

Experimental evolution in *Heterandria formosa*, a livebearing fish: group selection on population size

CHARLES F. BAER^{1*}, JOSEPH TRAVIS¹ AND KEVIN HIGGINS²

¹Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

²Ecology and Evolution Program, University of Oregon, Eugene, OR 97403-1210, USA

(Received 24 August 1999 and in revised form 28 February and 28 April 2000)

Summary

Group selection has historically been an important and controversial subject in evolutionary biology. There is now a compelling body of evidence, both theoretical and experimental, that group selection not only can be effective, but can be effective in situations when individual selection is not. However, experiments in which true population-level traits have been shown to evolve in response to group selection are currently limited to two species of flour beetle in the genus *Tribolium* and RNA viruses. Here we report the results of an experiment wherein we imposed group selection via differential extinction for increased and decreased population size at 6-week intervals, a true population-level trait, in the poeciliid fish *Heterandria formosa*. In contrast to most other group selection experiments, we observed no evolutionary response after six rounds of group selection in either the up- or down-selected lines. Populational heritability for population size was low, if not actually negative. Our results suggest that group selection via differential extinction may be effective only if population sizes are very small and/or migration rates are low.

1. Introduction

Group selection, whereby fitness is a function of group membership as well as of individual attributes, has been an important and controversial topic in evolutionary biology since Darwin's time (1859; also e.g. Wright, 1931, 1945; Maynard Smith, 1964; Wade, 1978; Wilson, 1983; Goodnight & Stevens, 1997). Group selection can operate in one or both of two ways: differential probability of sending out migrants and/or founding new colonies, or differential probability of extinction. The former is the mechanism by which Wright (1931, 1977) envisioned the shifting balance process of evolution operating (his 'Phase III') and has received substantial theoretical attention from population geneticists in recent years (Crow *et al.*, 1990; Phillips, 1993; Moore & Tonsor, 1994; Coyne *et al.* 1997). The latter has received less

attention from theoretical population geneticists, but has figured prominently in macroevolutionary thinking (Stanley, 1975; Gould, 1980; Eldredge, 1995).

The aspect of group selection that has caused controversy is the possibility that genotypes that have low fitness within a group can confer some property on the group as a whole that increases the probability that the group will survive and/or send out colonists. The theoretical possibility of such a situation has been convincingly demonstrated mathematically (Wright, 1945; reviewed in Wilson, 1983). The situation is best illustrated by a classic example from evolutionary epidemiology: the fitness of a pathogen depends both on its local fitness within its host and on its global fitness, i.e. the probability that the host conveys the pathogen to another host. A mutant pathogen genotype that reproduces much faster than the 'wild-type' genotype will have higher local fitness and will increase in frequency within the host. However, if its increased reproduction causes the host to die before transmitting the pathogen to another host, the mutant genotype will become extinct. In this case group selection among populations of pathogens opposes

* Corresponding author. Ecology and Evolution Program, 1210 University of Oregon, Eugene, OR 97403-1210, USA. Tel: +1 (541) 346 4697. Fax: +1 (541) 346 2364. e-mail: cbaer@darkwing.uoregon.edu

individual selection within populations, and the phenotypic optimum will be different from that favoured solely by individual selection (e.g. Miralles *et al.*, 1997).

An important property of evolution via group selection is that, unlike evolution due to individual selection, non-additive genetic variation due to dominance and epistasis can contribute to the response to selection (Goodnight, 1987). Perhaps more importantly, genetically based behavioural and/or ecological interactions between individuals can also contribute to the response to group selection (Griffing, 1977; Goodnight, 1991; Goodnight & Stevens, 1997). Such interactions have been implicated in several experimental studies in which traits that did not respond to individual selection did respond to group selection (Goodnight, 1985; Craig & Muir, 1996; Muir 1996).

In spite of its theoretical importance, the evolution of population-level traits (census size, intrinsic rate of increase and/or carrying capacity) via group selection has to our knowledge been investigated only in two species of flour beetle in the genus *Tribolium*, in the classic work of Wade and his colleagues (reviewed in Goodnight & Stevens, 1997) and in RNA viruses (Miralles *et al.*, 1997). The evolution of population size in *Tribolium* has been shown to be mediated by larval egg cannibalism, a genetically based behavioural interaction (Wade, 1979; McCauley & Wade, 1980). Here we report results from an experiment in which we imposed a group selection for population size (high and low) on laboratory populations of the poeciliid fish *Heterandria formosa*; we have previously documented genetically based variation in population-level demographic traits in experimental populations of this species (Leips *et al.* in press).

This study constitutes part of a larger research programme designed to infer the underlying genetic connections, if any, between life-history traits, thermal stress tolerance and demography in this species. Briefly, we hypothesized that genetically based trade-offs between thermal stress tolerance and certain life-history traits (fecundity, time to maturation) could result in differences in population-level traits (intrinsic rate of increase and/or carrying capacity) that could in turn affect group fitness, as discussed above. Populations of *H. formosa* vary from one another in many traits, including thermal stress tolerance (Forster-Blouin, 1989; Baer & Travis, 2000), body size and life-history traits (Leips & Travis, 1999; H. Rodd, unpublished data), and population demographic parameters (Leips *et al.*, in press). Because the natural history of the species is such that periodic thermal stress is of likely ecological relevance, such connections could influence the evolution of the species' range, our ultimate question of interest. Because of the nature of the results of this experiment, we do not report on individual phenotypic traits (life-

history traits and thermal stress tolerances) in this paper; we mention them simply to establish the broader context in which this study was conceived.

