Supporting Online Material for

Selection on Gamete Recognition Proteins Depends on Sex, Density, and Genotype Frequency

Don R. Levitan and David L. Ferrell

E-mail: levitan@bio.fsu.edu (D.R.L.); ferrell@bio.fsu.edu (D.L.F.)

Published 14 April 2006, *Science* **312**, 267 (2006)
DOI: 10.1126/science.1122183

This PDF file includes:

Materials, Methods and Discussion of Results

Discussion of fast and slow binding rates

We define fast binding rates as conditions where sperm and eggs have a high affinity for each other. Fast binding rates are manifested as relatively high levels of fertilization under low levels of sperm availability. Slow binding rates imply a lower affinity and eggs require higher concentrations of sperm to achieve high levels of fertilization. Among and within *Strongylocentrotus* species fertilization rates vary (*1-4*). This variation is correlated with the degree of sperm availability in nature resulting from species differences in abundance and distribution (*4*). Species that are more likely to experience sperm limitation have gamete traits and rates of fertilization that allow for high fertilization under sperm-limited conditions. Species that are more likely to experience sperm competition have gamete traits and rates of fertilization that allow for success under sperm competitive conditions (*1-4*). In addition, variation associated with fertilization rates from conspecific sperm are correlated with variation from heterospecific sperm; eggs that are easily fertilized under low conspecific sperm concentration are also more easily fertilized by heterospecific sperm (*5*). These results are consistent with the notion of density-dependent selection on gamete traits to maximize fitness as a function of sperm availability (*4*).

Definition and Methods of Estimating Reproductive Success

Reproductive success in females was defined as the fraction of eggs successfully developing and varies between 0-1. Reproductive success in males was determined using

microsatellite markers on all adults and 20 larvae per female and was defined as the product of paternity share (number of sired larvae out of 20 larvae analyzed per female) and that female's reproductive success (details see reference *6*). Male reproductive success can vary between 0 and n, where n is the number of spawning females. Pair-wise reproductive success is the proportion of eggs sired by a particular male and is simply the product of paternity share and female reproductive success.

Field Methods

Each adult was tagged with a numbered latex band and positions were mapped onto a 5 x 5 m grid. Individuals were monitored for the full experimental period and movement of sea urchins was minimal (e.g., they did not reposition themselves next to mates or switch to alternate mates). Fertilization success and movement patterns during KCl-induced spawning events produce similar levels of movement and reproductive success compared to natural spawning events at similar spawning densities (*4*). Water flow was measured as the mean advection of flow (m/s) over the experimental time period. Advection rate was calculated as the total straight-line distance a particle of water moved over the experimental time period, divided by that time. Flow was measured at half s intervals using an InterOcean S4 current meter placed 5 m from the spawning event and 0.25 m off the bottom.

Measuring Polyspermy

Estimates of female reproductive success were measured as the proportion of developing embryos three hours after experiments were conducted. Laboratory

experiments indicate that in this species at extreme sperm concentrations eggs fail to cleave or show a raised vitelline envelope. Field experiments were conducted to examine the degree to which eggs exhibited polyspermy as a function of male spawning density. Field methods were similar to the present study; however, collected eggs were fixed in formalin 4-8 minutes after exposure to sperm. These eggs were first examined for the fraction of eggs showing signs of development (raised vitelline envelope) and then examined with a confocal microscope to determine the number of eggs with one or more than one fused spermatozoan. The number of eggs fused with one spermatozoan first increased and then decreased with male spawning density. The number of eggs with multiple sperm fused to eggs increased with male spawning density. The decrease in monospermic fertilizations at high density was associated with the increase in polyspermic fertilizations and the density at which polyspermy was noted was similar to the densities at which reduced developmental success was noted in the present study (Details in *6*).

Molecular Methods

Tube feet samples (3-5 per individual) were digested in a solution of CTAB and proteinase K incubated in a 65^{o}C water bath for approximately 12 hours. DNA extractions were performed using a SprintPrep DNA Purification kit, a magnetic bead-based protocol. Extracts then were stored at –20^{o}C. We performed double-stranded PCR amplification of a 431 bp region of the bindin locus implementing the following primers developed by Debenham et al. (*7*): FNbindin5’ (5’-AGTCGACGTTCGACAGACGAC-3’) and FNbindin3’ (5’-TTACATGGTCCATTATAGTATGCC-3’). The PCR cocktail

