A rigorous comparison of sexual selection indexes via simulations of diverse mating systems

Jonathan M. Henshaw[a,1,2], Andrew T. Kahn[a], and Karoline Fritzsche[b,1]

[a]Division of Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra ACT 0200, Australia; and [b]Institute of Zoology, University of Graz, Graz 8010, Austria

Edited by Sarah P. Otto, University of British Columbia, Vancouver, BC, Canada, and approved December 7, 2015 (received for review September 10, 2015)

Sexual selection is a cornerstone of evolutionary theory, but measuring it has proved surprisingly difficult and controversial. Various proxy measures—e.g., the Bateman gradient and the opportunity for sexual selection—are widely used in empirical studies. However, we do not know how reliably these measures predict the strength of sexual selection across natural systems, and most perform poorly in theoretical worst-case scenarios. Here we provide a rigorous comparison of eight commonly used indexes of sexual selection. We simulated 500 biologically plausible mating systems, based on the templates of five well-studied species that cover a diverse range of reproductive life histories. We compared putative indexes to the actual strength of premating sexual selection, measured as the strength of selection on a simulated "mating trait." This method sidesteps a key weakness of empirical studies, which lack an appropriate yardstick against which proxy measures can be assessed. Our model predicts that, far from being useless, the best proxy measures reliably track the strength of sexual selection across biologically realistic scenarios. The maximum intensity of precopulatory sexual selection s'_{max} (the Jones index) outperformed all other indexes and was highly correlated with the strength of sexual selection. In contrast, the Bateman gradient and the opportunity for sexual selection were poor predictors of sexual selection, despite their continuing popularity.

opportunity for selection | index of resource monopolization | Morisita index | selection gradient | distributional selection differential

Our understanding of sexual selection is obstructed by lack of progress on a fundamental question: How do we measure sexual selection rigorously, and how do we compare its strength between the sexes or across taxa (1–3)? Sexual selection underlies some of the most extreme traits in nature, including spectacular armaments and ornaments, and affects macroevolution via both speciation and extinction (4–8). It is thought to vary systematically with a species' mating system (i.e., the degree of multiple mating by each sex) and with sex-specific investment in offspring and mating partners (i.e., parental investment and direct benefits) (9, 10). Testing such relationships requires a theoretically well-justified method to estimate the strength of sexual selection within a sex and taxon. Although many potential "proxy measures" have been proposed—e.g., the opportunity for sexual selection and the Bateman gradient (11–13)—none is widely accepted and most perform poorly under theoretical worst-case scenarios (1, 14). Here we use simulation models to show that, far from being useless, the best proxy measures track the strength of premating sexual selection faithfully under biologically relevant conditions. Our model explicitly specifies a focal "mating trait" that is under sexual selection. This allows us to compare proxy measures against the actual strength of sexual selection on the trait. We thereby sidestep a key weakness of empirical studies, which lack an appropriate yardstick against which proxy measures can be assessed.

Sexual Selection on Traits

The theoretical basis for measuring sexual selection on phenotypic traits is well established. Selection and mating differentials measure how strongly a trait is associated with reproductive and mating success, respectively, whereas the use of partial selection gradients allows one to control for indirect selection via correlated traits (13, 15). Unfortunately, trait-based estimates serve poorly in comparative studies of sexual selection, because there is no a priori method to determine which traits are sexually selected (2, 16). Failure to include the primary targets of sexual selection in analyses will bias conclusions, particularly when the ease of identifying such traits covaries systematically with other factors of interest. As an example, human researchers may identify visually based targets of sexual selection more easily than olfactory traits, leading to biased inferences when taxa differ in primary sensory modalities. Choosing a common trait, such as body size, as the basis for comparison does nothing to escape this problem, because it cannot be assumed that the trait's relative importance to sexual selection is constant across taxa.

Measures Based on Variance in Mating or Reproductive Success Alone

Due to the limitations of trait-based measures for comparative studies, researchers have developed a host of proxy measures that depend only on generalized aspects of mating and reproduction (Table 1) (17). Most measures are functions of the variance in mating or reproductive success, under the assumption that higher variance correlates with higher sexual selection, all else being equal. For instance, the opportunity for selection I and opportunity for sexual selection I_s represent the variance in relative reproductive success and relative mating success, respectively (11, 13, 18). These two measures have a clear connection to selection theory because they are strict upper bounds on the linear selection and mating differentials on standardized

Significance

How does sexual selection differ between males and females? What is its role in the speciation process? Answering such questions requires a reliable method to measure sexual selection, so that we can compare its strength between the sexes and across taxa. The development of appropriate measures has led to sustained controversy, however, with no measure being widely accepted. We provide to our knowledge the first rigorous comparison of indexes of sexual selection. Our results unambiguously support the use of a recently defined measure, the Jones index, over more established indexes of sexual selection such as the Bateman gradient.

Author contributions: J.M.H., A.T.K., and K.F. designed research; J.M.H. and K.F. performed research; J.M.H., A.T.K., and K.F. contributed new reagents/analytic tools; J.M.H. analyzed data; and J.M.H. and K.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in Dryad Digital Repository, datadryad.org (doi:10.5061/dryad.3p2j0).

[1]J.M.H. and K.F. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: jonathan.henshaw@anu.edu.au.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518067113/-/DCSupplemental.

Table 1. Proposed measures of the strength of sexual selection

Measure	Description	Formula
Variance-based measures		
Opportunity for selection, I	Variance in relative reproductive success	$\frac{\sigma_R^2}{\mu_R^2}$
Opportunity for sexual selection, I_s	Variance in relative mating success	$\frac{\sigma_M^2}{\mu_M^2}$
Morisita index for reproductive success, $I_{\delta-\mathrm{RS}}$, or mating success, $I_{\delta-\mathrm{MS}}$	Observed variance in success relative to the expected variance if all individuals competed equally	$N\left(\frac{\sum_i X_i^2 - \sum_i X_i}{\left(\sum_i X_i\right)^2 - \sum_i X_i}\right)$, with $X = R, M$ as appropriate
Monopolization index for reproductive success, Q_{RS}, or mating success, Q_{MS}	Observed variance in success relative to the maximum possible variance	$\frac{\sigma_X^2 - \mu_X}{N\mu_X^2 - \mu_X}$, with $X = R, M$ as appropriate
Bateman gradient, β_{ss}	Slope of least-squares regression of relative reproductive success on relative mating success	$\frac{\mu_M \sigma_{R,M}}{\mu_R \sigma_M^2}$
Jones index, $s'_{\max}$	Maximum strength of premating sexual selection $m\beta_{ss}$ (see below)	$\beta_{ss}\sqrt{I_s}$
Trait-based measures		
Linear selection differential, s'	Covariance between standardized trait values and relative reproductive success	$\frac{\sigma_{R,Z}}{\mu_R}$
Quadratic selection differential, c'	Covariance between the square of standardized trait values Z^2 and relative reproductive success	$\frac{\sigma_{R,Z^2}}{\mu_R}$
Distributional selection differential, d'	Total change in the trait distribution due to selection (see SI)	$\int_{-\infty}^{\infty} \lvert F(z) - G(z) \rvert \, dz$
Mating differential, m'	Covariance between standardized trait values and relative mating success	$\frac{\sigma_{M,Z}}{\mu_M}$
Strength of premating sexual selection	Portion of the selection differential due to premating sexual selection	$\beta_{ss} m'$

Formulas are shown for absolute reproductive success R, absolute mating success M, and a standardized quantitative trait Z, in terms of the number of individuals N of the relevant sex, the mean μ_X and variance σ_X^2 of any variable X, and the covariance $\sigma_{X,Y}$ between X and Y. We write X_i for the ith individual's value of X. The cumulative distribution functions of trait values before and after selection are given by F and G, respectively.

phenotypic traits (10, 13). They have been criticized, however, because their expected values depend on the number of competing individuals and the mean success per individual, even when mating is entirely random (1, 19–22). The Morisita index (23, 24) and the monopolization index (20) aim to correct this problem by dividing the realized variance by the expected or maximum possible variance, respectively.

Measures based on variance in either mating or reproductive success alone provide at best a partial picture of sexual selection. Variance in mating success is overlooked by selection unless it translates into variance in reproductive success (25). For instance, suppose fecund females attract more males, but females do not benefit from additional mates after the first one. In this case, females could show high variance in mating success (i.e., high I_s) despite negligible selection to increase their number of mates. High variance in mate number therefore need not indicate sexual selection. Moreover, mating success is not the only contributor to reproductive success, which depends also on naturally selected fitness components such as fecundity. Measures based on reproductive success alone consequently incorporate variance that is not due to sexual selection.

Putting It All Together: The Bateman Gradient and the Jones Index

The desire to integrate the roles of mating and reproductive success led to the development of two new measures: the Bateman gradient β_{ss} (12) and the maximum intensity of precopulatory sexual selection, $s'_{\max}$ (hereafter the "Jones index" after ref. 13) (Table 1). The (standardized) Bateman gradient is the slope of the linear regression of relative reproductive success on relative mating success. It measures the extent to which individuals that mate more also tend to have more offspring. The Bateman gradient has often been used as a standalone proxy for sexual selection, but the information it provides is incomplete because it ignores the role of variance in mating success. The Jones index corrects for this by incorporating both the Bateman gradient and the opportunity for sexual selection. It measures both how much mating success varies and how this translates into variance in reproductive success. The Jones index represents the maximum strength of selection on a trait due to the trait's effect on mating success (i.e., if m' is the mating differential on any trait, then $|m'\beta_{ss}| \leq |s'_{\max}|$; ref. 13 and Table 1).

The Comparative Approach Using Mathematical Models

Despite the above theoretical considerations, we have little idea how well the various proxy measures perform in tracking the strength of sexual selection in natural systems. This is particularly problematic because most variance-based measures are known to perform poorly in theoretical worst-case scenarios (*Discussion*) (1, 14). One tempting approach is to estimate the correlation between proxy measures of sexual selection and observed selection differentials on putative target traits of sexual selection. This approach has indeed been followed qualitatively with populations of the same or closely related species (3, 26–30), but it suffers from two main drawbacks. First, as noted above, researchers may not accurately identify which traits are under sexual selection, and any systematic bias in this regard will affect the conclusions (2). Second, if sexual selection acts on multiple traits simultaneously, then the total strength of sexual selection may be only weakly related to the selection differential on any one trait. Third, and perhaps more importantly, relatively few studies have measured all three requisite

variables—mating success, reproductive success, and a putative target trait—in the same system. This limits our power to compare the effectiveness of proxy measures rigorously.

Here we import this comparative approach into a theoretical setting. We constructed mathematical models of mating and reproduction, incorporating individual variation in mate choice and competition, fecundity, direct benefits, and gamete competition (*Methods*). We tailored the models to the reproductive life histories of five well-studied species and parameterized them using data from the literature (details in *Supporting Information* and Table S1). We chose species that are diverse with respect to their sex roles, mating patterns, and the apparent strength of sexual selection in each sex (Fig. 1) (28, 31–59). For each species, we generated 100 new mating systems by modestly perturbing the original parameters at random (details in *Supporting Information* and Table S2). The resulting groups of 100 mating systems can readily be interpreted as populations of the same species or as sister taxa within clades. This procedure allowed us to increase the diversity of mating systems under consideration while remaining in the realm of the biologically plausible.

