Multilevel Selection on Offspring Size and the Maintenance of Variation

Hayley Cameron,[1],* Darren W. Johnson,[2] Keyne Monro,[1] and Dustin J. Marshall[1]

1. Centre for Geometric Biology, School of Biological Sciences, Monash University, Melbourne, Victoria 3800, Australia;   2. Department of Biological Sciences, California State University, Long Beach, California 90840

ABSTRACT: Multilevel selection on offspring size occurs when offspring fitness depends on both absolute size (hard selection) and size relative to neighbors (soft selection). We examined multilevel selection on egg size at two biological scales—within clutches and among clutches from different females—using an external fertilizing tube worm. We exposed clutches of eggs to two sperm environments (limiting and saturating) and measured their fertilization success. We then modeled environmental (sperm-dependent) differences in hard and soft selection on individual eggs as well as selection on clutch-level traits (means and variances). Hard and soft selection differed in strength and form depending on sperm availability—hard selection was consistently stabilizing; soft selection was directional and favored eggs relatively larger (sperm limitation) or smaller (sperm saturation) than the clutch mean. At the clutch level, selection on mean egg size was largely concave, while selection on within-clutch variance was weak but generally negative—although some correlational selection occurred between these two traits. Importantly, we found that the optimal clutch mean egg size differed for mothers and offspring, suggesting some antagonism between the levels of selection. We thus identify several pathways that may maintain offspring size variation: environmentally (sperm-)dependent soft selection, antagonistic multilevel selection, and correlational selection on clutch means and variances. Multilevel approaches are powerful but seldom-used tools for studies of offspring size, and we encourage their future use.

Keywords: broadcast spawner, contextual analysis, frequency-dependent selection, group selection, maternal provisioning, parent-offspring conflict.

Introduction

Life-history theory predicts that mothers maximize their fitness by optimizing the trade-off between offspring size and number. Larger offspring generally perform better, but smaller offspring are cheaper to make, so mothers can produce them in greater numbers (Smith and Fretwell 1974). Mothers therefore balance the relative benefits of better-performing offspring with those of increased fecundity, and this balance depends on the relationship between offspring size and performance. Foundational theory thus predicts that a single offspring size will be optimal in any given environment—in other words, selection on offspring size within a population is stabilizing and, if persistent, should act to erode variance over time. In nature, however, offspring sizes vary remarkably across time, space, and all scales of organization (reviewed in Marshall et al. 2018). For example, offspring size often varies with maternal traits, such as size and age—larger, older mothers typically produce larger offspring (Parker and Begon 1986). Even within the same reproductive bout (i.e., clutch, litter, brood, etc.), a given parent may provision one of its offspring with fivefold more resources than another (Lips 2001; Turnbull et al. 2006). Yet the drivers of offspring size variation across these biological scales are often unclear.

We now recognize that multiple processes can maintain variation in offspring size. For example, environmental variation may select for different optima in space and time (reviewed in Marshall et al. 2018). Unpredictable environments may also select for diversified bet-hedging strategies, whereby mothers produce offspring of variable size so that at least some of these offspring are suited to the prevailing conditions (Marshall et al. 2008). Interestingly, however, most branches of theory model offspring fitness as a function of absolute offspring size (Smith and Fretwell 1974)—but fitness may also depend on the size of an offspring relative to its neighbors. In other words, offspring size may be under both hard selection, which is frequency independent and determined by absolute phenotypic values, and soft selection, which is determined by relative phenotypic values and thus depends on the frequency of other phenotypes in the population (Wallace 1975). As such, some components of offspring fitness (e.g., trade-offs with fecundity: Smith and Fretwell 1974;

some types of predation: Hambright 1991) may depend on absolute offspring size, while other components (e.g., competition: Marshall et al. 2006; Cameron and Marshall 2019) may depend on relative size. Although many studies have estimated overall patterns of selection on offspring size, few have formally partitioned fitness into components of hard and soft selection (but see Johnson et al. 2011).

Multilevel selection theory offers a powerful but seldom-applied framework that can be used to estimate hard and soft selection on offspring size. Multilevel selection occurs in hierarchical-structured populations when individual- and group-level traits affect fitness simultaneously (Heisler and Damuth 1987; Goodnight et al. 1992). Under this definition, soft selection can be considered a form of multilevel selection because fitness depends in part on a group-level trait—that is, the phenotype of an individual relative to the mean of the group to which it belongs (Goodnight et al. 1992). Contextual analysis can be used to partition trait-fitness relationships into components of hard and soft selection by including both absolute and relative phenotypic values in models of multilevel selection (Heisler and Damuth 1987; Goodnight et al. 1992; Johnson et al. 2011). Under multilevel selection, selection at the level of groups may oppose or augment that at the level of individuals, and populations can evolve in response to selection operating at both levels simultaneously (Wade 1977; Goodnight et al. 1992; Agrawal et al. 2001). For example, antagonism between group and individual selection may reduce the efficacy of selection at any one level, providing a pathway for the maintenance of phenotypic variation (Goodnight et al. 1992; Weinig et al. 2007). Despite the importance of these effects, however, most studies of offspring size focus on a single level of selection alone (e.g., see data in Rollinson and Rowe 2015), while explicit considerations of multilevel selection are much rarer (but see Reed et al. 2009; Johnson et al. 2011; Bouwhuis et al. 2015).

There are good reasons to suspect that multilevel selection on offspring size is widespread. Certainly, offspring sizes are hierarchically structured across levels of organization—variation occurs within clutches of the same female, among females, and even among local populations separated in time and space—and selection may operate at one or more of these levels simultaneously (Bouwhuis et al. 2015). For many species, offspring also aggregate across scales where interactions are more likely within, rather than among, family groups (e.g., in amphibians: Blaustein and O'Hara 1987; plants: Cheplick 1993; Donohue 2003; fish: Selkoe et al. 2006; birds: Reed et al. 2009; marine invertebrates: Kamel et al. 2010). As such, clutch-level traits, such as the mean offspring size, degree of within-clutch variation, and number of offspring within a clutch, may alter selection on individuals (Parker and Begon 1986; Plaistow et al. 2007; Cameron et al. 2016, 2017). These

clutch-level traits not only form important components of the offspring environment but can also have a heritable basis and may in turn evolve themselves (Moore et al. 1997). Explicit considerations of multilevel selection thus represent an important next step for understanding the processes that maintain offspring size variation across biological scales.

Here, we focus specifically on whether multilevel selection acts on egg size during external fertilization—the predominate mode of reproduction in the sea (Monro and Marshall 2015). During external fertilization, parents spawn their gametes directly into the environment, where they must then meet and fuse. External fertilization is the ancestral mode of reproduction and thus has long been the subject of theory in regard to gamete size evolution (Parker et al. 1972; Bode and Marshall 2007; Henshaw et al. 2014; Parker and Lehtonen 2014; Lehtonen 2016). Traditionally, such theory has considered selection on absolute (rather than relative) egg size (Parker et al. 1972; Bode and Marshall 2007; Henshaw et al. 2014; Parker and Lehtonen 2014; Lehtonen 2016), but the phenotypic context of neighbors may also influence the relative fitness of eggs. Certainly, eggs compete for sperm during external fertilization (Marshall et al. 2004; Marshall and Evans 2005b; Okamoto 2016) and may do so in size-dependent ways. Larger eggs are more attractive to sperm, and because sperm adhere to the surface of eggs on contact (Vogel et al. 1982), they may reduce the local concentration of sperm available for nearby eggs (Marshall et al. 2004; Marshall and Evans 2005b; Okamoto 2016). Eggs can also remain concentrated near females for extended periods once they have been released (Yund and Meidel 2003)—generating the possibility for soft selection to occur within family groups. In other words, the probability that a focal egg becomes fertilized may depend on its size relative to other eggs in the clutch. Patterns of hard and soft selection within clutches might also combine to influence selection on egg size traits at the clutch (group) level (Goodnight et al. 1992), although these hypotheses remain to be tested.

