CONCEPTS & SYNTHESIS
EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

Ecology, 86(9), 2005, pp. 2344–2357

GENETIC VARIATION FOR CLUTCH SIZE IN NATURAL POPULATIONS OF BIRDS FROM A REACTION NORM PERSPECTIVE

ERIK POSTMA[1] AND ARIE J. VAN NOORDWIJK

Netherlands Institute of Ecology (NIOO-KNAW), Centre for Terrestrial Ecology, P.O. Box 40, 6666 ZG Heteren, The Netherlands

Abstract. Genetic variation for ecologically important traits determines the potential for evolutionary changes and should be measured directly. Such measurements of genetic variation based on quantitative genetic theory rely on assumptions of environmental constancy. These assumptions are not likely to hold in nature. Instead, natural environments are structured, and systematic variation in environmental conditions is an important determinant of phenotypic variation. Here we provide an introduction to quantitative genetics using a reaction norms approach, because we believe that this provides us with a good framework for combining ecology and genetics. We subsequently review the literature on genetic variation for clutch size of birds, and we show that, in spite of the inherent limitations of the methods employed, there is strong evidence that clutch size has a heritable component in natural populations of several species. However, the number of studies on the amount of genetic variation for clutch size in different species and across a range of environmental conditions is still far too small to study patterns in the relationship between heritable variation and properties of species and/or their environments. Furthermore, the role of both correlations and interactions with the environment in these estimates requires much more attention. Above all else, however, we need more information on the structure and magnitude of the environmental variation present in these studies. Future work should focus on how to obtain such data, and how to subsequently incorporate them into the proposed reaction norm framework. This requires the search for, and measurement of, relevant ecological variables. Also a more detailed investigation of the within-individual variation and the use of animal model methodology may prove to be valuable. Such additional data are essential for interpreting the amounts of genetic variation present for clutch size as a model system in the general problem of better understanding the maintenance of genetic variation in heterogeneous environments and the estimation of evolutionary potential.

Key words: bird populations; clutch size; environmental variation; genetic variation in clutch size; genotype–environment correlation; genotype–environment interaction; heritability; phenotypic plasticity; quantitative genetics; reaction norm; selection; variance components.

INTRODUCTION

A major goal in ecology and evolutionary biology is to gain a better understanding of the processes underlying evolutionary adaptation. This requires measurement of the amount of quantitative genetic variation in life-history traits directly, and cannot be substituted by estimates based on neutral markers. Such measurements are essential for predicting the impact of both local and global environmental changes on biodiversity, and will provide insight in the development and maintenance of genotypic and phenotypic variation (Frankel and Soulé 1981, Lande and Barrowclough 1987, Falconer and Mackay 1996, Lande and Shannon 1996, Frankham et al. 2002).

Historically, birds have contributed disproportionately to the development of theories concerning the development and maintenance of biodiversity. Uniquely marked individuals can be followed over extended periods in their natural habitat, phenotypic characters often can be measured accurately, and their morphological and ecological properties are exceptionally well known (Grant 2001, Bennett and Owens 2002). Moreover, nesting habits allow us to link offspring to parents. Consequently, genetic variation has been estimated for many traits in wild populations of birds (see

Manuscript received 20 February 2004; revised 14 December 2004; accepted 14 January 2005; final version received 7 February 2005. Corresponding Editor: J. R. Sauer.
[1] E-mail: e.postma@nioo.knaw.nl

reviews by Boag and Van Noordwijk [1987], Hailman [1986], and Merilä and Sheldon [2001]). Merilä and Sheldon (2001) document a total of 167 published quantitative genetic studies of natural bird populations since 1970.

The heritability of clutch size has been of particular interest, because clutch size plays an important role in theories concerning reproductive rates and population regulation (for a review of the determination of clutch size, see Klomp [1970] and Murphy and Haukioja [1986]). Lack (1948) postulated that the mean clutch size in a population is determined by the number of offspring that pairs are able to raise, the latter being determined by food availability. This view has been expanded to optimization at the individual level. The Individual Optimization Hypothesis states that birds adopt the strategy that optimizes their lifetime contribution to the next generation, i.e., their fitness, and that this optimum varies among individuals (Perrins and Moss 1975, Högstedt 1980, Pettifor et al. 1988, Pettifor et al. 2001).

For natural selection acting upon individual variation in clutch size to have an evolutionary effect, variation in clutch size should have a genetic basis. When Lack postulated his theory in 1948, he wrote that the most important gap in the knowledge of clutch size was its inheritance. Since then, much work has been done on the subject, but progress has been slow (see Haukioja 1970, Klomp 1970, Hailman 1986, Schluter and Gustafsson 1993). Although recent methodological developments (e.g., Kruuk 2004), together with the large data sets that are available for several populations, are able to resolve some of the difficulties, they have also opened up new ones, and many of the older problems still remain.

The goal of this review is threefold. First we will review the necessary quantitative genetic theory, mainly with regard to the trait of interest, clutch size. We will introduce and advocate a reaction norms approach, because we believe that it enhances understanding of the sometimes abstract subject, and it brings out previously underexposed aspects of the interplay of environmental and genetic variation. We will elaborate on the ideas first presented in Van Noordwijk (1989). Second, we will provide an overview of the research on genetic variation for clutch size, limiting ourselves to clutch size in wild bird populations. Third, we will detect gaps in our knowledge and point out challenges for future research, with special reference to the application of a reaction norm approach to data from natural populations.

Genetic and Environmental Variance Components

The genetics of quantitative traits

Organisms are complex and we therefore study them by summarizing their characteristics in a limited number of traits. Most of these traits are influenced by a large number (tens or even hundreds) of gene loci. For these so called quantitative traits, the well-developed single-locus theory of population genetics (see Hedrick 2000), which is mainly based on allele frequencies, does not apply directly because allele frequencies cannot be directly observed. Quantitative genetics theory, on the other hand, is based on phenotypic means and variances and was developed for the analysis of quantitative trait variation (Lynch and Walsh 1998). Quantitative traits are characterized by the fact that the environment often plays a prominent role in shaping the phenotype. In a laboratory or in an agricultural context, the environmental variation is limited as a consequence of efforts to keep it constant. The remaining environmental variation can then be treated as error variance, i.e., as unpredictable variance without structure. In natural environments, one cannot and should not ignore the structure in the environmental conditions. We therefore have to build a framework in which we can study the effects of genetic and environmental variation simultaneously, as first proposed by Van Noordwijk (1989).

Reaction norms

Here we advocate the use of reaction norms when studying quantitative traits in a natural setting. A reaction norm is the function that relates the phenotype to a value on an environmental axis for a given genotype (Woltereck 1909). Because a population typically consists of more than one genotype, the phenotypic variation for a quantitative trait observed within a population can be described by means of a bundle of reaction norms (Van Noordwijk 1989).

The use of reaction norms provides us with several interesting opportunities. First, both the environmental and the genetic component of the phenotypic variation gain equal attention. Second, information on the causes and mechanisms underlying the observed phenotypic variation can be visualized or inferred if unknown. Finally, reaction norms allow for the integration of phenotypic plasticity into quantitative genetics. More generally, they provide a useful approach to understanding the relationships among the different parts of complex systems, and thus are a tool for combining ecology and genetics into a single framework. It should be borne in mind, however, that the choice of what should be represented on the (single) environmental axis is far from trivial.

