BRIEF COMMUNICATIONS

Evolution, 59(3), 2005, pp. 682–685

EXPERIMENTALLY ENFORCED MONOGAMY: INADVERTENT SELECTION, INBREEDING, OR EVIDENCE FOR SEXUALLY ANTAGONISTIC COEVOLUTION?

WILLIAM R. RICE[1,2] AND BRETT HOLLAND[3,4]
[1]*Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106*
[2]*E-mail:rice@lifesci.ucsb.edu*
[3]*Department of Biological Sciences, California State University, Sacramento, California 95819*
[4]*E-mail: holland@csus.edu*

Abstract.—There has been recent criticism of experiments that applied enforced monogamous mating to species with a long history of promiscuity. These experiments indicated that the newly introduced monogamy reversed sexually antagonistic coevolution and caused males to evolve to be less harmful to their mates and females to evolve reduced resistance to harm from males. Several authors have proposed alternative interpretations of these experimental results based on qualitative analysis. If well-founded, these criticisms would invalidate an important part of the empirical foundation for sexually antagonistic coevolution between the sexes. Although these criticisms have a reasonable basis in principle, we find that after quantitative evaluation that they are not supported.

Key words.—Interlocus contest evolution, monogamy, sexual conflict, sexually antagonistic coevolution.

Received September 8, 2004. Accepted December 6, 2004.

One of the major components of the empirical evidence for antagonistic coevolution between the sexes is based on the evolution that occurs when a population with a long history of promiscuous mating is experimentally made to be monogamous. If there is an arms race between the sexes in promiscuous species, as is predicted by some theory, then experimentally enforced, life-long monogamy should lead to males evolving to become more benign to females, and females evolving to become less resistant to male-induced harm. These predictions were confirmed in recent experiments with two different model systems (Holland and Rice 1999; Hosken et al. 2001). However, the original interpretation of these experiments has recently been criticized by several authors, but most emphatically in this journal by Wigby and Chapman (2004) who wrote,

> ''A potential problem with this (Holland and Rice 1999) and similar studies (Hosken et al. 2001) is that the effective population size of monogamy lines was smaller than that of control, polyandry lines. Thus, some findings, for example, reduced male competitiveness and reduced male harm to females under monogamy, are also predicted by the higher inbreeding and reduced body size in these lines (Sharp 1984; Pitnick et al. 2001; Snook 2001; Pitnick and Garcia-Gonzalez 2002; Chapman et al. 2003b). This makes the conclusions of these studies equivocal.''

We appreciate the need to consider the alternative explanations that have been proposed. However, until these alternatives are rigorously and quantitatively evaluated, we think that it is premature to conclude that the original conclusions are equivocal. Moreover, below we provide evidence that differences in inbreeding and body size had no important consequences for the interpretation of past studies concerning the influence monogamy on sexually antagonistic coevolution.

Here we use quantitative analyses to address the qualitative arguments that underlie the recent criticisms of the experiments on experimentally enforced monogamy. There are two major issues: (1) inadvertent selection for small body size, rather than the monogamy treatment itself, is a more likely explanation for the observed reduction in the degree to which males harm their mates and the degree to which female resistance declined in the monogamy lines, and (2) the monogamy treatment was significantly confounded by increased inbreeding, and therefore inbreeding, rather than monogamy, can better explain the observed results. We will address these issues separately. Our focus will be on the data with which we are most familiar, that from our experiments with the *D. melanogaster* model system (Holland and Rice 1999), but our conclusions also are relevant to the criticisms of the work by Hosken et al. (2001) with the dung fly (*Scathopaga stercoraria*) model system.

The basic experimental design was to establish four lines that were propagated for 47 generations by combining the progeny from 100 females per line. The two monogamy lines were reared with one female and one male per vial, and the two control lines with one female and three males per vial. In the control lines both polyandry and male-male competition were present. Theses two attributes were present in the promiscuous base population used to begin the experiments and, in theory, sustain sexually antagonistic coevolution. In contrast, both polyandry and male-male competition were absent in the monogamy lines, so that sexually antagonistic coevolution should be reversed and mutualistic coevolution should ensue. At the end of the experiments we measured how harmful males were to females and how resistant females were to male-induced harm.