2. Materials and methods

(i) Natural history of *Heterandria formosa*

Heterandria formosa (Pisces: Poeciliidae) is a small (12–30 mm), prolific livebearing fish, common throughout the coastal plain of the southeastern United States, ranging from the western Mississippi river drainage in central Louisiana, throughout the

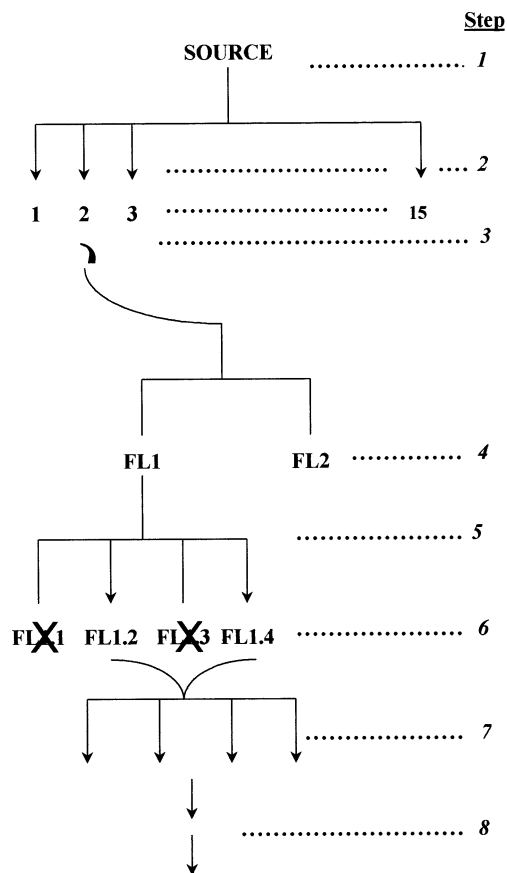


Fig. 1. Diagram of the selection protocol. *Step 1*: Collect ~ 1000 fish from each source population (FL in this example). *Step 2*: Divide collection into 15 (14 for SC) replicates. *Step 3*: Select the five replicates with the lowest census size to represent the Low treatment group (select the four replicates with the largest census sizes for the High treatment group). *Step 4*: Pool all fish collected in step 3 and divide them into two independently evolving lines (FL1 and FL2 in this example). *Step 5*: Divide each line into four replicate tanks (FL1.1, ... FL1.4 in this example). *Step 6*: After six weeks, census the four replicates. Keep the fish from the two replicates with the smallest (largest in High) census sizes (FL1.2 and FL1.4 in this example). *Step 7*: Divide the fish collected in step 6 into four replicate lines. *Step 8*: Go to step 6. Repeat six times.

Table 1. ANOVA table for census size at termination of selection

Effect	SS	d.f.	Mean square	F-ratio	P
Source population	0.007	1	0.007	0.303	0.587
Treatment	0.031	1	0.031	0.392	> 0.5
Source × Treatment	0.001	1	0.001	0.013	> 0.9
Line(Source)	0.087	2	0.044	1.866	0.176
Tr × Line(Source)	0.158	2	0.079	3.377	0.051
Error	0.560	24	0.023		

Dependent variables are residuals of the regression of log(census size) on day of census. See Section 2 for details. Treatment and Source × Treatment effects are tested over the Tr × Line(Source) mean square.

Florida peninsula, and north to the Cape Fear River, NC (Martin, 1980). It occurs almost exclusively in slow-moving, heavily vegetated waters and is often found in the very shallowest water present.

On the Florida peninsula, reproduction typically begins in March and continues through October, resulting in two to four generations per year. Females carry several clutches of embryos at different stages of development and give birth to small broods (1–8 offspring) at intervals of a few days to 2 weeks. Gestation time is approximately 5 weeks; females mature at approximately 5 weeks and males at 6–7 weeks (Leips, 1997). Demographic generation time in experimental populations is approximately 5 weeks. (J. Leips, personal communication).

The family Poeciliidae is neotropical; only a few representative species of three genera occur north of the Rio Grandé; one species (*Gambusia holbrooki*) is found as far north as New Jersey. *Heterandria formosa* is the only North American representative of its genus; all congeners occur from the Yucatan peninsula southward into Central America (Parenti & Rauchenberger, 1989).

(ii) Selection protocol

The selection protocol is outlined schematically in Fig. 1. In November and December 1994 we collected approximately 1000 adult individuals from two source populations: the Peace River, DeSoto County, FL (henceforth FL) and from the Savannah River, Jasper County, SC (henceforth SC). These populations were chosen to represent both the heart of the species' range on the Florida peninsula and the northern margin of its range. Populations north of the St Johns River drainage in northeastern Florida are genetically distinguishable from populations to the south and west, and are likely to be more recent in origin (Baer, 1998; C.F.B., unpublished data).