Here, we provide the first exploration of multilevel selection on egg size during external fertilization using a model marine tube worm. We used a split-clutch design to conduct in vitro fertilizations, where we orthogonally exposed clutches of eggs to sperm-limiting and sperm-saturating conditions. We then measured the size-specific fertilization success of eggs within these clutches—a major component of fitness, especially in the context of selection on gamete traits (Otto et al. 2015; Immler et al. 2018). Using our data, we then modeled multilevel selection on egg size in an analogous approach to contextual analysis, whereby we modeled fertilization success as a function of both absolute egg size and egg size relative to the clutch

mean (see Johnson et al. 2011). This approach allowed us to simultaneously explore whether sperm availability altered the strength and form of hard and soft selection on individual eggs, as well as selection on clutch-level egg size traits (clutch means and variances).

Material and Methods

Study Species and Specimen Collection

Galeolaria caespitosa (henceforth, *Galeolaria*) is an intertidal, broadcast-spawning serpulid polychaete from southern and eastern Australia. Individuals are dioecious and live in calcareous tubes that occur as mixed-sex clusters. *Galeolaria* is extensively used as a model for studies of gamete size evolution (Marshall and Evans 2005*b*; Johnson et al. 2013; Monro and Marshall 2016). The overall pattern of selection on egg size (as for other external fertilizers) is typically concave—smaller eggs are smaller targets and encounter sperm less frequently; larger eggs have a greater risk of unviable fertilizations by multiple sperm (Levitan 1996; Styan 1998; Marshall et al. 2002). Sperm availability also shifts optimal values of egg size—when sperm are highly abundant, directional selection typically favors smaller eggs that better avoid polyspermy, whereas sperm limitation selects for larger eggs that better attract sperm. Given these general patterns of selection, sperm availability may alter the relative strength and direction of hard and soft selection on eggs, and we explicitly test this prediction here.

We collected clusters of *Galeolaria* from St. Leonards Pier, Victoria, Australia (38°17′S, 144°71′E), from May to June 2017. We stored these clusters at 18°C in aerated, unfiltered seawater for up to 5 days. During this holding period, we fed the clusters commercially prepared phytoplankton (Seachem Reef Phytoplankton) and changed their water every 2 days. Importantly, egg sizes varied both within clutches and among females in our study population, and both levels of organization contributed equally to the total variation in egg size observed (app. S1; apps. S1, S2 are available online). In addition, maternal size was not strongly correlated with either of our clutch-level traits (means and variances in egg size; app. S1); thus, maternal size does not influence the effects we observe and model.

Experimental Design

To estimate multilevel selection on egg size, we conducted in vitro fertilizations on natural clutches of eggs spawned from multiple females. We used a split-clutch design, whereby each female's clutch was divided into four subsamples that we orthogonally exposed to two levels of sperm availability: two subsamples were exposed to a low sperm concentration (10^4 sperm mL^{-1}), while the other two were exposed to a high sperm concentration (10^5 sperm mL^{-1}). These sperm concentrations yield fertilization rates consistent with sperm limitation (~30%) and saturation (~90%), respectively, and are well within natural ranges observed in the field (Levitan and Petersen 1995). Importantly, our split-clutch design provides estimates of multilevel selection between sperm environments while controlling for differences in maternal genotypes and phenotypes (i.e., clutch means and variances).

To obtain gametes for our experiments, we induced spawning using standard techniques (Marshall and Evans 2005*b*; Johnson et al. 2013; Monro and Marshall 2016). Briefly, we removed the worms from their tubes and placed them in individual petri dishes with filtered seawater (0.22 μm)—methods that induce stress and cause the worms to release their stored gametes (identifiable as orange eggs and white sperm). Before splitting each female's clutch, we standardized the concentration of eggs among females (2.5×10^3 eggs mL^{-1}) to control for differences in maternal fecundity. We prepared our sperm treatments by pooling ejaculates from 10 individual males to avoid compatibility effects (Marshall and Evans 2005*a*). We standardized the concentration of sperm from each male before sperm mixing to ensure that each male was equally represented. We then performed a series of 10-fold dilutions on our pooled solution of sperm to achieve our desired experimental concentrations (10^4 and 10^5 sperm mL^{-1})—thus, paternal genetic backgrounds were identical between our sperm treatments. To avoid the effects of gamete aging, eggs and sperm were used within 1.5 h or 30 min after they had been released from the adults, respectively (Kupriyanova 2006).

We conducted the in vitro fertilizations in 48 well plates at 20°C. For each replicate fertilization (egg subsample), we exposed 0.1 mL of eggs (~125 eggs) to 0.9 mL of the desired sperm concentration in a single well. We randomized the position of each subsample of eggs, and the sperm treatments to which they were exposed, across the well plates. To control for errant fertilizations during gamete extraction, we included an additional subsample of eggs (0.1 mL) from each female's clutch, to which we added 0.9 mL of filtered seawater (instead of sperm). If control eggs showed >10% fertilizations, we excluded all replicates from that female's clutch ($n = 2$). Overall, our final data set included clutches from 13 females, which we split between two replicate fertilizations within each of our two sperm environments—yielding a total of 52 replicate fertilizations and ~7,000 eggs.

Measuring Fitness

We measured the size-specific fertilization success (fitness) of eggs within our clutch subsamples 2 h after the eggs were first exposed to sperm. To do this, we photographed each replicate fertilization (well) using a digital camera attached to an inverted compound microscope (Olympus IX73; ×10 magnification). For each replicate fertilization, we scored ~100 eggs as successfully or unsuccessfully fertilized and measured their diameter (μm) using Olympus cellSens Dimension software. We scored eggs as successfully fertilized if they had regularly cleaved (most were at the two- to eight-cell stage), while unsuccessful eggs were those that had not cleaved or had cleaved irregularly. We note that irregular cleavage can be a sign of polyspermy, but polyspermic eggs may also fail to cleave (Styan 1998)—thus, we could not distinguish between eggs that went unfertilized and those that were polyspermic. We estimated the initial (unfertilized) size of successful and unsuccessful eggs using equations that describe these relationships (initial egg size $= a + bx$; for successfully fertilized eggs (x): $a = 33.24 \pm 3.77$ SEM, $b = 0.46 \pm 0.06$ SEM; for unsuccessful eggs (x): $a = 5.59 \pm 5.13$ SEM, $b = 0.902 \pm 0.08$ SEM). We excluded immature eggs released during strip spawning from our data set. Importantly, our estimates of clutch means after fertilization were comparable to those obtained from photographs of each female's clutch before fertilization ($r^2 = 0.879$, $F_{1,22} = 152.398$, $P < .001$), and these estimates were also consistent between sperm environments ($F_{1,22} = 3.692$, $P = .068$).

Modeling Multilevel Selection

Multilevel selection analyses allowed us to partition fertilization success (our fitness proxy) into components of hard and soft selection. The basic modeling approach is analogous to contextual analysis and is outlined in Johnson et al. (2011), but we expand the description of this technique below.

