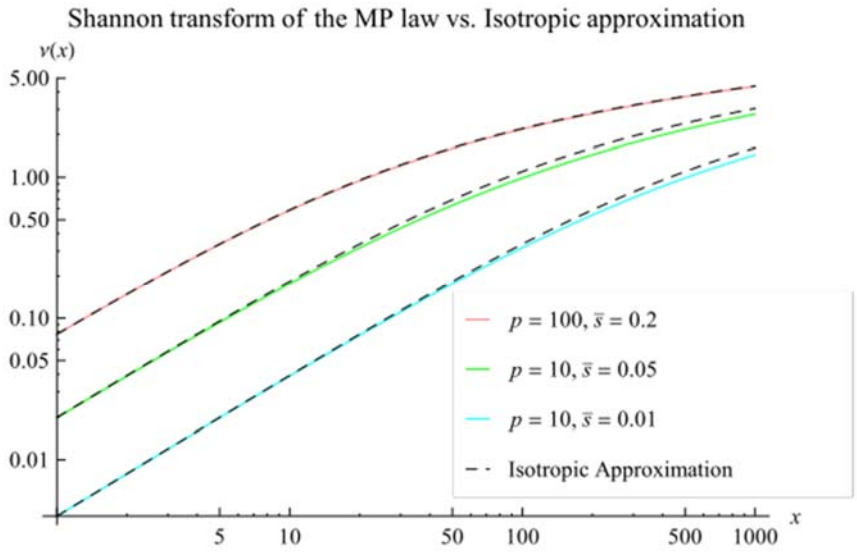

Figure S1 Convergence of the Shannon transform to the isotropic approximation

The Shannon transform of the M-P law ($v_{H.H^*}(x)$ with $\lambda_{\mathrm{H.H^*}} \sim MP(\beta,\zeta)$, see eq. A1.3) is compared to the isotropic approximation ($v(x) \approx \log(1+\tilde{\lambda}\,x)$) where $\tilde{\lambda} = \beta\zeta = 2\bar{s}/n$ where $\bar{s} = -E(s)$ is the mean (deleterious) effect of mutations. The parameters are indicated on the graph, with $n = 5$ and recalling that $\beta = p/n$ and $\zeta = 2\bar{s}/p$. The isotropic approximation proves accurate over a wide range of parameter values.

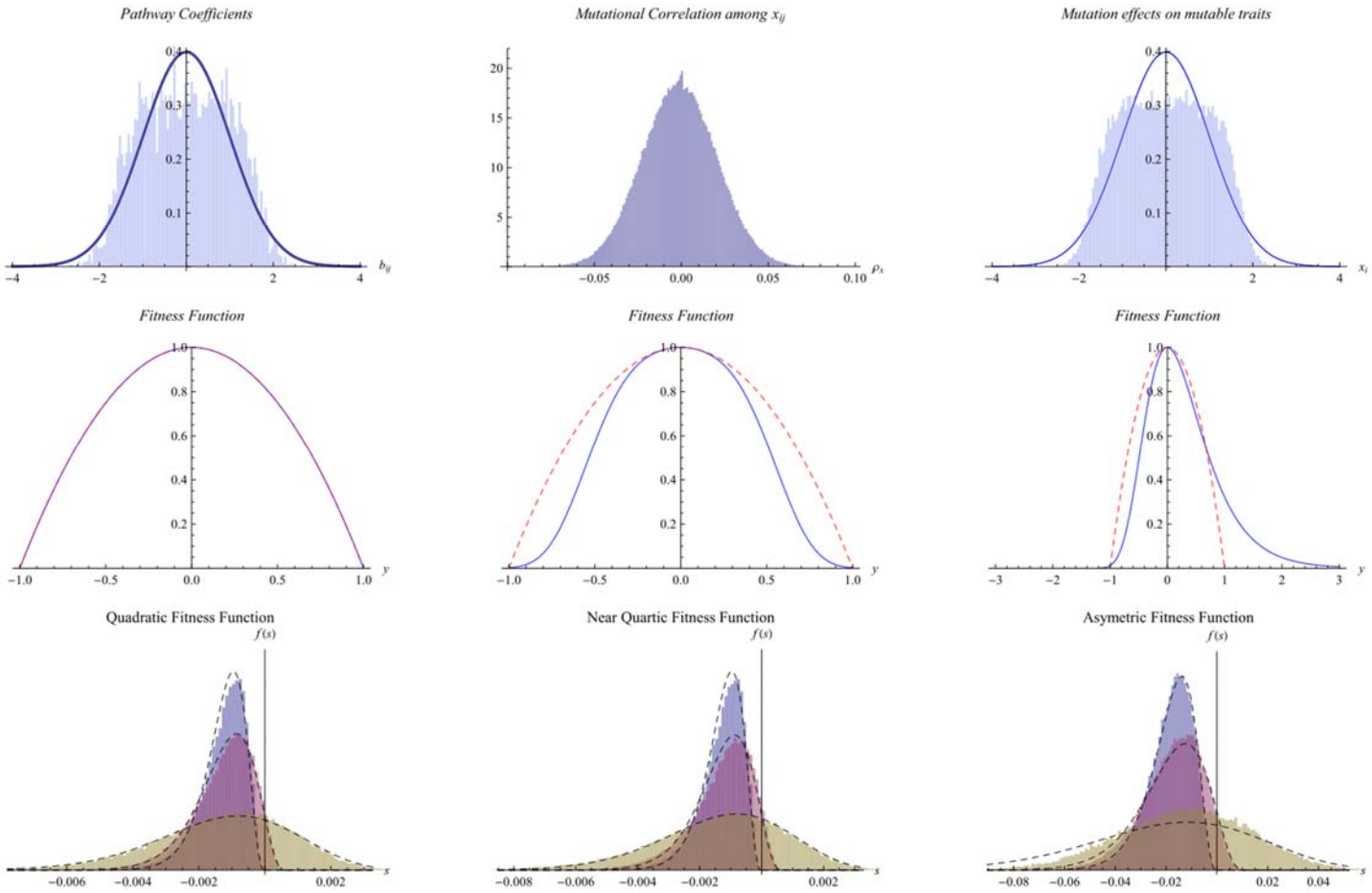

Figure S2 Distribution of mutation fitness effects with various fitness functions

The DFE in simulations (below phase transition, as in **Figure 5.c**) is compared to the analytic prediction (eq. 9) for three choices of fitness function for which the quadratic approximation applies (detailed in **Figure 2.a.**): quadratic, near quartic and asymmetric. The upper panels show various distributions common to all simulations: pathway coefficients, correlations between mutation effects on mutable traits $\mathbf{x}$ and mutation effects on $\mathbf{x}$ themselves. The middle panels show the three alternative fitness functions. The bottom panels show the corresponding DFE in simulations (histograms) and the corresponding prediction (dashed lines, eq. (9)) for various distances to the optimum (same as **Figure 5**: $s_o = 0$, blue, $s_o = \bar{s}$, red, $s_o = 5\bar{s}$, brown).

