Evolutionary Ecology Research, 2005, **7**: 643–650

A field test of the extent of bias in selection estimates after accounting for emigration

Benjamin H. Letcher,* Gregg E. Horton, Todd L. Dubreuil and Matthew J. O'Donnell

S.O. Conte Anadromous Fish Research Center, USGS/Leetown Science Center, One Migratory Way, Turners Falls, MA 01376, USA

ABSTRACT

Question: To what extent does trait-dependent emigration bias selection estimates in a natural system?

Organisms: Two freshwater cohorts of Atlantic salmon (*Salmo salar*) juveniles.

Field site: A 1 km stretch of a small stream (West Brook) in western Massachusetts, USA from which emigration could be detected continuously.

Methods: Estimated viability selection differentials for body size either including or ignoring emigration (include = emigrants survived interval, ignore = emigrants did not survive interval) for 12 intervals.

Results: Seasonally variable size-related emigration from our study site generated variable levels of bias in selection estimates for body size. The magnitude of this bias was closely related with the extent of size-dependent emigration during each interval. Including or ignoring the effects of emigration changed the significance of selection estimates in 5 of the 12 intervals, and changed the estimated direction of selection in 4 of the 12 intervals. These results indicate the extent to which inferences about selection in a natural system can be biased by failing to account for trait-dependent emigration.

Keywords: Atlantic salmon, emigration, selection, selection differentials, size dependence.

INTRODUCTION

Studies estimating the strength and form of phenotypic selection on quantitative traits have contributed important insights into evolutionary processes (Endler, 1986; Fairbairn and Reeve, 2001; Kingsolver *et al.*, 2001; Kingsolver and Pfenning, 2004). In these studies, selection estimates are obtained by calculating coefficients for regressions relating components of fitness (e.g. mating success, fecundity, survival) to quantitative traits (e.g. morphology, life history/phenology, behaviour) (Lande and Arnold, 1983; Brodie *et al.*, 1995; Kingsolver *et al.*, 2001). A number of caveats have been attached to these methods (Schluter, 1988; Kingsolver and Smith, 1995; Hersch and Phillips, 2004; Hereford *et al.*, 2004), and here we provide an empirical assessment of one caveat.

* Author to whom all correspondence should be addressed. e-mail: bletcher@forwild.umass.edu

Selection analyses do not require a geographically closed study population but they do assume that trait values are randomly distributed across individuals that enter or leave the study area (Lande and Arnold, 1983). Immigration or emigration that is unrelated to phenotypic traits will not influence selection coefficients for those traits, but a serious bias could result from trait-related emigration (Kingsolver and Smith, 1995). For example, unmeasured size-dependent emigration could be mistaken for size-dependent mortality, and thus alter the intensity and shape of a selection function. To our knowledge, the extent of this potential bias in selection coefficients has not been estimated in a natural system.

In an earlier assessment of selection on body size in stream-dwelling Atlantic salmon (*Salmo salar*) (Hendry *et al.*, 2003), we minimized the potential effects of emigration by limiting analyses to time intervals during which emigration should be low. Since then, we have installed remote sensing devices that can detect the movements of tagged individuals with previously measured trait values out of our study area. These new data allow a direct assessment of the potential bias in selection estimates owing to emigration. Here we report how the failure to account for emigration influences the strength and direction of viability selection estimates acting on fish body mass. Although it is no surprise that trait-dependent emigration can affect selection coefficients, the extent of the effect of ignoring emigration has not been reported from the wild. Data presented are for two cohorts of Atlantic salmon juveniles in a small stream, and span six sampling intervals for each cohort.

METHODS

We report data for Atlantic salmon juveniles that were tagged with passive integrated transponder (PIT) tags in a small stream in western Massachusetts, USA (West Brook) (for details, see Letcher *et al.*, 2002; Letcher and Gries, 2003). Our study area was a 1 km long stream reach that was bounded at the downstream end by a pair of PIT tag antennas that continuously recorded the passage of tagged fish. We have estimated the efficiency of the pair of antennas to be 91%. The present study accounts for downstream emigration from the study site. We ignore upstream emigration from the study site because routine sampling showed this was very rare. Individual fish whose last recorded location was a downstream movement on the antennas were considered to be permanent emigrants.

