

Hybrid vigor between native and introduced salamanders raises new challenges for conservation

Benjamin M. Fitzpatrick*[†] and H. Bradley Shaffer[‡]

*Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996; and [‡]Evolution and Ecology, University of California, Davis, CA 95616

Edited by John C. Avise, University of California, Irvine, CA, and approved August 10, 2007 (received for review May 22, 2007)

Hybridization between differentiated lineages can have many different consequences depending on fitness variation among hybrid offspring. When introduced organisms hybridize with natives, the ensuing evolutionary dynamics may substantially complicate conservation decisions. Understanding the fitness consequences of hybridization is an important first step in predicting its evolutionary outcome and conservation impact. Here, we measured natural selection caused by differential viability of hybrid larvae in wild populations where native California Tiger Salamanders (*Ambystoma californiense*) and introduced Barred Tiger Salamanders (*Ambystoma tigrinum mavortium*) have been hybridizing for 50–60 years. We found strong evidence of hybrid vigor; mixed-ancestry genotypes had higher survival rates than genotypes containing mostly native or mostly introduced alleles. Hybrid vigor may be caused by heterozygote advantage (overdominance) or recombinant hybrid vigor (due to epistasis or complementation). These genetic mechanisms are not mutually exclusive, and we find statistical support for both overdominant and recombinant contributions to hybrid vigor in larval tiger salamanders. Because recombinant homozygous genotypes can breed true, a single highly fit genotype with a mosaic of native and introduced alleles may eventually replace the historically pure California Tiger Salamander (listed as Threatened under the U.S. Endangered Species Act). The management implications of this outcome are complex: Genetically pure populations may not persist into the future, but average fitness and population viability of admixed California Tiger Salamanders may be enhanced. The ecological consequences for other native species are unknown.

Ambystoma | fitness | hybridization | invasive species | genetics

Hybridization (interbreeding between differentiated lineages) occurs in almost all sexually reproducing groups of organisms (1–3). Consequences of hybridization include fusion of previously distinct lineages, extinction or local extirpation of one or both lineages, evolution of reproductive isolation via reinforcement, and production of novel, highly fit hybrid phenotypes. These outcomes depend on the distribution and inheritance of fitness among hybrids and often have important implications for both evolutionary and conservation biology.

In evolutionary biology, hybrid fitness and its genetic basis are important for understanding the evolution of reproductive isolation and the consequences of horizontal gene flow on the diversity and complexity of life (4–7). In conservation biology, hybrid fitness is a key to the mechanistic basis of inbreeding and outbreeding depression, the potential loss of biodiversity due to genetic swamping, and the evolution of invasiveness (8–11). These considerations become particularly important when introduced organisms successfully hybridize with natives.

Unlike cases of natural hybridization, some outcomes of human-mediated hybridization are generally considered less desirable than others. Low fitness of hybrids can reduce mean population fitness, rendering populations more vulnerable to extinction (8, 12). Novel hybrid genotypes may contribute to *de novo* evolution of invasiveness (9), and admixed populations may be considered less valuable than an authentic native gene pool (10). On the other hand, admixture may rescue declining native

populations by alleviating inbreeding depression (13, 14) or facilitating adaptive evolution in modified or degraded habitats (15–17).

The long-term consequences of hybridization are strongly influenced by the genetic basis of hybrid fitness. In the case of hybrid vigor, genetic models fall into two classes: heterozygote advantage and recombinant hybrid vigor (18–20). Heterozygote advantage (overdominance) refers to beneficial interactions between heterospecific alleles of a single locus. Recombinant hybrid vigor depends on multilocus genotypes and may be caused by epistasis (beneficial interactions between heterospecific alleles from different loci) or by complementary effects of independent advantageous alleles from each parental population (19, 20). For example, if different loci have fixed deleterious alleles in each parental population, hybrids bearing the superior allele at each locus will have higher fitness than either parent.

