

ModEDI: An Extendable Software Architecture for Examining the Effects of Developmental Interactions on Evolutionary Trajectories

Elizabeth Brooks

Department of Computer Science
Western Washington University
Bellingham, WA 98225
Email: brookse5@wwu.edu

Graham Roberts

Department of Computer Science
Western Washington University
Bellingham, WA 98225
Email: robertg4@wwu.edu

Alison Scoville

Department of Biology
Central Washington University
Ellensburg, WA 98926
Email: scoville@cwu.edu

Filip Jagodzinski

Department of Computer Science
Western Washington University
Bellingham, WA 98225
Email: filip.jagodzinski@wwu.edu

Abstract—Quantitative genetics is the study of complex biological traits, or traits controlled by more than one gene. A primary goal of quantitative genetic studies is the development of computational models for predicting the evolution of such traits in response to selection. Most models for analyzing the evolution of multiple traits employ a constant genetic variance-covariance matrix (G-matrix) to describe the distribution of genetic variation. While G-matrix based models provide a reasonable approximation of short term evolution, they may not sufficiently capture all important associations between traits. For example, non-linear interactions between developmental factors underlying the production of traits can result in dramatic alteration of genetic variances and covariances as evolution proceeds. To aid in investigating this issue, we have developed an object-oriented code base, Models of Evolution with Developmental Interactions (ModEDI), that implements a powerful, general mathematical and conceptual framework developed by Sean Rice. This framework utilizes a phenotypic landscape to explicitly incorporate the effect of development in shaping heritable phenotypic variation. With our program, users can develop custom simulations for analyzing the evolution of multiple phenotypic traits that involve interactions between overlapping sets of developmental factors. Initial results from simulations performed with ModEDI indicate that developmental interactions may substantially alter evolutionary trajectories.

I. INTRODUCTION

The amount of heritable variation in a population determines its evolutionary response to selection. The statistical construct of genetic variances and covariances (the G-matrix), a central concept in quantitative genetics, is often used to summarize the distribution of heritable variation among multiple traits [1]–[3]. The geometry of the G-matrix quantifies genetic constraints that can hinder, facilitate, or bias evolution towards genetic lines of least resistance (the direction in which most genetic variation exists) [4]. Shared developmental processes represent a principal contributor to this geometry [5], and it is generally assumed that the G-matrix adequately captures aspects of developmental architecture that are relevant for predicting evolutionary trajectories.

A major limitation of G-matrix based models is that their predictive power depends upon the stability of the matrix [1], [4], [6]. Non-additive developmental interactions can dramatically alter genetic variances and covariances of traits [7], [8], rendering approaches based on a constant G-

matrix questionable [9], [10]. A general mathematical and conceptual framework derived by Sean Rice [2], [11]–[13] explicitly incorporates the effect of non-additive developmental interactions on the evolution of traits. **This framework encompasses quantitative genetics as a special case but is far more flexible, allowing for exploration of the effects of different types of selection, under essentially any form of genetic and developmental architecture.**

We present a novel object-oriented software architecture, Models of Evolution with Developmental Interactions (ModEDI), that incorporates the mathematical framework derived by Rice into an object-oriented design that is extendable for custom simulations of user specified populations. Our modular code base implements this framework as a re-usable software bundle to produce both numerical and graphical representations of evolutionary trajectories. As a case study, we analyze evolutionary change in melanic pigmentation and diel vertical migration (DVM) in *Daphnia melanica*, a species of freshwater zooplankton, in response to introduced fish predators. We compare evolutionary trajectories predicted by a constant G-matrix approach, and a model that explicitly recognizes the developmental entanglement between these two traits. Results from our *Daphnia* case study show that incorporating developmental interaction can greatly affect predicted evolutionary trajectories.

II. BACKGROUND

Quantitative genetic models recognize that two elements are needed to predict the evolution of traits due to selection: the strength and direction of selection, and the availability of heritable genetic variation. The strength and direction of selection is captured by the individual fitness surface, which describes the expected reproductive success of an individual as a function of its trait values. The fitness surface in Figure 1 depicts one plausible pattern of expected fitness of a *Daphnia melanica* in the presence of fish predators, based on its melanic pigmentation and tendency to move downward in the water column during the day (DVM). Thus, there exists a value of those two traits such that a maximum fitness is attained, depicted by the peak (plus sign). The partial first derivatives of this landscape represent the strength and direction of selection on each phenotypic trait at any given point in trait space.