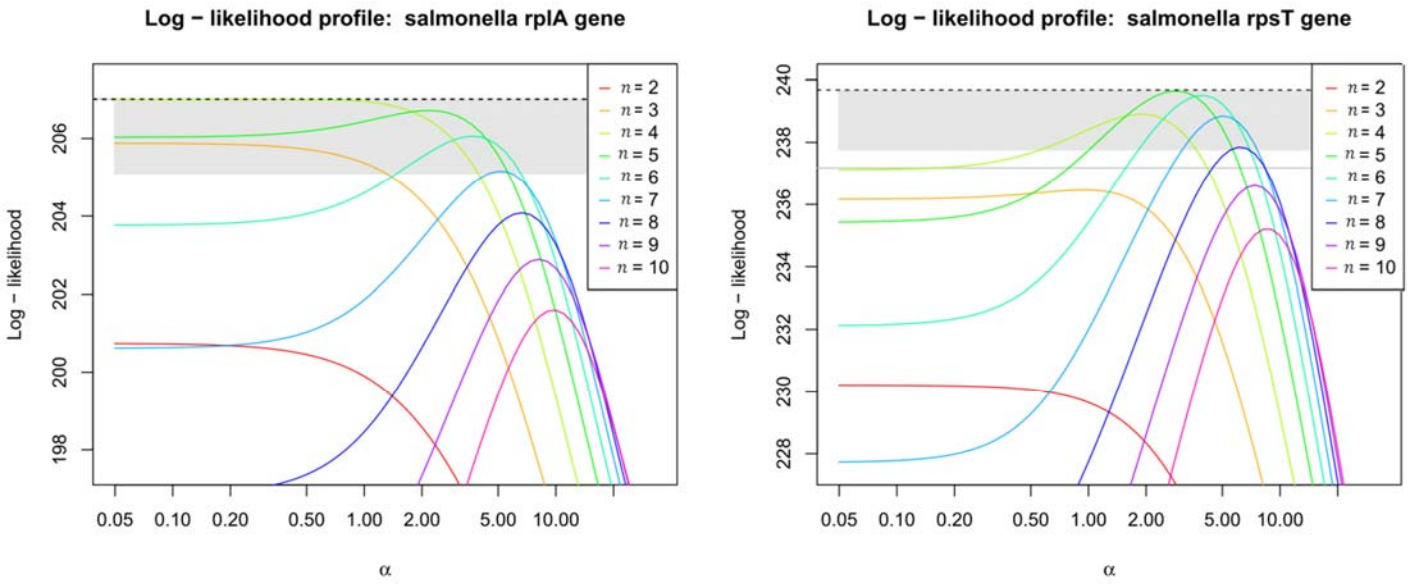

Figure S3 Log – likelihood profile for the parameter α from empirical DFEs

For each gene, the log-likelihood of the fitted model (**Figure 6** and **Table 2**) is given as a function of the parameter α, for various integer values of n (see legend) and with the scale λ set to match the mean of the distribution ($\lambda = 2\bar{s}/(n-1+\alpha)$). The dashed black line gives the log – likelihood of the best fitting gamma sum model (eq. (11) with n, α, λ jointly fitted) and the gray plain line gives the log – likelihood of the best fitting pure gamma model (i.e. setting $\alpha = 0$). The shaded area corresponds to log – likelihood lying within 1.92 points of the maximum log – likelihood for the best fitting model (with α fitted), implying no significant difference in the fit to the data. In the *rplA* gene, the pure gamma models with $n = 3,4,5$ or 6 are not statistically different from the best fitting model (**Table 3**): their maxima lie within the shaded area. For the *rpsT* gene, the 'gamma sum' model (eq. (11), with $\alpha > 0$) provides a significantly better fit than a pure gamma model with $\alpha = 0$ (the gray line lies below the gray shaded area). It further shows that the best fitting model (**Table 3**) and models with $n = 4,5,6$ or 7, with corresponding (increasing) values of $\alpha > 0$, cannot be distinguished in terms of goodness of fit (all their maxima lie within the shaded area).

File S1

Approximations to the LSD for some classes of random matrices

In this appendix, I first state several results on the Limit Spectral Distribution (LSD) of some large random covariance matrices. This LSD is the nonrandom limit of the distribution of the eigenvalues of a random matrix, as its dimensions get large (here as $n, p \to \infty$). The first section merely states various known results from Random matrix Theory: most are summarized in (TULINO and VERDÙ 2004). Then I derive an approximation for the LSD of $\mathbf{M}$ in the integrative phenotypic network model presented in the main text (**Figure 1**), and for key quantities related to the mutant fitness distribution. In what follows, I refer to the pdf of the spectral distribution of a given matrix $\mathbf{X}$ as $\rho_{\mathbf{X}}$ and to the corresponding Limit Spectral distribution (LSD) of this matrix as $\tilde{\rho}_{\mathbf{X}}$. I refer to any of the transforms of this LSD by an index referring to that matrix. All the derivations of this Appendix can be checked in a Mathematica® (WOLFRAM RESEARCH 2012) notebook (**file S3**, in freely readable [.cdf] format) available for download at **XXX**.

A) The Marchenko-Pastur law and its properties

Limiting Spectral Distribution (LSD): Central to our derivations is the Marchenko - Pastur (M-P) law, which provides the LSD of a covariance matrix with random entries. Consider a $n \times p$ matrix $\mathbf{H}$ whose entries h_{ij} are randomly and independently drawn from possibly different distributions. These distributions are arbitrary in nature but have mean zero, variance $1/n$ and fourth moments of order $O(1/n^2)$, i.e. they are 'standard' zero mean distributions, not too leptokurtic (normal, extreme value, uniform, etc.). Then, if $\zeta > 0$ is an arbitrary scale factor, the random $n \times n$ matrix $\zeta\, \mathbf{H}.\mathbf{H}^*$ is called a sample covariance matrix. Its name comes from the fact that $\zeta\, \mathbf{H}.\mathbf{H}^*$ can describe the covariance matrix in a sample (of size p) from a multivariate distribution with covariance matrix $\zeta \mathbf{I}_n$. When the entries of $\mathbf{H}$ are normally distributed, $n\, \mathbf{H}.\mathbf{H}^*$ follows the standard Wishart distribution: $n\, \mathbf{H}.\mathbf{H}^* \sim W_p(\mathbf{I}_n)$ with p degrees of freedom. As $(p, n) \to \infty$ with a fixed ratio ($p/n \to \beta$), the spectral distribution of $\mathbf{H}.\mathbf{H}^*$ converges to a nonrandom LSD whose pdf is given by eq. (1.12) in (TULINO and VERDÙ 2004)

$$\rho_{\mathbf{H}.\mathbf{H}^*} \Rightarrow \tilde{\rho}_{\mathbf{H}.\mathbf{H}^*}(x) = (1-\beta)^+ \delta(x) + \frac{\sqrt{(b-x)(x-a)}}{2\,\pi\,\zeta\,x}, \quad x \in [a, b] \qquad , \qquad \text{(A1.1)}$$
$$a = \zeta\left(1-\sqrt{\beta}\right)^2 \text{ and } b = \zeta\left(1+\sqrt{\beta}\right)^2$$

where $(x)^+ = \max(x, 0)$ and $\delta(.)$ is the Dirac delta function.

This pdf has a point mass at zero with weight $1-\beta$ whenever $0 < \beta < 1$, plus a bulk of positive eigenvalues with weight $\min(\beta, 1)$. When $\beta \geq 1$ (our case of interest), the left hand term vanishes and all eigenvalues are strictly positive: we then retrieve eq. (3) of the main text. The pdf of the bulk is $\rho_{+(x)} = \rho(\lambda|\lambda > 0) = \sqrt{(b-x)(x-a)}/(2\pi\,\zeta\,\lambda\,\min(\beta, 1))$. As stated in main text, I refer to this

distribution as the scaled M-P law with scale parameter ζ and ratio index β and denote the result by the stochastic representation $\lambda_{\zeta\,\mathbf{H}.\mathbf{H}^*} \sim MP(\beta,\zeta)$. The mean eigenvalue is $\mathrm{E}(\lambda) = \zeta\,\beta$, with variance $\mathrm{V}(\lambda) = \zeta^2\beta$. Note that the original (and more standard) statement of the M-P law (TULINO and VERDÙ 2004; BAI and SILVERSTEIN 2010) is in terms of the LSD of $\mathbf{H}^*.\mathbf{H}$, which has ratio index $1/\beta$ and corresponding scale $\zeta\beta$: $\lambda_{\mathbf{H}^*.\mathbf{H}} \sim MP(1/\beta,\zeta\beta)$ in our notation. Note also that, for notational simplicity, I drop the reference to the scaling factor ζ in the indexing when referring to the LSD of sample covariance matrices ($\tilde{\rho}_{\mathbf{H}.\mathbf{H}^*}(x), v_{\mathbf{H}.\mathbf{H}^*}(x)$ etc.).

