PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Life History Evolution in Heterogeneous Environments: A Review of Theory

Richard M. Sibly

Phil. Trans. R. Soc. Lond. B 1996 **351**, 1349-1359
doi: 10.1098/rstb.1996.0118

Life history evolution in heterogeneous environments: a review of theory

RICHARD M. SIBLY

School of Animal & Microbial Sciences, Whiteknights, P.O. Box 228, Reading, Berkshire RG6 6AJ, U.K.

SUMMARY

Analysis of life history evolution in spatially heterogeneous environments was revolutionized by the demonstration by Kawecki & Stearns (1993) and Houston & McNamara (1992) that earlier treatments had used incorrect fitness measures. The implications of this for the analysis of organisms with and without phenotypic plasticity are reviewed. It is shown that analyses ignoring age structure can give misleading results. The plausibility and implications of the assumptions are discussed, and suggestions are made for further work. The usefulness of reciprocal transplant and common garden experiments, in providing information relevant to the assumptions and predictions, is emphasized.

Two simulation studies show that life history evolution in temporally heterogeneous environments in which trade-offs are fixed are well predicted by Schaffer's (1974) model, with modification for asymmetric variations as necessary. Unfortunately the period of study needed to observe such effects is so long as to preclude experimental tests for most organsims.

1. INTRODUCTION

The world is variable in space and time, and this inevitably affects the evolutionary process. Quite how is harder to say. Theoreticians have addressed the question using the tools of population genetics, population dynamics and, more recently, life history theory, as follows.

Population genetic models take full account of differences in performance of different genotypes, and a primary interest is in how level of dominance affects the maintenance of genetic polymorphism. The two principal schools modelling spatial heterogeneity differ as to whether or not they allow density dependence within habitats. The two extreme cases are referred to as 'soft' and 'hard' selection respectively. Soft selection models date from a paper by Levene (1953) (see also Wallace 1968), and hard selection models from Dempster (1955). Temporal variation has also been considered following the discovery by Haldane & Jayakar (1963) that differences in genotype performance during occasional catastrophes can also promote the persistence of polymorphisms. Recent reviews can be found in Hartl & Clark (1989) and Barton & Clark (1990). The life history is implicity assumed to be semelparous in these models and the generations are assumed not to overlap.

Levins (1962, 1963, 1968) introduced 'fitness sets' as an alternative way to think about the evolutionary effects of heterogeneity without worrying about levels of dominance, and evolutionary ecologists have long found this a useful approach. The relationship with population genetics is, however, not always clear (Endler 1977; Ricklefs 1990). One topic of special interest has been Levins' analysis of the effects of varying the frequency of two habitats, and his demonstration that a gradual cline of habitat frequency can produce an abrupt change in phenotype.

The study of life history evolution owes much to three influential papers (Charnov & Shaffer 1973; Schaffer 1974*a*, *b*) which set the scene for much subsequent analysis of the evolutionary implications of trade-offs. In one of these Schaffer (1974*b*) considered the impact of temporal environmental variation on the optimal life history, and the strategy of spreading one's investment between different broods in a variable environment has since become known as bet-hedging (Stearns 1976). Life history analysis of evolution in spatially heterogeneous environments has been revolutionized by the discovery by Kawecki & Stearns (1993) and Houston & McNamara (1992) that earlier treatments had used incorrect fitness measures. This has important implications for the analysis of phenotypic plasticity (the ability of inviduals in heterogeneous environments to modify their phenotypes according to the kind of habitat they find themselves in). In particular correct identification of the optimal 'reaction norm' depends on use of an appropriate fitness measure. The quantitative genetics of the evolution of reaction norms has been described by Via & Lande (1985) and de Jong (1990).

Identification of appropriate fitness measures has also been a focus of research on temporally varying environments. Using matrix representation of the life history, environmentally induced variations in the life history can be represented by sequences of matrices. Diploidy is generally ignored, and the prime interest is in what happens to the population in the long run, and

Phil. Trans. R. Soc. Lond. B (1996) **351**, 1349–1359
Printed in Great Britain

in how to assess this in terms of the population's growth rate and its probability of extinction. Little has been achieved in incorporating the effects of density dependence. These matrix analyses were initiated by Cohen (1976) and Tuljapurkar (1982), and are reviewed in Caswell (1989).

Here I review what is known of life history evolution constrained by trade-offs, in heterogeneous environments. After a brief outline of the life history process, spatial heterogeneity is considered, then temporal variation. The variation between habitats may be of two kinds. The shape or position of the trade-off may vary between habitats. Alternatively the trade-off may be fixed but other life history traits may vary. The applicability of existing results to plants may be limited by the assumptions made in the analysis. This problem and its possible resolutions are discussed in the final section.

2. LIFE HISTORY EVOLUTION

Life history evolution consists in the selection of alleles coding for some trait or combination of traits, at the expense of competitor alleles coding for other combinations of traits. What is meant when an allele is said to be 'selected' is that its numbers increase at the expense of those of competitor alleles, if there are competitor alleles, and if there are not – in which case the allele is at fixation – then the allele resists invasion by such competitor alleles as may arise through mutation. The key question is as to whether the numbers of each allele increase or decrease in the whole population. This suggests studying the rates of increase or decrease of alleles, and these are most conveniently calculated on a 'per copy' basis. Thus the 'per copy rate of increase' of an allele in a specified environment is one of its key attributes, and this will be referred to here as the *fitness* of the allele, following Sibly & Curnow (1993). Although this definition of fitness is closely related to that used in classical population genetics, the two are not identical (see Sibly & Curnow (1993) for further disussion).

The fitness, or per copy rate of increase in the numbers, of an allele can also be thought of as the population growth rate of the allele. There is a good analogy between the study of populations of individuals and that of populations of alleles, and the analytic techniques appropriate to the study of one are often also applicable to the other. In particular this is true in calcuating population growth rate from knowledge of individual life histories. In population dynamics, population growth rate, sometimes referred to as the 'Malthusian parameter', here designated F, is related to the life histories of the individuals in the population by the Euler–Lotka equation, as follows. Suppose the individuals reproduce at ages t_1, t_2, $t_3, \ldots$, then producing n_1, n_2, $n_3, \ldots$, offspring, and suppose that offspring survive from birth to ages t_1, t_2, $t_3, \ldots$, with probabilities l_1, l_2, $l_3, \ldots$ respectively. Then the Euler–Lotka equation is:

$$1 = \sum_i e^{-Ft_i} l_i n_i. \qquad (1)$$

This equation is used to calculate F from knowledge of the t_is, l_is and n_is. The equation is derived on the supposition that population growth rate is constant during the lifetime of the individuals, or alternatively on the supposition that during individual lifetimes the proportion of individuals that are in each age class is constant. The l_is and n_is can be average values provided that the averages are calculated over all those individuals alive at age 0. However it is assumed that there is no variation in the t_is.

With modified interpretation, the Euler–Lotka equation can also be used to calculate the population growth rate (i.e. fitness) of an allele, where now the life history parameters (the l_is and n_is) are the average values of the carriers of the allele (the averages being computed over all copies of the allele present in individuals alive at age 0). A formal derivation and proof is provided by Sibly & Curnow (1993).

