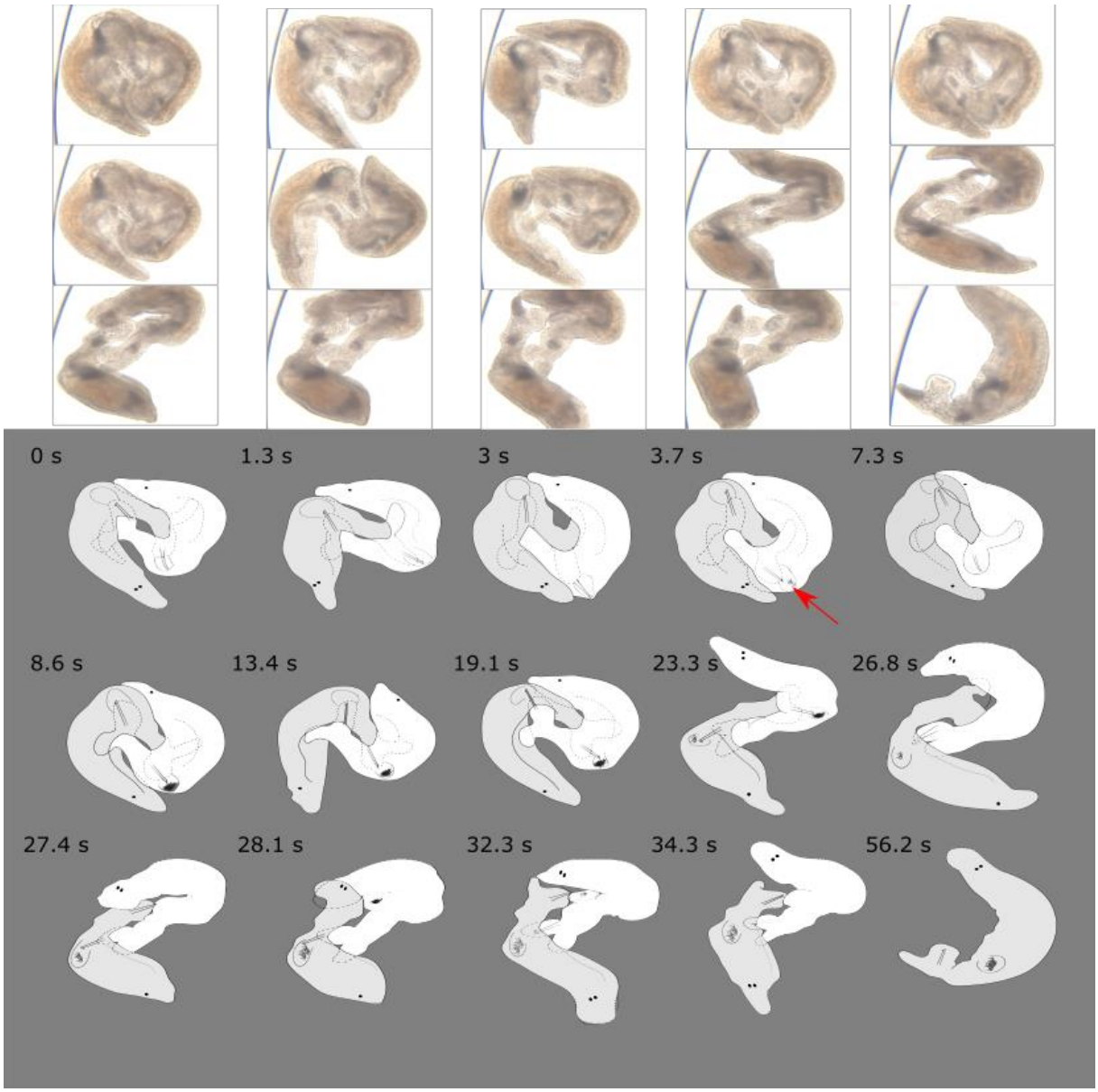


Figure 1. Reciprocal mating and a postcopulatory suck in *M. hamatum*: Illustration of ejaculate deposition and removal during the suck behaviour in P1 (white worm on right at 0 s). Ejaculate (dark mass indicated by red arrow) can be seen being deposited starting from 3.7 s in P1's female antrum, followed by the worm pushing its female antrum out (27.4 s) and sucking. There is a visible reduction in the quantity of ejaculate in the female antrum of focal worm after suck. Note that we call the frame from where we start describing the movie as t = 0 s, but the mating had started before.

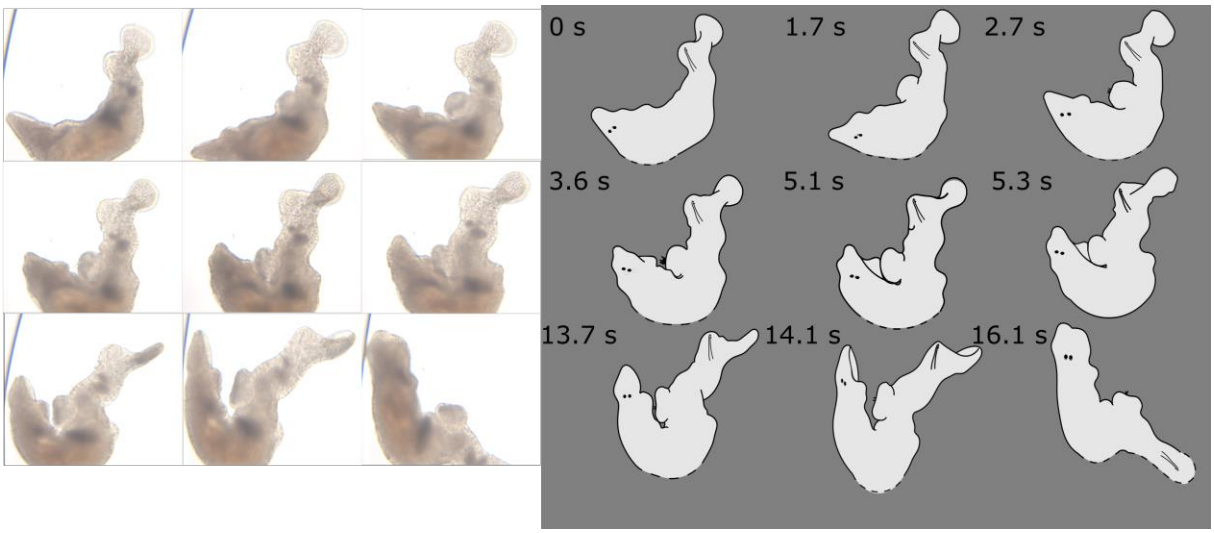


Figure 2. Postcopulatory suck: Following the reciprocal mating (in Figure 1), the individual that had not previously sucked, pushed out its female antrum (which leads to some sperm appearing at the female genital opening) and then put its pharynx on top of the female genital opening and sucked out most of the previously deposited ejaculate. Moreover, some sperm can be seen sticking out of the female genital opening after the suck. Note that the timing of this panel starts from 0, but this is a continuation of the movie from Figure 1.

Evolution of the mating and suck behaviour across Macrostomum genus

We observed a total of 2796 worms across 64 *Macrostomum* species, with a mean of 44 worms and 76.7 hours observed per species, for a total observation time of 4908 hours. Of the 64 species, 30 species exhibited reciprocal mating behaviour, 31 species exhibited the suck behaviour, and 25 species exhibited both reciprocal mating and the suck behaviour (Figure 3). Among the species that exhibited reciprocal mating, the average mating frequency was 0.84 hr^{-1} (range: 0.02-7.82 hr^{-1}) and the average mating duration was 283.66 s (range: 5.25-4609 s) (Figure 4). Among the species that did suck, the average suck frequency was 0.54 hr^{-1} (range: 0.01-3.7 hr^{-1}) and the average suck duration was 9.58 s (range: 4.75-16.08 s) respectively (Figure 4).

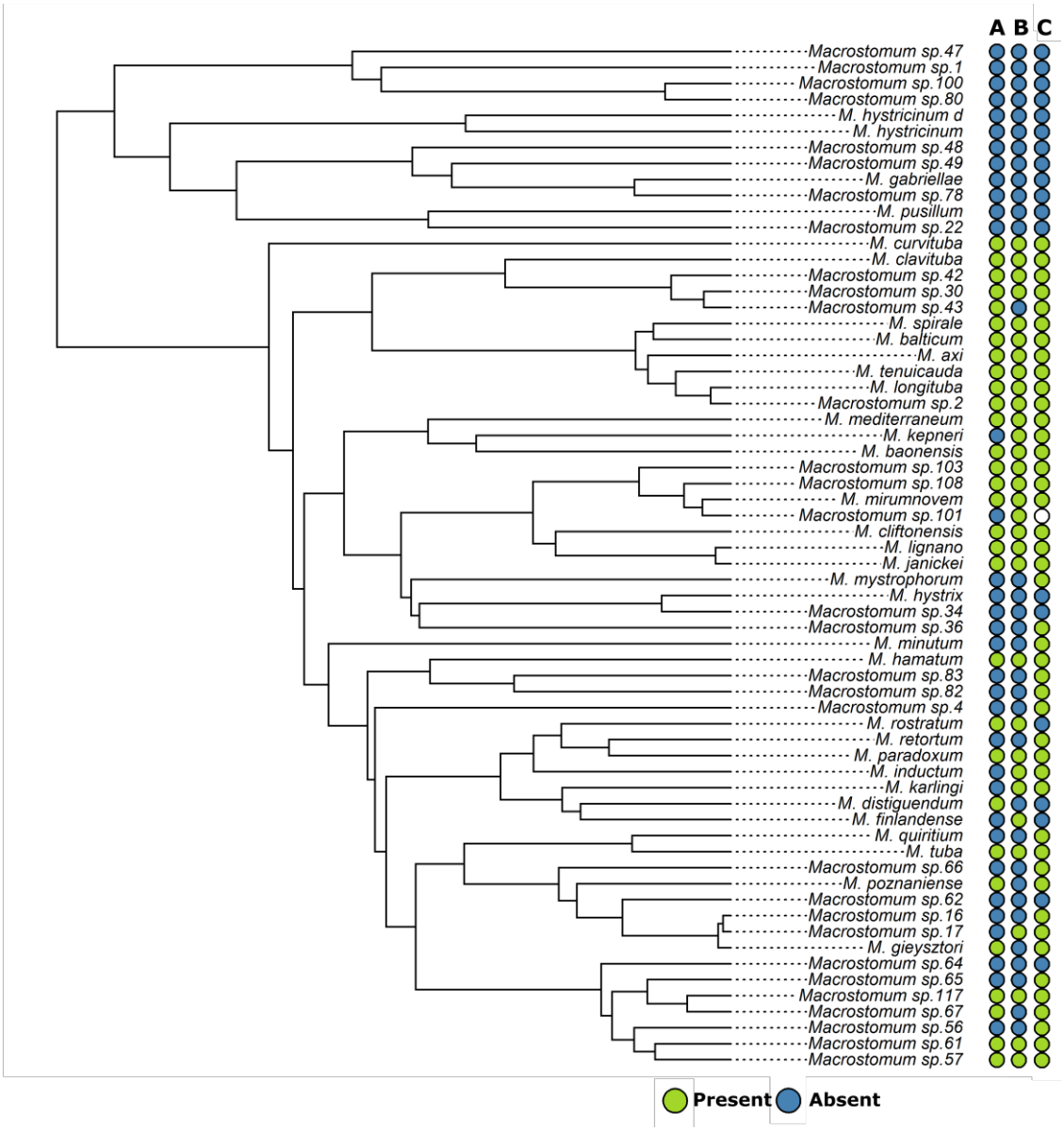


Figure 3. Behaviourally inferred presence or absence of reciprocal mating (A) and the suck behaviour (B), and morphologically inferred presence or absence of the reciprocal mating syndrome across the *Macrostomum* phylogeny (for a total of 64 *Macrostomum* species). Note that for the behaviourally inferred traits an absence may be due to a lack of sufficient data, and that the absence of the reciprocal mating syndrome represents the hypodermic insemination syndrome, except for *Macrostomum* sp. 101, which showed an intermediate syndrome (white dot).

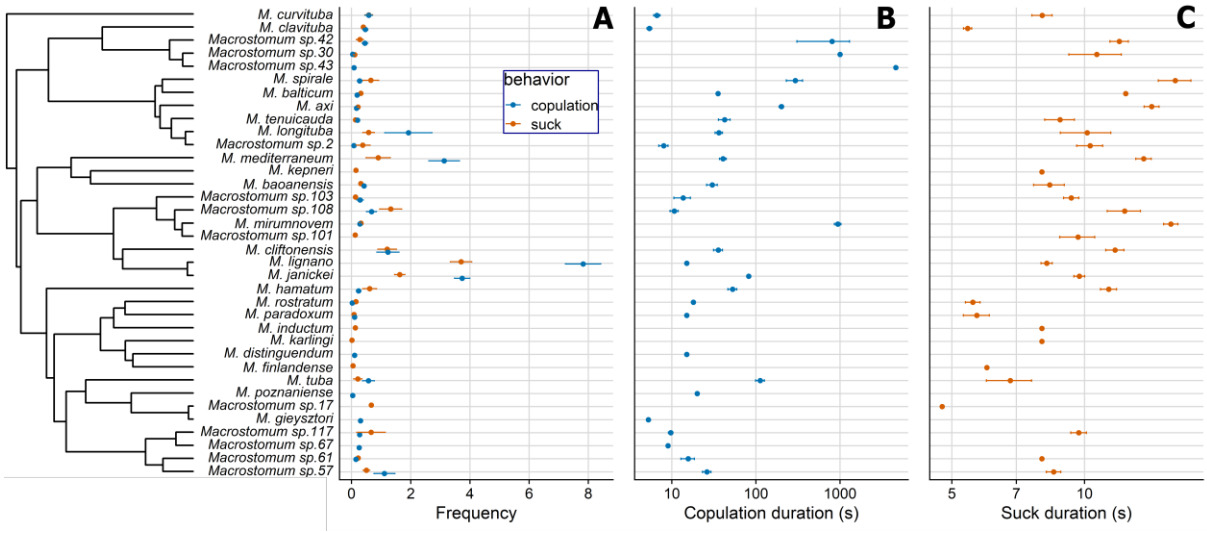


Figure 4. Trimmed *Macrostomum* phylogeny of the 36 species that showed either reciprocal mating behaviour and/or the suck behaviour alongside data on means and standard errors of (a) reciprocal mating and suck frequency, (b) log-transformed mating duration (s), and (c) suck duration (s) of the species. Note that some species exhibited only reciprocal mating or the suck behaviour. Also note that for the species in which a behaviour had been observed in only 1 replicate, we report only that single value.

Testing for coevolution between components of reproductive behaviour

Correlation between reciprocal mating and the suck behaviour: We found very strong support for the dependent model of evolution over the independent model for the correlation between reciprocal mating and suck behaviour, with all three runs for each model providing similar values (average marginal likelihood, independent = -88.67, dependent =-83.3, BF: 10.75; see also Supplementary Table S5). This showed that the presence of reciprocal mating and the suck behaviour are strongly correlated. This result was robust to observation time, since excluding the 6 species that were observed for less than 21 hours gave similar results (Supplementary Table S5).

The transition from the absence of both reciprocal mating and the suck behaviour to presence of either of these traits is the most unlikely, as evident from the very low evolutionary transition rates and the very high z-values (Figure 5). Interestingly, the other transitions, including losing reciprocal mating or the suck behaviour from both present states, are all equally likely.

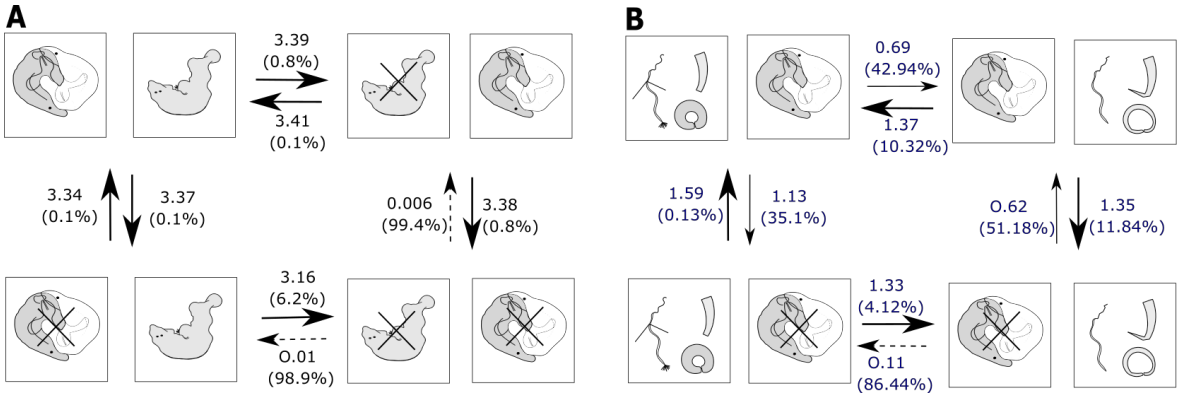


Figure 5. Evolutionary transition rates and the z-value % (in parenthesis): We examine the correlation between the presence of reciprocal mating and a) the suck behaviour, and b) the morphologically inferred mating syndrome. For the evolutionary transition rates, the mean of the posterior distributions across all runs is given. The z-value can be understood as the percentage of times the parameter of the relevant transition rate was zero, amongst all the sampled parameters. The different arrows represent different probabilities of transitions between the states: high probability (thick black arrows, z-value < 15%), moderate probability (thin black arrows, z-value 30-55%), and low probability (dashed black arrows, z-value > 85%).

Reciprocal mating and the morphologically inferred mating syndrome: There was a correlation between the presence of the reciprocal mating behaviour and the morphologically inferred mating syndrome, as evident from the strong support for the dependent model of evolution over the independent model, with similar values for the three independent runs of each model (average marginal likelihood, independent = -68.31, dependent = -65.12, BF: 6.38; see also Supplementary Table S5). Similar to earlier, our result was robust to observation time, as the subsetted data gave us similar results (Supplementary Table S5).

The transition from reciprocal mating behaviour and reciprocal mating morphology to absence of either of these traits was moderately unlikely, while the converse was very likely (Figure 5). Similarly, the transition from absence of reciprocal mating behaviour, and hypodermic morphology to either presence of reciprocal mating behaviour or reciprocal mating morphology was moderately unlikely, while the converse was very likely.

Correlation between the frequency and the duration of behaviours: Partially in line with our predictions, there is a significant correlation between both mating and suck frequency, and mating and suck duration; while there is no significant correlation between mating duration and frequency, and suck duration and frequency (Table 1, Figure S1). The subsetted dataset also gave qualitatively similar results (Supplementary Table S6).

Table 1. Summary of the PGLS results for the association between aspects of reciprocal mating and suck behaviour for the entire dataset. Note that PGLS was done on log-transformed variables.

Dependent variable	Predictor variable	λ	β±s.e.	t, df	P
Suck frequency	Reciprocal mating frequency	0.68	0.45±0.1	4.43, 23	0.0002
Suck duration	Reciprocal mating duration	0.75	0.1±0.03	3.83, 23	0.0008
Reciprocal mating frequency	Reciprocal mating duration	0.51	-0.09±0.15	-0.61,28	0.54
Suck frequency	Suck duration	0.68	0.68±0.68	0.99,29	0.33

Discussion

Sexual conflict can give rise to antagonistic coevolution across all sexual systems (Brennan and Prum 2015; Perry and Rowe 2015; Schärer et al. 2015; Hosken et al. 2019). Here we documented the presence and evolution of a female resistance trait, the suck behaviour, in the hermaphroditic flatworm genus *Macrostomum*. Our findings corroborated our hypothesis that the suck functions as a female resistance trait to remove received ejaculate, given the significant correlation between different aspects of reciprocal mating and the suck behaviour. We also showed that the reproductive morphology is a good proxy for inferring the mating strategy, presumably also as a result of coevolution.

While multiple studies in *Macrostomum* have examined different aspects of the suck behaviour, generally assuming that it is involved in removing received ejaculate components, our detailed movie of *M. hamatum* provides the first direct evidence that ejaculate is indeed removed during this postcopulatory behaviour. Interestingly, compared to *M. lignano* (Schärer et al. 2004a), *M. hamatum* has a more angular mating posture (possibly due to the angular position of the tail plate), a larger erection around the stylet and the worms prominently evert the female antrum just before the suck behaviour. This could result from differences in the female antrum morphology, which has a strong musculature and an inner second chamber connecting to the main female antrum (Luther 1947), in contrast to *M. lignano*, which has a simpler female antrum with a single chamber (Ladurner et al. 2005; Vizoso et al. 2010). Similarly, the prominent erection of the male antrum could result from a muscular morphology that is similar to the muscular cirrus seen in species of the sister genus, *Psammomacrostomum* (Ax 1966; Janssen et al. 2015). This combination of a prominent female antrum and relatively transparent specimens, potentially helped us visualise the suck

better in *M. hamatum* than in other *Macrostomum* species observed to date. Remarkably, in *M. hamatum*, worms sometimes suck while still copulating, unlike in *M. lignano* where a worm sucks only after mating (Schärer et al. 2004a; Marie-Orleach et al. 2013).

While we see ejaculate being removed during the suck behaviour, we are not sure whether it is ingested. While sperm-digestion is widespread in hermaphrodites (Sluys 1989; Baur 1998; Michiels 1998; Westheide 1999; Angeloni 2003; Dillen et al. 2009; Koene et al. 2009), it usually occurs inside an organ connected to the individual's reproductive system, unlike in the case of the suck behaviour. To our knowledge, there have been only two earlier cases of sperm being orally taken up in hermaphrodites, one in the arrowworm *Spadella cephaloptera* (John 1933) and the other in the leech *Placobdella parasitica* (Myers 1935). Thus, suck behaviour seems to be a novel trait, which to date has only been observed in a large subclade of the genus *Macrostomum* (see the deep split in the phylogeny in Figure 3). Similar to the suck, there is sperm dumping in many gonochoristic species, in which the female physically ejects sperm from males from her sperm storage organ, and this, at least in some cases, is thought to be a mechanism of cryptic female choice (Eberhard 1996; Pizzari and Birkhead 2000; Snook and Hosken 2004; Peretti and Eberhard 2010; Firman et al. 2017). If the suck behaviour also functions in cryptic female choice, we might expect individuals to remove or retain sperm of certain partners more frequently. This has indeed been observed in *M. lignano*, where the propensity to suck in an individual is affected by the genotype of its partners (Marie-Orleach et al. 2017).

We found a significant correlation between the presence of reciprocal mating and the suck behaviour. In reciprocally mating species, ejaculate is deposited in the female antrum allowing its removal during the suck behaviour, while in hypodermically inseminating species, sperm is injected potentially anywhere in the body and probably rarely in the female antrum (Schärer et al. 2011). Thus, we do not expect to see the suck behaviour in hypodermically inseminating species. Interestingly, the transition rates showed that while it is unlikely for a species that lacks both reciprocal mating and the suck behaviour to gain either of these traits, the loss of either reciprocal mating or the suck behaviour is more likely from any state. This loss of either of these traits could represent a transitional step towards hypodermic insemination, which might arise as a means to circumvent the female control and allow access to the eggs. Moreover, this is supported by the results that there are multiple independent origins of hypodermic insemination, indicating multiple transitions away from

reciprocal mating, while no such transition has been detected in the converse direction (Schärer et al. 2011).

Surprisingly, there were 11 species that showed either only reciprocal mating or the suck behaviour, but not both. This might either result from the species not exhibiting the behaviour under our laboratory conditions, or it might indicate that a species lacks the behaviour. If a species mates rarely, they might not need to remove the sperm received. In our study the species that copulated but did not suck had a low or intermediate mating frequency (see species with low mating frequency in Figure 4A). Alternatively, species might lack reciprocal mating, but losing a resistance trait like the suck behaviour might take longer, e.g. if suck does not impose costs on the fecundity or has an alternate function. Species are predicted to lose defensive or resistance traits only after the male traits have become substantially less harmful, leading to a time lag (Parker 1979). For example, a study on the seed beetle, *Callosobruchus maculatus*, showed that while males evolved relatively reduced length of genital spines in large males under monogamy, there was no detectable evolution in female genitalia in the same time period (Cayetano et al. 2011).

We found a significant correlation between frequency and duration of reciprocal mating and suck. If a frequent mating or longer mating duration implies more sperm transfer, then we expect selection for a frequent or longer duration suck (particularly if ejaculate receipt is associated with fitness costs). Alternatively, it might be the female resistance trait driving the evolution of reciprocal mating frequency and duration, e.g. frequent sucking might select for a higher mating frequency (Karlsson Green and Madjidian 2011). While a longer mating duration does imply more ejaculate transferred in some species (Engqvist and Sauer 2003) and given that it is often used as a proxy for ejaculate size (Kelly and Jennions 2011), it could also correlate with the complexity of genitalia, such that more complex genitalia might have longer mating duration (King et al. 2009). In the latter case, we could expect the suck duration to also correlate with complexity of genitalia or sperm. While we expected species that copulate longer to copulate infrequently, e.g. if it took longer to replenish the ejaculate transferred, we did not observe such a trend. Similarly, there was no significant correlation between suck frequency and duration, which we might have expected assuming frequent sucks and longer suck duration serve a similar function.