One powerful property of the M-P law (and of random matrix theory in general) is that the actual spectral distributions of finite random covariance matrices converge quickly to their limit: for any single draw of a matrix $\mathbf{H}$ with dimensions, say, $n = 10$ and $p = 100$ the spectral distribution of $\mathbf{H}.\mathbf{H}^*$ is already well described by the M-P law given above, in that it is bounded within the predicted domain $[a,b]$ in eq. (A1.1), and that the pdf $\rho(x)$ is close to the M-P law. However, it is difficult to represent this pdf with small $\min(n,p)$ as there are then only few eigenvalues to show. This is why I show examples with larger dimensions (e.g. $n = 100, p = 500$). However, when simulating the DFE I will use smaller parameter values and show the convergence of the DFE to the predicted distribution based on Random Matrix Theory.

Transforms of the LSD: Various transforms of the LSD of random matrices have been defined in Random Matrix Theory, I propose here a quick overview, drawn from section 2.2 of (TULINO and VERDÙ 2004). The purpose of these transforms is akin to that of generating functions in standard probability theory. These transforms can be derived from one another and allow to derive various properties of the LSD, or to compute the LSD of the sum or product of random matrices, which will prove useful in our case. Each transform fully characterizes a given LSD, just as the pdf $\tilde{\rho}(x)$ does. The mutual relationships between these transforms are illustrated in the notebook **file S3**.

Let $\tilde{\rho}(x)$ be the pdf of the LSD of a random matrix, defined on some finite or infinite range $[\lambda_{min},\lambda_{max}]$. The first important transform is the Stieltjes transform: $S(z) = \int_{\lambda_{min}}^{\lambda_{max}} 1/(x-z)\rho(x)dx$ with $z \in \mathbb{C}$ the set of complex numbers. It provides the range and pdf of the eigenvalues of the matrix, in particular $\tilde{\rho}(x) = \lim_{\omega\to 0+} S(x + \mathbf{i}\,\omega)$ where $\mathbf{i}$ is the complex unit number. The η transform: $\eta(z) = \int_{\lambda_{min}}^{\lambda_{max}} 1/(1+z\,x)\tilde{\rho}(x)dx$, with $z \in \mathbb{R}^+$(positive real numbers) is defined for positive semi-definite matrices only (with eigenvalues all positive or zero, $[\lambda_{min},\lambda_{max}] \subset \mathbb{R}^+$). This function is a generating function for the raw moments of the LSD: we have $\eta(0) = 1$ and $\eta[-z] = \sum_{k=0}^{\infty} z^k \mathrm{E}(\lambda^k)$. The η transform is related to the Stieltjes transform via $\eta(z) = S(-1/z)/z$. For the scaled M-P law in eq. (A1.1), the η transform is given by

$$\begin{aligned} \eta_{\mathbf{H}.\mathbf{H}^*}(z) &= 1 - \phi(z)/(4\,z\,\zeta) \\ \phi(z) &= \left(\sqrt{b\,z+1} - \sqrt{a\,z+1}\right)^2 \end{aligned} \quad , \qquad \text{(A1.2)}$$

where (a, b) are given in eq. (A1.1). Taking the limit at infinity provides the weight of the point mass at zero (the proportion of zero eigenvalues) $\lim_{z\to\infty} \eta_{\mathbf{H}.\mathbf{H}^*}(z) = \max(1-\beta, 0)$.

Central to our applications is the Shannon transform $v(z) = \int_{\lambda_{min}}^{\lambda_{max}} \log(1 + x\, z)\, \tilde{\rho}(x) dx$, also defined only for positive semi-definite matrices ($[\lambda_{min}, \lambda_{max}] \subset \mathbb{R}^+$), with $z \in \mathbb{R}^+$. It is related to the η transform via $v(z) = \int \big(1 - \eta(z)\big)/z\, dz$. For the scaled M-P law the Shannon transform is

$$v_{\mathbf{H}.\mathbf{H}^*}(z) = \beta \log\left(1 + \zeta\, z - \frac{\phi(z)}{4}\right) + \log\left(1 + \zeta\, z\, \beta - \frac{\phi(z)}{4}\right) - \frac{\phi(z)}{4\,\zeta\, z} \quad , \qquad \text{(A1.3)}$$

Where $\phi(.)$ is given in eq.(A1.2).

The two next transforms are key to compute the LSD of sums or products of random matrices whose entries are drawn independently, more specifically matrices that are asymptotically free, meaning that their LSD are independent (see details on free probability in section 2.4 p. 77 of TULINO and VERDÙ 2004). In the next section, we will use these transforms to compute approximations of the spectral distribution of various covariance matrices in terms of an M-P law with modified parameters. The R transform is related to the Stieltjes transform via $R(z) = S^{-1}(z) - 1/z$ where $S^{-1}(.)$ is the functional inverse of S such that $S^{-1}\big(S(z)\big) = z$. The R transform of the LSD of the sum of two matrices ($\mathbf{A}$ and $\mathbf{B}$) with entries independently drawn is the sum of the components' R transforms: $R_{\mathbf{A}+\mathbf{B}}(z) = R_{\mathbf{A}}(z) + R_{\mathbf{B}}(z)$. For the scaled M-P law, the R transform is

$$R_{\mathbf{H}.\mathbf{H}^*}(z) = \frac{\beta\, \zeta}{1 - z\, \zeta} \quad . \qquad \text{(A1.4)}$$

The S transform plays the exact same role as the R transform for products of semi-definite matrices. It is related to the η transform via $\Sigma(z) = -(z+1)/z\, \eta^{-1}(z+1)$ where $\eta^{-1}(.)$ is the functional inverse of η. The S transform of the LSD of the product of two positive semi-definite matrices ($\mathbf{A}$ and $\mathbf{B}$), with entries independently drawn, is $\Sigma_{\mathbf{A}.\mathbf{B}}(x) = \Sigma_{\mathbf{A}}(x)\Sigma_{\mathbf{B}}(x)$. For the scaled M-P law, the S transform is

$$\Sigma_{\mathbf{H}.\mathbf{H}^*}(z) = \frac{1}{(\beta + z)\, \zeta} \quad . \qquad \text{(A1.5)}$$

Another formula will prove particularly useful in what follows: for any $n \times n$ positive semi-definite matrix $\mathbf{T}$ whose LSD exists and has S transform $\Sigma_{\mathbf{T}}(z)$, we have (adapted from 2.216 p. 91 of TULINO and VERDÙ 2004):

$$\Sigma_{\mathbf{H}.\mathbf{T}.\mathbf{H}^*}(z) = \frac{z+1}{z+\beta} \Sigma_{\mathbf{H}^*.\mathbf{H}}\left(\frac{z}{\beta}\right) \Sigma_{\mathbf{T}}\left(\frac{z}{\beta}\right) = \frac{1}{\zeta(z+\beta)} \Sigma_{\mathbf{T}}\left(\frac{z}{\beta}\right) \quad . \qquad \text{(A1.6)}$$