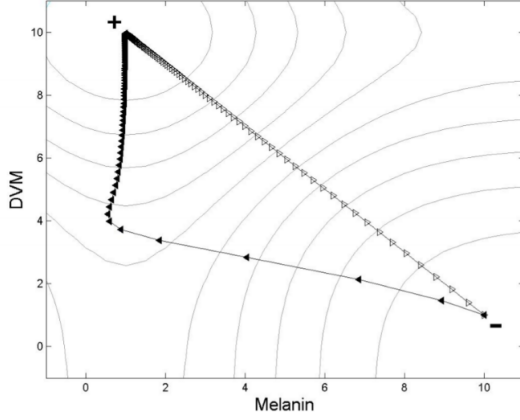


Fig. 1. A fitness surface displays levels of high (+) and low (-) expected reproductive success as a function of trait values. Contour lines indicate fitness isoclines. Arrow heads indicate the mean trait values of a population as it evolves according to a constant G-matrix model (open arrows) or a model that incorporates developmental entanglement (closed arrows).

The G-matrix is commonly employed as the primary descriptor of heritable variation and co-variation in a set of traits. However, evolution of the G-matrix is poorly understood. Long-term empirical evolution studies that examine stability of the G-matrix are rare, and at least some show rapid and unpredictable change over time [6]. Meanwhile, theoretical studies show that stability of the G-matrix depends upon the tendency for single mutations to have a predictable pattern of pleiotropy, i.e., correlated effects on multiple traits [14]. Until recently, approaches to a general understanding of the stability and evolution of the G-matrix have met with limited success [1].

A key insight, that developmental processes structure the geometry of the G-matrix, promises to transform our understanding of evolutionary patterns. The importance of development in shaping phenotypic variation has been recognized for some time [15], and inspired the emergence of evolutionary developmental biology (evo-devo) [16]. The effect of developmental processes has more recently been made analytically

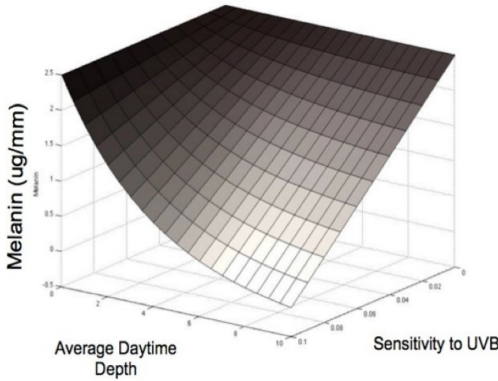


Fig. 2. A phenotypic landscape describing the degree of melanic pigmentation as a function of two underlying developmental factors: average daytime depth (DVM), which influences exposure to UVB radiation, and the degree to which an individual alters melanin production when protected from UVB radiation (sensitivity to UVB).

tractable by the concept of a phenotypic landscape, which is constructed by plotting the value of a phenotype against the values of the underlying genetic and environmental factors that interact in its development [12]. For example, *D. melanica*, a species of freshwater zooplankton, generally reacts to UV exposure by increasing melanin production. However, individuals can differ in UV exposure due to variation in their tendency to take refuge from UV and predation in deep water during the day (DVM). In addition, individuals vary in their ability to respond to changes in UV by altering melanin production. The interaction between these two traits is captured in the phenotypic landscape in Figure 2, which shows the melanic pigmentation in *D. melanica* as a function of average DVM and sensitivity to UVB radiation.

In the mathematical framework derived by Rice [2], the phenotypic landscape describes the distribution of heritable variation at any point in developmental space. In effect, the G-matrix serves as a local linear approximation of the phenotypic landscape at the current population mean [17]. Any non-additive interactions between developmental factors, such as the case of DVM and melanin in *D. melanica*, results in curvature of the landscape and concomitant change in the G-matrix as the population evolves. This change cannot be predicted from estimation of the initial G-matrix, but can be understood as a function of the specific form of developmental interaction underlying a set of traits [17].