We will now explore how the (additive) genetic and environmental variance components appear in a reaction norm framework (Fig. 1). Individual reaction norms represent the different genotypes in a population, and the height of a reaction norm relative to the average height in a given environment gives the genotypic value. In general, the genotypic values are assumed to be normally distributed within the population with a variance V_G. Thus, the amount of genetic vari-

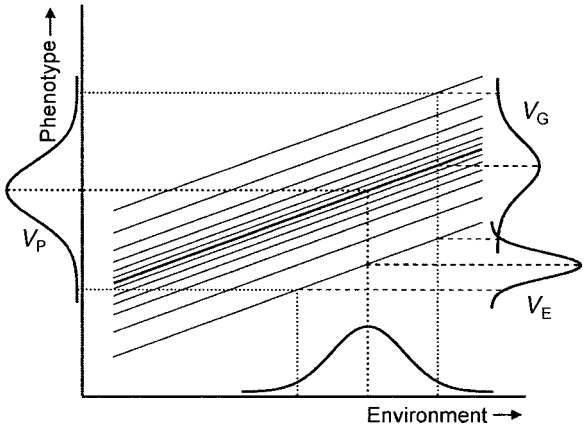

FIG. 1. The distribution of phenotypes as a result of variation in environment and genotypes. The curve along the *y*-axis gives the frequency distribution for a quantitative trait within a population. Its variance is the phenotypic variance V_P, which is the sum of variance due to genetic variation, V_G, and environmental variation, V_E. The curve along the *x*-axis gives the frequency distribution for the environments encountered by the population.

ation V_G, or V_A in the absence of dominance or epistasis, is related to the width of the reaction norm bundle, or the variance between reaction norms. The contribution of the environmental variation to the total phenotypic variation is given by V_E. Here we must take two aspects of the role of the environment into account. First, it is the distribution of the environmental conditions as they are experienced by the genotypes within a population (given on the *x*-axis) that shapes V_E. Second, there is the responsiveness of a genotype toward environmental variability, or the amount of phenotypic plasticity. This is depicted by the slope of the reaction norms. On the whole, V_E is given by the variance expressed within genotypes or reaction norms, across the range of environmental conditions encountered by the population.

Measures of genetic variance for quantitative traits

The most straightforward measure of the amount of genetic variance for a trait would be the genetic variance divided by the phenotypic variance, or V_G/V_P, which gives the broad-sense heritability or H^2. However, because the gene is the hereditary unit of transmission, the rate and amount of response to selection is only a function of the amount of additive genetic variation, V_A. It is for this reason that most evolutionary biologists are mainly concerned with V_A/V_P, known as the narrow-sense heritability and referred to as h^2 (Lynch and Walsh 1998). Note that h^2 is a relative measure of the size of V_A, and by itself it provides little information on the size of the variance components. From a reaction norm perspective, the heritability is expressed in the width or the slope of the bundle of reaction norms, relative to the width of the phenotypic distribution. The smaller the slope, the higher is h^2 (Fig. 2). Notice that this is opposite to the situation in a parent–offspring regression, in which a steep slope represents a high heritability (Lynch and Walsh 1998).

Although the narrow-sense heritability might be a good measure of the relative size of V_A, Houle (1992) has suggested that it might not be the best measure for the potential response to selection or evolvability of a trait, which depends on the absolute size of V_A. He therefore proposed to use the coefficient of additive genetic variation, CV_A, which is given by

$$CV_A = \sqrt{V_A}/\bar{X}$$

where $\bar{X}$ is the mean of the trait. The CV_A can be en-

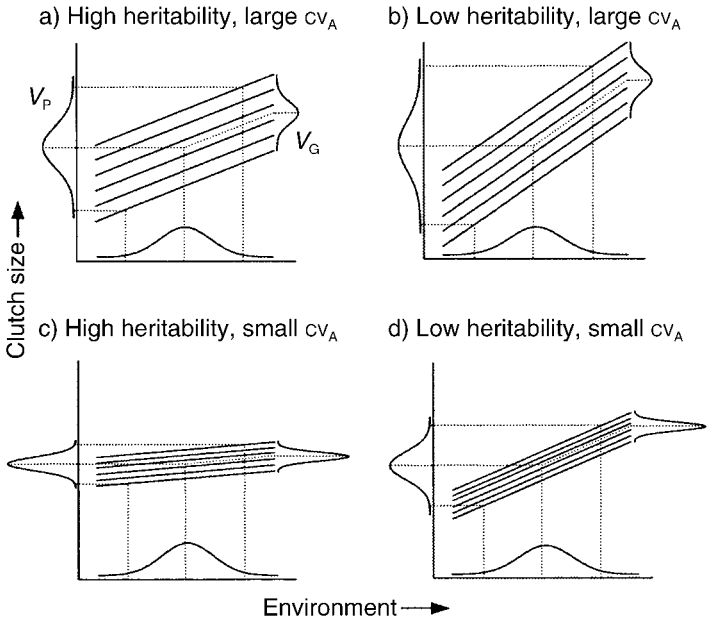

FIG. 2. The relationship between phenotypic variance V_P, narrow-sense heritability h^2, and the coefficient of additive genetic variation CV_A for a standard distribution of environmental conditions. Here, h^2 is high in both (a) and (c), and low in (b) and (d), but the CV_A is much larger in (a) than in (c). Similarly, h^2 is low in both (b) and (d), but the CV_A is much larger in (b) than in (d).

visaged as the width of the bundle relative to the phenotypic mean (Fig. 2). However, the coefficient of variation lacks a simple biological interpretation, as opposed to the heritability (Lynch and Walsh 1998).

In a meta-analysis comparing heritabilities and coefficients of additive genetic variation for different traits and their relationship with fitness, Houle (1992) found a negative relationship between the heritability of a trait and its association with fitness. Such a relationship had been observed previously by Falconer (in Falconer and Mackay 1996) and Gustafsson (1986). Initially this result was interpreted as support for Fisher's (1930) fundamental theorem of natural selection, namely, that selection should deplete genetic variation in life-history traits that are more closely associated with fitness (such as clutch size) at a more rapid rate relative to those traits less closely associated with fitness. However, Houle found a positive relationship between both the amount of additive genetic variance (V_A), and the coefficient of additive genetic variation (CV_A) for a trait and its association with fitness. This suggests that it is the residual variance that is higher in traits closely related to fitness, as a result of a higher environmental or nonadditive genetic component (Crnokrak and Roff 1995). Since then, a similar pattern has been found in the Collared Flycatcher (*Ficedula albicollis*) population on Gotland (Merilä and Sheldon 2000), and the Great Tit (*Parus major*) population in Wytham Woods (McCleery et al. 2004). This illustrates the importance of comparing variance components, in addition to heritabilities.

Which measure of genetic variation is most appropriate depends on the context and the questions to be answered. All measures are interchangeable, however, provided that information on phenotypic variance, additive genetic variance, and the trait mean is available. Unfortunately, as has been pointed out by Van Noordwijk (1990) and later by Houle (1992) and Merilä and Sheldon (2001), this is often not the case.

The environmental axis

The framework presented here makes many assumptions and is, of course, a simplification of reality because (1) the factor of time is ignored, and (2) all aspects of the environment affecting the trait under investigation are represented on a single axis. We can distinguish two different approaches that may lead to an accurate description of the environmental variation relevant for the trait under investigation. The first involves a bottom-up approach, in which we try to describe the characteristics of the environment and nongenetic characteristics of the individual that are relevant for the trait of interest, in this case clutch size. This requires a very detailed understanding of which environmental cues a bird uses before and during egg-laying to predict, for example, food availability during chick feeding, and which factors are limiting clutch size, both on a within- and a between-year basis. For example, a relationship has been shown between caterpillar density and clutch size in Great Tits (*Parus major*), both among years (Perrins 1965), and within years (Verboven et al. 2001). A relationship also exists between the amount of rainfall and mean clutch size in several species of Darwin's finches, *Geospiza* spp. (Grant et al. 2000). Individual characteristics that have been shown to correlate with clutch size include juvenile body mass in Great Tits (Haywood and Perrins 1992), parasite load and immunocompetence in Barn Swallows (*Hirundo rustica*; Martin et al. 2001), and condition, measured as the amount of glycosylated hemoglobin in the blood of Collared Flycatchers (*Ficedula albicollis*; Andersson and Gustafsson 1995). Unfortunately, however, the observed correlations are often weak, and many different environmental variables will have to be combined to get to a reliable description of the environment.

Instead of first identifying and then measuring as many environmental variables as possible, we can take a top-down approach and use the observed clutch size as the outcome of all relevant environmental effects. The traditional method for this is to describe each set of environmental conditions by the mean phenotypic value of a particular set of genotypes (e.g., Mather and Jinks 1982). This approach, however, assumes an experimental context in which genotypes can be divided over environments. This assumption may well be violated in natural populations, both in space and in time, due to habitat choice. We will discuss this in more detail in the section on genotype–environment correlations. The recent application of animal model methodology (Lynch and Walsh 1998, Kruuk 2004) provides another tool for tackling this problem in pedigreed populations, because it allows for the separation of an individual's phenotype into its genetic and environmental component at an individual level. Although most studies so far have mainly focused on the estimation of genetic effects (see Kruuk 2004 for an overview), it is also possible to estimate environmental effects. This would make it possible, on a yearly basis, to obtain a description of the environment that is independent of the genetic composition of the population in each year, but also at the level of, for example, a nest box or an individual bird.