In March 1995 15 replicate tanks from FL and 14 from SC were established in 120 litre aquaria in a climate-buffered greenhouse. Each replicate tank was begun with 15 males and 15 females; females store

sperm over the winter, so the effective population size was probably larger than the census size. These tanks were censused in early May 1995. The four tanks from each population with the largest number of individuals were selected to be the parental stock for the High lines; the five tanks with the fewest fish were selected as parental stock for the Low lines. Two replicate lines from each source population of each treatment were established, for a total of eight independently evolving lines of four tanks per line (designated by the one-letter initials for Source/Treatment/Line, e.g. FH1, ..., SL2). There are thus two levels of replication: tanks within line and lines within source population. Tanks were censused at 6 week intervals throughout the 1995 and 1996 breeding seasons for a total of six rounds of selection (= 'generation'); fish from the two tanks with the most individuals (adults and juveniles) for the high lines and the fewest individuals for the low lines were randomly distributed into the four tanks of a line to propagate the next generation. This selection protocol constitutes 'migrant pool' rather than 'propagule pool' (both terms *sensu* Slatkin, 1977; Wade, 1978) selection. Although the response to propagule pool group selection is expected to be stronger than to migrant pool selection (Wade, 1978), inbreeding is greater with a propagule

Table 2. Mean census size at termination of selection

Line	N	Residual
FH1	143.75 (6.933)	-0.078 (0.047)
FH2	173.25 (10.934)	0.199 (0.065)
FL1	130.75 (9.595)	-0.022 (0.071)
FL2	159.75 (11.982)	-0.009 (0.075)
SH1	179.50 (12.606)	0.078 (0.070)
SH2	142.00 (15.149)	-0.044 (0.114)
SL1	122.75 (10.371)	-0.120 (0.087)
SL2	159.75 (9.707)	0.056 (0.064)

'N' is the actual census size, 'Residual' is the residual of the regression of log(census size) on day of census. SEM is given in parentheses.

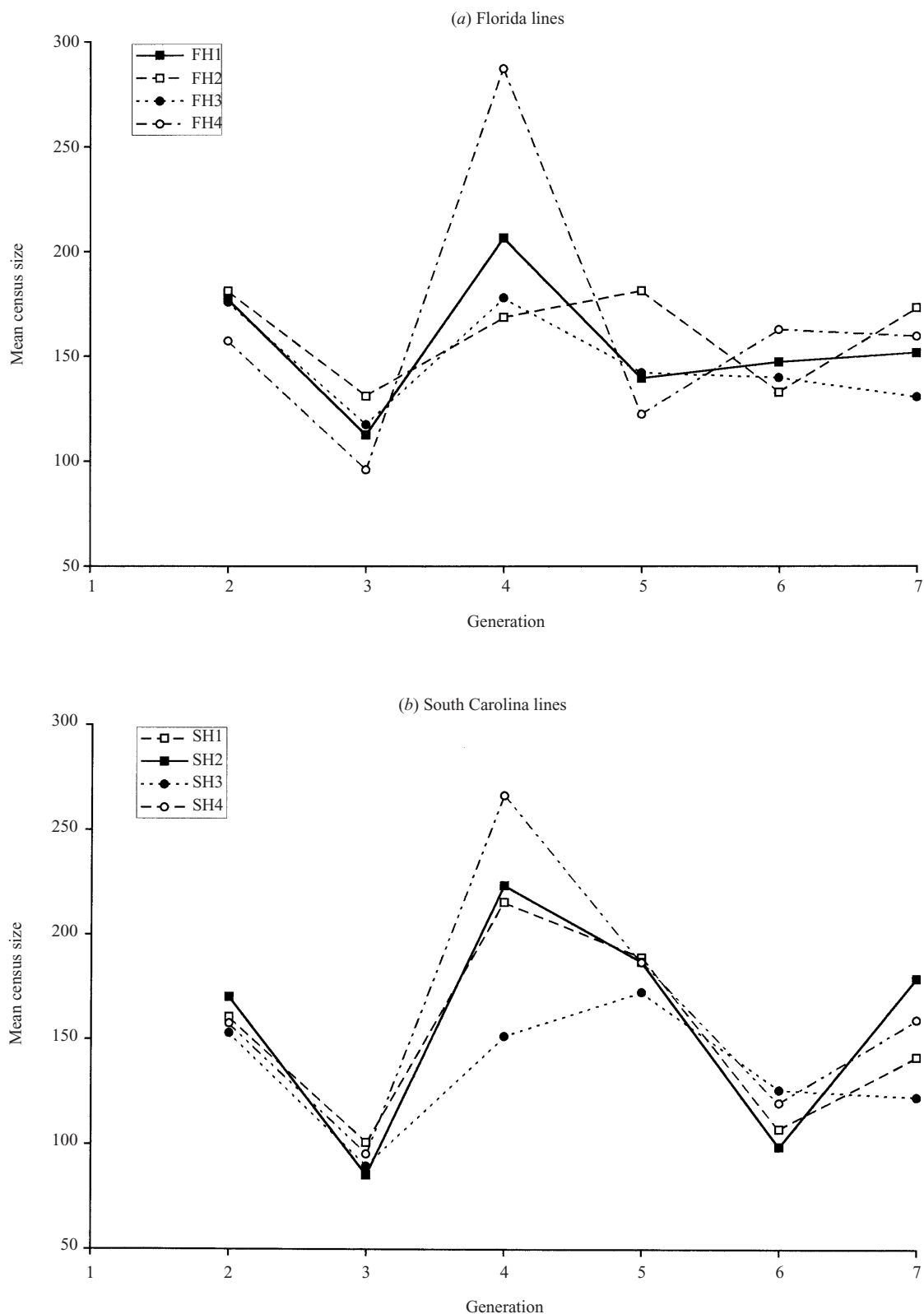


Fig. 2. For legend see facing page.

pool. Because the conceptual motivation for this experiment concerns adaptive evolution and because a response due to inbreeding depression is arguably not an adaptive response (but see Lynch *et al.*, 1999), we

wanted to minimize inbreeding, accordingly, we chose to use migrant pool selection.