Inadvertent Selection on Body Size

Pitnick and Garcia-Gonzalez (2002) suggested that the protocol used by Rice and Holland (1999) to construct experimental monogamous lines would indirectly select for small body size, and that smaller males would be expected to be

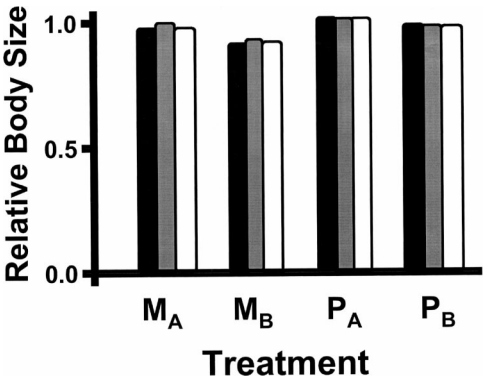

FIG. 1. The average weight of the two experimental monogamy (M) and two control populations (C). Males (black), females (gray), and both sexes combined (open).

less harmful to females. The indirect selection for smaller males was suggested to occur in the monogamy lines because of the removal of male-male competition coupled with selection for rapid development in both treatments (Pitnick, et al., 2001; Pitnick and Garcia-Gonzalez 2002). We have reservations about the conclusion that smaller males are less harmful to females because it is based on phenocopied males (made small through nutritional stress) and hence the stress rather than small body size may be responsible for the observed reduction in harm by a male to his mate. However, we do agree that this alternative hypothesis should be considered. Fortunately this can be done from data on body weight collected from the original monogamy experiments. These body size measurements are the most appropriate data because they were collected at the same time that the flies were assayed for male-induced harm to females and female resistance to this harm.

Average body weights of males and females from the four experimental lines are shown in Figure 1. It can be seen that average body size did not vary substantially between the experimental treatments, as indicated by the small coefficient of variation (4.1% or less for both sexes). Also, the average body weights among the two monogamy and two promiscuous lines overlapped for females and nearly overlapped for males. Using populations as the units of replication, as was done in all tests within the original manuscript, there is no significant difference between the monogamy and control treatments for average body size of males or females. Nonetheless, the smallest estimated mean body size for males and females was from the B replicate of the monogamy line.

If inadvertent selection on body size, rather than the monogamy treatment, was the cause for the observed experimental results, then body size should be a stronger predictor (compared to the experimental treatments of monogamy vs. control) in the statistical models used to analyze the data from the experiment. We first tested to see if body size was a statistically significant covariate to the treatment effect in any of the tests reported in Holland and Rice (1999), and found that in no case was this true. However, multicolinearity between treatment and body size may have reduced statistical power to detect body size as a significant covariate. As an additional check we have reanalyzed all of the statistical tests reported in Holland and Rice (1999) with both the original statistical models and then again with body size replacing the experimental monogamy/control treatment. The original statistical models were all of the form, $Y_{ij} = u + \text{treatment}_i + \text{error}_{ij}$, and the alternative model using body weight as a covariate is, $Y_i = b_0 + b_1(\text{average weight})_i + \text{error}_i$. The number of parameters in these two statistical models, as well as the degrees of freedom for the mean square error, are identical, thus their intrinsic statistical power should be the same.

No tests that were not statistically significant became so when body size replaced the experimental monogamy/control treatment. There were seven statistically significant tests reported in our original article. Results of these same tests using both experimental treatment (monogamy vs. control) and body size are listed in Table 1. Of the seven significant patterns reported in the original article, body size was a significant predictor in only one case, that is, net reproductive rate—a measure that was not central to the conclusion that monogamy lead to less harmful males and less resistant females. Because these tests were independent, the probability of getting at least one spurious significant association with body size by chance alone is 30.1%, thus the single significant relationship between experimental outcome and body size is not unexpected. In summary, quantitative analysis indicates that there is no empirical corroboration for the idea that inadvertent selection on body size, rather than the experimental treatment of monogamy versus promiscuity, was an important factor associated with the patterns observed in Holland and Rice (1999).

TABLE 1. Comparison of statistical test with models based on treatment (monogamy vs. control) and average body size. The covariate *body size* was average male weight except in test 5 where it was average female weight and tests 6–7 where it was the average weight across both sexes.

Population parameter	Significance	
	Treatment model	Body size model
1. Reproductive rate of test females	*	ns ($P = 0.293$)
2. Survival of tester females	*	ns ($P = 0.100$)
3. Courtship rate	*	ns ($P = 0.197$)
4. Survival of experimental females	**	ns ($P = 0.153$)
5. Egg mass after mating ancestral males	**	ns ($P = 0.365$)
6. Net reproductive rate	*	*
7. Development rate	***	ns ($P = 0.094$)

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$.

Differences in Inbreeding Depression

The second issue is whether or not the higher level of inbreeding in the monogamy treatment is a reasonable alternative explanation for the patterns observed in both Holland and Rice (1999) and Hosken et al. (2001). The qualitative observation that inbreeding was higher in the monogamy lines is certainly correct, but an evaluation of the biological significance of this observation requires quantitative analysis.