Selection on Reaction Norms

In the absence of any limits to natural selection (see Bradshaw 1991, Barton and Partridge 2000), a reaction norm is expected to be shaped to produce the phenotype with the highest fitness in all environments (Via and Lande 1985, Stearns and Koella 1986, De Jong 1990) (Fig. 3). The two aspects of a reaction norm, the (mean) genotypic value or height, and the slope (the amount of plasticity), however, are shaped by different aspects of the fitness curve. The average height is given by the mean fitness of the phenotype over all environments ''encountered'' by the reaction norm. The slope, on the

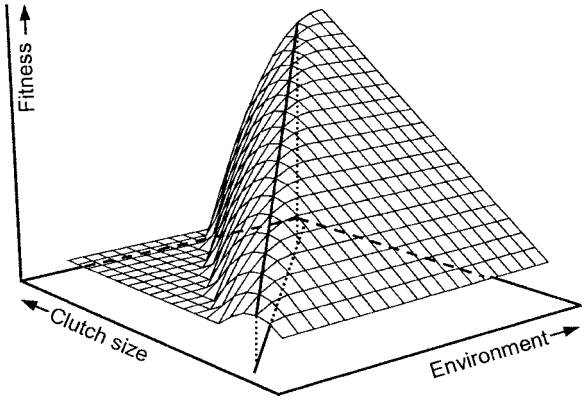

FIG. 3. Fitness consequences as a function of phenotype and environment. Every pairwise combination of phenotype and environment will result in a certain fitness value, which will create a typical (hypothetical) fitness landscape. From this landscape, we can derive the reaction norm that produces the phenotype with the highest fitness in every environment, given by the thick line.

other hand, is given by a trade-off between the gain in fitness by having the optimum phenotype over all environments, and the decrease in fitness due to the costs of plasticity, as well as its limits (see DeWitt et al. 1998). An important factor in this trade-off is the amount of environmental variation encountered by a reaction norm in both time and space, and the predictability of this variation (Scheiner 1993, Sultan and Spencer 2002).

There is a direct relationship between the selection differential and the response to selection, as given by the breeders' equation (Falconer and Mackay 1996). The strength of this relationship is given by the heritability of the trait under selection. If the heritability is close to zero and the reaction norms are thus steep, selection will mostly act on variation within reaction norms (environmental variation). Although the phenotypic mean of the selected group will be different from the population mean (selection differential different from zero), this difference does not have a genetic basis and will not be transmitted to future generations. On the other hand, if the heritability is close to unity and the reaction norms are almost horizontal, the phenotypic deviation of the selected group from the population mean will have a major genetic component and thus will be transmitted to their offspring, resulting in a response to selection (also see Fig. 4). This relationship is well accepted and has been confirmed in many selection experiments in captivity (see, e.g., Falconer and Mackay 1996, Lynch and Walsh 1998).

Clutch size has been shown to be under directional selection to increase in several populations of birds, based on persistently positive correlations between components of fitness and clutch size, or differences between the most productive clutch size and the average clutch size (references in Price and Liou [1989], Gibbs [1988], and Larsson et al. [1998]; but see Van Noordwijk et al. [1981] and Sheldon et al. [2003]). Taking this and the fact that there is general agreement that clutch size has a heritable component into account, clutch size is generally expected to increase with time. However, several long-term studies on wild populations have failed to find any evidence for this. The same phenomenon has been observed for traits such as laying date, tarsus length, and condition. For an extensive review of this subject, see Merilä et al. (2001*b*).

One of the possible causes of ''evolutionary stasis'' is related to the interaction of environmental and genetic change, and should thus be studied in a reaction norm approach. One of the basic assumptions of the breeders equation is that there are no systematic changes in the environment, which is unlikely to hold in nature. In fact, natural selection and environmental change are closely interleaved. Selection often is the direct result of environmental change (e.g., Visser et al. 1998), but selection also may be accompanied by a change in the environment that is not directly related to the observed selection pressures (e.g., Conover and Schultz 1995, Larsson et al. 1998, Merilä et al. 2001*a*). Finally, genetic change due to selection may result in a change of the environment (Cooke et al. 1990, Frank and Slatkin 1992). These environmental changes subsequently may result in phenotypic change by means of plasticity. For example, if females with large clutches recruit more offspring, this will increase population density. Negative density-dependent processes subsequently may reduce the environmental component of clutch size, counteracting the effect of selection (Cooke et al. 1990). On the whole, the interplay of environ-

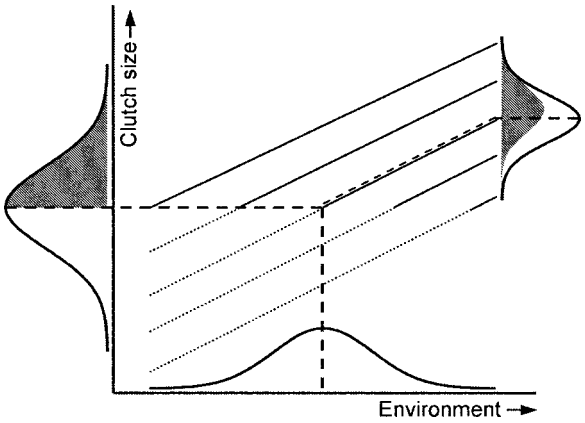

FIG. 4. The effect of phenotypic selection on genotypic values. Birds with a low genotypic value may produce an above-average clutch size when they are in a relatively good environment. The reverse will also be true. Selecting only the shaded part of the curve on the left (representing the phenotypic variation) will include some ''small clutch'' genotypes and miss some ''large clutch'' genotypes. The shape and mean of the distribution of selected genotypes (which shows the response to selection), given by the shaded area under the curve on the right (representing the genotypic variation), depends on the slope of the reaction norms and thus on the heritability.

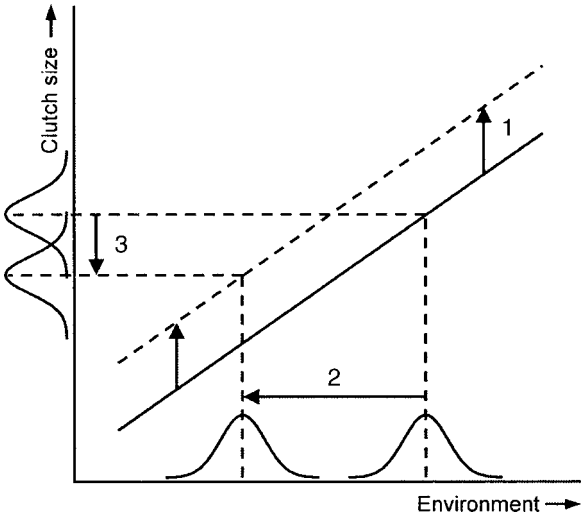

FIG. 5. Phenotypic response may be opposite to genotypic response. In the situation presented here, selection increases the genotypic value for clutch size (1), so the mean reaction norm shifts upward. Simultaneously, environmental deterioration will result in a leftward shift along the mean reaction norm (2), and, consequently, a decrease in mean clutch size (3). Thus, even though the genotypic value has increased, the mean phenotype has changed in the reverse direction.

mental change and selection may either obscure or reinforce a response to selection (Fig. 5).

Estimation of Genetic Variance for Clutch Size and Its Complications

For a thorough account of the different approaches to estimating heritabilities from the phenotypic resemblance among relatives, and quantitative genetics theory in general, see Falconer and Mackay (1996) and Lynch and Walsh (1998). For an overview, see Hedrick (2000: Chapter 11), Frankham et al. (2002: Chapter 5), or the reviews by Boag and Van Noordwijk (1987) and Merilä and Sheldon (2001).

General problems

Regression techniques and analyses of variance play prominent roles in the estimation of heritabilities of clutch size. Several assumptions are made in these analyses, many of which are not likely to be met in natural populations (Lynch and Walsh 1998). For example, it is assumed that populations are in Hardy-Weinberg equilibrium, that there is no inbreeding or selection, that parents are randomly sampled from the population, and that both environmental and genetic means and variances are constant over generations. Furthermore, least square methods ideally require a balanced design. One should also keep in mind that most regression techniques are highly sensitive to outliers. Another point of note is that low heritabilities are only detectable with very large sample sizes (Lynch and Walsh 1998). Finally, pedigrees based on social interactions may not always provide reliable information on genetic relationships if extra-pair paternity occurs within a population, resulting in conservative heritability estimates. Variation in the estimates of additive genetic variance among populations may reflect different levels of extra-pair paternity (Merilä and Sheldon 2001).