We simulated reproduction in each of the 500 mating systems and recorded the mating success, reproductive success, and trait values of each individual in the simulated population. We then calculated all commonly used proxy measures for the strength of sexual selection (Table 1) and compared them to the actual strength of sexual selection in each sex, measured by the linear selection differential on a simulated mating trait. Proxy measures that are more highly correlated with the strength of sexual selection are more reliable predictors. We also assessed the ability of proxy measures to predict differences in sexual selection between males and females by calculating the correlation between sex differences in proxy measures and sex differences in the strength of sexual selection. To mimic variation in empirical studies we analyzed both short-term and long-term datasets, corresponding to approximately one reproductive bout or to several reproductive bouts per individual, respectively (see *Methods* for details).

Our main analyses assume that an individual's expected success in mate competition increases monotonically with its mating trait value. The linear selection differential is an appropriate measure of the strength of selection in this case. However, we also consider the possibility that mating success peaks at an intermediate trait value, which results in stabilizing selection. In this case, the linear selection differential no longer captures the full strength of selection on the mating trait. Consequently, we instead assessed proxy measures against the "distributional selection differential" d', a measure that integrates both linear and nonlinear effects of selection on a trait distribution (Table 1; details in *Supporting Information*). Equivalent metrics have been used to quantify differences in trait distributions in other contexts (60, 61).

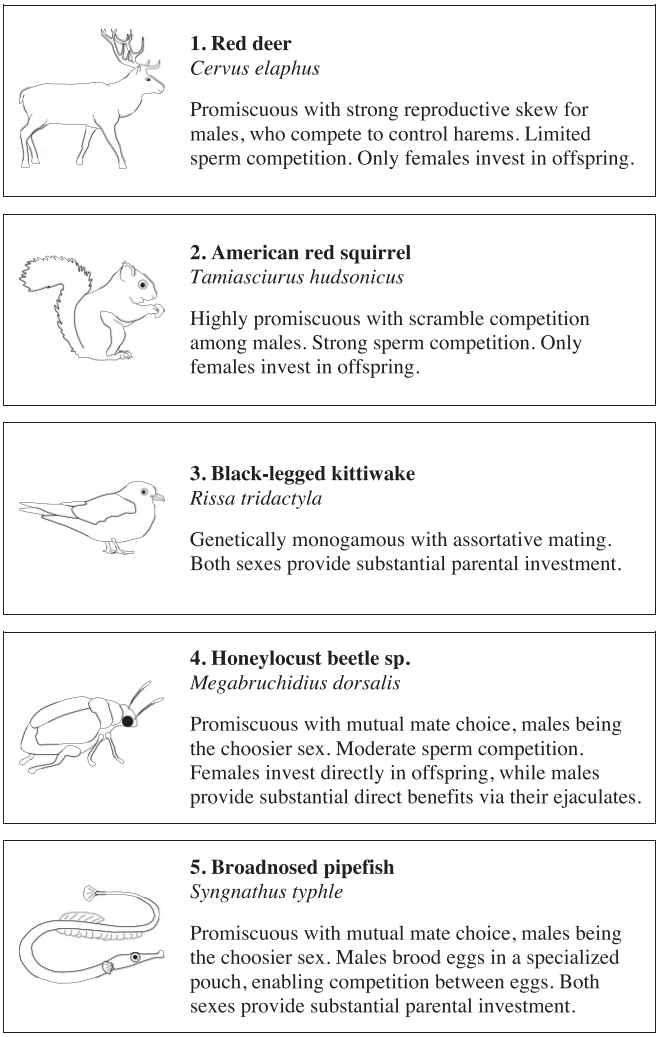

Fig. 1. The five mating systems that form the basis for our models: red deer (31–36), American red squirrel (37–42), black-legged kittiwake (43–46), a honeylocust beetle (28, 47–50), and broadnosed pipefish (51–59) (see *Supporting Information* for details).

Results

Our results support previous work showing that the relationship between variance-based proxy measures and the intensity of sexual selection depends on details of the mating system (Table 2 and Figs. 2 and 3). The predictive power of all measures differed significantly among the five mating system types for both sexes (all tests: $\chi^2_4 > 12.9$, $P < 0.02$ for both short- and long-term datasets), with the exception of the Bateman gradient, which showed uniformly low correlations for both sexes (Table S3). Moreover, again with the exception of the Bateman gradient, all measures predicted the actual strength of sexual selection significantly more accurately for males than for females (all tests: $|Z| > 4.8$, $P < 10^{-5}$ for both short- and long-term datasets) (Table 2 and Fig. 2). Unsurprisingly, predictive power was generally higher for long-term than for short-term datasets, as the former provide more precise estimates of both proxy measures and selection differentials (Table 2) (62).

Despite substantial variation in the predictive power of each individual measure, some measures did consistently better than others (Table 2 and Figs. 2 and 3). The Jones index $s'_{\max}$ was the best performer overall, with the highest power to predict actual sexual selection across all six simulated datasets (i.e., for males, females, and sex differences in both short- and long-term datasets). For sex differences in sexual selection, the Jones index was especially highly correlated with the strength of sexual selection ($r^2 = 0.85$ for short-term and $r^2 = 0.96$ for long-term datasets) and significantly outperformed all other measures (Williams' test: $|T_2| > 4.8$, $P < 10^{-5}$ for all pairwise comparisons; Fig. 3). Two rarely used measures also did reasonably well across all datasets: the Morisita index $I_{\delta-\mathrm{RS}}$ and the monopolization index Q_{RS} for reproductive success. In contrast, the Bateman gradient β_{ss} and the opportunity for sexual selection I_s performed poorly, despite being the most commonly used measures in the literature (1).

Our main analyses assume that the strength of sexual selection is represented by the linear selection differential s' on the mating trait. Our results are qualitatively unchanged if we instead use the linear selection gradient β' with the other traits as covariates (i.e., fecundity, direct benefits, and/or gamete competition ability; *Methods*) or the strength of premating sexual selection $m'\beta_{ss}$ (proposed in ref. 13). When selection has a strong stabilizing component, the linear selection differential is not an appropriate yardstick as it does not capture the effects of nonlinear

Table 2. Performance of proxy measures in predicting the linear selection differential s' on the mating trait (i.e., the strength of premating sexual selection)

Proxy measure	Short-term Male	Short-term Female	Short-term Sex difference	Long-term Male	Long-term Female	Long-term Sex difference
Jones index						
s'_{max}	0.80^{ab}	0.39^{e}	0.85	0.95	0.75^{q}	0.96
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.76^{cd}	0.33^{f}	0.79^{g}	0.89^{mn}	0.71	0.87
Q_{RS}	0.77^{cd}	0.38^{e}	0.79^{g}	0.84^{p}	0.73^{q}	0.82^{s}
I	0.63	0.12	0.35	0.81	0.07	0.83^{s}
Measures based on variance in mating success						
$I_{\delta-MS}$	0.77^{ac}	0.00	0.65^{hk}	0.90^{m}	0.05	0.79
Q_{MS}	0.80^{bd}	0.00	0.65^{h}	0.88^{n}	0.04	0.76
I_s	0.70	0.08	0.62^{k}	0.85^{p}	0.41^{r}	0.81^{s}
Bateman gradient						
β_{ss}	0.31	0.34^{ef}	0.62^{hk}	0.24	0.40^{r}	0.64

Shown is the coefficient of determination r^2 between the proxy measure and the linear selection differential (for the "sex difference" columns, we calculated r^2 based on the sex differences in proxy measures and selection differentials). Short-term data represent approximately one reproductive bout, whereas long-term data represent several reproductive bouts. Values within a column are marked with the same letter if the unsquared coefficients r do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.

selection. In this case, our results were qualitatively similar when we assessed proxy measures against the distributional selection differential d' on the mating trait (*Supporting Information*, Fig. S1, and Tables S4 and S5).

Because some relationships appear nonlinear (Figs. 2 and 3), we also performed our analyses using Spearman's rank correlation coefficient ρ in place of Pearson's correlation coefficient r. In this case, differences in performance between the Jones, Morisita, and monopolization indexes were less marked, although these three measures still outperformed all others (Tables S6–S8).

Discussion

Our model predicts strong and consistent differences in the effectiveness of various proxy measures in tracking the actual strength of premating sexual selection. The recently defined Jones index s'_{max} (13) outperformed all other measures and was especially highly predictive of sex differences in the strength of sexual selection. In contrast, the most frequently used proxy measures—the opportunity for sexual selection and the Bateman gradient—fared poorly. We therefore recommend that researchers use the Jones index in preference to these older measures, while keeping in mind the weaknesses of all variance-based measures of sexual selection (see below). We believe that the theoretical advance represented by the Jones index is currently underappreciated and that this index deserves more widespread use (see, e.g., refs. 59 and 63–66 for recent examples).

The Jones index requires data on the mating and reproductive success of each individual in a population (i.e., the same data that are used to calculate the Bateman gradient). Incomplete sampling of offspring or adults may lead to biased estimates, which can, however, be corrected for statistically (67, 68). When only reproductive success is known, our model suggests that the Morisita index $I_{\delta-RS}$ or the monopolization index Q_{RS} for reproductive success could provide passable substitutes. These measures quantify how unevenly reproductive success is distributed among individuals of the same sex, corrected for by the expected (for $I_{\delta-RS}$) or maximum (for Q_{RS}) variance if all individuals have equal propensities for reproductive success (17, 20, 24). However, we stress that this conclusion depends on the design and parameterization of our model. If non-sexually selected fitness components vary more than our model allows, then the Morisita and monopolization indexes will perform worse. This could occur, for example, in taxa with large intraspecific variance in female body size and hence fecundity (e.g., ref. 69).

Mate Quality and "Mating Success": The Problem with Kittiwakes. We follow most work on sexual selection by defining an individual's "mating success" as its number of matings, while ignoring variation in the quality of mates (13). However, selection for mate quality may act on similar traits and via similar mechanisms to selection for mate number (70, 71). The consequences of this are visible in our model results. For instance, the correlation between the Jones index and the actual strength of sexual selection is significantly weaker for the "kittiwake" mating systems than for those based on other species (Figs. 2 and 3 and Table S3). This reflects strong selection on mate quality in kittiwakes, which is not captured by the Jones index. A more precise definition of mating success would take the quality of mating partners into account (72, 73). There are considerable conceptual difficulties in devising such a definition, however, because the "true" (i.e., expected) quality of a mate is difficult to separate from stochastic processes that influence its realized reproductive success.

What Is a Bateman Gradient? Experimental vs. Observational Definitions. The poor performance of the Bateman gradient may surprise many researchers, given this measure's continued popularity. Our measure of the Bateman gradient resembles that used in observational studies, in which mating and reproductive success covary without explicit manipulation of specific individuals' mating opportunities. In contrast, some studies experimentally manipulate mating success (e.g., via sex ratios) and then measure the slope of the resultant change in reproductive success (e.g., ref. 74).

These two approaches lead to "observational" and "experimental" Bateman gradients that should not be lumped together. Experimental Bateman gradients are better suited to reveal causal relationships between mating and reproductive success (75). They should not, however, be used to calculate the Jones index, which is defined in terms of the observational Bateman gradient. Only under this definition does the Jones index represent the maximum strength of premating sexual selection (13). As proxies for the strength of sexual selection, both types of Bateman gradient are limited by their failure to account for variance in mating success, which places a fundamental limit on selection to increase the number of mates (*Introduction*).