Unlike heterozygote advantage, where the most fit genotype cannot breed true, recombinant hybrid vigor favors the fixation of a single recombinant genotype bringing together beneficial alleles from each parent. This new, true breeding genotype may be considered a modified version of one of the parental lineages or a new hybrid species. In conservation terms, fixation of a recombinant genotype may represent a minor evolutionary modification or total genetic “extinction” of a native form (this may depend more on one’s perspective than on the genetic details). Perhaps more important, the new form may have undesirable ecological effects on native communities (9).

Recent, human-mediated hybridization between native California Tiger Salamanders (*Ambystoma californiense*) and introduced Barred Tiger Salamanders, (*Ambystoma tigrinum mavortium*) constitutes an important case study of the evolutionary and conservation consequences of hybridization (21–23). Before the introduction 50–60 years ago (by bait dealers interested in selling salamander larvae to bass fishermen), the two lineages had been geographically separated for ≈ 5 my (24, 25). Despite such a long period of divergence, they hybridize readily and have formed a hybrid swarm in the Salinas Valley of California (21–23). Multiple introduction points, combined with subsequent dispersal and movements of salamanders across the landscape (23) have led to broad-scale admixture across 15–20% of the natural range of the native California Tiger Salamander. The recent listing of *A. californiense* as Threatened under the US Endangered Species Act (26) recognized hybridization as a major conservation concern.

In this study, we estimated viability of hybrid tiger salamander larvae over the first few weeks after hatching. This period of high mortality represents an enormous opportunity for natural se-

Author contributions: B.M.F. and H.B.S. designed research; B.M.F. and H.B.S. performed research; B.M.F. analyzed data; and B.M.F. and H.B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[†]To whom correspondence should be addressed. E-mail: benfitz@utk.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704791104/DC1.

© 2007 by The National Academy of Sciences of the USA

Table 1. Multivariate analyses of the difference between hatchlings and survivors

Pond	BW1		CVP		Pond F		Pond G		JCL2	
	Coefficient	P	Coefficient	P	Coefficient	P	Coefficient	P	Coefficient	P
Intercept	1.151		-0.598		2.777		1.188		3.376	
θ_S	1.911	0.002	13.744	0.000	4.246	0.000	-4.641	0.002	2.816	0.267
θ_H	3.508	0.000	2.076	0.000	5.190	0.000	3.079	0.000	5.924	0.000
θ_S^2	-1.756	0.088	-12.368	0.002	-6.276	0.002	-7.788	0.002	-6.335	0.042

A positive coefficient for θ_S indicates selection favoring nonnative alleles, a positive coefficient for θ_H indicates selection favoring heterozygous hybrid genotypes, and negative coefficients for θ_S^2 indicate selection favoring mixed genotypes regardless of heterozygosity. These heuristic interpretations of the coefficients are correct when all else is equal; refer to Fig. 1 to see actual changes in genotype distributions.

lection to operate on the diverse genotypes of a hybrid population (27). Observations were made in five wild breeding populations including the three major types of breeding habitats found in the Salinas Valley: natural vernal pools, seasonal cattle ponds, and perennial ponds (22). We categorized individual hybrids using multilocus genotypes based on nine physically unlinked molecular markers. Each individual was characterized as having native or introduced mtDNA and as being homozygous native, heterozygous, or homozygous introduced at each nuclear marker (22). Thus, our approach has the advantages of considering the fitness of genotypes defined more finely than simple pedigree categories (F_1 , backcross, etc.) and of estimating fitness in the wild across the range of larval habitats found within the hybrid swarm (28, 29). We focused on how early larval survival varies among hybrid genotypes in the wild. Specifically, is the pattern of survival consistent with hybrid dysfunction or hybrid vigor, and does fitness variation depend more on additive, dominant (heterozygous), or recombinant genetic effects?