As in other contextual models of multilevel selection (Goodnight et al. 1992), we modeled hard selection as a component in which individual fitness depends on absolute egg size and soft selection as a component in which fitness depends on relative egg size (as deviations from the clutch mean). Including a component of fitness that depends on deviations from the clutch mean provides a summary of frequency-dependent interactions among eggs within a clutch. If the fitness of a focal individual with phenotype z' is influenced by interactions with other individuals with phenotype z and those interactions are a function (ε) of the differences in phenotypes, then the expected fitness of phenotype z' can be calculated by adding the effects of all pairwise interactions (which are dependent on the differences in phenotypes), weighted by the frequency of other phenotypes in the population, $p(z)$:

$$\int \varepsilon(z' - z)p(z)dz.$$

In this study, we envision $\varepsilon(z' - z)$ to represent eggs competing for sperm, since the relative ability to attract sperm will depend on egg size (Vogel et al. 1982). If ε is a linear function, then the net effect on a focal individual can be summarized as follows:

$$\int \varepsilon(z' - z)p(z)dz = \varepsilon\left(z'\int p(z)dz - \int zp(z)dz\right)$$
$$= \varepsilon(z' - \bar{z}).$$

In our study, we describe the logit of fertilization success as a linear function of relative phenotypic value ($z' - \bar{z}$), with notation simplified to ($z - \bar{z}$) below. Note that this approach could be modified to include a component of fitness that depends on both the frequency ($p(z)$) and density (N) of eggs explicitly; for example,

$$\int \varepsilon(z' - z)Np(z)dz = \varepsilon N(z' - \bar{z}).$$

However, our experimental design held the number of eggs (N) constant among trials (recall that egg concentrations were standardized among females), and thus the results we present describe frequency-dependent selection at a constant density of eggs. In general, competition for sperm can depend on egg density in external fertilizers, although this does not necessarily result in negative effects—for example, eggs at high densities can actually shield each other from polyspermy when sperm are oversaturating (Okamoto 2016). Interactions among eggs are thus likely to depend on both the frequency and density of surrounding eggs (i.e., soft selection sensu stricto; Wallace 1975), and future experiments could explore these processes simultaneously.

Because both hard and soft selection can take several functional forms (i.e., directional, concave, or convex), we first fitted relatively flexible and complex models to the data and compared their fit with simplified models. We evaluated model fits using both Akaike information criterion (AIC) and corrected AIC (AICc), where degrees of freedom were adjusted to represent the number of individual eggs ($n = 7,008$) or the number of females ($n = 13$) in our data set, respectively. This conservative approach accounts for the fact that our unit of replication is somewhat ambiguous given the hierarchical nature of our study (i.e., the number of eggs when predicating fertilization success based on egg size and the number of females when inferring selection on clutch means and variances; see below). In both environments, hard and

soft selection were best represented by quadratic and linear logistic functions, respectively (app. S2). As such, we modeled the size-specific fertilization success (fitness) of eggs as

$$W(z) = \left[\frac{1}{1 + e^{-(\alpha + \gamma(z-\beta)^2)}}\right] \times \left[\frac{1}{1 + e^{-(\delta + \varepsilon(z-\bar{z}))}}\right].$$

The first term represents hard selection and describes the expected (average) probability that an egg of a given size (z) is fertilized, regardless of the frequency of other egg sizes in the clutch. The second term represents soft selection and describes how the expected probability of fertilization is modified by the size of a focal egg relative to the mean egg size of the clutch ($\bar{z}$). The term $W(z)$ was fitted to the data via maximum likelihood estimation, and parameter values were estimated using the package bbmle in R (Bolker and R Development Core Team 2020).

Using the above model for $W(z)$, we compared models in which hard and soft selection components were either allowed to vary or remained constant between sperm environments. For the model that allowed selection to vary, each parameter of $W(z)$ was expanded to two components: a baseline parameter and a deviation that depended on sperm environment (e.g., $\alpha_0 + \alpha_1 X$, where X is a dummy variable indicating sperm environment). The fit of this model was compared with the fit of the simpler five-parameter version of $W(z)$ described above, again using both AIC and AICc. Model comparisons showed overwhelming support for a model where both hard and soft selection varied between sperm environments (table 1). Finally, we compared models of $W(z)$ within each sperm environment, whereby $W(z)$ was determined by hard or soft selection alone, or included both selection compo-

nents simultaneously. Crucially, both AIC and AICc provided strong support for models that included both hard and soft selection simultaneously (table 1); thus, we present the selection coefficients from these models (see "Results").

The function $W(z)$ describes the expected fitness of an individual egg, but in a multilevel context, $W(z)$ can be combined with information about phenotypic distributions to describe variation in fitness among groups—in our case, clutches. We thus used our estimates of $W(z)$ to explore selection on clutch-level traits (means and variances) that arise from hard and soft selection on egg size acting within clutches. To do this, we calculated the average fitness (as a proportion of fertilized eggs within the clutch) of a given female in each sperm environment by integrating our estimates of $W(z)$ with phenotypic distributions of egg sizes. For simplicity, we assumed normal egg size distributions for each clutch, with mean and variance specified by ($\bar{z}_i$) and var(z)$_j$, respectively—a valid assumption given that natural clutches from our population were largely normally distributed. We explored the fitness surface of females across a range of plausible values of means (i) and variances (j). In particular, we calculated mean fitness for all combinations of clutch means and variances across the ranges observed in our study population (clutch means: 60–66; variances: 0.2–3.6; each evaluated at intervals of 0.05; app. S1). We calculated the average fitness of a given female for a given mean-variance combination ($\bar{W}_{i,j}$) as

$$\bar{W}_{i,j} = \int p(z|\bar{z}_i, \mathrm{var}(z)_j) W(z) dz,$$

where $p(z|\bar{z}_i, \mathrm{var}(z)_j)$ describes the distribution of egg sizes for a specified mean and variance. Other symbols are described above.

Table 1: Model comparisons for $W(z)$ testing differences in hard selection (HS) and soft selection (SS) among sperm environments, as well as whether selection is best described by models that include components of hard and soft selection alone or in combination within sperm environments

Model	K	AIC	ΔAIC	AICc	ΔAICc
Between environments:					
HS and SS varies	10	5,771.05	.00	5,881.05	.00
HS and SS constant	5	8,009.36	2,238.31	8,017.93	2,136.88
Sperm saturation:					
HS and SS	5	1,828.77	.00	1,837.35	.00
HS	3	1,938.73	109.953	1,941.40	104.05
SS	2	2,227.33	398.554	2,228.53	391.18
Sperm limitation:					
HS and SS	5	3,942.27	.00	3,950.85	.00
HS	3	4,130.65	188.37	4,133.31	182.47
SS	2	4,535.45	593.17	4,536.65	585.80

Note: K shows the number of parameters estimated by each model. Akaike information criterion (AIC) and corrected AIC (AICc) were calculated assuming degrees of freedom were either the number of individual eggs ($n = 7{,}008$) or the number of females ($n = 13$) in our data set, respectively. In all cases, ΔAIC and ΔAICc provide congruent support for models that provide the best fit to our data.

Note that the above model assumes that all mothers produce the same number of offspring, but when considering maternal fitness, one must also account for a trade-off between the number of offspring a mother can produce given the size of those offspring (Smith and Fretwell 1974). We thus incorporated a simple size-number trade-off into our model of ($\bar{W}_{i,j}$). We assumed that all mothers have a constant amount of resources for reproduction (e) and that the number of eggs a mother can produce (n_i) is inversely related to the mean volume of her eggs ($\bar{v}_i$), where $\bar{v}_i$ is a spherical approximation of volume based on the mean egg diameter of a given mother ($\bar{z}_i$):

$$n_i = \frac{e}{\bar{v}_i}.$$

Given this size-number trade-off, we can now calculate the average fitness of a given female for a given mean-variance combination ($\bar{W}_{i,j}$) as the product of egg number (n_i) and the cumulative fertilization success of all eggs within her clutch:

$$\bar{W}_{i,j} = \left[\int p(z|\bar{z}_i, \text{var}(z)_j) W(z) dz \right] n_i.$$

Results

Patterns of Hard and Soft Selection on Individual Eggs

The model that provided the best fit to our data was one that included components of hard (frequency-independent) and soft (frequency-dependent) selection simultaneously, and both components varied as a function of sperm environment (table 1). We thus provide strong evidence that sperm availability alters patterns of multilevel selection on egg size during external fertilization (Wade 1977; Goodnight et al. 1992).