The Best or the Luckiest? Limitations of Variance-Based Measures of Sexual Selection. The effectiveness of all variance-based proxy measures is sensitive to the role of stochasticity in determining mating and reproductive success (1, 14). For instance, although the opportunity for sexual selection provides a theoretical upper bound on the mating differential for any trait (i.e., $|m'| \leq I_s$; Table 1), actual mating differentials may fall far below this maximum, even for traits that are primary targets of sexual selection (1, 14, 76). This is because variance in mating success stems both from trait variation and from environmental effects that are independent of individual phenotype. If the relative contributions of phenotype and the environment differ between the sexes or across taxa, then I_s will be poorly correlated with mating differentials on sexually selected traits. Although subtle, this point is important: In worst-case scenarios, it can lead to almost any imaginable relationship between I_s and mating differentials, rather than a monotonically increasing relationship (1).

The Jones index inherits this problem, because it contains the square root of I_s as a factor (Table 1). The actual strength of premating sexual selection may thus be substantially lower than the Jones index (13). Consequently, the reliability of the Jones index is sensitive to variation in the proportion of mating success

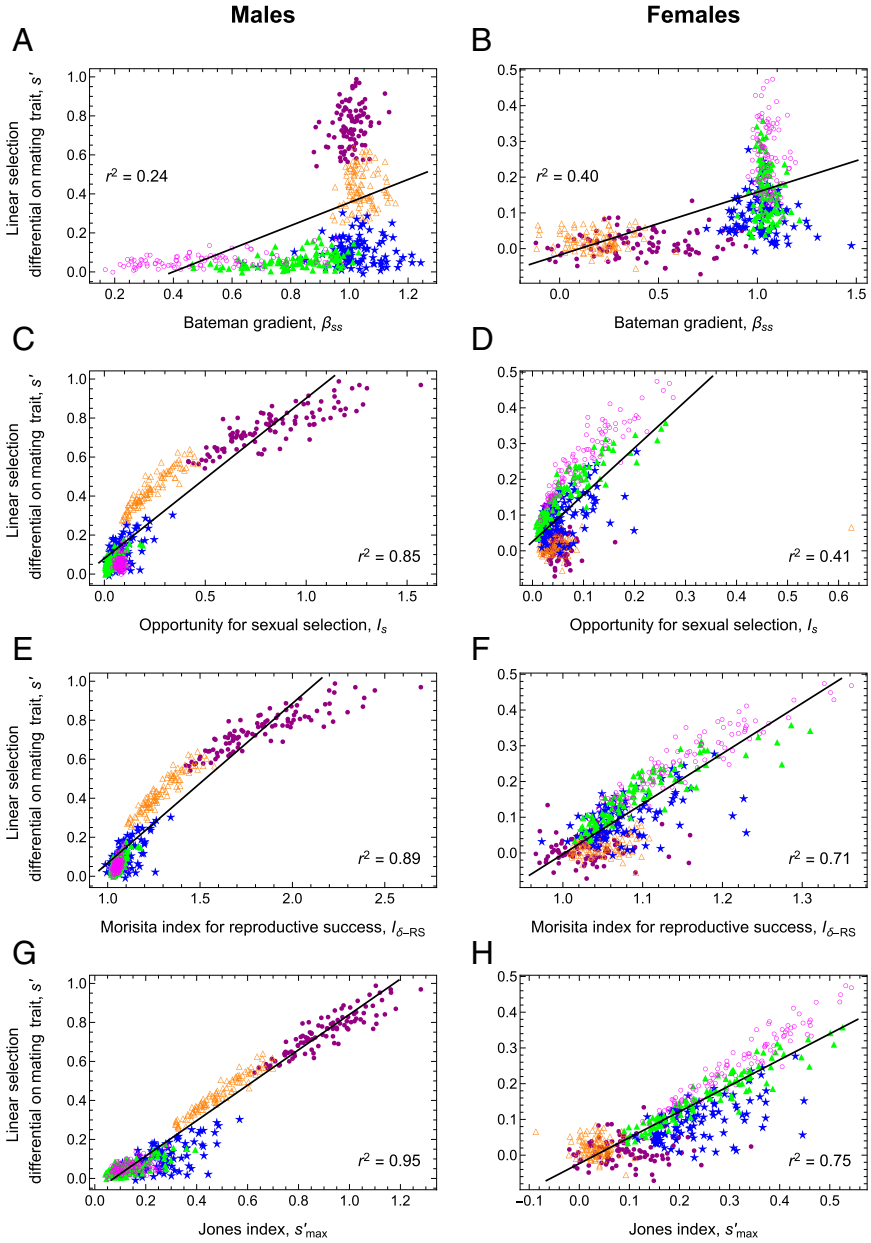

Fig. 2. Performance of selected proxy measures in predicting the linear selection differential s' on the mating trait (i.e., the strength of premating sexual selection). Shown is the regression of the linear selection differential against (*A* and *B*) the Bateman gradient, β_{ss}; (*C* and *D*) the opportunity for sexual selection, I_s; (*E* and *F*) the Morisita index for reproductive success, $I_{\delta-RS}$; and (*G* and *H*) the Jones index, s'_{max}, for males (*A*, *C*, *E*, and *G*) and for females (*B*, *D*, *F*, and *H*). Data points are based on red deer (purple solid circles), American red squirrels (orange open triangles), black-legged kittiwakes (blue stars), honeylocust beetles (green solid triangles), and broadnosed pipefish (pink open circles), using long-term data representing multiple reproductive bouts.

that is due to phenotypic traits, rather than environmental stochasticity. This is evidenced in our model results by higher predictive power for males than for females (Table 2 and Fig. 2). This indicates that the relative strength of environmental effects on mating success varies more across mating systems for females than for males in our datasets. The superiority of the Jones index relative to other indexes was robust across our simulated datasets, however, despite substantial variation in the level of stochasticity (Table 2 and Figs. 2 and 3).

How much the contribution of environmental variance differs between the sexes, or across taxa, is an empirical question, albeit a difficult one. When multiple traits are measured, it is possible to calculate their total contribution toward the opportunity for sexual selection (77). This can provide useful lower bounds on the proportion of I_s that is due to trait variation, even though the contribution of unmeasured traits will of course remain unknown. In practice, however, most empirical studies measure very few sexually selected traits and so this approach is of limited value with currently available data. Our simulations assume for simplicity that mating probabilities are determined by a single trait. However, because the Jones index is based on the total variance in mating success, it will capture the contributions of any number of traits, and so we expect our qualitative results to generalize to this case.

In summary, our results indicate that the Jones index is an ideal choice for comparative analyses of sexual selection and is preferable to other variance-based measures. The Jones index provides a theoretical upper bound for selection on mating success and our results suggest that it is highly correlated with

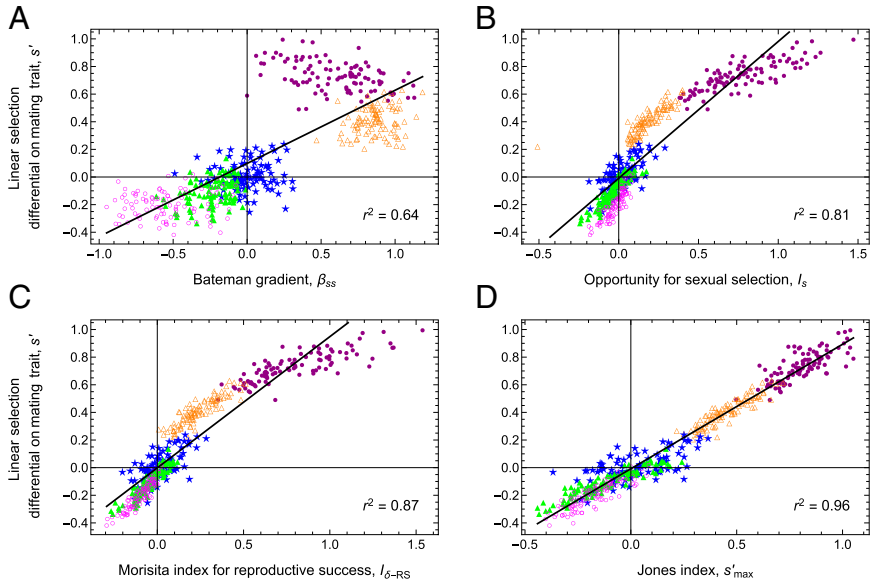

Fig. 3. Performance of selected proxy measures in predicting sex differences in the linear selection differential s' (i.e., the strength of premating sexual selection). Shown is the regression of sex differences in the linear selection differential against sex differences in (*A*) the Bateman gradient, β_{ss}; (*B*) the opportunity for sexual selection, I_s; (*C*) the Morisita index for reproductive success, $I_{\delta-RS}$; and (*D*) the Jones index, s'_{max}. Data points are based on red deer (purple solid circles), American red squirrels (orange open triangles), black-legged kittiwakes (blue stars), honeylocust beetles (green solid triangles), and broadnosed pipefish (pink open circles), using long-term data representing multiple reproductive bouts.

selection differentials on mating traits under a wide range of biologically realistic conditions. We remind readers nonetheless that the Jones index is not a direct measure of the strength of sexual selection. Like all indexes, it should be interpreted with caution and an eye to factors that might weaken its relationship with sexual selection (1, 14). For studies within species, we recommend that the Jones index be combined with trait-based measures where possible to give a fuller picture of sexual selection (1, 2, 28).

Methods

Model Structure. We constructed a general model of reproduction that we tailored to better represent the reproductive life histories of our five species (Fig. 1). At any point in time, each individual is either in the "mating pool" or in "time out." The mating pool is subdivided into "unmated individuals" (those that have not mated since they were last in time out) and "mated individuals" (those that have mated at least once). Our simulations consist of a series of events that can occur at any point in continuous time. There are three types of events:

i) Mate encounter: A male and female from the mating pool encounter each other and potentially mate (depending on their mating-trait values and the degree of mate choice or competition in each sex, see below). If mating occurs, and if either individual was previously "unmated," then that individual is now classified as "mated." It will continue to search for mates for an additional fixed period of S or $\tilde{S}$ (for females or males, respectively), after which it is scheduled to enter time out.

ii) Enter time out: A mated individual enters time out. The individual is then scheduled to return to the mating pool as an unmated individual after a fixed period of T or $\tilde{T}$ (for females or males, respectively).

iii) Return to the mating pool: An individual in time out returns to the mating pool. If the individual is male (for pipefish) or female (for all other species), then offspring are produced and parentage of each offspring is determined according to species-specific patterns (see below).

We determine which event will occur next by first generating the time until the next mate encounter as a random variable $M \sim \mathrm{Exp}(\lambda P_m \tilde{P}_m)$, where P_m and $\tilde{P}_m$ are the numbers of females and males, respectively, in the mating pool and λ is the population-wide rate of mate encounters. We then compare $t+M$ (where t is the current time) to the next time that an individual is scheduled to enter time out or return to the mating pool. If $t+M$ is smaller, then a mate encounter occurs. Otherwise, the next scheduled event occurs. Note that some of the time parameters S, $\tilde{S}$, T, and $\tilde{T}$ may be set to zero.

Traits. Each individual is endowed with up to three quantitative traits, depending on its sex and the mating system. First, the mating trait (Z_c or $\tilde{Z}_c$ for females or males, respectively) determines an individual's ability to attract or compete for mates. Second, the "fecundity trait" Z_f in females and the "direct benefits trait" $\tilde{Z}_b$ in males determine the amount of resources an individual invests in offspring, whether directly as parental investment or indirectly via nuptial gifts. Third, in species where the gametes of one sex compete for fertilization or survival, the competitive ability of an individual's gametes is determined by its "sperm competition trait" $\tilde{Z}_s$ or "egg competition trait" Z_e. We assume that all trait values follow standard normal distributions but allow for sex-specific correlations between traits [using Cholesky decomposition (78); *Supporting Information*].