Genotype–environment correlations

So far we have assumed that genotypes are randomized over environments. We will now relax this assumption and investigate the importance of genotype–environment correlations. If there is a correlation between environment and genotype, then

$$V_P = V_G + V_E + 2\mathrm{Cov}_{G,E}.$$

A positive correlation between genotypic and environmental value may arise when, for example, genotypes with higher values are also in better environments, and the poorer genotypes have poorer environments. This could be due to either competition or dispersal (e.g., Garant et al. 2005). A genotype–environment correlation at a higher level can arise when analyzing phenotypic trends over time, or when we ignore population substructure (e.g., Postma and Van Noordwijk 2005).

According to Falconer and Mackay (1996), such correlations are seldom an important complication in experimental populations where the randomization of environments is a major aspect of the experimental design, and they suggest that the correlations are analyzed as if they were part of V_G. However, in natural populations, genotype–environment correlations are biologically interesting, and we will lose information by considering them as part of V_G in advance.

When genotype–environment correlations are present, each reaction norm is confined to a part of the environmental range in the population, resulting in a steeper slope at the population level than on an individual level (Fig. 6), but also the reverse may be true. This will lead to an incorrect inference about the role of both environment and genes in shaping the observed phenotype–environment relationship. Also, the heritabilities estimated from parent–offspring regression may be misleading (Lynch and Walsh 1998). Finally, genotype–environment interactions can only be estimated when genotype–environment correlations are absent (Gustafsson and Merilä 1994).

Genotype–environment interactions

So far we have assumed that the amount of phenotypic variation is the sum of genetic and environmental variation, and possibly a covariance term, and we thus assumed that the individual reaction norms are parallel. However, if an interaction between environment and genotype were present, then

$$V_P = V_G + V_E + V_{G\times E}.$$

Genotype–environment interactions result in nonparallelism, and thus variation in the slopes of the individual reaction norms. There are several ways in which

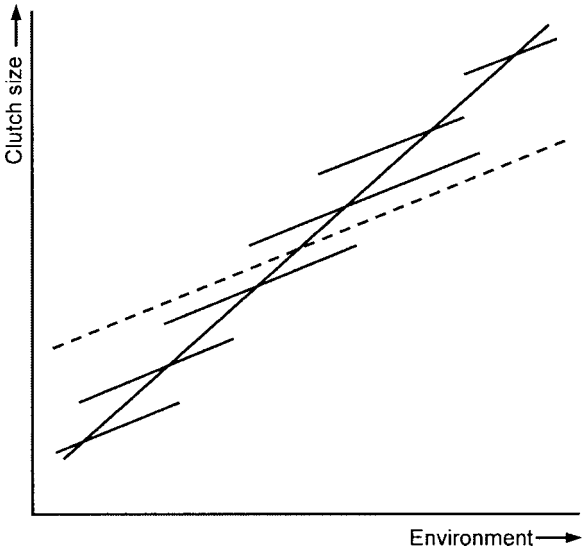

FIG. 6. Nonrandom distribution of genotypes over environments leads to a discrepancy between average reaction norm and mean phenotype as a function of environment. With a genotype–environment correlation, genotypes or individual reaction norms are restricted to a specific part of the environmental range. In this case, the mean phenotype (continuous line) across all environments will not give the mean reaction norm (dashed line).

nonparallelism may arise, as illustrated in Fig. 7. First, the expression of genetic variance may vary across environments, which is reflected in variation in the width of the reaction norm bundle. In Fig. 7a, the bundle is narrower and the heritability is thus smaller in poor environments. However, the different reaction norms do not intersect. Therefore, the order of the different genotypes does not change and the genetic correlation across environments is equal to one. Such an interaction often can be removed by a scale transformation of the phenotype axis. Second, the genetic correlation between environments may be different from unity (Fig. 7b), due to either different genes controlling the expression of genetic variation, or to a difference in the expression of a single gene under different environmental conditions (Scheiner 1993). Although the width of the reaction norm bundle, and thus V_A, is variable across environments in Fig. 7a, this is not the case in Fig. 7b, even though genotype–environment interactions are present here.

Genotype–environment interactions may have important evolutionary consequences. First, because the reaction norms differ in their slope and thus show different amounts of plasticity, they may enable the evolution of phenotypic plasticity and permit populations to evolve to have a more optimal phenotypic mean in different environments (Via and Lande 1985). In the long run, this will allow for adaptation to heterogeneous environments. In the short run, however, genotype–environment interactions may slow down local adaptation to a variable environment. On the other hand, according to Gillespie and Turelli (1989), they can lead to the maintenance of genetic variation, and thereby at least partly help to explain the high genetic variance in many fitness-related traits. On the whole, genotype–environment interactions may be a strong determinant of the fitness, and ultimately the survival, of a population (Merilä and Fry 1998).

Heritabilities may vary across environments due to differences in recombination and mutation rates between environments, different selection regimes, or genetic drift, for example. Also, the amount of environmental variation may vary across environments (Hoffmann and Merilä 1999). However, from our reaction norm approach, it becomes clear that none of these will lead to nonparallelism of reaction norms, and thus they are not the result of genotype–environment interactions (Fig. 8). Only if the variation in h^2 across environments is due to variation in the expression of additive genetic variance across environments, while keeping the genotypic composition equal, do genotype–environment interactions play a role.

If genotype–environment interactions lead to variation in the expression of genetic variance across environments, the amount of evolutionary change is related to the heritability under the conditions during selection (Gebhardt-Henrich and Van Noordwijk 1991). For example, Gebhardt-Henrich and Van Noordwijk (1994) showed that the amount of genetic variation for body mass was higher after growth under good conditions than under poor conditions, and Charmantier et al. (2004) found that the additive genetic variance for tarsus length was significantly higher in deparasitized broods than in parasitized control broods.

The shape of individual reaction norms

So far we have described a reaction norm by its average height (the genotypic value) and its slope (the amount of plasticity). In all of these cases, the reaction norm is considered to be linear. However, there also may be variation in the slope within reaction norms, leading to nonlinearity, and possibly to deviations from normality of the phenotypic distribution (Fig. 9). Here we only focus on the shape of the mean reaction norm,

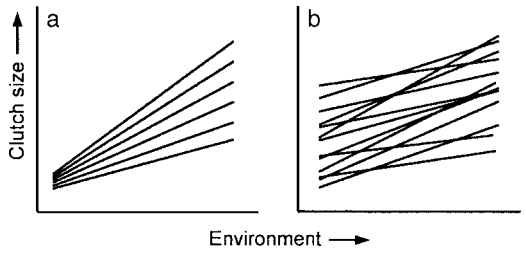

FIG. 7. Genotype–environment interactions are represented by nonparallelism of the reaction norms. In (a) the heritability varies across environments, but the genetic correlation is equal to 1, whereas in (b) the heritability is constant, but the genetic correlation across environments is less than 1.

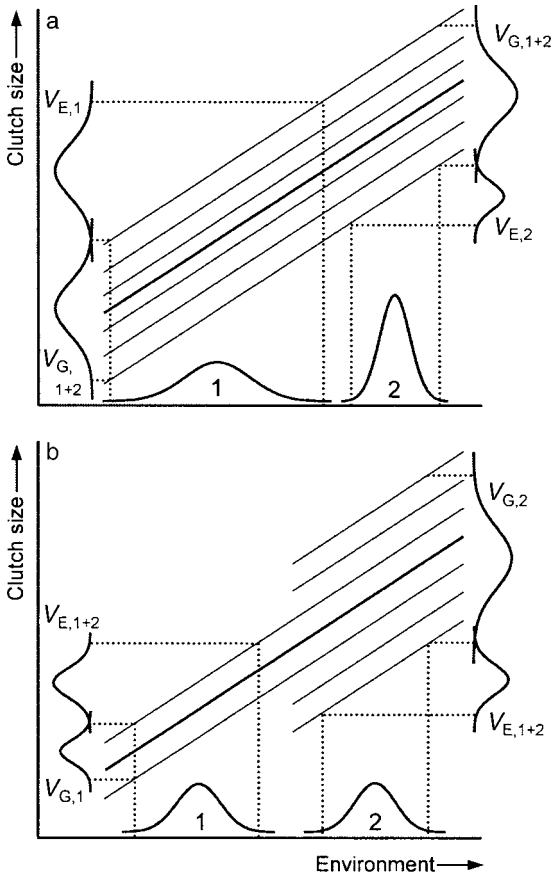

FIG. 8. Two scenarios for differences in heritability among environments in the absence of genotype–environment interactions. In (a) the amount of environmental variation experienced is larger in environment 1. $V_{E,1}$, and thus $V_{P,1}$, will be larger, leading to a lower heritability estimate in the more variable environment. In (b) we do have variation in the amount of V_G. However, this variation is due to strong selection or drift leading to a reduction of the amount of genetic variation in environment 1, or an increase of genetic variation in environment 2 due to higher mutation and/or recombination rates, for example.

and we ignore the variation around it. Although data often can be linearized by means of transformation to facilitate analysis, it is the deviation from linearity that may provide insight into the underlying physiological and developmental mechanisms, and that allows for a link between quantitative genetics and ecophysiology or developmental biology.

