File S2

Approximations for the DFE

We know that the limit spectral distribution of $\mathbf{M}$ is approximately given by the M-P law under reasonable assumptions (see **Appendix file S1**). Let us see the implications for the distribution of the fitness effect of mutations. Several of the analytical treatments can be checked using the Mathematica® (Wolfram Research 2012) notebook **file S3** (in freely readable [.cdf] format).

Stochastic representation of the DFE in the anisotropic FGM: Eq. (2) provides the stochastic representation of s as a function of each individual value of λ_i. Its derivation can be found elsewhere (Mathai and Provost 1992; Jaschke et al. 2004). It can be obtained in the basis of phenotype space where the mutational covariance is diagonal, via the change of basis $\mathbf{z} = \mathbf{Q}.\mathbf{y}$, where $\mathbf{Q}$ is the eigenbasis of $\mathbf{M}$. In this basis, the random variables dy_i in eq. (1) become dz_i, a set of n independent normal deviates: $dz_i \sim N\left(0, \sqrt{\lambda_i}\right)$. Eq. (1) can thus be rearranged into

$$s = \sum_{i=1}^{n} \frac{z_i^2}{2} - \sum_{i=1}^{n} \frac{\lambda_i}{2} \left(\frac{z_i}{\sqrt{\lambda_i}} + \frac{dz_i}{\sqrt{\lambda_i}} \right)^2 \quad , \qquad \textbf{(A2.1)}$$

where the variables $dz_i/\sqrt{\lambda_i} \sim N(0,1)$ are independent standard normal deviates. By definition of the non-central chi-square distribution: $\left(z_i/\sqrt{\lambda_i} + dz_i/\sqrt{\lambda_i}\right)^2 \sim \chi_1^2[z_i^2/\lambda_i]$, which leads to eq. (2). From this stochastic representation, one can directly obtain the approximations in eqs. (8) in the limit where all $\lambda_i = \tilde{\lambda}$ (or eq. (10) when only the $n-1$ lowest eigenvalues are equal to $\tilde{\lambda}$).

However, it is useful to study the distribution in more details, to understand why this isotropic approximation is in fact robust, considering that the eigenvalues are never exactly all equal. I do so via a generating function of the DFE.

Generating function of the DFE: the stochastic representation in eq. (2) directly yields a closed form expression for the cumulant-generating-function (CGF) of s : $\kappa_s(u) \equiv \log(E_s(e^{u\,s}))$. This CGF fully characterizes the DFE (all its cumulants). It is derived easily from the CGF of the non-central chi-square $\chi_1^2[v]$, which is $\kappa_1(u,v) = u\,v/(1-2\,u) - \log(1-2\,u)/2$. The CGF of the sum of independent variables is the sum of their CGFs, so eq. (2) yields

$$\kappa_s(u) = \frac{1}{2}\left(\sum_{i=1}^{n} \frac{u^2\, z_i^2 \lambda_i}{1+u\,\lambda_i} - \sum_{i=1}^{n} \log(1+u\,\lambda_i) \right) \quad . \qquad \textbf{(A2.2)}$$

As the CGF fully characterizes a distribution, the DFE is fully determined by the joint distribution of the eigenvalues of $\mathbf{M}$ (the λ_i's) and the position of the parental phenotype (the y_i's).

Link to the Shannon transform of the M-P law: At this point, we can note that a central quantity in eq. (A2.2) is $\nu_n(u) = 1/n \sum_{i=1}^{n} \log(1 + u\,\lambda_i)$, the average of $\log(1 + u\,\lambda)$ over the n eigenvalues, namely the Shannon transform of the spectral distribution of $\mathbf{M}$. For an optimal genotype ($s_o = 0$) all $z_i = 0$ and $\kappa_s(u) = -n/2\,\nu_n(u)$. For a suboptimal genotype, it is impossible to derive an equivalent expression. However, below phase transition ($\alpha \approx 1$), we may ignore any potential correlation, across traits i, between the maladaptation terms z_i^2 and the λ_i. Then, we can, first, introduce the derivatives of $\log(1 + u\,\lambda_i)$ with respect to u into $\kappa_s(u)$, then approximate

$$\kappa_s(u) = -\frac{n}{2}\,\nu_n(u) + \sum_{i=1}^{n} \frac{u^2\,z_i^2}{2}\,\partial_u \log(1 + u\,\lambda_i) \approx -\frac{n}{2}\,\nu_n(u) + u^2 \left(\frac{1}{n}\sum_{i=1}^{n} \frac{z_i^2}{2}\right)\left(\partial_u \sum_{i=1}^{n} \log(1 + u\,\lambda_i)\right)$$

Noting that $1/n \sum_{i=1}^{n} z_i^2/2 = s_o/n$ and that $\partial_u(\sum_{i=1}^{n} \log(1 + u\,\lambda_i)) = n\,\nu_n'(u)$, we thus get

$$\kappa_s(u) \approx u^2 s_o \nu_n{}'(u) - \frac{n}{2}\nu_n(u) \quad , \qquad \textbf{(A2.3)}$$

provided that $cov(z_i^2, \lambda_i) \approx 0$ and $\alpha \approx 1$. Therefore, the DFE obtains in terms of the spectral distribution of $\mathbf{M}$ (via $\nu_n(u)$), plus two parameters: s_o and n. The effect of maladaptation ($\mathbf{y}$) on the DFE is thus fully determined by the distance to the optimum ($s_o = \|\mathbf{y}\|^2/2$), not by the actual direction to the optimum: the model behaves *de facto* as an isotropic one. Of course the independence assumed between z_i^2 and λ_i across traits i, is never guaranteed, but it must become an accurate approximation once the λ_i become close to each other (convergence towards more isotropy as $\beta \to \infty$). Note also that (A2.3) is always exact for an optimal genotype ($s_o = 0$).