To see why qualitative analysis is insufficient, consider an extreme case where the effective size of one treatment was one million and that of the other was two million—a twofold difference in effective population sizes. Because of the large effective population sizes of *both* treatments, very little inbreeding depression would occur in populations of both sizes, so that large differences in male harm or female resistance could not feasibly be attributed to differences in their effective sizes. The differences between monogamy and control treatments reported in Holland and Rice (1999) were all sufficiently large to be detected despite low statistical power owing to the fact that the sample sizes were only two for each treatment (to avoid pseudoreplication, Holland and Rice [1999] analyzed the two line means for each treatment), and the degrees of freedom in the mean square errors for all statistical models were only two. Below we use quantitative analysis to show that the large differences that we observed between monogamy and control lines cannot feasibly be attributed to difference in the effective population size of the two treatments, because effective populations for both treatments were sufficiently large.

To estimate and compare effective population sizes of the monogamous and control populations we used the general relationship, $N_e = 4N_MN_F/(N_M + N_F)$, where N_e is the effective population size, N_M is the number of breeding males and N_F is the number of breeding females. Because in the monogamy lines every female mated a unique male, and because she was guaranteed to contribute to the pool of offspring that was randomly sampled to begin the next generation, $N_M = N_F = 100$ (i.e., the number of fertile females used propagate each generation) and $N_e = 200$. Note that extra vials of males and females were made as replacements in both experimental treatments to guarantee that infertility did not reduce below 100 the number of families contributing to each generation below 100.

In the control lines females were housed separately with three males per female, polyandry was possible, and the effective population size would be expected to increase. The influence of polyandry on N_e depends on the number of mates per female and the degree of sperm displacement.

Displacement by the last male to mate a female *D. melanogaster* averages 81% in the LH_M base population that was used to start the enforced monogamy experiment (unpubl. data). Preliminary results from an experimental examination of sperm displacement in multiply mated females, indicate that triple matings are rare in the LH_M base population (and presumably would have been rare in the monogamy experiments) and that triply mated females use stored sperm from three sequential mates in quantities that closely approximate those predicted based on multiplicative, mass-action sperm displacement (A. Stewart, E Morrow, and WRR, unpubl. data).

To calculate the effective size of the control lines we: (1) conservatively assume that all females mate with all three males, and (2) used the empirical estimate from the LH_M-base population that sperm displacement averages 81%. In this case a proportion P_1 (0.036) of fertilizations comes from a pool of 100 first males to mate females, P_2 (0.154) of fertilizations comes from a pool of 100 second males to mate females, and P_3 (0.810) of fertilizations comes from a pool of 100 third males that are last to mate females.

We next extend the general formula for effective population size with separate sexes by incorporation polyandry. The probability that two randomly selected alleles are derived from the same female in the previous generation is $1/4N_F$, the probability that the two alleles come from the same male is $P_1^2*(1/4N_F)$ for the first sires, $P_2^2*(1/4N_F)$ for the second sires, and $P_3^2*(1/4N_F)$ for last sires, where N_F is the number of females used to propagate the control line, and first, second and third sires are a female's three consecutive mates. In general $(1/N_e)$ = [probability that two randomly selected alleles are derived from the same female] + [probability that two randomly selected alleles are derived from the same male]. Expressing this same relationship for the experimental design of the control treatment, $(1/N_e) = [(1/4N_F) + [(P_1^2/4N_F) + (P_2^2/4N_F) + (P_3^2/4N_F)]$. Solving for N_e we obtain, $N_e = 4N_FN_F/[N_F(P_1^2 + P_2^2 + P_3^2) + N_F]$ in general, and specifically $N_e = 238.7$ for our control lines. Put another way, the effective population size of the monogamy lines was at most only 16.2% smaller than that of the control populations.

Inbreeding could only contribute to the observed patterns in the experiments of Holland and Rice (1999) if inbreeding depression, after 47 generations, was substantially larger in the monogamy lines. Inbreeding depression accrues due to the accumulation of deleterious mutations because of sampling error (genetic drift) overpowering selection. To a first approximation, harmful mutations will accumulate only when the absolute value of their selection coefficient is less than the reciprocal of the effective population size, that is, when $|s| < 1/N_e$. (Kimura 1983, pp. 44–48). This implies that harmful mutations with a selection coefficient of $|s| < 1/200 = 0.005$ will accumulate in the monogamy lines whereas those with $|s| < 1/238.7 = 0.0042$ will accumulate in the promiscuous lines. Put another way the difference in the upper-bound of the spectrum of deleterious mutations that would be expected to accumulate in the monogamy versus control lines is only $0.005 - 0.0042 = 0.0008$, which is quite small. This quantification makes it clear that virtually all mutation that would accumulate in the monogamy lines, due to drift overpowering selection, also would accumulate in the promiscuous lines.