The last transform used here is the D transform, introduced in (BENAYCH-GEORGES and NADAKUDITI 2011). It will prove useful to predict the behavior of the maximal eigenvalue when the entries h_{ij} have non zero mean. In this article, it is defined in terms of the distribution of singular values of $\mathbf{H}$ (when $\mathbf{H} = \mathbf{H}_0$ has entries with a zero mean). As the singular values of $\mathbf{H}$ are simply the square roots of the non zero eigenvalues of $\mathbf{H}.\mathbf{H}^*$, the D transform in (BENAYCH-GEORGES and NADAKUDITI 2011) can also be expressed in terms of the spectral distribution of $\mathbf{H}.\mathbf{H}^*$. More precisely, let $\varphi(z) = \int_a^b z/(z^2 - x)\rho(x)dx = \eta(-1/z^2)/z$, the D transform is $D(z) = \varphi(z)(\varphi(z)/\beta + (1 - 1/\beta)/z)$ which, for the M-P law, yields

$$D_{\mathbf{H}.\mathbf{H}^*}(z) = \frac{z^2}{2\beta\zeta^2}\left(\left(1 - \sqrt{1 + \frac{\zeta\left((\beta-1)^2\zeta - 2z^2(1+\beta)\right)}{z^4}}\right) - (1+\beta)\zeta\right) \quad . \qquad \text{(A1.7)}$$

In what follows we use these transforms and their approximations to derive an M-P law approximation for the spectral distribution of $\mathbf{M}$ when $1 \ll n \ll p$.

B) Spectral distribution of $\mathbf{M}$ with high phenotypic integration

In this section, I study the matrix $\mathbf{M}$ of mutational covariance among optimized traits, under the model of integrative phenotypic network described in **Figure 1**. I first describe its structure in detail, then derive an approximation of its LSD in terms of an M-P law.

***Structure of the mutational covariance matrix M*:** Matrix $\mathbf{B} = \{b_{ij}\}_{i\in[1,n],j\in[1,p]}$ is an $n \times p$ matrix of pathway coefficients given by the first derivatives of the developmental function about the parent phenotype. As explained in the main text, the $1 \times p$ vector $\mathbf{b}_i = \{b_{ij}\}_{j\in[1,p]}$ of the p pathway coefficients determining a given trait y_i is a single draw from a multivariate distribution. This distribution has mean vector $\boldsymbol{\mu}_\mathbf{B}$ and positive-definite $p \times p$ covariance matrix $\mathbf{C}_\mathbf{B}$. Otherwise the nature of these distributions is unspecified. As its entries are randomly distributed, $\mathbf{M} = \mathbf{B}.\mathbf{V}.\mathbf{B}^*$ has the structure of a sample covariance matrix (BAI and SILVERSTEIN 2010).

In order to go any further, we must characterize the structure of $\mathbf{M}$ in more mathematical detail. Let us first ignore any potential bias in the b_{ij} ($\boldsymbol{\mu}_\mathbf{B} = \mathbf{0}$) and denote by $\mathbf{B}_\mathbf{0}$ the matrix of pathway coefficients in this case. By assumption (8), the matrix $\mathbf{B}_0$ can be decomposed into the product $\mathbf{B}_\mathbf{0} = \mathbf{H}.\mathbf{A}$ where $\mathbf{H}$ is an $n \times p$ matrix with independent entries h_{ij} with mean 0 and variance $V(h_{ij}) = 1/n$, and $\mathbf{A}$ is a $p \times p$ matrix introduced to generate the suitable covariance among b_{ij}'s (matrix bending). By this definition, $\mathbf{C}_\mathbf{B} = E(\mathbf{B}.\mathbf{B}^*) = \mathbf{A}^*.\mathbf{A}/n$, which is positive-definite as required. The matrix $\mathbf{A}$ can thus be set as $\mathbf{A} = \sqrt{n}\,\mathbf{C}_\mathbf{B}^{1/2}$, the Cholesky decomposition of $n\,\mathbf{C}_\mathbf{B}$, but there will typically be many other possibilities. By definition, a given draw of the matrix $\mathbf{B}_0$ corresponds to a given draw of the matrix $\mathbf{H}$. This matrix $\mathbf{H}$ is

the building block of most models of Random Matrix Theory, and we derive the structure of $\mathbf{B}.\mathbf{V}.\mathbf{B}^*$ in terms of this building block.

From now on, we take the expectation $E(.)$ to mean the expected outcome of a given draw of the random coefficients b_{ij} (and the corresponding h_{ij}). Importantly, the random entries in matrix $\mathbf{H}.\mathbf{H}^*$ are independent of $\mathbf{C_B}$. If we now consider the general model where $\boldsymbol{\mu_B} \neq \mathbf{0}$, the bias in the distribution of the pathway coefficients b_{ij} boils down to adding an $n \times p$ matrix $\mathbf{U_B}$ to $\mathbf{B_o}$. We can write $\mathbf{B} = \mathbf{B_o} + \mathbf{U_B}$ where $\mathbf{B_o} = \mathbf{H}.\mathbf{A}$ and matrix $\mathbf{U_B}$ has all its n line vectors equal to $\boldsymbol{\mu_B}$, so by construction $\mathbf{U_B}$ has rank 1.

We can now derive the structure of the random matrix $\mathbf{B}.\mathbf{V}.\mathbf{B}^*$: it can be decomposed into $\mathbf{B}.\mathbf{V}.\mathbf{B}^* = \mathbf{K}.\mathbf{K}^*$ where $\mathbf{K} = \mathbf{K}_\mathrm{o} + \mathbf{U}_\mathrm{K}$ with $\mathbf{K}_\mathrm{o} = \mathbf{H}.\mathbf{A}.\mathbf{V}^{1/2}$ and $\mathbf{U_K} = \mathbf{U_B}.\mathbf{V}^{1/2}$. The singular values of $\mathbf{K}_\mathrm{o}$ are the square roots of the eigenvalues of $\mathbf{K}_\mathrm{o}.\mathbf{K}_\mathrm{o}^* = \mathbf{H}.\mathbf{W}.\mathbf{H}^*$, where $\mathbf{W} = \mathbf{A}.\mathbf{V}.\mathbf{A}^*$ is a positive-definite $p \times p$ matrix. The spectral distribution of $\mathbf{W}$ is equal to that of $n\,\mathbf{V}.\mathbf{C}_\mathrm{B}$. Like $\mathbf{U_B}$, matrix $\mathbf{U_K}$ is a rank 1 matrix by construction. Indeed, basic properties of the rank of matrix products imply that $0 < rank(\mathbf{U_K}) \leq \min\big(rank(\mathbf{U_B}), rank(\mathbf{V})\big) = \min(1, p) = 1$. Therefore, $\mathbf{U_K}$ has a single non-zero singular value θ which can be computed by using the fact that $\mathbf{U_K}.\mathbf{U_K^*}$ then has a single non-zero eigenvalue equal to θ^2 by definition. Therefore, its matrix trace $Tr(.)$ must satisfy $\theta^2 = Tr(\mathbf{U_K}.\mathbf{U_K^*})$. By construction of matrix $\mathbf{U_K} = \mathbf{U_B}.\mathbf{V}^{1/2}$, we also have $Tr(\mathbf{U_K}.\mathbf{U_K^*}) = \boldsymbol{\mu}_\mathbf{B}^*.\mathbf{V}.\boldsymbol{\mu_B}$ so the unique singular value of $\mathbf{U_K}$ is $\theta = \sqrt{\boldsymbol{\mu}_\mathbf{B}^*.\mathbf{V}.\boldsymbol{\mu_B}}$.

