

COMMENTARY

A fable of four functions

R. GOMULKIEWICZ* & J. G. KINGSOLVER†

School of Biological Sciences and Department of Mathematics, Washington State University, Pullman, WA, USA†Department of Biology, University of North Carolina, Chapel Hill, NC, USA*

The review by Blows (Blows, 2007) highlights a perspective with which we whole-heartedly agree: greater biological insight is obtained by focusing on fundamental structures of the two matrices \mathbf{G} and γ than by analysing the individual entries of those matrices. Here, we advocate extending this point of view further for any 'function-valued trait', i.e. a trait that is inherently a mathematical function, $z(t)$. The function $z(t)$ can be viewed as an extension of a multivariate trait vector \mathbf{z} (Kirkpatrick & Heckman, 1989). We will introduce three more functions $G(t,s)$, $\gamma(t,s)$ and $\beta(t)$ that are, respectively, extensions of the matrices \mathbf{G} and γ , and the vector β discussed in the review.

Function-valued traits arise naturally in many contexts. For example, Blows discusses a study of stabilizing selection on male song in the cricket *Teleogryllus commodus*, characterizing the song as a vector of five song attributes (chip number, interpulse interval, trill number, intercall interval and dominant frequency). However, the song itself is energy (or power) as a function of time and is thus a function-valued trait, $z(t)$. Reducing this function-valued trait to a vector of five measurements allows the 'song' to be analysed in a familiar quantitative genetics framework for multivariate traits, but doing so has important drawbacks (Kirkpatrick & Heckman, 1989; Ramsay & Silverman, 1997). First and foremost, information is necessarily lost in the data reduction. (Consider the futility of grasping Debussy's *Clair de Lune* from just its dominant frequency, duration, average inter-note interval, and a dozen other such elements!) Second, treating such data as functional *ab initio* leads to data reduction methods that retain functional information more efficiently (like *Clair de Lune*'s sheet music or its digital representation on a CD) than do *ad hoc* collections of attributes (Ramsay & Silverman, 1997).

There are many other types of function-valued traits, including growth trajectories (size as a function of age), gene expression profiles (product as a function of time), reaction norms (phenotype as a function of environment) and morphological shapes (radial distance as a function of angle). Typically, these traits are measured at a finite number of values of the independent variable and the finite collection of trait measurements is treated as a

multivariate trait. There are two problems with this. First, there may be more effective ways to combine the measurements taken (or more effective ways to take the finite set of measurements). Second, if measurements are taken at different index values on different individuals (e.g. sizes measured at different sets of ages for different individuals) then multivariate approaches would require some sort of binning procedure, usually improvised. By treating the trait at the outset as a function, however, these issues can be avoided altogether (Kirkpatrick *et al.*, 1994).

A function-centric quantitative genetics framework has been developed for describing function-valued traits, including their genetic variability, selection, and evolution (Kirkpatrick & Heckman, 1989; Kirkpatrick *et al.*, 1990; Gomulkiewicz & Kirkpatrick, 1992; Kirkpatrick & Lofsvold, 1992; Kirkpatrick *et al.*, 1994; Meyer, 1998; Pletcher & Geyer, 1999; Kingsolver *et al.*, 2001). In this framework, which is a direct extension of multivariate quantitative genetics, a trait function $z(t)$ is usually represented mathematically by an 'orthogonal function series expansion'. These expansions can describe practically any biologically plausible function shape, including nonsmooth and discontinuous functions that cannot be written in terms of the perhaps more familiar Taylor series expansion. Orthogonal function expansions can be used with least-squares or likelihood statistical methods to develop estimates of $z(t)$ based on the data collected rather than some preconceived notion about the form of the function. These estimates provide natural interpolations of the original data.

The extension of the genetic variance–covariance matrix \mathbf{G} is the genetic variance–covariance function $G(t,s)$. When $t \neq s$, the G -function gives the additive genetic covariance of the trait expressed at index values t and s (like the off-diagonal elements of \mathbf{G}) and when $s = t$ it equals the additive genetic variance of the trait expressed at index value t (like the diagonal elements of \mathbf{G}).

What does a G -function look like? Consider the example of a growth trajectory, size as a function of age, and suppose all individuals in a study were measured at the same five ages. Given an appropriate breeding design, one could estimate a five-dimensional genetic covariance matrix \mathbf{G} , the elements of which can be visualized as a three-dimensional histogram. The corresponding G -function would be a smoothed surface enveloping the tops of the histogram. Figures 1 and 2 of Kirkpatrick *et al.* (1990) show these visualizations of a genetic covariance matrix \mathbf{G} and the corresponding G -function.

Like \mathbf{G} , the G -function can be 'diagonalized' and understood in terms of its eigenstructure. The G -function can be decomposed into a series of eigenvalues, each of which is the variance associated with a principal component. Principal components are represented by loading functions (eigenfunctions) rather than loading vectors

Correspondence: Joel Kingsolver, Department of Biology, CB-3280, University of North Carolina, Chapel Hill, NC 27599, USA.
Tel.: 919 843 6291; fax: 919 962 1625; e-mail: jgking@bio.unc.edu

(eigenvectors). This decomposition facilitates interpretation of the evolutionary potential of a function-valued trait, to explore such issues as tradeoffs and specialists vs. generalists (Gilchrist, 1996; Kingsolver *et al.*, 2001, 2004; Izem & Kingsolver, 2005). Principal components with zero eigenvalues represent absolute genetic constraints whereas the leading principal components correspond to genetic lines of least resistance (Kirkpatrick & Lofsvold, 1992).

