Complexity and Diversity

Michael Doebeli* and Iaroslav Ispolatov*

The mechanisms for the origin and maintenance of biological diversity are not fully understood. It is known that frequency-dependent selection, generating advantages for rare types, can maintain genetic variation and lead to speciation, but in models with simple phenotypes (that is, low-dimensional phenotype spaces), frequency dependence needs to be strong to generate diversity. However, we show that if the ecological properties of an organism are determined by multiple traits with complex interactions, the conditions needed for frequency-dependent selection to generate diversity are relaxed to the point where they are easily satisfied in high-dimensional phenotype spaces. Mathematically, this phenomenon is reflected in properties of eigenvalues of quadratic forms. Because all living organisms have at least hundreds of phenotypes, this casts the potential importance of frequency dependence for the origin and maintenance of diversity in a new light.

There are two major aspects underlying the general question of how phenotypic diversity originates and is maintained within a population or species. First, standing genetic variation within a single species is typically much greater than would be expected if evolution were solely due to an adaptive optimization process (*1*–*3*). Second, the mechanisms that diversify lineages into multiple diverging descendent lineages remain elusive (*4*, *5*). Much pertinent evolutionary theory has been developed (*5*–*11*), including neutral evolution, mutation-selection balance, and frequency-dependent selection (also referred to as balancing selection). With negative frequency dependence, rare types have an advantage, which can contribute to the maintenance of genetic variation (*3*) and the origin of new species (*12*, *13*). Nevertheless, whether frequency-dependent selection is of general importance for biological diversity remains unclear. The vast majority of models have considered evolution in simple one-dimensional (1D) phenotype spaces, in which frequency dependence needs to be strong to generate diversity.

We extended classical models for frequency-dependent competition from one to many phenotypic dimensions. In the basic 1D model, which has a long tradition in ecological and evolutionary theory (*12*, *14*–*18*), individuals have a scalar (that is, 1D) phenotype x that determines resource preferences. If individuals have similar preferences, they will consume similar types of food and hence experience strong competition. The competitive impact between two individuals with phenotypes x and y is described by the competition kernel $\alpha(x,y)$, which is assumed to be a unimodal function of the phenotypic distance $|x - y|$ that has its maximum at $|x - y| = 0$. This generates a frequency-dependent component of selection: Individuals with a rare phenotype have only a few similar individuals in the population, and hence experience large competitive impacts from only a few individuals, whereas those with common phenotypes experience large competitive impacts from many other individuals.

If $\varphi(x)$ is the density distribution of phenotypes, its dynamics is given by the logistic partial differential equation

$$\frac{\partial\varphi(x)}{\partial t} = r\varphi(x)\left(1 - \frac{\int\alpha(x,y)\varphi(y)dy}{K(x)}\right) \quad (1)$$

Here $\int\alpha(x,y)\varphi(y)dy$ is the effective density experienced by x individuals, which is a weighted sum over the whole population $\varphi(y)$, with the weights equal to the competitive impact of y individuals on x individuals. The function $K(x)$ is the carrying capacity function, giving the equilibrium density of a hypothetical population that only contains x individuals. $K(x)$ is assumed to be unimodal with a maximum at $x = 0$ and represents a stabilizing component of selection for trait value $x = 0$. (The intrinsic growth rate r is assumed to be independent of the phenotype x and is set to 1.) We assume the following functional forms for the competition kernel and the carrying capacity

$$\alpha(x,y) = \exp(-a|x-y|^{n_\alpha}) \quad (2)$$

$$K(x) = K_0\exp(-k|x|^{n_K}) \quad (3)$$

These functions are symmetric because of the assumptions that they depend on the absolute value $|\ldots|$ of their argument. The parameters $a > 0$ and $k > 0$ measure the relative importance of the frequency-dependent component and the stabilizing component of selection, and n_α and n_K are shape parameters, which we assume to be real numbers ≥ 2 to ensure that the functions are at least twice differentiable at 0. K_0 scales the overall population size. The dynamics of the model in Eq. 1 is well known in the so-called Gaussian case, in which the shape parameters are $n_\alpha = n_K = 2$. If $k > a$, the stabilizing component of selection dominates and hence the dynamics of Eq. 1 converges to a delta peak centered at $x = 0$. In particular, all variation is lost from the population in this case. Conversely, if $k < a$, competition is more localized in phenotype space as compared to the width of the carrying capacity curve, so that the frequency-dependent component of selection dominates. In such cases, the dynamics of Eq. 1 converge to an equilibrium distribution that is proportional to a normal distribution with a positive variance, and hence variation is maintained. In essence, this is the basis for the claim that strong frequency dependence can maintain diversity (Fig. 1).

It is important to note that for competition kernels and carrying capacities that are non-Gaussian, the maintenance of diversity due to