For each cohort (2000 and 2001), individuals were tagged with 12 mm PIT tags (Digital Angel Corp.) during their first year of life (age-0+), once they had reached 60 mm fork length. A total of 797 (2000) and 747 (2001) fish were tagged. For each cohort, we conducted six sampling occasions (Table 1) that spanned the time of residence for most salmon in the stream. Sampling involved standard electrofishing techniques with unpulsed DC current at 400 V. We conducted two electrofishing passes within each of the 47 contiguous 20 m long sections of the study area. Each section was first bounded by blocking nets at the upstream and downstream ends, and was then sampled in an upstream direction. It took approximately 8 days to sample all of the sections during each sampling occasion. During sampling, any untagged fish were tagged and any tagged fish had their unique PIT tag code recorded. We also measured the wet mass ($\pm$ 0.1 g) of each individual each time it was captured. All fish were then returned to their location of capture in the stream.

To estimate the extent of size-dependent emigration between adjacent pairs of sampling occasions (each pair delineates an 'interval'), we conducted simple logistic regressions with emigration during the interval (yes = 1 or no = 0) as the dependent variable and body mass at the beginning of the interval as the independent variable.

Table 1. Schedule of sampling intervals for the two cohorts, numbers of fish captured at the beginning of each interval and the number of fish that emigrated during the interval

Cohort	Sampling interval	Median date	*N* at beginning	*N* emigrated	Size-dependent emigration		
					Intercept	Slope	*P*-value
2000	0+ fall to 0+ prewinter	27 Sept 2000	527	56	−1.97 ± 0.57	−0.03 ± 0.12	0.77
2000	0+ prewinter to 1+ spring	8 Dec 2000	306	36	−1.38 ± 0.75	−0.13 ± 0.15	0.39
2000	1+ spring to 1+ summer	13 Mar 2001	261	38	−1.50 ± 0.74	−0.05 ± 0.14	0.71
2000	1+ summer to 1+ fall	9 June 2001	293	52	−2.64 ± 0.74	0.06 ± 0.04	0.12
2000	1+ fall to 1+ prewinter	9 Sept 2001	298	57	−3.07 ± 0.67	0.09 ± 0.04	0.01
2000	1+ prewinter to 2+ summer	6 Dec 2001	157	27	1.25 ± 0.21	0.83 ± 0.25	0.00
2001	0+ fall to 0+ prewinter	9 Sept 2001	432	47	−2.28 ± 0.68	0.05 ± 0.21	0.79
2001	0+ prewinter to 1+ spring	6 Dec 2001	368	52	−2.84 ± 0.57	0.29 ± 0.15	0.05
2001	1+ spring to 1+ summer	31 Mar 2002	360	46	−2.41 ± 0.56	0.12 ± 0.13	0.36
2001	1+ summer to 1+ fall	12 June 2002	323	48	−2.05 ± 0.62	0.02 ± 0.05	0.61
2001	1+ fall to 2+ spring	25 Sept 2002	281	44	−2.87 ± 0.62	0.09 ± 0.04	0.04
2001	2+ spring to 2+ summer	10 Apr 2003	111	26	0.10 ± 0.22	1.08 ± 0.28	0.00

Note: Parameters for estimates of the intensity of size-dependent emigration are from logistic regressions for emigration for each interval.

To estimate selection on body mass during each interval, we calculated linear selection differentials in two ways – either ignoring or accounting for emigration. When ignoring emigration, fish that actually emigrated during an interval were treated as though they had died during that interval (absolute fitness = 0). When accounting for emigration, fish that actually emigrated during an interval were treated as though they had lived through that interval (absolute fitness = 1). For each combination of cohort and sampling period, we first standardized body mass to a mean of zero and standard deviation of unity. We then regressed body mass at the beginning of each interval on absolute fitness during that interval. Because fitness was dichotomous, we used logistic rather than linear regressions, and then converted the logistic coefficients to their linear equivalents (Janzen and Stern, 1998). Linear selection differentials were then converted to a relative fitness scale by multiplying by 1/(relative fitness). Relative fitness was calculated as the proportion of survivors for each interval (Janzen and Stern, 1998). The resulting coefficients estimate the total effect of selection on body mass – that is, the direct effect of selection on body mass plus any indirect effects owing to selection on correlated traits (Lande and Arnold, 1983; Kingsolver *et al.*, 2001).