To summarize, matrix $\mathbf{M}$ has the same eigenvalues as a (non-standard) sample covariance matrix $\mathbf{K}.\mathbf{K}^*$ where $\mathbf{K} = \mathbf{K}_\mathrm{o} + \mathbf{U_K}$ with $\mathbf{K}_\mathrm{o} = \mathbf{H}.\mathbf{W}^{1/2}$ a 'standard' random matrix $\mathbf{H}$ multiplied by a positive-definite 'constant' matrix $\mathbf{W} = \mathbf{A}.\mathbf{V}.\mathbf{A}^*$. Matrix $\mathbf{U_K}$ is of rank 1 and its unique singular value is $\theta = \sqrt{\boldsymbol{\mu}_\mathbf{B}^*.\mathbf{V}.\boldsymbol{\mu_B}}$. Both components are $n \times p$ matrices. This whole decomposition argument is exemplified in the notebook **file S3**.

Existence of the LSD of M: A cornerstone result of RMT is that the spectral distribution of $\mathbf{H}.\mathbf{H}^*$ admits a limit when $n\,p \to \infty$, given by the M-P law: $\lambda_{H.H^*} \sim MP(p/n, 1)$. In the absence of a bias among pathway coefficients ($\boldsymbol{\mu_B} = \mathbf{0}$), $\mathbf{M} = \mathbf{B_o}.\mathbf{B_o^*} = \mathbf{H}.\mathbf{W}.\mathbf{H}^*$ is a non-standard sample covariance matrix: as the entries of $\mathbf{H}$ are independent of $\mathbf{W}$ **,** the existence of an LSD for this matrix is also certain (chapter 4 of BAI and SILVERSTEIN 2010). This LSD is fully determined by that of $\mathbf{W}$ and $\mathbf{H}.\mathbf{H^T}$ taken separately, and is still independent of the nature of the distributions of the entries in $\mathbf{H}$. Tools from RMT can then be used to compute explicit approximations for the LSD of such a matrix product. The effect of bias, as we will see, does not affect the existence of an LSD, as it only modifies the leading eigenvalue of $\mathbf{M}$.

Approximation to the LSD of matrix $\mathbf{M}$ when $\boldsymbol{\mu_B} = \mathbf{0}$: Here, I describe how the LSD of the sample covariance matrix $\zeta\,\mathbf{H}.\mathbf{H}^*$ is modified by inner multiplication by the matrix $\mathbf{W}$, yielding the LSD of $\mathbf{M} = \mathbf{H}.\mathbf{W}.\mathbf{H}^*$.

As shown above, the matrix $\mathbf{W} = \mathbf{A}.\mathbf{V}.\mathbf{A}^*$ is an unspecified positive-definite matrix, whose eigenvalues are the same as those of $n\,\mathbf{V}.\mathbf{C_B}$. Overall, $\mathbf{W}$ gathers all the correlation among mutation effects on mutable traits and among pathway coefficients within $\mathbf{B}$. It is impossible to have a general *a priori* knowledge on the structure of $\mathbf{W}$, so we must rely on an approximate treatment under less general conditions, in order to keep as much generality as possible regarding $\mathbf{W}$. This is possible if we assume that the phenotypic integration from mutable to optimized traits is high (our assumption (4) : $\beta = p/n \gg 1$).

First, let us note that the LSD of $\mathbf{H}.\mathbf{W}.\mathbf{H}^*$ always exists as long as $\mathbf{W}$ is positive-definite and has a bounded spectral distribution (Chapter 4 in BAI and SILVERSTEIN 2010). Under the additional assumption that $n/p \to 0$ (or $\beta \to \infty$), we can approximate eq. (A1.6) to obtain a simple limit for this LSD. Let $\eta_{\mathbf{W}}(z)$ be the (unspecified) η transform of the spectral distribution of $\mathbf{W}$, and let $\eta_{\mathbf{W}}^{-1}(x)$ be the functional inverse of this transform. From the relationship between η and S transforms ($\Sigma(z) = -(z+1)/z\,\eta^{-1}(z+1)$) we can express eq. (A1.6) as $\Sigma_{\mathbf{H}.\mathbf{W}.\mathbf{H}^*}(x) = -\eta_{\mathbf{W}}^{-1}(1+x/\beta)/x$. Assuming that β is large, we can then take a series expansion of $\eta_{\mathbf{W}}^{-1}(1+x/\beta)$ for small x/β which gives

$$\eta_{\mathbf{W}}^{-1}\left(1+\frac{x}{\beta}\right) = \eta_{\mathbf{W}}^{-1}(1) + \frac{x}{\beta}\eta_{\mathbf{W}}^{-1\prime}(1) + \frac{x}{2\beta^2}\,\eta_{\mathbf{W}}^{-1\prime\prime}(1) + o\left(\frac{x}{\beta^2}\right) \quad . \qquad \text{(A1.8)}$$

This Taylor series can be expressed in terms of the derivatives of $\eta_{\mathbf{W}}(z)$ taken at $z = \eta_{\mathbf{W}}^{-1}(1) = 0$ (e.g. applying the method in KOEPF 1994), and we find that $\eta_{\mathbf{W}}^{-1\prime}(1) = 1/\eta_{\mathbf{W}}{}'(0)$ and $\eta_{\mathbf{W}}^{-1\prime\prime}(1) = -\eta_{\mathbf{W}}{}''(0)/(\eta_{\mathbf{W}}{}'(0)^3)$. Recalling that the η transform is a moment generating function for the LSD ($\eta[-z] = \sum_{k=0}^{\infty} z^k \mathrm{E}(\lambda^k)$), we have $\eta_{\mathbf{W}}^{-1\prime}(1) = -1/\zeta_{\mathbf{W}}$ and $\eta_{\mathbf{W}}^{-1\prime\prime}(1) = 2(1+cv_{\mathbf{W}}^2)/\zeta_{\mathbf{W}}$, where $\zeta_{\mathbf{W}}$ is the mean of the eigenvalues of $\mathbf{W}$ and $cv_{\mathbf{W}}$ is their coefficient of variation. Plugging these expressions into eq. (A1.8), we obtain a simple expression for the inverse η transform, which depends only on the mean and coefficient of variation of the eigenvalues of $\mathbf{W}$:

$$\eta_{\mathbf{W}}^{-1}\left(1+\frac{x}{\beta}\right) \approx \frac{x^2\left(1+cv_{\mathbf{W}}^2\right) - x\,\beta}{\zeta_{\mathbf{W}}\,\beta^2} + o\left(\frac{x}{\beta}\right)^2 \quad . \qquad \text{(A1.9)}$$

Plugging eq. (A1.9) into the formula $\Sigma_{\mathbf{H}.\mathbf{W}.\mathbf{H}^*}(x) = -\eta_{\mathbf{W}}^{-1}(1+x/\beta)/x$, we obtain a linear function of x: introducing the parameters $\zeta_e = \zeta_{\mathbf{W}}\left(1+cv_{\mathbf{W}}^2\right)$ and $\beta_e = \beta/\left(1+cv_{\mathbf{W}}^2\right)$:

$$\Sigma_{\mathbf{M}}(x) = \Sigma_{\mathbf{H}.\mathbf{W}.\mathbf{H}^*}(x) = -\frac{\eta_{\mathbf{W}}^{-1}\left(1+\frac{x}{\beta}\right)}{x} = \frac{1}{\zeta_e\beta_e}\left(1-\frac{x}{\beta_e}\right) + o\left(\frac{x}{\beta^2}\right) \quad , \qquad \text{(A1.10)}$$

which can be rearranged into $\Sigma_{\mathbf{M}}(x) = \zeta_e^{-1}\left(x+\beta_e\right)^{-1} + \mathrm{o}(x\beta^{-2})$. To first order in x/β^2, we retrieve the S transform of an M-P law (eq.(A1.5)) with parameters ζ_e and β_e.