In both sperm environments, hard selection was best described by a three-parameter logistic function (sperm saturation: $\alpha = 3.342 \pm 0.132$ SE, $\beta = 63.79 \pm 0.080$, $\gamma = -0.282 \pm 0.022$; sperm limitation: $\alpha = 0.578 \pm 0.079$, $\beta = 63.76 \pm 0.043$, $\gamma = -0.379 \pm 0.021$; see also app. S2)—suggesting that hard selection is consistently stabilizing but is relatively stronger under sperm limitation (compare the degree of curvature in fig. 1A with that in 1B). Importantly, we found no evidence for directional hard selection in either sperm environment—that is, fitness optima were relatively consistent between sperm environments and did not substantially deviate from the current mean egg size in our source population (63.38 μm $\pm$ 0.12 SE).

Soft selection, on the other hand, was best modeled as a two-parameter logistic function (app. S2)—suggesting that soft selection was, on average, directional (saturation: $\delta = 4.824 \pm 0.397$ SE, $\varepsilon = -1.473 \pm 0.175$; limitation: $\delta = 2.625 \pm 0.387$, $\varepsilon = 1.437 \pm 0.188$). Note that sperm availability altered the direction of soft selection (fig. 1C, 1D). When sperm were saturating, soft selection was (on average) negative, such that eggs relatively smaller than the clutch mean had the highest fertilization success (fig. 1C). In contrast, when sperm were limiting, soft selection was (on average) positive—that is, eggs relatively larger than the clutch mean were favored (fig. 1D).

Overall, combined patterns of hard and soft selection show that fitness surfaces varied with mean egg size. In other words, the relative fitness (fertilization success) of a given egg depended on both its phenotype and the mean phenotype of the clutch to which it belonged (fig. 1E, 1F). Notably, under sperm saturation, soft selection disfavored relatively larger eggs within a clutch, and eggs whose size was near the global average (62–68 μm) had higher relative fitness when they were part of a clutch with a larger mean size (fig. 1E). This occurred because eggs within this size range were closer to or comparatively smaller than the clutch mean. In addition, within a given clutch, the egg size with the highest fertilization success (i.e., optimal egg size) systematically shifted right with increasing values of clutch mean egg size. Because soft selection had a greater influence when mean egg size was small, the range of fertilizable egg sizes was narrower for clutches with smaller means. Importantly then, we found that when sperm were abundant, the position of the optima as well as the heights and widths of the fitness surfaces increased with increasing values of mean egg size (fig. 1E).

When sperm were limiting, fertilization was again a concave function of absolute egg size (fig. 1B), but this time, soft selection disfavored relatively smaller eggs (fig. 1D). Consequently, eggs closer to the global average (63.5–64.5 μm) had higher relative fitness when they were part of a clutch with a smaller mean size (fig. 1F). As mean egg size increased, the overall fitness surface became laterally compressed, reflecting the greater influence of soft selection. Thus, when sperm were limiting, the fitness optima shifted right with increasing mean egg size (as under sperm saturation), but both the height of the fitness peak and the width of the fitness surface decreased (in contrast to sperm saturation; compare fig. 1E with 1F). Note that soft selection was relatively more influential to the overall patterns of selection under sperm limitation (cf. saturation), as shown by the more dramatic decrease in the height of fitness surfaces in this environment (compare fig. 1E with 1F).