The significant correlation between the observed reciprocal mating and the morphologically inferred reciprocal mating syndrome confirms the previous findings (Schärer et al. 2011), and shows that persistence and resistance are generally not limited to single traits, but are

presumably composite entities involving a suite of behavioural, morphological and physiological traits acting together (Arnqvist and Rowe 2005). For example, the female antrum wall and the suck behaviour might be different components of resistance. While a thickened female antrum wall might prevent injury resulting from the male genitalia when reciprocally mating, the suck behaviour removes the received ejaculate. Adaptations of the female reproductive tracts are also seen in the seed beetle *C. maculatus,* where a thicker female tract lining served as a resistance trait against harm by male genitalia (Dougherty et al. 2017). Moreover, resistance and persistence traits can also be at the proteomic level. A study in *M. lignano* identified two seminal fluid transcripts that caused mating partners to suck less often, with the authors suggesting that these proteins might be male counter-adaptations to suck behavior (Patlar et al. 2020).

Thus, our study, while providing evidence for sperm removal during the suck behaviour, also provides support for the coevolution of reciprocal mating and the suck behaviour, and provides information in a phylogenetic context on the occurrence and interspecific variation of the suck behaviour. Moreover, we show that reproductive morphology can be a good proxy for the mating strategy. Future studies should try to experimentally manipulate the suck behaviour to elucidate its function and effect on the fitness of the sperm donor and sperm recipient, respectively.

Acknowledgements

We would like to thank Gudrun Viktorin, Lukas Zimmermann, Jürgen Hottinger, Urs Stiefel and Daniel Lüscher for technical support. The many people that have been involved in the collection and maintenance of the worms we used here will be thanked in the study that presents the phylogeny. This research was supported by grants 31003A_162543 and 310030_184916 of the Swiss National Science Foundation (SNSF).

References

Abbott, J. K. 2011. Intra-locus sexual conflict and sexually antagonistic genetic variation in hermaphroditic animals. Proc. R. Soc. B Biol. Sci. 278:161–169.

Ah-King, M. 2011. Female sexual selection in light of the Darwin-Bateman paradigm.

Angeloni, L. 2003. Multiple mating, paternity, and body size in a simultaneous hermaphrodite, *Aplysia californica*. Behav. Ecol. 14:554–560.

Anthes, N. 2010. Mate choice and reproductive conflict in simultaneous hermaphrodites. Pp. 329–357 *in* Animal Behaviour: Evolution and Mechanisms. Springer Berlin Heidelberg, Berlin, Heidelberg.

Arnqvist, G., and L. Rowe. 2005. Sexual conflict. Princeton University Press, Princeton, NJ.

Ax, P. 1966. Die bedeutung der interstitiellen sandfauna für allgemeine probleme der systematik, ökologie und biologie. verhandlungen des institutes für meeresforsch. Bremerhaven Sonderband:15–65.

Bateman, A. J. 1948. Intra-sexual selection in *Drosophila*. Heredity (Edinb). 2:349–368.

Baur, B. 1998. Sperm Competition in Molluscs. P. *in* Sperm Competition and Sexual Selection.

Beekman, M., B. Nieuwenhuis, D. Ortiz-Barrientos, and J. P. Evans. 2016. Sexual selection in hermaphrodites, sperm and broadcast spawners, plants and fungi. Philos. Trans. R. Soc. B Biol. Sci. 371:20150541.

Brennan, P. L. R., and R. O. Prum. 2015. Mechanisms and Evidence of Genital Coevolution: The Roles of Natural Selection, Mate Choice, and Sexual Conflict. Cold Spring Harb. Perspect. Biol. 7:a017749.

Brooks, S. P., and A. Gelman. 1998. General Methods for Monitoring Convergence of Iterative Simulations. J. Comput. Graph. Stat. 7:434–455.

Cayetano, L., A. A. Maklakov, R. C. Brooks, and R. Bonduriansky. 2011. Evolution of male and female genitalia following release from sexual selection. Evolution (N. Y). 65:2171–2183.

Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends Ecol. Evol. 18:41–47.

Charnov, E. L. 1979. Simultaneous hermaphroditism and sexual selection. Proc. Natl. Acad. Sci. 76:2480–2484.

Collet, J. M., R. F. Dean, K. Worley, D. S. Richardson, and T. Pizzari. 2014. The measure and significance of Bateman's principles. Proc. R. Soc. B Biol. Sci. 281:20132973.

Dillen, L., K. Jordaens, and T. Backeljau. 2009. Sperm transfer, sperm storage, and sperm digestion in the hermaphroditic land snail *Succinea putris* (Gastropoda, Pulmonata). Invertebr. Biol. 128:97–106.

Dougherty, L. R., E. van Lieshout, K. B. McNamara, J. A. Moschilla, G. Arnqvist, and L. W. Simmons. 2017. Sexual conflict and correlated evolution between male persistence and female resistance traits in the seed beetle *Callosobruchus maculatus*. Proc. R. Soc. B Biol. Sci. 284:20170132.

Eberhard, W. G. 1996. Female control: sexual selection by cryptic female choice. Princeton University Press.

Engqvist, L., and K. P. Sauer. 2003. Determinants of sperm transfer in the scorpionfly *Panorpa cognata*: male variation, female condition and copulation duration. J. Evol. Biol. 16:1196–1204.

Firman, R. C., C. Gasparini, M. K. Manier, and T. Pizzari. 2017. Postmating Female Control: 20 Years of Cryptic Female Choice. Trends Ecol. Evol. 32:368–382.

Frank, S. A. 2000. Sperm competition and female avoidance of polyspermy mediated by

sperm-egg biochemistry. Evol. Ecol. Res. 2:613–625.
Gelman, A., and D. B. Rubin. 1992. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 7:457–472.
Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. Geiger: investigating evolutionary radiations. Bioinformatics 24:129–131.
Hosken, D. J., C. R. Archer, and J. E. Mank. 2019. Sexual conflict. Curr. Biol. 29:R451–R455.
Janicke, T., I. K. Häderer, M. J. Lajeunesse, and N. Anthes. 2016. Evolutionary Biology: Darwinian sex roles confirmed across the animal kingdom. Sci. Adv., doi: 10.1126/sciadv.1500983.
Janicke, T., and E. H. Morrow. 2018. Operational sex ratio predicts the opportunity and direction of sexual selection across animals. Ecol. Lett. 21:384–391.
Janicke, T., and L. Schärer. 2009. Sex allocation predicts mating rate in a simultaneous hermaphrodite. Proc. R. Soc. B Biol. Sci. 276:4247–4253.
Janssen, T., D. B. Vizoso, G. Schulte, D. T. J. Littlewood, A. Waeschenbach, and L. Schärer. 2015. The first multi-gene phylogeny of the Macrostomorpha sheds light on the evolution of sexual and asexual reproduction in basal Platyhelminthes. Mol. Phylogenet. Evol. 92:82–107. Elsevier Inc.
John, C. 1933. Habits, structure and development of *Spadella cephaloptera*. Q J Microsc Sci 75:625–696.
Karlsson Green, K., and J. A. Madjidian. 2011. Active males, reactive females: stereotypic sex roles in sexual conflict research? Anim. Behav. 81:901–907.
Kelly, C. D., and M. D. Jennions. 2011. Sexual selection and sperm quantity: Meta-analyses of strategic ejaculation. Biol. Rev. 86:863–884.
King, R. B., R. C. Jadin, M. Grue, and H. D. Walley. 2009. Behavioural correlates with hemipenis morphology in New World natricine snakes. Biol. J. Linn. Soc. 98:110–120.
Koene, J. M., K. Montagne-Wajer, D. Roelofs, and A. Ter Maat. 2009. The fate of received sperm in the reproductive tract of a hermaphroditic snail and its implications for fertilisation. Evol. Ecol. 23:533–543.
Kokko, H., H. Klug, and M. D. Jennions. 2012. Unifying cornerstones of sexual selection: operational sex ratio, Bateman gradient and the scope for competitive investment. Ecol. Lett. 15:1340–1351.
Kokko, H., and J. Mappes. 2013. Multiple mating by females is a natural outcome of a null model of mate encounters. Entomol. Exp. Appl. 146:26–37.
Ladurner, P., L. Schärer, W. Salvenmoser, and R. M. Rieger. 2005. 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.
Lange, R., K. Reinhardt, N. K. Michiels, and N. Anthes. 2013. Functions, diversity, and evolution of traumatic mating. Biol. Rev. 88:585–601.
Leonard, J. L. 1991. Sexual Conflict and the Mating Systems of Simultaneously Hermaphroditic Gastropods. Am. Malacol. Bull. 9:45–58.
Luther, A. 1947. Untersuchungen an rhabdocoelen Turbellarien VI. Macrostomiden aus Finnland. ACTA Zool. Fenn. 49:1–38.
Marie-Orleach, L., T. Janicke, and L. Schärer. 2013. Effects of mating status on copulatory and postcopulatory behaviour in a simultaneous hermaphrodite. Anim. Behav. 85:453–461.
Marie-Orleach, L., N. Vogt-Burri, P. Mouginot, A. Schlatter, D. B. Vizoso, N. W. Bailey, and L. Schärer. 2017. Indirect genetic effects and sexual conflicts: Partner genotype influences multiple morphological and behavioral reproductive traits in a flatworm.

Evolution (N. Y). 71:1232–1245.
Meade, A., and M. Pagel. 2016. Bayes Traits.
Michiels, N. K. 1998. Mating Conflicts and Sperm Competition in Simultaneous Hermaphrodites. Pp. 219–254 *in* Sperm Competition and Sexual Selection. Elsevier.
Mundry, R. 2014. Statistical issues and assumptions of phylogenetic generalized least squares. P. *in* Modern Phylogenetic Comparative Methods and their Application in Evolutionary Biology.
Myers, R. J. 1935. Behavior and morphological changes in the leech, *Placobdella parasitica*, during hypodermic insemination. J. Morphol. 57:617–653.
Nakadera, Y., E. M. Swart, J. N. A. Hoffer, O. den Boon, J. Ellers, and J. M. Koene. 2014. Receipt of Seminal Fluid Proteins Causes Reduction of Male Investment in a Simultaneous Hermaphrodite. Curr. Biol. 24:859–862.
Orme, D., R. Freckleton, G. Thomas, T. Petzoldt, S. Fritz, N. Isaac, and W. Pearse. 2014. Caper: Comparative analyses of phylogenetics and evolution in R.
Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proc. R. Soc. London. Ser. B Biol. Sci. 255:37–45.
Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.
Pagel, M., and A. Meade. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump markov chain monte carlo. Am. Nat. 167:808–825.
Parker, G. . 2006. Sexual conflict over mating and fertilization: an overview. Philos. Trans. R. Soc. B Biol. Sci. 361:235–259.
Parker, G. A. 1979. Sexual selection and sexual conflict. Pp. 123–166 *in* Sexual Selection and Reproductive Competition in Insects. Elsevier.
Parker, G. A. 1970. Sperm competition and its evolutionary consequences in the insects. Biol. Rev. 45:525–567.
Patlar, B., M. Weber, T. Temizyürek, and S. A. Ramm. 2020. Seminal Fluid-Mediated Manipulation of Post-mating Behavior in a Simultaneous Hermaphrodite. Curr. Biol. 30:143-149.e4.
Pélissié, B., P. Jarne, and P. David. 2012. Sexual selection without sexual dimorphism: Bateman gradients in a simultaneously hermaphrodite. Evolution (N. Y). 66:66–81.
Peretti, A. V., and W. G. Eberhard. 2010. Cryptic female choice via sperm dumping favours male copulatory courtship in a spider. J. Evol. Biol. 23:271–281.
Perry, J. C., and L. Rowe. 2018. Sexual conflict in its ecological setting. Philos. Trans. R. Soc. B Biol. Sci. 373:20170418.
Perry, J. C., and L. Rowe. 2015. The Evolution of Sexually Antagonistic Phenotypes. Cold Spring Harb. Perspect. Biol. 7:a017558.
Pizzari, T., and T. R. Birkhead. 2000. Female feral fowl eject sperm of subdominant males. Nature 405:787–789.
Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. {CODA}: Convergence Diagnosis and Output Analysis for {MCMC}. R News.
R Core Team. 2017. R: A language and environment for statistical computing. http://www.R-project.org/.
Ramm, S. a, A. Schlatter, M. Poirier, and L. Schärer. 2015. Hypodermic self-insemination as a reproductive assurance strategy. Proc. R. Soc. B Biol. Sci. 282:20150660.
Reinhardt, K., N. Anthes, and R. Lange. 2015. Copulatory Wounding and Traumatic Insemination. Cold Spring Harb. Perspect. Biol. 7:a017582.
Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.
Rice, W. R., and A. K. Chippindale. 2001. Intersexual ontogenetic conflict. J. Evol. Biol.

14:685–693.
Rice, W. R., A. D. Stewart, E. H. Morrow, J. E. Linder, N. Orteiza, and P. G. Byrne. 2006. Assessing sexual conflict in the *Drosophila melanogaster* laboratory model system. Philos. Trans. R. Soc. B Biol. Sci. 361:287–299.
Sayadi, A., A. Martinez Barrio, E. Immonen, J. Dainat, D. Berger, C. Tellgren-Roth, B. Nystedt, and G. Arnqvist. 2019. The genomic footprint of sexual conflict. Nat. Ecol. Evol. 3:1725–1730.
Schärer, L., T. Janicke, and S. A. Ramm. 2015. Sexual Conflict in Hermaphrodites. Cold Spring Harb. Perspect. Biol. 7:a017673.
Schärer, L., G. Joss, and P. Sandner. 2004. Mating behaviour of the marine turbellarian Macrostomum sp.: these worms suck. Mar. Biol. 145:373–380.
Schärer, L., D. T. J. Littlewood, A. Waeschenbach, W. Yoshida, and D. B. Vizoso. 2011. Mating behavior and the evolution of sperm design. Proc. Natl. Acad. Sci. 108:1490–1495.
Schenkel, M. A., I. Pen, L. W. Beukeboom, and J. Billeter. 2018. Making sense of intralocus and interlocus sexual conflict. Ecol. Evol. ece3.4629.
Singh, P., D. Ballmer, M. Laubscher, and L. Schärer. Submitted. Successful mating and hybridisation in two closely related flatworm species despite significant differences in reproductive morphology and behaviour.
Sluys, R. 1989. Sperm resorption in triclads (Platyhelminthes, Tricladida). Invertebr. Reprod. Dev. 15:89–95.
Snook, R. R., and D. J. Hosken. 2004. Sperm death and dumping in *Drosophila*. Nature 428:939–941.
Svensson, E. I. 2019. Eco-evolutionary dynamics of sexual selection and sexual conflict. Funct. Ecol. 33:60–72.
Vizoso, D. B., G. Rieger, and L. Schärer. 2010. Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. Biol. J. Linn. Soc. 99:370–383.
Weber, M., A. Giannakara, and S. A. Ramm. 2019. Seminal fluid-mediated fitness effects in the simultaneously hermaphroditic flatworm *Macrostomum lignano*. Ecol. Evol., doi: 10.1002/ece3.5825.
Westheide, W. 1999. Ultrastructure and functional significance of intestinojunctional spermathecae in enchytraeids (Oligochaeta, Annelida). Hydrobiologia 406:199–211.

Supporting information for "Evolution of the suck behaviour, a postcopulatory female resistance trait in a hermaphroditic flatworm genus".

Table S1. Data on observation records of reproductive behaviour in 64 species in the genus *Macrostomum*. For some species and drops, we did not observe any matings or suck behaviours.

Species	No. of worms	No. of hours	No. of replicates	No. of replicates with mating	Mating frequency (h^{-1}) (Mean ± sd)	Mating duration (s) (Mean ± sd)	No. of replicates with suck	Suck frequency (h^{-1}) (Mean ± sd)	Suck duration (s) (Mean ± sd)
M. axi	180	491.34	67	23	0.16 ± 0.08	200.87 ± 58.77	26	0.22 ± 0.15	14.21 ± 2.83
M. balticum	18	48.00	9	1	0.19	35.33	1	0.31	12.4
M. baoanensis	20	26.54	10	6	0.42 ± 0.11	30.28 ± 10.65	3	0.31 ± 0.06	8.33 ± 1.15
M. clavituba	88	235.03	38	26	0.47 ± 0.44	5.41 ± 1.76	21	0.4 ± 0.30	5.43 ± 0.52
M. cliftonensis	36	42.00	18	12	1.23 ± 1.29	35.59 ± 14.82	11	1.20 ± 1.03	11.74 ± 1.85
M. curvituba	114	62.00	35	15	0.58 ± 0.49	6.64 ± 2.24	13	0.58 ± 0.47	8.02 ± 1.53
M. distinguendum	26	31.00	9	1	0.1	15	-	-	-
M. finlandense	50	108.00	18	-	-	-	1	0.05	6
M. gabriellae	80	80.00	40	-	-	-	-	-	-
M. gieysztori	13	20.00	4	1	0.3	5.25	-	-	-
M. hamatum	63	192.00	28	9	0.24 ± 0.23	52.71 ± 19.22	11	0.61 ± 0.77	11.35 ± 1.58
M. hystricinum	7	8.00	1	-	-	-	-	-	-
M. hystricinum_d	24	24.00	12	-	-	-	-	-	-
M. hystrix	145	210.98	64	-	-	-	-	-	-
M. inductum	16	32.00	6	-	-	-	1	0.125	8
M. janickei	120	58.73	60	56	3.74 ± 1.92	81.94 ± 30.1	39	1.63 ± 1.08	9.74 ± 1.71
M. karlingi	5	16.00	1	-	-	-	1	0.01	8
M. kepneri	6	12.00	2	-	-	-	2	0.14 ± 0.03	8 ± 0
M. lignano	114	29.95	57	54	7.82 ± 4.37	14.99 ± 4.11	50	3.70 ± 2.54	8.21 ± 1.71
M. longituba	42	65.43	16	9	1.92 ± 2.38	36.31 ± 11.08	8	0.57 ± 0.55	10.14 ± 3.51
M. mediterraneum	26	20.88	11	9	3.13 ± 1.56	40.60 ± 11.14	5	0.9 ± 0.91	13.62 ± 1.11
M. minutum	33	29.00	7	-	-	-	-	-	-
M. mirumnovem	273	560.20	117	11	0.28 ± 0.11	940.5 ± 236.51	15	0.32 ± 0.23	15.71 ± 2.28
M. mystrophorum	9	21.69	4	-	-	-	-	-	-

M. paradoxum	138	194.00	40	1	0.1	15	13	0.08 ± 0.06	5.69 ± 1.33
M. poznaniense	36	178.31	13	1	0.03	20	-	-	-
M. pusillum	98	84.99	27	-	-	-	-	-	-
M. quiritium	18	25.78	3	-	-	-	-	-	-
M. retortum	10	37.12	4	-	-	-	-	-	-
M. rostratum	28	26.75	8	1	0.02	18	3	0.14 ± 0.09	5.58 ± 0.37
M. spirale	249	644.38	109	11	0.27 ± 0.25	294.12 ± 203.84	10	0.65 ± 0.84	16.08 ± 4.32
M. tenuicauda	32	59.26	15	9	0.20 ± 0.16	42.5 ± 17.67	6	0.13 ± 0.07	8.8 ± 1.52
M. tuba	43	146.41	18	12	0.57 ± 0.68	112.42 ± 46.32	5	0.21 ± 0.31	6.78 ± 1.59
Macrostomum sp. 1	12	10.01	6	-	-	-	-	-	-
Macrostomum sp. 2	8	12.00	2	2	0.08 ± 0.07	8 ± 1.41	2	0.37 ± 0.35	10.3 ± 0.99
Macrostomum sp. 4	16	48.89	4	-	-	-	-	-	-
Macrostomum sp. 16	2	1.33	1	-	-	-	-	-	-
Macrostomum sp. 17	6	6.00	2	-	-	-	1	0.67	4.75
Macrostomum sp. 22	12	12.00	6	-	-	-	-	-	-
Macrostomum sp. 30	18	27.92	6	1	0.04	1002	3	0.11 ± 0.08	10.67 ± 2.52
Macrostomum sp. 34	3	3.00	1	-	-	-	-	-	-
Macrostomum sp. 36	8	8.00	3	-	-	-	-	-	-
Macrostomum sp. 42	91	172.74	40	6	0.45 ± 0.23	802.55 ± 1212.71	3	0.28 ± 0.21	12 ± 1
Macrostomum sp. 43	2	6.02	1	1	0.08	4609	-	-	-
Macrostomum sp. 47	10	13.09	5	-	-	-	-	-	-
Macrostomum sp. 48	10	9.67	5	-	-	-	-	-	-
Macrostomum sp. 49	10	10.91	5	-	-	-	-	-	-
Macrostomum sp. 56	14	26.20	5	-	-	-	-	-	-
Macrostomum sp. 57	12	12.00	4	4	1.11 ± 0.70	26.14 ± 6.02	4	0.5 ± 0.23	8.51 ± 0.64
Macrostomum sp. 61	12	12.00	4	3	0.15 ± 0.06	15.67 ± 4.93	1	0.22	8
Macrostomum sp. 62	18	15.35	6	-	-	-	-	-	-
Macrostomum sp. 64	18	18.00	6	-	-	-	-	-	-
Macrostomum sp. 65	12	12.00	4	-	-	-	-	-	-
Macrostomum sp. 66	12	9.33	4	-	-	-	-	-	-
Macrostomum sp. 67	3	2.58	1	1	0.26	9	-	-	-
Macrostomum sp. 78	55	38.33	13	-	-	-	-	-	-
Macrostomum sp. 80	45	57.00	13	-	-	-	-	-	-
Macrostomum sp. 82	10	12.00	4	-	-	-	-	-	-

Macrostomum sp. 83	12	12.00	4	-	-	-	-	-	-
Macrostomum sp. 100	28	36.00	9	-	-	-	-	-	-
Macrostomum sp. 101	47	75.09	20	-	-	-	3	0.12 ± 0.05	9.67 ± 1.53
Macrostomum sp. 103	55	116.93	27	3	0.29 ± 0.19	13.61 ± 5.25	7	0.13 ± 0.06	9.33 ± 0.99
Macrostomum sp. 108	30	46.36	15	9	0.67 ± 0.50	10.69 ± 3.51	8	1.32 ± 1.06	12.331 ± 3.02
Macrostomum sp. 117	45	243.77	17	3	0.27 ± 0.11	9.69 ± 1.03	5	0.66 ± 1.08	9.70 ± 0.88

Table S2. Reciprocal mating behaviour for different species.

Species	Reciprocal mating behaviour
M. distinguendum	Usually the anterior dorsal side of both worms faced the same direction, while the posterior body was twisted at an angle such that the tail plates of the worms were in ventral contact with the female antrum accessible to the partner's stylet and vice-versa, allowing reciprocal transfer of ejaculate.
M. paradoxum	The entire body of the worms was intertwined with two turns of the body, with the tail plates in ventral contact such that the female antrum was accessible to the partner's stylet and vice-versa, which would allow reciprocal transfer of ejaculate, after some time they untwine slightly. At this point they remain attached at their tail plates ventrally, but the rest of the body is unattached, facing in opposite directions.
M. tuba	Similar to *M. distinguendum,* although the anterior ventral side of each worm remains attached to the substrate.
Macrostomum sp. 57	Similar to *M. paradoxum.*
Macrostomum sp. 61	The worms intertwine similar to M. *paradoxum*, but towards the end they only partially untwine, such that they resemble a pretzel with both worms' anterior part facing in the same direction.
Macrostomum sp. 67	The entire body of the worms was intertwined, with the tail plates in ventral contact such that the female antrum was accessible to the partner's stylet and vice-versa, which would allow reciprocal transfer of ejaculate,

Table S3. Phylogenetic signal (Pagel's λ) for the frequency and duration of both mating and suck (all log-transformed). For the λ-statistic, a log-likelihood ratio test was used to assess if the maximum-likelihood estimate of λ was significantly different from a λ = 0 model.

Trait	λ	P
Mating frequency	0.51	**0.05**
Mating duration	0.44	0.08
Suck frequency	0.68	**0.02**
Suck duration	0.75	**0.005**

Table S4. For duration and frequency of both mating and suck, we determined the model fit of different character evolution models (i.e. the Brownian motion, Ornstein–Uhlenbeck, Early-burst and Lambda models). The values of σ^2 (Brownian rate parameter), α (selection strength parameter), and *a* (rate of evolutionary change parameter) for the different models are given.