Results

Generalized linear models relating survival probability to molecular marker-based hybrid indices clearly support hybrid vigor (Table 1 and Fig. 1). Survival was higher for individuals with high heterozygosity, θ_H , and intermediate ancestry, θ_S (see *Methods*). Negative coefficients for θ_S^2 indicate that surviving larvae tended to be of less pure and more mixed ancestry. A more complex model (with θ_H^2 and $\theta_S \times \theta_H$ terms) did not fit the data better than the reduced model shown in Table 1 [supporting information (SI) Table 2]. The significant effect of mixed ancestry over and above the effect of heterozygosity suggests that recombinant genotypes have enhanced survival that is not predicted by heterozygote advantage alone.

Details of the multivariate outcome differed among ponds (Table 1), but this is attributable to differences in initial genotype frequencies. Ponds that started highly introduced became more native and ponds that started highly native became more introduced (Fig. 1). This is consistent with selection favoring mixed ancestry over either pure native or pure introduced genotypes. Fig. 1 shows triangular regions of genotype space defined by the interdependence of ancestry and heterozygosity. Pure native or introduced genotypes, by definition, have no heterozygous loci. Individuals with all heterozygous loci must have evenly mixed ancestry, but mixed ancestry can also be achieved when some loci are homozygous for native alleles and others are homozygous for introduced alleles. Fig. 1 shows the density of individuals across this triangular genotype space before (hatchlings) and after (larvae) natural selection. In every pond, there is a clear shift toward a higher frequency of mixed-ancestry salamanders. That is, in all five ponds, the multilocus ancestry index shifts toward a more intermediate value and away from initially more extreme values. This was true regardless of whether the before-selection ancestry indices were fairly evenly distributed (Pond F) or skewed toward native (Pond G) or nonnative (BW1, CVP, JCL 2) ancestry values (Fig. 1).

Changes in single-locus allele frequencies were idiosyncratic among markers and ponds (SI Table 3). However, most genotype frequencies (28/40 marker-pond combinations) showed a trend in the direction of increased frequency of heterozygotes after

