doi:10.1111/evo.12022

THE MAINTENANCE OF SPERM VARIABILITY: CONTEXT-DEPENDENT SELECTION ON SPERM MORPHOLOGY IN A BROADCAST SPAWNING INVERTEBRATE

Darren W. Johnson,[1,2] Keyne Monro,[3,4] and Dustin J. Marshall[3,4]

[1]*National Center for Ecological Analysis and Synthesis, Santa Barbara, California 93101*

[2]*E-mail: Johnson@nceas.ucsb.edu*

[3]*School of Biological Sciences, The University of Queensland, QLD 4072, Australia*

[4]*School of Biological Sciences, Monash University, VIC 3800, Australia*

Received July 19, 2012
Accepted November 1, 2012
Data Archived: Dryad doi:10.5061/dryad.c43f1

Why are sperm so variable despite having a singular, critical function and an intimate relationship with fitness? A key to understanding the evolution of sperm morphology is identifying which traits enable sperm to be successful fertilizers. Several sperm traits (e.g., tail length, overall size) are implicated in sperm performance, but the benefits of these traits are likely to be highly context dependent. Here, we examined phenotypic selection on sperm morphology of a broadcast spawning tube worm (*Galeolaria gemineoa*). We conducted laboratory experiments to measure the relationship between average sperm morphology and relative fertilization success across a range of sperm environments that were designed to approximate the range of sperm concentrations and ages encountered by eggs in nature. We found that the strength and form of multivariate selection varied substantially across our environmental gradients. Sperm with long tails and small heads were favored in high-concentration environments, whereas sperm with long heads were favored at low concentrations and old ages. We suggest variation in the local fertilization environment and resulting differences in selection can preserve variability in sperm morphology both within and among males.

KEY WORDS: **Anisogamy, canonical analysis, disruptive selection, fertilization success, fitness, sperm competition, sperm limitation.**

Some of the most fundamental and enduring questions in evolutionary biology concern the evolution of gamete size and morphology (e.g., Kalmus 1932; Scudo 1967). The evolutionary forces shaping gamete morphology may be nuanced, complex, and are incompletely understood (see reviews by Randerson and Hurst 2001; Snook 2005). There has been a sustained and intense effort by theoreticians to model the evolution of anisogamy (small sperm and large eggs; Parker et al. 1972, Parker 1982; Levitan 1993; Dusenbery 2000; Lehtonen and Kokko 2011), but the assumptions underlying these models have received less empirical attention. In particular, although it now seems clear that males that produce more sperm may often secure more fertilizations, the role of sperm quality has been less explored. Indeed, a recent focus on sperm quality has highlighted the importance of understanding the evolutionary forces that shape not only sperm size, but also fertilization performance and its morphological underpinnings (reviewed by Snook 2005).

Although formal analyses of selection on sperm morphology are rare, one line of evidence to suggest an adaptive role for sperm morphology is the theoretical link between morphological traits (e.g., size and shape of sperm components) and sperm function, especially swimming speed (reviewed by Humphries et al. 2008).

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

Sperm move through a fluid environment that is dominated by viscosity, rather than inertia (i.e., the hydrodynamic regime is characterized by a low Reynolds number). It is thought that in this environment, swimming speed of sperm is determined by the balance between thrust, which is expected to be proportional to tail length, and drag, which is thought to be largely influenced by the surface area of the sperm head (Humphries et al. 2008). Indeed, among and within species, swimming speed is often correlated with sperm morphology, and speed has been associated with increased fertilization success in both internally and externally fertilizing species (e.g., Birkhead et al. 1999; Levitan 2000; Kupriyanova and Havenhand 2002; Gage et al. 2004; Malo et al. 2005). Sperm morphology may also be linked to other traits that have functional consequences for fertilization. Sperm longevity is often hypothesized to be important for fertilization success, especially under sperm-limited conditions when the average time until sperm-egg encounters is low (Levitan 1993; Williams and Bentley 2002; Johnson and Yund 2004; Pizzari et al. 2008). Although the morphological correlates of longevity are often unclear, there is some evidence that longevity may be related to energy reserves and associated storage structures such as mitochondria (Afzelius and Mohri 1966; Mita et al. 1995).

Comparisons among species and populations provide more evidence for selection on sperm morphology. Several among-species comparisons have indicated that species with mating systems that result in greater sperm competition (e.g., mating systems with a higher degree of promiscuity) often exhibit a corresponding increase in sperm size (e.g., Gomiendo and Roldan 1991; Gage 1994; Byrne et al. 2003; Anderson et al. 2005; Fitzpatrick et al. 2009). In addition, within-species Q_{ST}–F_{ST} comparisons suggest that among-population divergence in sperm morphology is more likely to have resulted from selection rather than neutral processes such as drift (Manier and Palumbi 2008). These authors also suggested that some form of disruptive selection may be responsible for maintaining the abundant, within-population variation in sperm morphology. Although these studies suggest that sperm morphology may be under selection, direct demonstrations of selection on sperm morphology within populations are far rarer.

Unequivocal demonstrations of selection on sperm traits are logistically challenging for many species. Cryptic female choice can obscure the relationship between sperm traits and fertilization success in internal fertilizers (Eberhard 1996). External fertilizers, where eggs and sperm are shed into an aquatic medium, offer the opportunity to study selection on sperm traits in the absence of the potentially confounding issues associated with internal fertilizers. A focus on external fertilizers for understanding selection on sperm has the additional advantage of providing much needed tests of the assumptions made in foundational theory on the evolution of isogamy in a hypothetical ancestral external fertilizer. Accordingly, Fitzpatrick et al. (2012) formally analyzed selection on a suite of ejaculate traits (including some measures of sperm morphology) in a broadcast-spawning mussel (*Mytilus galloprovincialis*). They demonstrated significant selection on several sperm traits, including sperm head size, swimming speed, and swimming trajectory. An important next step toward understanding the role of selection in maintaining observed distributions of sperm phenotypes is to evaluate the degree to which environmental context shapes patterns of selection.

Strong, consistent selection is expected to erode phenotypic and genetic variation (e.g., Fisher 1930; Falconer and Mackay 1996), yet sperm morphology can exhibit substantial variability (e.g., Ward 1998; Morrow and Gage 2001; Kleven et al. 2008; Manier and Palumbi 2008), including a sizeable amount of genetic variability (Ward 1998; Evans 2011). These observations prompt the question of what maintains the observed variability in sperm morphology. One contributing cause may be fluctuations in selection that are driven by variation in fertilization environments. For example, for some internal fertilizers the degree of promiscuity and sperm competition may determine the intensity of selection for faster sperm and associated morphologies (e.g., Gomiendo and Roldan 1991; Anderson et al. 2005; Fitzpatrick et al. 2009). Similarly, selection on sperm morphology may also depend on a male's reproductive role (Calhim et al. 2011). Some sperm morphologies appear to be better suited for storage (and advantageous if a male is the first to reproduce) whereas other morphologies are better suited for competition (and advantageous if a male is trying to outcompete previously stored sperm). For external fertilizers (particularly those whose adults are sedentary), gamete concentrations may range over many orders of magnitude, even across remarkably small scales (e.g., Pennington 1985; Denny and Shibata 1989; Yund 1990; Levitan 1991; Benzie et al. 1994; Babcock and Keesing 1999). Moreover, the relative age of gametes can range over seconds to several hours, depending on asynchrony in spawning and water motion (reviewed by Lotterhos and Levitan 2010). Variation in age and concentration of sperm may therefore have substantial effects on the environmental milieu in which fertilization takes place. It is conceivable that sperm performance, and the sperm phenotypes that improve performance, also depend on this environmental context. If certain sperm phenotypes are better fertilizers in certain environments, then variation in the sperm environment may be an important source of variation in selection on sperm phenotypes. Variation in selection could then contribute to the maintenance of genetic and phenotypic variation in sperm (reviewed by Bell 2010).

