All males are not created equal: Fertility differences depend on gamete recognition polymorphisms in sea urchins

Stephen R. Palumbi*

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138

Edited by Roy J. Britten, California Institute of Technology, Corona del Mar, CA, and approved August 10, 1999 (received for review June 18, 1999)

Behaviors, morphologies, and genetic loci directly involved in reproduction have been increasingly shown to be polymorphic within populations. Explaining how such variants are maintained by selection is crucial to understanding the genetic basis of fertility differences, but direct tests of how alleles at reproductive loci affect fertility are rare. In the sea urchin genus *Echinometra*, the protein bindin mediates sperm attachment to eggs, evolves quickly, and is polymorphic within species. Eggs exposed to experimental sperm mixtures show strong discrimination on the basis of the males' bindin genotype. Different females produce eggs that nonrandomly select sperm from different males, showing that variable egg–sperm interactions determine fertility. Eggs select sperm with a bindin genotype similar to their own, suggesting strong linkage between female choice and male trait loci. These experiments demonstrate that alleles at a single locus can have a strong effect on fertilization and that reproductive loci may retain functional polymorphisms through epistatic interactions between male and female traits. They also suggest that positive selection at gamete recognition loci like bindin involves strong selection within species on mate choice interactions.

Genes directly involved in reproduction are tightly associated with organismal fitness, and the simple evolutionary expectation is that selection should act quickly to weed out all but the best variant at these loci (1, 2). Despite this, there is a surprising degree of natural polymorphism in reproductive tactics and in reproductive genes within many species. Polymorphic mating tactics or mate preferences exist in insects, fish, birds, lizards, and amphibians (3–8). Among insects, accessory gland proteins are passed to females during insemination, and different alleles are associated with differential sperm success (1). The ability to displace sperm from a prior mating varies widely among males in insects, crustacea, molluscs, and mammals (9–12) and appears to be under genetic control (13). In humans, alleles that reduce fertility are generally treated as genetic diseases, but some variants associated with reduced reproduction are widespread (14). Is it possible that such variation is maintained by selection? Or are alleles at reproductive loci always subject to powerful selective sweeps (2) in which the best allele quickly comes to dominate?

It has been difficult to answer in detail how selection operates at most polymorphic loci involved in mating strategies because the genetic determinants of these traits are generally unknown (15). In some cases, however, adult interactions during mating are simplified and may rely on fewer gene products, especially in many terrestrial plants and marine taxa in which gamete recognition is pivotal (16). Study of molecular evolution of genes involved in gamete recognition has shown that positive selection for amino acid divergence occurs between species (17–21), but whether this selective regimen also operates within species is unknown.

In sea urchins, external fertilization is mediated by attachment to the egg of the sperm protein bindin (18, 22, 23). In the genera *Echinometra* and *Strongylocentrotus*, bindin varies greatly between species and, like many proteins involved in gamete recognition (24) or other reproductive interactions (25, 26), evolves by positive selection for amino acid divergence (21, 27). In species for which positive selection has been reported, bindin is also highly polymorphic, with multiple allelic polymorphisms that are generated by amino acid replacements or coding region insertion/deletions (21, 28). Although high replacement/silent site variation in these alleles suggests that they are under selection (21), there have been no direct tests of functional differences among bindin alleles. One alternative is that bindin allele differences are functionally neutral and that the marked bindin polymorphism is a reflection of overall high genetic variation in the sea urchin nuclear genome (29). Understanding functional differences among bindin alleles may clarify the mechanisms underlying positive selection on gamete recognition loci and allow characterization of selection rules governing evolution of mating system polymorphisms.

