Fitness consequences of male size in polyandrous livebearing fishes (*Gambusia* spp.): a systematic review and meta-analysis

Bora Kim[1$], Nicholas Patrick Moran[1,2]*, Klaus Reinhold[1], Alfredo Sánchez-Tójar[1]*

* Both authors contributed equally to this work.

1. Department of Evolutionary Biology, Bielefeld University, Bielefeld, Germany

2. Centre for Ocean Life DTU-Aqua, Technical University of Denmark, Lyngby, Denmark

[$] **Author for correspondence**: Bora Kim (bora.kim.kb@gmail.com)

ORCIDs and email addresses:

Bora Kim: 0000-0001-7957-9032; bora.kim.kb@gmail.com

Nicholas P. Moran: 0000-0002-7331-0400; nicholaspatrickmoran@gmail.com

Klaus Reinhold: 0000-0002-0249-8346; klaus.reinhold@uni-bielefeld.de

Alfredo Sánchez-Tójar: 0000-0002-2886-0649; alfredo.tojar@gmail.com

Article type: Review article

Short running title: Fitness consequences of male size

ABSTRACT

1. The genus *Gambusia* represents approximately 45 species of polyandrous livebearing fishes with reverse sexual size dimorphism (i.e. males smaller than females) and with copulation predominantly via male coercion. Male body size has been suggested as an important sexually selected trait, but despite abundant research, evidence for sexual selection on male body size in this genus is mixed.

2. Studies have found that large males have an advantage in both male-male competition and female choice, but that small males perform sneaky copulations better and at higher frequency and thus may sire more offspring in this coercive mating system. Here, we synthesized this discrepant body of evidence in the primary literature.

3. Using pre-registered methods and hypotheses, we performed a systematic review and meta-analysis combining published (n = 19 studies, k = 106 effect sizes) and unpublished data (n = 17, k = 242) to test whether there is overall selection on male body size across studies in *Gambusia*. We also tested several specific hypotheses to understand sources of heterogeneity across effects.

4. Meta-analysis revealed an overall positive correlation between male size and fitness proxies (r = 0.23, 95% confidence interval: 0.10 – 0.35, n = 36, k = 348, 4514 males, three *Gambusia* species). Despite high heterogeneity, the large male advantage appeared robust across all fitness proxies studied (i.e. female choice, mating success, paternity, sperm quantity and quality), but was considerably larger when the fitness proxy was female choice (r = 0.43, 95% confidence interval: 0.28 – 0.59, n = 14, k = 43). Meta-regressions found several important factors explaining heterogeneity across effects, including type of sperm characteristic, male-to-female ratio, female reproductive status, and environmental conditions. We found evidence of publication bias; however, its effect on our estimates was attenuated by including a substantial

amount of unpublished data, highlighting the importance of unpublished (open) data for accurate meta-analytic estimates.

5. Our study underscores the need to rethink the role and form of sexual selection in *Gambusia* and, more broadly, to consider the ecological factors that affect reproductive behaviour in livebearing fishes.

Keywords: *Gambusia affinis*, *Gambusia geiseri*, *Gambusia holbrooki*, intersexual selection, mate choice, mosquitofish, reproductive success, sexual coercion

1. INTRODUCTION

Body size is one of the most important traits affecting the fitness of organisms (Roff, 2002). Larger females are often more fecund than smaller females, while larger males may outcompete smaller males for access to females and are preferred by females in many species (Andersson, 1994; Roff, 2002). An outstanding example of large male advantage can be found in pinnipeds, where selection has led to males of some species being up to seven times heavier than females (Lindenfors et al., 2002). Nonetheless, the largest are not always the most successful. For example, an intermediate-sized male advantage has been documented in midges, leading to stabilizing selection (*Chironomus plumosus*; Neems et al., 1998), and negative selection on male body size has been found in several fly species (McLachlan & Allen, 1987). In most species, we do not yet understand if and how body size is selected for and how intraspecific variation in body size is maintained.

An extreme case of reverse sexual size dimorphism (i.e. males smaller than females) is observed in a family of livebearing fishes, *Poeciliidae*, in which males of some species are among the smallest living vertebrates (Pilastro et al., 1997; Bisazza, 1993). Within this family, the genus *Gambusia* contains approximately 45 species of promiscuous fishes with generally non-descript appearance (Froese & Pauly, 2000). Unlike most fishes, they show internal fertilization with males using a gonopodium, an intromittent organ that transfers sperm into the female gonopore (Constanz, 1989). Whether courtship occurs is unclear (Martin, 1975; Bisazza & Marin, 1991); however, it appears that males commonly bypass female cooperation and forcibly inseminate females via coercive mating tactics (i.e. ‘gonopodial thrusting’; Itzkowitz, 1971; Martin, 1975; McPeek, 1992; Bisazza, 1993; Bisazza & Marin, 1995). Males perform about one gonopodial thrust per minute (Wilson, 2005), and this incessant male harassment seemingly lowers female fitness by reducing foraging efficiency as well as increasing predation risk and energy expenditure (Dadda et al., 2005; Iglesias-Carrasco et al., 2019). *Gambusia* shows considerable inter- and intraspecific male size variation, making them an often-used model to study male body size selection (Zulian et al., 1995; Deaton, 2008). However, despite abundant research, evidence of size-dependent sexual selection is mixed.

Low detection and increased agility in performing gonopodial thrusts have been proposed as explanations for the apparent mating advantage of small males, and thus, for the existence of reverse sexual size dimorphism in *Gambusia* (Hughes, 1985). Laboratory experiments have found that smaller males perform thrusts at higher frequency (Bisazza & Marin, 1995), are more likely to inseminate females (Pilastro et al., 1997; but see Head et al., 2015b), and may sire more offspring than larger males in eastern mosquitofish (*Gambusia holbrooki*; Head et al., 2017). However, large male size may confer an advantage in intrasexual competition. For instance, large males have been observed to monopolize access to females and prevent other males from attempting gonopodial thrusting in both eastern and western mosquitofish (*Gambusia affinis*; Bisazza & Marin, 1995; Hughes, 1985) and to be more likely to sire offspring than small males in eastern mosquitofish (Booksmythe et al., 2016). It has also been observed that female presence can incite aggressive behaviour among eastern mosquitofish males and that larger males were more likely to be aggressive and dominant (Itzkowitz, 1971).

There is also evidence that *Gambusia* females may still exercise some control via pre- and postcopulatory female choice (Bisazza, 1993). At the precopulatory level, eastern and western mosquitofish females were found to preferentially associate with large males (Hughes, 1985; McPeek, 1992; Chen et al., 2018). At the postcopulatory level, *Gambusia* females can store sperm for months, and a single brood may have multiple paternity (Constanz, 1989; Zane et al., 1999; Head et al., 2017), suggesting that sperm competition is likely intense. Larger males have been found to produce more sperm in a number of poeciliid species, including eastern mosquitofish (Locatello et al., 2008; O'Dea et al., 2014; Vega-Trejo et al., 2019). However, Head et al. (2015b) found a nonlinear selection on male sperm count in eastern mosquitofish, where males with an intermediate sperm count were more successful at insemination than those with higher or lower sperm counts. Furthermore, sperm quality might trade off with sperm quantity (Head et al., 2007). Sperm quality traits such as longevity, viability, morphology, and velocity influence fertilization success under sperm competition in many species (García-González & Simmons, 2005; Boschetto et al., 2011; Birkhead & Pizzari, 2002). Although body size may be negatively correlated with sperm quality due to trade-offs between body growth/ maintenance and sperm quality (Evans et al., 2003; Locatello et al., 2008), the relationship between male size and sperm quality in *Gambusia* is unclear (Locatello et al., 2008; Vega-Trejo et al., 2019).

Several environmental factors have been suggested to mediate the body size-fitness relationship in *Gambusia*, leading to context-dependency. The operational sex ratio (i.e. the ratio of sexually receptive males to females) is often proposed as an important factor mediating sexual selection across species by altering the opportunity for selection (Emlen & Oring, 1977; Kvarnemo & Ahnesjö, 1996; but see Klug et al., 2010; Jennions et al., 2012; meta-analysis: Rios Moura & Peixoto, 2013). In coercive mating systems, male-biased operational sex ratios can be particularly costly to males and lead to increased opportunity for selection on male traits (Cureton et al., 2010). For instance, more male-biased ratios resulted in elevated male-male interference (e.g. chasing) and reduced number of gonopodial thrusts in western mosquitofish (Smith & Sargent, 2006). Furthermore, male-biased ratios have been suggested both to benefit large males (Bisazza & Marin, 1995) and to play no role in the relationship between male body size and reproductive success in eastern mosquitofish (Head et al., 2017).

In sum, there is conflicting evidence for male body size selection in *Gambusia*. Frequency-dependent selection may maintain male body size polymorphism (Pilastro et al., 1997). Nonetheless, environmental and ecological factors such as population density, sex ratio, habitat complexity, photoperiod, and temperature are at play, and could exert different selective pressures, leading to context-dependency. Here, we performed a systematic review and meta-analysis combining published and unpublished data to test whether (and how) there is sexual selection on male body size in *Gambusia*, and to understand the sources of heterogeneity. Our hypotheses and predictions, which we pre-registered prior to data collection (Kim et al., 2019), are:

1. Since most copulations in *Gambusia* seemingly involve forcible inseminations that bypass female cooperation and small males seem to be more successful at it, we expect that overall, small males achieve higher fitness than large males. Thus, we predict that male size and fitness are negatively correlated across studies, but we expect this overall effect to be small and uncertain with high heterogeneity in effect sizes.
2. We expect the association between male size and fitness to be context-dependent. Specifically, we predict a positive correlation between male size and fitness when: (a) females can choose between males without physical interaction (e.g. in dichotomous female mate choice test); (b) experimental density is low, allowing large males to physically dominate small males; (c) habitat complexity is high, allowing females to avoid or reduce sexual harassment, and thus, to be preferentially choosy; (d) sex ratio is male-biased due to increased male-male competition. Regarding postcopulatory selection,

we predict (e) a negative correlation between male size and sperm quality due to a trade-off between growth and reproductive investment, but (f) a positive correlation between male size and sperm quantity.

3. Since we expect that female reproductive potential plays a role in male reproductive behaviour, (a) we predict larger effect sizes when females are either virgin or postpartum than when they are gravid. Additionally, we expect the association between male size and fitness to be strengthened by male reproductive motivation. Therefore, (b) we predict larger effect sizes when males are kept separated from females prior to the experiment than when they are kept with females. Last, since the mating system is similar across *Gambusia* species, (c) we do not predict large differences in how male body size and fitness are associated among species.