Although sperm environments are likely to be highly variable, they may not be entirely unpredictable. If males can anticipate the environment in which their sperm will be released, then selection could favor some degree of plasticity in sperm phenotypes. There is increasing evidence for such plasticity. For

example, Crean and Marshall (2008) demonstrated that broadcast-spawning ascidians placed in high-density environments (i.e., those likely to experience greater sperm competition) produced larger, more motile sperm. Similarly, recent experiments on Gouldian finches indicate that individuals may adjust various aspects of sperm morphology in response to the degree of competitiveness in the local environment (Immler et al. 2010). Finally, in cichlid fishes that are plastic with respect to their reproductive tactics (i.e., males can hold breeding territories or adopt various types of "sneaker" strategies, depending on relative body size) individuals that were more likely to experience intense numerical competition for fertilizations (i.e., sneaker males) produced sperm with increased longevity (Ota et al. 2010). These results suggest that adults can to some extent adjust sperm characteristics based on local environmental cues. Moreover, in some cases a male's own phenotype may be a successful predictor of the local fertilization environment. Although studies have shown that sperm morphology can vary with adult traits (e.g., testis weight, body condition; Simmons and Kotiaho 2002), the degree to which this covariation is adaptive remains unclear as too few studies have addressed how selection on sperm traits varies with environmental conditions.

In this study we examined phenotypic selection on sperm morphology of a broadcast spawning tube worm (*Galeolaria geminoea*). We conducted laboratory experiments to measure the relationship between average sperm morphology and relative fertilization success across a range of sperm environments (i.e., various combinations of sperm age and concentration). Our goals were (1) to identify any relationships between sperm morphology (i.e., independent and combined measures of tail length, head length, and head width) and relative fertilization success, and examine whether selection on these traits varied among sperm environments; and (2) to examine the relationships between sperm morphology and two adult traits—body size and total sperm number—that may have strong effects on sperm environment.

Methods

STUDY SPECIES

Galeolaria gemineoa is a tube-dwelling, serpulid polychaete that inhabits the rocky intertidal zone along the eastern coast of Australia. It has a morphologically indistinguishable congener (*G. caespitosa*) that is found along the southern and south-western coasts (Halt et al. 2009). Much research has been done on the reproductive biology of *G. caespitosa* and we assume that the reproductive biology of *G. gemineoa* is very similar. Population densities of *Galeolaria* may range from solitary encrustations to large, extremely dense clusters in which calcareous tubes intertwine to form mats up to 4-cm thick (Minchinton 1997). *G. gemineoa* is a dioecous, broadcast spawner. Fertile specimens can be found throughout the year, and observations of clusters of "spent" individuals with depleted gametes suggest that there is at least some degree of local synchrony in natural spawning behavior (Kupriyanova 2006). The wide range of local densities suggests that eggs of *Galeolaria gemineoa* are likely to experience a substantial range of sperm concentrations. Males and females may be located immediately next to one another (suggesting eggs can be sperm-saturated and fertilization success can approach 100%), but experimental simulations of field spawning of *G. caespitosa* have indicated that at a distance of 1 cm from the sperm source, fertilization success was approximately 20% to 30% and declined to 5% to 10% at a distance of 26 cm (Hollows et al. 2007). Similarly, studies of other broadcast spawning intertidal polychaetes (e.g., Williams et al. 1997) suggest that spawning may occur during periods of low tide and that eggs may encounter sperm of varying ages, possibly up to a few hours old.

EXPERIMENTAL DESIGN

During early October 2011, mature specimens of *G. gemineoa* were collected from intertidal habitats in Southeast Queensland (27°56'12"S, 153°29'39"E) and held in recirculating aquaria at the University of Queensland. To measure selection on sperm morphology, we spawned males and females in the lab each day, measured sperm phenotypes (described below), and conducted fertilization trials in which we examined fertilization success of individual males across a range of six sperm concentrations and two sperm ages: "new" sperm that was used immediately and "old" sperm that was used in fertilization trials 3 h after initial release. Because sperm in seawater have a finite lifespan (for *G. caespitosa*, average sperm half-life is approximately 3 h; Kupriyanova 2006), old sperm treatments had a somewhat lower effective concentration than their new counterparts. Sperm characteristics associated with differences in fertilization success between old and new sperm treatments may therefore be associated with sperm longevity (Levitan 2000).

Fertilization trials consisted of sperm from single males fertilizing eggs from multiple females. We did not have the capacity to identify parentage of zygotes (e.g., via molecular analyses), so we did not conduct fertilization trials with multiple males at a time. Although we acknowledge that when logistically possible, combining fertilization trials with parentage analysis can provide a more direct examination of sperm competition, we believe that comparing separate fertilization trials among males can provide useful estimates of selection because in nature, many fertilizations may occur when eggs encounter sperm from only a single male (see Fitzpatrick et al. 2012 for a similar approach applied to a species of broadcast-spawning mussel).

To induce spawning, individuals were removed from their tubes and placed in their own 40 mm Petri dish containing 2.5 mL of 0.45 μm filtered seawater. After males completed spawning,

the sperm solution was collected in a 5 mL syringe and gently mixed to ensure a homogeneous solution. One milliliter of the sperm solution was immediately allocated to the first round of fertilization trials (new sperm treatment—fertilization trials are described in detail below). A total of 0.1 mL of the sperm solution was used to measure sperm concentration by conducting three replicate counts using a hemocytometer. The remaining sperm–seawater solution was kept for 3 h in a constant temperature room at 21°C before being used in the second round of fertilization trials (old sperm treatment). Eggs were collected from > 10 females to minimize the effects of male–female compatibility on our estimates of a male's average fertilization success (Kupriyanova and Havenhand 2002). The egg–seawater solution was diluted or supplemented to maintain a concentration of approximately 2×10^4 eggs per mL.

Within each fertilization trial, 0.9 mL of the initial sperm solution ($\sim 10^8$ sperm per mL) was placed in the first well of a six-well plate. This sperm solution was serially diluted 10-fold in each of the five successive plates. A total of 0.1 mL of egg solution was added to each well to yield final sperm concentrations of approx. 10^8, 10^7, 10^6, 10^5, 10^4, and 10^3 sperm per mL. These fertilization environments ranged from extreme sperm saturation ($\sim 5 \times 10^4$ sperm per egg) to sperm limitation ($\sim$20 eggs per sperm) and encompassed much of the range of concentrations likely to be encountered in nature. A single, pooled egg solution was used during each round of fertilization trials (4–12 males per round) and an additional treatment containing eggs only was included to control for any errant fertilizations (e.g., that may have occurred when worms were extracted from their tubes). Fertilized eggs were left for 90 min in a constant temperature room (21°C) before fertilization success was measured by recording the proportion of eggs undergoing normal development. Two observers examined samples of 100 eggs from each well and the number of fertilized eggs was averaged. Observers' counts of fertilized eggs were highly correlated (intraclass correlation = 0.97) suggesting our measures of average fertilization were robust.

For the new sperm treatment, fertilization trials were conducted within 30 min of initial gamete release. For the old sperm treatment, fertilization trials were conducted 3 h after initial sperm release, but with freshly extracted eggs. A total of 77 males were used in the fertilization trials, but due to logistic constraints, trials with old sperm treatments were not conducted for 10 of the 77 individuals.

SPERM MEASUREMENTS

To measure average sperm morphology for each male, we captured digital images of sperm under a compound microscope. For each male, we photographed at least 15 sperm under oil-immersed, 1000× magnification (Fig. 1). The tip of the sperm tail was identified by adjusting the fine focus and digitally marking its location on the image. Images were imported into the program Image J (http://imagej.nih.gov/ij) and head width, head length, and tail length were recorded. Head width and length were measured as straight line distances, whereas tail length was measured along the curvature of sperm tails. Images were enlarged and length was measured as the sum of a series of line segments superimposed upon the image of the tail. Note that our measure of head length includes both the acrosome and a ring of mitochondria at the base of the head (Fig. 1). The resolution of our images was not sufficient to accurately measure these features separately (but see Grant 1981 for scanning electron microscope images and Jamieson and Rouse 1989 for detailed description of sperm microstructure). We only measured normal sperm that had intact tails and that had not undergone an acrosome reaction. Measurements of sperm morphology and fertilization success were made by different observers and were double-blind.