During the evolutionary process, alleles increase in numbers in populations if they have higher fitness than their competitor alleles, and once at fixation this ability implies the ability to resist invasion by competitors. The outcome of the evolutionary process is, therefore, that there remain in the population only those alleles with highest fitness. The life histories of the carriers of these alleles are termed *optimal life histories* because they confer maximum fitness in the study environment, and the alleles can there be thought of as *optimal alleles*. Optimal life histories are, therefore, expected outcomes of the evolutionary process.

The main limitation of the above approach is the assumption that the average life history of the carriers of an allele remains constant over time. This may be approximately true in a tightly regulated population in which the optimal alleles have reached fixation, so that their fitness is zero and remains at zero. However it cannot be exactly true of an allele invading a population as then its fitness is positive, but when it approaches fixation, its fitness must decline to zero. Thus fitnesses may change when new alleles invade populations, and this will be associated with changes in individual life histories. Population growth rates, and thus fitnesses, may also change as a result of density-dependent processes.

This in outline summarizes the operation of the life history evolutionary process in a uniform environment. The central concept is the fitness of an allele, measured by its population growth rate, and related to the life history of its carriers by eqn (1). In the next section we consider how the process is modified if the population lives in a spatially-heterogeneous environment.

2. SPATIALLY HETEROGENEOUS ENVIRONMENTS

An attractive starting point is the case of a spatially heterogeneous environment consisting of a number of habitats in which the life histories of genetically identical individuals differ. Let the habitats be labelled 0, 1, 2, 3, ..., H. Consider the simple case that:

Assumption 1: gametes disperse randomly between habitats. Eventually some gametes pair up to form zygotes. Thereafter:

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

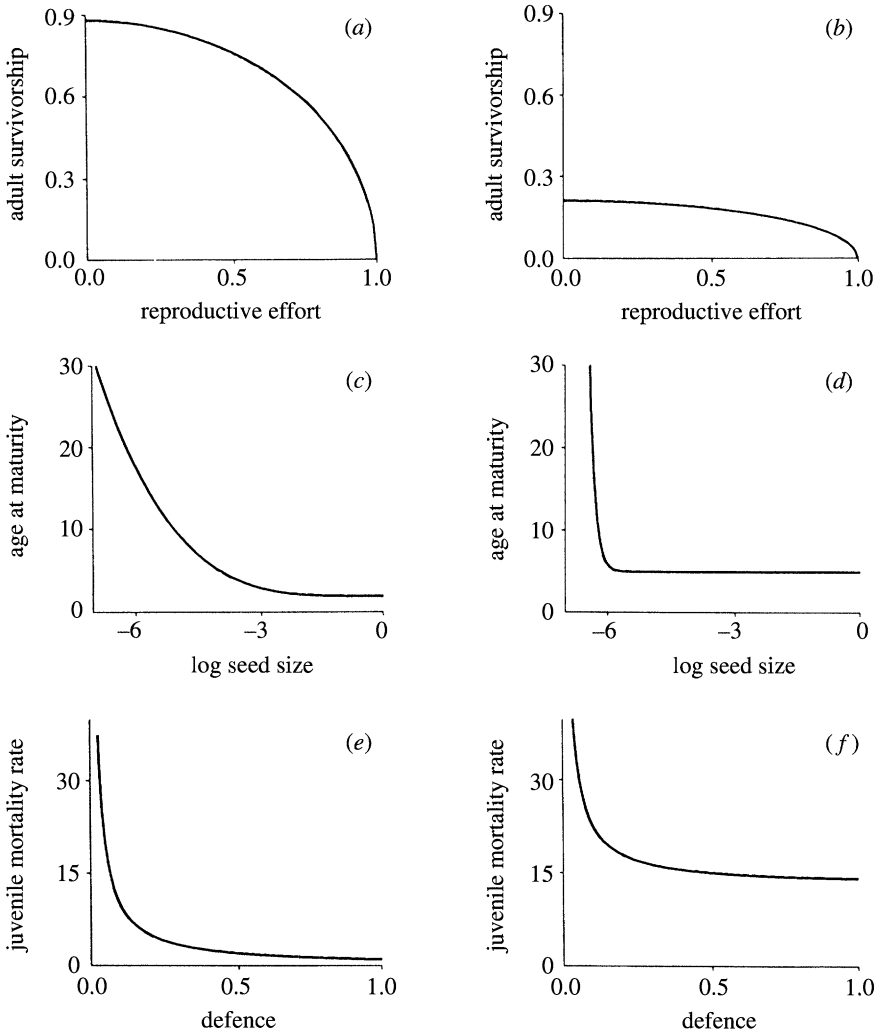

Figure 1. The forms of three possible trade-offs in each of two habitats. The cost of reproduction trade-off in habitat 0 (*a*) and habitat 1 (*b*), in which as reproductive effort increases, adult survivorship decreases. Note that the form of the relationship differs between the two habitats. A seed-size trade-off in habitat 0 (*c*) and habitat 1 (*d*). Because seed size is plotted on a log scale, -6 refers to very small seeds and 0 to relatively large seeds. It is assumed that it takes longer for individuals from smaller seeds to reach maturity, hence the negative slope of the trade-off.A trade-off between mortality rate and expenditure on defence in habitat 0 (*e*) and habitat 1 (*f*). Increasing expenditure on defence results in a reduction in mortality rate, giving the negative slope of the trade-off. After Sibly (1995).

Assumption 2: zygotes stay in their initial habitat until they die.

If a proportion $q(h)$ of newly-formed zygotes are in habitat h, and if $b_i(h)$ represents the product of l_i and n_i in habitat h, then the Euler–Lotka equation, eqn (1), takes the form:

$$1 = \sum_{h,i} q(h)\, e^{-Ft_i} b_i\,(h) \qquad (2)$$

(Kawecki & Stearns, 1993; Sibly, 1995). This equation can be used to calculate the fitness of each allele and to identify optimal alleles/life histories, just as was done above in the case of uniform environments.

It may happen that the optimal life histories change if one or more habitats are isolated from the others. In this case an isolated habitat optimum (*isolated optimum*) can be identified, which may be distinct from the global optimum. When the isolated and global optima differ, the position of the global optimum depends on the relative frequencies of the different habitat types – the $q(h)$s in eqn (2). The form of the dependency has been addressed in some stimulating papers by Levins (1963, 1963, 1968), but as mentioned earlier no satisfactory method exists by which to include Levins' approach within the framework of conventional evolutionary genetics. In a recent paper, however, Sibly (1995) tackled some simple cases numerically, as described below.

To gain some initial insight into the operation of selection in heterogeneous envirnoments Sibly (1995) analysed a simple five-parameter life cycle characterized by juvenile and adult survivorships, age at maturity, interval between breeding attempts and fecundity, which was assumed to be the same at each breeding attempt. The model can also be formulated in terms of mortality rates rather than survivorships, if appropriate. Adult size is implicity assumed constant because fecundity is the same at each breeding attempt, and adult mortality rate is assumed constant.