Effect	Brownian motion			Ornstein-Uhlenbeck				Early-burst				Lambda			
size	σ^2	AICc	ω_t	σ^2	α	AICc	ω_t	σ^2	a	AICc	ω_t	λ	σ^2	AICc	ω_t
Mating frequency	15.11	123.74	0.0002	19.47	2.72	116.17	0.007	15.11	-0.000001	126.22	0.00005	0.51	3.94	106.24	**0.99**
Mating duration	20.73	133.23	0.0017	26.75	2.72	125.7	0.075	20.73	-0.000001	135.71	0.00050	0.44	6.1	120.67	**0.92**
Suck frequency	6.15	101.63	0.0440	9.23	2.72	97.8	0.299	6.15	-0.000001	104.09	0.01284	0.68	2.87	96.26	**0.64**
Suck duration	0.41	17.9	0.0890	0.65	2.72	15.37	0.315	0.41	-0.000001	20.36	0.02602	0.75	0.22	14.19	**0.57**

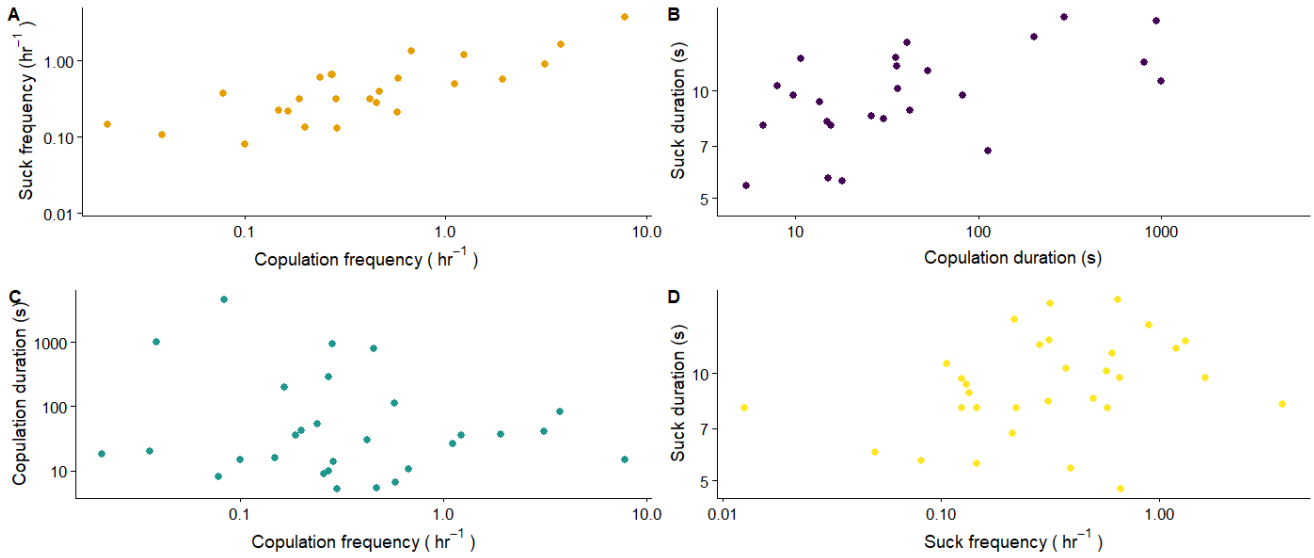
Table S5. Marginal likelihoods and Bayes Factor values of the independent and dependent models (three independent runs each) for examining the correlated evolution between a) the presence of reciprocal mating and the suck behaviour, and b) the behaviorally observed mating strategy and the morphologically inferred mating strategy syndrome, for the entire dataset and for the subsetted dataset.

		Entire dataset			Subsetted dataset		
		Independent	Dependent	Bayes Factor	Independent	Dependent	Bayes Factor
a) Presence of reciprocal mating and suck	Run 1	-88.67	-83.34	10.67	-75.35	-67.04	16.60
	Run 2	-88.67	-83.27	10.79	-75.34	-67.07	16.54
	Run 3	-88.66	-83.28	10.78	-75.36	-67.06	16.62
	Average	-88.67	-83.3	10.75	-75.35	-67.06	16.59
b) Behaviorally observed and morphologically inferred mating strategy	Run 1	-68.31	-65.13	6.35	-64.08	-61.22	5.73
	Run 2	-68.30	-65.08	6.44	-64.07	-61.24	5.67
	Run 3	-68.32	-65.14	6.36	-64.08	-61.18	5.81
	Average	-68.31	-65.12	6.38	-64.08	-61.21	5.74

Table S6. Summary of the PGLS results for the association between reciprocal mating and aspects of the suck behaviour for subset dataset. Note that PGLS was done on log-transformed variables.

Dependant variable	Predictor variable	λ	β±s.e.	t, df	P
Suck frequency	Reciprocal mating frequency	0.68	0.44±0.14	3.24, 18	0.004
Suck duration	Reciprocal mating duration	0.75	0.1±0.03	3.66, 18	0.002
Reciprocal mating frequency	Reciprocal mating duration	0.51	-0.02±0.16	-0.13, 19	0.89
Suck frequency	Suck duration	0.68	1.04±0.68	1.53, 23	0.14

Figure S1. Relationship between a) copulation frequency and suck frequency, and b) copulation duration and suck duration, c) copulation duration and frequency, and d) suck duration and frequency across 25 *Macrostomum* species. Note that log-transformed values have been plotted.



Chapter III

Variation in sex allocation plasticity in three closely related flatworm species

Published as
Singh, P., N. Vellnow, and L. Schärer. 2019.
Variation in sex allocation plasticity in three
closely related flatworm species.
Ecol. Evol. ece3.5566.

Received: 19 June 2019 | Revised: 15 July 2019 | Accepted: 28 July 2019

DOI: 10.1002/ece3.5566

ORIGINAL RESEARCH

Ecology and Evolution Open Access WILEY

Variation in sex allocation plasticity in three closely related flatworm species

Pragya Singh[1] | Nikolas Vellnow[1,2] | Lukas Schärer[1]

[1]Evolutionary Biology, Zoological Institute, University of Basel, Basel, Switzerland

[2]Evolutionary Biology Department, Bielefeld University, Bielefeld, Germany

Correspondence
Pragya Singh, Evolutionary Biology, Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland.
Email: pragya.singh42019@gmail.com

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A-162543

Abstract

Sex allocation (SA) theory for simultaneous hermaphrodites predicts an influence of group size on SA. Since group size can vary within an individual's lifetime, this can favor the evolution of phenotypically plastic SA. In an emerging comparative context, we here report on SA plasticity in three closely related *Macrostomum* flatworm species, namely *Macrostomum janickei*, *Macrostomum cliftonensis*, and *Macrostomum mirumnovem*. For each species, we experimentally raised worms in three group sizes (isolated, pairs, and octets) and two enclosure sizes (small and large) in all factorial combinations and studied the effects of these factors on different estimates of SA. In addition, we also evaluated whether isolated worms engage in self-fertilization. We found that all species have plastic SA, with *M. cliftonensis* being more plastic than the other two species, as assessed by comparing standardized effect sizes of (a) the presence/absence of mating partners and (b) the strength of sexual competition. Moreover, we found that sperm production rate—but not sperm morphology—is plastic in *M. cliftonensis*, and that only *M. mirumnovem* self-fertilized during our observation period. Our study suggests that both SA and SA plasticity can diverge even between closely related species.

KEYWORDS
local mate competition, local sperm competition, self-fertilization, simultaneous hermaphrodites, sperm morphology, sperm production rate

1 | INTRODUCTION

Sex allocation (SA) theory in simultaneous hermaphrodites (hermaphrodites henceforth) predicts an influence of mating group size on SA, with small mating group sizes favoring a female-biased SA and larger mating group sizes favoring a more equal SA (Charnov, 1982). This effect of mating group size on SA can be understood in terms of local sperm competition (LSC; Schärer, 2009; Schärer & Pen, 2013), which occurs when related sperm from the same sperm donor (or related sperm donors) compete for access to a partner's eggs. LSC can be thought of as the inverse of sperm competition, which in turn occurs when unrelated sperm compete for access to a partner's eggs (Parker, 1970, 1998). In the presence of high LSC—for example, under monogamy or obligate self-fertilization—the competing sperm are maximally related to each other, and thus, fitness returns for male investment increase at a strongly diminishing rate (Charlesworth & Charlesworth, 1981; Charnov, 1982; Schärer & Pen, 2013). This happens because any investment into the male function to produce more sperm than are required to fertilize the partner's eggs is a waste of resources, as the donor-related sperm simply compete among themselves. Under these conditions, a hermaphroditic individual can instead

maximize its fitness by reallocating investment from sperm production (male function) to the production of its own eggs (female function), thus favoring a female-biased SA under monogamy or obligate self-fertilization (Charlesworth & Charlesworth, 1981; Charnov, 1982; Schärer, 2009; Schärer & Pen, 2013). But as the mating group size increases, leading to more donors contributing sperm to each recipient, an individual's sperm competes more and more with unrelated sperm. Under these conditions, it pays off to invest into the male function to gain more fertilizations, favoring a subsequent shift toward a more equal SA (Charnov, 1982; Schärer, 2009; Schärer & Pen, 2013).

Most SA models predict the effect of mating group size and LSC in evolutionary terms, with SA representing an adaptation to the average mating group size experienced by a species over many generations and thus being modeled as a genetically fixed trait (but see Charnov, 1986). In contrast, most empirical tests of SA theory for hermaphrodites have evaluated these predictions based on the premise that SA can also be plastically adjusted, in response to environmental conditions experienced by the organism, either during their ontogeny (e.g., Janicke et al., 2013; Janicke, Sandner, Ramm, Vizoso, & Schärer, 2016; Schärer & Ladurner, 2003; Schärer & Wedekind, 2001; Tan, Govedich, & Burd, 2004) or during adulthood (e.g., Brauer, Schärer, & Michiels, 2007; Hart, Svoboda, & Cortez, 2011; Hoch & Levinton, 2012; Lorenzi, Sella, Schleicherová, & Ramella, 2005; Santi, Picchi, & Lorenzi, 2018; Schleicherová et al., 2014).

In our study, we also primarily focus on plastic adjustments of SA in response to the social environment, and more specifically, to different mating group sizes. The effect of mating group size on SA can be viewed in terms of reaction norms, with a steep slope of the reaction norm representing greater plasticity in SA compared to a shallower slope. Moreover, the reaction norm can be thought of as an evolved trait that can vary between species, and it can serve as a useful tool for studying evolution of plasticity (Stearns, 1992).

To estimate SA in hermaphrodites, resource allocation toward the male and female function needs to be quantified, and such allocation can be considered to have both static and dynamic components. While the static component involves allocation toward the establishment and maintenance of reproductive organs, such as testes and ovaries, the dynamic component involves allocation toward the actual production of gametes (and possibly secretion products of other reproductive glands), leading to the output of the reproductive organs (Schärer, 2009). Studies generally tend to focus on estimating the relatively easy to measure static reproductive organs and then simply assume a tight correlation between the static measures (e.g., testis size) and the dynamic measures (e.g., sperm production rate). While this assumption is likely often warranted to some degree, it clearly does not always hold perfectly, with dynamic measures sometimes exceeding the output expected from static measures (e.g., Giannakara, Schärer, & Ramm, 2016; Lüpold, Linz, Rivers, Westneat, & Birkhead, 2009; Ramm & Stockley, 2010; Schärer & Vizoso, 2007). Thus, it is interesting to measure both static and dynamic components of SA (and the link between them). For time reasons, we do only the former for two of the species studied here and both the former and latter for a third species.

Furthermore, phenotypically plastic changes in sperm production rate could, at least in theory, affect sperm morphology, since sperm size might trade-off with sperm quantity, in which case sperm competition might select more numerous and thus also smaller sperm (Immler et al., 2011; Kelly & Jennions, 2011; Snook, 2005). Additionally, an increase in sperm competition risk and/or intensity might under some conditions directly select for plasticity in sperm traits that increase the paternity share of a sperm donor (Pizzari & Parker, 2009; Snook, 2005). Several recent studies have shown phenotypic plasticity in sperm morphology and function (e.g., ascidians, Crean & Marshall, 2008; insects, Morrow, Leijon, & Meerupati, 2008; birds, Immler, Pryke, Birkhead, & Griffith, 2010; and flatworms, Janicke et al., 2016; but see Janicke & Schärer, 2010). Whether such sperm plasticity is widespread is unclear to date, which is why we explore this question in one of the species analyzed here (again for time reasons).

While SA plasticity in response to mating group size is widespread in hermaphroditic animals (reviewed in Schärer, 2009), the strength and nature of the plasticity vary both among and within species, with some studies showing plasticity only in the male function (Baeza, 2007; Hoch & Levinton, 2012; Janicke et al., 2013; Schärer & Janicke, 2009; Schärer & Ladurner, 2003; Winkler & Ramm, 2018) or the female function (Lorenzi et al., 2005; Schleicherová et al., 2014), while other studies show plasticity in either both functions (Janicke & Schärer, 2009, 2010) or no plasticity at all (Giannakara & Ramm, 2017). Although a part of this variation in these estimates of plasticity may reflect biases due to the ease with which male and female allocation can be measured in different study systems, it might also reflect different reproductive modes (with different postcopulatory processes, Schärer & Pen, 2013), species-specific costs of plasticity (Schleicherová et al., 2014), or different evolutionary histories (Schleicherová, Sella, & Lorenzi, 2013). Comparative investigations of SA plasticity, within a phylogenetic framework, can be used to infer the extent to which SA plasticity evolves between related species and to identify the correlates that might be shaping its evolution. This can be done by comparing reaction norms of SA in response to mating group size across related species. Ideally, such comparative studies should be done in a consistent experimental paradigm, which is what we report on here.

The free-living flatworm genus *Macrostomum* is an excellent model system for studying variation in SA plasticity, since reproductive organs can be measured accurately, repeatedly, and noninvasively in multiple species in the genus (e.g., Giannakara & Ramm, 2017; Schärer & Ladurner, 2003; Winkler & Ramm, 2018), allowing us to quantify and compare SA across closely related species. Earlier studies on the model organism *Macrostomum lignano* have shown that it has plastic SA, with SA increasing (i.e., becoming more male-biased) with increasing group size (e.g., Schärer & Ladurner, 2003). In addition, both sperm production rate (Schärer & Vizoso, 2007) and sperm length (Janicke et al., 2016) increase with group size (though the latter only by about 3%; see also Janicke & Schärer, 2010).

In light of an emerging comparative context, we here studied SA plasticity in three *Macrostomum* species, namely *M. janickei*, *M. cliftonensis*, and *M. mirumnovem*, which are all close relatives of the established model species *M. lignano* (Schärer et al., submitted). To manipulate the mating group size, we raised worms in three different social group sizes—isolated, pairs, or octets—given that social group size is a good proxy for the mating group size, as previously shown for *M. lignano* (Janicke et al., 2013). Since manipulating the group size also changes the density at which the worms live—which could itself affect the reproductive allocation through various routes (e.g., resource competition or higher concentration of metabolites, as discussed by Schärer & Ladurner, 2003)—we explored this by manipulating density independently of the group size, by having two enclosure sizes (small and large) for each of the three group sizes. We expected that, if there was an effect of the density, both an increase in group size and a decrease in enclosure size will have similar effects on SA. In addition, in order to facilitate comparisons between species, we also determined standardized effect sizes on SA plasticity. And finally, we asked whether the treatments had any effect on sperm morphology (for *M. cliftonensis*) and we checked whether isolated worms were able to self-fertilize.

2 | MATERIALS AND METHODS

2.1 | Study organisms

The three study species used here are all recently described free-living flatworms of the genus *Macrostomum* (Macrostomorpha, Platyhelminthes), and they include *M. janickei* (see also Zadesenets, Schärer, & Rubtsov, 2017; Zadesenets et al., 2016), *M. cliftonensis*, and *M. mirumnovem* (Schärer et al., submitted). Briefly, the worms used in the experiments were from laboratory cultures that were established using individuals collected from Palavas-les-Flots, near Montpellier, France, for *M. janickei*; from Lake Clifton, South of Perth, Western Australia, for *M. cliftonensis*; and from Port Phillip Bay, Queenscliff, Victoria, Australia, for *M. mirumnovem* (Schärer et al., submitted). Moreover, these species are all close relatives of our main model species, *M. lignano*, with *M. janickei* being the closest, as shown by molecular phylogenetic analyses (Schärer et al., submitted), using *M. hystrix* as an out-group species (see also Schärer, Littlewood, Waeschenbach, Yoshida, & Vizoso, 2011). All species are kept in mass cultures in the laboratory at 20°C in Petri dishes containing 32‰ artificial seawater (ASW) and fed with the diatom *Nitzschia curvilineata*.

2.2 | Experimental design and objective

In our experiments, we raised worms in different group sizes (i.e., isolated, pairs or octets) and enclosure sizes (i.e., 24- or 6-well tissue culture plates containing 1.5 or 6 ml of ASW, respectively) in all factorial combinations with ad libitum food and studied the effects of these treatments on reproductive allocation. The three treatments differ in their expected level of sperm competition, with the isolated treatment representing a nonreproductive state (or maximal LSC if worms self-fertilize), with the paired treatment representing maximal LSC (or somewhat lower LSC if worms also self-fertilize when in pairs, see below), and with the octet treatment representing low LSC (assuming worms engage in multiple mating; as shown for *M. lignano*, Janicke et al., 2013). Since the capability for self-fertilization could potentially affect the level of LSC, we assessed for all species whether isolated worms could self-fertilize. For all species, we measured body size (an estimate of overall size), ovary size (an estimate of female reproductive allocation), testis size (an estimate of male reproductive allocation), seminal vesicle size (an estimate of the amount of sperm available for transfer), and finally an estimate of SA (defined as testis size/(testis size + ovary size)). For *M. cliftonensis*, we further checked if the treatments had any effect on sperm production rate (by measuring the change in seminal vesicle size once a worm was isolated; cf. Schärer & Vizoso, 2007), and on different aspects of sperm morphology (cf. Janicke & Schärer, 2010).

Finally, using the obtained SA estimates, we compared the SA reaction norms between the different species. Specifically, we calculated standardized effect sizes to compare the effect of (a) the presence of mating partners (i.e., absent in isolated worms vs. present in worms in pairs and octets) and (b) the strength of sperm competition (i.e., low in pairs vs. high in octets) on SA plasticity in the different species.

2.3 | Experimental procedures

To obtain hatchlings of the same age, we allowed adult worms to lay eggs for periods of 3, 2, and 3 days for *M. janickei*, *M. cliftonensis*, and *M. mirumnovem*, respectively (note that for *M. mirumnovem* we used three consecutive batches to obtain sufficient offspring). Six days after removing the adults, the resulting hatchlings of each species were randomly assigned to a specific group size and enclosure size. We replicated each factor combination at least 30 times (yielding n = 180 replicates per species at least). Worms were transferred to fresh culture plates every week and allowed to mature. Starting from days 44, 48, and 57 after egg laying for *M. janickei*, *M. cliftonensis*, and *M. mirumnovem*, respectively, we, over a period of 3 days, morphometrically measured a range of traits (see below) in one randomly chosen worm per replicate. For *M. mirumnovem*, we measured each of the three batches on its respective day 57, so as to measure all the worms at the same age. For *M. cliftonensis*, we, after measurement, isolated the worms for a period of 4 days to obtain an estimate of the sperm production rate and sperm morphology (see below). Then, starting from day 52, we took images of the body size, seminal vesicle size, and the sperm, again over a period of three days.

Across all species, some replicates had to be excluded due to the worm being malformed (n = 17), one worm missing in isolated (n = 7) or paired (n = 7) treatment replicates, or some mortality while handling worms (n = 12). Furthermore, for the sperm production rate estimates in *M. cliftonensis*, we lost an additional 34 worms, either due to a missing worm or some mortality while handling the worms. For the sperm morphology of *M. cliftonensis*, we only measured a

subset of the worms for logistic reasons, chosen randomly such that all treatments were represented and culture plate effects avoided (see Figures 1 and 2 for final sample sizes).

2.4 | Morphometry

For morphometry, we anaesthetized worms using a 7:5 mixture of 71.4 g/L $MgCl_2$ solution and ASW and dorsoventrally squeezed them between a glass slide and a haemocytometer cover glass using standardized spacers (i.e., 40 µm for *M. janickei* and *M. cliftonensis*, and 45 µm for *M. mirumnovem* due to its somewhat larger body size; cf. Schärer & Ladurner, 2003). We then took images of the worms under a DM 2500 microscope (Leica Microsystems) using a digital camera (DFK41BF02; The Imaging Source) connected to a computer running BTV Pro 6.0b7 (Ben Software). The body size was imaged at 40×, while the testis, ovary, and seminal vesicle were imaged at 400× magnification. We analyzed these images using ImageJ (http://imagej.nih.gov/ij/) to obtain measures of body size, testis size (sum of both testes), ovary size (sum of both ovaries), and seminal vesicle size.

For each species, we obtained estimates of the repeatability of the above morphometric measurements, as given in Table S1, by imaging and measuring the same set of ~30 worms twice—randomly chosen from across all treatments—and then estimating the intraclass correlation coefficient (r_I) using the variance components from a one-way ANOVA with the R package "ICC" (Wolak, Fairbairn, & Paulsen, 2012).

For *M. cliftonensis*, we further estimated the sperm production rate, by measuring the increase in the seminal vesicle size in isolation between days 48 and 52 (see also Schärer & Vizoso, 2007). Moreover, we measured the sperm morphological traits of *M. cliftonensis* on days 52–55 using a procedure described in detail elsewhere (Janicke & Schärer, 2010). In brief, the tail plate of the worm was amputated with a scalpel and ruptured by transferring it in 1 µl of 32‰ ASW onto a glass slide and then covered with a coverslip (21 × 26 mm), thus rupturing the seminal vesicle and causing sperm to flow out where they could be imaged and measured approximately in a 2D plane. Similar to *M. lignano*, the sperm of *M. cliftonensis* has a complex design with a rapidly undulating anterior extension (termed feeler), followed by the sperm body, the sperm shaft, and the sperm brush at the end. In addition, there is a pair of stiff lateral bristles—anchored at the junction of the sperm body and shaft—which is hypothesized to represent a male persistence trait that allows the sperm to remain anchored in the female genitalia (see also Janicke et al., 2016; Schärer et al., 2011; Vizoso, Rieger, & Schärer, 2010). We imaged the sperm at 1,000× as above and measured the total sperm length (sum of length of sperm body and sperm shaft), sperm body length, and sperm bristle length of 10 sperm per individual (Janicke & Schärer, 2010), using the mean trait values per individual in the subsequent analyses.

2.5 | Self-fertilization

We assessed if the species can self-fertilize by looking for offspring in the wells where the worms were being kept before being measured. We did this 10 days after the adults had been removed in order to give the eggs sufficient time to hatch. If a species could self-fertilize, we expected the isolated treatments to also contain hatchlings.

2.6 | Statistical analyses

To evaluate whether group size and enclosure size had any effect on body size, we calculated a two-way (3 × 2) ANOVA with group size, enclosure size, and their interaction as fixed factors (and for *M. mirumnovem*, we also included batch number as a fixed factor). Furthermore, for testis, ovary, and seminal vesicle size we calculated the corresponding ANCOVAs, including body size as a covariate (since these traits were correlated with body size). In contrast, SA and the different sperm traits (i.e., total sperm length, sperm body length, and sperm bristle length) were analyzed without that covariate, since SA already represents a relative measure and the sperm traits were not found to be correlated with body size. To assess effects on sperm production rate in *M. cliftonensis*, we calculated the residuals of a linear regression of seminal vesicle size on the corresponding body size for days 48 and 52 (as estimates for the relative seminal vesicle size on that day), and then used the increase in these residuals (i.e., day 53 minus day 48) as the dependent variable, and performed the same two-way ANOVA as above.

For all reproductive parameters for which the analyses suggested a significant group size effect, we calculated post hoc Tukey HSD tests to determine which levels differed significantly from each other. The data were graphically checked for the assumptions of parametric test statistics and transformed where needed, using log-transformation for testis and ovary size (in all species); body size (in *M. mirumnovem*); and total sperm length, seminal vesicle size, and SA (in *M. cliftonensis*).

Finally, in order to facilitate comparisons of the level of SA plasticity in different *Macrostomum* species, we calculated standardized effect sizes using Cohen's *d* (Cohen, 1988) and confidence intervals (Howell, 2011) with the R package "effsize" (Torchiano, 2017). Specifically, in order to assess the effect of the presence of mating partners on SA we compared SA in isolated worms with that in pairs and octets (since worms in both pairs and octets have mating partners). Additionally, we assessed the effect of the strength of sperm competition by comparing SA in pairs with that in octets, with pairs representing high LSC and octets representing low LSC. All statistical analyses on our data were carried out using R, version 3.4.1 (R Core Team, 2017).