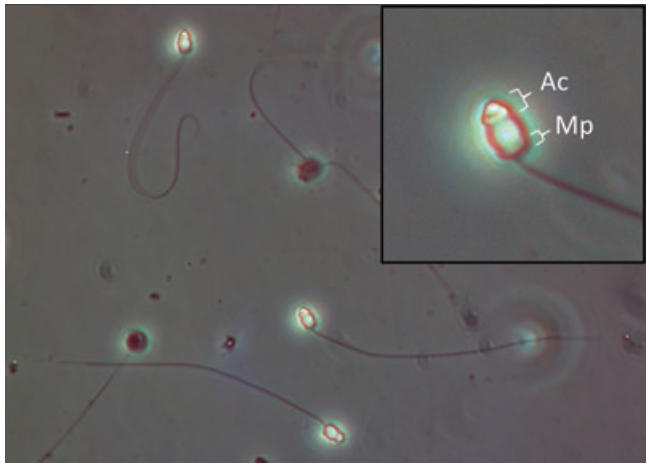

Figure 1. A close-up photograph (1000× power) illustrating sperm morphology in *Galeolaria gemineoa*. Ends of the sperm tails are marked by the bright dots superimposed onto the image. Head length and width were measured as the longest and widest axes of the head. Note that head length includes a measure of both the acrosome (AC) and the mitochondrial packet (MP). For scale, length of sperm heads are approximately 3.6 μm.

DATA ANALYSIS

Because of variation in ejaculate size among males, fertilization success was not assayed at the exact same sperm concentrations. To standardize fertilization success, an individual's mean, percent fertilization values were transformed to a logit scale and fertilization success at each of the target concentrations (10^8, 10^7, 10^6, 10^5, 10^4, and 10^3 sperm per mL) was linearly interpolated from nearby values. Relative success was defined as an individual's fertilization success divided by the average fertilization success of all males within that particular sperm environment. Because average fertilization success increased over the course of the experiment, we included the number of days since the beginning of

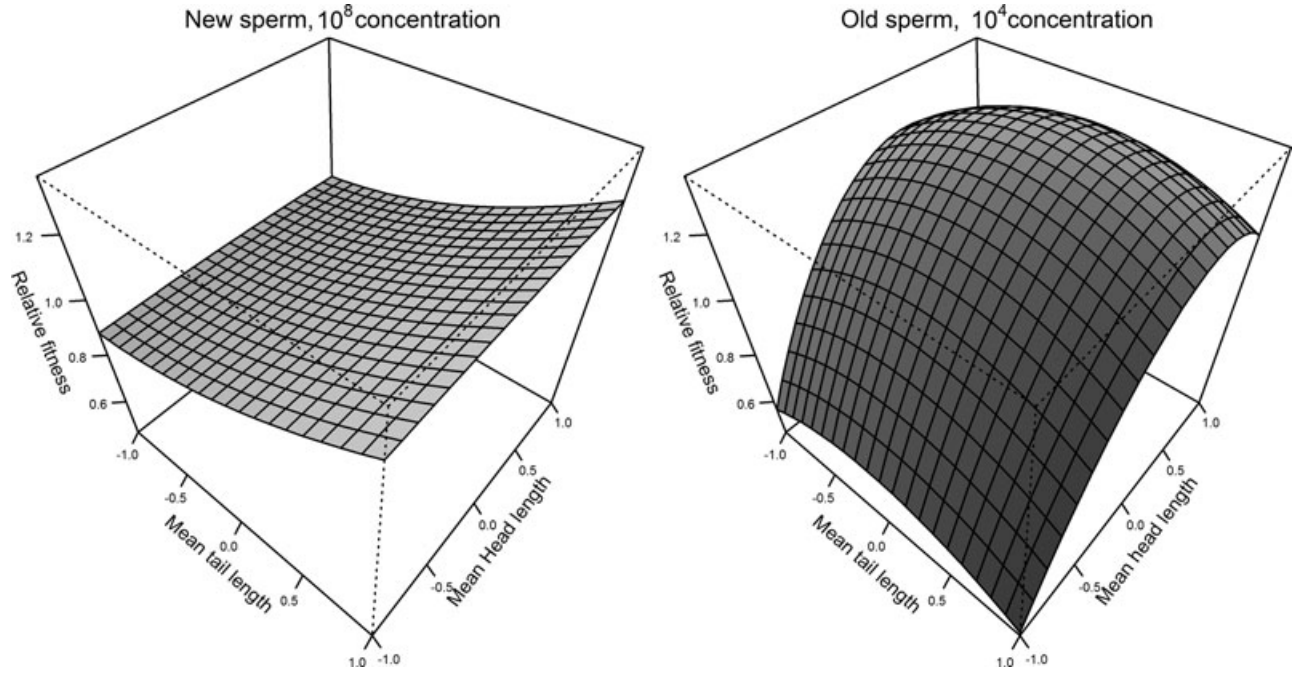

Figure 2. An illustration of the differences in estimated selection surfaces across our sperm age and concentration gradient. Surfaces depict the relationship between relative fertilization success and both mean tail length and mean head length (the two traits we detected significant selection for). Surfaces represent best quadratic approximations to the selection surfaces and used the parameter values in Table 1. Trait values are expressed as standard deviations from the mean.

the experiment as a covariate in all our selection analyses. The increase in average fertilization success over time within the lab has been observed in previous studies of this, and other species (D. Marshall, unpubl. data; Benzie and Dixon 1994; Levitan 2004) and may be related to natural, cyclic spawning behavior of this species (Kupriyanova 2006). Regardless of the exact cause of temporal variation in average fertilization success, there was still ample variation in success among males within each round of fertilization trials and accounting for the temporal trend provided a more direct measure of relative fertilization success.

To measure selection on sperm morphology, we used multiple regression to estimate the relationships between relative fitness (measured as fertilization success) and standardized morphological traits (Lande and Arnold 1983). To begin with, we conducted an overall, formal test of whether selection on sperm morphology (tail length, head length, and head width) varied among sperm environments. We compared models in which selection coefficients varied with sperm age and concentration to models in which selection was assumed to be constant across these environments. Following the approach outlined by Chenoweth and Blows (2005), we used a sequence of model comparisons to (1) establish a baseline model to account for environmental effects on relative fertilization success; (2) test whether linear selection gradients systematically differed with sperm age and concentration; and (3) test whether nonlinear selection gradients systematically differed with sperm age and concentration. In this study, the model selection procedure was extended to include tests of variation in selection across multiple environmental attributes, including both discrete (age) and continuous (sperm concentration) descriptors. Full details are described in the Appendix. Because these analyses suggested that selection did systematically vary with sperm environment (see Results section and Appendix), we illustrated these changes in selection by presenting the details of selection analyses in six sperm environments (concentrations of 10^8, 10^6, 10^4 sperm per mL for each of the new and old sperm treatments). This reduced subset of environments was chosen to simplify data presentation. Our overall conclusions are based on analyses of the entire dataset.

In each sperm environment, we used multiple linear regression to estimate standardized selection gradients (β values) for head length, head width, and tail length (traits were standardized as z-scores). We then conducted a second regression that included linear, quadratic, and correlational terms to estimate components of nonlinear selection (i.e., γ values; Lande and Arnold 1983; Phillips and Arnold 1989). Estimated coefficients of the quadratic terms were doubled to produce corresponding estimates of terms in the γ matrix (Lande and Arnold 1983). Used in this way, multiple regression provides a quadratic approximation of the selection surface in multivariate space. The number of dimensions in this space (n) is defined by the number of traits measured (three in our study). Although this method of presentation provides a useful description of selection on each of the measured traits, it does

not describe the full extent of nonlinear selection, particularly when nonlinear selection is strongest on certain combinations of traits (Phillips and Arnold 1989; Blows and Brooks 2003). We also included a canonical rotation of the γ matrix to examine nonlinear selection, including selection on trait combinations, in more detail. Canonical analysis provides a complement to multiple regression by rotating the trait axes to match the major axes of nonlinear selection, as estimated by the γ matrix (Phillips and Arnold 1989). The matrix rotation shifts the frame of reference such that an individual's multivariate phenotype is viewed along n axes that are linear combinations of the original trait values (the new axes are the eigenvectors of the γ matrix). In this form, nonlinear selection is described by n eigenvalues, rather than by $n(n-1)$ coefficients of the γ matrix, and correlational selection is described by the loadings of the original traits onto each canonical axis (Blows 2007). Canonical analysis may therefore provide a more efficient and convenient description of nonlinear selection on the multivariate phenotype as a whole. Examining components of linear selection on the canonical axes can also be useful for detecting directional selection on trait combinations. In our analyses, P values for canonical analyses were calculated using the permutation method described by Reynolds et al. (2010). The combination of linear regression and canonical rotation was used to describe how the selection surfaces changed among different combinations of sperm concentration and age.