Different trade-offs were assumed to operate in different habitats. To keep things simple only two habitat types were considered, labelled 0 and 1. The

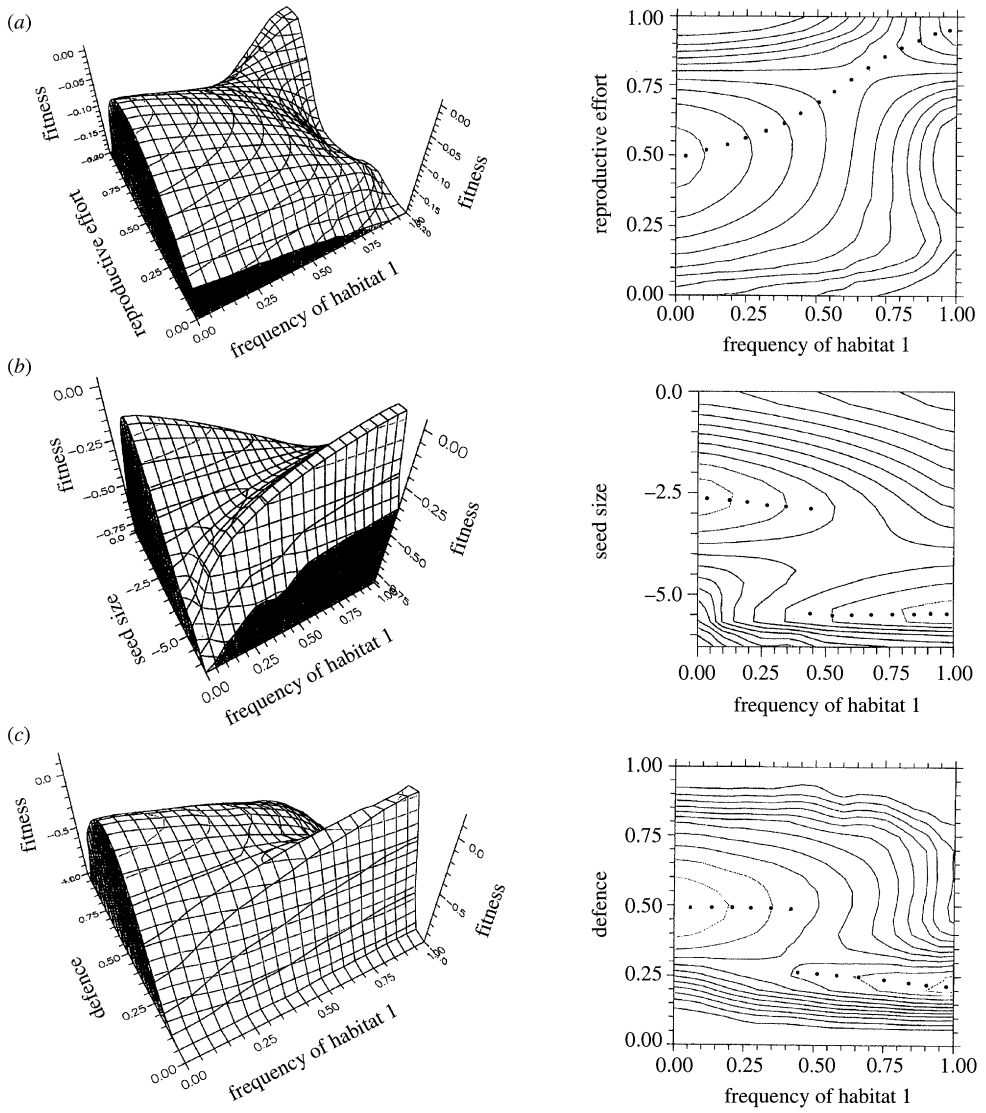

Figure 2. The fitness landscapes for a spatially heterogeneous environment consisting of two habitat types specified by the trade-offs shown in figure 1. Each row refers to a different trade-off as in figure 1. (*a*) The trade-off involving reproductive effort; (*b*) the seed-size trade-off; (*c*) the defence trade-off. Dotted lines on the contour plots indicate optimal aplastic strategies. Optimal isolated strategies occur when the frequency of habitat 1 is 0 or 1. The graphical procedure (using UNIMAP) involved smoothing and this occasionally produced minor misrepresentations. See text for further discussion. After Sibly (1995).

trade-offs operating in each habitat type are shown in figure 1. The forms of these trade-offs were chosen so that whilst broadly plausible (Sibly & Calow 1986; Sibly & Antonovics 1992) they give clear fitness landscapes (figure 2) with the isolated optima being clearly distinct and each having fitness ≈ 0.

The fitness landscapes that result from the trade-offs of figure 1 are shown in figure 2. For each trade-off a three-dimensional representation of the fitness landscape is shown on the left and a two-dimensional fitness-contour map on the right. In each case the z-axis (the 'vertical' axis) is fitness and the x-axis represents the frequency of habitat 1, i.e. the proportion of patches that are of the type of habitat 1. The y-axis represents a trait affecting the life history, and this varies according to the trade-off being considered. In constructing the landscapes shown in figure 2 it was assumed that individuals are not phenotypically plastic, i.e. do not modify their phenotypes according to the kind of habitat they find themselves in. Following Sibly (1995) we refer to a trait of this type as being 'aplastic'.

If the frequency of habitat 1 is 0, then the environment consists entirely of habitat 0, so habitat 0 is effectively isolated, and the fitness optimum is the isolated fitness optimum. In the reproductive effort trade-off (top row of figure 2) the isolated optimum in habitat 0 is a reproductive effort of 0.50. Conversely the isolated optimum in habitat 1 is observed when the frequency of habitat 1 is 1, and this occurs at a reproductive effort close to 1.

The aplastic optima (i.e. the optimal aplastic alleles) for spatially heterogeneous environments are shown by the dotted lines on the contour plots on the right-hand side of figure 2. In the reproductive effort trade-off (top row) the aplastic optima switch gradually from one isolated optimum to the other, as the frequency of habitat 1 increases. However, in the other two trade-

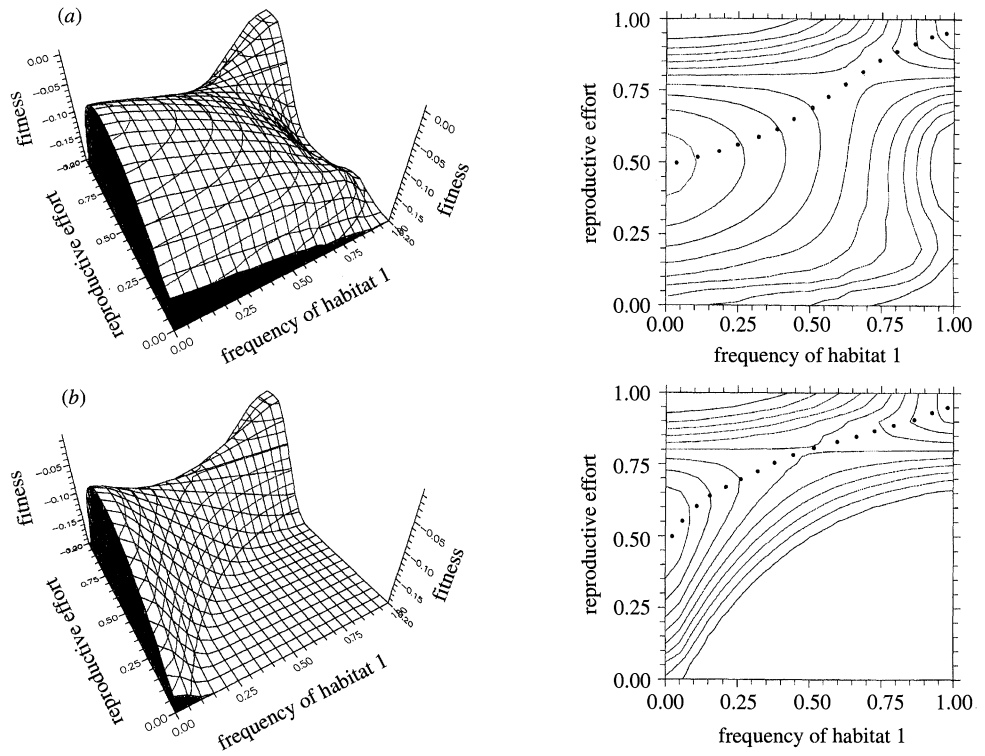

Figure 3. Fitness landscapes obtained from spatially heterogeneous models (*a*) with and (*b*) without age structure, for the trade-offs shown in figures 1*a* and *b* and 2*a*. Dotted lines on the contour plots represent the optimal aplastic strategies. The landscape in the bottom row also has an interpretation for age-structured populations in temporally-heterogeneous environments (see eqn (8)).