Curve a in Fig. 9 provides an example of a trait in which the responsiveness to the environment, and thus the slope of the reaction norm, increases along the environmental axis. Such a relationship, for example, may be found when a change in a spatial character such as egg volume is mediated through a change in a linear character such as egg width. Alternatively, a reaction norm may be S-shaped (Fig. 9, curve b), in which case there is a certain threshold value above which the phenotypic value rapidly increases and then remains relatively constant and/or linear again. Such S-shaped curves are typical for growth processes or biochemical reactions with saturation. However, we also may encounter the interesting and probably common situation in which the phenotype remains relatively constant around average environmental conditions and only responds to extreme environmental deviations, resulting from canalization (Fig. 9, curve c). In the presence of canalization, deviations from the preferred path of development, due to disturbances in either the internal or the external environment, are corrected by means of regulatory processes. This may be especially important in traits that show most variation arising during development and that remain constant in the adult stage (e.g., many morphological traits), and that are under strong stabilizing selection. Canalization was first described by Waddington (1953) and it has been reviewed by Scharloo (1991).

An attempt to visualize canalization and integrating developmental biology and quantitative genetics patterns, in general, was the use of so called gene–environmental factor/phenotype mapping functions, GEPM (Scharloo 1987). A GEPM relates the phenotypic distribution to the distribution of gene products and shows an obvious resemblance to a reaction norm as it is described here. In the absence of canalization, the GEPM (and the mean reaction norm) will be linear. When canalization occurs, the GEPM will be less steep around average amounts of gene product, and deviations in the amount of gene product consequently will be buffered. The strength of the buffering depends on the steepness of the slope of the GEPM around the average phenotype. At present, however, we lack in-

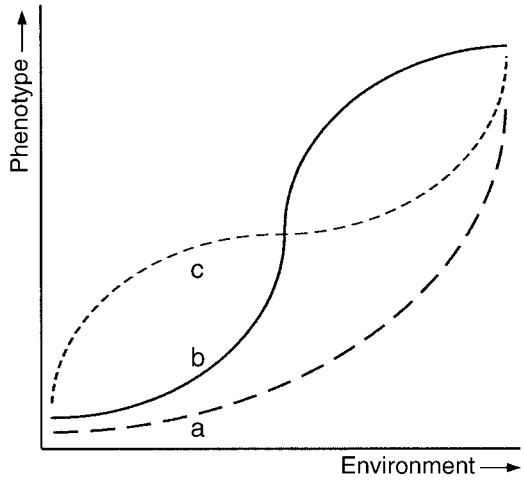

FIG. 9. Three possible scenarios leading to nonlinear reaction norms. In (a) the responsiveness to the environment increases along the environmental axis, whereas in the case of (b) there is a certain threshold value above which the phenotypic value rapidly increases and then levels off again. For (c) we have the opposite situation in which the phenotype remains relatively constant around average environmental conditions as the result of canalization.

TABLE 1. Overview of studies that have estimated heritabilities for clutch size in a natural population.

Species and location	Years	N	Mean (SD)	V_{within}	V_{among}	V_P	CV_P
Lesser Snow Goose							
La Pérouse Bay, Manitoba, Canada	1969–1988†	489	4.24 (1.15)	1.15	0.052	1.20	0.26
Collared Flycatcher							
Gotland, Sweden	1980–1999‡	NA	6.0 (0.78)	0.60	0.016	0.61	0.13
Great Tit							
Wytham Wood, Oxford, UK	1960–1998	NA	8.55 (1.69)	2.46	0.36	2.47	0.18
Hoge Veluwe, The Netherlands	1955–1979	336	9.60 (1.76)	2.62	0.47	3.09	0.18
Vlieland, The Netherlands	1958–1979	362	9.22 (1.36)	1.63	0.20	1.84	0.15
Darwin's Medium Ground Finch							
Isla Daphne Major, Galapagos	1976–1984	231	3.57 (0.93)	0.55	0.31	0.86	0.26
Starling							
Belmont, New Zealand	1973–1979	NA	4.45 (0.89)	0.79	0.010	0.80	0.20

Notes: Only estimates with $N > 200$ are included. If more than one estimate is available for the same population, the estimate based on the largest sample size is included. Column heads are defined as follows: Years, the years that were used for the heritability estimate; V_{within} and V_{among}, the amount of phenotypic variance within and among years, respectively; N, the number of data points in the mother–daughter regression; Mean, the mean clutch size across all years on which the heritability estimate was based. References are: 1, Lessells et al. (1989); 2, Sheldon et al. (2003); 3, McCleery et al. (2004); 4, Van Noordwijk et al. (1981); 5, Gibbs (1988); 6, Flux and Flux (1982). ''NA'' indicates ''not applicable.''

† Data on yearly means and variances were available for 1973–1987, derived from Cooch et al. (1989: Table 1, Fig. 1).

‡ Data on yearly means and variances were available from 1981 onward.

formation to construct such mapping functions for clutch size, as well as many other traits.

OVERVIEW OF STUDIES THAT ESTIMATED GENETIC VARIATION FOR CLUTCH SIZE

Repeatability estimates

Until 1970, the investigation of the genetic basis of clutch size in natural bird populations relied mainly upon the study of the intra- and interindividual variation in clutch size, and thus the estimation of repeatabilities (see Lessells and Boag 1987). For an overview of all repeatability estimates, see the Appendix. The pioneering work on this subject has been done by both Lack and Kluyver. Lack believed that clutch size at any time is determined by heredity and that environmental influences are only of minor significance. In their study on the breeding biology of the Common Swift (*Apus apus*), Lack and Lack (1951) found that clutch size was relatively constant within females (our calculations using their data: $r = 0.26 \pm 0.058$, mean $\pm$ SE; $F_{7,44} = 3.42$, $P = 0.005$). Interestingly, they also suggested that some individuals are more susceptible to environmental factors than others, and that this difference might have a hereditary basis. Kluyver (1951) found a repeatability of 0.28 for the clutch size of Great Tits (*Parus major*) (in Van Noordwijk et al. 1981). However, Kluyver considered this to be a low value and, in contrast to Lack, found it improbable that the individual constancy observed in the Great Tit had a genetic basis. He believed that this constancy was because each individual is permanently exposed to certain constant environmental influences that determine clutch size.

The main advantage of repeatabilities over heritabilities is that they require smaller sample sizes and do not require a pedigree. They can thus widen the range of populations and species in which the genetic basis of clutch size can be investigated. They also make it possible to compare different age classes and other subsets of a population, for example (e.g., Van Noordwijk et al. 1981, Kennedy and White 1991, Hochachka 1993). It would be reasonable not to go through the effort of estimating heritabilities directly, in the absence of a significant repeatability (but see Dohm 2002). Thus, focusing solely on estimated heritabilities may lead to incorrect conclusions about the heritability of clutch size and its generality.

Heritability estimates

Even though it had by then been shown many times that clutch size in natural populations is relatively constant within females, Prince and coworkers (1970) were the first to attempt to estimate a heritability for clutch size in a population of captive Mallards (*Anas platyrhynchos*) that originated from wild or semiwild populations. Perrins and Jones (1974) published the first heritability estimate made directly in a natural population. Since then, a total of 17 publications have reported heritability estimates for clutch size estimated in natural populations (see the Appendix). Some, however, contain heritability estimates for more than one population, different subsets of the same population, or make use of different methods (e.g., Van Noordwijk et al. 1981, Gibbs 1988).