Mate Encounters. When a mate encounter occurs, we choose a male and a female at random from the mating pool. Not all such encounters lead to successful matings. Rather, the probability of mating depends on the mating-trait values of both individuals, as well as the population-wide strength of mate choice or competition in each sex. We model both mate choice and competition, using flexible functions that allow for a wide range of scenarios. We treat choice and competition using the same mathematical formalism and so the mating trait might be interpreted differently (e.g., as an ornament, an armament, or a searching trait), depending on the mating system.

If the female has mating trait Z_c and the male has mating trait $\tilde{Z}_c$, then the pair mates with probability

$$p = q(Z_c, \tilde{Z}_c) \cdot \tilde{q}(Z_c, \tilde{Z}_c).$$

The first term represents the probability that the female is accepted by, or successfully competes for, the male. It is given by the logistic function

$$q(Z_c, \tilde{Z}_c) = \frac{1}{1+\exp(\tilde{m}_c + \tilde{a}_c \cdot \tilde{Z}_c - Z_c)}.$$

Similarly, the second term represents the probability that the male is accepted by, or can compete for, the female. It is given by

$$\tilde{q}(Z_c, \tilde{Z}_c) = \frac{1}{1+\exp(m_c + a_c \cdot Z_c - \tilde{Z}_c)}.$$

According to these functions, a female requires a mating trait of $\tilde{m}_c$ or greater to be accepted by a median male with a probability of at least one-half. Higher values of $\tilde{m}_c$ represent stronger male choice or female–female

competition. The parameter $\tilde{a}_c$ determines how a male's quality affects his choosiness or the strength of female–female competition to mate with him. If $\tilde{a}_c = 0$, then it is equally difficult to mate with any male, whereas positive values of $\tilde{a}_c$ mean that higher-quality males are choosier or attract more competition. Analogously, the parameter m_c represents the strength of female choice or male–male competition, whereas a_c determines how these increase with individual female quality. These functions allow us to model scenarios such as strong competition in either sex (large m_c or $\tilde{m}_c$), random mating ($m_c = \tilde{m}_c = -\infty$), and assortative mating (m_c and $\tilde{m}_c$ finite and $a_c, \tilde{a}_c > 0$).

Offspring Production. Offspring production occurs when an individual of the relevant sex (males for pipefish and females for all other species) returns to the mating pool. The number of offspring produced depends on the fecundity traits of females and, for some species, on the direct benefits traits of males, according to species-specific patterns.

A female's expected fecundity is determined by her standardized fecundity trait Z_f according to the function

$$f(Z_f) = m_f \exp(a_f Z_f).$$

The parameter m_f is the median fecundity in the population, whereas a_f determines how quickly a female's fecundity increases with her trait value Z_f. Higher values of a_f mean that fecundity is more variable in a population. Similarly, the expected direct benefits provided by a male increase with his direct benefits trait $\tilde{Z}_b$ according to the function

$$b(\tilde{Z}_b) = \tilde{m}_b \exp(\tilde{a}_b \tilde{Z}_b).$$

Here $\tilde{m}_b$ is the median level of direct benefits and $\tilde{a}_b$ determines its variability in the population.

In red deer and squirrels, males do not contribute to offspring production. Consequently, we assume that the number of offspring produced is a Poisson-distributed random integer with mean $f(Z_f)$. For kittiwakes, we assume serial monogamy (46), and the number of offspring fledged by the combined care of both parents follows a Poisson distribution with mean $f(Z_f) + b(\tilde{Z}_b)$.

For honeylocust beetles, we assume that only a female's most recent mate contributes materially to her egg production, although previous mates can still fertilize her eggs (see below) (48). The number of eggs she produces is thus a Poisson-distributed random integer with mean $f(Z_f) + b(\tilde{Z}_b)$, where $\tilde{Z}_b$ is the direct benefits trait of her most recent mate.

For pipefish, the number of offspring produced is limited by both the male's care and the number of eggs provided to him by females (54). Suppose that a male mates with n females with fecundity traits $Z_{f,1}, \ldots, Z_{f,n}$. The number of eggs he receives from the ith female is a Poisson-distributed integer F_i with mean $f(Z_{f,i})$. The maximum number of offspring B that the male can incubate similarly follows a Poisson distribution with mean $b(\tilde{Z}_b)$. The number of offspring produced is then given by $\min(B, \sum_{i=1}^{n} F_i)$, which is the smaller of the male's caring capacity and the total number of eggs he receives.

Parentage of Offspring. In species with sperm competition (i.e., red deer, squirrels, and honeylocust beetles), the competitive ability of a male's sperm increases with his standardized sperm competition trait $\tilde{Z}_s$ according to the function

$$s(\tilde{Z}_s) = \exp(\tilde{a}_s \tilde{Z}_s).$$

The parameter $\tilde{a}_s$ determines how variable sperm competition ability is in the population.

For red deer and squirrels, paternity of each offspring is decided by a skewed raffle among all of the males that a female has mated with since her previous time out. Writing these males' sperm competition traits as $\tilde{Z}_{s,1}, \ldots, \tilde{Z}_{s,m}$, the probability that the jth male fertilizes any particular offspring is given by

$$\frac{s(\tilde{Z}_{s,j})}{\sum_{k=1}^{m} s(\tilde{Z}_{s,k})}.$$

Honeylocust beetles can store sperm for long periods, and so a female's eggs may be fertilized by any of her previous mates, even those from before her previous time out. We assume, however, that the viability of sperm falls over time, and so more recent mates are more likely to father her offspring (48). Suppose that a female's recent mates had sperm competition trait values $\tilde{Z}_{s,1}, \ldots, \tilde{Z}_{s,m}$ and that the matings occurred at times $t_1, \ldots, t_m$, respectively (for simplicity, we assume that only the 10 most recent mates can gain a share of paternity, so that $m \leq 10$). If sperm competitiveness falls exponentially over time with rate constant r, then the ith male's sperm are discounted by a factor of

$$r_i = \exp(-r(t - t_i)).$$

The probability that the jth male fathers any particular offspring is then

$$\frac{r_j \cdot s(\tilde{Z}_{s,j})}{\sum_{k=1}^{m} r_k \cdot s(\tilde{Z}_{s,k})}.$$

In pipefish, the survival rate of eggs inside the male's brood pouch depends on maternal identity, which can be thought of as a type of egg competition (57). The competitive ability of a female's eggs increases with her standardized egg competition trait Z_e according to the function

$$e(Z_e) = \exp(a_e Z_e).$$

As above, the parameter a_e determines how variable egg competition ability is in the population. Suppose a male mates with n females with egg competition trait values $Z_{e,1}, \ldots, Z_{e,n}$. The reproductive success of each female is determined by sampling without replacement from the total pool of eggs until either all eggs have been sampled or the male's caring capacity B is reached (see above). For each sample, the probability of choosing any particular egg of the ith female is weighted by $e(Z_{e,i})$.

Running the Model. We parameterized the model for each species based on data in the literature. We then generated 500 new mating systems with a "family resemblance" to the template species by modestly perturbing the original model parameters at random (see *Supporting Information* for details of parameter choice and randomization).

For each simulation, we ran the model until the current time was $t \geq 100$. A typical individual of the offspring-producing sex (i.e., males in pipefish, females in all other species) completed a reproductive bout (i.e., one cycle of mating and time out) in ~1.5 time units, although this depended on the randomly chosen parameter values (see *Supporting Information* for details). We consequently analyzed the last 1.5 time units (short-term data) or 4.5 time units (long-term data), which correspond to roughly one reproductive bout or several reproductive bouts per individual, respectively.

Mating and Reproductive Success. We kept track of mating and reproductive success in each simulation via the full parental tables (12). For consistency, we counted both mating and reproductive success when an individual of the offspring-producing sex (i.e., males for pipefish or females for the other species) returned to the mating pool. This means that matings were counted only if there was a subsequent opportunity to produce offspring, thus ensuring comparability with field studies that are conducted over fixed breeding seasons. We calculated the number and parentage of offspring as above and then attributed one unit of reproductive success to the parents of each offspring. We then determined all of the offspring-producing individual's mates since its last time out. For each mating, we attributed one unit of mating success to both partners. Multiple matings between the same individuals were counted separately.

Assessment of Proxy Measures. We calculated each proxy measure separately for males and females over the two time spans (Table 1). We also calculated the strength of sexual selection, defined as the linear selection differential on the mating trait. For both proxy measures and selection differentials, we calculated the sex difference in sexual selection by subtracting the female value from the male value. This resulted in six simulated datasets (i.e., males, females, and sex differences for short-term and long-term data). All datasets are available from the Dryad Digital Repository at dx.doi.org/10.5061/dryad.3p2j0.

For each dataset, we calculated the predictive power of proxy measures as their Pearson's correlation coefficients r with the actual strength of sexual selection. Some proxy measures appeared to show nonlinear relationships with the strength of sexual selection, and so we also repeated our analyses, using Spearman's rank correlation coefficient ρ (Table S6). Despite the presence of nonlinearity, and even apparent nonmonotonicity, in some of our datasets, we believe that Pearson's and Spearman's correlation coefficients are informative. The aim of our analyses is not to measure dependency between variables per se, but rather to assess the effectiveness of proxy measures as practical indexes of sexual selection. Putative indexes that display nonmonotonic relationships with the strength of sexual selection are

not useful for empirical applications, especially if the precise shape of the relationship is sensitive to other variables. Indeed, we believe that a linear relationship is highly desirable. Consequently, we use correlation coefficients that penalize nonlinearity (Pearson's) and nonmonotonicity (Spearman's).

We tested for differences in the predictive power of a single measure among the five template mating systems, using the χ^2 test for the homogeneity of independent correlation coefficients (79). To compare the predictive power of a single measure between males and females, we used the Z test for differences between two independent correlation coefficients (79). We compared the predictive power of proxy measures on the same dataset by pairwise application of Williams' test for differences between two dependent correlation coefficients that share a common variable (80, 81).

ACKNOWLEDGMENTS. We thank Michael Jennions for valuable discussions and feedback. We also thank Hanna Kokko and Göran Arnqvist for their input and support. Two anonymous reviewers and an editor provided feedback that greatly improved the manuscript. Funding was provided by the EBC Graduate School on Genomes and Phenotypes, Stiftelsen för Zoologisk Forskning, the Gwendolyn Woodroofe PhD Scholarship in the Biological Sciences, and an Australian Postgraduate Award.