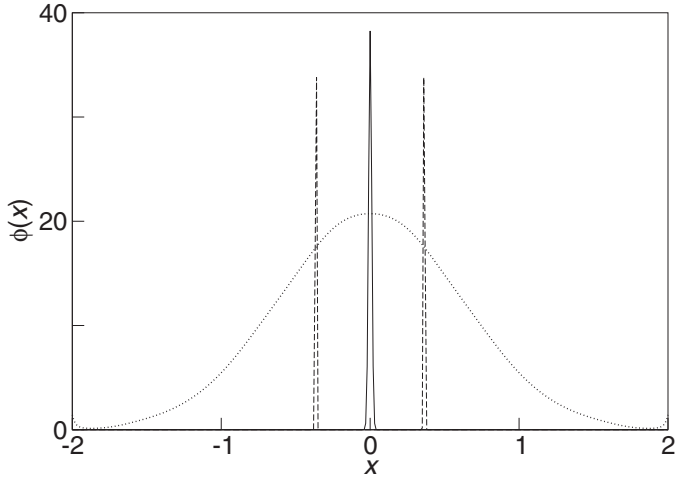

Fig. 1. Equilibrium population density for the Gaussian model (Eq. 1), with $n_\alpha = n_K = 2$ in Eqs. 2 and 3, and $a = k = 1$ (solid line) and $a = 2$, $k = 1$ (dotted line). In the long time limit, the former case converges to a delta peak centered at $x = 0$ and hence corresponds to loss of genetic variation, whereas the latter distribution retains a positive variance. The two narrow peaks given by the dashed lines represent maintenance of variation in the form of a bimodal equilibrium distribution for platykurtic competition and carrying capacity functions, $n_\alpha = n_K = 4$ in Eqs. 2 and 3, with $a = 2$ and $k = 1$. In the long time limit, this distribution converges to two delta peaks. It is straightforward to introduce mutations into the model given by Eq. 1; for example, by adding a diffusion term. This would result in qualitatively very similar results, except that the delta peaks would be replaced by narrow peaks with a small but positive variance due to mutation.

Department of Zoology and Department of Mathematics, University of British Columbia, 6270 University Boulevard, Vancouver, BC V6T 1Z4, Canada.

*To whom correspondence should be addressed. E-mail: doebeli@zoology.ubc.ca (M.D.); slava@math.ubc.ca (I.I.)

strong frequency dependence in Eq. 1 may not result in a unimodal equilibrium distribution (*19*–*22*). For example, with $n_a = n_k = 4$, and hence with platykurtic functions that fall off less sharply from the maximum than Gaussian functions, the case where the frequency-dependent component is stronger than the stabilizing component of selection leads to multimodal equilibrium distributions (Fig. 1), in which the different modes represent different phenotypic clusters. But regardless of the exact form in which variation is maintained, the basic model (Eq. 1) illustrates that in 1D phenotype spaces, strong frequency dependence is needed to maintain diversity.