Simulations of evolution using the framework developed by Rice promise to aid in the study of phenotypic evolution, multilevel selection, and social evolution for the fields of evolutionary developmental biology, quantitative genetics, and ecology [2], [17]. This framework combines quantitative genetics and evo-devo to incorporate the effects of developmental interactions on the evolution of phenotypes [13]. Thus, models developed with this framework allow for testing hypotheses about how developmental interactions among multiple traits affect their covariances and subsequent evolutionary trajectories.

III. METHODS - MODELS FOR *Daphnia*

To demonstrate the utility of ModEDI, we have developed a case study based on two traits in the species *D. melanica*. The first model in this case study relies on a constant G-matrix approach, while the second incorporates developmental interactions between the two traits. Our object oriented code base implements the mathematical framework derived by Rice [2] for predicting the evolution of phenotypic traits that share at least some underlying developmental factors. Our second model relies on a phenotypic landscape to specify trait values as a function of underlying developmental factors [17]. Additionally, ModEDI is generalized to allow users to easily develop custom models for studying the evolution of multiple phenotypic traits.

The simulations produced by ModEDI may be used to analyze the effect of selection on the means and (co)variances of traits, allowing for comparison with constant G-matrix models, as well as for investigation of the constancy of the G-matrix itself. The initial parameter values specific to *D. melanica* that were implemented in the custom models used in our simulation experiments are shown in Table I. Values

for UVB dose at the water surface and attenuation coefficient were obtained from Sequoia National Park in the Sierra Nevada Range [18]. Values specific to *Daphnia* were obtained from empirical study ([19], unpublished data). Starting trait values match mean values from populations that have not been subject to fish predation. Fitness optima were estimated from mean trait values in populations that have been subject to selection by human-introduced fish predators for approximately 100 years. Variances of the fitness surface and heritability were chosen to represent plausible values.

TABLE I. INITIAL VALUES OF SPECIES VARIABLES ENCODED IN OUR MODELS FOR *Daphnia*

Symbol	Parameter	Value
\bar{m}	Mean Melanin	2.5
\bar{d}	Mean DVM	1.0
σ_m	Phenotypic variance of Melanin	0.04
σ_d	Phenotypic variance of DVM	0.2
ω_m	Optimum Melanin	0.4
ω_d	Optimum DVM	12
v_m	Variance of the Gaussian function relating Melanin to fitness	100
v_d	Variance of the Gaussian function relating DVM to fitness	500
K	Attenuation Coefficient	0.26
u_0	UVB dose at water surface	27.9
h^2	Heritability matrix (diagonal values)	0.5

A. Mathematical Framework

We model the evolution of phenotypic traits due to selection by specifying the strength and direction of selection, as well as the availability of heritable genetic variation. The individual fitness surface (1) describes the strength and direction of selection for *D. melanica*, which is defined by the sum of fitness values associated with the individual m and d values of an individual. The fitness values for m and d are each described by a Gaussian function with respective optima at ω_m and ω_d and variances of v_m and v_d . The results of our models are visualized by graphing the trajectory of the mean bivariate phenotype, superimposed over the individual fitness surface (Figures 4-8).

$$w = \frac{1}{\sqrt{v_m 2\pi}} e^{-\frac{(m-\omega_m)^2}{2v_m}} + \frac{1}{\sqrt{v_d 2\pi}} e^{-\frac{(d-\omega_d)^2}{2v_d}} \quad (1)$$

In our **Classic**, G-matrix based model, evolution of melanin and DVM over time is simulated based on constant genetic covariance of these traits and the natural selection imposed by the introduction of a novel fish predator, as specified by our fitness function (1). Our Classic model predicts the evolution of mean melanin (2) and mean DVM (3) from generation t to generation $t+1$, starting at the initial values given in Table I. At each point in time t , the partial derivatives of the individual fitness surface are evaluated at \bar{m}_t and \bar{d}_t . Heritability is defined by the diagonal values of the heritability matrix h^2 .

$$\bar{m}_{t+1} = \bar{m}_t + h^2 \frac{1}{\bar{w}} \frac{\partial w}{\partial m} \sigma_m \quad (2)$$

$$\bar{d}_{t+1} = \bar{d}_t + h^2 \frac{1}{\bar{w}} \frac{\partial w}{\partial d} \sigma_d \quad (3)$$

Mean population fitness for our models is calculated as the product of individual fitness, w , and the joint distribution of m and d , integrated over m and d . The joint distribution of m and d is described by a bivariate Gaussian distribution with means \bar{m} and \bar{d} , variances σ_m and σ_d , and a covariance of 0.

