Evolutionary Ecology Research, 2018, **19**: 1–13

Site-specific plasticity and demography determine litter size in a population of white-footed mice

Douglas W. Morris

Department of Biology, Lakehead University, Thunder Bay, Ontario, Canada

ABSTRACT

Background: Despite heritable variation and positive selection gradients, the mean number of young that a female produces often fails to increase through time. This paradox can be resolved by mapping fitness as a manifestation of heritable variation, direct and indirect selection on traits, and environmental interactions with demography.

Questions: Does environmental variation in time and space determine the outcome of selection on litter size? Is the pattern of variation consistent with our understanding of the fitness map?

Study organism: White-footed mouse, *Peromyscus leucopus*, reproducing in nest boxes in southern Ontario, Canada.

Methods: I analysed 35-years' data on litter-size variation at two nearby study sites and interpreted the results in the context of the fitness-mapping function.

Analyses: Linear regression, mixed models, generalized linear models, model selection, mean-standardized selection gradients.

Results: There was no detectable cost of reproduction. Mean annual litter sizes did not vary at one site where shallow selection gradients were associated with small spring litters. Those litters yielded higher recruitment than larger litters born in autumn. The other site was different: mean litter size increased marginally from year to year and recruitment was higher for large autumn litters than for smaller litters produced in spring.

Conclusions: Spatial variation in apparent selection on white-footed mouse litter size is associated with site-specific, demographically induced seasonal differences in recruitment success. Failure to incorporate such relevant environmental clines can lead to biased interpretations of selection and misunderstanding of the lock-step connection between ecology and evolution.

Keywords: fitness-mapping equation, habitat, litter size, *Peromyscus*, phenotypic plasticity, predation, reaction norm.

INTRODUCTION

Despite clear evidence of heritable variation in clutch size (Postma, 2014) and frequent directional selection (Charmantier *et al.*, 2006), the mean number of offspring that a parent produces mysteriously and repeatedly fails to increase over time (Merilä *et al.*, 2001; Postma and van Noordwijk,

---

Correspondence: D.W. Morris, Department of Biology, Lakehead University, Thunder Bay, ON P7B 5E1, Canada. email: douglas.morris@lakeheadu.ca

2005; Charmantier *et al.*, 2006). Adaptive explanations for this paradox tend to hinge on costs of reproduction (Williams, 1966), annual variation in fitness (Boyce and Perrins, 1987), 'errors' in producing the optimum clutch size (Mountford, 1968), and state-dependent optimization of brood size (Smith and Fretwell, 1974; Morris, 1985; Pettifor *et al.*, 1988). These arguments assume implicitly that low survival of sibs, reductions in survival or reproduction of those parents producing clutches larger than the population mean, or parents that minimize the costs of producing individual offspring (Sikes, 1998), de-value the apparently higher fitness of large clutches.

There is a tantalizing alternative: mean clutch size might be a plastic response to environmental variation (Postma and van Noordwijk, 2005). If so, then clutch size might increase in some sites and not in others where an environmentally induced increase in clutch size is of small effect, associated with fluctuating environments, or neutralized as individuals move from one environment to another (Pemberton, 2010).

Environmental effects complicate interpretations of heritable variation in offspring production because the underlying plasticity correlates with traits other than clutch size (Pemberton, 2010). Selection on state-dependent traits such as territory quality, for example, will frequently be associated with environmental gradients, but linked only indirectly to clutch size (Cooke *et al.*, 1990). Even here, however, heritable clutch sizes should evolve in response to state-dependent selection gradients. Detecting such gradients requires that we can associate potential differences in production of offspring with environmental variability, as might occur among different habitats, and with underlying models of selection that target both the appropriate traits and their connection to environmental variation. Fitness-mapping equations (Coulson *et al.*, 2006; Morris and Lundberg, 2011) provide just such a novel and effective conceptual framework.

I begin by demonstrating how one can use fitness-mapping concepts to understand the relative effects of heritable versus environmental variability on fitness. I use those concepts to assess the long-term evolution of litter sizes produced by white-footed mice living in three habitats. I do so by evaluating changes in mean litter size through time, calculating environment-free estimates of heritability, and by associating selection gradients with the temporal pattern of litter-size variation. The results provide a consistent and coherent explanation for litter-size variation in this species.