2.7 | Prediction for SA response to group size

We expect that, with an increase in group size, sperm competition increases (and LSC decreases), as a result of which worms either mate more often or transfer more sperm per mating, leading to a higher sperm spending and a lower filling grade of the seminal vesicle. This then leads to a higher sperm production, in order to make up for the sperm depletion, which in turn favors an increased testis size and/or sperm production rate. Moreover, we expect that sperm production is likely not high enough to completely make up for sperm depletion,

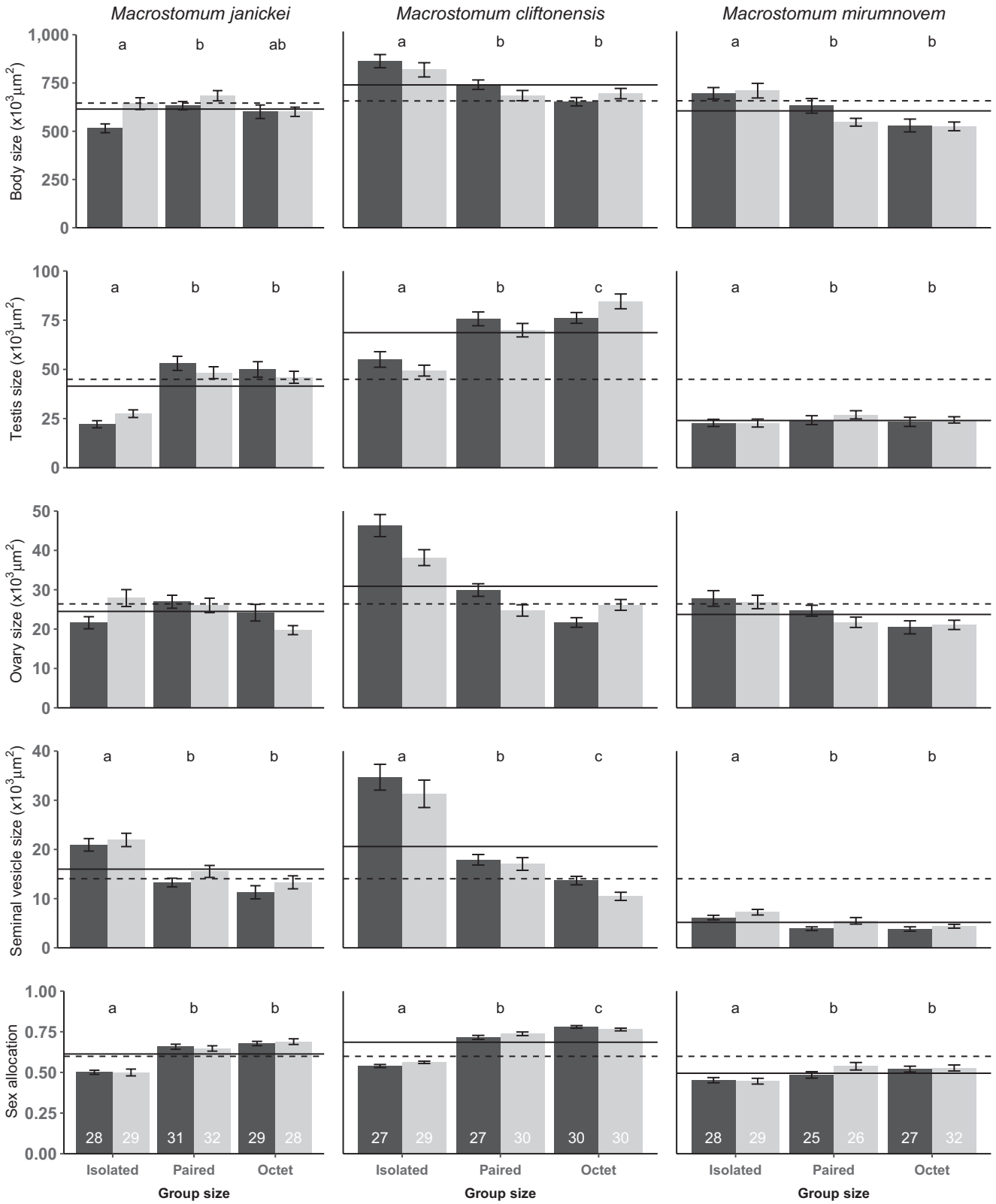


FIGURE 1 Effect of group size (*x*-axis) and enclosure size (black bars: large; gray bars: small) on body size, testis size, ovary size, seminal vesicle size, and sex allocation in *Macrostomum janickei*, *Macrostomum cliftonensis*, and *Macrostomum mirumnovem*. Different letters denote significantly different group size effects inferred from Tukey HSD post hoc tests (not done for ovary size of *M. janickei* and *M. mirumnovem*, and done separately for small and large enclosures for ovary size of *M. cliftonensis*, see text). The solid line represents the mean for each species, and the dashed line represents the grand mean across all species. The plots show mean and standard error of raw data, but note that log-transformed data were used for statistical analysis of testis and ovary size (in all species), body size (in *M. mirumnovem*), seminal vesicle size, and SA (in *M. cliftonensis*). Sample sizes for each treatment are given at the bottom

which should lead to a lower seminal vesicle size with increased group size, reflecting higher spending. Furthermore, under the assumption of a trade-off between male and female reproductive allocation, we might expect a corresponding decrease in the ovary size as a result of the predicted increase in testis size, and thus an increasingly male-biased SA as group size increases (although this SA trade-off can be difficult to observe, cf. Picchi & Lorenzi, 2019; Schärer, 2009; Schärer, Sandner, & Michiels, 2005). Finally, we explore whether an increase in group size affects sperm morphology, either leading to larger sperm, since a large size could be advantageous in the context of high sperm competition, or potentially smaller sperm, since numerical competition might become more important as group size increases.

3 | RESULTS

As we show in the following, all the studied *Macrostomum* species show phenotypic plasticity in SA, although they clearly do so to varying degrees. We first present the results for sex allocation plasticity in the three species separately, followed by the results about self-fertilization and a comparison of the observed effect sizes.

3.1 | Sex allocation plasticity in *M. janickei*

In *M. janickei*, body size was significantly affected by group size, with paired worms being significantly larger than isolated worms, while octet worms were intermediate in size and did not differ significantly from the other two groups (Table 1, Figure 1). In addition, worms in small enclosures were bigger compared to those in large enclosures, although care should be taken in interpreting these post hoc results, considering that the *p*-value of the interaction term is relatively close to the significance level. In accordance with our expectations, testis and seminal vesicle size were strongly and significantly affected by group size, with isolated worms having significantly smaller testes and larger seminal vesicles compared to pairs and octets, while the enclosure size and interaction term had no significant effects, indicating that worms in larger groups invested more in the male function. The ovary area was not significantly affected by either group size, enclosure size, or their interaction. As seen for the testis size, SA was significantly affected by only the group size, with isolated worms being significantly different from pairs and octets. However, pairs and octets did not differ significantly in either testis size or SA, both of which we had expected based on predictions of SA theory.

3.2 | Sex allocation and sperm morphology plasticity in *M. cliftonensis*

In *M. cliftonensis*, body size was significantly affected only by the group size, with isolated worms being significantly larger than worms in pairs and octets (Table 2, Figure 1). As predicted by SA theory, testis size increased significantly with group size, here differing significantly also between pairs and octets. The ovary size was significantly affected by both the group size and the interaction term, making it somewhat difficult to interpret the results biologically. We therefore performed the Tukey HSD post hoc tests separately for the small enclosures (isolated vs. pair, $p = .01$; pair vs. octet, $p = .98$; isolated vs. octet, $p = .07$) and large enclosures (isolated vs. pair, $p = .01$; pair vs. octet, $p = .04$; and isolated vs. octet, $p < .001$). In general, investment into ovaries was higher in smaller groups and larger enclosures, suggesting a trade-off between the male and female allocation. Both group size and enclosure size had significant effects on seminal vesicle size (with the effect of the former being greater than that of the latter), with worms in larger groups and smaller enclosures having smaller seminal vesicles. And matching our expectation that increased sperm competition favors increased sperm production, sperm production rate (the increase in seminal vesicle size during the four days of isolation) was also significantly higher in larger groups, although care should be taken in interpreting these post hoc results, as the *p*-value of the interaction term is relatively close to the significance level (Figure 2a). And finally, SA was affected by group size, with worms in larger groups having a higher SA in general, though the post hoc tests should be interpreted with some caution considering the marginally significant interaction term, which may stem from the

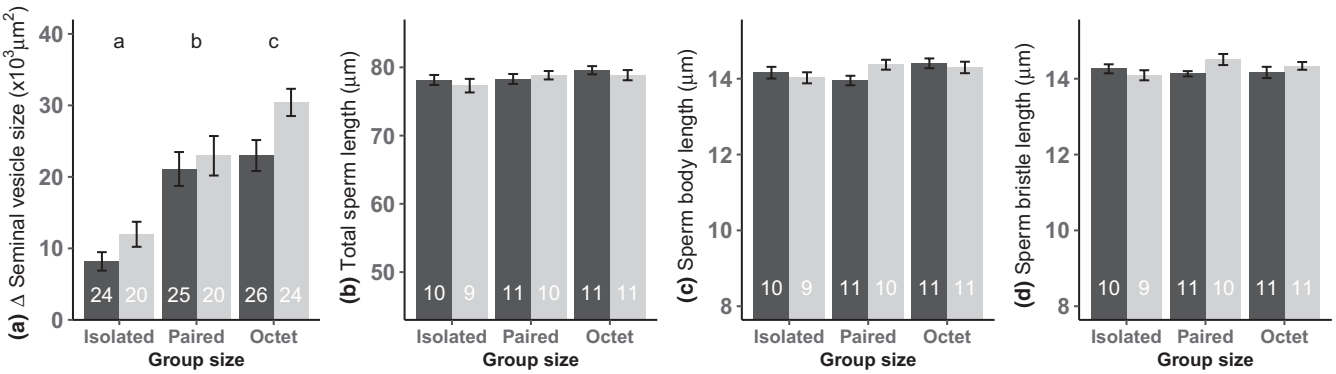


FIGURE 2 Effect of group size (*x*-axis) and enclosure size (black bars, large; gray bars, small) on (a) increase in seminal vesicle size, (b) total sperm length, (c) sperm body length, and (d) sperm bristle length in *Macrostomum cliftonensis*. Different letters denote significantly different group size effects inferred from Tukey HSD post hoc tests. The plots show means and standard errors of raw data, though log-transformed data were used for statistical analysis of total sperm length. Sample sizes for each treatment are given at the bottom

above-mentioned interaction term for ovary size (Table 2, Figure 1). None of the sperm morphological traits in *M. cliftonensis* were significantly affected by the group size, enclosure size, or their interaction (Table 2, Figure 2b–d), suggesting that phenotypic plasticity in testis size and sperm morphology do not correlate.

3.3 | Sex allocation plasticity in *M. mirumnovem*

Similar to the results in the previous species, body size in *M. mirumnovem* was significantly affected only by the group size, with isolated worms being larger than worms in pairs and octets (Table 3, Figure 1). Again, and in agreement with our expectations, group size significantly affected testis size, with isolated worms having smaller testes than worms in pairs and octets, while pairs and octets did not differ significantly, though the interaction was relatively close to the significance level, so the post hoc results should be interpreted with some caution. The batch number also significantly affected testis size, with worms from batch 3 having significantly smaller testis (Tukey HSD test: Batch 1 vs. 2, $p = .28$; Batch 2 vs. 3, $p = .0008$; and Batch 3 vs. 1, $p < .0001$). Isolated worms had larger seminal vesicles than the other two treatments, in accordance with our expectations; however, pairs and octets did not differ significantly in this parameter. Moreover, enclosure size also had a significant effect, with worms in larger enclosures having smaller seminal vesicles, which was contrary to our expectations, since, if anything, we expected worms in smaller enclosures to donate more sperm. The ovary size was not significantly affected by any of the factors. Note that although a visual inspection appears to suggest a trend in *M. mirumnovem* for ovary size decreasing with increasing group size (Figure 1), this effect was not significant when correcting for body size, which shows a similar pattern. Similar to the testis, SA was significantly affected by group size and batch number, with isolated worms being significantly different from pairs and octets and batch 3 having significantly lower SA (Tukey HSD test: Batch 1 vs. 2, $p = .19$; Batch 2 vs. 3, $p = .003$; and Batch 3 vs. 1, $p < .0001$, Table 3). Worms from pairs and octets also did not differ significantly in SA.

3.4 | Self-fertilization

Only *M. mirumnovem* was found to self-fertilize during our observation period, with offspring being produced in 17 out of the 56 isolated replicates, while in a comparable number of isolated replicates the other two species never had any offspring (Table S2). Conversely, we found offspring in nearly all replicates containing pairs and in all replicates containing octets, irrespective of the species (Table S2), clearly suggesting that our holding conditions were generally favorable for reproduction.

3.5 | Effect of presence of mating partner and local sperm competition

Overall, the presence of mating partners had a larger effect on SA than the strength of sperm competition in all the studied species, and of these species, *M. cliftonensis* was the most plastic, followed by *M. janickei* and *M. mirumnovem* (Table 4). This implies that the shift in our estimate of SA from the isolated to the pair/octet treatment is of a higher magnitude compared to the increase from pairs to octets.

4 | DISCUSSION

In agreement with earlier results obtained for the model organism *M. lignano* (Brauer et al., 2007; Janicke et al., 2013; Janicke & Schärer, 2009; Schärer & Ladurner, 2003; Schärer et al., 2005), we found that all the *Macrostomum* species analyzed here exhibit SA plasticity, with testes being more plastic than ovaries. Interestingly, despite the fact that *M. cliftonensis* is more distantly related to *M. lignano* than *M. janickei*, it showed plasticity patterns that were more similar to those of *M. lignano*. Specifically, SA increased significantly from pairs to octets—as previously seen in *M. lignano*—while the much more closely related *M. janickei* was less plastic, with pairs and octets having similar SA. This result suggests that SA plasticity can evolve readily, even in closely related species. Moreover, as in the earlier studies on *M. lignano*, it is group size rather than enclosure size that is primarily responsible for the observed shifts in SA, with the latter either not being significant or having smaller effects compared to the former on SA (as is evident from the much higher *F*-values for group size compared to enclosure size across the species; Tables 1–3). This suggests that the observed shifts in SA are not simply caused by density effects, but that they largely result from changes in the number of

TABLE 1 Effects of the fixed factors (group size, enclosure size, and their interaction) and the covariate (body size) on the dependent variables for *Macrostomum janickei*

	Group size			Enclosure size			Interaction			Body size		
Trait	***dfs***	***F***	***p***	***dfs***	***F***	***p***	***dfs***	***F***	***p***	***dfs***	***F***	***p***
Body size	2,171	4.61	**.01**	1,171	7.2	**.01**	2,171	2.73	.07	–	–	–
Testis size	2,170	63.2	**<.001**	1,170	1.36	.25	2,170	1.23	.3	1,170	96.9	**<.001**
Ovary size	2,170	1.95	.15	1,170	0.57	.45	2,170	1.19	.31	1,170	33.5	**<.001**
Seminal vesicle size	2,170	30.2	**<.001**	1,170	2.31	.13	2,170	0.25	.78	1,170	0.22	.64
SA	2,171	70.4	**<.001**	1,171	0.001	.97	2,171	0.24	.79	–	–	–

Note: Significant *p*-values are highlighted in bold.
Results of ANOVA (body size and SA) and ANCOVA (all others) are shown.

TABLE 2 Effects of the fixed factors (group size, enclosure size, and their interaction) and the covariate (body size) on the dependent variables for *Macrostomum cliftonensis*

	Group size			Enclosure size			Interaction			Body size		
Trait	***dfs***	***F***	***p***	***dfs***	***F***	***p***	***dfs***	***F***	***p***	***dfs***	***F***	***p***
Body size	2,167	18.15	**<.001**	1,167	0.62	.43	2,167	1.79	.17	–	–	–
Testis size	2,166	143.55	**<.001**	1,166	0.03	.85	2,166	0.82	.44	1,166	31.77	**<.001**
Ovary size	2,166	27.33	**<.001**	1,166	1.15	0.28	2,166	5.99	**.003**	1,166	204.62	**<.001**
Seminal vesicle size	2,166	45.41	**<.001**	1,166	5.03	**.03**	2,166	1.38	.26	1,166	36.52	**<.001**
Increase in seminal vesicle size	2,133	78.08	**<.001**	1,133	2.8	.1	2,133	2.63	.08	–	–	–
SA	2,167	306.47	**<.001**	1,167	2.02	.16	2,167	2.81	.06	–	–	–
Total sperm length	2,56	1.95	.15	1,56	0.31	.58	2,56	0.56	.57	–	–	–
Sperm body length	2,56	1.9	.16	1,56	0.29	.6	2,56	2.48	.09	–	–	–
Sperm bristle length	2,56	0.55	.58	1,56	1.87	.18	2,56	2.47	.09	–	–	–

Note: Significant *p*-values are highlighted in bold.
Results of ANOVA (body size and SA) and ANCOVA (all others) are shown.

interacting individuals (cf. Schärer & Ladurner, 2003). In the following, we discuss these results in some more detail.

In our study, we found that the response of body size to the treatments differed across species, with paired worms being larger than isolated worms for *M. janickei*, while for *M. cliftonensis* and *M. mirumnovem* the isolated worms were larger than paired and octet worms. While a potential reason for body size decreasing in larger groups could be resource competition, in our experiment the worms were fed ad libitum, which is why we consider this an unlikely explanation. Interestingly, also in earlier studies on *M. lignano* different trends for body size have been observed. Some studies showed that worms grow larger with increased group size (Brauer et al., 2007; Janicke et al., 2013), while another showed that isolated worms were smaller than worms in pairs (Schärer & Janicke, 2009), and still others showed that neither group nor enclosure size had an effect on body size (Janicke & Schärer, 2009; Schärer & Ladurner, 2003). While this could potentially complicate the interpretation of our results, since body size has been shown affecting SA in simultaneous hermaphrodites (Schärer, 2009; Vizoso & Schärer, 2007), in *M. lignano* SA (and particularly the testis size component of the SA estimate) was similarly plastic and increased with group size across all studies, irrespective of the trend for body size that was observed, pointing toward an effect of group size on SA that is largely independent of body size.

Furthermore, in our study testis size—our primary estimate for male reproductive allocation—increased with an increase in group size in a way similar to *M. lignano*, and in line with the prediction of SA theory (Charnov, 1982), though not to the same extent across the three species. In contrast, in all three species the ovary size did not appear to be as plastic as testis size. More specifically—and in contrast to the prediction of SA theory—there was no significant effect of group size on ovary size in *M. janickei* and *M. mirumnovem*. And while group size had a significant effect on ovary size for *M. cliftonensis*, the significant interaction term means that some care is needed to interpret that main effect. Briefly, although there seemed to be a trend for decrease in ovary size as the group size increased, isolated and paired treatment worms had a smaller ovary size in smaller enclosures, while octet treatment worms had a larger ovary size in smaller enclosures. One potential explanation for not observing a corresponding decrease in ovary size, in spite of the often marked increase in testis size, could be that ad libitum feeding masks the trade-off between male and female reproduction function (cf. Schärer et al., 2005). Furthermore, male reproduction may not trade-off with female reproduction, but with other life-history traits (Schärer, 2009). However, there are several lines of evidence to suggest that a trade-off between male and female allocation exists, at least in *M. lignano* (Janicke & Schärer, 2009; Schärer et al., 2005). Alternatively, ovary size may be an inferior proxy for measuring investment into the female function compared to testis size for the male function. This seems plausible, since at least part of the egg development, such as the addition of yolk and shell granules (Gremigni, 1988; Gremigni, Falleni, & Lucchesi, 1987), occurs outside the transparent portion of the ovary, which is what we generally use to estimate ovary size in *Macrostomum*.

We predicted that seminal vesicle size would decrease with increased group size (as worms likely mate more often and/or transfer more sperm per mating), and this is indeed what we found in all three species, though again to varying degrees. In all species, seminal vesicle size decreased significantly going from isolated worms to worms in pairs and octets, while only in *M. cliftonensis* did it further decrease significantly going from pairs to octets. Interestingly, these results mirror the observed differences in testis size between pairs and octets across the three species, which may support our notion that sperm production and testis size respond to the fill grade of the seminal vesicle. Moreover, in *M. cliftonensis* the enclosure size had a significant effect, with worms in smaller enclosures having smaller seminal vesicles. This could potentially be explained by considering that worms in smaller enclosures may encounter and mate with other individuals at a higher rate. Interestingly, we see the converse effect in *M. mirumnovem*, with worms in larger enclosures having smaller seminal vesicles, but we currently do not know of a plausible hypothesis for this observation.

TABLE 3 Effects of the fixed factors (group size, enclosure size, and their interaction) and the covariate (body size) on the dependent variables for *Macrostomum mirumnovem*. Note that we here also included the batch effect as a fixed factor

Trait	Group size			Enclosure size			Interaction			Batch effect			Body size		
	dfs	*F*	*p*	*dfs*	*F*	*p*	*dfs*	*F*	*p*	*dfs*	*F*	*p*	*dfs*	*F*	*p*
Body size	2,158	17.33	**<.001**	1,158	0.59	.44	2,158	0.98	.38	2,158	1.7	.19	–	–	–
Testis size	2,157	11.95	**<.001**	1,157	2.72	.1	2,157	2.71	.07	2,157	12.36	**<.001**	1,157	69.23	**<.001**
Ovary size	2,157	0.28	.76	1,157	0.08	.78	2,157	0.82	.44	2,157	0.38	.69	1,157	227	**<.001**
Seminal vesicle size	2,157	8.08	**<.001**	1,157	9.16	**.003**	2,157	1.17	.31	2,157	1.56	.21	1,157	27.8	**<.001**
SA	2,158	10.1	**<.001**	1,158	1.59	.21	2,158	1.71	.19	2,158	11.02	**<.001**	–	–	–

Note: Significant *p*-values are highlighted in bold.
Results of ANOVA (body size and SA) and ANCOVA (all others) are shown.

TABLE 4 Standardized effect sizes (Cohen's *d*) for the effect of the presence of mating partners and sperm competition on SA

Species	Mating partner	Sperm competition
Macrostomum janickei	1.88 (1.51 to 2.26)	0.36 (0 to 0.72)
Macrostomum cliftonensis	3.73 (3.22 to 4.24)	0.82 (0.44 to 1.2)
Macrostomum mirumnovem	0.70 (0.36 to 1.03)	0.12 (−0.26 to 0.5)

Note: The 95% confidence intervals are given in brackets.

As predicted by sex allocation theory (Charnov, 1982), SA increased with increasing group size, as worms start investing more in their male function and possibly somewhat less in their female function. However, the SA effect we see here likely stemmed primarily from the above-mentioned variation in testis size, since the ovaries did not appear to be very plastic.

Generally, changes in SA should always be interpreted in the light of the changes observed in the different components used to estimate it (Schärer, 2009). In our study, we estimated SA as the proportion of testis size to overall gonad size (cf. Janicke et al., 2013; Janicke et al., 2016; Janicke & Schärer, 2009; Schärer & Janicke, 2009; Vellnow, Vizoso, Viktorin, & Schärer, 2017; Vizoso & Schärer, 2007), and we thus assume that testes and ovaries are comparably useful proxies for investment into male and female reproductive function. While estimating SA in this way allows us to compare relative differences in allocation to testes and ovaries between the different group sizes, it does not provide an absolute estimate of SA for a number of reasons. Firstly, while both testes and ovaries are involved in gamete production, the energetic expenditure per unit tissue could certainly differ between the two organs. Secondly, while testis size has been experimentally well-verified to be quite a good proxy for male expenditure on sperm production rate (Giannakara et al., 2016; Schärer, Ladurner, & Rieger, 2004; Schärer & Vizoso, 2007), the support for ovary size being a good proxy for female reproductive investment is somewhat more limited (Schärer et al., 2005), as already mentioned above. And finally, there are certainly other components of both male and female reproduction, for example, sex-specific behaviors, prostate glands, seminal fluids, copulatory organs, egg yolk production, eggshell glands, and eggshell material, that could potentially be plastic as well, and which we did not quantify here (Janicke et al., 2016; Patlar, Weber, & Ramm, 2019; Picchi & Lorenzi, 2019; Schärer, 2009; Schärer & Pen, 2013). Thus, our measures of male and female reproduction are not absolute, which is also reflected by the fact that our estimates of SA sometimes exceed 0.5 (Figure 1), which is not necessarily suggestive of a male-biased SA. However, it appears likely that a greater sperm production will tend to go along with a greater expenditure on other male components (i.e., more ejaculate requires both more sperm and more seminal fluid), and similarly for egg production and female components (i.e., more eggs require both more oocytes and more yolk and shell material).

In one species, *M. cliftonensis*, we also looked at the sperm production rate, by measuring the increase in seminal vesicle size after isolating the worms, and found that it was significantly higher in larger group sizes. Thus, worms in larger groups produced sperm at a higher rate, likely to replenish the greater amounts of sperm that was being transferred under these conditions. Furthermore, despite the considerable plasticity in testis size and sperm productivity, the sperm morphology in *M. cliftonensis* did not show significant plasticity in response to variation in group size (or enclosure size). A recent study in *M. lignano* (with about 90 replicates per group size, Janicke et al., 2016) showed that worms raised in octets produce (about 3%) longer sperm than worms raised in pairs, while an earlier study in *M. lignano* (n = 24 replicates per group size, Janicke & Schärer, 2010) found no significant plasticity in sperm morphology in response to group size, possibly since its statistical power was only expected to reliably detect effect sizes above 5%. Similar to the latter study, we here have a sample size of about 21 replicates per group size. Thus, it is possible that we may have failed to detect very small differences in sperm morphology between the treatments, though how biologically relevant such minor differences in sperm morphology are, is still an open question (see Janicke et al., 2016). In summary, the results of this study suggest that the genus has at most slight plasticity in sperm morphology in response to group size, particularly in comparison with the marked plasticity in a trait like testis size.

With respect to our estimates of standardized effect sizes, there was overall a stronger effect of the presence of a mating partner (i.e., isolated vs. pairs and octets) than the strength of sperm competition (i.e., pairs vs. octets) on SA for all three species (Table 4). This points toward relatively few resources being invested in sperm production (testes) in the absence of partners when no mating or fertilization can occur (unless by self-fertilization), similar to *M. lignano* (Schärer & Janicke, 2009), while individuals upregulate sperm production in presence of partners either in response to the potential for mating or due to sperm expenditure occurring during mating. Another possibility could be that in the absence of partners there is little/no sperm expenditure such that the seminal vesicle fills up (Figure 1), which in turn may act as a signal to downregulate sperm production in isolated worms. Furthermore, at least in *M. lignano*, worms seem to be unable to distinguish between familiar and novel partners (Sandner & Schärer, 2010). This could mean that in pairs, even the single partner may be perceived as representing some sperm competition risk (though sperm competition is absent unless the species self-fertilizes), leading to an increase in allocation into the male function.