2. METHODS

2.1 PROTOCOL

The study protocol was pre-registered on the Open Science Framework prior to data collection (Kim et al., 2019). The pre-registration specified our a priori hypotheses, search methods, and confirmatory and exploratory analysis plan. Unless stated otherwise, we adhered to these plans. We followed the relevant PRISMA guidelines (Moher et al., 2009; Figure S2.3). All data processing, analysis, and presentation were conducted using R v.3.6.3 (R Core Team, 2020).

2.2 INFORMATION SOURCES AND SEARCH

We performed a systematic literature search to find published studies in English from all years. Three blocks of search keywords were designed to search for the genus (i.e. *Gambusia*), the predictor (i.e. body size estimates), and the response of interest (i.e. fitness proxies) in titles, abstracts, and keywords. Searches were conducted on 21st January 2019. See Supporting Information S1 for full details about the search.

2.3 STUDY SELECTION & ELIGIBILITY CRITERIA

Our searches on Web of Science Core Collection and PubMed yielded 278 and 97 records, respectively, which were combined and deduplicated using the R package 'revtool' v.0.3.0 (Westgate, 2018). The titles and abstracts of 310 unique records were screened using Rayyan (Ouzzani et al., 2016). Ninety records passed the title-and-abstract screening and were subjected to full-text screening. Full-text records varied in their specific research question, but studies were included as long as they fulfilled the criteria of measuring male size (standard length, total length, body mass) and any fitness proxy (see below) for any species in genus *Gambusia* (see decision trees in Figure S2.1 and S2.2; more information below). Full-text screening identified 55 studies meeting our inclusion criteria (PRISMA diagram in Figure S2.3). All titles/abstracts, and full-texts were double-screened to reduce potential individual biases, with the primary screener (BK) screening all records and secondary screeners (NPM, AST) each independently screening 50%. Conflicting decisions were collectively discussed and resolved.

Studies where animals were exposed to environmental pollutants and/or pharmaceuticals (e.g. endocrine disrupting chemicals such as fluoxetine) were excluded because even very low levels of exposure can affect morphology and reproductive behaviour (Saaristo et al., 2013); however, data from non-exposed control groups from those studies were included, if available. Studies where male fish were size-matched in trials were excluded because potential effects of male body size were effectively eliminated, whereas studies testing non-size related hypotheses were included as long as males were not size-matched.

Four categories of outcome measures were considered proxies for male fitness: female choice, mating success, sperm characteristics (quantity and quality), and paternity (number of offspring sired). In some cases, female choice was measured as the number of approaches made toward males or the number of arching displays by females ($n = 3$ studies, $k = 12$ effects), but the predominant female choice measure was association time in a dichotomous mate choice test ($n = 13$, $k = 31$). Female association preferences have been shown to be indicative of the likelihood of reproducing with preferred males in a poeciliid (Walling et al., 2010). Likewise, the number of mating attempts (gonopodial thrusts), the predominant measure of male mating success, has been shown to be a good predictor of successful copulation (Bisazza, 1993) and paternity (Deaton, 2008) in mosquitofish. Outcome measures not considered as male fitness proxies and excluded were male mate choice, male aggressive behaviour, and male gonadal size or mass.

2.4 DATA COLLECTION AND EXTRACTION

One observer (BK) performed all data extraction and secondary observers (NPM, AST) each independently extracted data from 27% ($n = 15$, 54% total) of records to verify extraction and enhance reproducibility. Data were extracted from text, tables, or figures in published datasets and supplementary materials. R package 'metaDigitise' v.1.0.1 (Pick et al., 2019) was used to extract data from figures. Complete data extraction from published material was possible for 18 studies, and partial extraction from seven additional studies. Requests for missing or partially reported data were sent to 24 authors of 37 studies via a standardized e-mail template, from which we obtained data for 11 studies (from nine authors). Six authors communicated that data were lost, and the remaining nine did not reply. During author correspondence, it was revealed that Head et al. (2015b) re-analysed a subset of data from another study (Head et al., 2015a), so the former was excluded from analyses.

2.5 EXTRACTED VARIABLES

Information was extracted regarding the study (publication year, journal, author information), study subject (species, collection site, fish considered native or invasive at the collection site, wild or laboratory born, female reproductive status), laboratory maintenance conditions (fish kept with/without the opposite sex, temperature, photoperiod), experimental condition (dimension of experimental aquarium, number of female and male fish within experimental trials, presence/absence of physical interaction among experimental fish, habitat complexity), and type/unit of experimental variable. The type of male body size trait (standard length, total length, body mass) and the type of fitness proxy were also recorded. The complete list of continuous and categorical moderators is in Table S3.1 and Table S3.2.

2.6 EFFECT SIZE CALCULATION

We extracted all necessary statistical information to quantify the association between male size and fitness proxies using Pearson's correlation coefficients (hereafter *r*). Following Jacobs and Viechtbauer (2017), mean differences between small and large fish in studies that compared male size categories (e.g. dichotomous female choice trials) were transformed to biserial correlations using the function 'escalc' from the R package 'metafor' v.2.4-0 (Viechtbauer, 2010). Biserial correlations are conceptually equivalent to and directly comparable to *r* (Jacobs & Viechtbauer, 2017). When there were more than two male size groups, we specified in the pre-registration that all pairwise correlations would be calculated; however, this was not a common issue in our studies (i.e. only two such designs), so instead, only data from the smallest and the largest groups were extracted to calculate the biserial correlation.

Where more than one effect size could be calculated from the same data due to the reporting of multiple statistical outputs, we chose one using the following order of preference:

(1) r; (2) other correlation coefficients (e.g. Spearman's rho); (3) mean differences between small and large males (used to calculate biserial correlations as above); (4) R^2 from simple or multiple regression; and (5) inferential statistics (e.g. t-statistic, F-statistic). Effect sizes other than r and biserial correlations were converted into r using the equations provided in Lajeunesse (2013) and Nakagawa and Cuthill (2007; see Table S4). Sampling variances of r (V_r) were calculated as $(1 - r^2)^2/(n - 1)$ (Borenstein et al., 2009), and that of biserial correlations was calculated using the function 'escalc' from the R package 'metafor' v.2.4-0 (Viechtbauer, 2010). The sample size of each effect size reflected the number of replicates rather than the number of males. These two numbers were the same except for dichotomous mate choice trials, in which one female chose between two males, and we assigned the number of females as the sample size rather than the number of males to avoid artificially inflating sample size. Effect sizes were coded so that a negative effect size denoted a negative correlation between male size and fitness, and vice versa.

2.7 MAIN EFFECT MODEL

A multilevel intercept-only meta-analytic model was fitted to estimate the overall effect size or meta-analytic mean for the association between male size and fitness proxies using the R package 'metafor' v.2.4-0 (Viechtbauer, 2010). Estimates (i.e. means) are presented with their 95% confidence intervals (CI) in square brackets throughout. Furthermore, we also estimated 95% prediction intervals (PI), which incorporate heterogeneity (IntHout et al., 2016). Whereas confidence intervals show the range in which the overall effect is likely to be found, prediction intervals estimate the likely range in which effects are expected to occur in similar future (or unknown) studies (IntHout et al., 2016).

All models, including the meta-regressions (see below), included the following random effects: (i) study ID, which encompasses effect sizes extracted from the same study, (ii) group ID, which encompasses effect sizes obtained from the same group of fish, (iii) experiment ID, which encompasses effect sizes derived from the same experiment, and (iv) effect ID, which represents residual variance among-effect sizes. Our models included one more random effect (i.e. Group ID) than planned in our pre-registration, but this was considered necessary to account for this source of non-independence among effect sizes.

For the intercept-only meta-analytic model, we calculated Cochran's Q and I^2_{total} (Higgins & Thompson, 2002) and the equivalent for each random effect, as measures of absolute and relative heterogeneity, respectively. Heterogeneity refers to the unexplained variation among effect sizes after accounting for sampling variance.

2.8 META-REGRESSIONS FOR TESTING HYPOTHESES

We fitted multilevel meta-regressions to investigate potential effects of moderators on the relationship between male size and fitness proxies. To test if physical interaction among individual fish affected the results (Hypothesis 2a), we fitted a meta-regression including the moderator 'physical interaction' (levels: yes, no) for the subset of studies in which female choice was measured. For experiments where fish could physically interact, we fitted a meta-regression including the following moderators: experimental density (i.e. total number of fish in the trial divided by the aquarium volume (L); Hypothesis 2b), habitat complexity (levels: low, high; Hypothesis 2c), and male-to-female ratio (Hypothesis 2d) as well as the interaction between experimental density and habitat complexity, and the interaction between male-to-female ratio and habitat complexity. Since the latter two meta-regressions tested hypotheses related to precopulatory mechanisms, they did not include effect sizes on sperm quantity nor quality. For the subset of studies that measured sperm quantity and/or quality, we fitted a meta-regression including the type of sperm characteristic as a moderator (levels: quantity, quality; Hypotheses 2e and 2f).

Due to limited reporting on female reproductive status and male housing conditions in the literature, we deviated from our pre-registration for hypotheses 3a and 3b (details in Supporting Information S8). Instead, to test for effects of female reproductive status (Hypothesis 3a), we fitted a meta-regression with four female status levels (virgin, gravid, male-deprived, and non-deprived). To test for male housing condition effects (Hypothesis 3b), we fitted a meta-regression including a moderator with two levels (mixed-sex: kept with females, same-sex: kept separated from females). Last, we fitted a meta-regression including a moderator ‘species’ with three levels (*G.affinis*, *G.geiseri,* and *G.holbrooki*) to test if the effect differs among species (Hypothesis 3c).

2.9 META-REGRESSIONS FOR EXPLORATORY ANALYSES

Five additional pre-registered exploratory meta-regressions were performed to test hypotheses related to methodological design, but for which no specific direction was predicted (Kim et al., 2019). We tested if results differed: (1) depending on the type of male size proxy used (levels: standard length, total length, body mass); (2) between native and invasive populations (levels: native, invasive); (3) depending on the fish’s rearing environment (levels: wild, laboratory); (4) depending on temperature (°C) and photoperiod (i.e. number of daylight hours per day); and (5) depending on the type of outcome variable (i.e. type of fitness proxy; levels: female choice, mating success, sperm quality, sperm quantity, paternity).