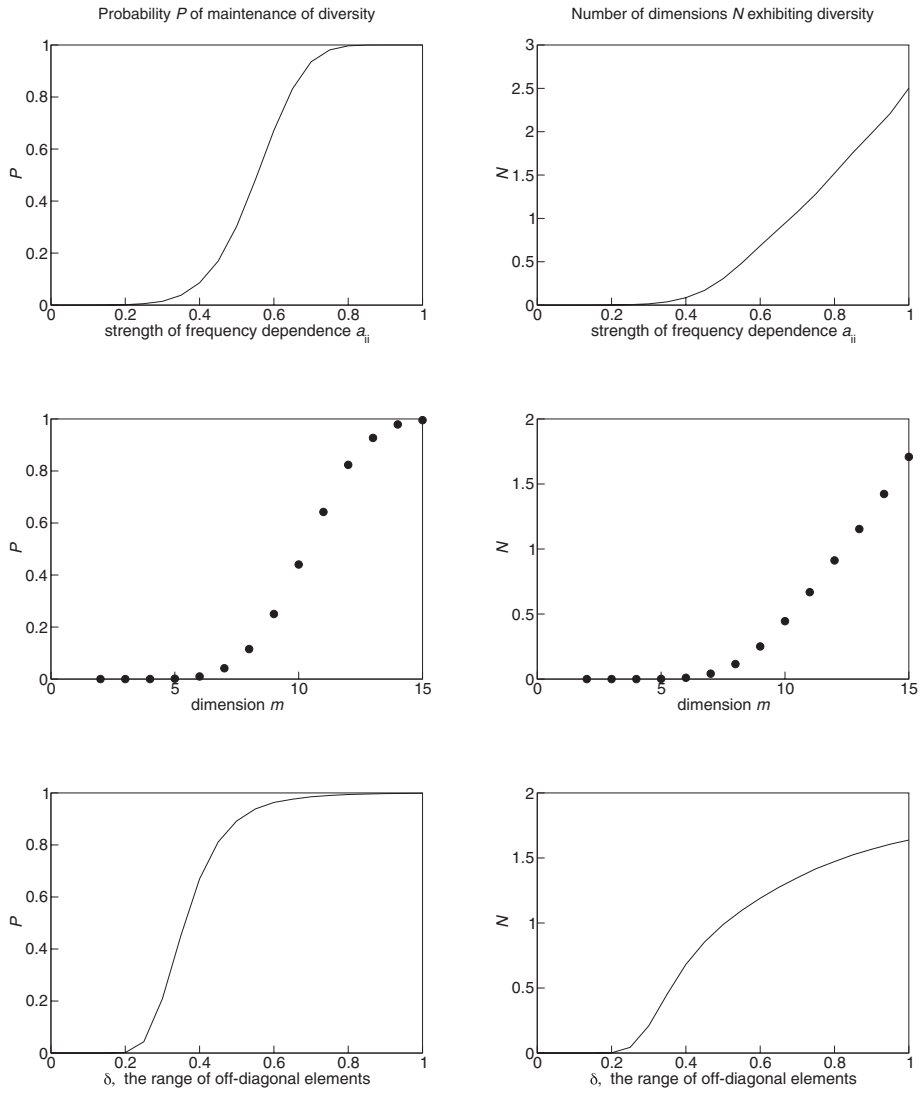

Fig. 2. Effects of interactions between phenotypic components on diversity. The starting points are diagonal matrices K with identical diagonal elements $k_{ii} = 1$, and A with identical diagonal elements $a_{ii} < 1$, in the Gaussian model (Eq. 6), so that the condition $a > k$ for the maintenance of diversity is not satisfied in any phenotypic direction. The panels shows results due to the addition of uniformly randomly chosen off-diagonal elements in the range [–δ,δ] at each off-diagonal position in both matrices. Only realizations when both A and K are positive definite are considered. The first column shows the probability that this results in the condition for diversification $a > k$ satisfied in at least one direction after simultaneous diagonalization of the resulting quadratic forms (*23*). This probability is shown as a function of the diagonal elements a_{ii} in the starting matrix A (first row), as a function of the dimension of phenotype space m (second row), and as a function of the parameter δ determining the range from which the off-diagonal elements are drawn (third row). The right-hand column shows the average number of phenotypic dimensions in which diversification occurs after the addition of random off-diagonal elements. Results were obtained from 10^5 random samplings of k_{ij} and a_{ij}, with standard diagonalization techniques applied to the resulting quadratic forms. The parameters that are fixed are the dimension of phenotype space $m = 5$ and the range $\delta = 0.2$ for the first row, $a_{ii} = 0.2$ and $\delta = 0.2$ for the second row, and $m = 5$ and $a_{ii} = 0.2$ for the third row. The probability of diversification becomes large even if $k_{ii} - a_{ii}$ is substantially larger than δ (top left panel). Also, even for k_{ii} much larger than a_{ii} (that is, even when stabilizing selection is overwhelming frequency dependence in each component direction), the probability of diversification due to off-diagonal elements approaches 1 for high dimensions m (middle left panel).

We extend the Gaussian model (Eq. 1) with $n_a = n_k = 2$ to multiple phenotypic dimensions $x_1,\ldots,x_m$ by assuming that in each direction x_i there is a frequency-dependent component and a stabilizing component of selection, described by parameters a_{ii} and k_{ii}, respectively. The competition kernel is

$$\alpha(x_1,\ldots,x_m;y_1,\ldots,y_m) = \exp(-(x-y)\cdot A\cdot(x-y)^T) \quad (4)$$

where A is an $m \times m$ matrix and where $x = (x_1,\ldots,x_m)$, $y = (y_1,\ldots,y_m)$ [so that $(x - y) = (x_1 - y_1,\ldots,x_m - y_m)$], and y^T is the transposed vector. Here the various scalars x_i should be thought of as observable phenotypic traits, such as body size or some other morphological feature, behavior, coloration, or a metabolic property. Such traits are not only biologically identifiable but are also often controlled by distinct sets of genes and/or signaling pathways. If there are no interactions between the different components x_i of the m-dimensional phenotype (where interaction is understood in the statistical sense), A is a diagonal matrix with entries $a_{11}, \ldots, a_{mm}$. In this case, $x \cdot A \cdot y^T = a_{11}(x_1 - y_1)^2 + \ldots + a_{mm}(x_m - y_m)^2$, and the competitive impact between individuals with phenotype vectors $x = (x_1,\ldots,x_m)$ and $y = (y_1,\ldots,y_m)$ is simply a product of the impacts of the phenotypic components. However, in reality, different phenotypic components are rarely ecologically independent and will typically interact to affect competition in a complicated nonmultiplicative way. For example, the competitive impact of an individual with beak length y_1 and wing span y_2 on an individual with beak length x_1 and wing span x_2 may be larger than the product of the competitive impacts measured in each phenotypic direction separately, if birds with both a large beak and large wings tend to be more aggressive. This would be reflected by an off-diagonal element $a_{12} < 0$ in the matrix A. Similarly, the competitive impact of bacteria secreting amounts y_1 and y_2 of two toxins may be less than the product of the competitive impacts induced by y_1 and y_2 separately, if the effectiveness of one toxin is reduced in the presence of another toxin, which would be reflected by an off-diagonal element $a_{12} > 0$. Without loss of generality, we assume A, representing the quadratic form (Eq. 4), to be a symmetric matrix.

For the stabilizing component of selection, we may assume that in each phenotypic direction, the location of the maximum of the carrying capacity is at $x_i = 0$. The carrying capacity is then given by a function

$$K(x_1,\ldots,x_m) = K_0 \exp(-x\cdot K\cdot x^T) \quad (5)$$

where K is an $m \times m$ matrix and where $x = (x_1,\ldots,x_m)$. Again, if there are no interactions

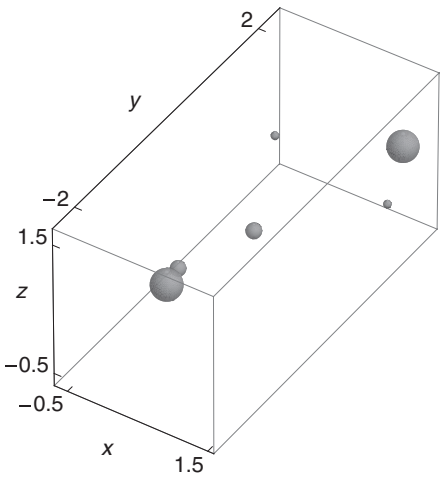

Fig. 3. Equilibrium distribution of the 3D non-Gaussian model (Eqs. 9 and 10) with matrices

$$A = \begin{pmatrix} 0.32 & 0.55 & 0.28 \\ 0.55 & 0.32 & -0.21 \\ 0.28 & -0.21 & 0.32 \end{pmatrix} \text{ and}$$

$$K = \begin{pmatrix} 1 & -0.018 & 0.686 \\ -0.018 & 1 & 0.552 \\ 0.686 & 0.552 & 1 \end{pmatrix}$$

In all three phenotypic dimensions, stabilizing selection dominates frequency-dependent selection (taken in each direction separately, the above parameters correspond to $a = 0.32$ and $k = 1$ in Eqs. 2 and 3, with $n_\alpha = n_K = 4$), but phenotypic interactions in the form of nonzero off-diagonal elements generate diversification into a sum of delta peaks whose location in the 3D phenotype space is indicated by the spheres. The size of the spheres represents the weight of the corresponding delta peak. Projections onto axes *x*, *y*, and *z* show diversification in all three phenotypic dimensions [fig. S1 in (*23*)].

between the different phenotypic components, K is a diagonal matrix with entries $k_{11}, \ldots, k_{mm}$, and $x \cdot K \cdot x^T = k_{11}x_1^2 + \ldots + k_{mm}x_m^2$, so that the carrying capacity of phenotype $x = (x_1,\ldots,x_m)$ is simply a product of the carrying capacities of the phenotypic components. In general, phenotypes affect the carrying capacity in a more complicated way, which is again reflected in nonzero off-diagonal elements in the matrix K. For example, a predator feeding on two different species with efficiencies determined by two traits x_1 and x_2 could have a nonzero off-diagonal matrix element k_{12} in its carrying capacity due to the fact that more effective predation on one species affects the availability of the other species through the dynamics of the entire food web. Without loss of generality, we choose K to be a symmetric matrix. Also, we assume that the carrying capacity has a unique maximum at $x = 0$ and that the competition kernel has a line of maxima at $y = x$. This implies that both matrices A and K are positive definite (that is, all their eigenvalues are real and > 0).