Interestingly, we found that *M. mirumnovem* could self-fertilize, since hatchlings were present in close to one third of the wells with isolated worms (who had grown up in isolation from hatchlings and were therefore virgins). The genus *Macrostomum* contains species exhibiting at least two different modes of reproduction (Schärer et al., 2011). The hypodermic mating syndrome involves presumably unilateral hypodermic mating, with sperm being injected via a sharp needle-like male copulatory organ (stylet) into the parenchyma of the partner worm. In contrast, the reciprocal mating syndrome involves reciprocal copulation (e.g., *M. lignano*, see Schärer, Joss, & Sandner, 2004), during which both partners reciprocally transfer sperm into the female antrum (sperm receiving organ) of the partner. Both earlier reports of self-fertilization in *Macrostomum* are in species that are hypodermically inseminating (e.g., *M. hystrix*, see Ramm, Vizoso, & Schärer, 2012; Ramm, Schlatter, Poirier, & Schärer, 2015; and *M. pusillum*, Giannakara & Ramm, 2017), with both species showing a similar needle-like stylet morphology that potentially facilitates self-fertilization by allowing self-injection of sperm. In contrast, the current study documents the first occurrence of self-fertilization in a reciprocally copulating species, *M. mirumnovem*, which has a large blunt-ending stylet (Schärer et al., submitted), and for which we currently do not understand how selfing is actually achieved.

Being able to self-fertilize would lead to high local sperm competition in isolated worms, thus favoring a lower allocation into the testis if self-fertilization occurs frequently (Charlesworth & Charlesworth, 1981; Charnov, 1982; Schärer, 2009). In support of this, *M. mirumnovem* indeed has the lowest absolute testis and seminal vesicle size, as well as the lowest SA, compared to the other species (i.e., compare the species means to the grand mean across all species in Figure 1). A potential explanation for this could be that we used slightly thicker spacers for the standardized measurement of *M. mirumnovem* (i.e., 45 µm compared to 40 µm in *M. janickei* and *M. cliftonensis*), thus potentially leading to an underestimation of trait sizes. However, if that were the case, then all traits, including ovary size, should have been underestimated, but ovary size of *M. mirumnovem* was similar to that in the other species. Furthermore, the trend still holds even when looking at estimates of trait (testis, ovary, and seminal vesicle) size relative to body size for all species (data not shown). Thus, the low allocation into the male function may possibly be due to the fact that self-fertilization in *M. mirumnovem* occurs with appreciable frequency. Interestingly, an earlier study in *M. pusillum*—a species that regularly self-fertilizes—has shown that the species does not exhibit any plasticity in SA in response to group size (Giannakara & Ramm, 2017). In contrast, *M. hystrix*—a facultatively self-fertilizing species—does adjust its SA depending on group size (Winkler & Ramm, 2018). This suggests that the degree of self-fertilization could also influence sperm competition and therefore SA and its plasticity. In conclusion, to better understand the evolution of SA and the nature and extent of SA plasticity in the genus *Macrostomum*, more species need to be investigated, thus facilitating comparative studies of the evolution of SA and SA plasticity. This will allow us to understand how different factors, such as group size and the degree of self-fertilization, can affect SA plasticity.

ACKNOWLEDGMENTS

We would like to thank Christian Felber for help with the morphological measurements of *M. janickei*, Tim Janicke, Georgina Rivera Ingraham, and Jeremias Brand for help with collecting specimens of the study species, and Gudrun Viktorin, Jürgen Hottinger, Urs Stiefel and Lukas Zimmermann for technical support. This research was supported by Grant 31003A-162543 of the Swiss National Science Foundation (SNSF) to LS.

CONFLICT OF INTERESTS

None declared.

AUTHOR CONTRIBUTIONS

PS and LS designed the experiments. PS performed the experiments with some help from NV. PS analyzed the images. PS, NV, and LS statistically analyzed the data. PS and LS wrote the manuscript. All authors have read and approved the final version.

DATA AVAILABILITY STATEMENT

All data in this manuscript will be deposited on Zenodo (https://doi.org/10.5281/zenodo.3333630).

ORCID

Pragya Singh https://orcid.org/0000-0002-7411-3206

REFERENCES

Baeza, J. A. (2007). Male mating opportunities affect sex allocation in a protrandric- simultaneous hermaphroditic shrimp. *Behavioral Ecology and Sociobiology*, *61*(3), 365–370. https://doi.org/10.1007/s00265-006-0265-2

Brauer, V. S., Schärer, L., & Michiels, N. K. (2007). Phenotypically flexible sex allocation in a simultaneous hermaphrodite. *Evolution*, *61*(1), 216–222. https://doi.org/10.1111/j.1558-5646.2007.00018.x

Charlesworth, D., & Charlesworth, B. (1981). Allocation of resources to male and female functions in hermaphrodites. *Biological Journal of the Linnean Society*, *15*(1), 57–74. https://doi.org/10.1111/j.1095-8312.1981.tb00748.x

Charnov, E. L. (1982). The theory of sex allocation. *Monographs in Population Biology*, *18*, 1–355.

Charnov, E. L. (1986). An optimisation principle for sex allocation in a temporally varying environment. *Heredity*, *56*(1), 119–121. https://doi.org/10.1038/hdy.1986.16

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

Crean, A. J., & Marshall, D. J. (2008). Gamete plasticity in a broadcast spawning marine invertebrate. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(36), 13508–13513. https://doi.org/10.1073/pnas.0806590105

Giannakara, A., & Ramm, S. A. (2017). Self-fertilization, sex allocation and spermatogenesis kinetics in the hypodermically inseminating flatworm *Macrostomum pusillum*. *The Journal of Experimental Biology*, *220*(9), 1568–1577. https://doi.org/10.1242/jeb.149682

Giannakara, A., Schärer, L., & Ramm, S. A. (2016). Sperm competition-induced plasticity in the speed of spermatogenesis. *BMC Evolutionary Biology*, *16*(1), 60. https://doi.org/10.1186/s12862-016-0629-9

Gremigni, V. (1988). A comparative ultrastructural study of homocellular and heterocellular female gonads in free-living Platyhelminthes-Turbellaria. *Forschritte der Zoologie*, *36*, 245–261.

Gremigni, V., Falleni, A., & Lucchesi, P. (1987). An ultrastructural study of oogenesis in the turbellarian Macrostomum. *Acta Embryologiae Et Morphologiae Experimentalis*, *8*, 257–262.

Hart, M. K., Svoboda, A., & Cortez, D. M. (2011). Phenotypic plasticity in sex allocation for a simultaneously hermaphroditic coral reef fish. *Coral Reefs*, *30*(2), 543–548. https://doi.org/10.1007/s00338-011-0737-3

Hoch, J. M., & Levinton, J. S. (2012). Experimental tests of sex allocation theory with two species of simultaneously hermaphroditic acorn barnacles. *Evolution*, *66*(5), 1332–1343. https://doi.org/10.1111/j.1558-5646.2011.01548.x

Howell, D. C. (2011). *Confidence intervals on effect size*. University of Vermont. https://doi.org/10.1134/S0362119713030110

Immler, S., Pitnick, S., Parker, G. A., Durrant, K. L., Lupold, S., Calhim, S., & Birkhead, T. R. (2011). Resolving variation in the reproductive tradeoff between sperm size and number. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(13), 5325–5330. https://doi.org/10.1073/pnas.1009059108

Immler, S., Pryke, S. R., Birkhead, T. R., & Griffith, S. C. (2010). Pronounced within-individual plasticity in sperm morphometry across social environments. *Evolution*, *64*(6), 1634–1643. https://doi.org/10.1111/j.1558-5646.2009.00924.x

Janicke, T., Marie-Orleach, L., De Mulder, K., Berezikov, E., Ladurner, P., Vizoso, D. B., & Schärer, L. (2013). Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. *Evolution*, *67*(11), 3233–3242. https://doi.org/10.1111/evo.12189

Janicke, T., Sandner, P., Ramm, S. A., Vizoso, D. B., & Schärer, L. (2016). Experimentally evolved and phenotypically plastic responses to enforced monogamy in a hermaphroditic flatworm. *Journal of Evolutionary Biology*, *29*(9), 1713–1727. https://doi.org/10.1111/jeb.12910

Janicke, T., & Schärer, L. (2009). Sex allocation predicts mating rate in a simultaneous hermaphrodite. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1676), 4247–4253. https://doi.org/10.1098/rspb.2009.1336

Janicke, T., & Schärer, L. (2010). Sperm competition affects sex allocation but not sperm morphology in a flatworm. *Behavioral Ecology and Sociobiology*, *64*(9), 1367–1375. https://doi.org/10.1007/s00265-010-0951-y

Kelly, C. D., & Jennions, M. D. (2011). Sexual selection and sperm quantity: Meta-analyses of strategic ejaculation. *Biological Reviews*, *86*(4), 863–884. https://doi.org/10.1111/j.1469-185X.2011.00175.x

Lorenzi, M. C., Sella, G., Schleicherová, D., & Ramella, L. (2005). Outcrossing hermaphroditic polychaete worms adjust their sex allocation to social conditions. *Journal of Evolutionary Biology*, *18*(5), 1341–1347. https://doi.org/10.1111/j.1420-9101.2005.00916.x

Lüpold, S., Linz, G. M., Rivers, J. W., Westneat, D. F., & Birkhead, T. R. (2009). Sperm competition selects beyond relative testes size in birds. *Evolution*, *63*(2), 391–402. https://doi.org/10.1111/j.1558-5646.2008.00571.x

Morrow, E. H., Leijon, A., & Meerupati, A. (2008). Hemiclonal analysis reveals significant genetic, environmental and genotype x environment effects on sperm size in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, *21*, 1692–1702. https://doi.org/10.1111/j.1420-9101.2008.01585.x

Parker, G. A. (1970). Sperm competition and its evolutionary consequences in the insects. *Biological Reviews*, *45*(4), 525–567. https://doi.org/10.1111/j.1469-185X.1970.tb01176.x

Parker, G. A. (1998). Sperm competition and the evolution of ejaculates: Towards a theory base. In T. R. Birkhead & A. P. Møller (Eds.), *Sperm competition and sexual selection* (pp. 3–54). San Diego, CA: Academic Press.

Patlar, B., Weber, M., & Ramm, S. A. (2019). Genetic and environmental variation in transcriptional expression of seminal fluid proteins. *Heredity*, *122*(5), 595–611. https://doi.org/10.1038/s41437-018-0160-4

Picchi, L., & Lorenzi, M. C. (2019). Gender-related behaviors: Evidence for a trade-off between sexual functions in a hermaphrodite. *Behavioral Ecology*, *30*(3), 770–784. https://doi.org/10.1093/beheco/arz014

Pizzari, T., & Parker, G. A. (2009). Sperm competition and sperm phenotype. In T. R. Birkhead, D. J. Hosken, & S. Pitnick (Eds.), *Sperm biology* (pp. 207–245). London, UK: Academic Press.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Ramm, S. A., Schlatter, A., Poirier, M., & Schärer, L. (2015). Hypodermic self-insemination in the head as a reproductive assurance strategy. *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20150660. https://doi.org/10.1098/rspb.2015.0660

Ramm, S. A., & Stockley, P. (2010). Sperm competition and sperm length influence the rate of mammalian spermatogenesis. *Biology Letters*, *6*(2), 219–221. https://doi.org/10.1098/rsbl.2009.0635

Ramm, S. A., Vizoso, D. B., & Schärer, L. (2012). Occurrence, costs and heritability of delayed selfing in a free-living flatworm. *Journal of Evolutionary Biology*, *25*(12), 2559–2568. https://doi.org/10.1111/jeb.12012

Sandner, P., & Schärer, L. (2010). No plastic responses to experimental manipulation of sperm competition per se in a free-living flatworm. *Ethology*, *116*, 292–299. https://doi.org/10.1111/j.1439-0310.2010.01746.x

Santi, M., Picchi, L., & Lorenzi, M. C. (2018). Dynamic modulation of reproductive strategies in a simultaneous hermaphrodite and preference for the male role. *Animal Behaviour*, *146*, 87–96. https://doi.org/10.1016/j.anbehav.2018.10.004

Schärer, L. (2009). Tests of sex allocation theory in simultaneously hermaphroditic animals. *Evolution*, *63*(6), 1377–1405. https://doi.org/10.1111/j.1558-5646.2009.00669.x

Schärer, L., Brand, J. N., Singh, P., Zadesenets, K. S., Stelzer, C.-P., & Viktorin, G. (submitted). A phylogenetically-informed search for an alternative *Macrostomum* model species, with notes on taxonomy, mating behaviour, karyology and genome size. *Journal of Zoological Systematics and Evolutionary Research*.

Schärer, L., & Janicke, T. (2009). Sex allocation and sexual conflict in simultaneously hermaphroditic animals. *Biology Letters*, *5*(5), 705–708. https://doi.org/10.1098/rsbl.2009.0100

Schärer, L., Joss, G., & Sandner, P. (2004). Mating behaviour of the marine turbellarian *Macrostomum* sp.: These worms suck. *Marine Biology*, *145*(2), 373–380. https://doi.org/10.1007/s00227-004-1314-x

Schärer, L., & Ladurner, P. (2003). Phenotypically plastic adjustment of sex allocation in a simultaneous hermaphrodite. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1518), 935–941. https://doi.org/10.1098/rspb.2002.2323

Schärer, L., Ladurner, P., & Rieger, R. M. (2004). Bigger testes do work more: Experimental evidence that testis size reflects testicular cell proliferation activity in the marine invertebrate, the free-living flatworm *Macrostomum* sp. *Behavioral Ecology and Sociobiology*, *56*(5), 420–425. https://doi.org/10.1007/s00265-004-0802-9

Schärer, L., Littlewood, D. T. J., Waeschenbach, A., Yoshida, W., & Vizoso, D. B. (2011). Mating behavior and the evolution of sperm design. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1490–1495. https://doi.org/10.1073/pnas.1013892108

Schärer, L., & Pen, I. (2013). Sex allocation and investment into pre- and post-copulatory traits in simultaneous hermaphrodites: The role of polyandry and local sperm competition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1613), 20120052. https://doi.org/10.1098/rstb.2012.0052

Schärer, L., Sandner, P., & Michiels, N. K. (2005). Trade-off between male and female allocation in the simultaneously hermaphroditic flatworm *Macrostomum* sp. *Journal of Evolutionary Biology*, *18*(2), 396–404. https://doi.org/10.1111/j.1420-9101.2004.00827.x

Schärer, L., & Vizoso, D. B. (2007). Phenotypic plasticity in sperm production rate: There's more to it than testis size. *Evolutionary Ecology*, *21*(3), 295–306. https://doi.org/10.1007/s10682-006-9101-4

Schärer, L., & Wedekind, C. (2001). Social situation, sperm competition and sex allocation in a simultaneous hermaphrodite parasite, the cestode *Schistocephalus solidus*. *Journal of Evolutionary Biology*, *14*, 942–953. https://doi.org/10.1046/j.1420-9101.2001.00350.x

Schleicherová, D., Sella, G., & Lorenzi, M. C. (2013). Do stable environments select against phenotypic plasticity in hermaphroditic sex allocation? *Italian Journal of Zoology*, *80*(3), 358–363. https://doi.org/10.1080/11250003.2013.805826

Schleicherová, D., Sella, G., Meconcelli, S., Simonini, R., Martino, M. P., Cervella, P., & Lorenzi, M. C. (2014). Does the cost of a function affect its degree of plasticity? A test on plastic sex allocation in three closely related species of hermaphrodites. *Journal of Experimental Marine Biology and Ecology*, *453*, 148–153. https://doi.org/10.1016/j.jembe.2014.01.010

Snook, R. R. (2005). Sperm in competition: Not playing by the numbers. *Trends in Ecology and Evolution*, *20*(1), 46–53. https://doi.org/10.1016/j.tree.2004.10.011

Stearns, S. C. (1992). *The evolution of life histories*. Oxford, UK: Oxford University Press.

Tan, G. N., Govedich, F. R., & Burd, M. (2004). Social group size, potential sperm competition and reproductive investment in a hermaphroditic leech, *Helobdella papillornata* (Euhirudinea: Glossiphoniidae). *Journal of Evolutionary Biology*, *17*(3), 574–580. https://doi.org/10.1111/j.1420-9101.2004.00692.x

Torchiano, M. (2017). *Package "effsize": Efficient effect size computation*. CRAN Package. https://doi.org/10.5281/ZENODO.196082

Vellnow, N., Vizoso, D. B., Viktorin, G., & Schärer, L. (2017). No evidence for strong cytonuclear conflict over sex allocation in a simultaneously hermaphroditic flatworm. *BMC Evolutionary Biology*, *17*(1), 103. https://doi.org/10.1186/s12862-017-0952-9

Vizoso, D. B., Rieger, G., & Schärer, L. (2010). Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. *Biological Journal of the Linnean Society*, *99*(2), 370–383. https://doi.org/10.1111/j.1095-8312.2009.01363.x

Vizoso, D. B., & Schärer, L. (2007). Resource-dependent sex-allocation in a simultaneous hermaphrodite. *Journal of Evolutionary Biology*, *20*(3), 1046–1055. https://doi.org/10.1111/j.1420-9101.2007.01294.x

Winkler, L., & Ramm, S. A. (2018). Experimental evidence for reduced male allocation under selfing in a simultaneously hermaphroditic animal. *Biology Letters*, *14*(10), 20180570. https://doi.org/10.1098/rsbl.2018.0570

Wolak, M. E., Fairbairn, D. J., & Paulsen, Y. R. (2012). Guidelines for estimating repeatability. *Methods in Ecology and Evolution*, *3*(1), 129–137. https://doi.org/10.1111/j.2041-210X.2011.00125.x

Zadesenets, K. S., Schärer, L., & Rubtsov, N. B. (2017). New insights into the karyotype evolution of the free-living flatworm *Macrostomum lignano* (Platyhelminthes, Turbellaria). *Scientific Reports*, *7*(1), 6066. https://doi.org/10.1038/s41598-017-06498-0

Zadesenets, K. S., Vizoso, D. B., Schlatter, A., Konopatskaia, I. D., Berezikov, E., Schärer, L., & Rubtsov, N. B. (2016). Evidence for karyotype polymorphism in the free-living flatworm, macrostomum lignano, a model organism for evolutionary and developmental biology. *PLoS ONE*, *11*(10), e0164915. https://doi.org/10.1371/journal.pone.0164915

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Singh P, Vellnow N, Schärer L. Variation in sex allocation plasticity in three closely related flatworm species. *Ecol Evol*. 2019;00:1–12. https://doi.org/10.1002/ece3.5566

Supplementary file to 'Variation in sex allocation plasticity in three closely related flatworm species'

Table S1. Repeatability estimates (r_I) of morphometric measurements for all species, with confidence intervals in brackets (all $p < 0.05$).

	Macrostomum janickei		*Macrostomum cliftonensis*		*Macrostomum mirumnovem*	
	$F_{34,35}$	r_I	$F_{30,31}$	r_I	$F_{26,27}$	r_I
Body size	3.7	0.57 (0.34-0.79)	2.9	0.48 (0.21-0.75)	7.1	0.75 (0.59-0.92)
Testis size	12.5	0.85 (0.76-0.94)	7.5	0.76 (0.62-0.91)	9	0.8 (0.66-0.94)
Ovary size	10.9	0.83 (0.72-0.93)	4.2	0.61 (0.39-0.84)	5.8	0.7 (0.51-0.9)
Seminal vesicle size	11.6	0.84 (0.74-0.94)	15.2	0.88 (0.79-0.96)	16	0.88 (0.8-0.97)

Table S2. No. of replicates in which offspring were present or absent for each group size.

Group size	*Macrostomum janickei*		*Macrostomum cliftonensis*		*Macrostomum mirumnovem*	
	Present	Absent	Present	Absent	Present	Absent
Isolated	0	57	0	56	17	39
Pair	60	3	56	1	45	6
Octet	57	0	60	0	59	0

Chapter IV

Evolution of sex allocation plasticity and its predictors in a flatworm genus

Singh, P., and L. Schärer. To be submitted
Evolution of sex allocation plasticity
and its predictors in a flatworm genus.

Abstract

Sex allocation (SA) theory in simultaneous hermaphrodites predicts an influence of group size on SA, mediated via changes in the strength of local sperm competition (LSC), which occurs when related sperm compete for access to a given set of eggs. Since the group size experienced by individuals can vary, temporally or spatially, this can favour the evolution of SA plasticity. Furthermore, mating strategy (such as reciprocal vs. unilateral mating) and the ability to self-fertilize could also influence the strength of LSC, and hence the optimal SA, with self-fertilizing species predicted to have lower SA. An ideal model system for testing the above predictions is the flatworm genus, *Macrostomum* that exhibits interspecific differences in the mating strategy and the ability to self-fertilize. In our study, we first assessed the effect of three group sizes (isolated, pairs or octets) and two enclosure sizes (small and large) in all factorial combinations, on estimates of SA in a new and currently still undescribed species, *Macrostomum* sp. 22, and examined if isolated worms can self-fertilize. We found that this species exhibits SA plasticity in response to group size and that it can self-fertilize. Second, we estimated standardized effect sizes for SA plasticity, across a total of seven *Macrostomum* species, in response to i) the presence of mating partners and ii) the strength of LSC, and tested if the mating strategy or the ability to self-fertilize predicted the standardized effect sizes, while controlling for phylogeny. Furthermore, we assessed how consistent these effect sizes were across multiple published experiments of *M. lignano*. Although there is considerable variation between species in their SA plasticity, we found a significant effect only of self-fertilization, such that self-fertilizing species had a lower SA plasticity with respect to presence of mating partners. Finally, we found considerable overlap between the SA plasticity estimates of the different experiments in *M. lignano*, giving us confidence in our interspecific comparisons.

Introduction

Sex allocation (SA) theory in simultaneous hermaphrodites predicts the optimal allocation of resources towards the male versus female function. Specifically, Charnov's mating group size model (Charnov 1980, 1982) predicts the influence of the mating group size on SA, such that the optimal proportion of resources allocated to the male function in an individual is determined by the number of sperm donors/competitors present in a local mating group. Specifically, the mating group size is the number of sperm donors from which a sperm recipient receives sperm at the time its eggs are fertilized, plus 1. The effect of mating group size on SA is mediated via local sperm competition (LSC) (Schärer 2009; Schärer and Pen 2013), which can be seen as the inverse of sperm competition (Parker 1970, 1998) and which occurs when related sperm, usually from the same individual, compete for access to a given set of eggs. The relationship between optimal SA and LSC can be visualised in terms of the male fitness gain curve, which describes how much fitness is gained through the male function per unit resource investment (Schärer 2009).

In small groups, e.g. under monogamy or self-fertilisation, LSC is necessarily very high, since the competing sperm are maximally related to each other, resulting in the male fitness gain curve saturating quickly. Thus, any additional investment into the male function (testis) for production of more sperm than what is required to fertilise the partner's eggs is a waste of resources, as the related sperm simply compete amongst themselves. Instead, these resources could be more profitably invested into the individual's female function to get more fitness returns (Charnov 1982), as the female fitness gain curve is often assumed to be non-saturating and linear. Consequently, under monogamy or self-fertilisation, the organism is expected to have a female-biased SA. But as the mating group size increases, with more competitors contributing sperm, an individual's sperm are more and more competing with unrelated sperm. So it now pays off to invest in the male function to gain a greater share of the fertilisations, and the male fitness gain curve is predicted to linearize, with a subsequent shift towards a more equal SA (Schärer 2009; Schärer and Pen 2013).

An interesting question that arises then is whether and how interspecific variation in reproductive biology, including the mating strategy or the presence of self-fertilization, could influence the strength of LSC and hence the optimal SA. For example, hermaphrodites often exhibit different mating strategies in response to sexual conflict over mating roles (i.e. mating as a sperm donor or as a sperm recipient) (Charnov 1979; Michiels 1998; Schärer et al. 2015). One such mating strategy is reciprocal mating (also called reciprocal copulation), in which

both partners mate in both male and female roles and donate and receive sperm simultaneously (Michiels 1998; Schärer et al. 2015). This strategy can lead to both partners being quite willing to engage in matings in order to get an opportunity to donate sperm, resulting in reduced pre-copulatory sexual selection. The resulting higher mating rate may in turn result in increased sperm competition (conversely decrease in LSC) and more intense post-copulatory sexual selection, including via processes (like sperm digestion) that allow manipulation of received ejaculate. The presence of postcopulatory processes could also alter LSC, e.g. by removing sperm of one or multiple donors. Indeed, theoretical studies have predicted that differences in post-copulatory sexual selection processes, such as cryptic female choice and sperm displacement, can result in changes in LSC and hence, SA (Pen and Weissing 1999; van Velzen et al. 2009; Schärer and Pen 2013).