Effective population size also influences the accumulation of inbreeding depression in a population by its effect on the speed with which harmful mutations accumulate. The expected time to fixation of a mildly deleterious mutation that is destined to be fixed is minimally $4N_e$ generations, which is the expected time to fixation of a neutral mutation (Kimura 1983, p. 49). Because the effective size of the monogamy populations was approximately 16.2% smaller, those muta-

tions that do fix would be expected to do so 16.2% faster. Because the experiment lasted 47 generations and because the expected time to fixation is much longer (≥800 generations), very little accumulation of harmful mutations would be expected over the course of the experiment in either the monogamy or control populations. Furthermore, the level accumulation of any harmful mutations that did occur would be nearly identical between the monogamy and control populations. To put this into perspective, assume that a rare mildly deleterious mutation in a monogamy or promiscuous line was drifting toward fixation because drift was overpowering selection. To a conservative first approximation for a rare mutation, its expected progress toward fixation (Δp) would be no more than $\Delta p = 47/4N_e = 47/954 = 0.492$ in one of the control populations, whereas that of the same mutation in one of the monogamous populations would be $\Delta p = 47/4N_e = 47/800 = 0.0588$. The approximation is conservative because we have assumed a constant rate of accumulation whereas Δp varies with gene frequency and it is much slower when the mutation is rare, as would be expected for harmful alleles in mutation-selection balance at the start of our experiments. The difference is only $0.0588 - 0.0492 = 0.0096$. In other words, we expect virtually identical, and very small, levels of accumulation of rare harmful mutations in both the monogamous and control populations.

These calculations are based on many simplifying assumptions but they nonetheless illustrate that the spectrum of mutations accumulating, and the degree of accumulation in the time span of the experiments, would be expected to be nearly identical in the two experimental treatments. As a consequence, different levels of inbreeding are not reasonable explanations for the large difference observed in the published studies of experimentally enforced monogamy. This interpretation is empirically corroborated by the observation that the net reproductive rates of the monogamy lines at the end of the experiment were significantly higher than those of the control lines (not lower as predicted by excess accumulation inbreeding depression in the monogamy lines). Lastly, the same logic regarding inbreeding apples to the experiments on dung flies by Hosken et al. (2001).

Conclusions

In principle, it is reasonable to suggest that inadvertent selection on body size and different levels of inbreeding provide potential alternative explanations for the results of recent experiments with experimentally enforced monogamy. However, before concluding that ''This makes the conclusions of these studies equivocal'' the qualitative logic underpinning these suggestions needed to be quantitatively evaluated. Doing so, we find no support for either suggestion and conclude that results from the monogamy experiments provide valuable evidence for the operation of sexually antagonistic coevolution.

Acknowledgements

We thank the two anonymous referees for comments on an earlier draft of the manuscript. This work was supported by three grants from the National Science Foundation to WRR (DEB-9996164, DEB-0128780, and DEB-0410112).

Literature Cited

Chapman, T., G. Arnqvist, J. Bangham, and L. Rowe. 2003. Sexual conflict. Trends Ecol. Evol. 18:41–47.

Holland, B. and W. R. Rice. 1999. Experimental removal of sexual selection reverses intersexual antagonistic coevolution and removes a reproductive load. Proc. Natl Acad. Sci. USA 96: 5083–5088.

Hosken, D. J., T. W. J. Garner, and P. I. Ward. 2001. Sexual conflict selects for male and female reproductive characters. Curr. Biol. 11:489–493.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, New York.

Pitnick, S., G. T. Miller, J. Reagan, and B. Holland. 2001. Males' evolutionary responses to experimental removal of sexual selection. Proc. R. Soc. Lond. B 268:1071–1080.

Pitnick, S., and F. García-González. 2002. Harm to females increases with male body size in *Drosophila melanogaster*. Proc. R. Soc. Lond. B 269:1821–1828.

Sharp, P. M. 1984. The effect of inbreeding on competitive male mating ability in *Drosophila melanogaster*. Genetics 106: 601–612.

Snook, R. R. 2001. Sexual selection: conflict, kindness and chicanery. Curr. Biol. 11:R337–R341.

Wigby, S., and T. Chapman. 2004. Female resistance to male harm evolves in response to manipulation of sexual conflict. Evolution 58:1028–1037.

Corresponding Editor: R. Harrison