THEORETICAL CONTEXT

Fitness is accrued through the nested mapping of how genetics influence traits, how traits are translated into demography, and how demography is converted into fitness, such that

$$w = f(D(T(G))), \tag{1}$$

where w = fitness, D represents demographic rates of survival and reproduction, T = the trait(s) of interest, and G = genotype(s) (Coulson *et al.*, 2006). Equation (1) becomes tractable by specifying the separate effects of genetics, traits, and demography on the marginal change in fitness:

$$\frac{\partial w}{\partial G} = \frac{\partial w}{\partial D}\frac{\partial D}{\partial T}\left[\left(\frac{\partial T}{\partial G} + \frac{\partial T}{\partial E}\right)\right] + \frac{\partial D}{\partial E}, \tag{2}$$

where $\partial w/\partial G$ represents natural selection on genes, $\partial w/\partial D$ is the opportunity for selection, $\partial D/\partial T$ is 'selection' (marginal effect of trait value on demography), $\partial T/\partial G$ is the response to selection emerging from heritable trait values, $\partial T/\partial E$ is phenotypic plasticity, and $\partial D/\partial E$ represents the influence of environmental variation on demography [e.g. density and frequency dependence (Coulson *et al.*, 2006; Morris and Lundberg, 2011)]. It is rather important not to equate the $\partial T/\partial E$ term with the differential expression of a heritable trait (reaction norm). Environmental influences can induce phenotypic changes at numerous stages and scales of development. Although such developmental recombination can result in heritable trait change via genetic accommodation (West-Eberhard, 2005; Laland *et al.*, 2015), it is clearly not a necessary outcome of the interaction of traits with environment. Nor is genetic change a necessary outcome of selection that acts on variation in function whether or not that variation has a genetic basis (Morris and Lundberg, 2011). Thus, I refer to $\partial T/\partial E$ as a cline in the trait × environment interaction. That cline may, or may not, be a norm of reaction.

Figure 1 illustrates implications for a change in trait value along such a cline (the solid sloped line). In Fig. 1(a), the distributions represent the frequencies of individuals with different trait values in each of two environments and the corresponding means (horizontal lines). Figure 1(b) illustrates directional selection on the cline (dashed line) and trait value when individuals successfully recruited into the population (red) have larger trait values than those that were not recruited (open; the difference between weighted means of the red versus open distributions are roughly equivalent to the selection differential). If the environment deteriorates over time, the mean trait value can remain constant despite increased elevation of the cline (environment 1 in Fig. 1c). The mean value will increase if the environment improves (environment 2 in Fig. 1c). Figure 1(d) depicts a change in the trait × environment cline associated with putatively constant environments (the environments are the same as in Fig. 1a and 1b). Weak directional selection in environment 1 yields an undetectable increase in the mean trait value. A modified trait × environment interaction increases the mean trait value in environment 2 even though direct selection favours the modal phenotype.

In order to test these conjectures, one needs to explore long-term temporal dynamics in trait values at different sites, account for environmental co-variation with traits, evaluate selection gradients, and assess heritability of both genes and environment (Morris and Lundberg, 2011). If selection gradients are similar between environments that are constant through time, then only the elevation of the cline will be subject to direct evolutionary change. If the gradient is greater in one environment than it is in another, then direct selection on the trait will alter the cline's slope. Equation (2) reminds us that the cline can also evolve through indirect effects imposed by the environment.

METHODS

Study system

I evaluated changes in mean litter size with 35 years of data on litter size and recruitment by white-footed mice, *Peromyscus leucopus*, living in wooden nest boxes in forest, edge, and fencerow habitats in an agricultural landscape in southern Ontario, Canada. These semi-arboreal nocturnal mice occupy a variety of wooded habitats in this northern part of their geographic range. Reproduction is concentrated in spring (April–June, overwintered animals only) and autumn (September–October, includes the spring-born cohort). Few females produce litters during the hot summer breeding hiatus. Adults survive only one winter season (Morris, 1989), so my data encapsulating 70 reproductive seasons represent, at a minimum, 35 generations of evolution.