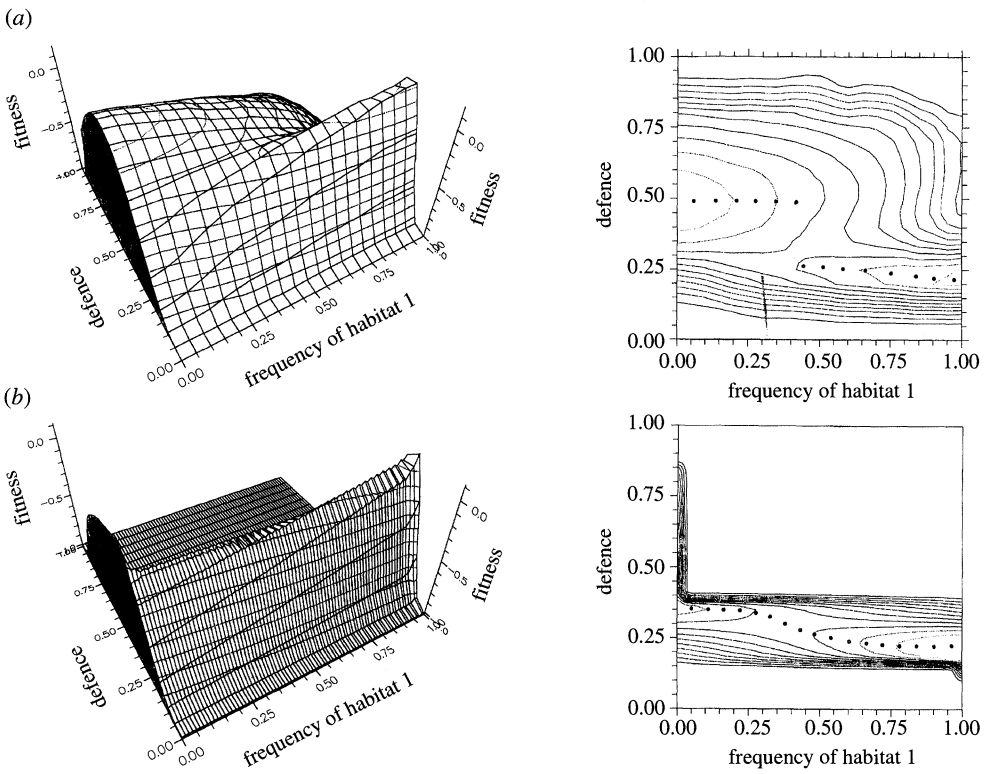

Figure 4. Fitness landscapes obtained from models (*a*) with and (*b*) without age structure, for the trade-offs shown in figures 1*e* and *f* and 2*c*. Other details as in figure 3.

offs the optimal aplastic strategy switches abruptly between the isolated optima. The abruptness or otherwise of the switch appears to depend on how sharply the fitness peaks are differentiated. In the top row of figure 2 the peaks in the isolated habitats are relatively rounded, whereas in the other cases the peaks

are sharp and clearly differentiated. This sharp differentiation produces a ridges and valley system parallel to the x-axis in the fitness landscape and it is the existence of the valley between the ridges that makes the optimal aplastic strategy switch abruptly between the isolated optima as habitat 1 increases. By contrast the top row of figure 2 lacks a valley parallel to the x-axis, so the transition is gradual.

In all three cases, then, the optimal aplastic strategy is between or almost between the isolated optima, and the global optimum switches between the isolated optima as the relative frequencies of the habitats change. The abruptness or otherwise of the switch depends on the position and structure of the valleys in the fitness landscape. Sharp differentiation of the fitness peaks leads to deep valleys between the fitness ridges, and in this case the global optimum switches abruptly from one to the other as the relative frequencies of the habitats change. However, if fitness peaks are not sharply differentiated, the switch may be gradual.

It would be attractive if it were possible to derive the fitness landscape from knowledge of the shape of the fitness peaks in isolated habitats. Some progress can be made by considering the following simple model lacking age structure. The argument will be phrased in terms of individuals, but 'allele copies' can be substituted for individuals to obtain an allelic version if desired. Suppose there are $N_h(t)$ individuals (or allele copies) in habitat h at time t, and suppose that this population would increase by δN_h between t and $t+\delta t$ if habitat h was isolated. δN_h is the difference between the number of new individuals produced in habitat h, δB_h, say, and the number that die, δM_h. Thus $\delta N_h = \delta B_h - \delta M_h$. Let F_h be the fitness of individuals in habitat h if it were isolated. Thus:

$$\frac{dN_h}{dt} = F_h N_h\,(t). \tag{3}$$

Consider now what would happen if the habitats were not isolated. The total population at time t is still $\Sigma_h\, N_h\,(t)$. The total number of new individuals produced between t and $t+\delta t$ is $\Sigma_h\, \delta B_h$ and the total number of deaths is $\Sigma_h\, \delta M_h$. Hence:

$$\delta N = \sum_h \delta B_h - \sum_h \delta M_h = \sum_h \delta N_h, \tag{4}$$

and since global fitness, F, is defined by:

$$\frac{dN}{dt} = FN, \tag{5}$$

it follows from eqns (3) – (5) that

$$\delta N = FN\delta t = \sum_h \delta N_h = \sum_h F_h N_h \delta t.$$

Hence:

$$F = \sum_h F_h \frac{N_h}{N}. \tag{6}$$

This shows that F is simply the weighted average of the F_hs, each F_h being weighted by N_h/N, i.e. by the proportion of individuals that live in habitat h. This may provide a good approximation to the age-structured model when the isolated-habitat fitnesses are similar (Sibly 1995). However, when isolated-habitat fitnesses differ markedly, ignoring age structure can give very different results from the age-structured model, as will now be shown. In comparing the models with and without age structure, note first that the proportion of individuals that live in habitat h, N_h/N (eqn (6)), is not the same as q_h (eqn (2)), which is defined as the proportion of newly formed zygotes that occur in habitat h. The two quantities are the same only if the death rates in all habitats are the same. Nevertheless the two are obviously related, and, like q_h, N_h/N can be regarded as a measure of the 'frequency of habitat 1'. It is then of interest to plot the analogue of figure 2 for the model lacking age structure, with the x-axis ('frequency of habitat 1') representing N_h/N instead of q. A comparison is made in figures 3 and 4 between the model with age structure (and $x = q_h$, top rows of figures 3 and 4) and the model without age structure (and $x = N_h/N$, bottom rows of figures 3 and 4). Figure 3 shows the reproductive effort model of figure 1 *a* and *b*. Note that dependence of fitness on reproductive effort in the isolated habitats (i.e. at frequencies of habitat 1 of 0 or 1) is the same whether age structure is included (top row) or not (bottom row). Taking a cross-section through the landscape parallel to the x-axis results in a straight line in the model without age structure (bottom row), and this is the result of the linear dependence of fitness on N_h/N in eqn (6). By contrast a cross-section through the age-structured landscape (top row) is far from a straight line. Despite this, the way the aplastic optima change as the frequency of habitat 1 increases is rather similar in the two models (compare the dotted lines in the contour plots on the right-hand side of figure 3). Figure 4 shows that qualitatively similar conclusions hold for the defence trade-off, except that the way the aplastic optima change as the frequency of habitat 1 increases differs drastically between the two models (compare the dotted lines in the contour plots on the right-hand side of figure 4).