Census sizes were established at the beginning of each generation by the number of individuals of a

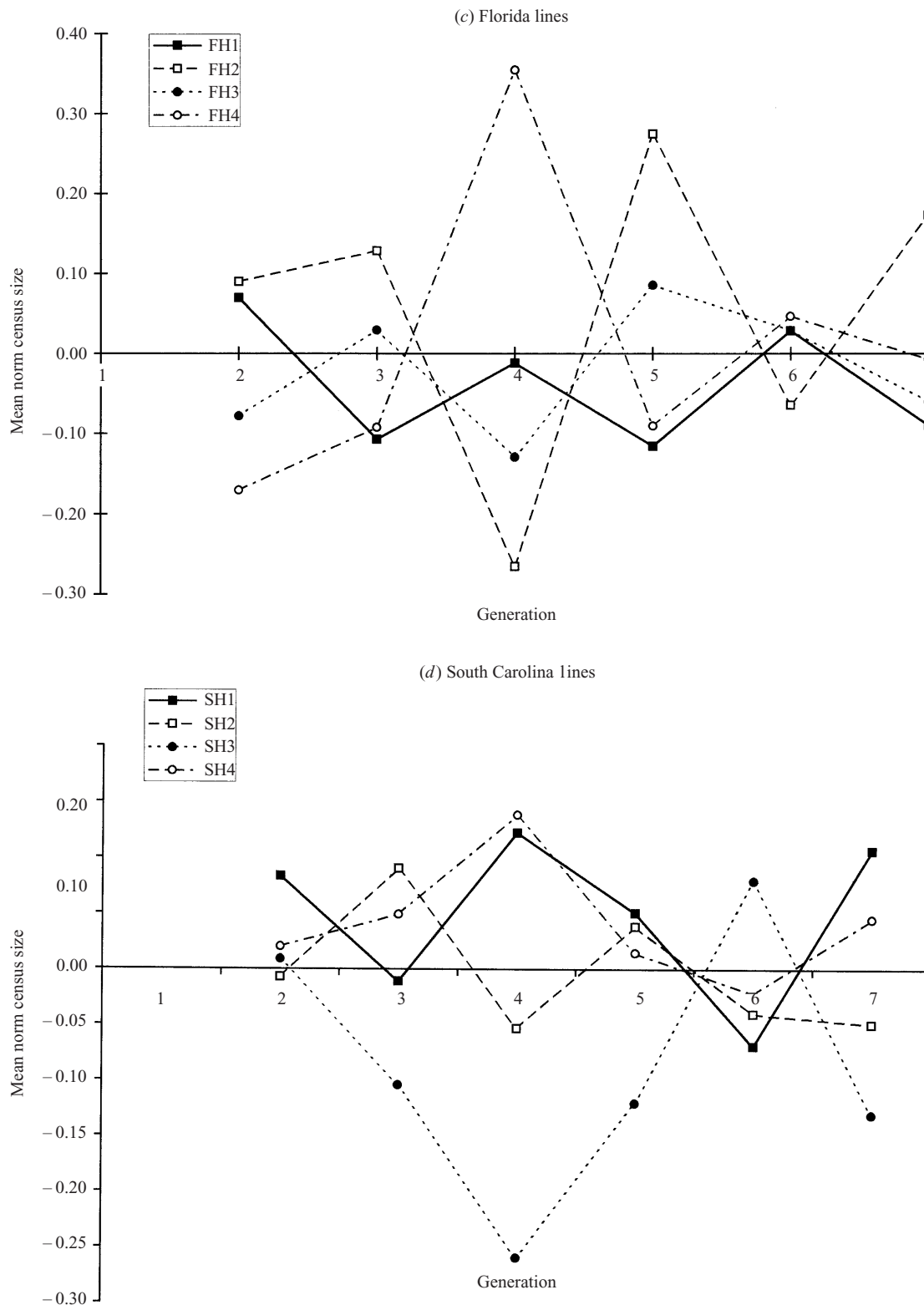


Fig. 2. Line mean census sizes at each generation. Squares represent High lines, circles represent Low lines. (a) FL lines, raw data. (b) SC lines, raw data. (c) FL lines, data normalized by source population generation mean (see Section 2 for details of normalization procedure). (d) SC lines, normalized data.

class ('class' = male, female, juvenile) present in the line with the fewest individuals of that class. For example, if the two replicates used to seed the next generation contained 8 and 10 males, respectively,

each replicate got 4 males in the next generation; each tank was seeded with the same number of males, females and juveniles. Harmonic mean beginning census sizes over the six generations were approx-

imately 9 males, 12 females and 23 juveniles. Because individuals could potentially contribute genetically to more than one round of selection, generations were overlapping. Tank positions were systematically rotated at each round of selection to minimize the effects of environmental heterogeneity.