Patterns of Selection on Clutch-Level Egg Size Traits

Selection surfaces of clutch-level traits (means and variances in egg size) varied slightly between sperm environments in the absence of a size-number trade-off (fig. 2A, 2B). In both environments, selection on mean egg size was stabilizing—that is, mothers that produced clutches with a
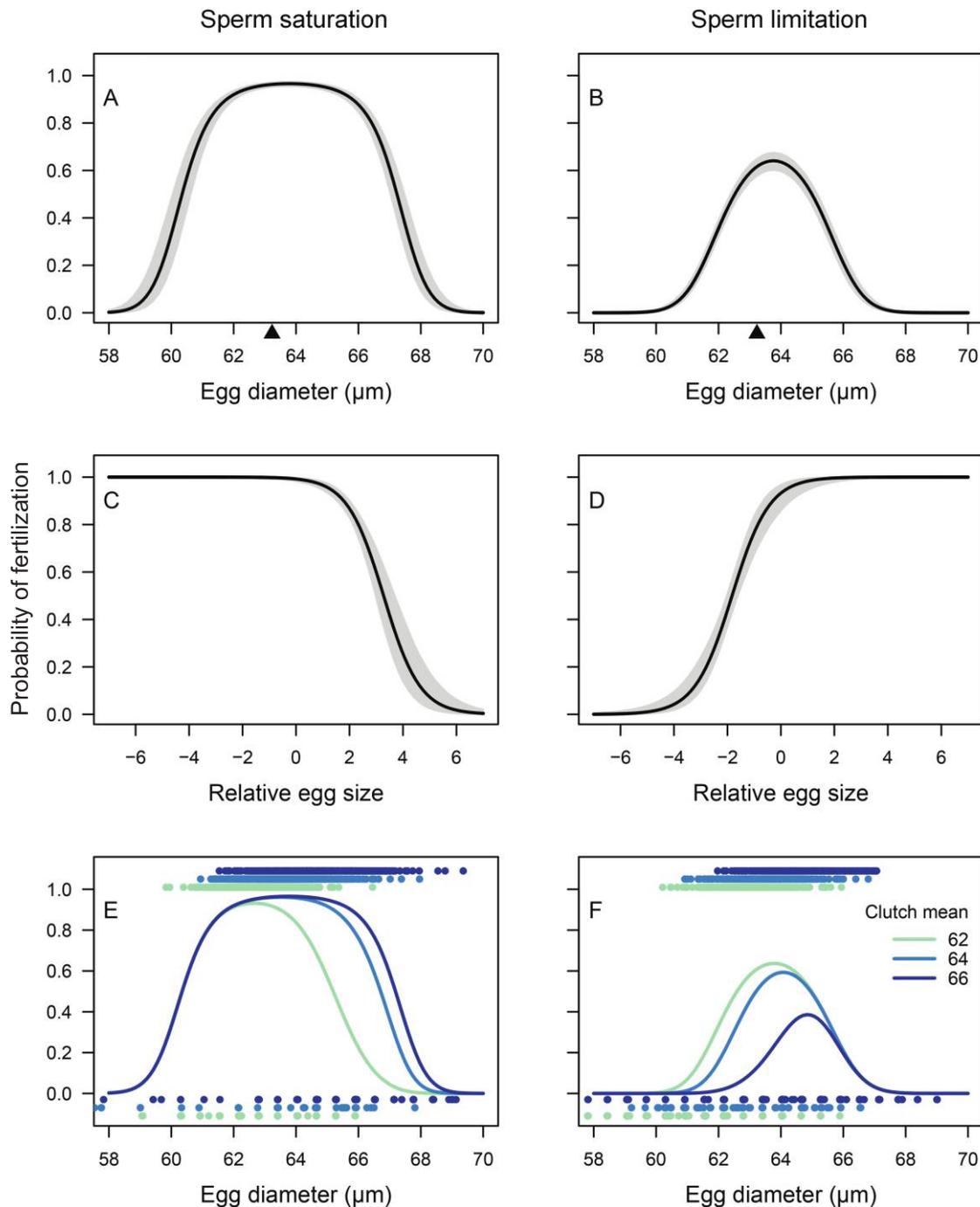
Figure 1: Components of hard and soft selection on egg size during external fertilization under sperm saturation (*A, C, E*) or sperm limitation (*B, D, F*). *A* and *B* depict the average nature of hard selection, where the probability of fertilization (95% confidence interval [CI]) depends on absolute egg size (as diameter; μm) regardless of neighbor frequency. Note that the triangles on the *X*-axes show the current average value of egg size in our study population. *C* and *D* depict the average nature of soft selection, where the probability of fertilization (95% CI) depends on relative values of egg size (as deviations from the clutch mean). *E* and *F* show the combined effects of hard and soft selection on the predicted probability that eggs of a given size will be fertilized as a function of the mean egg size of the clutch to which they belong. For simplicity, we show predicted egg size fitness functions for three representative females whose clutch means represent small (62 μm), average (64 μm), and large (66 μm) values for our study population. Data points show raw values of fertilization success (1 = successful, 0 = unsuccessful) for all eggs in our data set (*n* = 7,008), binned according to the mean egg size of the clutch to which they belonged (small: 60–63 μm; intermediate: 63–65 μm; large: 65–68 μm). The underlying data have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.vq83bk3r6; Cameron et al. 2020).
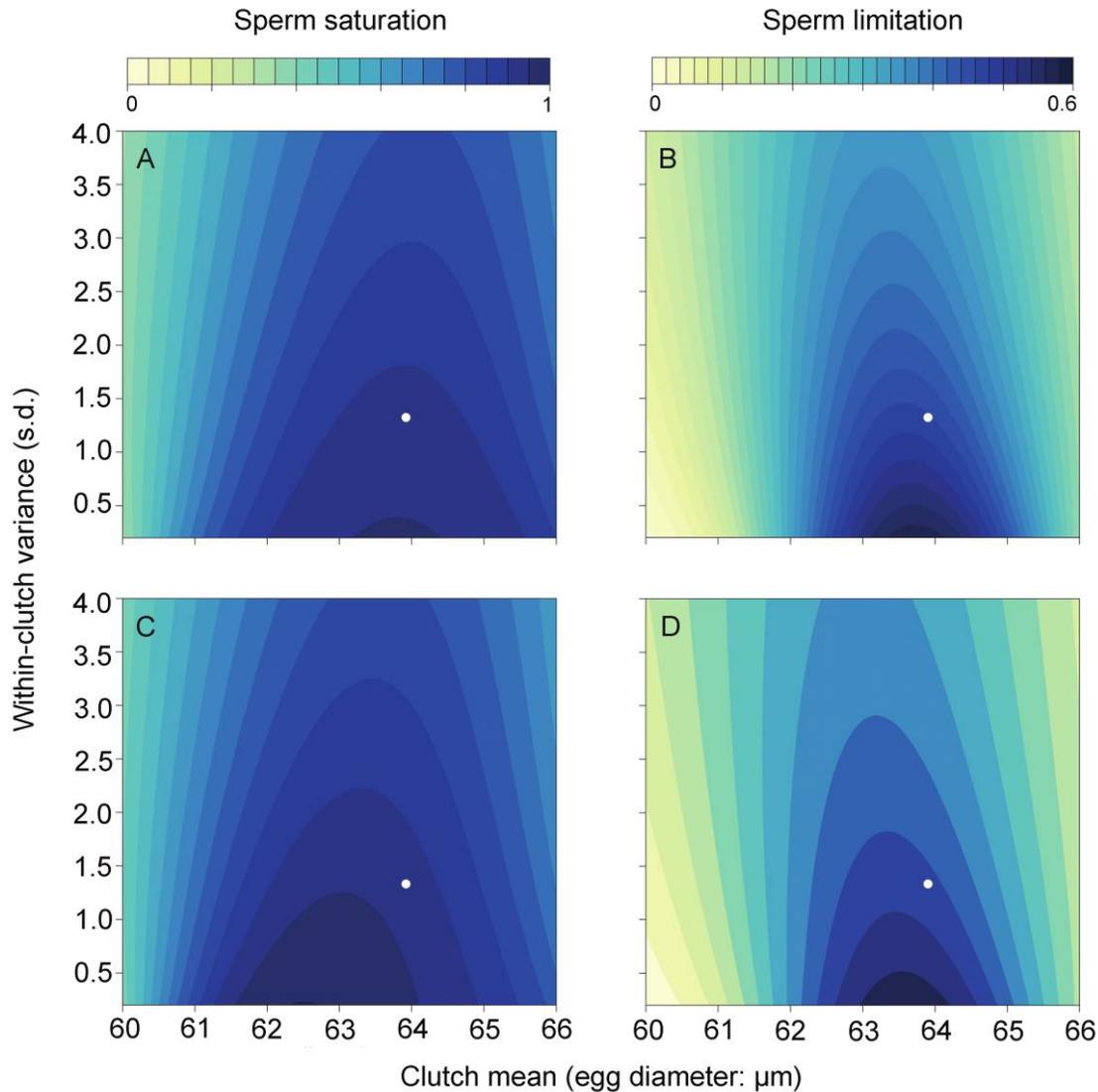
Figure 2: Clutch-level fitness surfaces (as the predicted proportion of fertilized eggs; scale bars) for females that produce clutches that differ in mean egg size (as diameter; μm) and within-clutch variance in egg size (as standard deviation) under sperm saturation (*A, C*) and sperm limitation (*B, D*). *A* and *B* show predictions based on our experimental data in the absence of a trade-off between offspring size and number. *C* and *D* show predictions after accounting for a theoretical trade-off between offspring size and number. Dots represent the current population mean values for clutch means and variances in our study population. Note the change in scale between sperm environments.

mean that deviated from the population grand average generally suffered reduced fitness. Note, however, that stabilizing selection was strongest under sperm limitation. Selection on within-clutch variation also varied slightly in strength depending on sperm availability. Under sperm saturation, selection on within-clutch variance was weak, but overall, larger variances were disfavored (fig. 2A). Under sperm limitation, selection again generally disfavored larger variances, and the fitness surface was (slightly) steeper along the variance axis (fig. 2B). Variance was thus penalized relatively more around the

optimum for mean egg size under sperm limitation compared with sperm saturation (compare fig. 2A with 2B).

Finally, sperm availability altered correlational selection on clutch means and within-clutch variances (fig. 2A, 2B). Under sperm saturation, there was no apparent correlational selection between brood means and variances across the phenotypic range explored (fig. 2A). In contrast, when sperm were limiting, some (albeit weak) negative correlational selection occurred across restricted ranges of parameter space—fertilizations were enhanced by high variance in clutches with very small mean egg sizes

(60.5–61.5 μm), while clutches with very large mean egg sizes (65–66 μm) had higher fitness when within-clutch variance was low (fig. 2*B*). Importantly, however, correlational selection was weak relative to stabilizing selection on mean egg size. As such, in both sperm environments, fitness was consistently maximized for mothers that produced clutches with an intermediate mean egg size and low within-clutch variance (fig. 2*A*, 2*B*).

The above patterns of clutch-level selection were relatively unchanged after accounting for a trade-off between offspring size and number (fig. 2*C*, 2*D*). In both sperm environments, selection on clutch mean egg size remained strongly concave, but incorporating the effects of a trade-off caused a slight rightward shift in optimal values of mean egg size (i.e., negative directional selection), and this was most pronounced under sperm saturation. As such, selection favored relatively smaller clutch mean egg sizes under sperm saturation (62.4–62.6 μm; fig. 2*C*) compared with sperm limitation (63–64.2 μm; fig. 2*D*). Accounting for a size-number trade-off also (slightly) intensified correlational selection between clutch means and variances. Under sperm saturation, some evidence for positive correlational selection emerged, while negative correlational selection was slightly strengthened under sperm limitation—as shown by the increase in leftward and rightward slant in fitness surfaces under sperm saturation (compare fig. 2*A* with 2*C*) and limitation (compare fig. 2*B* with 2*D*), respectively. Recall, however, that our experiments held egg number constant—thus, we do not estimate the effects of egg density on patterns of multilevel selection. Our predictions after accounting for a trade-off may therefore not hold if egg competition depends strongly on egg density (i.e., clutch size). Future experiments that manipulate egg density across clutches that vary in clutch means and variances are thus an important next step in verifying these predictions.