In summary, it seems essential to include age structure when calculating the effects of selection in spatially heterogeneous environments, when the fitness peaks are sharply differentiated as in figure 4. However if the fitness peaks are relatively rounded, as in figure 3, the two models give fairly similar results, at least in computing aplastic optima.

So far we have not allowed the possibility of phenotypic plasticity, which allows phenotypes the ability to adjust their life histories according to the habitat in which they find themselves.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

4. PHENOTYPIC PLASTICITY IN SPATIALLY STRUCTURED ENVIRONMENTS

We consider first the case of 'complete' or 'infinite' plasticity (Sibly 1995; Houston & McNamara 1992) meaning that strategy in any one habitat is completely independent of strategy in any other. With complete plasticity the expected evolutionary outcome is an optimal reaction norm, a concept discussed by Stearns & Koella (1986). Complete plasticity cannot evolve if habitats are isolated and identification of the optimal reaction norm must therefore take account of the flow of alleles between habitats (Houston & McNamara 1992; Kawecki & Stearns 1993). Earlier analyses did not take this into account and as a result their analyses of optimal reaction norms were flawed.

Kawecki & Stearns (1993) introduced the important theorem that the optimal plastic strategy maximizes the reproductive value of newborns in each habitat (a proof is provided in the Appendix). This theorem is most easily understood in the case that the global population is stable and all alleles are at evolutionary equilibrium, so that $F = 0$. Reproductive value at birth is then the same as lifetime reproductive success (LRS), that is, the expected number of offspring that a newborn will produce over its lifetime. In the case that $F = 0$, the theorem states that the optimal plastic strategy is to produce as many offspring as possible in each habitat. The theorem is intuitive when one remembers the assumptions that gametes disperse randomly between habitats, and that newborn stay in their natal habitats. Nevertheless, the validity of the theorem was not generally appreciated until it was pointed out by Kawecki & Stearns (1993).

When $F \neq 0$ the reproductive value of newborns corresponds to discounted LRS, in which case offspring born at age t_i are discounted by e^{-Ft_i}. Thus in growing populations ($F > 0$) later offspring are discounted, and the faster the population grows, the higher the discount. This is because later offspring have less genetic value, because they are inserted into an expanded gene pool. Their contribution to parental reproductive value is therefore lower. Conversely in declining populations ($F < 0$) later offspring are at a premium and contribute more to parental reproductive value.

If $F = 0$, the optimal plastic strategy in each habitat is the same as the isolated optimum in that habitat (Sibly 1995), but if $F \neq 0$ this is generally not the case.

5. EFFECTS OF TEMPORAL HETEROGENEITY

In the last two sections we considered some of the evolutionary effects of making environments patchy in space. Here we consider the temporal equivalent. In both cases we suppose that life histories vary among patches of different habitat type. In the last section we supposed that these patches were distributed in the spatial environment but were constant over time. In this section we consider the converse, that patches vary over time but are spatially uniform.

Temporal heterogeneity only has evolutionary effects if the optimal strategy changes with time. Such variation may be a result of variation in the shape or position of trade-offs – as in figure 1 – or of variation in other traits that affect the optimal strategy. The simplest case is that in which the changes in the optimal strategy happen relatively infrequently, so that the transient effects of change from one optimum to another can be neglected. In this case, if a population spends a proportion $p(h)$ of its time t in habitat h, obtaining fitness F_h, then its long-term increase is given by:

$$N(t) = N(0)\exp(F_1 p(1)t + F_2 p(2)t + \ldots) = N(0)\exp(\Sigma_h F_h p(h)t), \quad (7)$$

so its long-term rate of increase, $\bar{F}$, is:

$$\bar{F} = \frac{\sum_h F_h p(h)t}{t} = \sum_h F_h p(h), \quad (8)$$

which is of the form of eqn (6), with N_h/N replaced by $p(h)$. Hence $\bar{F}$ is simply the weighted average of the F_hs, each F_h being weighted by $p(h)$, where $p(h)$ represents the proportion of time that the population spends living in habitat h. The relationship is illustrated for the case of two habitats by the bottom rows of figures 3 and 4, where now the x-axis indicates the proportion of its time that the population spends in habitat 1. Note that for any given value of x (e.g. for any given sequence of habitats) there is only one optimum. In these examples fitness falls off sharply on either side of the optimum at all values of x (except at $x = 0$ in figure 4). Thus there is never any indication in these cases that temporal heterogeneity might allow polymorphisms to persist.

Another obvious case to consider is that of random fluctuations in life history traits. The analytic theory uses matrix representation of populations and has been reviewed by Tuljapurkar (1990), Caswell (1989), and Metz *et al.* (1992). The sequence of habitats encountered is represented by a sequence of matrices, one for each time unit. The general conclusion is that, under certain assumptions, the long-term rate of increase of the population is best measured by a generalization of F known as the dominant Lyapunov exponent. In the case of small independent fluctuations in life history traits this can be expanded about its value (F_c, say) in a constant environment. Remember that the sequence of habitats encountered is represented by a sequence of matrices, one for each time unit. For each matrix a value of F can be calculated, let the variance of these numbers be written $V(F)$. Then it turns out that the Lyapunov exponent is approximately equal to

$$F_c - \frac{V(F)}{2\exp(F_c)}$$

(Caswell 1989, p. 224). Autocorrelation between successive environments can be dealt with similarly provided that environmental fluctuations are small. However, for some life histories small fluctuations in life history traits have no effect on long-term fitness,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