Such post-copulatory processes can lead to selection for traits or mating strategies that allow the ejaculate to bypass these processes and fertilize the eggs. One such mating strategy is forced unilateral insemination, in which one of the partner mates in the male role and donates sperm while the other mates in the female role (usually against its interests) (Charnov 1979; Michiels 1998; Schärer et al. 2015), e.g. traumatic or hypodermic insemination (Lange et al. 2013; Reinhardt et al. 2015). While the exact direction and magnitude of change in SA will depend on the particular post-copulatory process, e.g. removal of fixed proportion versus fixed amount of sperm can lead to different optimal SA (van Velzen et al. 2009), we could in general expect differences in SA between the mating strategies that differ in such processes. Moreover, in reciprocal mating while there can be some control over who or how many partners you receive sperm from, this might not always be possible in unilateral insemination.

Similarly, species that indulge in self-fertilization might differ from species that outcross, as self-fertilization will generally increase the strength of LSC and hence reduce the fitness returns of investing in the male function, thus leading to a lower optimal SA compared to outcrossing species (Charlesworth and Charlesworth 1981; Charnov 1982, 1987). This is analogous to the effect of local mate competition on sex ratio in separate-sex species (Hamilton 1967). Empirical work in both plants and animals has shown that investment in the male function is inversely correlated with self-fertilizing rate (Lemen 1980; Schoen 1982; Charnov 1987; McKone 1987; Johnston et al. 1998). In general, we would expect any process that has an effect on LSC to affect the shape of the male fitness gain curve, and hence the optimum SA (Schärer 2009).

While most models predict the effect of LSC on SA over evolutionary timescales, LSC can vary temporally or/and spatially within an individuals' lifetime. In hermaphrodites, altering the SA can influence the immediate reproductive success of an individual, favouring the evolution of SA plasticity. Indeed, plastic SA has been suggested to be one of the advantages of hermaphroditism over separate-sex species (Charnov 1982; Michiels 1998). However, for species that do not often experience variation in mating group size during their lifetime, we may not expect high levels of plasticity, particularly if there are costs to the plasticity (DeWitt et al. 1998; Auld et al. 2010; Siljestam and Östman 2017). Interestingly, while SA plasticity has been documented in many hermaphroditic species (Schärer 2009), its evolution is still comparatively poorly understood, with few studies having investigated variation in SA plasticity across species in animals in a controlled experimental layout (Schleicherová et al. 2014).

An excellent model system for testing the effects of the mating strategy and self-fertilisation on SA plasticity is the free-living flatworm genus *Macrostomum* (Macrostomorpha, Platyhelminthes). This genus contains species exhibiting at least two different mating strategies; one involving reciprocal mating and the other hypodermic insemination (Brand et al., in prep.; Schärer et al. 2011). In reciprocally mating species, a facultative postcopulatory 'suck' behaviour has been observed that is hypothesised to remove received sperm (Singh et al., in prep.; Schärer et al. 2004a, 2011; Vizoso et al. 2010; Marie-Orleach et al. 2013). Remarkably, no such postcopulatory 'suck' behaviour has been documented in five hypodermically inseminating species (Schärer et al. 2011). Interestingly, hypodermic insemination is also associated with a suite of morphological traits that potentially allow self-fertilisation (Ramm et al. 2012, 2015; Giannakara and Ramm 2017), although self-fertilization has recently also been documented in at least one reciprocally mating species, *M. mirumnovem* (Singh et al. 2019).

Here, we first report on SA plasticity in response to group size in a new and currently still undescribed species, *Macrostomum* sp. 22. In our experiment, we raised worms in three different group sizes (isolated, pairs or octets) and at two different densities, and measured the effects of these factors on different estimates of SA. This also allowed us to evaluate whether isolated worms can engage in self-fertilization. Next, we collected data on SA plasticity in different *Macrostomum* species (resulting from studies with a similar experimental layout as above). To facilitate comparisons between species, we calculated standardised effect sizes for SA plasticity due to i) the presence of mating partners (i.e.

isolated worms vs. worms with partners) and ii) the strength of LSC (i.e. paired worms vs. octet worms). Using these effect sizes, we then examined if the mating strategy and ability to self-fertilize predicted the SA plasticity effect sizes, while accounting for the phylogeny. Finally, to understand how much variation there can be in these effect sizes between experiments, and thus how consistently they can be estimated for a species, we calculated these effect sizes for multiple published experiments in *M. lignano*.

Materials and Methods

Study organism

Macrostomum sp. 22 is a new and currently still undescribed species, which falls into a large clade of hypodermically inseminating *Macrostomum* species, with the closest known relative being the North Sea form of *M. pusillum* (Brand et al., in prep.; Schärer et al. 2011), and with the Mediterranean form of *M. pusillum* (studies by Giannakara and Ramm 2017; see also below) being a bit more distantly related. The specimens used to initiate the laboratory cultures were collected in 2014 in the lower intertidal zone on two beaches at Cold Spring Harbor, New York, USA (N 40.8591, W 73.4646; N 40.8667, W 73.4688). Since then we maintained this species in mass cultures in the laboratory at 20°C in Petri dishes containing 20‰ artificial sea water (ASW) and fed them with the diatom algae *Nitzschia curvilineata*, until we performed this experiment in December 2016. However, this culture has in the meantime been lost.

SA plasticity and self-fertilization in Macrostomum sp. 22

Experimental setup

To obtain hatchlings of similar age, we allowed 600 adult worms to lay eggs for a period of 3 days. Six days after removing the adults, the resulting hatchlings were randomly allocated to a group size (isolated, pairs or octets) and enclosure size (small or large; i.e. 24-hole or 6-hole tissue culture plates containing 1.5 or 6 ml of ASW, respectively), while providing *ad libitum* algae. We assume that manipulating the group size changes the mating group size, as shown in previous studies in *M. lignano* (Schärer and Ladurner 2003; Janicke and Schärer 2009; Janicke et al. 2013). Moreover, we manipulated the density independently of the group size

by using two enclosure sizes, since changing the group size also changes the density and density itself could have an effect on SA (for a more detailed discussion see Schärer and Ladurner 2003). Each enclosure size combination was replicated 33 times for both isolated and pairs, and 30 times for octets. Worms were transferred to fresh culture plates every week and allowed to mature. Starting from day 25, we then measured the body, testis, ovary and seminal vesicle size for one randomly chosen worm per replicate over a period of 3 days.

Across all treatments, we excluded some replicates, either due to mortality induced while handling worms (n = 4), worms being malformed (n = 9), or a worm missing in the isolated (n = 3) or paired (n = 2) treatment replicates (see Figure 1 for final sample sizes). We did not exclude octet replicates that were missing one or two worms, as we considered a group size of 7 or 6 worms to still be considerably larger than isolated or paired treatments.

Morphometry

For measurement, we anaesthetised worms using a mixture of 300 µl of 71.4g/l $MgCl_2$ and 500 µl 20‰ ASW and dorsoventrally squeezed them on a microscope slide with a haemocytometer cover glass and standardised spacers (of 40 µm thickness) (Schärer and Ladurner 2003). We then took images of the worms with a digital camera (DFK41BF02, The Imaging Source, Bremen, Germany) under a DM 2500 microscope (Leica Microsystems, Heerbrugg, Switzerland) connected to a computer running BTV Pro 6.0b7 (Bensoftware) software. The testis, ovary and seminal vesicle were imaged at 400x magnification, while the body size was imaged at 40x. We analysed these images using ImageJ (available at http://imagej.nih.gov/ij/) and measured the body size, testis size (sum of both testes), ovary size (sum of both ovaries), seminal vesicle size, egg area (area occupied by all developing or developed eggs in the body of the worm) and SA (defined as testis size/(testis size + ovary size)).

We estimated the repeatability for the morphometric measurements by measuring the same set of 37 worms twice in short sequence, randomly chosen from across all treatments, and then estimating the intraclass correlation coefficient (r_I) using the variance components from a one-way ANOVA with the R package 'ICC' (Wolak et al. 2012). The resulting estimates (with confidence intervals) are for body size $F_{35,36} = 4.4$, $P < 0.0001$, and $r_I = 0.63$ (0.43 to 0.83); for testis size $F_{35,36} = 4.3$, $P < 0.0001$, and $r_I = 0.62$ (0.42 to 0.82); for ovary size $F_{35,36} = 1.7$, $P = 0.06$, and $r_I = 0.26$ (-0.05 to 0.56); for seminal vesicle size, $F_{35,36} = 8.3$, $P < 0.0001$,

and $r_I = 0.78$ (0.66 to 0.91); and for egg area, $F_{35,36} = 11.49$, $P < 0.0001$, and $r_I = 0.84$ (0.74 to 0.94).

Self-fertilization

We assessed if *Macrostomum* sp. 22 can self-fertilise, by checking for offspring production on day 22 in the isolated treatments and comparing it with offspring production in the pairs and octets. If *Macrostomum* sp. 22 can self-fertilize, we expected the isolated treatments to also contain hatchlings.

Statistical Analyses

To evaluate if group size and enclosure size had any effect on the dependent traits in *Macrostomum* sp. 22, we took group size and enclosure size as fixed factors and calculated a two-way ANOVA for body size and SA, and a two-way ANCOVA with body size as covariate for the reproductive morphology traits (testis size, ovary size, seminal vesicle size and egg area; since these traits usually correlate with body size). Significant tests were then followed by post-hoc Tukey HSD tests. The data was graphically checked to see if they fulfil the assumptions of parametric test statistics and transformed where needed. Specifically, to account for data skewness, we log-transformed body, testis and seminal vesicle size.

All statistical analyses were carried out using R, version 3.4.1 (R Core Team 2017) unless stated otherwise.

SA and SA plasticity effect sizes across species

To examine how SA plasticity evolves across the *Macrostomum* genus, we gathered data on SA plasticity in different *Macrostomum* species. In total, we collected data from seven species, three that are hypodermically inseminating, *M. pusillum* (Giannakara and Ramm 2017), *M. hystrix* (Winkler and Ramm 2018), and *Macrostomum* sp. 22 (current study); and four that are reciprocally mating, namely *M. janickei*, *M. cliftonensis*, and *M. mirumnovem* (Singh et al. 2019), plus multiple studies from *M. lignano* (see below). *M. mirumnovem* and all three hypodermically inseminating species can self-fertilize (Ramm et al. 2012, 2015; Giannakara and Ramm 2017; Singh et al. 2019). We classified species as reciprocally mating or hypodermically inseminating using behavioural (Singh et al., in prep.) and morphological (Brand et al., in prep.) data. Behavioural data showed the presence of both reciprocal mating and the postcopulatory 'suck' behaviour in *M. lignano* (Schärer et al. 2004a), *M. janickei*, *M. cliftonensis*, and *M. mirumnovem*, while neither of these behaviours were observed in

M. pusillum, *M. hystrix*, and *Macrostomum* sp. 22 (Singh et al., in prep.). In addition, the classification of species using morphological traits, which is known to be correlated with the mating strategy in *Macrostomum* (Schärer et al. 2011), corroborated the classification using behaviour. Briefly, morphologically, species were classified as reciprocal mating or hypodermically inseminating depending on their male and female genital morphology, sperm morphology, and the location of (received) allosperm (Brand et al., in prep.).

The experimental procedure for most species was similar to what we described above for *Macrostomum* sp. 22, except that in some experiments, only group size was varied without independently varying the density via the enclosure size (for details and sample sizes see Supplementary Table S1). We think that the experiments are nevertheless comparable, since all experiments where density was included as a factor showed that it did not have a significant effect on our estimate of SA.

For each experiment, we calculated standardised effect sizes of plasticity in SA, using Cohen's d (Cohen 1988) with Hedges correction for small sample size (Hedges 1981), including confidence intervals (Howell 2011) of the effect sizes with the R package 'effsize' (Torchiano 2017), to facilitate interspecific comparisons. For each species, we calculated an effect size for SA plasticity to assess the effect on presence of mating partners (i.e. isolated worms vs. worms with partners) and strength of LSC (i.e. worms in pairs vs. octets). Pairs represent conditions with high LSC whereas octets represent conditions with low LSC.

To examine how effect size varies across experiments for the same species, we collected datasets from multiple published experiments on *M. lignano*, and calculated effect sizes for each of them separately. In total, we found two and six studies, respectively, where we could extract effect of the presence of mating partners (Schärer and Janicke 2009; Ramm et al. 2019) and the strength of LSC (Schärer and Ladurner 2003; Schärer et al. 2004b; Janicke and Schärer 2009; Janicke et al. 2013; Marie-Orleach et al. 2014; Ramm et al. 2019). For Schärer and Ladurner (2003), we calculated SA as testis size/(testis size+ovary size), since that study only gives values for testis and ovary size, but not SA. We excluded one previous study (Janicke and Schärer 2010) from our analysis, since the worms used in that study came from the same experiment as Janicke and Schärer (2009), and so these were not independent datasets (we used the latter study since it had more replicates). In addition, in Marie-Orleach et al. (2014) we combined the data from the two different inbred lines (HUB1 and DV1), as they did not differ in any of the traits measured. Also, note that in Schärer and Janicke (2009), there are only isolated and paired worms and hence the estimated effect size for the

presence of mating partners does not include the effect of octets. This should not have a large effect on the calculated effect size, since in most cases the difference between isolated and pairs is much larger than the difference between paired and octet treatments (Figure 2).

For *M. lignano*, a weighted average is used to pool the estimated effect sizes across experiments for both the presence of mating partners and the strength of LSC (Turner and Bernard 2006). To weight each effect size, it is multiplied by the inverse of its variance (inverse variance weight - IVW), which allows us to take into consideration the differences in sample size across experiments. The formula for the average effect size is:

$$Hedges's\ g^* = \frac{\sum_{i=1}^{n} Hedges's\ g_i \times IVW_{Hedges's\ g_i}}{\sum_{i=1}^{n} IVW_{Hedges's\ g_i}}$$

And its standard error is calculated as:

$$SE_{g^*} = \sqrt{\frac{1}{\sum_{i=1}^{n} IVW_{Hedges's\ g_i}}}$$

Association between mating strategy and self-fertilization and SA plasticity effect sizes

For the phylogenetically corrected analyses of the SA plasticity effect sizes, we used a trimmed (*ape* package, Paradis et al. 2004) ultrametric phylogenetic tree of the *Macrostomum* genus, which is based on large-scale phylogenomic analyses of *de novo* transcriptomes (Brand et al., in prep.). We estimated the phylogenetic signal using Blomberg's K (*picante* package, Blomberg et al. 2003; Kembel et al. 2010) and Pagel's λ (*phytools* package, Pagel 1999; Revell 2012), for both effect sizes. For both estimates of phylogenetic signal, a value close to zero is suggestive of phylogenetic independence, while a value close to 1 suggests that the traits evolve under Brownian motion. In our case, the estimates of phylogenetic signal for the two traits did not differ significantly from 0 (Supplementary Table S2). Since our sample sizes are small, the likelihood ratio tests used to assess Pagel's λ can be unreliable (Boettiger et al. 2012), as the asymptotic properties of maximum-likelihood estimation may not hold for small sample sizes. Thus, we decided to present both phylogenetically corrected analyses and analyses without correcting for phylogeny.

For the phylogenetically corrected analyses, we compared the fit of different character evolution models (i.e. the Brownian motion, Ornstein–Uhlenbeck, and Early-burst models) (*geiger* package, Harmon et al. 2008) for each effect size, and found that Brownian motion was the best model for character evolution (i.e. having the lowest AICc; Supplementary

Table S3). Thus, we used Brownian motion as the preferred model for the subsequent analysis. We tested if the mating strategy (reciprocal mating vs. hypodermic insemination) or self-fertilization (presence or absence) predict the SA plasticity effect sizes (presence of mating partners and strength of LSC) via phylogenetic generalized least squares (PGLS) regressions, using a Brownian motion model of character evolution (*nlme* package, Pinheiro et al. 2014). PGLS allows us to account for the phylogenetic non-independence of observations due to common evolutionary history of species.

And for the analyses without correcting for phylogeny, we calculated Wilcoxon rank-sum tests to assess if the mating strategy or self-fertilization predicted the differences in SA plasticity effect sizes.

Results

SA plasticity and self-fertilization in Macrostomum sp. 22

In *Macrostomum* sp. 22, body size was not significantly affected by either group or enclosure size (Figure 1, Table 1). And, in accordance with predictions from SA theory, testis size was significantly affected by the group size, with octet worms having significantly larger testis size than isolated and paired worms. In contrast, the ovary size and egg area were significantly affected only by the enclosure size, with worms in larger enclosures (low density) having larger ovary size and egg area compared to worms in smaller enclosures. Similarly, the seminal vesicle size was significantly affected by the enclosure size with worms in larger enclosures (low density) having a smaller seminal vesicle size than worms in smaller enclosures (high density), contrary to our expectations. Similar to what we observed for testis size (and likely largely driven by that trait), SA was significantly affected by the group size, with octet worms having a significantly higher SA than isolated and paired worms.

Macrostomum sp. 22 could self-fertilize with offspring being produced in all but four (out of 58) isolated replicates. Moreover, two (out of 59) of the paired replicates also failed to produce offspring, suggesting that they were incompatible or that at least one worm was infertile and the other failed to self. All octet replicates had offspring in the wells.

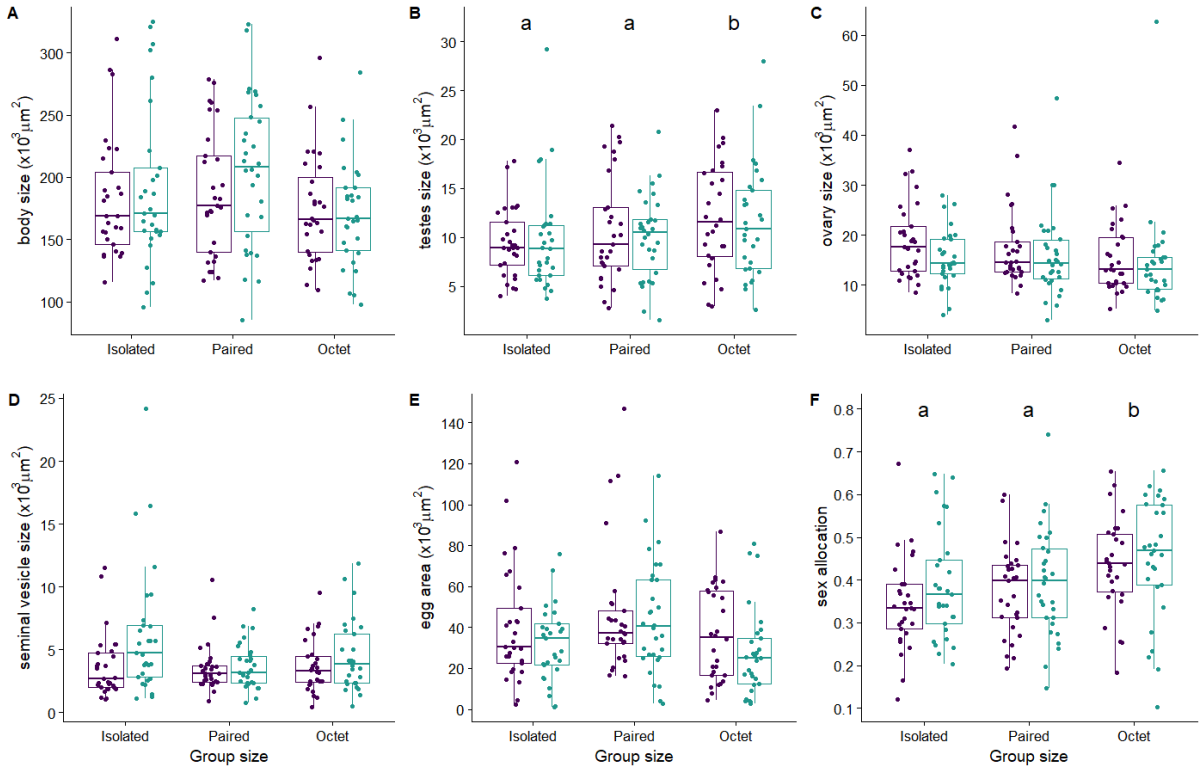


Figure 1. Effect of group size (x-axis) and enclosure size (purple boxplot: large; green boxplot: small) on a) body size, b) testis size, c) ovary size, d) seminal vesicle size, e) egg area, and f) sex allocation in *Macrostomum* sp. 22. The boxplots display the 25th percentile, median, and 75th percentile and the whiskers represent the 5th and the 95th percentiles of the raw data, though note that log-transformed data was used for statistical analysis of body, testis and seminal vesicle size. Different letters denote significantly different group size effects inferred from Tukey HSD post-hoc tests (not done for body, ovary, and seminal vesicle size and egg area, see text).

Table 1. Effects of the fixed factors (group size, enclosure size, and their interaction) and the covariate (body size) on the dependent variables for *Macrostomum* sp. 22. Results of ANOVA (body size and SA) and ANCOVA (all others) are shown.

Trait	Group size			Enclosure size			Interaction			Body size		
	DFs	F	p	DFs	F	p	DFs	F	p	DFs	F	p
Body size	2,168	2.6	0.08	1,168	0.43	0.51	2,168	0.58	0.56	-	-	-
Testis size	2,167	6.4	**0.002**	1,167	1.53	0.22	2,167	0.43	0.65	1,167	63.41	**<0.001**
Ovary size	2,167	0.52	0.6	1,167	4.11	**0.04**	2,167	0.81	0.45	1,167	22.66	**<0.001**
Egg area	2,167	2.33	0.10	1,167	5.187	**0.02**	2,167	0.1	0.91	1,167	40.56	**<0.001**
Seminal vesicle size	2,167	2.39	0.09	1,167	4.11	**0.04**	2,167	2.4	0.09	1,167	25.33	**<0.001**
SA	2,168	6.67	**0.002**	1,168	2.170	0.14	2,168	0.23	0.79	-	-	-

SA and SA plasticity effect sizes across species

There was substantial variation in overall SA across the seven species (Figure 2), both with respect to the mating strategy (solid vs. stippled lines) and self-fertilization (circles vs. triangles). Similarly, there was interspecific variation in the effect size estimates of SA plasticity, even among relatively closely related species (Figure 3, Supplementary Table S1). *M. cliftonensis* and *M. lignano*, respectively, exhibited the highest SA plasticity in response to the presence of mating partners and the strength of LSC.

In addition, *M. lignano* exhibited SA plasticity across all experiments and while the estimate varied across the experiments, the confidence intervals overlapped (Figure 3, Supplementary Table S1).

Association between mating strategy and self-fertilization and SA plasticity effect sizes

The PGLS models (see Supplementary TableS2 and S3 for more details) showed that mating strategy did not predict the SA plasticity effect size either due to the presence of mating partners ($t_{7,5}$ = 0.94, P = 0.39) or the strength of LSC ($t_{7,5}$ = 0.23, P = 0.82). In contrast, self-fertilization predicted SA plasticity with respect to the presence of mating partners ($t_{7,5}$ = -3.07, P = 0.03), but not due to the strength of LSC ($t_{7,5}$ = -0.57, P = 0.59). SA plasticity with respect to the presence of mating partners was lower for selfing species (compare selfing vs. non-selfing species, given by green lines, in Figure 3). Essentially the same was observed in the analyses without correcting for phylogeny using the Wilcoxon rank-sum tests (Supplementary Table S4).

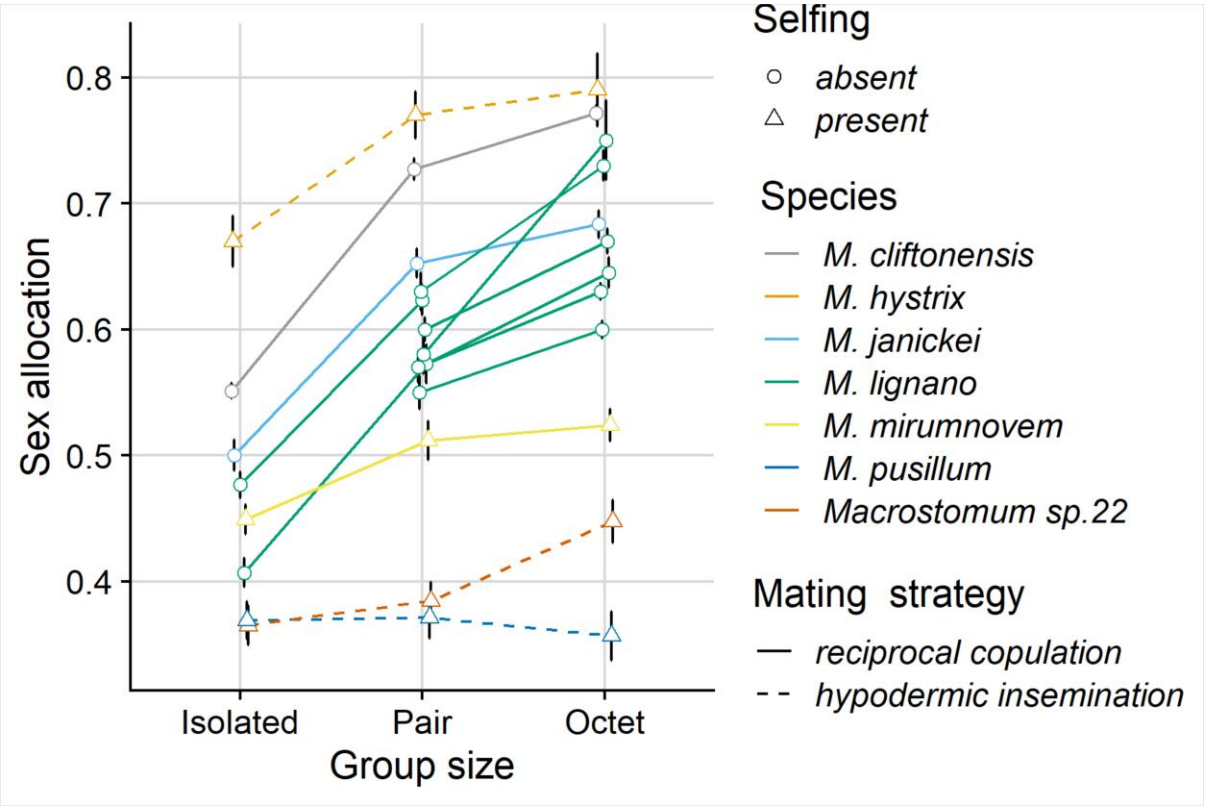


Figure 2. Effect of group size on estimates of sex allocation in seven different *Macrostomum* species (given by different colours). The line types represent the two mating strategies we observe in *Macrostomum*, with solid lines representing reciprocal mating and dashed lines representing hypodermic mating. The symbols represent the ability to self-fertilize, with triangles denoting the presence and circles denoting the absence of self-fertilization, respectively. The plots show means and standard errors of the raw (untransformed) data. Note that for *M. lignano*, data from seven independent experiments are shown.