consisted of 12.95 μl double distilled water, 2.5 μl 10X PCR buffer, 1.0 μl 2mM MgCl, 2.5 μl 2mM dNTPs, 0.15 μl Taq, 1.2 μl 0.5 μM FNbindin5' primer, 1.2 μl μM FNbindin3' primer, 1.0 μl 10 μM bovine serum albumin, and 1.0 μl DNA (25 ng/μl). The PCR program was as follows: 5 minutes at 95^{o}C; 30 cycles of 1 minute at 94^{o}C, 1 minute at 59^{o}C, and 2 minutes at 72^{o}C; and 7 minutes at 72^{o}C. Internal sequencing primer KTseq3' (5'-ATACACACGATGGTCAAG-3') was used to obtain initial sequence data and determine whether individuals were homozygous or heterozygous with respect to a 273 bp region (*7*) within the bindin locus (bp 944 to 1217 in reference *8*). This subset of the locus is one of two variable regions and is known to exhibit both interspecific and intraspecific sequence variation in *Echinometra* (*9, 10*) and *Strongylocentrotus* species (*7*). Additional internal sequencing primers, KTseq5' (GGAGCGCGTAAGAAGCGTTAT) or KTseqn5' (ACGTTCGACAGACGACGAC), were used to confirm sequence data of homozygotes and individuals exhibiting polymorphism at only a single site. We obtained re-amplified PCR product from all other heterozygote individuals, which were subcloned into pCR2.1-TOPO vector from Invitrogen using the manufacturer's protocol. The inserts from at least four recombinant clones for each sample were PCR-amplified directly from the bacterial colonies using M13 forward and reverse primers. The resulting DNA fragments were then gel purified. All PCR products were purified with a QIAquick purification column prior to sequencing on an Applied Biosystems 3100 Genetic Analyzer.

Testing for the Genetic Signature of Selection

We tested for conformance to Hardy-Weinberg equilibrium by using GenePop (http://wbiomed.curtin.edu.au/genepop/genepop_op1.html) to perform an exact test, and then conducting contingency χ^2analysis with categorical pooling into common ($\geq$ 10 individuals) and rare (< 10 individuals) genotypes. Overall genotype frequencies were very similar to Hardy-Weinberg expectations; for example, observed/expected values for genotypes 'aa,' 'ab,' 'ac,' and 'ad' were 41/42, 24/20, 23/23, and 12/12, respectively. Neither test indicated deviation from Hardy-Weinberg equilibrium (exact test: $p = 0.45$; contingency test: $\chi^2_{calc, df=1} = 0.20$, $p = 0.65$)

We explored potential differences between our observed bindin allele frequencies and those reported by Debenham et al. (*7*) with two contingency chi-square tests. The first included overall allele frequencies whereas common ($\geq$ 0.10) and rare (< 0.10) alleles were pooled in the second test. Neither test detected differences in allele frequencies between the two studies (overall test: $\chi^2_{calc, df=4} = 3.26$, $p = 0.52$; pooled test: $\chi^2_{calc, df=1} = 0.03$, $p = 0.86$).

Using the McDonald and Kreitman (*11*) test for neutral evolution, we compared the number of polymorphic *S franciscanus* synonymous and non-synonymous substitutions to the number of fixed synonymous and non-synonymous substitutions in the bindin gene of three other urchin species. We used the same three sequences used by Debenham et al. (*7*) to test for deviations from neutral evolution: *S. purpuratus*, GenBank accession number M14487 (*12*); *S. droebachiensis*, GenBank accession number AF133804 (*13*); *Lytechinus variegatus*, GenBank accession number M59489, (*8*). Sequences were imported and aligned using MEGA 3.1 software

(http://www.megasoftware.net/mega.html) after which DNAsp 4.10.3 software (http://www.ub.es/dnasp/) was used to apply a G test of independence with the Williams correction for continuity (*11*). We conducted neutrality tests on sequence data from the 15 alleles observed in this study only as well as a more inclusive data set generated by combining additional *S. franciscanus* alleles reported by Debenham et al. (*7*). Three related urchin species (*S. droebachiensis, S. purpuratus,* and *Lytechinus variegatus*) were used as outgroups and only a 123-bp subset of the variable region that aligned unambiguously was included in this analysis, as in Debenham et al. (*7*). All tests failed to detect deviations from a neutral model whether including only alleles identified in the present study (*S. droebachiensis*: G = 0.00, p = 0.98; *S purpuratus*: G = 0.00, p = 0.96; *L. variegatus*: G = 0.07, p = 0.79) or all known alleles (*S. droebachiensis*: G = 0.53, p = 0.46; *S purpuratus*: G = 0.47, p = 0.49; *L. variegatus*: G = 0.09, p =0.76).