Therefore, we obtain a simple approximation to the LSD of the mutational covariance, accounting for arbitrary covariance matrices among mutation effects on metabolic traits or among

pathway coefficients. This approximation is to the second order in $1/\beta$ and is thus relatively robust: it remains valid even when p is not much larger than n. All along, we made no assumption on the mean or variance of the spectral distribution of $\mathbf{W}$ (on $\zeta_{\mathbf{W}}$ or $cv^2_{\mathbf{W}}$). Therefore, eq. (A1.10) remains valid even with high correlations and heterogeneity within $\mathbf{V}$ or $\mathbf{B}$. As expected, we retrieve the original M-P law when $\mathbf{W} = \zeta_{\mathbf{W}}\mathbf{I}_p$: in this case $cv^2_{\mathbf{W}} = 0$ (all the eigenvalues of $\mathbf{W}$ are equal to their mean $\zeta_{\mathbf{W}}$) so that $\beta_e = \beta$ and $\zeta_e = \zeta_{\mathbf{W}}$, and multiplication by $\mathbf{W}$ results in a mere scaling.

Obviously, this derivation cannot claim full mathematical rigor: caution might be necessary on how the approximation of the S transform implies convergence in distribution to the corresponding LSD. Yet, **Figure 4.a** and other extensive simulations (not shown) do suggest that this approximation is valid: the convergence of the spectral distribution of $\mathbf{M}$ to this modified M-P law is good, even with very high heterogeneity in the eigenvalues of $\mathbf{W}$ (here $cv^2_{\mathbf{W}} = 2.1$).

Effect of a non-zero mean in b_{ij}: The above treatment describes the LSD of $\mathbf{K_o}.\mathbf{K_o^*}$ that equals that of $\mathbf{M}$ whenever the $\mathbf{b}_i$'s are unbiased ($\boldsymbol{\mu}_{\mathrm{B}} = \mathbf{0}$), see main text. However, in general we want to allow for arbitrary $\boldsymbol{\mu}_{\mathrm{B}} \neq \mathbf{0}$. To do so, we rely on Theorem 2.9 of Benaych-Georges & Nadakuditi (2011) on the effect of small rank perturbations on the singular value distribution of random rectangular matrices (such as $\mathbf{K_o}$). As the singular values of $\mathbf{K}$ are simply the square roots of the eigenvalues of $\mathbf{K}.\mathbf{K}^*$ and thus of $\mathbf{M}$, this provides the required result.

Let us order the eigenvalues of $\mathbf{M}$ in decreasing order ($\lambda_1 \geq \lambda_2 .. \geq \lambda_n > 0$). To generate $\mathbf{K}$, the random matrix $\mathbf{K}_{\mathrm{o}}$ (the standard one with zero mean entries) is "additively perturbed" by a small rank matrix $\mathbf{U}_{\mathrm{K}}$: $\mathbf{K} = \mathbf{K}_{\mathrm{o}} + \mathbf{U}_{\mathrm{K}}$ has a number of singular values that differ from those of $\mathbf{K}_{\mathrm{o}}$. When $\mathbf{K}$ is a large random matrix, these differing eigenvalues are the r largest in general, where r is the rank of the perturbation matrix (BENAYCH-GEORGES and NADAKUDITI 2011). The perturbation $\mathbf{U_K}$ is of rank 1 with a unique singular value $\theta = \sqrt{\boldsymbol{\mu}_{\mathbf{B}}^*.\mathbf{V}.\boldsymbol{\mu}_{\mathbf{B}}}$ (see above). Therefore, only the leading singular value $\sqrt{\lambda_1}$ is affected by the perturbation, while all the lower singular values $\sqrt{\lambda_i}$ of $\mathbf{K}$ retain the same distribution as those of $\mathbf{K}_{\mathrm{o}}$. The leading singular value $\sqrt{\lambda_1}$ shows a phase transition behavior that is determined by what the Benaych-Georges and Nadakuditi dubbed the D-transform of the singular value distribution, which is the expectation $D(z) = \mathrm{E}_{\sigma}(z/(z^2 - \sigma_o))$ over the distribution of the singular values σ_o of the unperturbed matrix, in their notation. In our context, the unperturbed matrix is $\mathbf{K_o}$, and it is useful to express $D(z)$ in terms of the eigenvalues $\lambda_o = \sigma_o^2$ of $\mathbf{K_o}$. This gives: $D(z) = \mathrm{E}_{\lambda_o}(z/(z^2 - \lambda_o))$ over the spectral distribution of $\mathbf{K_o}.\mathbf{K_o^*}$. Now as we have seen above, the spectral distribution of $\mathbf{K_o}.\mathbf{K_o^*}$ can be approximated by an M-P law: $\lambda_o \sim MP(\beta_e, \zeta_e)$, so the corresponding D transform is $D_{\mathbf{H}.\mathbf{H}^*}(z)$ given above (eq.(A1.7)) for the M-P law.

To summarize: the effect of the bias in the b_{ij} is (i) no effect on the bulk of lower eigenvalues ($\lambda_{i\neq 1} \sim MP(\beta_e, \zeta_e)$) and (ii) a "phase transition" behavior for the leading eigenvalue λ_1. Denote $\theta_{th} = 1/\sqrt{D_{\mathbf{H}.\mathbf{H}^*}(\max \sigma_o)}$ the functional inverse of the D transform, taken at $\max \sigma_o = \max \sqrt{\lambda_o}$. Whenever $0 \leq \theta^2 < \theta_{th}^2$, all eigenvalues (including λ_1) pertain to the M-P law so that $\lambda_1 \to \max \lambda_o$

which is the expected maximum of the spectral distribution of the unperturbed matrix. However, whenever $\theta^2 > \theta_{th}^2$, λ_1 rises above the bulk of smaller eigenvalues, to a higher value $\sqrt{\lambda_1} \rightarrow D_{\mathbf{H}.\mathbf{H}^*}^{-1}(1/\theta^2)$.

Using the M-P law approximation to the LSD of $\mathbf{K_o}.\mathbf{K_o^*}$ ($\lambda_o \sim MP(\beta_e, \zeta_e)$), the maximum of the M-P law is $\max \sigma_o^2 = \max \lambda_o = \left(1 + \sqrt{\beta_e}\right)^2 \zeta_e$, so the threshold for the phase transition is

$$\theta_{th} = \frac{1}{\sqrt{D_{\mathbf{H}.\mathbf{H}^*}\left(\max \sqrt{\lambda_o}\right)}} = \beta_e^{1/4} \sqrt{\zeta_e} \quad . \qquad \text{(A1.11)}$$

The limit reached by the leading eigenvalue beyond the phase transition depends on the functional inverse of the D transform of λ_o, which, for the M-P law, is (from eq. (A1.7)):

$$D_{\mathbf{H}.\mathbf{H}^*}^{-1}(z) = \sqrt{\frac{(1 + z\,\zeta_e)(1 + z\,\zeta_e\,\beta_e)}{z}} \quad . \qquad \text{(A1.12)}$$

This yields the limit:

$$\lambda_1 \underset{\theta > \theta_{th}}{\rightarrow} \left(D_{\mathbf{H}.\mathbf{H}^*}^{-1}\left(\frac{1}{\theta^2}\right)\right)^2 = \frac{(\zeta_e + \theta^2)(\beta_e \zeta_e + \theta^2)}{\theta^2} \quad . \qquad \text{(A1.13)}$$