Discussion

We found that multilevel selection targets egg size during external fertilization. At the individual level, fertilization depended on both absolute egg size (hard selection) and egg size relative to the clutch mean (soft selection). Hard and soft selection differed in strength and form—hard selection was stabilizing and strongest under sperm limitation; soft selection was directional and favored eggs relatively larger (under sperm limitation) or smaller (sperm saturation) than the clutch mean egg size. Soft selection thus tempered the intensity of stabilizing hard selection (particularly under sperm limitation), such that within clutches, both the range of fertilizable egg sizes and fitness optima depended on the mean egg size of the clutch (compare fig. 1*E* with 1*F*). At the clutch level, selection on mean egg size was consistently concave and favored clutches with intermediate mean egg sizes, whereas selection on within-clutch variance was weak but generally negative—although some (albeit weak) correlational selection did occur between these two traits. Importantly, we provide one of the first formal tests of multilevel selection on offspring size and, by doing so, reveal several unanticipated pathways via which offspring size variation may be maintained in natural populations.

Patterns of Hard and Soft Selection on Individual Eggs

That hard selection was stabilizing reflects previous findings in external fertilizers—eggs that are too small fail to encounter sperm; eggs that are too large suffer polyspermy (Levitan 1996; Styan 1998; Marshall et al. 2002). In contrast, our detection of soft selection provides novel evidence that interactions among eggs at fertilization are frequency dependent. The patterns of soft selection that we observed may occur if larger eggs exploit sperm more efficiently than smaller eggs. For example, soft selection favored relatively smaller eggs when sperm were saturating, such that eggs closer to the global optimum (imposed by hard selection) had higher fitness when they belonged to a clutch with a larger mean size. Such dynamics may occur if larger eggs remove more sperm and, by doing so, buffer their relatively smaller siblings from polyspermy. In contrast, soft selection favored relatively larger eggs when sperm were limiting, such that eggs closer to the global optimum had lower fitness when they belonged to clutches with a larger mean size. Again, this may occur if larger eggs overexploit sperm at the expense of their smaller siblings. That larger eggs exploit more sperm has some support. Certainly, sperm bind to the egg surface after contact even if those sperm do not enter the egg (Marshall and Evans 2005*b*; Okamoto 2016). Given that sperm-egg collisions increase (at least) proportionately with egg cross-sectional area (Vogel et al. 1982), larger eggs are more likely to be "wasteful" of sperm and rob their smaller siblings of fertilizations (Marshall et al. 2004). Importantly, theory on gamete size evolution extensively considers soft selection on sperm but not eggs (Parker et al. 1972; Bode and Marshall 2007; Henshaw et al. 2014; Parker and Lehtonen 2014; Otto et al. 2015; Lehtonen 2016; Immler and Otto 2018). In light of our findings, we suggest that theory now account for these effects.

We also found that sperm availability altered the relative importance of hard and soft selection—and thus the opportunity for multilevel selection. When sperm were saturating, hard selection largely overwhelmed soft selection, while soft selection was relatively more influential

when resources (sperm) were limited and sibling eggs competed more intensely for fertilizations—congruent with other studies of multilevel selection (Donohue 2004; Weinig et al. 2007). That sperm availability so profoundly altered hard and soft selection implies that selection on egg size may depend on the density of males in this system—although the degree to which sperm concentration is synonymous with male density in broadcast spawners is unclear. Certainly, higher densities of males can yield higher sperm concentrations, but the two are not always correlated. Rather, environmental factors, such as small-scale hydrodynamics and boundary currents, often affect the rate of transport and dilution of sperm independently of population density (Denny and Shibata 1989). As such, local sperm conditions often vary widely over small spatial and temporal scales under field conditions (Levitan and Petersen 1995). Given the sperm-dependent reversals in hard and soft selection that we observe, such local fluctuations in sperm may weaken selection for any one phenotype overall (e.g., Johnson et al. 2013), maintaining at least some egg size variation.

More broadly, patterns of hard and soft selection reported here are unlikely to be restricted to external fertilizers. For example, larger offspring often disproportionately acquire resources (Weiner 1990), such that clutches with larger means may exhaust local resources to the detriment of the entire clutch, particularly when resources are limited (as observed under sperm limitation; fig. 1*F*). Conversely, in subsocial insects, smaller offspring may buffer relatively larger siblings from predation (Kudo 2006), while in colonial marine invertebrates, larger offspring ameliorate water flows to reduce resource stress for their smaller siblings (Cameron et al. 2017; Cameron and Marshall 2019). Frequency-dependent interactions among siblings may thus be cooperative under certain circumstances (e.g., sperm saturation; fig. 1*E*). It is worth noting that our experiments focused on interactions among eggs within family groups (clutches), but in nature, eggs may encounter and interact with clutches spawned by adjacent females. Future experiments could investigate whether higher levels of population structure (i.e., interactions both within and among family groups) alter the patterns of multilevel selection reported here.

Patterns of Selection at the Clutch Level

In contrast to individual-level selection, selection at the clutch level was relatively consistent between sperm environments: selection on clutch mean egg size was largely concave; selection on within-clutch variance was weak but generally negative. Importantly, these findings were relatively consistent even after accounting for a trade-off between offspring size and number (fig. 2). Overall then, these results appear to contradict a major tenet of life-history theory—that is, within any single environment, concave selection on offspring size should strongly disfavor within-clutch variance (Smith and Fretwell 1974). In our study, selection against within-clutch variance (despite concave selection on clutch means) was tempered by frequency-dependent soft selection acting within clutches. Similarly, in colonial marine invertebrates, frequency-dependent selection on offspring size appears to ameliorate sibling competition and counteract selection against within-clutch variance (Cameron et al. 2017). Few other studies have formally estimated selection on within-clutch variance—but a rare example in Soay sheep found variation in the birth weights of twins (i.e., within litters) to be under negligible selection (Childs et al. 2011), despite a stabilizing component of selection acting on the average birth weight of lambs produced by females in this population (Wilson et al. 2005). Thus, the patterns of selection that we report may be more widespread, and we encourage further tests.

We also detected some evidence for correlational selection between clutch means and variances. Correlational selection between these two traits has been previously hypothesized to maintain offspring size variation (Crean and Marshall 2009), but our study is the first empirical demonstration. In our study, such correlational selection emerged from the dynamics of hard and soft selection acting within clutches. To illustrate, recall that under sperm limitation the range of fertilizable egg sizes was narrower for clutches with larger means (see fig. 1*F*)—a pattern that, when integrated across all eggs in a clutch and compared across clutches, yields negative correlational selection (see fig. 2*B*, 2*D*). Conversely, recall that the range of fertilizable egg sizes increased with clutch means under sperm saturation (fig. 1*E*)—a pattern that is congruent with positive correlational selection. Positive correlational selection was extremely weak under sperm saturation, however, and was detectable only after accounting for a theoretical trade-off between offspring size and number (compare fig. 2*A* with 2*C*). Differences in the strength and direction of correlational selection between sperm environments arose as a result of the relatively weaker influence of soft selection under sperm saturation (cf. limitation), as well as differences in the direction of soft selection acting within these two environments. While correlational selection was relatively weak in our study overall, we suspect that these effects may become more pronounced under more extreme levels of resource stress (i.e., sperm limitation or oversaturation) than explored here—that is, conditions where competitive and cooperative interactions are expected to be most intense.

Importantly, our analyses of multilevel selection revealed that trait optima differed between mothers and offspring, albeit in nuanced ways—providing evidence that multilevel selection is somewhat antagonistic. Maternal

(clutch-level) fitness was maximized at an intermediate mean egg size (or a slightly smaller mean under sperm saturation after accounting for a size-number trade-off; fig. 2), while at the individual level, eggs closest to the global optimum had highest fitness when included in a clutch with a larger or smaller mean egg size under sperm saturation and limitation, respectively (fig. 1*E*, 1*F*). In other words, there is a mismatch between the clutch mean egg size that maximizes maternal and offspring fitness (which is exacerbated by the size-number trade-off). Such parent-offspring conflict may result in an evolutionary tug-of-war that maintains offspring size variation (Otto et al. 2015; Immler et al. 2018). Intriguingly, our finding of correlational selection between clutch means and variances also provides a pathway for mothers to recover some of the costs associated with clutch means that favor offspring (rather than maternal) fitness. Ultimately, however, the winner of this parent-offspring conflict may depend on the relative heritability of offspring and maternal egg size traits (Wolf and Wade 2001). Heritability has rarely been estimated at both levels simultaneously (but see Mojonnier 1998), and this represents an important next step.

