INTRODUCTION
Some traits of interest to breeders can be quantified by a single measurement, for example the number of offspring in an animal’s first litter. Many other traits, however, can only be described by a curve or function that relates an individual’s age to its phenotype; body weight and lactation rate are two obvious examples. Because they are quantified by an infinite number of measurements, that is, a function, these kinds of traits are called “infinite-dimensional” or “function-valued” (Kirkpatrick and Heckman, 1989; Pletcher and Geyer, 1999; Kingsolver et al., 2001). This paper briefly outlines a framework for quantifying and selecting on function-valued traits that has been developed over the last 15 years.

THE FUNCTION-VALUED APPROACH
How can we describe individuals and populations for function-valued traits? We might denote the phenotypic value for a simple univariate trait as $z$. For a vector-valued record, say a set of several metric measurements, an individual’s phenotype can be described by a vector $z = (z_1, z_2, ...)^T$. With an age-dependent trait, the analogous quantity is the function $z(a)$, which is just the individual’s size at each age $a$. The mean growth trajectory, which simply gives the average size of individuals at each age, is written $\bar{z}(a)$. Throughout this paper, italics is used for scalar values, bold for vectors and matrices, and script for function-valued quantities; the abbreviation “FV” is used as shorthand for “function-valued”.

How does the function-valued approach quantify variability in a population? Variation for a univariate trait is quantified by the phenotypic variance, $P$, and additive genetic variance, $G$. When multiple traits are under selection, standard quantitative genetics uses the phenotypic covariance matrix $P$ whose elements $P_{ij}$ give the covariance between traits $i$ and $j$. The FV approach uses the phenotypic covariance function, defined so that $P(a_1, a_2)$ is the covariance between an individual’s size at age $a_1$ and its size at age $a_2$. Similarly, $G(a_1, a_2)$ is the additive genetic covariance function whose value gives the additive genetic covariance between sizes at ages $a_1$ and $a_2$.

I am intentionally emphasizing the similarity between the function-valued framework and standard multivariate approaches. The similarity extends beyond the descriptive notation to results such as the breeder’s equation for response to selection. Given these similarities, a reasonable person will ask if there is any advantage to using the FV approach rather than standard multivariate methods. In fact, FV methods offer several benefits over the two standard alternatives that are in widespread use (Kirkpatrick and Heckman, 1989).
One standard approach is to focus attention on measurements at a set of several landmark ages; these values can then be treated as separate characters using standard multivariate methods. Perhaps the most important advantage that the FV approach offers here is that it makes use of the information about the ages at which the measurements are taken, data that is otherwise discarded. By using information about age, FV methods can smooth the data in a way that reduces sampling error and causes the estimates to converge more rapidly on the true parameter values. Comparison of FV and conventional analyses using simulated data sets shows that the function-valued approach is indeed more efficient with data and so can make more accurate predictions about the response to selection (Kirkpatrick and Heckman, 1989; Kirkpatrick et al., 1994). Other advantages are that the FV framework can accommodate measurements taken at different ages on different individuals, and it gives predictions for selection response at all ages rather than at just a finite set of landmark ages.

The function-valued approach also offers advantages over another conventional method that is commonly used. Growth trajectories, lactation curves, and other FV traits are often described using some parametric function chosen by the investigator (for example, the logistic function). The parameters of these functions are then analyzed with standard multivariate methods. An argument in favor of using the FV approach is that it makes no a priori (and perhaps biologically arbitrary) constraint on the form of the growth trajectory. A second advantage of FV methods is that they make explicit and consistent assumptions about phenotypic and genetic variation, as we will see below. In contrast, it is not clear what phenotypic distribution results from assuming, for example, that the genetic parameters of a logistic function are normally distributed.

**SELECTION RESPONSE AND SELECTION INDICES**

Any method for predicting selection response requires assumptions about inheritance. Standard quantitative genetics assumes that phenotypes and breeding values are normally distributed. This then leads to the celebrated breeder’s equation for selection response:

$$\Delta \bar{z} = G \left( \frac{1}{P} \right) \mathbf{s},$$

where $\Delta \bar{z}$ is the change in the trait mean across generations and $\mathbf{s}$ is the selection differential (Falconer and Mackay, 1996). When several traits are under selection, the selection response of trait $i$ is

$$\Delta \bar{z}_i = \sum_j \sum_k G_{ij} P_{jk}^{-1} \mathbf{s}_k,$$

where now $G$ and $P$ are the additive genetic and phenotypic covariance matrices, and $\mathbf{s}$ is a vector of selection differentials.