The extension of the Gaussian model (Eq. 1) to multiple dimensions is thus determined by the two quadratic forms A and K, and the dynamics of the multidimensional phenotype density distribution $\varphi(x_1,\ldots,x_m)$ are given by

$$\frac{\partial \varphi(x_1,\ldots,x_m)}{\partial t} = \varphi\left(1 - \frac{\int \alpha(x_1,\ldots,x_m;y_1,\ldots,y_m)\varphi(y_1,\ldots,y_m)dy_1\cdots dy_m}{K(x_1,\ldots,x_m)}\right) \quad (6)$$

where α and K are given by Eqs. 4 and 5.

If there are no interactions between different phenotypic dimensions, and hence the matrices A and K are diagonal, the conditions for the maintenance of diversity in the multidimensional model are essentially the same as in the 1D model (Eq. 1): Diversity is maintained if and only if $a_{ii} > k_{ii}$ for at least one phenotypic component i. This shows that the dimensionality alone has no effect on the maintenance of diversity. This changes, however, if there are interactions between phenotypic components, so that A and K are not diagonal, as is most easily illustrated in the case of two phenotypic components. Assuming that $a_{11} = k_{11} = a_{22} = k_{22} = 1$, so that in each dimension, the system is exactly on the boundary of the parameter region in which diversity is maintained, with interactions between the two phenotypes, the matrices A and K have the form

$$A = \begin{pmatrix} 1 & a_{12} \\ a_{12} & 1 \end{pmatrix} \quad K = \begin{pmatrix} 1 & k_{12} \\ k_{12} & 1 \end{pmatrix} \quad (7)$$

where a_{12} and k_{12} may be positive or negative. The matrices A and K can be diagonalized simultaneously by the linear change of coordinates $y_1 = (x_1 - x_2)/\sqrt{2}$ and $y_2 = (x_1 + x_2)/\sqrt{2}$. In these new, diagonal coordinates, the quadratic forms A and K become

$$A = \begin{pmatrix} 1 - a_{12} & 0 \\ 0 & 1 + a_{12} \end{pmatrix} K = \begin{pmatrix} 1 - k_{12} & 0 \\ 0 & 1 + k_{12} \end{pmatrix} \quad (8)$$

Hence the dynamical system (Eq. 6) is equivalent to a system where the phenotypic components do not interact, but where the parameters a_{ii} and k_{ii} have the values $1 \pm a_{12}$ and $1 \pm k_{12}$ instead of 1. In general, the interaction effects between phenotypes are different for competition and for the carrying capacity (that is, $a_{12} \neq k_{12}$), and it immediately follows that in the new coordinate system, the condition for diversification is satisfied in one of the two dimensions (we either have $1 - a_{12} > 1 - k_{12}$, or $1 + a_{12} > 1 + k_{12}$). The direction satisfying the diversification condition is a linear combination of the original phenotypic axes (more precisely, the direction of diversification is either the diagonal or the antidiagonal in the original phenotype space). Nevertheless, if diversification occurs along this composite axis, it also occurs in the two original axes. This shows that no matter what phenotypic interactions are assumed, they always generate robust diversification in 2D systems that are only on the brink of diversification without interactions.

It turns out that this is a very general and powerful mechanism for the maintenance of diversity. In dimensions higher than 2, the arguments are a bit more complicated, but it is also true that starting out with coordinates $x_1,\ldots,x_m$ such that $a_{ii} = k_{ii}$ for all i (that is, such that in each isolated phenotypic direction, the system is on the brink of diversification), the addition of interactions between phenotypes generates robust diversification. More precisely, with the addition of off-diagonal elements in A and K, there is a change of coordinates such that the corresponding quadratic forms $\hat{A}$ and $\hat{K}$ are diagonal and such that $\hat{a}_{ii} > \hat{k}_{ii}$ for at least one i, so that diversity is generated in at least one direction (and often many) of the new phenotypic coordinate system. The analytical arguments for the diversifying effects of interactions between phenotypic components provided in (*23*) are valid for the case $a_{ii} = k_{ii}$, but similar diversifying effects also occur in more general situations. These effects are independent of the sign of the off-diagonal elements in the matrices A and K and independent of whether only A, only K, or both A and K have off-diagonal elements.

More generally, if one starts out with phenotypic dimensions $x_1,\ldots,x_m$ for which $a_{ii} = k_{ii} - \varepsilon$ with $\varepsilon > 0$, so that variation is not maintained in any of the phenotypic components, the addition of off-diagonal elements in A and K can still lead to the maintenance of diversity (Fig. 2). For example, if $m = 2$, and if we assume randomly distributed off-diagonal elements in A and K in the range $(-\delta,\delta)$, then the probability that the system given by Eq. 6 exhibits diversity at equilibrium is $(1 - \varepsilon/2\delta)^2$ (*23*). Thus, even if the stabilizing component of selection dominates the frequency-dependent component of selection in each of the phenotype directions x_i, interactions between the x_i directions often generate diversity. Moreover, the higher the dimension m of phenotype space, the stronger is this effect, so that for very large m, interactions between the phenotypes can be weak and yet still generate diversity even if frequency dependence is very weak; and hence net selection is strongly stabilizing, in each of the phenotypic components (Fig. 2). Increasing the dimension of phenotype space not only makes the maintenance of diversity more likely, but also increases the amount of diversity maintained; that is, the variance of the equilibrium distributions.