It is worth noting that average sperm phenotypes were not known exactly, and that variability in our estimates of average phenotype values will affect our estimates of selection. When predictor variables are subject to measurement error, regression slopes will be biased toward zero (reviewed by McArdle 1988), and within our experiment the magnitude of selection will likely be underestimated. We proceeded to analyze selection using ordinary least squares regression because exact values of selection gradients are relatively unimportant in the context of this laboratory experiment, and because our main question of interest was whether selection differed across fertilization environments. Because sperm from the same set of males was used throughout the experiment, any biases would be constant across environments. Differences in selection gradients across sperm environments will not be biased by variation in estimated phenotype values.

RELATIONSHIPS BETWEEN ADULT AND SPERM PHENOTYPES

The results of our fertilization experiment are best interpreted in the light of natural covariation among sperm characteristics and adult phenotypes. Adult body size and number of sperm produced are two related traits that are expected to have a major influence on the concentration of sperm in the environment. In our experiment, after sperm were extracted and measured, males were preserved in 95% ethanol, blotted dry using filter paper, and weighed to the nearest 0.01 g. Number of sperm produced was calculated as the mean concentration of sperm multiplied by the volume of the sperm / seawater solution collected from each male after they completed spawning. In our analysis we used simple linear regression to examine the relationship between body size and the number of sperm produced by the males in our experiment. Next, we examined the direct relationships between adult characteristics (body size and sperm number) and each of the three aspects of sperm morphology (tail length, head width, and head length). For each morphological trait, we used multiple linear regression to examine the partial relationship between adult characteristics and sperm morphology. Because sperm phenotypes may be correlated, we also included the other two sperm traits as predictor variables in these analyses.

Table 1. Summary statistics describing variation in sperm phenotype values (N = 1200 sperm). Measurements are in micrometers. Coefficients of variation were calculated both within males (CV_W; calculated within males then averaged across males), and between males (CV_B). CVs are expressed as percentages.

Trait	Mean	SD	CV_W	CV_B
Tail length	51.989	4.453	7.5	3.4
Head length	3.568	0.202	5.3	1.7
Head width	2.181	0.155	6.6	2.9

Results

VARIATION IN SPERM MORPHOLOGY

Within the total population, sperm phenotypes exhibited moderate variability (Table 1), with coefficients of variation ([SD / mean] $\times$ 100%) ranging from 5.7% to 8.7%. Average values for the coefficients of variation calculated within males (CV_W) were greater than the coefficients of variation between males (CV_B). Among-male variation contributed significantly to the total variation (ANOVA for tail length: $F_{76,1123} = 3.806$, $P < 2 \times 10^{-16}$; head length: $F_{76,1123} = 2.461$, $P = 3.08 \times 10^{-10}$; head width: $F_{76,1123} = 4.237$, $P < 2 \times 10^{-16}$), and average sperm phenotypes were measured with a reasonable degree of precision (Appendix S1). Sperm phenotypes exhibited varying degrees of correlation. Among the 77 males in our study, mean tail length was not significantly correlated with either mean head length ($r = 0.163$, $P = 0.156$) or mean head width ($r = 0.129$, $P = 0.263$). In contrast, head width and head length were appreciably correlated ($r = 0.328$, $P = 0.004$), suggesting that selection on sperm morphology was best analyzed within a multivariate framework.

SELECTION ON SPERM MORPHOLOGY

In our overall analysis of selection on sperm traits, a model in which linear selection gradients systematically changed with

sperm age and concentration (Appendix S2, eq. S4) provided a much better fit to the data than a model in which linear selection was constant across sperm environments (eq. S3; partial F-test: $F_{6,846} = 2.697$, $P = 0.0134$). In particular, selection on tail length decreased as sperm aged ($\beta_1 = -0.096$, SE $= 0.054$, $P = 0.075$) and increased with the log of sperm concentration ($\beta_2 = 0.0325$, SE $= 0.0153$, $P = 0.034$). Predicted selection gradients on tail length ranged from $\beta = -0.0385$ (SE $= 0.0903$) for old sperm at 10^3 sperm per mL, to $\beta = 0.220$ (SE $= 0.110$) for new sperm at 10^8 sperm per mL. Selection on head length increased as sperm aged ($\beta_1 = 0.090$, SE $= 0.053$, $P = 0.091$) and decreased with the log of sperm concentration ($\beta_2 = -0.048$, SE $=$ 0.0155, $P = 0.004$). Predicted selection gradients on head length ranged from $\beta = 0.148$ (SE $= 0.091$) for old sperm at 10^3 sperm per mL treatment, to $\beta = -0.171$ (SE $= 0.116$) for new sperm at 10^8 sperm per mL. Individual coefficients suggested no significant changes in linear selection on head width across sperm age or concentration (P values were > 0.70). Similar model comparisons suggested little evidence that nonlinear selection changed substantially with sperm age and concentration (partial F-tests for nonlinear selection: $F_{12,828} = 1.428$, $P = 0.147$; see Appendix S2 for full details).

Because selection on sperm traits varied among environments, we examined patterns of selection in greater detail by conducting selection analyses in a subset of six sperm environments (concentrations of 10^8, 10^6, 10^4 for each of the old and new sperm treatments). These analyses confirmed several significant associations between average sperm traits and relative fitness (as measured by relative fertilization success). In high-concentration environments that resulted in high fertilization success (new 10^8, old 10^8, and new 10^6), males whose sperm tails were, on average, longer had higher fertilization success (positive β values; Table 2). In contrast, in the environment with old sperm and a low concentration (old 10^4), average fertilization success was low and relative fitness was strongly associated with sperm head length. When the entire dataset was analyzed, a model including terms for nonlinear selection (constant across environments) fit much better than a model without nonlinear selection (comparing S5 to S4, partial F-test: $F_{6,840} = 5.479$, $P = 1.415 \times 10^{-5}$), although when analyzed on an environment-by-environment basis, few of the nonlinear coefficients were significant at the $\alpha = 0.05$ level. Regression coefficients (γ values) were reasonably consistent across environments and confirm that there was little evidence for systematic changes in nonlinear selection.

As expected, canonical analyses revealed slightly stronger evidence for nonlinear selection. In particular, the negative eigenvalues along the second axes in the old 10^8 and 10^6 treatments indicated convex selection on sperm traits (mainly on head length in old 10^8 and on a combination of all traits in new 10^4 and old 10^6; Table 3). For new sperm at concentrations of 10^8 and 10^6, there was little evidence for nonlinear selection. Rather, the analyses suggested that in these environments sperm that had a combination of long tails and small heads were more successful (linear selection on the first canonical axis was significant in each environment though note the differences in the signs of the coefficients defining the axes; Table 3). Canonical analysis in the oldest, lowest-concentration environment indicated linear selection on a combination of head length and head width. Here the coefficients suggest that sperm with slightly longer, skinnier heads had greater fertilization success (Table 2). Overall, these results suggest that the characteristics of sperm that improve fertilization success under sperm saturated conditions when fertilization success is high are not the same characteristics that confer an advantage under sperm limited conditions when fertilizations are low. Figure 2 illustrates the degree to which the selection surface can change among these sperm environments. Note that because canonical rotation redefines the traits in each environment, we visualized differences in the selection surface among environments by comparing traits on their original axes.

RELATIONSHIPS BETWEEN ADULT AND SPERM PHENOTYPES

On average, larger males produced more sperm (linear regression: coeff. $= 1.87 \times 10^{10}$, SE $= 3.36 \times 10^9$, $P = 5.02 \times 10^{-7}$, $r^2 =$ 0.302), and were therefore more likely to create environments in which fertilization success was high. In addition, both body size and sperm number exhibited direct relationships with some, but not all sperm traits (Fig. 4). Mean tail length was unrelated to both body size ($P = 0.717$) and sperm number ($P = 0.540$; Fig. 4, top row). Head width strongly decreased with adult body size ($P = 0.023$) but was essentially unrelated to sperm number ($P =$ 0.576; Fig. 3, middle row). There was marginal evidence that head length increased with body weight ($P = 0.083$), and strong evidence that head length decreased with sperm number ($P = 0.013$; Fig. 3, bottom row). Overall, these results suggest that individuals that produced relatively few sperm tended to produce sperm with longer heads. Separate from this effect, larger individuals tended to produce sperm that had longer, skinnier heads.