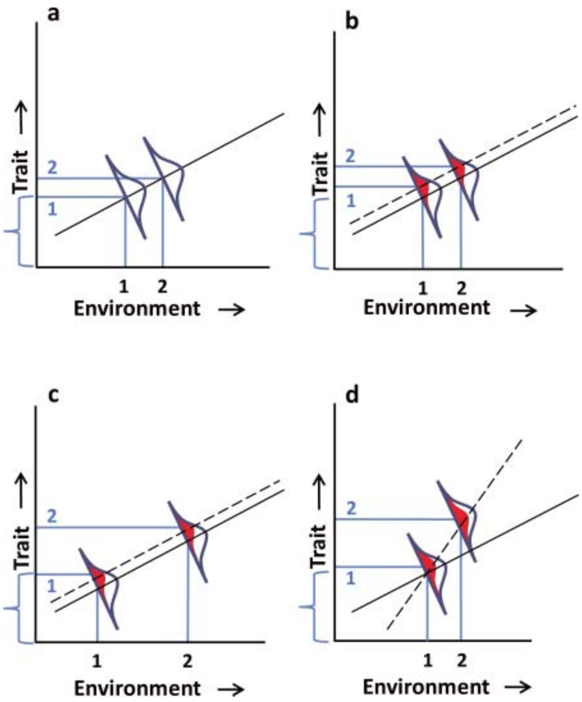

Fig. 1. Examples of the effects of environment and selection on the temporal change in mean trait value. Panel (a) represents the original trait distributions along an environmental gradient. The bracket identifies the original value of the trait in environment 1. Panel (b) illustrates a large selection differential (difference between open and red-filled distributions) that yields a new trait × environment cline (dashed line). In (c), the cline is the same as in (b) but environment 1 has deteriorated such that the mean value of the trait is the same as in panel (a) [compare brackets] while environment 2 has improved [increased mean trait value]. In (d) the environments are the same as in panels (a) and (b). Weak directional selection in environment 1 yields a minor and statistically undetectable increase in the mean trait value while, despite stabilizing selection, a dramatic change in the interaction between the trait and environment has increased the mean trait value in environment 2. Whether a change in mean trait value will occur and be observable thus depends on selection, constancy of the environment, and the trait × environment interaction [see Postma and van Noordwijk (2005) for an earlier but somewhat different version with respect to reaction norms for clutch size].

My assistants and I checked boxes during daylight hours at approximately monthly intervals and marked all juvenile and adult individuals with unique metal ear tags (for full details of the field procedures, see Morris, 1989). I placed boxes at permanent locations with a linear spacing of approximately 30 m. The two major research sites (8^{th} and Mary40) are isolated by crop fields and separated by a straight-line distance of 3.2 km. None of our marked animals moved between sites. I placed boxes at the 8^{th} research site during three timeframes (1981: 79 boxes; 1984: 4 additional boxes; 1986: a further 50 boxes). I placed those at the Mary40 site in 1992 (17 boxes). The 8^{th} site includes three habitats (fencerow, edge, forest) occupied by white-footed mice. Only one habitat (fencerow) exists at the Mary40 site. No other rodent species is abundant in the habitats containing nest boxes and, with very rare exceptions, the white-footed mouse is the only mammal that resides in the boxes.

We checked boxes three times in spring and twice in autumn between 1981 and 1986 (except autumn 1982 when no data were collected), and once monthly (April through October) thereafter (again with one exception, October 2011). I did not include data from occasional checks during other months outside of the breeding seasons.

Weaning in white-footed mice begins at about three weeks of age (Layne, 1968). Thus, I included litters in the data set only if their development (based on Layne, 1968) was consistent with animals aged 20 days or less. I deemed a littermate as 'recruited' if we recaptured it in a nest box during a subsequent check. Littermates < 5 days old were too small to mark so I excluded them from the recruitment analyses unless none survived (unsuccessful litter, all found dead in the box).

Earlier studies using data from the 8th research site document: (1) a lottery in which a minority of litters successfully recruit one or more offspring; (2) that overall recruitment success is greater in spring than autumn; and (3) that success is higher in forest than in edge or fencerow habitats (Morris, 1989). The odds associated with low litter success explain why these small mammals adopt a bet-hedging strategy of post-partum oestrus and frequent iterated reproductions (Morris, 1986). Litter size declines with population density and trends towards a larger mean size in fencerows than in either forest or edge (Morris, 1989). Despite directional selection revealed by consistently positive selection differentials, mean litter size at the 8th site is static through time and the most productive litter size of five offspring is not as common as expected (Morris, 1992). There is no relationship between litter size and either maternal survival or future fecundity that could otherwise account for the failure of mean litter size to respond to selection through time. Very low recruitment from litter sizes larger than the optimum number of offspring that mothers can care for appears to account for why mean litter size is less than the most productive in this population (Morris, 1996).

Analyses

I began by cumulating successful and unsuccessful litter sizes for each area, habitat, and season. I searched for an annual change in mean litter size with linear regression.

I estimated heritability of litter size in a three-step process. First, I assessed environmental variation in litter sizes with mixed models [dependent variable = litter size; fixed effects = season, habitat, the season × habitat interaction, and study site (when appropriate); random effect = year, denominator *df* adjusted with Satterthwaite's approximation]. I eliminated non-significant fixed effects from each model before calculating residuals and used a variance components analysis to assess the relative magnitude of annual variation on litter size. I repeated the mixed model using litters only from those boxes in place at the 8th research site for the entire 35-year period and saved the residuals in order to obtain 'environment-free' estimates of heritability for litter size with daughter-mother regressions. I used the results from the mixed models to infer patterns in litter size associated with habitat, season, and site. I calculated the best estimate of heritability and its standard error from 10,000 bootstrap samples (Brown and Shine, 2007). I used only those litters with uniquely marked mothers.

I used a similar series of generalized linear models (binomial distribution with a logit link; covariate = litter size) to assess selection gradients by first evaluating whether litter success varied by site, habitat, or season. I used Akaike's Information Criterion for finite sample sizes (AICc) to choose the most parsimonious model [lowest AICc, if ΔAICc < 2 I chose the model with fewest parameters (Burnham and Anderson, 2002)]. I restricted each analysis to years with equal sampling efforts (1987–2015 at the 8th research site; 1993–2015 for the Mary40 data).