In the Gaussian case, in which the functions appearing in the exponents of the competition kernel α and the carrying capacity K are quadratic, interactions between phenotypic components generate diversity that is represented as multivariate, unimodal Gaussian equilibrium distributions of the dynamics (Eq. 6), formu-

lated in coordinates that are linear combinations of the original phenotypic axes. These distributions have a single maximum and hence represent the maintenance of standing genetic variation, but not of distinct phenotypic clusters (Fig. 1). Projections of such equilibrium distributions onto the original phenotype axes would result in unimodal distributions with positive variance of at least two, and generally of many, of the original phenotypic components.

Adding interactions between phenotypes in non-Gaussian models, in which the functions appearing in the exponents of the competition kernel α and the carrying capacity K are not quadratic, has very similar effects and can also strongly promote the maintenance of diversity. For example, replacing the competition kernel (Eq. 4) and the carrying capacity (Eq. 5) by functions with quartic exponents

$$\alpha(x_1,\ldots,x_m;y_1,\ldots,y_m) = \exp\left[-\left((x-y)\cdot A\cdot(x-y)^T\right)^2\right] \quad (9)$$

$$K(x_1,\ldots,x_m) = K_0\exp\left[-\left(x\cdot K\cdot x^T\right)^2\right] \quad (10)$$

results in qualitatively very similar effects on the maintenance of diversity as shown in Fig. 2. However, in contrast to Gaussian models, in which diversity is maintained in the form of unimodal distributions with positive variance, in the non-Gaussian model resulting from Eqs. 9 and 10, diversification induced by phenotype interactions takes the form of multimodal equilibrium distributions representing different phenotypic clusters (Fig. 3).

The message from this analysis is simple and general. Suppose there are a number of phenotypes, each of which is under both stabilizing selection and frequency-dependent selection, so that frequency dependence is too weak to induce the maintenance of diversity along any of the phenotypic components in isolation. Then any interaction between phenotypes (that is, any departure from multiplicative determination of ecological properties) strongly increases the propensity for diversification. In essence, this is a consequence of the mathematics of simultaneous diagonalization of quadratic forms: When two quadratic forms with identical diagonal elements and arbitrary off-diagonal elements of comparable scale are simultaneously diagonalized so that one of the forms becomes the identity matrix, then the other quadratic form will have some diagonal elements that are larger than 1 and some that are smaller than 1. The sign of the off-diagonal elements, and hence the sign of the phenotypic interactions, does not matter for this effect, and the higher the dimension of phenotype space, the stronger the effect, so that for very high-dimensional phenotype spaces, phenotypic interactions can generate diversity even if frequency dependence is very weak in each component direction. The models we used are related to competition models with multidimensional resources (*24*), and the results are reminiscent of earlier findings that adding a spatial dimension to models of frequency-dependent selection can enhance the maintenance of diversity (*25*, *26*). Our results remain valid if the model given by Eq. 6 is altered to include mutations and/or sexual reproduction [for example, based on (*20*)]. In that case, the maintenance of substantial variation also requires frequency dependence, and hence phenotypic interactions in high-dimensional phenotype spaces also promote diversity. However, in sexual populations, the generation of diversity as multimodal phenotype distributions requires assortative mating (*20*). We have derived our results for classical competition models, but in essence the analysis only depends on the relative local curvatures at the maxima of the carrying capacity and the competition kernel and should therefore extend to any functions having a nondegenerate Taylor expansion at those maxima. It has been argued that competition models of the type considered here may be locally representative of a large class of models (*27*), and we conjecture that the effects observed here occur whenever frequency-dependent selection on high-dimensional phenotypes is determined by complex interactions between phenotypic components. This is different from the classical explanation for the maintenance of variation in continuous characters determined by multiple loci (*10*), which is based on mutation-selection balance, and which has been challenged repeatedly (*7*, *28*, *29*).

In real systems, phenotype space is often very high-dimensional, and interactions between phenotypic components in determining ecological properties must be ubiquitous. Therefore, the strength of frequency-dependent selection along each phenotypic component may be much less important for the maintenance of diversity than previously thought, and the conditions for the origin of diversity may be much more generally satisfied. According to our models, such diversity should often manifest itself through segregation along a composite direction in phenotype space by simultaneous and coordinated diversification of two or more phenotypic traits. The splitting of bacterial populations into subpopulations with anticorrelated metabolic networks (*30*) and of insect species into different host races with many correlated trait differentiations (*31*, *32*) are potential examples of such diversification in non-separable phenotype spaces. Our theory may help explain the extraordinary amounts of diversity found in marine microbial ecosystems (*33*–*35*) and in some microbial evolution experiments (*36*). Indeed, we think that microbes are good candidates for experimental tests of our models, as it is technically feasible to manipulate many metabolic pathways, such as those for resource acquisition and the secretion of toxins and other metabolites. In particular, it would be interesting to test whether at the genetic level, pathways controlling different traits become co-regulated to enable inheritability of composite, diversifying traits. In conclusion, if the complexity of high-dimensional phenotypes with interacting components is taken into account, the origin and maintenance of diversity may not be such a riddle after all.