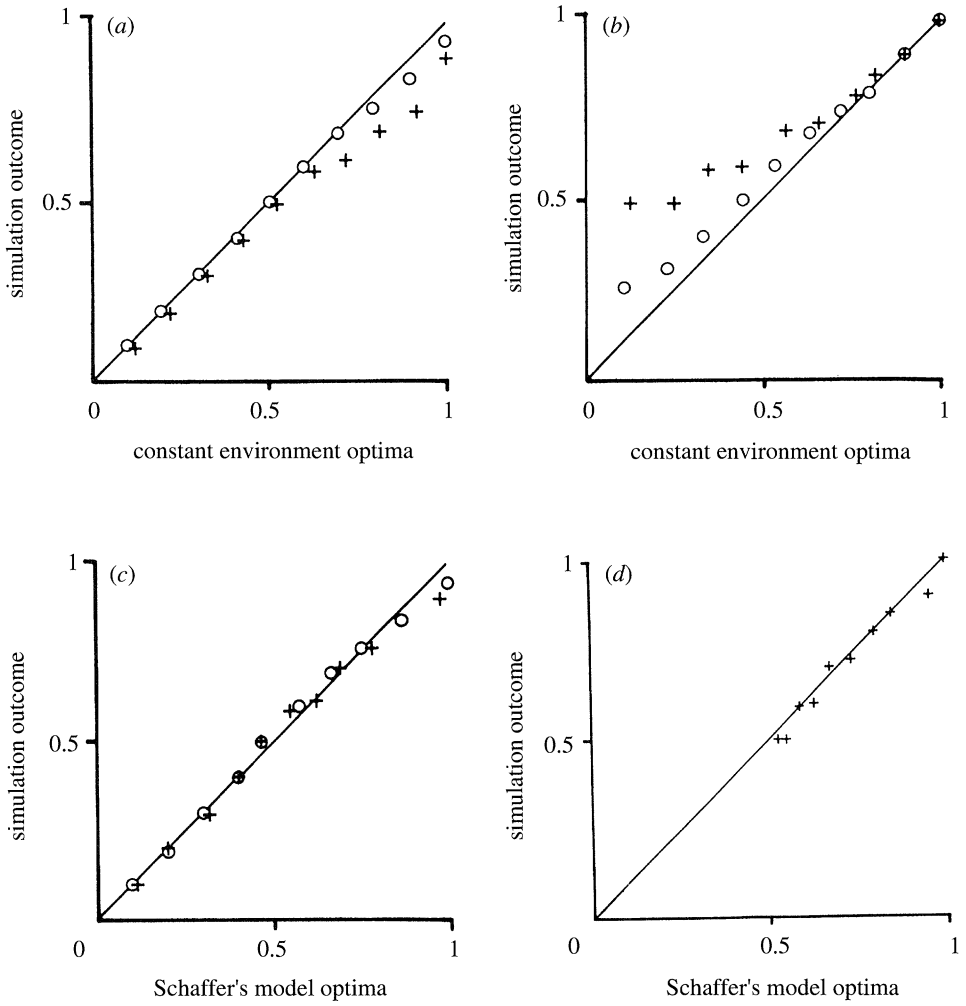

Figure 5. Effects of temporal variability in juvenile (*a*) and (*c*) or adult (*b*) and (*d*) survivorship on optimal reproductive effort. A constant trade-off of the general form of figure 1*a* was assumed. Outcomes after 5000 generations of simulated evolution in a variable environment are plotted against predictions from a modified version of Schaffer's analytical model. Variations were chosen randomly from normal distributions, with the standard deviation of the variations set equal to their mean (denoted by crosses) or half the mean (denoted by open circles). Choosing variations from uniform distributions gave very similar results. From Sibly *et al.* (1991).

so optimal strategies/evolutionary outcomes are unaffected (Sibly *et al.* 1991).

The evolutionary effects of large fluctuations in life history traits have not been widely examined, except in the case of the reproductive effort trade-off. Even here it has been assumed that the trade-off curve itself is not affected by the fluctuations. The original analysis was due to Schaffer (1974*b*). Using a simple analytical model Schaffer showed that optimal reproductive effort should decrease in an environment in which juvenile survivorship varied from year to year. In other words, in a variable world not too many eggs should be placed in one 'basket' (year). This strategy was termed 'bet-hedging' by Stearns (1976). However, if adult survivorship varies, then optimal reproductive effort should increase. Thus environmental variation leads to an increase or decrease in optimal reproductive effort according to whether it is adult or juvenile survivorship which is variable.

Bet-hedging has often been invoked by field ecologists as a possible factor affecting the interpretation of their results (see for example Godfray *et al.* 1991). However, its true significance has been hard to evaluate, partly because it is difficult to obtain the necessary measurements of survivorship variances and of the shape of the trade-off curve, and partly because of restrictive assumptions made in the original analysis. However, Sibly *et al.* (1991) and Cooch & Ricklefs (1994), using computer simulations, have shown that Schaffer's conclusions are valid beyond the range of the restrictive assumptions used in the original analysis. Sibly *et al.* (1991) showed for sexually reproducing diploid populations that whether random variation occurred in juvenile or adult survivorship from either a normal or uniform distribution, there was relatively little effect on optimal reproductive effort, even for large variations with standard deviation equal to the mean (figure 5 top row). Where optimal reproductive effort was affected, it moved as predicted by Schaffer (1974*b*), particularly when asymmetries in variation were accounted for (figure 5). Perhaps surprisingly, temporal heterogeneity had little effect on the speed of convergence to the eventual evolutionary outcome (Sibly *et al.* 1991). Cooch & Ricklefs (1994) showed similarly that large random variations in fecundity had relatively little effect on the evolution of reproductive

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

effort. In addition they showed that density-dependence and temporal environmental autocorrelation acted to reduce the effects of variation. They also found that the shape of the trade-off curve did not affect the conclusions, provided the trade-off curve was sufficiently convex.

These results may however require modification if the population dynamics are non-linear, for instance chaotic (Ferriere & Fox 1995).

6. DISCUSSION

The results pertaining to temporal heterogeneity are the most straightforward and will be discussed first. It seems that if the trade-off is fixed, then even large stochastic variations in life history traits have little effect on evolutionary outcomes or rates. Such effects as do occur are well predicted by Schaffer's (1974*b*) model, with modification for asymmetric variations as necessary. Unfortunately the period of study needed to observe such effects is so long as to preclude experimental test for most organisms. Reviews of such animal evidence as is available can be found in Roff (1992) and Stearns (1992).

To date no analysis has been made of the effects of temporal variation in the trade-off. A problem in designing such an analysis is that it is difficult to know what sorts of variation are plausible. Data on this point are hard to obtain because it is difficult to establish experimentally shapes of trade-offs even in constant environments (with the possible exception of resource allocation trade-offs; see papers in Bazzaz & Grace 1996).

The results on spatial heterogeneity have many ramifications. The three-dimensional method presented here (figure 2) supercedes Levins' (1962, 1963, 1968) two-dimensional fitness set analysis, as it is based on a rigorous definition of allelic fitness. Contrary to the general understanding of Levins' position (e.g. Rickleffs 1990), fitness in a spatially heterogeneous environment is not simply the average of the fitnesses in individual habitats, at least for age-structured populations. This can be seen by comparing figures 2, 3 and 4. Nevertheless Levins' interesting prediction that a gradual cline in habitat frequency can produce an abrupt change in phenotype is preserved (figures 2*b* and *c*). Hopefully the present more detailed theoretical treatment of the phenomenon will suggest new ways of testing this prediction. The starting point might be a series of reciprocal transplant experiments measuring the life histories and fitnesses of different strains/genotypes in each of a number of habitats (e.g. Stratton 1994). The ways in which allele frequencies actually change along environmental gradients and the mechanisms responsible for those changes have been reviewed in depth by Endler (1977). Extension of the theoretical results of figure 2 to the cases of sources and sinks, transition zones and density-dependence are briefly considered in Sibly (1995) (see also Holt 1985; McGinley 1987; Pulliam 1988; Pulliam & Danielson 1991; van Tienderen 1991; Holt & Gaines 1992).