Materials and Methods

Characterization of Bindin Alleles in Adults. Adults used in this study were collected from Pidi Reef, Guam, and Oahu, HI. Only Hawaiian animals were used in crosses. These were shipped to Boston and maintained in recirculating sea water tables until use. Adult genotypes were obtained by PCR of tube foot DNA isolated from stabled animals. PCR was conducted with the error-correcting enzyme mix *rTth* (Perkin–Elmer) by using primers listed in ref. 21. PCR products from the first 105 codons of the mature bindin gene (showing the strongest signal of positive selection (21)) were cloned into plasmid vectors. Between 6 and 10 positive clones were sequenced for each individual to obtain diploid gene sequences for each adult used as a parent in this study. Phylogenetic trees from aligned allele sequences were generated by using parsimony (PAUP 3.1.1). Gaps were treated as single characters, weighted five times as much as nucleotide substitutions. Sequences from the sister species to *Echinometra mathaei*, E. sp. nov. A, were used as outgroups.

Crosses. Gametes were obtained by minimal KCl injection. Eggs were washed three times in filtered, artificial sea water. Dry sperm collected at the gonopores were diluted 1,000-fold before use. Sperm from different males were diluted, and concentrations were measured by using a Coulter cell counter. Sperm were mixed in approximately equal concentrations and aliquoted into separate culture dishes. Eggs from different females were added to each dish, and fertilization was allowed to proceed for 30 min. All crosses were done within 10 min of

**This paper was submitted directly (Track II) to the PNAS office.**

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF186274 and AF186358).

*To whom reprint requests should be addressed: Department of Organismic and Evolutionary Biology, 16 Divinity Avenue, Harvard University, Cambridge, MA 02138. E-mail: spalumbi@oeb.harvard.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
gamete dilution. Crosses with the same sperm mixtures were started within 1 min of one another. Fertilized eggs were washed in filtered sea water and allowed to develop for 48 h at room temperature, at which time larvae were at an early pluteus stage. Eggs that did not show at least 95% fertilization were discarded.

Characterization of Bindin Alleles in Larvae. Single larvae were collected in 5 μl of filtered sea water, 5 μl of 4 mg/ml proteinase K was added, and the mixture was incubated at 70°C for 15 min and then 95°C for 5 min. The most variable section of the mature bindin coding region was amplified from 0.5 μl of the DNA extraction from each individual larva tested. Bindin allele clus-

Fig. 1. Allele genealogy for bindins in the sea urchin _E. mathaei_. Sequences from the first 105 codons of the mature bindin coding region show multicodon insertions and deletions that define major clades. Alleles not labeled “Guam” are from Hawaiian individuals. The tree is rooted to the sister species _E_. sp. nov. A and is a minimum length, strict consensus phylogram (PAUP 3.1.1). All polymorphic substitutions and insertions/deletions have a consistency index of 1.0. Polymorphic amino acid changes are denoted as asterisks.
ters A and B (Fig. 1) can be positively identified with DdeI or Apol restriction digestions of the associated PCR products, respectively. PCR product from each larva was digested with DdeI or Apol, and the restriction fragment length polymorphism patterns were visualized on ethidium bromide-stained 3% agarose gels. In each cross, individual genotypes of 40–70 larvae were recorded.

### Sperm Utilization Analysis

The percentage of larvae sired by each male was inferred from the restriction fragment length polymorphism patterns of the 40–70 larvae assayed. These percentages were compared with random expectation by a χ² test of the observed number of offspring of each male versus the number expected if sperm use was random. The expectations were based on the measured concentrations of sperm from each male used in each cross. Sperm fertilization probabilities corrected for differences in sperm density were calculated as SR/(SR + 1) where SR = (L1/L2)²(S2/S1). L1 and L2 are the number of larvae sired by males 1 and 2, respectively. S1 and S2 are the sperm concentrations from the two males.