For all meta-regressions, we estimated the percentage of heterogeneity explained by the moderators using R^2_{marginal} (Nakagawa and Schielzeth, 2013). Missing and unreported data were not included in the meta-regressions (i.e. we ran complete-case analyses). Continuous and categorical moderators involved in interactions terms (e.g. habitat complexity) were mean-centred to aid interpretation (Schielzeth, 2010). Results of the main effect model and meta-regressions with categorical moderators were graphically represented as orchard plots using the R package 'orchaRd' v.0.0.0.9000 (Nakagawa et al., 2020). Meta-regressions with continuous moderators were plotted with the R package 'ggplot2' v.3.3.2 (Wickham, 2016).

2.10 PUBLICATION BIAS TESTS

To test for small-study bias, we fitted a multilevel meta-regression with sample size as a moderator (Nakagawa & Santos, 2012). Likewise, to test for time-lag bias in the published literature (Jennions & Møller, 2002; Koricheva & Kulinskaya, 2019), we fitted a multilevel meta-regression including the year of publication as a moderator (Sánchez-Tójar et al., 2018). Furthermore, the source of data was included as a moderator (levels: published, unpublished) in a meta-regression to test whether effect sizes were larger in published than unpublished data (Sánchez-Tójar et al., 2018; Moran et al., 2020). We categorized supplementary data as unpublished whenever the specific research question/hypothesis of the original study did not involve male size per se, but male size was nevertheless measured and provided. Additionally, whether results were reported completely or incompletely (e.g. missing effect sizes, relationships reported as simply 'non-significant', etc.) was included as a moderator (levels: complete, incomplete) in a meta-regression to test whether effect sizes were larger in studies that incompletely reported results. Last, we originally intended to test whether data collected by observers blind to male size led to smaller effect sizes than data collected by observers not blind to male size (see Holman et al., 2015), but we did not encounter any study using blind data collection regarding male size, so this pre-registered hypothesis was not tested.

3. RESULTS

Overall, 348 effect sizes were obtained from 36 studies including 179 groups of fish tested in 216 experiments (4514 male fish in total). Median and mean sample sizes were 16 and 35, respectively (range: 3 – 294; only three data points had a sample size of three). Data were available only for three species: *G. affinis* ($n = 7$ studies, $k = 29$ effects), *G. geiseri* ($n = 1$, $k = 5$), and *G. holbrooki* ($n = 29$, $k = 314$; map of collection sites shown in Figure S5.1).

3.1 MAIN EFFECT MODEL (HYPOTHESIS 1)

Contrary to our hypothesis, the intercept-only model revealed a positive association between male size and fitness proxies ($r = 0.23$ [0.10 – 0.35], 95% PI = -0.69 – 1.15, $p < 0.001$, $n = 36$, $k = 348$; Figure 1). That is, our meta-analysis suggests that there is positive selection on male size in *Gambusia*. Nonetheless, absolute ($Q = 5484$, $p < 0.001$) and relative heterogeneity (I^2_{total} = 92.2% [85.3 – 95.7]) were high. When I^2_{total} was partitioned, 33.0% [23.7 – 41.2] was attributed to study ID, 53.1% [40.8 – 60.9] to group ID, 6.2% [0.8 – 11.9] to experiment ID, and 0.0% [0.0 – 1.8] to effect ID (i.e. residual variance).

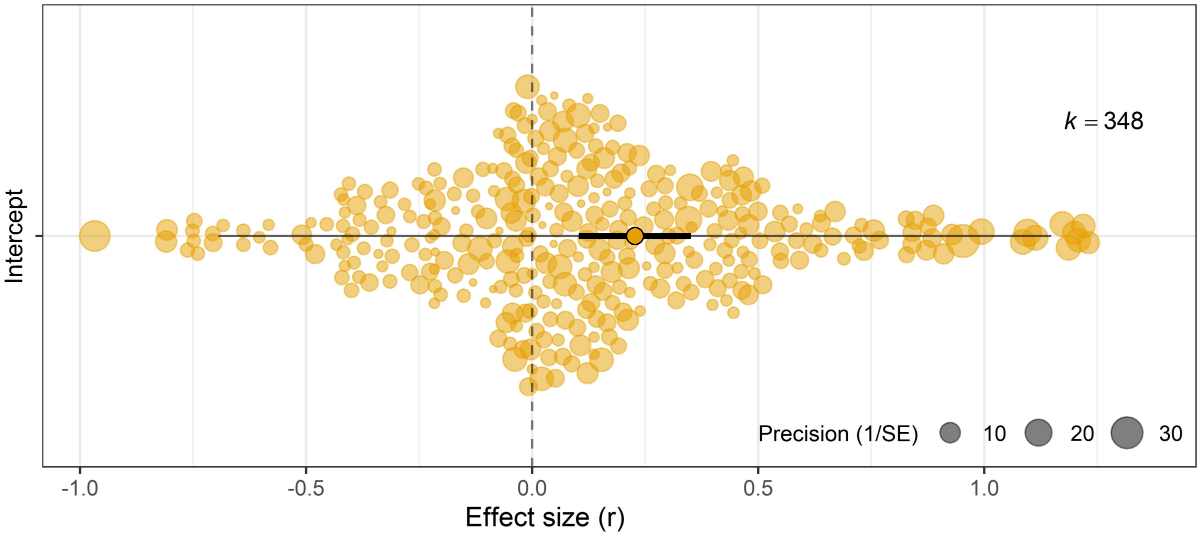

Figure 1. Male size appears positively selected across included effects. Orchard plot of the meta-analytic model, showing the meta-analytic mean, 95% CI (thick whisker), 95% PI (thin whisker), and individual effect sizes scaled by their precision (circles).

3.2 META-REGRESSIONS FOR TESTING HYPOTHESES

3.2.1 *Physical Interaction (Hypothesis 2a)*

The size-fitness correlation was positive in both presence ($r = 0.18$ [0.01 – 0.35], $p = 0.015$, $n = 19$, $k = 171$) and absence ($r = 0.38$ [0.16 – 0.59], $p < 0.001$, $n = 14$, $k = 37$) of physical interaction between males and females during mate choice tests. Effect sizes tended to be larger in absence than in presence, but that difference was not statistically significant ($p = 0.105$). The moderator explained 2.3% of heterogeneity ($R^2_{\text{marginal}} = 0.023$).

3.2.2 *Experimental Density (Hypothesis 2b), Habitat Complexity (Hypothesis 2c) and Male-to-Female Ratio (Hypothesis 2d)*

For experiments where fish were allowed to physically interact, the size-fitness correlation did not seem to be affected by experimental density, male-to-female ratio or the interaction between those and habitat complexity (Table S6.1). Effect sizes tended to be stronger in complex habitats, but a subsequent non-pre-registered meta-regression including habitat complexity as the only moderator showed that the difference between low ($r = 0.10$ [-0.11 – 0.30], $p = 0.354$, $n = 15$, $k = 144$) and high habitat complexity ($r = 0.23$ [-0.05 – 0.50], $p = 0.115$; $n = 6$, $k = 27$) was not statistically significant ($p = 0.383$; $R^2_{\text{marginal}} = 0.008$). In contrast, an additional non-pre-registered meta-regression that included male-to-female ratio as the only moderator showed that, as predicted, the more male-biased the population, the greater the fitness advantage of large males (intercept = 0.14 [-0.05 – 0.33], $p = 0.137$; slope = 0.13 [0.02 – 0.25], $p = 0.022$; $n = 19$, $k = 171$; $R^2_{\text{marginal}} = 0.104$; Figure 2). Bear in mind, since the latter two meta-regressions were not pre-registered, the results should be interpreted cautiously.

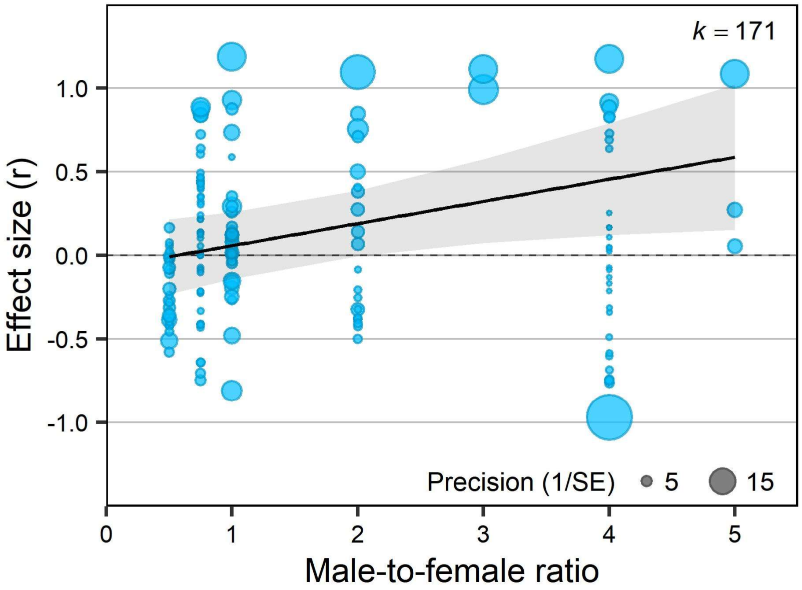

Figure 2. Large males showed greater fitness with more male-biased populations. The solid line represents the model estimate, shading represents the 95% CI, and individual effect sizes are scaled by their precision.

3.2.3 *Sperm Quantity and Quality (Hypotheses 2e & 2f)*

Male size and sperm quantity were positively correlated ($r = 0.17$ [0.09 – 0.24], $p < 0.001$, $n = 10$, $k = 74$), while the estimate for sperm quality was small and its 95% CI overlapped zero ($r = 0.04$ [-0.04 – 0.12], $p = 0.316$, $n = 8$, $k = 66$). Indeed, the difference between quantity and quality was statistically significant ($p < 0.001$; Figure 3A), and the type of sperm characteristic as a moderator explained 8.8% of the heterogeneity ($R^2_{\text{marginal}} = 0.088$).