Discussion

Our study demonstrated that selection on sperm morphology can be highly context dependent: the magnitude and form of selection strongly depended on the age and concentration of sperm in the fertilization environment. For example, sperm with longer tails performed better at high sperm concentrations, but tail length was relatively unimportant at low concentrations. Selection on head length exhibited even more variation and actually reversed directions along our age and concentration gradient. Sperm with long heads were selected for when fertilization success was very

Table 2. Summary statistics for multiple regression analyses relating relative fertilization success to sperm morphology in each of six sperm environments. EC is egg concentration (number per mL), TAI is time after the experiment began (in days). Coefficients with *P*-values less than 0.05 are highlighted in bold font.

Sperm conc. = 10^8					Sperm conc. = 10^6					Sperm conc. = 10^4				
	Beta		Gamma			Beta		Gamma			Beta		Gamma	
NEW														
meanTL	**0.16**	0.06			meanTL	**0.23**	−0.08			meanTL	0.11	−0.07		
meanHL	0.00	0.00	−0.08		meanHL	−0.20	−0.07	−0.14		meanHL	0.05	−0.02	−0.18	
meanHW	−0.04	−0.09	0.08	−0.03	meanHW	0.04	0.02	0.12	−0.18	meanHW	0.11	−0.02	0.18	−0.24
	coeff.	*P*				coeff.	*P*				coeff.	*P*		
TAI	0.15	0.00			TAI	0.08	0.03			TAI	0.08	0.06		
EC	0.00	0.00			EC	0.00	0.00			EC	0.00	0.48		
OLD														
meanTL	0.06	0.02			meanTL	0.07	0.02			meanTL	−0.01	−0.14		
meanHL	0.06	0.02	−0.17		meanHL	0.03	−0.08	−0.22		meanHL	**0.28**	0.04	−0.35	
meanHW	−0.01	−0.12	−0.01	−0.12	meanHW	0.00	−0.07	0.03	**−0.23**	meanHW	0.02	−0.01	0.14	−0.15
	coeff.	*P*				coeff.	*P*				coeff.	*P*		
TAI	0.10	0.00			TAI	0.05	0.21			TAI	0.01	0.77		
EC	0.00	0.15			EC	0.00	0.14			EC	0.00	0.48		

low (old sperm treatment at low concentrations), but were selected against when fertilization success was very high (the high concentration environments; see Results and Tables 1, 2). Such reversals in the direction of selection among local reproductive environments can result in a type of disruptive selection when combined within the population at large (e.g., Felsenstein 1976; Sappington and Taylor 1990). Even if there are no strong reversals in the strength of selection, variation in selection that leads to traits being positively associated with fitness in some environments but more-or-less neutral in others can still maintain genetic variation in traits (Fry 1996). Although the basic phenomenon of anisogamy may be largely driven by a trade-off between sperm number and size (Parker et al. 1972), the abundance of within-species variation in sperm morphology may be the result of differential success of sperm morphologies in variable fertilization environments. As broadcast spawning is the ancestral mode of reproduction and how sex evolved in the first place, these effects are likely to be important in shaping the evolution of sperm morphology.

For many broadcast spawners, local spawning environments are likely to be substantially heterogeneous with respect to sperm concentration. Such variation may be in response to several factors, including variation in local population density (e.g., Pennington 1985; Yund 1990; Brazeau and Lasker 1992; Babcock et al. 1992), depth (Levitan 1998), and water flow (Levitan et al. 1992; Petersen et al. 1992, 2001; Babcock and Keesing 1999). Moreover, spawning asynchrony is frequently observed for many taxa (reviewed by Lotterhos and Levitan 2010) implying that eggs of many broadcast spawners routinely encounter sperm of varying

Table 3. Summary of canonical analyses relating relative fertilization success to sperm morphology in each of six sperm environments. Trait axes (eigenvalues of the γ matrix) that were significantly associated with either nonlinear or linear selection (i.e., respective *P*-values were less than 0.05) are highlighted in bold font.

Sperm concentration = 10^8						Sperm concentration = 10^6						Sperm concentration = 10^4					
			Traits						Traits						Traits		
	θ	λ	meanTL	meanHL	meanHW		θ	λ	meanTL	meanHL	meanHW		θ	λ	meanTL	meanHL	meanHW
NEW																	
M1	**0.15**	0.13	**0.79**	**−0.23**	**−0.56**	M1	**−0.26**	−0.01	**−0.57**	**0.70**	**0.43**	M1	−0.04	−0.01	0.49	−0.67	−0.56
M2	−0.07	0.02	−0.55	−0.68	−0.49	M2	0.15	−0.09	0.77	0.27	0.58	M2	0.15	−0.08	0.87	0.38	0.32
M3	0.02	−0.16	0.27	−0.69	0.67	M3	0.09	−0.30	−0.29	−0.66	0.70	M3	0.05	−0.39	0.00	−0.64	0.76
OLD																	
M1	0.06	0.09	0.86	0.10	−0.49	M1	0.06	0.06	0.92	−0.29	−0.25	M1	0.15	−0.08	0.09	0.47	0.88
M2	0.06	**−0.17**	**−0.06**	**0.99**	**0.11**	M2	−0.06	**−0.24**	**−0.38**	**−0.74**	**−0.55**	M2	0.01	−0.14	0.99	0.08	−0.15
M3	0.02	−0.19	0.50	−0.07	0.86	M3	−0.02	−0.26	0.03	−0.60	0.80	M3	**−0.24**	−0.42	**0.14**	**−0.88**	**0.46**

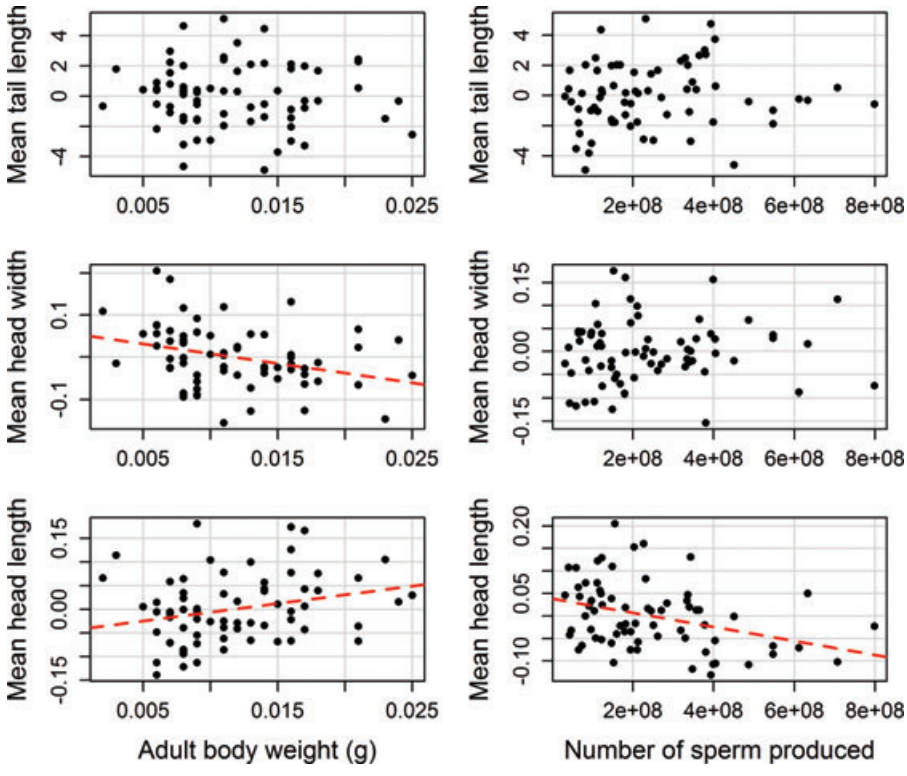

Figure 3. **Partial residual plots displaying direct relationships between adult phenotypes (body weight and number of sperm produced) and sperm morphology (tail length, head width, and head length). Values on the Y-axes are components plus residuals, expressed in micrometers. Dashed lines represent least-squares smoothers and were included if *P* values were near 0.05.**