### Results

#### Allele Genealogies

Among 87 bindin sequences from 53 individuals of the tropical sea urchin *E. mathaei*, alleles fall into four main classes that differ enormously in the 105-codon region of the protein under positive selection (Fig. 1). Allele clusters A and B are the most common, appear to be the most recently diverged, and make up >30–35% of the alleles in each population. The major clades show strong amino acid differentiation but few silent substitutions. Clades A and B differ by two separate deletions (of 5 and 8 codons, respectively, at positions 307–321 and 244–273) plus an insertion (5 codons, positions 94–108). These two clades differ at 18 of 105 aligned amino acid positions because of the insertion or deletion of groups of repeated amino acid motifs (Fig. 2). Of these 18 differences, 11 involve charged or polar residues, and clade B sequences have two more negatively charged amino acids than clade A sequences. There are no silent substitutions between clades A and B.

Within clusters A and B, most sequences have zero or one difference, suggesting they are very recently evolved. Of the three substitutions seen in more than one individual, two are amino acid replacements, and one is silent. Of the 63 individual sequences in these two clades, there were 41 singleton substitutions (not seen in any other individual). This rate of sequence heterogeneity (0.15%) is about that expected from incorporation of errors by *rTh* polymerase (30).

Clade C is defined by a deletion of a proline and a single silent substitution. Within this clade, all five phylogenetically informative nucleotide substitutions are replacement changes (Fig. 1, asterisks). Clades A, B, and C have diverged significantly from one another but cluster together in a clade defined by four substitutions, all of which are amino acid replacements. Finally, clade D is defined by two amino acid replacements and includes a single phylogenetically informative substitution, which is also a replacement.

In total, there are 15 polymorphic amino acid replacements and four polymorphic insertions/deletions that occur in >1 individual (Fig. 1). By contrast, there are only two silent polymorphisms that occur in more than one individual. Ignoring the insertions, the ratio of replacement to silent polymorphisms is slightly higher than neutral expectations (P = 0.04), consistent with the action of positive selection on divergence of alleles within species.

#### Functional Differences Among Alleles

To test whether bindin alleles in clades A and B are functionally different, sperm competition/egg choice experiments were conducted in which sperm from males of known bindin genotype were mixed and used to fertilize eggs of individual females. The resulting 48-h larvae were individually subjected to PCR and restriction digestion to identify their sires and to test paternity against expectations based on concentrations of sperm from each male in the original mixtures.

These experiments show that eggs are often fertilized non-randomly by sperm from different males and that males have different fertilization characteristics based on their bindin ge-
Figure 3. Relationship between fertilization and female genotype. Females with BB genotypes tended to be fertilized best by sperm from BB males, whereas AA females tended to use sperm of AA males. Points represent separate experiments from which 50–70 larvae were typed. Horizontal lines represent 95% confidence limits for detecting significantly nonrandom sperm choice in an experiment with an average number of typed larvae. Sample sizes: experiments with AA females = 7, 6, 5 (females, AA males, BB males); BB females = 4, 4, 4.

Table 2. Sperm use patterns depend on the similarity of male and female bindin genotypes

<table>
<thead>
<tr>
<th>Female genotype</th>
<th>Most similar</th>
<th>Least similar</th>
<th>Ties</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>16</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>AB</td>
<td>4</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>BB</td>
<td>8</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