Six studies find only nonsignificant estimates, but these are based on sample sizes that are too small to

TABLE 1. Extended.

V_{among}/V_P	h^2	V_A	V_E	CV_A	CV_E	Ref.
0.043	0.17	0.20	1.00	0.11	0.24	1
0.026	0.29	0.18	0.44	0.070	0.11	2
0.13	0.33	0.80	1.67	0.10	0.15	3
0.15	0.29	0.90	2.20	0.099	0.15	4
0.11	0.50	0.92	0.92	0.10	0.10	4
0.36	0	0	0.86	0	0.26	5
0.012	0.33	0.26	0.54	0.11	0.17	6

detect a heritability smaller than about 0.4 (which requires a standard error < 0.2), but they are all greater than zero. These estimates are included in the Appendix for completeness, but in the discussion, we will focus on the 11 remaining studies. Of the 11 studies that had sufficient statistical power, 10 found significant heritabilities for clutch size, ranging from 0.21 to 0.48. Only Gibbs (1988) found very small, nonsignificant repeatabilities and heritabilities for first clutch sizes in Darwin's Medium Ground Finches (*Geospiza fortis*). Except for Van Noordwijk et al. (1981), Lessells et al. (1989), and the more recent studies mentioned, all of these studies use fewer than 10 years of data. However, sample sizes were generally on the order of several hundreds. These sample sizes are large enough to have the power to detect heritabilities of about 0.2 and higher, but they do not allow for quantitative comparisons within and between studies. The Appendix shows that this may be possible only with the heritability estimates based either on an animal model analysis using restricted maximum likelihood techniques (REML), or on the response to selection, given their superiority in terms of the precision of the estimate (as measured by the size of their standard error).

Although most of these studies were carried out between 1980 and 1990, currently there is renewed interest in the genetic basis of clutch size. This is partly due to the large amount of data currently available in several long-term studies, and the implementation of new methods (Merilä and Sheldon 2000, Van der Jeugd and McCleery 2002, Sheldon et al. 2003, McCleery et al. 2004). However, 11 relatively precise studies is still a surprisingly limited number for a trait that has been under such scrutiny in other fields of biology. The number of species (five) and populations (seven) involved is even smaller, with almost three- fourths of the studies being performed on populations of Great Tits (*Parus major*) (five) and Collared Flycatchers (*Ficedula albicollis*) (three).

Most estimated heritabilities are based on the amount of resemblance between mother and daughter. Van der Jeugd and McCleery (2002) probably used the most rigid form of mother–daughter regression in their study on the Great Tit population in Wytham Woods. They used lifetime mean clutch size of both mother and daughter, with mean daughter clutch sizes averaged across all daughters. Using this approach, every mother is included only once, avoiding the effects of pseudoreplication. However, this approach requires very large sample sizes, which were not available at the time most studies were carried out. As a consequence, most studies employed some sort of compromise, for example, comparing clutch size of the mother and the mean clutch size of individual daughters that were recruited from that clutch. Consequently, mothers that produced recruits from multiple clutches are included more than once. Although new, more powerful techniques may allow for the estimation of heritabilities in smaller populations of a wide range of species, the accurate estimation of variance components and the investigation of genotype–environment correlations and interactions, for example, still require very large data sets. Therefore we do not expect the bias toward the few long-term studies on nest box populations of tits and flycatchers to disappear soon.

DEALING WITH AND QUANTIFYING ENVIRONMENTAL STRUCTURE

Variation among years

One of the key features in a reaction norm approach is that it uses an axis to portray the environmental effects on the phenotype. Which aspects of the environment are most informative depends on what we know about the ecophysiology and on what information is available. A posteriori, we have only limited information about the environmental conditions under which the data for the published heritability estimates were collected. We can, however, use the fact that for several important aspects of the environment (e.g., weather, population density, food availability), the among-year variation may be greater than the within-year variation. In order to quantify at least one aspect of the structure of environmental variation represented in the published heritability estimates for clutch size in natural populations, we have retrieved data on the between- and within-year phenotypic variance (Table 1).

These data provide us with interesting insights. All Great Tit populations show average amounts of variation among years (10–15%), and mean clutch sizes of both Snow Goose and the Collared Flycatcher populations show relatively little variation between years (3–5%). Clutch size in the Starling population studied by Flux and Flux (1982) showed remarkably little variation between years, with only 1.2% of the total phenotypic variation being present among years. At the

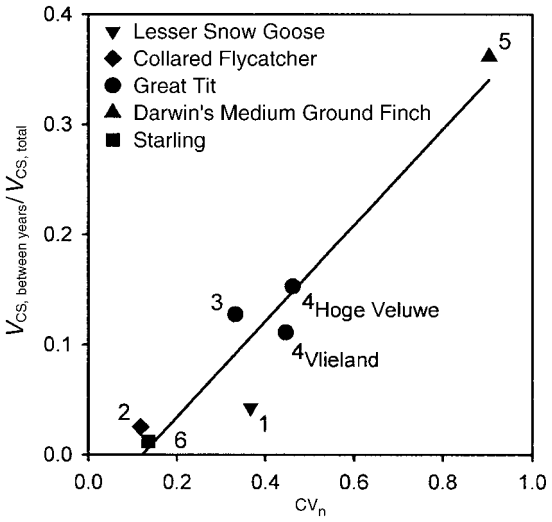

FIG. 10. The relationship between the coefficient of variation for population density across years (CV_n), and the proportion of the total phenotypic variance for clutch size that is present between years ($V_{CS,yr}/V_{CS,tot}$) ($R^2 = 0.92$, $P < 0.001$). Number labels refer to references in Table 1.

other extreme, there is the population of Darwin's Medium Ground Finches in which not less than 36% of the total phenotypic variation is present among years. Although the environment is relatively constant within years, it is extremely variable across years. This implies automatically that mothers and daughters are measured in quite different environments. In principle, one could compare those mothers still breeding in the same years as their daughters, but this would be a small data set potentially biased by differential survival among mothers and by age effects on clutch size.

Although we have pointed out that it may be difficult to identify the relevant biological variable responsible for variation in clutch size among years, an easily obtainable proxy is the amount of variation in population size among years. This is illustrated in Fig. 10, where we see a clear relationship between the coefficient of variation for population size and the proportion of the total phenotypic variation that is present among years.

As noted earlier, it is likely that our assumption of the linearity of reaction norms is no longer valid across a very large environmental range. This violation provides a likely explanation for the small and nonsignificant heritabilities found in the population of Darwin's Medium Ground Finches. Apart from the technical aspect of violating the assumption of reaction norm linearity more seriously when the among-year component of V_P is greater, there is also the biological aspect that there may exist stronger selection for plasticity in fluctuating and unpredictable environments. Across the (limited number of) studies in Table 1, there is, however, no evidence for a negative correlation between heritability and the amount of between-year phenotypic variance (including Gibbs (1988), Pearson $r = -0.60$, $P = 0.16$; excluding Gibbs (1988), Pearson $r = 0.34$, $P = 0.50$).

Environmental structure within years

Heritability estimates are likely to be biased because females may stay in the same area throughout their life. Such biases are often referred to as permanent or general environment effects V_{Eg} (and opposed to special environment effects V_{Es}). Comparing repeatability and heritability can give us an estimate of the importance of permanent environment effects (but see Dohm 2002). Although only a limited number of studies estimate both heritability and repeatability for clutch size, in all these studies, both estimates are remarkably similar. From this similarity, Van Noordwijk et al. (1981) concluded in their investigation of clutch size in several Dutch populations of Great Tit that the effect of common environments is probably of minor importance. The largest difference was found by Van der Jeugd and McCleery (2002). They corrected their heritability estimate based on a mother–daughter regression for the resemblance between the environments of mothers and daughters. Such a resemblance is the result of females staying in the area where they were born, and results in spatial autocorrelation of clutch sizes. After correction, they found a heritability of 0.34, whereas the repeatability was 0.45. Interestingly, the estimated heritability before correction was indeed very similar to the observed repeatability. Furthermore, the heritability estimate from the animal model analysis in McCleery et al. (2004), using the same data, is almost equal to the corrected estimate in Van der Jeugd and McCleery, suggesting that heritability estimates from an animal model approach are less biased by spatial autocorrelations.