A worrying feature of the analysis of spatial heterogeneity is its dependence on assumptions that are implausible for many plants. In particular it was assumed that:

Assumption 1: gametes disperse randomly between habitats;

Assumption 2: zygotes stay in their initial habitat until they die.

In the analysis underlying figure 2 it was assumed that:

Assumption 3: within each habitat the life history is simple and regular.

The possibility that within-individual variation might be adaptive is neglected, even though various authors have suggested that within-individual variation in seed size may have evolved as a bet-hedging strategy in a variable environment (see e.g. Geritz 1995). Of the above assumptions, only Assumption 2 is generally plausible and does not need further discussion. The implications of Assumption 3 will be discussed before those of Assumption 1.

Plant life histories are, with some exceptions, generally not simple and regular (Assumption 3). Although the present results may provide a useful starting point, it is obviously necessary to explore more realistic model life histories. The problem here is to identify appropriate models. Perhaps breeding should occur at regular intervals as above, but fecundity should increase exponentially with age, reflecting an increase in somatic size.

For Assumption 1 to be plausible, it would be necessary that dispersal distances be much greater than patch 'diameters'. Otherwise many gametes will not disperse outside the parental patch. It is known that seed dispersal is in many species of the order of metres or more (Howe & Westley 1986), and some recent work suggests that patch diameters may be less than this. For instance Stratton (1994), studying an early successional weed *Erigeron annuus*, showed that almost all G × E interactions for fitness occurred at the smallest spatial scale studied (10 cm). According to the information given, over half the seeds dispersed more than 1.5 m. In consequence progeny were likely to disperse to other patches (strictly, to microsites where their relative fitness was unpredictable). Similarly Lechowicz & Bell (1991) and Bell *et al.* (1991), studying understorey herbs in an old-growth forest, demonstrated environmental variance at scales (2 m) relevant to seed dispersal (cf Mitchell-Olds 1992).

Even if Assumption 1 is untrue, however, the fitness landscapes of figure 2 give some insight into the evolutionary process. For any given frequency of habitat 1, x say, more habitat 0 alleles end up in habitat 0 than expected on a chance basis. This means that for alleles doing well in habitat 0, the effective frequency of habitat 1 is less than x. The performance of habitat 0 alleles can then be judged by moving along the habitat 0 ridge from x towards the habitat 0 optimum. A converse shift is needed for habitat 1 alleles. Notice that this shows how traits restricting gamete dispersal are selected for in heterogeneous environments (see also De Meeûs *et al.* 1993). It is a small step to see in this process one of the mechanisms that may lead to sympatric speciation (cf Brown & Pavlovic 1992).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The fitness effects of limited dispersal also affect the form of the optimal reaction norm (the evolutionary outcome as regards phenotypic plasticity). With random dispersal the optimal within-habitat strategy maximizes $\Sigma_i e^{-Ft_i}b_i$, where F is global fitness. With no dispersal the optimal within-habitat strategy maximizes $\Sigma_i e^{-F_h t_i}b_i$, where F_h is within-habitat fitness. These strategies differ in the discounts given to later reproduction. With random dispersal, the discount is e^{-Ft_i}, with no dispersal, $e^{-F_h t_i}$. This suggests that with limited dispersal a discount $e^{-F_d t_i}$ might be applied, for some value F_d intermediate between F and F_h. This further suggests that, under plausible assumptions, the optimal reaction norm with limited dispersal would be intermediate between those for the other two cases. Further discussion of the effects of limited dispersal can be found in Sibly (1995).

Also relevant to the theory of optimal reaction norms are the effects of incomplete plasticity. Existing theory handles the cases of aplasticity and complete plasticity, but incomplete plasticity has not been considered. The difficulty is in knowing what kinds of incomplete plasticity might be relevant.

Many of the ambiguities about assumptions can be resolved by reciprocal transplant and common garden experiments. These give information about the extent of plasticity, and about its variation with genotype (G × E interaction) if any. In addition, as noted above, reciprocal transplant experiments can give information about genetic options and associated fitnesses in each environment. It might even be possible to design common-garden experiments that would produce experimentally fitness landscapes like those shown in figure 2.

REFERENCES

Barton, N. & Clark, A. 1990 Population structure and processes in evolution. In *Population biology: ecological and evolutionary viewpoints* (ed. K. Wöhrmann & S. K. Jain), pp.115–73. Berlin: Springer-Verlag.

Bazzaz, F. A. & Grace, J. (eds) 1996 *Allocation in plants*. Academic Press. (In the press.)

Bell, G., Lechowicz, M. J. & Schoen, D. J. 1991 The ecology and genetics of fitness in forest plants. III. Environmental variance in natural populations of *Impatiens pallida*. *J. Ecol.* **79**, 697–713.

Brown, J. S. & Pavlovic, N. B. 1992 Evolution in heterogeneous environments – effects of migration on habitat specialization. *Evol. Ecol.* **6**, 360–382.

Caswell, H. 1989 *Matrix population models*. Sunderland, MA: Sinauer.

Charnov, E. L. & Schaffer, W. M. 1973 Life history consequences of natural selection: Cole's result revisited. *Am. Nat.* **107**, 791–793.

Cohen, J. E. 1976 Ergodicity of age structure in populations with Markovian vital rates, I: countable states. *J. Am. statist. Ass.* **71**, 335–339.

Cooch, E. G. & Ricklefs, R. E. 1994 Do variable environments significantly influence optimal reproductive effort in birds? *Oikos*, **69**, 447–459.

De Meeûs, T., Michalakis, Y., Renaud, F. & Olivieri, I. 1993 Polymorphism in heterogeneous environments, evolution of habitat selection and sympatric speciation: soft and hard-selection models. *Evol. Ecol.* **7**, 175–198.

Dempster, E. R. 1955 Maintenance of genetic heterogeneity. *Cold Spring Harbor Symp. Quant. Biol.* **20**, 25–32.

Endler, J. A. 1977 *Geographic variation, speciation, and clines*. Princeton University Press.

Ferriere, R. & Fox, G. A. 1995 Chaos and evolution. *Trends Ecol. Evol.* **10**, 480–485,

Geritz, S. A. H. 1995 Evolutionarily stable seed polymorphism and small-scale spatial variation in seedling density. *Am. Nat.* **146**, 685–707.

Godfray, H. C. J., Partridge, L. & Harvey, P. H. 1991 Clutch size. *A. Rev. Ecol. Syst.* **22**, 409–429.

Haldane, J. B. S. & Jayakar, S. D. 1963 Polymorphism due to selection of varying direction. *J. Genet.* **58**, 237–242.

Hartl, D. L. & Clark, A. G. 1989 *Principles of population genetics*. Sunderland, MA: Sinauer.

Holt, R. D. 1985 Population dynamics in two-patch environments; some anomalous consequences of an optimal habitat distribution. *Theor. Popul. Biol.* **28**, 181–208.

Holt, R. D. & Gaines, M. S. 1992 Analaysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. *Evol. Ecol.* **6**, 433–447.

Houston, A. I. & McNamara, J. M. 1992 Phenotypic plasticity as a state dependent life-history decision. *Evol. Ecol.* **6**, 243–253.

Howe, H. F. & Westley, L. C. 1986 Ecology of pollination and seed dispersal. In *Plant ecology* (ed. M. J. Crawley). Oxford: Blackwell Scientific Publications.

de Jong, G. 1990 Quantitative genetics of reaction norms. *J. evol. Biol.* **3**, 447–468.