(iii) Data analysis

The most straightforward analysis of data from a selection experiment is to examine the line means in the ultimate generation of selection; a statistically significant treatment effect (either as a main effect or an interaction) is indicative of a response to selection. The experiment was designed as a nested, split-plot mixed model ANOVA, with treatment (High/Low) and source population (FL/SC) as fixed effects and line nested within source population (plot) as a random effect.

There was a complicating effect of day of census on census size. It took an average of 8 days to census all 32 tanks, and the four replicate tanks of a line had to be censused consecutively in order to accurately assess which tanks had the largest/smallest census sizes; effects of day of census were therefore not independent of line effects within a generation. When averaged over all generations, the overall effect was a slight, but highly significant, positive relationship between day of census and census size ($y = 0.031x + 4.857$, $F_{1,187} = 10.06$, $P < 0.002$). We therefore analysed residuals of the regression of census size on day of census rather than raw data.

Our experimental design can be considered as analogous to truncation selection in an individual selection experiment with a selection intensity of 0.5; in our case we imposed a 50% extinction rate. However, the selection in our design is not centred around the mean census size of the four tanks, but rather the median. It is possible to calculate a weighted selection differential, S , defined here as $[(z_i/z_s)z_i + (z_j/z_s)z_j] - \bar{z}$, where i and j represent the two surviving replicates, z_s is the mean census size of the two surviving replicates and \bar{z} is the mean census size of the four replicates in a line. The response to selection, R , is then determined by the equation $R = h_g^2 S$, where R is the change in mean census size across generations and h_g^2 is a measure of group heritability (*sensu* Wade & McCauley, 1980; Slatkin, 1981). The realized group heritability can be determined by the regression of generation means, R , on the cumulative selection differential, S_{cum} (Falconer, 1989, chapter 11). Note that weighting the selection differential by the contribution of each replicate to the next generation has the effect of increasing the selection differential in the High treatment and reducing the selection differential in the Low treatment relative to the unweighted selection differentials. Individual

selection thus reinforces group selection in up-selected lines and opposes group selection in down-selected lines.

In calculating the realized group heritability, there are two complicating factors that need to be considered. Mean census sizes averaged over all replicates varied from generation to generation and, in some cases, between source population (see Fig. 2). To standardize data among generations and source populations, we transformed raw data to deviations from the source population mean, expressed as a fraction of the source population mean in that generation. The normalized deviation is $x_i = (z_i - \bar{z}_{j,k}) / \bar{z}_{j,k}$, where z_i is the census size of replicate i and $\bar{z}_{j,k}$ is the mean census size of all replicates derived from source population j in generation k . This calculation is the same in principle as calculating realized heritability from the deviation between up- and down-selected lines (Falconer 1989, pp. 197–8), except it allows explicit consideration of the variation between lines within a treatment/source group. If the response to selection is asymmetrical (and there is reason to expect it would be; see Wade & McCauley, 1980), the heritability in one direction will be overestimated and the heritability in the other direction will be underestimated. To account for this possibility, we also calculated realized heritabilities from the non-normalized data; comparison of realized heritabilities calculated in the two different ways provides a heuristic way to identify asymmetrical responses to selection.

As we discuss below, the census size of the experimental populations exhibited marked oscillations that weaken the ANOVA approach. We developed a Monte Carlo simulation as an alternative method of analysis. The simulation explored the range of responses to selection we could have expected, given our sample sizes and a range of population heritabilities. This procedure provides an alternative way of assigning confidence limits to the realized heritability and also serves as a heuristic power analysis if we interpret the results as addressing the question: ‘Given a particular average heritability and our sample sizes, how often would we observe a response less than or equal to the observed response to selection?’

The simulation faithfully reproduced the experimental design with the exception of ignoring the effect of source population. Because there was no significant effect of source population on census size at either the beginning (two-tailed unpaired t -test, $t = 0.452$, d.f. = 27, $P > 0.45$) or the end of the experiment (Table 1), this assumption is statistically justified, and it simplifies the subsequent analysis considerably. In the initial generation, we randomly drew 14 populations (as in the SC lines) from a normal distribution with mean and variance equal to those calculated

from the data (bootstrapping the mean and variance changed the initial values only very slightly), and repeated this eight times, so that there were eight independently evolving lines, four selected for high census size and four for low census size. We then allowed the system to evolve for six generations, calculating selection differentials exactly as above and holding the coefficient of variation of the distribution constant. At the seventh generation, we calculated line means, and interpreted the difference of the mean of the High and Low lines divided by the mean of all eight lines as the response to selection. The simulation was repeated 1000 times for heritabilities from 0 to 1 at intervals of 0.1. We interpret the upper 95% confidence limit of the heritability as that heritability for which we observed a response less than or equal to the actual response in 5% of simulations or fewer. Results of simulations and annotated C++ source code are available from K.H. upon request.

3. Results

Line mean census sizes at each generation for each source population are presented in Fig. 2. ANOVA of census sizes at the ultimate round of selection reveals no main effect of treatment and no significant treatment \times source interaction (Table 1). There is a marginally significant ($P = 0.051$) effect of $\text{Tr} \times \text{Line}(\text{Source})$, but examination of line means (Table 2) indicates that that result cannot easily be interpreted as a partial response to selection whereby some lines responded in the expected way to selection and some lines did not respond to selection. More importantly, the large fluctuations in rank of line means across generations (Fig. 2) is not consistent with a partial response to selection.