Developmental entanglement arises in our second model, **Tanning**, due to the fact that DVM influences exposure to UVR, which decays exponentially with depth in the water, and UVR affects production of melanin. This model predicts the evolution of mean trait values both **Analytically** (4) and by **Simulation** of individuals (5). For simplicity, we assume that the relationship between UV exposure and melanin production is linear. Melanin is calculated for each individual *Daphnia* as $m = z + pa$, where z is the concentration of melanin produced in response to exposure to UV light available at the top of the water column, a is the change in UVB exposure compared to the top of the water column, and p is the slope of the reaction norm indicating sensitivity of melanin production to UVB. Thus, the phenotypic function for melanin production by an individual (4) is calculated by describing the change in UVB exposure, a , as UVB decreases exponentially with increased depth in the water column.

$$m = z + p (\mu_0 e^{-dk} - \mu_o) \quad (4)$$

Mean melanin is calculated analytically (5) in our Tanning model as the product of m and the joint distribution of z , p , and d , integrated analytically over z and p and numerically over d .

$$\bar{m} = \bar{z} + \left(-1 + e^{\frac{1}{2}k(-2\bar{d}+k\sigma_d)} \right) \bar{p}\mu_0 \quad (5)$$

Both mean melanin and DVM are determined in our Tanning model by treating the underlying developmental factors, z and p , as quantitative genetic traits, following the polynomial approach to modeling the evolution of phenotypic plasticity [20], [21]. Thus, change in trait values from generation t to generation $t+1$ due to directional selection is given by (6) when developmental entanglement is taken into account [2], [12].

$$\begin{pmatrix} \bar{z} \\ \bar{p} \\ \bar{d} \end{pmatrix}_{t+1} = \begin{pmatrix} \bar{z} \\ \bar{p} \\ \bar{d} \end{pmatrix}_t + H \left(\frac{1}{\bar{w}} \frac{\partial w}{\partial d} \begin{bmatrix} 0 \\ 0 \\ \sigma_d \end{bmatrix} + \frac{1}{\bar{w}} \frac{\partial w}{\partial m} \begin{bmatrix} \mu_0 e^{-dk} - \mu_o \\ Kp\mu_0 e^{-dk} \sigma_p \end{bmatrix} \right) \quad (6)$$

B. Software Architecture

Object-oriented programming is centered around the concept of modular objects, which represent items with related states and behaviors. In our case study, the species *D. melanica* is represented as an object with phenotypes as states and phenotypic functions as behaviors. Using a well defined class hierarchy, related states and behaviors are described in connected bundles of software. The code architecture of ModEDI, Figure 3, may be easily extended by developing new derived classes that reuse already existing software bundles of phenotypic states and functions.

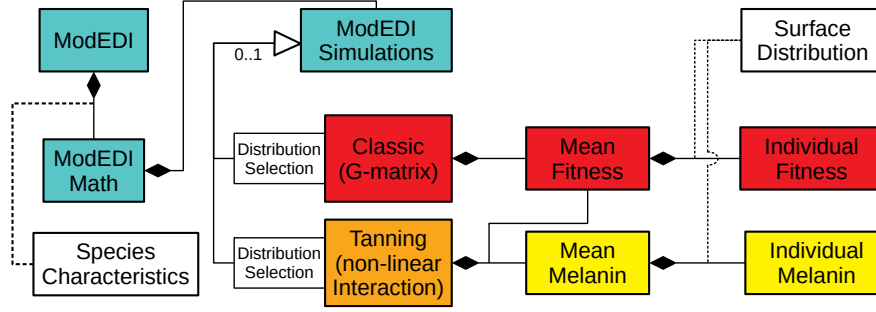


Fig. 3. A UML diagram of the ModEDI software architecture with two colored derived classes (Classic and Tanning). Interface classes are connected with dashed lines, while derived classes with solid lines and filled diamonds. The multiple generalization relationship is designated with a hollow arrowhead.