References and Notes

1. J. L. Hubby, R. C. Lewontin, *Genetics* **54**, 577 (1966).
2. R. C. Lewontin, J. L. Hubby, *Genetics* **54**, 595 (1966).
3. B. C. Clarke, *Proc. R. Soc. London Ser. B* **205**, 453 (1979).
4. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer, Sunderland, MA, 2004).
5. U. Dieckmann, M. Doebeli, D. Tautz, J. A. J. Metz, Eds., *Adaptive Speciation* (Cambridge Univ. Press, Cambridge, 2004).
6. M. Kimura, *Proc. Natl. Acad. Sci. U.S.A.* **54**, 731 (1965).
7. M. Turelli, *Theor. Popul. Biol.* **25**, 138 (1984).
8. N. H. Barton, P. D. Keightley, *Nat. Rev. Genet.* **3**, 11 (2002).
9. S. Gavrilets, *Fitness Landscapes and the Origin of Species* (Princeton Univ. Press, Princeton, NJ, 2004).
10. R. Lande, *Genet. Res.* **26**, 221 (1975).
11. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, New York, 1983).
12. U. Dieckmann, M. Doebeli, *Nature* **400**, 354 (1999).
13. A. S. Kondrashov, F. A. Kondrashov, *Nature* **400**, 351 (1999).
14. R. H. Macarthur, R. Levins, *Am. Nat.* **101**, 377 (1967).
15. J. Roughgarden, *Theory of Population Genetics and Evolutionary Ecology: An Introduction* (Macmillan, New York, 1979).
16. M. L. Taper, T. J. Chase, *Ecology* **66**, 355 (1985).
17. M. Kirkpatrick, N. H. Barton, *Am. Nat.* **150**, 1 (1997).
18. C. Matessi, K. A. Schneider, *Theor. Popul. Biol.* **76**, 1 (2009).
19. M. Gyllenberg, G. Meszéna, *J. Math. Biol.* **50**, 133 (2005).
20. M. Doebeli, H. J. Blok, O. Leimar, U. Dieckmann, *Proc. Biol. Sci.* **274**, 347 (2007).
21. E. Hernandez-García, C. Lopez, S. Pigolotti, K. Andersen, *Theor. Ecol.* 10.1007/s12080-009-0056-2 (2009).
22. S. Pigolotti, C. Lopez, E. Hernandez-García, K. H. Andersen, *Philos. Trans. R. Soc. London Ser. A* **367**, 183 (2009).
23. See supporting material on *Science* Online.
24. F. B. Christiansen, V. Loeschcke, *Theor. Popul. Biol.* **31**, 33 (1987).
25. M. Doebeli, U. Dieckmann, *Nature* **421**, 259 (2003).
26. R. Bürger, *Theor. Popul. Biol.* **76**, 214 (2009).
27. M. Durinx, J. A. J. Metz, G. Meszéna, *J. Math. Biol.* **56**, 673 (2008).
28. R. Bürger, *J. Math. Biol.* **24**, 341 (1986).
29. P. C. Phillips, *Genet. Res. Camb.* **89**, 371 (2007).
30. M. Le Gac *et al.*, *Genetics* **178**, 1049 (2008).
31. M. C. Singer, C. S. McBride, *Evolution* 10.1111/j.1558-5646.2009.00866.x (2009).
32. A. Frantz, V. Calcagno, L. Mieuzet, M. Plantagenest, J. Simon, *Biol. J. Linn. Soc.* **97**, 718 (2009).
33. S. Yooseph *et al.*, *PLoS Biol.* **5**, e16 (2007).
34. J. A. Huber *et al.*, *Science* **318**, 97 (2007).
35. D. E. Hunt *et al.*, *Science* **320**, 1081 (2008).
36. R. Maharjan, S. Seeto, L. Notley-McRobb, T. Ferenci, *Science* **313**, 514 (2006).

37. The order of authors is alphabetical. We thank K. Schneider for improving some of the mathematical arguments and D. Weigel, H. Metz, and S. Otto for discussions and comments. This research was supported by the Natural Sciences and Engineering Research Council (Canada) and by the Human Frontier Science Program (European Union).

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5977/494/DC1
SOM Text
Fig. S1
References

25 January 2010; accepted 12 March 2010
10.1126/science.1187468

Supporting Online Material for

Complexity and Diversity

Michael Doebeli* and Iaroslav Ispolatov*

*To whom correspondence should be addressed. E-mail: doebeli@zoology.ubc.ca (M.D.); slava@math.ubc.ca (I.I.)

Published 23 April 2010, *Science* **328**, 494 (2010)
DOI: 10.1126/science.1187468

This PDF file includes:

Complexity and Diversity: Supplementary Online Material

Michael Doebeli & Iaroslav Ispolatov

Department of Zoology and Department of Mathematics

University of British Columbia, 6270 University Boulevard

Vancouver B.C. Canada, V6T 1Z4

1. Different forms of maintenance of variation in 1-dimensional systems

It is well-known that if $a > k$ in the Gaussian model (1) with $n_\alpha = n_k = 2$ in (2) and (3), then the dynamics given by (1) has a Gaussian equilibrium distribution of the form

$$\phi^*(x) = \frac{aK_0}{\sqrt{\pi(a-k)}} \exp\left[-\frac{ak}{a-k}(x-y)^2\right]. \tag{S.1}$$

This follows directly from

$$K(x) = \int \alpha(x,y)\phi^*(y)dy, \tag{S.2}$$

where $\alpha(x,y)$ and $K(x)$ are given by (2) and (3). The equilibrium (S.1) is stable *(1)*, and it disappears for $a < k$, in which case the dynamics (1) converges to a Dirac Delta function centred at $x = 0$, i.e., at the maximum of the carrying capacity function K.

If $n_\alpha = n_K = 4$ in (2) and (3), model (1) again converges to a Delta function centred at $x = 0$ if a is small enough compared to k. Again, variation is maintained as a increases above a critical threshold, but not in the form of a continuous unimodal equilibrium distribution as in the Gaussian case, but instead in the form of multiple Delta peaks corresponding to distinct phenotypic clusters, as illustrated in Figure 1. Interestingly, the threshold value of a for which the transition between regimes with and without variation occurs is also different from the Gaussian case. With $n_\alpha = n_K = 4$, this threshold is approximately $a = 0.25k$, so that diversity is in fact

easier to maintain in the non-Gaussian model (i.e., already occurs for values of a that are smaller than k, in contrast to the Gaussian case, where diversity occurs if and only if $a > k$).