We can derive the equivalent expression beyond phase transition by simply separating the leading eigenvalue ($\lambda_1 = \alpha\,\tilde{\lambda}$) from the others: $\kappa_s(u) \approx \kappa_{n-1}(u) + \kappa_1(u)$ with

$$\begin{cases} \kappa_{n-1}(u) = u^2 s_{n-1} \nu_{n-1}{}'(u) - \dfrac{n-1}{2}\nu_{n-1}(u) \\ \kappa_1(u) = u^2 s_1 \nu_1{}'(u) - \dfrac{1}{2}\nu_1(u) \end{cases} \quad . \qquad \textbf{(A2.4)}$$

where $s_1 = z_1^2/2$ and $s_{n-1} = \sum_{i=2}^{n} z_i^2/2 = s_o - s_1$, while $\nu_{n-1}(u)$ is the average $\log(1 + u\,\lambda)$ over the $n-1$ smallest eigenvalues and $\nu_1(u) = \log(1 + \lambda_1 u)$ with $\lambda_1 = \alpha\,\tilde{\lambda}$ (eq. (6)).

The general formula above is not fully determined because $\nu_n(u)$ is an average, which varies as n eigenvalues vary randomly. To overcome this, we simply take a limit when $n, p \to \infty$, in which case (i) the distribution of λ_i converges to a non-random limit (the M-P law approximately) and (ii) the expectations $\nu_n(u)$ and$\nu_{n-1}(u)$both converge to the Shannon transform of this M-P law as $n \to \infty$:

$$\nu_n(u) \underset{n,p\,\to\,\infty}{\to} E(\log(1 + u\,\lambda)) = \nu_{\mathbf{M}}(u) = \int_0^{\infty} \tilde{\rho}_{\mathbf{H}.\mathbf{H}^*}(\lambda)\,\log(1 + u\,\lambda)\,d\lambda \approx \nu_{\mathbf{H}.\mathbf{H}^*}(u) \quad , \qquad \textbf{(A2.5)}$$

in the context of the M-P law approximation, $\nu_{\mathbf{M}}$ is approximately given by eq. (A1.3), with effective ratio index $\beta_e = p_e/n$ and effective scale $\zeta_e = \tilde{\lambda}/\beta_e$ as given by eq. (4).

Moments of the DFE: Based on the limit obtained for $\nu_n(u)$ (eq. (A2.5)), we obtain a non-random limit for the CGF via eq. (A2.4). Taking the derivatives of $\kappa_s(u)$ with respect to u taken at $u = 0$ yields the cumulants of the distribution. In particular, defining $(n - 1 + \alpha)\,\tilde{\lambda}/2 = \bar{s}$, $\theta = n/(n - 1 + \alpha) - 1$, $\epsilon_o = s_o/\bar{s}$ and $\epsilon_1 = s_1/\bar{s}$, we obtain, for the mean and variance of the DFE:

$$\begin{cases} E(s) = \kappa_s{}'(0) = -\bar{s} = -\,\zeta_{\mathbf{W}}\,p/2 \\ V(s) = \kappa_s''(0) = \dfrac{2\,\bar{s}^2}{n}\,\dfrac{(n-1)(1+\theta)^2 + p_e(1 + (n-1)\theta^2 - 2\,n\,\theta\,\epsilon_1 + 2(1+\theta)\epsilon_o)}{p_e} \end{cases} , \quad \textbf{(A2.6)}$$

We obtain a simpler expression below phase transition ($\alpha = 1$ so that $\theta = 0$ and $\bar{s} = n\,\tilde{\lambda}/2$), especially when taking the leading order in $n \gg 1$:

$$V(s) \underset{\alpha \to 1}{\to} \frac{2\,\bar{s}^2}{n}\left(\frac{1}{\beta_e} + (1 + 2\,\epsilon_o)\right) + o\left(\frac{1}{n}\right) \qquad \cdot \qquad \textbf{(A2.7)}$$

It can be checked that eq. (A2.7) yields the exact expression from the purely isotropic FGM (see Martin and Lenormand 2006) whenever $\beta_e \gg 1$.