The FV approach extends standard quantitative genetics in a natural way by assuming that phenotypic and additive genetic variation is normally distributed. This implies (among other
things) that the phenotypes and breeding values at each age are normally distributed. The selection response of the mean size at age $a$ is then

$$\Delta z(a) = \int \int G(a,x)P^{-1}(x,y)s(y)dydx,$$

(3)

where $P^{-1}()$ is the “generalized inverse” of $P()$, $s(a)$ is the selection differential at age $a$, and the integrals range over the interval of ages during which selection is applied (see Kirkpatrick and Heckman, 1989; Gomulkiewicz and Beder, 1996; Beder and Gomulkiewicz, 1998). The parallels between the univariate, multivariate, and function-valued versions of the selection response shown in Equations (1 – 3) are obvious.

We can calculate the selection response for a function-valued trait using Equation (3) by numerical integration if we have estimates for $G()$ and $P()$ and if we know the difference between the mean size of selected and unselected individuals at all ages. (The question of how these quantities can be estimated is briefly discussed below.) Equation (3) has been used to study variation, selection, and evolution of age-dependent traits such as growth trajectories (Kirkpatrick and Lofsvold, 1989; Kirkpatrick and Lofsvold, 1992; Kirkpatrick, 1993; Meyer 2000), condition (Jones et al., 1999), and lactation curves (Kirkpatrick et al., 1994; White et al., 1999).

Selection on age-dependent traits is sometimes practiced on the records of individuals taken at a single culling age. Denoting that age as $a_c$, Equation (3) can be used to show that the selection response at any other age $a$ is

$$\Delta z(a) = G(a,a_c)\left[ \frac{s(a_c)}{P(a_c,a_c)} \right].$$

(4)

The quantity inside the square brackets is sometimes referred to as the selection gradient acting at age $a_c$, denoted $\beta(a_c)$ (Kirkpatrick and Heckman, 1989; Gomulkiewicz and Beder, 1996; Beder and Gomulkiewicz, 1998). Multiplying this number by the phenotypic standard deviation at that age gives the intensity of selection, written $\delta(a)$, which is a nondimensional number often used by breeders to quantify the strength of selection on a trait (Falconer and Mackay, 1996 chapter 11).

A powerful use of the function-valued approach is to visualize patterns of genetic variation. This is done by calculating the eigenfunctions and eigenvalues from the genetic covariance function. This analysis allows us to identify the changes in a growth trajectory that can be easily achieved by selection, and those that are not possible (see Kirkpatrick and Lofsvold, 1992; Kirkpatrick et al., 1994; Kingsolver et al., 2001).

**Selection indices for age-dependent traits.** A standard problem in quantitative genetics is how to best combine measurements from several traits into a single selection index so as to maximize the rate of economic gain (Falconer and Mackay, 1996 chapter 19). A variant on this
problem appears with age-dependent traits. We might, for example, want to maximize final carcass weight based on body weight measured at two or more earlier ages. How should the measurement ages be chosen and how should those data be used?

Given a vector \( z \) of \( m \) measurements on each individual, the optimal selection index is \( I = b^T z \), where \( b \) is a vector of weights given by

\[
b = P^{-1} g .
\]

Here \( g \) is the vector of additive genetic covariances between the trait of economic interest (e.g. final carcass weight) and the \( m \) measured traits, and \( P \) is the \( m \times m \) phenotypic covariance matrix for the measured traits (see Falconer and Mackay, 1996 Eq. 19.13). Using these weights, the rate of economic gain is proportional to

\[
R = g^T P^{-1} g .
\]

To find the optimal vector of ages \( a = (a_1, a_2, \ldots, a_m) \) at which the measurements should be taken, we need to maximize \( R \). A numerical algorithm for doing that proceeds in three steps. First, obtain estimates of the phenotypic covariance function \( P() \) and the function \( g(a) \) whose value is the additive genetic covariance between the measured trait (e.g. body weight) at age \( a \) and the trait of interest (e.g. carcass weight). Second, choose a trial vector of ages \( a_{(1)} \) and evaluate \( P() \) and \( g() \) at these ages to form the \( n \times n \) phenotypic covariance matrix \( P_{(1)} \) and the \( n \times 1 \) genetic covariance vector \( g_{(1)} \). Third, calculate \( R_{(1)} \) from them using Equation (6). These three steps are repeated with successive trial vectors \( a_{(2)}, a_{(3)}, \ldots \) using an appropriate numerical search procedure until an optimal age vector \( a_{opt} \) that maximizes \( R \) is found. From this vector, corresponding values for \( P_{opt} \) and \( g_{opt} \) are used in Equation (5) to calculate the optimal selection index \( b_{opt} \).