We note that if the matrices A and K are diagonal in the general, multi-dimensional model (6) with Gaussian competition kernel and carrying capacity, model (6) becomes separable into a set of independent 1-dimensional Gaussian models, for in this case, the dynamics (6) admits a multivariate Gaussian equilibrium distributions, in which the variance in each phenotypic dimension is given as in the 1-dimensional model described above, i.e., the variance is 0 in directions i in which $a_{ii} < k_{ii}$, and that variance is $(a_{ii} - k_{ii})/(2a_{ii}k_{ii})$ in directions i in which $a_{ii} > k_{ii}$ (where a_{ii} and k_{ii} are the diagonal elements of A and K).

2. Interactions between phenotypic components generate robust diversification in systems that are poised on the brink of diversification without interaction

Here we show that for the Gaussian model (6) given by $m \times m$ positive definite symmetric matrices A and K for the competition kernel and the carrying capacity with diagonal elements $a_{ii} = k_{ii}$, $i = 1, \ldots, m$, and arbitrary off-diagonal elements of which at least one is non-zero, there is a (linear) transformation of the coordinates such that in the transformed coordinates, model (6) is given by diagonal matrices $\hat{A}$ and $\hat{K}$ such that $\hat{a}_{ii} > \hat{k}_{ii}$ for at least one index i.

According to standard theory *(2)*, the quadratic forms A and K can be simultaneously diagonalized as follows. First, there is an orthogonal transformation Q that diagonalizes A, i.e., $QAQ^T = \tilde{A}$, where $\tilde{A}$ is a diagonal matrix with entries the eigenvalues $\lambda_1, \ldots, \lambda_m$ of A. (and where Q^T denotes the transposed matrix). Again according to general rules of linear algebra, we have $w \equiv tr(A) = tr(\tilde{A}) = \sum_{i=1}^{m} \lambda_i$, where tr denotes the trace, i.e., the sum of the diagonal element. Let $\tilde{K} = QKQ^T$ be the symmetric matrix determining the carrying capacity in the new coordinates,

and let ρ_i be the diagonal elements of $\tilde{K}$ (note that $\tilde{K}$ is in general not diagonal). We again have $tr(\tilde{K}) = \sum_{i=1}^{m} \rho_i = w$.

By dividing the ith new coordinate axis by $\sqrt{\lambda_i}$, $i = 1, \ldots, m$, we get a new coordinate system in which the competition kernel is determined by a diagonal matrix $\hat{A}$ with all diagonal entries equal to 1, i.e., in the new coordinate system the competition kernel is now given by the identity matrix. On the other hand, the carrying capacity is determined by a new matrix $\bar{K}$ with diagonal entries ρ_i/λ_i, $i = 1, \ldots, m$. Generically (i.e., for almost all choices of the off-diagonal elements in the original matrices A and K), we have $\rho_i \neq \lambda_i$ for at least one index i, and it follows that there is at least one index i_0 such that $\rho_{i_0}/\lambda_{i_0} < 1$. For if $\rho_i/\lambda_i \geq 1$ for all i with at least one strict inequality, then $\rho_i \geq \lambda_i$ for all i with at least one strict inequality, and hence $\sum_{i=1}^{m} \rho_i > \sum_{i=1}^{m} \lambda_i = w$, in contradiction to what was said above. Let i_0 be the index for which $\rho_{i_0}/\lambda_{i_0} < 1$ is smallest.

We then chose an orthogonal transformation S that diagonalizes the matrix $\bar{K}$, i.e., $S\bar{K}S^T = \hat{K}$ is a diagonal matrix with entries the (ordered) eigenvalues $\eta_1 \leq \ldots \leq \eta_m$ of $\bar{K}$. Consider the map $\boldsymbol{x} \mapsto \boldsymbol{x}\hat{K}\boldsymbol{x}^T = \sum_{i=1}^{n} \eta_i x_i^2$. This map is continuous on the unit sphere and therefore attains a minimum. Clearly, the minimum is attained at unit vector $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$. Since $\|S\boldsymbol{e}_j\| = \|\boldsymbol{e}_j\| = 1$ for all j, we have in particular $\eta_1 = \boldsymbol{e}_1\hat{K}\boldsymbol{e}_1^T \leq (S\boldsymbol{e}_{i_0})\hat{K}(S\boldsymbol{e}_{i_0})^T = \boldsymbol{e}_{i_0}\bar{K}\boldsymbol{e}_{i_0}^T = \rho_{i_0}/\lambda_{i_0} < 1$. On the other hand, in the new coordinate system defined by S, the competition kernel is still given by the identity matrix. Thus, we have found a coordinate system in which the competition matrix is given by the identity matrix, and the carrying capacity is given by a diagonal matrix with at least one diagonal element < 1. In this coordinate system, model (6) will therefore generate diversification in at least that direction. This shows that starting from a model in which each phenotypic component is exactly on the boundary of diversification (so that $a_{ii} = k_{ii}$ in the original quadratic forms A and K), adding any type of interactions between the phenotypic components in

determining the competition kernel and the carrying capacity by introducing non-zero off-diagonal elements in A and K generically leads to robust diversification in at least one (composite) phenotypic direction.