RESULTS AND DISCUSSION

Overall, we found that the strength and direction of linear selection differentials can change when emigrants are assumed to live or assumed to die, and that the extent of this change is related to the magnitude of size-dependent emigration. Significant size-dependent emigration occurred in 5 of the 12 sampling intervals (Table 1), with larger fish more likely to emigrate in each case (positive slope parameters). The mean percentage of fish emigrating over an interval was $15.2 \pm 3.7\%$ ($\pm$ standard error) and ranged from 10.6% to 23.4%. Size-dependent emigration was most prevalent during the age-1+ fall and the age-2+ spring. Emigration in the spring (age-2+ spring to age-2+ summer) is primarily caused by individuals, termed 'smolts', starting their seaward migration, and these smolts are known to be larger than non-smolts (see Metcalfe, 1998). Emigration in the fall (age-1+ fall to age-1+ prewinter interval) may represent an early smolt emigration (Riddell and Leggett, 1981; Youngson *et al.*, 1994) or increased local movement of mature male parr (Buck and Youngson, 1982; Youngson *et al.*, 1983). The latter, however, is unlikely to explain the tendency for larger individuals to emigrate because mature parr are generally smaller than immature parr at this time (Letcher and Gries, 2003).

Overall, mean ($\pm$ standard deviation) selection differentials were stronger ($F_{1,22} = 5.5$, $P = 0.022$) when emigrants were assumed to die (-0.12 ± 0.16) than when they were assumed to live (0.003 ± 0.048). Moreover, the range of estimates among intervals was larger in the former case (–0.39 to 0.031) than in the latter case (–0.090 to 0.075) (Table 2). Assuming emigrants died yielded six intervals with P-values for selection ≤ 0.05, whereas assuming emigrants lived resulted in only one (Table 2). For the five intervals where significance was lost, the strength of selection decreased and, in two cases, changed in sign from negative to positive (Table 2). For the interval where selection was significant under both assumptions (last interval cohort 2000), it went from significantly negative to significantly positive (Table 2). These changes occurred because migrants were generally larger than non-migrants (Table 1), and assuming they died thus gave the impression of negative selection on body size. Our results demonstrate the extent to which the intensity and direction of selection can change in a natural system after accounting for trait-dependent emigration.

Table 2. Linear selection differentials (± standard error) for two cohorts of Atlantic salmon sampled on six intervals

		Emigrants die		Emigrants live		Change	
Cohort	Sample interval	β	*P*-value	β	*P*-value	Significant to non-significant?	Change sign?
2000	0+ fall to 0+ prewinter	−0.068 ± 0.034	0.05	−0.058 ± 0.032	0.07	Yes	No
	0+ prewinter to 1+ spring	−0.038 ± 0.056	0.50	0.045 ± 0.049	0.36	No	Yes
	1+ spring to 1+ summer	−0.004 ± 0.039	0.92	−0.036 ± 0.033	0.27	No	No
	1+ summer to 1+ fall	−0.006 ± 0.037	0.86	−0.023 ± 0.034	0.51	No	No
	1+ fall to 1+ prewinter	−0.121 ± 0.059	0.04	−0.090 ± 0.048	0.06	Yes	No
	1+ prewinter to 2+ summer	−0.322 ± 0.108	0.00	0.075 ± 0.037	0.03	No	Yes
2001	0+ fall to 0+ prewinter	0.019 ± 0.034	0.57	0.023 ± 0.029	0.42	No	No
	0+ prewinter to 1+ spring	−0.116 ± 0.037	0.00	0.029 ± 0.022	0.17	Yes	Yes
	1+ spring to 1+ summer	0.012 ± 0.034	0.73	0.021 ± 0.029	0.45	No	No
	1+ summer to 1+ fall	0.031 ± 0.033	0.34	0.041 ± 0.022	0.06	No	No
	1+ fall to 2+ spring	−0.393 ± 0.081	0.00	−0.022 ± 0.037	0.56	Yes	No
	2+ spring to 2+ summer	−0.396 ± 0.082	0.00	0.029 ± 0.039	0.43	Yes	Yes

Note: Emigrants were assumed to die or live (see text for details). Assuming migrants live resulted in a change in β from significant to non-significant and a change of sign of β.

When emigrants were assumed to die, the average of the absolute values of selection differentials (0.13) was not atypical for viability selection in nature. Deriving a distribution of viability differentials from Kingsolver *et al.* (2001), we estimated that the value of 0.13 fell into the 58th percentile. Some of the greatest values for particular intervals (~0.39), however, were quite high, falling into the 88th percentile. When emigrants were assumed to live, the average of the absolute values of selection differentials dropped to 0.041, which fell into the 33rd percentile and the greatest differential (0.09) yielded a value in the 53rd percentile. Thus, assuming emigrants died did not yield unreasonable results except perhaps in the most extreme cases. This suggests that the effects of emigration would not have been obvious based on inflated estimates alone.