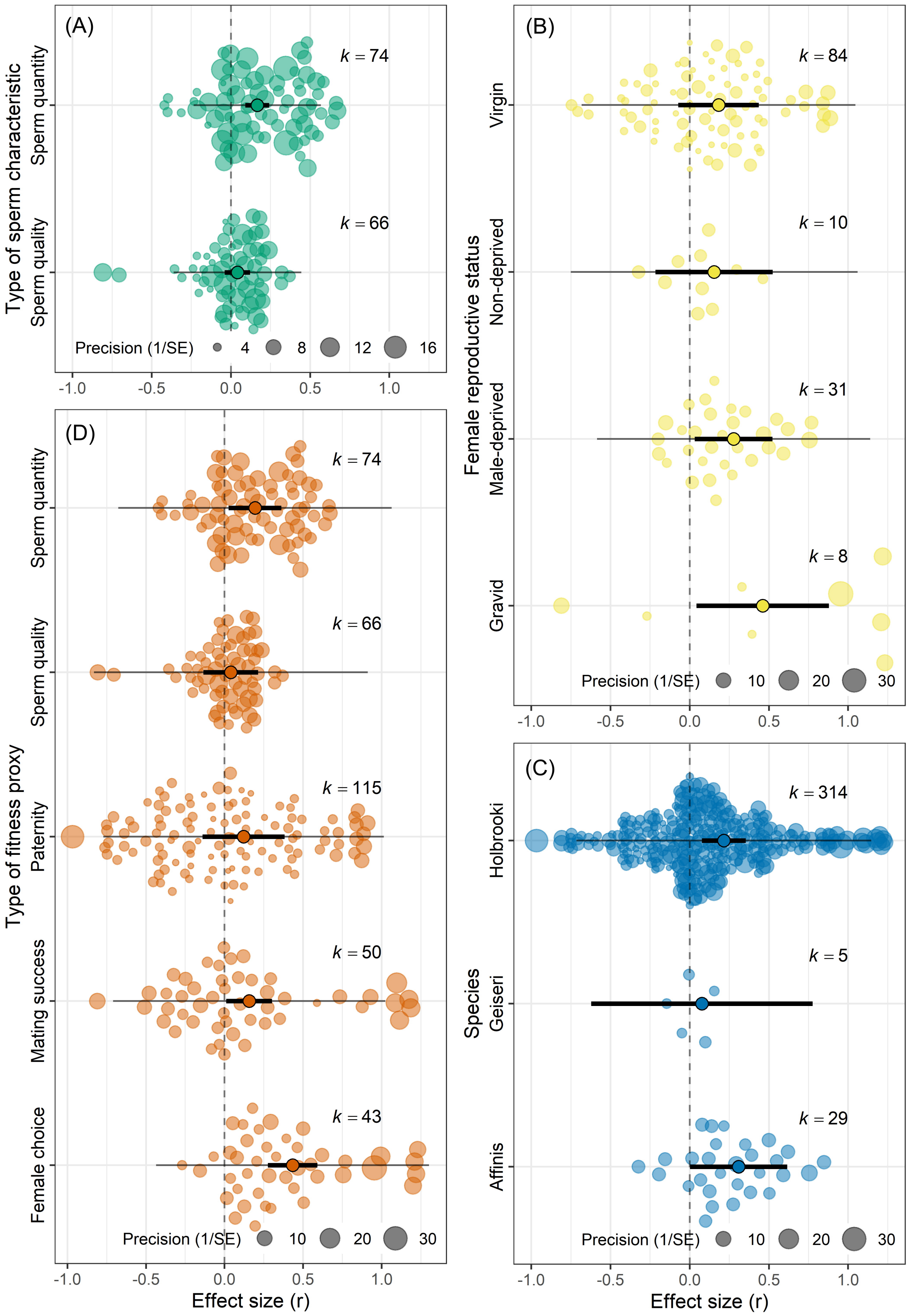
(A)
Type of sperm characteristic
Sperm quantity
Sperm quality
k = 74
k = 66
Precision (1/SE) 4 8 12 16
-1.0 -0.5 0.0 0.5 1.0
(B)
Female reproductive status
Virgin
Non-deprived
Male-deprived
Gravid
k = 84
k = 10
k = 31
k = 8
Precision (1/SE) 10 20 30
-1.0 -0.5 0.0 0.5 1.0
(D)
Type of fitness proxy
Sperm quantity
Sperm quality
Paternity
Mating success
Female choice
k = 74
k = 66
k = 115
k = 50
k = 43
Precision (1/SE) 10 20 30
-1.0 -0.5 0.0 0.5 1.0
Effect size (r)
(C)
Species
Holbrooki
Geiseri
Affinis
k = 314
k = 5
k = 29
Precision (1/SE) 10 20 30
-1.0 -0.5 0.0 0.5 1.0
Effect size (r)

Figure 3. Orchard plots showing that A) Male size was positively correlated with sperm quantity but not quality; B) Female reproductive status did not strongly influence the correlation; C) The correlation did not differ substantially across *Gambusia* species; D) The correlation was generally positive across male fitness proxies in *Gambusia* species. Note that, although paternity contains more effect sizes than the other levels, only four studies measured paternity. Plots show means, 95% CI (thick whisker), 95% PI (thin whisker), and individual effect sizes scaled by their precision (circles).

3.2.4 *Female Reproductive Status (Hypothesis 3a)*

The size-fitness correlation was positive in all four levels of female reproductive status, but the 95% CIs overlapped zero in virgin ($r = 0.18$ [-0.07 – 0.44], $p = 0.160$, $n = 7$, $k = 84$) and non-deprived females ($r = 0.15$ [-0.22 – 0.52], $p = 0.414$, $n = 3$, $k = 10$), while they did not in gravid ($r = 0.46$ [0.04 – 0.88], $p = 0.031$, $n = 3$, $k = 8$) and male-deprived females ($r = 0.28$ [0.03 – 0.52], $p = 0.026$, $n = 8$, $k = 31$; Figure 3B). Post-hoc Wald tests revealed no statistically significant differences between those four levels of female reproductive status ($p > 0.282$ in all cases), and the moderator explained 3.0% of heterogeneity ($R^2_{\text{marginal}} = 0.030$).

3.2.5 *Male Housing Condition (Hypothesis 3b)*

The size-fitness correlation was positive in both mixed-sex ($r = 0.38$ [0.18 – 0.57], $p < 0.001$, $n = 10$, $k = 98$) and same-sex housing conditions (r = 0.16 [0.01 – 0.32], $p = 0.038$, $n = 17$, $k = 164$). Contrary to our hypothesis, effect sizes tended to be larger in mixed-sex than in same-sex conditions ($p = 0.091$). Male housing conditions explained 5.3% of heterogeneity ($R^2_{\text{marginal}} = 0.053$).

3.2.6 *Species (Hypothesis 3c)*

The size-fitness correlation was positive in all three species, although the 95% CI substantially overlapped zero in *G. geiseri* (*G. affinis*: $r = 0.31$ [0.00 – 0.62], $p = 0.048$, $n = 7$, $k = 29$; *G. geiseri*: $r = 0.08$ [-0.62 – 0.78], $p = 0.829$, $n = 1$, $k = 5$; *G. holbrooki*: $r = 0.22$ [0.08 – 0.35], $p = 0.002$, $n = 29$, $k = 314$). As predicted, the differences across species were not statistically significant ($p > 0.515$ in all cases; Figure 3C), and the moderator explained only 0.4% of heterogeneity ($R^2_{\text{marginal}} = 0.004$).

3.3 META-REGRESSIONS FOR EXPLORATORY ANALYSES

3.3.1 *Type of Male Size Proxy*

The size-fitness correlation was positive and similar regardless of the type of male size proxy used ($p > 0.949$ in all cases; $R^2_{\text{marginal}} = 0.000$): standard length ($r = 0.22$ [0.09 – 0.35], $p < 0.001$, $n = 32$, $k = 263$), total length ($r = 0.23$ [0.06 – 0.39], $p = 0.008$, $n = 4$, $k = 31$), and body mass ($r = 0.23$ [0.09 – 0.36], $p = 0.001$, $n = 7$, $k = 43$).

3.3.2 *Origin of Population*

The size-fitness correlation was positive for both invasive ($r = 0.21$ [0.07 – 0.36], $p = 0.004$, $n = 27$, $k = 274$) and native populations ($r = 0.26$ [-0.02 – 0.53], $p = 0.069$, $n = 8$, $k = 73$). That difference was not statistically significant ($p = 0.784$), and the moderator explained only 0.1% of heterogeneity ($R^2_{\text{marginal}} = 0.001$).

3.3.3 *Rearing Environment*

The size-fitness correlation was positive for wild fish ($r = 0.27$ [0.13 – 0.41], $p < 0.001$, $n = 28$, $k = 222$), but not statistically significantly so for laboratory-bred fish ($r = 0.08$ [-0.17 – 0.32], $p = 0.551$, $n = 7$, $k = 125$); however, that difference was not statistically significant ($p = 0.181$). Rearing environment explained 3.9% of heterogeneity ($R^2_{\text{marginal}} = 0.039$).

3.3.4 *Temperature and Photoperiod*

Neither temperature nor photoperiod seemed to strongly influence the size-fitness correlation (intercept = 0.26 [0. 12 – 0.41], $p < 0.001$; temperature = -0.03 [-0.10 – 0.04], $p = 0.359$; photoperiod = 0.11 [-0.02 – 0.24], $p = 0.101$; $n = 26$, $k = 250$). However, there was a tendency for the correlation to be greater with longer hours of daylight, and both moderators combined explained 5.2% of heterogeneity ($R^2_{\text{marginal}} = 0.052$).

3.3.5 *Type of Proxies for Male Fitness*

The size-fitness correlation was positive regardless of the type of male fitness proxy. However, it was only statistically significant for female choice ($r = 0.43$ [0.28 – 0.59], $p < 0.001$, $n = 14$, $k = 43$), mating success ($r = 0.16$ [0.01 – 0.30], $p = 0.035$, $n = 14$, $k = 50$), and sperm quantity ($r = 0.19$ [0.03 – 0.36], $p = 0.024$, $n = 10$, $k = 74$), whereas the estimates for paternity ($r = 0.12$ [-0.14 – 0.38], $p = 0.362$, $n = 4$, $k = 115$) and sperm quality ($r = 0.04$ [-0.13 – 0.21], $p = 0.651$, $n = 8$, $k = 66$) were not statistically significant (Figure 3D). Post-hoc Wald tests showed that the estimate for female choice was statistically significantly larger than those of the other proxies ($p < 0.041$ in all cases), and the estimate for sperm quantity was statistically significantly larger than that of sperm quality ($p < 0.001$). The type of fitness proxy explained 6.3% of heterogeneity ($R^2_{\text{marginal}} = 0.063$).