I estimated annual and cumulative selection differentials (S) for each research site over the same time intervals. I subtracted the mean litter size of successful litters, weighted by the number of recruits, from the mean litter size (μ_L) of all litters produced (Ahola *et al.*, 2009). I calculated mean annual standardized selection gradients by dividing the product ($\mu_L \cdot S$) by the population variance in litter size ($\beta_\mu = \mu_L \cdot S/\sigma_L^2$) (Hereford *et al.*, 2004; Matsumura *et al.*, 2012). I bootstrapped these values with the same protocol that I used for heritability.

I used the selection gradients, estimates of heritability, and their association with phenotypic variance to assess whether they were compatible with the patterns of litter-size variation through time. I conducted all analyses with SPSS (Version 24) or MINITAB (Versions 17/18) software.

RESULTS

My screening requirements for inclusion yielded 2006 litters. Of these, 1760 qualified for the analysis interval between 1987 and 2015. Litter sizes varied between 1 and 11 offspring. Despite a reasonably large range of annual mean litter sizes at the 8th research site (1987–2015 = 3.98–5.54, grand mean = 4.50, $N = 1323$), there was no significant relationship between annual mean litter size and time ($F_{1,27} = 0.64$, $P = 0.43$, Fig. 2a; the same was true when I analysed the subset of nest boxes in place for the entire 35-year period). The pattern was different at the Mary40 fencerows where mean litter size increased marginally through time (litter size = 4.2 + 0.028 year, $F_{1,21} = 4.13$, $P = 0.055$, $N = 437$, Fig. 2b). This pattern was not simply a 'fence effect'. Analysis of only fencerow boxes at the 8th site for the same time period was non-significant ($N = 404$, $F_{1,21} = 0.11$, $P = 0.74$).

Mean litter size at the 8th research site (29 years, 1987–2015, $N = 1323$) was larger in fencerows (mean ± SD = 4.81 ± 1.54) than in either edge (4.42 ± 1.44) or forest (4.19 ± 1.51) habitats ($F_{2,1318.8} = 22.19$, $P < 0.001$). Mean litter size in all habitats was larger in autumn (4.62 ± 1.59) than in spring (4.33 ± 1.44, $F_{1,1290.8} = 9.49$, $P = 0.002$). Mean litter size was also larger in autumn (4.94 ± 1.70) than in spring (4.31 ± 1.64) at the Mary40 site ($N = 437$, $F_{1,433.8} = 15.66$, $P < 0.001$).

Random annual variation accounted for only 1–4% of the total variation in litter size [variance components analyses (REML): 34 years at the 8th research site (3.5%), 29 years including Wylie (3.6%), 23 years for fencerows (0.09%), 23 years for the Mary40 site (2.1%)].

Litter size varied among habitats but not between seasons in the reduced data set that I used to calculate residuals for the heritability analyses ($N = 936$; Habitat main effect, $F_{2,927.2} = 9.86$, $P < 0.001$; Season main effect, $F_{1,929,9} = 2.23$, $P = 0.14$). Even so, the mean litter size calculated in autumn (4.81) was larger than during spring (4.51).

The heritability analysis assessing residual variation in litter size in 51 daughter/mother pairs was similarly non-significant ($F_{1,49} = 1.48$, $P = 0.23$). The bootstrapped estimate (0.232, SE = 0.214) was nevertheless similar to the mean value of 0.25 revealed in Postma's (2014) survey of 37 published estimates of heritability for avian clutch size.

A total of 839 juvenile mice representing 530 different litters, from 1564 eligible litters, were successfully recruited into the nest-box population. These data reiterate my earlier conclusion (Morris, 1986) that the probability of one or more recruits is the main indicator of fitness in white-footed mice, and that the lottery of litter success is the underlying reason for frequent repeated reproduction in this species. Success at the 8th research site was

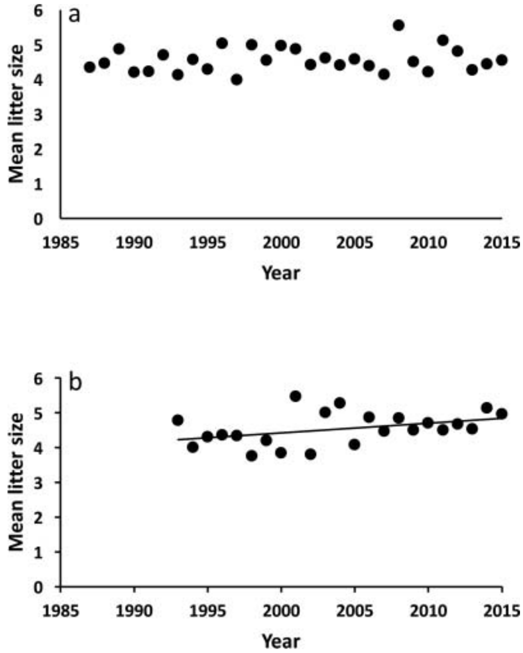

Fig. 2. Annual variation in mean litter size. Panel (a) includes litters observed across all 8th site habitats ($N = 1323$, 1987–2015), (b) includes only litters from the Mary40 research site ($N = 437$, 1993–2015).

not different between seasons or among habitats (Table 1). This was not the case for comparisons between research sites (fencerow habitat only) where litter success was higher in spring at the 8th site but higher in autumn at the Mary40 site (Table 1, Fig. 3). Litter success at both sites depended on litter size (Table 1).