1. Klug H, Heuschele J, Jennions MD, Kokko H (2010) The mismeasurement of sexual selection. *J Evol Biol* 23(3):447–462.
2. Krakauer AH, Webster MS, Duval EH, Jones AG, Shuster SM (2011) The opportunity for sexual selection: Not mismeasured, just misunderstood. *J Evol Biol* 24(9):2064–2071.
3. Fitze PS, Le Galliard J-F (2011) Inconsistency between different measures of sexual selection. *Am Nat* 178(2):256–268.
4. Doherty PF, Jr, et al. (2003) Sexual selection affects local extinction and turnover in bird communities. *Proc Natl Acad Sci USA* 100(10):5858–5862.
5. Seddon N, Merrill RM, Tobias JA (2008) Sexually selected traits predict patterns of species richness in a diverse clade of suboscine birds. *Am Nat* 171(5):620–631.
6. Kraaijeveld K, Kraaijeveld-Smit FJL, Maan ME (2011) Sexual selection and speciation: The comparative evidence revisited. *Biol Rev Camb Philos Soc* 86(2):367–377.
7. Wagner CE, Harmon LJ, Seehausen O (2012) Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* 487(7407):366–369.
8. Gavrilets S (2014) Is sexual conflict an "engine of speciation"? *Cold Spring Harb Perspect Biol* 6(12):a017723.
9. Trivers RL (1972) Parental investment and sexual selection. *Sexual Selection and the Descent of Man 1871-1971*, ed Campbell B (Aldine, Chicago), pp 136–179.
10. Shuster SM, Wade MJ (2003) *Mating Systems and Strategies* (Princeton Univ Press, Princeton).
11. Wade MJ, Arnold SJ (1980) The intensity of sexual selection in relation to male sexual behaviour, female choice, and sperm precedence. *Anim Behav* 28(2):446–461.
12. Arnold SJ, Duvall D (1994) Animal mating systems: A synthesis based on selection theory. *Am Nat* 143(2):317–348.
13. Jones AG (2009) On the opportunity for sexual selection, the Bateman gradient and the maximum intensity of sexual selection. *Evolution* 63(7):1673–1684.
14. Jennions MD, Kokko H, Klug H (2012) The opportunity to be misled in studies of sexual selection. *J Evol Biol* 25(3):591–598.
15. Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. *Evolution* 37(6):1210–1226.
16. Parker GA, Birkhead TR (2013) Polyandry: The history of a revolution. *Philos Trans R Soc Lond B Biol Sci* 368(1613):20120335.
17. Kokko H, Mackenzie A, Reynolds JD, Lindström J, Sutherland WJ (1999) Measures of inequality are not equal. *Am Nat* 154(3):358–382.
18. Crow JF (1958) Some possibilities for measuring selection intensities in man. *Hum Biol* 30(1):1–13.
19. Sutherland WJ (1985) Chance can produce a sex difference in variance in mating success and explain Bateman's data. *Anim Behav* 33:1349–1352.
20. Ruzzante DE, Hamilton DC, Kramer DL, Grant JWA (1996) Scaling of the variance and the quantification of resource monopolization. *Behav Ecol* 7(2):199–207.
21. Baena ML, Macías-Ordóñez R (2012) Phenology of scramble polygyny in a wild population of chrysomelid beetles: The opportunity for and the strength of sexual selection [corrected]. *PLoS One* 7(6):e38315.
22. Moura RR, Peixoto PEC (2013) The effect of operational sex ratio on the opportunity for sexual selection: A meta-analysis. *Anim Behav* 86(4):675–683.
23. Morisita M (1962) I-Index, a measure of dispersion of individuals. *Res Popul Ecol* 4(1): 1–7.
24. Fairbairn DJ, Wilby AE (2001) Inequality of opportunity: Measuring the potential for sexual selection. *Evol Ecol Res* 3(6):667–686.
25. Lorch PD (2005) Using upper limits of "Bateman gradients" to estimate the opportunity for sexual selection. *Integr Comp Biol* 45(5):924–930.
26. Mills SC, Grapputo A, Koskela E, Mappes T (2007) Quantitative measure of sexual selection with respect to the operational sex ratio: A comparison of selection indices. *Proc Biol Sci* 274(1606):143–150.
27. Croshaw DA (2010) Quantifying sexual selection: A comparison of competing indices with mating system data from a terrestrially breeding salamander. *Biol J Linn Soc* 99(1):73–83.
28. Fritzsche K, Arnqvist G (2013) Homage to Bateman: Sex roles predict sex differences in sexual selection. *Evolution* 67(7):1926–1936.
29. Martin AM, Festa-Bianchet M, Coltman DW, Pelletier F (2015) Comparing measures of breeding inequality and opportunity for selection with sexual selection on a quantitative character in bighorn rams. *J Evol Biol* 28(1):223–230.
30. Janicke T, David P, Chapuis E (2015) Environment-dependent sexual selection: Bateman's parameters under varying levels of food availability. *Am Nat* 185(6):756–768.
31. Clutton-Brock TH, Guinness FE, Albon SD (1982) *Red Deer: Behavior and Ecology of Two Sexes* (Univ of Chicago Press, Chicago).
32. Pemberton JM, Albon SD, Guinness FE, Clutton-Brock TH, Dover GA (1992) Behavioral estimates of male mating success tested by DNA fingerprinting in a polygynous mammal. *Behav Ecol* 3(1):66–75.
33. Clutton-Brock TH, Lonergan ME (1994) Culling regimes and sex ratio biases in highland red deer. *J Appl Ecol* 31(3):521–527.
34. Clutton-Brock TH, Rose KE, Guinness FE (1997) Density-related changes in sexual selection in red deer. *Proc Biol Sci* 264(1387):1509–1516.
35. Bonenfant C, Gaillard J-M, Klein F, Maillard D (2004) Variation in harem size of red deer (*Cervus elaphus* L.): The effects of adult sex ratio and age-structure. *J Zool* 264(1): 77–85.
36. Malo AF, Roldan ERS, Garde J, Soler AJ, Gomendio M (2005) Antlers honestly advertise sperm production and quality. *Proc Biol Sci* 272(1559):149–157.
37. Rusch DA, Reeder WG (1978) Population ecology of Alberta red squirrels. *Ecology* 59(2):400–420.
38. Hurly TA (1987) Male-biased sex ratios in a red squirrel population. *Can J Zool* 65(5): 1284–1286.
39. Vernes K (2004) Breeding biology and seasonal capture success of northern flying squirrels (*Glaucomys sabrinus*) and red squirrels (*Tamiasciurus hudsonicus*) in Southern New Brunswick. *Northeast Nat* 11(2):123–136.
40. McAdam AG, Boutin S, Sykes AK, Humphries MM (2007) Life histories of female red squirrels and their contributions to population growth and lifetime fitness. *Ecoscience* 14(3):362–369.
41. Lane JE, Boutin S, Gunn MR, Slate J, Coltman DW (2008) Female multiple mating and paternity in free-ranging North American red squirrels. *Anim Behav* 75(6):1927–1937.
42. Lane JE, Boutin S, Gunn MR, Coltman DW (2009) Sexually selected behaviour: Red squirrel males search for reproductive success. *J Anim Ecol* 78(2):296–304.
43. Coulson JC, Wooller RD (1984) Incubation under natural conditions in the kittiwake gull, *Rissa tridactyla*. *Anim Behav* 32(4):1204–1215.
44. Hatch SA, Byrd GV, Irons DB, Hunt GLJ (1993) Status and ecology of kittiwakes (*Rissa tridactyla* and *R. brevirostris*) in the North Pacific. *The Status, Ecology, and Conservation of Marine Birds of the North Pacific*, eds Vermeer K, Briggs KT, Morgan KH, Siegel-Causey D (Canadian Wildlife Service, Ottawa, Ontario), pp 140–153.
45. Helfenstein F, Danchin E, Wagner RH (2004) Assortative mating and sexual size dimorphism in black-legged kittiwakes. *Waterbirds* 27(3):350–354.
46. Wagner RH, Helfenstein F, Danchin E (2004) Female choice of young sperm in a genetically monogamous bird. *Proc Biol Sci* 271(Suppl 4):S134–S137.
47. Takakura K-I (1999) Active female courtship behavior and male nutritional contribution to female fecundity in *Bruchidius dorsalis* (Fahraeus) (Coleoptera: Bruchidae). *Res Popul Ecol* 41(3):269–273.
48. Takakura K-I (2001) Courtship-role-reversal in the bean weevil, *Bruchidius dorsalis* (Coleoptera: Bruchidae): Interplay between male-male competition and cryptic female choice. *Appl Entomol Zool* 36(3):311–316.
49. Takakura K-I (2006) Estimation of relative reproductive expenditure in the courtship-role-reversed bean weevil, *Bruchidius dorsalis* (Fahraeus). *J Ethol* 24(1):33–36.
50. Salehialavi Y, Fritzsche K, Arnqvist G (2011) The cost of mating and mutual mate choice in 2 role-reversed honey locust beetles. *Behav Ecol* 22(5):1104–1113.
51. Berglund A, Rosenqvist G, Svensson I (1986) Reversed sex roles and parental energy investment in zygotes of two pipefish (Syngnathidae) species. *Mar Ecol Prog Ser* 29: 209–215.
52. Berglund A, Rosenqvist G, Svensson I (1986) Mate choice, fecundity and sexual dimorphism in two pipefish species (Syngnathidae). *Behav Ecol Sociobiol* 19(4):301–307.
53. Berglund A, Rosenqvist G, Svensson I (1989) Reproductive success of females limited by males in two pipefish species. *Am Nat* 133(4):506–516.
54. Berglund A, Rosenqvist G (1990) Male limitation of female reproductive success in a pipefish: Effects of body-size differences. *Behav Ecol Sociobiol* 27(2):129–133.
55. Ahnesjö I (1992) Consequences of male brood care: Weight and number of newborn in a sex-role reversed pipefish. *Funct Ecol* 6(3):274–281.
56. Berglund A, Rosenqvist G (1993) Selective males and ardent females in pipefishes. *Behav Ecol Sociobiol* 32(5):331–336.
57. Ahnesjö I (1996) Apparent resource competition among embryos in the brood pouch of a male pipefish. *Behav Ecol Sociobiol* 38(3):167–172.
58. Partridge C, et al. (2009) The effect of perceived female parasite load on post-copulatory male choice in a sex-role-reversed pipefish. *Behav Ecol Sociobiol* 63(3):345–354.
59. Mobley KB, Abou Chakra M, Jones AG (2014) No evidence for size-assortative mating in the wild despite mutual mate choice in sex-role-reversed pipefishes. *Ecol Evol* 4(1): 67–78.
60. Gregorius H-R, Gillet EM, Ziehe M (2003) Measuring differences of trait distributions between populations. *Biometrical J* 8:959–973.
61. Rudnicki R, Zwoleński P (2015) Model of phenotypic evolution in hermaphroditic populations. *J Math Biol* 70(6):1295–1321.
62. Sefc KM (2011) Mating and parental care in Lake Tanganyika's cichlids. *Int J Evol Biol* 2011:470875.
63. Barreto FS, Avise JC (2010) Quantitative measures of sexual selection reveal no evidence for sex-role reversal in a sea spider with prolonged paternal care. *Proc R Soc B* 277(1696):2951–2956.