Several authors have estimated heritabilities after correcting for environmental variables such as habitat quality or area. For example, Perrins and Jones (1974) performed corrections for age, area, year, and habitat quality, by adding the mean differences for these factors. The use of additive corrections for environmental factors, however, may introduce more problems than are resolved, for three reasons (Van Noordwijk et al. 1981). First, it is essential that the same corrections are used in measurements of selection. Second, by making corrections, we assume that genotypes are randomized over environments. Furthermore, we impose a model of how the various components of the correction interact. If this model is wrong, we create a structure that can confound the patterns for which we are searching. Third, if we had complete knowledge of all environmental effects and could successfully correct phenotypic values for these factors, we would end up with a meaningless heritability of one. Note that if we have a measure of, e.g., habitat quality that explains part of the variation in clutch size, we can directly incorporate it in a reaction norm approach instead of simply removing it.

Because natural selection mainly acts upon variation within years, it is reasonable to remove the variation between years. Not removing variation among years can lead to an underestimate of the response to selection. Whereas most heritabilities for clutch sizes have indeed been estimated using clutch size corrected for between-year variation, this usually is not the case for repeatability estimates. This makes the comparison of repeatability and heritability estimates problematic.

Comparing Levels of Genetic Variation

When we want to compare the amount of phenotypic variation or additive genetic variation for the same trait between species or for different traits within species, we have to use a standardized value such as the narrow-sense heritability or the coefficient of additive genetic variation. However, if we are comparing the same trait between populations of the same species, it is most informative to compare variance components directly.

Although it would be very interesting to test for ecological or phylogenetic predictors of heritabilities, the number of accurate estimates for different species and populations is currently too small. The only species for which we have a reasonable comparison is the Great Tit, for which heritabilities have been estimated in three different populations (Table 1). If we would limit ourselves to comparing heritabilities, we would conclude that the Vlieland population ($h^2 \approx 0.5$) has more additive genetic variation for clutch size than both the Hoge Veluwe ($h^2 \approx 0.3$) and the Wytham population ($h^2 = 0.33$). However, the amount of phenotypic variation in the Vlieland population is only ~60% of that observed in the two other populations. If we calculate V_A from $h^2 \times V_P$, we find that, in fact, V_A in all three populations is remarkably similar. The large difference in heritability, in this case, can be fully explained by V_E being much larger in the Hoge Veluwe and the Wytham population. Notice that the CV_A gives a much better indication of V_A in this case, because mean clutch sizes are very similar in all three populations.

Conclusions

Although there may be a bias toward publishing large and/or significant heritabilities, and the estimates are based on only a limited number of populations and species, there is good evidence for the presence of genetic variation for clutch size, with narrow-sense heritabilities on the order of 0.2–0.4 and coefficients of additive genetic variance of ~0.1. These observed amounts of genetic variation for clutch size imply that clutch size could change rapidly due to selection. However, genotype–environment interactions may be common in natural populations. Because they are key elements in the evolutionary potential of populations, their quantification, as well the elucidation of the underlying mechanisms, are among the most important subjects that need further study.

In order to see the general patterns in genetic aspects of phenotypic plasticity, we require studies over a wide range of species and environments. For such a comparison, it is essential that studies of heritability report much more information, particularly with regard to the environmental components. Mean phenotypes and sample sizes, within- and among-year phenotypic variances, as well as any patterns in phenotypic variances need to be reported. Any environmental factor that can partly explain V_P should be quantified.

Theoretically, the reaction norm concept allows for an integration of patterns in the environment with genetic variation. It allows us to integrate genetics, ecophysiology, and environmental variation to arrive at realistic estimates of evolutionary potential. Unfortunately, the data required for a practical application are as yet scarce. Studies of the mechanism underlying the determination of clutch size and other traits, extensive descriptive work of relevant ecological variables, and the application of advanced statistical methods could soon change this. This will depend on the ability to choose the most informative environmental axes. In the absence of data on food availability, weather conditions, or other direct measurements of environmental variables, indirect measurements such as population density or the breeding success of competitors might be used.

The formulation of the appropriate environmental axis requires two additional research lines: first, a better understanding of the ecophysiological processes underlying trait formation; and second, a major effort in data collection to quantify the environment at the level of the individual. Thus, the additional information needed should come from ecologists.

Acknowledgments

We are grateful to Evan Cooch, Lars Gustafsson, and Robin McCleery for providing the additional data on yearly means and variances for clutch size, and to anonymous reviewers for their comments on previous versions of the manuscript. E. Postma is financed by ALW-NWO, project 811.34.004.

Literature Cited

Andersson, M. S., and L. Gustafsson. 1995. Glycosylated hemoglobin: a new measure of condition in birds. Proceedings of the Royal Society of London Series B, Biological Sciences **260**:299–303.

Barton, N., and L. Partridge. 2000. Limits to natural selection. Bioessays **22**:1075–1084.

Bennett, P. M., and I. P. F. Owens. 2002. Evolutionary ecology of birds: life histories, mating systems and extinction. Oxford University Press, New York, New York, USA.

Boag, P. T., and A. J. Van Noordwijk. 1987. Quantitative genetics. Pages 45–78 *in* F. Cooke and P. A. Buckley, editors. Avian genetics: a population and ecological approach. Academic Press, London, UK.

Bradshaw, A. D. 1991. The Croonian Lecture, 1991: genostasis and the limits to evolution. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences **333**:289–305.

Charmantier, A., L. E. B. Kruuk, and M. M. Lambrechts. 2004. Parasitism reduces the potential for evolution in a wild bird population. Evolution **58**:203–206.