ages. Natural variation in local fertilization environments, combined with context-specific selection on sperm morphology may therefore produce a spatial and temporal mosaic of selection that contributes to preserving within-species variation in sperm morphology. Although we know of no studies that have examined directly such effects on sperm morphology of broadcast spawners, patterns of context-dependent selection on sea urchin sperm bindin proteins were recently reported by Levitan (2012). In red sea urchins (*Stronglyocentrotus franciscanus*), sperm bindin exhibits a genetic polymorphism in which certain genotypes have a higher affinity for binding to eggs. The high-affinity genotypes have greater relative fertilization success at low sperm concentrations, but when sperm concentrations are high these genotypes are more likely to induce lethal polyspermy, leading to relatively low success (Levitan 2012). Moreover, these differences in performance were suggested to be driving the observed changes in gene frequencies over time as urchins transitioned from a low-to-high density regime (Levitan 2012). Links between variation in sperm morphology and variation in selection have been observed for several groups of internal fertilizers. Comparative studies have revealed that species with a high incidence of sperm competition, and greater consistency in postcopulatory selection for large sperm, exhibited significantly less variation in sperm length (a trait related to sperm velocity and success under competition). Such patterns have been observed for groups of passerine birds (Kleven et al. 2008; Immler et al. 2008), and eusocial ants and bees (Fitzpatrick and Baer 2011).

Data on within-species variation in sperm morphology are rare for broadcast spawners, but available studies suggest that such variation is substantial (e.g., Crean and Marshall 2008; Manier and Palumbi 2008; Fitzpatrick et al. 2010, 2012), similar to the level of variation we observed in this study, and worthy of further evaluation. In this study we observed considerable variability in sperm morphology both within- and among-males. Variation in selection may promote variation among males for all of the reasons described above. Context-dependent selection in an unpredictable environment may also favor within-male variability. Producing variable sperm morphologies may be a way of increasing an individual's geometric mean fitness in variable environments (e.g., Marshall et al. 2008).

MECHANISMS LINKING SPERM MORPHOLOGY AND RELATIVE FERTILIZATION SUCCESS

Our experiment revealed that the characteristics of sperm are that important for successful fertilization systematically varied with sperm concentration and age. At high sperm concentrations, males whose sperm had long tails and smaller heads had relatively high fertilization success. In contrast, at low concentrations, males

whose sperm had longer heads had higher relative fertilization success. We believe that these changes in the morphological correlates of relative fertilization success may indicate shifts in which mechanisms are important for improving fertilization efficacy and subsequent fitness.

Both theoretical and empirical studies suggest that sperm morphology may be tightly linked to swimming speed (reviewed by Humphries et al. 2008). In the high-viscosity environment that sperm experience, swimming speed may be strongly influenced by the balance between thrust, which is thought to be proportional to tail length, and drag, which is thought to be proportional to surface area of the sperm head (Humphries et al. 2008). Indeed, several within-species comparisons have demonstrated positive relationships between tail length and speed (e.g., Mossman et al. 2009; Fitzpatrick et al. 2010) and head shape and speed (Malo et al. 2006). Our selection analyses indicated that, at high sperm concentrations, males that produce sperm with a combination of long tails and small, thin heads had higher success. Although we did not measure swimming speed directly, a previous study on this species demonstrated a strong relationship between swimming speed and fertilization success under lab conditions at similar relative concentrations of eggs and sperm (Kupriyanova and Havenhand 2002). We believe there is a strong possibility that the positive relationship between sperm morphology (long tails and small heads) and relative fertilization success at high densities is mediated by functional links between morphology and speed, but we recognize that long tails and small heads may be associated with other aspects of sperm performance that affect fertilization success (e.g., greater ability to fuse with eggs), although we know of no studies that have examined these effects directly.

At low concentrations, tail length was relatively unimportant, and head length emerged as a trait that was important for fertilization success in the old sperm treatment. Our measurements of head length provide a measure of head size but also reflect the size of both the acrosome and the mitochondrial packet (cf. Fig. 1), two features of sperm morphology that may influence sperm performance (SeGall and Lennarz 1981; Mita et al. 1995). The volume of mitochondria and their associated energy stores have been demonstrated to be important for long-term activity of sperm in broadcast spawning invertebrates (Afzelius and Mohri 1966; Mita et al. 1995), suggesting that sperm with larger mitochondrial packets may live longer and thereby increase their chance of participating in fertilizations. Larger acrosomes could also confer a fertilization advantage. It is expected that very few of the sperm–egg collisions actually result in fertilization (e.g., Vogel et al. 1982). Sperm with larger acrosomes may contain more of the enzymes involved in the fertilization process (e.g., Menkveld et al. 2003) and may be more likely to successfully fuse with an egg during a potentially fertilizing collision. Although we regard both of these mechanisms as possible explanations for the observed relationship between head length and fertilization success, because head length had appreciable effects on fertilization success in the old sperm treatment, we believe that the association between head length and fertilization success was driven by effects on sperm longevity. Along similar lines, Fitzpatrick et al. (2012) examined selection on ejaculate traits for broadcast-spawning mussels at relatively low sperm concentrations. Their results also suggested that traits associated with sperm longevity (and other traits that increased the probability of sperm encountering eggs) increased relative fertilization success under sperm-limited conditions.

The overall patterns of selection we observed share some similarities with Levitan's (1993) study of sperm characteristics and fertilization success in broadcast-spawning sea urchins. Levitan (1993) compared sperm characteristics among three congeneric sea urchins that differed substantially with respect to their natural population densities. Of the three species, green urchins (*Stronglyocentrotus droebachiensis*) were the species with the lowest population density, were most often sperm-limited, and had sperm with elongated heads and high longevity. In contrast, purple urchins (*S. purpuratus*) were highly abundant, often sperm-saturated, and had sperm with rounded heads and high swimming speed. Red urchins (*S. franciscanus*) were intermediate with respect to all three characteristics. These among-species comparisons suggest that concentration-dependent selection like we observed in our study may also contribute to variation in sperm morphology among species.

RELATIONSHIPS BETWEEN ADULT AND SPERM PHENOTYPES

Variation in the concentration of sperm in the local fertilization environment is a ubiquitous feature for many broadcast spawners (Levitan 1998). Consequently, one might expect that, if possible, adjusting sperm phenotype to increase fertilization success within the local environment would confer a substantial fitness advantage. Indeed, recent studies have revealed considerable plasticity in gamete morphology (Crean and Marshall 2008; Luttikhuizen et al. 2011). For example, Crean and Marshall (2008) demonstrated that ascidians placed in high-density environments (i.e., those likely to experience greater sperm competition) produced larger, more motile sperm, which were expected to perform better in high competition environments. In this study, we observed a similar phenomenon. We found that total number of sperm produced was correlated with sperm morphology. This pattern was more subtle than a size-number tradeoff. Rather, males that produced relatively few sperm produced sperm with longer heads. We also observed phenotypic correlations between adult mass and both the average width and length of sperm heads. Separate from the effects of sperm number, adults with larger body sizes tended to produce sperm with longer, skinnier heads. Interestingly, our selection analyses suggest that a negative covariance

between sperm head length and sperm number is adaptive—males that are most likely to produce high sperm concentration environments are predicted to have higher fitness if they produce sperm with shorter heads. We hypothesize that this pattern may represent an adaptive, plastic response to endogenous cues that signal a high probability of sperm saturation in the local fertilization environment, although we acknowledge that an unequivocal demonstration of this phenomenon would require further study. Such adaptive, transgenerational plasticity whereby parents use their own phenotype as a cue for the environment their offspring are likely to experience is predicted by theory and supported by data (e.g., Begon and Parker 1986; Plaistow et al. 2006). Thus far however, most considerations of such plasticity have been restricted to maternal effects and offspring size only; our results suggest that similar, adaptive paternal effects could also be possible.