The extent of changes in selection estimates was related to the strength of size-dependent emigration (Figure 1). For most intervals during which the size dependence of emigration was weak (P-value > 0.05), the difference between differential estimates was small (symbols close to the 1:1 line in Figure 1). When the size dependence of emigration was strong, the difference between estimates was often large (symbols away from the 1:1 line in Figure 1). Plotting the difference in differentials against the P-value for the logistic function describing size-dependent emigration yielded a significant ($P = 0.019$) relationship (difference = $0.0098 + 0.10 \cdot \log_{10}(P\text{-value})$), explaining 43% of the variance.

Selection differentials changed when known emigrants were assumed to die or to live. Thus, the nature of the 'bias' owing to emigration depends on which of these two possibilities is more likely. Three lines of evidence suggest that the more appropriate estimate is obtained when emigrants are assumed to live to the end of the interval. First, timing of

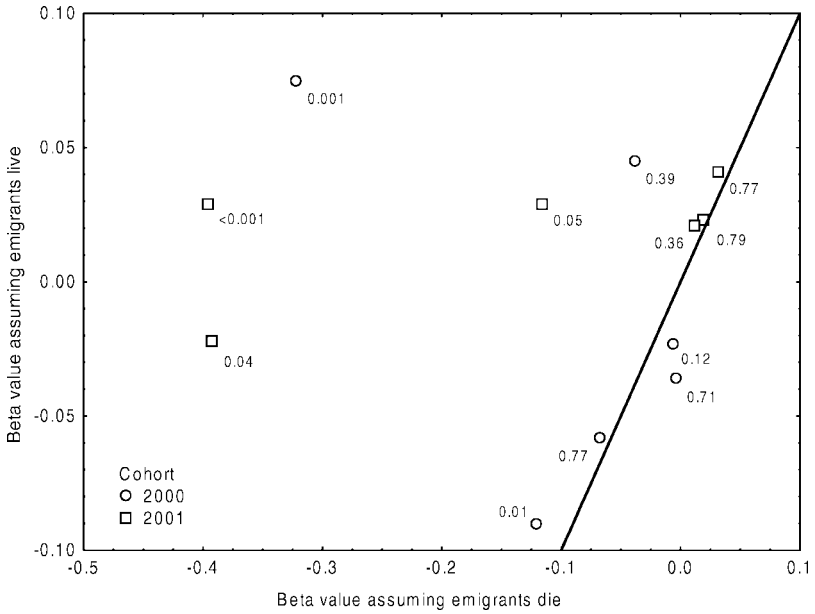

Fig. 1. Linear selection differentials (β-values from Table 2) assuming migrants die or assuming they live. Each point represents estimates of β over one of the sampling intervals in Table 2. The line is the 1:1 line and text associated with each point is the P-value for the logistic regression (from Table 1) estimating the strength of size-dependent emigration for each interval.

emigration was evenly spread throughout each interval, such that many emigrants had already lived through most of the interval before emigrating, leaving little time to die before the end of the interval. Second, approximately 5 km of suitable habitat is found below our study area, within which survival rates should be similar to those in our study area. Third, survival rates within the study site are relatively high and typically range between 0.6 and 0.8 over an interval (Letcher *et al.*, 2002). If we assume that emigrants spent one-half of the interval in the study area and survived at a rate of either 0.8 or 0.6 for the remainder of the interval outside the study area, we would overestimate survival by only 10% or 20% if the emigrants actually died following emigration. Finally, these overestimates may themselves be overestimates because the probability of capture within the study area (0.6–0.8) (Letcher *et al.*, 2004) was lower than the probability of detecting an emigrant (0.91). Overall, the actual strength of selection is probably closer to that estimated when emigrants are assumed to live but the estimate under this assumption may be slightly too high.

Two other size-dependent processes have the potential to bias results of selection analyses: size-dependent probabilities of capture and growth variation among individuals. Based on detailed analysis of mark–recapture data, we have found no evidence of such a bias in size-dependent capture for each of the sampling periods reported here (Letcher *et al.*, 2004). Bias might also arise if size at the start of an interval is not a good predictor of size at the time of selection (i.e. there was variation in growth rate among individuals). Fortunately, the correlation between size at the first and second occasion of each interval is generally greater than 0.9 (B.H. Letcher *et al.*, unpublished data). Thus, the rank order of fish body sizes does not change dramatically over an interval and initial size is a reasonable predictor of size throughout an interval. For these reasons, we ignored the possibility of size-dependent probability of capture and individual growth variation influencing bias and focused exclusively on the effects of size-dependent emigration.