There was weak selection for increased mean litter size at the 8th site (Fig. 4a; 1987–2015 data, grand mean annual standardized selection gradient = 0.59; 95% confidence interval = 0.222 to 0.923; $N = 1323$) but no significant long-term directional selection at the

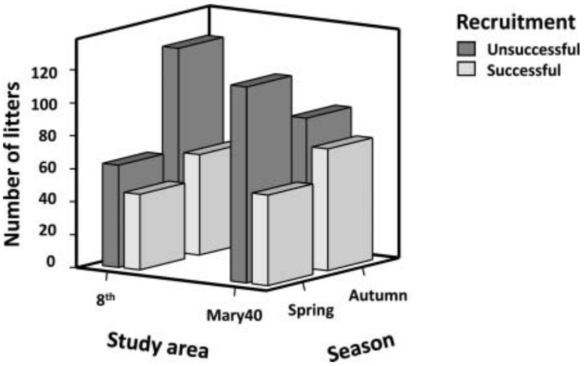

Fig. 3. The proportion of successful fencerow litters was higher in the spring reproductive season than it was in autumn at the 8th research site and the opposite at the Mary40 research site.

Table 1. Results of the generalized linear models assessing habitat and seasonal effects in the successful recruitment of juvenile mice living in nest boxes in southern Ontario, Canada

Source	Wald chi-square	*df*	*P*
8th research site, 1987–2015, $N = 1008$			
Intercept	45.24	1	<0.001
Litter size	12.31	1	<0.001
Habitat	3.62	2	0.16
Season	4.80	1	0.028
Fencerows at 8th and Mary40 research sites, 1993–2015, $N = 632$			
Intercept	34.51	1	<0.001
Litter size	18.26	1	<0.001
Season	0.24	1	0.63
Research site	0.05	1	0.83
Season × Site	8.76	1	0.003
Season × Litter size	0.46	1	0.5

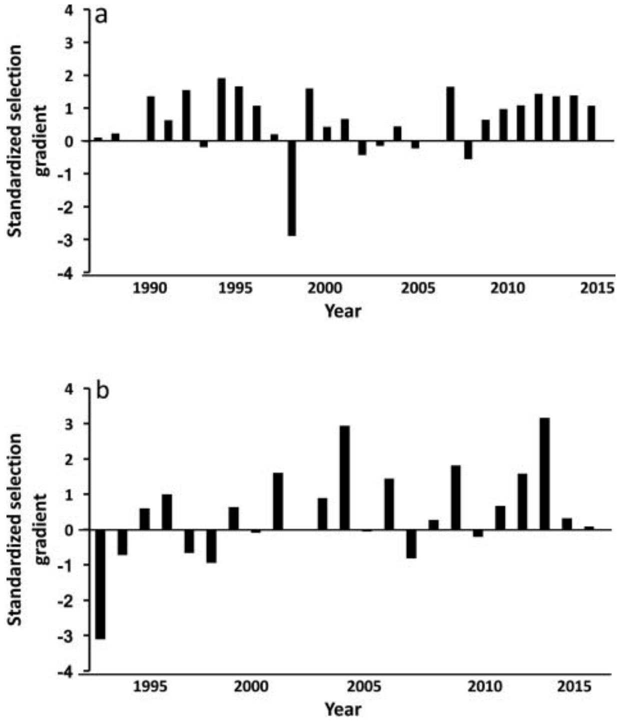

Fig. 4. Annual mean standardized selection gradients on litter size at the 8th (a) and Mary40 (b) research sites.

Mary40 site (Fig. 4b; grand mean = 0.48; 95% confidence interval = –0.107 to 1.041; $N = 437$). The difference between sites was retained when I restricted the 8th data to the same time interval as at Mary40 (grand mean = 0.58; 95% confidence interval = 0.118 to 0.971; $N = 870$). Recruitment was biased towards small litters at the Mary40 site in 1993 (Fig. 4b) and may misinform later selection for increased litter size (grand mean omitting 1993 = 0.64; 95% confidence interval = 0.185 to 1.137; $N = 428$).

DISCUSSION

Directional selection is not a sufficient explanation for patterns of litter-size variation in this population of white-footed mice. Mean annual litter size did not increase at the 8th site despite a long-term positive selection gradient. Meanwhile, mean annual litter size increased marginally at the Mary40 site where the long-term gradient was not different from zero. Selection remains insufficient even if I discard the anomalously low Mary40 value for 1993. True, selection could then account for increasing litter size at that site, but why then did litter size fail to increase at the 8th site?