3.4 PUBLICATION BIAS TESTS

Overall, we found some evidence of publication bias in the published literature, the influence of which was seemingly ameliorated by our approach of including both published and unpublished data. Effect sizes tended to become slightly smaller as sample size increased (i.e. small-study effect; intercept = 0.23 [0.11 – 0.35], $p < 0.001$; slope = -0.001 [-0.002 – 0.000], $p = 0.082$; $n = 36$, $k = 348$; $R^2_{\text{marginal}} = 0.010$; Figure 4). This small-study effect became prominent when only the effect sizes from published papers were considered (Figure S7.1). There was no clear evidence of time-lag bias (i.e. decline effect) in published effect sizes (intercept = 0.32 [0.05 – 0.59], $p = 0.017$; slope = -0.002 [-0.024 – 0.020], $p = 0.834$; $n = 19$, $k = 106$; $R^2_{\text{marginal}} = 0.003$). However, published effect sizes ($r = 0.33$ [0.16 – 0.51], $p < 0.001$, $n = 19$, $k = 106$) tended to be larger than unpublished ones ($r = 0.12$ [-0.05 – 0.29], $p = 0.157$, $n = 17$, $k = 242$), although not statistically significantly so ($p = 0.086$; $R^2_{\text{marginal}} = 0.043$; Figure 5). Finally, as expected, studies reporting data incompletely ($r = 0.53$ [0.12 – 0.95], $p < 0.012$, $n = 5$, $k = 29$) tended to show larger effect sizes than studies reporting data in full ($r = 0.27$ [0.02 – 0.51], $p < 0.032$, $n = 14$, $k = 77$), but that difference was not statistically significant ($p = 0.284$; $R^2_{\text{marginal}} = 0.049$).

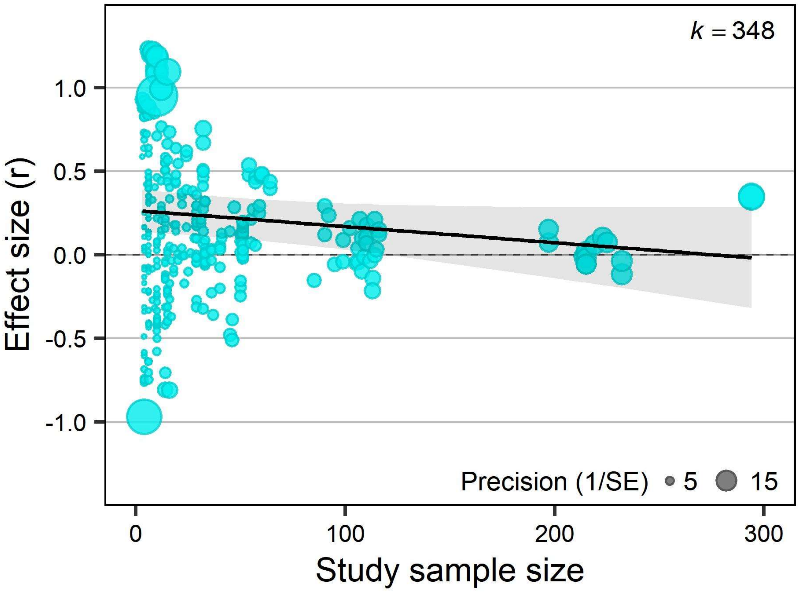

Figure 4. Effect sizes became slightly smaller as sample size increased, demonstrating some evidence of small-study effect. The solid line represents the model estimate, shading represents the 95% CI, and circles represent individual effect sizes scaled by their precision.

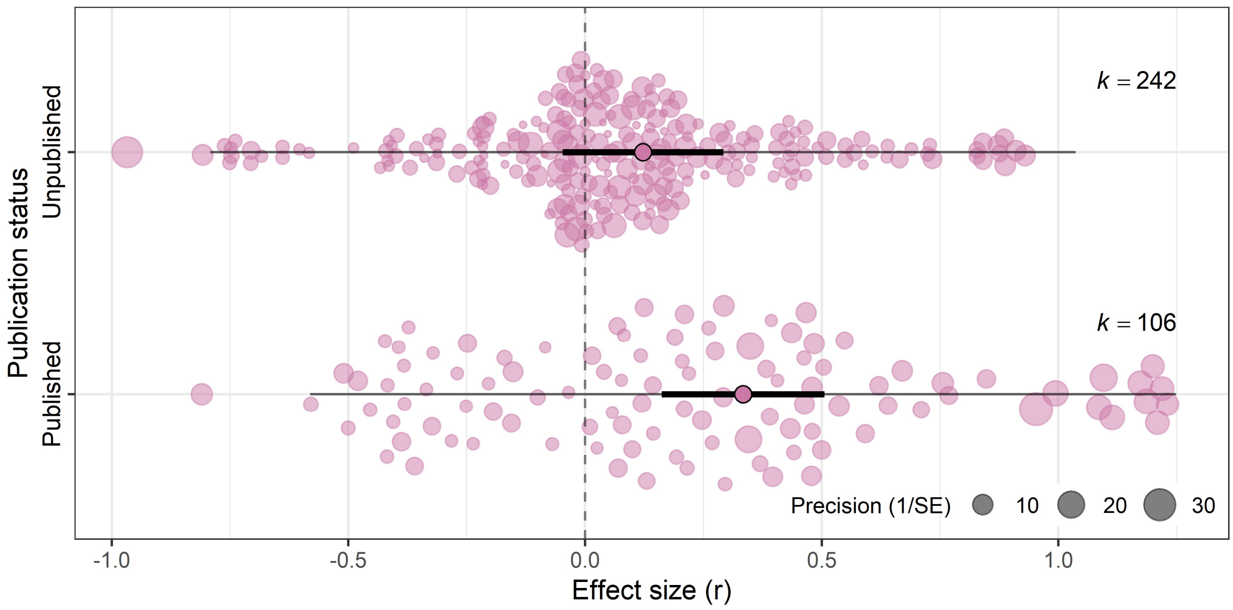

Figure 5. Published effect sizes tended to be larger than unpublished effect sizes for the correlation between male size and fitness in *Gambusia*. Orchard plot showing means, 95% CI (thick whisker), 95% PI (thin whisker), and individual effect sizes scaled by their precision (circles).

4. DISCUSSION

We found that male size and fitness are positively correlated across studies on *Gambusia*. Throughout, all mean effect estimates were positive, including the overall effect and the category-specific meta-regression effects, which suggests that evidence for large male advantage is robust. Whether this seemingly directional and positive selection is driving evolution of male body size in *Gambusia* is unclear, in part because the heritability of body size appears negligible in the most studied *Gambusia* species, eastern mosquitofish (Zulian et al., 1993; Booksmythe et al., 2016; Vega-Trejo et al., 2018). Indeed, environmental effects, including maternal effects, have been found to be important components of male body size in the eastern mosquitofish

(Vega-Trejo et al., 2018). Furthermore, differential selection on the age/size at maturity (e.g. Carmona-Catot et al., 2011; Reznick et al., 2006; Hughes, 1985) is likely a key component explaining why variation in male body size is commonplace in this genus. The overall positive effect contrasts with our prediction and with earlier studies that found a small male advantage, mostly when focusing on gonopodial thrusting as fitness proxy (Bisazza & Marin, 1995). Nonetheless, the high heterogeneity found and consequently wide prediction interval for our main effect highlights that our results do not preclude a small male advantage being the 'true' effect in certain contexts.

Meta-regressions revealed that the type of fitness proxy, the male-to-female ratio, and the type of sperm characteristic are important moderators explaining a sizable amount of heterogeneity. The five categories of fitness proxies we used could be associated with different aspects of sexual selection: Female choice is associated with precopulatory intersexual selection, mating success presumably with both male-male competition (intrasexual selection) and precopulatory intersexual selection, sperm quality and quantity with postcopulatory sexual selection, and paternity with overall reproductive success. The category-specific estimates were generally positive, suggesting large males have an advantage at each level of sexual selection. However, there are reasons to interpret this cautiously. The estimate for paternity, arguably the most direct proxy for fitness in our data, was positive but small and uncertain. The paternity category had the highest number of effect sizes (k = 115) among all five categories, but all of those effect sizes were based on a few males (range: 4 – 36) and came from only four studies. Furthermore, according to the literature, *Gambusia* shows a coercive mating system where small males outperform large males at gonopodial thrusting (e.g. Bisazza & Marin, 1995; Pilastro et al., 1997), and thus, we expected a negative estimate for the mating success category.

Surprisingly, the estimate was still slightly positive, even though this category included many effect sizes for which individual males were tested singly, which potentially benefitted smaller males due to the absence of competitors. As the number of males tested together increased, larger males generally prevailed and performed more gonopodial thrusting (Figure 2; as in Bisazza & Marin, 1995; Deaton, 2008; Booksmythe et al., 2013). The inconspicuousness and manoeuvrability that give smaller males an edge in gonopodial thrusting (Bisazza & Marin, 1995; Pilastro et al., 1997) may be eclipsed by larger males' competitive dominance, and thus, this category may have underestimated the influence of male-male competition.

As predicted, the association between male size and sperm quantity was positive, while the relationship between male size and sperm quality was virtually non-existent. The latter finding contrasted with our prediction for a trade-off between sperm quality and male size/growth. It is possible that sperm competition in this genus is so intense irrespective of male size that no clear association exists between male size and sperm quality (Zane et al., 1999). Moreover, *Gambusia* males may facultatively adjust how much sperm they spend depending on the perceived sperm competition risk instead of altering the quality of their ejaculate (Evans et al., 2003). Future studies are needed to understand the role and mechanism of sperm competition and to disentangle the effect of male size, sperm quantity, and sperm quality, especially since internal fertilization and livebearing make poeciliids an ideal model organism for studying sperm competition.

The female choice category showed a greater estimate than the other categories, challenging us to rethink the role that female choice may play in *Gambusia* and also the way female choice is measured in the laboratory. Of 13 studies that investigated female choice, 11 confirmed female preference for large males, so it is possible that there is a latent female preference whose expression is hindered in the wild but is detectable in the artificial settings of dichotomous mate choice tests. However, it is unclear whether the female association preference represents a preference to reproduce with large males. In the laboratory, *G. holbrooki* females were shown to aggregate with other females to dilute the costs of excessive male sexual harassment such as increased predation risk and reduced foraging efficiency (Dadda et al., 2005). Similarly, females associated with a larger male when a harassing male was present, which may be a strategy to curtail harassment via the larger male monopolizing access to the female and fending off smaller males (Dadda et al., 2005; Searcy, 1982). In nature, *G. holbrooki* females tended to shoal with similar-sized females (Bisazza & Marin, 1995), so female preference for large males may also be a by-product of female schooling behaviour. Future studies into the role of female choice in *Gambusia* should consider the effect of this gregarious tendency in females.