For example, the Surface Distribution interface class provides a template for describing functions that generate random values following Gaussian, sigmoid, and linear distributions. These Distribution Selections are used by the Mean Fitness class to create the Gaussian fitness surface for both models, and the Mean Melanin class to create the melanin phenotype surface. The derived ModEDI Simulations class is qualified by the input Distribution Selection to each model, where up to one evolutionary trajectory may be simulated for each set of input model values.

Our models for *Daphnia* describe the evolution of mean melanin and DVM over time by first referencing the initial phenotypic states declared in the Species Characteristic interface class. Each model projects the evolutionary trajectories of mean melanin and DVM by iteratively updating states and passing the updated values to intermediate functions and classes. Finally, the calculated mean values for each model are output by ModEDI Math to separate text files for parsing and plotting upon simulation completion.

Underlying developmental factors are incorporated in our Tanning model to project changes from one generation to the next in the presence of directional selection. The dynamics depend upon the partial derivatives of the phenotypic landscape. Mean trait values for each generation are passed to phenotypic functions that calculate new mean values as the phenotypic states for the next generation. This iterative process is used to predict the evolutionary response of both traits.

Both models are deterministic and can be approached analytically. However, we also estimate mean population values and derivatives of the phenotypic landscape for our Tanning model via simulation, for comparison with the analytic results. This serves as a proof of concept for use of simulation in cases where the system of equations describing a set of traits is too complex to be analytically tractable.

IV. RESULTS - A CASE STUDY

An extensive sweep was performed for each model to determine the sensitivity to starting values for each parameter. Starting at the initial values given in Table I, each parameter was independently altered prior to running the program. The results of our models are visualized by graphing the trajectory of the mean bivariate phenotype, superimposed over the contour plot of the individual fitness landscape.

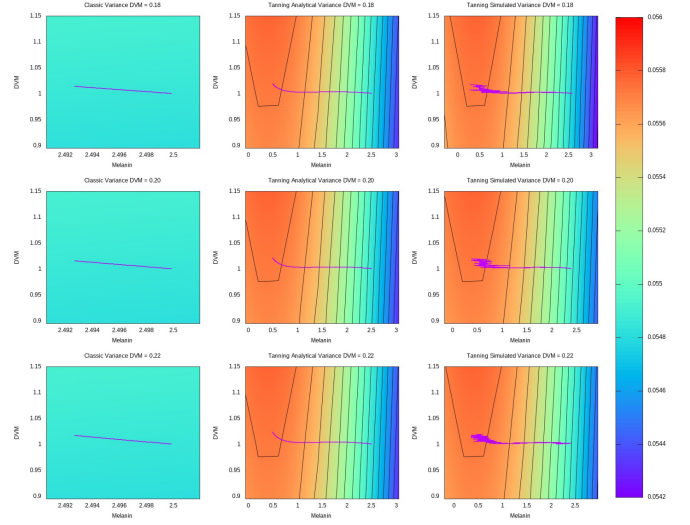


Fig. 4. A focused parameter sweep of the phenotypic variance of DVM, $\sigma_d = 0.2 \pm 0.02$. The value of σ_d is constant across rows (0.18, 0.20, and 0.22, Top-to-Bottom), while the columns indicate modeling method (Classical G-matrix, Tanning Analytical, and Tanning Simulated, Left-to-Right). The colors among the contoured fitness surface specify high (red) and low (blue) expected reproductive success.

The results of two focused parameter sweeps are shown in Figure 4 and Figure 5. The first shows the effect of changing the phenotypic variance of DVM around 10% of the value quoted in Table I. Displayed are three runs on each of the models, Classic, Analytic Tanning, and Simulated Tanning. These experiments show similar results, suggesting that trajectories are robust to uncertainties in the variance of at least one developmental factor. The stochastic nature of the simulated model does appear to decrease as the variance decreases, which is to be expected.