Seasonally-dependent predation

Seasonal differences in recruitment success suggest a sensible explanation. Litter success at the 8th site was greatest in spring when mean litter sizes were smaller than in autumn. Large autumn litters bias the annual mean upwards. Any consistent increase in the size of small spring litters would thus have relatively little effect on the annual mean size. Trends were different at the Mary40 site, where large autumn litters yielded high recruitment success compared with smaller spring litters. But why should litter success be higher in spring at the 8th and higher in autumn at the Mary40 site?

A probable candidate is predation. Our nest-box records include observations of weasels (*Mustela frenata*), weasel scats, and weasel-predated mouse cadavers. We have never observed any of these indicators at the Mary40 site since we began monitoring boxes there in 1993. We recorded 40 observations during the same timeframe at the 8th site where it appears that sustainable weasel populations require forest habitat. Fully 30% (12) of all weasel records occurred at the end of the breeding season in October. Autumn-born mice must run this predation gauntlet without reproduction until the following spring. It is thus likely that high predation during late fall and winter yields lower recruitment from the autumn cohort than does the shorter period of pre-reproduction predation faced by spring-born litters.

Predation may also be responsible for the smaller range of litter sizes observed at the 8th research site than at Mary40. Mice foraging under risk of predation must balance the marginal cost of that risk against the marginal gain that could be achieved by more tenacious foraging (Brown, 1988, 1992; Oksanen and Lundberg, 1995; Brown and Kotler, 2004). Mice foraging in edge and forest habitats at the 8th site reveal the trade-off (Morris and Davidson 2000). The difference in quitting-harvest rates of mice foraging under cover (safety) versus in the open (risky) is less in the safe forest habitat than in the riskier edge (Morris and Davidson, 2000). We can thus anticipate that mice foraging at the safe Mary40 site will also forage more intensely, and reap the reward through larger maximum litter sizes, than is possible at the much riskier 8th location. Large litters thus gain an indirect advantage over smaller litter sizes.

Could reduced predation risk account for the gradual increase in mean litter size at the Mary40 site? The answer depends on when weasels (and perhaps other predators) disappeared from that site, how much of a reduction in litter size is associated with apprehensive foraging and elevated stress, and whether their effects have dissipated through time. Mothers transmitting stress hormones to offspring, for example, can pre-program their descendants' stress response through the hypothalamic-pituitary-adrenal axis [HPA (Sheriff *et al.*, 2009)]. Although this indirect effect can persist for multiple generations (Sheriff *et al.*, 2015), it is unlikely to do so over the 23-year duration of our field studies at the Mary40 site. If the gradual increase in litter size is a stress response, it is more likely associated with indirect evolution acting to alter the HPA axis than it is to a multi-year carry-over of pre-programmed stress responses.

Three alternatives and parsimony

You are likely, dear reader, to wonder whether one needs even to invoke an evolutionary explanation. Might the patterns be determined solely by environmental change, such as might occur if environmental quality is increasing at the Mary40 site, and constant or decreasing at the 8th? The two sites simply diverged along a static trait × environment cline. My data cannot rule out this possibility, but they cannot easily account for why selection was directional at the 8th and stabilizing at the Mary40 site. If mean litter size has indeed responded to directional selection, and nevertheless failed to increase, then the gradient itself must have shifted through time.

There are two additional candidates that can explain why consistent directional selection failed to increase mean litter sizes at the 8th research site. First, the expected increase is likely constrained by seasonally determined dispersal among the three habitats [reciprocating dispersal (Morris and Diffendorfer, 2004)]. Reciprocating dispersal emerges through seasonal differences in density-dependent population growth ($\partial D/\partial E$) that alter habitat preference (Morris *et al.*, 2004). High fitness in forest habitat in spring increases abundance in that habitat that then causes a density-dependent reduction in fitness later in the breeding season. The reduced fitness in forest motivates dispersal towards the edge habitat where slower population growth has depressed fitness less than in the forest (Morris and Diffendorfer, 2004). The opposite occurs during population decline through late autumn and winter when animals move from the lower-fitness edge to the high-fitness forest. The two effects act to increase success from small (forest) relative to larger (edge) litters, and thereby reduce the selection differential on mean litter size.

Second, heritability for litter size may be, as my data suggest, too small to yield a consistently significant increase across years. Detecting significant additive genetic variance amid the large phenotypic variation that typifies clutch and litter sizes requires much larger samples than in my analysis, even if one uses the animal model to mine detailed pedigrees (Kruuk, 2004). The requirement for large samples should give pause to those eager to hurl brickbats denouncing the low power of my heritability estimate. Low genetic variance signals a predominant role for environmental 'control' or constraint (Wilson, 2014) on trait values, and for non-genetic and indirect genetic inheritance additional to that of maternal effects (e.g. McAdam *et al.*, 2014).