**Selection on multiple age-dependent traits.** Breeders are sometimes interested in improving several age-dependent traits simultaneously. The FV approach for a single trait is easily adapted to this situation because, under the standard assumptions of quantitative genetics, the effects of selection on several traits are additive. For example, Equation (3) for the selection response in the mean of age-dependent trait \( i \) at age \( a \) becomes

\[
\Delta \bar{z}_i(a) = \sum_j \int \frac{G_j(a,x)P_j^{-1}(x,y)\gamma_j(y)dy}{\int P_j(x,x)} dx,
\]

\[
= \sum_j \int G_j(a,x) \left[ \frac{\nu_j(x)}{\sqrt{P_j(x,x)}} \right] dx ,
\]

where \( G_j(), P_j(), \gamma_j(), \) and \( \nu_j() \) are the genetic covariance, phenotypic covariance, selection differential, and selection intensity functions for trait \( j \) (Kirkpatrick, 1993). Other results, such
as optimal selection indices, can be generalized in a similar way when multiple age-dependent traits are selected.

ESTIMATION
The function-value approach involves quantities like the population’s mean function, phenotypic covariance function, and additive genetic covariance function. How can these be estimated? Estimation of FV quantities is a rapidly developing field that I have only the space here to discuss briefly.

Estimates of variances, particularly genetic variances, often include large sampling errors. The FV framework takes advantage of the ordering of the ages at which measurements were taken to reduce these errors by smoothing the data. To smooth, we must choose a mathematical representation of the covariance function. Two families of methods have been proposed for this task. The first is nonparametric in that it makes no assumption about the form of the covariance function. This approach begins with a mathematical representation (such as a polynomial or spline) that is flexible enough to represent any form of the covariance function. That function is then fit to the data (Kirkpatrick and Heckman, 1989; Kirkpatrick et al., 1990; Kirkpatrick et al., 1994; see also Rice and Silverman, 1991; Green and Silverman, 1994)

The second family of methods to estimate the genetic covariance function and other FV quantities takes a parametric approach. It begins by assuming that the covariance function of interest can be described by a function that has a small number of parameters (Pletcher and Geyer, 1999; Jaffrézic and Pletcher, 2000). A closely related approach are methods based on “random regression” (Schaeffer and Dekkers, 1994; Meyer, 1998; Jones et al., 1999; Meyer, 2000). The parameters are estimated by fitting the function to the genetic data.

These two families of methods have both advantages and disadvantages. A strength of the parametric approach is that only a small number of parameters need to be estimated. A disadvantage is that by making an assumption about the form of the covariance function, it makes assumptions about the types of genetic changes in the mean function that are possible. Other tradeoffs between the two approaches are discussed in Kirkpatrick and Bataillon (1999). Both families of methods require that an appropriate method be used for testing the fit of the estimate to the data, for example a likelihood framework (Meyer and Hill, 1997).

The issue of how best to model covariance functions and how to fit them to data is the focus of active research. Integrating FV methods with statistical paradigms used in animal breeding such as the animal model has begun (e.g. Schaeffer and Dekkers, 1994), but there much more yet to do.

SELECTION ON OTHER KINDS OF FUNCTION-VALUED TRAITS
I have introduced the function-valued framework here by discussing how they can be applied to age-dependent traits. The FV approach has uses in at least two other domains that may be of interest to breeders. One is selection on an animal’s physiological response to its environment, for example growth rate or metabolic rate as a function of nutrition level. The animal’s reaction norm can be represented by a function of the environmental variable, just as size is represented...
as a function of age for a growth trajectory or lactation curve. Thus the FV suite of methods can be applied to reaction norms just as it is to age-dependent traits (Gomulkiewicz and Kirkpatrick, 1992; Gilchrist, 1996; Kingsolver and Huey, 1998; Kingsolver et al., 2001). Selection on the reaction norm introduces some new issues. For example, for some kinds of traits each individual can only be measured in one environment. A start on developing selection methods for reaction norms was made by Kirkpatrick and Bataillon (1999).

A second area where FV methods may be useful to breeders is for selection on morphological shape (Kirkpatrick and Heckman, 1989). For example, an animal’s body conformation in two (or more dimensions) can be described by a mathematical function. New questions raised in this context are how to register the functions for different individuals. Little work has been done here, but it certainly deserves development.

CONCLUSION

This paper briefly outlines an integrated framework for describing traits such as growth trajectories that are a function of age. This approach has several advantages over other methods currently used in quantitative genetics for these “function-valued” traits. Equations predicting the response to selection on one or more age-dependent traits are reviewed, emphasizing parallels with the classic breeder’s equation for selection response. Selection indices can be constructed using (for example) size measurements at several ages, and an algorithm for choosing the optimal ages for measurement is outlined. The framework also provides several families of methods for estimating genetic and phenotypic variation of function-valued traits, and can be used to design selection programs for reaction norms and morphological shapes as well as growth trajectories.

REFERENCES