The second sweep was performed about the initial value for DVM of $d = 1.0$. There is a fairly dramatic change in the early generations of the evolutionary trajectory for the Tanning models, resulting in a pronounced curve as this initial value is decreased. This suggests that the starting point of a population can significantly affect its initial evolutionary trajectory, although each of these curves converges to approximately the same trajectory over later generations.

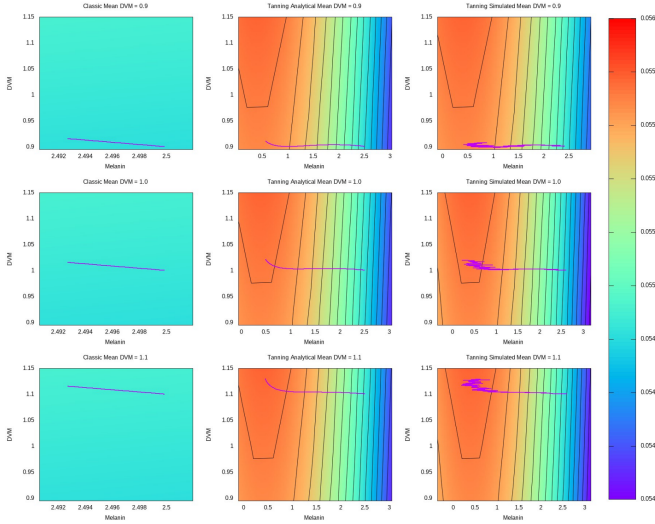


Fig. 5. A focused parameter sweep of the initial value for DVM. The rows show 10% changes in the value (0.9, 1.0, and 1.1, Top-to-Bottom), while the columns are for each modeling method (Classical G-matrix, Tanning Analytical, and Tanning Simulated, Left-to-Right). The colors among the contoured fitness surface specify high (red) and low (blue) expected reproductive success.

Our experiments show that incorporating non-additive interactions between the two traits substantially influences evolutionary trajectories. While both models approach the same optimum values, the developmental entanglement incorporated in the Tanning model shows increased genetic covariance in the direction of selection. This results in a substantially faster rate of adaption than displayed in the results of the Classic model. In addition, DVM changes much more rapidly relative to melanin in Tanning. As we see in Figure 6, even though both parameters evolve toward their optimum values, the line in the Classic model displays a direct path up the fitness landscape, rather than the sharp “L” shape seen in the Tanning models. In addition, parameter sweeps reveal that the Tanning model displays greater responsiveness to minor changes in initial values and phenotypic variances.

V. CONCLUSIONS - FUTURE WORK

The class structure of our ModEDI code base is conducive to future extension and generalization. This will be particularly important for allowing users to input custom functions describing trait relationships within a species, while maintaining efficiency. For example, a surface distribution needs to be generated and evaluated based on the user supplied functions for every trait under consideration. By maintaining an object oriented structure for our code base, the same functions may be used to draw and sample from each distribution.

With our program, users will be able to develop custom simulations for analyzing the evolution of multiple phenotypic traits. We expect that these simulations will enable users to efficiently test hypotheses regarding the consequences of non-linear developmental interactions on the evolution of specific traits of a species.

A. User Interface

A streamlined web interface is in development. It allows a user to quickly input the parameters they wish to study.

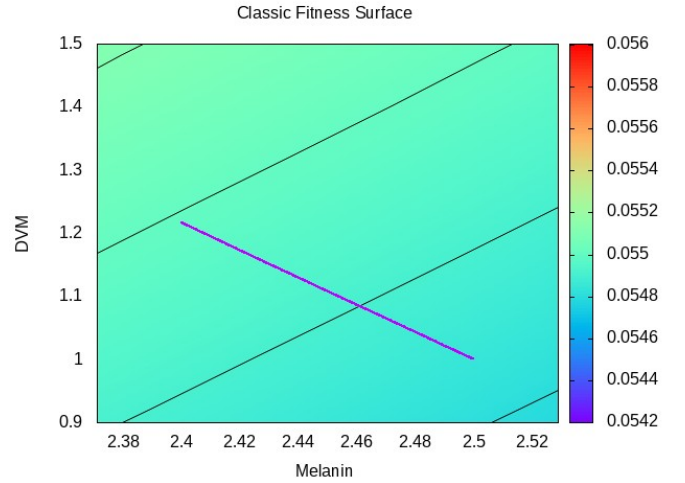


Fig. 6. Classic model results with melanin plotted against DVM over a contour plot of the fitness surface. Colors specify high (red) and low (blue) expected reproductive success.