We can express this result in more intuitively amenable terms. Define $cv = \sqrt{Tr(\mathbf{C}_\mathrm{B}.\mathbf{V})/(\boldsymbol{\mu}_\mathrm{B}^*.\mathbf{V}.\boldsymbol{\mu}_\mathrm{B})}$, which is analogous to a coefficient of variation of the means μ_j but modified by the mutational covariance $\mathbf{V}$. When $\mathbf{V} \propto \mathbf{I}_p$, this is exactly the mean coefficient of variation of the linear coefficients b_{ij} across mutable traits x_j. Recall that $\zeta_e = \zeta_\mathbf{W}\, p/p_e$ and that the $p \times p$ matrix $\mathbf{W}$ has the same spectrum as $n\,\mathbf{V}.\mathbf{C}_\mathrm{B}$: its mean eigenvalue is therefore $\zeta_\mathbf{W} = Tr(n\,\mathbf{V}.\mathbf{C}_\mathrm{B})/p = n/p\; Tr(\mathbf{C}_\mathrm{B}.\mathbf{V})$. Putting this together, we get

$$\frac{\theta^2}{\theta_{th}^2} = \frac{\theta^2}{\sqrt{\beta_e}\zeta_e} = \frac{\sqrt{n\,p_e}}{cv^2} \quad . \qquad \text{(A1.14)}$$

We see that when the bias in the coefficients b_{ij} is small enough, cv^2 is large enough to outweigh $\sqrt{n\,p_e}$. By a "small enough bias", we mean specifically that the cumulated variance of mutational effects on mutable traits, and of the linear coefficients is larger than $\boldsymbol{\mu}_\mathrm{B}^*.\mathbf{V}.\boldsymbol{\mu}_\mathrm{B}$. Otherwise, a phase transition appears and the leading eigenvalue rises above the bulk of lower eigenvalues, by a factor

$$\frac{\lambda_1}{\mathrm{E}(\lambda_o)} \underset{\theta > \theta_{th}}{\rightarrow} \left(1 + \frac{n}{cv^2}\right)\left(1 + \frac{cv^2}{p_e}\right) \underset{p_e \rightarrow \infty}{\rightarrow} 1 + \frac{n}{cv^2} \quad . \qquad \text{(A1.15)}$$

This is our eq. (6), except that we replaced λ_o (all eigenvalues of $\mathbf{K}.\mathbf{K}^*$ when $\boldsymbol{\mu}_\mathbf{B} = \mathbf{0}$) by $\lambda_{i\neq 1}$ (the set of $n-1$ smallest eigenvalues, namely the "bulk" eigenvalues in the general situation (arbitrary $\boldsymbol{\mu}_\mathbf{B}$). The latter is distributed as λ_o, except that it is depleted of its maximal value λ_1, so it is very slightly (unnoticeably if $n \gg 1$) biased downwards.

The accuracy of this result is checked **Figure 4b**. In this figure, the actual LSD of $\mathbf{M}$ is not exactly the M-P law because of the covariances in mutable traits ($\mathbf{V} \neq \mathbf{I}_p$) and pathway coefficients ($\mathbf{C_B} \neq \mathbf{I}_p$), so that $\mathbf{H}.\mathbf{W}.\mathbf{H}^* \neq \zeta_W \mathbf{H}.\mathbf{H}^*$. However, the prediction for the phase transition behavior, which is based on the M-P law approximation (derivation above), is accurate. This is simply because the behavior of λ_1 is entirely determined by the spectral distribution of the unperturbed matrix (λ_o), which was shown to be accurately captured by the M-P law approximation as long as $n \ll p$ (e.g. **Figure 4a**).

Extensions to allow for correlations among the rows of B: In the end of the Discussion, I stress the fact that in the model as formulated, it is impossible to have substantial correlations among the rows of $\mathbf{B}$, namely among the b_{ij}, across i, for a given j. This is because the covariance $\mathbf{C_B}$ is of full rank p. A possible way to incorporate such correlations was suggested by one of the reviewers, by letting $\mathbf{B} = \mathbf{A_1}.\mathbf{H}.\mathbf{A_2}$ still $n \times p$, with $\mathbf{A_1}$ an $n \times n$ matrix and $\mathbf{A_2}$ a $p \times p$ matrix, both invertible (our model so far corresponds to $\mathbf{A_1} = \mathbf{I}_n$ and $\mathbf{A_2} = \mathbf{A}$). I provide a quick analysis of this case, but merely to illustrate how extensions can be made, and mostly in the form of conjectures rather than proofs. A full treatment of this more general model is beyond the scope of this article.

The covariance of b_{ij} among the rows i is now given by the $p \times p$ matrix $\mathbf{C_2} = E(\mathbf{B}^*.\mathbf{B}) = E\big((\mathbf{A_2^*}.\mathbf{H}^*.\mathbf{A_1^*}).(\mathbf{A_1}.\mathbf{H}.\mathbf{A_2})\big)$. When $\mathbf{A_1} = \mathbf{I}_n$ and $\mathbf{A_2} = \mathbf{A}$ (our former model), this gives $\mathbf{C_2} = \mathbf{C_B} = \mathbf{A}^*.E(\mathbf{H}^*.\mathbf{H}).\mathbf{A} = \mathbf{A}^*.\mathbf{A}/n$, as $E(\mathbf{H}^*.\mathbf{H}) = 1/n$. Even in the general case ($\mathbf{A_1} \neq \mathbf{I}_n$), with $n \ll p$ and $\mathbf{A_1}.\mathbf{A_1^*}$ of full rank n, we have convergence to $\mathbf{C_2} \propto \mathbf{A_2^*}.\mathbf{A_2}/n$. Conversely, the covariance of the b_{ij} among the columns j is now given by the $n \times n$ matrix $\mathbf{C_1} = E(\mathbf{B}.\mathbf{B}^*) = E\big((\mathbf{A_1}.\mathbf{H}.\mathbf{A_2}).(\mathbf{A_2^*}.\mathbf{H}^*.\mathbf{A_1^*})\big)$. In our former model $\mathbf{C_1} = E(\mathbf{H}.\mathbf{A_2}.\mathbf{A_2^*}.\mathbf{H}^*)$: as $n/p \to 0$ with $\mathbf{A_2}.\mathbf{A_2^*}$ of full rank p, we get $\mathbf{C_1} \propto \mathbf{I}_n$, and in the general case, we get $\mathbf{C_1} \propto \mathbf{A_1}.\mathbf{A_1^*}$. Therefore the introduction of matrix $\mathbf{A_1}$ does introduce a potential for correlations among the rows i, which is effectively negligible otherwise (provided $n \ll p$ and $\mathbf{C_B}$ is of full rank p), as conjectured.

Let us now see how the mutational covariance $\mathbf{M}$ is affected by multiplication of $\mathbf{B}$ on the left by $\mathbf{A_1}$. In our extension, if we let $\mathbf{M} = \mathbf{H}.\mathbf{A_2}.\mathbf{V}.\mathbf{A_2^*}.\mathbf{H}^*$ be the former form of the mutational covariance, we now have $\mathbf{M'} = \mathbf{A_1}.\mathbf{M}.\mathbf{A_1^*}$, as the new mutational covariance. The LSD of $\mathbf{M'}$ is the same as that of $\mathbf{C_1}.\mathbf{M}$ where we define $\mathbf{C_1} = \mathbf{A_1}.\mathbf{A_1^*}$ which is approximately the covariance matrix of b_{ij} among rows i. As the elements in $\mathbf{C_1}$ are independent of those in $\mathbf{H}$, $\mathbf{M'}$ has S transform given by $\Sigma_{\mathbf{M'}}(x) = \Sigma_{\mathbf{M}}(x)\Sigma_1(x)$ where Σ_1 is the S transform of the LSD of $\mathbf{C_1}$.