Conclusions

Overall, our results suggest that there is no single, best, sperm phenotype. Fertilization success was linked to morphological traits, but sperm morphology that confers a fitness advantage in one set of environmental conditions may be unimportant, or even at a disadvantage in others. Variation among local sperm environments may generate disruptive selection in the population at large, and for many broadcast spawners, this may be a persistent force that maintains variability in sperm morphology. In this study we examined a range of sperm concentrations and ages in the lab. In nature, where fertilization environments are affected by additional factors, including turbulence and advective flow (Denny and Shibata 1989), sperm environments are likely to be even more variable (reviewed by Levitan 1998) and may result in even stronger variation in selection. Equally important to understanding this variation in selection is estimating the natural distribution of sperm environments in the field (Lotterhos and Levitan 2010). Future studies that evaluate both of these phenomena will go a long way toward understanding how selection within a mosaic of local environments combines to ultimately shape the evolution of sperm morphology.

ACKNOWLEDGMENTS

We are extremely grateful to Hannah Ritchie for her tremendous assistance in the lab and field. J. Evans, J. Fitzpatrick, and D. Levitan provided helpful comments on the article. This work was conducted while D. W. Johnson was a Postdoctoral Associate at the National Center for Ecological Analysis and Synthesis, a Center funded by National Science Foundation (Grant No. EF-0553768), the University of California, Santa Barbara, and the State of California. D. J. Marshall and K. Monro were supported by funding from the Australian Research Council.

LITERATURE CITED

Afzelius, B. A., and H. Mohri. 1966. Mitochondria respiring without exogenous substrate: a study of aged sea urchin spermatazoa. Exp. Cell Res. 42:10–17.

Anderson, M. J., J. Nyholt, and A. F. Dixson. 2005. Sperm competition and the evolution of sperm midpiece volume in mammals. J. Zool. 267: 135–142.

Babcock, R., and J. Keesing. 1999. Fertilization biology of the abalone *Haliotis laevigata*: laboratory and field studies. Can. J. Fisheries Aquat. Sci. 56:1668–1678.

Babcock, R., C. Mundy, J. Keesing, and J. Oliver. 1992. Predictable and unpredictable spawning events: in situ behavioral data from free-spawning coral reef invertebrates. Invertebrate Reprod. Develop. 22:213–228.

Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philos. Trans. R. Soc. B-Biol. Sci. 365:87–97.

Benzie, J. A. H., and P. Dixon. 1994. The effects of sperm concentration, sperm:egg ratio, and gamete age on fertilization success in crown-of-thorns starfish (*Acanthaster planci*). Biol. Bull. 186:139–152.

Benzie, J. A. H., K. P. Black, P. J. Moran, and P. Dixon. 1994. Small-scale dispersion of eggs and sperm of the crown-of-thorns starfish (*Acanthaster planci*) in a shallow coral-reef habitat. Biol. Bull. 186:153–167.

Birkhead, T. R., J. G. Martinez, T. Burke, and D. P. Froman. 1999. Sperm mobility determines the outcome of sperm competition in the domestic fowl. Proc. R. Soc. Lond. Series B-Biol. Sci. 266:1759–1764.

Blows, M. W. 2007. A tale of two matrices: multivariate approaches in evolutionary biology. J. Evolut. Biol. 20:1–8.

Blows, M. W., and R. Brooks. 2003. Measuring nonlinear selection. Amer. Nat. 162:815–820.

Brazeau, D. A., and H. R. Lasker. 1992. Reproductive success in the caribbean octocoral *Briareum asbestinum*. Marine Biol. 114:157–163.

Byrne, P. G., L. W. Simmons, and J. D. Roberts. 2003. Sperm competition and the evolution of gamete morphology in frogs. Proc. Roy. Soc. Lond. Series B-Biol. Sci. 270:2079–2086.

Calhim, S., M. C. Double, N. Margraf, T. R. Birkhead, and A. Cockburn. 2011. Maintenance of sperm variation in a highly promiscuous wild bird. PLoS One 6:e28809. doi:10.1371/journal.pone.0028809

Chenoweth, S. F., and M. W. Blows. 2005. Contrasting mutual sexual selection on homologous signal traits in *Drosophila serrata*. Amer. Nat. 165: 281–289.

Crean, A. J., and D. J. Marshall. 2008. Gamete plasticity in a broadcast spawning marine invertebrate. Proc. Natl. Acad. Sci. USA 105:13508–13513.

Denny, M. W., and M. F. Shibata. 1989. Consequences of surf-zone turbulence for settlement and external fertilization. Amer. Nat. 134:859–889.

Dusenbery, D. B. 2000. Selection for high gamete encounter rates explains the success of male and female mating types. J. Theor. Biol. 202:1–10.

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

Evans, J. P. 2011. Patterns of genetic variation and covariation in ejaculate traits reveal potential evolutionary constraints in guppies. Heredity 106: 869–875.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman, Essex.

Felsenstein, J. 1976. Theoretical population genetics of variable selection and migration. Ann. Rev. Genet. 10:253–280.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, London, U. K.

Fitzpatrick, J. L., and B. Baer. 2011. Polyandry reduces sperm length variation in social insects. Evolution 65:3006–3012.

Fitzpatrick, J. L., R. Montgomerie, J. K. Desjardins, K. A. Stiver, N. Kolm, and S. Balshine. 2009. Female promiscuity promotes the evolution of faster sperm in cichlid fishes. Proc. Natl. Acad. Sci. USA 106:1128–1132.

Fitzpatrick, J. L., F. Garcia-Gonzalez, and J. P. Evans. 2010. Linking sperm length and velocity: the importance of intramale variation. Biol. Lett. 6:797–799.

Fitzpatrick, J. L., L. W. Simmons, and J. P. Evans. 2012. Complex patterns of multivariate selection on the ejaculate of a broadcast spawning marine invertebrate. Evolution. 66:2451–2460.

Fry, J. D. 1996. The evolution of host specialization: are trade-offs overrated? Amer. Nat. 148:S84–S107.

Gage, M. J. G. 1994. Associations between body size, mating pattern, testis size and sperm lengths across butterflies. Proc. R. Soc. Lond. Series B-Biol. Sci. 258:247–254.

Gage, M. J. G., C. P. Macfarlane, S. Yeates, R. G. Ward, J. B. Searle, and G. A. Parker. 2004. Spermatozoal traits and sperm competition in Atlantic salmon: Relative sperm velocity is the primary determinant of fertilization success. Curr. Biol. 14:44–47.

Gomendio, M., and E. R. S. Roldan. 1991. Sperm comptetition influences sperm size in mammals. Proc. R. Soc. Lond. Series B-Biol. Sci. 243: 181–185.

Humphries, S., J. P. Evans, and L. W. Simmons. 2008. Sperm competition: linking form to function. BMC Evolut. Biol. 8:319.

Immler, S., S. Calhim, and T. R. Birkhead. 2008. Increased postcopulatory sexual selection reduces the intramale variation in sperm design. Evolution 62:1538–1543.

Immler, S., S. R. Pryke, T. R. Birkhead, and S. C. Griffith. 2010. Pronounced within-individual plasticity in sperm morphometry across social environments. Evolution 64:1634–1643.

Johnson, S. L., and P. O. Yund. 2004. Remarkable longevity of dilute sperm in a free-spawning colonial ascidian. Biol. Bull. 206:144–151.

Hollows, C. F., E. L. Johnston, and D. J. Marshall. 2007. Copper reduces fertilisation success and exacerbates Alee effects in the field. Marine Ecol. Progress Series 333:51–60.

Kalmus, H. 1932. Uber den Erhaltungswert der phanotypischen (morphologischen) anisogamie und die Entstehung der ersten geschlechtsunterschiede. Biol. Zentralbl. 52:716–736.

Kleven, O., T. Laskemoen, F. Fossoy, R. J. Robertson, and J. T. Lifjeld. 2008. Intraspecific variation in sperm length is negatively related to sperm competition in passerine birds. Evolution 62:494–499.

Kupriyanova, E., and J. N. Havenhand. 2002. Variation in sperm swimming behaviour and its effect on fertilization success in the serpulid polychaete *Galeolaria caespitosa*. Invertebrate Reprod. Develop. 41:21–26.