We examined selection on a single trait, but many analyses use multiple regression techniques to explore independent effects of multiple, potentially interacting traits (Endler, 1986; Kingsolver *et al.*, 2001). The bias in selection differentials that we report here should also apply to selection gradients on multiple traits. Gradient estimates for traits that are not related to emigration (or any other loss of individuals from the population that could be confused with mortality) should not be affected by ignoring emigration. We have demonstrated here, however, that the intensity of selection estimates for traits that are related to emigration can be influenced substantially by accounting for emigration.

ACKNOWLEDGEMENTS

We thank all the people who helped sample the West Brook and Bob Duda for gracious access to private property. Andrew Hendry, Stephanie Carlson and Jason Coombs provided helpful comments on the manuscript. Partial funding was provided by the US Forest Service, Northeast Research Unit.

REFERENCES

Brodie, E.D., Moore, A.J. and Janzen, F.J. 1995. Visualizing and quantifying natural-selection. *Trends Ecol. Evol.*, **10**: 313–318.

Buck, R.J.G. and Youngson, A.F. 1982. The downstream migration of precociously mature Atlantic salmon, *Salmo salar* L. parr in autumn: its relation to the spawning migration of mature adult fish. *J. Fish Biol.*, **20**: 279–288.

Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton, NJ: Princeton University Press.

Fairbairn, D.J. and Reeve, J.P. 2001. Natural selection. In *Evolutionary Ecology: Concepts and Case Studies* (C.J. Fox, D.A. Roff and D.J. Fairbairn, eds.), pp. 29–43. Oxford: Oxford University Press.

Hendry, A.P., Letcher, B.H. and Gries, G. 2003. Natural selection in a restoration program for an extirpated population: size, condition and growth of Atlantic salmon in the Connecticut River. *Conserv. Biol.*, **17**: 795–805.

Hereford, J., Hansen, T.F. and Houle, D. 2004. Comparing strengths of natural selection: how strong is strong? *Evolution*, **58**: 2133–2143.

Hersch, E.I. and Phillips, P.C. 2004. Power and potential bias in field studies of natural selection. *Evolution*, **58**: 479–485.

Janzen, F.J. and Stern, H.S. 1998. Logistic regression for empirical studies of multivariate selection. *Evolution*, **52**: 1564–1571.

Kingsolver, J.G. and Pfenning, D.W. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evolution*, **58**: 1608–1612.

Kingsolver, J.G. and Smith, S.G. 1995. Estimating selection on quantitative traits using capture–recapture data. *Evolution*, **49**: 384–388.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M. *et al.* 2001. The strength of phenotypic selection in natural populations. *Am. Nat.*, **157**: 245–261.

Lande, R. and Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution*, **37**: 1210–1226.

Letcher, B.H. and Gries, G. 2003. Effects of life history variation on size and growth in stream-dwelling Atlantic salmon. *J. Fish Biol.*, **62**: 97–114.

Letcher, B.H., Gries, G. and Juanes, F. 2002. Survival of stream-dwelling Atlantic salmon: effects of life history variation, season and age. *Trans. Am. Fish. Soc.*, **131**: 838–854.

Letcher, B.H., Dubreuil, T.D., O'Donnell, M.J. *et al.* 2004. Long-term consequences of variation in timing and manner of fry introduction on juvenile Atlantic salmon growth, survival and life history expression. *Can. J. Fish. Aquat. Sci.*, **61**: 2288–2301.

Metcalfe, N.B. 1998. The interaction between behavior and physiology in determining life history patterns in Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.*, **55**(suppl. 1): 93–103.

Riddell, B.E. and Leggett, W.C. 1981. Evidence of an adaptive basis for geographic variation in body morphology and time of downstream migration of juvenile Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.*, **38**: 308–320.

Schluter, D. 1988. Estimating the form of natural-selection on a quantitative trait. *Evolution*, **42**: 849–861.

Youngson, A.F., Buck, R.J.G., Simpson, T.H. and Hay, D.W. 1983. The autumn and spring emigrations of juvenile Atlantic salmon, *Salmo salar* L., from the Girnock Burn, Aberdeenshire, Scotland: environmental release of migration. *J. Fish Biol.*, **23**: 625–639.

Youngson, A.F., Jordan, W.C. and Hay, D.W. 1994. Homing of Atlantic salmon (*Salmo salar* L.) to a tributary spawning stream in a major river catchment. *Aquaculture*, **121**: 259–267.