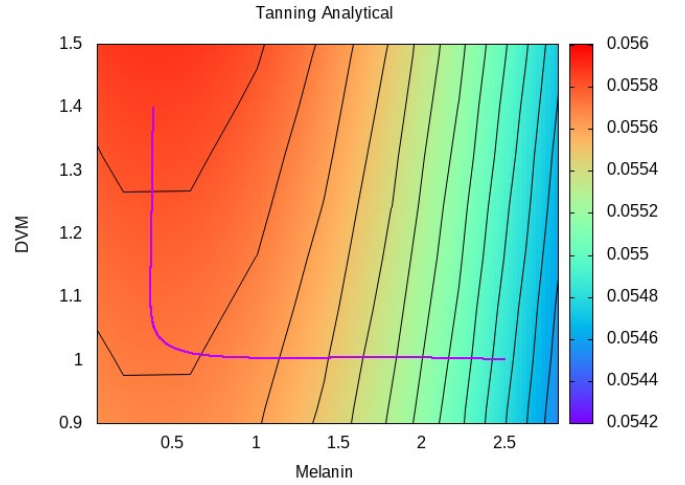


Fig. 7. Tanning model, analytical results with melanin plotted against DVM over a contour plot of the fitness surface. Colors specify high (red) and low (blue) expected reproductive success.

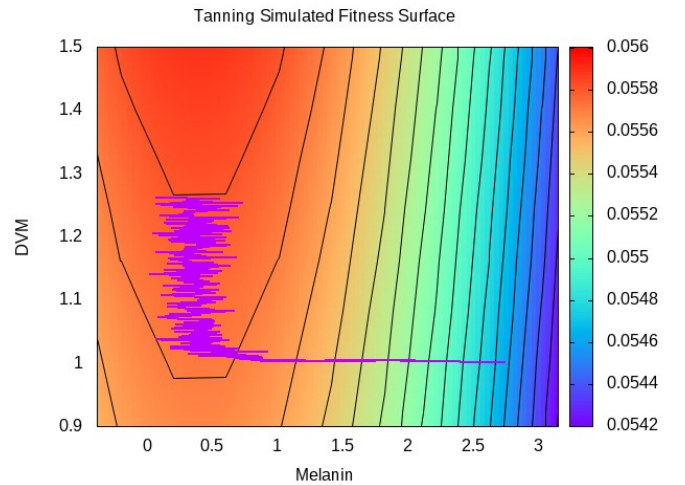


Fig. 8. Tanning model, simulated results with melanin plotted against DVM over a contour plot of the fitness surface. Colors specify high (red) and low (blue) expected reproductive success.

The output numerical data is visualized, enabling end users to rapidly gauge the evolutionary trajectory of an organism. Detailed information of the generation-by-generation distribution of phenotypic traits are displayed and available for download. Previous simulations are stored allowing the user to reference past simulations and compare results in a streamlined manner.

Our online interface also provides users with the ability to sweep parameter ranges, provided minimum, maximum, and step distances for one or more parameters. Through an iterative process, multiple evolutionary trajectories are efficiently modeled and results displayed by-by-side. The output data can then be compressed and downloaded for a more detailed analysis. This allows users to quantify the degree to which initial parameter values affect their conclusions, providing insight into the uncertainty of evolutionary trajectories. In cases where a user wishes to sweep multiple parameters, trials are conducted for each set of possible values.