64. Serbezov D, Bernatchez L, Olsen EM, Vøllestad LA (2010) Mating patterns and determinants of individual reproductive success in brown trout (*Salmo trutta*) revealed by parentage analysis of an entire stream living population. *Mol Ecol* 19(15): 3193–3205.
65. Pélissié B, Jarne P, David P (2012) Sexual selection without sexual dimorphism: Bateman gradients in a simultaneous hermaphrodite. *Evolution* 66(1):66–81.
66. Wacker S, Amundsen T, Forsgren E, Mobley KB (2014) Within-season variation in sexual selection in a fish with dynamic sex roles. *Mol Ecol* 23(14):3587–3599.
67. Mobley KB, Jones AG (2013) Overcoming statistical bias to estimate genetic mating systems in open populations: A comparison of Bateman's principles between the sexes in a sex-role-reversed pipefish. *Evolution* 67(3):646–660.
68. Jones AG (2015) BATEMANATER: A computer program to estimate and bootstrap mating system variables based on Bateman's principles. *Mol Ecol Resour* 15(6): 1396–1402.
69. Hines AH (1991) Fecundity and reproductive output in nine species of *Cancer* crabs (Crustacea, Brachyura, Cancridae). *Can J Fish Aquat Sci* 48(2):267–275.
70. Tobias JA, Montgomerie R, Lyon BE (2012) The evolution of female ornaments and weaponry: Social selection, sexual selection and ecological competition. *Philos Trans R Soc Lond B Biol Sci* 367(1600):2274–2293.
71. Cain KE (2014) Mates of competitive females: The relationships between female aggression, mate quality, and parental care. *Adv Zool* 2014:319567.
72. Webster MS, Pruett-Jones S, Westneat DF, Arnold SJ (1995) Measuring the effects of pairing success, extra-pair copulations and mate quality on the opportunity for sexual selection. *Evolution* 49(6):1147–1157.
73. Fitzpatrick CL (2015) Expanding sexual selection gradients; a synthetic refinement of sexual selection theory. *Ethology* 121(3):207–217.
74. Jones AG, Rosenqvist G, Berglund A, Arnold SJ, Avise JC (2000) The Bateman gradient and the cause of sexual selection in a sex-role-reversed pipefish. *Proc Biol Sci* 267(1444):677–680.
75. Collet JM, Dean RF, Worley K, Richardson DS, Pizzari T (2014) The measure and significance of Bateman's principles. *Proc R Soc B* 281(1782):2013.2973.
76. Lorch PD, Bussière LF, Gwynne DT (2008) Quantifying the potential for sexual dimorphism using upper limits on Bateman gradients. *Behaviour* 145(1):1–24.
77. Moorad JA, Wade MJ (2013) Selection gradients, the opportunity for selection, and the coefficient of determination. *Am Nat* 181(3):291–300.
78. Asmussen S, Glynn PW (2007) *Stochastic Simulation: Algorithms and Analysis* (Springer, New York).
79. Kenny DA (1987) *Statistics for the Social and Behavioral Sciences* (Little, Brown & Co, Boston).
80. Williams EJ (1959) The comparison of regression variables. *J R Stat Soc B* 21(2):396–399.
81. Steiger JH (1980) Tests for comparing elements of a correlation matrix. *Psychol Bull* 87(2):245–251.

Supporting Information

Henshaw et al. 10.1073/pnas.1518067113

Model Initialization

For each simulation run, we generated a population of $1{,}000 \cdot \mathrm{ASR}$ males and $1{,}000(1-\mathrm{ASR})$ females, each rounded to the nearest integer, where ASR is the adult sex ratio. To ensure that initial fluctuations in the mating pool quickly evened out, we began simulations with all individuals in time out and scheduled each individual to return to the mating pool at a random time chosen uniformly from [0,1].

We initially chose trait values for each individual according to a standard normal distribution and then transformed these values using Cholesky decomposition. The transformation was performed in such a way that female traits were correlated with coefficient ρ and male traits with coefficient $\tilde{\rho}$ (78).

Choice of Parameter Values

We chose parameters based on descriptions of the five species' mating systems in the literature (summarized in Table S1). For species with seasonal breeding, our parameterization is based only on behavior during the breeding season and not at other times. To ensure that reproductive bouts were of approximately equal length between species, we scaled the mate search and time-out parameters so that they added to one for the sex with the lower potential reproductive rate (i.e., either $T+S=1$ or $\tilde{T}+\tilde{S}=1$). Where possible, we based these parameters on explicit estimates of the proportion of time during the breeding season that each sex is available for mating. If these data were unavailable, we used data on the operational sex ratio. We set the mate encounter rate λ to ensure that the average number of mates per reproductive bout was realistic for the offspring-producing sex.

Red Deer (*Cervus elaphus*). Most studies found a moderately to strongly female-biased adult sex ratio, although this varied by population, culling regime, and the age cutoff used to define reproductively mature males (33–35). We chose $\mathrm{ASR}=0.4$ as our base parameter, representing 1.5 reproductively mature females to each reproductively mature male. Females are fertile for approximately 1 d during a breeding season of approximately 30 d (31), so we set $T=29/30$ and $S=1/30$. Males are reproductively active during the whole breeding season (i.e., $\tilde{T}=0$). The prevalence of polyandry in red deer is believed to be limited, although it is noteworthy that red deer males have large testes relative to other cervid species (31, 36). We consequently set the rate of mate encounters to $\lambda=0.1$, which meant that on average, given our other parameters, females mated with slightly more than one male per reproductive bout. Males compete fiercely for control of harems and generally only older, larger males are successful. We set $m_{\mathrm{c}}=2$, which means that a male must have a mating-trait value 2 SD above average to have a 0.5 probability of mating with a female he encounters. We assumed there is no male mate choice or female–female competition for mates ($\tilde{m}_{\mathrm{c}}=-\infty$). We assumed that females give birth to one calf every two breeding seasons on average ($m_{\mathrm{f}}=0.5$) and, because twinning is unusual, assumed a shallow fecundity slope of $a_{\mathrm{f}}=0.1$ (31). Males do not provide direct benefits.

American Red Squirrel (*Tamiasciurus hudsonicus*). Sex ratios in American red squirrel vary strongly among years and populations, but sex ratios during the breeding season are usually close to parity (37, 39) (but see ref. 38). We consequently assumed $\mathrm{ASR}=0.5$. Females enter estrus asynchronously, resulting in male-biased operational sex ratios of 25:1 during the breeding season (42). We consequently set $T=0.96$ and $S=0.04$ for females and assumed that males were always reproductively active during the breeding season (i.e., $\tilde{T}=0$). Females are highly polyandrous, mating with almost seven males during estrus (41). We set the mate encounter rate to $\lambda=0.5$, which meant that females mated with approximately six to seven males per reproductive bout. Given the scramble competition mating system, we assumed moderate male mate competition ($m_{\mathrm{c}}=0$) and no female–female competition or male choice ($\tilde{m}_{\mathrm{c}}=-\infty$). On average, females produce one litter of about three offspring per breeding season (40), so we set $m_{\mathrm{f}}=3$. Males do not provide direct benefits.

Black-Legged Kittiwake (*Rissa tridactyla*). In the absence of other evidence we assumed an adult sex ratio of $\mathrm{ASR}=0.5$. Black-legged kittiwakes are almost entirely genetically monogamous (46) and parental care is shared fairly equally between the sexes (43). We consequently assumed that the length of time out was equal for the sexes ($T=\tilde{T}=1$) and that neither sex engaged in mate search after accepting a mate ($S=\tilde{S}=0$). We arbitrarily set the mate encounter rate to $\lambda=0.1$. There is strong assortative mating in kittiwakes (45), so we set the median mate choice to $m_{\mathrm{c}}=\tilde{m}=0$ and slope to $a_{\mathrm{c}}=\tilde{a}_{\mathrm{c}}=0.5$. This means that an average kittiwake of either sex has a probability of 0.5 of accepting a mate of the same quality, but choosiness increases with an individual's own quality. Pairs lay an average of 1.5 eggs per breeding season (44) and we assumed that males and females contribute equally to productivity ($m_{\mathrm{f}}=\tilde{m}_{\mathrm{b}}=0.75$).

Honeylocust Beetles (*Megabruchidius dorsalis* and *Megabruchidius tonkineus*). Adult sex ratios in both species are approximately equal (K. Fritzsche, personal observation), so we set $\mathrm{ASR}=0.5$. Both sexes have refractory periods after each mating, with females becoming receptive slightly faster than males (K. Fritzsche, personal observation). We consequently set $T=0.8$ for females, $\tilde{T}=1$ for males, and $S=\tilde{S}=0$. Although females mate with only one male per reproductive bout, they store sperm from previous mates (48). Because the last mate fertilizes the majority of eggs (48), we assumed that the competitiveness of sperm declines exponentially over time with a rate constant of $r=1$ (main text). Males are moderately choosy, but females less so (50), so we set $m_{\mathrm{c}}=-2$ and $\tilde{m}_{\mathrm{c}}=0$. Males provide substantial direct benefits through their large ejaculates, although studies differ greatly in their estimates of the relative resource contributions of males and females toward offspring (28, 47, 49). We assumed values for median female fecundity per reproductive bout ($m_{\mathrm{f}}=20$) and median direct benefits provided by males ($\tilde{m}_{\mathrm{c}}=10$) that lie in the middle of the range of estimates presented in these studies.

Broadnosed Pipefish (*Syngnathus typhle*). We assumed equal adult sex ratios of $\mathrm{ASR}=0.5$ (51, 59). Estimates of female fecundity and male pouch capacity varied substantially between studies (52–55). We assumed that females transfer a median of $m_{\mathrm{f}}=50$ eggs per mating and that median male caring capacity is $\tilde{m}_{\mathrm{b}}=75$. We set the length of time out to $T=0.3$ for females and $\tilde{T}=0.9$ for males. Given that females transfer approximately two-thirds as many eggs per mating as a male can incubate, this means that overall females produce eggs approximately twice as fast as males can incubate (53, 54). We assume that females enter time out directly after mating ($S=0$), which ensures that a female's egg production is not directly determined by her number of mates. Males search for additional mates for $\tilde{S}=0.1$, which, at a mate encounter

rate of $\lambda = 0.1$, means that males have between two and three mates per pregnancy on average. Males are choosier than females (56), so we assumed $m_c = -2$ and $\tilde{m}_c = 0$.

Parameters Applying to all Mating Systems. For many traits, information on the magnitude of interindividual variance was not available. We consequently assumed a standard slope of 0.2 between individuals' realized fecundity, direct benefits, or gamete competition ability and their standard trait values (respectively Z_f, $\tilde{Z}_b$, and Z_e or $\tilde{Z}_s$) unless stated otherwise above. This means that individuals whose standardized trait values were 3 SD above the mean ($Z = 3$) had realized trait values almost twice as large as those of average individuals ($Z = 0$), whose realized trait values were in turn almost twice those of individuals who were 3 SD below the mean ($Z = -3$).

There was generally little information about correlations among different fitness components across individuals and the data we did find suggested modest positive correlations [e.g., mate competition and sperm competition ability in male red deer (32) and fecundity and egg competition ability in female pipefish (57, 58)]. We consequently used trait correlations ρ and $\tilde{\rho}$ that were chosen randomly from the range [0,0.3] for all species.

We found no evidence of assortative mating in any species other than kittiwakes. Mobley et al. (59) explicitly looked for assortative mating pipefish and found none. We consequently set the slopes of mate choice and competition to $a_c = \tilde{a}_c = 0$ for all species other than kittiwakes.

Randomization

For each simulation run, we perturbed the values of each parameter by a random amount from its original value for that mating system (summarized in Table S2). For most parameters, new values were chosen uniformly from the interval $[0.5x, 1.5x]$, where x is the original value of the parameter. For some parameters (e.g., the ASR and thresholds for mate choice and competition) this method would have resulted in biologically unrealistic values, so we used a different procedure. For the ASR, we chose new values uniformly from within ± 0.1 of the original value. The median thresholds for mate choice and competition (m_c and $\tilde{m}_c$) were similarly chosen uniformly from within ± 1 of their original values. This corresponds to differences of up to 1 SD in the quality of mating trait that an individual would require to achieve a given level of mating success. For the black-legged kittiwakes, slopes of the mate choice/competition thresholds (a_c and $\tilde{a}_c$) were chosen uniformly from $[0.25, 0.75]$.