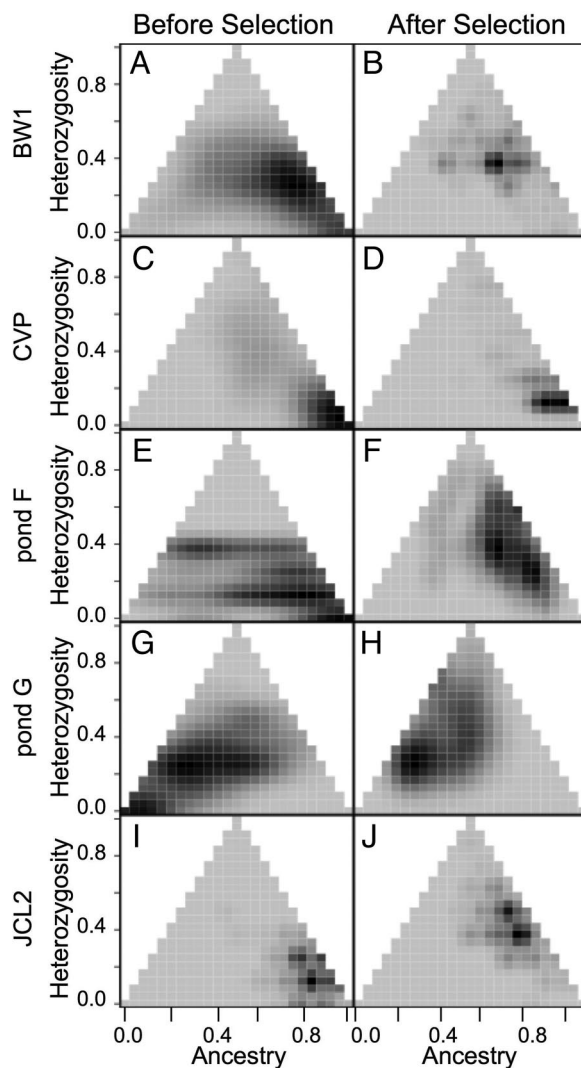


Fig. 1. Approximate densities of multilocus genotypes in hatchlings (before selection) and larvae (after selection) in five populations of hybrid tiger salamanders were estimated by using a two-dimensional density-estimation procedure with a bivariate normal kernel [function `kde2d` in R (30, 31)]. Gray triangles represent potential genotype space described by ancestry (the fraction of introduced alleles in an individual’s marker genotype) and heterozygosity (fraction of markers heterozygous for native and introduced alleles). Darker areas are those with larger numbers of observed individuals.

selection, supporting the interpretation of hybrid vigor (SI Table 3). One possible exception is the *GNAT1* marker, for which the dominance coefficient was negative (albeit not statistically significant) in all five ponds, suggesting low fitness of heterozygotes for this genomic region despite the overall pattern of hybrid vigor (SI Table 3).

Discussion

Our study of early-larval survival in hybrid tiger salamanders reveals higher fitness of hybrid genotypes with evenly mixed ancestry than genotypes composed predominantly of alleles from one parental lineage. Such strong evidence of hybrid vigor was unexpected based on previous studies suggesting constraints on admixture, inferred from high levels of linkage disequilibrium and habitat-dependent invasion success (21, 22). Our statistical analyses suggest that hybrid survival is influenced by both heterozygote advantage and recombinant hybrid vigor. Correct interpretation of the underlying genetics is crucial to predicting the evolutionary outcome of admixture between native California Tiger Salamanders and introduced Barred Tiger Salamanders. From a conservation perspective, hybrid vigor may be viewed as good or bad for the California Tiger Salamander, depending on whether one chooses to emphasize fitness of extant populations or genetic purity as a primary conservation goal.

The Genetic Basis of Hybrid Vigor. When considering the genetic basis of hybrid vigor, it is important to recall that our markers represent a very sparse sample of the genome. Our ability to derive information from this set of markers is enhanced by extensive linkage disequilibrium due to admixture in hybrid populations (22). Admixture linkage disequilibrium decreases the likelihood that a particular marker–trait association represents close physical linkage between the marker and a quantitative trait locus (32). By the same token, admixture linkage disequilibrium increases the accuracy of our marker-based estimates of genomic ancestry (θ_S) and heterozygosity (θ_H) because the sampled loci are correlated with many other regions of the genome (33, 34).

Our analyses show a strong positive relationship between heterozygosity and probability of survival (Table 1 and Fig. 1). This result is consistent with heterozygote advantage (overdominance for fitness) contributing to the genetic basis of hybrid vigor. The coefficient on the “admixture” term (θ_S^2) is consistently negative, implying that recombinant hybrid vigor also plays a role. Recombinant hybrid vigor may be caused by synergistic interactions between heterospecific alleles at different loci (epistasis) or complementary effects of superior introduced alleles at some loci and superior native alleles at other loci (19, 20). That is, θ_S^2 is better viewed as an indicator of overall genomic admixture rather than additive \times additive interaction *per se*.

Hybrid vigor in other systems has often been attributed to heterozygote advantage (35–39). However, heterozygote advantage is lost each generation because of segregation among gametes. If the most-fit genotype is heterozygous at many loci, it cannot breed true and will rarely, if ever, reappear in generations beyond the F_1 (40). In contrast, if hybrid vigor is caused by complementary or epistatic genes, the most-fit genotype is homozygous for alleles derived from each lineage at different loci (19). Such recombinant homozygotes appear only in later generations, and they can breed true. True breeding recombinant genotypes can therefore establish new lineages with genomes that are mosaics of the two ancestral genomes (41–44).