Conclusions and Future Directions

Within-population variation in offspring size has long perplexed life-history theoreticians (reviewed in Marshall et al. 2018). Our analyses of multilevel selection revealed multiple non–mutually exclusive pathways through which offspring size variation may be maintained across levels of organization—that is, antagonistic multilevel selection, context-dependent (i.e., sperm- or resource-dependent) soft selection, and correlational selection between clutch means and variances. These effects are unlikely to be restricted to external fertilizers given that interactions among siblings are ubiquitous (Blaustein and O'Hara 1987; Cheplick 1993; Donohue 2003; Selkoe et al. 2006; Reed et al. 2009; Kamel et al. 2010). Failure to consider multilevel selection risks an incomplete and potentially misleading view of the selective forces acting across levels of organization. Multilevel approaches are thus powerful inferential tools for exploring selection on offspring size, and we encourage their future use.

Acknowledgments

We thank E. Chirgwin and A. Guillaume for technical advice and M. Amor for graphical assistance and proofreading. We are also grateful to S. Alonzo, R. Bassar, R. Bonduriansky, V. Rudolf, and one anonymous reviewer for insightful comments that greatly improved the manuscript.

Statement of Authorship

H.C., D.J.M., and K.M. conceptualized and designed the experiments; H.C. collected the data; D.W.J. and H.C. analyzed the data; H.C. wrote the first draft of the manuscript; and all authors contributed substantially to revisions.

Data and Code Availability

Data have been deposited in the Dryad Digital Repository (https://doi.org/10.5061/dryad.vq83bk3r6; Cameron et al. 2020).

Literature Cited

Agrawal, A. F., E. D. Brodie III, and M. J. Wade. 2001. On indirect genetic effects in structured populations. American Naturalist 158:308–323.

Blaustein, A. R., and R. K. O'Hara. 1987. Aggregation behaviour in *Rana cascadae* tadpoles: association preferences among wild aggregations and responses to non-kin. Animal Behaviour 35:1549–1555.

Bode, M., and D. J. Marshall. 2007. The quick and the dead? sperm competition and sexual conflict in sea. Evolution 61:2693–2700.

Bolker, B., and R Development Core Team. 2020. bbmle: tools for general maximum likelihood estimation. https://cran.r-project.org/package=bbmle.

Bouwhuis, S., O. Vedder, C. J. Garroway, and B. C. Sheldon. 2015. Ecological causes of multilevel covariance between size and first-year survival in a wild bird population. Journal of Animal Ecology 84:208–218.

Cameron, H., D. W. Johnson, K. Monro, and D. J. Marshall. 2020. Data from: Multilevel selection on offspring size and the maintenance of variation. American Naturalist, Dryad Data Repository, https://doi.org/10.5061/dryad.vq83b3r6.

Cameron, H., and D. J. Marshall. 2019. Can competitive asymmetries maintain offspring size variation? a manipulative field test. Evolution 73:1663–1671.

Cameron, H., K. Monro, M. Malerba, S. Munch, and D. J. Marshall. 2016. Why do larger mothers produce larger offspring? a test of classic theory. Ecology 97:3452–3459.

Cameron, H., K. Monro, and D. J. Marshall. 2017. Should mothers provision their offspring equally? a manipulative field test. Ecology Letters 20:1025–1033.

Cheplick, G. P. 1993. Sibling competition is a consequence of restricted dispersal in an annual cleistogamous grass. Ecology 74:2161–2164.

Childs, D. Z., T. N. Coulson, J. M. Pemberton, T. H. Clutton-Brock, and M. Rees. 2011. Predicting trait values and measuring selection in complex life histories: reproductive allocation decisions in Soay sheep. Ecology Letters 14:985–992.

Crean, A. J., and D. J. Marshall. 2009. Coping with environmental uncertainty: dynamic bet hedging as a maternal effect. Philosophical Transactions of the Royal Society B 364:1087–1096.

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

Donohue, K. 2003. The influence of neighbor relatedness on multilevel selection in the Great Lakes sea rocket. American Naturalist 162:77–92.

———. 2004. Density-dependent multilevel selection in the Great Lakes sea rocket. Ecology 85:180–191.

Goodnight, C. J., J. M. Schwartz, and L. Stevens. 1992. Contextual analysis of models of group selection, soft selection, hard selection

and the evolution of altruism. American Naturalist 140:743–761.

Hambright, K. D. 1991. Experimental analysis of prey selection by largemouth bass: role of predator mouth width and prey body depth. Transactions of the American Fisheries Society 120:500–508.

Heisler, I. L., and J. Damuth. 1987. A method for analyzing selection in hierarchically structured populations. American Naturalist 130:582–602.

Henshaw, J. M., D. J. Marshall, M. D. Jennions, and H. Kokko. 2014. Local gamete competition explains sex allocation and fertilization strategies in the sea. American Naturalist 184:32–49.

Immler, S., and S. P. Otto. 2018. The evolutionary consequences of selection at the haploid gametic stage. American Naturalist 192:241–249.

Johnson, D. W., M. R. Christie, J. Moye, and M. A. Hixon. 2011. Genetic correlations between adults and larvae in a marine fish: potential effects of fishery selection on population replenishment. Evolutionary Applications 4:621–633.

Johnson, D. W., K. Monro, and D. J. Marshall. 2013. The maintenance of sperm variability: context-dependent selection on sperm morphology in a broadcast spawning invertebrate. Evolution 67:1383–1395.

Kamel, S. J., R. K. Grosberg, and D. J. Marshall. 2010. Family conflicts in the sea. Trends in Ecology and Evolution 25:442–449.

Kudo, S. I. 2006. Within-clutch egg-size variation in a subsocial bug: the positional effect hypothesis. Canadian Journal of Zoology 84:1540–1544.

Kupriyanova, E. K. 2006. Fertilization success in *Galeolaria caespitosa* (Polychaeta: Serpillidae): gamete characteristics, role of sperm dilution, gamete age, and contact time. Scientia Marina 70:9.

Lehtonen, J. 2016. Multilevel selection in kin selection language. Trends in Ecology and Evolution 31:752–762.

Levitan, D. R. 1996. Effects of gamete traits on fertilization in the sea and the evolution of sexual dimorphism. Nature 382:153–155.

Levitan, D. R., and C. Petersen. 1995. Sperm limitation in the sea. Trends in Ecology and Evolution 10:228–231.

Lips, K. R. 2001. Reproductive trade-offs and bet-hedging in *Hyla calypsa*, a Neotropical treefrog. Oecologia 128:509–518.

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

Marshall, D. J., C. N. Cook, and R. B. Emlet. 2006. Offspring size effects mediate competitive interactions in a colonial marine invertebrate. Ecology 87:214–225.

Marshall, D. J., and J. P. Evans. 2005a. The benefits of polyandry in the free-spawning polychaete *Galeolaria caespitosa*. Journal of Evolutionary Biology 18:735–741.

———. 2005b. Does egg competition occur in marine broadcast-spawners? Journal of Evolutionary Biology 18:1244–1252.

Marshall, D. J., A. K. Pettersen, and H. Cameron. 2018. A global synthesis of offspring size variation, its eco-evolutionary causes and consequences. Functional Ecology 32:1436–1446.

Marshall, D. J., P. D. Steinberg, and J. P. Evans. 2004. The early sperm gets the good egg: mating order effects in free spawners. Proceedings of the Royal Society B 271:1585–1589.

Marshall, D. J., C. A. Styan, and M. J. Keough. 2002. Sperm environment affects offspring quality in broadcast spawning marine invertebrates. Ecology Letters 5:173–176.