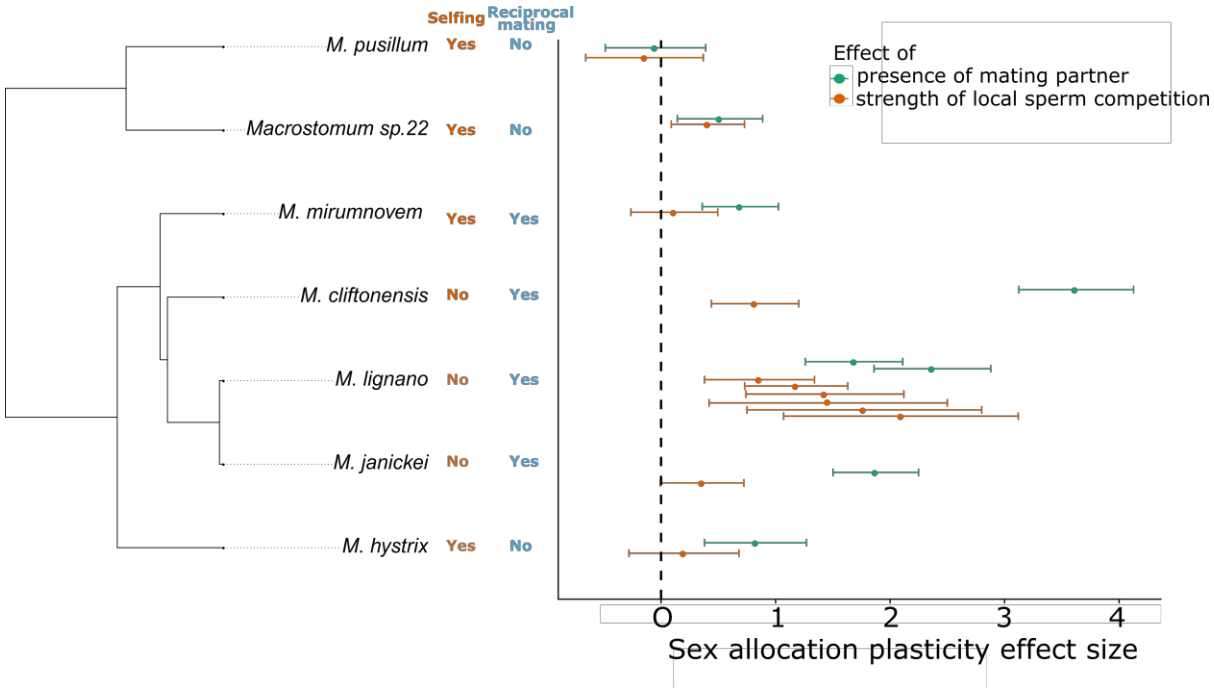


Figure 3. Standardized SA plasticity effect sizes for the effect of the presence of mating partners (i.e. isolated worms vs. worms with partners) and the strength of LSC (i.e. worms in pairs vs. octets) in seven *Macrostomum* species. Black circles or squares represent presence of self-fertilization and reciprocal mating respectively, while grey circles or squares denote their absence (hypodermic insemination in case of mating strategy). Note that for *M. lignano*, data from multiple experiments are shown. The error bars represent 95% confidence intervals.

Discussion

Our study showed that there was substantial interspecific variation in both SA and SA plasticity, also among relatively closely related species, suggesting that both SA and SA plasticity can evolve rapidly in *Macrostomum* genus. Moreover, although mating strategy did not predict either of the SA plasticity effect sizes, self-fertilizing species had a lower SA plasticity in response to presence of mating partner. We discuss these results in more detail below.

SA plasticity and self-fertilization in Macrostomum sp. 22

While *Macrostomum* sp. 22 exhibited plasticity in SA in response to group size, this primarily stemmed from its plasticity in testis size, while ovary size was not significantly plastic. A similar trend has been seen in other *Macrostomum* species (Singh et al. 2019). A

possible reason for not seeing plasticity in ovary size could be that ovary size does not capture the female allocation as well as testis size does male allocation, and there could be other components of female allocation that might be more plastic (see Schärer 2009; Singh et al. 2019 for detailed discussion). *Ad libitum* feeding might also make it harder to uncover trade-off patterns (Schärer et al. 2005). Alternatively, there might not be a trade-off between male allocation and female allocation (Schärer 2009), though we do have some evidence for such a SA trade-off in *M. lignano* (Schärer et al. 2005; Janicke and Schärer 2009). In addition, the repeatability of ovary size was low relative to the repeatability of the other measured traits; which could also have affected our ability to detect plasticity if present, particularly if it were slight. However, we currently do not have an explanation for this low repeatability. Interestingly, the seminal vesicle size was larger in small enclosures (high density), while we expected the converse assuming that worms in smaller enclosures encounter and mate with other individuals at a higher rate. This is similar to what was observed in *M. mirumnovem* (Singh et al. 2019), though we currently do not have a plausible hypothesis for this observation. *Macrostomum* sp. 22 also showed extensive self-fertilization, which was similar in magnitude to its congener species, *M. pusillum* (Giannakara and Ramm 2017).

SA and SA plasticity effect sizes across species

Very few studies have explored predictions of SA theory across multiple species in hermaphroditic animals in a phenotypic plasticity framework (Hoch and Levinton 2012; Schleicherová et al. 2014), with most studies focussing on intraspecific comparisons (Schärer 2009). Hoch and Levinton (2012) tested Charnov's model in two species of acorn barnacles, *Semibalanus balanoides* and *Balanus glandula,* by varying both the number and density of competitors in a natural setting. They showed that both species exhibit increased allocation in absolute, but not relative, male allocation at higher densities. Interestingly, the species differed in how they responded to the treatments in terms of their female allocation, with both the number and density of competitors having an effect in *S. balanoides*—although in the opposite direction than predicted by the model, while there was no effect in *B. glandula.* The authors proposed that this might result from interspecific differences in life-history traits. In contrast to the acorn barnacles, a study in three related polychaete worms, *Ophryotrocha diadema, O. adherens* and *O. gracilis*, showed that while there were interspecific differences in female allocation plasticity in response to different numbers of mating partners, there was no plasticity in the male function across species (Schleicherová et

al. 2014). They proposed that the magnitude of plasticity depended on the species-specific costs of the sex function, and also on the mating system of the species. So, while both of these studies document interspecific variation in SA plasticity, they do not explore whether interspecific differences in reproductive biology traits may affect standardised SA plasticity effect sizes and also lacked the statistical power to do so.

Our results showed that SA and SA plasticity varied across *Macrostomum*. Hypodermically inseminating species (*Macrostomum* sp. 22 and *M. pusillum)*, except *M. hystrix,* exhibited a low SA. Similarly, and in line with our predictions, self-fertilizing species (*Macrostomum* sp. 22, *M. pusillum* and *M. mirumnovem*), except for *M. hystrix*, exhibited a relatively low SA, likely indicating female-biased SA. Interestingly, *M. hystrix* has the highest SA. This species represents an independent origin of hypodermic insemination in the reciprocally mating clade (Schärer et al. 2011), and thus it might still retain characteristics of its reciprocally mating ancestral species that usually have higher SA. Moreover, hypodermic insemination is postulated to allow sperm to compete in a fair-raffle-type scenario, which can lead to an SA that closely matches the group size (Schärer and Janicke 2009). In additon, *M. hystrix* is known to reach very high densities in the field (Ramm et al. 2012). Although we did not statistically test for effect of mating strategy or self-fertilization ability on SA, a more extensive study will examine the evolution of SA and its predictors in *Macrostomum* (Brand et al., in prep.).

M. lignano and *M. janickei,* which are congeneric species capable of hybridization (Singh et al., in prep.), show clear differences in SA plasticity in response to the strength of LSC (although the 95% confidence intervals overlap compared to two of the *M. lignano* studies). Variation in SA plasticity in species with similar reproductive biology could stem from different environmental conditions experienced by the species. For example, evolution of SA plasticity might not be favoured in species that inhabit stable environments, especially if maintenance of plasticity is costly (DeWitt et al. 1998; Auld et al. 2010; Siljestam and Östman 2017). A study in the polychaete worm, *Ophryotrocha diadema,* showed that worms from a laboratory population that had been kept under constant crowded conditions for a long period, were less plastic than worms from a population recently collected from the wild (Schleicherová et al. 2013). The authors suggested that the loss of plasticity could potentially result from the constant conditions experienced by the laboratory worms (Schleicherová et al. 2013). Even in the absence of maintenance costs, there could still be production costs of plasticity though if the benefits of adjusting SA outweigh these costs, then phenotypic

plasticity could still be adaptive. This cost-benefit ratio of SA plasticity can vary across species (Van Buskirk 2002; Steiner 2007) and environments (Ratikainen and Kokko 2019), leading to its retention or loss in a species. Moreover, such costs of SA phenotypic plasticity might play an important role in both the evolution and the maintenance of simultaneous hermaphroditism (St. Mary 1997), and could also constrain changes in SA leading to a suboptimal SA. In *M. lignano* production costs of plasticity have been documented (Sandner 2013), with an alternating environment (group size) leading to a lower hatchling production compared to a stable environment. Although we currently do not have data on costs of plasticity in *M. janickei,* note that not all hermaphrodites exhibit costs of sex adjustment (Lorenzi et al. 2008).

Across species, the estimates of plasticity effect sizes are larger for presence of mating partners than for strength of LSC (number of mating partners), and for *M. cliftonensis* and *M. janickei*, the effect sizes confidence intervals don't overlap. This suggests that the increase in SA going from paired to octet groups is not as drastic as going from isolated worms to larger groups. This situation could potentially arise if in the presence of multiple partners in octets, worms do not only increase their testis size, but also increase their sperm production per unit testis size, (Schärer and Vizoso 2007) due to an increased speed of spermatogenesis (Giannakara et al. 2016) (in which case we might underestimate male allocation when solely measuring testis size). Alternatively, there could be increased investment in other components of the male function, e.g. seminal fluid production (Patlar and Ramm 2019; Patlar et al. 2019), or mating behaviour (Janicke and Schärer 2009; Santi et al. 2018). A similar phenomena has been seen in a freshwater snail, *Lymnaea stagnalis*, where paired snails showed higher expression of six seminal fluid protein (SFP) genes than isolated snails, while the SFP expression of snails in larger (6 individuals) groups was similar to that of paired snails (Nakadera et al. 2019; Ramm et al. 2019).

Association between mating strategy and self-fertilization and SA plasticity effect sizes

The SA plasticity effect sizes did not differ significantly between the reciprocal and hypodermic mating strategies, though our findings did partly confirm our expectation of self-fertilizing species having a lower SA plasticity (in response to presence of mating partner, but not the strength of LSC). Interesting questions that arise then are about the causes of the variation in SA and SA plasticity between the self-fertilizing species. While there has been both theoretical and empirical work exploring the effect of self-fertilization on SA in plants

(Lemen 1980; Charlesworth and Charlesworth 1981; Schoen 1982; Charnov 1987; McKone 1987; Brunet 1992), there have been few such studies in animals (Johnston et al. 1998; Winkler and Ramm 2018). One possible explanation for the variation could be if the species differ in their rate and pattern of self-fertilization, which could have an effect on SA and its plasticity. For example, in obligatorily self-fertilizing species, we would expect that they do not increase their SA (investment in the male allocation) with group size, since LSC would be expected to remain high irrespective of group size. On the other hand, for species that exhibit facultative self-fertilization, we might expect SA plasticity despite low overall SA.

Interestingly, our results do conform somewhat to this pattern, with *M. pusillum*, which has been hypothesised to be obligatorily self-fertilizing, showing both the lowest overall SA and no SA plasticity (Giannakara and Ramm 2017). Its congener, *Macrostomum* sp. 22, showed a similarly female-biased SA in isolated and paired worms, but significant SA plasticity between worms in pairs and octets. If worms shifted from self-fertilization in isolation to only outcrossing in pairs, then the strength of LSC would not be expected to change, while the strength of LSC would clearly be expected to drop from worms in pairs to octets, a scenario that would match the observed SA patterns. This could suggest that *Macrostomum* sp. 22 may only self in isolation. Similarly, *M. hystrix,* which is thought to be a preferentially outcrossing species with a recent study showing delayed selfing and potential costs of self-fertilization (Ramm et al. 2012), was also more plastic than *M. pusillum.* Collectively, our results suggest that SA and SA plasticity can evolve rapidly in *Macrostomum*, with self-fertilization being an important predictor. Future studies should investigate if there is a correlation between SA plasticity and self-fertilization rates in *Macrostomum*.

Acknowledgements:

We would like to thank Christian Felber for help with the morphological measurements of *Macrostomum* sp. 22, Tim Janicke and Steven A. Ramm for providing access to their datasets, Dita Vizoso and Kaja Wasik for help with collecting specimens of *Macrostomum* sp. 22, Jeremias N. Brand, 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 LS.

References

Auld, J. R., A. A. Agrawal, and R. A. Relyea. 2010. Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proc. R. Soc. B Biol. Sci. 277:503–511.

Blomberg, S. P., T. Garland, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution (N. Y). 57:717–745.

Boettiger, C., G. Coop, and P. Ralph. 2012. Is your phylogeny informative? Measuring the power of comparative methods. Evolution (N. Y). 66:2240–2251.

Brunet, J. 1992. Sex allocation in hermaphroditic plants. Trends Ecol. Evol. 7:79–84.

Charlesworth, D., and B. Charlesworth. 1981. Allocation of resources to male and female functions in hermaphrodites. Biol. J. Linn. Soc. 15:57–74.

Charnov, E. L. 1987. On sex allocation and selfing in higher plants. Evol. Ecol. 1:30–36.

Charnov, E. L. 1980. Sex allocation and local mate competition in barnacles. Mar. Biol. Lett. 1:269–272.

Charnov, E. L. 1979. Simultaneous hermaphroditism and sexual selection. Proc. Natl. Acad. Sci. 76:2480–2484.

Charnov, E. L. 1982. The theory of sex allocation. Monogr. Popul. Biol. 18:1–355. Princeton University Press, Cambridge.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences.

DeWitt, T. J., A. Sih, and D. S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends Ecol. Evol. 13:77–81.

Giannakara, A., and S. A. Ramm. 2017. Self-fertilization, sex allocation and spermatogenesis kinetics in the hypodermically inseminating flatworm *Macrostomum pusillum*. J. Exp. Biol. 220:1568–1577.

Giannakara, A., L. Schärer, and S. A. Ramm. 2016. Sperm competition-induced plasticity in the speed of spermatogenesis. BMC Evol. Biol. 16:60.

Hamilton, W. D. 1967. Extraordinary Sex Ratios. Science (80-.). 156:477–488.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hedges, L. V. 1981. Distribution Theory for Glass's Estimator of Effect size and Related Estimators. J. Educ. Stat. 6:107–128.

Hoch, J. M., and J. S. Levinton. 2012. Experimental tests of sex allocation theory with two species of simultaneously hermaphroditic acorn barnacles. Evolution (N. Y). 66:1332–1343.

Howell, D. C. 2011. Confidence Intervals on Effect Size. Univresity of Vermont, doi: 10.1134/S0362119713030110.

Janicke, T., L. Marie-Orleach, K. De Mulder, E. Berezikov, P. Ladurner, D. B. Vizoso, and L. Schärer. 2013. Sex allocation adjustment to mating group size in a simultaneous hermaphrodite. Evolution (N. Y). 67:3233–3242.

Janicke, T., and L. Schärer. 2009. Sex allocation predicts mating rate in a simultaneous hermaphrodite. Proc. R. Soc. B Biol. Sci. 276:4247–4253.

Janicke, T., and L. Schärer. 2010. Sperm competition affects sex allocation but not sperm morphology in a flatworm. Behav. Ecol. Sociobiol. 64:1367–1375.

Johnston, M. O., B. Das, and W. R. Hoeh. 1998. Negative correlation between male allocation and rate of self-fertilization in a hermaphroditic animal. Proc. Natl. Acad. Sci. 95:617–620.

Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg, and C. O. Webb. 2010. Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464.

Lange, R., K. Reinhardt, N. K. Michiels, and N. Anthes. 2013. Functions, diversity, and

evolution of traumatic mating. Biol. Rev. 88:585–601.
Lemen, C. 1980. Allocation of reproductive effort to the male and female strategies in wind-pollinated plants. Oecologia 45:156–159.
Lorenzi, M. C., D. Schleicherova, and G. Sella. 2008. Sex adjustments are not functionally costly in simultaneous hermaphrodites. Mar. Biol. 153:599–604.
Marie-Orleach, L., T. Janicke, and L. Schärer. 2013. Effects of mating status on copulatory and postcopulatory behaviour in a simultaneous hermaphrodite. Anim. Behav. 85:453–461.
Marie-Orleach, L., T. Janicke, D. B. Vizoso, M. Eichmann, and L. Schärer. 2014. Fluorescent sperm in a transparent worm: validation of a GFP marker to study sexual selection. BMC Evol. Biol. 14:148.
McKone, M. J. 1987. Sex Allocation and Outcrossing Rate: A Test of Theoretical Predictions Using Bromegrasses (Bromus). Evolution (N. Y). 41:591.
Michiels, N. K. 1998. Mating Conflicts and Sperm Competition in Simultaneous Hermaphrodites. Pp. 219–254 *in* Sperm Competition and Sexual Selection. Elsevier.
Nakadera, Y., A. Giannakara, and S. A. Ramm. 2019. Plastic expression of seminal fluid protein genes in a simultaneously hermaphroditic snail. Behav. Ecol. 30:904–913.
Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.
Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20:289–290.
Parker, G. A. 1970. Sperm competition and its evolutionary consequences in the insects. Biol. Rev. 45:525–567.
Parker, G. A. 1998. Sperm Competition and the Evolution of Ejaculates: Towards a Theory Base. Pp. 3–54 *in* Sperm competition and Sexual selection.
Patlar, B., and S. A. Ramm. 2019. Genotype-by-environment interactions for seminal fluid expression and sperm competitive ability. J. Evol. Biol. jeb.13568.
Patlar, B., M. Weber, and S. A. Ramm. 2019. Genetic and environmental variation in transcriptional expression of seminal fluid proteins. Heredity (Edinb). 122:595–611.
Pen, I., and F. J. Weissing. 1999. Sperm competition and sex allocation in simultaneous hermaphrodites: A new look at Charnov's invariance principle. Evol. Ecol. Res. 517–525.
Pinheiro, J., D. Bates, S. DebRoy, and D. Sarkar. 2014. R Core Team (2014). nlme: linear and nonlinear mixed effects models. R package version 3.1–117. URL http//cran. r-project. org/web/packages/nlme/index. html.
Ramm, S. A., B. Lengerer, R. Arbore, R. Pjeta, J. Wunderer, A. Giannakara, E. Berezikov, P. Ladurner, and L. Schärer. 2019. Sex allocation plasticity on a transcriptome scale: Socially sensitive gene expression in a simultaneous hermaphrodite. Mol. Ecol. mec.15077.
Ramm, S. A., D. B. Vizoso, and L. Schärer. 2012. Occurrence, costs and heritability of delayed selfing in a free-living flatworm. J. Evol. Biol. 25:2559–2568.
Ramm, S. a, A. Schlatter, M. Poirier, and L. Schärer. 2015. Hypodermic self-insemination as a reproductive assurance strategy. Proc. R. Soc. B Biol. Sci. 282:20150660.
Ratikainen, I. I., and H. Kokko. 2019. The coevolution of lifespan and reversible plasticity. Nat. Commun. 10:538.
Reinhardt, K., N. Anthes, and R. Lange. 2015. Copulatory Wounding and Traumatic Insemination. Cold Spring Harb. Perspect. Biol. 7:a017582.
Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.
Sandner, P. 2013. Impacts of sperm competition on mating behaviour and life history traits in a simultaneous hermaphrodite. University of Basel.

Santi, M., L. Picchi, and M. C. Lorenzi. 2018. Dynamic modulation of reproductive strategies in a simultaneous hermaphrodite and preference for the male role. Anim. Behav. 146:87–96.
Schärer, L. 2009. Tests of sex allocation theory in simultaneously hermaphroditic animals. Evolution (N. Y). 63:1377–1405.
Schärer, L., and T. Janicke. 2009. Sex allocation and sexual conflict in simultaneously hermaphroditic animals. Biol. Lett. 5:705–708.
Schärer, L., T. Janicke, and S. A. Ramm. 2015. Sexual Conflict in Hermaphrodites. Cold Spring Harb. Perspect. Biol. 7:a017673.
Schärer, L., G. Joss, and P. Sandner. 2004a. Mating behaviour of the marine turbellarian *Macrostomum* sp.: these worms suck. Mar. Biol. 145:373–380.
Schärer, L., and P. Ladurner. 2003. Phenotypically plastic adjustment of sex allocation in a simultaneous hermaphrodite. Proc. R. Soc. B Biol. Sci. 270:935–941.
Schärer, L., P. Ladurner, and R. M. Rieger. 2004b. Bigger testes do work more: experimental evidence that testis size reflects testicular cell proliferation activity in the marine invertebrate, the free-living flatworm *Macrostomum* sp. Behav. Ecol. Sociobiol. 56:420–425.
Schärer, L., D. T. J. Littlewood, A. Waeschenbach, W. Yoshida, and D. B. Vizoso. 2011. Mating behavior and the evolution of sperm design. Proc. Natl. Acad. Sci. 108:1490–1495.
Schärer, L., and I. Pen. 2013. Sex allocation and investment into pre- and post-copulatory traits in simultaneous hermaphrodites: the role of polyandry and local sperm competition. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368:20120052.
Schärer, L., P. Sandner, and N. K. Michiels. 2005. Trade-off between male and female allocation in the simultaneously hermaphroditic flatworm *Macrostomum* sp. J. Evol. Biol. 18:396–404.
Schärer, L., and D. B. Vizoso. 2007. Phenotypic plasticity in sperm production rate: There's more to it than testis size. Evol. Ecol. 21:295–306.
Schleicherová, D., G. Sella, and M. C. Lorenzi. 2013. Do stable environments select against phenotypic plasticity in hermaphroditic sex allocation? Ital. J. Zool. 80:358–363.
Schleicherová, D., G. Sella, S. Meconcelli, R. Simonini, M. P. Martino, P. Cervella, and M. C. Lorenzi. 2014. Does the cost of a function affect its degree of plasticity? A test on plastic sex allocation in three closely related species of hermaphrodites. J. Exp. Mar. Bio. Ecol. 453:148–153.
Schoen, D. J. 1982. Male reproductive effort and breeding system in an hermaphroditic plant. Oecologia 53:255–257.
Siljestam, M., and Ö. Östman. 2017. The combined effects of temporal autocorrelation and the costs of plasticity on the evolution of plasticity. J. Evol. Biol. 30:1361–1371.
Singh, P., D. Ballmer, M. Laubscher, and L. Schärer. Submitted. Successful mating and hybridisation in two closely related flatworm species despite significant differences in reproductive morphology and behaviour.
Singh, P., J. N. Brand, and L. Schärer. In preparation. Evolution of the suck behaviour, a postcopulatory female resistance trait in a hermaphroditic flatworm genus.
Singh, P., N. Vellnow, and L. Schärer. 2019. Variation in sex allocation plasticity in three closely related flatworm species. Ecol. Evol. ece3.5566.
St. Mary, C. M. 1997. Sequential patterns of sex allocation in simultaneous hermaphrodites: do we need models that specifically incorporate this complexity? Am. Nat. 150:73–97.
Steiner, U. K. 2007. Investment in defense and cost of predator-induced defense along a resource gradient. Oecologia 152:201–210.
Torchiano, M. 2017. Package "effsize": efficient effect size computation. CRAN Packag.,

doi: 10.5281/ZENODO.196082.
Turner, H. M., and R. M. Bernard. 2006. Calculating and Synthesizing Effect Sizes. Contemp. Issues Commun. Sci. Disord. 33:42–55.
Van Buskirk, J. 2002. A comparative test of the adaptive plasticity hypothesis: relationships between habitat and phenotype in anuran larvae. Am. Nat. 160:87–102.
van Velzen, E., L. Schärer, and I. Pen. 2009. The effect of cryptic female choice on sex allocation in simultaneous hermaphrodites. Proc. R. Soc. B Biol. Sci. 276:3123–3131.
Vizoso, D. B., G. Rieger, and L. Schärer. 2010. Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. Biol. J. Linn. Soc. 99:370–383.
Winkler, L., and S. A. Ramm. 2018. Experimental evidence for reduced male allocation under selfing in a simultaneously hermaphroditic animal. Biol. Lett. 14:20180570.
Wolak, M. E., D. J. Fairbairn, and Y. R. Paulsen. 2012. Guidelines for estimating repeatability. Methods Ecol. Evol. 3:129–137.