Isotropic approximation: The isotropic approximation that we use in the main text consists in equating the $n - 1$ lowest eigenvalues $\lambda_{i>1}$ to a constant $\tilde{\lambda}$, the mean of the M-P law, which amounts to setting $p_e \to \infty$. The expressions above do not make such an assumption; they only rely on the convergence to the M-P law for the spectral distribution of $\mathbf{M}$ (eq.(A2.5)), and on ignoring any potential correlation between z_i^2 and λ_i (eqs. (A2.3) and (A2.4)). However, to characterize the DFE more explicitly (stochastic representation or pdf), it proves critical to further rely on the isotropic approximation. This approximation proves accurate even though the actual system is clearly anisotropic, as illustrated on **Figure 5.a**, where the spectrum of $\mathbf{M}$ is quite spread. A tentative explanation for this robustness can be proposed, below phase transition (when $\alpha = 1$). Beyond phase transition ($\alpha > 1$), the problem boils down to whether the isotropic approximation is accurate in the eigenspace associated with the $n - 1$ lower eigenvalues, so it is an equivalent issue.

Even when $\alpha = 1$, the actual model is of course never isotropic. The ratio index β must be finite, and because of metabolic correlations in $\mathbf{W}$, the equivalent ratio β_e can be substantially smaller than β (eq. (4)). **Figures 3 and 4** confirm that even with relatively large β, the spectral distribution of $\mathbf{M}$ shows substantial variance, in a manner captured by the M-P law approximation. More precisely, the coefficient of variation of the eigenvalues of $\mathbf{M}$ is approximately $CV(\lambda_i) = 1/\sqrt{\beta_e}$. Therefore, the eigenvalues λ_i are not equal and phenotypic directions are not equivalent. However, this anisotropy remains mild, and proves to have approximately no influence on the DFE, as long as p_e is large enough. This can be understood by looking at the CGF and its approximate expression in eq. (A2.3).

The pdf of a distribution can be obtained as an inverse Fourier transform of the characteristic function of this distribution. In our context, this characteristic function is given by $\psi(t) = e^{\kappa_s(\mathbf{i}\,t)}$, where $\kappa_s(.)$ is the CGF given in eq. (A2.2) when $\alpha = 1$ and $\mathbf{i}$ is the unit complex number ($\mathbf{i}^2 = -1$). Therefore, to find a suitable approximation for the pdf of s one must approximate $e^{\kappa_s(\mathbf{i}\,t)}$. When we can ignore correlations between z_i^2 and λ_i or at the optimum, $\kappa_s(.)$ is approximately given by eq. (A2.3). Part of the anisotropy then vanishes already: only the distance to the optimum s_o has an impact, not its direction. However, to obtain the isotropic approximation exactly (n traits all equivalent) still requires to seek an approximation for $\nu_{\mathbf{H}.\mathbf{H}^*}(u)$, more precisely for $\psi(t) = e^{-n/2\,\nu_{\mathbf{H}.\mathbf{H}^*}(\mathbf{i}\,t)}$at the optimum. The corresponding expression in the isotropic model (all $\lambda_i = \tilde{\lambda}$) is simply $\psi_{iso}(t) = \left(1 + \tilde{\lambda}\,\mathbf{i}t\right)^{-n/2}$ which is the characteristic of a negative gamma distribution $s \sim -\Gamma(n/2, \tilde{\lambda})$. The characteristic function, at the optimum ($s_o \to 0$), is equal to this isotropic approximation, to leading order in $\zeta_e = \tilde{\lambda}/\beta_e$. Indeed, recalling that $\tilde{\lambda} = 2\bar{s}/n$ and $\beta_e = p_e/n$, the ratio between the exact and approximate characteristic functions satisfies

$$\frac{\psi(t)}{\psi_{iso}(t)} \approx \left(1 - \frac{t^2\,\bar{s}^2}{p_e}\right) + o\left(\frac{\bar{s}^2}{p_e}\right) \quad , \qquad \textbf{(A2.8)}$$

The relative error in equating $\varphi(t) \approx \varphi_{iso}(t)$ is thus small under fairly mild conditions. Even when $\beta_e = p_e/n$ is not very large, so that variation across λ_i is substantial, it suffices that p_e be large enough and that mutation effects be mild enough ($\bar{s}^2/p_e \ll 1$) for the isotropic approximation to perform satisfyingly. This accuracy of the isotropic approximation is illustrated in **Figure S1** where $\nu_{\mathbf{H}.\mathbf{H}^*}(u)$ is compared to its equivalent in the isotropic approximation $\nu_{iso}(u) = \log(1 + u\,\tilde{\lambda})$.

This whole argument is merely intuitive as it relies on approximate results, but it does give an intuition on why, even when the spectral distribution of $\mathbf{M}$ is fairly spread (e.g. **Figure 5**.**a**), eqs. (8-11) prove accurate.

References:

Jaschke, S., C. Kluppelberg, and A. Lindner. 2004. Asymptotic behavior of tails and quantiles of quadratic forms of Gaussian vectors. J. Multivar. Anal. 88:252-273.

Martin, G. and T. Lenormand. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60:893-907.

Mathai, A. M. and S. B. Provost. 1992. Quadratic forms in random variables. Marcel Dekker, New York.

Wolfram Research, I. 2012. Mathematica Edition: Version 9.0. Wolfram Research, Inc., Champaign, Illinois.