Kupriyanova, E. K. 2006. Fertilization success in *Galeolaria caespitosa* (Polychaeta : Serpulidae): gamete characteristics, role of sperm dilution, gamete age, and contact time. Sci. Marina 70:309–317.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lehtonen, J., and H. Kokko. 2011. Two roads to two sexes: unifying gamete competition and gamete limitation in a single model of anisogamy evolution. Behav. Ecol. Sociobiol. 65:445–459.

Levitan, D. R. 1991. Influence of body size and population density on fertilization success and reproductive output in a free-spawning invertebrate. Biol. Bull. 181:261–268.

Levitan, D. R. 1993. The importance of sperm limitation to the evolution of egg size in marine invertebrates. Amer. Nat. 141:517–536.

———. 1998. Sperm limitation, gamete competition, and sexual selection in external fertilizers. Pp. 173–215 *in* T. R. Birkhead, and A. P. Moller, eds. Sperm competition and sexual selection. Academic Press, San Diego, CA.

———. 2000. Sperm velocity and longevity trade off each other and influence fertilization in the sea urchin *Lytechinus variegatus*. Proc. R. Soc. Lond. Series B-Biol. Sci. 267:531–534.

———. 2004. Density-dependent sexual selection in external fertilizers: variances in male and female fertilization success along the continuum from sperm limitation to sexual conflict in the sea urchin Stronglyocentrotus franciscanus. Amer. Nat. 164:298–309.

———. 2012. Contemporary evolution of sea urchin gamete-recognition proteins: experimental evidence of density-dependent gamete performance predicts shifts in allele frequencies over time. Evolution 66:1722–1736.

Levitan, D. R., M. A. Sewell, and F. S. Chia. 1992. How distribution and abundance influence fertilization success in the sea urchin *Stronglyocentrotus franciscanus*. Ecology 73:248–254.

Lotterhos, K., and D. R. Levitan. 2010. Gamete release and spawning behavior in broadcast spawning marine invertebrates. Pp. 99–120 *in* J. Leonard, ed. The evolution of primary sexual characters. Oxford Univ. Press, Oxford, U. K.

Luttikhuizen, P. C., P. J. C. Honkoop, and J. Drent. 2011. Intraspecific egg size variation and sperm limitation in the broadcast spawning bivalve *Macoma balthica*. J. Exp. Marine Biol. Ecol. 396:156–161.

Malo, A. F., J. J. Garde, A. J. Soler, A. J. Garcia, M. Gomendio, and E. R. S. Roldan. 2005. Male fertility in natural populations of red deer is determined by sperm velocity and the proportion of normal spermatozoa. Biol. Reprod. 72:822–829.

Malo, A. F., M. Gomendio, J. Garde, B. Lang-Lenton, A. J. Soler, and E. R. S. Roldan. 2006. Sperm design and sperm function. Biol. Lett. 2:246–249.

Manier, M. K., and S. R. Palumbi. 2008. Intraspecific divergence in sperm morphology of the green sea urchin, *Strongylocentrotus droebachiensis*: implications for selection in broadcast spawners. BMC Evolut. Biol. 8:283–296.

Marshall, D. J., R. Bonduriansky, and L. F. Bussiere. 2008 Offspring size variation within broods as a bet-hedging strategy in unpredictable environments. Ecology 89:2506–2517.

Menkveld, R., Y. El-Garem, W. B. Schill, and R. Henkel. 2003. Relationship between human sperm morphology and acrosomal function. J. Assist. Reprod. Genet. 20:432–438.

Minchinton, T. E. 1997. Life on the edge: conspecific attraction and recruitment of populations to disturbed habitats. Oecologia 111:45–52.

Mita, M., I. Yasumasu, and M. Nakamura. 1995. Endogenous substrate for energy metabolism and ultrastructural correlates in spermatozoa of the sea urchin *Diadema setosum*. Mol. Reprod. Develop. 40:103–109.

Morrow, E. H., and M. J. G. Gage. 2001. Consistent significant variation between individual males in spermatozoal morphometry. J. Zool. 254: 147–153.

Mossman, J., J. Slate, S. Humphries, and T. Birkhead. 2009. Sperm morphology and velocity are genetically codetermined in the zebra finch. Evolution 63:2730–2737.

Ota, K., D. Heg, M. Hori, and M. Kohda. 2010. Sperm phenotypic plasticity in a cichlid: a territorial male's counterstrategy to spawning takeover. Behav. Ecol. 21:1293–1300.

Parker, G. A. 1982. Why are there so many tiny sperm? Sperm competition and the maintenance of two sexes J. Theor. Biol. 96:281–294.

Parker, G. A., R. R. Baker, and V. G. F. Smith. 1972. The origin and evolution of gamete dimorphism and the male-female phenomenon. J. Theor. Biol. 36:529–553.

Parker, G. A., and M. Begon. 1986. Optimal egg size and clutch size. Effects of environment and maternal phenotype. Amer. Nat. 128:573–592.

Pennington, J. T. 1985. The ecology of fertilization of echinoid eggs: The consequences of sperm dilution, adult aggregation, and synchronous spawning. Biol. Bull. 169:417–430.

Petersen, C. W., R. R. Warner, S. Cohen, H. C. Hess, and A. T. Sewell. 1992. Variable pelagic fertilization success: implications for mate choice and spatial patterns of mating. Ecology 73:391–401.

Petersen, C. W., R. R. Warner, D. Y. Shapiro, and A. Marconato. 2001. Components of fertilization success in the bluehead wrasse, *Thalassoma bifasciatum*. Behav. Ecol. 12:237–245.

Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. Evolution 43:1209–1222.

Pizzari, T., R. Dean, A. Pacey, H. Moore, and M. B. Bonsall. 2008. The evolutionary ecology of pre- and post-meiotic sperm senescence. Trends Ecol. Evolut. 23:131–140.

Plaistow, S. J., C. T. Lapsley, and T. G. Benton. 2006. Context-dependent intergenerational effects: The interaction between past and present environments and its effect on population dynamics. Amer. Nat. 167: 206–215.

Randerson, J. P., and L. D. Hurst. 2001. The uncertain evolution of the sexes. Trends Ecol. Evolut. 16:571–579.

Reynolds, R. J., D. K. Childers, and N. M. Pajewski. 2010. The distribution and hypothesis testing of eigenvalues from the canonical analysis of the gamma matrix of quadratic and correlational selection gradients. Evolution 64:1076–1085.

Sappington, T. W., and O. R. Taylor. 1990. Disruptive sexual selection in *Colias eurytheme* butterflies. Proc. Natl. Acad. Sci. USA 87: 6132–6135.

Scudo, F. M. 1967. Adaptive value of sexual dimorphism: I. Anisogamy. Evolution 21:285–291.

Segall, G. K., and W. J. Lennarz. 1981. Jelly coat and induction of the acrosome reaction in echinoid sperm. Develop. Biol. 86:87–93.

Simmons, L. W., and J. S. Kotiaho. 2002. Evolution of ejaculates: patterns of phenotypic and genotypic variation and condition dependence in sperm competition traits. Evolution 56:1622–1631.

Snook, R. R. 2005. Sperm in competition: not playing by the numbers. Trends Ecol. Evolut. 20:46–53.

Vogel, H., G. Czihak, P. Chang, and W. Wolf. 1982. Fertilization kinetics of sea urchin eggs. Math. Biosci. 58:189–216.

Ward, P. I. 1998. Intraspecific variation in sperm size characters. Heredity 80:655–659.

Williams, M. E., and M. G. Bentley. 2002. Fertilization success in marine invertebrates: the influence of gamete age. Biol. Bull, 202: 34–42.

Williams, M. E., M. G. Bentley, and J. D. Hardege. 1997. Assessment of field fertilization success in the infaunal polychaete Arenicola marina (L). Invertebrate Reprod. Develop. 31:189–197.

Yund, P. O. 1990. An in situ measurement of sperm dispersal in a colonial marine hydroid. J. Exp. Zool. 253:102–106.

Associate Editor: M. Hart

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Variation in sperm phenotypes

Appendix S2. Testing for overall variation in selection

Figure S1. Variation in mean sperm phenotypes.

Table S1. ANOVA summaries for among-male variation in sperm phenotypes.