Supplementary Information for "Evolution of sex allocation plasticity and its predictors in a flatworm genus"

Table S1. Details of sex allocation plasticity experiment(s) for each species, including their sample size, resulting SA plasticity effect sizes (for presence of mating partners and strength of LSC), inclusion of enclosure size in the experimental design, and reference. For *Macrostomum lignano*, data from seven independent experiments and the weighted average effect sizes across all experiments are given. Note that the largest estimates are printed in bold type.

Species	Sample sizes for each group size			SA plasticity effect size (mean ± sd)		Enclosure size included	Reference
	Isolated	Pair	Octet	Presence of mating partner	Strength of LSC		
M. janickei	57	63	57	1.875±0.188	0.36±0.183	Yes	Singh et al. 2019
M. cliftonensis	56	57	60	**3.62±0.253**	0.82±0.191	Yes	Singh et al. 2019
M. mirumnovem	57	51	59	0.692±0.168	0.115±0.190	Yes	Singh et al. 2019
Macrostomum sp. 22	58	59	57	0.514±0.161	0.41±0.187	Yes	This study
M. pusillum	31	32	29	-0.05±0.219	-0.14±0.253	No	Giannakara and Ramm 2017
M. hystrix	30	48	27	0.83±0.222	0.2±0.239	No	Winkler and Ramm 2018
M. lignano	-	11	12	-	1.77±0.474	Yes	Schärer & Ladurner, 2003
	-	9	12	-	1.46±0.475	Yes	Schärer *et al.*, 2004
	-	46	46	-	1.18±0.224	No	Janicke & Schärer, 2009
	56	62	-	1.69±0.212	-	No	Schärer & Janicke, 2009
	-	16	10	-	2.09±0.482	No	Janicke *et al.*, 2013
	-	19	25	-	1.43±0.335	No	Marie-Orleach *et al.*, 2014
	37	38	37	2.37±0.253	0.86±0.239	No	Ramm *et al.*, 2019
M. lignano (weighted average)	-	-	-	1.97±0.173	**1.25±0.130**		

Table S2. Phylogenetic signal for estimates of the SA plasticity effect sizes. Significance testing shows whether Blomberg's K > 0, using 10000 random permutations of the data. For Pagel's λ a log-likelihood ratio test was used to assess if the maximum-likelihood estimate of λ was significantly different from λ = 0.

Effect size	Blomberg's K		Pagel's λ	
	K	P	λ	P
Presence of mating partner	0.75	0.07	0.29	0.71
Strength of LSC	0.09	0.78	<0.0001	1

Table S3. For each SA plasticity effect size, we determined the fit of different character evolution models (i.e. the Brownian motion, Ornstein–Uhlenbeck and Early-burst models), and evaluated them using the Akaike Information Criterion corrected for small sample size (AICc). The values of σ^2 (Brownian rate parameter), α (selection strength parameter), and *a* (rate of evolutionary change parameter) for the different models are given.

Effect size	Brownian motion			Ornstein-Uhlenbeck				Early-burst			
	σ^2	AICc	ω_t	σ^2	α	AICc	ω_t	σ^2	a	AICc	ω_t
Presence of mating partner	3.78	28.11	0.92	7.76	2.72	33.94	0.05	3.78	-0.000001	35.11	0.03
Strength of LSC	4.21	28.86	0.68	4.72	2.72	30.46	0.3	4.21	-0.000001	35.86	0.02

Table S4. Effect of mating strategy and presence of self-fertilization on estimates of SA plasticity effect size.

Predictor variable	Effect size	Wilcoxon rank-sum test statistic (W)	P
Mating strategy	Presence of mating partner	1	0.11
	Strength of LSC	3	0.38
Presence of self-fertilization	Presence of mating partner	12	0.05
	Strength of LSC	11	0.11

Discussion

In my PhD project, I examined the evolution of mating behaviour and sex allocation (SA) plasticity in the simultaneously hermaphroditic flatworm genus, *Macrostomum*, with a special focus on the species that are closely related to the model organism, *M. lignano*.

In **chapter I**, I show that genital morphology and mating behaviour can evolve rapidly in *Macrostomum*, with *M. lignano* and its congener *M. janickei*, differing significantly in these reproductive traits. Interestingly, despite these considerable differences, the worms were readily able to engage in heterospecific matings and produced fertile hybrid offspring (albeit only few). Similar to our study, a previous study on two genetically similar populations of a neotropical dung fly, *Archisepsis diversiformis*, had shown that despite interpopulation differences in reproductive behaviour and strong premating isolation, extended exposure to individuals of the other population in the absence of individuals of same population, led to interpopulation mating and production of offspring (Puniamoorthy 2014). Moreover, while there is some sparse information about hybridization in free-living flatworms in general (Pala et al. 1982; Bullini 1985), this is, to the best of my knowledge, the first study to have documented hybridization between species of the genus *Macrostomum*.

In our study, very few heterospecific replicates produced offspring, despite most pairs having copulated successfully, presumably due to postmating-prezygotic or postzygotic reproductive barriers. It would be interesting to explore what mechanisms lead to the postmating barriers in our study. A follow up study can examine the female antrum of worms after mating, under the microscope, to check if heterospecific mating leads to transfer of sperm and if yes, whether the number of sperm transferred is similar to conspecific pairs. A study on hybridizations in three species in the *Drosophila simulans* species complex had shown that despite heterospecific matings, the three species exhibited postmating-prezygotic reproductive isolation (Price et al. 2001). Interestingly, each of the three species-pair hybridizations had a different set of cryptic barriers to heterospecific fertilization, such that either the matings were too short to allow sperm transfer; or very few heterospecific sperm were transferred even during long matings; or that despite abundant sperm being transferred, these sperm were lost rapidly from the female's reproductive tract (Price et al. 2001).

Our conspecific pairs of the two species did not differ significantly in their offspring production, suggesting that the nearly two-fold higher mating rate of *M. lignano* did not translate into higher offspring production. Predictions for hermaphrodites do suggest that reciprocal mating might be motivated more from the desire to donate sperm rather than to

receive sperm for fertilizing the eggs, which could lead to a higher mating rate but not necessarily higher offspring production. The higher mating rate of *M. lignano* also led to it engaging predominantly in conspecific matings, thereby exhibiting assortative mating, while *M. janickei* ended up mating more often with heterospecific individuals. The occurrence of more heterospecific than conspecific matings for *M. janickei*, is suggestive of asymmetric reproductive interference between the species, with there likely being a larger fitness cost for *M. janickei* in contact zones. In summary, this study extends our understanding of two central topics in speciation, namely reproductive barriers and hybridization, particularly in the understudied sexual system of simultaneous hermaphrodites. It also highlights the role of mating rate in shaping heterospecific interactions.

It would be interesting to explore if hybridization is possible between more distantly related *Macrostomum* species, particularly using species like *M. cliftonensis* and *M. mirumnovem*, which though collected from a geographically very distant location, are also genetically quite close to *M. lignano.* Both of these species have different stylet morphology and both can be cultured in the laboratory.

In **chapter II**, I provide conclusive evidence that the suck behaviour removes ejaculate out of the antrum. Moreover, I show that there is a correlation between the presence, frequency and duration of reciprocal mating and the suck behaviour, providing evidence for the hypothesis that the suck behaviour co-evolves with reciprocal mating. Future studies should try to experimentally manipulate the suck behaviour to examine its effect on the fitness of both sperm donor and sperm recipient. In addition, I expect there to be correlations between aspects of reproductive behaviour (reciprocal mating and suck) and reproductive morphology (male and female genitalia and sperm). I would particularly expect the suck (female resistance trait) to correlate with presence of sperm bristles (potential male persistence trait), presumably as a result of sexually-antagonistic coevolution. Future studies should examine this using the available data on behaviour and morphology.

Moreover, while I could document reciprocal mating in many species, I do not have convincing videographic evidence for how hypodermic insemination occurs. Possible reasons for this could be that hypodermic mating is very short and instantaneous (Michiels 1998), with no particularly evident mating posture, except the posterior region of one worm being in contact with the partner worm (which we have observed in certain species with hypodermic mating syndrome/morphology). This would make it easy to miss hypodermic insemination

while scoring movies at our available movie resolution and playback speed. Another possible reason could be if these species mate comparatively rarely, as might be expected from their generally relatively low sex allocation, assuming that matings are motivated by the desire to donate sperm (Brand et al. in preparation). Future studies should aim at recording the mating movies at higher frame rates and under the microscope, to increase our chances of observing a hypodermic mating. Moreover, looking at the distribution of the allosperm sperm in the body of a hypodermically inseminating species, can give us an idea of where the sperm was injected during a hypodermic mating (Ramm et al. 2012, 2015). Having a GFP-expressing line, like for *M. lignano* (Marie-Orleach et al. 2014), for a hypodermically inseminating species like *M. hystrix*, will allow us to more easily visualise the transfer and location of received allosperm, and differentiate it from the autosperm in the body of a wildtype recipient.

In **chapter III**, I showed that there is variation in sex allocation (SA) and SA plasticity in three closely related *Macrostomum* species. Moreover, I found that sperm production rate—but not sperm morphology—is plastic in *M. cliftonensis*. Interestingly, self-fertilization had been previously documented in two hypodermically inseminating species of *Macrostomum* (Ramm et al. 2012, 2015; Giannakara and Ramm 2017), and it was postulated that self-fertilization might be associated with hypodermic insemination, since the sharp needle-like stylet and simple sperm presumably made it easy to self-fertilise. In this study, I showed that the reciprocally mating *M. mirumnovem* also exhibited self-fertilization. It also had the lowest SA of the three species, suggesting that self-fertilization could lead to a reduced SA.

Finally, in **chapter IV**, I showed that while the mating strategy does not have a strong effect on SA plasticity effect sizes, self-fertilization has an effect on SA plasticity with respect to the presence of mating partners. While I do not currently have data on the selfing rate of the different species, it would be interesting to examine its relation to allocation to the male function. In addition to the above, there could be species that exhibit mixed-mating, i.e. they could exhibit both self-fertilization and outcrossing in the presence of mating partners. The mating group size experienced by such a species would be different relative to a species that does not self-fertilize in presence of mating partners. The prediction for how such species would modulate their SA in response to increase in group size is not intuitively clear, but I expect there would be a negative correlation between the self-fertilizing rates and the male function (Charlesworth and Charlesworth 1981). Future studies should explore the effect of pattern and rate of self-fertilization on SA plasticity. Moreover, my project is part of a larger

comparative study on the evolution of SA and its correlates in *Macrostomum,* in which SA of field-collected worms is measured. My study allows us to reliably interpret interspecific variation in SA of field collected worms, even assuming that they come from different social environments, keeping the SA plasticity estimates in mind.

In conclusion, I show that mating conflict in *Macrostomum* can give rise to different mating strategies like reciprocal mating and hypodermic insemination, and that a combination of reproductive morphological traits, defined as syndromes, can be used as a proxy for inferring the mating strategy. Even within the reciprocally mating strategy, sexual conflict over the usage of received sperm can give rise to female resistance traits like the suck behaviour. Moreover, both reciprocal mating and the suck behaviour co-evolve rapidly across the genus. Finally, while the mating strategy does not affect the SA directly, self-fertilization (which may be more common in hypodermic species) can lead to lower SA. Thus, my study examines predictions of sexual conflict and SA theory, and elucidates the evolution of reproductive behaviour and SA plasticity and their links, in simultaneously hermaphroditic animals.

References (for Introduction and Discussion)

Altenburg, E. 1934. A Theory of Hermaphroditism. Am. Nat., doi: 10.1086/280526.

Angeloni, L. 2003. Multiple mating, paternity, and body size in a simultaneous hermaphrodite, *Aplysia californica*. Behav. Ecol. 14:554–560.

Anthes, N. 2010. Mate choice and reproductive conflict in simultaneous hermaphrodites. Pp. 329–357 *in* Animal Behaviour: Evolution and Mechanisms. Springer Berlin Heidelberg, Berlin, Heidelberg.

Anthes, N., J. Koene, A. Staikou, H. Kokko, P. Jarne, P. David, J. Auld, D. Sprenger, B. Pelissie, L. Scharer, J. Hoffer, and C. Lorenzi. 2010. Bateman Gradients in Hermaphrodites: An Extended Approach. Am. Nat., doi: 10.1086/655218.

Arnqvist, G., and L. Rowe. 2005. Sexual conflict. Princeton University Press, Princeton, NJ.

Bateman, A. J. 1948. Intra-sexual selection in *Drosophila*. Heredity (Edinb). 2:349–368.

Baur, B. 1998. Sperm Competition in Molluscs. P. *in* Sperm Competition and Sexual Selection.

Borgia, G., and J. Blick. 1981. Sexual competition and the evolution of hermaphroditism. J. Theor. Biol. 89:523–532.

Bullini, L. 1985. Speciation by hybridization in animals. Bolletino di Zool. 52:121–137.

Charlesworth, D., and B. Charlesworth. 1981. Allocation of resources to male and female functions in hermaphrodites. Biol. J. Linn. Soc. 15:57–74.

Charnov, E. L. 1980. Sex allocation and local mate competition in barnacles. Mar. Biol. Lett. 1:269–272.

Charnov, E. L. 1979. Simultaneous hermaphroditism and sexual selection. Proc. Natl. Acad. Sci. 76:2480–2484.

Charnov, E. L. 1996. Sperm competition and sex allocation in simultaneous hermaphrodites. Evol. Ecol. 10:457–462.

Charnov, E. L. 1982. The theory of sex allocation. Monogr. Popul. Biol. 18:1–355. Princeton University Press, Cambridge.

Darwin, C. 1876. The effects of cross and self fertilisation in the vegetable kingdom. D. Appleton, New York :

Dillen, L., K. Jordaens, and T. Backeljau. 2009. Sperm transfer, sperm storage, and sperm digestion in the hermaphroditic land snail *Succinea putris* (Gastropoda, Pulmonata). Invertebr. Biol. 128:97–106.

Eberhard, W. G. 1996. Female Control: Sexual Selection by Cryptic Female Choice. Princeton University Press.

Eppley, S. M., and L. K. Jesson. 2008. Moving to mate: the evolution of separate and combined sexes in multicellular organisms. J. Evol. Biol. 21:727–736.

Ghiselin, M. T. 1969. The Evolution of Hermaphroditism Among Animals. Q. Rev. Biol. 44:189–208.

Giannakara, A., and S. A. Ramm. 2017. Self-fertilization, sex allocation and spermatogenesis kinetics in the hypodermically inseminating flatworm *Macrostomum pusillum*. J. Exp. Biol. 220:1568–1577.

Jarne, P., and J. R. Auld. 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution (N. Y). 60:1816.

Jennions, M. D., and H. Kokko. 2010. Sexual selection. Pp. 343–364 *in* F. C. Westneat DF, ed. Evolutionary behavioral ecology. Oxford University Press, Oxford.

Kimura, K., and S. Chiba. 2015. The direct cost of traumatic secretion transfer in hermaphroditic land snails: Individuals stabbed with a love dart decrease lifetime fecundity. Proc. R. Soc. B Biol. Sci., doi: 10.1098/rspb.2014.3063.

Koene, J. M., K. Montagne-Wajer, D. Roelofs, and A. Ter Maat. 2009. The fate of received

sperm in the reproductive tract of a hermaphroditic snail and its implications for fertilisation. Evol. Ecol. 23:533–543.
Kokko, H., and M. D. Jennions. 2014. The Relationship between Sexual Selection and Sexual Conflict. Cold Spring Harb. Perspect. Biol. 6:a017517–a017517.
Kokko, H., H. Klug, and M. D. Jennions. 2012. Unifying cornerstones of sexual selection: operational sex ratio, Bateman gradient and the scope for competitive investment. Ecol. Lett. 15:1340–1351.
Ladurner, P., B. Egger, K. De Mulder, D. Pfister, G. Kuales, W. Salvenmoser, and L. Schärer. 2008. The Stem Cell System of the Basal Flatworm *Macrostomum lignano*. Pp. 75–94 *in* Stem Cells. Springer Netherlands, Dordrecht.
Lange, R., K. Reinhardt, N. K. Michiels, and N. Anthes. 2013. Functions, diversity, and evolution of traumatic mating. Biol. Rev. 88:585–601.
Leonard, J. L. 2005. Bateman's principle and simultaneous hermaphrodites: A paradox. P. *in* Integrative and Comparative Biology.
Leonard, J. L., and K. Lukowiak. 1984. Male-Female Conflict in a Simultaneous Hermaphrodite Resolved by Sperm Trading. Am. Nat. 124:282–286.
Lodi, M., and J. M. Koene. 2015. The love-darts of land snails: integrating physiology, morphology and behaviour. J. Molluscan Stud. eyv046.
Manning, J. T. 1983. Hermaphroditism and the sampling-error model. Heredity (Edinb)., doi: 10.1038/hdy.1983.10.
Marie-Orleach, L., T. Janicke, and L. Schärer. 2013. Effects of mating status on copulatory and postcopulatory behaviour in a simultaneous hermaphrodite. Anim. Behav. 85:453–461.
Marie-Orleach, L., T. Janicke, D. B. Vizoso, M. Eichmann, and L. Schärer. 2014. Fluorescent sperm in a transparent worm: validation of a GFP marker to study sexual selection. BMC Evol. Biol. 14:148.
Michiels, N. K. 1998. Mating Conflicts and Sperm Competition in Simultaneous Hermaphrodites. Pp. 219–254 *in* Sperm Competition and Sexual Selection. Elsevier.
Michiels, N. K., P. H. Crowley, and N. Anthes. 2009. Accessory male investment can undermine the evolutionary stability of simultaneous hermaphroditism. Biol. Lett., doi: 10.1098/rsbl.2009.0280.
Michiels, N. K., and J. M. Koene. 2006. Sexual selection favors harmful mating in hermaphrodites more than in gonochorists. Integr. Comp. Biol., doi: 10.1093/icb/icj043.
Mouton, S., M. Willems, B. P. Braeckman, B. Egger, P. Ladurner, L. Schärer, and G. Borgonie. 2009. The free-living flatworm *Macrostomum lignano*: A new model organism for ageing research. Exp. Gerontol. 44:243–249. Elsevier Inc.
Pala, M., S. Casu, and N. G. Lepori. 1982. Stabilized natural interspecific hybrid population of the fresh water planarians *Dugesia Gonocephala* S. L. (Turbellaria, Tricladida). Caryologia 35:247–256.
Parker, G. . 2006. Sexual conflict over mating and fertilization: an overview. Philos. Trans. R. Soc. B Biol. Sci. 361:235–259.
Parker, G. A. 1979. Sexual selection and sexual conflict. Pp. 123–166 *in* Sexual Selection and Reproductive Competition in Insects. Elsevier.
Preece, T., Y. Mao, J. P. Garrahan, and A. Davison. 2009. Harmful Mating Tactics in Hermaphrodites. Am. Nat. 173:632–639.
Price, C. S. C., C. H. Kim, C. J. Gronlund, and J. A. Coyne. 2001. Cryptic reproductive isolation in the *Drosophila simulans* species complex. Evolution (N. Y). 55:81–92.
Puniamoorthy, N. 2014. Behavioural barriers to reproduction may evolve faster than sexual morphology among populations of a dung fly (Sepsidae). Anim. Behav. 98:139–148. Elsevier Ltd.

Ramm, S. A., D. B. Vizoso, and L. Schärer. 2012. Occurrence, costs and heritability of delayed selfing in a free-living flatworm. J. Evol. Biol. 25:2559–2568.
Ramm, S. a, A. Schlatter, M. Poirier, and L. Schärer. 2015. Hypodermic self-insemination as a reproductive assurance strategy. Proc. R. Soc. B Biol. Sci. 282:20150660.
Reinhardt, K., N. Anthes, and R. Lange. 2015. Copulatory Wounding and Traumatic Insemination. Cold Spring Harb. Perspect. Biol. 7:a017582.
Schärer, L. 2009. Tests of sex allocation theory in simultaneously hermaphroditic animals. Evolution (N. Y). 63:1377–1405.
Schärer, L., J. N. Brand, P. Singh, K. S. Zadesenets, C.-P. Stelzer, and G. Viktorin. 2019. 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. jzs.12344.
Schärer, L., T. Janicke, and S. A. Ramm. 2015. Sexual Conflict in Hermaphrodites. Cold Spring Harb. Perspect. Biol. 7:a017673.
Schärer, L., G. Joss, and P. Sandner. 2004. Mating behaviour of the marine turbellarian Macrostomum sp.: these worms suck. Mar. Biol. 145:373–380.
Schärer, L., and P. Ladurner. 2003. Phenotypically plastic adjustment of sex allocation in a simultaneous hermaphrodite. Proc. R. Soc. B Biol. Sci. 270:935–941.
Schärer, L., D. T. J. Littlewood, A. Waeschenbach, W. Yoshida, and D. B. Vizoso. 2011. Mating behavior and the evolution of sperm design. Proc. Natl. Acad. Sci. 108:1490–1495.
Schärer, L., and I. Pen. 2013. Sex allocation and investment into pre- and post-copulatory traits in simultaneous hermaphrodites: the role of polyandry and local sperm competition. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 368:20120052.
Singh, P., N. Vellnow, and L. Schärer. 2019. Variation in sex allocation plasticity in three closely related flatworm species. Ecol. Evol. ece3.5566.
Sluys, R. 1989. Sperm resorption in triclads (Platyhelminthes, Tricladida). Invertebr. Reprod. Dev. 15:89–95.
Tomlinson, J. 1966. The advantages of hermaphroditism and parthenogenesis. J. Theor. Biol., doi: 10.1016/0022-5193(66)90038-5.
van Velzen, E., L. Schärer, and I. Pen. 2009. The effect of cryptic female choice on sex allocation in simultaneous hermaphrodites. Proc. R. Soc. B Biol. Sci. 276:3123–3131.
Vizoso, D. B., G. Rieger, and L. Schärer. 2010. Goings-on inside a worm: Functional hypotheses derived from sexual conflict thinking. Biol. J. Linn. Soc. 99:370–383.
Westheide, W. 1999. Ultrastructure and functional significance of intestinojunctional spermathecae in enchytraeids (Oligochaeta, Annelida). Hydrobiologia 406:199–211.
Winkler, L., and S. A. Ramm. 2018. Experimental evidence for reduced male allocation under selfing in a simultaneously hermaphroditic animal. Biol. Lett. 14:20180570.
Zadesenets, K. S., L. Schärer, and N. B. Rubtsov. 2017. New insights into the karyotype evolution of the free-living flatworm Macrostomum lignano (Platyhelminthes, Turbellaria). Sci. Rep. 7:6066.
Zadesenets, K. S., D. B. Vizoso, A. Schlatter, I. D. Konopatskaia, E. Berezikov, L. Schärer, and N. B. Rubtsov. 2016. Evidence for Karyotype Polymorphism in the Free-Living Flatworm, Macrostomum lignano, a Model Organism for Evolutionary and Developmental Biology. PLoS One 11:e0164915.

Acknowledgements

While I have tried to thank people involved in particular parts of my project in each chapter, I take the opportunity here to thank everyone, especially people who might not have been part of my individual thesis chapters but have helped me in different ways.

I would like to thank my supervisor, Lukas Schärer, for his great supervision and for giving me the opportunity to know these amazing worms. It took time to understand the hermaphroditic viewpoint, but it's been an interesting, and enlightening journey. His knowledge and experience have been invaluable for my PhD.

A special thank you to my Schärer lab members, both current and past, for helping in so many ways. Particular thanks to Dita Vizoso, Gudrun Viktorin, Nikolas Vellnow and Jeremias Brand who helped me get acquainted with these wondrous worms. Also, my office mates Axel Wiberg and Vitor Faria for their pep talks. I would also like to take this opportunity to thank everyone at the Zoological institute, the members of different labs, technical and administrative staff, for making my time here a pleasant experience. Thank you to Lukas Zimmermann, Jürgen Hottinger, Urs Stiefel and Daniel Lüscher for helping me keep my computer and lab space functional. A special thank you to Yasmin Picton for helping me sort out my permit issues, time and again, over these past four years.

I am grateful to Hanna Kokko who kindly accepted to be the co-referee of my thesis (and was so accommodating with dates given my permit constraints), and to Dieter Ebert for being the head of my thesis committee (and for helping with the documents on such short notice).

(Second) Last and the most, I am filled with gratitude for my parents, didi, bhaiya and Yuvraj for all the love and words of encouragement from 7000 km away; and Gaurav, for being my home away from home in Switzerland. Words cannot do justice. Also thank you Neha and Priyom for being my support system.

I am also thankful to all the money sources for my PhD, including grants of the Swiss National Science Foundation to Lukas Schärer (31003A_162543 and 310030_184916), and bridging fund from University of Basel and Dieter Ebert.