The conclusion of no response to selection is confirmed by examination of the realized group heritabilities (Table 3). When census sizes are normalized for among-generation variation, the mean realized group heritability of the eight lines is estimated

as -0.023 , with 95% confidence limits of $(-0.113, 0.068)$ (determined empirically; Hill 1971). There was no difference in heritability between the two source populations; the mean of the four FL lines is -0.015 , the mean of the four SC lines is -0.030 . When realized group heritabilities are calculated from the non-normalized data rather than from deviations from generation means, the overall results are almost identical, although the realized heritabilities of some lines differ qualitatively between the two analyses and there is substantially more variance both within and among lines (Table 3). Also, note that the mean non-normalized heritabilities for High ($h^2 = -0.014$) and Low ($h^2 = -0.109$) treatments are similar and negative, not suggestive of a greatly asymmetrical response (or any response) to selection. We consider negative group heritabilities as potentially meaningful (also see Goodnight, 1985), because, unlike individual heritabilities, the genetic contribution to group heritability need not be non-negative (see Section 4). Expressed as percentages of the line mean, per-generation weighted selection differentials for normalized data averaged 12.1% in Low lines and 16.2% in High lines; weighted selection differentials for non-normalized data averaged 14.1% in Low lines and 15.3% in High lines (Table 3).

Results from the Monte Carlo simulations allow us to formally reject the hypothesis ($\alpha = 0.05$) that the realized group heritability is ≥ 0.4 . With a heritability of 0.4, 48/1000 simulations resulted in a response to selection less than 0.108, the value observed in the experiment. With a heritability of 0.3, the observed value of 0.108 or less occurred in 130/1000 simulations; with true realized heritability of 0.5, responses less than observed occurred in only 23 of 1000 runs.

4. Discussion

To our disappointment, the results from this experiment constitute the first documented general

Table 3. Cumulative weighted selection differentials and realized group heritabilities

Line	S_{cum}	h^2	$S_{\text{cum}}(\text{Norm})$	$h^2(\text{Norm})$	$\bar{S}(\text{Norm})$
FH1	1.170	-0.161	1.176	-0.105	0.196
FH2	0.951	0.076	1.032	0.178	0.172
FL1	-0.884	-0.274	-0.709	-0.090	-0.118
FL2	-0.962	-0.096	-0.835	-0.102	-0.139
SH1	0.508	-0.533	0.619	-0.089	0.103
SH2	1.033	0.561	1.058	-0.062	0.176
SL1	-0.867	-0.110	-0.704	-0.044	-0.117
SL2	-0.666	0.045	-0.668	0.076	-0.111

'Norm' represents values normalized for the source population generation mean. $\bar{S}(\text{Norm})$ is the mean normalized selection differential per generation, expressed as a fraction of the line mean. (See Section 2 for details.)

failure to detect a response to group selection on population size (Goodnight & Stevens, 1997). Moreover, our results strongly suggest that the realized group heritability for population size must be low, if not actually negative. However, examination of Wade & McCauley's (1980) data (their table 6) also reveals an important consistency between our results and theirs. In their experiment, the realized heritability (their 'relative response') of *increased* population size in lines with beginning census sizes of 24 and 48 adults was zero; they only observed positive responses in up-selected lines with beginning census sizes of 6 or 12. The harmonic mean beginning census size in our experiment was 44 individuals (adults and juveniles), and since we did not change the demographic structure from the end of one generation to the beginning of the next, evolution (or lack thereof) would have occurred in populations with overlapping generations, an average census size of 44 individuals and free migration between the two surviving populations.

We can think of three possible reasons why populations of *Tribolium* consistently respond to group selection for *decreased* population size whereas our similarly sized populations of *Heterandria formosa* apparently did not. First, populations will diverge due to genetic drift more slowly under migrant pool selection than they will under propagule pool selection, resulting in a slower response to selection. Second, with migrant pool selection, inbreeding depression will be lower for a given population size and number of generations than it will under propagule pool selection; it is known that inbreeding depression is responsible for a general decrease in mean population size in the *Tribolium* experiments (Wade & McCauley, 1980). Third, and we suspect more likely, it may be that our populations of *H. formosa* do not exhibit the same sort of genetically based behavioural interactions that are largely responsible for the evolution of population size via group selection in *Tribolium*.

Because the *Tribolium* experiments were substantially larger than ours, an obvious question concerns statistical power, i.e. could we have detected a response to selection given a particular heritability? We will use Wade & McCauley's (1980) results as a guideline. They determined that for populations of *Tribolium* with census sizes on the order of our own (24 and 48), the across-generation correlation coefficient (their surrogate for heritability) between unselected populations of *Tribolium* was approximately 0.7 after six generations (their figure 3). In our study, the average across-generation correlation coefficient was approximately -0.1. After one generation of selection, they observed a realized h^2 of about 0.7 for populations of 24 individuals selected for decreased population size and an h^2 of 1.0 for down-selected populations of 48 individuals. Their lines selected for increased population size had realized heritabilities of 0 or less

(although this result might have been anomalous; see their table 1). Taking the results of our simulations as an approximation for statistical power, we had good power to detect a response to selection if the true realized heritability had been about 0.4 or greater ($1 - \alpha > 0.95$). We can make another estimate of the realized heritability in our experiments through these simulation results by using the value of heritability that maximizes the probability density in the neighborhood of the observed values (a pseudo-maximum likelihood estimate). When we do this, a realized heritability of about 0.1 provides the best fit to the data.