3. Probability of diversification in 2-dimensional phenotype spaces when $\mathbf{a_{ii} = 1 - \epsilon}$ and $\mathbf{k_{ii} = 1}$

Consider quadratic forms A and K of the form

$$A = \begin{pmatrix} 1-\epsilon & u \\ u & 1-\epsilon \end{pmatrix} \qquad K = \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix} \tag{S.4}$$

where u and v are uniformly distributed between $-\delta$ and δ, and where $\epsilon > 0$. Without interactions between phenotypes, i.e., when $u = v = 0$, the equilibrium distribution of the corresponding dynamics (6) has zero variance in both phenotypic directions. The matrices A and K have eigenvalues $1-\epsilon \pm u$ and $1 \pm v$ and can be simultaneously diagonalized by the linear transformation of coordinates,

$$\begin{aligned} y_1 &= (x_1 - x_2)/\sqrt{2} \\ y_2 &= (x_1 + x_2)/\sqrt{2}. \end{aligned} \tag{S.5}$$

It immediately follows that when $v = 0$, one eigenvalue of A is larger than the corresponding eigenvalue of K (so that diversification occurs in that direction) whenever $|u| > \epsilon$. Similarly, when $u = 0$, one eigenvalue of A is larger than the corresponding eigenvalue of K whenever $|v| > \epsilon$. More generally, the condition for diversification that one eigenvalue of A is greater than the corresponding eigenvalue of K translates into satisfying one of the two inequalities,

$$\begin{aligned} u &> v + \epsilon, \\ u &< v - \epsilon. \end{aligned} \tag{S.6}$$

Geometrically, the probability that a uniformly randomly chosen pair of u, v in the interval $[-\delta, \delta]$ satisfies the first inequality in (S.6) is the area of the right triangle

which lies below the line $v = u - \epsilon$ in the $[-\delta, \delta; -\delta, \delta]$ square, divided by normalization, that is, $(2\delta - \epsilon)^2/8\delta^2$. Similarly, the probability to satisfy the second inequality in (S.6) is the area of the right triangle which lies above the line $v > u + \epsilon$ in the same $[-\delta, \delta; -\delta, \delta]$ square, also divided by normalization, which is equal to $(2\delta - \epsilon)^2/8\delta^2$ as well. Consequently, the total probability for diversification is

$$\left(1 - \frac{\epsilon}{2\delta}\right)^2 \tag{S.7}$$

We note that the probability that a given pair (u, v) satisfies at least one of the inequalities (S.6) can be calculated for more general probability distributions $\rho(U)$ and $\sigma(V)$ on the interval $[-\delta, \delta]$ as

$$P = \int_{-\delta}^{\delta} \sigma(V) dV \left(\int_{min[\delta, v+\epsilon]}^{\delta} \rho(U) dU + \int_{-\delta}^{max[-\delta, v-\epsilon]} \rho(U) dU \right). \tag{S.8}$$

For uniform distributions $\rho(U) = \sigma(V) = 1/(2\delta)$, (S.8) recovers (S.7). Clearly, $(1 - \epsilon/2\delta)^2$ is thus the probability that uniformly randomly chosen u and v in the interval $[-\delta, \delta]$ lead to diversification in model (6) if incorporated as off-diagonal elements in the quadratic forms A and K as in (S.4). In particular, if one starts out with two phenotypic components along each of which stabilizing selection dominates frequency-dependent selection (by an amount determined by ϵ), then, independent of the sign of the interactions, the stronger the interactions between the phenotypic components determining the competition kernel and the carrying capacity (i.e., the larger δ), the higher the probability that such interactions lead to diversification.

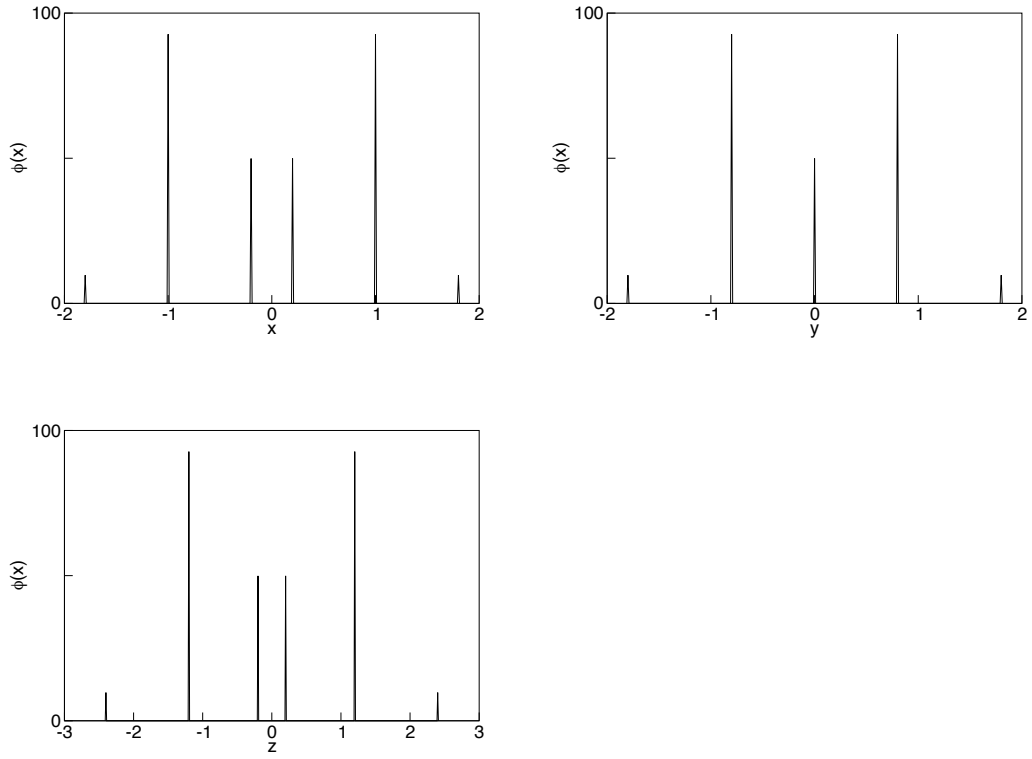

Figure S1: Projection of the equilibrium distribution of the 3-dimensional non-Gaussian model (9), (10) shown in Figure 3 (main text) onto the three axes.

References

1. Pigolotti, S., Lopez, C., Hernandez-García, E., and Andersen, K. H.Species competition: coexistence, exclusion and clustering. Philosophical Transactions of the Royal Society A 367, 183-3195, 2009.

2. Greub, W. H. Linear Algebra (Fourth Edition). Springer, New York, 1981.