Supplementary Simulations with Stabilizing Selection

We assume in the main text that an individual's probability of mating is an increasing function of its mating trait value. Here we instead consider the possibility that mating probabilities peak at intermediate trait values, which leads to stabilizing selection on the mating trait. As before, when a female with mating trait Z_c encounters a male with mating trait $\tilde{Z}_c$, the probability that they mate is given by

$$p = q(Z_c, \tilde{Z}_c) \cdot \tilde{q}(Z_c, \tilde{Z}_c).$$

The probability that the female is accepted by, or successfully competes for, the male is now given by the Gaussian function

$$q(Z_c, \tilde{Z}_c) = \exp\left(-\frac{1}{2}\left(\frac{Z_c - z_{opt}}{\tilde{\sigma}(1 + \tilde{a}_\sigma|\tilde{Z}_c - \tilde{z}_{opt}|)}\right)^2\right).$$

Similarly, the probability that the male is accepted by, or can compete for, the female is

$$\tilde{q}(Z_c, \tilde{Z}_c) = \exp\left(-\frac{1}{2}\left(\frac{\tilde{Z}_c - \tilde{z}_{opt}}{\sigma(1 + a_\sigma|Z_c - z_{opt}|)}\right)^2\right).$$

The parameters z_{opt} and $\tilde{z}_{opt}$ are the trait values (for females and males, respectively) at which mating success is maximized. The parameters σ and $\tilde{\sigma}$ determine how quickly mating success drops off on either side of these trait values (i.e., the "peak width"). Smaller values of σ or $\tilde{\sigma}$ represent strong mate choice or competition, in which only individuals close to the optimal trait values are likely to mate. The slopes a_σ and $\tilde{a}_\sigma$ determine how an individual's quality affects its choosiness or the strength of competition to mate with it. For example, if $a_\sigma = 0$, then all females are equally likely to mate with a given male, whereas if $a_\sigma > 0$, then females with trait values closer to the optimum are choosier or more difficult to compete for.

We ran two sets of supplementary simulations. In the first set, the optimal mating trait values were fixed at $z_{opt} = \tilde{z}_{opt} = 0$ for both sexes. This leads to pure stabilizing selection on the mating trait. In the second set, we chose the optimal mating-trait value uniformly from the interval $[0, 2]$, independently for each sex (Table S2). This leads to a mixture of directional and stabilizing selection on the mating trait. We chose the peak width parameters so that species with stronger mate choice or competition had narrower peaks (i.e., smaller σ or $\tilde{\sigma}$: Tables S1 and S2). For kittiwakes, we chose the slopes of the peak widths a_σ and $\tilde{a}_\sigma$ uniformly from $[0, 1]$ for each sex. The supplementary simulations were otherwise identical to those in the main text.

The Total Effect of Selection on a Trait

When fitness increases monotonically with trait values, the linear selection differential s' is a good measure of the strength of selection. However, when selection has considerable nonlinear components (e.g., stabilizing or disruptive selection), these will be overlooked by the linear selection differential, which may consequently underestimate the strength of selection. Similarly, the quadratic selection differential c' describes the curvature of the univariate selection surface, but does not account for the effects of directional selection. To estimate the full strength of selection on a trait, we must therefore look at the total change in the trait's distribution, including both linear and nonlinear effects.

We define the "distributional selection differential" d' on a standardized univariate trait as

$$d' = \int_{-\infty}^{\infty} |F(z) - G(z)| dz,$$

where $F(z)$ and $G(z)$ are the cumulative distribution functions of the trait distribution before and after selection. This is simply the area between the cumulative distribution curves before and after selection. Equivalent metrics have been used to quantify differences in trait distributions in other contexts (60, 61).

When fitness is a monotonic function of trait values, the distributional selection differential is equal to the magnitude of the linear selection differential (i.e., $d' = |s'|$). However, d' also reflects nonlinear components of selection. For instance, we might define pure stabilizing or disruptive selection to occur when standardized trait values are symmetrically distributed before selection and expected fitness $w(z)$ is a symmetric monotonic function of $|z|$. In this case, the linear selection differential is zero, but $d' = |\mathrm{cov}(w(z), |z|)|$. The distributional selection differential thus captures both linear and nonlinear components of selection.

Results of Supplementary Simulations

The results of our supplementary simulations were qualitatively similar to those in the main text. The Jones index s'_{max} was the best predictor of the strength of sexual selection (as measured by its

correlation with the distributional selection differential d') both under pure stabilizing selection (Table S4) and under a mixture of stabilizing and directional selection (Table S5). The Morisita index $I_{\delta-\mathrm{RS}}$ and the monopolization index Q_{RS} for reproductive success also did reasonably well, whereas the Bateman gradient β_{ss} and opportunity for sexual selection I_s did poorly. When we repeated our analyses using Spearman's rank correlation ρ instead of Pearson's correlation coefficient r, the differences in performance between the Jones, Morisita, and monopolization indexes were less marked, although these three measures still outperformed all others (Tables S7 and S8).

The linear and quadratic selection differentials (s' and c') showed weaker and less consistent relationships with the Jones index than the distributional selection differential d', because the former measures do not capture the full effect of selection on a trait distribution (Fig. S1). Under pure stabilizing selection, where nonlinear effects dominated, the Jones index was highly correlated with the quadratic selection differential, but showed no relationship with the linear selection differential. In contrast, under a mixture of directional and stabilizing selection, the Jones index was highly correlated with the linear selection differential, but not with the quadratic selection differential. In both cases, the Jones index predicted the distributional selection differential significantly better than either the linear or the quadratic selection differential (Williams' test: $|T_2| > 2.9$, $P < 0.004$ for all pairwise comparisons across the six datasets under both selection regimes).

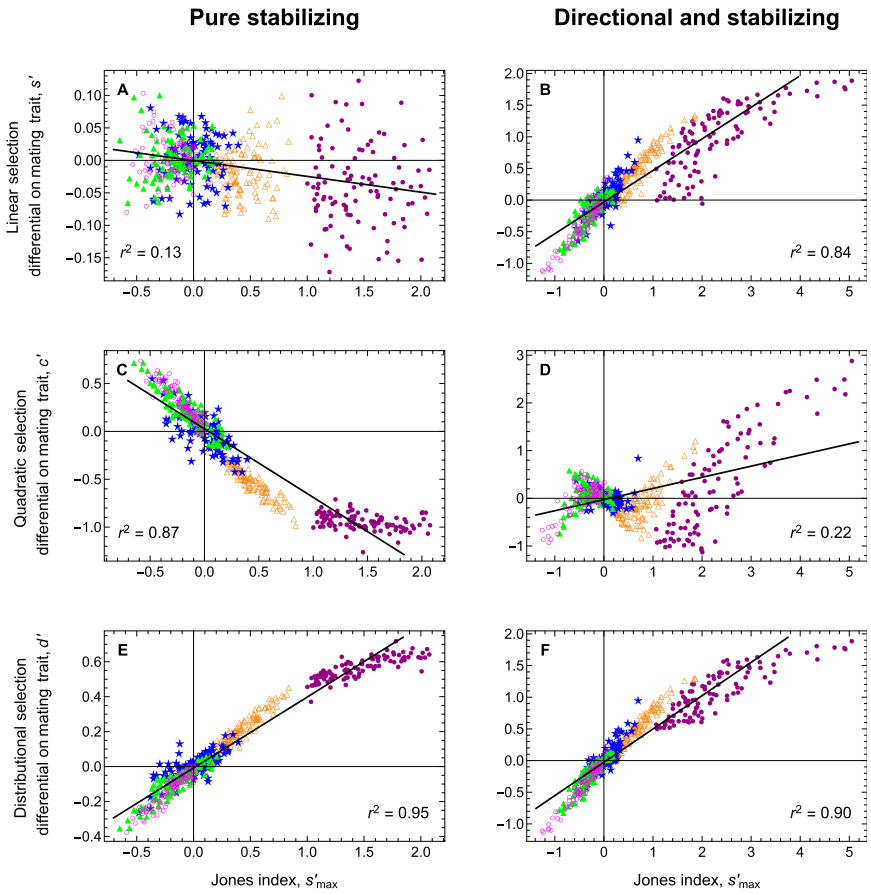

Fig. S1. Performance of the Jones index $s'_{\max}$ in predicting sex differences in selection on the mating trait. Shown is the regression of sex differences in (*A* and *B*) the linear selection differential, s'; (*C* and *D*) the quadratic selection differential, c'; and (*E* and *F*) the distributional selection differential, d', against sex differences in the Jones index under either pure stabilizing selection (*A*, *C*, and *E*) or a mixture of directional and stabilizing selection (*B*, *D*, and *F*). Data points are based on red deer (purple solid circles), American red squirrels (orange open triangles), black-legged kittiwakes (blue stars), honeylocust beetles (green solid triangles), and broadnosed pipefish (pink open circles), using long-term data representing multiple reproductive bouts.

Table S1. Default parameter values for the five species

Parameter	Deer	Squirrel	Kittiwake	Honeylocust beetle	Pipefish
ASR	0.4	0.5	0.5	0.5	0.5
Rate of mate encounters λ	0.1	0.5	0.1	0.1	0.1
Length of time out T and mate search S, females	$\left(\frac{29}{30}, \frac{1}{30}\right)$	$(0.96, 0.04)$	$(1,0)$	$(0.8, 0)$	$(0.3, 0)$
Length of time out $\widetilde{T}$ and mate search $\widetilde{S}$, males	No time out	No time out	$(1,0)$	$(1,0)$	$(0.9, 0.1)$
Median thresholds m_c and $\widetilde{m}_c$ for male/female mate competition	$(2, -\infty)$	$(0, -\infty)$	$(0,0)$	$(-2,0)$	$(-2,0)$
Fecundity median m_f and slope a_f	$(0.5, 0.1)$	$(3,0.2)$	$(0.75, 0.2)$	$(20,0.2)$	$(50,0.2)$
Direct benefits median $\widetilde{m}_b$ and slope $\widetilde{a}_c$	No direct benefits	No direct benefits	$(0.75, 0.2)$	$(10,0.2)$	$(75,0.2)$
Egg competition slope a_e	NA	NA	NA	NA	0.2
Sperm competition slope $\widetilde{a}_s$	0.2	0.2	NA	0.2	NA
Supplementary simulations with stabilizing selection					
Width of peak σ and $\widetilde{\sigma}$ around optimal mating-trait values	$(0.25, \infty)$	$(1, \infty)$	$(1,1)$	$(2,1)$	$(2,1)$

NA, not applicable.