Analysis of Reproductive Success Using Event-by-Event Genotype Frequencies

Chance at small population sizes alters the genotype frequencies among spawning events, an issue that is exacerbated at low densities. For instance, in one low density event there were no AA males (most common genotype at an overall frequency of 0.33) and the usually rare genotype (CC, overall frequency of 0.02) was common (0.17). To adjust for this, and to determine if the genotype frequency of matched mates in an event explains reproductive success (the probability of encountering a like or unlike individual), we calculated the genotype frequency within each sex and spawning event. Reproductive success of an individual was plotted as a function of the proportion of individuals in the other sex that shared that genotype. This was done independently for both males and females within each spawning event. For each event, the slope of this relationship (adjusted by the covariates of average distance to mates and advection) was calculated. A positive slope indicates that individuals that share a matched genotype with a high proportion of mates have higher reproductive success than individuals that share genotypes with a lower proportion of mates. This is positive frequency dependence (a high frequency of matching is favored). Negative slopes are the reciprocal situation of negative frequency dependence (a low frequency of matching is favored). Results indicate no significant difference between the sexes ($P > 0.5$), but a significant effect of density (log transformed) on the direction and intensity of frequency dependence ($R^2 = 0.27$, $P < 0.039$).

Analysis of Pair-wise Reproductive Success

Figure 4A presents the least-squared means of the pair-wise analysis shown in Table S3. Least-square means are the means adjusted to the mean covariate level. To highlight the complex set of interactions that characterize this analysis, we generated Figure 4B which shows the pair-wise reproductive success of the most common and rare male genotypes. Comparisons are made between AA males with matched (fully and partially) and unmatched (non-A) females and non-A males with matched and unmatched females at both low and high density.

The overall higher reproductive success at high density is because male-female distance is reduced and sperm limitation is rare (Distance effect). The overall higher success of common (AA) males compared to rare (non-A) males is because the likely distance between matched individuals varies as a function of genotype frequency; females are more likely to be close to common males. This sets up the complex interactions between spawning density, genotype frequency, nearest mate distance and matching rules.

The male genotype effect is stronger among common males at high density (male genotype frequency by density effect and male genotype frequency effect at high density, but not at low density). Genotype frequency is less likely to be important at low density, because individuals are further apart and sperm diffusion effects will swamp genetic identity effects (distance effect, matching by distance effect, and matching by distance by density effects), and because at low densities overall genotype frequency is a poorer predictor of the probability of matching than at higher densities, where population size is larger (See Fig.3).

Common males have high reproductive success at low densities because matches are common and result in high gametic affinities under sperm-limited conditions and in the absence of polyspermy (matching effect at low density and matching by density interaction). Common males also have high reproductive success at high densities because they do exceptionally well with unmatched females which minimizes polyspermy (matching effect and matching by density interaction), and although common males may cause polyspermy with matched females, they still out-compete rare males because of their higher affinity for these common females (male genotype effect at high density).

Rare males have generally reduced reproductive success at low density because they match with few mates and the mates they do match with are not likely to be close (distance effect, matching by distance effect and matching by distance by density effect). Rare males at high density have reduced reproductive success compared to common males because the females they do match with are not likely to be close (matching by distance, matching by distance by density) and they will be out-competed for fertilizations in the common females they do not match (density by male genotype frequency).

Superimposed on these effects, advection plays a major role in reproductive success. When flow is high, fertilization is reduced (advection effect). Advection is more important at low densities, where distances are greater and the rapid dissipation of sperm has a more dramatic effect (advection by density). Advection also influences the effect of matching; under low density, high flow environments gamete collisions are rare and competition based on affinity rates become less critical (advection by matching at

low density). There was no significant difference between the level of advection at low and high densities (student's 't' test, $P>0.75$).

Supporting Online Text

Examination of Positive Selection in Other Sea Urchin Species

Biermann (*14*) conducted pair-wise significance tests of an excess of non-synonymous substitutions compared with synonymous substitutions in five closely related species of Strongylocentrotidae. All significant pair-wise tests involved *S. purpuratus* or *S. polyacanthus* in two polymorphic regions. *Strongylocentrotus purpuratus* was significant in four out of five comparisons at one polymorphic site and one out of three at the other site. *Strongylocentrotus polyacanthus* was significant in two out of four comparisons at both sites. All the other species examined (*S. droebachiensis*, *S. pallidus*, and *Allocentrotus fragilis*) failed to show significant excesses in non-synonymous substitutions when paired with one another.

Tables

Tables S1 and S2. A general linear model (SAS) was used to test for the effects of bindin genotype frequency on reproductive success. We included the level of advection into the model to account for variation in reproductive success induced by water flow. For males (S1) and females (S2) there was a significant genotype frequency effect.