If both heterozygote advantage and recombinant hybrid vigor are operating in the *Ambystoma* hybrid zone, then one likely outcome is a population of tiger salamanders that is fixed for native alleles at some loci, advantageous introduced alleles at other loci, and segregating native and introduced alleles as

balanced polymorphisms at still other loci. However, the statistical pattern of high mean fitness of highly heterozygous genotypes may be transient, a temporary consequence of the current levels of admixture and linkage disequilibrium that may abate as the population genetic background changes from a highly variable hybrid swarm to a more stable mosaic of native and introduced alleles in different parts of the genome. As the population becomes fixed for a single admixed genetic background, the heterozygote advantage at a given locus may disappear (SI Text).

Variation, linkage disequilibrium, and divergent allele frequencies between habitats have been maintained over 12–24 generations of admixture between *A. californiense* and *A. t. mavortium* in the Salinas Valley (22, 23). However, hybrid vigor tends to maintain genetic variation, decrease linkage disequilibrium between conspecific alleles, and homogenize allele frequencies across habitats. All of these processes are illustrated in Fig. 1. Three processes are likely working against the tendency of hybrid vigor to bring about a stable distribution of mixed-ancestry genotypes. First, genetic drift contributes to variance among breeding ponds (45). Second, higher frequencies of introduced alleles are maintained in perennial than in seasonal breeding ponds (22). Finally, immigration from pure native populations outside of the hybrid swarm tends to increase native allele frequencies and contribute to a broad-scale geographic gradient in allele frequencies (23).

The high frequencies of introduced alleles in perennial ponds remains one of the strongest genetic patterns in the tiger salamander hybrid system; it may be maintained by habitat choice or selection in later life history stages. Fitzpatrick and Shaffer (22) hypothesized that introduced genotypes gain an advantage in perennial water bodies by extending their larval period or even breeding in the larval (paedomorphic) condition. Pure native California Tiger Salamanders must metamorphose to reproduce (46), but Barred Tiger Salamanders regularly forgo metamorphosis in perennial ponds and breed as paedomorphs (47, 48). Paedomorphs often reach sexual maturity earlier than metamorphs, produce larger clutches, and may breed earlier in a given season; any of these life-history factors may provide introduced genotypes with an advantage in perennial ponds (49, 50). The perennial pond populations studied here (BW1 and JCL2) show strong viability selection favoring hybrids during the early part of the larval period (Table 1 and Fig. 1). Antagonism between hybrid vigor for larval survival in all ponds and later selection favoring introduced genotypes in perennial ponds may explain the persistent pattern of high introduced allele frequencies in perennial ponds and relatively even native and introduced allele frequencies in seasonal ponds throughout the range of the hybrid swarm (23).

Conservation Consequences of Hybrid Vigor. Hybridization between an introduced form (*A. t. mavortium*) and a declining native protected by the Endangered Species Act (*A. californiense*) raises several difficult conservation issues (8, 10, 23, 51). These include, (i) what are the consequences of hybridization for population viability? (ii) what are the consequences for other native species and the communities in which they occur? (iii) which individuals or populations of salamanders should be protected? and (iv) should any individuals or populations be eradicated? The answers to *iii* and *iv* depend, in part, on the answers to *i* and *ii* but may also depend on practical considerations and perspectives on the merit of genetic purity as a conservation goal. Our results have important implications for population viability and genetic purity of California Tiger Salamanders and for potential impacts on other organisms native to California ponds and grasslands.

Hybrid vigor is probably not detrimental to population viability. However, some authors have raised the concern that negative effects of introgressive hybridization on fitness may not

additive, dominance, and epistatic effects on hybrid fitness as terms in a linear statistical model. In our case,

$$\log\left(\frac{P}{1-P}\right) = \text{intercept} + \alpha_1\theta_S + \delta_1\theta_H + \alpha_2\theta_S^2 + \delta_2\theta_H^2 + (\alpha\delta)_1\theta_S\theta_H + \text{error}, \quad [2]$$

where individual genotypes are described by θ_S , a linear ancestry index ranging from -1 (all native alleles) to $+1$ (all introduced alleles) and θ_H , a heterozygosity index ranging from -1 (all markers homozygous) to $+1$ (all markers heterozygous). The fitted coefficients α_1 and δ_1 describe simple additive and dominance effects, whereas α_2 , δ_2 , and $(\alpha\delta)_1$ describe additive \times additive, dominance \times dominance, and additive \times dominance epistasis, respectively (18). This model was fitted to the data from each pond by using the `glm()` function with a binomial error distribution and observations weighted by stratum, as for the single-locus case (see below).