Table S2. Randomization of parameters

Parameter	Random value chosen uniformly from (where original value is denoted x)
Rate of mate encounters λ	$[0.5x, 1.5x]$
Length of time out T and $\widetilde{T}$	$[0.5x, 1.5x]$
Length of mate search S and $\widetilde{S}$	$[0.5x, 1.5x]$
Median m_f and slope a_f of fecundity	$[0.5x, 1.5x]$
Median $\widetilde{m}_b$ and slope $\widetilde{a}_b$ of direct benefits	$[0.5x, 1.5x]$
Slope of sperm competition $\widetilde{a}_s$ and egg competition a_e	$[0.5x, 1.5x]$
Rate of sperm discounting r	$[0.5x, 1.5x]$
ASR	$[x-0.1, x+0.1]$
Median thresholds for mate competition m_c and $\widetilde{m}_c$	$[x-1, x+1]$
Slope of thresholds for mate competition a_c and $\widetilde{a}_c$ (kittiwakes only)	$[0.25, 0.75]$
Trait correlations ρ and $\widetilde{\rho}$	$[0, 0.3]$
Supplementary simulations with stabilizing selection	
Width of peak σ and $\widetilde{\sigma}$ around optimal mating trait values	$[0.5x, 1.5x]$
Slope of peak width a_σ and $\widetilde{a}_\sigma$ around optimal mating trait values (kittiwakes only)	$[0,1]$
Optimal mating trait values z_{opt} and $\widetilde{z}_{opt}$ in simulations with a mixture of directional and stabilizing selection	$[0,2]$

Table S3. Performance of proxy measures in predicting the strength of premating sexual selection within the five mating system types

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index, s'_{max}						
D	0.60	0.00	0.21	0.81	0.01	0.58
S	0.74	0.05	0.39	0.93	0.01	0.82
K	0.24	0.14	0.23	0.41	0.33	0.42
HB	0.25	0.66	0.73	0.55	0.89	0.86
P	0.12	0.73	0.48	0.22	0.93	0.76
Overall	0.80	0.39	0.85	0.95	0.75	0.96
Morisita index for reproductive success, $I_{\delta-RS}$						
D	0.56	0.00	0.35	0.79	0.02	0.68
S	0.63	0.00	0.53	0.88	0.06	0.79
K	0.19	0.07	0.21	0.37	0.26	0.41
HB	0.29	0.62	0.72	0.53	0.80	0.81
P	0.10	0.62	0.53	0.22	0.86	0.83
Overall	0.76	0.33	0.79	0.89	0.71	0.87
Monopolization index for reproductive success, Q_{RS}						
D	0.60	0.00	0.48	0.76	0.03	0.68
S	0.68	0.00	0.58	0.87	0.06	0.78
K	0.23	0.08	0.27	0.43	0.28	0.48
HB	0.27	0.65	0.75	0.56	0.81	0.86
P	0.10	0.65	0.52	0.23	0.84	0.81
Overall	0.77	0.38	0.79	0.84	0.73	0.82
Opportunity for selection, I						
D	0.19	0.00	0.02	0.45	0.00	0.42
S	0.38	0.03	0.42	0.72	0.00	0.74
K	0.07	0.01	0.12	0.14	0.07	0.26
HB	0.26	0.61	0.72	0.51	0.80	0.83
P	0.09	0.61	0.52	0.19	0.86	0.83
Overall	0.63	0.12	0.35	0.81	0.07	0.83
Morisita index for mating success, $I_{\delta-MS}$						
D	0.65	0.00	0.42	0.79	0.01	0.66
S	0.86	0.01	0.47	0.92	0.02	0.66
K	0.14	0.15	0.28	0.25	0.24	0.43
HB	0.13	0.10	0.33	0.17	0.36	0.52
P	0.00	0.33	0.28	0.02	0.78	0.68
Overall	0.77	0.00	0.65	0.90	0.05	0.79
Monopolization index for mating success, Q_{MS}						
D	0.61	0.00	0.47	0.71	0.01	0.65
S	0.84	0.01	0.47	0.87	0.02	0.65
K	0.17	0.14	0.28	0.30	0.24	0.49
HB	0.21	0.20	0.73	0.29	0.48	0.85
P	0.00	0.43	0.44	0.02	0.79	0.74
Overall	0.80	0.00	0.65	0.88	0.04	0.76
Opportunity for sexual selection, I_s						
D	0.53	0.01	0.35	0.72	0.00	0.65
S	0.6	0.01	0.42	0.87	0.02	0.65
K	0.24	0.11	0.23	0.39	0.33	0.42
HB	0.34	0.61	0.70	0.48	0.81	0.84
P	0.02	0.65	0.49	0.01	0.85	0.76
Overall	0.70	0.08	0.62	0.85	0.41	0.81
Bateman gradient, β_{ss}						
D	0.06	0.00	0.06	0.07	0.01	0.16
S	0.02	0.05	0.01	0.00	0.01	0.05
K	0.00	0.00	0.00	0.00	0.01	0.00
HB	0.00	0.00	0.01	0.19	0.01	0.13
P	0.13	0.02	0.01	0.16	0.01	0.01
Overall	0.31	0.34	0.62	0.24	0.40	0.64

Shown is the coefficient of determination r^2 between the proxy measure and the linear selection differential s' on the mating trait (for the "sex difference" columns, we calculate r^2 based on the sex differences in proxy measures and selection differentials). Short-term data represent approximately one reproductive bout, whereas long-term data represent several reproductive bouts. Mating systems are given as: D, red deer; HB, honeylocust beetles; K, black-legged kittiwakes; Overall, correlation across all five groups as shown in Table 2; P, broadnosed pipefish; S, American red squirrels.

Table S4. Performance of proxy measures in predicting the distributional selection differential d' on the mating trait (i.e., the total effect of both linear and nonlinear selection) under pure stabilizing selection

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index						
s'_{max}	0.90	0.47^{c}	0.89	0.94	0.80^{p}	0.95
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.79^{a}	0.47^{c}	0.75	0.80^{m}	0.83^{p}	0.75
Q_{RS}	0.76^{b}	0.52	0.71^{f}	0.77	0.84	0.71st
I	0.78^{ab}	0.00^{d}	0.71^{f}	0.82	0.01	0.75
Measures based on variance in mating success						
$I_{\delta-MS}$	0.84	0.00^{d}	0.67^{gh}	0.83	0.05^{q}	0.70^{s}
Q_{MS}	0.82	0.00^{d}	0.65^{k}	0.79^{mn}	0.05^{q}	0.67^{u}
I_s	0.77^{b}	0.24^{e}	0.67^{g}	0.80^{n}	0.32^{r}	0.71^{t}
Bateman gradient						
β_{ss}	0.26	0.20^{e}	0.60^{hk}	0.21	0.25^{r}	0.64^{u}

Shown is the coefficient of determination r^2 between the proxy measure and the distributional selection differential (for the "sex difference" columns, we calculate r^2 based on the sex differences in proxy measures and selection differentials). Values within a column are marked with the same letter if the unsquared coefficients r do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.

Table S5. Performance of proxy measures in predicting the distributional selection differential d' on the mating trait (i.e., the total effect of both linear and nonlinear selection) under a mixture of directional and stabilizing selection

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index						
s'_{max}	0.86	0.76^{e}	0.87	0.88	0.92	0.90
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.64^{ab}	0.76^{e}	0.57^{g}	0.63^{mn}	0.85	0.56^{s}
Q_{RS}	0.64^{ac}	0.79	0.56^{gh}	0.63^{mp}	0.84^{r}	0.55st
I	0.66^{d}	0.02	0.55^{hk}	0.64^{pq}	0.43	0.56^{tu}
Measures based on variance in mating success						
$I_{\delta-MS}$	0.66^{d}	0.24^{f}	0.55^{k}	0.64^{pq}	0.70	0.55^{tu}
Q_{MS}	0.65^{bd}	0.23^{f}	0.54^{k}	0.64^{nq}	0.69	0.54^{u}
I_s	0.63^{c}	0.72	0.54^{k}	0.63^{mn}	0.83^{r}	0.55^{tu}
Bateman gradient						
β_{ss}	0.25	0.23^{f}	0.45	0.17	0.25	0.43

Shown is the coefficient of determination r^2 between the proxy measure and the distributional selection differential (for the "sex difference" columns, we calculate r^2 based on the sex differences in proxy measures and selection differentials). Values within a column are marked with the same letter if the unsquared coefficients r do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.

Table S6. Performance of proxy measures in predicting the linear selection differential s' on the mating trait (i.e., the strength of premating sexual selection) under pure directional selection, measured using the square ρ^2 of Spearman's rank correlation coefficient

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index						
s'_{max}	0.68^{a}	0.43^{c}	0.81^{ef}	0.82^{h}	0.69	0.93^{mn}
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.68^{a}	0.36^{d}	0.80^{e}	0.81^{h}	0.57	0.93^{m}
Q_{RS}	0.69^{a}	0.40^{c}	0.82^{f}	0.81^{h}	0.59	0.92^{n}
I	0.55^{b}	0.09	0.48^{g}	0.71^{k}	0.04	0.87^{p}
Measures based on variance in mating success						
$I_{\delta-MS}$	0.6^{b}	0.01	0.49^{g}	0.63	0.01	0.81
Q_{MS}	0.6^{b}	0.01	0.53^{g}	0.64	0.00	0.85^{pq}
I_s	0.47	0.04	0.53^{g}	0.74^{k}	0.26	0.85^{q}
Bateman gradient						
β_{ss}	0.36	0.42^{cd}	0.66	0.32	0.45	0.66

All data are the same as in Table 2. Values within a column are marked with the same letter if the unsquared coefficients ρ do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.

Table S7. Performance of proxy measures in predicting the distributional selection differential d' on the mating trait (i.e., the total effect of both linear and nonlinear selection) under pure stabilizing selection, measured using the square ρ^2 of Spearman's rank correlation coefficient

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index						
s'_{max}	0.71^{a}	0.46^{f}	0.83^{k}	0.83^{p}	0.65^{r}	0.94^{u}
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.68^{bc}	0.42^{f}	0.82	0.83^{p}	0.63^{r}	0.94^{u}
Q_{RS}	0.69^{ac}	0.42^{f}	0.84^{k}	0.83^{p}	0.63^{r}	0.95
I	0.64^{cd}	0.00	0.77	0.75^{q}	0.01^{s}	0.91
Measures based on variance in mating success						
$I_{\delta-MS}$	0.60^{d}	0.04^{g}	0.46^{m}	0.66	0.00^{s}	0.70^{v}
Q_{MS}	0.61^{d}	0.04^{g}	0.50^{n}	0.68	0.00^{s}	0.74
I_s	0.46^{e}	0.15^{h}	0.50^{mn}	0.76^{q}	0.27^{t}	0.72^{v}
Bateman gradient						
β_{ss}	0.40^{e}	0.24^{h}	0.65	0.36	0.27^{t}	0.65^{v}

All data are the same as in Table S4. Values within a column are marked with the same letter if the unsquared coefficients ρ do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.

Table S8. Performance of proxy measures in predicting the distributional selection differential d' on the mating trait (i.e., the total effect of both linear and nonlinear selection) under a mixture of directional and stabilizing selection, measured using the square ρ^2 of Spearman's rank correlation coefficient

	Short-term			Long-term		
Proxy measure	Male	Female	Sex difference	Male	Female	Sex difference
Jones index						
s'_{max}	0.83[a]	0.69[bc]	0.89	0.90[g]	0.79	0.96[n]
Measures based on variance in reproductive success						
$I_{\delta-RS}$	0.82[a]	0.67[b]	0.90	0.90[g]	0.74[h]	0.96[n]
Q_{RS}	0.82[a]	0.68[c]	0.91	0.90[g]	0.75[h]	0.96[n]
I	0.70	0.03[d]	0.87	0.79	0.16[km]	0.95[p]
Measures based on variance in mating success						
$I_{\delta-MS}$	0.59	0.00[d]	0.80	0.70	0.20[k]	0.94
Q_{MS}	0.60	0.00	0.83[f]	0.71	0.19[m]	0.95[p]
I_s	0.74	0.48[e]	0.82[f]	0.87	0.68	0.95[p]
Bateman gradient						
β_{ss}	0.32	0.42[e]	0.59	0.24	0.40	0.55

All data are the same as in Table S5. Values within a column are marked with the same letter if the unsquared coefficients ρ do not differ significantly in pairwise comparisons. Unmarked values are significantly different from all others in the same column.