It thus appears that the most parsimonious explanation for site-specific differences in white-footed mouse litter size lies in the influence of environmental variation on phenotypes and demography ($\partial T/\partial E$ and $\partial D/\partial E$ terms in equation 2). Mechanistically, the consequences of environmental variability are likely to emerge through a state-dependent version of Mountford's hypothesis that operates in white-footed mice (Morris, 1996). Females that produce a litter smaller than their optimum have less likelihood of successful recruitment than they would have achieved by producing a larger litter size. Offspring born by females that produce a litter substantially larger than their optimum have almost no chance of successful recruitment (Morris, 1996). Environmental and random variation in litter size ($\partial T/\partial E$) thus ramify their influence through a correlated influence on demography ($\partial D/\partial E$).

Reaction norms or fitness mapping?

Postma and van Noordwijk (2005) were the first to champion reaction norms as an explanation for adaptive variation in mean clutch sizes. Their approach imagined that the reaction norm was a more-or-less fixed feature of environment. The reaction norm evolved as heritable variation in clutch size increased the mean relative to its original environmentally determined value. The fitness-mapping equation (eq. 1) directs our attention instead to any trait(s) that effects an environmentally induced change in clutch size.

The two viewpoints differ more by degree than kind. Each can account for the presence or absence of an apparent response to selection. The Postma-van Noordwijk model's assumption of consistent selection on mean clutch size most likely applies along single environmental gradients of relatively small magnitude. The more general fitness-mapping approach is free of this assumption because it can account for multiple gradients along which reaction norms (or clines in the trait × environment interaction) evolve from selection on a panoply of traits.

A large and spatially variable environmental effect has at least two rather significant implications for the study of life-history variation and evolution. First, an understanding of sources of variability must include detailed assessments of environmental components most likely to influence trait values. Second, the same understanding may often be required in order to unravel how such traits do, or do not, evolve in the context of clines responding to other, often non-obvious, sources and scales of selection. That context is crucial if we are to comprehend fully the reciprocity of interactions between ecology and evolution (Laland *et al.*, 2015).

One should not mistake environmental interactions as a substitute for directly heritable variation in trait values, or that such variation has negligible influence on trait evolution. The proper interpretation is that environmentally induced variation may often reflect numerous indirect sources of selection that will be difficult to disentangle without clear attention to the ecological context of evolution.

ACKNOWLEDGEMENTS

Morridell Farms Ltd. and Ronald Wiley generously provided access to research sites. S. Vijayan and M.J. Moses helped with data management. J. Barten, C. Learn, E. Gajda, J. Enright, and K. Morris assisted with fieldwork. J. Millar collected all field data during the past two decades. Comments from P. Lundberg and anonymous reviewers helped to improve this contribution. Canada's Natural Sciences and Engineering Research Council (NSERC) sponsored the research with a series of Discovery Grants to D.W.M.

REFERENCES

Ahola, M.P., Laaksonen, T., Tapio, E. and Lehikoinen, E. 2009. Great tits lay increasingly smaller clutches than selected for: a study of climate- and density-related changes in reproductive rates. *J. Anim. Ecol.*, **78**: 1298–1306.

Boyce, M.S. and Perrins, C.M. 1987. Optimizing great tit clutch size in a fluctuating environment. *Ecology*, **68**: 142–153.

Brown, G.P. and Shine, R. 2007. Repeatability and heritability of reproductive traits in free-ranging snakes. *J. Evol. Biol.*, **20**: 588–596.

Brown, J.S. 1988. Patch use as an indicator of habitat preference, predation risk, and competition. *Behav. Ecol. Sociobiol.*, **22**: 37–47.

Brown, J.S. 1992. Patch use under predation risk: I. Models and predictions. *Ann. Zool. Fenn.*, **29**: 301–309.

Brown, J.S. and Kotler, B.P. 2004. Hazardous duty pay and the foraging cost of predation. *Ecol. Lett.*, **7**: 999–1014.

Burnham, K.P. and Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information-theoretical Approach*. New York: Springer.

Charmantier, A., Perrins, C., McCleery, R.H. and Sheldon, B.C. 2006. Evolutionary response to selection on clutch size in a long-term study of the mute swan. *Am. Nat.*, **167**: 453–465.

Cooke, F., Taylor, P.D., Francis, C.M. and Rockwell, R.F. 1990. Directional selection and clutch size in birds. *Am. Nat.*, **136**: 261–267.

Coulson, T., Benton, T.G., Lundberg, P., Dall, S.R.X. and Kendall, B.E. 2006. Putting evolutionary biology back in the ecological theatre: a demographic framework mapping genes to communities. *Evol. Ecol. Res.*, **8**: 1155–1171.