Two additional issues needed to be addressed in our statistical analyses. First, an interesting problem was created by the temporally extended breeding season in three of our study ponds (BW1, G, and JCL2). In these ponds, breeding consisted of three to five major episodes corresponding to significant rain storms; each episode potentially involved a different set of adults with different genotype frequencies. Thus, our before-selection samples were stratified simple random samples. Each weekly sample (stratum sample) was weighted by stratum size when estimating before-selection genotype frequencies and their moments (62). We used the total number of eggs laid on our artificial oviposition sites each week as that week's stratum size (we assume this number is proportional to the total number of eggs produced in the pond that week). Each before-selection observation (hatchling) was weighted according to stratum size [(stratum size \div stratum sample size) \times (total sample size \div sum of stratum sizes)] and those weights passed to the `glm()` function. Under this weighting scheme, the after-selection individuals each get a weight of 1.0 because they come from a single stratum, and the averages of the individual weights before and after selection are equal (62).

Second, traditional estimates of confidence intervals for regression coefficients do not account for error in the independent variables. However, θ_S and θ_H were estimated with error because our markers are samples of the genome ($\theta_S = 2P_T - 1$, and $\theta_H = 2P_{CT} - 1$, where P_T is the fraction of an individual's marker alleles derived from *A. t. mavortium*, and P_{CT} is the fraction of markers heterozygous for native and introduced alleles). Therefore, we used a bootstrap and randomization procedure to incorporate this sampling variance into hypothesis tests regarding the fitted coefficients described in Eq. 2. For each of 10,000 replicates, we performed a two-step Monte Carlo procedure.

First, we bootstrapped across nuclear markers. Bootstrapping across markers addresses error in individual estimates of θ_S and θ_H due to sampling only a fraction of each salamander's genome. This is in the same spirit as bootstrapping across characters in phylogenetic analysis (63) and results in each individual having associated with it a distribution rather than a single point estimate of each θ . mtDNA was treated as a fixed effect; it was included once and only once in every replicate, because there is only a single mitochondrial allele to sample. Excluding mtDNA slightly affected parameter estimates but not their signs or P values. Including mtDNA in the pool of markers to resample is inappropriate because it does not contribute to θ_H . The second step in each replicate was to randomize observations to simulate the null hypothesis that hatchlings and surviving larvae were random samples from a single distribution of genotypes. We recorded the fraction of replicates giving coefficients larger (f_{Larger}) and smaller (f_{Smaller}) than the estimate from the original data and estimated two-tailed P values as $2 \times \text{minimum}(f_{\text{Larger}}, f_{\text{Smaller}})$. The procedure was applied separately to each of our five study ponds.

The full Lynch (18) model (Eq. 2) may overfit some data sets (particularly those with highly skewed allele frequencies), leading to problems with colinearity of variables and high variances around coefficient estimates. Therefore, we also fitted a reduced model:

$$\log\left(\frac{P}{1-P}\right) = \text{intercept} + \alpha_1\theta_S + \delta_1\theta_H + \alpha_2\theta_S^2 + \text{error}. \quad [3]$$

This model focuses on fitness variation as a function of individual ancestry (θ_S), individual heterozygosity (θ_H), and as a function of individual admixture (θ_S^2) decoupled from ancestry and heterozygosity (θ_S^2 equals 1 in both pure native and pure introduced genotypes, and it equals 0 in individuals with evenly mixed ancestry, i.e., half native and half introduced alleles regardless of heterozygosity). The coefficient α_2 represents effects of mixed ancestry not depending on heterozygosity but does not distinguish true epistasis from complementary effects. We compared the likelihoods of the full and reduced models by means of Akaike's Information Criterion (AIC) (64).