Table S1

Male

Source of Variation	df	Type III SS	MS	F	P
Genotype Frequency	1	1.309	1.309	4.70	0.034
Advection	1	1.692	1.692	6.07	0.016
Error	66	18.405	0.279		
Corrected Total	68	20.925			

Table S2

Female

Source of Variation	df	Type III SS	MS	F	P
Genotype Frequency	1	1.215	1.215	13.24	0.0006
Advection	1	0.130	0.130	1.41	0.2391
Error	52	4.775	0.092		
Corrected Total	54	6.120			

Tables S3-S5. An analysis of covariance (General Linear Model, SAS) was used to explain variation in pair-wise reproductive success with the main effect of spawning density (Density: high or low) and matching rule (Matching: full, partial or full allelic matching with female) and the covariates of distance between the pair (Distance: log transformed), water flow (Advection: log transformed), the genotype frequency of the male (Male) and the female (Female) and the number of male competitors (Competitors). Type three sums of squares were used to test for significance, which adjusts the level of significance by the other factors in the model. For example a significant matching and advection effect means matching is significant after taking into account the effect of water flow on fertilization (and vice versa). Table S3 presents the ANCOVA with significant and marginally significant ($P < 0.10$) three-way and all two-way interactions involving main effects kept in the model. Because of the significant interactions with spawning density and in particular an interaction between matching rules and density, separate ANCOVAs were conducted at each density. Within each density (Tables S4 and S5) significant factors were left in the model, and there were significant matching rule effects.

Table 3

Source of Variation	df	Type III SS	MS	F	P
Matching	2	0.00074	0.00037	0.03	0.9714
Density	1	0.00341	0.00341	0.27	0.6061
Distance	1	0.11077	0.11077	8.64	0.0034
Advection	1	0.12641	0.12641	9.86	0.0018
Competitors	1	0.00369	0.00369	0.29	0.5917
Male	1	0.09384	0.09384	7.32	0.0071
Female	1	0.00409	0.00409	0.32	0.5722
Matching by Density	2	0.08016	0.04008	3.13	0.0447
Matching by Distance	2	0.06140	0.03070	2.40	0.0922
Matching by Advection	2	0.00040	0.00020	0.02	0.9846
Matching by Competitors	2	0.04788	0.02394	1.87	0.1556
Matching by Male	2	0.01223	0.00611	0.48	0.6208
Matching by Female	2	0.02332	0.01166	0.91	0.4033
Density by Distance	1	0.07114	0.07114	5.55	0.0189
Density by Advection	1	0.00024	0.00024	0.02	0.8912
Density by Competitors	1	0.08128	0.08128	6.34	0.0121
Density by Male	1	0.05713	0.05713	4.46	0.0353
Density by Female	1	0.03090	0.03090	2.41	0.1212
Matching by Density by Distance	2	0.06200	0.03100	2.42	0.0901
Matching by Density by Advection	2	0.13512	0.06756	5.27	0.0054
Error	484	6.20346	0.01282		
Corrected Total	513	8.14500			

Table 4 (Low Density)

Source of Variation	df	Type III SS	MS	F	P
Matching	2	0.04926	0.02463	3.29	0.0392
Competitors	1	0.04997	0.04997	6.68	0.0105
Advection	1	0.09317	0.09317	12.46	0.0005
Matching by Advection	2	0.05527	0.02764	3.70	0.0266
Error	194	1.45031	0.00747		
Corrected Total	200	1.81548			

Table 5 (High Density)

Source of Variation	df	Type III SS	MS	F	P
Matching	2	0.12792	0.06396	3.88	0.0216
Competitors	1	0.45595	0.45595	27.69	0.0001
Distance	1	0.20977	0.20977	12.74	0.0004
Advection	1	0.07391	0.07391	4.49	0.0349
Male	1	0.21838	0.21838	13.26	0.0003
Error	194	1.45031	0.00747		
Corrected Total	200	1.81548			

Online Supplement References

1. D.R. Levitan, *Am. Nat.* **141**, 517 (1993).

2. D.R. Levitan, *Nature* **382**, 153 (1996).

3. D.R. Levitan, *Evolution* **52**, 1043 (1998).

4. D.R. Levitan, *Ecology* **83**, 464 (2002a).

5. D.R. Levitan, *Evolution* **56**, 1599 (2002b).

6. D.R. Levitan, *Am. Nat*. **164**, 298 (2004).

7. P. Debenham, M.A. Brzezinski, K.R. Foltz, *J. Mol. Evol*. **51**, 481 (2000).

8. J.E. Minor, D.R. Fromson, R.J. Britten, E.H. Davidson, *Mol. Biol. Evol*. **8**, 781 (1991).

9. V.D. Vacquier, W.J. Swanson, M.E. Hellberg, *Dev. Growth Differ*. **37**, 1 (1995).

10. E.C. Metz, S.R. Palumbi, *Mol. Biol. Evol*. **13**, 397 (1996).

11. J.H. McDonald, M. Kreitman, *Nature* 351, **652** (1991).

12. B. Gao, L.E. Klein, R.J. Britten, E.H. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8634 (1986).

13. C. Biermann, W. Eanes, unpublished data.

14. C.H. Biermann, *Mol. Biol. Evol*. **15**, 1761 (1998).