ACKNOWLEDGMENT

The authors would like to thank the Science Honors program at Central Washington University, and the College of Sciences and Engineering at Western Washington University, for supporting this work. Brooks designed and developed the prototype and subsequent versions of ModEDI, Roberts conducted the parameter sweep experiments and aided in developing the code for numerical calculations, while Jagodzinski and Scoville supervised the work, focusing on the software development and testing, and biological and mathematical foundations, respectively. All authors contributed to writing the manuscript.

REFERENCES

- [1] S. J. Arnold, R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones, "Understanding the evolution and stability of the G-matrix," *Evolution*, vol. 62, no. 10, pp. 2451–2461, 2008.
- [2] S. H. Rice, "A general population genetic theory for the evolution of developmental interactions," *Proceedings of the National Academy of Sciences*, vol. 99, no. 24, pp. 15 518–15 523, 2002.
- [3] R. Lande and S. J. Arnold, "The measurement of selection on correlated characters," *Evolution*, pp. 1210–1226, 1983.
- [4] D. Schluter, "Adaptive radiation along genetic lines of least resistance," *Evolution*, pp. 1766–1774, 1996.
- [5] J. M. Smith, R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert, "Developmental constraints and evolution: a perspective from the mountain lake conference on development and evolution," *Quarterly Review of Biology*, pp. 265–287, 1985.
- [6] M. Björklund, A. Husby, and L. Gustafsson, "Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years," *Journal of Evolutionary Biology*, vol. 26, no. 1, pp. 1–13, 2013.
- [7] A. J. Carter, J. Hermisson, and T. F. Hansen, "The role of epistatic gene interactions in the response to selection and the evolution of evolvability," *Theoretical population biology*, vol. 68, no. 3, pp. 179–196, 2005.
- [8] I. Salazar-ciudad and J. Jernvall, "How different types of pattern formation mechanisms affect the evolution of form and development," *Evolution & development*, vol. 6, no. 1, pp. 6–16, 2004.
- [9] F. Eroukhanoff, "Just how much is the G-matrix actually constraining adaptation?" *Evolutionary Biology*, vol. 36, no. 3, pp. 323–326, 2009.
- [10] T. F. Hansen, "The evolution of genetic architecture," *Annual Review of Ecology, Evolution, and Systematics*, pp. 123–157, 2006.
- [11] S. Rice, "Developmental associations between traits: Covariance and beyond," *Genetics*, vol. 166, no. 1, pp. 513–526, 2004.
- [12] S. H. Rice, *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates Sunderland, MA, 2004.
- [13] J. B. Wolf, "The geometry of phenotypic evolution in developmental hyperspace," *Proceedings of the National Academy of Sciences*, vol. 99, no. 25, pp. 15 849–15 851, 2002.
- [14] A. G. Jones, S. J. Arnold, and R. Bürger, "Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift," *Evolution*, vol. 57, no. 8, pp. 1747–1760, 2003.
- [15] W. Bateson, *Materials for the study of variation: treated with especial regard to discontinuity in the origin of species*. Cambridge University Press, 2012.
- [16] J. T. Streebman, *Advances in Evolutionary Developmental Biology*. John Wiley & Sons, 2013.
- [17] S. H. Rice, "The G-matrix as one piece of the phenotypic evolution puzzle," *Evolutionary Biology*, vol. 35, no. 2, pp. 106–107, 2008.
- [18] J. Diamond, J. Butcher, S. Klaine, J. Bearr, M. Bowersox, H. Latimer, T. Hoang, and H. Zahner, "Validation of understanding assumption for integrating frequency, magnitude, and duration in permit conditions," *Water Environmental Research Foundation Report*, 2005.
- [19] A. Scoville and M. Pfrender, "Phenotypic plasticity facilitates recurrent rapid adaptation to introduced predators," *Proceedings of the National Academy of Sciences*, vol. 107, no. 9, pp. 4260–4263, 2010.
- [20] S. M. Scheiner and R. F. Lyman, "The genetics of phenotypic plasticity i. heritability," *Journal of Evolutionary Biology*, vol. 2, no. 2, pp. 95–107, 1989.
- [21] S. Gavrillets and S. M. Scheiner, "The genetics of phenotypic plasticity. vi. theoretical predictions for directional selection," *Journal of Evolutionary Biology*, vol. 6, no. 1, pp. 49–68, 1993.