Conover, D. O., and E. T. Schultz. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. Trends in Ecology and Evolution **10**:248–252.
Cooch, E. G., D. B. Lank, R. F. Rockwell, and F. Cooke. 1989. Long-term decline in fecundity in a snow goose population: evidence for density dependence. Journal of Animal Ecology **58**:711–726.
Cooke, F., P. D. Taylor, C. M. Francis, and R. F. Rockwell. 1990. Directional selection and clutch size in birds. American Naturalist **136**:261–267.
Crnokrak, P., and D. A. Roff. 1995. Dominance variance: associations with selection and fitness. Heredity **75**:530–540.
De Jong, G. 1990. Quantitative genetics of reaction norms. Journal of Evolutionary Biology **3**:447–468.
DeWitt, T. J., A. Sih, and D. S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends in Ecology and Evolution **13**:77–81.
Dohm, M. R. 2002. Repeatability estimates do not always set an upper limit to heritability. Functional Ecology **16**: 273–280.
Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Fourth edition. Longman, Harlow, UK.
Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, UK.
Flux, J. E. C., and M. M. Flux. 1982. Artificial selection and gene flow in wild starlings, *Sturnus vulgaris*. Naturwissenschaften **69**:96–97.
Frank, S. A., and M. Slatkin. 1992. Fisher's fundamental theorem of natural selection. Trends in Ecology and Evolution **7**:92–95.
Frankel, O. H., and M. E. Soulé. 1981. Conservation and evolution. Cambridge University Press, Cambridge, UK.
Frankham, R., J. D. Ballou, and D. A. Briscoe. 2002. Introduction to conservation genetics. Cambridge University Press, Cambridge, UK.
Garant, D., L. E. B. Kruuk, T. A. Wilkin, R. H. McCleery, and B. C. Sheldon. 2005. Evolution driven by differential dispersal within a wild bird population. Nature **433**:60–65.
Gebhardt-Henrich, S. G., and A. J. Van Noordwijk. 1991. Nestling growth in the great tit 1. Heritability estimates under different environmental conditions. Journal of Evolutionary Biology **4**:341–362.
Gebhardt-Henrich, S. G., and A. J. Van Noordwijk. 1994. The genetic ecology of nestling growth in the Great Tit: environmental influences on the expression of genetic variances during growth. Functional Ecology **8**:469–476.
Gibbs, H. L. 1988. Heritability and selection on clutch size in Darwin's medium ground finches (*Geospiza fortis*). Evolution **42**:750–762.
Gillespie, J. H., and M. Turelli. 1989. Genotype–environment interactions and the maintenance of polygenic variation. Genetics **121**:129–138.
Grant, P. R. 2001. Reconstructing the evolution of birds on islands: 100 years of research. Oikos **92**:385–403.
Grant, P. R., B. R. Grant, L. F. Keller, and K. Petren. 2000. Effects of El Niño events on Darwin's finch productivity. Ecology **81**:2442–2457.
Gustafsson, L. 1986. Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. American Naturalist **128**:761–764.
Gustafsson, L., and J. Merilä. 1994. Foster parent experiment reveals no genotype–environment correlation in the external morphology of *Ficedula albicollis*, the collared flycatcher. Heredity **73**:124–129.
Hailman, J. P. 1986. The heritability concept applied to wild birds. Current Ornithology **4**:71–95.
Haukioja, E. 1970. Clutch size of the reed bunting *Emberiza schoeniclus*. Ornis Fennica **47**:101–135.
Haywood, S., and C. M. Perrins. 1992. Is clutch size in birds affected by environmental conditions during growth? Proceedings of the Royal Society of London Series B, Biological Sciences **249**:195–197.
Hedrick, P. W. 2000. Genetics of populations. Second edition. Jones and Barlett, Sudbury, Massachusetts, USA.
Hochachka, W. M. 1993. Repeatable reproduction in song sparrows. Auk **110**:603–613.
Hoffmann, A. A., and J. Merilä. 1999. Heritable variation and evolution under favourable and unfavourable conditions. Trends in Ecology and Evolution **14**:96–101.
Högstedt, G. 1980. Evolution of clutch size in birds: adaptive variation in relation to territory quality. Science **210**:1148–1149.
Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics **130**:195–204.
Kennedy, E. D., and D. W. White. 1991. Repeatability of clutch size in House Wrens. Wilson Bulletin **103**:552–558.
Klomp, H. 1970. The determination of clutch-size in birds; a review. Ardea **58**:1–124.
Kluyver, H. N. 1951. The population ecology of the great tit, *Parus m. major* L. Ardea **39**:1–135.
Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the ''animal model.'' Philosophical Transactions of the Royal Society of London Series B, Biological Sciences **359**:873–890.
Lack, D. 1948. Natural selection and family size in the starling. Evolution **2**:95–110.
Lack, D., and E. Lack. 1951. The breeding biology of the swift *Apus apus*. Ibis **93**:501–546.
Lande, R., and G. F. Barrowclough. 1987. Effective population size, genetic variation, and their use in population management. Pages 87–123 *in* M. E. Soulé, editor. Viable populations for conservation. Cambridge University Press, Cambridge, UK.
Lande, R., and S. Shannon. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution **50**:434–437.
Larsson, K., H. P. Van der Jeugd, I. T. Van der Veen, and P. Forslund. 1998. Body size declines despite positive directional selection on heritable size traits in a barnacle goose population. Evolution **52**:1169–1184.
Lessells, C. M., and P. T. Boag. 1987. Unrepeatable repeatabilities: a common mistake. Auk **104**:116–121.
Lessells, C. M., F. Cooke, and R. F. Rockwell. 1989. Is there a trade-off between egg weight and clutch size in wild lesser snow geese (*Anser c. caerulescens*)? Journal of Evolutionary Biology **2**:457–472.
Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, Massachusetts, USA.
Martin, T. E., A. P. Møller, S. Merino, and J. Clobert. 2001. Does clutch size evolve in response to parasites and immunocompetence? Proceedings of the National Academy of Sciences (USA) **98**:2071–2076.
Mather, K., and J. L. Jinks. 1982. Biometrical genetics. Third edition. Chapman and Hall, London, UK.
McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins. 2004. Components of variance underlying fitness in a natural population of the great tit *Parus major*. American Naturalist **164**:E1–E11.
Merilä, J., and J. D. Fry. 1998. Genetic variation and causes of genotype–environment interaction in the body size of blue tit (*Parus caeruleus*). Genetics **148**:1233–1244.
Merilä, J., L. E. B. Kruuk, and B. C. Sheldon. 2001*a*. Cryptic evolution in a wild bird population. Nature **412**:76–79.
Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. American Naturalist **155**: 301–310.

Merilä, J., and B. C. Sheldon. 2001. Avian quantitative genetics. Current Ornithology **16**:179–255.
Merilä, J., B. C. Sheldon, and L. E. B. Kruuk. 2001*b*. Explaining stasis: microevolutionary studies in natural populations. Genetica **112**:199–222.
Murphy, E. C., and E. Haukioja. 1986. Clutch size in nidicolous birds. Current Ornithology **4**:141–180.
Perrins, C. M. 1965. Population fluctuations and clutch size in the great tit. Journal of Animal Ecology **34**:601–647.
Perrins, C. M., and P. J. Jones. 1974. The inheritance of clutch size in the Great Tit (*Parus major*). Condor **76**:225–228.
Perrins, C. M., and D. Moss. 1975. Reproductive rates in the great tit. Journal of Animal Ecology **44**:695–706.
Pettifor, R. A., C. M. Perrins, and R. H. McCleery. 1988. Individual optimization of clutch size in great tits. Nature **336**:160–162.
Pettifor, R. A., C. M. Perrins, and R. H. McCleery. 2001. The individual optimization of fitness: variation in reproductive output, including clutch size, mean nestling mass and offspring recruitment, in manipulated broods of great tits *Parus major*. Journal of Animal Ecology **70**:62–79.
Postma, E., and A. J. Van Noordwijk. 2005. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. Nature **433**:65–68.
Price, T. D., and L. Liou. 1989. Selection on clutch size in birds. American Naturalist **134**:950–959.
Prince, H. H., P. B. Siegel, and G. W. Cornwell. 1970. Inheritance of egg production and juvenile growth in mallards. Auk **87**:342–352.
Scharloo, W. 1987. Constraints in selection response. Pages 125–149 *in* V. Loeschcke, editor. Genetic constraints on adaptive evolution. Springer-Verlag, Berlin, Germany.
Scharloo, W. 1991. Canalization: genetic and developmental aspects. Annual Review of Ecology and Systematics **22**: 65–93.
Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. Annual Review of Ecology and Systematics **24**: 35–68.
Schluter, D., and L. Gustafsson. 1993. Maternal inheritance of condition and clutch size in the collared flycatcher. Evolution **47**:658–667.
Sheldon, B. C., L. E. B. Kruuk, and J. Merilä. 2003. Natural selection and the inheritance of breeding time and clutch size in the collared flycatcher. Evolution **57**:406–420.
Stearns, S. C., and J. Koella. 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.
Sultan, S. E., and H. G. Spencer. 2002. Metapopulation structure favors plasticity over local adaptation. American Naturalist **160**:271–283.
Van der Jeugd, H. P., and R. H. McCleery. 2002. Effects of spatial autocorrelation, natal philopatry and phenotypic plasticity on the heritability of laying date. Journal of Evolutionary Biology **15**:380–387.
Van Noordwijk, A. J. 1989. Reaction norms in genetical ecology. BioScience **39**:453–458.
Van Noordwijk, A. J. 1990. The methods of genetical ecology applied to the study of evolutionary change. Pages 291–319 *in* K. Wöhrmann and S. K. Jain, editors. Population biology: ecological and evolutionary viewpoints. Springer-Verlag, Berlin, Germany.
Van Noordwijk, A. J., J. H. Van Balen, and W. Scharloo. 1981. Genetic and environmental variation in clutch size of the great tit (*Parus major*). Netherlands Journal of Zoology **31**:342–372.
Verboven, N., J. M. Tinbergen, and S. Verhulst. 2001. Food, reproductive success and multiple breeding in the Great Tit *Parus major*. Ardea **89**:387–406.
Via, S., and R. Lande. 1985. Genotype–environment interaction and the evolution of phenotypic plasticity. Evolution **39**:505–522.
Visser, M. E., A. J. van Noordwijk, J. M. Tinbergen, and C. M. Lessells. 1998. Warmer springs lead to mistimed reproduction in great tits (*Parus major*). Proceedings of the Royal Society of London Series B, Biological Sciences **265**:1867–1870.
Waddington, C. H. 1953. Genetic assimilation of an acquired character. Evolution **7**:118–126.
Woltereck, R. 1909. Weitere experimentelle Untersuchungen über Art-veränderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. Verhandlungen der Deutschen Tschechieschen Zoologischen Gesellschaft **1909**:110–172.

APPENDIX

An overview of all studies that estimated repeatabilities and heritabilities for clutch size in natural populations of birds is available in ESA's Electronic Data Archive: *Ecological Archives* E086-123-A1.