Female choice was mostly measured in dichotomous mate choice tests with no physical interaction between the sexes, which does not reflect the ecological reality of male-female interactions. Instead, researchers could make use of recent advances in tracking technology to study female choice in this group (e.g. Pérez-Escudero et al., 2014; Sridhar et al. 2019). Our analyses revealed a larger effect in the absence than in the presence of physical interaction, so it is possible that female preference for large males was somewhat artificially inflated. When experimental fish did freely interact, experimental density, male-to-female ratio, and the interaction between these and habitat complexity explained a substantial percentage of heterogeneity. When considered singly, male-to-female ratio had a positive effect on the relationship between male size and fitness, explaining the second greatest amount of heterogeneity in this meta-analysis (10.4%). That is, our results suggest that male size is a stronger predictor of fitness when male-male competition is high. It should be kept in mind that separating the effects of male-to-female ratio from the effects of male and female density is difficult; for example, male and female density under varying sex ratios was shown to exert different influence on patterns of male behaviour change in *G. affinis* (Smith, 2007).

Some of the limitations of our meta-analysis reside in the experimental conditions of the included studies. First, all included studies were conducted in the laboratory where *Gambusia* mating behaviour was often measured in unrealistically low complexity settings, making it difficult to draw connections between the results of our meta-analysis and reproductive dynamics in natural populations. Furthermore, even the ‘high complexity’ category in our meta-analysis (small rocks and/or natural or artificial plants) did not reflect the true complexity of natural habitats and was heavily underrepresented (k = 27), which could explain the lack of a clear statistical effect in our meta-regression. Visual field observations revealed that male chases of females in *G. affinis* mostly came to a halt when the chased female dashed into dense vegetation in shallow water (Martin, 1975). Thus, it is likely that females use vegetation to escape from, and selectively not escape from, the males, and this aspect of *Gambusia* mating behaviour was largely overlooked. In addition, most trials were conducted at 28 °C with 14 hours of light period, which does not reflect the natural variation since *Gambusia* can occupy icy lakes and ponds as well as hot springs and thermally elevated lakes reaching 42-44 °C (Meffe & Snelson, 1989). Importantly, *G. holbrooki* males have been observed to reproduce across the entire test temperature range of 14 °C to 38 °C in laboratory (Wilson, 2005). Since temperature and photoperiod are generally regarded as the two most vital environmental factors in fish reproductive cycle, how photoperiod and temperature interact to control *Gambusia* reproduction requires further investigation. Specifically, attention should be paid to seasonal and daily fluctuations, which might have greater influence than the test temperature and photoperiod.

Female reproductive status is another important factor to consider when studying *Gambusia* mating behaviour. Although females try to thwart male copulatory attempts at all stages of their reproductive cycle (Bisazza & Marin, 1995), mosquitofish females have been suggested to more likely associate with males when virgin, postpartum, or male-deprived (Hughes, 1985; Pilastro et al., 2003; Bisazza et al., 2001). Thus, we hypothesized larger effect sizes for virgin or postpartum females than for gravid females. Unfortunately, there were insufficient effect sizes to calculate an estimate for postpartum females because many studies excluded postpartum females due to heightened male interest (Constanz, 1989), which was deemed a confounding variable for some research questions. If female receptivity and male interest are at their peak 1-2 days after parturition, future sexual selection studies may benefit from focusing more on postpartum females, not less, which would help avoid a systematic design issue that underestimates the role of female behaviour and mate choice.

Our systematic review and meta-analysis also underscored evidence of publication bias in the published literature. First, our analyses showed some evidence of small-study bias, suggesting that some low-precision studies might still remain unavailable despite our efforts to include both published and unpublished studies. Evidence for small-study bias is often found in meta-analysis in ecology and evolution and needs to be considered when interpreting meta-analytic results (e.g. Parker, 2013; Wang et al., 2018; Sánchez-Tójar et al., 2020a; 2020b). The existence of publication bias was further demonstrated since published effect sizes tended to be larger than unpublished effect sizes, and studies reporting data incompletely also tended to show larger effect sizes than studies reporting data in full. Similar patterns have been shown in recent meta-analyses in the field (Sánchez-Tójar et al., 2018; Moran et al., 2020), and we expect these patterns to be more and more commonly uncovered since meta-analysts have started to make use of open data (Culina et al., 2018). Despite the evidence of publication bias in the published literature, our approach of combining both published and unpublished data largely mitigated its effect (Figure S7.1). However, some caution should still be taken when interpreting the results of our meta-analysis.

In sum, our meta-analysis found evidence of positive selection on male body size in *Gambusia* that was seemingly robust across contexts. We found gaps and limitations in experimental designs used to study *Gambusia* mating behaviour, which should help guiding the necessary future research on this topic, particularly since our meta-analysis revealed a large proportion of unexplained heterogeneity across effect sizes. Our study also identified the need to rethink the role and form of female choice in this genus and how it is measured in the laboratory. Female choice may play a subtle and underestimated part, and association preference for large males for protection could be a means through which females may exert some amount of choice in an ostensibly coercive mating system.

ACKNOWLEDGEMENTS

We thank Andrea S. Aspbury, Michael G. Bertram, Isobel Booksmythe, Thea M. Edwards, Megan L. Head, Andrew T. Kahn, Jake M. Martin, Rose E. O'Dea, and Regina Vega-Trejo for kindly sharing their raw data. We are also grateful to Michael G. Bertram for commenting on our pre-registration, and to Megan L. Head for commenting on an earlier version of this manuscript. This research received funding from the German Research Foundation (DFG) as part of the SFB TRR 212 (NC^3; project numbers 316099922, 396782608) and the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 836937.

AUTHORS' CONTRIBUTIONS

BK, NPM, and AST were involved in conceptualization, methodology, software, formal analysis and investigation; BK performed data curation, and project administration; BK and NPM performed visualization; NPM and AST performed supervision and validation; KR was involved in funding acquisition; BK and AST performed the writing – original draft preparation, and BK, NPM, KR and AST performed writing – review and editing.

DATA AVAILABILITY

All data and code are available on the Open Science Framework (http://doi.org/10.17605/OSF.IO/2QXT5). Upon acceptance, all data and code will be timestamped via Zenodo.

REFERENCES

The asterisk (*) indicates articles used in the meta-analysis.

Andersson, M. B. (1994). *Sexual selection*. Princeton University Press.

*Bertram, M. G., Ecker, T. E., Wong, B. B., O'Bryan, M. K., Baumgartner, J. B., Martin, J. M., & Saaristo, M. (2018). The antidepressant fluoxetine alters mechanisms of pre- and postcopulatory sexual selection in the eastern mosquitofish (*Gambusia holbrooki*). *Environmental Pollution*, *238*, 238-247. https://doi.org/10.1016/j.envpol.2018.03.006

Birkhead, T. R., & Pizzari, T. (2002). Postcopulatory sexual selection. *Nature Reviews Genetics*, *3*(4), 262–273. https://doi.org/10.1038/nrg774

*Bisazza, A., & Marin, G. (1991). Male size and female choice in the eastern mosquitofish (*Gambusia holbrooki*, Poeciliidae). *Copeia, 1991*(3), 730-735. https://doi.org/10.2307/1446400

*Bisazza, A., & Marin, G. (1995). Sexual selection and sexual size dimorphism in the eastern mosquitofish *Gambusia holbrooki* (Pisces Poeciliidae). *Ethology Ecology & Evolution*, *7*(2), 169–183. https://doi.org/10.1080/08927014.1995.9522963

*Bisazza, A. (1993). Male competition, female mate choice and sexual size dimorphism in poeciliid fishes. *Marine Behaviour and Physiology*, *23*(1–4), 257–286. https://doi.org/10.1080/10236249309378869

*Bisazza, A., Vaccari, G., & Pilastro, A. (2001). Female mate choice in a mating system dominated by male sexual coercion. *Behavioral Ecology*, *12*(1), 59–64. https://doi.org/10.1093/oxfordjournals.beheco.a000379

*Booksmythe, I., Backwell, P. R. Y., & Jennions, M. D. (2013). Competitor size, male mating success and mate choice in eastern mosquitofish, *Gambusia holbrooki*. *Animal Behaviour*, *85*(2), 371–375. https://doi.org/10.1016/j.anbehav.2012.11.009

*Booksmythe, I., Head, M. L., Keogh, J. S., & Jennions, M. D. (2016). Fitness consequences of artificial selection on relative male genital size. *Nature Communications*, *7*(1), 11597. https://doi.org/10.1038/ncomms11597

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470743386

Boschetto, C., Gasparini, C., & Pilastro, A. (2011). Sperm number and velocity affect sperm competition success in the guppy (*Poecilia reticulata*). *Behavioral Ecology and Sociobiology*, *65*(4), 813–821. https://doi.org/10.1007/s00265-010-1085-y

Carmona-Catot, G., Benito, J., & García-Berthou, E. (2011). Comparing latitudinal and upstream-downstream gradients: life history traits of invasive mosquitofish: comparing latitudinal and upstream-downstream gradients. *Diversity and Distributions*, *17*(2), 214–224. https://doi.org/10.1111/j.1472-4642.2011.00743.x

*Chen, B., Liu, K., Zhou, L., Gomes-Silva, G., Sommer-Trembo, C., & Plath, M. (2018). Personality differentially affects individual mate choice decisions in female and male Western mosquitofish (*Gambusia affinis*). *PLoS ONE*, *13*(5), e0197197. https://doi.org/10.1371/journal.pone.0197197

Constanz, G. D. (1989). Reproductive biology of poeciliid fishes. In *Ecology and evolution of livebearing fishes (Poeciliidae)* (pp. 33–50). Prentice Hall.