Kawecki, T. J. & Stearns, S. C. 1993 The evolution of life histories in spatially heterogeneous environments: optimal reaction norms revisited. *Evol. Ecol.* **7**, 155–174.

Lechowicz, M. J. & Bell, G. 1991 The ecology and genetics of fitness in forest plants. II. Microspatial heterogeneity of the edaphic environment. *J. Ecol.* **79**, 687–696.

Levene, H. 1953 Genetic equilibrium when more than one niche is available. *Am. Nat.* **87**, 331–3.

Levins, R. 1962 Theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function. *Am. Nat.* **96**, 361–373.

Levins, R. 1963 Theory of fitness in a heterogeneous environment. II. Developmental flexibility and niche selection. *Am. Nat.* **97**, 75–90.

Levins, R. 1968 *Evolution in changing environments. Some theoretical explorations*. Princeton University Press.

McGinley, M. A., Temme, D. H. & Geber, M. A. 1987 Parental investment in offspring in variable environments: theoretical and empirical considerations. *Am. Nat.* **130**, 370–398.

Metz, J. A. J., Nisbet, R. M. & Geritz, S. A. H. 1992 How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol.* **7**, 198–202.

Mitchell-Olds, T. 1992 Does environmental variation maintain genetic variation? A question of scale. *Trends Ecol. Evol.* **7**, 397–398.

Pulliam, H. R. 1988 Sources, sinks, and population regulation. *Am. Nat.* **132**, 652–661.

Pulliam, H. R. & Danielson, B. J. 1991 Sources, sinks, and habitat selection: a landscape perspective on population dynamics. *Am. Nat.* **137**, S50–S66.

Ricklefs, R. E. 1990 *Ecology*. New York: W.H. Freeman.

Roff, D. A. 1992 *The evolution of life histories: theory and analysis*. London: Chapman & Hall.

Schaffer, W. M. 1974*a* Selection for optimal life histories: the effects of age structure. *Ecology* **55**, 291–303.

Schaffer, W. M. 1974*b* Optimal reproductive effort in fluctuating environments. *Am. Nat.* **108**, 783–790.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Sibly, R. M. 1995 Life-history evolution in spatially heterogeneous environments, with and without phenotypic plasticity. *Evol. Ecol.* **9**, 242–257.
Sibly, R. M. & Antonovics, J. 1992 Life-history evolution. In *Genes in ecology* (ed. R. J. Berry, T. J. Crawford & G. M. Hewitt), pp. 87–122. Oxford: Blackwell Scientific Publications.
Sibly, R. M. & Calow, P. 1986 *Physiological ecology of animals*. Oxford: Blackwell Scientific Publications.
Sibly, R. M. & Curnow, R. N. 1993 An allelocentric view of life-history evolution. *J. theor. Biol.* **160**, 533–546.
Sibly, R. M., Linton, L. & Calow, P. 1991 Testing life-cycle theory by computer simulation – II. Bet-hedging revisited. *Computer Bio. Med.* **21**, 345–355.
Stearns, S. C. 1976 Life-history tactics: a review of the ideas. *Q. Rev. Biol.* **51**, 3–47.
Stearns, S. C. 1992 *The evolution of life histories*. Oxford: Oxford University Press.
Stearns, S. C. & Koella, J. 1986 The evolution of phenotypic plasticity in life-history traits: predictions for norms of reaction for age- and size-at-maturity. *Evolution* **40**, 893–913.
Stratton, D. A. 1994 Genotype-by-environment interactions for fitness of *Erigeron annus* show fine-scale selective heterogeneity. *Evolution* **48**, 1607–1618.
Tuljapurkar, S. D. 1982 Population dynamics in variable environments. II. Correlated environments, sensitivity analysis and dynamics. *Theor. Popul. Biol.* **21**, 114–140.
Tuljapurkar, S. D. 1990 *Population dynamics in variable environments*. New York: Springer-Verlag.
van Tienderen, P. H. 1991 Evolution of generalists and specialists in spatially heterogeneous environments. *Evolution* **45**, 1317–1331.
Via, S. & Lande, R. 1985 Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **39**, 505–522.
Wallace, B. 1968 *Topics in population genetics*. New York: W. W. Norton.

APPENDIX

Proof that, in each habitat, the optimal plastic strategy maximizes reproductive value at birth, calculated using global fitness

Let $\underline{b}(h)$ represent the life history in habitat h of the carriers of a particular allele. Let $\underline{B}$ represent the set of life histories of the carriers of the allele, i.e. the set of $\underline{b}(h)$s for all habitats. The fitness of the allele, designated $\overline{F}$, is defined by eqn (2), i.e.:

$$1 = \sum_{h,i} q(h) e^{-F t_i} b_i(h). \tag{A1}$$

Note that to each $\underline{B}$ there corresponds one and only one value of F. The optimal $\underline{B}$, $\underline{B}^*$, is that maximizing F in eqn (A1). The habitat-specific life histories comprising $\underline{B}^*$ will be described as optimal life histories, and, individually, these will be designated $\underline{b}^*(h)$. The maximum value of F will be written F^*. In other words F^* by definition satisfies the inequality:

$$F^* \geqslant F \tag{A2}$$

for all possible $\underline{B}$s.

Reproductive value at birth in habitat h when the life history is $\underline{b}(h)$ and fitness is F is defined as:

$$V(h, \underline{b}(h), F) = \sum_i e^{-F t_i} b_i(h) \tag{A3}$$

Note that $V(h, \underline{b}(\mathrm{h}), F)$ is a monotonic decreasing function of F, and that from (A1),

$$\sum_h q(h) V(h, \underline{b}(h), F) = 1.$$

We now prove the main theorem.

Theorem: The optimal life history in habitat h maximizes reproductive value at birth in that habitat, i.e. $\underline{b}^*(h)$ maximizes $V(h, b(h), F)$, where it is assumed that, outside habitat h, optimal life histories are used.

Proof: We seek to show there does not exist a life history $\underline{b}'$ in habitat h with the property that:

$$V(h, \underline{b}'(h), F') > V(h, \underline{b}^*(h), F^*), \tag{A4}$$

where F' is the fitness associated with $\underline{b}'(h)$ when in habitats other than h, the optimal life history is used.

From (A2), $F^* \geqslant F$, and since reproductive value at birth in habitat h is a monotonic decreasing function of F, it follows that:

$$\sum_{h' \neq h} q(h') V(h', \underline{b}^*(h'), F') \geqslant \sum_{h' \neq h} q(h') V(h', \underline{b}^*(h'), F^*) \tag{A5}$$

and from (A4):

$$q(h) V(h, \underline{b}'(h), F') > q(h) V(h, \underline{b}^*(h), F^*). \tag{A6}$$

Combining (A5) and (A6):

$$\sum_{h' \neq h} q(h') V(h', \underline{b}^*(h'), F') + q(h) V(h, \underline{b}'(h), F') > \sum_{h'} q(h') V(h', \underline{b}^*(h'), F^*). \tag{A7}$$

By eqn (A1) the right-hand side of eqn (A7) equals 1, so the left-hand side is greater than 1. This, however, contradicts the definition of F', which is that the left-hand side of eqn (A7) equals 1 (from eqn (A1)). Hence the supposition that there exists a life history satisfying (A4) leads to a contradiction. It follows that no such life history exists, which is what we set out to prove.