Mojonnier, L. 1998. Natural selection on two seed size traits in the common morning glory *Ipomoea purpurea* (Convolvulaceae): patterns and evolutionary consequences. American Naturalist 152:188–203.

Monro, K., and D. J. Marshall. 2015. The biogeography of fertilization mode in the sea. Global Ecology and Biogeography 24:1499–1509.

———. 2016. Unravelling anisogamy: egg size and ejaculate size mediate selection on morphology in free-swimming sperm. Proceedings of the Royal Society B 283:20160671.

Moore, A. J., E. D. Brodie III, and J. B. Wolf. 1997. Interacting phenotypes and the evolutionary process. 1. Direct and indirect genetic effects of social interactions. Evolution 51:1352–1362.

Okamoto, D. K. 2016. Competition among eggs shifts to cooperation along a sperm supply gradient in an external fertilizer. American Naturalist 187:129–142.

Otto, S. P., M. F. Scott, and S. Immler. 2015. Evolution of haploid selection in predominantly diploid organisms. Proceedings of the National Academy of Sciences of the USA 112:15952–15957.

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

Parker, G. A., and M. Begon. 1986. Optimal egg size and clutch size: effects of environment and maternal phenotype. American Naturalist 128:573–592.

Parker, G. A., and J. Lehtonen. 2014. Gamete evolution and sperm numbers: sperm competition versus sperm limitation. Proceedings of the Royal Society B 281:20140836.

Plaistow, S. J., J. J. H. St. Clair, J. Grant, and T. G. Benton. 2007. How to put all your eggs in one basket: empirical patterns of offspring provisioning throughout a mother's lifetime. American Naturalist 170:520–529.

Reed, W. L., M. E. Clark, and C. M. Vleck. 2009. Maternal effects increase within family variation in offspring survival. American Naturalist 174:685–695.

Rollinson, N., and L. Rowe. 2015. Persistent directional selection on body size and a resolution to the paradox of stasis. Evolution 69:2441–2451.

Selkoe, K. A., S. D. Gaines, J. E. Caselle, and R. R. Warner. 2006. Current shifts and kin aggregation explain genetic patchiness in fish recruits. Ecology 87:3082–3094.

Smith, C. C., and S. D. Fretwell. 1974. Optimal balance between size and number of offspring. American Naturalist 108:499–506.

Styan, C. A. 1998. Polyspermy, egg size, and the fertilization kinetics of free spawning marine invertebrates. American Naturalist 152:290–297.

Turnbull, L. A., L. Santamaria, T. Martorell, J. Rallo, and A. Hector. 2006. Seed size variability: from carob to carats. Biology Letters 2:397–400.

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

Wade, M. J. 1977. Experimental study of group selection. Evolution 31:134–153.

Wallace, B. 1975. Hard and soft selection revisited. Evolution 29:465–473.

Weiner, J. 1990. Asymmetric competition in plant populations. Trends in Ecology and Evolution 5:360–364.

Weinig, C., J. A. Johnston, C. G. Willis, and J. N. Maloof. 2007. Antagonistic multilevel selection on size and architecture in variable density settings. Evolution 61:58–67.

Wilson, A. J., J. G. Pilkington, J. M. Pemberton, D. W. Coltman, A. D. Overall, K. A. Byrne, and L. E. Kruuk. 2005. Selection on mothers and offspring: whose phenotype is it and does it matter? Evolution 59:451–463.

Wolf, J. B., and M. J. Wade. 2001. On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? Journal of Evolutionary Biology 14:347–356.

Yund, P. O., and S. K. Meidel. 2003. Sea urchin spawning in benthic boundary layers: are eggs fertilized before advecting away from females? Limnology and Oceanography 48:795–801.

Literature Cited Only in the Online Enhancements

Graham, M. H. 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84:2809–2815.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, New York.

Associate Editor: Volker H. W. Rudolf
Editor: Russell Bonduriansky
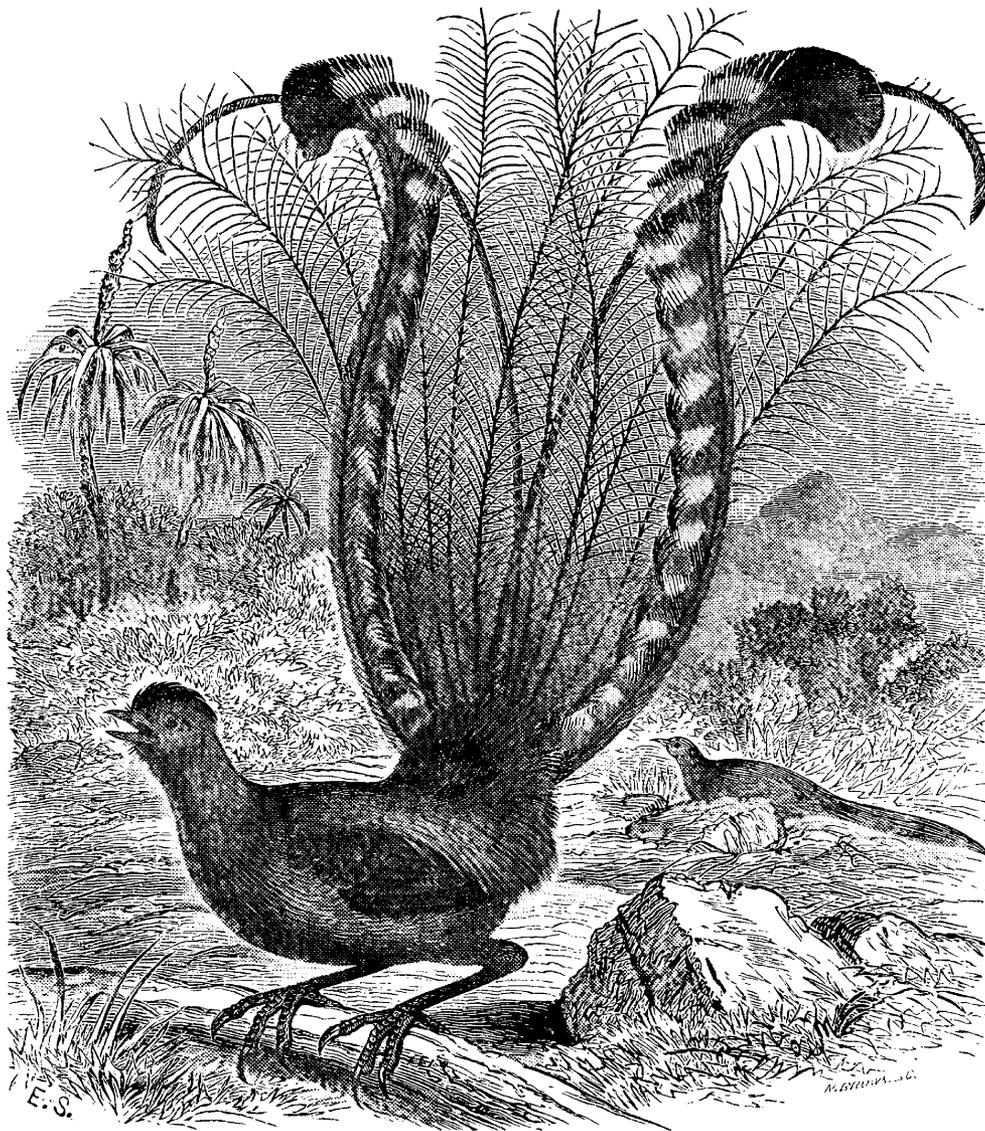
"But that which gives character to the whole is the arrangement of the external feathers. These curve in such a manner that the two together form the outline of an ancient lyre, an appearance so striking as to confer on the birds their popular name. These two feathers contrast with the middle ones by presenting vanes, *wide* on the inner side, on the whole length of the shaft." From "The Lyre Bird" by Grace Anna Lewis (*The American Naturalist*, 1870, 4:321–331).