We must now approximate each S transform. We have seen above that the LSD of $\mathbf{M}$ is approximately the M-P law so that $\Sigma_{\mathbf{M}}(x) = 1/(\zeta_e(x+\beta_e))$ with ζ_e, β_e given above. I could not derive a general result for $\Sigma_1(.)$ without making an additional assumption. I assume that the spectral distribution

of $\mathbf{C_1}$ is not widely spread. More precisely, define$\lambda_{\mathbf{C_1}}$ and set, without loss of generality, that $E(\lambda_{\mathbf{C_1}}) = 1$ (all the scaling can be absorbed into ζ_e). I assume that (i) $v_1 = V(\lambda_{\mathbf{C_1}}) \ll 1$ and that (ii) the higher raw moments of the LSD of $\mathbf{C_1}$ scale with this variance $E(\lambda_{\mathbf{C_1}}^k) = 0(v_1^k)$ for all $k \geq 2$. This simply implies that the eigenvalues of $\mathbf{C_1}$ are not too spread apart, corresponding to mild correlations among the rows of $\mathbf{B}$. In this case, we can find a first order approximation for the S transform of the LSD of $\mathbf{C_1}$ and a corresponding approximation for the LSD of $\mathbf{M'}$ again in terms of an M-P law with modified parameters. As for the rest of the Appendix, details of the computations can be found in the notebook **file S3**.

Define $\eta_1(x) = E(1/(1 + \lambda_{C_1} x))$ the η transform of the LSD of $\mathbf{C_1}$ and $\eta_1^{-1}(.)$ the functional inverse such that $z = \eta_1\left(\eta_1^{-1}(z)\right)$. First, take the Taylor series expansion for ratio: $1/(1 + \eta_1^{-1}(z)\,\lambda_{\mathbf{C_1}})$ to order $\lambda_{\mathbf{C_1}}^2$; then take expectations with respect to the distribution of $\lambda_{\mathbf{C_1}}$: this yields

$$z = \eta_1\left(\eta_1^{-1}(z)\right) = E\left(\frac{1}{1 + \eta_1^{-1}(z)\,\lambda_{\mathbf{C_1}}}\right) = \frac{\left(v_1\,\eta_1^{-1}(z)\right)^2}{\left(1 + \eta_1^{-1}(z)\right)^3} + \frac{1}{1 + \eta_1^{-1}(z)} + O\left(v_1^{3/2}\right) \quad , \quad \text{(A1.16)}$$

Solving for $\eta_1^{-1}(z)$, we retrieve four solutions, but only one has the correct behavior when $v_1 \to 0$: namely, it converges to the solution when all $\lambda_{\mathbf{C_1}} = E(\lambda_{\mathbf{C_1}}) = 1$, which is $\eta_1^{-1}(z) = (1 - z)/z$. Taking this solution, we can compute the S transform $\Sigma_1(x)$ of $\mathbf{C_1}$ under the approximation. The expression is analytic but unpractical. Yet, after taking a series of $1/\Sigma_1(x)$ to leading order in v_1 , we obtain an approximate expression:

$$\Sigma_1(x) = -\frac{x+1}{x}\eta_1^{-1}(x+1) \underset{v_1 \ll 1}{\approx} \frac{1}{1 + v_1\,x} \quad . \quad \text{(A1.17)}$$

Interestingly, this approximation is exact when $\mathbf{C_1}$ is a sample covariance matrix, with LSD given by the M-P law ($\lambda_{\mathbf{C_1}} \sim MP(\beta_1, 1/\beta_1)$ and arbitrary $\beta_1 > 0$). Note that, in this particular case, this result should thus be valid for an arbitrary level of variation in the eigenvalues of $\mathbf{C_1}$ (arbitrary $\beta_1 > 0$).

Using eq. (A1.17), we can compute the S transform of the LSD of $\mathbf{M'}$: $\Sigma_{\mathbf{M'}}(x) = \Sigma_{\mathbf{M}}(x)\Sigma_1(x) \approx 1/\left((x + \beta_e)\,\zeta_e\,(1 + v_1 x)\right)$. In general, this S transform is not related to any specific form of random matrix. However, when the ratio index of $\mathbf{M} = \mathbf{H}_e.\mathbf{H}_e^*$ is large ($\beta_e \gg 1$), we retrieve a simple approximation in terms of an MP law. Taking the leading order when $v_1 = \mathrm{o}(1)$ while $\beta_e v_1 = 0(1)$ ($v_1 \ll 1$ and $\beta_e \gg 1$), we get the S transform of an M-P law with modified parameters $\zeta_e' = \zeta_e\,(1 + \beta_e v_1)$ and $\beta_e' = \beta_e/(1 + \beta_e v_1)$:

$$\Sigma_{\mathbf{M'}}(x) = \Sigma_{\mathbf{M}}(x)\Sigma_1(x) \underset{v_1 \ll 1}{\approx} \frac{1}{\zeta_e(x + \beta_e)\,(1 + v_1 x)} \underset{\beta_e \gg 1}{\approx} \frac{1}{\zeta_e'(1 + \beta_e'\,x)} \quad . \quad \text{(A1.18)}$$

As expected, we retrieve the original MP law when $v_1 \to 0$ (our former model: $\mathbf{M'} = \mathbf{M}$, with $\beta_e' = \beta_e$). The approximation in eq. (A1.18) must break whenever v_1 is large enough or β_e small enough that $0 <$

$\beta'_e < 1$. Indeed, in this case, the approximation would predict that a portion $(1 - \beta'_e)$ of eigenvalues should be zero whereas in fact there are none ($\beta_e > 1$ and $\mathbf{C_1}$ is positive-definite). This sets a limit for the validity of eq. (A1.18): we must have $0 < v_1 < (\beta_e - 1)/\beta_e$, i.e. $v_1 \leq 1$ in the best case scenario ($\beta_e \to \infty$).

To conclude, the effect of correlations in the rows of $\mathbf{B}$ is thus to further reduce the shape parameter of the LSD of the mutational covariance, while retaining the M-P law structure. When $n/p \to \infty$, so that $\beta_e \propto \beta \to \infty$, we retrieve a finite shape parameter this time:

$$\beta'_e = \frac{\beta_e}{(1 + \beta_e v_1)} \underset{\beta_e \to \infty}{\approx} \frac{1}{v_1} \quad . \qquad \text{(A1.19)}$$

Then, there is mild anisotropy, all the more as v_1 gets larger, namely as the rows of $\mathbf{B}$ get more correlated. The effective dimensionality in the sense of the matching moment approximation in (MARTIN and LENORMAND 2006), is $n_e = n/(1 + cv(\lambda)^2)$ where $cv(\lambda)$ is the coefficient of variation of the eigenvalues of $\mathbf{M}'$. With our M-P law approximation we have $cv(\lambda)^2 = 1/\beta'_e$ so that

$$n_e \underset{n \ll p}{\approx} \frac{n}{(1 + v_1)} \quad . \qquad \text{(A1.20)}$$

In the main text, I refer directly to $v_1 = cv_b^2$ as the coefficient of variation of the eigenvalues of the covariance matrix of the b_{ij} among rows i, to avoid stating the scaling $E(\lambda_{C_1}) = 1$ that was made here for mere notational simplicity. Simulations (not shown) suggest that this new M-P law approximation is indeed accurate as long as v_1 is small and β_e is large ($0 < v_1 < (\beta_e - 1)/\beta_e$).

References:

BAI, Z., and J. W. SILVERSTEIN, 2010 *Spectral Analysis of Large Dimensional Random Matrices, 2nd Edition*.

BENAYCH-GEORGES, F., and R. R. NADAKUDITI, 2011 The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. Advances in Mathematics **227:** 494-521.

KOEPF, W., 1994 Taylor polynomials of implicit functions, of inverse functions, and of solutions of ordinary differential equations. Complex Variables, Theory and Application: An International Journal **25:** 23-33.

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.

TULINO, A. M., and S. VERDÙ, 2004 *Random Matrix Theory and Wireless Communications*. Now Publishers.

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

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.

File S3

Supplementary Notebook

Available for download as a .cdf file at
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.160325/-/DC1