Culina, A., Crowther, T. W., Ramakers, J. J. C., Gienapp, P., & Visser, M. E. (2018). How to do meta-analysis of open datasets. *Nature Ecology & Evolution*, *2*(7), 1053–1056. https://doi.org/10.1038/s41559-018-0579-2

Cureton II, J., Deaton, R., & Martin, R. (2010). Short term changes in sex ratio and density alter coercive male mating tactics. *Behaviour*, *147*(11), 1431–1442. https://doi.org/10.1163/000579510X519495

*Dadda, M. (2015). Female social response to male sexual harassment in poeciliid fish: a comparison of six species. *Frontiers in Psychology, 6*, 1-9. https://doi.org/10.3389/fpsyg.2015.01453

*Dadda, M., Pilastro, A., & Bisazza, A. (2005). Male sexual harassment and female schooling behaviour in the eastern mosquitofish. *Animal Behaviour*, *70*(2), 463–471. https://doi.org/10.1016/j.anbehav.2004.12.010

*Dadda, M., Pilastro, A., & Bisazza, A. (2008). Innate responses to male sexual harassment in female mosquitofish. *Behavioral Ecology and Sociobiology, 63*, 53-62. https://doi.org/10.1007/s00265-008-0635-z

*Deaton, R. (2008a). Factors influencing male mating behaviour in *Gambusia affinis* (Baird & Girard) with a coercive mating system. *Journal of Fish Biology*, *72*(7), 1607–1622. https://doi.org/10.1111/j.1095-8649.2008.01827.x

*Deaton, R. (2008b). Use of microsatellite paternity analysis to determine male mating success in the western mosquitofish, *Gambusia affinis*. *Behaviour*, *145*(6), 795–814. https://doi.org/10.1163/156853908783929151

*Edwards, T. M., & Guillette Jr., L. J. (2007). Reproductive characteristics of male mosquitofish (*Gambusia holbrooki*) from nitrate-contaminated springs in Florida. *Aquatic Toxicology, 85*(1), 40-47. https://doi.org/10.1016/j.aquatox.2007.07.014

*Edwards, T. M., Miller, H. D., Toft, G., & Guillette Jr., L. J. (2013). Seasonal reproduction of male *Gambusia holbrooki* (eastern mosquitofish) from two Florida lakes. *Fish Physiology and Biochemistry, 39,* 1165-1180. https://doi.org/10.1007/s10695-013-9772-z

Emlen, S., & Oring, L. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*, *197*(4300), 215–223. https://doi.org/10.1126/science.327542

*Espinedo, C. M., Gabor, C. R., & Aspbury, A. S. (2010). Males, but not females, contribute to sexual isolation between two sympatric species of *Gambusia*. *Evolutionary Ecology, 24*, 865-878. https://doi.org/10.1007/s10682-009-9343-z

Evans, J. P., Pierotti, M., & Pilastro, A. (2003). Male mating behavior and ejaculate expenditure under sperm competition risk in the eastern mosquitofish. *Behavioral Ecology*, *14*(2), 268–273. https://doi.org/10.1093/beheco/14.2.268

Froese, R., & Pauly, D. (Eds.). (2000). *FishBase 2000: concepts, design and data sources*. The International Center for Living Aquatic Resources Management. https://www.fishbase.de/identification/SpeciesList.php?genus=Gambusia

García-González, F., & Simmons, L. W. (2005). Sperm viability matters in insect sperm competition. *Current Biology*, *15*(3), 271–275. https://doi.org/10.1016/j.cub.2005.01.032

*Head, M. L., Jacomb, F., Vega-Trejo, R., & Jennions, M. D. (2015a). Male mate choice and insemination success under simultaneous versus sequential choice conditions. *Animal Behaviour*, *103*, 99–105. https://doi.org/10.1016/j.anbehav.2015.02.011

*Head, M. L., Kahn, A. T., Henshaw, J. M., Keogh, J. S., & Jennions, M. D. (2017). Sexual selection on male body size, genital length and heterozygosity: consistency across habitats and social settings. *Journal of Animal Ecology*, *86*(6), 1458–1468. https://doi.org/10.1111/1365-2656.12742

Head, M. L., Lindholm, A. K., & Brooks, R. (2007). Operational sex ratio and density do not affect directional selection on male sexual ornaments and behavior: Demographic effects on sexual selection. *Evolution*, *62*(1), 135–144. https://doi.org/10.1111/j.1558-5646.2007.00277.x

Head, M. L., Vega-Trejo, R., Jacomb, F., & Jennions, M. D. (2015b). Predictors of male insemination success in the mosquitofish (*Gambusia holbrooki*). *Ecology and Evolution*, *5*(21), 4999–5006. https://doi.org/10.1002/ece3.1775

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Holman, L., Head, M. L., Lanfear, R., & Jennions, M. D. (2015). Evidence of experimental bias in the life sciences: why we need blind data recording. *PLoS Biology*, *13*(7), e1002190. https://doi.org/10.1371/journal.pbio.1002190

*Hughes, A. L. (1985). Male size, mating success, and mating strategy in the mosquitofish *Gambusia affinis* (Poeciliidae). *Behavioral Ecology and Sociobiology*, *17*(3), 271–278. https://doi.org/10.1007/BF00300146

Iglesias-Carrasco, M., Fox, R. J., Vega-Trejo, R., Jennions, M. D., & Head, M. L. (2019). An experimental test for body size-dependent effects of male harassment and an elevated copulation rate on female lifetime fecundity and offspring performance. *Journal of Evolutionary Biology*, *32*(11), 1262–1273. https://doi.org/10.1111/jeb.13526

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Itzkowitz, M. (1971). Preliminary study of the social behavior of male *Gambusia affinis* (Baird and Girard) (Pisces: Poeciliidae) in aquaria. *Chesapeake Science*, *12*(4), 219. https://doi.org/10.2307/1350908

Jacobs, P., & Viechtbauer, W. (2017). Estimation of the biserial correlation and its sampling variance for use in meta-analysis: biserial Correlation. *Research Synthesis Methods*, *8*(2), 161–180. https://doi.org/10.1002/jrsm.1218

Jennions, M. D., Kokko, H., & Klug, H. (2012). The opportunity to be misled in studies of sexual selection: the opportunity to be misled in studies of sexual selection. *Journal of Evolutionary Biology*, *25*(3), 591–598. https://doi.org/10.1111/j.1420-9101.2011.02451.x

Jennions, Michael D., & Møller, A. P. (2002). Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1486), 43–48. https://doi.org/10.1098/rspb.2001.1832

*Kahn, A. T., Livingston, J. D., & Jennions, M. D. (2012). Do females preferentially associate with males given a better start in life? *Biology Letters, 8*(3), 362-364. https://doi.org/10.1098/rsbl.2011.1106

*Kahn, A. T., Mautz, B., & Jennions, M. D. (2010). Females prefer to associate with males with longer intromittent organs in mosquitofish. *Biology Letters*, *6*(1), 55–58. https://doi.org/10.1098/rsbl.2009.0637

Kim, B., Moran, N. P., Reinhold, K., & Sánchez-Tójar, A. (2019). *The fitness consequences of male size in* Gambusia*: a systematic review and meta-analysis* [Preregistration]. Open Science Framework. https://doi.org/10.17605/OSF.IO/MWCS4

Klug, H., Heuschele, J., Jennions, M. D., & Kokko, H. (2010). The mismeasurement of sexual selection. *Journal of Evolutionary Biology*, *23*(3), 447–462. https://doi.org/10.1111/j.1420-9101.2009.01921.x

Koricheva, J., & Kulinskaya, E. (2019). Temporal instability of evidence base: a threat to policy making? *Trends in Ecology & Evolution*, *34*(10), 895–902. https://doi.org/10.1016/j.tree.2019.05.006

Kvarnemo, C., & Ahnesjö, I. (1996). The dynamics of operational sex ratios and competition for mates. *Trends in Ecology & Evolution*, *11*(10), 404–408. https://doi.org/10.1016/0169-5347(96)10056-2

Lajeunesse, M. J. (2013). Recovering missing or partial data from studies: a survey of conversions and imputations for meta-analysis. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution* (pp. 195–206). Princeton University Press.

Lindenfors, P., Tullberg, B., & Biuw, M. (2002). Phylogenetic analyses of sexual selection and sexual size dimorphism in pinnipeds. *Behavioral Ecology and Sociobiology*, *52*(3), 188–193. https://doi.org/10.1007/s00265-002-0507-x

*Locatello, L., Rasotto, M. B., Adriaenssens, B., & Pilastro, A. (2008). Ejaculate traits in relation to male body size in the eastern mosquitofish *Gambusia holbrooki*. *Journal of Fish Biology*, *73*(7), 1600–1611. https://doi.org/10.1111/j.1095-8649.2008.02034.x

*Lynch, K. S., Ramsey, M. E., & Cummings, M. E. (2012). The mate choice brain: comparing gene profiles between female choice and male coercive poeciliids. *Genes, Brain and Behavior, 11*(2), 222-229. https://doi.org/10.1111/j.1601-183X.2011.00742.x

*Marsh, J. N., Vega-Trejo, R., Jennions, M. D., & Head, M. L. (2017). Why does inbreeding reduce male paternity? Effects on sexually selected traits. *Evolution*, *71*(11), 2728–2737. https://doi.org/10.1111/evo.13339

*Martin, J. M., Bertram, M. G., Saaristo, M., Ecker, T. E., Hannington, S. L., Tanner, J. L., Michelangeli, M., O'Bryan, M. K., & Wong, B. B. M. (2019). Impact of the widespread pharmaceutical pollutant fluoxetine on behaviour and sperm traits in a freshwater fish. *Science of the Total Environment, 650*(2), 1771-1778. https://doi.org/10.1016/j.scitotenv.2018.09.294

Martin, R. G. (1975). Sexual and aggressive behavior, density and social structure in a natural population of mosquitofish, *Gambusia affinis holbrooki*. *Copeia*, *1975*(3), 445. https://doi.org/10.2307/1443641

McLachlan, A. J., & Allen, D. F. (1987). Male mating success in Diptera: advantages of small size. *Oikos*, *48*(1), 11. https://doi.org/10.2307/3565681

*McPeek, M. A. (1992). Mechanisms of sexual selection operating on body size in the mosquitofish (*Gambusia holbrooki*). *Behavioral Ecology*, *3*(1), 1–12. https://doi.org/10.1093/beheco/3.1.1

Meffe, G. K., & Snelson, F. F. (1989). An ecological overview of poeciliid fishes. In F. F. Snelson & G. K. Meffe (Eds.), *Ecology and evolution of livebearing fishes (Poeciliidae)*. Prentice Hall.