In each cross, the male that shared the largest number of bindin alleles with a female is considered the most similar. Values represent the number of crosses in which most similar or least similar males showed significantly enhanced fertilization. Females with different bindin genotypes consistently tended to use sperm from males with the greatest bindin similarity.
What mechanisms could explain persistence of functional fertilization differences in *E. mathaei*? In the case of the bindin egg receptor system, at least three possible mechanisms exist for the maintenance of multilocus polymorphism. Heterozygote advantage may operate in bindin and would slow the loss of alleles (39). Although heterozygote males are frequently involved in ties, they are more successful than homozygote males in ~40% of crosses completed to date (Table 1). Mild heterozygote advantage may account for the maintenance of polymorphism and the fact that there is only a slight excess of homozygote bindin genotypes within populations (43% instead of the 37% predicted under Hardy–Weinberg assumptions; $n = 137, P = 0.07$). Heterozygote advantage at other loci under positive selection like MHC or the s-locus of plants is known to enhance polymorphism but also results in transspecies allele clades (26, 40–42). Such sharing of allele clades among species is not observed in bindin and is not predicted for loci thought to be intrinsic to reproductive isolation among species (16). However, it is possible that only certain bindin heterozygotes are at a reproductive advantage and that general models of balancing selection (39), in which all alleles are equally advantageous when ally different alleles at loci directly linked to fitness (2, 32). This view has been increasingly challenged as the male gamete interaction loci is associated with strong differences in reproductive success may be maintained without rapid selective sweeps. In this case, variation at female loci are intrinsic to fertilization choice patterns. Because male and female mating strategies are in conflict (43). When sperm densities are high, selection may favor eggs with rare receptor mutations that slow sperm entry to limit polyspermy (20). Selection on males to overcome this barrier may in turn lead to selection for mutant bindins. This molecular coevolution could generate positive selection of gamete recognition proteins (20), monophyly of species at recognition loci, and intraspecific polymorphism.

Eggs have often been thought of as passive acceptors of sperm (15), and fertilization has been characterized as a race among males (9). This view has been increasingly challenged as the genes involved in reproduction have become more widely understood (52). The current results show that positive selection on male gamete interaction loci is associated with strong differences among alleles in fertilization properties, but that both male and female loci are intrinsic to fertilization choice patterns. Because different bindin alleles are favored in interactions with eggs of different females, strong polymorphisms at loci intrinsic to reproductive success may be maintained without rapid selective sweeps. In this case, variation at female loci can help stabilize polymorphisms and allow the persistence of multiple, functionally different alleles at loci directly linked to fitness (2, 32). This population level sorting depends on both male and female genotype and may be the basis for some observed variation in reproductive function between individuals.

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence*</th>
<th>No. of sequences</th>
<th>No. of alleles</th>
<th>No. of variable amino acids</th>
<th>No. of variable indels</th>
<th>Positive selection?</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Echinometra mathaei</em></td>
<td>5'</td>
<td>85</td>
<td>15</td>
<td>16</td>
<td>4</td>
<td>Y</td>
<td>This study</td>
</tr>
<tr>
<td><em>Echinometra oblonga</em></td>
<td>5'</td>
<td>16</td>
<td>9</td>
<td>11</td>
<td>1</td>
<td>Y</td>
<td>21</td>
</tr>
<tr>
<td><em>Echinometra sp. A</em></td>
<td>5'</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>0</td>
<td>Y</td>
<td>21</td>
</tr>
<tr>
<td><em>Strongylocentrotus pallidus</em></td>
<td>FL</td>
<td>10</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>Y</td>
<td>48</td>
</tr>
<tr>
<td><em>Strongylocentrotus droebachiensis</em></td>
<td>FL</td>
<td>13</td>
<td>5</td>
<td>19</td>
<td>5</td>
<td>Y</td>
<td>48</td>
</tr>
<tr>
<td><em>Strongylocentrotus franciscanus</em></td>
<td>5'</td>
<td>134</td>
<td>14</td>
<td>9</td>
<td>0</td>
<td>N</td>
<td>49</td>
</tr>
<tr>
<td><em>Arbacia lixula</em></td>
<td>FL</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>N</td>
<td>46</td>
</tr>
<tr>
<td><em>Arbacia punctulata</em></td>
<td>FL</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>N</td>
<td>46</td>
</tr>
<tr>
<td><em>Arbacia incisa</em></td>
<td>FL</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>N</td>
<td>46</td>
</tr>
<tr>
<td><em>Arbacia dufresnei</em></td>
<td>FL</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>N</td>
<td>46</td>
</tr>
</tbody>
</table>

*Sequences compared are either the coding region 5' of the intron (5') or the complete coding region of the mature bindin gene (FL).
David Plouffe provided technical assistance. The manuscript was improved by discussions with P. Barber, S. Cohen, T. Duda, D. Haig, M.