What are the advantages of a functional perspective? First, it preserves information about the ordering and spacing of trait measurements taken at different values of the independent variable. To get a feel for the importance of this, consider an individual's growth trajectory. Clearly, sizes must become increasingly correlated as the ages at which they are measured become increasingly similar. Treating a growth trajectory as a function preserves this information whereas treating a series of measurements as a multivariate trait does not. This advantage can produce practical benefits in statistical analyses. Indeed, function-valued based approaches reduce *P*-values associated with hypothesis testing compared with multivariate tests applied to the same data (N. Heckman, C.K. Griswold, and R. Gomulkiewicz, unpublished).

Another advantage of the function-valued approach is that measurements of a function-valued trait need not be taken at the same index values for different individuals. For example, the sizes of different individuals could be collected at different ages to estimate growth trajectories utilizing standard function-valued statistical methods (Kirkpatrick *et al.*, 1994; Meyer & Hill, 1997; Meyer, 1998). Multivariate analyses, in contrast, would require either arbitrary binning of the same data or demanding laboratory protocols that ensured that sizes were measured at common ages.

Methods have been developed for estimating and modelling selection on function-valued traits (Kirkpatrick & Heckman, 1989; Gomulkiewicz & Kirkpatrick, 1992; Kirkpatrick & Lofsvold, 1992; Kingsolver *et al.*, 2001; Kingsolver & Gomulkiewicz, 2003). There are functional extensions of the linear selection gradient vector β and the quadratic selection matrix γ : the linear selection gradient function $\beta(t)$ and the bivariate quadratic selection function $\gamma(t,s)$, respectively. The meanings of these functions are completely analogous to their multivariate counterparts. For instance, $\beta(t)$ describes linear selection on the trait z when expressed at index value t , holding expression of the trait at all other index values constant, $\gamma(t,s)$ describes pairwise correlational selection between $z(t)$ and $z(s)$, and $\gamma(t,t)$ indicates stabilizing or disruptive quadratic selection on $z(t)$. Like the gamma matrix described in this Target Review, diagonalization or a principal component analysis of the gamma function reveals directions (the PCs) of quadratic selection on a

function-valued trait. We emphasize that gamma alone—whether matrix or function—doesn't completely define the type of selection acting in a population. For example, negative values of γ are necessary but not sufficient for having stabilizing selection on a trait: values of both β and γ are needed to identify stabilizing (or disruptive) selection.

Acknowledgments

This study was supported by NSF grants IBN-0212798 to JGK and EF-0328594 to RG and JGK.

References

- Blows, M.W. 2007. A tale of two matrices: a multivariate approach to evolutionary biology. *J. Evol. Biol.* **20**: 1–8.
- Gilchrist, G.W. 1996. A quantitative genetic analysis of thermal sensitivity in the locomotor performance curve of *Aphidius ervi*. *Evolution*. **50**: 1560–1572.
- Gomulkiewicz, R. & Kirkpatrick, M. 1992. Quantitative genetics and the evolution of reaction norms. *Evolution*. **46**: 390–411.
- Izem, R. & Kingsolver, J.G. 2005. Variation in continuous reaction norms: quantifying directions of biological interest. *Am. Nat.* **166**: 277–289.
- Kingsolver, J.G. & Gomulkiewicz, R. 2003. Environmental variation and selection on performance curves. *Integr. Comp. Biol.* **43**: 470–477.
- Kingsolver, J.G., Gomulkiewicz, R. & Carter, P.A. 2001. Variation, selection and evolution of function-valued traits. *Genetica*. **112/113**: 87–104.
- Kingsolver, J.G., Ragland, G.J. & Shlichta, J.G. 2004. Quantitative genetics of continuous reaction norms: thermal sensitivity of caterpillar growth rates. *Evolution*. **58**: 1521–1529.
- Kirkpatrick, M. & Heckman, N. 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* **27**: 429–450.
- Kirkpatrick, M. & Lofsvold, D. 1992. Measuring selection and constraint in the evolution of growth. *Evolution*. **46**: 954–971.
- Kirkpatrick, M., Lofsvold, D. & Bulmer, M. 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics*. **124**: 979–993.
- Kirkpatrick, M., Hill, W.G. & Thompson, R. 1994. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genet. Res.* **64**: 57–69.
- Meyer, K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol.* **30**: 221–240.
- Meyer, K. & Hill, W.G. 1997. Estimation of genetic and phenotypic covariance functions for longitudinal or repeated records by restricted maximum likelihood. *Livest. Prod. Sci.* **47**: 185–200.
- Pletcher, S.D. & Geyer, C.J. 1999. The genetic analysis of age-dependent traits: modeling the character process. *Genetics*. **151**: 825–835.
- Ramsay, J.O. & Silverman, B.W. 1997. *Functional Data Analysis*. Springer-Verlag, New York.