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

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the ‘animal model’. *Phil. Trans. R. Soc. Lond. B: Biol. Sci.*, **359**: 873–890.

Laland, K.N., Uller, T., Feldman, M.W., Sterelny, K., Müller, G.B., Moczek, A. *et al.* 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. Lond. B: Biol. Sci.*, **282**: 20151019.

Layne, J.N. 1968. Ontogeny. In *Biology of* Peromyscus *(Rodentia)* (J.A. King, ed.), pp. 148–253. Special Publication #2. Stillwater, OK: American Society of Mammalogists.

Matsumura, S., Arlinghaus, R. and Dieckmann, U. 2012. Standardizing selection strengths to study selection in the wild: a critical comparison and suggestions for the future. *BioScience*, **62**: 1039–1054.

McAdam, A.G., Garant, D. and Wilson, A.J. 2014. The effects of others’ genes: maternal and other indirect genetic effects. In *Quantitative Genetics in the Wild* (A. Charmantier, D. Garant and L.E.B. Kruuk, eds.), pp. 84–103. New York: Oxford University Press.

Merilä, J., Sheldon, B.C. and Kruuk, L.E.B. 2001. Explaining stasis: microevolutionary studies in natural populations. *Genetica*, **112/113**: 199–222.

Morris, D.W. 1985. Natural selection for reproductive optima. *Oikos*, **45**: 290–292.

Morris D.W. 1986. Proximate and ultimate controls on life-history variation: the evolution of litter size in white-footed mice (*Peromyscus leucopus*). *Evolution*, **40**: 169–181.

Morris, D.W. 1989. Density-dependent habitat selection: testing the theory with fitness data. *Evol. Ecol.*, **3**: 80–94.

Morris D.W. 1992. Optimal brood size: tests of alternative hypotheses. *Evolution*, **46**: 1848–1861.

Morris, D.W. 1996. State-dependent life histories, Mountford’s hypothesis, and the evolution of brood size. *J. Anim. Ecol.*, **65**: 43–51.

Morris, D.W. and Davidson, D.L. 2000. Optimally foraging mice match patch use with habitat differences in fitness. *Ecology*, **81**: 2061–2066.

Morris, D.W. and Diffendorfer, J.E. 2004. Reciprocating dispersal by habitat-selecting white-footed mice. *Oikos*, **107**: 549–558.

Morris, D.W. and Lundberg, P. 2011. *Pillars of Evolution*. Oxford: Oxford University Press.

Morris, D.W., Diffendorfer, J.E. and Lundberg, P. 2004. Dispersal among habitats varying in fitness: reciprocating migration through ideal habitat selection. *Oikos*, **107**: 559–575.

Mountford, M.D. 1968. The significance of litter-size. *J. Anim. Ecol.*, **37**: 363–367.

Oksanen, L. and Lundberg, P. 1995. Optimization of reproductive effort and foraging time in mammals: the influence of resource level and predation risk. *Evol. Ecol.*, **9**: 45–56.

Pemberton, J.M. 2010. Evolution of quantitative traits in the wild: mind the ecology. *Phil. Trans. R. Soc. Lond. B: Biol. Sci.*, **365**: 2431–2438.

Pettifor, R.A., Perrins, C.M. and McCleery, R.H. 1988. Individual optimization of clutch size in great tits. *Nature*, **336**: 160–162.

Postma, E. 2014. Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates? In *Quantitative Genetics in the Wild* (A. Charmantier, D. Garant and L.E.B. Kruuk, eds.), pp. 16–33. New York: Oxford University Press.

Postma, E. and van Noordwijk, A.J. 2005. Genetic variation for clutch size in natural populations of birds from a reaction norm perspective. *Ecology*, **86**: 2344–2357.

Sheriff, M.J., Krebs, C.J. and Boonstra, R. 2009. The sensitive hare: sublethal effects of predator stress on reproduction in snowshoe hares. *J. Anim. Ecol.*, **78**: 1249–1258.

Sheriff, M.J., McMahan, E.K., Krebs, C.J. and Boonstra, R. 2015. Predator-induced maternal stress and population demography in snowshoe hares: the more severe the risk, the longer the generational effect. *J. Zool.*, **296**: 305–310.

Sikes, R.S. 1998. Unit pricing: economics and the evolution of litter size. *Evol. Ecol.*, **12**: 179–190.

Smith, C.C. and Fretwell, S.D. 1974. The optimal balance between size and number of offspring. *Am. Nat.*, **108**: 499–506.

West-Eberhard, M.J. 2005. Developmental plasticity and the origin of species differences. *Proc. Natl. Acad. Sci. USA*, **102**: 6543–6549.

Williams, G.C. 1966. Natural selection, the costs of reproduction, and a refinement of Lack's hypothesis. *Am. Nat.*, **100**: 687–690.

Wilson, A.J. 2014. Competition as a source of constraint on life history evolution in natural populations. *Heredity*, **112**: 70–78.