We thank A. Chang, M. Fujita, O. J. Abramyan, P. C. Trenham, and W. K. Savage for assistance in the field and laboratory; the H.B.S. laboratory for discussion; and many ranchers for access to their properties. Our work was funded by Environmental Protection Agency Grants U 91572401 and R 828896, the U.S. Department of Agriculture (Exotic/Invasive Pests and Diseases Research Program Grant 04XN022 and National Science Foundation Grants DEB 0516475 and DEB 0213155), CAL FED Grant 01-N43, and the University of California, Davis, CA, Agricultural Experiment Station.

- Arnold ML (2006) *Evolution Through Genetic Exchange* (Oxford Univ Press, Oxford).
- Dowling TE, Secor CL (1997) *Annu Rev Ecol Syst* 28:593–619.
- Mallet J (2005) *Trends Ecol Evol* 20:229–237.
- Barton NH (2001) *Mol Ecol* 10:551–568.
- Rivera MC, Lake JA (2004) *Nature* 431:152–155.
- Simonson AB, Servin JA, Skophammer RG, Herbold CW, Rivera MC, Lake JA (2005) *Proc Natl Acad Sci USA* 102:6608–6613.
- Arnold ML, Meyer A (2006) *Zoology* 109:261–276.
- Rhymer JM, Simberloff D (1996) *Annu Rev Ecol Syst* 27:83–109.
- Ellstrand NC, Schierenbeck KA (2000) *Proc Natl Acad Sci USA* 97:7043–7050.
- Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) *Trends Ecol Evol* 16:613–622.
- Allendorf FW, Luikart G (2007) *Conservation and the Genetics of Populations* (Blackwell, Malden, MA).
- Reisenbichler RR, Rubin SP (1999) *ICES J Mar Sci* 56:459–466.
- Mansfield KG, Land ED (2002) *J Wildl Dis* 38:693–698.
- Hedrick PW (1995) *Conservation Biol* 9:996–1007.
- Lewontin RC, Birch LC (1966) *Evolution (Lawrence, Kans)* 20:315–336.
- Grant BR, Grant PR (1996) *Ecology* 77:500–509.
- Anderson E (1948) *Evolution (Lawrence, Kans)* 2:1–9.
- Lynch M (1991) *Evolution (Lawrence, Kans)* 45:622–629.
- Rieseberg LH, Archer MA, Wayne RK (1999) *Heredity* 83:363–372.
- Lippman ZB, Zamir D (2007) *Trends Genet* 23:60–66.
- Riley SPD, Shaffer HB, Voss SR, Fitzpatrick BM (2003) *Ecol App* 13:1263–1275.
- Fitzpatrick BM, Shaffer HB (2004) *Evolution (Lawrence, Kans)* 58:1282–1293.
- Fitzpatrick BM, Shaffer HB (2007) *Ecol App* 17:598–608.
- Shaffer HB, McKnight ML (1996) *Evolution (Lawrence, Kans)* 50:417–433.
- Shaffer HB, Pauly GB, Oliver JC, Trenham PC (2004) *Mol Ecol* 13:3033–3049.
- US Fish and Wildlife Service (2004) *Federal Register* 69:47211–47248.
- Anderson JD, Hassinger DD, Dalrymple GH (1971) *Ecology* 52:1107–1112.
- Lexer C, Randell RA, Rieseberg LH (2003) *Ecology* 84:1688–1699.
- Rieseberg LH, Linder CR (1999) *Ecology* 80:361–370.
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S-Plus*, 4th Ed (Springer, New York).