Moher, D., Liberati, A., Tetylaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Annals of Internal Medicine*, *151*(4), 264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moran, N. P., Sánchez-Tójar, A., Schielzeth, H., & Reinhold, K. (2020). Poor nutritional condition promotes high-risk behaviours: a systematic review and meta-analysis. *Biological Reviews, 96*(1), 269-288. https://doi.org/10.1111/brv.12655

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, *82*(4), 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Nakagawa, S., Lagisz, M., O’Dea, R. E., Rutkowska, J., Yang, Y., Noble, D. W. A., & Senior, A. M. (2020). The orchard plot: cultivating a forest plot for use in ecology, evolution, and beyond. *Research Synthesis Methods*, jrsm.1424. https://doi.org/10.1002/jrsm.1424

Nakagawa, S., & Poulin, R. (2012). Meta-analytic insights into evolutionary ecology: an introduction and synthesis. *Evolutionary Ecology*, *26*(5), 1085–1099. https://doi.org/10.1007/s10682-012-9593-z

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, *26*(5), 1253–1274. https://doi.org/10.1007/s10682-012-9555-5

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Neems, R. M., Lazarus, J., & McLachlan, A. J. (1998). Lifetime reproductive success in a swarming midge: trade-offs and stabilizing selection for male body size. *Behavioral Ecology*, *9*(3), 279–286. https://doi.org/10.1093/beheco/9.3.279

*O'Dea, R. E., Jennions, M. D., & Head, M. L. (2014). Male body size and condition affects sperm number and production rates in mosquitofish, *Gambusia holbrooki*. *Journal of Evolutionary Biology*, *27*(12), 2739–2744. https://doi.org/10.1111/jeb.12534

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Parker, T. H. (2013). What do we really know about the signalling role of plumage colour in blue tits? A case study of impediments to progress in evolutionary biology: case study of impediments to progress. *Biological Reviews*, *88*(3), 511–536. https://doi.org/10.1111/brv.12013

Pérez-Escudero, A., Vicente-Page, J., Hinz, R. C., Arganda, S., & Polavieja, G. G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nature Methods, 11*, 743-748. https://doi.org/10.1038/nmeth.2994

Pick, J. L., Nakagawa, S., & Noble, D. W. A. (2019). Reproducible, flexible and high-throughput data extraction from primary literature: the metaDigitise R package. *Methods in Ecology and Evolution*, *10*(3), 426–431. https://doi.org/10.1111/2041-210X.13118

*Pilastro, A., Benetton, S., & Bisazza, A. (2003). Female aggregation and male competition reduce costs of sexual harassment in the mosquitofish *Gambusia holbrooki*. *Animal Behaviour*, *65*(6), 1161–1167. https://doi.org/10.1006/anbe.2003.2118

*Pilastro, A., Giacomello, E., & Bisazza, A. (1997). Sexual selection for small size in male mosquitofish (*Gambusia holbrooki*). *Proceedings of the Royal Society B: Biological Sciences*, *264*(1385), 1125–1129. https://doi.org/10.1098/rspb.1997.0155

Pyke, G. H. (2005). A review of the biology of *Gambusia affinis* and *G. holbrooki*. *Reviews in Fish Biology and Fisheries*, *15*(4), 339–365. https://doi.org/10.1007/s11160-006-6394-x

R Core Team. (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Reznick, D., Schultz, E., Morey, S., & Roff, D. (2006). On the virtue of being the first born: the influence of date of birth on fitness in the mosquitofish, *Gambusia affinis*. *Oikos*, *114*(1), 135–147. https://doi.org/10.1111/j.2006.0030-1299.14446.x

Rios Moura, R., & Peixoto, P. E. C. (2013). The effect of operational sex ratio on the opportunity for sexual selection: a meta-analysis. *Animal Behaviour*, *86*(4), 675–683. https://doi.org/10.1016/j.anbehav.2013.07.002

Roff, D. A. (2002). *Life history evolution*. Sinauer Associates.

*Saaristo, M., Myers, J., Jacques-Hamilton, R., Allinson, M., Yamamoto, A., Allinson, G., Pettigrove, V., & Wong, B. B. M. (2014). Altered reproductive behaviours in male mosquitofish living downstream from a sewage treatment plant. *Aquatic Toxicology, 149*, 58-64. https://doi.org/10.1016/j.aquatox.2014.02.001

*Saaristo, M., Tomkins, P., Allinson, M., Allinson, G., & Wong, B. B. M. (2013). An androgenic agricultural contaminant impairs female reproductive behaviour in a freshwater fish. *PLoS ONE*, *8*(5), e62782. https://doi.org/10.1371/journal.pone.0062782

Sánchez-Tójar, A., Lagisz, M., Moran, N. P., Nakagawa, S., Noble, D. W. A., & Reinhold, K. (2020). The jury is still out regarding the generality of adaptive 'transgenerational' effects. *Ecology Letters*, *23*(11), 1715–1718. https://doi.org/10.1111/ele.13479

Sánchez-Tójar, A., Moran, N. P., O'Dea, R. E., Reinhold, K., & Nakagawa, S. (2020). Illustrating the importance of meta-analysing variances alongside means in ecology and evolution. *Journal of Evolutionary Biology*, 33(9): 1216-1223. https://doi.org/10.1111/jeb.13661

Sánchez-Tójar, A., Nakagawa, S., Sánchez-Fortún, M., Martin, D. A., Ramani, S., Girndt, A., Bókony, V., Kempenaers, B., Liker, A., Westneat, D. F., Burke, T., & Schroeder, J. (2018). Meta-analysis challenges a textbook example of status signalling and demonstrates publication bias. *ELife*, *7*, e37385. https://doi.org/10.7554/eLife.37385

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients: interpretation of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Searcy, W. A. (1982). The evolutionary effects of mate selection. *Annual Review of Ecology and Systematics*, *13*(1), 57–85. https://doi.org/10.1146/annurev.es.13.110182.000421

Smith, C. C. (2007). Independent effects of male and female density on sexual harassment, female fitness, and male competition for mates in the western mosquitofish *Gambusia affinis*. *Behavioral Ecology and Sociobiology*, *61*(9), 1349–1358. https://doi.org/10.1007/s00265-007-0365-7

Smith, C. C., & Sargent, R. C. (2006). Female fitness declines with increasing female density but not male harassment in the western mosquitofish, *Gambusia affinis*. *Animal Behaviour*, *71*(2), 401–407. https://doi.org/10.1016/j.anbehav.2005.06.003

Snelson, F. F., & Meffe, G. K. (Eds.). (1989). *Ecology and evolution of livebearing fishes (Poeciliidae)*. Prentice Hall.

Sridhar, V. H., Roche, D. G., & Gingins, S. (2019). Tracktor: image-based automated tracking of animal movement and behaviour. *Methods in Ecology and Evolution, 10*(6), 815-820. https://doi.org/10.1111/2041-210X.13166

Vega-Trejo, R., Fox, R. J., Iglesias-Carrasco, M., Head, M. L., & Jennions, M. D. (2019). The effects of male age, sperm age and mating history on ejaculate senescence. *Functional Ecology*, *33*(7), 1267–1279. https://doi.org/10.1111/1365-2435.13305

*Vega-Trejo, R., Head, M. L., Jennions, M. D., & Kruuk, L. E. B. (2018). Maternal-by-environment but not genotype-by-environment interactions in a fish without parental care. *Heredity*, *120*(2), 154–167. https://doi.org/10.1038/s41437-017-0029-y

*Vega-Trejo, R., Head, M. L., Keogh, J. S., & Jennions, M. D. (2017). Experimental evidence for sexual selection against inbred males. *Journal of Animal Ecology, 86*(2), 394-404. https://doi.org/10.1111/1365-2656.12615

*Vega-Trejo, R., Jennions, M. D., & Head, M. L. (2016). Are sexually selected traits affected by a poor environment early in life? *BMC Evolutionary Biology, 16*, 1-12. https://doi.org/10.1186/s12862-016-0838-2

*Vega-Trejo, R., O'Dea, R. E., Jennions, M. D., & Head, M. L. (2014). The effects of familiarity and mating experience on mate choice in mosquitofish, *Gambusia holbrooki*. *Behavioral Ecology, 25*(5), 1205-1211. https://doi.org/10.1093/beheco/aru113

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Walling, C. A., Royle, N. J., Lindström, J., & Metcalfe, N. B. (2010). Do female association preferences predict the likelihood of reproduction? *Behavioral Ecology and Sociobiology*, *64*(4), 541–548. https://doi.org/10.1007/s00265-009-0869-4

Wang, D., Forstmeier, W., Ihle, M., Khadraoui, M., Jerónimo, S., Martin, K., & Kempenaers, B. (2018). Irreproducible text-book "knowledge": the effects of color bands on zebra finch fitness: Color bands have no effect on fitness in zebra finches. *Evolution*, *72*(4), 961–976. https://doi.org/10.1111/evo.13459

*Wang, S. M. T., Ramsey, M. E., & Cummings, M. E. (2014). Plasticity of the mate choice mind: Courtship evokes choice-like brain responses in females from a coercive mating system. *Gene, Brain and Behavior, 13*(4), 365-375. https://doi.org/10.1111/gbb.12124

Westgate, M. J. (2018). *revtools: tools to support evidence synthesis* (0.3.0) [Computer software]. The Comprehensive R Archive Network. https://CRAN.R-project.org/package=revtools

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wilson, R. S. (2005). Temperature influences the coercive mating and swimming performance of male eastern mosquitofish. *Animal Behaviour*, *70*(6), 1387–1394. https://doi.org/10.1016/j.anbehav.2004.12.024

Zane, L., Nelson, W. S., Jones, A. G., & Avise, J. C. (1999). Microsatellite assessment of multiple paternity in natural populations of a live-bearing fish, *Gambusia holbrooki*. *Journal of Evolutionary Biology*, *12*(1), 61–69. https://doi.org/10.1046/j.1420-9101.1999.00006.x

Zulian, E., Bisazza, A., & Marin, G. (1995). Variations in male body size in natural populations of *Gambusia holbrooki*. *Ethology Ecology & Evolution*, *7*(1), 1–10. https://doi.org/10.1080/08927014.1995.9522966

Zulian, Elsa, Bisazza, A., & Marin, G. (1993). Determinants of size in male eastern mosquitofish (*Gambusia holbrooki):* inheritance and plasticity of a sexual selected character. *Bolletino Di Zoologia*, *60*(3), 317–322. https://doi.org/10.1080/11250009309355831