Examination of the data in Fig. 2 indicates a general tendency for census size to oscillate between generations, i.e. lines that increased in census size in one generation tended to decrease in the next generation and vice versa. If the fluctuations in census size from one generation to the next were random (disregarding that census size will ultimately fluctuate around some carrying capacity), we would expect the census size to show changes in the same direction in successive transitions (two increases or two decreases) as often as changes in opposing directions (increase-decrease or vice versa). When all lines are considered together, normalized census size shows a significant oscillation (observed 'same' = 24, observed 'opposite' = 8, expected = 16/16; $\chi^2 = 8.0$, d.f. = 1, $P < 0.005$). When the data are broken down by source populations, the FL lines exhibit a highly significant oscillation (no. the same = 14, no. opposite = 2; $\chi^2 = 9.0$, d.f. = 1, $P < 0.005$), but the SC lines do not (no. the same = 10, no. opposite = 6; $\chi^2 = 1.0$, d.f. = 1, $P > 0.4$). Visual inspection of the data does not convincingly suggest a two-generation periodicity to the oscillation.

The oscillation explains the apparent absence of group heritability. We can think of two plausible scenarios that would produce such an oscillation. First, given the significant positive relationship between census size and day of census, it seems likely that the 6-week interval between census was not long enough to allow the replicate tanks to reach their equilibrium population size. Since census size in any real population at any given point in time will generally be an unknown function of both r and K , non-equilibrium population dynamics is not a theoretical problem in our experiment. However, if we happened to census a replicate just after a burst of parturition, census size would be large (and skewed toward juveniles), but at the next census the population would consist of a relatively large number of females that had not yet given birth and relatively few juveniles, resulting in a smaller overall census size. Thus, the oscillation in census size could be an incidental byproduct of the timing of our selection regime.

The second possible explanation for the apparent

oscillation is a consistent maternal effect resulting in a negative phenotypic correlation between mothers and daughters, such that daughters of a mother who produced many offspring (resulting in a population with a large census size, on average) tend themselves to have fewer offspring (resulting in a population with a small census size) and vice versa. Female *H. formosa* raised at high densities are known to produce fewer, larger offspring than females raised at low densities (H. Rodd, unpublished data). Such a negative correlation is theoretically expected to produce an oscillation in the approach to the asymptotic response to individual selection, and in the initial generation the response to selection will be in the opposite direction of selection (Kirkpatrick & Lande, 1989). In either case, the biological phenomena responsible for the oscillation in census size could have an underlying genetic basis, thereby resulting in a genetic component of negative group heritability. It is perhaps not coincidental that Wade & Goodnight (1991) observed the strongest response to group selection on population size in *Tribolium castaneum* when they imposed selection every second generation; responses were weaker when selection was imposed every generation or every third generation.

To conclude, the lesson that can be drawn from this experiment is that with population sizes on the order of a few dozen individuals and migrant pool selection, six (overlapping) generations of group selection with a 50% extinction rate produced at best a very weak evolutionary response in both up- and down-selected lines. Depending on the method of inference, realized group heritability for population size as measured here is likely to be somewhere between weakly positive and weakly negative. This result suggests that extinction, either random or due to selection on population-level traits, may be unlikely to produce an evolutionary response in the demographic properties of populations unless the populations involved are quite small for a number of generations and/or migration among surviving populations is low.

We thank G. Coffey, J. Poulton and H. Rodd for help collecting fish, and J. Grubich, C. Hays, T. Hoag, L. Horth, B. Shoplock, T. Spears and T. Waltzek for babysitting them. C. Goodnight, J. Herbers, P. Phillips, D. Simberloff, A. Winn and an insightful anonymous reviewer provided helpful discussions and/or comments on the manuscript. Support was provided by a Florida State University Dissertation Fellowship to C. F. B. and NSF grant DEB 92-20849 to J. T.

References

- Baer, C. F. (1998). Species-wide population structure in a southeastern US freshwater fish, *Heterandria formosa*: gene flow and biogeography. *Evolution* **52**, 183–193.
- Baer, C. F. & Travis, J. (2000). Direct and correlated responses to artificial selection on acute thermal stress tolerance in a livebearing fish. *Evolution* **54**, 238–244.
- Coyne, J. A., Barton, N. H. & Turelli, M. (1997). Perspective: A critique of Sewell Wright's shifting balance theory of evolution. *Evolution* **51**, 643–671.
- Craig, J. V. & Muir, W. M. (1996). Group selection for adaptation to multiple-hen cages: beak-related mortality, feathering, and body weight responses. *Poultry Science* **75**, 294–302.
- Crow, J. F., Engels, W. R. & Denniston, C. (1990). Phase three of Wright's shifting-balance theory. *Evolution* **44**, 233–243.
- Darwin, C. (1859). *On the Origin of Species*. London: J. Murray.
- Eldredge, N. (1995). *Reinventing Darwin*. New York: Wiley.
- Falconer, D. S. (1989). *Introduction to Quantitative Genetics*, 3rd edn. New York: Wiley.
- Forster-Blouin, S. (1989). Genetic and environmental components of thermal tolerance in the Least Killifish, *Heterandria formosa*. PhD dissertation, Florida State University, Tallahassee, FL.
- Goodnight, C. J. (1985). The influence of environmental variation on group and individual selection in a cress. *Evolution* **39**, 545–588.
- Goodnight, C. J. (1987). On the effect of founder events on the epistatic genetic variance. *Evolution* **41**, 80–91.
- Goodnight, C. J. (1990). Experimental studies of group selection. I. The response to selection at the community level. *Evolution* **44**, 1614–1624.
- Goodnight, J. (1991). Intermixing ability in two species communities of *Tribolium* flour beetles. *American Naturalist* **138**, 342–354.
- Goodnight, C. J. & Stevens, L. (1997). Experimental studies of group selection: what do they tell us about group selection in nature? *American Naturalist* **150**, S59–S79.
- Gould, S. J. (1980). Is a new and general theory of evolution emerging? *Paleobiology* **3**, 115–151.
- Griffing, B. (1977). Selection for populations of interacting genotypes. *Proceedings of the International Congress on Quantitative Genetics* (ed. E. Pollak, O. Kempthorne & T. B. Bailey), pp. 413–434. Ames, IA: Iowa State University Press.
- Hill, W. G. (1971). Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* **27**, 293–311.
- Kirkpatrick, M. & Lande, R. (1989). The evolution of maternal characters. *Evolution* **43**, 485–503.
- Leips, J. (1997). Genetic differentiation in life history traits: influence on population dynamics in the least killifish, *Heterandria formosa*. PhD dissertation, Florida State University, Tallahassee, FL.
- Leips, J. & Travis, J. (1999). The comparative expression of life-history traits and its relationship to the numerical dynamics of four populations of the Least Killifish, *Heterandria formosa*. *Journal of Animal Ecology* **68**, 595–616.
- Leips, J., Travis, J. and Rodd, F. H. (2000). Genetic differentiation in life-histories: effects on population dynamics in the Least Killifish. *Ecological Monographs*, in press.
- Lynch, M., Pfrender, M., Spitze, K., Lehman, N., Hicks, J., Allen, D., Latta, L., Ottene, M., Bogue, F. & Colbourne, J. (1999). The quantitative and molecular genetic architecture of a subdivided species. *Evolution* **53**, 100–110.
- Martin, F. D. (1980). *Heterandria formosa* Agassiz, Least Killifish. In *Atlas of North American Fishes* (ed. D. S. Lee, C. R. Gilbert, C. H. Hocutt, R. E. Jenkins, D. E. McAllister & J. R. Stauffer, Jr), p. 547. Raleigh, NC: North Carolina State Museum of Natural History.

- Maynard Smith, J. (1964). Group selection and kin selection. *Nature* **201**, 1145–1147.
- McCauley, D. M. & Wade, M. J. (1980). Group selection: the genetic and demographic basis of phenotypic differentiation of small populations of *Tribolium castaneum*. *Evolution* **34**, 813–821.
- Miralles, R., Moya, A. & Elena, S. F. (1997). Is group selection a factor modulating the virulence of RNA viruses? *Genetical Research* **69**, 165–172.
- Moore, F. B.-G. & Tonsor, S. J. (1994). A simulation of Wright's shifting-balance process: migration and the three phases. *Evolution* **48**, 69–80.
- Muir, W. M. (1996). Group selection for adaptation to multiple hen cages: selection program and direct responses. *Poultry Science* **75**, 447–458.
- Parenti, L. R. & Rauchenberger, M. (1989). Systematic overview of the Poeciliines. In: *Ecology and Evolution of Livebearing Fishes* (ed. G. K. Meffe & F. F. Snelson). Englewood Cliffs, NJ: Prentice Hall.
- Phillips, P. C. (1993). Peak shifts and polymorphism during phase three of Wright's shifting balance process. *Evolution* **47**, 1733–1743.
- Slatkin, M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology* **12**, 253–262.
- Slatkin, M. (1981). Populational heritability. *Evolution* **35**, 859–871.
- Stanley, S. M. (1975). A theory of evolution above the species level. *Proceedings of the National Academy of Sciences of the USA* **72**, 757–760.
- Wade, M. J. (1977). An experimental study of group selection. *Evolution* **31**, 134–153.
- Wade, M. J. (1978). A critical review of the models of group selection. *Quarterly Review of Biology* **53**, 101–114.
- Wade, M. J. (1979). The primary characters of *Tribolium* populations group selected for increased and decreased population size. *Evolution* **33**, 749–764.
- Wade, M. J. & Goodnight, C. J. (1991). Wright's shifting balance theory: an experimental study. *Science* **253**, 1015–1018.
- Wade, M. J. & McCauley, D. E. (1980). Group selection: the phenotypic and genotypic differentiation of small populations. *Evolution* **34**, 799–812.
- Wilson, D. S. (1983). The group selection controversy: history and current status. *Annual Review of Ecology and Systematics* **14**, 159–187.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.
- Wright, S. (1945). Tempo and mode in evolution: a critical review. *Ecology* **26**, 415–419.
- Wright, S. (1977). *Evolution and Genetics of Populations*, vol. 3, *Experimental Results and Evolutionary Deductions*. Chicago: University of Chicago Press.